



National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service Services des thèses canadiennes

Ottawa, Canada
K1A 0N4

CANADIAN THESES

THÈSES CANADIENNES

NOTICE

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

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

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

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

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30.

**THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED**

AVIS

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

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

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

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

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30.

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

**FILTERING AND IDENTIFICATION OF NONLINEAR
STOCHASTIC DIFFERENTIAL SYSTEMS**

By

TAYEL ESSAWY DABBOUS

A Thesis
Presented to the University of Ottawa
in Partial Fulfilment of the
Requirements for the Degree of
Doctor of Philosophy
In
Electrical Engineering
Department of Electrical Engineering
Faculty of Science and Engineering

Permission has been granted to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film.

The author (copyright owner) has reserved other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without his/her written permission.

L'autorisation a été accordée à la Bibliothèque nationale du Canada de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur (titulaire du droit d'auteur) se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation écrite.

ISBN 0-315-36478-5



UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

The University of Ottawa requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

TO
ALMIGHTY GOD
THE CREATOR
THE MERCIFUL
THE FORGIVER

ACKNOWLEDGEMENTS

The author expresses his deepest gratitude to his advisor Professor N.U. Ahmed for his enormous patience, generous encouragement, invaluable guidance and numerous hours of discussions through this work, without which this thesis would have not been possible.

The author would also like to thank Dr. D.F. Liang and Mr. J.C. McMillan, Defence Research Establishment Ottawa, for their help, support and suggestions during the modelling of Navigation System Errors.

Special thanks also are due to Dr. S.K. Biswas, Mr. S.S. Lim and Mr. Y.W. Lee for their encouragements and long discussions.

The author wishes to express his appreciation to the faculty and staff of the Department of Electrical Engineering, University of Ottawa, for their kindness and support.

Financial assistance provided by the Department of Electrical Engineering, University of Ottawa, and by the National Science and Engineering Research Council of Canada during the period of this research, is gratefully acknowledged.

Finally, the author is highly indebted to his wife Nevin Dabbous for her enormous patience, constant support and encouragement.

ABSTRACT

Filtering and identification problems of partially observable stochastic dynamical systems has been considered. A complete derivation of the Zakai equation for partially observed diffusions has been presented. Utilizing the properties of the likelihood ratio and using Ito differential rule, the Kushner result was obtained from that of Zakai and vice versa. Using a Radon-Nikodym approach and the properties of the semigroup of the state processes, the result of Zakai has been extended to wider classes of stochastic differential systems.

Modelling of the errors that arise in Dead Reckoning and Loran-C navigation sensors has been also investigated. It has been shown that these errors can be modelled as a linear stochastic differential equation driven by Wiener and Poisson processes. Based on these models, we have utilized the filter equations obtained for the nonlinear case, to derive a set of finite dimensional differential equations for the estimator and covariance matrix from which velocity and position of the ship can be estimated. The results are illustrated by a numerical example.

Techniques of optimal control Theory as well as nonlinear filter Theory have been utilized in identifying the drift parameters of nonlinear (partially observable) stochastic differential systems with and without Markov jump coefficients. Using the variational methods and the maximum likelihood approach, the necessary conditions for optimal identification have been derived. The results are illustrated by two (scalar) examples.

CONTENTS

ACKNOWLEDGEMENTS	v
ABSTRACT	vi
CHAPTER	<u>PAGE</u>
1. INTRODUCTION	
1.1 A Brief Review of Previous Studies	2
1.1.1 Filtering Problem	2
1.1.2 Modelling and Filtering of Navigation Errors	4
1.1.3 Identification Problem	5
1.2 Outline of the Thesis	6
2. NONLINEAR FILTERING OF DIFFUSION PROCESSES	
2.1 Introduction	8
2.2 Formulation of Filtering Problem, Assumptions and Notations	10
2.3 An Evolution Equation For Unnormalized Density	14
2.4 Kushner Equation and its Relation to Zakai Equation	29
2.5 Remarks on Existence and Uniqueness of Zakai Equation	34
2.6 Summary and Conclusion	37
3. NONLINEAR FILTERING OF SYSTEMS GOVERNED BY ITO EQUATIONS WITH JUMP PARAMETERS	
3.1 Introduction	38
3.2 Formulation of Filtering Problem, Assumptions and Notations	39
3.3 Stochastic Integral Equation For Unnormalized Density	44
3.4 Zakai Type Equation For Unnormalized Density	53
3.5 Remarks on The Existence of Solutions of Filter Equation	58
3.6 Examples	62
3.7 Summary and Conclusion	65
4. NONLINEAR FILTERING OF DISCONTINUOUS PROCESSES WITH DISCONTINUOUS OBSERVATIONS	
4.1 Introduction	66
4.2 Formulation of Filtering Problem, Assumptions and Notations	67
4.3 Stochastic Integral Equation For Unnormalized Density	72
4.4 An Evolution Equation For Unnormalized Density	82
4.5 Nonlinear Filtering Formula For Discrete Observations	87
4.6 Examples	89

8

4.7 Summary, Conclusion and Some Open Questions	94
5. MODELLING AND FILTERING OF ERRORS ARISING IN NAVIGATION SYSTEMS	
5.1 Introduction	96
5.2 Dead Reckoning, Loran-C and Observation Models	99
5.2.1 Dead Reckoning Error Model	99
5.2.2 Loran-C Error Model	99
5.2.3 Observation Model	105
5.3 Linear Unbiased Filter For Navigation System	107
5.4 Computational Algorithm and an Example	116
5.5 Summary and Conclusion	150
6. APPLICATION OF NONLINEAR FILTERING TO PARAMETER IDENTIFICATION PROBLEM	
6.1 Introduction	152
6.2 Formulation of Identification Problems, Assumptions and Notations ..	154
6.3 Existence Theory For Identification Problems	158
6.3.1 Identification Problem (I1)	158
6.3.2 Identification Problem (I2)	165
6.3.3 Identification Problem (I3)	169
6.4 Necessary Conditions For Optimal Identification	176
6.4.1 Identification Problem (I1)	175
6.4.2 Identification Problem (I2)	181
6.4.3 Identification Problem (I3)	185
6.5 Computational Algorithm and Examples	190
6.6 Summary and Conclusion	197
7. CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH	
7.1 Conclusions	198
7.2 Suggestions For Further Research	200
REFERENCES	201

CHAPTER 1

INTRODUCTION

Physical systems are designed and built to perform certain defined functions. For example, submarine, aircraft, and spacecraft must navigate in their respective environment to accomplish their objectives; whereas an electric power system network must meet the load demand. In order to determine that whether a system is performing properly, and ultimately one can control the system performance, the engineer must know the state of the system. In a navigation system, for example, the state consists of the position and velocity of the craft in question ; in an electric power system, the state may be taken as the number of generators in operation or the voltages and phase angles at network nodes.

In most cases, physical systems are subjected to random disturbances, so that the state of the system may itself be random. In order to determine the state of the system, the engineer should build a measuring device and take measurements (or observations) on his system. These measurements are generally contaminated with noise caused by the electronic and mechanical components of the measuring device. The problem of determining the state of the system from these noisy measurements is called estimation problem. There are three types of estimations, which can be stated as follows. When the time at which an estimate is desired falls within the span of available measurement data, the problem is termed smoothing. If the time of interest occurs after the last measurement, the problem is referred to as prediction. And if the time of interest coincides with the last available measurement point, the problem is called filtering, which is the main subject of this thesis. In fact the filtering problem is of central importance in many practical engineering problems, since the estimated states are required in monitoring, and for the control of the

systems. Furthermore, a large class of (system) identification problems can also be regarded as filtering problems [45].

1.1 A Brief Review of Previous Studies

1.1.1 Filtering Problem

Over the last two decades, much attention has been focused on the study of filtering problem of linear as well as nonlinear dynamical systems [19, 20, 26, 27, 29, 30, 33, 42, 45, 46, 47, 49, 49, 55, 61, 62, 67, 76]. In 1960, Kalman and Bucy [49] used least square method to derive a system of linear stochastic and ordinary differential equations describing the evolution of the estimated state and the corresponding covariance matrix, respectively. In fact these filter equations were extensively used in practice for many engineering problems (see for example [45, 59, 60, 64]). In 1960, Stratonovich [69] was also pioneering the development of the probabilistic approach to nonlinear filtering problem. The work of Stratonovich was not, however, immediately known in the West. From 1963-1966, nonlinear filtering and stochastic control problems were studied by Wonham [74], Kushner [53,54], Ho and Lee [41] and Jazwiniski [42,43]. In 1967, Kushner [55,56] obtained his equation, which is a nonlinear stochastic partial differential equation describing the flow of normalized conditional density of diffusion Markov processes, with the help of which the whole filtering problem is theoretically resolved. A more theoretically convenient filter equation was obtained by Zakai [76] in 1969. This equation is a linear stochastic partial differential equation describing the flow of unnormalized conditional density of a diffusion Markov process.

After this pioneering work of Kushner and Zakai, the nonlinear filtering problem for diffusion processes was treated by many authors. In references [18, 51, 57, 58],

the problem of existence of solutions of Zakai and Kushner equations was considered. In [47] Kallianpur and Karandikar used the Radon Nikodym derivative result obtained by Balakrishnan [18] and finitely additive white noise approach to derive a linear partial differential equation for the unnormalized conditional density in which the observation process (in finitely additive set up) occurs as a parameter. They also proved the existence and uniqueness of solutions of this filter equation. In [70] Takeuchi and Akashi used Girsanov transformation to derive the Kushner equation for a class of systems driven by continuous martingales whereas the observed process (discrete) is excited by white Gaussian noise.

The extension of Kushner and Zakai results for the case where the state as well as the observed processes are discontinuous, was also treated in [25, 26, 31, 63, 67, 71, 72, 75]. In [31] Di Masi and Runggaldier considered the filtering problem for diffusion processes with observations driven by Wiener and doubly stochastic Poisson processes. Using a Girsanov type transformation, they obtained Zakai type equation for the unnormalized conditional density, the solution of which can be only obtained if the Poisson noise is observable. In [63] Liptser and Shiriyayev obtained Kushner type equation for systems driven by right continuous martingales whereas the observation is doubly stochastic Poisson process. Pardoux [67] obtained Zakai type equation for the case where the state process is a diffusion Markov process and the observation is doubly stochastic Poisson process. In [25,26] Boel, Varaiya and Wong, used martingale Theoretic approach and Girsanov type transformation to obtain Zakai type equation when the state as well as the observed processes are both driven by two independent right continuous martingales. This result is very general and hence very abstract in nature. Vaca and Snyder [71,72] also used martingale approach to derive Kushner type equation when the state process is driven by a continuous martingale whereas the output (observed) process is driven

by Wiener as well as generalized Poisson processes. Again this filter equation is very abstract in nature and in order to compute the required estimate one has to observe the continuous as well as discontinuous noise of the output process which is practically difficult. In [73] Yavin considered the filtering problem for the case where the state process is of diffusion type and the observed process is driven only by generalized Poisson process. Under the assumption that the estimator is governed by Kushner type equation with unknown gain matrix, he used optimal control Theory to compute this gain matrix through which the required estimate can be determined.

In [40], Grigelionis considered the filtering problem for a more general class of stochastic systems. This result cover those obtained in [25, 26, 28, 31, 63, 67, 71, 72, 75]. However, for the case where the observed process is of discrete type (which is the usual practice), computing the required estimate, using Grigelionis result, is difficult if not impossible. This problem does not arise in the filter equation developed by the author in [28], as it will be indicated in chapter 4 of this thesis.

The filter equations of Kushner and Zakai have also been applied to stochastic control problems of partially observable diffusion process by many authors such as Benes [21], Fleming [34]-[36], Kushner [53,57] where the existence of optimal control was proved given the observed path.

1.1.2 Modelling and Filtering of Navigation Errors

Over the last few years the problem of modelling and filtering of errors that arise in Marine Integrated Navigation systems was considered by many authors (see for example [1, 59, 60, 64]) and their work was mainly concentrated on modelling these navigation sensor errors as a linear diffusion Markov processes. Then assuming that the observed process (discrete) is excited by a white Gaussian noise, they used

Kalman filter to obtain an estimate of ship (or aircraft) position and velocity. In most circumstances, these models, and hence the corresponding filter, give quite satisfactory results. However, there are some well known errors, particularly those of Loran-C and Omega radio aids, that do not conform to these standard model and therefore warrant further investigation.

1.1.3 Identification Problem

An important and essential aspect of modelling any physical system is the identification of parameters in the model equation. For distributed parameter systems this problem has received considerable attention in the past as indicated in [16] and [50]. Techniques of optimal control Theory have also been utilized in identification of parameters in distributed systems. Optimality conditions for controls in the coefficients of a class of hyperbolic systems have been developed by Ahmed [7]. Semigroup Theory and maximum likelihood ratio approach were used in [17] to identify unknown parameters (or operators) for deterministic distributed systems as well as stochastic systems. A general formulation for identification of operators in systems governed by parabolic, hyperbolic and structurally damped hyperbolic evolution equations has been discussed recently in [12]. Liptser and Shiriyayev [63] have also considered the identification problem of drift parameters of systems governed by Ito equation of the form

$$dx_t = \alpha h(x_t)dt + dW_t,$$

where W is a standard Wiener process and α is unknown. Using maximum likelihood approach, they obtained an explicit expression for the estimated parameter given $\{x_s; s \leq t\}$. Further, using the law of iterated logarithm of Brownian motion, they showed that as $t \rightarrow \infty$, the estimated parameter converges (in the almost sure sense) to the true underlying parameter.

In [39] Gland has considered the identification problem for drift parameters for a class of systems governed by nonlinear Ito equation of the form

$$dy_t = h(\alpha, x_t)dt + dV_t,$$

where $x_t; t \in [0, T]$, is a diffusion Markov process, V is a standard Wiener process and α is unknown. Using maximum likelihood approach and forward and backward Zakai equations, he developed a numerical scheme for computing the unknown parameter given the observed process $\{y_s; s \leq t\}$.

1.2 Outline of The Thesis

The thesis is organized as follows : Chapter 1 contains the motivation for filtering and identification problems and a brief review of the previous studies in these areas.

In chapter 2 we state the filtering problem and present a brief derivation for Zakai equation. Further, using the result of Zakai, we obtain Kushner equation and show their equivalence. Finally, we discuss the question of existence and uniqueness of solutions of Zakai equation using the standard result on existence of solutions of partial differential equations.

In chapter 3 we follow similar procedure as that of chapter 2 to extend the result of Zakai to the case where the drift and diffusion parameters of the state process are perturbed by a temporally homogeneous Markov chain. We also discuss the question of existence and uniqueness of solutions of this filter equation.

In chapter 4 we make use of Radon-Nikodym derivative of measures induced by discontinuous Markov processes, to derive Zakai type equation for the case where the state as well as the observed processes are driven by Wiener and generalized Poisson

processes. We also obtain the corresponding filter equation when the observed process (discrete) is only excited by white Gaussian noise.

In chapter 5 we consider the problem of modelling the errors that arise in Dead Reckoning and Loran-C navigation sensors. Then using the results of chapter 4, we obtain the corresponding filter equations.

In chapter 6 we consider identification problem for stochastic systems governed by nonlinear Ito differential equations with and without jump parameters. Then using the results of chapters 2 and 3, we prove the existence of solutions of these identification problems and obtain the corresponding optimality conditions, along the line of [12] and [13].

Concluding remarks and suggestions for further research are presented in chapter 7.

Original contributions in this thesis include :

- (i) A complete derivation of the filter equations for a class of systems governed by nonlinear Ito differential equations with drift and diffusion parameters perturbed by a jump process; sections 3.2, 3.3 and 3.4. This is independent of the general formulation in [40].
- (ii) A complete derivation of the filter equation for a class of systems driven by Wiener and generalized Poisson processes ; sections 4.2, 4.3 and 4.4.
- (iii) Modelling and filtering of jump processes arising in navigation system ; sections 5.2 and 5.3.
- (iv) Parameter identification of partially observed nonlinear diffusion processes with and without Markov jump coefficients ; sections 6.3 and 6.4 .

CHAPTER 2

NONLINEAR FILTERING OF DIFFUSION PROCESSES

2.1 Introduction

In this chapter we consider the filtering problem for a class of systems governed by Ito differential equations of the form

$$\begin{aligned} dx_t &= a(x_t)dt + b(x_t)dW_t, \\ x(0) &= x_0, \end{aligned} \tag{2.1}$$

where x, a are vectors in R^n ; b is $(n \times n)$ matrix and W is n -dimensional standard Wiener process independent of the initial state x_0 . The output (observed) process $y_t, t \geq 0$, is assumed to be related to the state process $x_t, t \geq 0$, through the following stochastic differential equation

$$\begin{aligned} dy_t &= h(x_t)dt + \sigma_0(y_t)dV_t, \\ y(0) &= 0, \end{aligned} \tag{2.2}$$

where y, h are vectors in R^m , σ_0 is an $(m \times m)$ matrix and V is an m -dimensional standard Wiener process independent of W and the initial state x_0 .

As indicated earlier in chapter 1, the filtering problem for the above system was considered in the literature by Stratonovich [69], Kushner [53-58], Bucy [27], Liptser and Shirayev [63], Zakai [76], Kunita [51,52], Davis [29,30], Kallianpur [46], and others and their work was concentrated mainly on finding an equation for the conditional density (normalized or unnormalized) of the process $x_t, t \geq 0$, given the observed path $\{y_s; s \leq t\}$, with the help of which the filtering problem is resolved.

Since the overall thesis is mainly based on the extension and application of the Zakai equation, we present in this chapter a brief derivation for this equation and show how it is related to Kushner equation.

In section 2.2, we state the filtering problem and introduce the assumptions and notations that have been used in the sequel. In section 2.3, we present some known results due to Girsanov and Ito, on the basis of which we derive a stochastic integral equation for the unnormalized conditional density. Using this integral equation and utilizing the properties of the semigroup corresponding to the state (process) x , we derive the Zakai equation. In section 2.4, we utilize the results of section 2.3 to derive the Kushner equation and show its equivalence to that obtained by Zakai. Finally, in section 2.5 we use the standard result on existence of solutions of partial differential equations given in [6,22], to discuss the question of existence and uniqueness of solutions of the Zakai equation.

2.2 Formulation of Filtering Problem, Assumptions and Notations

Let the state and observed processes be governed by the stochastic differential equations (2.1) and (2.2), respectively. Let f be any bounded measurable function on R^n with values in R^k . Let \mathcal{F}_t^y denote the σ -algebra generated by the process y up to time t . We wish to estimate $f(x_t)$ given the history of y up to time t , that is \mathcal{F}_t^y . Assuming that all the random processes and vectors described above are defined on a complete probability space $(\Omega_0, \mathcal{B}_0, P_0)$, we state the filtering problem as follows.

Given any bounded measurable R^k -valued function f on R^n , find \tilde{f}_t , which is \mathcal{F}_t^y -measurable, such that

$$E\{|f(x_t) - \tilde{f}_t|^2 \mid \mathcal{F}_t^y\}, \quad (2.3)$$

is minimum. Defining

$$\hat{f}_t \equiv E\{f(x_t) \mid \mathcal{F}_t^y\}, \quad (2.4)$$

equation (2.3) can be rewritten as

$$\begin{aligned} E\{|f(x_t) - \tilde{f}_t|^2 \mid \mathcal{F}_t^y\} &= E\{|f(x_t) - \hat{f}_t|^2 + 2(\hat{f}_t - \tilde{f}_t) \cdot (f(x_t) - \hat{f}_t) \\ &\quad + |\hat{f}_t - \tilde{f}_t|^2 \mid \mathcal{F}_t^y\}, \end{aligned} \quad (2.5)$$

where \cdot denotes the scalar product in R^k and $|\cdot|$ is the usual Euclidean norm. Using the definition (2.4) and the fact that the functions \tilde{f}_t and \hat{f}_t are \mathcal{F}_t^y -measurable, it follows from equation (2.5) that

$$E\{|f(x_t) - \tilde{f}_t|^2 \mid \mathcal{F}_t^y\} = E\{|f(x_t) - \hat{f}_t|^2 \mid \mathcal{F}_t^y\} + |\hat{f}_t - \tilde{f}_t|^2. \quad (2.6)$$

Clearly, it follows from (2.6) that the quantity, given by (2.3), attains its minimum for

$$\tilde{f}_t = \hat{f}_t \equiv E\{f(x_t) \mid \mathcal{F}_t^y\}. \quad (2.7)$$

This shows that the conditional expectation gives the best (optimal) estimate in the mean square sense. Since

$$\begin{aligned}\hat{f}_t &\equiv E\{f(x_t)|\mathcal{F}_t^y\} = \int_{R^n} f(x) \Pr\{x_t \in dx|\mathcal{F}_t^y\} \\ &\equiv \int_{R^n} f(x) P(t, dx|\mathcal{F}_t^y) \\ &= \int_{R^n} f(x) p(t, x|\mathcal{F}_t^y) dx,\end{aligned}\tag{2.8}$$

where $P(t, dx|\mathcal{F}_t^y)$ is a measure-valued \mathcal{F}_t^y -measurable random process, the whole filtering problem is resolved if the measure $P(t, dx|\mathcal{F}_t^y)$ (or the density $p(t, x|\mathcal{F}_t^y)$) can be computed. In fact Kushner [55] obtained a nonlinear stochastic partial differential equation for the conditional density $p(t, x|\mathcal{F}_t^y)$. This equation is given by

$$\begin{aligned}dp(t, x|\mathcal{F}_t^y) &= A^*p(t, x|\mathcal{F}_t^y)dt + (h - \hat{h}_t) \cdot \Gamma_0^{-1}\{dy_t - \hat{h}_t dt\}p(t, x|\mathcal{F}_t^y), \\ p(0, x) &= p_0(x),\end{aligned}\tag{2.9}$$

for all $t \in [0, T]$, where \cdot denotes the scalar product, $\hat{h}_t \equiv E\{h(x_t)|\mathcal{F}_t^y\}$ and A^* is the formal adjoint of the operator A given by

$$\begin{aligned}Af &= (a \cdot f_x) + \frac{1}{2}tr(bb'f_{xx}), \\ \Gamma_0 &\equiv \sigma_0\sigma_0',\end{aligned}$$

and p_0 is the initial density of x_0 . Further, $tr(A)$ denotes the trace of the matrix A .

Theoretically, a more convenient equation was obtained by Zakai [76] which is linear stochastic differential equation describing the flow of unnormalized conditional density, denoted $\varphi(t, \cdot)$, of x_t given \mathcal{F}_t^y , $t \geq 0$. This equation is given by

$$\begin{aligned}d\varphi(t, x) &= A^*\varphi(t, x)dt + \varphi(t, x)h(x) \cdot \Gamma_0^{-1}dy_t, \quad t \geq 0, \\ \varphi(0, x) &= p_0(x).\end{aligned}\tag{2.10}$$

Since the conditional density $p(t, x | \mathcal{F}_t^y)$ is related to the unnormalized conditional density $\varphi(t, x)$, $x \in R^n$, through the relation (see section 2.3)

$$p(t, x | \mathcal{F}_t^y) = \frac{\varphi(t, x)}{\int_{R^n} \varphi(t, x) dx}, \quad (2.11)$$

it is clear that solving (2.10) and using the relation (2.11), one obtains the conditional density $p(t, \cdot | \mathcal{F}_t^y)$, $t \geq 0$, and hence the estimate \hat{f}_t can be computed using equation (2.8). In the next section we shall use Girsanov transformation [38] of measures induced by systems (2.1) and (2.2) to derive Zakai equation (2.10). For this, we need the following assumptions and notations.

