NATIONAL LIBRARY
OTTAWA

BIBLIOTHÈQUE NATIONALE
OTTAWA

CANADA

NAME OF AUTHOR.........David F. Liang..

TITLE OF THESIS....Nonlinear Estimation of Systems with and without Delays

UNIVERSITY..........of Alberta..............

DEGREE FOR WHICH THESIS WAS PRESENTED.........Ph.D.........

YEAR THIS DEGREE GRANTED..........1974.................

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

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

(Signed)......David Liang

PERMANENT ADDRESS:

508 Kelvin Blvd

Winnipeg

Manitoba

DATED..Oct. 3.......1974

NL-91 (10-68)

THE UNIVERSITY OF ALBERTA

NONLINEAR ESTIMATION OF SYSTEMS WITH AND WITHOUT DELAYS

by

© DAVID FESENG LIANG

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

AND RESEARCH IN PARTIAL FULFILMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL ENGINEERING

EDMONTON, ALBERTA

FALL, 1974

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled " NONLINEAR ESTIMATION OF SYSTEMS WITH AND WITHOUT DELAYS " submitted by David Feseng LIANG in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

....

Supervisor

....

....

....

....

....

External Examiner

Date ..September 26, 1974..........

ABSTRACT

In this thesis, equations are derived for the state estimates and error-covariances of discrete and continuous nonlinear systems with and without delays, corrupted by white noise as well as non-white noise processes.

In the case of continuous systems, new filtering algorithms are derived for nonlinear systems without delays, imbedded in white noise, correlated noise and noise free processes. The results obtained for white noise sequences are exact and optimal, with respect to the constraints imposed on the filtering dynamic equations, minimizing the error-variance cost functionals. The main technique makes use of the matrix minimum principle together with the Kolmogorov and Kushner equations to derive the optimal values of the coefficients in the estimation algorithms under the requirements that the estimates be unbiased. For non-white noise processes the results obtained are suboptimal since an approximate assumption has been made.

The algorithms can be implemented in computer evaluation and are recursive in nature under the assumption that the conditional probability density functions of the estimator errors are Gaussian. Various nonlinear systems were simulated and compared with results obtained from some widely used finite dimensional approximate nonlinear filters. The results clearly indicate the superiority of the proposed minimum variance filter. Results pertaining to linear problems can be easily deduced from the nonlinear estimation algorithms, they agree well with those derived in the literature, using other optimization techniques.

In the case of discrete-time systems, new nonlinear estimation algorithms, that directly yield the fixed-lag, fixed-point and fixed-interval smoothing and the filtering algorithms, are derived for nonlinear delayed systems with non-delayed measurements and multi-channel time-delayed measurements, corrupted by white noise, correlated noise and colored noise processes. The derivation makes use of the concept of the gradient matrix to minimize the error-variance, taken to be the estimation criterion, under the condition that the estimates be unbiased. The derivation is straightforward and clearly indicate the close link between three different classification of smoothers and the filtering estimator.

For systems with product types and polynomial nonlinearities, the only assumption needed to implement the algorithms in computer evaluation, is to assume that the conditional probability density functions of the estimator errors are Gaussian. They are expected to be computationally efficient, since no augmentation of state variables are introduced. And the results obtained are exact and optimal with respect to the imposed constraints on the dynamic equations of the estimators, minimizing the error variance cost functionals.

The results can also be applied to various special cases of nonlinear as well as linear systems with and without delays. For linear systems, the results are identified with those in the literature.

ACKNOWLEDGEMENTS

The author acknowledges with thanks the encouragement and guidance received from his supervisor Dr. G. S. Christensen throughout the course of this research.

Thanks are due to the National Research Council of Canada, the Izaak Walton Killam Memorial Fellowship Trust, and the Department of Electrical Engineering for the financial support during the course of this work. The computational facilities provided by the Computing Centre of the University of Alberta were of great help to the author.

The author owes a debt of gratitude to his parents for their constant moral support and encouragement. Mrs. Barb Gallaiford is to be thanked for typing the manuscript. Also, the author would be remiss without acknowledging the help and grace of God, the Creator of the universe and the author of life.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER I

GENERAL INTRODUCTION

1.1 *INTRODUCTION*

Physical systems are designed and built to perform certain defined functions. In order to determine whether a system is performing properly and ultimately to control the system performance, the engineer would like to know the state of his system. To determine the state of his system, the engineer must take measurements or make observations on his system. These measurements are generally contaminated with noise caused by the electronic and mechanical components of the device used.

The problem of determining the state of a system from noisy measurements is called estimation and is the main subject of this research project. In general, it includes the problems of filtering, prediction and smoothing.

In many applications, it is meaningful to assign a cost function to an estimate representing a quantitative measure of how good "an estimate" is. With a knowledge of the cost function, the system states, and parameters, one can determine the best control which minimizes (or maximizes) the cost function.

The problem delineated above is a rather old problem, dating back almost two centuries to Legendre (1806) and Gauss (1809). Gauss was interested in determining the orbital elements of a celestial body from observations and developed the technique that is known today as least-squares estimation.

In the early 1940's, Wiener [1] and Kolmogorov[2] discussed problems of linear least-squares estimation for stochastic processes corrupted by additive noise. Their solution was dependent upon the assumption of stationarity, ergodicity, and knowledge of the entire past of observed processes. But in actual practice, the signal and noise processes may not be stationary, nor may the observation time always start at $t = -\infty$.

Since then, numerous attempts have been made to remove one or both of the above restrictions. The notable work by Kalman [3,4] and Kalman and Bucy [5] resulted in by far the most successful solution to the problem. They considered the nonstationary linear system from the state space point of view with measurements corrupted by white Gaussian noise processes, and obtained recursive solutions for both the continuous and the discrete linear cases. The original derivation was based upon the derivation of the Wiener-Hopf equation, using the Orthogonal Projection Lemma [3]. This theory has become known as Kalman-Bucy filtering, and because of the obvious computational advantages of the recursive algorithms, it has found numerous applications in the field of missile guidance, space-vehicle navigation, state estimation for state-vector control, plant identification for adaptive process control and orbit determination, etc.

After the appearance of Bucy's and Kalman's work the study of linear estimation problems was further generalized to the non-white noise problems, that is, colored or no-noise in one or more of the measurements. One of the first, and in many ways one of the most complete treatments of colored noise in continuous systems was presented by Bryson and Johansen [6]. The discrete colored noise problems was examined by Bryson and Henrikson [7]. More recently, Stubberud [8], Stear and Stubberud [9], and Sarachik [10] considered the problems for continuous systems in which each measurement has only colored noise.

Judging by the literature, recursive solutions for the smoothing problem have been thought to be harder to obtain than recursive filtering solutions. Previous methods have been proposed by Bryson and Frazier [11] utilizing the calculus of variations, Rauch, Tung and Striebel [12] adopting the maximum likelihood method, Meditch [13] using orthogonal projection, and most recently, Kailath [14] and Kailath and Frost [15] made use of the innovations.

However, dynamic message models and measurements models for the majority of realistic control problems are inherently nonlinear. Most of the work in nonlinear filtering is very theoretical, some are no more than a philosophy of approach rather than a procedure leading to the derivation of practical estimations.

One of the main lines of attack to the nonlinear estimation problem is the probability approach pioneered by Stratonovich [16], and subsequently taken up by Kushner [17, 18], Wonham [19], Bucy [20], and Mortensen [21]. The truly optimal nonlinear filters for systems corrupted with additive white noise, were given in Kushner [17, 22], however, their exact solutions required infinite dimensional systems which are impossible to realize except in trivially simple cases.

The problems of obtaining a good approximation to the exact solutions of the nonlinear estimation problems are, therefore very important. Extensive work has been carried out to approximate the optimal nonlinear filters by suboptimal finite dimensional filters; amongst them, Cox [23] used the dynamic programming method, Detchmendy and Sridhar [24] used the technique of invariant imbedding, Friedland and Bernstein [25] used the variational techniques, whereas Bass et al. [26] assumed that third and higher order central moments are negligible. As a result, the conditional means and covariance equations are approximated to either first or second-order finite dimensional filters.

The majority of the suboptimal finite dimensional filters introduced in literature involves the Taylor's series expansion of the nonlinear functions, but not all nonlinear functions can be expanded in Taylor's series, moreover, for systems with highly pronounced nonlinearities, Kushner [27] pointed out that the first order filter cannot estimate the states of the Van der Pol oscillator. Schwartz and Stear [28] concluded on the basis of some simulation results, that the added computational complexity may not improve the performance of the nonlinear filter.

Recently, Sunahara [29] proposed to replace the nonlinear functions by quasi-linear functions via stochastic linearization, but the resulting filters are yet to be tested and compared to all other approximate filters.

Sage and Ewing [30, 31] and Sage [32] utilized the methods of maximum likelihood and invariant imbedding to derive approximate filtering as well as smoothing algorithms for nonlinear continuous and discrete systems. The continuous-time algorithms follow from the discrete-time algorithms via the Kalman formal limiting argument [4].

Other works on nonlinear smoothing were presented by Lainiotis [33], Lee [34] and Leondes et al. [35]. In particular, Leondes et al. derived the exact functional differential equations for the smoothing density functions and the smoothed estimates; but their solutions are prohibitive except in trivially simple cases, and hence, approximations are developed for sequential nonlinear smoothing. An iterative technique is also suggested for cases where the nonlinearities may be severe.

However, the above nonlinear estimation problems are only concerned with nonlinear systems corrupted by additive white noise. To the author's knowledge optimal nonlinear filtering as well as smoothing for non-white noise processes have not been derived.

As far as state estimation problems for systems with time delays are concerned, Kwakernaak [36] used the method of orthogonal projection to derive filtering equations for linear continuous systems with multiple time delays, which, when solved, also yield smoothing estimates. Priemer and Vacroux [37, 38] later considered the estimation problems in linear discrete systems containing multiple delays in the message model. Farooq and Mahalanabis [39] rederived the estimation algorithms of Priemer and Vacroux using the state augmentation technique, however, it was pointed out in [37] that the augmentation of state vectors has the effect of increasing the dimension of the systems, and thus lead to a filter that is computationally inefficient. Recently, Biswas and Mahalanabis [40] presented smoothing results for the continuous systems with time delays by first discretizing the continuous systems and then employed the state augmentation technique.

To the author's knowledge nonlinear filtering and smoothing algorithms for nonlinear time-delayed systems corrupted by additive white noise as well as non-white noise processes have not been derived in literature.

1.2 *SCOPE OF THE THESIS*

In this thesis, exact and approximate, discrete and continuous state estimation algorithms are derived for nonlinear systems, with and without delays in the message models and measurement models, involving white Gaussian noise, correlated noise and colored noise processes.

In Chapter II, some basic concepts of state estimation theory and mathematical preliminaries essential to the work of this thesis are presented.

Chapter III deals with continuous-time nonlinear systems without delays corrupted with additive white noise as well as non-white noise processes. In Section 3.2, the noise processes are assumed to be Gaussian white, the basic approach makes use of the matrix minimum principle [44] together with the Kolmogorov [2] and Kushner [17, 18] equations to minimize the error variance, taken to be the estimation criterion. However, the exact algorithms derived in such a manner require infinite dimensional systems to realize, which is computationally impossible. In order that the estimation algorithms can be physically realized, it is assumed that the conditional probability density functions of the estimator errors are Gaussian. An example is included to show how one can evaluate the proposed algorithms under the above assumption.

For the purpose of assessing the performance of the proposed minimum variance filter and to compare it with various other approximate finite dimensional filters, in Section 3.3 three different types of nonlinear systems are selected and simulated on the digital computer. Discussions are included to explain the apparent discrepancies amongst the proposed minimum variance filter and all other approximate finite dimensional filters.

In practical situations, nonlinear dynamic systems are imbedded in non-white noise processes. Therefore, in Sections 3.4 and

3.5, we turn our attention to more general nonlinear estimation problems, they are respectively, correlated noise and noise free processes , and the colored noise problem can be considered as a special case of the noise free estimation problem.

In Chapter IV, discrete-time filtering and smoothing algorithms are derived for nonlinear time-delayed systems imbedded in white Gaussian noise processes. In Section 4.2, the measurement models do not involve any time delays whereas in Section 4.3, the measurements can be considered as the sum of signals propagated with different delays. The main technique makes use of the matrix minimum principle to derive the optimal values of the coefficients in the estimation algorithms under the requirements that the estimates be unbiased. The resulting algorithms can be recursively evaluated under the assumption that the probability density functions of the estimator errors are Gaussian. Examples are included to illustrate the use of the proposed estimation algorithms, in particular, they provide better insight as to, how one can properly substitute for the discrete-time indices, in order to arrive at the filtering, fixed-interval smoothing, fixed-point smoothing and the fixed-lag smoothing algorithms. The linear estimation problems are considered solved, since they can be treated as a special case of the nonlinear estimation problems.

Chapters V and VI deal with discrete nonlinear time-delayed systems involving, respectively, correlated noise and colored noise processes. The derivation as well as the presentation follow that of Chapter IV. Similar to that of Chapter IV, the linear estimation algorithms for these problems can be obtained from the nonlinear estimation algorithms.

Chapter VII summarizes most of the results presented in Chapters III to VI. Also outlined are some recommendations for further research.

CHAPTER II

BASIC CONCEPTS AND MATHEMATICAL PRELIMINARIES

2.1 *BASIC CONCEPTS*

Expectation: If x_i is a discrete random variable, the expectation of the function $h(x_i)$ is defined by

$$E\{h(x_i)\} = \sum h(x_i)\ f\ (x_i)$$

where E denotes the expectation operator, $f(x_i)$ is the density function associated with the random variable and the x_i are the values of the discrete random variable for which $f(x_i)$ is defined.

In the same way, if x is a continuous random variable, the expectation of the function h(x) is defined by

$$E\{h(x)\} = \int_{-\infty}^{\infty} h(x)\ f(x)\ dx$$

where f(x) is the density function associated with the random variable x.

The expectation operator $E\{\cdot\}$ has a number of rather obvious properties:

1. $E\{h_1(x) + h_2(x)\} = E\{h_1(x)\} + E\{h_2(x)\}$
2. $E\{C\ h(x)\} = C\ E\{h(x)\}$
3. $E\{C\} = C$, where C is a constant.

4. $E\{h(x)\} \geq 0$, if $h(x) \geq 0$ for all x.

The expectation is also called the mean, the average and the first moment.

Variance: The variance is defined for the discrete case and for the density function $f(x_i)$ as

$$V_x = \sum_i [x_i - E(x_i)]^2 f(x_i)$$

and for the continuous case

$$V_x = \int_{-\infty}^{\infty} [x - E(x)]^2 f(x)\, dx$$

The variance is also called the second moment.

Unbiased Estimate: $\hat{x}$ is said to be the unbiased estimate of x, if it satisfies the relation

$$E\{x - \hat{x}\} = 0.$$

Correlated noise: Two random white-noise processes v(t) and w(t) are said to be correlated if they satisfy the relation

$$E\{v(t)\, w^T(s)\} = \Psi_{vw}\, \delta(t-s)$$

for all t and s. Where T is the matrix transpose, Ψ_{vw} is not equal to zero, and δ is the Dirac delta function.

Colored noise: A colored noise process is a time-correlated process, that is, it exhibits correlation between different instants of time. Consider v(t) as a colored noise process, then

$$E\{v(t)\ v^T(\tau)\} \neq 0 \quad \text{for } \tau \neq t.$$

In general, one tends to describe the colored noise by a shaping filter [42] given by the equation

$$\dot{v}(t) = A(t)\ v(t) + w(t)$$

where A(t) and the white noise process w(t) must be chosen so that v(t) has the prescribed statistics.

Linear estimation problems: Assume that the measurement process of a dynamical system is given by

$$z(t) = H(t)\ x(t) + v(t)$$

in the continuous-time case, and

$$z(k) = H(k)\ x(k) + v(k)$$

in the discrete-time case, where t denotes the continuous time with

$t \geq t_0$, t_0 is fixed; and $k = 0,1,\ldots$ is the discrete time index. Here, z and v are m-vector measurement and measurement error vectors, respectively.

Any estimate of $x(k)$, $k = 0, 1, \ldots$, which is based on the sequence of measurements

$$\{z(0), z(1), \ldots, z(j)\},$$

$j \geq 0$ is denoted by $\hat{x}(k/j)$. Similarly, in the continuous-time case, $\hat{x}(t/s)$ is the estimate of $x(t)$ derived from the measurement set $\{z(\tau), t_0 \leq \tau \leq s\}$. These estimates are classified [3] in the following according to the value of k or t relative to that of j or s, respectively:

	Discrete Process	Continuous Process
1. Prediction :	$k > j$	$t > s$
2. Filtering :	$k = j$	$t = s$
3. Smoothing :	$k < j$	$t < s$

In general, the state estimation problem is to obtain an estimate of the state $x(k)$ or $x(t)$, such that it is "best" or optimal according to some estimation error functional.

Smoothing: The smoothed estimates can be divided into three separate classes [43]

1. Fixed-interval smoothed estimate $\hat{x}(k/N)$, $k = 0,1,\ldots, N-1$, N = fixed positive integer.
2. Fixed-point smoothed estimate $\hat{x}(k/j)$, $j = k + 1, k + 2, \ldots$, k = fixed integer.
3. Fixed-lag smoothed estimate $\hat{x}(k/k+N)$, $k = 0,1,\ldots$, N = fixed

positive integer.

Similar classification can be made in the continuous-time case.

2.2 *THE KALMAN FILTER*

Kalman and Bucy [3,4] considered the minimum mean-square filtering problem of linear time-varying systems and driven by a white noise process. The message model and the measurement model are given by

$$\dot{x}(t) = F(t)\, x(t) + G(t)\, w(t)$$

and

$$z(t) = H(t)\, x(t) + v(t)$$

respectively. Where $w(t)$ and $v(t)$ are independent, zero-mean Gaussian white noise processes with positive definite variances $\Psi_w(t)$ and $\Psi_v(t)$, respectively.

Their block diagrams are presented in Figure 2.1.

The optimum linear filter is obtained, in the form of an algorithm suitable for direct computer evaluation

$$\dot{\hat{x}}(t) = F(t)\, \hat{x}(t) + K(t)\, [z(t) - H(t)\, \hat{x}(t)]$$

$$\dot{V}_{\tilde{x}}(t) = F(t)\, V_{\tilde{x}}(t) + V_{\tilde{x}}(t)\, F^T(t) - K(t)\, \Psi_v(t)\, K^T(t) + G(t)\, \Psi_w(t)\, G^T(t)$$

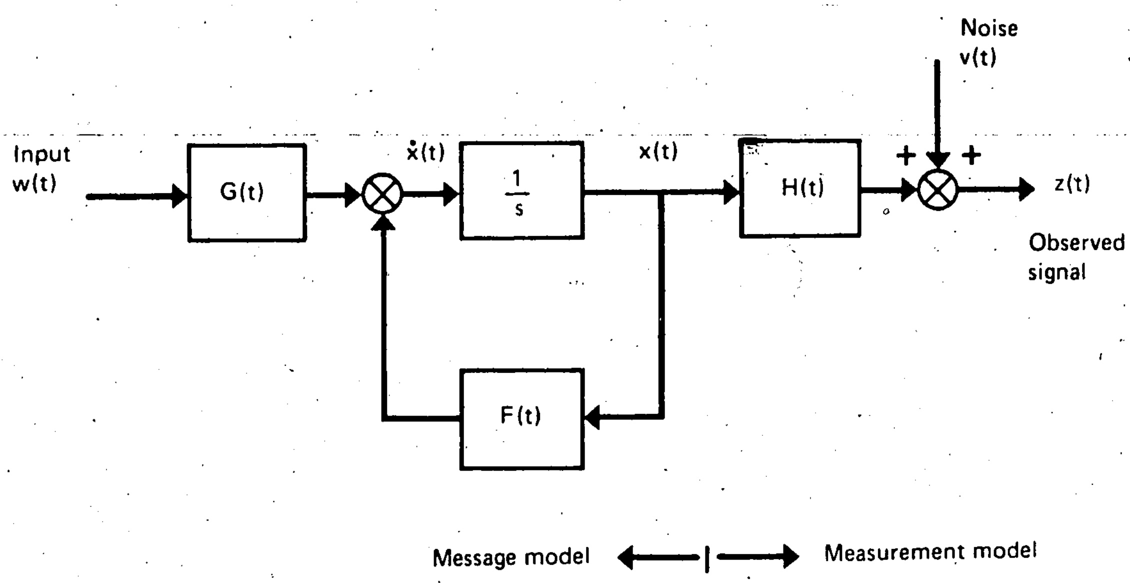

Figure 2.1 Block diagram of the message and measurement models

where $K(t) = V_{\tilde{x}}(t)\, H^T(t)\, \Psi_v^{-1}(t),\ \hat{x}(t_0) = 0$

The block diagram for the above optimum linear filter is presented in Figure 2.2.

Because of its practicality in solving the estimation problem, the Kalman filter is immensely popular in aerospace applications, such as in navigation and guidance control, and also in industrial process control applications.

2.3 *GRADIENT MATRIX*

In this section, the basic concept of a gradient matrix is defined and some identity equations are presented. The basic idea and more complete treatment of this topic can be found in Athans and Schweppe [44].

Let X be an nxn matrix with elements x_{ij}. Let $f(\cdot)$ be a scalar-valued function of the x_{ij}

$$f(X) = f(x_{11}, x_{12}, \ldots, x_{1n}, x_{21}, \ldots, x_{2n}, \ldots).$$

As an example, the trace of the matrix X is a scalar-valued function

$$f(X) = \text{trace}\,[X] = x_{11} + x_{22} + \ldots + x_{nn}\ .$$

The gradient matrix of f(X) with respect to the matrix X is

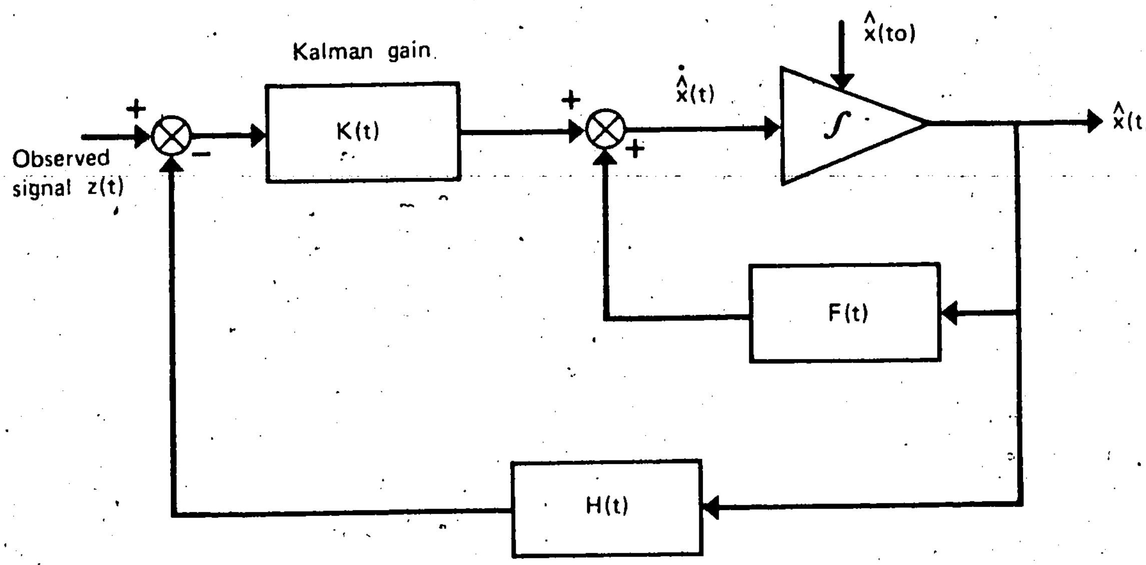

Figure 2.2 Block diagram of the Kalman filter

defined to be the nxn matrix whose ij^{th} element is the following partial derivative

$$\frac{\partial f(X)}{\partial x_{ij}} \quad , \quad i,j = 1,2, \ldots, n.$$

The gradient matrix is denoted by

$$\frac{\partial f(X)}{\partial X}$$

and the gradient of the trace of X satisfies the following relations:

$$\frac{\partial}{\partial X} \text{ trace } [X] = I$$

$$\frac{\partial}{\partial X} \text{ trace } [AX] = A^T$$

$$\frac{\partial}{\partial X} \text{ trace } [AX^T] = A$$

$$\frac{\partial}{\partial X} \text{ trace } [AXB] = A^T B^T$$

$$\frac{\partial}{\partial X} \text{ trace } [AX^T B] = BA$$

$$\frac{\partial}{\partial X} \text{ trace } [AXBX] = A^T X^T B^T + B^T X^T A^T$$

$$\frac{\partial}{\partial X} \text{ trace } [AXBX^T] = A^T XB^T + AXB$$

Additional relations can be obtained through the use of the following trace identities

$$\text{trace } [AB] = \text{trace } [BA]$$
$$\text{trace } [AB^T] = \text{trace } [BA^T].$$

2.4 *MATRIX MINIMUM PRINCIPLE*

2.4.1 *Continuous-Time Matrix Minimum Principle*

Consider a system with "state matrix" $X(t)$, "Control matrix" $U(t)$, whose elements are respectively, x_{ij} and $u_{\alpha\beta}$. Assume that $U(t) \in \Omega$, described by the matrix differential equation

$$\dot{X}(t) = F[X(t), U(t), t] ;$$

and the cost functional is defined as

$$J = K[X(T), T] - K[X(t_o), t_o] + \int_{t_o}^{T} L[X(t), U(t), t]\, dt ;$$

t_o and T are fixed

where $L[\cdot]$ and $K[\cdot]$ are scalar-valued functions of their argument satisfying the usual differentiability conditions.

Let $P(t)$ denote the costate matrix. Define the scalar Hamiltonian function H by

$$H[X(t), P(t), t, U(t)] = L[X(t), U(t), t] + \text{trace}\,\{F[X(t), U(t), t]\, P^T(t)\}$$

If * is used to indicate optimal quantities, then there exists a costate matrix $P^*(t)$ such that the following necessary conditions are satisfied:

(1) First Necessary Conditions

$$\dot{X}^*(t) = \frac{\partial H}{\partial P(t)}\Big|_* = F[X^*(t), U^*(t), t]$$

$$\dot{P}^*(t) = -\frac{\partial H}{\partial X(t)}\Big|_*$$

$$\frac{\partial H}{\partial U(t)}\Big|_* = 0$$

(2) Boundary Conditions

At the initial time $X^*(t_o) = \dfrac{\partial K[X^*(t_o), t_o]}{\partial X^*(t_o)}$

At the terminal time

$$P^*(T) = \frac{\partial}{\partial X^*(T)} K[X^*(T), T]$$

2.4.2 *Discrete-Time Matrix Minimum Principle*

Consider the discrete system defined by the equation

$$X(k+1) = F[X(k), U(k), k], \quad k = 0,1, \ldots, N-1$$

with scalar cost functional defined by

$$J = K[X(N), N] - K[X(0), 0] + \sum_{k=0}^{N-1} L[X(k), U(k), k]$$

Then the Hamiltonian function can be defined by the relation

$$H = L[X(k), U(k), k] + \text{trace}\,\{F[X(k), U(k), k]\, P^T(k+1)\}$$

where P(k) is the costate matrix.

It is assumed that F[·], K[·] and L[·] satisfy the conditions required by the discrete minimum principle [45].

Then the discrete matrix minimum principle states that there exists a costate matrix P*(k), k = 0,1, ..., N, such that the following equations are satisfied

(1) First Necessary Conditions

$$X^*(k+1) = \frac{\partial H}{\partial P(k+1)} \Big|_*$$

$$P^*(k+1) = \frac{\partial H}{\partial X(k)} \Big|_*$$

$$\frac{\partial H}{\partial U^*(k)} = 0$$

(2) Boundary Conditions

When k=0, $P^*(0) = \frac{\partial}{\partial X^*(0)} K[X^*(0),0]$

When k=N, $P^*(N) = \frac{\partial}{\partial X^*(N)} K[X(N), N]$

CHAPTER III

EXACT AND APPROXIMATE MINIMUM VARIANCE FILTERING FOR NONLINEAR CONTINUOUS SYSTEMS

3.1 *INTRODUCTION*

This chapter is devoted to the estimation problems of nonlinear continuous systems without delays corrupted by white Gaussian noise as well as non-white noise processes.

Research in the area of nonlinear filtering has been mushrooming during the past few years as evidenced by the large number of publications and reports on this topic. In general, there are two distinct methods of approach; one such approach extends the Kalman-Bucy filters to the nonlinear dynamics systems using the so-called first-order, quasilinear filter, or extended-Kalman filter [11, 23, 25, etc.]. Another method is developed by Kushner [22], which is based upon the determination of the exact equations satisfied by the conditional probability density functions and conditional expectations. This approach uses the stochastic Ito [46] calculus, the results obtained are optimal, however, they cannot be realized by finite dimensional systems. In order that the solutions be physically realizable, the exact equations are approximated to derive suboptimal finite dimensional filters [17, 18, 20, 22, 26, 27, etc.]. However, in spite of the abundance of papers on optimal and suboptimal nonlinear filters corrupted by white noise processes, to the author's knowledge optimal nonlinear filtering algorithms for non-white noise processes have not been derived in literature.

Therefore, this chapter presents a modified approach to solve nonlinear state estimation problems involving (1) additive white noise, (2) correlated noise, and (3) noise-free processes. The basic approach makes use of the matrix minimum principle to minimize the error variance cost function, which is obtained from the exact conditional probability density functions, as presented in Kushner and Kolmogorov's equations. Therefore, it is not unexpected that the exact nonlinear filtering equations obtained in this thesis for white noise problems closely resemble those derived by Bass et al. [26]. The only difference is that the argument of the expectations have been transformed from x(t) into its estimator error $\tilde{x}(t)$. As a consequence, for filtering problems with polynomial,product types or piecewise continuous nonlinearities, the proposed algorithms can be evaluated without any other approximations under the assumption that the conditional probability density functions of the estimator errors are Gaussian.

Here, the estimation algorithms derived for non-white noise processes are suboptimal, since the change in probability density function due to the differential measurements δz is not known, and is neglected in our derivation. To the author's knowledge the filtering algorithms presented for non-white noise processes are new and are reported in [47, 48]. In the particular case that the dynamical systems are linear, the resulting algorithms are optimal and agree well with those of the literature [6, 49].

Since there is a lack of published papers that give clear comparisons between the performance of various approximate nonlinear filters, and in order to test the performance of the proposed minimum variance filters, extensive simulation results accompanied with discussions are presented to compare the performance characteristics of these filters.

3.2 *OPTIMAL MINIMUM VARIANCE CONTINUOUS NONLINEAR FILTERING WITH WHITE NOISE PROCESSES*

Consider the class of nonlinear systems described by the stochastic differential equation [46]

$$\frac{dx(t)}{dt} = f[x(t), t] + G[x(t), t]\, w(t) \tag{3.1}$$

with the measurement given by

$$z(t) = h[x(t), t] + v(t) \tag{3.2}$$

Where x(t) and z(t) are the n-dimensional state and m-dimensional measurement vectors, f and h are, respectively, n and m-dimensional nonlinear vector valued functions, and G is a vector valued matrix.

The random vectors w(t) and v(t) are, statistically independent zero-mean white Gaussian noise processes such that for all $t, \tau \geq t_0$

$$\mathrm{Cov}\{w(t),\, w(\tau)\} = \Psi_w(t)\ \delta(t-\tau)$$

$$\mathrm{Cov}\{v(t),\, v(\tau)\} = \Psi_v(t)\ \delta(t-\tau) \qquad (3.3)$$

and $$\mathrm{Cov}\{w(t),\, v(\tau)\} = 0$$

Where the variances $\Psi_w(t)$ and $\Psi_v(t)$ are non-negative definite and positive definite, respectively.

The initial state vector $x(t_o) = x_o$ is a zero-mean Gaussian random process, independent of $w(t)$ and $v(t)$ for $t \geq t_o$, with a positive definite variance matrix

$$\mathrm{Var}\{x(t_o),\, x(t_o)\} = V_x(t_o)$$

In the typical filtering problem, it is required to compute $\hat{x}(t)$, the unbiased estimate of $x(t)$ conditioned on the set of measurements

$$Z(t) = [z(s)/t_o \leq s \leq t],$$

such that the cost function for $\tau > t_o$

$$J(\tau) = E\{[x(\tau) - \hat{x}(\tau)]^T M(\tau)\ [x(\tau) - \hat{x}(\tau)]\} \qquad (3.4)$$

is minimized. Here $M(\tau)$ is an arbitrary symmetric positive definite matrix.

The filtering algorithm is assumed to satisfy the general nonlinear differential equation

$$\dot{\hat{x}}(t) = \ell[\hat{x}(t), t] + K(t)\, z(t) \tag{3.5}$$

where $\ell[\hat{x}(t), t]$ and $K(t)$ are yet unknown. Then the estimation problem is to determine the time varying nonlinear vector valued function $\ell[\hat{x}(t), t]$ and the gain algorithm $K(t)$ such that the cost function of Equation (3.4) is minimized.

In fact, the nonlinear filtering equation may be assumed to take various forms, however, once enough information concerning $\hat{x}(t)$ and $z(t)$ are included in the estimator model, the resulting filtering algorithms would be unique. For example, other dynamic filtering equations such as

$$\dot{\hat{x}}(t) = \ell[\hat{x}(t), t] + K(t)\, \{z(t) - E[z(t)]\} \tag{3.6}$$

can also be assumed. Comparing the structures of Equations (3.5) and (3.6), it is obvious that an extra term $-K(t)\, E[z(t)]$ is included in Equation (3.6), however, the resulting filtering algorithm using either one of the above estimator models, would result in exactly the same nonlinear filtering algorithm.

Now, let $\tilde{x}(t)$ denote the estimator error defined by

$$\tilde{x}(t) = x(t) - \hat{x}(t) \tag{3.7}$$

using Equations (3.1), (3.2) and (3.5), the derivatives of Equation (3.7) becomes

$$\dot{\tilde{x}}(t) = f[\tilde{x}(t) + \hat{x}(t), t] + G[\tilde{x}(t) + \hat{x}(t), t]\, w(t)$$

$$- \ell[\hat{x}(t), t] - K(t)\, \{h[\tilde{x}(t) + \hat{x}(t), t] + v(t)\} \quad (3.8)$$

Since $\hat{x}(t)$ is required to be an unbiased estimate, it therefore requires the expectations of both $\tilde{x}(t)$ and $\dot{\tilde{x}}(t)$ be zero. Hence, if the expectations of both sides of Equation (3.8) are taken, it is necessary that

$$\ell[\hat{x}(t), t] = \hat{f}[\tilde{x}(t) + \hat{x}(t), t] - K(t)\, \{\hat{h}[\tilde{x}(t) + \hat{x}(t), t]\} \quad (3.9)$$

where

$$\hat{f}[\tilde{x}(t) + \hat{x}(t), t] = E\{f[\tilde{x}(t) + \hat{x}(t), t]/Z(t)\}$$

$$\hat{h}[\tilde{x}(t) + \hat{x}(t), t] = E\{h[\tilde{x}(t) + \hat{x}(t), t]/Z(t)\}$$

Then the estimator error can be shown to satisfy the relation

$$\dot{\tilde{x}}(t) = f^*[\tilde{x}(t), \hat{x}(t), t] + G^*[\tilde{x}(t), \hat{x}(t), t]\, w^*(t) \quad (3.10)$$

where

$$f^*[\tilde{x}(t), \hat{x}(t), t] = f[\tilde{x}(t) + \hat{x}(t), t] - \hat{f}[\tilde{x}(t) + \hat{x}(t), t]$$

$$+ K(t)\ \{\hat{h}[\tilde{x}(t) + \hat{x}(t), t] - h[\tilde{x}(t) + \hat{x}(t), t]\}$$

and

$$G^*[\tilde{x}(t), \hat{x}(t), t] = [G[\tilde{x}(t) + \hat{x}(t), t], - K(t)] \begin{bmatrix} w(t) \\ v(t) \end{bmatrix} ,$$

Furthermore, Equation (3.2) can be rewritten as

$$z(t) = h[\tilde{x}(t) + \hat{x}(t), t] + v(t) \qquad (3.11)$$

now the filtering problem of Equations (3.1) and (3.2) has been transformed into that of Equations (3.10) and (3.11).

Let $\phi[\tilde{x}(t)]$ be a twice continuously differentiable function of the vector $\tilde{x}(t)$; by definition of the conditional expectation operator

$$d\ E\{\phi[\tilde{x}(t)]\} = \int \phi[\tilde{x}(t)]\ dp[\tilde{x}(t), t/Z(t)]\ d\ \tilde{x}(t) \qquad (3.12)$$

Where $p[\tilde{x}(t), t/Z(t)]$ is the conditional probability density function.

Next, the change in $p[\tilde{x}(t), t/Z(t)]$ due to the dynamic equations of (3.10) and (3.11) must be computed. It can be shown that [46]

$$\delta p = p[\tilde{x}(t + \delta t), t + \delta t/Z(t), \delta z] - p[\tilde{x}(t), t/Z(t)]$$

$$= p[\tilde{x}(t + \delta t), t + \delta t/Z(t), \delta z] - p[\tilde{x}(t), t/Z(t), \delta z]$$

$$+ p[\tilde{x}(t), t/Z(t), \delta z] - p[\tilde{x}(t), t/Z(t)] \quad (3.13)$$

where the first two terms are simply the change due to the dynamic equation of (3.10) and the last two terms are due to the differential measurements δz. These two changes are given by Kolmogorov and Kushner's equation [20, 22]. Therefore,

$$\frac{dE\{\phi[\tilde{x}(t)]\}}{dt} = E\{\mathcal{L}\ \phi[\tilde{x}(t)]\} + E\{\phi[\tilde{x}(t)]\{h[\tilde{x}(t) + \hat{x}(t), t]$$

$$- \hat{h}[\tilde{x}(t) + \hat{x}(t), t]\}\} \ \Psi_v(t)^{-1}\{z(t) - \hat{h}[x(t)$$

$$+ \hat{x}(t), t]\} \quad (3.14)$$

where

$$\mathcal{L}\phi[\tilde{x}(t)] = \sum_{i=1}^{n} f_i^*[\tilde{x}(t), x(t), t] \frac{\partial \phi[\tilde{x}(t)]}{\partial \tilde{x}_i}$$

$$+ \frac{1}{2} \sum_{i,j=1}^{n} \{G^*[\tilde{x}(t), \hat{x}(t), t] \ \Psi_{w^*}(t) \ G^{*T}[\tilde{x}(t), \hat{x}(t), t]\}_{ij}$$

$$\cdot \frac{\partial^2 \phi[\tilde{x}(t)]}{\partial \tilde{x}_i \ \partial \tilde{x}_j} \quad (3.15)$$

Then the error-variance equation is obtained from Equations

(3.10), (3.11), (3.14) and (3.15) by setting $\phi[\tilde{x}(t)] = \tilde{x}(t)\,\tilde{x}^T(t)$. For which

$$\frac{dV_{\tilde{x}}(t)}{dt} = E\{\tilde{x}(t)\, f^{*T}[\tilde{x}(t), \hat{x}(t), t] + f^{*}[\tilde{x}(t), \hat{x}(t), t]\, \tilde{x}^T(t)\}$$

$$+ E\{G[\tilde{x}(t) + \hat{x}(t), t]\, \Psi_w(t)\, G^T[\tilde{x}(t) + \hat{x}(t), t]$$

$$+ K(t)\, \Psi_v(t) \cdot K^T(t)\} + E\{\tilde{x}(t)\, \tilde{x}^T(t)\, h^T[\tilde{x}(t)$$

$$+ \hat{x}(t), t] - V_{\tilde{x}}(t)\, h^T[\tilde{x}(t) + \hat{x}(t), t]\}$$

$$\cdot\, \Psi_v^{-1}(t)\, \{z(t) - \hat{h}[\tilde{x}(t) + \hat{x}(t), t]\} \qquad (3.16)$$

where it can be shown that

$$E\{\tilde{x}(t)\, f^{*T}[\tilde{x}(t), \hat{x}(t), t]\} = E\{\tilde{x}(t)\, f^T[\tilde{x}(t) + \hat{x}(t), t]$$

$$- \tilde{x}(t)\, h^T[\tilde{x}(t) + \hat{x}(t), t]\, K^T(t)\}$$

On the other hand, the initial condition is given by

$$V_{\tilde{x}}(t_o) = E\{[x(t_o) - \hat{x}(t_o)][x(t_o) - \hat{x}(t_o)]\} = V_x(t_o)$$

Now the estimation problem takes the form of an optimal control problem in which K(t) is the only variable available for manipulation such that the cost function of Equation(3.4) is minimized.