Assumptions:

- (A1) The functions $a(x)$ and $b(x)$ are continuous on R^n and there exists a constant $\gamma > 0$ such that

$$\sigma \equiv (bb') \geq \gamma I,$$

where I denotes the identity matrix.

- (A2) There exists a constant $k > 0$ such that

$$T|a(x) - a(y)|^2 + \|b(x) - b(y)\|^2 \leq k|x - y|^2,$$

for any $x, y \in R^n$.

- (A3) The functions $a_i, \sigma_{ij}, \frac{\partial}{\partial x_i} a_i$, $\frac{\partial}{\partial x_i} \sigma_{ij}$ and $\frac{\partial^2}{\partial x_i \partial x_j} \sigma_{ij}$; $i, j = 1, 2, \dots, n$, are bounded and satisfy Holder condition on R^n .

- (A4) The function h is continuous on R^n and there exists a constant $k > 0$ such that

$$|h(x) - h(y)|^2 \leq k|x - y|^2, \quad x, y \in R^n,$$

and

$$E \int_0^t |h(x_s)|^2 ds < \infty, \quad P_0 - a.s. \text{ (almost surely),}$$

along any solution $x_t; t \geq 0$, of equation (2.1).

(A5) The matrix function σ_0 is continuous, bounded and satisfies uniform Lipschitz condition on R^m (see assumption (A2)). Further, there exist constants $\delta > 0$ such that $\sigma_0(y) \geq \delta I$, $y \in R^m$, and

$$\gamma_1 |\xi|^2 \leq ((\sigma_0 \sigma_0'(y) \xi, \xi) \leq \gamma_0 |\xi|^2, \quad \xi \in R^m.$$

Notations:

Let $\{\eta_t, t \geq 0\}$ be any random process and let $\sigma\{\eta_s, s \leq t\}$ denote the σ -field generated by the process η up to time t . Define $\mathcal{F}_t^y \equiv \sigma\{y_s, s \leq t\}$, $\mathcal{F}_t^x \equiv \sigma\{x_s, s \leq t\}$, $\mathcal{F}_t^W \equiv \sigma\{W_s, s \leq t\}$, $\sigma(x_t) \equiv \sigma\{x_t\}$ and $\mathcal{F} \equiv \mathcal{F}_T^W \vee \mathcal{F}_T^y \vee \sigma(x_0) \subset \mathcal{B}_0$. Let $\mathcal{F}_t, t \geq 0$, be an increasing family of sub- σ -fields contained in \mathcal{F} such that for each $t \in [0, T]$, the processes x_t and y_t are \mathcal{F}_t -measurable. Let Ω denote the space of continuous functions on $[0, T]$ with values in R^{n+m} , and let \mathcal{A} denote the Borel σ -algebra on Ω . We call (Ω, \mathcal{A}) the canonical sample space. We denote by $\mathcal{B}(R^n)$ the Borel σ -field generated by subsets of R^n . Let $L_2(I; R^n); I \equiv [0, T]$, denote the equivalence classes of measurable functions f on I with values in R^n such that

$$\int_I |f(t)|^2 dt < \infty,$$

where $|\cdot|$ denotes the usual Euclidean norm. Clearly, the space $L_2(I; R^n)$ furnished with the norm topology

$$\|f\|_{L_2} \equiv \left(\int_I |f(t)|^2 dt \right)^{\frac{1}{2}},$$

is a Banach space. For any Banach space E , we denote by $C(I; E)$ the space of strongly continuous E -valued functions on I , furnished with the uniform topology

$$\|f\|_C = \sup_{t \in I} \|f(t)\|_E.$$

Note that the space $C(I; E)$ furnished with the norm topology $\|\cdot\|_C$, is a Banach space. For any pair of Banach spaces E and F , we use $\mathcal{L}(E, F)$ to denote the space of bounded linear operators from E to F . For any vectors $a, b \in R^n$, we use $(a \cdot b)$ to denote the scalar product in R^n . Further notations will be introduced in the sequel as required.

2.3 An Evolution Equation For Unnormalized Density

In this section we use Girsanov transformation (or Radon-Nikodym derivative) of measures induced on the canonical sample space (Ω, \mathcal{A}) by systems (2.1) and (2.2) to derive a stochastic integral equation for the unnormalized conditional density leading to Zakai equation. Before presenting our derivation, which is basically same as that given in [76], we first present the following known results due to Girsanov and Ito. The proofs of these results are known and can be found, for example, in [37].

Lemma 2.1 (Girsanov Transformation [37] Theorem 11, P.250)

Let μ_1 and μ_2 (which are defined on the space (Ω, \mathcal{A})) denote the the measures induced, respectively, by the systems

$$\begin{aligned} dx_t &= a(x_t)dt + b(x_t)dW_t, \\ dy_t &= h(x_t)dt + \sigma_0(y_t)dV_t, \end{aligned} \tag{2.12}$$

and

$$\begin{aligned} dx_t &= a(x_t)dt + b(x_t)d\bar{W}_t, \\ dy_t &= \sigma_0(y_t)dV_t. \end{aligned} \tag{2.13}$$

For each $t \in [0, T]$, define

$$\begin{aligned} \frac{d\mu_1}{d\mu_2} &\equiv L_t \\ &\equiv \exp \left\{ -\frac{1}{2} \int_0^t |\sigma_0^{-1}(y_s)h(x_s)|^2 ds + \int_0^t (\sigma_0^{-1}(y_s)h(x_s) \cdot \sigma_0^{-1}(y_s)dy_s) \right\}, \end{aligned} \quad (2.14)$$

where h and σ_0 satisfy assumptions (A4) and (A5). Then if $E_2 L_T = 1$, the process $\{(W_t, \xi_t); \mathcal{F}_t, t \geq 0\}$, where E_2 denotes the expectation with respect to the measure μ_2 and

$$\xi_t \equiv \int_0^t \sigma_0^{-1}(y_s) dy_s, \quad t \in [0, T],$$

is a standard Wiener process on the probability space $(\Omega, \mathcal{A}, \mu_2)$.

Remark 2.1

Note that under the measure μ_2 , the σ -fields \mathcal{F}_t^x and $\mathcal{F}_t^y, t \geq 0$, are independent. This fact is very important and it will be used in the derivation of Zakai equation.

Lemma 2.2 (Ito differential, [37] corollary 1, P.70)

Let the R^n -valued process $\{\eta(t), t \geq 0\}$ be the solution of the following stochastic differential equation

$$d\eta(t) = a(\eta(t))dt + b(\eta(t))dW_t,$$

with initial state $\eta(0) = \eta_0$, where W is n -dimensional standard Wiener process independent of the initial state η_0 and the functions a and b satisfy our basic assumptions. Let f be any twice continuously differentiable function on R^n . Then the Ito differential of f is given by

$$df(\eta(t)) = (Af)(\eta(t))dt + (b'f_x)(\eta(t)) \cdot dW_t, \quad (2.15)$$

where f_x denotes the partial derivative of f with respect to x and A denotes the (backward) Kolmogorov operator corresponding to the process $\{\eta(t), t \geq 0\}$, given by

$$Af = \sum_{i=1}^n a_i(x) f_{x_i} + \frac{1}{2} \sum_{i,j=1}^n \sigma_{ij}(x) f_{x_i x_j}, \quad (2.16)$$

and $\sigma \equiv bb'$. ■

Let f be any bounded measurable function on R^n . Then using the fact that the measures μ_1 and μ_2 (see Lemma 2.1) are absolutely continuous and that

$$\int_{\Omega} \gamma d\mu_1 = \int_{\Omega} \gamma L_T d\mu_2,$$

where L_T is given by (2.14) and γ is any G -measurable random variable, $G \subset \mathcal{A}$, it follows from (2.7) and Baye's formula that the conditional expectation \hat{f}_t can be written as

$$\begin{aligned} \hat{f}_t &\equiv E_1\{f(x_t) | \mathcal{F}_t^y\} = \frac{E_2\{L_t f(x_t) | \mathcal{F}_t^y\}}{E_2\{L_t | \mathcal{F}_t^y\}} \\ &= \frac{E_2\{E_2\{L_t | \mathcal{F}_t^y \vee \sigma(x_t)\} f(x_t) | \mathcal{F}_t^y\}}{E_2\{E_2\{L_t | \mathcal{F}_t^y \vee \sigma(x_t)\} | \mathcal{F}_t^y\}}, \end{aligned} \quad (2.17)$$

where E_1 and E_2 denote the conditional expectations with respect to the measures μ_1 and μ_2 , respectively.

Define

$$\tilde{L}(t, x_t) \equiv E_2\{L_t | \mathcal{F}_t^y \vee \sigma(x_t)\}, \quad (2.18)$$

and

$$Q(t, dx) \equiv \Pr\{x_t \in dx\}. \quad (2.19)$$

Taking f as the indicator function on $\Gamma, \Gamma \in \mathcal{B}(R^n)$, and using the fact that the σ -fields \mathcal{F}_t^x and \mathcal{F}_t^y are independent, under μ_2 , it follows from the above definitions, Fubini's Theorem and equation (2.17) that the measure

$$\Pr\{x_t \in \Gamma | \mathcal{F}_t^y\} \equiv P(t, \Gamma | \mathcal{F}_t^y),$$

is given by

$$P(t, \Gamma | \mathcal{F}_t^y) = \frac{\int_{\Gamma} \tilde{L}(t, x) Q(t, dx)}{\int_{R^n} \tilde{L}(t, x) Q(t, dx)}, \quad (2.20)$$

for all $t \in [0, T]$. We now present the following result which shows that the process $\tilde{L}(t, \cdot), t \geq 0$, is governed by two stochastic integral equations.

Theorem 2.1

The process $\tilde{L}(t, \cdot), t \geq 0$, satisfies, μ_2 -a.s., the following stochastic integral equations

$$\tilde{L}(t, x) = 1 + \int_0^t \int_{R^n} \tilde{L}(s, \eta) h(\eta) P_{t,s}(x; d\eta) \cdot \Gamma_0^{-1}(y_s) dy_s, \quad (2.21)$$

and

$$\begin{aligned} \tilde{L}(t, x) = & \int_{R^n} \tilde{L}(s, \eta) P_{t,s}(x; d\eta) \\ & + \int_s^t \int_{R^n} \tilde{L}(\theta, \eta) h(\eta) P_{t,\theta}(x; d\eta) \cdot \Gamma_0^{-1}(y_\theta) dy_\theta, \end{aligned} \quad (2.22)$$

for all $(t, x) \in [0, T] \times R^n$. Here

$$P_{t,s}(x; d\eta) \equiv \Pr\{x_s \in d\eta | x_t = x\}, \quad s \leq t, \quad (2.23)$$

denotes the backward transition probability of the \mathcal{F}_t -Markov process $x_t, t \geq 0$, and $\Gamma_0 \equiv \sigma_0 \sigma_0'$.

Proof

Define

$$\alpha(t) \equiv -\frac{1}{2} \int_0^t |\sigma_0^{-1}(y_s) h(x_s)|^2 ds + \int_0^t (\sigma_0^{-1}(y_s) h(x_s) \cdot \sigma_0^{-1}(y_s) dy_s). \quad (2.24)$$

Then it is clear from equation (2.14) that

$$L_t = \exp(\alpha(t)).$$

Applying Lemma 2.2, the Ito differential of $L_t, t \geq 0$, is given by

$$dL_t = \exp(\alpha(t)) \left[-\frac{1}{2} |\sigma_0^{-1}(y_t)h(x_t)|^2 dt + \sigma_0^{-1}(y_t)h(x_t) \cdot \sigma_0^{-1}(y_t)dy_t \right] + \frac{1}{2} \exp \alpha(t) |\sigma_0^{-1}(y_t)h(x_t)|^2 dt,$$

which implies that

$$dL_t = L_t h(x_t) \cdot \Gamma_0^{-1}(y_t) dy_t,$$

$$L_0 = 1,$$

and hence

$$L_t = 1 + \int_0^t L_s h(x_s) \cdot \Gamma_0^{-1}(y_s) dy_s. \quad (2.25)$$

Thus

$$\begin{aligned} E_2\{L_t | \mathcal{F}_t^y \vee \sigma(x_t)\} &\equiv \tilde{L}(t, x_t) \\ &= 1 + E_2\left\{ \int_0^t L_s h(x_s) \cdot \Gamma_0^{-1}(y_s) dy_s | \mathcal{F}_t^y \vee \sigma(x_t) \right\} \\ &= 1 + \int_0^t E_2\{L_s h(x_s) | \mathcal{F}_t^y \vee \sigma(x_t)\} \cdot \Gamma_0^{-1}(y_s) dy_s. \end{aligned} \quad (2.26)$$

Using the independence of the σ -fields \mathcal{F}_s^x and \mathcal{F}_s^y under μ_2 and the fact that $\{x_t; \mathcal{F}_t, t \geq 0\}$ is a Markov process, it follows that $\mathcal{F}_s^x \vee \mathcal{F}_s^y$ and $\sigma\{y_\theta - y_s; s \leq \theta \leq t\}$ are independent given $\mathcal{F}_s^y \vee \sigma(x_t)$. Hence

$$\begin{aligned} \tilde{L}(t, x_t) &= 1 + \int_0^t E_2\{L_s h(x_s) | \mathcal{F}_s^y \vee \sigma(x_t)\} \cdot \Gamma_0^{-1}(y_s) dy_s \\ &= 1 + \int_0^t E_2\{E_2\{L_s | \mathcal{F}_s^y \vee \sigma(x_s) \vee \sigma(x_t)\} h(x_s) | \mathcal{F}_s^y \vee \sigma(x_t)\} \cdot \Gamma_0^{-1}(y_s) dy_s. \end{aligned} \quad (2.27)$$

By the Markov property of the process $\{x_t; \mathcal{F}_t, t \geq 0\}$ and the independence of the σ fields \mathcal{F}_t^x and \mathcal{F}_t^y under μ_2 , it follows that $\mathcal{F}_s^y \vee \mathcal{F}_s^x$ and $\sigma(x_t)$ are independent given $\mathcal{F}_s^y \vee \sigma(x_s)$. Hence

$$E_2\{L_s | \mathcal{F}_s^y \vee \sigma(x_s) \vee \sigma(x_t)\} = E_2\{L_s | \mathcal{F}_s^y \vee \sigma(x_s)\} \equiv \tilde{L}(s, x_s). \quad (2.28)$$

Utilizing equation (2.28) in equation (2.27), we obtain

$$\tilde{L}(t, x_t) = 1 + \int_0^t E_2\{\tilde{L}(s, x_s)h(x_s)|\mathcal{F}_s^y \vee \sigma(x_t)\} \cdot \Gamma_0^{-1}(y_s) dy_s. \quad (2.29)$$

Using the independence of the σ -fields \mathcal{F}_t^x and \mathcal{F}_t^y under μ_2 , it follows from Fubini's Theorem and the definition of $P_{t,s}$; $s \leq t$, (see equation (2.23)) that equation (2.29) can be written as

$$\tilde{L}(t, x) = 1 + \int_0^t \int_{R^n} \tilde{L}(s, \eta)h(\eta)P_{t,s}(x; d\eta) \cdot \Gamma_0^{-1}(y_s) dy_s,$$

which is equation (2.21). Using equation (2.25), we obtain

$$L_t = L_s + \int_s^t L_\theta h(x_\theta) \cdot \Gamma_0^{-1}(y_\theta) dy_\theta,$$

and hence

$$\begin{aligned} \tilde{L}(t, x_t) &= E_2\{L_s|\mathcal{F}_t^y \vee \sigma(x_t)\} \\ &\quad + \int_s^t E_2\{L_\theta h(x_\theta)|\mathcal{F}_t^y \vee \sigma(x_t)\} \cdot \Gamma_0^{-1}(y_\theta) dy_\theta. \end{aligned}$$

Following similar arguments as given above, one can easily verify that equation (2.22) follows from the above equation. This completes the proof. ■

Based on the above Theorem we now derive a stochastic integral equation for the unnormalized density, from which Zakai equation can be obtained. For this, we need the semigroup generated by the Markov process $\{x_t; \mathcal{F}_t, t \geq 0\}$. Let Ψ be any bounded measurable function on R^n and let $\{T_{s,t}, s \leq t\}$ denote the (backward) semigroup given by

$$(T_{s,t}\Psi)(\eta) \equiv E\{\Psi(x_t)|x_s = \eta\}, \quad s \leq t.$$

Let

$$P_{s,t}(\eta; \Gamma) \equiv \Pr\{x_t \in \Gamma|x_s = \eta\}, \quad \eta \in R^n, \Gamma \in \mathcal{B}(R^n),$$

be the kernel corresponding to the semigroup $T_{s,t}$ so that

$$(T_{s,t}\Psi)(\eta) = \int_{R^n} \Psi(\gamma) P_{s,t}(\eta; d\gamma), \quad (2.30)$$

satisfying

$$S.\lim_{t \rightarrow s} (T_{s,t}\Psi)(\eta) = \Psi(\eta),$$

for any bounded measurable function Ψ . From Lemma 2.2, it is clear that the infinitesimal generator corresponding to the (backward) semigroup $T_{s,t}, s \leq t$, is given by

$$(A\Psi)(x) = \sum_{i=1}^n a_i(x) \Psi_{x_i}(x) + \frac{1}{2} \sum_{i,j=1}^n \sigma_{ij}(x) \Psi_{x_i x_j}(x). \quad (2.31)$$

Define

$$(T_{s,t}^* \beta)(t, \gamma) \equiv \int_{R^n} P_{s,t}(d\eta; \gamma) \beta(s, \eta), \quad (2.32)$$

where β denotes the density corresponding to the measure induced by the process $x_t, t \geq 0$, on $\mathcal{B}(R^n)$. Suppose that the measure $\Gamma \rightarrow P_{s,t}(\eta; \Gamma), \eta \in R^n$, is absolutely continuous with respect to Lebesgue measure on R^n admitting a density so that

$$P_{s,t}(\eta; \Gamma) = \int_{\Gamma} p_{s,t}(\eta; \gamma) d\gamma. \quad (2.33)$$

It is known (see for example [33] P.276) that the density $p_{s,t}(\eta; \gamma), \eta, \gamma \in R^n$, satisfies the following Kolmogorov (or Fokker-Plank) forward equation

$$\frac{\partial}{\partial t} p_{s,t}(\eta; \gamma) = A^* p_{s,t}(\eta; \gamma), \quad (2.34)$$

for all $s \leq t$, and $\eta, \gamma \in R^n$, where A^* denotes the formal adjoint of the operator A , given by equation (2.31). Under the assumptions (A1) and (A3), it follows from equations (2.32), (2.33), and (2.34) that the semigroup $T_{s,t}^*, s \leq t$, satisfies the following properties

(Pi) For all $s \leq t \in [0, T]$,

$$(T_{s,t}^* f)(t, \gamma) = (T_{s,\theta}^* f)(\theta, \gamma) + \int_{\theta}^t A^*(T_{\theta,\tau}^* f)(\tau, \gamma) d\tau, \quad (2.35)$$

and

(Pii) There exists a constant $\delta > 0$ such that

$$\int_{R^n} |(T_{s,t}^* g)(x)|^2 dx \leq \delta \int_{R^n} |g(x)|^2 dx,$$

for all $g \in H$ and $0 \leq s \leq t \leq T$, where H denotes the family of functions $\{g\} \equiv \{g(x), x \in R^n\}$, satisfying the following integrability condition

$$\int_{R^n} |g(x)|^2 dx < \infty.$$

With this preparation we now present the following corollary which shows that the unnormalized density satisfies two stochastic integral equations.

Corollary 2.1

Suppose for each $\eta \in R^n$, the measure $\Gamma \rightarrow P_{s,t}(\eta; \Gamma), s \leq t$, is absolutely continuous with respect to Lebesgue measure admitting the density $\gamma \rightarrow p_{s,t}(\eta; \gamma)$ and the semigroup $T_{s,t}^*, s \leq t$, satisfies the properties (Pi) and (Pii). Then the measure $\Gamma \rightarrow P(t, \Gamma | \mathcal{F}_t^y), t \geq 0$, (see equation (2.20)) is also absolutely continuous with respect to Lebesgue measure admitting a density $p(t, \gamma | \mathcal{F}_t^y)$ and that

$$p(t, x | \mathcal{F}_t^y) = \frac{\varphi(t, x)}{\int_{R^n} \varphi(t, x) dx}, \quad (2.36)$$

where $\varphi(t, x), (t, x) \in [0, T] \times R^n$, satisfies (μ_2 -a.s) the following stochastic integral equations

$$\varphi(t, x) = q(t, x) + \int_0^t (T_{s,t}^*(\varphi h))(t, x) \cdot \Gamma_0^{-1}(y_s) dy_s, \quad (2.37)$$

and

$$\varphi(t, x) = (T_{s,t}^* \varphi)(t, x) + \int_s^t (T_{\theta,t}^* (\varphi h))(t, x) \cdot \Gamma_0^{-1}(y_\theta) dy_\theta, \quad (2.38)$$

for all $(t, x) \in [0, T] \times R^n$, where q denotes the density corresponding to the measure Q (see equation (2.19)) and $T_{s,t}^*$ is given by equation (2.32).