However, Equation(3.4)can be expressed in a more convenient form by observing that for any two column vectors x and y, the following identity equation holds

$$x^T y = \text{trace } [y\ x^T]$$

it follows that

$$J(\tau) = E\{\text{trace}[M(\tau)\ \tilde{x}(\tau)\ \tilde{x}^T(\tau)]\}$$

$$= \text{trace } \{M(\tau)\ E[\tilde{x}(\tau)\ \tilde{x}^T(\tau)]\}$$

$$= \text{trace } \{M(\tau)\ V_{\tilde{x}}(\tau)\}$$

To use the matrix minimum principle, to solve this problem, a costate matrix P(t) is defined, and the Hamiltonian is given by

$$H = \text{trace } \{\dot{V}_{\tilde{x}}(t)\ P^T(t)\}$$

Using the concept of gradient matrices [44] the necessary condition which the optimal K(t) must satisfy is obtained as

$$K(t) = E\{\tilde{x}(t)\, h^T[\tilde{x}(t) + \hat{x}(t), t]\}\, \Psi_V^{-1} \tag{3.17}$$

note that K(t) is independent of the costate matrix P(t) and the weighting factor M(t).

Then the filtering estimate and the error-variance equation become

$$\dot{\hat{x}}(t) = \hat{f}[\tilde{x}(t) + \hat{x}(t), t] + E\{\tilde{x}(t)\, h^T[\tilde{x}(t) + \hat{x}(t), t]\}\, \Psi_V^{-1}(t)\, \{z(t) - \hat{h}[\tilde{x}(t) + \hat{x}(t), t]\} \tag{3.18}$$

and

$$\frac{dV_{\tilde{x}}(t)}{dt} = E\{\tilde{x}(t)\, f^T[\tilde{x}(t) + \hat{x}(t), t] + f[\tilde{x}(t) + \hat{x}(t), t]\, \tilde{x}^T(t)\} + E\{G[\tilde{x}(t) + \hat{x}(t), t]\, \Psi_W(t)\, G^T[\tilde{x}(t) + \hat{x}(t), t]\} - E\{\tilde{x}(t)\, h^T[\tilde{x}(t) + \hat{x}(t), t]\}\, \Psi_V^{-1}(t)\, E\{h[\tilde{x}(t) + \hat{x}(t), t]\, \tilde{x}^T(t)\} + E\{\tilde{x}(t)\, \tilde{x}^T(t)\, h^T[\tilde{x}(t) + \hat{x}(t), t] - V_{\tilde{x}}(t)\, h^T[\tilde{x}(t) + \hat{x}(t), t]\}\, \Psi_V^{-1}(t)\, \{z(t) - \hat{h}[\tilde{x}(t) + \hat{x}(t), t]\} \tag{3.19}$$

respectively.

It should be noted that before the above algorithms can be physically realized, a number of difficult expectations must be evaluated. They involve infinite dimensional systems except in trivially simple cases. Also notice that the presented algorithms closely resemble the exact equations due to Bass et al [26]. The only difference is that the argument of the expectations have been transformed into $\tilde{x}(t)$. Such a transformation is particularly significant for filtering problems with polynomial, product types or piecewise continuous nonlinearities. In such cases, the filtering algorithms can be obtained without any further approximations, under the assumption that the probability density functions of the estimator errors are Gaussian.

For example, consider a nonlinear system with message and measurement models described by

$$\dot{x}(t) = -\sin x(t) + u(t)$$

and $$y(t) = x^3(t) + v(t)$$

respectively. Where u(t) and v(t) are zero-mean white Gaussian noise processes, with variances $\Psi_u(t)$ and $\Psi_v(t)$, respectively.

Assuming that x(t) is a Gaussian process, then the expectations of the nonlinear functions can be evaluated as follows

$$E[x^3(t)] = E\{[\tilde{x}(t) + \hat{x}(t)]^3\} = 3V_{\tilde{x}}(t)\,\hat{x}(t) + \hat{x}^3(t)$$

$$E[\sin x(t)] = E\{\sin[\tilde{x}(t) + \hat{x}(t)]\} = \sin \hat{x}(t) \cdot e^{-V_{\tilde{x}}(t)/2}$$

$$E[\tilde{x}(t) \sin x(t)] = - \cos \hat{x}(t)\, V_{\tilde{x}}(t)\, e^{-V_{\tilde{x}}(t)/2}$$

and

$$E[\tilde{x}^2(t)\, x^3(t)] = 9V_{\tilde{x}}^2(t)\, \hat{x}(t) + V_{\tilde{x}}(t)\, \hat{x}^3(t)$$

Hence, Equations (3.18) and (3.19) become

$$\dot{\hat{x}}(t) = -\sin \hat{x}(t) \cdot e^{-V_{\tilde{x}}(t)/2} + [3V_{\tilde{x}}^2(t) + 3V_{\tilde{x}}(t)\, \hat{x}^2(t)]\Psi_v^{-1}[y(t) - 3V_{\tilde{x}}(t)\, \hat{x}(t) - \hat{x}^3(t)]$$

and

$$\dot{V}_{\tilde{x}}(t) = -2V_{\tilde{x}}(t)\, e^{-V_{\tilde{x}}(t)/2} \cos \hat{x}(t) + \Psi_u(t) - [3V_{\tilde{x}}^2(t) + 3V_{\tilde{x}}(t)\, \hat{x}^2(t)]^2\, \Psi_v^{-1}(t) + [6V_{\tilde{x}}^2(t)\, \hat{x}(t)]\, \Psi_v^{-1}(t)[y(t) - 3V_{\tilde{x}}(t)\, \hat{x}(t) - \hat{x}^3(t)]$$

respectively.

Note that in the evaluation of the expectations the only assumption needed is the Gaussian assumption of the estimator error, whereas all other finite dimensional algorithms delete some of the higher order terms, in particular the higher order terms of the error-variance. When the nonlinear functions in Equations (3.18) and (3.19) are approximated by Taylor series expansions, the results obtained can be identified with various other approximate nonlinear filtering algorithms in literature [23, 24, 26, 28].

Furthermore, the above derivation can be extended to filtering problems with non-white noise processes. However, in such cases, the change in $p[\tilde{x}(t), t/Z(t)]$ due to the differential measurements δz is not available in the literature, and to the author's knowledge optimal nonlinear filtering algorithms for such problems have not yet been derived.

As an approximation, the effect due to the differential measurements δz is neglected, and Equation (3.14) becomes simply

$$\frac{dE\{\phi[\tilde{x}(t)]\}}{dt} = E\{ L\ \phi[\tilde{x}(t)]\} \tag{3.20}$$

It is noteworthy to mention that when such an approximation is made, the results obtained for the filtering problem corrupted by white noise processes are equivalent to that of the stochastic linearization due to Sunahara [29]. In such a case the random forcing term in the error-variance equation of Equation (3.19) is neglected, whereas the filtering algorithm of Equation (3.18) remains unchanged.

When Equation (3.20) is used in place of Equation (3.14) to obtain the error-variance equation, and the nonlinear functions are approximated by second order Taylor series expansions, the resulting algorithms are commonly called modified minimum variance filter [28].

3.3 *ILLUSTRATIVE EXAMPLES AND DISCUSSIONS*

In order to test the performance of the proposed minimum variance filter and to compare it with various other approximate nonlinear filters [23-24, 26-28]. Three different types of nonlinear systems are selected, the stochastic equations are transformed to Stratonovich's forms, and then simulated on an IBM 360 computer using CSMP. The simulation results presented only show the estimates for one simulation run . The integration was performed by the rectangular and fixed fourth-order Runga-Kutta method, the step size chosen ranged from 0.001 to 0.00005.

Meanwhile, various structures of dynamic finite dimensional filters and their error-variance equations are, respectively, listed in Tables 3.1 and 3.2, where the scalar case is considered.

It can be observed that amongst the truncated, the quasi-moment and the modified minimum variance filters, the only difference is the random forcing term that appears in the error-variance equations. This is also true in comparing the stochastic linearization filter and the minimum variance filter proposed in Section 3.2.

For Figures (3.1) - (3.10), the system models corresponding to the example of Section 3.2 were simulated, the random forcing term for such systems is given by

Table 3.1 Dynamic Equations of Various Nonlinear Filters

System Dynamics $\dot{x} = f(x) + g(x)\,w$
Measurement Model $z = h(x) + v$

Filter Nomenclature	Filter Dynamics	
Modified Minimum Variance Filter	$\dot{\hat{x}} = f(\hat{x}) + \frac{1}{2} V_{\tilde{x}} f''(\hat{x}) + V_{\tilde{x}} h'(\hat{x}) \Psi_v^{-1}\{z - h(\hat{x}) - \frac{1}{2} V_{\tilde{x}} h''(\hat{x})\}$	(A)
Truncated Minimum Variance Filter	Same as Eq. (A)	
Quasi-Moment Minimum Variance Filter	Same as Eq. (A)	
Stochastic Linearization Filter (Sunahara)	$\dot{\hat{x}} = \hat{f}(x) + E[\tilde{x}\, h(x)]\, \Psi_v^{-1}[z - \hat{h}(x)]$	(B)
Proposed Minimum Variance Filter	Same as Eq. (B)	

The prime "'" denotes the first derivative with respect to the argument.

Table 3.2 Error-Variance Equations of Various Nonlinear Filters

Systems Dynamics $\dot{x} = f(x) + g(x)\, w$
Measurement Model $z = h(x) + v$

Filter Nomenclature	Error-Variance Equations	
Modified Minimum Variance Filter	$\dot{V}_{\tilde{x}} = 2V_{\tilde{x}} f'(\hat{x}) + g^2(\hat{x})\, \Psi_w - V_{\tilde{x}}^2\, h'^2(\hat{x})\, \Psi_v^{-1}$	(C)
Truncated Minimum Variance Filter	$\dot{V}_{\tilde{x}} =$ Right-hand-side of Eq. (C) $- \frac{1}{2} V_{\tilde{x}}^2 h''(\hat{x}) \Psi_v^{-1} \{z - h(\hat{x}) - \frac{1}{2} V_{\tilde{x}}\, h''(\hat{x})\}$	
Quasi-Moment Minimum Variance Filter	$\dot{V}_{\tilde{x}} =$ Right-hand-side of Eq. (C) $+ V_{\tilde{x}}^2 h''(\hat{x}) \Psi_v^{-1} \{z - h(\hat{x}) - \frac{1}{2} V_{\tilde{x}} h''(\hat{x})\}$	
Stochastic Linearization Filter (Sunahara)	$\dot{V}_{\tilde{x}} = 2E[\tilde{x} f(x)] + E[g^2(x)\, \Psi_w] - \{E[\tilde{x}\, h(x)]\}^2\, \Psi_v^{-1}$	(D)
Proposed Minimum Variance Filter	$\dot{V}_{\tilde{x}} =$ Right-hand-side of Eq. (D) $+ E[\tilde{x}^2 h(x) - V_{\tilde{x}} h(x)]\, \Psi_v^{-1} [z - \hat{h}(x)]$	

$$6 V_{\tilde{x}}^{2}(t)\, \hat{x}(t)\, \Psi_v^{-1}(t)\{y(t) - 3 V_{\tilde{x}}(t)\, \hat{x}(t) - \hat{x}^3(t)\}$$

Four sets of prior statistics were assumed, the results are very interesting.

In Figures (3.1) and (3.2), the statistics are given by

$$x(0) = -3.0, \quad \hat{x}(0) = 0.0,$$
$$V_{\tilde{x}}(0) = 2.25, \quad \Psi_u = \Psi_v = 1.0$$

whereas in Figures (3.3) and (3.4)

$$x(0) = 1.5, \quad \hat{x}(0) = 0.0$$
$$V_{\tilde{x}}(0) = 2.25, \quad \Psi_u = \Psi_v = 0.1$$

For these two sets of statistics, theoretically, it can be shown that the contribution due to the random forcing term is not significant at all, therefore, it can be expected that the results from the truncated, the quasi-moment and the modified minimum variance filters are not considerably different from one another, and also the stochastic linearization filter would not deviate from the proposed minimum variance filter in any significant way. Both are verified experimentally and are obvious from Figures (3.1) - (3.4).

However, from Figures (3.1) - (3.4), it is obvious that the performance characteristics of the stochastic linearization and the proposed minimum variance filters are significantly superior to all other

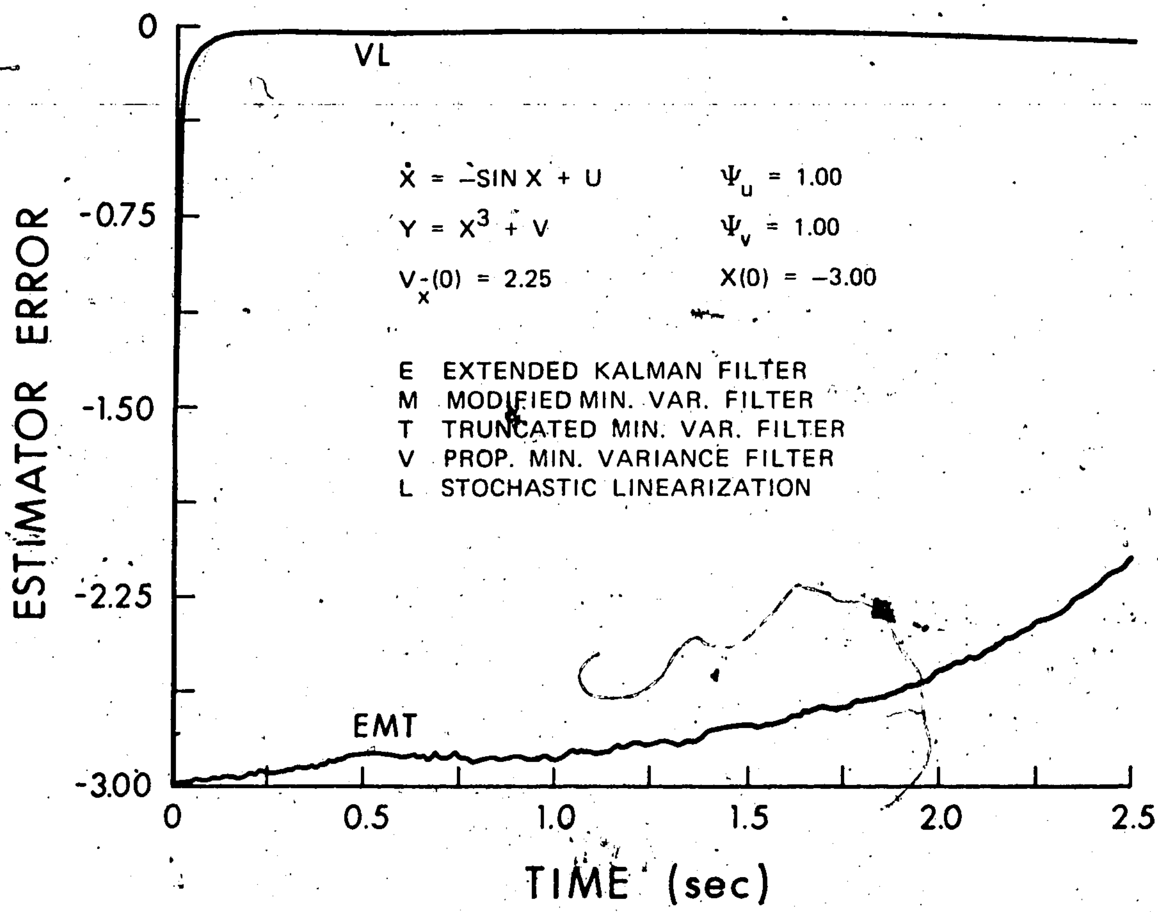

Figure 3.1 The output run of the estimator error for systems with $\hat{x}(0)=0.0$, $x(0)=-3.0$, $\Psi_u=\Psi_v=1.0$ and $V_{\tilde{x}}=2.25$

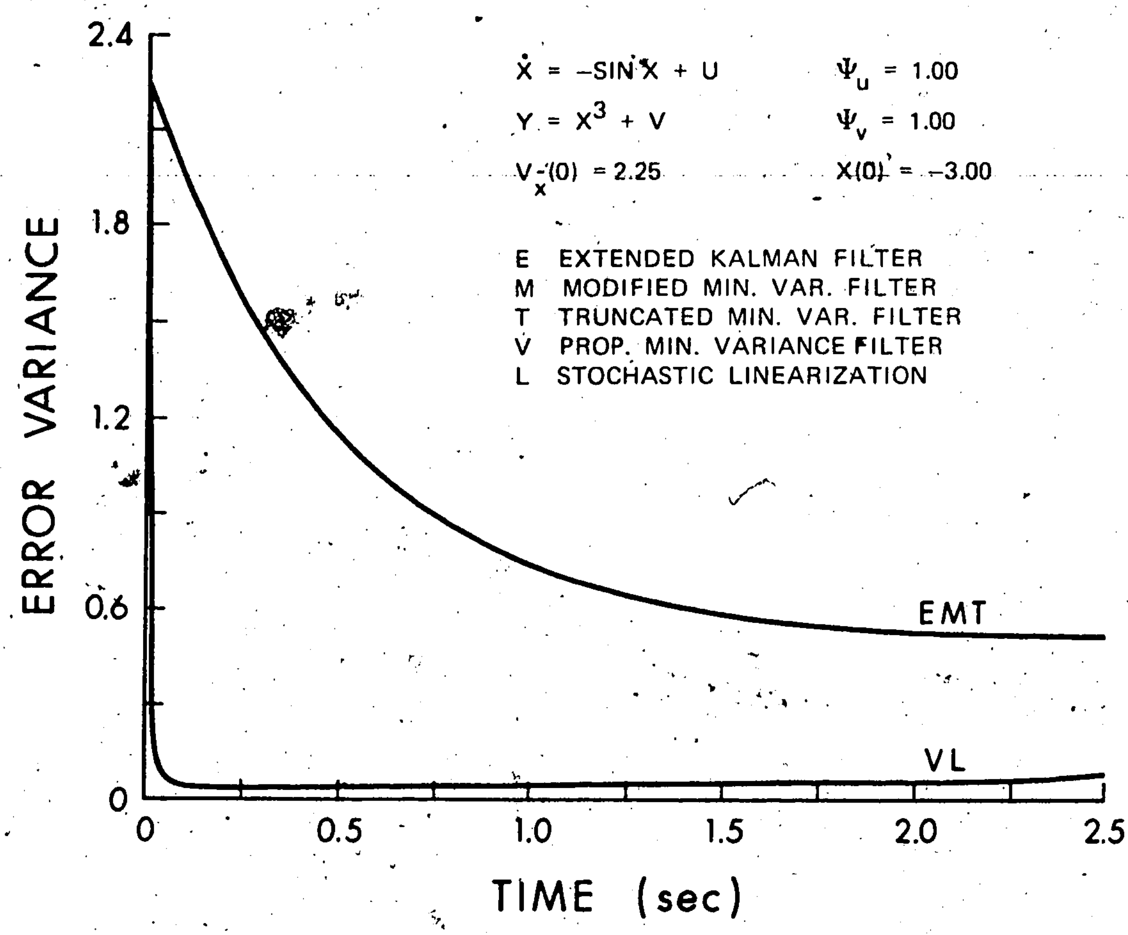

Figure 3.2 The output run of the error-variance for systems with $\hat{x}(0)=0.0$, $x(0)=-3.0$, $\Psi_u=\Psi_v=1.0$, and $V_{\tilde{x}}=2.25$

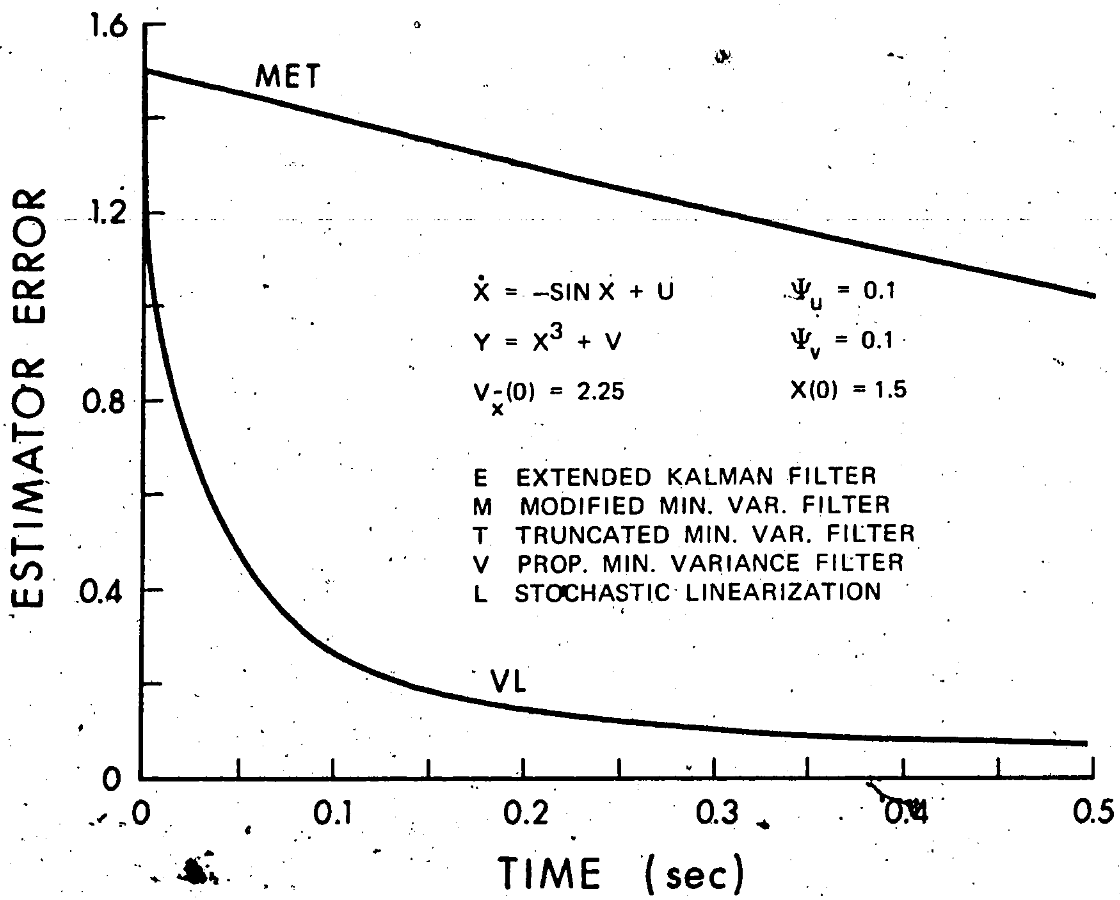

Figure 3.3 The output run of the estimator error for systems with $\hat{x}(0)=0.0$, $x(0)=1.5$, $\Psi_u=\Psi_v=0.1$ and $V_{\tilde{x}}=2.25$

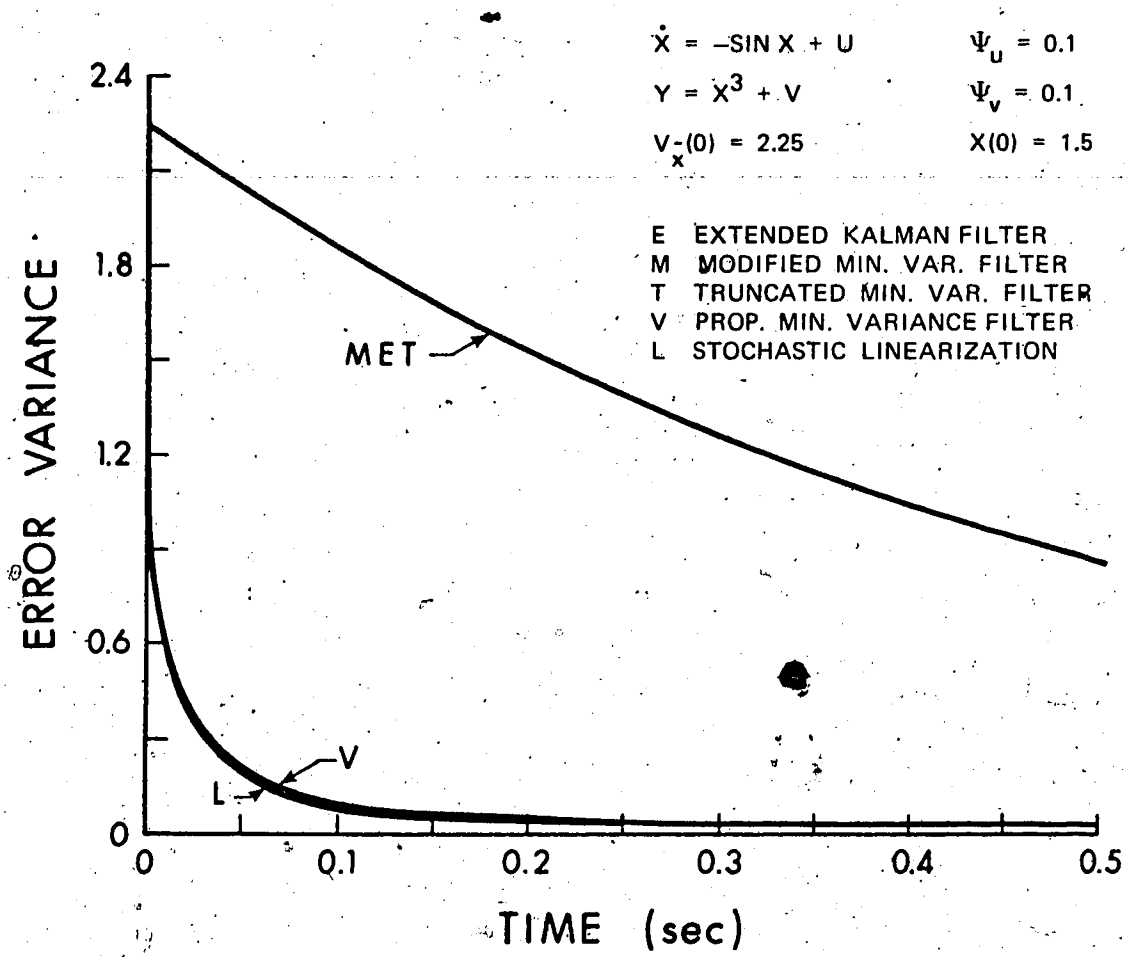

Figure 3.4 The output run of the error-variance for systems with $\hat{x}(0)=0.0$, $x(0)=1.5$, $\Psi_u=\Psi_v=0.1$ and $V_{\tilde{x}}=2.25$

first or second order approximate filters, this is due to the fact that $\{E[\tilde{x}(t)h(x)]\}$ is significantly different from that of $\{V_{\tilde{x}}h'(\hat{x})\}$.

Figures (3.5) and (3.6) presents the estimator errors and their error-variances for various structures of filters with statistics given as

$$x(0) = -3.0, \quad \hat{x}(0) = -1.0$$

$$V_{\tilde{x}}(0) = 0.1, \quad \Psi_u = \Psi_v = 1.0$$

It is interesting to observe that the results obtained from the proposed minimum variance filter comes very close to that of the quasi-moment minimum variance filter, whereas, the stochastic linearization filter comes very close to those of the modified filter and the extended Kalman filter. Moreover, the error-variances of the last three filters are considerably lower than those of the first two filters, while the estimator errors for the last three filters are considerably greater than those of the first two filters. The fact that the last three filters underestimate their error-variances is quite undesirable, in some cases, it may produce detrimental effects, since the filters think they are performing very well, but in reality they are not.

Using this particular set of statistics, it can be shown that the random forcing term has very significant effect, whereas, the difference between $\{E[\tilde{X}(t) h(x)]\}$ and $\{V_{\tilde{x}} h'(\hat{x})\}$ is very insignificant. The conclusion drawn from these theoretically considerations agree well

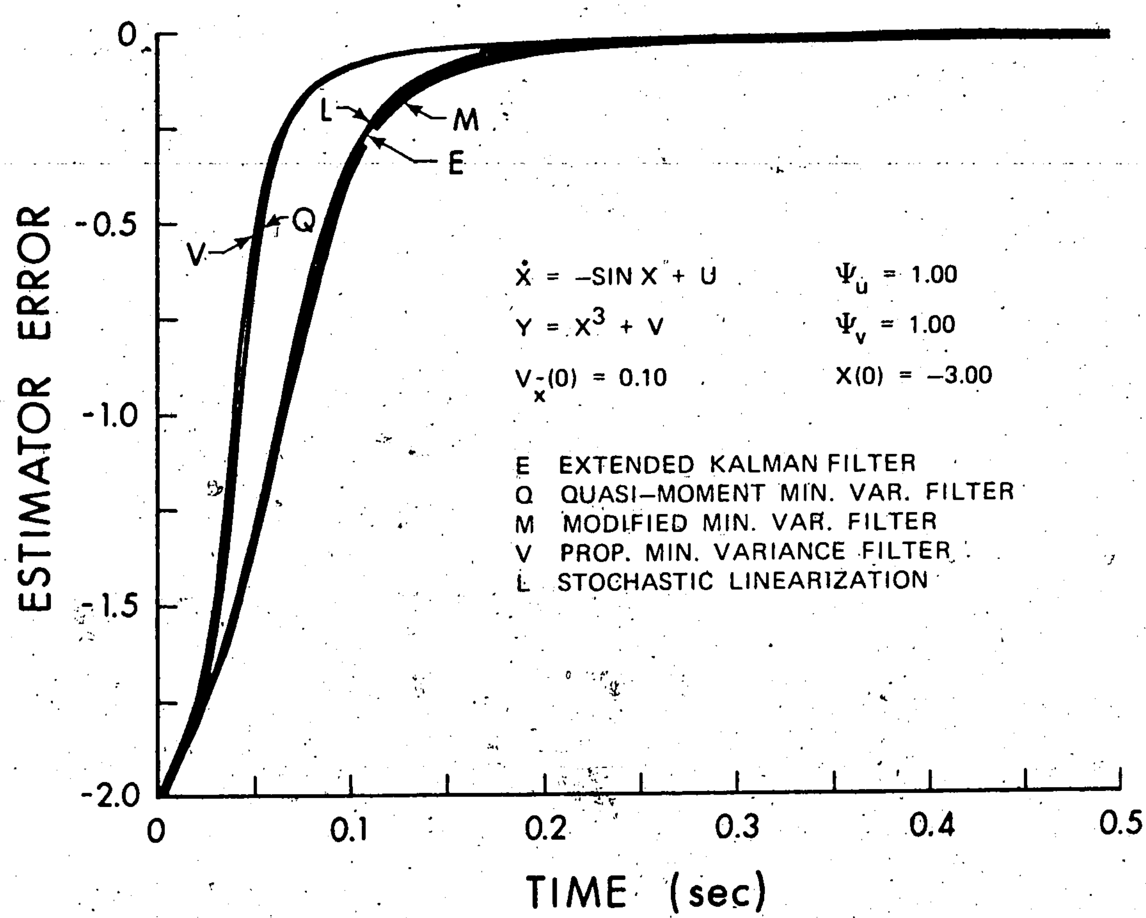

Figure 3.5 The output run of the estimator error for systems with $\hat{x}(0)=-1.0$, $x(0)=-3.0$, $\Psi_u=\Psi_v=1.0$ and $V_{\tilde{x}}=0.1$

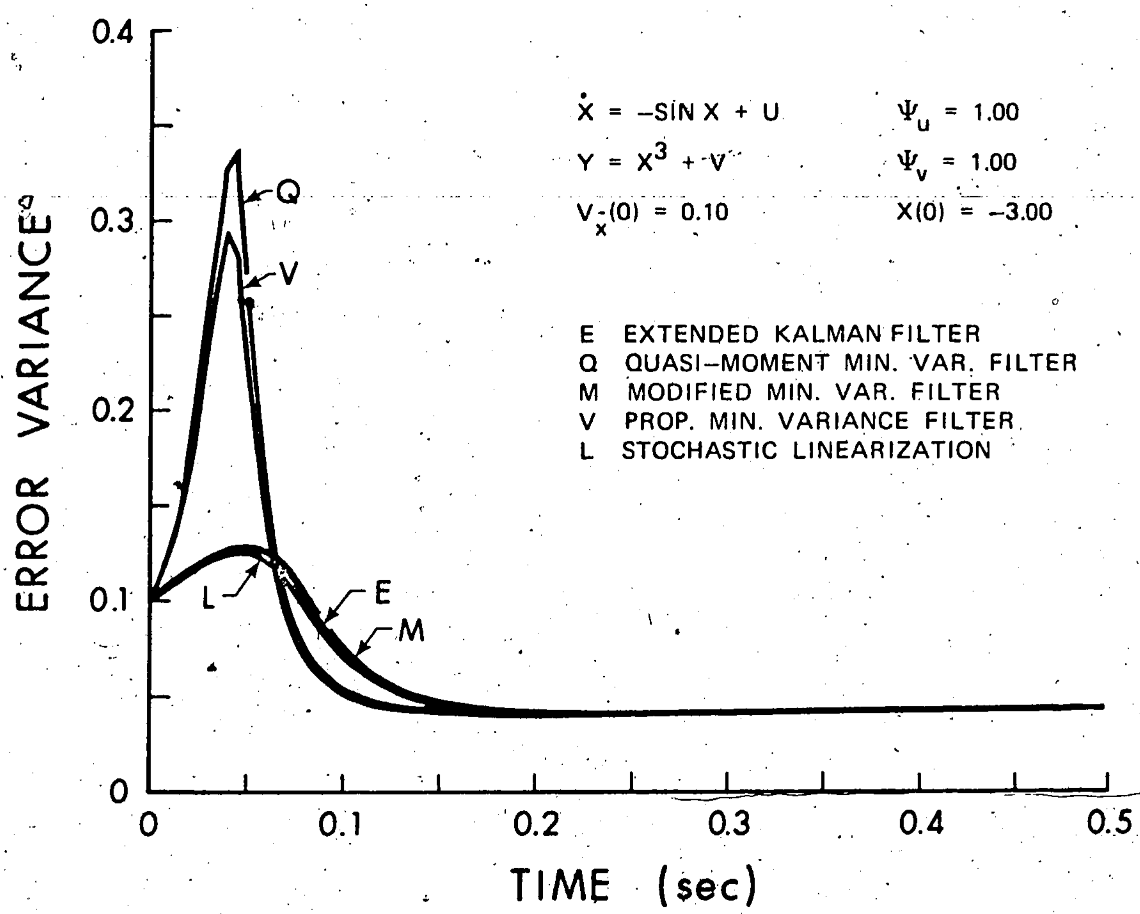

Figure 3.6 The output run of the error-variance for systems with $\hat{x}(0)=-1.0$, $x(0)=-3.0$, $\Psi_u=\Psi_v=1.0$ and $V_{\tilde{x}}=0.1$

with the preceeding experimental observations.

In Figures (3.7) and (3.8), the prior statistics are chosen as

$$x(0) = -0.5, \quad \hat{x}(0) = 0.5$$

$$V_{\tilde{x}}(0) = 5.0, \quad \Psi_u = \Psi_v = 1.0$$

The estimator errors obtained from various structures of finite dimensional filters are moderately different from one another. In particular, the first-order extended Kalman filter is almost as good as the second order modified minimum variance filter,and is moderately better than the second-order quasi-moment minimum variance filter.

However, the error-variances appear to cluster into two significantly different results, the results taken from the proposed minimum variance filter, the stochastic linearization filter and the quasi-moment minimum variance filter cluster together, and are significantly smaller than those of the modified minimum filter and the extended Kalman filter.