Proof

Let Q_0 denote the initial distribution of x_0 . Then one can easily verify that the measure $Q(t, \Gamma)$ (see equation (2.19)) is given by

$$Q(t, \Gamma) = \int_{\Gamma} P_{0,t}(z; \Gamma) Q_0(dz).$$

Since by assumption, the measure $P_{0,t}(z; \Gamma)$ is absolutely continuous with respect to Lebesgue measure, it is clear that the measure $Q(t, \cdot)$ admits a density $q(t, \cdot)$ so that

$$Q(t, \Gamma) = \int_{\Gamma} q(t, x) dx, \quad \Gamma \in \mathcal{B}(R^n).$$

Therefore, it follows from equation (2.20) the the measure $\Gamma \rightarrow P(t, \Gamma | \mathcal{F}_t^y)$ is also absolutely continuous with respect to Lebesgue measure and hence it admits a density which we denote by $p(t, x | \mathcal{F}_t^y)$. Defining

$$\varphi(t, x) \equiv \tilde{L}(t, x) q(t, x),$$

and using equation (2.20), one obtains equation (2.36). Multiplying equations (2.21) and (2.22) by $q(t, x)$, we have

$$\begin{aligned} \tilde{L}(t, x) q(t, x) &\equiv \varphi(t, x) \\ &= q(t, x) + \int_0^t \int_{R^n} \tilde{L}(s, \eta) h(\eta) P_{t,s}(x; d\eta) q(t, x) \cdot \Gamma_0^{-1}(y_s) dy_s, \end{aligned}$$

and

$$\begin{aligned} \varphi(t, x) &= \int_{R^n} \tilde{L}(s, \eta) P_{t,s}(x; d\eta) q(t, x) \\ &\quad + \int_s^t \int_{R^n} \tilde{L}(\theta, \eta) h(\eta) P_{t,\theta}(x; d\eta) q(t, x) \cdot \Gamma_0^{-1}(y_\theta) dy_\theta. \end{aligned}$$

Using the duality

$$P_{t,s}(x; d\eta)q(t, x) = P_{s,t}(d\eta; x)q(s, \eta), \quad s \leq t,$$

and the definition $\varphi(t, x) \equiv \tilde{L}(t, x)q(t, x)$, it follows from the above expressions that

$$\begin{aligned} \varphi(t, x) &= q(t, x) + \int_0^t \int_{R^n} \tilde{L}(s, \eta)q(s, \eta)h(\eta)P_{s,t}(d\eta; x) \cdot \Gamma_0^{-1}(y_s)dy_s \\ &= q(t, x) + \int_0^t \int_{R^n} \varphi(s, \eta)h(\eta)P_{s,t}(d\eta; x) \cdot \Gamma_0^{-1}(y_s)dy_s, \end{aligned}$$

and similarly

$$\begin{aligned} \varphi(t, x) &= \int_{R^n} \varphi(s, \eta)P_{s,t}(d\eta; x) \\ &\quad + \int_s^t \int_{R^n} \varphi(\theta, \eta)h(\eta)P_{\theta,t}(d\eta; x) \cdot \Gamma_0^{-1}(y_\theta)dy_\theta. \end{aligned}$$

Using the definition of the semigroup $T_{s,t}^*$, $s \leq t$, (see equation (2.32)), it is easy to verify that equations (2.37) and (2.38) follow from the above expressions. This completes the proof. ■

Based on the above corollary, we now derive Zakai equation. For this we need the following property of the semigroup $T_{s,t}^*$.

(Piii) Let f be any bounded, continuous function on R^n whose first and second derivatives are bounded, continuous and satisfy Holder condition on R^n . Then $(T_{s,t}^*f)(t, x)$ and its first and second derivatives, with respect to x , are continuous on $(0, T) \times R^n$ and bounded on $[0, T] \times G$, where G is any bounded subset of R^n .

Theorem 2.2 (Zakai Equation)

Suppose that assumptions (A1)-(A5) hold and the semigroup $T_{s,t}^*$, $s \leq t$, satisfies the properties (Pi)-(Piii). Let φ be the solution of (2.38), and a.s. the functions

φ and φh together with their first and second derivatives, with respect to x , are continuous, bounded and satisfy Holder condition on R^n , uniformly in $t \in [0, T]$.

Then φ satisfies also the following linear stochastic partial differential equation

$$\begin{aligned} d\varphi(t, x) &= (A^*\varphi)(t, x)dt + \varphi(t, x)h(x) \cdot \Gamma_0^{-1}(y_t)dy_t, & \mu_2 - a.s, \\ \varphi(0, x) &= p_0(x), \end{aligned} \quad (2.41)$$

for all $(t, x) \in [0, T] \times R^n$, where the operator A^* is the formal adjoint of A given by equation (2.31).

Proof

From equations (2.35) and (2.38), we have

$$(T_{s,t}^*\varphi)(t, x) = \varphi(s, x) + \int_s^t A^*(T_{s,\theta}^*\varphi)(\theta, x)d\theta, \quad (2.42)$$

and

$$\varphi(\theta, x) = (T_{s,\theta}^*\varphi)(\theta, x) + \int_s^\theta (T_{\tau,\theta}^*(\varphi h))(\theta, x) \cdot \Gamma_0^{-1}(y_\tau)dy_\tau. \quad (2.43)$$

From equation (2.42) it follows that

$$(T_{\alpha,t}^*(\varphi h))(t, x) = (\varphi h)(\alpha, x) + \int_\alpha^t A^*(T_{\alpha,\tau}^*(\varphi h))(\tau, x)d\tau, \quad (2.44)$$

and hence taking the Ito integral with respect to the \mathcal{F}_t^y -martingale $\int_0^t \Gamma_0^{-1}dy_s$, we obtain

$$\begin{aligned} \int_s^t (T_{\alpha,t}^*(\varphi h))(t, x) \cdot \Gamma_0^{-1}(y_\alpha)dy_\alpha &= \int_s^t (\varphi h)(\alpha, x) \cdot \Gamma_0^{-1}(y_\alpha)dy_\alpha \\ &+ \int_s^t \int_\alpha^t A^*(T_{\alpha,\tau}^*(\varphi h))(\tau, \alpha)d\tau \cdot \Gamma_0^{-1}(y_\alpha)dy_\alpha. \end{aligned} \quad (2.45)$$

Substituting equation (2.38) into equation (2.45), we obtain

$$\begin{aligned} \varphi(t, x) &= (T_{s,t}^* \varphi)(t, x) + \int_s^t (\varphi h)(\alpha, x) \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha \\ &\quad + \int_s^t \int_\alpha^t \mathbf{A}^*(T_{\alpha,\tau}^*(\varphi h))(\tau, \alpha) d\tau \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha. \end{aligned} \quad (2.46)$$

Substituting $(T_{s,t}^* \varphi)(t, x)$ from (2.42) into (2.46), we have

$$\begin{aligned} \varphi(t, x) &= \varphi(s, x) + \int_s^t \mathbf{A}^*(T_{s,\theta}^* \varphi)(\theta, x) d\theta + \int_s^t (\varphi h)(\alpha, x) \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha \\ &\quad + \int_s^t \int_\alpha^t \mathbf{A}^*(T_{\alpha,\tau}^*(\varphi h))(\tau, \alpha) d\tau \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha. \end{aligned} \quad (2.47)$$

From equations (2.43) and (2.47), it follows that

$$\begin{aligned} \varphi(t, x) &= \varphi(s, x) + \int_s^t (\mathbf{A}^* \varphi)(\theta, x) d\theta \\ &\quad - \int_s^t \mathbf{A}^* \int_s^\theta (T_{\tau,\theta}^*(\varphi h))(\theta, x) \cdot \Gamma_0^{-1}(y_\tau) dy_\tau d\theta \\ &\quad + \int_s^t (\varphi h)(\alpha, x) \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha \\ &\quad + \int_s^t \int_\alpha^t \mathbf{A}^*(T_{\alpha,\tau}^*(\varphi h))(\tau, x) d\tau \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha. \end{aligned} \quad (2.48)$$

Comparing equations (2.41) and (2.48), it is clear that in order to prove the Theorem, it is enough to show

$$\begin{aligned} &\int_s^t \int_\alpha^t \mathbf{A}^*(T_{\alpha,\tau}^*(\varphi h))(\tau, x) d\tau \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha \\ &= \int_s^t \mathbf{A}^* \int_s^\theta (T_{\tau,\theta}^*(\varphi h))(\theta, x) \cdot \Gamma_0^{-1}(y_\tau) dy_\tau d\theta. \end{aligned} \quad (2.49)$$

Since by assumption φh is continuous and bounded and the semigroup $T_{s,t}^*$ satisfies the properties (Pi)-(Piii), it follows from the properties of stochastic integral (based on continuous martingales) and the mean value Theorem that equation (2.49) can be written as [76]

$$\int_s^t \int_\alpha^t \mathbf{A}^*(T_{\alpha,\tau}^*(\varphi h))(\tau, x) d\tau \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha$$

$$\begin{aligned}
 &= \int_s^t \int_s^\theta \mathbf{A}^*(T_{\tau,\theta}^*(\varphi h))(\theta, x) \cdot \Gamma_0^{-1}(y_\tau) dy_\tau d\theta \\
 &= \int_s^t \int_s^\alpha \mathbf{A}^*(T_{\tau,\alpha}^*(\varphi h))(\alpha, x) \cdot \Gamma_0^{-1}(y_\tau) dy_\tau d\alpha. \quad (2.50)
 \end{aligned}$$

Hence it remains only to justify the formal interchange of the order of integration in (2.50).

For a fixed $x \in R^n$, define

$$\Psi(\alpha, \tau) \equiv \mathbf{A}^*(T_{\alpha,\tau}^*(\varphi h))(\tau, x), \quad \tau \geq \alpha.$$

Then we must show that

$$\int_s^t \int_\alpha^t \Psi(\alpha, \tau) d\tau \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha = \int_s^t \int_s^\alpha \Psi(\tau, \alpha) \cdot \Gamma_0^{-1}(y_\tau) dy_\tau d\alpha. \quad (2.50.a)$$

Define

$$f(\alpha) \equiv \int_\alpha^t \Psi_M(\alpha, \tau) d\tau,$$

and

$$\tilde{f}(\alpha) \equiv \int_s^\alpha \Psi_M(\tau, \alpha) \cdot \Gamma_0^{-1}(y_\tau) dy_\tau,$$

where

$$\Psi_M \equiv \begin{cases} \Psi, & \text{if } |\Psi| \leq M, \\ 0, & \text{otherwise.} \end{cases}$$

Let

$$\beta(t) \equiv E \left(\int_s^t f(\alpha) \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha - \int_s^t \tilde{f}(\alpha) d\alpha \right)^2, \quad (2.51)$$

and

$$\tilde{\beta}_M(t) \equiv \tilde{\Psi}(t) - \tilde{\Psi}_M(t), \quad (2.52)$$

for all $t \in [0, T]$, where

$$\begin{aligned}
 \tilde{\Psi}(t) \equiv & \left\{ \int_s^t \int_\alpha^t \Psi(\alpha, \tau) d\tau \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha \right. \\
 & \left. - \int_s^t \int_s^\tau \Psi(\alpha, \tau) \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha d\tau \right\}, \quad (2.52.1)
 \end{aligned}$$

and

$$\begin{aligned} \tilde{\Psi}_M(t) \equiv & \left\{ \int_s^t \int_\alpha^t \Psi_M(\alpha, \tau) d\tau \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha \right. \\ & \left. - \int_s^t \int_s^\tau \Psi_M(\alpha, \tau) \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha d\tau \right\}. \end{aligned} \quad (2.52.2)$$

To prove the equality (2.50.a), we proceed as follows. First we show that the function β , as given, by (2.51), is identically zero. Then with the help of the property (Piii) of the semigroup $T_{s,t}^*$, we show that $\tilde{\beta}_M \rightarrow 0$ in probability as $M \rightarrow \infty$. Define

$$I_1(t) \equiv E \left(\int_s^t f(\alpha) \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha \right)^2, \quad (2.53)$$

$$I_2(t) \equiv -2E \left\{ \left(\int_s^t f(\alpha) \cdot \Gamma_0^{-1}(y_\alpha) dy_\alpha \right) \left(\int_s^t \tilde{f}(\alpha) d\alpha \right) \right\}, \quad (2.54)$$

and

$$I_3(t) \equiv E \left(\int_s^t \tilde{f}(\alpha) d\alpha \right)^2. \quad (2.55)$$

Clearly, from equations (2.51), (2.53), (2.54) and (2.55), we have

$$\beta(t) = I_1(t) + I_2(t) + I_3(t). \quad (2.56)$$

Using the definition $\Gamma_0 \equiv \sigma_0 \sigma_0'$, the fact that $\int_0^t \sigma_0^{-1} dy_s$ is standard Wiener process under μ_2 , and the properties of stochastic integral, it follows from equations (2.53)-(2.55) that the quantities I_1, I_2 and I_3 are given, respectively, by

$$I_1(t) = E \left\{ \int_s^t \int_\alpha^t \Psi_M(\alpha, \tau) \cdot \int_\alpha^\tau \Gamma_0^{-1}(y_\xi) \Psi_M(\xi, \tau) d\xi d\tau d\alpha \right\}, \quad (2.57)$$

$$\begin{aligned} I_2(t) &= -2E \left\{ \int_s^t \tilde{f}(\alpha) \int_s^\alpha f(\xi) \cdot \Gamma_0^{-1}(y_\xi) dy_\xi d\alpha \right\} \\ &= -2E \left\{ \int_s^t \int_s^\alpha \Psi_M(\tau, \alpha) \cdot \Gamma_0^{-1}(y_\tau) f(\tau) d\tau d\alpha \right\} \end{aligned} \quad (2.58)$$

$$= -2E \left\{ \int_s^t \int_s^\alpha \int_\tau^t \Psi_M(\tau, \alpha) \cdot \Gamma_0^{-1}(y_\tau) \Psi_M(\tau, \xi) d\xi d\tau d\alpha \right\},$$

and

$$\begin{aligned} I_3(t) &= E \left\{ \int_s^t \tilde{f}(\alpha) \int_s^\alpha \tilde{f}(\xi) d\xi d\alpha \right\} \\ &= E \left\{ \int_s^t \int_s^\alpha \Psi_M(\tau, \alpha) \cdot \Gamma_0^{-1}(y_\tau) dy_\tau \int_s^\tau \int_s^\xi \Psi_M(\tau, \xi) \cdot \Gamma_0^{-1}(y_\tau) dy_\tau d\xi d\alpha \right\} \\ &= E \left\{ \int_s^t \int_s^\alpha \int_s^\xi \Psi_M(\tau, \alpha) \cdot \Gamma_0^{-1}(y_\tau) \Psi_M(\tau, \xi) d\xi d\tau d\alpha \right\}. \end{aligned} \quad (2.59)$$

Defining

$$\chi_u(v) \equiv \begin{cases} 0, & \text{if } u > v, \\ \frac{1}{2}, & \text{if } u = v, \\ 1, & \text{if } u < v, \end{cases} \quad (2.60)$$

it follows from (2.57), (2.58) and (2.59) that

$$I_1(t) = E\left\{\int_s^t \int_s^t \int_s^t \chi_r(\alpha)\chi_r(\xi)\Psi_M(\alpha, \tau) \cdot \Gamma_0^{-1}(y_\xi)\Psi_M(\xi, \tau)d\xi d\tau d\alpha\right\}, \quad (2.61)$$

$$I_2(t) = -2E\left\{\int_s^t \int_s^t \int_s^t \chi_r(\alpha)\chi_r(\xi)\Psi_M(\alpha, \tau) \cdot \Gamma_0^{-1}(y_\xi)\Psi_M(\xi, \tau)d\tau d\xi d\alpha\right\}, \quad (2.62)$$

and

$$I_3(t) = E\left\{\int_s^t \int_s^t \int_s^t \chi_r(\alpha)\chi_r(\xi)\Psi_M(\alpha, \tau) \cdot \Gamma_0^{-1}(y_\xi)\Psi_M(\xi, \tau)d\tau d\xi d\alpha\right\}. \quad (2.63)$$

Using (2.56), (2.61), (2.62) and (2.63), it is clear that $\beta \equiv 0$. Let $\tilde{\beta}_M$ be given by (2.52). Then for any $\delta > 0$, we have

$$\Pr\{|\tilde{\beta}_M| > \delta\} \leq \Pr\left\{\sup_{s \leq \alpha \leq \tau \leq t} |\Psi(\alpha, \tau)| \geq M\right\}.$$

Since by assumption, the function $\{\Psi(\alpha, \tau); s \leq \alpha \leq \tau \leq t\}$ is bounded, it follows from the above inequality that $\tilde{\beta}_M \rightarrow 0$ in probability as $M \rightarrow \infty$, and the proof of the Theorem follows. ■

Remark 2.2

Note that the strong assumptions in the above Theorem, which have been imposed by Zakai in his original work, were mainly needed for the justification of the interchange of the operator \mathbf{A}^* with the stochastic integral in the right hand side of equation (2.49).

In the next section shall use the result of Theorem 2.2 to obtain Kushner equation (2.9). Further, using Ito's Lemma (Lemma 2.2) we show that Zakai equation can also be obtained from Kushner's equation.

2.4 Kushner Equation and Its Relation to Zakai Equation

Let the (state) process $x_t, t \geq 0$, and the output (observed) process $y_t, t \geq 0$, be governed by equations (2.1) and (2.2), respectively. As indicated earlier in this chapter, the filtering problem of the systems (2.1) and (2.2), is completely resolved provided that one is able to compute the normalized (or unnormalized) conditional density of the process x_t given the σ -field $\mathcal{F}_t^y, t \geq 0$. In fact the conditional density can be computed by two methods, which are basically the same. The first method is to solve Zakai equation (2.41), which is linear stochastic partial differential equation describing the flow of the unnormalized density, from which the (normalized) density can be obtained via equation (2.36). The second method is to solve directly Kushner equation (2.9), which is a nonlinear stochastic partial differential equation.

Although these two equations are well known and have been used extensively in both theoretical and practical applications. We are interested in this section not only in deriving Kushner equation from that of Zakai, but also to show their equivalence using Ito's Lemma and the properties of the likelihood ratio.

Let f be any bounded measurable function on R^n . Then, as indicated earlier in section 2.3, the optimal estimate (in the mean square sense) of the function f given $\mathcal{F}_t^y, t \geq 0$, is given by

$$\hat{f}_t \equiv E_1\{f(x_t)|\mathcal{F}_t^y\} = \frac{E_2\{L_t f(x_t)|\mathcal{F}_t^y\}}{E_2\{L_t|\mathcal{F}_t^y\}}, \quad (2.64)$$

where L_t is given by (2.14) and E_1 and E_2 denote the conditional expectations with respect to the measures μ_1 and μ_2 , respectively. Using equation (2.36), it is easy to verify that the estimate \hat{f}_t is also given by

$$\hat{f}_t = \frac{\langle \varphi(t, \cdot), f(\cdot) \rangle}{\langle \varphi(t, \cdot), 1 \rangle}, \quad t \in [0, T], \quad (2.65)$$

where

$$\langle \varphi(t, \cdot), f(\cdot) \rangle \equiv \int_{R^n} \varphi(t, x) f(x) dx,$$

and

$$\langle \varphi(t, \cdot), 1 \rangle \equiv \int_{R^n} \varphi(t, x) dx.$$

Comparing equations (2.64) and (2.65), it is clear that

$$E_2\{L_t f(x_t) | \mathcal{F}_t^y\} = \langle \varphi(t, \cdot), f(\cdot) \rangle, \quad (2.66)$$

and the likelihood ratio is given by

$$E_2\{L_t | \mathcal{F}_t^y\} = \langle \varphi(t, \cdot), 1 \rangle \equiv \tilde{\varphi}(t). \quad (2.67)$$

For notational convenience, from now on, we shall use $\varphi(t)$ to denote $\varphi(t, x)$ and f to denote $f(x)$, $x \in R^n$.

Let f be any function in $C_0^2(R^n)$, where $C_0^2(R^n)$ denotes the class of twice continuously differentiable functions on R^n with compact support. Then multiplying equation (2.41) by $f \in C_0^2(R^n)$ and integrating over R^n , we obtain

$$d \langle \varphi(t), f \rangle = \langle \varphi(t), A f \rangle dt + \langle \varphi(t), h f \rangle \cdot \Gamma_0^{-1}(y_t) dy_t, \quad (2.68)$$

$$\langle \varphi(0), f \rangle = \langle p_0, f \rangle,$$

for all $t \in [0, T]$. Taking $f \equiv 1$ (this can be justified by taking a sequence of functions $\{f_n\} \in C_0^\infty(R^n)$ so that $\{f_n\} \rightarrow 1$), it follows from (2.68) that

$$d \langle \varphi(t), 1 \rangle = \langle \varphi(t), h \rangle \cdot \Gamma_0^{-1} dy_t, \quad (2.69)$$

$$\langle \varphi(0), 1 \rangle = \langle p_0, 1 \rangle = 1.$$

Using the fact that

$$\hat{h}_t \equiv \frac{\langle \varphi(t), h \rangle}{\langle \varphi(t), 1 \rangle} \equiv \frac{\langle \varphi(t), h \rangle}{\tilde{\varphi}(t)}, \quad (2.70)$$

it follows from (2.69) that

$$\begin{aligned} d\tilde{\varphi}(t) &= \tilde{\varphi}(t)\hat{h}_t \cdot \Gamma_0^{-1}(y_t)dy_t, \\ \tilde{\varphi}(0) &= 1. \end{aligned} \tag{2.71}$$

We now present the following result which shows that the (normalized) conditional density (see (2.36)) is governed by the Kushner equation (2.9).

Lemma 2.3 (Kushner From Zakai)

Let the conditional density $p(t, \cdot | \mathcal{F}_t^y) \equiv p(t)$ be related to the (unnormalized) density $\varphi(t)$ through the following relation

$$p(t) = \varphi(t)\tilde{\varphi}^{-1}(t), \quad t \geq 0, \tag{2.72}$$

where $\varphi(t)$ and $\tilde{\varphi}(t), t \geq 0$, are the solutions of equations (2.41) and (2.71), respectively. Then the density $p(t), t \geq 0$, satisfies the Kushner equation (2.9).

Proof

Define

$$\varphi(f)(t) \equiv \langle \varphi(t), f \rangle \equiv \int_{R^n} \varphi(t, x) f(x) dx. \tag{2.73}$$

Then multiplying equation (2.41) by $f \in C_0^2(R^n)$, one obtains

$$\begin{aligned} d\varphi(f)(t) &= \varphi(Af)(t)dt + \varphi(hf)(t) \cdot \Gamma_0^{-1}(y_t)dy_t, \\ \varphi(f)(0) &= \langle p_0, f \rangle. \end{aligned} \tag{2.74}$$

Clearly, equation (2.74) is the weak form of Zakai equation (2.41). Applying Ito's differential rule (see for example [37], Theorem 2 P.72) and using the fact that

$\int_0^t \sigma_0^{-1} dy_s$ is a standard Wiener process under the measure μ_2 , it follows that the Ito differential of $\varphi(f)(t)\tilde{\varphi}^{-1}(t)$ is given by

$$\begin{aligned} d(\varphi(f)(t)\tilde{\varphi}^{-1}(t)) &= \tilde{\varphi}^{-1}(t)d\varphi(f)(t) - \tilde{\varphi}^{-2}(t)\varphi(f)(t)d\tilde{\varphi}(t) \\ &\quad + \tilde{\varphi}^{-1}(t)\varphi(f)(t)(\Gamma_0^{-1}(y_t)\hat{h}_t \cdot \hat{h}_t)dt \\ &\quad - \tilde{\varphi}^{-1}(t)(\Gamma_0^{-1}(y_t)\hat{h}_t \cdot \varphi(hf)(t))dt. \end{aligned} \quad (2.75)$$

Since

$$\tilde{\varphi}^{-1}(t)\varphi(f)(t) \equiv \hat{f}_t \equiv \langle p(t), f \rangle, \quad (2.75.1)$$

using equations (2.71) and (2.74), it follows from (2.75) that

$$d\hat{f}_t = (\widehat{A}f)_t dt + ((\widehat{h}f)_t - \hat{h}_t \hat{f}_t) \cdot \Gamma_0^{-1}(y_t)[dy_t - \hat{h}_t dt]. \quad (2.76)$$

Using equation (2.75.1), it follows from (2.76) that

$$\begin{aligned} d \langle p(t), f \rangle &= \langle A^* p(t) dt \\ &\quad + (h - \hat{h}_t)p(t) \cdot \Gamma_0^{-1}(y_t)[dy_t - \hat{h}_t dt], f \rangle. \end{aligned} \quad (2.77)$$

Since the above equality holds for any $f \in C_0^2(\mathbb{R}^n)$, it follows from (2.77) that the conditional density $p(t, \cdot | \mathcal{F}_t^y)$, $t \geq 0$, satisfies Kushner equation (2.9). This completes the proof. ■

Corollary 2.2 (Zakai From Kushner)

Suppose that the normalized conditional density $p(t, \cdot | \mathcal{F}_t^y) \equiv \hat{p}(t)$, $t \geq 0$, is governed by Kushner equations (2.9). Let the unnormalized density $\varphi(t)$, $t \geq 0$, be given by

$$\varphi(t) = p(t)\tilde{\varphi}(t), \quad t \geq 0,$$

where $\tilde{\varphi}(t)$, $t \geq 0$, is the solution of (2.71). Then $\varphi(t)$, $t \geq 0$, satisfies the Zakai equation (2.41).

Proof

The proof is similar to that of Lemma 2.3. Let $f \in C_0^2(R^n)$, then multiplying (2.9) by f and integrating over R^n one obtains (2.76) which is the weak form of (2.9). Again applying Ito's differential rule and using the fact that $\int_0^t \sigma_0^{-1} dy_s$ is a standard Wiener process under the measure μ_2 , it follows that the Ito differential of $\tilde{\varphi}(t)\hat{f}_t$ is given by

$$d(\tilde{\varphi}(t)\hat{f}_t) = \tilde{\varphi}(t)d\hat{f}_t + \hat{f}_t d\tilde{\varphi}(t) + \tilde{\varphi}(t)\Gamma_0^{-1}(y_t)\hat{h}_t \cdot [(\widehat{hf})_t - \hat{h}_t\hat{f}_t]dt. \quad (2.78)$$

Substituting equations (2.71) and (2.76) into (2.78), we obtain

$$\begin{aligned} d(\tilde{\varphi}(t)\hat{f}_t) &= \tilde{\varphi}(t)(\widehat{Af})_t dt + \tilde{\varphi}(t)[(\widehat{hf})_t - \hat{h}_t\hat{f}_t] \cdot \Gamma_0^{-1}(y_t) dy_t \\ &\quad - \tilde{\varphi}(t)[(\widehat{hf})_t - \hat{h}_t\hat{f}_t] \cdot \Gamma_0^{-1}(y_t)\hat{h}_t dt \\ &\quad + \tilde{\varphi}(t)\hat{h}_t\hat{f}_t \cdot \Gamma_0^{-1}(y_t) dy_t + \tilde{\varphi}(t)\Gamma_0^{-1}(y_t)\hat{h}_t \cdot [(\widehat{hf})_t - \hat{h}_t\hat{f}_t]dt, \end{aligned}$$

and hence

$$d(\tilde{\varphi}(t)\hat{f}_t) = \tilde{\varphi}(t)(\widehat{Af})_t dt + \tilde{\varphi}(t)(\widehat{hf})_t \cdot \Gamma_0^{-1}(y_t) dy_t. \quad (2.79)$$

Using the definition (2.75.1), it follows from (2.79) that

$$d\varphi(f)(t) = \varphi(Af)(t)dt + \varphi(hf)(t) \cdot \Gamma_0^{-1}(y_t) dy_t,$$

which is precisely Zakai equation in the weak form (see equation (2.74)). Since the above equation holds for any $f \in C_0^2(R^n)$, the proof follows. ■

In the next section we will use the standard results on existence of solutions of partial differential equations, as given in [6] (Theorem 5.1.1 P.278); to show that Zakai equation (2.41) has a unique (weak) solution. This result will be used later in chapter 6 for the identification problems.

2.5 Remarks On Existence and Uniqueness of Solutions of Zakai Equation

In order to complete our brief overview of the filtering problem considered in this chapter, we discuss in this section the question of existence and uniqueness of solutions of Zakai equation (2.41). In fact this question was considered in the literature by Baras [20], Davis [29], Pardoux [66, 67], and others. We shall follow the work of Pardoux to show that Zakai equation has a unique (weak) pathwise solution. Since the proof of this result is known, we shall only show how one can treat Zakai equation as an ordinary linear (parabolic) partial differential equation whose coefficients are parametrized by the observed process y . Then using the standard result on existence of solutions of partial differential equations [6,22], we show that Zakai equation has a unique solution. For this we need the following definitions.

Let $L_2(R^n)$ denote the equivalence classes of measurable functions on R^n such that $\int_{R^n} |f(x)|^2 dx < \infty$. Consider the Sobolev space

$$H^1 \equiv \{f \in L_2(R^n) : \frac{\partial f}{\partial x_i} \in L_2(R^n); 1 \leq i \leq n\},$$

with $H^{-1} \equiv (H^1)'$ being the dual space. Let $\mathcal{L}(H^1, H^{-1})$ denote the class of bounded linear operators from H^1 to H^{-1} . Hence under assumption (A3), one can write the operator A (or A^*) (see (2.31)) in the following (divergence) form

$$\begin{aligned} \langle Av, u \rangle &= \langle v, A^*u \rangle \\ &= -\frac{1}{2} \int_{R^n} \sum_{i,j=1}^n \sigma_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx + \int_{R^n} \sum_{i=1}^n \tilde{a}_i \frac{\partial u}{\partial x_i} v dx \end{aligned} \quad (2.80)$$

for any $u, v \in H^1$, where

$$\begin{aligned} \tilde{a}_i(x) &\equiv a_i(x) - \frac{1}{2} \sum_{j=1}^n \frac{\partial \sigma_{ij}}{\partial x_j}, \\ \sigma(x) &\equiv (bb')(x), \end{aligned}$$

and $\langle \cdot, \cdot \rangle$ denotes the pairing of H^1 and H^{-1} . Clearly, both A and $A^* \in \mathcal{L}(H^1, H^{-1})$ and under the assumptions (A1) and (A3), it is known (see [36] or [66, 67]) that the operator A^* satisfies the following properties

(i) For each $u, v \in H^1$, there exists a constant $c > 0$ so that

$$| \langle A^* u, v \rangle | \leq c \|u\|_{H^1} \|v\|_{H^1}.$$

(ii) There exist constants $\gamma > 0$ and $\eta \geq 0$ such that

$$- \langle A^* u, u \rangle + \eta \|u\|_{L_2}^2 \geq \gamma \|u\|_{H^1}^2.$$

Define

$$\varphi(t) \equiv \Psi(t) \exp(h \cdot Z_t), \quad t \in [0, T], \quad (2.81)$$

where ' \cdot ' denotes the scalar product in R^m , $\varphi(t) \equiv \varphi(t, x)$, $h \equiv h(x)$, $x \in R^n$, and $Z_t, t \in [0, T]$, is given by

$$Z_t \equiv \int_0^t \Gamma_0^{-1}(y_s) dy_s, \quad t \in [0, T], \quad (2.82)$$

Multiplying equation (2.41) by $\exp(-h \cdot Z_t)$ and using equation (2.82), we have

$$\exp(-h \cdot Z_t) d\varphi(t) = \exp(-h \cdot Z_t) A^* \varphi(t) dt + \exp(-h \cdot Z_t) h \cdot dZ_t. \quad (2.83)$$

Since by Ito's Lemma

$$\begin{aligned} d(\exp(-h \cdot Z_t) \varphi(t)) &= \exp(-h \cdot Z_t) d\varphi(t) - \exp(-h \cdot Z_t) \varphi(t) h \cdot dZ_t \\ &\quad + \frac{1}{2} \exp(-h \cdot Z_t) \varphi(t) h \cdot \Gamma_0^{-1}(y_t) h dt, \end{aligned}$$

it follows from (2.83) that

$$\frac{d}{dt} (\exp(-h \cdot Z_t) \varphi(t)) = \exp(-h \cdot Z_t) \left\{ A^* - \frac{1}{2} (\Gamma_0^{-1}(y_t) h \cdot h) \right\} \varphi(t). \quad (2.83.a)$$

Using (2.81) and defining

$$F^* u \equiv \exp(-h \cdot Z_t) A^* (u \exp(h \cdot Z_t)) - \frac{1}{2} (\Gamma_0^{-1}(y_t) h \cdot h) u, \quad (2.84)$$

it follows from equation (2.83.a) that the function $\Psi(t), t \in [0, T]$, satisfies the following linear (parabolic type) partial differential equation

$$\begin{aligned} \frac{d}{dt} \Psi(t) &= F^*(t) \Psi(t), \\ \Psi(0) &= p_0, \end{aligned} \quad (2.85)$$

where p_0 is the initial density of x_0 .

For any $u, v \in H^1$, the operator F^* , as given by equation (2.84), gives rise to the following bilinear form

$$\begin{aligned} \langle F^* u, v \rangle &= \frac{1}{2} Z_t \cdot \sum_{i,j=1}^n \int_{R^n} \sigma_{ij} \frac{\partial h}{\partial x_i} \frac{\partial u}{\partial x_j} v dx - \frac{1}{2} \sum_{i,j=1}^n \int_{R^n} \sigma_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx \\ &\quad + \sum_{i=1}^n \int_{R^n} \tilde{a}_i \frac{\partial v}{\partial x_i} u dx + \sum_{i=1}^n \int_{R^n} \check{a}_i u v dx, \end{aligned} \quad (2.86)$$

where

$$\tilde{a}_i(t, x) \equiv a_i(x) - \frac{1}{2} \sum_{j=1}^n \frac{\partial \sigma_{ij}}{\partial x_j}(x) - \frac{1}{2} Z_t \cdot \sum_{j=1}^n \sigma_{ij}(x) \frac{\partial h}{\partial x_j}(x), \quad (2.87)$$

and

$$\begin{aligned} \check{a}_i(t, x) &\equiv \left(\frac{1}{2} \sum_{j=1}^n \frac{\partial \sigma_{ij}}{\partial x_j}(x) - a_i(x) \right) \frac{\partial h}{\partial x_j}(x) \cdot Z_t \\ &\quad + \frac{1}{2} \sum_{j=1}^n \sigma_{ij} \left(\frac{\partial h}{\partial x_i} \cdot Z_t \right) \left(\frac{\partial h}{\partial x_j} \cdot Z_t \right) - \frac{1}{2} (\Gamma_0^{-1}(y_t) h \cdot h). \end{aligned} \quad (2.88)$$

Under the assumptions (A1), (A3), (A5) and that

(A6) The function h is bounded and twice continuously differentiable on R^n , that is $h \in C_b^2(R^n)$,

(A7) $\Pr\{\sup_{0 \leq t \leq T} |Z_t| < \infty\} = 1$,

the operator F^* , as given by equation (2.86), satisfies also the properties (i) and (ii) given above. With this preparation, we now present the following existence result for the initial value problem (2.85).

Lemma 2.4 ([6] Theorem 5.1.1 P.278)

Suppose the operator F^* , as given by (2.86), satisfies the properties (i) and (ii); then for every initial density $p_0 \in L_2(R^n)$, the initial value problem (2.85) has a unique (weak) solution $\Psi \in L_2((0, T); H^1) \cap C([0, T]; L_2(R^n))$. ■

In the next chapter we shall use similar arguments as those given in section 2.3, to extend the result of Zakai (see Theorem 2.2) to the case where the parameters $\{a, b\}$ of the diffusion process $x_t, t \geq 0$, (see equation (2.1)), are perturbed by a temporally homogeneous continuous time finite state Markov chain.

2.6 Summary and Conclusion

In this chapter, an overview of the filtering problem of diffusion processes, which was considered by Kushner [55] and Zakai [76], has been presented. Using a Girsanov transformation (Radon-Nikodym derivative) of measures, induced on the canonical space (Ω, \mathcal{A}) by the systems (2.1) and (2.2) and the properties of the semigroup, we have presented a complete derivation for the Zakai equation. Further, using Ito's Lemma and the properties of the likelihood ratio, we have obtained the Kushner equation from the Zakai equation and show their equivalence. Finally, under the given assumptions, we have shown that Zakai equation can be treated as a linear parabolic partial differential equation, parametrized by the output process. The existence of solutions of this equation can be proved by standard partial differential equations arguments.

CHAPTER 3

NONLINEAR FILTERING OF SYSTEMS GOVERNED BY ITO EQUATIONS WITH JUMP PARAMETERS

3.1 Introduction

In this chapter we consider the filtering problem for stochastic systems governed by a class of nonlinear stochastic differential equations with drift and diffusion parameters perturbed by a temporally homogeneous Markov chain. The observation process (continuous) is assumed to be governed by an Ito differential equation with parameters dependent only on the state of the system. We use Girsanov transformation (see Lemma 2.1) and similar arguments as those of Zakai [76] (see chapter 2), to derive filter equations for the cases where the drift and diffusion parameters are perturbed by :

- (i) A temporally homogeneous Markov chain, or
- (ii) Deterministic processes, or
- (iii) Periodically observable Markov Chain.

In section 3.2 we formulate the filtering problem and introduce the necessary notations and assumptions that will be used in the derivation of the filter equations. In section 3.3 we use similar arguments as those of section 2.3 to derive coupled system of stochastic integral equations for the unnormalized density. In section 3.4 we use the results of section 3.3 to obtain a coupled system of stochastic partial differential equations of Zakai-type for the unnormalized density. In section 3.5 we discuss the question of existence of solutions of the filter equation. Finally, in section 3.6 we present some examples indicating the possible applications of our results.

3.2 Formulation of Filtering Problem, Assumptions and Notations

We consider the filtering problem for the process $\{x_t; t \geq 0\}$ governed by the following stochastic evolution equation

$$dx_t = a(x_t, \xi_t)dt + b(x_t, \xi_t)dW_t \quad (3.1)$$

with initial states x_0, ξ_0 , where the drift and dispersion parameters $\{a, b\}$ are perturbed by a temporally homogeneous continuous time finite state Markov chain $\{\xi_t; t \geq 0\}$. The observed process $\{y_t; t \geq 0\}$ is assumed to be governed by the following Ito equation

$$\begin{aligned} dy_t &= h(x_t)dt + dV_t, \\ y(0) &= 0, \end{aligned} \quad (3.2)$$

where x and a are vectors in R^n ; y and h are vectors in R^k ; b is $(n \times n)$ matrix and ξ takes values from a finite set $\Sigma \equiv \{e_1, e_2, \dots, e_N\}$. It is assumed that all the random processes and vectors described above are defined on a complete probability space $(\Omega_0, \mathcal{B}_0, P_0)$. Further, we assume that $\{W_t, V_t; t \geq 0\}$ are independent $(n+k)$ -dimensional standard Wiener processes defined on $(\Omega_0, \mathcal{B}_0, P_0)$ and independent of the initial states $\{x_0, \xi_0\}$. The transition probability matrix of the process $\{\xi_t; t \geq 0\}$ is defined by $S(t) \equiv \{S_{ij}(t), t \geq 0; i, j = 1, 2, \dots, N\}$ with infinitesimal transition rates $\{\lambda_{ij}; i, j = 1, 2, \dots, N\}$ given by

$$\lambda_{ij} = \begin{cases} \lim_{t \rightarrow 0} \frac{S_{ij}(t)}{t}, & \text{for } i = j \\ \lim_{t \rightarrow 0} \frac{S_{ij}(t) - 1}{t}, & \text{for } i \neq j. \end{cases} \quad (3.3)$$

The parameters $\{\lambda_{ij}\}$ satisfy the following properties

(P1) $\sum_{i \neq j} \lambda_{ij} + \lambda_{ii} = 0$, and

(P2) $\lambda_{ij} \geq 0$ for $i \neq j$.

A similar class of systems governed by stochastic differential equations of the form (3.1) was also considered in the literature [2, 24, 73] for the control problem (for the case where the process $x_t; t \geq 0$ is completely observable) where the existence of optimal control was proved and the necessary conditions of optimality were obtained. Stochastic linear evolution equations on infinite dimensional spaces with random operator valued coefficients perturbed by Markov chain, was considered by Ahmed [5] for the control problem. In [5], the author presented the existence Theorem of optimal controls and the corresponding necessary conditions of optimality for the case where the objective functional is quadratic.

In this chapter we are interested in the filtering problem which consists of estimating x_t or any function of x_t given the observation $\{y_s; s \leq t\}$. Note that the process $\{x_t; t \geq 0\}$ is not a Markov process, however, $\{(x_t, \xi_t); t \geq 0\}$ is a Markov process. For the derivation of the filter equation, we need the following notations and assumptions.

Notations :

Let $\{\eta_t; t \geq 0\}$ be any random process and let $\sigma\{\eta_s; s \leq t\}$ denote the σ -field generated by the process η up to time t . Define $\xi_0^t \equiv \{\xi_s; s \leq t\}$, $\sigma(\xi_0^t) \equiv \sigma\{\xi_s; s \leq t\}$, $\sigma(x_0^t) \equiv \sigma\{x_s; s \leq t\}$, $\sigma(y_0^t) \equiv \sigma\{y_s; s \leq t\}$, $\sigma(x_t) \equiv \sigma\{x_t\}$ and $\mathcal{F} = \sigma(\xi_0^T, W_0^T, V_0^T) \vee \sigma(x_0) \subset \mathcal{B}_0$. Let $\{\mathcal{F}_t; t \geq 0\}$ be an increasing family of sub- σ -fields contained in \mathcal{F} such that for each $t \in R_0 \equiv [0, \infty)$, ξ_t, x_t and y_t are \mathcal{F}_t -measurable. Let Ω denote the space of continuous functions on R_0 with values in $R^{(n+k)}$ and

let \mathcal{A} denote the Borel σ -algebra on Ω . We denote by $\mathcal{D}(R_0; \Sigma)$, the space of right continuous functions on R_0 , with values in Σ , having left limits. Let $L_2(I; R^n)$ denote the equivalence classes of measurable functions f on $I \equiv [0, T]$ with values in R^n such that $\int_I |f(t)|^2 < \infty$. For any Banach space E we use $C(I; E)$ to denote the space of strongly continuous E valued functions on I furnished with the uniform topology $\|f\|_C = \sup\{\|f(t)\|_E; t \in I\}$. For any pair of Banach spaces E and F , we use $\mathcal{L}(E, F)$ to denote the space of bounded linear operators from E to F . For notational convenience, we use $\varphi(t)$ and h to denote $\varphi(t, x)$ and $h(x)$, respectively, $x \in R^n$. Further notations will be introduced in the sequel as required.

Assumptions

- (A1) For every $e \in \Sigma$, the mappings $x \rightarrow a(x, e)$ and $x \rightarrow b(x, e)$ are continuous on R^n . Further, there exists a constant $\gamma > 0$ (independent of e) such that

$$\sigma \equiv (bb') \geq \gamma I,$$

where I denotes the identity matrix.

- (A2) For every $e \in \Sigma$, there exists a constant $k > 0$ (independent of e) such that

$$T|a(x, e) - a(y, e)|^2 + \|b(x, e) - b(y, e)\|^2 \leq k|x - y|^2,$$

for any $x, y \in R^n$.

- (A3) The functions $a_i(x, e), \sigma_{ij}(x, e)$, $\frac{\partial}{\partial x_i} a_i(x, e), \frac{\partial}{\partial x_i} \sigma_{ij}(x, e)$ and $\frac{\partial^2}{\partial x_i \partial x_j} \sigma_{ij}(x, e)$; $i, j = 1, 2, \dots, n$, are bounded and satisfy Holder condition on R^n , for all $e \in \Sigma$.

- (A4) The function h is continuous on R^n and there exists a constant $k > 0$ such that

$$|h(x) - h(y)|^2 \leq k|x - y|^2.$$

Further,

$$E \int_0^t |h(x_s)|^2 ds < \infty, \quad P_0 \text{ - a.s. (almost surely);}$$

along any solution $x_t; t \geq 0$, of (3.1).

(A5) The σ -field $\sigma(\xi_0^t); t \geq 0$, is right continuous independent of the σ -fields generated by the Wiener processes W_t, V_t and the initial state x_0 .

Note that the assumptions (A1), (A2), and (A4) are required to guarantee the existence of solutions and to prove the absolute continuity of measures induced by the systems (3.1)-(3.2).

Given any bounded measurable function f on R^n , our main goal is to compute the optimal estimate (in the mean square sense) of the process $f(x_t)$ given $\sigma(y_0^t)$. As indicated earlier in chapter 2, the solution of this filtering problem is given by

$$\hat{f}_t \equiv E\{f(x_t) | \sigma(y_0^t)\}. \quad (3.4)$$

In this chapter we will compute the above conditional expectation for the following cases :

Case (i)

The process $\{\xi_t; t \geq 0\}$ is a Markov chain with transition matrix $S(t)$ and infinitesimal (transition) rates λ_{ij} satisfying the properties (P1) and (P2).

Case (ii)

The process $\xi \in D(R_0, \Sigma)$ is deterministic.

Case (iii)

The process $\{\xi_t; t \geq 0\}$ is a Markov chain periodically observable. For example, during the intervals $(t_i, t_{i+1}]$; $i = 0, 2, 4, \dots$, the process ξ_t is observable while during the intervals $(t_i, t_{i+1}]$; $i = 1, 3, 5, \dots$, no measurement of ξ is available and hence the estimate \hat{f}_t of $f(x_t)$ must be based on its priori statistical information and the observed process $\{y_s; s \leq t\}$.

We use the approach involving absolute continuity of measures (see Lemma 2.1) to derive a system of coupled stochastic differential equations of Zakai-type for the unnormalized conditional density of the process $\{x_t; t \geq 0\}$ from which the estimate (3.4) can be computed. We present a complete derivation of the filter equation for case (i). The filter equations corresponding to the cases (ii) and (iii) can be easily obtained from that of case (i) as it will be indicated at the end of section 3.4.

3.3 Stochastic Integral Equation For The Unnormalized Density

In this section we use Girsanov transformation of measures induced on the canonical sample space (Ω, \mathcal{A}) by the systems (3.1) and (3.2) (see Lemma 2.1), to derive a system of coupled integral equations for the unnormalized density leading to the filter equation. In our derivations we will use similar arguments as those of chapter 2.

Let μ_1 and μ_2 denote the measures induced on the canonical sample space (Ω, \mathcal{A}) , respectively, by the system (3.1)-(3.2) and the system

$$\begin{aligned} dx_t &= a(x_t, \xi_t)dt + b(x_t, \xi_t)dW_t, \\ dy_t &= dV_t. \end{aligned}$$

For each $t \in R_0 \equiv [0, T]$, define

$$L_t \equiv \exp \left\{ -\frac{1}{2} \int_0^t |h(x_s)|^2 ds + \int_0^t h(x_s) \cdot dy_s \right\}, \quad (3.5)$$

where h satisfies assumption (A4). Using the results of Lemma 2.1, it is clear that if $EL_T = 1$, the process $\{W_t, y_t; t \geq 0\}$ is a Wiener process on the probability space $(\Omega, \mathcal{A}, \mu_2)$ with respect to the current of σ -algebra $\mathcal{F}_t; t \geq 0$, where

$$d\mu_1 = L_T d\mu_2,$$

and

$$\int_{\Omega} \gamma d\mu_1 = \int_{\Omega} \gamma L_T d\mu_2, \quad (3.5.a)$$

for any G -measurable random variable $\gamma, G \subset \mathcal{A}$.