It is obvious that some of the various second-order filters and the stochastic linearization filter, may have very distinct performance characteristics, in some cases, their performance characteristics may be almost as good as those of the proposed minimum variance filter; but in other cases, they may produce terribly poor estimates, even worse than the first-order extended Kalman filter, worst of all,

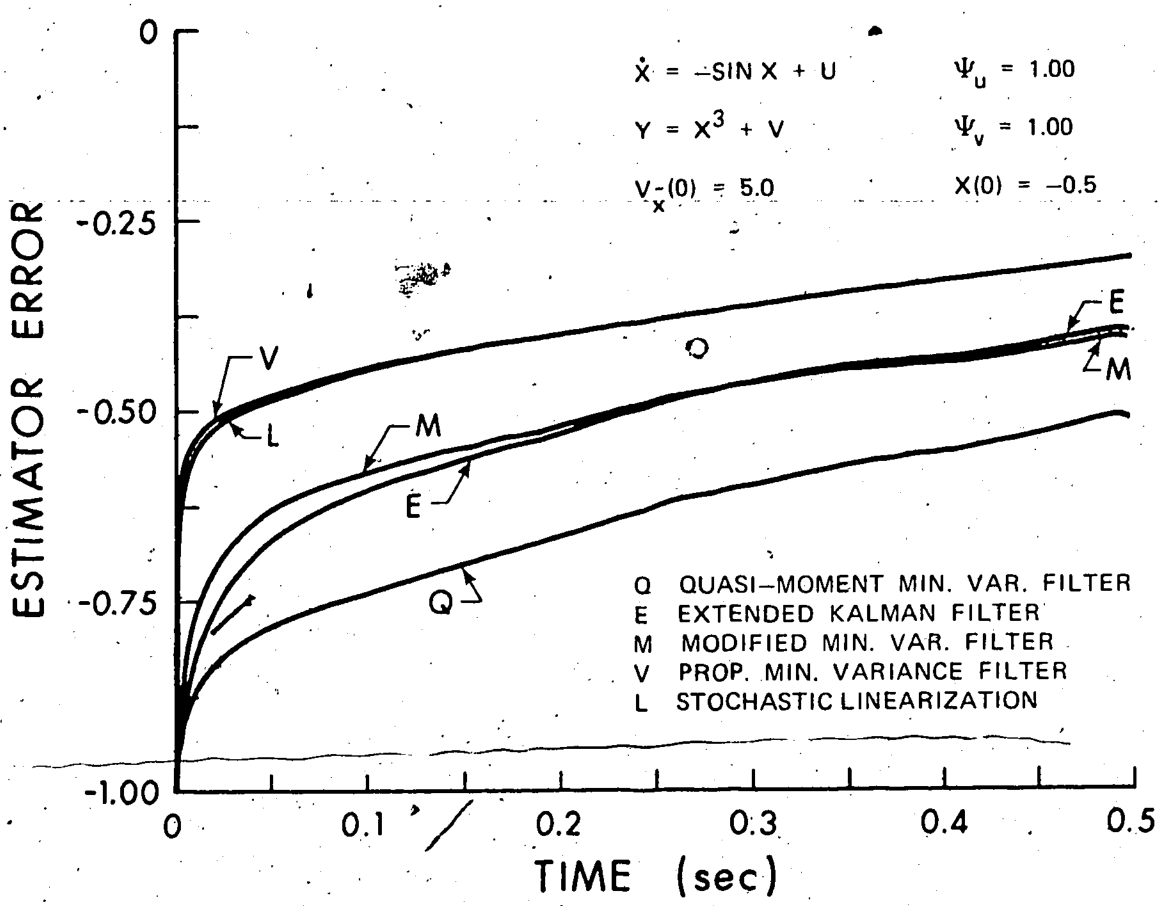

Figure 3.7 The output run of the estimator error for systems with $\hat{x}(0)=0.5$, $x(0)=-0.5$, $\Psi_u=\Psi_v=1.0$ and $V_{\tilde{x}}=5.0$

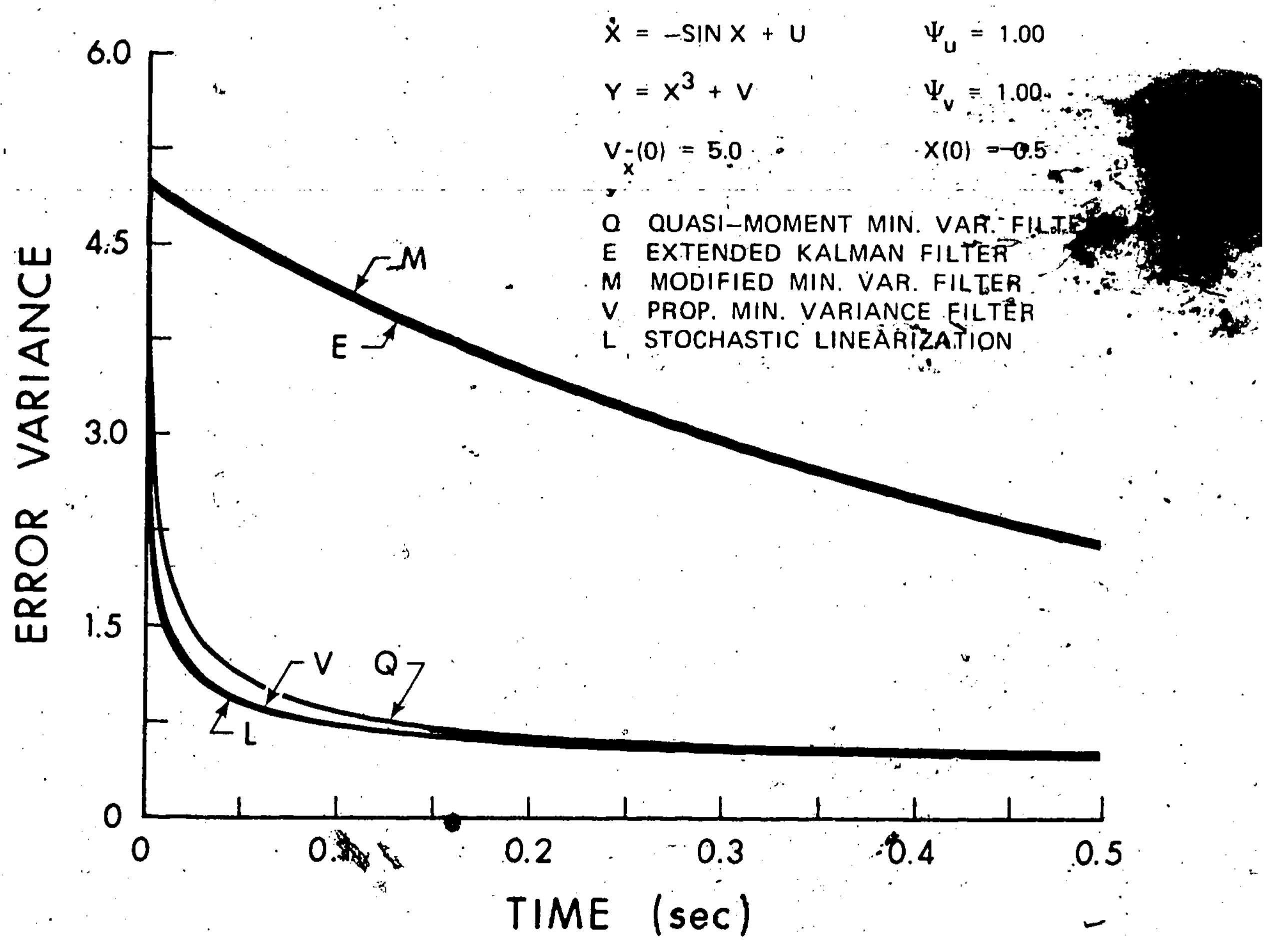

Figure 3.8 The output run of the error-variance for systems with $\hat{x}(0)=0.5$, $x(0)=-0.5$, $\Psi_u=\Psi_v=1.0$ and $V_{\tilde{x}}=5.0$

they may underestimate their error-variances. This makes the second-order filters and the stochastic linearization filter very undesirable and unreliable.

Figures (3.9) - (3.11) are concerned with another system given by

$$\dot{x} = - x + 5 + u$$
$$y = x^3 + v$$

and theoretical discussion and experimental observation similar to those of Figures (3.1) to ([illegible] can be carried out. They indicate that the performance characteristics of the stochastic linearization filter and the proposed minimum variance filter are much superior to those of the first and the second order approximate filters.

Figures (3.12) - (3.23) are concerned with the Van der Pol oscillator with systems described by

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = - x_1 + 3x_2(1 - x_1^2)$$

and

$$y = x_1 - 0.1\, x_1^3 + v.$$

Figures (3.12) and (3.13) are, respectively, the measurements y for $\Psi_v = 5.0$ and 0.5.

From Figures (3.14) - (3.21), it is obvious for both cases of prior statistics, the performance characteristics of the proposed minimum variance filter are much superior to those of the stochastic linearization

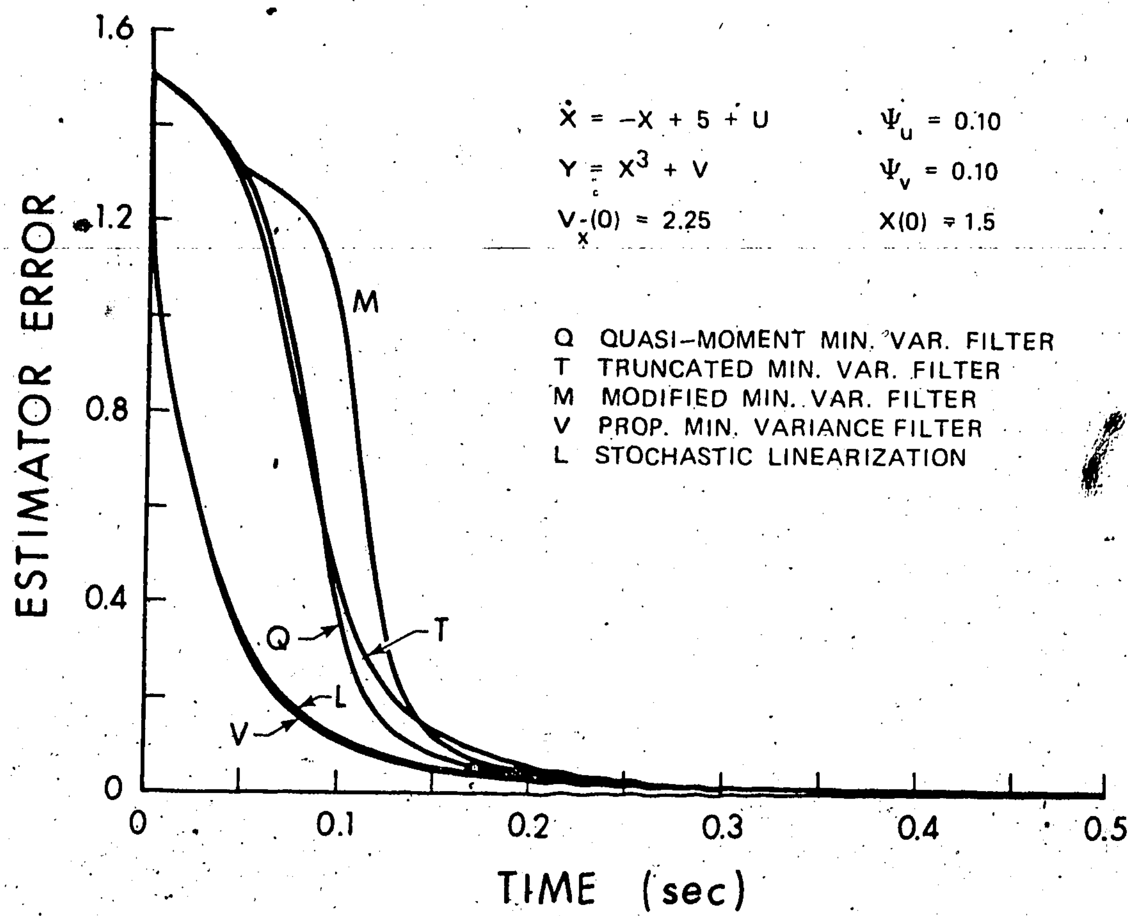

Figure 3.9 The output run of the estimator error for systems with $\hat{x}(0)=0.0$, $x(0)=1.5$, $\Psi_u=\Psi_v=0.1$ and $V_{\tilde{x}}=2.25$

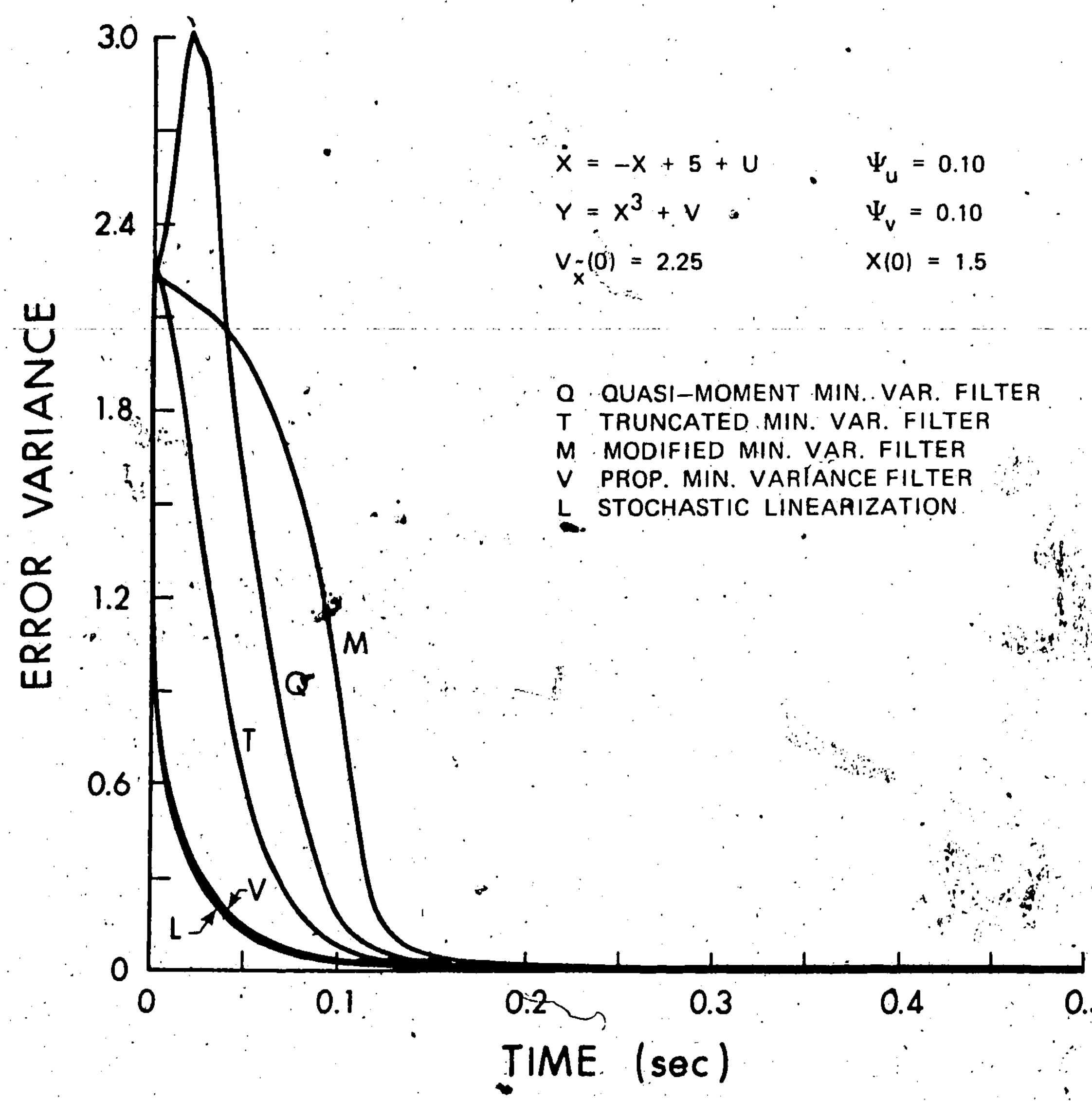

Figure 3.10 The output run of the error-variance for systems with $\hat{x}(0)=0.0$, $x(0)=1.5$, $\Psi_u=\Psi_v=0.1$ and $V_{\tilde{x}}=2.25$

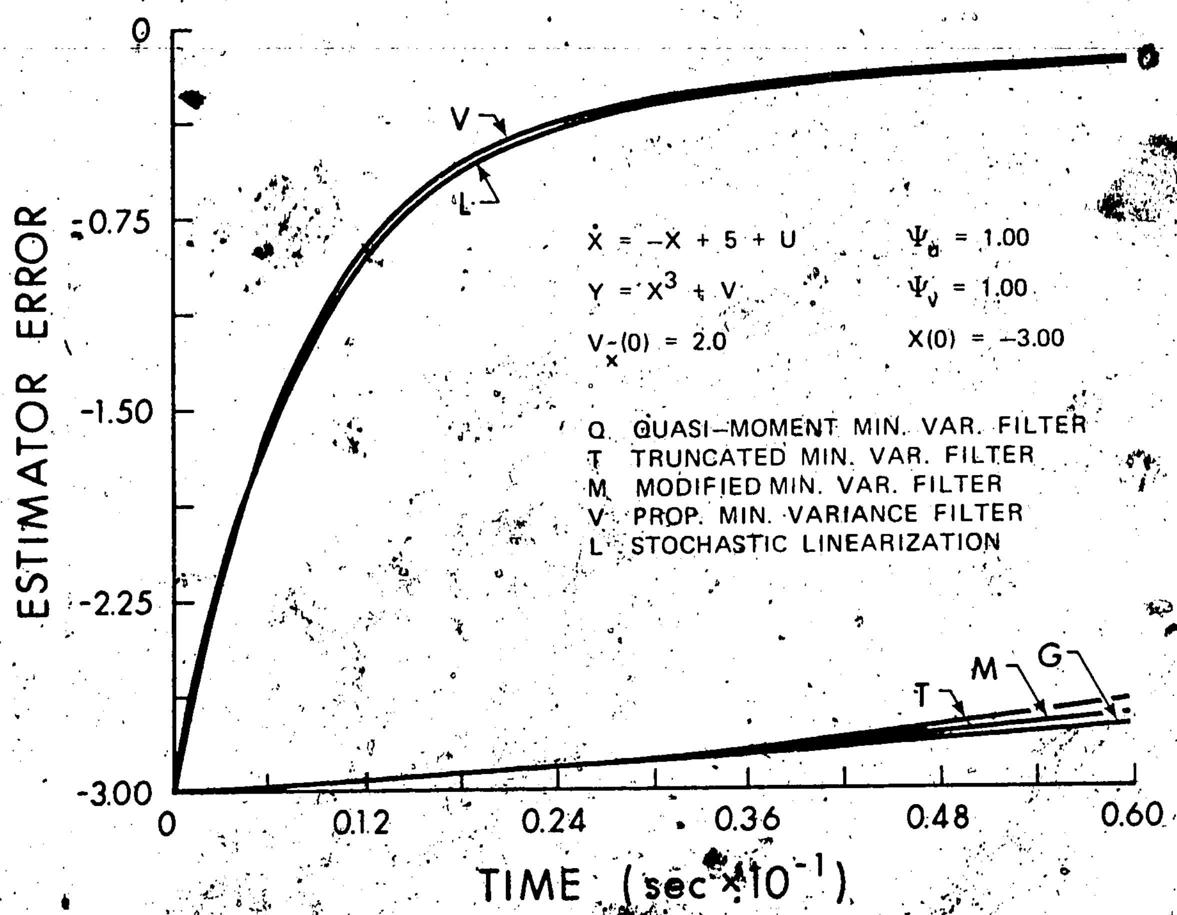

Figure 3.11 The output run of the estimator error for systems with $\hat{x}(0)=0.0$, $x(0)=-3.0$, $\Psi_u=\Psi_v=1.0$ and $V_x=2.0$

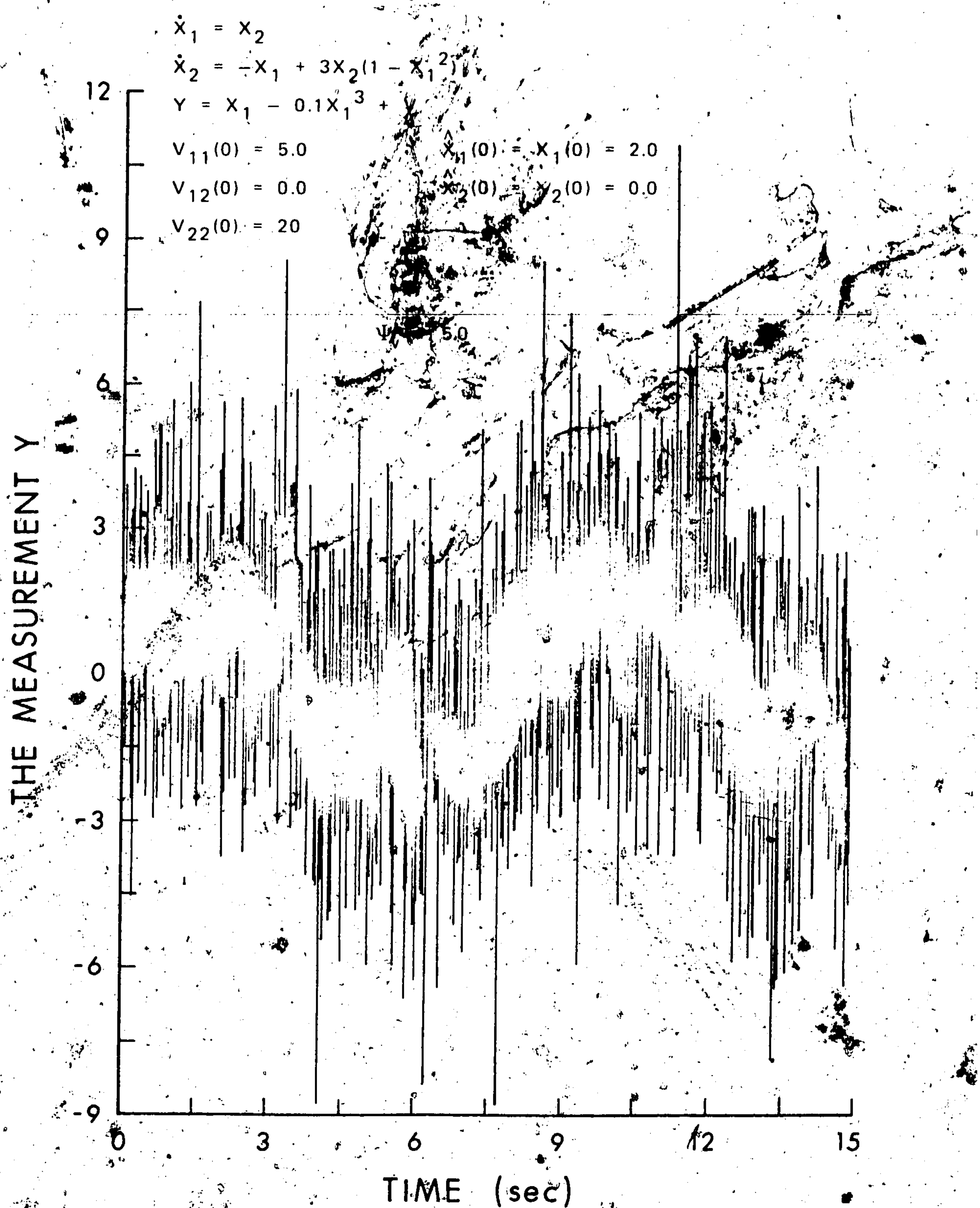

Figure 3.12 The output run of the measurement Y for systems with $\Psi = 5.0$

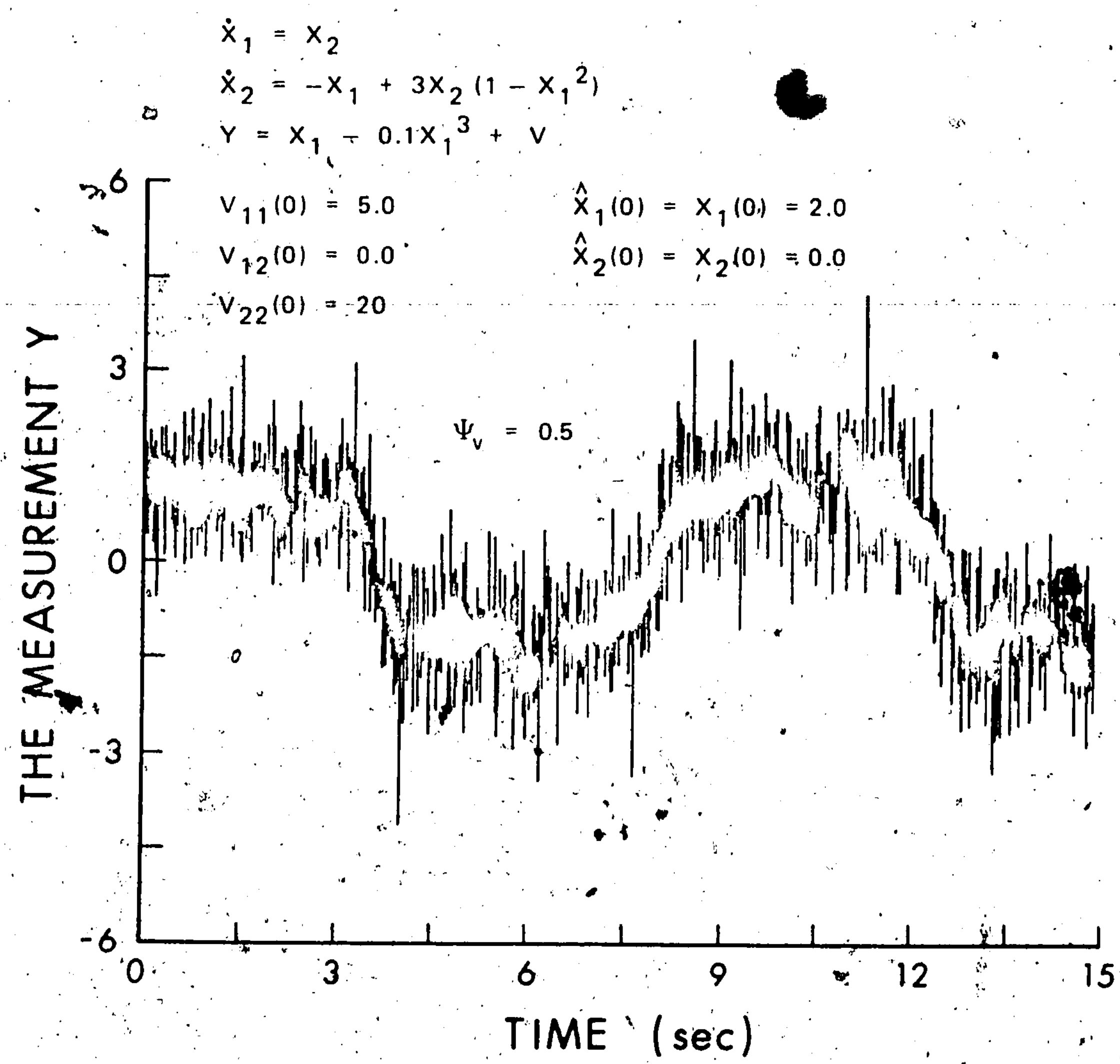

Figure 3.13 The output run of the measurement y for systems with Ψ_V=0.5

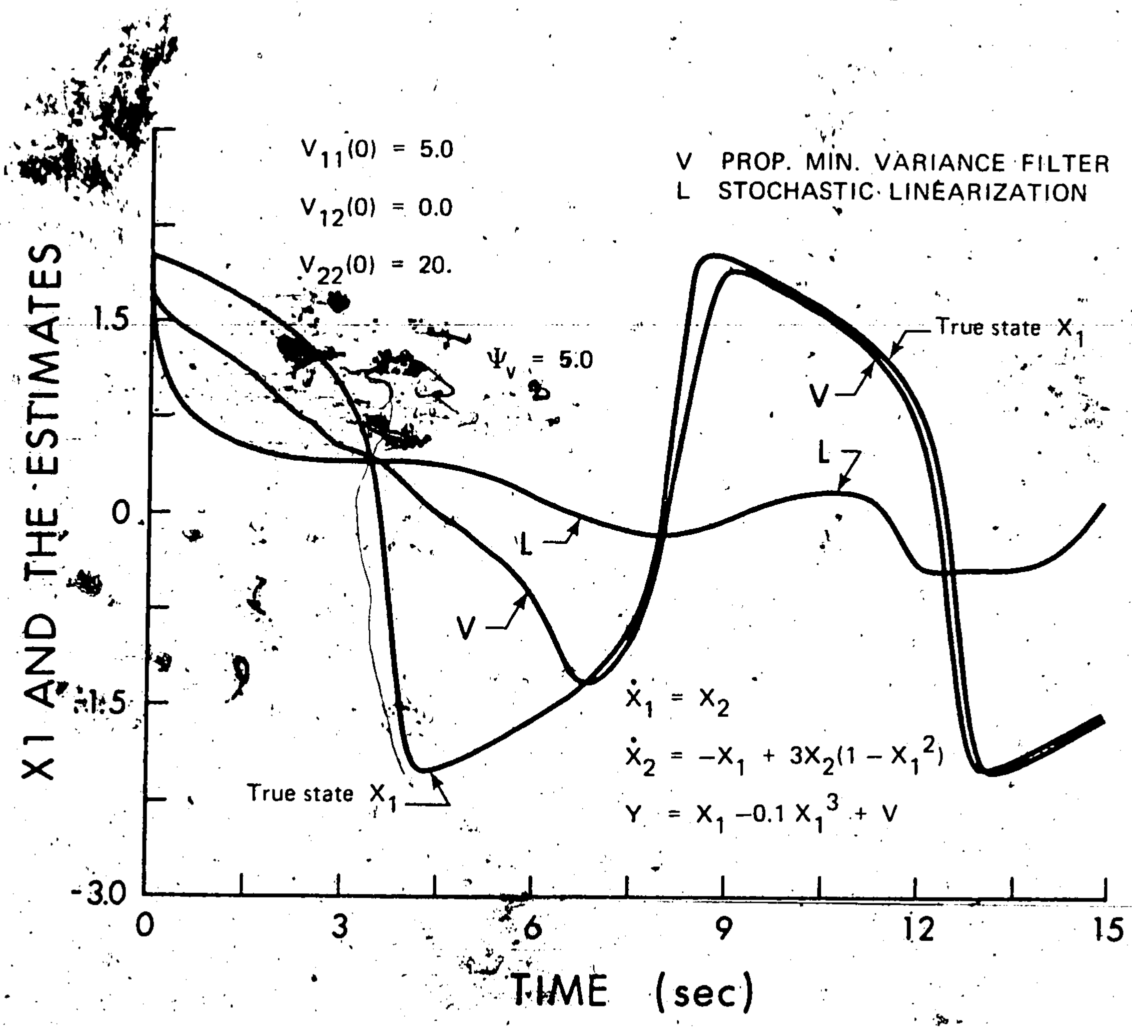

Figure 3.14 The output run of x_1 and the estimates for Ψ_v=5.0, V_{11}=5.0, V_{12}=0.0 and V_{22}=20

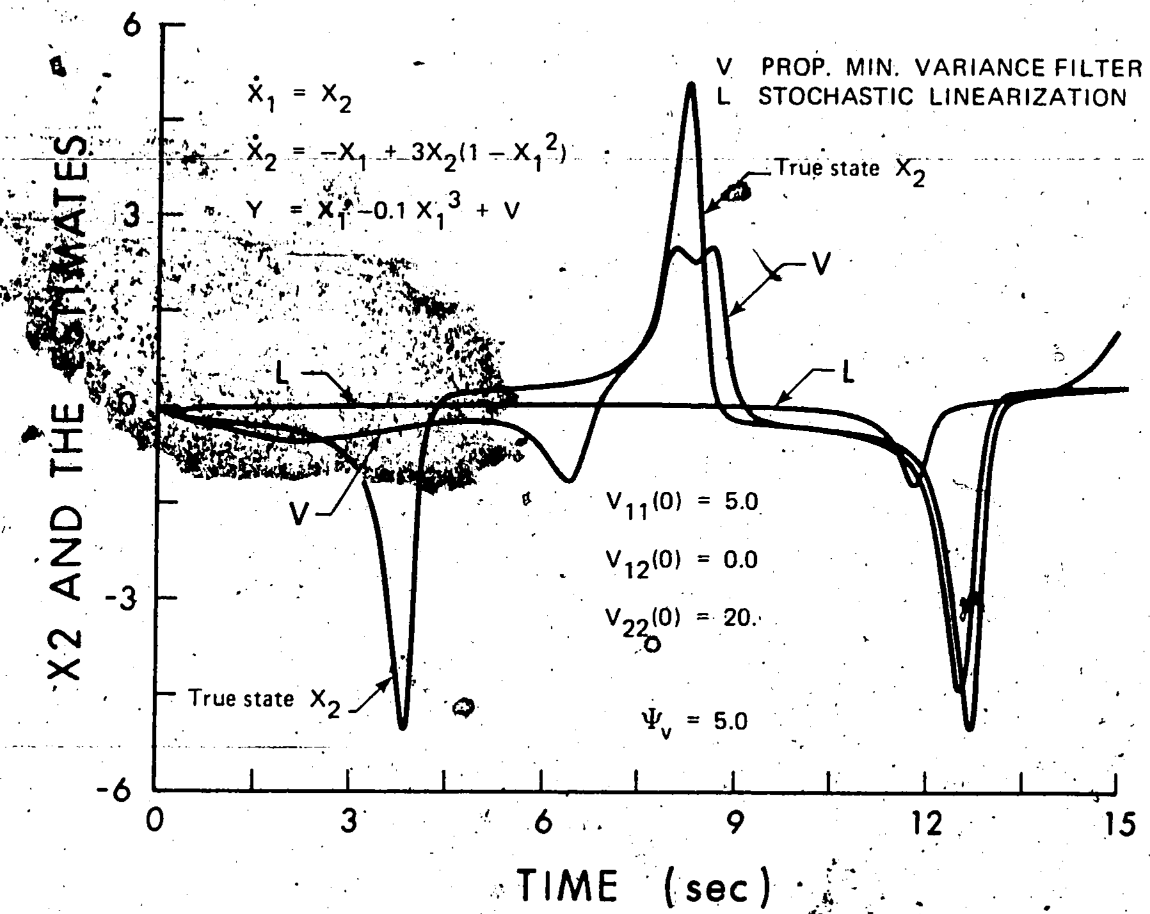

Figure 3.15 The output run of x_2 and the estimates for Ψ_v=5.0, V_{11}=5.0, V_{12}=0.0 and V_{22}=20

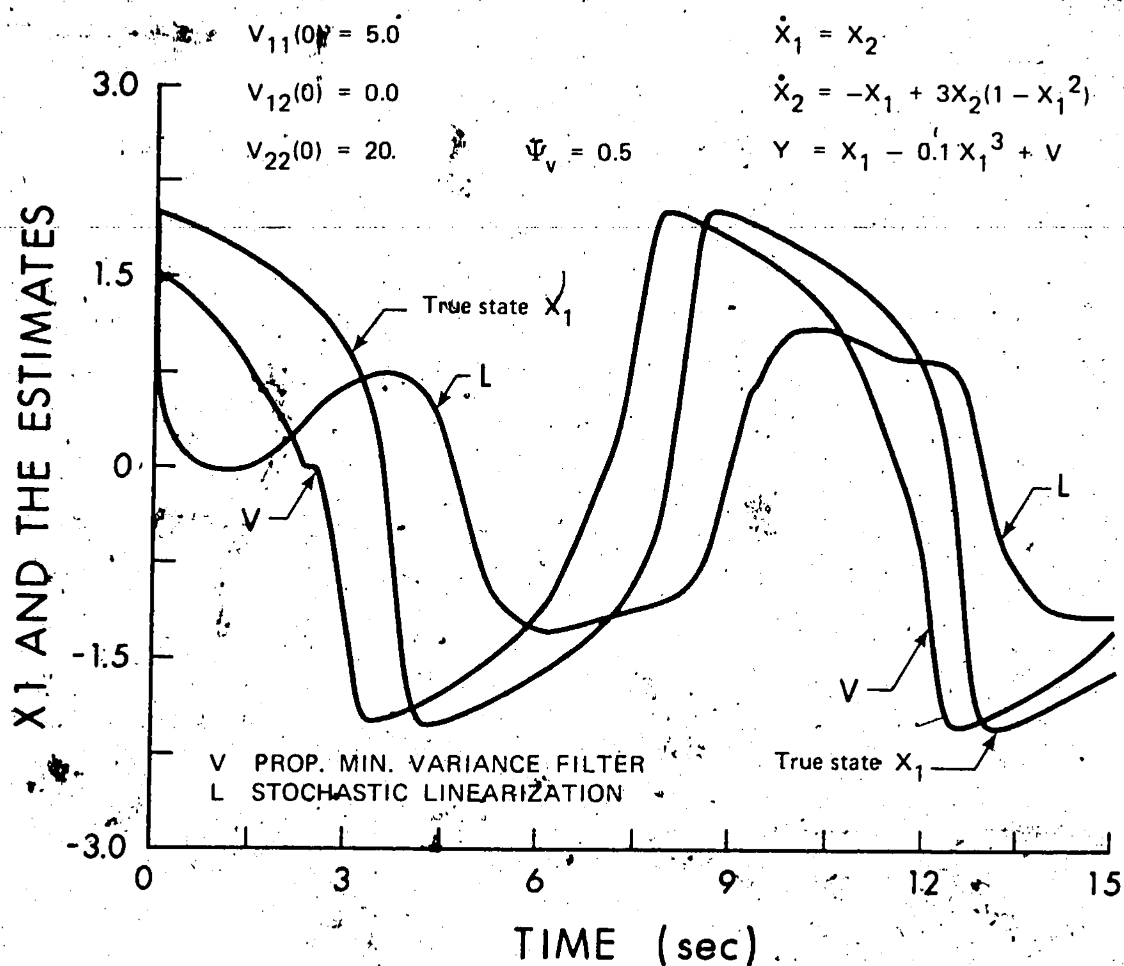

Figure 3.16 The output run of x_1 and the estimates for $\Psi_v=0.5$, $V_{11}=5.0$, $V_{12}=0.0$ and $V_{22}=20$

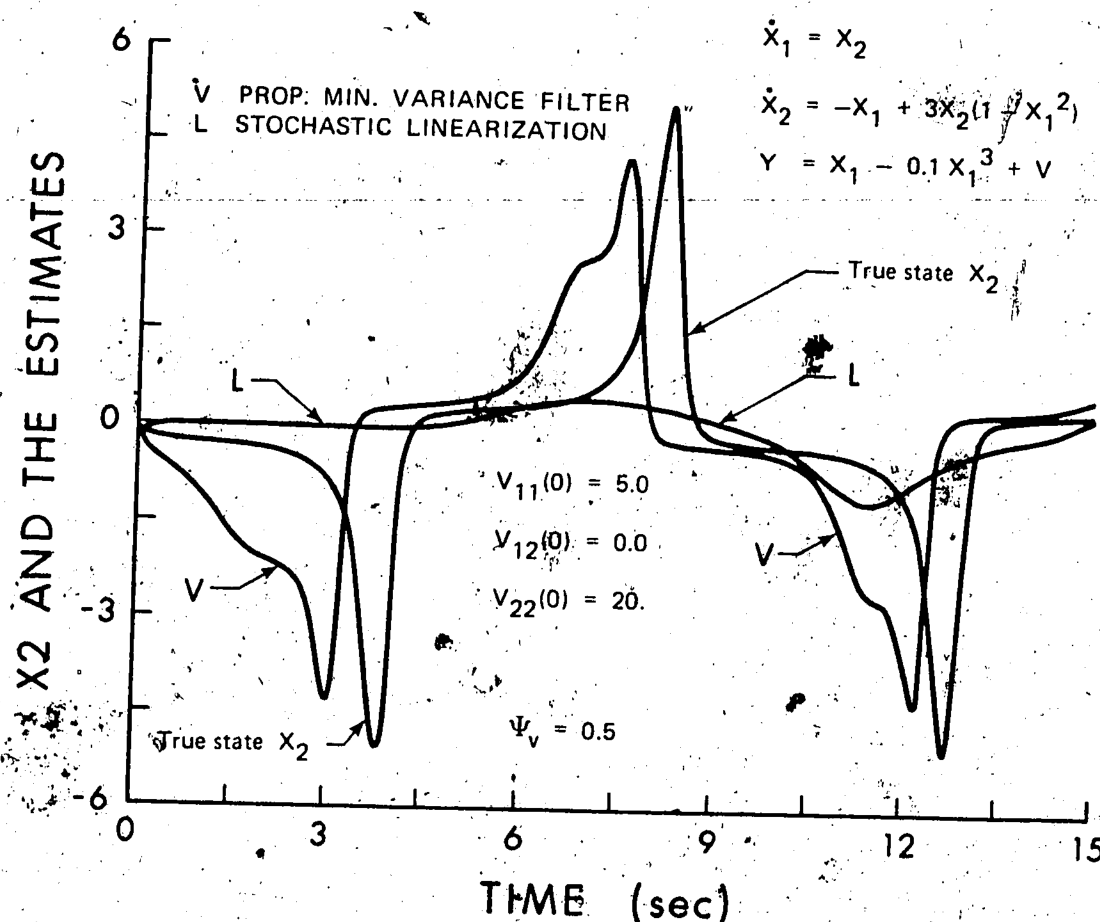

Figure 3.17 The output run of x_2 and the estimates for $\Psi_v=0.5$, $V_{11}=5.0$, $V_{12}=0.0$ and $V_{22}=20$