For any bounded measurable function f on R^n , the conditional expectation (3.4) can be written as

$$E_1\{f(x_t)|\sigma(y_0^t)\} = E_1\{E_1\{f(x_t)|\sigma(\xi_t, y_0^t)\}|\sigma(y_0^t)\}, \quad (3.6)$$

where E_1 denotes the expectation with respect to the measure μ_1 . Using the fact that the measures μ_1 and μ_2 are absolutely continuous with respect to each other, it follows from (3.5.a) and Baye's formula that

$$E_1\{f(x_t)|\sigma(y_0^t)\} = \frac{E_2\{E_2\{L_t|\sigma(x_t, \xi_t, y_0^t)\}f(x_t)|\sigma(\xi_t, y_0^t)\}}{E_2\{E_2\{L_t|\sigma(x_t, \xi_t, y_0^t)\}|\sigma(\xi_t, y_0^t)\}}, \quad (3.7)$$

where E_2 denotes the expectation with respect to the measure μ_2 . Define

$$\tilde{L}_{\xi_t}(t, x_t) \equiv E_2\{L_t|\sigma(x_t, \xi_t, y_0^t)\},$$

and

$$Q_{\xi_t}(t, dx) \equiv Pr\{x_t \in dx|\sigma(\xi_t)\},$$

where the process \tilde{L}_{ξ_t} is $\sigma(y_0^t)$ -measurable given x_t and ξ_t . Note that the measure Q_{ξ_t} is well defined since

$$Q_k(t, \Gamma) \equiv Pr\{x_t \in \Gamma|\xi_t = e_k\} = \int_{A_k} Pr\{x_t \in \Gamma|\xi_0^t = \eta\} d\nu_{\xi_t}(\eta), \quad (3.8)$$

where ν_{ξ_t} is the measure induced by the process $\{\xi_t; t \geq 0\}$ on the Borel subsets of $D(R_0, \Sigma)$, $A_k \equiv \{\xi \in D : \xi_t = e_k\}$, and $\Gamma \in \mathcal{B}(R^n)$.

Taking f as the indicator function on Γ , and using the independence of the σ -fields $\sigma(x_0^t)$ and $\sigma(y_0^t)$ under μ_2 , it follows from (3.7), Fubini's Theorem and the above definitions that the conditional measure $Pr\{x_t \in \Gamma|\xi_t = e_k; \sigma(y_0^t)\} \equiv P_k(t, \Gamma|\sigma(y_0^t))$ is given by

$$P_k(t, \Gamma|\sigma(y_0^t)) = \frac{\int_{\Gamma} \tilde{L}_k(t, x) Q_k(t, dx)}{\int_{R^n} \tilde{L}_k(t, x) Q_k(t, dx)} \quad (3.9)$$

for each $t \in R_0, 1 \leq k \leq N$.

With this preparations we now introduce the following result which shows that the process $\tilde{L}_k; 1 \leq k \leq N$, which is $\sigma(y_0^t)$ measurable, can be obtained by solving a system of coupled stochastic integral equations driven by the y process. This result is similar to that of Theorem 2.1 except that the process \tilde{L} is now dependent on the state of the process $\xi_t, t \geq 0$.

Theorem 3.1

The process $\{\tilde{L}_k; 1 \leq k \leq N\}$ satisfies the following system of stochastic integral equations

$$\tilde{L}_k(t, \gamma) = 1 + \int_0^t \sum_{\ell=1}^N \int_{R^n} \tilde{L}_\ell(s, \eta) h(\eta) P_{t,s}(\gamma, e_k; d\eta, e_\ell) \cdot dy_s, \quad (3.10)$$

and

$$\begin{aligned} \tilde{L}_k(t, \gamma) = & \sum_{\ell=1}^N \int_{R^n} \tilde{L}_\ell(s, \eta) P_{t,s}(\gamma, e_k; d\eta, e_\ell) \\ & + \int_s^t \sum_{\ell=1}^N \int_{R^n} \tilde{L}_\ell(\theta, \eta) h(\eta) P_{t,\theta}(\gamma, e_k; d\eta, e_\ell) \cdot dy_\theta, \end{aligned} \quad (3.11)$$

for $(t, \gamma) \in R_0 \times R^n; 1 \leq k \leq N$, where $P_{t,s}(\gamma, e_k; d\eta, e_\ell)$ denotes the backward transition probability $Pr\{x_s \in d\eta, \xi_s = e_\ell | x_t = \gamma, \xi_t = e_k\}, s \leq t$, of the \mathcal{F}_t -Markov process $\{x_t, \xi_t; t \geq 0\}$.

Proof

Define

$$\beta(t) = -\frac{1}{2} \int_0^t |h(x_s)|^2 ds + \int_0^t h(x_s) \cdot dy_s.$$

Then it is clear from equation (3.5) that

$$L_t = \exp(\beta(t)).$$

Applying Ito's Lemma (Lemma 2.2), the Ito differential of the process $L_t, t \geq 0$ is given by

$$\begin{aligned} dL_t = & \exp(\beta(t)) \left[-\frac{1}{2} |h(x_t)|^2 dt + h(x_t) \cdot dy_t \right] \\ & + \frac{1}{2} \exp(\beta(t)) |h(x_t)|^2 dt, \end{aligned}$$

which implies that

$$dL_t = L_t h(x_t) \cdot dy_t,$$

$$L_0 = 1,$$

and hence

$$L_t = 1 + \int_0^t L_s h(x_s) \cdot dy_s. \quad (3.12)$$

Thus

$$\begin{aligned} E_2\{L_t | \sigma(x_t, \xi_t, y_0^t)\} &\equiv \tilde{L}_{\xi_t}(t, x_t) = 1 + E_2\left\{\int_0^t L_s h(x_s) \cdot dy_s | \sigma(x_t, \xi_t, y_0^t)\right\} \\ &= 1 + \int_0^t E_2\{L_s h(x_s) | \sigma(x_t, \xi_t, y_0^t)\} \cdot dy_s. \end{aligned} \quad (3.13)$$

Using the independence of the σ -fields $\sigma(x_0^s, \xi_0^s)$ and $\sigma(y_0^s)$ under μ_2 , and the fact that (x_t, ξ_t) is an \mathcal{F}_t -Markov process, it follows that $\sigma(x_0^s, \xi_0^s, y_0^s)$ and $\sigma\{y_\theta - y_s; s \leq \theta \leq t\}$ are independent given $\sigma(x_t, \xi_t, y_0^s)$. Therefore,

$$\begin{aligned} \tilde{L}_{\xi_t}(t, x_t) &= 1 + \int_0^t E_2\{L_s h(x_s) | \sigma(x_t, \xi_t, y_0^s)\} \cdot dy_s \\ &= 1 + \int_0^t E_2\{E_2\{L_s | \sigma(x_t, \xi_t, x_s, \xi_s, y_0^s)\} \times \\ &\quad \times h(x_s) | \sigma(x_t, \xi_t, y_0^s)\} \cdot dy_s. \end{aligned} \quad (3.14)$$

By the Markov property of the process $(x_t, \xi_t; \mathcal{F}_t)$ and the independence of $\sigma(x_0^t, \xi_0^t)$ and $\sigma(y_0^t)$ under μ_2 , it follows that $\sigma(x_0^s, \xi_0^s, y_0^s)$ and $\sigma(x_t, \xi_t)$ are independent given $\sigma(x_s, \xi_s, y_0^s)$. Hence

$$E_2\{L_s | \sigma(x_t, \xi_t, x_s, \xi_s, y_0^s)\} = E_2\{L_s | \sigma(x_s, \xi_s, y_0^s)\} \equiv \tilde{L}_{\xi_s}(s, x_s). \quad (3.15)$$

Substituting equation (3.15) into (3.14) we obtain

$$\tilde{L}_{\xi_t}(t, x_t) = 1 + \int_0^t E_2\{\tilde{L}_{\xi_s}(s, x_s) h(x_s) | \sigma(x_t, \xi_t, y_0^s)\} \cdot dy_s. \quad (3.15.a)$$

Using the fact that the σ -fields $\sigma(x_0^t, \xi_0^t)$ and $\sigma(y_0^t)$ are independent under the measure μ_2 , it follows from Fubini's Theorem and the definition of $P_{t,s}; s \leq t$, that (3.15.a) can be written as

$$\tilde{L}_{\xi_t}(t, \gamma) = 1 + \int_0^t \sum_{\ell=1}^N \int_{R^n} \tilde{L}_{\ell}(s, \eta) h(\eta) P_{t,s}(\gamma, \xi_t; d\eta, e_{\ell}) \cdot dy_s,$$

which is equation (3.10). Using equation (3.12) we obtain

$$L_t = L_s + \int_s^t L_{\theta} h(x_{\theta}) \cdot dy_{\theta},$$

and hence

$$\tilde{L}_{\xi_t} = E_2\{L_s | \sigma(x_t, \xi_t, y_0^t)\} + \int_s^t E_2\{L_{\theta} h(x_{\theta}) | \sigma(x_t, \xi_t, y_0^t)\} \cdot dy_{\theta}.$$

Following similar arguments as above it is easy to verify that equation (3.11) follows from the above equation. This completes the proof. ■

On the basis of the above Theorem, we now derive a system of coupled stochastic integral equations for the unnormalized density (see equation (3.9)) from which a Zakai-type equation can be obtained. In the sequel we need the semigroups generated by the process $\{x_t, \xi_t; \mathcal{F}_t\}$. Let Ψ be any bounded measurable function on $R^n \times \Sigma$ and let $\{T_{s,t}; s \leq t\}$ denote the (backward) semigroup given by [2,4,73]

$$(T_{s,t}\Psi)(\eta, e_t) \equiv E\{\Psi(x_t, \xi_t) | x_s = \eta, \xi_s = e_t\}.$$

Let

$$P_{s,t}(\eta, e_t; \Gamma, e_k) \equiv Pr\{x_t \in \Gamma, \xi_t = e_k | x_s = \eta, \xi_s = e_t\},$$

$\eta \in R^n, \Gamma \in \mathcal{B}(R^n), e_t, e_k \in \Sigma$, be the kernel corresponding to the semigroup $T_{s,t}$ so that

$$(T_{s,t}\Psi)(\eta, e_t) = \sum_{k=1}^N \int_{R^n} \Psi(\gamma, e_k) P_{s,t}(\eta, e_t; d\gamma, e_k), \quad (3.16)$$

and

$$S \cdot \lim_{t \rightarrow s} (T_{s,t}\Psi)(\eta, e_t) = \Psi(\eta, e_t),$$

for any bounded measurable function Ψ and $\ell = 1, 2, \dots, N$.

It is known that (see for example [2,4,73]) the infinitesimal generator corresponding to the semigroup $\{T_{s,t}; s \leq t\}$, is given by

$$(\mathbf{L}\Psi)(x, e_\ell) = (\mathbf{A}\Psi)_\ell(x) + \sum_{k=1}^N \lambda_{k,\ell} \Psi_k(x); \quad (3.17)$$

for all $(x, e_\ell) \in R^n \times \Sigma$, where $\Psi_k(x) \equiv \Psi(x, e_k)$ and

$$\begin{aligned} (\mathbf{A}\Psi)_\ell(x) &= (\mathbf{A}\Psi)(x, e_\ell) \\ &= \sum_{i=1}^n a_i(x, e_\ell) \Psi_{x_i}(x, e_\ell) + \frac{1}{2} \sum_{i,j=1}^n (bb')_{ij}(x, e_\ell) \Psi_{x_i x_j}(x, e_\ell). \end{aligned} \quad (3.17.a)$$

Suppose for each $\eta \in R^n, e_\ell, e_k \in \Sigma$ and $0 \leq s \leq t$, the measure $\Gamma \rightarrow P_{s,t}(\eta, e_\ell; \Gamma, e_k)$ is absolutely continuous with respect to Lebesgue measure on R^n admitting a density so that

$$P_{s,t}(\eta, e_\ell; \Gamma, e_k) = \int_{\Gamma} p_{s,t}(\eta, e_\ell; \gamma, e_k) d\gamma,$$

for each $\Gamma \in \mathcal{B}(R^n)$. For each $t \geq 0$ let ν_t denote the measure induced by the process $\{x_t, \xi_t\}$ on $\mathcal{B}(R^n \times \Sigma)$ and suppose for each $e_\ell \in \Sigma$, the measure $E \rightarrow \nu_t(E, e_\ell), E \in \mathcal{B}(R^n)$, is absolutely continuous with respect to Lebesgue measure admitting a density $\beta_t(\eta, e_\ell)$ so that

$$\nu_t(E, e_\ell) = \int_E \beta_t(\eta, e_\ell) d\eta.$$

For notational convenience, from now on, we shall use the $\beta_\ell(t, \eta)$ to denote $\beta_t(\eta, e_\ell)$.

Define

$$(T_{s,t}^* \beta)_k(t, \gamma) \equiv \sum_{\ell=1}^N \int_{R^n} p_{s,t}(\eta, e_\ell; \gamma, e_k) \beta_\ell(s, \eta) d\eta, \quad (3.18)$$

where $T_{s,t}^*$ is the semigroup corresponding to the (forward) Kolmogorov operator L^* of the process $\{(x_t, \xi_t); t \geq 0\}$, and L^* is the formal adjoint of the operator L (see equation (3.17)). Then under our assumptions, the kernel $p_{s,t}(\eta, e_\ell; \gamma, e_k)$ of the semigroup $T_{s,t}^*$, as defined above, satisfies the following Kolmogorov (or Fokker-Plank) forward equation

$$\frac{\partial}{\partial t} p_{s,t}(\eta, e_\ell; x, e_k) = L^* p_{s,t}(\eta, e_\ell; x, e_k), \quad (3.18.a)$$

for $s \leq t$ and all $x, \eta \in R^n; e_\ell, e_k \in \Sigma$. Further, under the assumptions (A1) and (A3) it follows from (3.18.a) that the semigroup $T_{s,t}^*, s \leq t$, satisfies the following properties

$$\begin{aligned} \text{(Si)} \quad (T_{s,t}^* \beta)_k(t, \gamma) &= (T_{s,\theta}^* \beta)_k(\theta, \gamma) + \int_\theta^t A^*(T_{\theta,\tau}^* \beta)_k(\tau, \gamma) d\tau \\ &+ \sum_{m=1}^N \int_\theta^t \lambda_{mk}(T_{\theta,\tau}^* \beta)_m(\tau, \gamma) d\tau, \end{aligned} \quad (3.19)$$

for all $0 \leq s \leq t$. Let \mathbf{H} denote the family of functions $\{\beta\} = \{\beta(\gamma, e_k) \equiv \beta_k(\gamma), \gamma \in R^n, e_k \in \Sigma\}$ satisfying the integrability condition

$$\sum_{k=1}^N \int_{R^n} |(T_{s,t}^* \beta)_k(\gamma)|^2 d\gamma < \infty.$$

Then there exist a constant $M > 0$ such that

$$\text{(Sii)} \quad \sum_{k=1}^N \int_{R^n} [(T_{s,t}^* f)_k(x)]^2 dx \leq M \sum_{k=1}^N \int_{R^n} f_k^2(x) dx,$$

for all $f \in \mathbf{H}$ and $0 \leq s \leq t < \infty$.

With this preparation, we now show in the following corollary that the unnormalized density satisfies a system of coupled stochastic integral equations.

Corollary 3.1

Suppose for each $\eta \in R^n, e_k, e_l \in \Sigma$, the measure $\Gamma \rightarrow P_{s,t}(\eta, e_l; \Gamma, e_k), s \leq t$, is absolutely continuous with respect to Lebesgue measure admitting the density $\gamma \rightarrow p_{s,t}(\eta, e_l; \gamma, e_k)$ and the semigroup $T_{s,t}^*$ satisfy the properties (Si) and (Sii). Then the measure $\Gamma \rightarrow P_k(t, \Gamma | \sigma(y_0^t)), t \geq 0$, (see equation (3.9)) is also absolutely continuous with respect to Lebesgue measure admitting a density $p_k(t, \gamma | \sigma(y_0^t))$ and that

$$p_k(t, \gamma | \sigma(y_0^t)) = \frac{\varphi_k(t, \gamma)}{\int_{R^n} \varphi_k(t, \gamma) d\gamma}, \quad (3.20)$$

for $k = 1, 2, \dots, N$, where $\{\varphi_k; 1 \leq k \leq N\}$ satisfies the following system of coupled stochastic integral equations

$$\varphi_k(t, \gamma) = q_k(t, \gamma) + \int_0^t (T_{s,t}^*(\varphi h))_k(t, \gamma) \cdot dy_s, \quad (3.21)$$

and

$$\varphi_k(t, \gamma) = (T_{s,t}^*(\varphi))_k(t, \gamma) + \int_s^t (T_{\theta,t}^*(\varphi h))_k(t, \gamma) \cdot dy_\theta, \quad (3.21)$$

almost surely (a.s.) for $1 \leq k \leq N$, and $(t, \gamma) \in R_0 \times R^n$, where q_k is the density corresponding to the measure Q_k (see equation (3.8)) and $T_{s,t}^*$ is given by (3.18).

Proof

Let Q_0 denote the initial distribution of x_0 and $\Theta(t)$ denote the probability distribution of the process $\{\xi_t; t \geq 0\}$ on Σ . Then one can easily verify that the measure $Q_k(t, \Gamma)$ (see equation (3.8)) is given by

$$Q_k(t, \Gamma) = \sum_{l=1}^N \frac{\Theta_l(0)}{\Theta_k(t)} \int_{R^n} P_{0,t}(z, e_l; \Gamma, e_k) Q_0(dz).$$

Since by assumption the measure $\Gamma \rightarrow P_{0,t}(z, e_l; \Gamma, e_k)$ is absolutely continuous with respect to Lebesgue measure, it is clear that the measure $Q_k(t, \cdot)$ admits a density $q_k(t, \cdot)$ so that

$$Q_k(t, \Gamma) = \int_{\Gamma} q_k(t, \gamma) d\gamma, \quad 1 \leq k \leq N,$$

for $\Gamma \in \mathcal{B}(R^n)$. Therefore, it follows from equation (3.9) that the measure $\Gamma \rightarrow P_k(t, \Gamma | \sigma(y_0^t))$ is also absolutely continuous with respect to Lebesgue measure and hence it admits a density which we denote by $p_k(t, \gamma | \sigma(y_0^t))$, $1 \leq k \leq N$. Defining

$$\varphi_k(t, \gamma) \equiv \tilde{L}_k(t, \gamma) q_k(t, \gamma),$$

and using equation (3.9), one obtains equation (3.20). Multiplying equations (3.10) and (3.11) by $q_k(t, \gamma)$ and using the above definition, we have

$$\begin{aligned} \varphi_k(t, \gamma) &= q_k(t, \gamma) + \int_0^t \sum_{\ell=1}^N \int_{R^n} \tilde{L}_\ell(s, \eta) h(\eta) \times \\ &\quad \times P_{t,s}(\gamma, e_k; d\eta, e_\ell) q_k(t, \gamma) \cdot dy_s, \end{aligned}$$

and

$$\begin{aligned} \varphi_k(t, \gamma) &= \sum_{\ell=1}^N \int_{R^n} \tilde{L}_\ell(s, \eta) P_{t,s}(\gamma, e_k; d\eta, e_\ell) q_k(t, \gamma) \\ &\quad + \int_s^t \sum_{\ell=1}^N \int_{R^n} \tilde{L}_\ell(\theta, \eta) h(\eta) P_{t,\theta}(\gamma, e_k; d\eta, e_\ell) q_k(t, \gamma) \cdot dy_\theta. \end{aligned}$$

Using equation (3.18) and the duality

$$P_{t,s}(\gamma, e_k; d\eta, e_\ell) q_k(t, \gamma) = P_{s,t}(d\eta, e_\ell; \gamma, e_k) q_\ell(s, \eta),$$

for $s \leq t$, it follows from similar arguments as those of Corollary 2.1 that equations (3.21) and (3.21.a) follow from the above expressions. This completes the proof of the corollary. ■

3.4 Zakai-Type Equation For Unnormalized Density

In this section, we use equation (3.21) and the properties of the semigroup $T_{s,t}^*$ to derive a system of coupled stochastic differential equations for the unnormalized density $\{\varphi_k; 1 \leq k \leq N\}$ from which the conditional expectation $E_1\{f(x_t)|\sigma(y_0^t)\}$ can be computed using the following relation

$$\begin{aligned} E_1\{f(x_t)|\sigma(y_0^t)\} &= \sum_{k=1}^N \Theta_k(t) \frac{\langle \varphi_k(t), f \rangle}{\langle \varphi_k, 1 \rangle} \\ &= \sum_{k=1}^N \Theta_k(t) \int_{R^n} f(x) p_k(t, x | \sigma(y_0^t)) dx, \end{aligned} \quad (3.23)$$

where $\varphi_k(t) \equiv \varphi_k(t, \cdot)$ and

$$\langle \varphi_k(t), f \rangle \equiv \int_{R^n} \varphi_k(t, x) f(x) dx.$$

Further, for $0 \leq \tau \leq t$, the probability distribution $\Theta(t); t \geq 0$, is given by

$$\Theta_k(t) = \sum_{i=1}^N S_{ik}(t - \tau) \Theta_i(\tau), \quad (3.23.a)$$

where $\{S_{ik}; i, k = 1, 2, \dots, N\}$ are the elements of the transition matrix $S(t)$.

Definition 3.1

The semigroup $T_{s,t}^*; s \leq t$, given by (3.18), is said to be of class \mathcal{M} if it satisfies the properties (Si)-(Sii) and

(Siii) Let f be any bounded, continuous function on R^n whose first and second derivatives are bounded, continuous and satisfy Holder condition on R^n . Then $(T_{s,t}^* f)(t, x)$ and its first and second derivatives, with respect to x , are continuous on $(0, T) \times R^n$ and bounded on $[0, T] \times G$, where G is any bounded subset of R^n and $s \leq t$.

Theorem 3.2 (Zakai Type Equation)

Suppose that assumptions (A1)-(A4) hold and the semigroup $T_{s,t}^*$ is of class \mathcal{M} . Let $\varphi_k; 1 \leq k \leq N$, satisfy equation (3.21), and a.s. the functions φ_k and $\varphi_k h; 1 \leq k \leq N$, together with their first and second partial derivatives with respect to $x \in R^n$ are continuous, bounded and satisfy Holder condition on R^n , uniformly in $t \in [0, T]$. Then $\varphi_k; 1 \leq k \leq N$, satisfies the following system of coupled stochastic differential equations

$$d\varphi_k(t, x) = \{ (A^* \varphi)_k(t, x) + \sum_{m=1}^N \lambda_{m,k} \varphi_m(t, x) \} dt + \varphi_k(t, x) h(x) \cdot dy_t \quad (3.24)$$

$$\varphi_k(0, x) = p_0(x), \quad 1 \leq k \leq N,$$

a.s. $(t, x) \in (R_0 \times R^n)$.

Proof

The proof essentially follows from similar arguments as those of Theorem 2.2 and hence omitted. ■

Based on the above result, we use Lemma 2.3 to obtain the corresponding differential equation for the normalized density.

Let the operator L be given by (3.17) and let L^* be the formal adjoint of L . Then the filter equation (3.24) can be written as

$$d\varphi_k(t) = (L^* \varphi)_k(t) dt + \varphi_k(t) h \cdot dy_t,$$

$$\varphi_k(0) = p_0, \quad (3.25)$$

for all $t \in [0, T]$ and $1 \leq k \leq N$, where $\varphi_k(t) \equiv \varphi_k(t, \cdot)$ and $h \equiv h(\cdot)$. Let f be any twice continuously differentiable function on R^n having compact support, that is

$f \in C_0^2(R^n)$. Then multiplying (3.25) by $f \in C_0^2(R^n)$ and integrating over R^n , we have

$$\begin{aligned} d \langle \varphi_k(t), f \rangle &= \langle \varphi_k(t), (Lf)_k \rangle dt + \langle \varphi_k(t), hf \rangle \cdot dy_t, \\ \langle \varphi_k(0), f \rangle &= \langle p_0, f \rangle, \end{aligned} \quad (3.26)$$

for all $f \in C_0^2(R^n)$, where

$$\langle \varphi_k(t), f \rangle \equiv \int_{R^n} \varphi_k(t, x) f(x) dx, \quad 1 \leq k \leq N.$$

Setting $f \equiv 1$, it follows from (3.26) and the fact $\sum_{i \neq j} \lambda_{ij} + \lambda_{ii} = 0$, that

$$\begin{aligned} d \langle \varphi_k(t), 1 \rangle &= \langle \varphi_k(t), h \rangle \cdot dy_t, \\ \langle \varphi_k(0), 1 \rangle &= 1, \end{aligned} \quad (3.27)$$

for all $1 \leq k \leq N$, which implies that

$$\tilde{\varphi}_k(t) \equiv \langle \varphi_k(t), 1 \rangle = 1 + \int_0^t \langle \varphi_k(t), h \rangle \cdot dy_t. \quad (3.28)$$

Since

$$E_1\{f(x_t) | \sigma(y_0^t); \xi_t = e_k\} \equiv \hat{f}_k(t) = \frac{\langle \varphi_k(t), f \rangle}{\langle \varphi_k(t), 1 \rangle},$$

for all $1 \leq k \leq N$, it follows by similar arguments as those of section 2.4 that the likelihood ratio $E_2\{L_t | \sigma(y_0^t); \xi_t = e_k\}$, is given by

$$\begin{aligned} E_2\{L_t | \sigma(y_0^t); \xi_t = e_k\} &\equiv \tilde{\varphi}_k(t) \\ &= 1 + \int_0^t \langle \varphi_k(t), h \rangle \cdot dy_t. \end{aligned} \quad (3.29)$$

Since the (normalized) conditional density $p_k(t, \cdot | \sigma(y_0^t)) \equiv p_k(t), t \geq 0$, is related to the unnormalized conditional density $\varphi_k(t), t \geq 0$, through the relation

$$p_k(t) = \tilde{\varphi}_k^{-1}(t) \varphi_k(t), \quad 1 \leq k \leq N, \quad t \geq 0,$$

it follows from the differential equations (3.26) and (3.27) and Lemma 2.3 that the density $p_k(t), 1 \leq k \leq N, t \geq 0$, is governed by the following system of coupled nonlinear stochastic partial differential equations

$$\begin{aligned} dp_k(t) &= (L^*p)_k(t)dt + (h - \hat{h}_k(t))p_k(t) \cdot [dy_t - \hat{h}_k(t)dt], \\ p_k(0) &= p_0, \quad \text{for all } 1 \leq k \leq N, \end{aligned} \quad (3.30)$$

where $\hat{h}_k(t) \equiv E_1\{h(x_t) | \sigma(y_0^t); \xi_t = e_k\}$, and L^* is the formal adjoint of the operator L given by (3.17).

In the remaining part of this section, we shall use the result of Theorem 3.2 to obtain the filter equations corresponding to the cases (ii) and (iii) as stated in section 3.2.

Case (ii) (Deterministic Case)

Suppose that the process $\{\xi_t; t \geq 0\}$ is deterministic and $\xi_t = \eta_t; t \geq 0$ with probability one (W.P.1), where η is a given element of $D(R_0; \Sigma)$. In this case, the transition matrix $\{S_{ij}(t); i, j = 1, 2, \dots, N\}$ is given by

$$S_{ij}(t) \equiv Pr\{\xi_{s+t} = e_j | \xi_s = e_i\} = \delta_{e_i, \eta_s} \delta_{e_j, \eta_{s+t}}$$

where δ is the dirac measure. Using equation (3.3) and the fact that η is right continuous, it follows that

$$\lambda_{ij} = \lim_{t \rightarrow 0} \frac{S_{ij}(t)}{t} = 0, \quad \text{for } i \neq j,$$

and hence by the property (P1), $\lambda_{ij} = 0$ for all $i, j = 1, 2, \dots, N$. Therefore, in this case, equation (3.24) reduces to

$$d\varphi_\eta(t, x) = (A^*\varphi)_\eta(t, x)dt + (h\varphi)_\eta(t, x) \cdot dy_t, \quad (3.31)$$

giving the usual Zakai equation (2.41). Hence from equation (3.23), it follows that the conditional expectation $E_1\{f(x_t)|\sigma(y_0^t)\}$ is given by

$$E_1\{f(x_t)|\sigma(y_0^t)\} = \frac{\langle \varphi_\eta(t), f \rangle}{\langle \varphi_\eta(t), 1 \rangle}$$

Case (iii) (Periodically Observable)

Suppose that the process $\{\xi_i; t \geq 0\}$ is a Markov chain which is observable during the time intervals $(t_i, t_{i+1}]$; $i = 0, 2, 4, 6, \dots$, while no observation is possible during the intervals $(t_i, t_{i+1}]$; $i = 1, 3, 5, 7, \dots$. In this case, using similar arguments as in case (ii), it follows from (3.24), for $t \in (t_i, t_{i+1}]$; $i = 0, 2, 4, \dots$, and $\xi_i^{t_{i+1}} \equiv \eta$, that

$$d\Psi_\eta(t, x) = (A^*\Psi)_\eta(t, x)dt + (h\Psi)_\eta(t, x) \cdot dy_t, \quad (3.32)$$

with

$$\Psi(t_i, x) = \begin{cases} \sum_{m=1}^N \Theta_m(t_i - 0) \varphi_m(t_i, x), & \text{for } i = 2, 4, 6, \dots, \\ p_0(x), & \text{for } i = 0, \end{cases}$$

where p_0 is the initial density of x_0 and $\varphi_m(t_i, x)$ is the solution of the following system of coupled stochastic partial differential equations

$$\begin{aligned} d\varphi_m(t, x) &= \{(A^*\varphi)_m(t, x) + \sum_{\ell=1}^N \lambda_{\ell m} \varphi_\ell(t, x)\}dt + (h\varphi)_m(t, x) \cdot dy_t, \\ \varphi_{\xi_i}(t_i, x) &= \Psi(t_i, x), \end{aligned} \quad (3.33)$$

for $t \in (t_i, t_{i+1}]$; $i = 1, 3, 5, 7, \dots$ and $1 \leq m \leq N$. Clearly, from equations (3.23), (3.32) and (3.33), the conditional expectation $E_1\{f(x_t)|\sigma(y_0^t)\}$ is given by

$$E_1\{f(x_t)|\sigma(y_0^t)\} = \begin{cases} \frac{\langle \Psi_\eta(t), f \rangle}{\langle \Psi_\eta(t), 1 \rangle}, & \text{for } t \in (t_i, t_{i+1}] ; i = 0, 2, 4, \dots \\ \sum_{k=1}^N \Theta_k(t) \frac{\langle \varphi_k(t), f \rangle}{\langle \varphi_k(t), 1 \rangle}, & \text{for } t \in (t_i, t_{i+1}] ; i = 1, 3, 5, \dots \end{cases}$$

where Ψ and φ are the solutions of (3.32) and (3.33), respectively.

3.5 Remarks On The Existence Of Solutions Of Filter Equation

In this section, we use similar arguments as those of chapter 2 and the standard result on existence of solutions of partial differential equations (see Lemma 2.4), to show that the filter equation (3.24) has a unique (weak) solution.

Define

$$B^* = \begin{pmatrix} A_1^* & 0 & \cdots \\ 0 & A_2^* & \cdots \\ \vdots & \vdots & \ddots \\ \vdots & \vdots & \vdots \\ 0 & \cdots & A_N^* \end{pmatrix}, \quad (3.34)$$

and

$$\Lambda^* = \begin{pmatrix} \lambda_{11} & \lambda_{21} & \cdots & \lambda_{N1} \\ \lambda_{12} & \lambda_{22} & \cdots & \lambda_{N2} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{1N} & \lambda_{2N} & \cdots & \lambda_{NN} \end{pmatrix}, \quad (3.35)$$

where the elements $\{\lambda_{ij}; i, j = 1, 2, \dots, N\}$ satisfy the properties (P1) and (P2) given in section 3.2, and A_ℓ^* is given by

$$A_\ell^* f \equiv - \sum_{i=1}^n \frac{\partial}{\partial x_i} (a_i(x, e_\ell) f) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} (\sigma_{ij}(x, e_\ell) f), \quad (3.36)$$

for all $1 \leq \ell \leq N$. Defining

$$\Gamma^* f \equiv B^* f + \Lambda^* f, \quad (3.37)$$

and

$$\Phi(t) \equiv (\varphi_1(t), \varphi_2(t), \dots, \varphi_N(t))',$$

it follows from equation (3.24) that the density $\Phi(t); 0 \leq t \leq T < \infty$ satisfies the following stochastic differential equation

$$\begin{aligned} d\Phi(t) &= \Gamma^* \Phi(t) dt + \Phi(t) (h \cdot dy_t), \\ \Phi(0) &= \Phi_0, \end{aligned} \quad (3.38)$$

for all $t \in [0, T]$, where $h \equiv h(x)$ and $\Phi_0 \equiv (p_0, \dots, p_0)'$. Using similar arguments as those of chapter 2 and Lemma 2.4, we show that equation (3.38) has a unique (weak) solution. For this we need the following definitions.

Let $L_2(R^n)$ denote the equivalence classes of measurable functions f on R^n such that $\int_{R^n} |f(x)|^2 dx < \infty$. Consider the Sobolev space

$$H^1 \equiv \{f \in \mathcal{X} : \frac{\partial f}{\partial x_i} \in \mathcal{X}; 1 \leq i \leq n\},$$

where $\mathcal{X} \equiv (L_2(R^n))^N$ denotes N -copies of $L_2(R^n)$. Let $H^{-1} \equiv (H^1)'$ be the dual space of H^1 and let $\mathcal{L}(H^1, H^{-1})$ denote the class of bounded linear operators from H^1 to H^{-1} . Hence for any $u, v \in H^1$, it follows from equation (3.37) that

$$\sum_{m=1}^N \sum_{\ell=1}^N \langle \Gamma_{m\ell}^* u_\ell, v_m \rangle = \sum_{m=1}^N \langle A_{m\ell}^* u_m, v_m \rangle + \sum_{m=1}^N \sum_{\ell=1}^N \Lambda_{m\ell}^* \langle u_\ell, v_m \rangle, \quad (3.39)$$

where $\langle \cdot, \cdot \rangle$ denotes the pairing of H^1 and H^{-1} ; $\Gamma_{m\ell}^*, \Lambda_{m\ell}^*$ denote the components of the operators Γ^*, Λ^* , respectively, and x_ℓ denotes the ℓ -th component of the vector x . Integrating equation (3.39) by parts and using equation (3.36), one obtains

$$\begin{aligned} \sum_{m,\ell=1}^N \langle \Gamma_{m\ell}^* u_\ell, v_m \rangle &= -\frac{1}{2} \sum_{m=1}^N \int_{R^n} \sum_{i,j=1}^n \sigma_{ij}^m \frac{\partial u_m}{\partial x_i} \frac{\partial v_m}{\partial x_j} dx \\ &\quad + \sum_{m=1}^N \int_{R^n} \sum_{i=1}^n \tilde{a}_i^m \frac{\partial u_m}{\partial x_i} v_m dx \\ &\quad + \sum_{m,\ell=1}^N \Lambda_{m\ell}^* \int_{R^n} u_\ell v_m dx. \end{aligned} \quad (3.40)$$

where

$$\tilde{a}_i^m \equiv \tilde{a}_i(x, e_m) \equiv a_i(x, e_m) - \frac{1}{2} \sum_{j=1}^n \frac{\partial \sigma_{ij}}{\partial x_j}(x, e_m), \quad (3.41)$$

and

$$\sigma_{ij}^m \equiv \sigma_{ij}(x, e_m) \equiv (bb')_{ij}(x, e_m).$$

Under the assumptions (A1), (A3) and $h \in C_b^2(R^n)$, it is clear that $\Gamma^* \in \mathcal{L}(H^1, H^{-1})$ and the following properties hold

(Pi) For any $u, v \in H^1$ there exists a constant $c > 0$ such that

$$| \langle \Gamma^* u, v \rangle | \leq c \|u\|_{H^1} \|v\|_{H^1}.$$

(Pii) There exist constants $\beta > 0$ and $\eta \geq 0$ such that

$$- \langle \Gamma^* u, u \rangle + \eta \|u\|_{L^2}^2 \geq \beta \|u\|_{H^1}^2.$$

With this preparation, we now follow similar arguments as those given in section 2.5 to show that equation (3.38) has a unique (weak) solution.

Define

$$\Phi(t) \equiv \exp(h \cdot y_t) V(t), \quad (3.42)$$

and

$$\tilde{\Gamma}^*(t)f \equiv \exp(-h \cdot y_t) \Gamma^*(f \exp(h \cdot y_t)) - \frac{1}{2} |h|^2 f, \quad (3.43)$$

where $\Phi(t)$ satisfies the stochastic partial differential equation (3.38) and ' \cdot ' denotes the scalar product. Multiplying equation (3.38) by $\exp(-h \cdot y_t)$ we have

$$\exp(-h \cdot y_t) d\Phi(t) = \exp(-h \cdot y_t) \tilde{\Gamma}^* \Phi(t) dt + \exp(-h \cdot y_t) \Phi(t) (h \cdot dy_t). \quad (3.44)$$

Since by Ito differential rule [38]

$$d(\exp(-h \cdot y_t) \Phi(t)) = \exp(-h \cdot y_t) \{ d\Phi(t) - \Phi(t) (h \cdot dy_t) - \frac{1}{2} |h|^2 \Phi(t) dt \},$$

it follows from equations (3.42), (3.43) and (3.44) that the process

$$V(t) \equiv (V_1(t), V_2(t), \dots, V_N(t))' ; t \geq 0,$$

satisfies the following ordinary partial differential equation

$$\begin{aligned} \frac{d}{dt}V(t) &= \tilde{\Gamma}^*(t)V(t), \quad t \in [0, T], \\ V(0) &= \Phi_0. \end{aligned} \quad (3.45)$$

For any $u, v \in H^1$, equation (3.43) can be written as

$$\begin{aligned} \sum_{m, \ell=1}^N \langle \tilde{\Gamma}_{m\ell}^* u_\ell, v_m \rangle &= \frac{1}{2} y_t \cdot \sum_{m=1}^N \sum_{i, j=1}^n \int_{R^n} \sigma_{ij}^m \frac{\partial h}{\partial x_i} \frac{\partial u_m}{\partial x_j} v_m dx \\ &\quad - \frac{1}{2} \sum_{m=1}^N \sum_{i, j=1}^n \int_{R^n} \sigma_{ij}^m \frac{\partial u_m}{\partial x_i} \frac{\partial v_m}{\partial x_j} dx \\ &\quad + \sum_{m=1}^N \sum_{i=1}^n \int_{R^n} \tilde{a}_i^m \frac{\partial v_m}{\partial x_i} u_m dx \\ &\quad + \sum_{m=1}^N \sum_{i=1}^n \int_{R^n} \check{a}_i^m u_m v_m dx, \end{aligned} \quad (3.46)$$

where

$$\tilde{a}_i^m \equiv a_i^m - \frac{1}{2} \sum_{j=1}^n \frac{\partial \sigma_{ij}^m}{\partial x_j} - \frac{1}{2} y_t \cdot \sum_{j=1}^n \sigma_{ij}^m \frac{\partial h}{\partial x_j}, \quad (3.47)$$

and

$$\begin{aligned} \check{a}_i^m &\equiv \left(\frac{1}{2} \sum_{j=1}^n \frac{\partial \sigma_{ij}^m}{\partial x_j} - a_i^m \right) \left(\frac{\partial h}{\partial x_i} \cdot y_t \right) + \frac{1}{2} \sum_{j=1}^n \sigma_{ij}^m \left(\frac{\partial h}{\partial x_i} \cdot y_t \right) \left(\frac{\partial h}{\partial x_j} \cdot y_t \right) \\ &\quad - \frac{1}{2} |h|^2 + \sum_{\ell=1}^N \Lambda_{m\ell}^*. \end{aligned} \quad (3.48)$$

Again under the assumptions (A1), (A3), $h \in C_b^2(R^n)$ and

$$\Pr\left\{ \sup_{t \in [0, T]} |y_t| < \infty \right\} = 1,$$

the operator $\tilde{\Gamma}^*$ satisfies the properties (Pi) and (Pii). Hence by Lemma 2.4, the initial value problem of equation (3.45) has a unique (weak) solution $V(t) \in L_2((0, T); H^1) \cap C([0, T]; (L_2(R^n))^N)$.

3.6 Examples

Example 3.1

Suppose a firm has m -machines which are used for production of homogeneous goods. The rate of production depends on the number of machines in operation. Define $\Sigma \equiv \{0, 1\}^m$ and let the process $\xi_t; t \geq 0$, be a Markov chain with values in Σ , where '0' denotes the failed state and '1' the operating state. The flow of the inventory and total sales are assumed to be governed by the following set of stochastic differential equations

$$\begin{aligned} dI_t &= (a(\xi_t) - h(I_t, \alpha))dt + \sigma(t)dW(t), \\ dS(t) &= h(I_t, \alpha)dt + dV(t), \end{aligned} \tag{3.49}$$

where 'a' denotes the production rate depending on the state of the process ξ (more machines means more production and conversely), 'h' represents the instantaneous sales rate which is dependent on the current inventory of goods and their unit price α which is assumed to be fixed. The martingale terms σdW and dV represent the spoilage of inventory and sales return, respectively. The firm is interesting in estimating the cost of holding inventories $C(I_t)$ given $\sigma(S_0^t)$, where $\sigma(S_0^t)$ denotes the σ -algebra generated by the process S up to time t . Using equation (3.23) it follows that the estimated cost of holding inventories \widehat{C}_t is given by

$$\widehat{C}_t \equiv E\{C(I_t)|\sigma(S_0^t)\} = \sum_{k=1}^N \Theta_k(t) \frac{\langle C(\cdot), \varphi_k(t, \cdot) \rangle}{\langle 1, \varphi_k(t, \cdot) \rangle},$$

where $\{\varphi_k; 1 \leq k \leq N\}$ is the solution of (3.24), $N = 2^m$, and $\Theta_k, 1 \leq k \leq N; t \geq 0$, is given by (3.23.a).

Example 3.2

Let $\{a_1, b_1\}$ and $\{a_2, b_2\}$ be two pair of functions satisfying our basic assumptions (A1),(A2), and (A3) of section 3.2. Let τ be a non-negative random variable with density ν satisfying $\nu(\tau \in dt) = \lambda \exp(-\lambda t)dt, \lambda > 0$. Let $\{\Gamma_t, t \geq 0\}$ be an increasing family of σ -algebras generated by the indicator function of the set $\{t < \tau\}$. Clearly, $\Gamma_t = \sigma(1_{\{t \wedge \tau\}})$ and it is right continuous and τ is totally inaccessible stopping time.

Consider the system

$$dX_t = 1_{\{t < \tau\}} a_1(X_t)dt + 1_{\{t < \tau\}} b_1(X_t)dW_t \\ + 1_{\{t \geq \tau\}} a_2(X_t)dt + 1_{\{t \geq \tau\}} b_2(X_t)dW_t,$$

where

$$1_{\{t < \tau\}} \equiv \begin{cases} 1 & \text{if } t < \tau, \\ 0 & \text{if } t \geq \tau. \end{cases}$$

Let the observed process $\{y_s; s \leq t\}$ be governed by the following stochastic differential equation

$$dy_t = h(X_t)dt + dV_t, \quad y \in R^k.$$

Again our problem is to determine the conditional expectation $E\{f(X_t) | \sigma(y_0^t)\}$.

Using the properties (P1) and (P2) of the infinitesimal transition rates $\{\lambda_{ij}; i, j = 1, 2, \dots, N\}$ and setting $N = 2$, it is clear that $\lambda_{11} = -\lambda_{12} = -\lambda$ and $\lambda_{21} = \lambda_{22} = 0$. Hence it follows from equation (3.24) that the unnormalized densities $\{\varphi_k; k = 1, 2\}$ are governed by the following stochastic partial differential equations

$$d\varphi_1(t, x) = \{(A^* \varphi)_1(t, x) - \lambda \varphi_1(t, x)\}dt + h(x)\varphi_1(t, x) \cdot dy_t, \\ \varphi_1(0, x) = p_0(x), \tag{3.50}$$

and

$$\begin{aligned} d\varphi_2(t, x) &= \{(A^*\varphi)_2(t, x) + \lambda\varphi_1(t, x)\}dt + h(x)\varphi_2(t, x) \cdot dy_t, \\ \varphi_2(0, x) &= 0. \end{aligned} \tag{3.51}$$

Hence using (3.23) it follows that

$$E\{f(X_i) | \sigma(y_0^t)\} = \sum_{k=1}^2 \Theta_k(t) \frac{\langle \varphi_k(t, \cdot), f(\cdot) \rangle}{\langle \varphi_k(t, \cdot), 1 \rangle},$$

where $\{\varphi_k; k = 1, 2\}$ are the solutions of (3.50) and (3.51), respectively, and $\Theta_k; k = 1, 2$ is given by (3.23.a).

The above example can be interpreted to present a situation where a machine wears out to a state of lower productivity (or complete retirement) with no possible return. In fact this example is a finite dimensional version of that given in [5] for control problem.