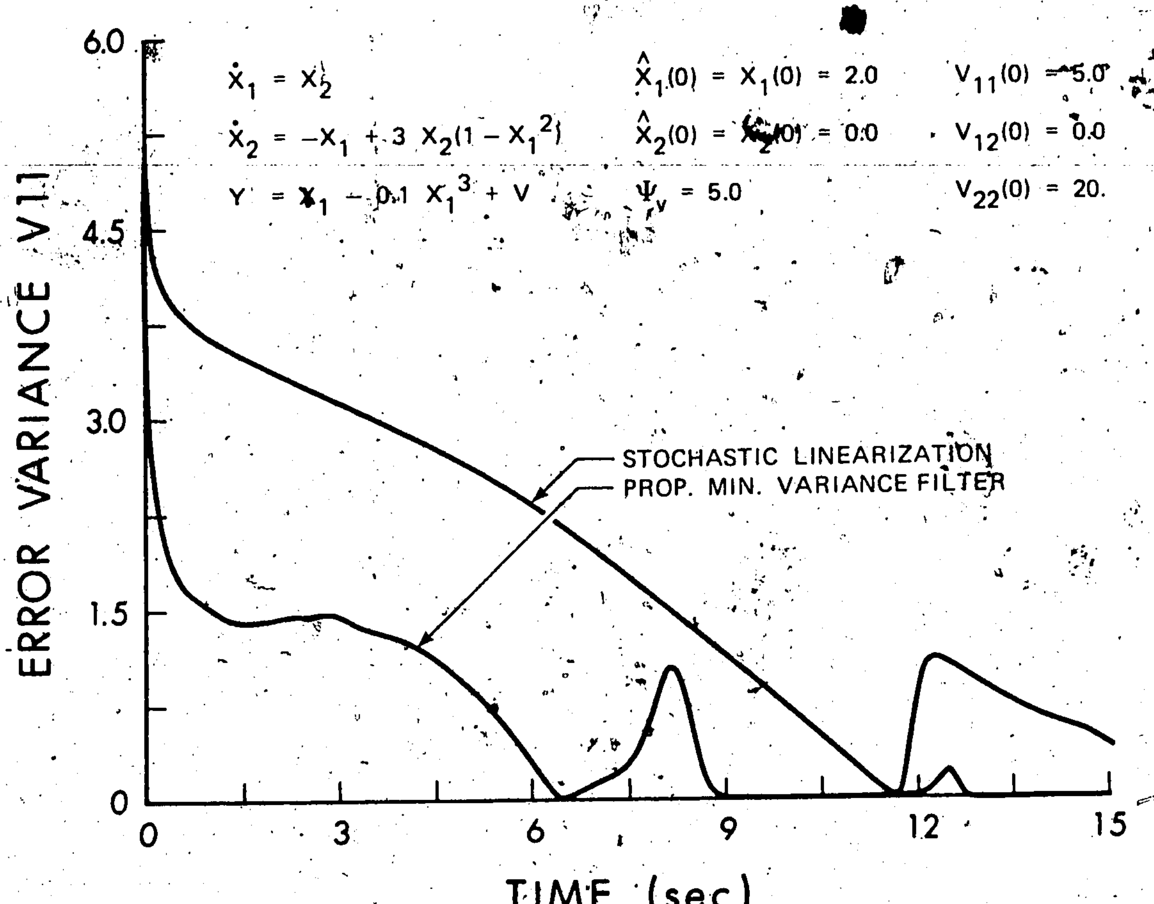

Figure 3.18 The output run of the error-variance V_{11} for systems with Ψ_u=5.0, V_{11}=5.0, V_{12}=0.0 and V_{22}=20

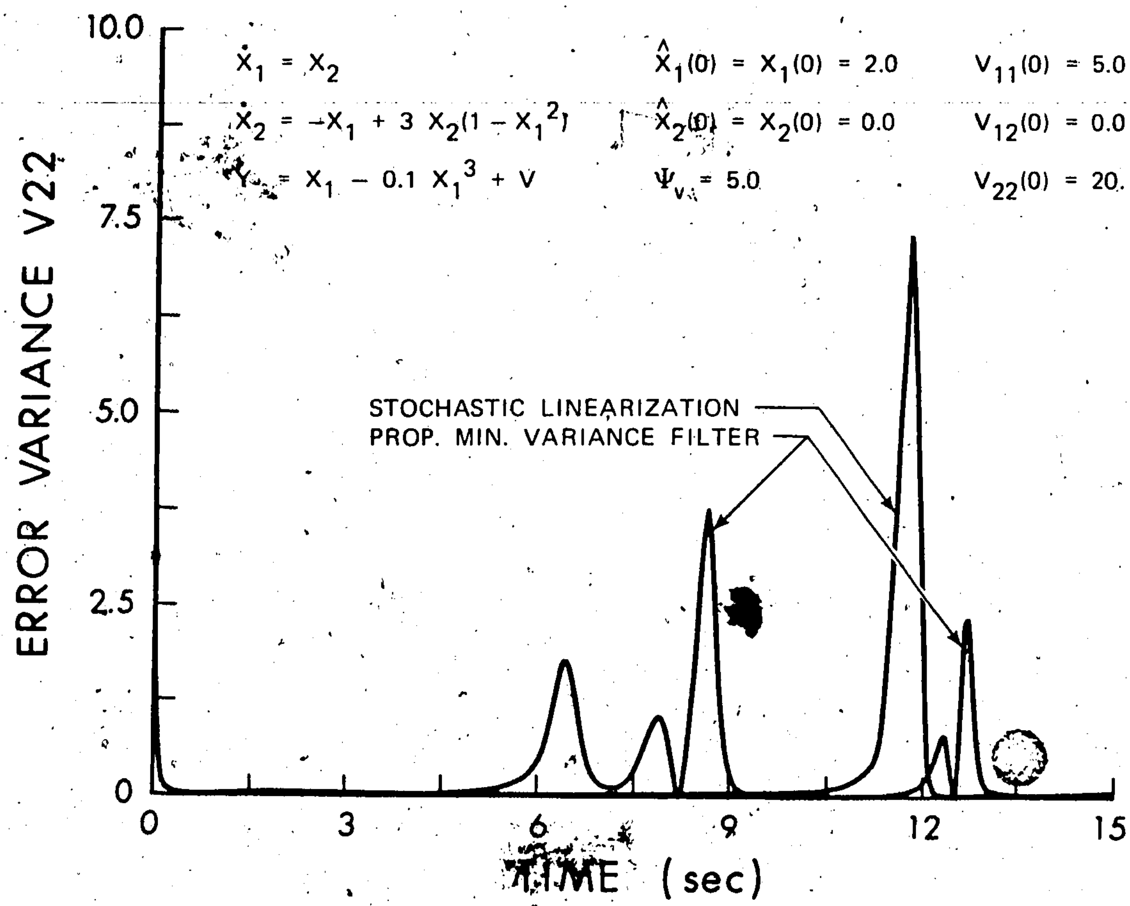

Figure 3.19 The output run of the error-variance V_{22} for systems with Ψ_u=5.0, V_{11}=5.0, V_{12}=0.0 and V_{22}=20

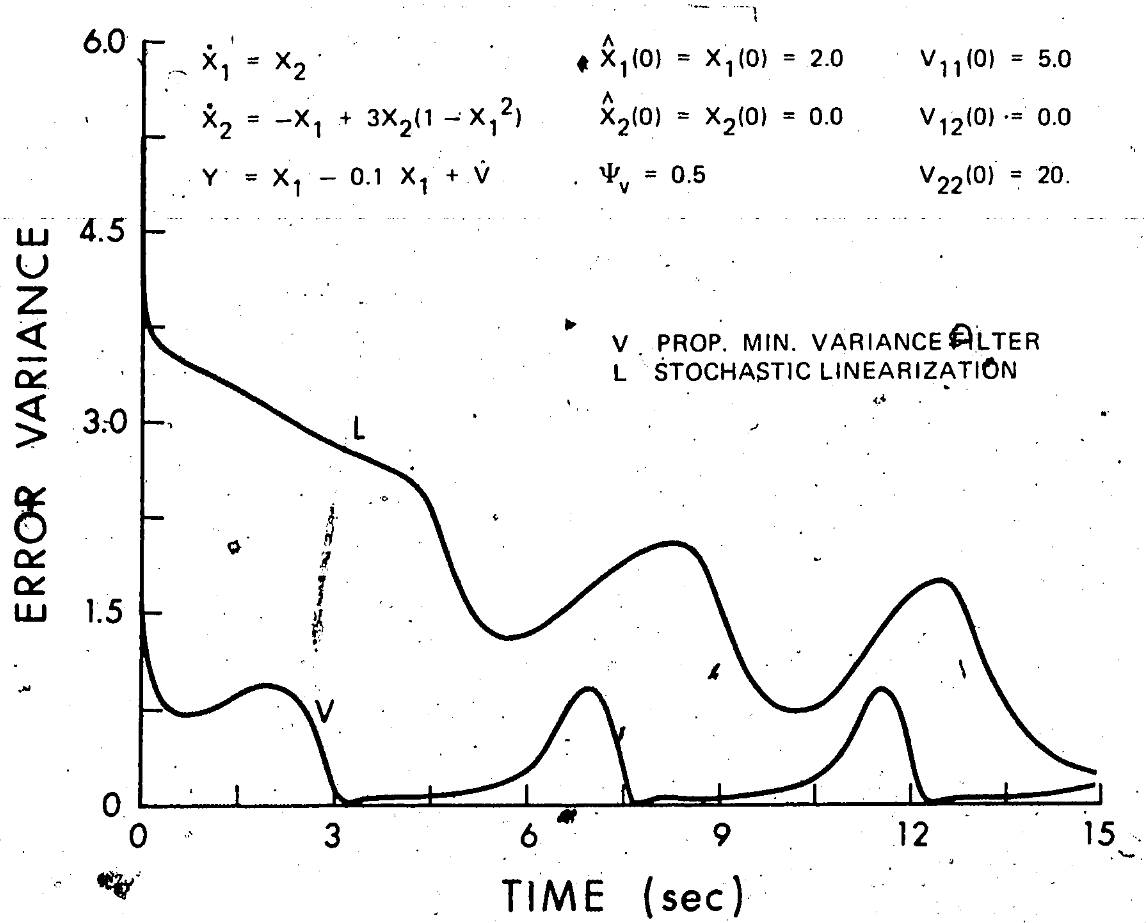

Figure 3.20 The output run of the error-variance V_{11} for systems with Ψ_v=0.5, V_{11}=5.0, V_{12}=0.0 and V_{22}=20

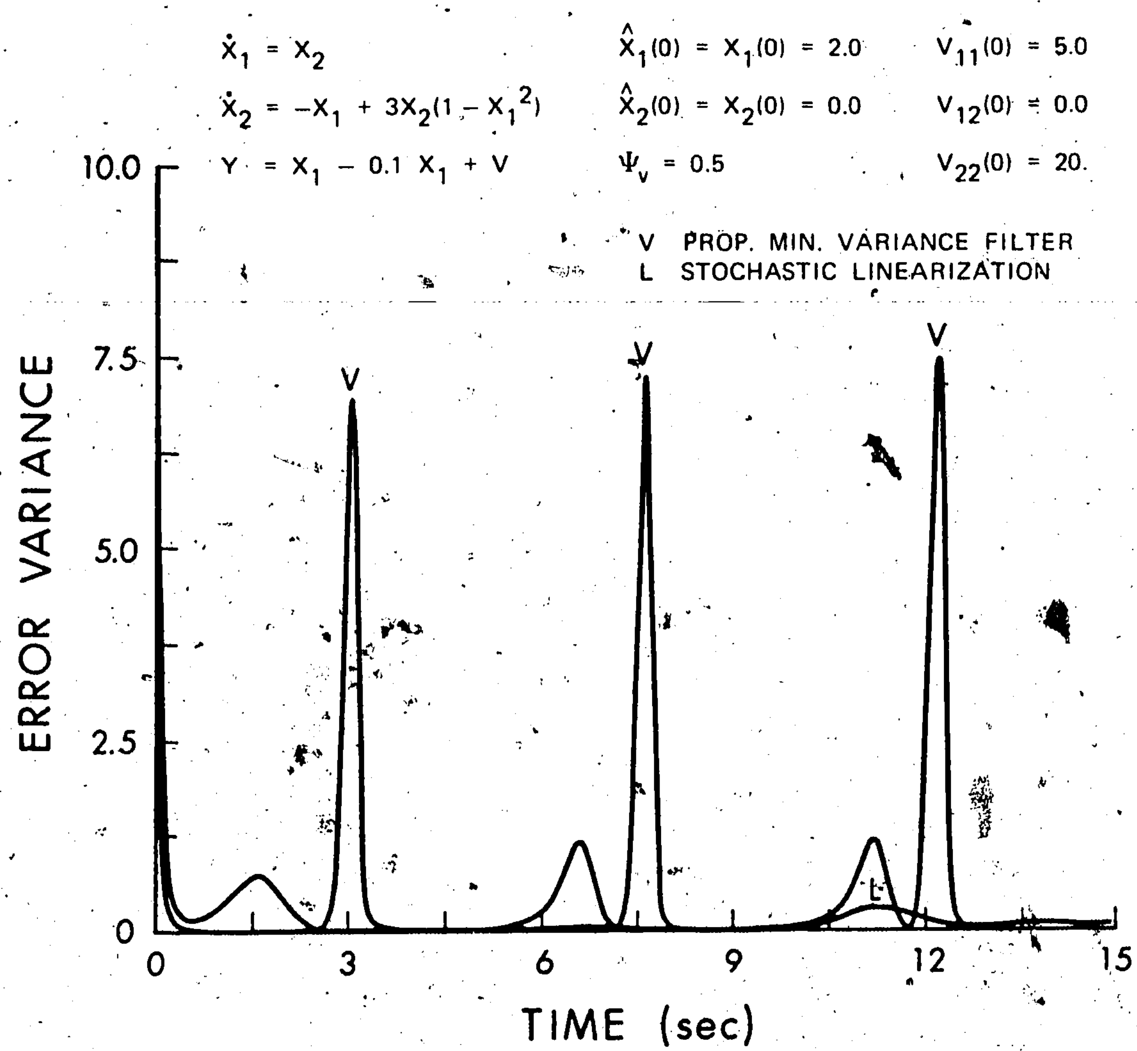

Figure 3.21 The output run of the error-variance V_{22} for systems with Ψ_v=0.5, V_{11}=5.0, V_{12}=0.0 and V_{22}=20

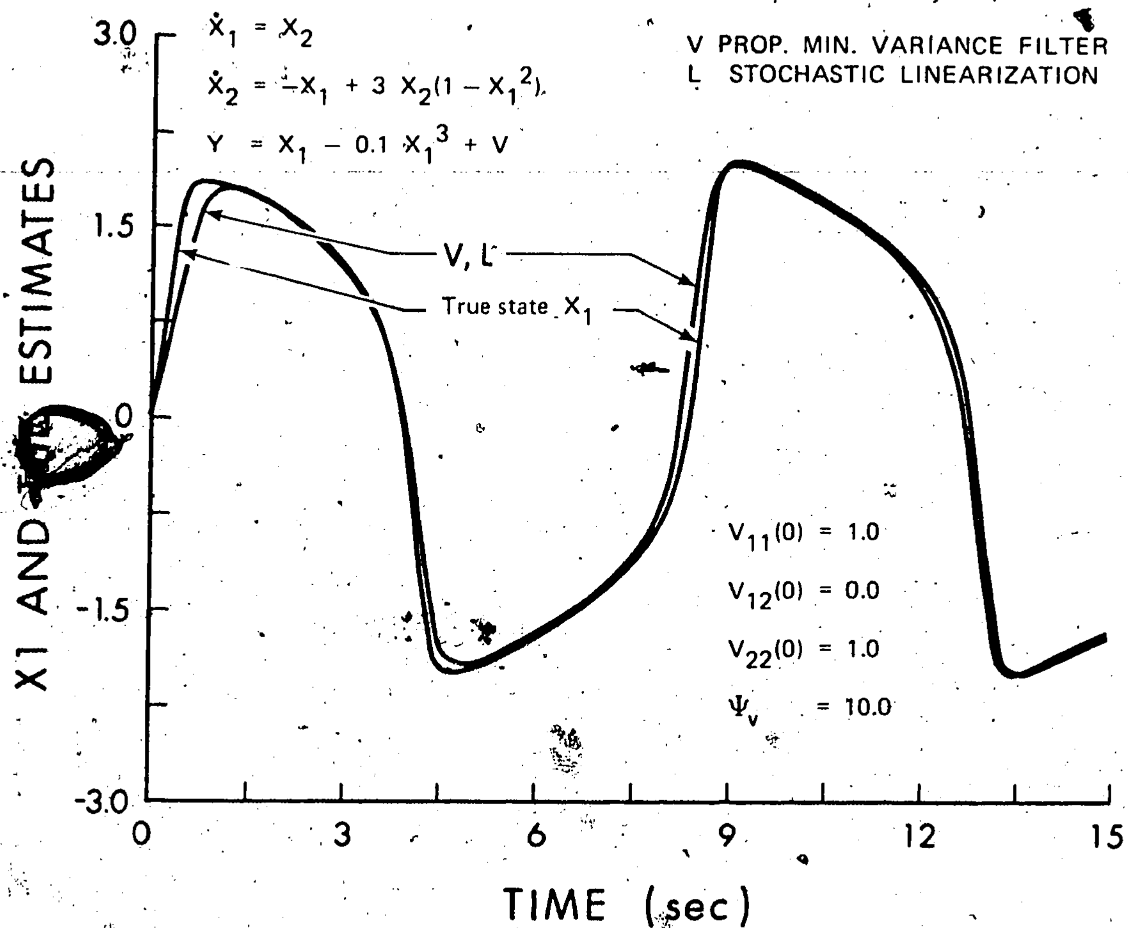

Figure 3.22 The output run of x_1 and the estimates for systems with Ψ_v=10.0, V_{11}=V_{22}=1.0 and V_{12}=0.0

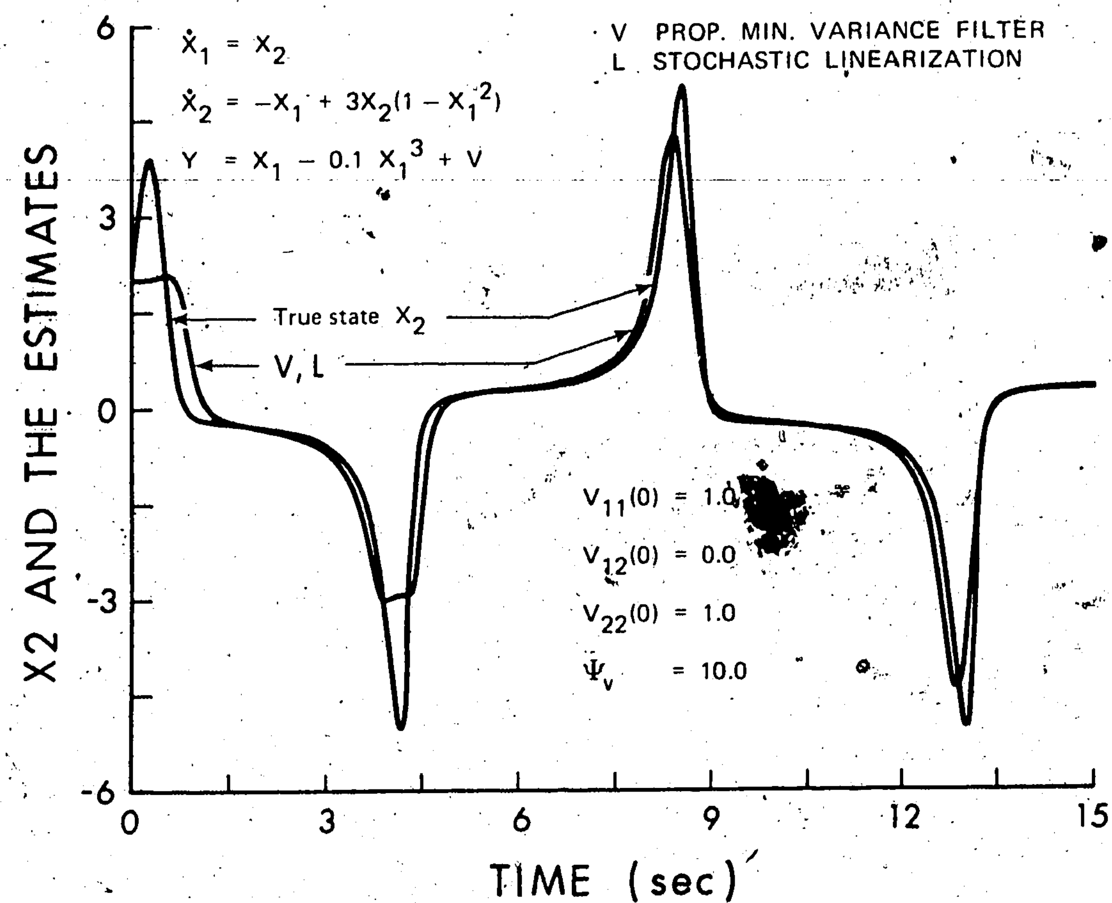

Figure 3.23 The output run of x_2 and the estimates for systems with Ψ_V=10.0, V_{11}=V_{22}=1.0 and V_{12}=0.0

filter. Also, the comparison of Figures (3.14) to (3.17) with respectively, those of Figures (3.18) to (3.21) shows that the stochastic linearization filter underestimate its error-variances, whereas the error-variances obtained from the proposed filter clearly indicate the performance of its filtering estimates.

The above observations can also be explained by the fact that, in these cases, the random forcing terms are very significant. Therefore, the deletion of the random forcing terms may result in misleading estimators.

In Figures (3.22) and (3.23), another set of statistics was chosen, in this case, the random forcing term is not significant at all, therefore, it can be expected that the two different filters yield much the same results.

On the basis of Figures (3.1) - (3.23) as well as many more runs not presented in this thesis, it can be concluded that the performance of the proposed minimum variance filter is much superior to any other type of filter investigated. It also points out the danger of adopting any other finite dimensional filter , since the random forcing terms as well as the approximation introduced to its non-linearities may be so significant that its deletion results in unacceptable as well as misleading system estimates.

3.4 *MINIMUM VARIANCE GENERAL CONTINUOUS NONLINEAR FILTER*

Consider the general nonlinear message model

$$\dot{x}(t) = f[x(t),t] + G[x(t),t]\, w(t) + E[x(t),t]\, u(t) \tag{3.21}$$

with measurement given by

$$z(t) = h[x(t),t] + v(t) + y(t) \tag{3.22}$$

where $u(t)$ and $y(t)$ are known input time functions, $w(t)$ and $v(t)$ are correlated noise processes with mean $\mu_w(t)$ and $\mu_v(t)$, respectively, and also

$$\mathrm{Cov}\{w(t), v(\tau)\} = \Psi_{wv}(t)\, \delta(t-\tau) \tag{3.23}$$

and all other prior statistics follow that of Section 3.2.

Following the development of Section 3.2, and neglecting the change of $p[\tilde{x}(t), t/Z(t)]$ due to the differential measurements $\delta z(t)$, then the filtering algorithm is given by

$$\begin{aligned}\dot{\hat{x}}(t) = {} & \hat{f}[\tilde{x}(t) + \hat{x}(t), t] + \hat{G}[\tilde{x}(t) + \hat{x}(t), t]\, \mu_w(t) \\ & + \hat{E}[\tilde{x}(t) + \hat{x}(t), t]\, u(t) + K(t)\, \{z(t) - \mu_v(t) - y(t) \\ & - \hat{h}[\tilde{x}(t) + \hat{x}(t), t]\}\end{aligned} \tag{3.24}$$

where the gain algorithm

$$K(t) = E\{\tilde{x}(t)\, h^T[\tilde{x}(t) + \hat{x}(t),t] + G[\tilde{x}(t) + \hat{x}(t),t] \text{[illegible]} \}\, \Psi_v^{-1}(t) \tag{3.25}$$

and the error-variance equation is given by

$$\dot{V}_{\tilde{x}}(t) = E\{\tilde{x}(t)\, f^T[\tilde{x}(t) + \hat{x}(t),t] + f[\tilde{x}(t) + \hat{x}(t),t]\, \tilde{x}^T(t)\}$$

$$\text{[illegible]}\ E\{G[\tilde{x}(t) + \hat{x}(t),t]\Psi_w(t)G^T[\tilde{x}(t) + \hat{x}(t),t]\}$$

$$- K(t)\, \Psi_{\text{[illegible]}} \text{[illegible]}\, K^T(t) \tag{3.26}$$

To the author's knowledge, the results presented in this section are new. In the particular case that f and h are linear, G and E are independent of the state variable, namely

$$f[x(t),t] = F(t)\, x(t)$$
$$h[x(t),t] = H(t)\, x(t)$$
$$G(x(t),t] = G(t) \quad \text{and}$$
$$E[x(t),t] = E(t)$$

Then Equations (3.24) - (3.26) are, respectively,

$$\hat{x}(t) = F(t)\,\hat{x}(t) + G(t)\mu_w(t) + E(t)\,u(t) + K(t)\,[\,z(t)$$

$$- \mu_v(t) - y(t) - H(t)\,\hat{x}(t)]$$

$$K(t) = [V_{\tilde{x}}(t)\,H^T(t) + G(t)\,\Psi_{wv}(t)]\,\Psi_v^{-1}(t)$$

and

$$\dot{V}_{\tilde{x}}(t) = F(t)\,V_{\tilde{x}}(t) + V_{\tilde{x}}(t)\,F^T(t) + G(t)\,\Psi_w(t)\,G^T(t)$$

$$- K(t)\,\Psi_v(t)\,K^T(t)$$

These algorithms agree well with the general continuous Kalman filter [49].

3.5 *MINIMUM VARIANCE CONTINUOUS NONLINEAR NOISE-FREE FILTERING*

Consider the continuous nonlinear message model of Equation (3.1) with the noise-free measurement model given by

$$z(t) = h[x(t),t] \tag{3.27}$$

Where w(t) is zero-mean, white noise with non-negative definite variance $\Psi_w(t)$, h[x(t),t] is considered to be of full rank, otherwise an equivalent z(t) of lower dimension can be used.

Since z(t) is noise free $\Psi_w(t)$ is non-negative definite, when z(t) is differentiated, some of its element may not contain any white noise. Therefore, each element of z(t) has to be differentiated and

Equation (3.1) is used to substitute for the derivative of the state variable, until white noise is obtained in the derivatives of each element in $z(t)$.

The signals obtained can be arranged into two sets: In the first set,

$$z_1(t) = h_1[x(t),t]$$

which comprises Equation (3.27) and all linearly independent derivatives of $z(t)$ that are noise-free.

In the second set

$$z_2(t) = h_2[x(t),t] + N[x(t),t]\,w(t)$$

which comprises all derivatives of $z(t)$ that contain linearly independent white noise.

It is assumed that the filtering estimate $\hat{x}(t)$ is given by

$$\dot{\hat{x}}(t) = \ell[\hat{x}(t),t] + K_1(t)\,z_1(t) + K_2(t)\,z_2(t)$$

since in place of Equation (3.27), there are two sets of measurements.

Here, $z_1(t)$ is considered as a known input, it does not contain any new information, and following the development of Section 3.2, the following filtering algorithm is obtained

$$\dot{\hat{x}}(t) = \hat{f}[\tilde{x}(t) + \hat{x}(t), t] + K_2(t)\{z_2(t) - \hat{h}_2[\tilde{x}(t) + \hat{x}(t), t]\}$$

with the gain algorithm

$$K_2(t) = E\{\tilde{x}(t)\, h_2^T[\tilde{x}(t) + \hat{x}(t), t] + G[\tilde{x}(t) + \hat{x}(t), t]\Psi_w(t)\, N^T[\tilde{x}(t) + \hat{x}(t), t]\}[E\{N[\tilde{x}(t) + \hat{x}(t), t]\Psi_w(t)\, N^T[\tilde{x}(t) + \hat{x}(t), t]\}]^{-1}$$

and the error-variance equation is given as

$$V_{\tilde{x}}(t) = E\{\tilde{x}(t)\, f^T[\tilde{x}(t) + \hat{x}(t), t] + f[\tilde{x}(t) + \hat{x}(t), t]\, \tilde{x}^T(t)\} + E\{G[\tilde{x}(t) + \hat{x}(t), t]\, \Psi_w(t)\, G^T[\tilde{x}(t) + \hat{x}(t), t]\} - K_2(t)E\{N[\tilde{x}(t) + \hat{x}(t), t]\, \Psi_w(t) N^T[\tilde{x}(t) + \hat{x}(t), t]\}K_2^T(t)$$

To the author's knowledge, the results presented here are also new. In the case that the vector valued functions are linear, the results presented here agree well with that of Bryson and Johansen [6].

3.6 *CONCLUSION*

In this chapter, nonlinear filtering algorithms for correlated, uncorrelated or noise-free processes were derived. To the author's knowledge the filtering algorithms presented for non-white noise processes are new. In order that the systems be physically realizable, the assumption that the probability density functions of the estimator error are Gaussian is needed to expand the expectations.

In Section 3.3, various nonlinear systems were simulated, the noise processes assumed are white Gaussian, they are compared with results obtained from various other finite dimensional approximate nonlinear filters. The results clearly indicate the superiority of the proposed [illegible] variance filter over those of other filters investigated. [illegible] theoretical explanation are also given for the apparent poor performance characteristics of various other filters considered.

In the case that the systems are linear, the proposed minimum variance filters agree well with those of the optimal filters in literature [6, 49].

CHAPTER IV

MINIMUM-VARIANCE FILTERING AND SMOOTHING FOR DISCRETE NONLINEAR DELAYED SYSTEMS WITH ADDITIVE WHITE NOISE

4.1 *INTRODUCTION*

This chapter deals with discrete-time filtering and smoothing estimation of nonlinear systems with multiple delays imbedded in additive white noise processes. In general, the filtering algorithm enables one to estimate present values of the variables of interest using present data, whereas the smoother allows [illegible]mate past values. A typical smoothing problem is the post-fl[illegible]mation of the flight path of a missile based on tracking system measurements during the entire duration of the flight. If the estimates of the missile's position and velocity at one particular flight point are desired, the estimates can be based upon all the measurements recorded, including those made before and after that particular flight point.

In the estimation problems for linear systems without delays involving additive white noise processes, numerous papers have been written to deal with filtering and smoothing estimation, however, most of them are merely rederivation of earlier recursive algorithms presented by Carlton [50], Rauch [51] and Bryson and Frazier [11], or reformulation of problems using various estimation techniques.

However, Kelly and Anderson [52] pointed out that the algorithms for both discrete and continuous-time, linear, fixed-lag smoothing given in [11, 15, 43, 51] may be unstable, and therefore impractical. To be more explicit, although the fixed-lag smoothing equations are bounded-input and bounded-output stable, realizations of these in [43, 51] contain a subsystem which is unstable in the sense of Lyapunov. In [53], it is pointed out that the apparent culprit is an uncontrollable and unstable block in the smoother state equations which can be removed without affecting the input-output characteristics.

In [38], a computationally stable smoothing algorithm is derived for linear discrete systems containing time delays, using the method of orthogonal projection. The smoother for linear discrete systems without delays can be considered as a special case of the above problem, with time delay index setting to zero. The results of [38] are rederived in a simple manner in [39], using the state augmentation technique. However, the smoothers derived in such a manner are of nN dimensions, where n is the order of the message model and N is the amount of the fixed-lag.

In this chapter, the matrix minimum principle is applied to nonlinear discrete-time multiple delays systems, resulting in new dynamic discrete estimation algorithms; they are reported in [53, 54]. In Section 4.2, the message model involves multiple time delays but the measurement is without delay. In Section 4.3, more general nonlinear system models are considered, where both the message and measurement models contain multiple time delays.

For nonlinear estimation problems, the results obtained in this chapter are new, recursive in nature and directly yield the fixed-interval, fixed-lag, fixed point smoothing and the filtering algorithms. The derivation is straightforward and shows the close link between the smoothing and filtering estimation algorithms.

The estimation algorithms presented in this chapter, can be easily applied to various special cases of nonlinear estimation problems; for example, the estimation problem for nonlinear systems without delays, the estimation of nonlinear systems with multiple delays only appearing in the measurements, the linear estimation of systems with or without delays, etc.

In the particular case of linear message and measurement models with non-delayed measurements, the proposed filtering and fixed-lag smoothing algorithms agree well with stable estimation algorithms presented by Priemer and Vacroux [37-38], and later extended by Biswas and Mahalanabis [55], and Shukla and Srinath [56], to fixed point smoothing of the same systems and fixed-lag smoothing for systems with multiple delays in the measurements, respectively.

4.2 *NONLINEAR DISCRETE DELAYED SYSTEMS WITH NON-DELAYED MEASUREMENTS*

This section is concerned with nonlinear discrete-time dynamic systems containing multiple time delays with non-delayed measurement sequence imbedded in additive white noise processes.

In Section 4.2.1, the problem statement is presented. Section

4.2.2 presents the derivation of the nonlinear [illegible]. Sections 4.2.3 to 4.2.6, provide handy sets of reference equations respectively, for the fixed-lag, fixed-point, fixed-interval smoothing and the filtering algorithms. Section 4.2.7 shows the applicability of the presented algorithms to linear estimation problems and also various other special cases of estimation problems they were identified with existing algorithms in literature.

4.2.1 *The Problem Statement*

The processes considered in this section are described by the following vector difference equation

$$x(k+1) = \sum_{j=o}^{L} f_j[x(k-j), k-j] + G[x(k),k] w(k) \qquad (4.2.1)$$

where the measurement is given by

$$y(k) = h[x(k),k] + v(k) \qquad (4.2.2)$$

Here, x the state is an n-vector; y the measurement, an m-vector; w the random input, an r-vector; v the measurement noise, an m-vector; G, a nonlinear state dependent n x r matrix; k = 0,1,... is the discrete time index. The nonlinear vector valued functions f_j and h are, respectively, n and m - dimensional.

The random vectors w(k) and v(k) are independent zero-mean white Gaussian sequences, for which

$$E\{w(k)\ w^T(j)\} = \Psi_w(k)\ \delta_{kj}$$

$$E\{v(k)\ v^T(j)\} = \Psi_v(k)\ \delta_{kj}$$

and

$$E\{w(k)\ v^T(j)\} = 0$$

for all integers k and j, where $E\{\cdot\}$ denotes the expectation operator, T the matrix transpose, δ_{jk} the Kronecker delta, and Ψ_v and Ψ_w are m x m and r x r positive definite matrices, respectively.

The initial states x(0) and x(-j), for j = 1, ..., L are zero-mean Gaussian random vectors, independent of v(k) and w(k), for which

$$E\{x(-j)\ w^T(k)\} = 0$$

and

$$E\{x(-j)\ v^T(k)\} = 0$$

for j = 0,1 ..., L, and k = 0,1, ..., with a positive definite covariance matrix

$$E\{x(-j)\ x^T(-\ell)\} = V_x(j,\ell)$$

for j, ℓ = 0, ..., L.

The smoothing problem is to obtain $\hat{x}(k-\ell+1/k+1)$, the unbiased smoothing estimate of $x(k-\ell+1)$, with $0 \leq \ell \leq k+1$, conditioned on the set of measurements

$$Y(k+1) = \{y(0), y(1), \ldots, y(k+1)\}$$

such that the cost function

$$J = \sum_{j=o}^{k+1} \text{Trace}\,[M(j)\,V_{\tilde{x}}(j-\ell+1/j+1)] \tag{4.2.3}$$

is minimized. Here $M(j)$ is some non-negative definite weighting matrix, and $V_{\tilde{x}}(j-\ell+1/j+1)$ is defined by

$$V_{\tilde{x}}(j-\ell+1/j+1) = E_{j+1}[\{x(j-\ell+1) - \hat{x}(j-\ell+1/j+1)\}$$
$$\{x(j-\ell+1) - \hat{x}(j-\ell+1/j+1)\}^T] \tag{4.2.4}$$

where $E_{j+1}\{\cdot\}$ denotes the expectation operation conditioned on the set of measurements $Y(j+1)$.

Since $M(j)$ is known, minimizing Equation (4.2.3) is equivalent to minimizing the trace of $V_{\tilde{x}}(j-\ell+1/j+1)$ at each individual sample point j [49].

4.2.2 *The Derivation of the Smoother*

It is assumed that the smoothing algorithm is described by the nonlinear differential equation

$$\hat{x}(k-\ell+1/k+1) = \sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] + K^{\ell}_{k+1}y(k+1) \quad (4.2.5)$$

Here, the assumption of linearity in innovations is made. The nonlinear functions $\sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j]$ and K^{ℓ}_{k+1} are yet to be determined.

In fact, the smoothing equation may take various forms, however, it is essential that enough information concerning $\hat{x}(k-\ell-j/k)$ and $y(k+1)$ are included in the estimator model. Since the message model considered in this section consists of the sum of nonlinear functions containing multiple time delays, it is realistic to assume that the smoothing estimate is a linear combination of the sum of nonlinear functions of $\hat{x}(k-\ell-j/k)$, to account for the time delay characteristics, and the present measurement $y(k+1)$. Here $\hat{x}(k-\ell-j/k)$ is assumed to have made optimum use of all the measurements up to $y(k)$.

The problem formulated in such a manner may therefore lead to a smoother, optimum with respect to the imposed constraints, but not identical to the truly optimal one.

In the meantime, it is interesting to note that even though the smoothing algorithm may be described by the dynamic equations such as

$$\hat{x}(k-\ell+1/k+1) = \sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] + K^{\ell}_{k+1}\tilde{y}(k+1/k) \quad (4.2.6)$$

where $\tilde{y}(k+1/k) = y(k+1) - E_k\{y(k+1)\}$. It represents the "new" information contained in $y(k+1)$, since the best estimate of $y(k+1)$ given $Y(k)$, namely,

$E_k\{y(k+1)\}$ is subtracted from $y(k+1)$ to obtain $\tilde{y}(k+1/k)$, which is generally known as the innovation. The resulting smoother using either Equation (4.2.5) or (4.2.6) would yield exactly the same dynamic smoother. This can be explained by the fact that the two equations contain exactly the same amount of information, that is, $y(k+1)$ contains as much information as that of its innovation $\tilde{y}(k+1/k)$ [14].

Now the estimation problem is to determine the time-varying nonlinear vector function $\sum_{j=0}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j]$, and the algorithm K_{k+1}^{ℓ} such that the trace of $V_{\tilde{x}}(k-\ell+1/k+1)$ is minimized.

Let $\tilde{x}(k-\ell+1/k+1)$ denote the smoothing error defined by

$$\tilde{x}(k-\ell+1/k+1) = x(k-\ell+1) - \hat{x}(k-\ell+1/k+1) \tag{4.2.7}$$

then

$$\begin{aligned}\tilde{x}(k-\ell+1/k+1) = & \sum_{j=0}^{L} f_j[x(k-\ell-j), k-\ell-j] + G[x(k-\ell),k-\ell]w(k-\ell) \\ & - \sum_{j=0}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] - K_{k+1}^{\ell}\{h[x(k+1),k+1] \\ & + v(k+1)\}\end{aligned} \tag{4.2.8}$$

In order that $\hat{x}(k-\ell+1/k+1)$ is an unbiased smoothing estimate, it is necessary that

$$\begin{aligned}\sum_{j=0}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] = & \sum_{j=0}^{L} \hat{f}_j[x(k-\ell-j), k-\ell-j/k] \\ & - K_{k+1}^{\ell}\hat{h}[x(k+1), k+1/k]\end{aligned} \tag{4.2.9}$$

where

$$\hat{f}_j[x(k-\ell-j),\ k-\ell-j/k] = E_k\{f_j[x(k-\ell-j),\ k-\ell-j]\}$$

and

$$\hat{h}[x(k+1),\ k+1/k] = E_k\{h[x(k+1),\ k+1]\}$$

Substituting Equations (4.2.9) into (4.2.5), the smoothing algorithm becomes

$$\hat{x}(k-\ell+1/k+1) = \hat{x}(k-\ell+1/k) + K^{\ell}_{k+1}\ \{y(k+1) - \hat{h}[x(k+1),\ k+1/k]\} \qquad (4.2.10)$$

for k = 0,1, ..., and $0 \leq \ell \leq k+1$.