3.7 Summary and Conclusion

In this chapter we have considered the filtering problem of stochastic processes governed by a class of nonlinear stochastic differential equations with drift and diffusion parameters perturbed by a temporally homogeneous Markov chain. We have assumed that the observation process (continuous) is governed by Ito differential equation with drift parameter dependent only on the state of the system. Using Girsanov transformation, we have derived a coupled system of stochastic integral equations for the unnormalized density from which the filter equations were obtained. Indeed, as indicated in section 3.4 (case (ii)), the filter equation (3.24) covers that of Zakai. Further, we have also noticed that when no measurement of the Markov chain ξ is possible, the estimate \hat{f}_t of $f(x_t)$ is based on the a priori statistical information of ξ and the observed process $\{y_s; s \leq t\}$. On the other hand, when ξ is observable, the estimate \hat{f}_t can be computed by solving Zakai equation (3.32). Moreover, we have also shown that the filter equation (3.24) has a unique (weak) solution. Finally, two examples were presented to illustrate some of the results of this chapter.

CHAPTER 4

NONLINEAR FILTERING OF DISCONTINUOUS PROCESSES WITH DISCONTINUOUS OBSERVATIONS

4.1 Introduction

In chapter 3 we have considered the filtering problem for a class of systems governed by nonlinear (continuous) Ito differential equations with jump parameters and continuous observations. In this chapter, we are interested in the filtering problem of a wider class of stochastic processes where the state as well as the observation processes are governed by a nonlinear stochastic differential equations driven by Wiener processes and generalized Poisson random counting measures. Utilizing the results obtained in [68] for the Radon-Nikodym derivatives and generalized Ito's Lemma (see [37]), we use similar arguments as those of chapter 2 (see sections 2.3) to derive the corresponding filter equations.

In section 4.2 we formulate the filtering problem and introduce the necessary notations and assumptions which are used in the sequel. In section 4.3 we present some known results due to Skorokhod and Ito on the basis of which we derive a stochastic integral equation for the unnormalized conditional density. Using the results of section 4.3 and the properties of the semigroup of the state process, we derive in section 4.4 a stochastic differential equation for the unnormalized conditional density. In section 4.5, we use the results of section 4.4 to obtain the filter equation for the case where the state process is governed by nonlinear Ito differential equation driven by Wiener and generalized Poisson processes whereas the observed process (discrete) is excited by zero mean white Gaussian noise. Finally, in section 4.6 we present few examples to illustrate some of the results of this chapter.

4.2 Formulation Of Filtering Problem, Assumptions and Notations

We consider the filtering problem for a class of systems governed by stochastic differential equation of the form

$$\begin{aligned} dx_t &= a(x_t)dt + b(x_t)dW_t + \int_{R_0^r} g(x_t, v)\tilde{\gamma}(dv \times dt) \\ x(0) &= x_0, \end{aligned} \quad (4.1)$$

where x, a and g are vectors in R^n ; b is an $(n \times n)$ matrix and $R_0^r \equiv R^r \setminus \{0\}$. The measure $\tilde{\gamma}(t, \Gamma), t \in [0, T], \Gamma \in \mathcal{B}(R_0^r)$ (where $\mathcal{B}(R_0^r)$ denotes the σ field generated by Borel subsets of R_0^r) is a random counting measure giving the number of jumps, of sizes lying in Γ , of the process $\{x_\theta; \theta \leq t\}$ over the interval $[0, t]$. We assume that $\tilde{\gamma}$ is independent of the initial state x_0 and the Wiener process W and it obeys generalized Poisson distribution with mean given by

$$E\{\tilde{\gamma}(\Gamma, I)\} = \tilde{\Pi}(\Gamma)\ell(I), \quad (4.2)$$

where $(\Gamma \times I)$ is any Borel subset of $R_0^r \times [0, T]$ and $\ell(I)$ denotes the Lebesgue measure of the set $I \subset [0, T]$. In other words, for each nonnegative integer m , and $\Gamma \in \mathcal{B}(R_0^r)$ the probability law of the measure $\tilde{\gamma}$, is given by

$$Pr\{\tilde{\gamma}(\Gamma, I) = m\} = \frac{\exp\{-\ell(I)\tilde{\Pi}(\Gamma)\}(\ell(I)\tilde{\Pi}(\Gamma))^m}{m!} \quad (4.3)$$

The process W is an n -dimensional standard Wiener process independent of $\tilde{\gamma}$ and the initial state x_0 . Further, we assume that the output (observed) process $\{y_s; s \leq t\}$ is governed by the following stochastic differential equation

$$\begin{aligned} dy_t &= h(x_t)dt + dV_t + \int_{R_0^m} g_1(x_t, y_t, v)\gamma(dv \times dt) \\ y(0) &= 0, \end{aligned} \quad (4.4)$$

where h and g_1 are vectors in R^k ; V is k -dimensional standard Wiener process independent of W and $\tilde{\gamma}$, and $R_0^m \equiv R^m \setminus \{0\}$. Again the measure γ is a random

counting measure on the Borel subset of $R_0^m \times [0, T]$ independent of the processes $W, \tilde{\gamma}, V$ and the initial state x_0 . We assume also that the measure γ obeys generalized Poisson distribution with mean

$$E\{\gamma(A, I)\} = \Pi(A)\ell(I), \quad (4.5)$$

where $(A \times I)$ is any Borel subset of $R_0^m \times [0, T]$, and $\ell(I)$ denotes the Lebesgue measure of the interval $I \in [0, T]$.

Define

$$q(A, I) \equiv \gamma(A, I) - \Pi(A)\ell(I). \quad (4.6)$$

Then it is easy to verify (see [37] and [68]) that the measure q is a square integrable martingale (with respect to the σ -field generated by the measure γ denoted \mathcal{F}_t^γ) satisfying the following properties

$$E\{q(A, I)\} = 0,$$

$$E(q(I, A))^2 = \ell(I)\Pi(A),$$

for any Borel $(A \times I) \subset R_0^m \times [0, T]$.

Assuming that all the random processes and vectors described above are defined on a complete probability space $(\Omega_0, \mathcal{B}_0, P_0)$, our problem is to compute the conditional expectation

$$\hat{f}_t \equiv E\{f(x_t) | \mathcal{F}_t^y\},$$

where f is any bounded measurable function on R^n and \mathcal{F}_t^y denotes the σ -field generated by the process y up to time t .

Remark 4.1

As indicated earlier in chapter 1, the system (4.1) and (4.2) is indeed a special class of that considered in [40] for the filtering problem. In fact the result obtained in [40] is based on the fact that the random measure q (see equation (4.6)) is observable. However, for the case where the observed process y is given by

$$y(t_k) = h(x(t_k)) + V(t_k) + \int_{R_0^n} g_1(x(t_k), y(t_k), v) \gamma(dv \times (t_{k-1}, t_k]),$$

where $0 = t_0 < t_1 \cdots t_n = T$, observing q is almost impossible. Hence our main goal in this chapter is to develop a filter equation, for system (4.1) and (4.4), which is only driven by the observed process y and the compensator Π corresponding to the measure γ (see equation (4.6)). In fact our filter equation is also valid even when the observed process y is of discrete type as given above.

For the derivation of the filter equation, we need the following notations and assumptions.

Notations:

Let $\mathcal{D}_1([0, T]; R^{n+k})$ denote the space of right continuous R^{n+k} -valued functions having left limits and suppose this is furnished with the Skorokhod topology [68]. Let $\mathcal{B}(\mathcal{D}_1)$ denote the σ -field generated by the Borel subsets of \mathcal{D}_1 . We denote this measurable space by $(\mathcal{D}_1, \mathcal{B}(\mathcal{D}_1))$. Let $\{\eta_t; t \geq 0\}$ be any random process and let $\sigma\{\eta_s; s \leq t\}$ denote the σ -field generated by the process η up to time t . Let $\mathcal{B}(R^n)$ denote the Borel σ -field generated by subsets of R^n and $\mathcal{F}_t^x \equiv \sigma\{x_s; s \leq t\}$, $\mathcal{F}_t^y \equiv \sigma\{y_s; s \leq t\}$, $\mathcal{F}_t^\gamma \equiv \sigma\{\gamma((\theta, s], A), A \in \mathcal{B}(R_0^m); I \equiv [0, T]; 0 \leq \theta \leq s \leq t\}$ and $\mathcal{G}_t \equiv \mathcal{F}_t^W \vee \mathcal{F}_t^{\tilde{\gamma}} \vee \mathcal{F}_t^V \vee \mathcal{F}_t^\gamma \vee \sigma\{x_0\}$. The σ -fields \mathcal{F}_t^W , $\mathcal{F}_t^{\tilde{\gamma}}$, \mathcal{F}_t^V , \mathcal{F}_t^γ and $\sigma\{x_0\}$ are assumed to be independent. For any vectors $a, b \in R^n$ we denote by $(a \cdot b)$ the scalar product in R^n . Further notations will be introduced in the sequel as required.

Assumptions:

- (A1) The functions a and b are continuous and satisfy uniform Lipschitz and growth conditions on R^n . That is, there exists a constant $k > 0$ such that

$$T|a(x) - a(y)|^2 + \|b(x) - b(y)\|^2 \leq k|x - y|^2,$$

and

$$T|a(x)|^2 + \|b(x)\|^2 \leq k(1 + |x|^2),$$

for all $x, y \in R^n$.

- (A2) The functions $a_i(x), \sigma_{ij}(x), \frac{\partial}{\partial x_i} a_i(x), \frac{\partial}{\partial x_i} \sigma_{ij}(x)$ and $\frac{\partial^2}{\partial x_i \partial x_j} \sigma_{ij}$ where $\sigma_{ij} \equiv (bb')_{ij}(x); i, j = 1, 2, \dots, n$ are bounded and satisfy Holder condition on R^n .

- (A3) The function g satisfies uniform Lipschitz and growth conditions on R^n and

$$\int_{R_0^n} |g(x, v)|^2 \tilde{\Pi}(dv) < \infty.$$

- (A4) The function h is continuous and satisfies uniform Lipschitz and growth conditions on R^n and

$$E \int_0^t |h(x_\theta)|^2 d\theta < \infty,$$

along any solution $x_t; t \geq 0$, of equation (4.1).

- (A5) The function g_1 satisfies uniform Lipschitz and growth conditions on R^{n+k} and

$$\int_{R_0^{n+k}} |g_1(x, u)|^2 \Pi(du) < \infty,$$

for all $x \in R^{n+k}$. Note that under the above assumptions it is known that (see [68] Theorem 1 P.47) equations (4.1) and (4.4) have unique strong solutions with finite second moments.

In the next section we present (without proof) some known results due to Skorokhod [68] on the basis of which we derive a stochastic integral equation for the unnormalized density leading to the filter equation.

4.3 Stochastic Integral Equation For Unnormalized Density

In this section we use the Radon-Nikodym derivative of measures induced on the space $(\mathcal{D}_1, \mathcal{B}(\mathcal{D}_1))$ by the systems (4.1) and (4.4) and generalized Ito's Lemma, to derive a stochastic integral equation for the unnormalized density. First we present the following known results.

Lemma 4.1 (Skorokhod [68] section 4 P.117)

Suppose that assumptions (A1)-(A5) hold and there exist functions $g_2 : R^k \times R_0^m \rightarrow R^k$ and $\varrho : R^n \times R^k \times R_0^m \rightarrow (0, \infty)$ satisfying the following properties :

(i) There exists a constant $K > 0$ such that

$$\int_{R_0^m} |g_2(x, u) - g_2(y, u)|^2 \Pi(du) \leq K|x - y|^2,$$

for all $x, y \in R^k$ and

$$\int_{R_0^m} |g_2(y, u)|^2 \Pi(du) < \infty, \quad y \in R^k.$$

(ii) For all $x \in R^n$ and $y \in R^k$,

$$\int_{R_0^m} |\beta(x, y, u)|^2 \Pi(du) < \infty,$$

with β being any of the functions $\varrho, \ln \varrho, \tilde{\varrho} \equiv \frac{1}{\varrho}$ and $\ln \tilde{\varrho}$.

(iii) For any $A \in \mathcal{B}(R_0^m)$ and for all $(x, y) \in R^{n+k}$,

$$\int_{\{u: g_2(y, u) \in A\}} \Pi(du) = \int_{\{u: g_1(x, y, u) \in A\}} \varrho(x, y, g_1(x, y, u)) \Pi(du).$$

Then the measures μ_1 and μ_2 defined on $(\mathcal{D}_1, \mathcal{B}(\mathcal{D}_1))$ and induced, respectively, by the systems

$$dx_t = a(x_t)dt + b(x_t)dW_t + \int_{R_0^m} g(x_t, v)\tilde{\gamma}(dv, dt),$$

$$dy_t = h(x_t)dt + dV_t + \int_{R_0^m} g_1(x_t, y_t, u)\gamma(du, dt),$$

and

$$\begin{aligned} dx_t &= a(x_t)dt + b(x_t)dW_t + \int_{R_0^n} g(x_t, v)\tilde{\gamma}(dv, dt), \\ dy_t &= dV_t + \int_{R_0^m} g_2(y_t, u)\gamma(du, dt), \end{aligned} \quad (4.7)$$

are absolutely continuous with respect to one another. Further, the Radon-Nikodym derivative of μ_1 with respect to μ_2 is given by

$$\begin{aligned} \frac{d\mu_1}{d\mu_2} \equiv \Lambda_t &= \exp\left\{ \int_0^t \Psi(x_s, y_s) ds - \int_0^t \alpha(x_s, y_s) \cdot dV_s \right. \\ &\quad \left. + \int_0^t \int_{R_0^m} \ln \tilde{q}(x_s, y_s, u) \tilde{q}(du, ds) \right\}, \end{aligned} \quad (4.8)$$

where \cdot denotes the scalar product in R^k ,

$$\Psi(x, y) \equiv \int_{R_0^m} [1 - \tilde{q}(x, y, u) + \ln \tilde{q}(x, y, u)] \Pi(du) - \frac{1}{2} |\alpha|^2, \quad (4.9)$$

$$\alpha(x, y) \equiv \int_{R_0^m} [g_2(y, u) - g_1(x, y, u)] \Pi(du) - h(x), \quad (4.10)$$

for all $x \in R^n, y \in R^k$, and

$$q(du, dt) \equiv \gamma(du, dt) - \Pi(du)dt. \quad (4.11)$$

Lemma 4.2 (Generalized Ito's Formula [37] Theorem 11 P.104)

Let the R^n -valued process $\{\xi(t); t \geq 0\}$ be the solution of the following stochastic differential equation

$$d\xi(t) = a(\xi(t))dt + b(\xi(t))dW_t + \int_{R_0^m} g(\xi(t), u)q(du, dt), \quad (4.11.a)$$

with $\xi(0) = \xi_0$, where W is an n -dimensional standard Wiener process independent of the measure q given by (4.11) and the function g satisfies the following property

$$\int_{R_0^m} |g(x, u)|^2 \Pi(du) < \infty, \quad x \in R^n, P_0 - a.s.$$

Let f be twice continuously differentiable function on R^n satisfying

$$\int_{R_0^m} |f(\xi(t) + g(\xi(t), u)) - f(\xi(t))|^2 \Pi(du) < \infty. \quad P_0 - a.s.$$

Then the Ito differential of the function f is given by

$$\begin{aligned}
 df(\xi(t)) = & \left\{ \sum_{i=1}^n a_i(\xi(t)) \frac{\partial f}{\partial x_i}(\xi(t)) + \frac{1}{2} \sum_{i,j=1}^n \sigma_{ij}(\xi(t)) \frac{\partial^2 f}{\partial x_i \partial x_j}(\xi(t)) \right. \\
 & + \int_{R_0^n} [f(\xi(t) + g(\xi(t), u)) - f(\xi(t)) - \sum_{i=1}^n g_i(\xi(t), u) \frac{\partial f}{\partial x_i}(\xi(t))] \times \\
 & \times \Pi(du) \Big\} dt + \sum_{i,j=1}^n \frac{\partial f}{\partial x_i}(\xi(t)) b_{ij}(\xi(t)) dW_t^j \\
 & + \int_{R_0^n} [f(\xi(t) + g(\xi(t), u)) - f(\xi(t))] q(du, dt), \quad (4.12)
 \end{aligned}$$

where $\sigma_{ij} \equiv (bb')_{ij}$.

Remark 4.2

Another convenient way for writing the Ito differential of the function f is

$$df(\xi(t)) = (Af)(\xi(t))dt + b \frac{\partial f}{\partial x} \cdot dW_t + \int_{R_0^n} \{f(\xi(t) + g(\xi(t), u)) - f(\xi(t))\} q(du, dt),$$

where A denotes the (backward) Kolmogorov operator of the process $\{\xi(t); t \geq 0\}$ given by [68]

$$\begin{aligned}
 (Af)(x) \equiv & \sum_{i=1}^n a_i(x) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n \sigma_{ij}(x) \frac{\partial^2 f}{\partial x_i \partial x_j} + \int_{R_0^n} \{f(x + g(x, u)) \\
 & - f(x) - \sum_{i=1}^n g_i(x, u) \frac{\partial f}{\partial x_i}\} \Pi(du). \quad (4.12.a)
 \end{aligned}$$

Utilizing the results of Lemmas 4.1 and 4.2, we follow similar procedure as that of chapter 2 to derive an integral equation for the unnormalized density. Let f be a bounded measurable function on R^n , then defining

$$\begin{aligned}
 \tilde{\Lambda}(t, x_t) & \equiv E_2\{\Delta_t | \mathcal{F}_t^y \vee \sigma(x_t)\}, \\
 Q(t, dx) & \equiv Pr\{x_t \in dx\}, \quad t \geq 0, \quad (4.13)
 \end{aligned}$$

and using similar arguments as those of section 2.3, we have

$$E_1\{f(x_t)|\mathcal{F}_t^y\} = \frac{\int_{R^n} \tilde{\Lambda}(t, x) f(x) Q(t, dx)}{\int_{R^n} \tilde{\Lambda}(t, x) Q(t, dx)} \quad (4.13.a)$$

Taking the indicator function 1_Γ for $f, \Gamma \in \mathcal{B}(R^n)$, it follows from (4.13.a) that

$$P(t, \Gamma | \mathcal{F}_t^y) \equiv Pr\{x_t \in \Gamma | \mathcal{F}_t^y\} = \frac{\int_\Gamma \tilde{\Lambda}(t, x) Q(t, dx)}{\int_{R^n} \tilde{\Lambda}(t, x) Q(t, dx)} \quad (4.14)$$

With the help of the above Lemmas and the transition probability of the state process $x_t, t \geq 0$, we show in the following Theorem that the process $\tilde{\Lambda}(t, \cdot), t \geq 0$, is governed by two stochastic integral equations.

Theorem 4.1 (Stochastic Integral Equations For $\tilde{\Lambda}$)

Suppose that the hypotheses of Lemma 4.1 hold and there exists a bounded measurable function $c : R^k \times R^n \rightarrow R^k$ such that

$$\int_0^t c(y_s, x_s) \cdot dM_s = \int_0^t \left\{ \int_{R_0^n} [\tilde{\varrho}(x_s, y_s, u) - 1] q(du, ds) - \alpha(x_s, y_s) \cdot dV_s \right\}, \quad (4.15)$$

μ_2 -a.s, where M_t is a square integrable \mathcal{G}_t -martingale, with respect to the measure μ_2 , given by

$$M_t \equiv V_t + \int_0^t \int_{R_0^n} g_2(y_s, u) q(du, ds). \quad (4.16)$$

Then the process $\tilde{\Lambda}(t, x_t); t \geq 0$, satisfies μ_2 -a.s, the following stochastic integral equations

$$\begin{aligned} \tilde{\Lambda}(t, x) = & 1 - \int_0^t ds \int_{R_0^n} \Pi(du) \int_{R^n} \tilde{\Lambda}(s, z) g_2(y_s, u) \cdot c(y_s, z) P_{t,s}(x; dz) \\ & + \int_0^t dy_s \cdot \int_{R^n} \tilde{\Lambda}(s, z) c(y_s, z) P_{t,s}(x; dz), \end{aligned} \quad (4.17)$$

and

$$\begin{aligned} \tilde{\Lambda}(t, x) &= \int_{R^n} \tilde{\Lambda}(s, z) P_{t,s}(x; dz) \\ &\quad - \int_s^t d\theta \int_{R_0^m} \Pi(du) \int_{R^n} \tilde{\Lambda}(\theta, z) g_2(y_\theta, u) \cdot c(y_\theta, z) P_{t,\theta}(x; dz) \\ &\quad + \int_s^t dy_\theta \cdot \int_{R^n} \tilde{\Lambda}(\theta, z) c(y_\theta, z) P_{t,\theta}(x; dz), \end{aligned} \quad (4.17.a)$$

where

$$P_{t,s}(x; \Gamma) \equiv Pr\{x_s \in \Gamma | x_t = x\}, \quad s \leq t, \Gamma \in \mathcal{B}(R^n). \quad (4.18)$$

Proof:

Defining

$$\beta(t) \equiv \int_0^t \Psi(x_s, y_s) ds - \int_0^t \alpha(x_s, y_s) \cdot dW_s + \int_0^t \int_{R_0^m} \ln \tilde{q}(x_s, y_s, u) q(du, ds),$$

it follows from equation (4.8) that

$$\Lambda_t \equiv \exp \{\beta(t)\}. \quad (4.19)$$

Utilizing the result of Lemma 4.2, the Ito differential of the process Λ_t is given by

$$\begin{aligned} d \exp \{\beta(t)\} &\equiv d\Lambda_t = \{\Psi(x_t, y_t) \exp(\beta(t)) + \frac{1}{2} |\alpha(x_t, y_t)|^2 \exp(\beta(t)) \\ &\quad + \int_{R_0^m} [\exp(\beta(t) + \ln \tilde{q}(x_t, y_t, u)) - \exp(\beta(t)) \\ &\quad - \ln \tilde{q}(x_t, y_t, u) \exp(\beta(t))] \Pi(du)\} dt - \exp(\beta(t)) \alpha(x_t, y_t) \cdot dV_t \\ &\quad + \int_{R_0^m} [\exp(\beta(t) + \ln \tilde{q}(x_t, y_t, u)) - \exp(\beta(t))] q(du, dt). \end{aligned} \quad (4.20)$$

Using equations (4.8), (4.9) and (4.19), it follows from (4.20) that

$$\begin{aligned} d\Lambda_t &= \Lambda_t \left\{ \int_{R_0^m} (\tilde{q}(x_t, y_t, u) - 1) q(du, dt) - \alpha(x_t, y_t) \cdot dV_t \right\}, \\ \Lambda_0 &= 1, \end{aligned} \quad (4.21)$$

from which one obtains

$$\begin{aligned} \Lambda_t = 1 + \int_0^t \int_{R_0^m} \Lambda_s \{ \tilde{q}(x_s, y_s, u) - 1 \} q(du, ds) \\ - \int_0^t \Lambda_s \alpha(x_s, y_s) \cdot dV_s, \quad \mu_2 - a.s. \end{aligned} \quad (4.22)$$

Since under the measure μ_2 the process $\{y_t; t \geq 0\}$ satisfies

$$dy_t = dM_t + \int_{R_0^m} g_2(y_t, u) \Pi(du) dt, \quad (4.23)$$

where M_t is given by (4.16), it follows from (4.15) and (4.22) that

$$\begin{aligned} \Lambda_t = 1 - \int_0^t \int_{R_0^m} \Lambda_s g_2(y_s, u) \cdot c(y_s, x_s) \Pi(du) ds \\ + \int_0^t \Lambda_s c(y_s, x_s) \cdot dy_s. \end{aligned} \quad (4.24)$$

Taking the conditional expectation, with respect to μ_2 , for both sides of (4.24) we have

$$\begin{aligned} E_2 \{ \Lambda_t | \mathcal{F}_t^y \vee \sigma(x_t) \} \equiv \tilde{\Lambda}(t, x_t) = 1 - \int_0^t \int_{R_0^m} E_2 \{ \Lambda_s \times \\ \times g_2(y_s, u) \cdot c(y_s, x_s) | \mathcal{F}_t^y \vee \sigma(x_t) \} \Pi(du) ds \\ + \int_0^t E_2 \{ \Lambda_s c(y_s, x_s) | \mathcal{F}_t^y \vee \sigma(x_t) \} \cdot dy_s. \end{aligned} \quad (4.25)$$

Since under the measure μ_2 the σ -fields \mathcal{F}_t^x and \mathcal{F}_t^y are independent it follows that the σ -fields $\mathcal{F}_s^x \vee \mathcal{F}_s^y$ and $\sigma\{y_\theta - y_s; s \leq \theta \leq t\}$ are independent given $\mathcal{F}_s^y \vee \sigma(x_t)$.