Where

$$\hat{x}(k-\ell+1/k) = \sum_{j=o}^{L} \hat{f}_j[x(k-\ell-j),\ k-\ell-j/k] \qquad (4.2.11)$$

Equation (4.2.8) can be rewritten as

$$\tilde{x}(k-\ell+1/k+1) = \tilde{x}(k-\ell+1/k) - K^{\ell}_{k+1}\{\tilde{h}[x(k+1)/k] + v(k+1)\} \qquad (4.2.12)$$

where

$$\tilde{h}[x(k+1)/k] = h[x(k+1),\ k+1] - \hat{h}[x(k+1),\ k+1/k] \qquad (4.2.13)$$

Therefore, the smoothing error-variance equation is given by

$$V_{\tilde{x}}(k-\ell+1/k+1) = V_{\tilde{x}}(k-\ell+1/k) - E_k\{\tilde{x}(k-\ell+1/k)\,\tilde{h}^T[x(k+1)/k]\}$$

$$K^{\ell}_{k+1} - K^{\ell}_{k+1}E_k\{\tilde{h}[x(k+1)/k]\,\tilde{x}^T(k-\ell+1/k)\}$$

$$+ K^{\ell}_{k+1}\,\Psi_v(k+1)\,K^{\ell^T}_{k+1} \qquad (4.2.14)$$

$$+ K^{\ell}_{k+1}E_k\{\tilde{h}[x(k+1)/k]\,\tilde{h}^T[x(k+1)/k]\}\,K^{\ell^T}_{k+1}$$

Here, K^{ℓ}_{k+1} is the only variable available for manipulation. The necessary condition for minimizing the trace of $V_{\tilde{x}}(k-\ell+1/k+1)$ and subject to the constraint of Equation (4.2.14) is provided by the matrix minimum principle, for which K^{ℓ}_{k+1} is considered as the control variable.

The necessary condition can now be obtained from the condition

$$\frac{\partial}{\partial K^{\ell}_{k+1}}\,\mathrm{Trace}[V_{\tilde{x}}(k-\ell+1/k+1)] = [0] \qquad (4.2.15)$$

where [0] is the null matrix, and the result is

$$K^{\ell}_{k+1} = E_k\{\tilde{x}(k-\ell+1/k)\,\tilde{h}^T[x(k+1)/k]\}\,[\Psi_v(k+1)$$

$$+ E_k\{\tilde{h}[x(k+1)/k]\,\tilde{h}^T[x(k+1)/k]\}^{-1} \qquad (4.2.16)$$

and Equation (4.2.14) becomes

$$V_{\tilde{x}}(k-\ell+1/k+1) = V_{\tilde{x}}(k-\ell+1/k)$$

$$- K^{\ell}_{k+1} E_k\{\tilde{h}[x(k+1)/k]\tilde{x}^T(k-\ell+1/k)\} \qquad (4.2.17)$$

And since, it is seen

$$\tilde{x}(k-\ell+1/k) = \sum_{j=0}^{L} \tilde{f}_j[x(k-\ell-j)/k] + G[x(k-\ell),k-\ell]w(k-\ell)$$

where

$$\tilde{f}_j[x(k-\ell-j)/k] = f_j[x(k-\ell-j), k-\ell-j]$$

$$- \hat{f}_j[x(k-\ell-j), k-\ell-j/k]$$

then

$$V_{\tilde{x}}(k-\ell+1/k) = \sum_{i=0}^{L} \sum_{j=0}^{L} E_k\{\tilde{f}_i(k-\ell-i)/k] \tilde{f}_j^T[x(k-\ell-j)/k]\}$$

$$+ G[x(k-\ell), k-\ell] \Psi_w(k-\ell) G^T[x(k-\ell), k-\ell] \qquad (4.2.18)$$

and one can also obtain

$$V_{\tilde{x}}(k-\ell-i, k-\ell-j/k) = V_{\tilde{x}}(k-\ell-i, k-\ell-j/k-1)$$

$$- K^{\ell+i}_{k} E_{k-1}\{\tilde{h}[x(k)/k-1]\tilde{x}^T(k-\ell-j/k-1)\}$$

(4.2.19)

for $0 \leq \ell + i$, $\ell + j \leq k$, where $V_{\tilde{x}}(k-\ell-i, k-\ell-j/k)$ is defined by the relation

$$V_{\tilde{x}}(k-\ell-i,\ k-\ell-j/k) = E_k\{\tilde{x}(k-\ell-i/k)\ \tilde{x}^T(k-\ell-j/k)\}$$

For $k<0$, there is no input to the smoother and therefore one can set $\hat{x}(-j/-1)$ to zero for $j = 0,1, \ldots, L$, which results in

$$\tilde{x}(-j/-1) = x(-j)$$

and

$$V_{\tilde{x}}(-j,-\ell/-1) = V_x(j,\ell)$$

for $j,\ \ell = 0,1, \ldots, L$.

Since the smoothing estimator is unbiased, the expectations that are in Equations (4.2.16) - (4.2.19) can be replaced by the following

$$E_k\{\tilde{x}(k-\ell+1/k)\ \tilde{h}^T[x(k+1)/k]\} = E_k\{\tilde{x}(k-\ell+1/k)h^T[x(k+1),\ k+1]\} \quad (4.2.20)$$

$$E_k\{\tilde{h}[x(k+1)/k]\tilde{h}^T[x(k+1)/k]\} = E_k\{h[x(k+1),\ k+1]h^T[x(k+1),\ k+1]\} - \hat{h}[x(k+1),k+1/k]\hat{h}^T[x(k+1),k+1/k] \quad (4.2.21)$$

and

$$E_k\{\tilde{f}_i[x(k-\ell-i)/k]\ \tilde{f}_j^T[x(k-\ell-j)/k]\}$$

$$= E_k\{f_i[x(k-\ell-i),\ k-\ell-i]\ f_j^T[x(k-\ell-j),\ k-\ell-j]\} \tag{4.2.22}$$

$$- \hat{f}_i[x(k-\ell-i),\ k-\ell-i/k]\ \hat{f}_j^T[x(k-\ell-j),\ k-\ell-j/k]$$

It should be noted that in the case of nonlinear systems the above expectations require infinite dimensional systems for their realization. The problem of obtaining a good approximation to the above expectations is, therefore, of practical importance.

For practical realization, it is assumed that the conditional probability density functions of the smoothing error $\tilde{x}$ are Gaussian, then the expectation can be obtained without any further approximation for systems with product types or polynomial nonlinearities.

To evaluate the above expectation, one can make use of Equation (4.2.7) to replace the state variables by the sum of their respective estimators and the estimator errors. For example, in the scalar case of

$$h[x(k+1),\ k+1] = x^3(k+1)$$

from Equations (4.2.20) and (4.2.7), it can be shown that

$$E_k\{\tilde{x}(k-\ell+1/k)\ \tilde{h}^T[x(k+1)/k]\} = E_k\{\tilde{x}(k-\ell+1/k)$$

$$\cdot[\tilde{x}^3(k+1/k) + 3\ \tilde{x}^2(k+1/k)\ \hat{x}(k+1/k) + 3\ \tilde{x}(k+1/k)$$

$$\cdot\ \hat{x}^2(k+1/k) + \hat{x}^3(k+1/k)]\} = 3\ V_{\tilde{x}}(k-\ell+1,\ k+1/k)\ \hat{x}^2(k+1/k)$$

$$+ 3\ V_{\tilde{x}}(k-\ell+1,k+1/k)\ V_{\tilde{x}}(k+1/k)$$

Notice that in the evaluation of the expectation, the only assumption needed is the Gaussian assumption of the estimator error. And it can be easily shown that if the nonlinearities were expanded in terms of first or second order Taylor's series, the last term in the preceeding equation would have been dropped. In such a case the resulting filter is equivalent to an intelligent application of an extended Kalman filter.

Similarly, it is obtained from Equations (4.2.21) and (4.2.7)

$$E_k\{\tilde{h}[x(k+1)/k]\ \tilde{h}^T[x(k+1)/k]\}$$

$$= E_k\{x^6(k+1)\} - E_k\{x^3(k+1)\}^2 = [15\ V^3_{\tilde{x}}(k+1/k)$$

$$+ 45\ V^2_{\tilde{x}}(k+1/k)\ \hat{x}^2(k+1/k) + 15\ V_{\tilde{x}}(k+1/k)$$

$$\cdot\ \hat{x}^4(k+1/k) + \hat{x}^6(k+1/k)] - [\hat{x}^3(k+1/k) + 3\ V_{\tilde{x}}(k+1/k)$$

$$\cdot\ \hat{x}(k+1/k)]^2 = 15\ V^3_{\tilde{x}}(k+1/k) + 36\ V^2_{\tilde{x}}(k+1/k)\ \hat{x}^2(k+1/k)$$

$$+ 9\ V_{\tilde{x}}(k+1/k)\ \hat{x}^4(k+1/k)$$

The preceeding evaluation is rather straightforward. In the particular case that higher-order terms do not improve system accuracy, one can delete the higher-order terms and approximate the expectations up to arbitrary order, as desired.

4.2.3 *The Nonlinear Fixed-Lag Smoothing*

Replace k+1 and ℓ by k and N, respectively, where N<K, from Equations (4.2.10), (4.2.11), (4.2.16) and (4.2.19) one would then obtain the following recursive nonlinear fixed-lag smoothing algorithms:

$$\hat{x}(k-N/k) = \hat{x}(k-N/k-1) + K^N_k\{y(k) - \hat{h}[x(k), k/k-1]\} \qquad (4.2.23)$$

$$\hat{x}(k-N/k-1) = \sum_{j=o}^{L} \hat{f}_j[x(k-N-j), k-N-1-j/k-1] \qquad (4.2.24)$$

$$K^N_k = E_{k-1}\{\tilde{x}(k-N/k-1)\ \tilde{h}^T[x(k)/k-1]\}[\Psi_v(k) + E_{k-1}\{\tilde{h}[x(k)/k-1]$$

$$\tilde{h}^T[x(k)/k-1]\}]^{-1} \qquad (4.2.25)$$

$$V_{\tilde{x}}(k-N/k) = V_{\tilde{x}}(k-N/k-1) \qquad (4.2.26)$$

$$- K^N_k E_{k-1}\{\tilde{h}[x(k)/k-1]\ \tilde{x}^T(k-N/k-1)\}$$

$$V_{\tilde{x}}(k-N/k-1) = \sum_{i=0}^{L} \sum_{j=0}^{L} E_{k-1}\{\tilde{f}_i[x(k-1-N-i)/k-1] \tag{4.2.27}$$

$$\cdot \tilde{f}_j^T[x(k-1-N-j)/k-1]\} + G[x(k-1-N), k-1-N]$$

$$\cdot \Psi_w(k-1-N)\, G^T[x(k-1-N), k-1-N]$$

$$V_{\tilde{x}}(k-1-N-i, k-1-N-j/k-1) = V_{\tilde{x}}(k-1-N-i, k-1-N-j/k-2) \tag{4.2.28}$$

$$- K_{k-1}^{N+i} E_{k-2}\{\tilde{h}[x(k-1)/k-2]$$

$$\cdot \tilde{x}^T(k-1-N-j/k-2)\}$$

Also

$$V_{\tilde{x}}(k-N, k+1/k) = E_k\{\tilde{x}(k-N/k) \sum_{j=0}^{L} \tilde{f}_j^T[x(k-j)/k]\}$$

and

$$V_{\tilde{x}}(k+1/k) = \sum_{i=0}^{L} \sum_{j=0}^{L} E_k\{\tilde{f}_i[x(k-i)/k]\, \tilde{f}_j^T[x(k-j)/k]\} \tag{4.2.29}$$

$$+ G[x(k), k]\, \Psi_w(k)\, G^T[x(k),k]$$

4.2.4 *The Nonlinear Fixed-Point Smoothing*

Setting $\ell = K - N+1$ where $k+1 > N$, from Equations (4.2.10), (4.2.11), (4.2.16)- (4.2.18), one would then have the following recursive nonlinear fixed-point smoothing algorithms.

$$\hat{x}(N/k+1) = \hat{x}(N/k) + K_{k+1}^{k-N+1}\{y(k+1) - \hat{h}[x(k+1), k+1/k]\}$$

$$\hat{x}(N/k) = \sum_{j=0}^{L} \hat{f}_j[x(N-1-j),\ N-1-j/k]$$

$$K_{k+1}^{k-N+1} = E_k\{\tilde{x}(N/k)\ \tilde{h}^T[x(k+1)/k]\}\ [\Psi_v(k+1)$$

$$+ E_k\{\tilde{h}[x(k+1)/k]\ \tilde{h}^T[x(k+1)/k]\}]^{-1}$$

$$V_{\tilde{x}}(N/k+1) = V_{\tilde{x}}(N/k) - K_{k+1}^{k-N+1}\ E_k\{\tilde{h}[x(k+1)/k]\ \tilde{x}^T(N/k)\}$$

$$V_{\tilde{x}}(N/k) = \sum_{i=0}^{L} \sum_{j=0}^{L} E_k\{\tilde{f}_i[x(N-1-i)/k]\ \tilde{f}_j^T[x(N-1-j)/k]\}$$

$$+ G[x(N-1),N-1]\ \Psi_w(N-1)\ G^T[x(N-1),N-1]$$

$$V_{\tilde{x}}(N-1-i,N-1-j/k) = V_{\tilde{x}}(N-1-i,N-1-j/k-1)$$

$$- K_k^{k-N+1+i}\ E_{k-1}\{\tilde{h}[x(k)/k-1]\ \tilde{x}^T(N-1-j/k-1)\}$$

and

$$V_{\tilde{x}}(N,\ k+1/k) = E_k\{\tilde{x}(N/k) \sum_{j=0}^{L} \tilde{f}_j^T[x(k-j)/k]\}$$

$V_{\tilde{x}}(k+1/k)$ is the same as Equation (4.2.29), and also from Equation (4.2.19), it can be shown that

$$V_{\tilde{x}}(N,\ k-j/k) = V_{\tilde{x}}(N,\ k-j/k-1) - K_k^{k-N}\ E_{k-1}\{\tilde{h}[x(k)/k-1]\tilde{x}^T(k-j/k-1)\}$$

4.2.5 *The Nonlinear Fixed-Interval Smoothing*

Setting k+1=N, ℓ=N-k, where N≥k, Equations (4.2.10), (4.2.11), (4.2.16) - (4.2.19), would yield the following recursive nonlinear fixed-interval smoothing algorithms.

$$\hat{x}(k/N) = \hat{x}(k/N-1) + K_N^{N-k} \{y(N) - \hat{h}[x(N), N/N-1]\}$$

$$\hat{x}(k/N-1) = \sum_{j=o}^{L} \hat{f}_j[x(k-1-j), k-1-j/N-1]$$

$$K_N^{N-k} = E_{N-1}\{\tilde{x}(k/N-1)\, \tilde{h}^T[x(N)/N-1]\}$$

$$[\Psi_v(N) + E_{N-1}\{\tilde{h}[x(N)/N-1]\, \tilde{h}^T[x(N)/N-1]\}]^{-1}$$

$$V_{\tilde{x}}(k/N) = V_{\tilde{x}}(k/N-1) - K_N^{N-k} E_{N-1}\{\tilde{h}[x(N)/N-1]\, \tilde{x}^T(k/N-1)\}$$

$$V_{\tilde{x}}(k/N-1) = \sum_{i=o}^{L} \sum_{j=o}^{L} E_{N-1}[\tilde{f}_i[x(k-1-i)/N-1]$$

$$\tilde{f}_j^T[x(k-1-j)/N-1]] + G[x(k-1),k-1]\Psi_w(k-1)G^T[x(k-1),k-1]$$

$$V_{\tilde{x}}(k-1-i,k-1-j/N-1] = V_{\tilde{x}}(k-1-i,k-1-j/N-2)$$

$$- K_{N-1}^{N-k+i} E_{N-2}\{\tilde{h}[x(N-1)/N-2]\tilde{x}^T(k-1-j/N-2)\}$$

and

$$V_{\tilde{x}}(k, N/N-1) = E_{N-1}\{\tilde{x}(k/N-1) \sum_{i=o}^{L} \tilde{f}_i^T[x(N-1-i)/N-1]\}$$

4.2.6 The Nonlinear Filtering

The nonlinear filtering algorithm can be easily obtained from Equations (4.2.10), (4.2.11), (4.2.16) - (4.2.19), by setting $\ell = 0$.

$$\hat{x}(k+1/k+1) = \hat{x}(k+1/k) + K^{o}_{k+1}\{y(k+1) - \tilde{h}[x(k+1), k+1/k]\}$$

$$\hat{x}(k+1/k) = \sum_{j=o}^{L} \hat{f}_j[x(k-j), k-j/k]$$

$$K^{o}_{k+1} = E_k\{\tilde{x}(k+1/k)\, \tilde{h}^T[x(k+1)/k]\}[\Psi_v(k+1)$$

$$+ E_k\{\tilde{h}[x(k+1)/k]\, \tilde{h}^T[x(k+1)/k]\}]^{-1}$$

and

$$V_{\tilde{x}}(k+1/k+1) = V_{\tilde{x}}(k+1/k) - K^{o}_{k+1}\{\tilde{h}[x(k+1)/k]\, \tilde{x}^T(k+1/k)\}$$

whereas

$V_{\tilde{x}}(k+1/k)$ is given by Equation (4.2.29).

4.2.7 *Estimation in Linear Discrete Systems*

In order to provide an insight into the structure of the smoothed estimate, consider the particular case of linear systems and measurements with

$$\sum_{j=o}^{L} f_j[x(k-j),k-j] = \sum_{j=o}^{L} F_j(k)\, x(k-j)$$

$$h[x(k),k] = H(k)\, x(k)$$

and

$$G[x(k),k] = G(k)$$

Then the linear fixed-lag smoothing algorithm can be directly obtained from Equations (4.2.23) to (4.2.28), respectively as the following

$$\hat{x}(k-N/k) = \hat{x}(k-N/k-1) + K_k^N\{y(k) - H(k)\,\hat{x}(k/k-1)\}$$

$$\hat{x}(k-N/k-1) = \sum_{j=o}^{L} F_j(k-N-1)\,\hat{x}(k-N-1-j/k-1)$$

$$K_k^N = V_{\tilde{x}}(k-N,k/k-1)\,H^T(k)\{H(k)\,V_{\tilde{x}}(k/k-1)\,H^T(k) + \Psi_v(k)\}^{-1}$$

$$V_{\tilde{x}}(k-N/k) = V_{\tilde{x}}(k-N/k-1) - K_k^N\,H(k)\,V_{\tilde{x}}(k,k-N/k-1)$$

$$V_{\tilde{x}}(k-N/k-1) = \sum_{i=o}^{L}\sum_{j=o}^{L} F_i(k-N-1)\,V_{\tilde{x}}(k-1-N-i,k-1-N-j/k-1)$$

$$F_j^T(k-N-1) + G(k-1-N)\,\Psi_w(k-1-N)\,G^T(k-1-N)$$

$$V_{\tilde{x}}(k-1-N-i,k-1-N-j/k-1) = V_{\tilde{x}}(k-1-N-i,k-1-N-j/k-2)$$

$$- K_{k-1}^{N+i}[H(k-1)V_{\tilde{x}}(k-1,k-1-N-j/k-2)]$$

Also $$V_{\tilde{x}}(k-N,k+1/k) = \sum_{j=o}^{L} V_{\tilde{x}}(k-N,k-j/k)\,F_j^T(k)$$

and finally,

$$V_{\tilde{x}}(k+1/k) = \sum_{i=o}^{L} \sum_{j=o}^{L} F_i(k)\, V_{\tilde{x}}(k-i,k-j/k)\, F_j^T(k)$$

$$+ G(k)\, \Psi_w(k)\, G^T(k)$$

It can be easily identified that the structure of the above linear fixed-lag smoother is simply the stable fixed-lag smoother introduced by Priemer and Vacroux [38]. In the same manner, one can identify the linear fixed-point smoother for linear systems without delays with that of Biswas and Mahalanabis [55]. Similarly, the results presented in Section 4.2.5 and 4.2.6 can be easily applied to yield the fixed-interval smoothing and the filtering algorithms, respectively.

Thus, a unified approach to obtain the filtering and smoothing algorithms for linear as well as nonlinear,delayed as well as non-delayed systems has been presented.

4.3 *NONLINEAR DISCRETE DELAYED SYSTEMS WITH MEASUREMENTS CONTAINING MULTIPLE DELAYS*

In this section, nonlinear smoothing algorithms are derived for discrete nonlinear delayed systems with multichannel measurements containing time delays, imbedded in additive white noise processes.

Section 4.3.1 presents the problem statement. Section 4.3.2 is presented with the derivation of the nonlinear smoother. Section 4.3.3 gives an illustrative example to show how one can easily obtain the fixed-lag, fixed-point, fixed-interval smoothing as well as the filtering algorithms. It also illustrates the use of the presented algorithms to arrive at physically realizable nonlinear estimators. Reference equations for various smoothing as well as filtering algorithms are not presented, since they can be easily obtained from Section 4.3.2 by properly substituting the discrete time indices.

4.3.1 *The Problem Statement*

Consider the discrete nonlinear message model of Equation 4.2.1 with the multi-channel measurements given by

$$y(k) = \sum_{j=o}^{N} h_j[x(k-j),\ k-j] + v(k) \tag{4.3.1}$$

where h_j is an m-dimensional nonlinear vector valued function. All other prior statistics and initial conditions follow that of Section 4.2.1.

The smoothing problem is to obtain the unbiased smoothing estimate $\hat{x}(k-\ell+1/k+1)$ of the delayed state $x(k-\ell+1)$, where $0 \leq \ell \leq k+1$, conditioned on the measurements

$$Y(k+1) = \{y(0),\ y(1),\ \ldots,\ y(k+1)\}$$

such that the cost function

$$J(k+1) = \text{Trace}[M(k)\ V_{\tilde{x}}(k-\ell+1/k+1)] \tag{4.3.2}$$

is minimized. Where $M(k)$ is also a non-negative definite weighting matrix.

Here, a different cost function is purposefully chosen. In outward appearance, the cost function of Equation (4.2.3) is different from that of (4.3.2). In real essence, the two cost functions are exactly the same, since $M(k)$ is known, it can be shown that minimizing Equation (4.2.3) is the same as minimizing (4.3.2), and is also equivalent to minimizing the trace of $V_{\tilde{x}}(k-\ell+1/k+1)$ [49].

4.3.2 *The Derivation of the Smoother*

With reasoning similar to that of Section 4.2.2, the smoothing algorithm is assumed to be constrained by the following dynamic equation*

$$\hat{x}(k-\ell+1/k+1) = \sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] + K_{k+1}^{\ell} y(k+1) \quad (4.3.3)$$

---

* It can be shown that other dynamic equations such as

$$\hat{x}(k-\ell+1/k+1) = \sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] + K_{k+1}^{\ell} \tilde{y}(k+1/k)$$

where $\tilde{y}(k+1/k) = y(k+1) - E[y(k+1)/Y(k)]$

would result in exactly the same smoothing algorithm.

Now the problem is to determine the time-varying nonlinear vector valued function $\sum_{j=0}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j]$ and the gain algorithm K_{k+1}^{ℓ} such that the trace of $V_{\tilde{x}}(k-\ell+1/k+1)$ is minimized.

Let $\tilde{x}(k-\ell+1/k+1)$ denote the smoothing error defined by

$$\tilde{x}(k-\ell+1/k+1) = x(k-\ell+1) - \hat{x}(k-\ell+1/k+1) \tag{4.3.4}$$

then one can obtain

$$\begin{aligned}\tilde{x}(k-\ell+1/k+1) = \sum_{j=0}^{L} f_j[x(k-\ell-j), k-\ell-j] + G[x(k-\ell), k-\ell] \\ \cdot w(k) - \sum_{j=0}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] \\ - K_{k+1}^{\ell} \{ \sum_{j=0}^{N} h_j[x(k-j+1), k-j+1] + v(k+1)\}\end{aligned}$$

In order that $\hat{x}(k-\ell+1/k+1)$ is an unbiased smoothing estimate, it is necessary that

$$\begin{aligned}\sum_{j=0}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] = \sum_{j=0}^{L} \hat{f}_j[x(k-\ell-j), k-\ell-j/k] \\ - K_{k+1}^{\ell} \sum_{j=0}^{N} \hat{h}_j[x(k-j+1), k-j+1/k]\end{aligned} \tag{4.3.5}$$

where

$$\hat{f}_j[x(k-\ell-j), k-\ell-j/k] = E_k\{f_j[x(k-\ell-j), k-\ell-j]\}$$

and

$$\hat{h}_j[x(k-j+1),\ k-j+1/k] = E_k\{h_j[x(k-j+1),\ k-j+1]\}$$

Substituting Equation (4.3.5) into (4.3.4), the smoothing algorithm becomes

$$\hat{x}(k-\ell+1/k+1) = \hat{x}(k-\ell+1/k) + K^{\ell}_{k+1}\{y(k+1) - \sum_{j=o}^{N} \hat{h}_j[x(k-j+1),k-j+1/k]\} \tag{4.3.6}$$

for $k = 0, 1, 2, \ldots,$ and $0 \leq \ell \leq k+1$,

where

$$\hat{x}(k-\ell+1/k) = \sum_{j=o}^{L} \hat{f}_j[x(k-\ell-j),\ k-\ell-j/k] \tag{4.3.7}$$

Substitution of Equation (4.3.6) into (4.3.4), yields

$$\tilde{x}(k-\ell+1/k+1) = \tilde{x}(k-\ell+1/k) - K^{\ell}_{k+1}\{\sum_{j=o}^{N} \tilde{h}_j[x(k-j+1)/k] + v(k+1)\} \tag{4.3.8}$$

where

$$\tilde{h}_j[x(k-j+1)/k] = h_j[x(k-j+1),\ k-j+1] - \hat{h}_j[x(k-j+1),k-j+1/k]$$

Therefore, the smoothing error-variance equation is given by

$$V_{\tilde{x}}(k-\ell+1/k+1) = V_{\tilde{x}}(k-\ell+1/k) - E_k\{\tilde{x}(k-\ell+1/k) \sum_{j=o}^{N} \tilde{h}^T_j[x(k-j+1)/k]\}$$

$$\cdot\ K_{k+1}^{\ell^T} - K_{k+1}^{\ell} E_k\{\sum_{j=o}^{N} \tilde{h}_j[x(k-j+1)/k]\ \tilde{x}^T(k-\ell+1/k)\}$$

$$+ K_{k+1}^{\ell}\ \Psi_v(k+1)\ K_{k+1}^{\ell^T} \qquad (4.3.9)$$

$$+ K_{k+1}^{\ell} E_k\{\sum_{i=o}^{N} \sum_{j=o}^{N} \tilde{h}_i[x(k-i+1)/k]\tilde{h}_j^T[x(k-j+1)/k]\}$$

$$\cdot\ K_{k+1}^{\ell^T}$$

The necessary condition for minimizing the trace of $[V_{\tilde{x}}(k-\ell+1/k+1)]$ is provided by the matrix minimum principle [7], for which K_{k+1}^{ℓ} is considered as the control variable, and is obtained from the condition

$$\frac{\partial}{\partial K_{k+1}^{\ell}}\ \text{Trace}[V_{\tilde{x}}(k-\ell+1/k+1)] = [0]$$

Hence, the gain algorithm is given by

$$K_{k+1}^{\ell} = E_k\{\tilde{x}(k-\ell+1/k) \sum_{j=o}^{N} \tilde{h}^T[x(k-j+1)/k]\}$$

$$[\Psi_v(k+1) + E_k\{\sum_{i=o}^{N} \sum_{j=o}^{N} \tilde{h}_i[x(k-i+1)/k] \qquad (4.3.10)$$

$$\tilde{h}_j^T[x(k-j+1)/k]\}]^{-1}$$

and Equation (4.3.9) is reduced to

$$V_{\tilde{x}}(k-\ell+1/k+1) = V_{\tilde{x}}(k-\ell+1/k) \qquad (4.3.11)$$

$$- K^{\ell}_{k+1} E_k\{\sum_{j=o}^{N} \tilde{h}_j[x(k-j+1)/k]\ \tilde{x}^T(k-\ell+1/k)\}$$

Using Equations (4.2.1) and (4.3.7), there is the relation

$$\tilde{x}(k-\ell+1/k) = \sum_{j=o}^{L} \tilde{f}_j[x(k-\ell-j)/k] + G[x(k-\ell),\ k-\ell]\ w(k-\ell)$$

where

$$\tilde{f}_j[x(k-\ell-j)/k] = f_j[x(k-\ell-j),\ k-\ell-j] - \hat{f}_j[x(k-\ell-j),\ k-\ell-j/k] \quad .$$

Then the following recursive smoothing algorithms can also be derived

$$V_{\tilde{x}}(k-\ell+1/k) = \sum_{i=o}^{L} \sum_{j=o}^{L} E_k\{\tilde{f}_i[x(k-\ell-i)/k]\ \tilde{f}_j^T[x(k-\ell-j)/k]\}$$

$$+ G[x(k-\ell),\ k-\ell]\ \Psi_w(k-\ell)\ G^T[x(k-\ell),\ k-\ell] \qquad (4.3.12)$$

and

$$V_{\tilde{x}}(k-\ell+1,\ k-m+1) = V_{\tilde{x}}(k-\ell+1,\ k-m+1/k) \qquad (4.3.13)$$

$$- K^{\ell}_{k+1}\ E_k\{\sum_{j=o}^{N} \tilde{h}_j[x(k-j+1)/k]\tilde{x}^T(k-m+1/k)\}$$

for $0 \leq \ell,\ m \leq k+1$.

For $k < 0$, there are no inputs to the smoother and therefore $\hat{x}(-j/-1)$ is set to zero for $j = 0, 1, \ldots, L$, which in turn leads to

$$\tilde{x}(-j/-1) = x(-j)$$

and

$$V_{\tilde{x}}(-j, -\ell/-1) = V_x(j,\ell)$$

for $j, \ell = 0, 1, \ldots, L$.

As shown in Section 4.2, the fixed-lag, fixed-point, fixed-interval smoothing and also the filtering algorithms together with their error-variance algorithms follow immediately from Equations (4.3.6), (4.3.10) to (4.3.13) by properly substituting the indices k and ℓ. And in the particular case of linear message and measurement models, the results presented for the fixed-lag smoothing estimate can be shown to be equivalent to those of Shukla and Srinath[56], whereas for linear problems with delays only appearing in the message model, the above algorithms agree well with those of Priemer and Vacroux [37,38].

Also notice that in the case of nonlinear systems, the preceeding algorithms require infinite dimensional systems to realize the various terms involving the expextations.

In order that the smoother can be physically realized, it is assumed that the conditional probability density functions of the smoothing errors are Gaussian, and Equation (4.3.4) is used to substitute for the state variables by the sum of their estimators and smoothing errors. This is particularly significant for systems with product types or polynomial nonlinearities, since in such cases, the smoothing algorithms can be evaluated without much difficulty and are physically realizable without any further approximation.

On the other hand, the majority of other finite dimensional filters require the Taylor's series expansion for the system nonlinearities, in general, they lead to the deletion of some of the higher order terms of the estimators and the error-covariances.

Moreover, the smoothing estimators are already required to be unbiased, therefore, the expectations that are in Equations (4.3.10) to (4.3.13) can be further simplified as

$$E_k\{\tilde{x}(k-\ell+1/k) \sum_{j=o}^{N} \tilde{h}_j^T[x(k-j+1)/k]\}$$

$$= E_k\{\tilde{x}(k-\ell+1/k) \sum_{j=o}^{N} h_j^T[x(k-j+1),\ k-j+1]\} \qquad (4.3.14)$$

$$E_k\{\sum_{i,j=o}^{N} \tilde{h}_i[x(k-i+1)/k]\ \tilde{h}_j^T[x(k-j+1)/k]\}$$

$$= E_k\{\sum_{i,j=o}^{N} h_i[x(k-i+1),\ k-i+1]\ h_j^T[x(k-j+1),\ k-j+1]\}$$

$$- \sum_{i,j=o}^{N} \hat{h}_i[x(k-i+1),\ k-i+1/k]\ \hat{h}_j^T[x(k-j+1),\ k-j+1/k] \qquad (4.3.15)$$

and

$$E_k\{\sum_{i,j=o}^{L} \tilde{f}_i[x(k-\ell-i)/k]\ \tilde{f}_j^T[x(k-\ell-j)/k]\}$$

$$= E_k\{\sum_{i,j=o}^{L} f_i[x(k-\ell-i),\ k-\ell-i]\ f_j^T[x(k-\ell-j),\ k-\ell-j]\}$$

$$- \sum_{i,j=o}^{L} \hat{f}_i[x(k-\ell-i),\ k-\ell-i/k]\ \hat{f}_j^T[x(k-\ell-j),\ k-\ell-j/k] \quad (4.3.16)$$

4.3.3 *Illustrative Example and Discussion*

In order to illustrate the use of the preceeding algorithms consider the following scalar nonlinear smoothing problem

$$x(k+1) = x^3(k) + 2x(k-1) + w(k)$$
$$y(k) = x^2(k) + x(k-1) + v(k)$$

where $w(k)$ and $v(k)$ are independent Gaussian noise processes with variances $\Psi_w(k)$ and $\Psi_v(k)$, respectively. The initial conditions $x(0)$ and $x(-1)$ are zero-mean Gaussian with respective moments $\hat{x}(0/0)$ and $\hat{x}(-1/0)$, and are independent of both $v(k)$ and $w(k)$, and also

$$E\{x(-j)\ x^T(-\ell)\} = V_{\tilde{x}}(j,\ell)$$

for $j,\ \ell = 0,\ 1$.

The expectations of Equations (4.3.14) to (4.3.16) can be obtained as follows,respectively

$$E_k\{\tilde{x}(k-\ell+1/k)\,[x^2(k+1) + x(k)]\}$$

$$= E_k\{\tilde{x}(k-\ell+1/k)\,[\tilde{x}^2(k+1/k) + 2\tilde{x}(k+1/k)\,\hat{x}(k+1/k)$$

$$+ \hat{x}^2(k+1/k) + \tilde{x}(k) + \hat{x}(k)]\}$$

$$= 2\,V_{\tilde{x}}(k-\ell+1,\ k+1/k)\,\hat{x}(k+1/k) + V_{\tilde{x}}(k-\ell+1,\ k/k). \qquad (4.3.17)$$

$$E_k\{[x^2(k+1) + x(k)]^2\} - [E_k\{x^2(k+1) + x(k)\}]^2$$

$$= E_k\{x^4(k+1) + 2\,x^2(k+1)\,x(k) + x^2(k)\}$$

$$- [V_{\tilde{x}}(k+1/k) + \hat{x}^2(k+1/k) + \hat{x}(k)]^2$$

$$= 3\,V^2_{\tilde{x}}(k+1/k) + 6\,\hat{x}^2(k+1/k)\,V_{\tilde{x}}(k+1/k) + \hat{x}^4(k+1/k)$$

$$+ 2[2V_{\tilde{x}}(k+1,k/k)\,\hat{x}(k+1/k) + V_{\tilde{x}}(k+1/k)\,\hat{x}(k) + \hat{x}^2(k+1/k)$$

$$\cdot\,\hat{x}(k)] + V_{\tilde{x}}(k) + \hat{x}^2(k) - [V_{\tilde{x}}(k+1/k) + \hat{x}^2(k+1/k) + \hat{x}(k)]^2$$

$$= 2\,V^2_{\tilde{x}}(k+1/k) + 4\,V_{\tilde{x}}(k+1/k)\,\hat{x}^2(k+1/k)$$

$$+ 4\,V_{\tilde{x}}(k+1,k/k)\,\hat{x}(k+1/k) + V_{\tilde{x}}(k) \qquad (4.3.18)$$

and:

$$E_k\{[x^3(k-\ell) + 2\,x(k-\ell-1)]^2\} - [E_k\{x^3(k-\ell) + 2\,x(k-\ell-1)\}]^2$$

$$= E_k\{x^6(k-\ell) + 4\,x^3(k-\ell)\,x(k-\ell-1) + 4\,x^2(k-\ell-1)\}$$

$$- [3\,V_{\tilde{x}}(k-\ell/k)\,\hat{x}(k-\ell/k) + \hat{x}^3(k-\ell/k) + 2\,\hat{x}(k-\ell-1/k)]^2$$

$$= 15\,V_{\tilde{x}}^3(k-\ell/k) + 36\,V_{\tilde{x}}^2(k-\ell/k)\,\hat{x}^2(k-\ell/k)$$

$$+ 9\,V_{\tilde{x}}(k-\ell/k)\,\hat{x}^4(k-\ell/k) + 12\,V_{\tilde{x}}(k-\ell,k-\ell-1/k)$$

$$+ V_{\tilde{x}}(k-\ell/k) + 12\,V_{\tilde{x}}(k-\ell,\,k-\ell-1/k)\,\hat{x}^2(k-\ell/k) \qquad (4.3.19)$$

$$+ 4\,V_{\tilde{x}}(k-\ell-1/k)$$

Then Equations (4.3.6), (4.3.10) to (4.3.13) can be obtained with the substitution of Equations (4.3.17) to (4.3.19). In particular when ℓ=0, Equations (4.3.6), (4.3.10) to (4.3.13) directly yield the nonlinear filtering algorithms. Namely,

$$\hat{x}(k+1/k+1) = \hat{x}(k+1/k) + K_{k+1}^0\{y(k+1) - V_{\tilde{x}}(k+1/k)$$

$$- \hat{x}^2(k+1/k) - \hat{x}(k)\}$$

where

$$\hat{x}(k+1)/k) = 3V_{\tilde{x}}(k)\ \hat{x}(k) + \hat{x}^3(k/k) + 2\ \hat{x}(k-1/k)$$

$$K^0_{k+1} = [2V_{\tilde{x}}(k+1/k)\ \hat{x}(k+1/k) + V_{\tilde{x}}(k+1,\ k/k)]$$

$$\cdot\ [\Psi_v(k+1) + 2V^2_{\tilde{x}}(k+1/k) + 4V_{\tilde{x}}(k+1/k)\ \hat{x}^2(k+1/k)$$

$$+ 4V_{\tilde{x}}(k+1,\ k/k)\ \hat{x}(k+1/k) + V_{\tilde{x}}(k)]^{-1}$$

Also,

$$V_{\tilde{x}}(k+1/k+1) = V_{\tilde{x}}(k+1/k) - K^0_{k+1}[2V_{\tilde{x}}(k+1/k)\ \hat{x}(k+1/k)$$

$$+ V_{\tilde{x}}(k+1,k/k)]$$

and

$$V_{\tilde{x}}(k+1/k) = 15V^3_{\tilde{x}}(k) + 36V^2_{\tilde{x}}(k)\ \hat{x}^2(k-\ell/k)$$

$$+ 9V_{\tilde{x}}(k)\ \hat{x}^4(k) + 12V_{\tilde{x}}(k,k-1/k)\ V_{\tilde{x}}(k)$$

$$+ 12V_{\tilde{x}}(k,k-1/k)\ \hat{x}^2(k/k) + 4V_{\tilde{x}}(k-1/k) + \Psi_w(k)$$

The fixed-lag smoothing algorithms are obtained from Equations (4.3.8), (4.3.10) to (4.3.13), with the substitution k and N for k+1 and ℓ, respectively.

$$\hat{x}(k-N/k) = \hat{x}(k-N/k-1) + K_k^N\{y(k) - V_{\tilde{x}}(k/k-1) - \hat{x}^2(k/k-1) - \hat{x}(k-1)\}$$

where

$$\hat{x}(k-N/k-1) = 3V_{\tilde{x}}(k-1-N/k-1)\,\hat{x}(k-1-N/k-1) + \hat{x}^3(k-1-N/k-1) + 2\,\hat{x}(k-2-N/k-1)$$

and

$$K_k^N = [2V_{\tilde{x}}(k-N,k/k-1)\,\hat{x}(k/k-1) + V_{\tilde{x}}(k-N,k-1/k-1)] \cdot [\Psi_V(k) + 2V_{\tilde{x}}^2(k/k-1) + 4V_{\tilde{x}}(k/k-1)\,\hat{x}^2(k/k-1) + 4V_{\tilde{x}}(k,k-1/k-1)\,\hat{x}(k/k-1) + V_{\tilde{x}}(k-1)]^{-1}$$

Also

$$V_{\tilde{x}}(k-N/k) = V_{\tilde{x}}(k-N/k-1) - K_k^N[2V_{\tilde{x}}(k-N,k/k-1)\,\hat{x}(k/k-1) + V_{\tilde{x}}(k-N,\ k-1/k-1)]$$

In a similar manner, the fixed-interval as well as the fixed-point smoothing algorithms can be obtained from Equations (4.3.6), (4.3.10) to (4.3.13), (4.3.17) to (4.3.19) by properly substituting the indices k and ℓ. As a consequence, the effectiveness of the proposed algorithms is illustrated.

Notice that in the above evaluation, the only assumption needed is the Gaussian assumption of the estimator error. And it can be easily shown that if the system nonlinearities were expanded in terms of the first or second-order Taylor's series, the higher order terms of the error-covariances as well as the estimators would have been dropped. Which when applied to estimation problems with relatively large initial error covariances, may lead to erroneous results.

To qualitatively assess the performance characteristics of the proposed nonlinear estimation algorithms, extensive numerical simulation and comparative studies must be carried out to compare the above estimators with those of other finite dimensional filters, however, this is not included in this section.

4.4 *CONCLUSION*

In this chapter, discrete estimation algorithms are derived for nonlinear multiple delayed systems with single-channel non-delayed measurement as well as measurements with multiple delays, imbedded in additive white noise processes. The results are new, recursive in nature and directly yield the fixed-interval, fixed-point, fixed-lag smoothing and the filtering algorithms. The derivation is straightforward and clearly shows the close link between the smoothing and filtering estimation algorithms.

The presented algorithms are physically realizable under the assumption that the probability density functions of the estimator errors are Gaussian, they can be obtained without any further approximation for systems with product types or polynomial nonlinearities. They are expected to be computationally faster than applying an extended

Kalman filter to an augmented system with no explicit delay times.

In the particular case of linear system models, the smoothing algorithms are stable, and for the nonlinear case the stability behavior needs to be further investigated; but it is easy to observe that the basic structure of the nonlinear estimator is the same as the linear counterpart, hence there is a good reason to believe that the corresponding nonlinear estimator will be stable provided the original message model is also stable.

The results derived in this thesis can be extended to continuous time problems through a formal limiting procedure [4]. The presented approach can also be extended to nonlinear distributive systems with or without delays.

CHAPTER V

MINIMUM VARIANCE FILTERING AND SMOOTHING FOR NONLINEAR SYSTEMS WITH CORRELATED NOISES

5.1 *INTRODUCTION*

In Chapter IV, the noise processes considered are assumed to be Gaussian white and mutually independent, however, in practical situations such assumptions are invalid, since in physical systems, independent white noise processes simply do not exist. Therefore, it is natural to extend the new estimation technique developed in Chapter IV to more realistic problems, where message noise processes are correlated with measurement noise processes.

The treatment of linear estimation problems for systems without delays imbedded in correlated noise processes can be handled by the uncorrelated algorithms if one can uncorrelate the message and measurement noise processes. Recently, Raja Rao and Mahalanabis [57] derived estimation algorithms for linear systems with delay imbedded in correlated noise processes, however, their results appear to have a fundamental mistake in the procedure given, which leads to self-contradictory results; this is reported in [58].

To the author's knowledge estimation algorithms have not been derived for nonlinear systems with or without delays imbedded in correlated noise processes, this is due to the inherent difficulty in finding the conditional-mean estimate equation and the associated error-variance, when the noise processes are correlated.

In Section 5.2, new recursive estimation algorithms are derived for discrete nonlinear multiple delays systems with non-delayed measurements imbedded in correlated noises. Handy sets of reference equations are provided for each specific case of smoothing and also the filtering estimation; general linear estimation algorithms are also provided.

In Section 5.3, new recursive estimation algorithms are derived for discrete nonlinear delayed systems with multi-channel measurements involving time delays and imbedded in correlated noise processes. An example is given to illustrate the use of the proposed smoothing algorithm, and also derived are the general linear estimation algorithms.

In this chapter, the basic technique makes use of the matrix minimum principle to derive the optimal values of the coefficients in the estimation algorithms under the requirements that the estimates be unbiased and the error-variance, which is taken to be the estimation criterion, is minimized; they are reported in [59,60].

5.2 *NONLINEAR DISCRETE DELAYED SYSTEMS WITH NON-DELAYED MEASUREMENTS IMBEDDED IN CORRELATED NOISE PROCESSES.*

In this section, estimation algorithms are derived for non-linear discrete delayed systems with non-delayed measurements imbedded in correlated noise processes.

Section 4.2.1 presents the problem statement, Section 4.2.2 deals with the derivation of the smoother. In Section 4.2.3 to 4.2.6, non-linear estimators together with their error-variance equations are obtained for the fixed-lag, fixed-point, fixed-interval smoothing and the filtering estimation, respectively. Section 4.2.7 presents the general smoothing algorithms for linear discrete delayed systems, which can be easily converted to yield the fixed-lag, fixed-point, fixed-interval smoothing and the filtering algorithms, by properly substituting the time indices.

5.2.1 *The Problem Statement*

The state equation of the discrete nonlinear time-delayed systems under consideration is described by

$$x(k+1) = \sum_{j=o}^{L} f_j[x(k-j), k-j] + G[x(k), k]\, w(k) \tag{5.2.1}$$

and the measurement is given by

$$y(k) = h[x(k), k] + v(k) \tag{5.2.2}$$

where $k = 0, 1, 2, \ldots,$ is the discrete time index, the state x is an n-vector; the measurement y an m-vector; the state noise sequence w an r-vector; the measurement noise v an m-vector; G, a nonlinear state dependent nxr matrix. The nonlinear vector valued functions f_j and h are, respectively, n and m dimensional.

The noise sequences w and v, are zero-mean white Gaussian with non-negative definite covariance Ψ_w and positive definite Ψ_v, respectively. Also

$$E\{v(k)\, w^T(j)\} = \Psi_{vw}(k)\, \delta_{k,j}$$

for all integers k and j, where E denotes the expected value, $\delta_{k,j}$ the Kronecker delta, and Ψ_{vw} is non-negative definite.

The initial states x(0) and x(-j), for j = 1, ..., L are zero-mean Gaussian random vectors, which are independent of v(k) and w(k), with a positive definite covariance matrix

$$E\{x(-j)\, x^T(-\ell)\} = V_x(j,\ell)$$

for j, ℓ = 0, 1, ..., L.

The smoothing problem is to obtain the unbiased smoothed estimate $\hat{x}(k-\ell+1/k+1)$ of the state $x(k-\ell+1)$, where $0 \leq \ell \leq k+1$, conditioned on the set of measurements

$$Y(k+1) = \{y(0), y(1), \ldots, y(k+1)\}$$

such that the cost function

$$J(k+1) = \mathrm{Trace}[M(k)\, V_{\tilde{x}}(k-\ell+1/k+1)] \tag{5.2.3}$$

is a minimum. Here M(k) is some non-negative definite weighting matrix, and $V_{\tilde{x}}(k-\ell+1/k+1)$ is defined by

$$V_{\tilde{x}}(k-\ell+1/k+1) = E_{k+1}[\{x(k-\ell+1) - \hat{x}(k-\ell+1/k+1)\}$$

$$\{x(k-\ell+1) - \hat{x}(k-\ell+1/k+1)\}^T] \qquad (5.2.4)$$

where $E_{k+1}\{\cdot\}$ denotes the expectation operation conditioned on the set of measurements Y(k+1).

5.2.2 *The Derivation of the Smoother*

With reasoning similar to that of Section 4.2.2, the smoothed estimate is assumed to be constrained by the nonlinear dynamic equation*

$$\hat{x}(k-\ell+1/k+1) = \sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] + K_{k+1}^{\ell}\, y(k+1) \qquad (5.2.5)$$

---

* It can be shown that when other dynamic equations such as

$$\hat{x}(k-\ell+1/k+1) = \sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] + K_{k+1}^{\ell}\{y(k+1) - E_k[y(k+1)]\}$$

is assumed, the resulting smoothing estimator would be exactly the same as those using Equation (5.2.5).

where $\sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k),k-\ell-j]$ and K_{k+1}^{ℓ} are yet to be determined.

Let $\tilde{x}(-\ell+1/k+1)$ denotes the smoothing error defined by

$$\tilde{x}(k-\ell+1/k+1) = x(k-\ell+1) - \hat{x}(k-\ell+1/k+1)$$

then we obtain,

$$\tilde{x}(k-\ell+1/k+1) = \sum_{j=o}^{L} f_j[x(k-\ell-j),\ k-\ell-j] + G[x(k-\ell),\ k-\ell] \cdot w(k-\ell) - \sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k),\ k-\ell-j] - K_{k+1}^{\ell}\{h[x(k+1),\ k+1] + v(k+1)\} \tag{5.2.6}$$

Since it is required that the smoothed estimate be unbiased, it is necessary that

$$\sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k),\ k-\ell-j] = \sum_{j=o}^{L} \hat{f}_j[x(k-\ell-j),\ k-\ell-j/k] - K_{k+1}^{\ell}\hat{h}[x(k+1),\ k+1/k] \tag{5.2.7}$$

where

$$\hat{f}_j[x(k-\ell-j),\ k-\ell-j/k] = E_k\{f_j[x(k-\ell-j),\ k-\ell-j]\}$$

and

$$\hat{h}[x(k+1),\ k+1/k] = E_k\{h[x(k+1),\ k+1]\}$$

With Equation (5.2.7) substituted into (5.2.5), the smoothed estimate becomes

$$\hat{x}(k-\ell+1/k+1) = \hat{x}(k-\ell+1/k) + K_{k+1}^{\ell}\{y(k+1) - \hat{h}[x(k+1),\ k+1/k]\} \quad (5.2.8)$$

for k = 0, 1, 2, ..., and $0 \leq \ell \leq k+1$.

Where

$$\hat{x}(k-\ell+1/k) = \sum_{j=0}^{L} \hat{f}_j[x(k-\ell-j),\ k-\ell-j/k] \quad (5.2.9)$$

Equation (5.2.6) is now simplified as

$$\tilde{x}(k-\ell+1/k+1) = \tilde{x}(k-\ell+1/k) - K_{k+1}^{\ell}\{\tilde{h}[x(k+1)/k] + v(k+1)\} \quad (5.2.10)$$

with

$$\tilde{h}[x(k+1)/k] = h[x(k+1),\ k+1] - \hat{h}[x(k+1),\ k+1/k]$$

Then, the error-variance equation is given as

$$V_{\tilde{x}}(k-\ell+1/k+1) = V_{\tilde{x}}(k-\ell+1/k) - K_{k+1}^{\ell}E_k\{\tilde{h}[x(k+1)/k]$$

$$\cdot\ \tilde{x}^T(k-\ell+1/k)\} - E_k\{\tilde{x}(k-\ell+1/k)\ \tilde{h}^T[x(k+1)/k]\}$$

$$\cdot\ (K_{k+1}^{\ell})^T + K_{k+1}^{\ell}\ \Psi_v(k+1)(K_{k+1}^{\ell})^T$$

$$+ K_{k+1}^{\ell} E_k\{\tilde{h}[x(k+1)/k]\ \tilde{h}^T[x(k+1)/k]\}(K_{k+1}^{\ell})^T \tag{5.2.11}$$

The necessary condition for minimizing the trace of $V_{\tilde{x}}(k-\ell+1/k+1)$, can now be obtained from the condition

$$\frac{\partial}{\partial K_{k+1}^{\ell}} \text{Trace}[V_{\tilde{x}}(k-\ell+1/k)] = [0]$$

Hence, the gain algorithm K_{k+1}^{ℓ} is given by

$$K_{k+1}^{\ell} = E_k\{\tilde{x}(k-\ell+1/k)\ \tilde{h}^T[x(k+1)/k]\}$$

$$[\Psi_v(k+1) + E_k\{\tilde{h}[x(k+1)/k]\ \tilde{h}^T[x(k+1)/k]\}]^{-1} \tag{5.2.12}$$

and Equation (5.2.11) becomes

$$V_{\tilde{x}}(k-\ell+1/k+1) = V_{\tilde{x}}(k-\ell+1/k) - K_{k+1}^{\ell} E_k\{\tilde{h}[x(k+1)/k]$$

$$\tilde{x}^T(k-\ell+1/k)\} \tag{5.2.13}$$

and if the matrix $V_{\tilde{x}}(k-\ell+1,\ k-m+1/k+1)$ is defined as

$$V_{\tilde{x}}(k-\ell+1,\ k-m+1/k+1) = E_{k+1}\{\tilde{x}(k-\ell+1/k+1)\tilde{x}^T(k-m+1/k+1)\}$$

then

$$V_{\tilde{x}}(k-\ell+1,\ k-m+1/k+1) = V_{\tilde{x}}(k-\ell+1,\ k-m+1/k)$$
$$-K_{k+1}^{\ell}E_k\{\tilde{h}[x(k+1)/k]\tilde{x}^T(k-m+1/k)\} \quad (5.2.14)$$

for $0 \leq \ell,\ m \leq k+1$.

Now, there remains the problem of evaluating $V_{\tilde{x}}(k-\ell+1/k)$ in Equation (5.2.13).

Even though substracting Equation (5.2.9) from Equation (5.2.1) would easily yield

$$\tilde{x}(k-\ell+1/k) = \sum_{j=o}^{L} \tilde{f}_j[x(k-\ell-j)/k] + G[x(k-\ell),\ k-\ell]\ w(k-\ell) \quad (5.2.15)$$

where it is defined that

$$\tilde{f}_j[x(k-\ell-j)/k] = f_j[x(k-\ell-j),\ k-\ell-j]$$
$$-\ \hat{f}_j[x(k-\ell-j),\ k-\ell-j/k]$$

it is seen that the error-variance equation can not be obtained by taking the expectation of Equation (5.2.15) multiplied by its own transpose, since the expectation of

$$\mathcal{E}_k\{\sum_{j=0} \tilde{f}_j[x(k-\ell-j)/k]\, w^T(k-\ell)\}$$

can not be explicitly evaluated.

On the other hand, following the estimation technique presented above, the unbiased estimate of the state $x(k-\ell+1)$ can be obtained as

$$\hat{x}(k-\ell+1/k) = \sum_{j=0}^{L} \hat{f}_j[x(k-\ell-j),\ k-\ell-j/k-1]$$
$$+ \kappa_k^{\ell}\{y(k) - \hat{h}[x(k),\ k/k-1]\}$$

which in turn leads to

$$\tilde{x}(k-\ell+1/k) = \sum_{j=0}^{L} \tilde{f}_j[x(k-\ell-j),\ k-\ell-j/k-1] \tag{5.2.16}$$
$$- \kappa_k^{\ell}\{\tilde{h}[x(k)/k-1] + v(k)\} + G[x(k-\ell),k-\ell]w(k-\ell)$$

The optimal value of the matrix κ_k^{ℓ} can be determined from the necessary condition that

$$\frac{\partial}{\partial \kappa_k^{\ell}} \text{ trace } [M(k)\ V_{\tilde{x}}(k-\ell+1/k)] = [0] \tag{5.2.17}$$

Since $V_{\tilde{x}}(k-\ell+1/k)$ is the expectation of Equation (5.2.16) multiplied by its own transpose, then using Equation (5.2.19), the result is

$$\kappa_k^\ell = [E_{k-1}\{\sum_{j=o}^{L} \tilde{f}_j[x(k-\ell-j)/k-1]\ \tilde{h}^T[x(k)/k-1]\}$$

$$+ G[x(k-\ell),\ k-\ell]\ \Psi_{wv}(k)\ \delta_{k,k-\ell}] \qquad (5.2.18)$$

$$\cdot\ [\Psi_v(k) + E_{k-1}\{\tilde{h}[x(k)/k-1]\ \tilde{h}^T[x(k)/k-1]\}]^{-1}$$

and $V_{\tilde{x}}(k-\ell+1/k)$ is simply

$$V_{\tilde{x}}(k-\ell+1/k) = \sum_{i,j=o}^{L} E_{k-1}\{\tilde{f}_i[x(k-\ell-i)/k-1]\ \tilde{f}_j^T[x(k-\ell-j)/k-1]\}$$

$$+ G[x(k-\ell),k-\ell]\Psi_w(k-\ell)G^T[x(k-\ell),k-\ell] \qquad (5.2.19)$$

$$- \kappa_k^\ell[\Psi_{vw}(k-\ell)G^T[x(k-\ell),k-\ell]\ \delta_{k,k-\ell}$$

$$+ E_{k-1}\{\tilde{h}[x(k)/k-1]\ \sum_{j=o}^{L} \tilde{f}_j^T[x(k-\ell-j)/k-1]\}]$$

also

$$V_{\tilde{x}}(k-\ell+1,\ k-m+1/k) = \sum_{i,j=o}^{L} E_{k-1}\{\tilde{f}_i[x(k-\ell-i)/k-1]\tilde{f}_j^T[x(k-m-j)/k-1]\}$$

$$+ G[x(k-\ell),k-\ell]\Psi_w(k-\ell)\delta_{k-\ell,k-m}G^T[x(k-m),k-m]$$

$$- \kappa_k^{\ell}[\Psi_{vw}(k-m)\delta_{k,k-m} G^T[x(k-m),k-m] \quad (5.2.20)$$

$$+ E_{k-1}\{\tilde{h}[x(k)/k-1] \sum_{j=0}^{L} \tilde{f}_j^T[x(k-m-j)/k-1]\}]$$

for $0 \leq \ell, m \leq k+1$.

When the discrete time index $k<0$, there is no input to the smoother and therefore $\hat{x}(-j/-1)$ is set to zero for $j = 0, 1, \ldots, L$, which results in

$$\tilde{x}(-j/-1) = x(-j)$$

and

$$V_{\tilde{x}}(-j,-\ell/-1) = V_x(j,\ell)$$

for $j, \ell = 0, 1, \ldots, L$.

On the other hand, since the smoothing estimator is unbiased, the expectations that are in Equations (5.2.12) to (5.2.14), (5.2.18) to (5.2.20) can be replaced by Equations (4.2.20) to (4.2.22) and also

$$E_{k-1}\{\tilde{f}_i[x(k-\ell-i)/k-1] \tilde{f}_j[x(k-m-j)/k-1]\} \quad (5.2.21)$$

$$= E_{k-1}\{f_i[x(k-\ell-i),k-\ell-i] f_j^T[x(k-m-j), k-m-j]\}$$

$$- \hat{f}_i[x(k-\ell-i), k-\ell-i/k-1]\ \hat{f}_j^T[x(k-m-j), k-m-j/k-1]$$

It must be noted that in the case of nonlinear systems the above expectations require infinite dimensional systems to realize.

Therefore, as an approximation, it is assumed that the conditional probability density functions of the smoothing error $\tilde{x}$ are Gaussian. It is important to note that under such an assumption, the algorithms presented in this chapter can be physically realized without any further approximation for systems with product types or polynomial nonlinearities.

The estimation algorithms presented here are not only new and recursive in nature, but also directly yield the fixed-lag, fixed-point, fixed-interval smoothing as well as filtering algorithms, this is done by the proper substitution of indices ℓ and m.

5.2.3 *The Nonlinear Fixed-Lag Smoothing*

When k and N are substituted in place of k+1, and ℓ, into Equations (5.2.8), (5.2.9), (5.2.12) and (5.2.13) and also Equations (5.2.18) to (5.2.20), where k>N, the results obtained are the recursive nonlinear fixed-lag smoothing algorithms; they are respectively, as follows

$$\hat{x}(k-N/k) = \hat{x}(k-N/k-1) + K_k^N\{y(k) - \hat{h}[x(k), k/k-1]\}$$

where

$$\hat{x}(k-N/k) = \sum_{j=0}^{L} \hat{f}_j[x(k-1-N-j),k-1-N-j/k]$$

$$K_k^N = E_{k-1}\{\tilde{x}(k-N/k-1)\ \tilde{h}^T[x(k)/k-1]\}$$

$$[\Psi_v(k) + E_{k-1}\{\tilde{h}[x(k)/k-1]\ \tilde{h}^T[x(k)/k-1]\}]^{-1}$$

$$V_{\tilde{x}}(k-N/k) = V_{\tilde{x}}(k-N/k-1) - K_k^N E_{k-1}\{\tilde{h}[x(k)/k-1]$$

$$\cdot\ \tilde{x}^T(k-N/k-1)\}$$

$$\kappa_{k-1}^N = E_{k-2}\{\sum_{j=0}^{L} \tilde{f}_j[x(k-1-N-j)/k-2]\ \tilde{h}^T[x(k-1)/k-2]\}$$

$$+ G[x(k-1-N),\ k-1-N]\ \Psi_{wv}(k-1)\ \delta_{k-1,k-1-N}]$$

$$\cdot\ [\Psi_v(k-1) + E_{k-2}\{\tilde{h}[x(k-1)/k-2]\ \tilde{h}^T[x(k-1)/k-2]\}]^{-1}$$

$$V_{\tilde{x}}(k-N/k-1) = \sum_{i,j=0}^{L} E_{k-2}\{\tilde{f}_i[x(k-1-N-i)/k-2]$$

$$\cdot\ \tilde{f}_j^T[x(k-1-N-j)/k-2]\} + G[x(k-1-N),k-1-N]$$

$$\cdot\ \Psi_w(k-1-N)G^T[x(k-1-N),k-1-N] - \kappa_{k-1}^N[\Psi_{vw}(k-1-N)$$

$$\cdot\ G^T[x(k-1-N),k-1-N]\delta_{k-1,k-1-N} + E_{k-2}$$

$$\{\tilde{h}[x(k-1)/k-2]$$

$$\cdot \sum_{j=o}^{L} \tilde{f}_j^T[x(k-1-N-j)/k-2]\}]$$

$$V_{\tilde{x}}(k-\ell,k-m/k-1) = \sum_{i,j=o}^{L} E_{k-2}\{\tilde{f}_i[x(k-\ell-i-1)/k-2]$$

$$\tilde{f}_j^T[x(k-m-j-1)/k-2]\}+G[x(k-1-\ell),k-\ell-1]\Psi_w(k-1-\ell)$$

$$\delta_{k-1-\ell,k-1-m}G^T[x(k-1-m),k-1-m]$$

$$- K_{k-1}^{\ell}[\Psi_{vw}(k-1-m)\delta_{k-1,k-1-m}G^T[x(k-1-m),k-1-m]$$

$$+ E_{k-2}\{\tilde{h}[x(k-1)/k-2] \sum_{j=o}^{L} \tilde{f}_j^T[x(k-1-m-j)/k-2]\}] \tag{5.2.22}$$

5.2.4 *The Nonlinear Fixed-Point Smoothing*

When $\ell = k - N+1$ (where $k+1>N$) is substituted into Equations (5.2.8), (5.2.9) - (5.2.12), (5.2.13) and (5.2.18) - (5.2.19), then the recursive fixed-point smoother is given by the following

$$\hat{x}(N/k+1) = \hat{x}(N/k) + K_{k+1}^{k-N+1} \{y(k+1)$$

$$- \hat{h}[x(k+1),k+1/k]\}$$

$$\hat{x}(N/k) = \sum_{j=o}^{L} \hat{f}_j[x(N-1-j), N-1-j/k]$$

$$K_{k+1}^{k-N+1} = E_k\{\tilde{x}(N/k)\,\tilde{h}^T[x(k+1)/k]\}$$

$$[\Psi_v(k+1) + E_k\{\tilde{h}[x(k+1)/k]\,\tilde{h}^T[x(k+1)/k]\}]^{-1}$$

$$V_{\tilde{x}}(N/k+1) = V_{\tilde{x}}(N/k) - K_{k+1}^{k-N+1}\, E_k\{\tilde{h}[x(k+1)/k]$$

$$\cdot\, \tilde{x}^T(N/k)\}$$

$$V_{\tilde{x}}(N/k) = \sum_{i,j=o}^{L} E_{k-1}\{\tilde{f}_i[x(N-1-i)/k-1]\,\tilde{f}_j^T[x(N-1-j)/k-1]\}$$

$$+\, G[x(N-1),N-1]\,\Psi_w(N-1)G^T[x(N-1),N-1]$$

$$-\, \kappa_k^{k-N+1}[\Psi_{vw}(N-1)G^T[x(N-1),N-1]\,\delta_{k,N-1}$$

$$+\, E_{k-1}\{\tilde{h}[x(k)/k-1]\sum_{j=o}^{L}\tilde{f}_j^T[x(N-1-j)/k-1]\}]$$

$$\kappa_k^{k-N+1} = [E_{k-1}\{\sum_{j=o}^{L}\tilde{f}_j[x(N-1-j)/k-1]\,\tilde{h}^T[x(k)/k-1]\} + G[x(N-1),N-1]]$$

$$\cdot\, \Psi_{wv}(k)\delta_{k,N-1}][\Psi_v(k)+E_{k-1}\{\tilde{h}[x(k)/k-1]$$

$$\cdot\, \tilde{h}^T[x(k)/k-1]\}]^{-1}$$

and $V_{\tilde{x}}(k-\ell,k-m/k-1)$ is the same as that of Equation (5.2.22).

5.2.5 *The Nonlinear Fixed-Interval Smoothing*

Setting k+1=N, ℓ=N-k, where N≥k, then Equations (5.2.8), (5.2.9), (5.2.12) - (5.2.13) and (5.2.18) - (5.2.19), are the recursive algorithms for fixed-interval smoothing. They are

$$\hat{x}(k/N) = \hat{x}(k/N-1) + K_N^{N-k}\{y(N) - \hat{h}[x(N),N/k]\}$$

$$\hat{x}(k/N-1) = \sum_{j=o}^{L} \hat{f}_j[x(k-1-j),\ k-1-j/N-1]$$

$$K_N^{N-k} = E_{N-1}\{\tilde{x}(k/N-1)\ \tilde{h}^T[x(N)/N-1]\}$$

$$[\Psi_v(N) + E_{N-1}\{\tilde{h}[x(N)/N-1]\tilde{h}^T[x(N)/N-1]\}]^{-1}$$

$$V_{\tilde{x}}(k/N) = V_{\tilde{x}}(k/N-1) - K_N^{N-k}E_{N-1}\{\tilde{h}[x(N)/N-1]$$

$$\tilde{x}^T(k/N-1)\}$$

$$V_{\tilde{x}}(k/N-1) = \sum_{i,j=o}^{L} E_{N-2}\{\tilde{f}_i[x(k-1-i)/N-2]$$

$$\cdot\ \tilde{f}_j^T[x(k-1-j)/N-2]\} + G[x(k-1),k-1]\ \Psi_w(k-1)$$

$$\cdot\ G^T[x(k-1),k-1] - K_{N-1}^{N-k}[\Psi_{vw}(k-1)G^T[x(k-1),k-1]$$

$$\delta_{N-1,k-1} + E_{N-2}\{\tilde{h}[x(N-1)/N-2] \sum_{j=o}^{L} \tilde{f}_j^T[x(k-1-j)/N-2]\}]$$

and $$\kappa_{N-1}^{N-k} = [E_{N-2}\{\sum_{j=o}^{L} \tilde{f}_j[x(k-1-j/N-2]\ \tilde{h}^T[x(N-1)/N-2]\}$$

$$+ G[x(k-1),k-1]\ \Psi_{wv}(N-1)\ \delta_{N-1,k-1}]$$

$$[\Psi_v(N-1) + E_{N-2}\{\tilde{h}[x(N-1)/N-2]\ \tilde{h}^T[x(N-1)/N-2]\}]^{-1}$$

5.2.6 *The Nonlinear Filtering*

When ℓ is set to zero, Equations (5.2.8), (5.2.9), (5.2.12), (5.2.13) and (5.2.18) - (5.2.19) become the filtering algorithms. Namely:

$$\hat{x}(k+1/k+1) = \hat{x}(k+1/k) + K_{k+1}^{0}\{y(k+1) - \hat{h}[x(k+1),k+1/k]\}$$

$$\hat{x}(k+1/k) = \sum_{j=o}^{L} \hat{f}_j[x(k-j),k-j/k]$$

$$K_{k+1}^{0} = E_k\{\tilde{x}(k+1/k)\ \tilde{h}^T[x(k+1)/k]\}$$

$$[\Psi_v(k+1) + E_k\{\tilde{h}[x(k+1)/k]\ \tilde{h}^T[x(k+1)/k]\}]^{-1}$$

$$V_{\tilde{x}}(k+1/k+1) = V_{\tilde{x}}(k+1/k) - K_{k+1}^{0}E_k\{\tilde{h}[x(k+1)/k]\ \tilde{x}^T(k+1/k)\}$$

$$V_{\tilde{x}}(k+1/k) = \sum_{i,j=o}^{L} E_{k-1}\{\tilde{f}_i[x(k-i)/k-1]\ \tilde{f}_j^T[x(k-j)/k-1]\}$$

$$+ G[x(k),k]\ \Psi_w(k)\ G^T[x(k),k]$$

$$- \kappa_k^0[\Psi_{vw}(k)\ G^T[x(k),k]$$

$$+ E_{k-1}\{\tilde{h}[x(k)/k-1] \sum_{j=o}^{L} \tilde{f}_j^T[x(k-j)/k-1]\}]$$

and finally

$$\kappa_k^0 = [E_{k-1}\{\sum_{j=o}^{L} \tilde{f}_j[x(k-j)/k-1]\ \tilde{h}^T[x(k)/k-1]\} + G[x(k),k]$$

$$\cdot\ \Psi_{wv}(k)][\Psi_v(k) + E_{k-1}\{\tilde{h}[x(k)/k-1]\ \tilde{h}^T[x(k)/k-1]\}]^{-1}$$

5.2.7 *Estimation in Linear Discrete Systems*

In this section, general linear estimation algorithms are obtained for discrete delayed systems with non-delayed measurements corrupted by correlated noise processes. The algorithms can be easily converted to yield the fixed-lag, fixed-point, fixed-interval smoothing and the filtering estimators together with their respective error-variance equations. This is done by making the proper choice of discrete indices, as shown in Sections 5.2.3 to 5.2.6.

For linear discrete systems, we have

$$\sum_{j=o}^{L} f_j[x(k-j),k-j] = \sum_{j=o}^{L} F_j(k)\ x(k-j)$$

$$h[x(k),k] = H(k)\ x(k)$$

and

$$G[x(k),k] = G(k)$$

Then the general linear estimation algorithms are obtained from Equations (5.2.8), (5.2.9), (5.2.12) - (5.2.13) and (5.2.18) - (5.2.20), by making the above replacement

$$\hat{x}(k-\ell+1/k+1) = \hat{x}(k-\ell+1/k) + K_{k+1}^{\ell}\{y(k+1) - H(k+1)\,\hat{x}(k+1/k)\}$$

$$\hat{x}(k-\ell+1/k) = \sum_{j=o}^{L} F_j(k-\ell)\,\hat{x}(k-\ell-j/k)$$

$$K_{k+1}^{\ell} = V_{\tilde{x}}(k-\ell+1,k+1/k)\,H^T(k+1)\,[\Psi_v(k+1) + H^T(k+1)\,V_{\tilde{x}}(k+1/k)\,H^T(k+1)]^{-1}$$

$$V_{\tilde{x}}(k-\ell+1/k+1) = V_{\tilde{x}}(k-\ell+1/k) - K_{k+1}^{\ell}\,H(k+1)\,V_{\tilde{x}}(k+1,k-\ell+1/k)$$

$$V_{\tilde{x}}(k-\ell+1/k) = \sum_{i,j=o}^{L} F_i(k-\ell)\,V_{\tilde{x}}(k-\ell-i,k-\ell-j/k-1)\cdot F_j^T(k-\ell) + G(k-\ell)\,\Psi_w(k-\ell)\,G^T(k-\ell) - K_k^{\ell}[\Psi_{vw}(k-\ell)\,G^T(k-\ell) + H(k)\sum_{j=o}^{L} V_{\tilde{x}}(k,k-\ell-j/k-1)\,F_j^T(k-\ell)]$$

$$\kappa_k^{\ell} = [\sum_{j=o}^{L} F_j(k-\ell)\ V_{\underset{\sim}{x}}(k-\ell-j,k/k-1)\ H^T(k)$$

$$+ G(k-\ell)\ \Psi_{wv}(k)\ \delta_{k,k-\ell}]$$

$$\cdot [\Psi_v(k) + H(k)\ V_{\underset{\sim}{x}}(k/k-1)\ H^T(k)]^{-1}$$

and

$$V_{\underset{\sim}{x}}(k-\ell+1,\ k-m+1/k) = \sum_{i,j=o}^{L} F_i(k-\ell) V_{\underset{\sim}{x}}(k-\ell-i,k-m-j/k-1)\ F_j^T(k-m)$$

$$+ G(k-\ell)\Psi_w(k-\ell)\delta_{k-\ell,k-m}G^T(k-m)$$

$$- \kappa_k^{\ell}[\Psi_{vw}(k-m)\delta_{k,k-m}G^T(k-m) + \sum_{j=o}^{L} H(k)$$

$$\cdot V_{\underset{\sim}{x}}(k,k-m-j/k-1)$$

$$\cdot F_j^T(k-m)]$$

It can be easily shown that in the special case of discrete systems without delay, the results obtained for the filtering algorithm agree well with those in the literature [49], but the smoothing results are new.

In the particular case of linear discrete delayed systems with uncorrelated noise processes, the filtering and fixed-lag smoothing algorithms obtained here can be shown to agree with those of Priemer and Vacroux [37,38].

5.3 *NONLINEAR DISCRETE DELAYED SYSTEMS WITH MEASUREMENTS CONTAINING MULTIPLE-DELAYS IMBEDDED IN CORRELATED NOISE PROCESSES*

In this section, estimation algorithms are derived for nonlinear discrete delayed systems with measurements containing multiple delays imbedded in correlated noises.

In Section 5.3.1, the problem statement is presented, Section 5.3.2 derives the discrete nonlinear smoothing algorithms, which can be directly applied to yield the fixed-lag, fixed-point, fixed-interval smoothing and the filtering algorithms, with the proper substitution of discrete indices.

Section 5.3.3, presents an example to indicate how one can apply the proposed algorithms to physical systems evaluation.

Section 5.3.4, presents the general estimation algorithms for the linear counterpart of the above problems; the results are identified with those of the literature.

5.3.1 *The Problem Statement*

Consider the discrete nonlinear multiple time delays message model of Equation (5.2.1) with the measurement model described by

$$y(k) = \sum_{j=o}^{N} h_j[x(k-j), k-j] + v(k) \tag{5.3.1}$$

where h_j is an m-dimensional nonlinear vector valued function, v is zero-mean white noise sequence correlated with w, with prior statistics and all initial conditions identical to that of Section 5.2.1.

The state estimation problem also follows that of Section 5.2.1.

5.3.2 *The Derivation of the Smoother*

With reasoning similar to that of Section 4.2.2, it is assumed that the smoothed estimate satisfies the dynamic equation*

$$\hat{x}(k-\ell+1/k+1) = \sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] + K_{k+1}^{\ell} y(k+1) \quad (5.3.2)$$

Since $\hat{x}(k-\ell+1/k+1)$ is required to be an unbiased estimate, it can be shown that

$$\sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] = \sum_{j=o}^{L} \hat{f}_j[x(k-\ell-j), k-\ell-j/k]$$

$$- K_{k+1}^{\ell} \sum_{j=o}^{N} \hat{h}_j[x(k-j+1), k-j+1/k]$$

where

---

* When other constraining equations such as

$$\hat{x}(k-\ell+1/k+1) = \sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] + K_{k+1}^{\ell}\{y(k+1)$$

$$- E_k[y(k+1)]\}$$

is used, the resulting smoothed estimate remains the same.

$$\hat{f}_j[x(k-\ell-j),\ k-\ell-j/k] = E_k\{f_j[x(k-\ell-j),\ k-\ell-j]\}$$

and

$$\hat{h}_j[x(k-j+1),\ k-j+1/k] = E_k\{h_j[x(k-j+1),\ k-j+1]\}$$

Also that,

$$\sum_{j=o}^{L} \hat{f}_j[x(k-\ell-j),\ k-\ell-j/k] = \hat{x}(k-\ell+1/k) \qquad (5.3.3)$$

Then

$$\hat{x}(k-\ell+1/k+1) = \hat{x}(k-\ell+1/k)+K_{k+1}^{\ell}\{y(k+1)-\sum_{j=o}^{N}\hat{h}_j[x(k-j+1),k-j+1/k]\} \qquad (5.3.4)$$

whereas the smoothing error satisfies the relation

$$\tilde{x}(k-\ell+1/k+1) = \tilde{x}(k-\ell+1/k) - K_{k+1}^{\ell}\{ \sum_{j=o}^{N} \tilde{h}_j[x(k-j+1)/k] + v(k+1)\} \qquad (5.3.5)$$

where it is defined that

$$\tilde{h}_j[x(k-j+1)/k] = h_j[x(k-j+1),\ k-j+1] - \hat{h}_j[x(k-j+1),k-j+1/k]$$

and the matrix $V_{\tilde{x}}(k-\ell+1/k+1)$ can be determined from Equations (5.2.4) and (5.3.5).

In order that the cost function of Equation (5.2.3) is minimized, the optimal value of the matrix K^{ℓ}_{k+1} is obtained by setting the gradient of J(k+1) equal to the null matrix. Hence,

$$K^{\ell}_{k+1} = E_k\{\tilde{x}(k-\ell+1/k) \sum_{j=o}^{N} \tilde{h}_j^T[x(k-j+1)/k]\} \tag{5.3.6}$$

$$[\Psi_v(k+1) + E_k\{\sum_{i,j=o}^{N} \tilde{h}_i[x(k-i+1)/k]\, \tilde{h}_j^T[x(k-j+1)/k]\}]^{-1}$$

and $V_{\tilde{x}}(k-\ell+1/k+1) = V_{\tilde{x}}(k-\ell+1/k) - K^{\ell}_{k+1}\, E_k\{\sum_{j=o}^{N} \tilde{h}_j[x(k-j+1)/k]$

$$\tilde{x}^T(k-\ell+1/k)\} \tag{5.3.7}$$

also if the matrix $V_{\tilde{x}}(k-n+1,\ k-m+1/k+1)$ is defined as

$$V_{\tilde{x}}(k-n+1,\ k-m+1/k+1) = E_{k+1}\{\tilde{x}(k-n+1/k+1)\ \tilde{x}^T(k-m+1/k+1)\}$$

then it can be shown that

$$V_{\tilde{x}}(k-n+1,\ k-m+1/k+1) = V_{\tilde{x}}(k-n+1,\ k-m+1/k)$$

$$- K^{n}_{k+1} E_k\{\sum_{j=o}^{N} \tilde{h}_j[x(k-j+1)/k]\tilde{x}^T(k-m+1/k)\} \tag{5.3.8}$$

for any integers n,m with $0 \leq n,m \leq k+1$.

As is presented in Section 5.2.2, the unbiased estimate

$\hat{x}(k-\ell+1/k)$ of the state vector $x(k-\ell+1)$ can also be obtained as

$$\hat{x}(k-\ell+1/k) = \sum_{j=o}^{L} \hat{f}_j[x(k-\ell-j), k-\ell-j/k-1]$$

$$+ \kappa_k^{\ell}\{y(k) - \sum_{j=o}^{N} \hat{h}_j[x(k-j), k-j/k-1]\} \quad (5.3.9)$$

and the optimal value of κ_k^{ℓ} is given by

$$\kappa_k^{\ell} = [E_{k-1}\{\sum_{i=o}^{L} \tilde{f}_i[x(k-\ell-i)/k-1] \sum_{j=o}^{N} \tilde{h}_j^T[x(k-j)/k-1]\}$$

$$+ G[x(k-\ell), k-\ell]\, \Psi_{wv}(k)\, \delta_{k,k-\ell}] \quad (5.3.10)$$

$$[\Psi_v(k) + E_{k-1}\{\sum_{i,j=o}^{N} \tilde{h}_i[x(k-i)/k-1]\, \tilde{h}_j^T[x(k-j)/k-1]\}]^{-1}$$

where

$$\tilde{f}_i[x(k-\ell-i)/k-1] = f_i[x(k-\ell-i), k-\ell-i] - \hat{f}_i[x(k-\ell-i),k-\ell-i/k-1]$$

and

$$V_{\tilde{x}}(k-n+1,k-m+1/k) = \sum_{i,j=o}^{L} E_{k-1}\{\tilde{f}_i[x(k-n-i)/k-1]$$

$$\cdot\, \tilde{f}_j^T[x(k-m-j)/k-1]\}$$

$$+ G[x(k-n),k-n]\Psi_w(k-n)\delta_{k-n,k-m}G^T[x(k-m),k-m]$$

$$- \kappa_k^n[\Psi_{wv}(k-m)\delta_{k-m,k}G^T[x(k-m),k-m] \quad (5.3.11)$$

$$+ E_{k-1}\{\sum_{i=o}^{N} \tilde{h}_i[x(k-i)/k-1]$$

$$\sum_{j=o}^{L} \tilde{f}_j^T[x(k-m-j)/k-1]\}]$$

whereas $V_{\tilde{x}}(k-\ell+1/k)$ can be obtained from the above equation by setting both n and m equal to ℓ.

It should be noted that three different types of smoothing and the filtering algorithms all follow immediately from Equations (5.3.3), (5.3.4), (5.3.6) - (5.3.8) and (5.3.10) - (5.3.11) with the following substitution:

	The replacement for k+1	The replacement for ℓ
The filtering estimation	k+1	0
The fixed-lag smoothing	k	N
The fixed-point smoothing	k+1	k-N+1
The fixed-interval smoothing	N	N-k

Also notice that in the case of nonlinear systems, the preceeding algorithms involve infinite dimensional systems for realization, which is practically impossible except in trivially simple cases.

In order that the smoother can be implemented in computer evaluation, it is assumed that the conditional probability density functions of the smoothing errors are Gaussian. This is very significant for systems with product types or polynomial nonlinearites, since in these cases, the smoothing algorithms can be evaluated or physically realized without any further approximation.

Moreover, the smoothing estimators are already required to be unbiased, therefore the expectation terms that are in Equations (5.3.6) - (5.3.8) and (5.3.10) - (5.3.11) can be further simplified by Equations (4.2.20) - (4.2.22) and (5.2.21) and also

$$E_{k-1}\{\tilde{h}_i[x(k-i)/k-1]\ \tilde{h}_j^T[x(k-j)/k-1]$$

$$= E_{k-1}\{h_i[x(k-i),k-i]\ h_j^T[x(k-j),k-j]\}$$

$$- \hat{h}_i[x(k-i)k-i/k-1]\ \hat{h}_j^T[x(k-j),k-j/k-1]$$

Therefore, the estimation algorithms presented in this chapter are not only new and recursive in nature, but also very effective computationally speaking and shows the close link between various types of state estimation problems.

5.3.3 *Illustrative Example*

To illustrate the use of the preceeding algorithms consider the nonlinear systems equations presented in Section 4.3.3. However, noise processes v(k) and w(k) are correlated with the following statistic :

$$E\, E\{v(k)\, w^T(j)\} = \Psi_{vw}(k)\, \delta_{k,j}$$

The initial conditions and prior statistics are the same as Section 4.3.3.

Then following the evaluation as presented in Section 4.3.3, and applying Equations (5.3.3), (5.3.4), (5.3.6), (5.3.7) and (5.3.10), with ℓ set to zero, the filtering algorithms are obtained as follows:

$$\hat{x}(k+1/k) = 3V_{\tilde{x}}(k)\, \hat{x}(k) + \hat{x}^3(k/k) + 2\hat{x}(k-1/k)$$

$$\hat{x}(k+1/k+1) = \hat{x}(k+1/k) + K^0_{k+1}\{y(k+1) - V_{\tilde{x}}(k+1/k) - \hat{x}^2(k+1/k) - \hat{x}(k)\}.$$

$$K^0_{k+1} = [2V_{\tilde{x}}(k+1/k)\, \hat{x}(k+1/k) + V_{\tilde{x}}(k+1,k/k)]$$

$$[\Psi_v(k+1) + 2V^2_{\tilde{x}}(k+1/k) + 4V_{\tilde{x}}(k+1/k)\, \hat{x}^2(k+1/k) + 4V_{\tilde{x}}(k+1,k/k)\, \hat{x}(k+1/k) + V_{\tilde{x}}(k)]^{-1}$$

$$V_{\tilde{x}}(k+1/k+1) = V_{\tilde{x}}(k+1/k) - K^0_{k+1}[2V_{\tilde{x}}(k+1/k)\, \hat{x}(k+1/k) + V_{\tilde{x}}(k+1,k/k)]$$

$$V_{\tilde{x}}(k+1/k) = 15V_{\tilde{x}}^3(k/k-1) + 36V_{\tilde{x}}^2(k/k-1)\,\hat{x}^2(k/k-1)$$

$$+ 9V_{\tilde{x}}(k/k-1)\,\hat{x}^4(k) + 12V_{\tilde{x}}(k,k-1/k-1)\,V_{\tilde{x}}(k/k-1)$$

$$+ 12V_{\tilde{x}}(k,k-1/k-1)\,\hat{x}^2(k/k-1)$$

$$+ 4V_{\tilde{x}}(k-1/k-1)$$

$$+ \Psi_w(k) - \kappa_k^0\{\Psi_v(k) + 2V_{\tilde{x}}^2(k/k-1)$$

$$+ 4V_{\tilde{x}}(k/k-1)\,\hat{x}^2(k/k-1) + 4V_{\tilde{x}}(k,k-1/k-1)\,\hat{x}(k/k-1)$$

$$+ V_{\tilde{x}}(k-1)\}\,\kappa_k^{0T}$$

where,

$$\kappa_k^0 = \{\Psi_{wv}(k)\,G^T[x(k),k]$$

$$+ 12V_{\tilde{x}}^2(k/k-1)\,\hat{x}(k/k-1) + 3V_{\tilde{x}}(k,k-1/k-1)\,V_{\tilde{x}}(k/k-1)$$

$$+ 10V_{\tilde{x}}(k/k-1)\,\hat{x}^3(k/k-1) + 3V_{\tilde{x}}(k/k-1)\,\hat{x}(k/k-1)\,\hat{x}(k-1/k-1)$$

$$+ 3V_{\tilde{x}}(k.k-1/k-1)\,\hat{x}^2(k/k-1) + \hat{x}^5(k/k-1) + \hat{x}^3(k/k-1)$$

$$\cdot\,\hat{x}(k-1/k-1) + 4V_{\tilde{x}}(k-1,k/k-1)\,\hat{x}(k/k-1) + 2V_{\tilde{x}}(k-1/k-1)$$

$$+ 2\hat{x}(k-1/k-1)\,V_{\tilde{x}}(k/k-1) + 2\hat{x}(k-1/k-1)\,\hat{x}^2(k/k-1)$$

$$+ 2\hat{x}^2(k-1/k-1)\} \cdot \{\Psi_v(k) + 2V^2_{\tilde{x}}(k/k-1)$$

$$+ 4V_{\tilde{x}}(k/k-1)\,\hat{x}^2(k/k-1) + 4V_{\tilde{x}}(k,k-1/k-1)$$

$$\cdot\, \hat{x}(k/k-1) + V_{\tilde{x}}(k-1)\}$$

In the same manner, the fixed-lag, fixed-point and fixed-interval smoothing algorithms can be obtained from Equations (5.3.3), (5.3.4), (5.3.6), (5.3.7), (5.3.10) and (5.3.11) by proper choice of discrete time indices k and ℓ, as listed in Section 5.3.2.

Notice that in the above evaluation, the only assumption needed is the Gaussian assumption of the estimator error.

5.3.4 *Estimation in Linear Discrete Systems*

In this section, the message and measurement models of the estimation problems are linear and discrete. It can be considered as a special case of the nonlinear estimation problems, satisfying the following relations:

$$\sum_{j=o}^{L} f_j[x(k-j),\ k-j] = \sum_{j=o}^{L} F_j(k)\ x(k-j)$$

$$\sum_{j=o}^{N} h_j[x(k-j),\ k-j] = \sum_{j=o}^{N} H_j(k)\ x(k-j)$$

and

$$G[x(k),\ k] = G(k)$$

Then the linear estimation algorithms can be obtained directly from results presented in Section 5.3.2. They are:

$$\hat{x}(k-\ell+1/k+1) = \hat{x}(k-\ell+1/k) + K_{k+1}^{\ell}\{y(k+1)$$

$$- \sum_{j=0}^{N} H_j(k+1)\ \hat{x}(k-j+1/k)\}$$

$$\hat{x}(k-\ell+1/k) = \sum_{j=0}^{L} F_j(k-\ell)\ \hat{x}(k-\ell-j/k)$$

$$K_{k+1}^{\ell} = \sum_{j=0}^{N} V_{\tilde{x}}(k-\ell+1,\ k-j+1/k)\ H_j^T(k+1)$$

$$[\Psi_v(k+1) + \sum_{i,j=0}^{N} H_i(k+1)\ V_{\tilde{x}}(k+1-i,k+1-j/k)\ H_j^T(k+1)]^{-1}$$

$$V_{\tilde{x}}(k-\ell+1/k+1) = V_{\tilde{x}}(k-\ell+1/k) - K_{k+1}^{\ell}\ \{\sum_{j=0}^{N} H_j(k+1)$$

$$\cdot\ V_{\tilde{x}}(k-j+1,k-\ell+1/k)\}$$

$$V_{\tilde{x}}(k-\ell+1/k) = \sum_{i,j=0}^{L} F_i(k-\ell)\ V_{\tilde{x}}(k-\ell-i,k-\ell-j/k-1)$$

$$\cdot\ F_j^T(k-\ell) + G(k-\ell)\ \Psi_w(k-\ell)\ G^T(k-\ell)$$

$$- K_k^{\ell}[\Psi_{wv}(k-\ell)\ G^T(k-\ell) +$$

$$\sum_{i=0}^{N}\ \sum_{j=0}^{L} H_i(k)\ V_{\tilde{x}}(k-i,k-\ell-j/k-1)\ F_j^T(k-\ell)]$$

and

$$\kappa_k^{\ell} = [\sum_{i=o}^{L} \sum_{j=o}^{N} F_i(k-\ell)\ V_{\tilde{x}}(k-\ell-i,k-j/k-1)$$

$$H_j(k) + G(k-\ell)\ \Psi_{wv}(k-\ell)]$$

$$\cdot\ [\Psi_v(k) + \sum_{i,j=o}^{N} H_i(k)\ V_{\tilde{x}}(k-i,k-j/k-1)\ H_j^T(k)]^{-1}$$

The three different types of smoothing and the filtering algorithms all follow immediately from the above equations with the proper choice of k and ℓ listed in Section 5.3.2.

For linear state and measurement models without delays, the results agree well with those in the literature [49]. And if Ψ_{wv} and N are set to zero, the results of this section become the stable estimation algorithms as was presented by Priemer and Vacroux [37, 38]. For linear systems with uncorrelated noise processes the results also agree well with that of Shukla and Srinath [56].

5.4 *CONCLUSION*

In this chapter a new unified approach has been presented to obtain the fixed-lag, fixed-point, fixed-interval smoothing and the filtering algorithms for discrete nonlinear delayed systems with non-delayed measurements as well as multichannel delayed measurements corrupted by correlated noise processes. The derivation is straightforward and shows the close link between the smoothing and filtering estimation problems.

The results are new, recursive in nature, and physically realizable under the assumption that the probability density functions of the estimator errors are Gaussian, they can be obtained without any further approximations for systems with product types or polynomial nonlinearities. They are expected to be computationally faster than applying an extended Kalman filter to an augmented system with no explicit delay terms. Since the augmentation of the state leads to higher dimension systems, this is particularly significant for systems with numerous delay parameters.

For linear estimation problems, the results agree well with those of the literature and are computationally stable, however, for nonlinear estimation algorithms the stability behavior of the systems needs to be further investigated.

The results presented can be extended to continuous time problems through a formal limiting procedure.

CHAPTER VI

MINIMUM VARIANCE FILTERING AND SMOOTHING FOR NONLINEAR SYSTEMS WITH COLORED NOISE

6.1 *INTRODUCTION*

This chapter extends the technique developed in preceeding chapters to another physically realistic estimation problem. That is, the state estimation of nonlinear discrete delayed systems with non-delayed measurement as well as measurements containing multiple delays imbedded in colored noise sequence.

In the past, for linear estimation problems without time delays, two methods have, in general, been employed to account for time-correlated measurement noise in a linear filter. The first method, proposed by Kalman [3] employed the state augmentation method, it involves the construction of a hypothetical shaping filter whose input is white noise and whose output is noise of the required correlations. Here, the dynamics of the shaping filter are included as a part of the system dynamics, and its states are estimated together with the system states. The second method was developed by Bryson and Henrikson [7], it employs output differencing technique, which devises a new modified measurement, that subtracts the most recent past value of the measurement, weighted by the noise correlation coefficient matrix, from the present measurement. As a result, the modified measurement only contains white noise processes.

Recently, Priemer and Vacroux [37, 38] reported filtering and fixed-lag smoothing algorithms for linear discrete time-delayed systems corrupted by white noise sequences, using the orthogonal projection method. Biswas and Mahalanabis, employed the state augmentation technique for linear discrete time-delayed estimation problems involving colored noise processes [55, 61].

To the author's knowledge, nonlinear estimation problems for systems with or without delays, imbedded in colored noise processes, are still outstanding. As an innovation, this thesis presents a direct derivation of estimation algorithms for discrete nonlinear time-delayed systems with non-delayed measurements as well as multi-channel measurements involving time delays and being imbedded in colored noise processes.

The derivation does not require the augmentation of the state vector, and hence does not require an increase in the order of the system and leads to a filter that is computationally efficient. The proposed technique makes use of the matrix minimum principle to derive the optimal values of the coefficients in the estimation algorithms under the requirements that the estimates be unbiased and the error-variance be minimized; they are reported in [62,63].

6.2 *NONLINEAR DISCRETE DELAYED SYSTEMS WITH NON-DELAYED MEASUREMENTS IMBEDDED IN COLORED NOISE*

In this section, estimation algorithms for all three cases of smoothing and the filtering estimation are derived for non-linear discrete time-delayed systems with non-delayed measurements imbedded in colored noise.

Section 6.2.1 presents the statement of the problem, Section 6.2.2 describes the derivation of the nonlinear smoother. Sections 6.2.3 and 6.2.4 respectively, present the fixed-lag smoothing and the fixed-point smoothing algorithms. Fixed-interval smoothing and filtering algorithms are not presented, they can be obtained from the results of Section 6.2.2, with the proper choice of discrete time indices.

For linear estimation problems the three cases of smoothing and the filtering algorithms follow immediately from the results presented in Section 6.2.5. They are identified with various results in literature.

6.2.1 *The Problem Statement*

The process considered in this section is described by

$$x(k+1) = \sum_{j=o}^{L} f_j[x(k-j), k-j] + G[x(k),k]\, w(k) \qquad (6.2.1)$$

where the measurement is given by

$$y(k) = H(k)\, x(k) + \gamma(k) \qquad (6.2.2)$$

and the measurement noise $\gamma(k)$ is a colored sequence given by

$$\gamma(k+1) = A(k+1,k)\, \gamma(k) + v(k) \qquad (6.2.3)$$

Here, k=0,1,2, ..., is the discrete time index, the state x is an n-vector; the measurement y an m-vector; the state noise sequence w an r-vector; the measurement noise γ an m-vector; G, a nonlinear state dependent nxr matrix; f_j, a nonlinear n-dimensional vector valued function; H and A are respectively, mxn and mxm matrices.

The noise sequences w and v are correlated zero-mean white Gaussian with known statistics

$$E\{w(k)\, v^T(j)\} = \Psi_{wv}\,(k)\,\delta_{k,j}$$

$$E\{w(k)\, w^T(j)\} = \Psi_{w}\,(k)\,\delta_{k,j}$$

and

$$E\{v(k)\, v^T\,(j)\} = \Psi_{v}(k)\,\delta_{k,j}$$

for all integers k and j, where E denotes the expected value, $\delta_{k,j}$, the Kronecker delta, the matrices Ψ_w and Ψ_{wv} are non-negative definite, whereas Ψ_v is positive definite.

The initial states are zero-mean Gaussian random vectors, which are independent of v(k) and w(k). And for all positive integers m and ℓ, of interest it is required that

$$E\{x(-m)\, x^T(-\ell)\} = V_x(m,\ell) \qquad (6.2.4)$$

The smoothing problem is to obtain $\hat{x}(k-\ell+1/k+1)$, the unbiased smoothing estimate of $x(k-\ell+1)$ for non-negative integer values of ℓ, conditioned on the set of measurements

$$Y(k+1) = \{y(0),\ y(1),\ \ldots,\ y(k+1)\}$$

such that the cost function

$$J(k+1) = \text{Trace}\ [V_{\tilde{x}}(k-\ell+1/k+1)] \tag{6.2.5}$$

is minimized.

6.2.2 *The Derivation of the Smoother*

Examination of Equations (6.2.2) and (6.2.3), indicates that a modified observation $z(k+1)$ can be defined such that it is a white Gaussian sequence. That is

$$z(k+1) = y(k+1) - A(k+1,k)\ y(k)$$

which can be further simplified as

$$z(k+1) = \sum_{j=o}^{L} d_j[x(k-j),\ k-j] + H(k+1)$$

$$\cdot\ G[x(k),\ k]\ w(k) + v(k) \tag{6.2.6}$$

where

$$\sum_{j=o}^{L} d_j[x(k-j), k-j] = H(k+1) \sum_{j=o}^{L} f_j[x(k-j), k-j]$$

$$- A(k+1,k) H(k) x(k) \qquad (6.2.7)$$

Note that the modified observation z(k+1) contains as much information as that of y(k+1).

It can be assumed that the smoothed estimate is constrained by the discrete nonlinear dynamic equation

$$\hat{x}(k-\ell+1/k+1) = \sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] + K_{k+1}^{\ell} z(k+1) \qquad (6.2.8)$$

Let $\tilde{x}(k-\ell+1/k+1)$ denote the smoothing error defined by

$$\tilde{x}(k-\ell+1/k+1) = x(k-\ell+1) - \hat{x}(k-\ell+1/k+1) \qquad (6.2.9)$$

In order that the smoothed estimate is unbiased, it can be shown that

$$\sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] = \sum_{j=o}^{L} \hat{f}_j[x(k-\ell-j), k-\ell-j/k]$$

$$- K_{k+1}^{\ell} \sum_{j=o}^{L} \hat{d}_j[x(k-j), k-j/k] \qquad (6.2.10)$$

where the following identities are introduced

$$\hat{f}_j[x(k-\ell-j),\ k-\ell-j/k] = E_k\{f_j[x(k-\ell-j),\ k-\ell-j]\}$$

and

$$\hat{d}_j[x(k-j),\ k-j/k] = E_k\{d_j[x(k-j),\ k-j]\}$$

with E_k denoting the expectation operation conditioned on the set of measurements $Z(k)$.

Using Equations (6.2.1), (6.2.8) and (6.2.10), the smoothing error equation of Equation (6.2.9) can be simplified as

$$\begin{aligned} \tilde{x}(k-\ell+1/k+1) = \sum_{j=o}^{L} \tilde{f}_j[x(k-\ell-j)/k] + G[x(k-\ell),\ k-\ell] \\ \cdot\ w(k-\ell) - K_{k+1}^{\ell}\ \{H(k+1)\ G[x(k),k]\ w(k) \\ + v(k) + \sum_{j=o}^{L} \tilde{d}_j[x(k-j)/k]\} \end{aligned} \qquad (6.2.11)$$

where

$$\tilde{f}_j[x(k-\ell-j)/k] = f_j[x(k-\ell-j),\ k-\ell-j] - \hat{f}_j[x(k-\ell-j), k-\ell-j/k]$$

and

$$\tilde{d}_j[x(k-j)/k] = d_j[x(k-j),\ k-j] - \hat{d}_j[x(k-j),\ k-j/k]$$

The error-variance matrix $V_{\tilde{x}}(k-\ell+1/k+1)$ can be obtained by taking the expectation of Equation (6.2.11) multiplied by it's own transpose.

Then the optimal value of the matrix K_{k+1}^{ℓ} can be obtained by setting the gradient of Equation (6.2.5) equal to the null matrix, therefore,

$$
\begin{aligned}
K_{k+1}^{\ell} = [E_k\{\sum_{i,j=0}^{L} & \tilde{f}_i[x(k-\ell-i)/k]\ \tilde{d}_j^T[x(k-j)/k] \\
& + G[x(k-\ell),\ k-\ell]\ \Psi_w(k-\ell)\ \delta_{k-\ell,k} G^T[x(k),k]\ H^T(k+1) \\
& + G[x(k-\ell),\ k-\ell]\ \Psi_{wv}(k-\ell)\ \delta_{k-\ell,k}] \qquad (6.2.12) \\
& \cdot [\Psi_v(k) + H(k+1)\ G[x(k),k]\ \Psi_w(k)\ G^T[x(k),k]\ H^T(k+1) \\
& + E_k\{\sum_{i,j=0}^{L} \tilde{d}_i[x(k-i)/k]\ \tilde{d}_j^T[x(k-j)/k]\} \\
& + H(k+1)\ G[x(k),k]\ \Psi_{wv}(k) + \Psi_{vw}(k)\ G^T[x(k),k]\ H^T(k+1)]^{-1}
\end{aligned}
$$

The smoothing estimate and its error-covariance equation are given by

$$
\begin{aligned}
\hat{x}(k-\ell+1/k+1) = \sum_{j=0}^{L} & \hat{f}_j[x(k-\ell-j),\ k-\ell-j/k] \\
& + K_{k+1}^{\ell}\{y(k+1) - A(k+1,k)\ y(k)\} \qquad (6.2.13)
\end{aligned}
$$

$$- H(k+1) \sum_{j=o}^{L} \hat{f}_j[x(k-j),\ k-j/k] - A(k+1,k)$$

$$\cdot H(k)\ \hat{x}(k)\}$$

and

$$V_{\tilde{x}}(k-m+1,\ k-\ell+1/k+1) = [E_k\{\sum_{i,j=o}^{L} \tilde{f}_i[x(k-m-i)/k]$$

$$\cdot \tilde{f}_j^T[x(k-\ell-j)/k]\}$$

$$+ G[x(k-m),\ k-m]\ \Psi_w(k-m)\delta_{k-m,k-\ell}$$

$$G^T[x(k-\ell),\ k-\ell]]$$

$$- K_{k+1}^m[E_k\{\sum_{i,j=o}^{L} d_i[x(k-i)/k]\ \tilde{f}_j^T[x(k-\ell-j)/k]\}$$

$$+ H(k+1)G[x(k),k]\Psi_w(k)\delta_{k,k-\ell}G^T[x(k-\ell),k-\ell]$$

$$+ \Psi_{vw}(k-\ell)\delta_{k,k-\ell}G^T[x(k-\ell),k-\ell]] \qquad (6.2.14)$$

respectively. And $V_{\tilde{x}}(k-\ell+1/k+1)$ can be obtained from Equation (6.2.14) by setting m equal to ℓ.

Also, it can be shown that

$$\sum_{j=0}^{L} \hat{f}_j[x(k-\ell-j),k-\ell-j/k] = \hat{x}(k-\ell+1/k) \qquad (6.2.15)$$

then Equations (6.2.13) and (6.2.14) can be respectively rewritten as:

$$\hat{x}(k-\ell+1/k+1) = \hat{x}(k-\ell+1/k) + K^{\ell}_{k+1}\{y(k+1)$$

$$- H(k+1)\,\hat{x}(k+1/k) \qquad (6.2.16)$$

$$- A(k+1,k)\,[y(k) - H(k)\,\hat{x}(k)]\}$$

and

$$V_{\tilde{x}}(k-m+1,k-\ell+1/k+1) = V_{\tilde{x}}(k-m+1,k-\ell+1/k)$$

$$- K^{m}_{k+1}[E_k \sum_{i,j=0}^{L} \tilde{d}_i[x(k-i)/k]$$

$$\cdot \tilde{f}^T_j[x(k-\ell-j)/k]\}$$

$$+ H(k+1)G[x(k),k]\Psi_w(k)\delta_{k,k-\ell}G^T[x(k-\ell),k-\ell]$$

$$+ \Psi_{vw}(k-\ell)\delta_{k,k-\ell}G^T[x(k-\ell),k-\ell]] \qquad (6.2.17)$$

where $V_{\tilde{x}}(k-m+1,k-\ell+1/k) = E_k\{\sum_{i,j=0}^{L} \tilde{f}_i[x(k-m-i)/k]$

$$\cdot \tilde{f}^T_j[x(k-\ell-j)/k]\} + G[x(k-m),k-m]\Psi_w(k-m)$$

$$\cdot\ \delta_{k-m,k-\ell} G^T[x(k-\ell),k-\ell] \tag{6.2.18}$$

Here, Equations (6.2.12), (6.2.15) - (6.2.18) form the recursive smoothing algorithm starting with initial condition $\hat{x}(-\ell/-1)$ for nonnegative integer values of m and ℓ. The above equations can be modified to give

$$\hat{x}(-\ell/o) = \hat{x}(-\ell/-1) + K_o^\ell[y(o) - H(o)\ \hat{x}(o/-1)]$$

and

$$V_{\tilde{x}}(-\ell,\ -\ell/o) = V_{\tilde{x}}(-\ell,\ -\ell/-1) - K_o^\ell H(o)\ V_{\tilde{x}}(o,\ -\ell/-1)$$

where

$$K_o^\ell = V_{\tilde{x}}(-\ell,\ 0/-1)\ H^T(o)\ [H(o)\ V_{\tilde{x}}(o,o/-1)\ H^T(o) + \Psi_\gamma(o)]^{-1}$$

Notice that in the above derivation the results obtained are exact and optimal with respect to the constraints of Equation (6.2.8) minimizing the error-variance cost functional. However, in the case of nonlinear systems, the preceeding algorithms require infinite dimensional systems to realize the various terms involving the expectations.

In order that the smoother can be physically realized, the conditional probability density functions of the smoothing errors are assumed to be Gaussian, and Equation (6.2.9) can be used to substitute for the state variables. This is particularly significant for systems with product types or polynomial nonlinearities, since in these cases, the smoothing algorithms can be evaluated without much difficulty and are physically realizable without any further need of approximation.

It must be noted that the results presented in this chapter are new, the smoothing algorithms can be implemented in the computer evaluation, the computations, are expected to be efficient, since no augmentation of state variables are involved. Another very significant contribution of the presented derivation is the fact that all three cases of the smoothing problems and the filtering estimation follow immediately from the above results, with the proper choice of discrete time indices k and ℓ. They are given as follows:

	Replacement for k	Replacement for ℓ
The filtering estimation		0
The fixed-lag smoothing	k-1	N
The fixed-point smoothing	k	k-N+1
The fixed-interval smoothing	N-1	N-k

It is clear that a unified approach has been developed to derive the three different classifications of smoothers and the filtering algorithms for discrete nonlinear delayed systems involving colored noise. The derivation clearly shows the close link between the various smoothers and filter.

6.2.3 *The Nonlinear Fixed-Lag Smoothing*

In this section, nonlinear fixed-lag smoother and it's error variance equations are presented. They follow immediately from Equations (6.2.12), (6.2.15) - (6.2.18), with substitution of k and N for k+1 and ℓ, respectively. The results are:

$$\hat{x}(k-N/k) = \hat{x}(k-N/k-1) + K_k^N\{y(k) - H(k)\,\hat{x}(k/k-1)$$
$$- A(k,k-1)[y(k-1) - H(k-1)\,\hat{x}(k-1)]\}$$

$$\hat{x}(k-N/k-1) = \sum_{j=o}^{L} \hat{f}_j[x(k-N-j-1),\ k-N-j-1/k-1]$$

$$K_k^N = [E_{k-1}\{\sum_{i,j=o}^{L} \tilde{f}_i[x(k-N-i-1)/k-1]\ \tilde{d}_j^T[x(k-1-j)/k-1]$$
$$+ G[x(k-1-N),k-1-N]\ \Psi_w(k-1-N)\ \delta_{k-1-N,k-1} G^T[x(k-1),k-1]$$
$$\cdot H^T(k) + G[x(k-1-N),k-1-N]\ \Psi_{wv}(k-1-N)\ \delta_{k-1-N,k-1}]$$
$$\cdot [\Psi_v(k-1) + H(k)\ G[x(k-1),k-1]\ \Psi_w(k-1)\ G^T[x(k-1),k-1]$$
$$\cdot H^T(k) + E_{k-1}\{\sum_{i,j=o}^{L} \tilde{d}_i[x(k-1-i)/k-1]\tilde{d}_j^T[x(k-1-j)/k-1]\}$$
$$+ H(k)\ G[x(k-1),k-1]\ \Psi_{wv}(k-1) + \Psi_{vw}(k-1)\ G^T[x(k-1),k-1]$$
$$\cdot H^T(k)]^{-1}$$

$$V_{\tilde{x}}(k-N/k) = V_{\tilde{x}}(k-N/k-1) - K_k^N[E_{k-1}\{\sum_{i,j=o}^{L} d_i[x(k-1-i)/k-1]$$

$$\cdot \tilde{f}_j^T[x(k-1-N-j)/k-1]\} + H(k-1)G[x(k-1),k-1]\Psi_w(k-1)$$

$$\cdot \delta_{k-1,k-1-N}G^T[x(k-1-N), k-1-N]$$

$$+ \Psi_{vw}(k-1-N)\ \delta_{k-1,k-1-N}G^T[x(k-1-N),k-1-N]]$$

and

$$V_{\tilde{x}}(k-N/k-1) = E_{k-1}\{\sum_{i,j=o}^{L} \tilde{f}_i[x(k-1-N-i)/k-1]$$

$$\cdot \tilde{f}_j^T[x(k-1-N-j)/k-1]\} + G[x(k-1-N),k-1-N]$$

$$\cdot \Psi_w(k-1-N)\ G^T[x(k-1-N),k-1-N]$$

6.2.4 *The Nonlinear Fixed-Point Smoothing*

In this section a nonlinear fixed-point smoother is obtained from Equations (6.2.12), (6.2.15) - (6.2.18), with k-N+1 substituted for ℓ. Then

$$\hat{x}(N/k+1) = \hat{x}(N/k) + K_{k+1}^{k-N+1}\ \{y(k+1)$$

$$- H(k+1)\ \hat{x}(k+1/k) - A(k+1,k)[y(k)-H(k)\ \hat{x}(k)]\}$$

$$\hat{x}(N/k) = \sum_{j=o}^{L} \hat{f}_j[x(N-1-j),\ N-1-j/k]$$

$$K_{k+1}^{k-N+1} = [E_k\{\sum_{i,j=o}^{L} \tilde{f}_i[x(N-1-i)/k]\ \tilde{d}_j^T[x(k-j)/k]$$

$$+ G[x(N-1),N-1]\ \Psi_w(N-1)\ \delta_{N-1,k} G^T[x(k),k]\ H^T(k+1)$$

$$+ G[x(N-1),N-1]\ \Psi_{wv}(N-1)\ \delta_{N-1,k}]$$

$$\cdot [\Psi_v(k) + H(k+1)G[x(k),k]\Psi_w(k)G^T[x(k),k]\ H^T(k+1)$$

$$+ E_k\{\sum_{i,j=o}^{L} \tilde{d}_i[x(k-i)/k]\ \tilde{d}_j^T[x(k-j)/k]\}$$

$$+ H(k+1)G[x(k),k]\ \Psi_{wv}(k) + \Psi_{vw}(k)\ G^T[x(k),k]$$

$$\cdot H^T(k+1)]^{-1}$$

$$V_{\tilde{x}}(N/k+1) = V_{\tilde{x}}(N/k) - K_{k+1}^{k-N+1}[E_k\{\sum_{i,j=o}^{L} \tilde{d}_i[x(k-i)/k]$$

$$\cdot \tilde{f}_j^T[x(N-1-j)/k]\} + H(k+1)G[x(k),k]\ \Psi_w(k)$$

$$\cdot \delta_{k,N-1}G^T[x(N-1),N-1] + \Psi_{vw}(N-1)\delta_{k,N-1}G^T[x(N-1),N-1]$$

$$V_{\tilde{x}}(N/k) = E_k\{\sum_{i,j=o}^{L} \tilde{f}_i[x(N-1-i)/k]\tilde{f}_j^T[x(N-1-j)/k]\}$$

$$+ G[x(N-1),N-1]\ \Psi_w(N-1)\ G^T[x(N-1),N-1]$$

6.2.5 *Estimation in Linear Discrete Systems*

The estimation problem of linear discrete systems can be considered as a special case of the nonlinear estimation problems satisfying the relation that

$$\sum_{j=o}^{L} f_j[x(k-j),k-j] = \sum_{j=o}^{L} F_j(k)\; x(k-j)$$

$$G[x(k),k] = G(k)$$

From Equations (6.2.12), (6.2.15) - (6.2.18), the resulting linear estimation algorithms are as follows:

$$\hat{x}(k-\ell+1/k+1) = \hat{x}(k-\ell+1/k) + K^{\ell}_{k+1}\{y(k+1)$$

$$- H(k+1)\;\hat{x}(k+1/k) - A(k+1,k)[y(k)-H(k)\;\hat{x}(k)]\}$$

where $\hat{x}(k-\ell+1/k) = \sum_{j=o}^{L} F_j(k-\ell)\;\hat{x}(k-\ell-j/k)$

$$K^{\ell}_{k+1} = [\sum_{i,j=o}^{L} F_i(k-\ell)\; V_{\tilde{x}}(k-\ell-i,k-j/k)\; F_j^T(k)$$

$$\cdot H^T(k+1) + \sum_{i=o}^{L} F_i(k-\ell)\; V_{\tilde{x}}(k-\ell-i,k/k)\; H^T(k)$$

$$\cdot A^T(k+1,k) + G(k-\ell)\Psi_w(k-\ell)\delta_{k-\ell,k}G^T(k-\ell)H^T(k+1)$$

$$+ G(k-\ell)\ \Psi_{wv}(k-\ell)\ \delta_{k-\ell,k}]$$

$$\cdot\ [\Psi_v(k) + H(k+1)G(k)\ \Psi_w(k)\ G^T(k)\ H^T(k+1)$$

$$+ H(k+1) \sum_{i,j=o}^{L} F_i(k)\ V_{\tilde{x}}(k-i,k-j/k)\ F_j^T(k)\ H^T(k+1)$$

$$- H(k+1) \sum_{i=o}^{L} F_i(k)\ V_{\tilde{x}}(k-i,k/k)\ H^T(k)\ A^T(k+1)$$

$$- \sum_{j=o}^{L} A(k+1)\ H(k)\ V_{\tilde{x}}(k,k-j/k)\ F_j^T(k)\ H^T(k+1)$$

$$+ A(k+1)\ H(k)\ V_{\tilde{x}}(k/k)\ H^T(k)\ A^T(k+1)$$

$$+ H(k+1)\ G(k)\ \Psi_{wv}(k) + \Psi_{vw}(k)\ G^T(k)\ H^T(k+1)]^{-1}$$

and

$$V_{\tilde{x}}(k-\ell+1/k+1) = V_{\tilde{x}}(k-\ell+1/k) - K_{k+1}^{\ell}[\sum_{i,j=o}^{L} H(k+1)\ F_i(k)$$

$$\cdot\ V_{\tilde{x}}(k-i,k-\ell-j/k)\ F_j^T(k-\ell) + \sum_{i=o}^{L} A(k+1,k)\ H(k)$$

$$\cdot\ V_{\tilde{x}}(k,k-\ell-j/k)\ F_j^T(k-\ell)$$

$$+ H(k+1)\ G(k-\ell)\ \delta_{k,k-\ell}\Psi_w(k-\ell)G^T(k-\ell)$$

$$+ \Psi_{vw}(k-\ell)\ G^T(k-\ell)\ \delta_{k,k-\ell}]$$

where, $V_{\tilde{x}}(k-\ell+1/k) = \sum_{i,j=o}^{L} F_i(k-\ell)\, V_{\tilde{x}}(k-\ell-i,k-\ell-j/k)$

$$\cdot\, F_j^T(k-\ell) + G(k-\ell)\, \Psi_w(k-\ell)\, G^T(k-\ell)$$

Three different classifications of smoothers and the filtering algorithms can be obtained from the preceeding algorithms with substitution for k and ℓ, identical to those of Section 6.2.2.

In the particular case of linear message model without delays, the results presented above agree well with algorithms developed in the literature [6,7]. For the case of linear message model with delays, the results presented for fixed-lag smoother and fixed-point smoother can be identified with those presented by Biswas and Mahalanabis [55,61].

6.3 *NONLINEAR DISCRETE DELAYED SYSTEMS WITH MEASUREMENTS CONTAINING MULTIPLE DELAYS IMBEDDED IN COLORED NOISE*

This section derives nonlinear estimation algorithms for nonlinear discrete delayed systems with measurements containing multiple time delays, corrupted by colored noise. The results are new, recursive in nature and directly yield the three different classification of smoothers and the filtering estimator.

Section 6.3.1 presents the statement of the problem, in Section 6.3.2, the nonlinear smoother are derived. Sections 6.3.3, 6.3.4 and 6.3.5 respectively, present the nonlinear fixed-lag, fixed-point smoothers and the filtering algorithms.

General linear estimation algorithms are presented in Section 6.3.6, the results can be directly applied to the fixed-lag, fixed-point, fixed-interval smoothers and the filtering estimator. The results obtained here for linear discrete systems are also new.

6.3.1 *The Problem Statement*

In this section, the discrete nonlinear time-delayed message model is described by Equation (6.2.1), the measurement noise sequence $\gamma(k)$ is modelled by Equation (6.2.3), and the measurement model is given by

$$y(k) = \sum_{j=0}^{M} H_j \, x(k-j) + \gamma(k) \qquad (6.3.1)$$

here, H_j is an mxn dimensional matrix. All initial conditions and priori statistics are the same as that of Section 6.2.1. The smoothing estimation problem is also identical to that of Section 6.2.1.

6.3.2 *The Derivation of the Smoother*

Examination of Equations (6.3.1) and (6.2.3) indicates that a modified measurement signal $z(k+1)$ consisting of an additive white noise sequence can be obtained by multiplying $y(k)$ by $A(k+1,k)$ and subtracting it from $y(k+1)$, for which

$$z(k+1) = y(k+1) - A(k+1, k)\, y(k) \qquad (6.3.2)$$

Notice that subtracting a function of the previous observation in no way affects the present observation, and hence the estimate conditioned on Y(k+1) is equivalent to that with Z(k+1).

Equation (6.3.2) can be further simplified as

$$z(k+1) = \sum_{i=o}^{L} \sum_{j=o}^{M} d_{i,j}[x(k-j-i), k-j-i]$$

$$+ \sum_{j=o}^{M} H_j(k+1)\, G[x(k-j), k-j]\, w(k-j) + v(k) \qquad (6.3.3)$$

where

$$\sum_{i=o}^{L} \sum_{j=o}^{M} d_{i,j}[x(k-j-i),k-j-i] = \sum_{i=o}^{L} \sum_{j=o}^{M} H_j(k+1)\, f_i[x(k-j-i),$$

$$k-j-i]-A(k+1,k) \sum_{j=o}^{M} H_j(k)x(k-j) \qquad (6.3.4)$$

It is assumed that the smoothed estimate is constrained by the discrete nonlinear dynamic equation

$$\hat{x}(k-\ell+1/k+1) = \sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] + K_{k+1}^{\ell} z(k+1) \qquad (6.3.5)$$

and since it is required that the smoothed estimate is unbiased, the following relation is obtained

$$\sum_{j=o}^{L} b_j[\hat{x}(k-\ell-j/k), k-\ell-j] = \sum_{j}^{L} f_j[\hat{x}(k-\ell-j), k-\ell-j/k]$$

$$- K_{k+1}^{\ell} \sum_{i=o}^{L} \sum_{j=o}^{M} \hat{d}_{i,j}[x(k-j-i),k-j-i/k]$$

where

$$\sum_{j=0}^{L} \hat{f}_j[x(k-\ell-j),k-\ell-j/k] = E\sum_{j=0}^{L} \{ f_j[x(k-\ell-j),k-\ell-j]/Z(k)\}$$

Also

$$\sum_{i=0}^{L} \sum_{j=0}^{M} \hat{d}_{i,j} [x(k-j-i), k-j-i/k]$$

$$= E\{\sum_{i=0}^{L} \sum_{j=0}^{M} d_{i,j}[x(k-j-i), k-j-i]/Z(k)\} \qquad (6.3.6)$$

On the basis of Equation (6.2.1), the following identify is obtained

$$\hat{x}(k-\ell+1/k) = \sum_{j=0}^{L} \hat{f}_j[x(k-\ell-j), k-\ell-j/k] \qquad (6.3.7)$$

then Equation (6.3.5) can be simplified as

$$\hat{x}(k-\ell+1/k+1) = \hat{x}(k-\ell+1/k) + K_{k+1}^{\ell} \{\underline{z}(k+1)$$

$$- \sum_{i=0}^{L} \sum_{j=0}^{M} \hat{d}_{i,j} [x(k-j-i), k-j-i/k]\} \qquad (6.3.8)$$

Substituting Equation (6.3.2) into (6.3.8), and using the identities of Equations (6.3.4) and (6.3.6), the smoothing error takes the form

$$\tilde{x}(k-\ell+1/k+1) = \tilde{x}(k-\ell+1/k) - K_{k+1}^{\ell}\{\sum_{j=0}^{M} H_j(k+1)\, \tilde{x}(k-j+1\,/k)$$

$$\doteq A(k+1,k) \sum_{j=0}^{M} H_j(k)\, \tilde{x}(k-j/k) + v(k) \qquad (6.3.9)$$

Now the error-variance equation for $V_{\tilde{x}}(k-\ell+1/k)$ can be obtained by taking the expectation of Equation (6.3.9) multiplied by its own transpose. Then the optimal value of the matrix K_{k+1}^{ℓ} can be obtained by setting the gradient of Equation (6.2.5) equal to the null matrix.

As a result, the following relations are found

$$K_{k+1}^{\ell} = \{\sum_{j=0}^{M} V_{\tilde{x}}(k-\ell+1, k-j+1/k)\, H_j^T(\text{[illegible]})$$

$$- \sum_{j=0}^{M} V_{\tilde{x}}(k-\ell+1, k-j/k)\, H_j^T(k)\, A^T(k+1,k) + G[x(k-\ell),k-\ell]$$

$$\cdot \Psi_{wv}(k-\ell)\, \delta_{k-\ell,k}\}$$

$$\cdot \{\sum_{i,j=0}^{M} H_i(k+1)\, V_{\tilde{x}}(k-i+1, k-j+1/k)\, H_j^T(k+1)$$

$$+ \sum_{i,j=0}^{M} A(k+1,k)\, H_i(k)\, V_{\tilde{x}}(k-i,k-j/k)\, H_j^T(k)\, A^T(k+1,k)$$

$$- \sum_{i,j=0}^{M} H_i(k+1)\, V_{\tilde{x}}(k-i+1,k-j/k)\, H_j^T(k)\, A^T(k+1,k)$$

$$- \sum_{i,j=0}^{M} A(k+1,k)\, H_i(k)\, V_{\tilde{x}}(k-i,k-j+1/k)\, H_j^T(k+1) \qquad (6.3.10)$$

$$+ \Psi_v(k) + \sum_{i=0}^{M} H_i(k+1) G[x(k-i),k-i] \Psi_{wv}(k-i) \delta_{k-i,k}$$

$$+ \sum_{j=o}^{M} \Psi_{vw}(k-j)\ G^T[\hat{x}(k-j), k-j]\ H_j^T(k+1)\ \delta_{k-j,k}\}^{-1}$$

and

$$V_{\underset{\sim}{x}}(k-m+1, k-\ell+1/k+1) = V_{\underset{\sim}{x}}(k-m+1, k-\ell+1/k)$$

$$- K_{k+1}^{m}\{\sum_{j=o}^{M} H_j(k+1)\ V_{\underset{\sim}{x}}(k-j+1, k-\ell+1/k)$$

$$- \sum_{j=o}^{M} A(k+1,k) H_j(k) V_{\underset{\sim}{x}}(k-j, k-\ell+1/k)$$

$$+ \Psi_{vw}(k-\ell)\delta_{k-\ell,k} G^T[x(k-\ell), k-\ell]\} \quad (6.3.11)$$

Furthermore, from Equations (6.2.[illegible]) and (6.3.7), one can obtain

$$\tilde{x}(k-\ell+1/k) = \sum_{j=o}^{L} \tilde{f}_j[x(k-\ell-j), k-\ell-j] + G[x(k-\ell), k-\ell]w(k-\ell) \quad (6.3.12)$$

therefore

$$V_{\underset{\sim}{x}}(k-m+1, k-\ell+1/k) = \sum_{i,j=o}^{L} E_k\{\tilde{f}_i[x(k-m-i), k-m-i/k]\tilde{f}_j^T[x(k-\ell-j), k-\ell-j/k]\} + G[x(k-m), k-m]\ \Psi_w(k-m)\ \delta_{k-m,k-\ell}\ G^T[x(k-\ell), k-\ell] \quad (6.3.13)$$

and

$$V_{\tilde{x}}(k-m+1, k-\ell/k) = \sum_{j=o}^{L} E_k\{f_j[\tilde{x}(k-m-j), k-m-j/k]$$

$$\tilde{x}^T(k-\ell/k)\} \qquad (6.3.14)$$

for all non-negative integers m and ℓ of interest. Note that $V_{\tilde{x}}(k-\ell+1/k+1)$ and $V_{\tilde{x}}(k-\ell+1/k)$ can be easily obtained from Equations (6.3.11) and (6.3.13) with m set equal to ℓ.

It should be noted that the above estimation algorithms are very general, they can be applied to all three cases of smoothing as well as the filtering estimation problems. The results obtained are exact, new and optimal with respect to the constraining equation of (6.2.6) minimizing the error-variance cost functional of Equation (6.2.5).

However, in the case of nonlinear systems, the preceeding algorithms require infinite dimensional systems to realize the various terms involving the expectations.

For practical realization, it can be assumed that the conditional probability density functions of the estimator errors are Gaussian. Under such an assumption, for systems with product types or polynomial nonlinearities, the presented algorithms can be physically realized without any further needs of approximation. The computer evaluation are expected to be efficient, since no augmentation of state variables are involved. Other significant advantages of the above technique include the facts that the derivation is straightforward and easy to apply to various other estimation problems, the results clearly show the close link between the various smoothers and filters, and also the algorithms can be computed recursively starting with initial condition $\hat{x}(-\ell/-1)$ and $V_{\tilde{x}}(-m,-\ell/-1)$ for non-negative values of m and ℓ.

6.3.3 *The Nonlinear Fixed-Lag Smoothing.*

With substitution of k+1 = k and ℓ = N, where N is a fixed integer and k(>N) is variable, Equations (6.3.7), (6.3.8), (6.3.10), (6.3.13) and (6.3.14) yield the discrete fixed-lag smoothing algorithms for nonlinear time-delayed systems with colored noise. They are

$$\hat{x}(k-N/k-1) = \sum_{j=o}^{L} \hat{f}_j[x(k-1-N-j), k-1-N-j/k-1]$$

$$\hat{x}(k-N/k) = \hat{x}(k-N/k-1) + K_k^N \{z(k)$$

$$- \sum_{i=o}^{} \sum_{j=o}^{M} \hat{d}_{i,j}[x(k-1-j-i), k-1-j-i/k-1]\}$$

$$K_k^N = \{\sum_{j=o}^{M} V_{\tilde{x}}(k-N,k-j/k-1)\, H_j^T(k)$$

$$-\sum_{j=o}^{M} V_{\tilde{x}}(k-N,k-j-1/k-1)\; H_j^T(k-1)A^T(k,k-1)+G[\hat{x}(k-N-1),k-N-1]$$

$$\cdot \Psi_{wv}(k-N-1)\delta_{k-N+1,k}\}\cdot\{\sum_{i,j=o}^{M} H_i(k)V_{\tilde{x}}(k-i,k-j/k-1)H_j^T(k)$$

$$+ \sum_{i,j=o}^{M} A(k,k-1)H_i(k-1)V_{\tilde{x}}(k-i-1,k-j-1/k-1)H_j^T(k-1)A^T(k,k-1)$$

$$- \sum_{i,j=0}^{M} H_i(k) V_{\tilde{x}}(k-i, k-j-1/k-1) H_j^T(k-1) A^T(k,k-1)$$

$$- \sum_{i,j=0}^{M} A(k,k-1)\, H_i(k-1)\, V_{\tilde{x}}(k-i-1, k-j/k-1)\, H_j^T(k)$$

$$+ \Psi_v(k-1) + \sum_{i=0}^{M} H_i(k) G[x(k-1-i), k-1-i] \Psi_{wv}(k-1-i) \delta_{k-1-i,k-1}$$

$$+ \sum_{j=0}^{M} \Psi_{vw}(k-1-j) G^T[x(k-1-j), k-1-j] H_j^T(k)\, \delta_{k-1-j,k-1} \}^{-1}$$

$$V_{\tilde{x}}(k-N/k) = V_{\tilde{x}}(k-N/k-1) - K_k^N \{ \sum_{j=0}^{M} H_j(k)$$

$$\cdot V_{\tilde{x}}(k-j, k-N/k-1) - \sum_{j=0}^{M} A(k,k-1) H_j(k-1)$$

$$\cdot V_{\tilde{x}}(k-j-1, k-N/k-1)$$

$$+ \Psi_{vw}(k-N-1)\, \delta_{k-N-1,k}\, G^T[x(k-N-1), k-N-1] \}$$

$$V_{\tilde{x}}(k-N/k-1) = \sum_{i,j=0}^{L} E_{k-1}\{ \tilde{f}_i[x(k-1-N-i), k-1-N-i/k-1]$$

$$\tilde{f}_j^T[x(k-1-N-j), k-1-N-j/k-1] \}$$

$$+ G[x(k-1-N), k-1-N]\, \Psi_w(k-1-N) G^T[x(k-1-N), k-1-N]$$

and

$$V_{\tilde{x}}(k-N, k-i-1/k-1) = \sum_{j=0}^{L} E_{k-1}\{ \tilde{f}_j[x(k-N-j-1), k-N-j-1/k-1]$$

$$\cdot \tilde{x}^T(k-i-1/k-1)\}$$

Whereas $V_{\tilde{x}}(k+1/k)$ can be directly obtained from Equation (6.3.13) with substitution of $m=\ell=0$, and $V_{\tilde{x}}(k-i+1,k-j+1/k)$ is the same as that of Equation (6.3.13).

6.3.4 *The Nonlinear Fixed-Point Smoothing*

If $\ell=k-N+1$ is substituted into Equations (6.3.7), (6.3.8), (6.3.10) - (6.3.11) and (6.3.13) - (6.3.14), where $k+1 > N$, and N is fixed, then the above equations represent the fixed-point smoothing algorithms for nonlinear time-delayed systems with colored noise. These equations are

$$\hat{x}(N/k) = \sum_{j=0}^{L} \text{[illegible]}(N-1-j), N-1-j/k]$$

$$\hat{x}(N/k+1) = \hat{x}(N/k) + K_{k+1}^{k-N+1}\{z(k+1)$$

$$-\sum_{i=0}^{L}\sum_{j=0}^{M} \hat{d}_{i,j}[x(k-j-i),\ k-j-i/k]\}$$

$$K_{k+1}^{k-N+1} = \{\sum_{j=0}^{M} V_{\tilde{x}}(N,k-j+1/k)\ H_j^T(k+1)$$

$$-\sum_{j=0}^{M} V_{\tilde{x}}(N,k-j/k)\ H_j^T(k)A^T(k+1,k) + G[x(N-1),N-1]$$

$$\cdot \Psi_{wv}(N-1)\ \delta_{N-1,k}\}$$

$$\cdot\Big\{\sum_{i,j=o}^{M} H_i(k+1)\, V_{\tilde{x}}(k-i+1,k-j+1/k)\, H_j^T(k+1)$$

$$+\sum_{i,j=o}^{M} A(k+1,k)\, H_i(k)\, V_{\tilde{x}}(k-i,k-j/k)\, H_j^T(k)\, A^T(k+1,k)$$

$$-\sum_{i,j=o}^{M} H_i(k+1)\, V_{\tilde{x}}(k-i+1,k-j/k)\, H_j^T(k)\, A^T(k+1,k)$$

$$-\sum_{i,j=o}^{M} A(k+1,k)\, H_i(k)\, V_{\tilde{x}}(k-i,k-j+1/k)\, H_j^T(k+1)$$

$$+\Psi_v(k)+\sum_{i=o}^{M} H_i(k+1)\, G[x(k-i),k-i]\, \Psi_{wv}(k-i)\, \delta_{N-i,k}$$

$$\sum_{j=o}^{M} \Psi_{vw}(k-j)\, G^T[x(k-j),k-j]\, H_j^T(k+1)\, \delta_{k-j,k}\Big\}^{-1}$$

$$V_{\tilde{x}}(N/k+1) = V_{\tilde{x}}(N/k) - K_{k+1}^{k+N+1}\Big\{\sum_{j=o}^{M} H_j(k+1)$$

$$\cdot V_{\tilde{x}}(k-j+1,N/k) - \sum_{j=o}^{M} A(k+1,k)\, H_j(k)$$

$$\cdot V_{\tilde{x}}(k-j,N/k) + \Psi_{vw}(N-1)\, \delta_{N-1,k}\, G^T[x(N-1),N-1]\Big\}$$

where

$$V_{\tilde{x}}(N/k) = \sum_{i,j=o}^{L} E_k\{\tilde{f}_i[x(N-i-1),N-i-1/k]\, \tilde{f}_j^T[x(N-j-1),N-j/k]\}$$

$$+\, G[x(N-1),N-1]\Psi_w(N-1)\, G^T[x(N-1),N-1]$$

and

$$V_{\tilde{x}}(N,k-i/k) = \sum_{j=o}^{L} E_k\{\tilde{f}_j[x(N-j-1),N-j-1/k]\;\; \tilde{x}^T(k-i/k)\}$$

Here, $V_{\tilde{x}}(k+1/k)$ and $V_{\tilde{x}}(k-i+1,k-j+1/k)$ are the same as that of

Section 6.3.3.

6.3.5 *The Nonlinear Filtering*

The nonlinear filtering algorithms are obtained from Equations (6.3.7), (6.3.8), (6.3.10), (6.3.11) (6.3.13) and (6.3.14) with ℓ and m set to zero. The results are:

$$\hat{x}(k+1/k+1) = \hat{x}(k+1/k) + K^{0}_{k+1} \{z(k+1)$$

$$- \sum_{i=o}^{L} \sum_{j=o}^{M} \hat{d}_{i,j}[x(k-j-i),k-j-i/k]\}$$

$$\hat{x}(k+1/k) = \sum_{j=o}^{L} \hat{f}_j[x(k-j),k-j/k]$$

$$K^{0}_{k+1} = \{\sum_{j=o}^{M} V_{\tilde{x}}(k+1,k-j+1/k) H_j^T(k+1)$$

$$- \sum_{j=o}^{M} V_{\tilde{x}}(k+1,k-j/k) H_j^T(k) A^T(k+1,k) + G[x(k),k]$$

$$\cdot \Psi_{?}(k)\} \cdot \{\sum_{i,j=o}^{M} H_i(k+1) V_{\tilde{x}}(k-i+1,k-j+1/k) H_j^T(k+1)$$

$$+ \sum_{i,j=o}^{M} A(k+1,k)H_i(k) V_{\tilde{x}}(k-i,k-j/k) H_j^T(k) A^T(k+1,k)$$

$$- \sum_{i,j=o}^{M} H_i(k+1) V_{\tilde{x}}(k-i+1,k-j/k) H_j^T(k) A^T(k+1,k)$$

$$- \sum_{i,j=o}^{M} A(k+1,k)H_i(k)V_{\tilde{x}}(k-i,k-j+1/k) H_j^T(k+1)$$

$$+ \Psi_v(k) + H_o(k+1)G[x(k),k] \Psi_{wv}(k)$$

$$+ \Psi_{vw}(k)\, G^T[x(k),k]\, H_o^T(k+1)\}^{-1}$$

$$V_{\tilde{x}}(k+1/k+1) = V_{\tilde{x}}(k+1/k) - K_{k+1}^{o} \{\sum_{j=o}^{M} H_j(k+1)$$

$$\cdot V_{\tilde{x}}(k-j+1,k+1/k) - \sum_{j=o}^{M} A(k+1,k)\, H_j(k)$$

$$\cdot V_{\tilde{x}}(k-j,k+1/k) + \Psi_{vw}(k)\, G^T[x(k),k]\}$$

$$V_{\tilde{x}}(k+1/k) = \sum_{i,j=o}^{L} E_k\{\tilde{f}_i[x(k-i),k-i/k]\, \tilde{f}_j^T[x(k-j),k-j/k]\}$$

$$+ G[x(k),k]\, \Psi_w(k)\, G^T[x(k),k]$$

and

$$V_{\tilde{x}}(k+1,k-i/k) = \sum_{j=o} E_k\{\tilde{f}_j[x(k-j),k-j/k]\, \tilde{x}^T(k-i/k)\}$$

6.3.6 *Estimation in Linear Discrete Systems*

In this section, general linear estimation algorithms are presented, they are directly obtained from results of Section 6.3.2, satisfying the following relations

$$\sum_{j=o}^{L} f_j[x(k-j), k-j] = \sum_{j=o}^{L} F_j(k)\, x(k-j)$$

$$G[x(k),k] = G(k)$$

The results are:

$$\hat{x}(k-\ell+1/k) = \sum_{j=0}^{L} F_j(k-\ell)\ \hat{x}(k-\ell-j/k)$$

$$\hat{x}(k-\ell+1/k+1) = \hat{x}(k-\ell+1/k) + K^{\ell}_{k+1}\ \{[\text{illegible}])$$

$$- \sum_{i=0}^{L} \sum_{j=0}^{M} \hat{d}_{i,j}\ [\text{illegible}]),\ k-j-i/k]\}$$

$$K^{\ell}_{k+1} = \{\sum_{j=0}^{M} V_{\tilde{x}}(k-\ell+1,k-j+1/k)\ H_j^T(k+1)$$

$$- \sum_{j=0}^{M} V_{\tilde{x}}(k-\ell+1,k-j/k)\ H_j^T(k)\ A^T(k+1,k)$$

$$+ G(k-\ell)\ \Psi_{wv}(k-\ell)\ \delta_{k-\ell,k}\}$$

$$\{\sum_{i,j=0}^{M} H_i(k+1)\ V_{\tilde{x}}(k-i+1,k-j+1/k)\ H_j^T(k+1)$$

$$+ \sum_{i,j=0}^{M} A(k+1,k)\ H_i(k)\ V_{\tilde{x}}(k-i,k-j/k)\ H_j^T(k) A^T(k+1,k)$$

$$- \sum_{i,j=0}^{M} H_i(k+1)\ V_{\tilde{x}}(k-i+1,k-j/k)\ H_j^T(k)\ A^T(k+1,k)$$

$$- \sum_{i,j=0}^{M} A(k+1,k)\ H_i(k)\ V_{\tilde{x}}(k-i,k-j+1/k)\ H_j^T(k+1)$$

$$+ \Psi_v(k) + H_o(k+1)\ G(k)\ \Psi_{wv}(k)$$

$$+ \Psi_{vw}(k)\ G^T(k)\ H_o^T(k+1)\}^{-1}$$

$$V_{\tilde{x}}(k-m+1,k-\ell+1/k+1) = V_{\tilde{x}}(k-m+1,k-\ell+1/k)$$

$$- K^m_{k+1}\{\sum_{j=0}^{M} H_j(k+1)\, V_{\tilde{x}}(k-j+1,k-\ell+1/k)$$

$$- \sum_{j=0}^{M} A(k+1,k)\, H_j(k)\, V_{\tilde{x}}(k-j,k-\ell+1/k)$$

$$+ \Psi_{vw}(k-\ell)\, \delta_{k,k-\ell} G^T(k-\ell)\}$$

$$V_{\tilde{x}}(k-m+1,k-\ell+1/k) = \sum_{i=0}^{L} F_i(k-m)\, V_{\tilde{x}}(k-m-i,k-\ell-j/k)$$

$$F_j^T(k-\ell) + G(k-m)\, \Psi_w(k-m)\, \delta_{k-m,k-\ell} G^T(k-\ell)$$

$$V_{\tilde{x}}(k-m+1,k-\ell/k) = \sum_{=0}^{L} F_j(k-m)\, V_{\tilde{x}}(k-m-j,k-\ell/k)$$

The three different classes of smoothers as well as the filtering algorithms can be obtained from the above equations with the following substitution.

	Substitution for k	Substitution for ℓ
The filtering estimation	k	0
The fixed-lag smoothing	$k-1$	N
The fixed-point smoothing	k	$k-N+1$
The fixed-interval smoothing	$N-1$	$N-k$

The results presented here for linear estimation are new and recursive in nature. If the stochastic processes are Gaussian, the estimator is optimal for all reasonable cost functionals, and if they are not, then the algorithm is the linear estimator minimizing the error-variance.

6.4 *CONCLUSION*

In this chapter, a new unified approach has been presented to obtain the fixed-lag, fixed-point, fixed-interval smoothing as well as the filtering algorithms for nonlinear as well as linear delayed systems with non-delayed measurements as well as multi-channel delayed measurements corrupted by colored noise. The derivation is straightforward and shows the close link between various smoothers and the filtering estimator.

Except the problem of linear estimation for delayed systems with non-delayed measurement, all the results obtained are new. In that case the results agree well with those in the literature.

The results obtained are exact and optimal with respect to the given constraint on the dynamic equation of the smoother, minimizing the error-variance cost functional.

The results can be recursively evaluated under the assumption that the probability density functions of the estimator errors are Gaussian, they can be obtained without any further approximations for systems with product types or polynomial nonlinearities. They are expected to be computationally efficient, since there is no augmentation of state variables.

The results can also be extended to continuous time problems through a formal limiting procedure.

CHAPTER VII

CONCLUDING REMARKS

7.1 *Conclusions*

This thesis is devoted to the derivation of nonlinear estimation algorithms for discrete and continuous nonlinear systems with and without delays, given non-delayed measurements, as well as multichannel time-delayed measurements, corrupted by white Gaussian noise, correlated noise, colored noise as well as noise-free processes. Even though concluding remarks have been made in Chapters III to VI, it is the purpose of this section to summarize the main results of this dissertation.

In Chapter III, nonlinear filtering algorithms are derived for continuous nonlinear systems without delays corrupted by white noise, correlated noise and noise-free processes. For systems corrupted with additive white noise, the derivation is new and exact, in the case of non-white noise processes, the results o[illegible] new but an approximation has been made. The main techniq[illegible] the matrix minimum principle together with the Kolmogorov and Kushner equations to minimize the error-variance, taken to be the estimation criterion.

In order that the systems be physically realizable, the assumption that the probability density functions of the estimator errors are Gaussian is needed to expand the expectations involved in the estimation algorithms. Various nonlinear systems were simulated, the noise processes are assumed to be additive white, they are compared with results obtained from various other finite dimensional approximate nonlinear filters. The results clearly indicate the superiority of the proposed minimum variance filter over those of other filters investigated, and theoretical explanations are given for the apparent poor performance characteristics of various other filters considered. In the particular case of linear systems, the results agree well with those of the literature.

New nonlinear estimation algorithms for discrete nonlinear delayed systems, with non-delayed measurements as well as multi-channel time-delayed measurements corrupted by white noise, correlated noise and colored noise processes are respectively derived in Chapters IV, V and VI. The estimation algorithms directly yield the fixed-lag, fixed-point and fixed-interval smoothing and the filtering algorithms, with proper substitution of the discrete time indices. The proposed technique makes use of the concept of the gradient matrix, the minimum principle to derive the optimal values of the coefficients in the estimation algorithms under the requirements that the estimates be unbiased and the error-variances minimized. The derivation is straight-forward and clearly indicate the close link between three different classifications of smoothers and the filtering estimator.

The results obtained are exact and optimal with respect to the imposed constraints on the dynamic equations of the estimators, minimizing the error-variances. The algorithms derived can be implemented in the computer evaluation under the assumption that the probability density functions of the estimator errors are Gaussian. For systems with product types and polynomial nonlinearities, no further approximation is needed to evaluate the expectations involved in the algorithms. They are expected to be computationally efficient, since no augmentation of state variables are involved.

The results can also be applied to various special cases of nonlinear as well as linear estimation problems; for example: The estimation problems for linear or nonlinear systems without delays, the estimation of linear or nonlinear systems with multiple delays only appearing in the measurements and the linear estimation problems of time-delayed systems corrupted by white noise as well as non-white noise processes, etc.

For linear estimation problems the results obtained are identified with published results in literature. In the particular case of linear time-delayed systems corrupted by additive white noise, the results are stable, but the stability behaviour of the nonlinear estimators are yet to be investigated.

The results obtained in Chapter IV, V and VI can be extended to continuous time systems with formal limiting argument [4].

7.2 *Suggestions for Further Research*

A unified approach has been developed to derive estimation algorithms for various types of nonlinear estimation problems. But many more nonlinear estimation algorithms need to be derived for various other nonlinear systems; for example: noise processes involving delays, signals measured through jump processes, multistage correlated processes. Systems with jump effect or discontinuous nonlinearities. Besides problems considered in this thesis, estimation for continuous systems with discrete measurements are yet to be developed. Nonlinear smoothers for continuous nonlinear delayed systems with white noise and non-white noise processes are yet to be extended using the formal limiting argument.

Besides deriving estimation algorithms, one also needs to know more about the performance characteristics of the estimators, the sensitivity, divergence and modelling problems are of extreme importance. The problems of error compensation, modelling error and error detection, data editing, data compression and data ordering in real systems application are yet to be considered. One very important problem which requires immediate attention is the stability behaviour of the estimators developed in this thesis.

Extensive computer simulation also needs to be undertaken to test the estimators developed in Chapters IV to VI, and they are to be compared it other finite dimensional estimators, where one would have used the extended Kalman Filters to approximate the nonlinear estimation problems.

Since the estimation technique developed here is simple, straightforward and effective in dealing with wide-ranging nonlinear systems, one would immediately think of its parallel development to parameter estimation, identification and nonlinear mean-square control problems. Once can also think of the extension to estimation problems for distributive parameter systems.

Last but not least, the nonlinear estimation algorithms are bound to replace the role of linear estimators in the field of communications, guidance and control. Therefore, extensive research and development work is yet to be carried out to employ the nonlinear estimators to real systems applications.

BIBLIOGRAPHY

[1] N. Wiener, *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series With Engineering Applications.* New York: Wiley, 1949.

[2] A.N. Kolmogorov, "Interpolation and extrapolation of stationary random sequence," *Bull. Acad. Sci. U.S.S.R.*, Math. Ser., vol. 5, 1941.

[3] R.E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME, J. Basic Engrg.*, vol. 82D, pp. 34-45, March 1960.

[4] R.E. Kalman, "New methods in Wiener Filtering Theory," *Proc. 1st Symp. on Engrg. Appl. of Random Function Theory and Probability*, New York: Wiley, 1963.

[5] R.E. Kalman and R.S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME, J. Basic Engrg.*, vol. 83D, pp. 95-107, Dec. 1961.

[6] A.E. Bryson and D.E. Johansen, "Linear filtering for time-varying systems using measurements containing colored noise," *IEEE Trans. Aut. Control*, vol. AC-10, pp. 4-10, Feb. 1965.

[7] A.E. Bryson and L.J. Henrikson, "Estimation using sampled-data containing sequentially correlated noise," *Harvard University Div. Engrg. Appl. Phys. Tech. Report 533*, Cambridge, Mass., June 1967.

[8] A.R. Stubberud, *Optimal Filtering for Gauss-Markov Noise*. Aerospace Report No. TR-0158 (3307-01)-10, Dec. 1967.

[9] E.B. Stear and A.R. Stubberud, "Optimal filtering for Gauss-Markov noise," *Int. J. Control*, vol. 8, no. 2, pp. 123-130, 1968.

[10] P. Sarachik, "State estimation from measurements with correlated noise without using differentiators," *Int. J. Control*, vol. 13, no. 1, pp. 53-63, 1971.

[11] A.E. Bryson and M. Frazier, "Smoothing for linear and nonlinear dynamic systems," *Proc. Optimum System Synthesis Conf.*, U.S. Air Force Tech. Report ASD-TDR-063-119, Feb. 1963.

[12] H.E. Rauch, F. Tung and C.T. Striebel, "Maximum likelihood estimate of linear dynamic systems," *AIAA J.*, vol. 3, pp. 1445-1455, August 1965.

[13] J.S. Meditch, *Stochastic Optimal Linear Estimation and Control*. New York: McGraw-Hill, 1969.

[14] T. Kailath, "An innovations approach to least-squares estimation-Part I: Linear filtering in additive white noise," *IEEE Trans. Aut. Control*, vol. AC-13, pp. 646-655, Dec. 1968.

[15] T. Kailath and P. Frost, "An innovations approach to least-squares estimation- Part II: Linear smoothing in additive white noise," *IEEE Trans. Aut. Control*, vol. AC-13, pp. 655-660, Dec. 1968.

[16] R.L. Stratonovich, "Conditional Markov process theory," *Theory Prob. Appl.* (U.S.S.R.), vol. 5, pp. 156-178, 1960.

[17] H.J. Kushner, "On differential equations satisfied by conditional probability densities of Markov processes," *SIAM J. Control*, vol. 2, pp. 106-119, 1964.

[18] H.J. Kushner, "On the dynamical equations of conditional probability density functions with applications to optimal stochastic control theory," *J. Math. Anal. Appl.*, vol. 8, pp. 332-344, 1964.

[19] W.M. Wonham, "Some applications of stochastic differential equations to optimal nonlinear filtering," *SIAM J. Control*, vol. 2, pp. 347-369, 1965.

[20] R.S. Bucy, "Nonlinear filtering," *IEEE Trans. Aut. Control*, vol. AC-10, p. 198, April 1965.

[21] R.E. Mortensen, "Optimal control of continuous time stochastic systems," Ph.D. dissertation, University of Calif., Berkeley, Calif., 1966.

[22] H.J. Kushner, "Dynamical equations for optimum nonlinear filtering," *J. Differential Equations*, vol. 3, pp. 179-190, 1967.

[23] H. Cox, "On the estimation of state variables and parameters for noisy dynamic systems," *IEEE Trans. Aut. Control*, AC-9, pp. 5-12, Feb. 1964.

[24] D.M. Detchmendy and R. Sridhar, "Sequential estimation of states and parameters in noisy nonlinear dynamical systems," *Trans. ASME, J. Basic Engrg.*, vol. 88D, pp. 362-368, June 1966.

[25] B. Friedland and I. Bernstein, "Estimation of the state of a nonlinear process in the presence of non-Gaussian noise and disturbances," *J. Franklin Inst.*, vol. 281, pp. 455-480, 1966.

[26] R.W. Bass, V.D. Norum and L. Schwartz, "Optimal multi-channel nonlinear filtering," *J. Math. Anal. Appl.*, vol. 16, pp. 152-164, 1966.

[27] H.J. Kushner, "Approximations to optimal nonlinear filters, *IEEE Trans. Aut. Control*, vol. AC-12, pp. 546-556, Oct. 1967.

[28] L. Schwartz and E.B. Stear, "A computational comparison of several nonlinear filters," *IEEE Trans. Aut. Control*, vol. 13, pp. 83-86, Feb. 1968.

[29] Y. Sunahara, "An approximate method of state estimation for nonlinear dynamical systems," *Trans. ASME, J. Basic Engrg.*, vol. 92D, pp. 385-393, June 1970.

[30] A.P. Sage and W.S. Ewing, "Approximate solution algorithms for nonlinear filter problems," Proc. Second Princeton Conf. Inf. Sci. Systems, pp. 323-327, March 1968.

[31] A.P. Sage and W.S. Ewing, "On filtering and smoothing algorithms for nonlinear state estimation," *Int. J. Control*, vol. 11, no. 1, pp. 1-18, 1970.

[32] A.P. Sage, "Maximum a-posteriori filtering and smoothing algorithms," *Int. J. Control*, vol. 11, no. 3, pp. 171-183, 1970.

[33] D.G. Lainiotis, "Optimal nonlinear estimation," *Int. J. Control*, vol. 14, no. 6, 1137-1148, 1971.

[34] G.M. Lee, "Nonlinear interpolation," *IEEE Trans. Inf. Theory*, vol. IT-17, pp. 45-49, January 1971.

[35] C.T. Leondes, J.B. Peller and E.B. Stear, "Nonlinear smoothing thoery," *IEEE Trans. System Sci. Cyb.*, vol. SSC-6, pp. 63-71, January 1970.

[36] H. Kwakernaak, "Optimal filtering in linear systems with time delays," *IEEE Trans. Aut. Control*, vol. AC-12, pp. 169-173, April 1967.

[37] R. Priemer and A.G. Vacroux, "Estimation in linear discrete systems with multiple time delays," *IEEE Trans. Aut. Control*, vol. AC-14, pp. 384-387, August 1969.

[38] R. Priemer and A.G. Vacroux, "Smoothing in linear discrete systems with time delays," *Int. J. Control*, vol. 13, no. 2, pp. 299-303, 1971.

[39] M. Farooq and A.K. Mahalanabis, "A note on the maximum-likelihood state estimation of linear discrete systems with multiple time delays," *IEEE Trans. Aut. Control*, vol. AC-16, pp. 104-105, Feb. 1971.

[40] K.K. Biswas and A.K. Mahalanabis, "Optimal smoothing for continuous-time systems with multiple time delays," *IEEE Trans. Aut. Control*, vol. AC-17, pp. 572-574, August 1972.

[41] M. Athan, "The matrix minimum principle," *J. Inf. Control*, vol. 11, pp. 592-606, 1968.

[42] E.B. Stear, "Synthesis of shaping filters for nonstationary stochastic processes and their uses," University of California, Report 61-50, Los Angeles, California, 1961.

[43] J.S. Meditch, "On optimal linear smoothing theory," *J. Inf. Control*, vol. 10, pp. 598-615, 1967.

[44] F.C. Schweppe, *Uncertain Dynamic Systems*, New Jersey : Prentice Hall, 1973.

[45] L.T. Fan and C.S. Wang, *The Discrete Maximum Principle*, New York: Wiley, 1964.

[46] K. Ito, "On stochastic processes," Lecture Notes, Tata Inst. for Fundamental Research, Bombay, 1961.

[47] D.F. Liang and G.S. Christensen, "Exact and approximate nonlinear filters involving additive noise, correlated noise or noise - free processes," Preprints of Canadian Conference on Automatic Control, Univ. of New Brunswick, Fredericton, 1973.

[48] D.F. Liang and G.S. Christensen, "An approach to exact and approximate state estimation for nonlinear dynamic systems," submitted to *IEEE Trans. Aut. Control*, 1974.

[49] A.P. Sage and J.L. Melsa, *Estimation Theory with Applications to Communications and Control*. New York: McGraw-Hill, 1971.

[50] A.G. Carlton, "Linear estimation in stochastic processes," John Hopkins University, Appl. Phys. Lab., Baltimore, Md., Internal Rept., 1962.

[51] H.E. Rauch, "Solutions to the linear smoothing problem," *IEEE Trans. Aut. Control*, vol. AC-8, pp. 371-372, Oct. 1963.

[52] C.N. Kelly and B.D.O. Anderson, "On the stability of fixed-lag smoothing algorithms," *J. Franklin Inst.*, vol. 291, pp. 271-281, 1971.

[53] D.F. Liang and G.S. Christensen, "New filtering and smoothing algorithms for discrete nonlinear systems with time delays," To appear in *Int. J. Control*.

[54] D.F. Liang and G.S. Christensen, "New estimation algorithms for discrete nonlinear delayed systems and observations," To appear in *IFAC Journal, Automatica.*

[55] K.K. Biswas and A.K. Mahalanabis, "An approach to fixed-point smoothing problems," *IEEE Trans. Aerospace and Elect. Systems,* vol. AES-8, pp. 676-680, Sept. 1972.

[56] V. Shukla and M.D. Srinath, "Sequential estimation in linear systems with multiple time delays," *J. Inf. Control,* vol. 22, pp. 471-486, 1973.

[57] B.V. Raja Rao and A.K. Mahalanabis, "Estimation in linear delayed discrete-time systems with correlated state and measurement noises," *IEEE Trans. Aut. Control,* vol. 16, p. 267, June 1971.

[58] D.F. Liang and G.S. Christensen, "Comments on estimation in linear delayed discrete-time systems with correlated state and measurement noises," submitted to *IEEE Trans. Aut. Control,* 1974.

[59] D.F. Liang and G.S. Christensen, "New estimation algorithms for discrete nonlinear delayed systems with correlated noises," submitted to *J. Franklin Institute,* 1974.

[60] D.F. Liang and G.S. Christensen, "Estimation of discrete nonlinear delayed systems and observations with correlated noises," submitted to *IEEE Trans. Aut. Control,* 1974.

[61] K.K. Biswas and A.K. Mahalanabis, "Optimal fixed-lag smoothing for time delayed systems with colored noise," *IEEE Trans. Aut. Control*, vol. 17, pp. 387-388, June 1972.

[62] D.F. Liang and G.S. Christensen, "Estimation of discrete non-linear delayed systems with colored noise," submitted to *IEEE Trans. Aut. Control*, 1974.

[63] D.F. Liang and G.S. Christensen, "New filtering and smoothing algorithms for discrete nonlinear delayed system with colored noise," submitted to *Int. J. Control*, 1974.