Therefore,

$$\begin{aligned} E_2 \{ \Lambda_s g_2(y_s, u) \cdot c(y_s, x_s) | \mathcal{F}_t^y \vee \sigma(x_t) \} &= E_2 \{ \Lambda_s g_2(y_s, u) \cdot c(y_s, x_s) | \mathcal{F}_s^y \vee \sigma(x_t) \} \\ &= E_2 \{ E_2 \{ \Lambda_s g_2(y_s, u) \cdot c(y_s, x_s) | \mathcal{F}_s^y \vee \sigma(x_s) \vee \sigma(x_t) \} | \mathcal{F}_s^y \vee \sigma(x_t) \} \\ &= E_2 \{ g_2(y_s, u) \cdot c(y_s, x_s) E_2 \{ \Lambda_s | \mathcal{F}_s^y \vee \sigma(x_s) \vee \sigma(x_t) \} | \mathcal{F}_s^y \vee \sigma(x_t) \}, \end{aligned} \quad (4.26)$$

and similarly ,

$$E_2\{\Lambda_s c(y_s, x_s) | \mathcal{F}_t^y \vee \sigma(x_t)\} = E_2\{c(y_s, x_s) E_2\{\Lambda_s | \mathcal{F}_s^y \vee \sigma(x_s) \vee \sigma(x_t)\} | \mathcal{F}_s^y \vee \sigma(x_t)\}. \quad (4.26.a)$$

Since the process $\{(x_s, y_s); s \leq t\}$, is \mathcal{G}_t -Markov process and \mathcal{F}_t^x and \mathcal{F}_t^y are independent under μ_2 , it follows that the σ -fields $\mathcal{F}_s^x \vee \mathcal{F}_s^y$ and $\sigma(x_t)$ are conditionally independent given $\sigma(x_s)$. Hence

$$E_2\{\Lambda_s | \mathcal{F}_s^y \vee \sigma(x_s) \vee \sigma(x_t)\} = E_2\{\Lambda_s | \mathcal{F}_s^y \vee \sigma(x_s)\} \equiv \tilde{\Lambda}(s, x_s). \quad (4.27)$$

By the independence of the σ -fields \mathcal{F}_t^x and \mathcal{F}_t^y , under μ_2 , it follows from Fubini's Theorem and equation (4.27) that the right hand sides of equations (4.26) and (4.26.a) can be written, respectively, as

$$\int_{R^n} g_2(y_s, u) \cdot c(y_s, z) \tilde{\Lambda}(s, z) P_{t,s}(x; dz), \quad (4.28)$$

and

$$\int_{R^n} c(y_s, z) \tilde{\Lambda}(s, z) P_{t,s}(x; dz), \quad (4.28.a)$$

where the transition probability $P_{t,s}(x; dz)$ is given by equation (4.18). From equations (4.25)-(4.28.a), equation (4.17) follows. It remains only to prove equation (4.17.a).

Using equation (4.23) one obtains

$$\begin{aligned} \Lambda_t &= \Lambda_s - \int_s^t \int_{R_0^n} \Lambda_\theta g_2(y_\theta, u) \cdot c(y_\theta, x_\theta) \Pi(du) d\theta \\ &\quad + \int_s^t \Lambda_\theta c(y_\theta, x_\theta) \cdot dy_\theta. \end{aligned} \quad (4.29)$$

Thus,

$$\begin{aligned} E_2\{\Lambda_t | \mathcal{F}_t^y \vee \sigma(x_t)\} &= E_2\{\Lambda_s | \mathcal{F}_t^y \vee \sigma(x_t)\} - \int_s^t \int_{R_0^n} E_2\{g_2(y_\theta, u) \cdot c(y_\theta, x_\theta) \times \\ &\quad \times \tilde{\Lambda}(\theta, x_\theta) | \mathcal{F}_t^y \vee \sigma(x_t)\} \Pi(du) d\theta \\ &\quad + \int_s^t E_2\{c(y_\theta, x_\theta) \times \\ &\quad \times \tilde{\Lambda}(\theta, x_\theta) | \mathcal{F}_t^y \vee \sigma(x_t)\} \cdot dy_\theta. \end{aligned} \quad (4.30)$$

Again by similar arguments as those given above, it is easy to verify that equation (4.17.a) follows from (4.30). This completes the proof. ■

On the basis of the above Theorem we now follow similar procedure as in section 2.3 to derive stochastic integral equations for the unnormalized density. For this we need the semigroup generated by the process $\{x_t; \mathcal{G}_t, t \geq 0\}$. Let f be any bounded continuous function on R^n and let $\{T_{s,t}; s \leq t\}$ denote the (backward) semigroup given by

$$(T_{s,t}f)(\eta) \equiv E\{f(x_t)|x_s = \eta\}, \quad s \leq t.$$

Let

$$P_{s,t}(\eta; \Gamma) \equiv Pr\{x_t \in \Gamma | x_s = \eta\}, \quad s \leq t,$$

for $\eta \in R^n, \Gamma \in \mathcal{B}(R^n)$, be the kernel corresponding to the semigroup $T_{s,t}$ so that

$$(T_{s,t}f)(\eta) = \int_{R^n} f(\gamma) P_{s,t}(\eta; d\gamma), \quad (4.31)$$

and

$$S. \lim_{t \rightarrow s} (T_{s,t}f)(\eta) = f(\eta),$$

for all $f \in C_b(R^n)$, where $C_b(R^n)$ denotes the space of bounded continuous functions on R^n . It is clear from Lemma 4.2 that the infinitesimal generator corresponding to the semigroup $\{T_{s,t}; s \leq t\}$ is given by

$$\begin{aligned} (Af)(x) \equiv & \sum_{i=1}^n a_i(x) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n \sigma_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j} + \int_{R_0^n} \{f(x+g(x,u)) \\ & - f(x) - \sum_{i=1}^n g_i(x,u) \frac{\partial f}{\partial x_i}\} \tilde{\Pi}(du), \end{aligned} \quad (4.31.a)$$

where $(bb')_{ij}(x) \equiv \sigma_{ij}(x)$ and g_i denotes the i -th component of g .

Suppose for each $\eta \in R^n$ and $0 \leq s \leq t$, the measure $\Gamma \rightarrow P_{s,t}(\eta; \Gamma)$ is absolutely continuous with respect to Lebesgue measure on R^n admitting a density so that

$$P_{s,t}(\eta; \Gamma) = \int_{\Gamma} \tilde{p}_{s,t}(\eta; \gamma) d\gamma, \quad \Gamma \in \mathcal{B}(R^n).$$

For each $t \geq 0$, let ν_t denote the measure induced by the process $\{x_t; t \geq 0\}$ on $\mathcal{B}(R^n)$ and suppose that $E \rightarrow \nu_t(E), E \in \mathcal{B}(R^n)$, is absolutely continuous with respect to Lebesgue measure admitting a density $\beta_t(\eta), \eta \in R^n$, so that

$$\nu_t(E) = \int_E \beta_t(\eta) d\eta.$$

Define

$$(T_{s,t}^* \beta)(x) \equiv \int_{R^n} p_{s,t}(\eta; x) \beta(\eta) d\eta, \quad (4.32)$$

where $T_{s,t}^*, s \leq t$, denotes the semigroup corresponding to the (forward) Kolmogorov operator A^* of the process $\{x_t; t \geq 0\}$, and A^* is the formal adjoint of the operator A (see equation (4.31.a)). Again the kernel $p_{s,t}(\eta; x)$ of the semigroup $T_{s,t}^*, s \leq t$, satisfies the Kolmogorov (or Fokker-Plank) forward equation

$$\frac{\partial}{\partial t} p_{s,t}(\eta; x) = A^* p_{s,t}(\eta; x),$$

and under our assumptions (A1)-(A3), the semigroup $T_{s,t}^*$ possesses the following properties

(Si) For all $0 \leq s \leq \theta \leq t$,

$$(T_{s,t}^* \beta)(x) = (T_{s,\theta}^* \beta)(x) + \int_{\theta}^t A^*(T_{\theta,\tau}^* \beta)(x) d\tau, \quad (4.33)$$

(Sii) Let \mathbf{H} denote the family of functions $\{\beta\} \equiv \{\beta(x), x \in R^n\}$, satisfying the integrability condition

$$\int_{R^n} |(T_{s,t}^* \beta)(x)|^2 dx < \infty.$$

Then there exists a constant $\delta > 0$ such that

$$\int_{R^n} |(T_{s,t}^* f)(x)|^2 dx \leq \delta \int_{R^n} |f(x)|^2 dx,$$

for all $f \in \mathbf{H}$ and $0 \leq s \leq t < \infty$.

With this background we now present the following corollary .

Corollary 4.1

Suppose for each $\eta \in R^n$, the measure $\Gamma \rightarrow P_{s,t}(\eta; \Gamma), s \leq t, \Gamma \in \mathcal{B}(R^n)$, is absolutely continuous with respect to Lebesgue measure admitting the density $\gamma \rightarrow p_{s,t}(\eta; \gamma)$ and let the semigroup $T_{s,t}^*$ satisfy the properties (Si) and (Sii). Then the measure $\Gamma \rightarrow P(t, \Gamma | \mathcal{F}_t^y), t \geq 0$, (see equation (4.14)) is also absolutely continuous with respect to Lebesgue measure admitting a density $p(t, \gamma | \mathcal{F}_t^y)$ and that

$$p(t, \gamma | \mathcal{F}_t^y) = \frac{\varphi(t, \gamma)}{\int_{R^n} \varphi(t, \gamma) d\gamma}, \quad (4.35)$$

where φ satisfies, μ_2 -a.s, the following stochastic integral equations

$$\begin{aligned} \varphi(t, \gamma) = & q(t, \gamma) - \int_0^t \int_{R_0^m} (T_{s,t}^*(\varphi c))(t, \gamma) \cdot g_2(y_s, u) \Pi(du) ds \\ & + \int_0^t T_{s,t}^*(\varphi c)(t, \gamma) \cdot dy_s, \end{aligned} \quad (4.36)$$

and

$$\begin{aligned} \varphi(t, \gamma) = & (T_{s,t}^* \varphi)(t, \gamma) - \int_s^t \int_{R_0^m} T_{\theta,t}^*(\varphi c)(t, \gamma) \cdot g_2(y_\theta, u) \Pi(du) d\theta \\ & + \int_s^t T_{\theta,t}^*(\varphi c)(t, \gamma) \cdot dy_\theta. \end{aligned} \quad (4.36.a)$$

Here q denotes the density corresponding to the measure Q (see equation (4.13)) and $c(y_t, x_t) \equiv c(t, x_t)$.

Proof

The proof essentially follows from similar arguments as those of Corollary 2.1 and hence omitted. ■

In the next section, we use the results of the above corollary to derive a Zakai type equation for the unnormalized density φ from which the conditional expectation $E_1\{f(x_t) | \mathcal{F}_t^y\}$ can be computed.

4.4 An Evolution Equation For The Unnormalized Density

In this section we use the results of the previous section and the properties (Si) and (Sii) of the semigroup $T_{s,t}^*$ to derive a stochastic differential equation for the unnormalized density $\varphi(t, x); x \in R^n, t \geq 0$, from which the conditional expectation $E_1\{f(x_t)|\mathcal{F}_t^y\}$ can be computed using the following relation

$$E_1\{f(x_t)|\mathcal{F}_t^y\} \equiv \langle p(t, \cdot | \mathcal{F}_t^y), f(\cdot) \rangle = \frac{\langle \varphi(t, \cdot), f(\cdot) \rangle}{\langle \varphi(t, \cdot), 1 \rangle}, \quad (4.37)$$

where

$$\langle \varphi(t, \cdot), f(\cdot) \rangle \equiv \int_{R^n} f(x)\varphi(t, x) dx.$$

Definition 4.1

The semigroup $T_{s,t}^*; s \leq t$, given by equation (4.32) is said to be of class \mathcal{M} if it satisfies the properties (Si) (Sii) and

- (Siii) Let f be any bounded continuous function on R^n whose first and second partial derivatives (with respect to x) are bounded, continuous, and satisfy Holder condition on R^n . Then $(T_{s,t}^* f)(t, x)$ and its first and second derivatives (with respect to x) are continuous on $(0, T) \times R^n$ and bounded on $[0, T] \times G$, where G is any bounded subset of R^n .

Theorem 4.2

Suppose that assumptions (A1)-(A4) hold and the function g_2 satisfies the hypotheses of Lemma 4.1 and the semigroup $T_{s,t}^*; s \leq t$, is of class \mathcal{M} . Let $\varphi(t, x), x \in R^n, t \geq 0$, satisfy equation (4.36.a) and μ_2 -a.s, the function φ with its first and second derivatives, with respect to x , are continuous, bounded and satisfy Holder

condition on R^n , uniformly in $t \in [0, T]$. Then the unnormalized density φ satisfies the following linear stochastic partial differential equation

$$d\varphi(t, x) = A^* \varphi(t, x) dt + \varphi(t, x) c(t, x) \cdot \left\{ dy_t - \int_{R_0^n} g_2(y_t, u) \Pi(du) dt \right\}, \quad (4.38)$$

for all $(t, x) \in [0, T] \times R^n$, where $c(t, x) \equiv c(y_t, x)$, $x \in R^n$, is given by equation (4.15).

Proof:

From equations (4.33) and (4.36.a) we have

$$(T_{s,t}^* \varphi)(t, x) = \varphi(s, x) + \int_s^t A^*(T_{s,\theta}^* \varphi)(\theta, x) d\theta, \quad (4.39)$$

and

$$\begin{aligned} \varphi(t, x) &= (T_{s,t}^* \varphi)(t, x) + \int_s^{\theta} (T_{\tau,\theta}^* (\varphi c))(\theta, x) \cdot \{ dy_\tau \\ &\quad - \int_{R_0^n} g_2(y_\tau, u) \Pi(du) d\tau \}. \end{aligned} \quad (4.40)$$

From equation (4.33), we have

$$(T_{\alpha,t}^* (\varphi c))(t, x) = (\varphi c)(\alpha, x) + \int_\alpha^t A^*(T_{\alpha,\tau}^* (\varphi c))(\tau, x) d\tau. \quad (4.41)$$

Since

$$\begin{aligned} M(t) &\equiv y_t - \int_0^t \int_{R_0^n} g_2(y_\theta, u) \Pi(du) d\theta, \\ &\equiv V_t + \int_0^t \int_{R_0^n} g_2(y_\theta, u) q(du, d\theta), \end{aligned} \quad (4.42)$$

integrating both sides of (4.41) with respect to the right continuous \mathcal{G}_t -martingale $M(t)$, we have

$$\begin{aligned} \int_s^t (T_{\alpha,t}^* (\varphi c))(t, x) \cdot dM(\alpha) &= \int_s^t (\varphi c)(\alpha, x) \cdot dM(\alpha) \\ &\quad + \int_s^t \int_\alpha^t A^*(T_{\alpha,\tau}^* (\varphi c))(\tau, x) d\tau \cdot dM(\alpha). \end{aligned} \quad (4.43)$$

Using equation (4.36.a) in (4.43), we obtain

$$\begin{aligned} \varphi(t, x) &= (T_{s,t}^* \varphi)(t, x) + \int_s^t (\varphi c)(\alpha, x) \cdot dM(\alpha) \\ &\quad + \int_s^t \int_{\alpha}^t A^*(T_{\alpha,r}^*(\varphi c))(\tau, x) d\tau \cdot dM(\alpha). \end{aligned} \quad (4.44)$$

Substituting the expression for $(T_{s,t}^* \varphi)(t, x)$ from equation (4.39) into (4.44), we have

$$\begin{aligned} \varphi(t, x) &= \varphi(s, x) + \int_s^t A^*(T_{s,\theta}^* \varphi)(\theta, x) d\theta + \int_s^t (\varphi c)(\alpha, x) \cdot dM(\alpha) \\ &\quad + \int_s^t \int_{\alpha}^t A^*(T_{\alpha,r}^*(\varphi c))(\tau, x) d\tau \cdot dM(\alpha). \end{aligned} \quad (4.45)$$

From equations (4.42) and (4.47), it follows that

$$\begin{aligned} \varphi(t, x) &= \varphi(s, x) + \int_s^t (A^* \varphi)(\theta, x) d\theta - \int_s^t A^* \int_s^{\theta} (T_{r,\theta}^*(\varphi c))(\theta, x) \cdot dM(\alpha) d\theta \\ &\quad + \int_s^t \int_{\alpha}^t A^*(T_{\alpha,r}^*(\varphi c))(\tau, x) d\tau \cdot dM(\alpha) \\ &\quad + \int_s^t (\varphi c)(\alpha, x) \cdot dM(\alpha). \end{aligned} \quad (4.46)$$

Again following similar arguments as those of Theorem 2.2, and using the following facts

(F1) For any \mathcal{G}_t -measurable R^k valued functions f_1 and f_2 ,

$$E \left\{ \int_0^t f_1(s) \cdot dM(s) \int_0^t f_2(s) ds \right\} = E \int_0^t f_2(s) \int_0^s f_1(\xi) \cdot dM(\xi), \quad (4.47)$$

where $M(t)$ is given by (4.42) and

(F2) For any \mathcal{G}_t -measurable functions f_1 and f_2 ,

$$E \int_0^t f_1(s) \cdot dM(s) \int_0^t f_2(s) \cdot dM(s) = E \int_0^t f_1(s) \cdot f_2(s) c(s) ds, \quad (4.48)$$

where

$$e(t) \equiv 1 + \int_{R_0^m} |g_2(y_t, u)|^2 \Pi(du), \quad (4.49)$$

one can verify that

$$\int_s^t \int_{\alpha} A^*(T_{\alpha, \tau}^*(\varphi c))(\tau, x) d\tau \cdot dM(\alpha) = \int_s^t A^* \int_s^{\theta} (T_{\tau, \theta}^*(\varphi c))(\theta, x) \cdot dM(\tau) d\theta. \quad (4.50)$$

Thus equation (4.46) reduces to

$$\varphi(t, x) = \varphi(s, x) + \int_s^t (A^* \varphi)(\theta, x) d\theta + \int_s^t (\varphi c)(\alpha, x) \cdot dM(\alpha), \quad (4.51)$$

where the \mathcal{G}_t -martingale $M(t), t \geq 0$, is given by (4.42). Clearly, (4.51) is the integral form of (4.38) and the proof of the Theorem follows. ■

Remark 4.3

For the case where the increments of the Wiener process V (see equation (4.4)) has the property

$$E\{(V_t - V_s)(V_t - V_s)'\} = \int_s^t R(\theta) d\theta,$$

where R is positive definite ($k \times k$) matrix, the filter equation (4.38), in this case, is given by

$$d\varphi(t, x) = A^* \varphi(t, x) dt + \varphi(t, x) c(t, x) \cdot R^{-1}(t) \left\{ dy_t - \int_{R_0^m} g_2(y_t, u) \Pi(du) dt \right\}. \quad (4.51.a)$$

Remark 4.4

For the case where the processes $x_t, y_t, t \geq 0$, are governed by

$$dx_t = a(x_t)dt + b(x_t)dW_t + \int_{R_0^m} g(x_t, u)\tilde{\gamma}(du, dt), \quad (4.52)$$

$$dy_t = h(x_t)dt + dV_t, \quad (4.53)$$

it is clear from Lemma 4.1 that $g_1 = g_2 = 0$, $\hat{\rho} = \tilde{\rho} = 1$, and $\alpha(x, y) = -h(x)$, and hence it follows from (4.15) that $c(y, x) = h(x)$. In this case equation (4.38) reduces to

$$\begin{aligned} d\varphi(t, x) &= \mathbf{A}^*\varphi(t, x)dt + h(x) \cdot dy_t, \\ \varphi(0, x) &= p_0, \end{aligned} \quad (4.54)$$

where \mathbf{A}^* is the formal adjoint of the operator \mathbf{A} given by equation (4.12.a). Using Lemma 2.3, it follows from equation (4.54) that the normalized density $p(t, x|\mathcal{F}_t^y)$ (see equation (4.14)) is governed by the following (Kushner type) nonlinear stochastic partial differential equation

$$\begin{aligned} dp(t, x|\mathcal{F}_t^y) &= \mathbf{A}^*p(t, x|\mathcal{F}_t^y)dt + (h(x) - \hat{h}_t)p(t, x|\mathcal{F}_t^y) \cdot [dy_t - \hat{h}_t dt], \\ p(0, x) &= p_0(x), \end{aligned}$$

and hence for any $f \in C_0^2(R^n)$, the estimate \hat{f}_t is given by the solution of the following stochastic differential equation

$$\begin{aligned} d\hat{f}_t &= (\widehat{\mathbf{A}f})_t dt + [(\widehat{hf})_t - \hat{h}_t \hat{f}_t] \cdot [dy_t - \hat{h}_t dt], \\ \hat{f}(0) &= Ef(x_0). \end{aligned} \quad (4.55)$$

In the next section, we will use the above equation to obtain the filter equation for the case where the process x is governed by (4.52) whereas the observed process (discrete) is excited by white Gaussian noise.

4.5 Nonlinear Filtering Formula For Discrete Observations

In this chapter we have considered, so far, the filtering problem for a class of systems governed by nonlinear Ito differential equations driven by Wiener and generalized Poisson processes with discontinuous (as well as continuous) observations. Although the filtering problem, for the case where the process y is continuous, is of theoretical interest, discrete observations are more convenient in many practical applications (see chapter 5).

In this section we consider the filtering problem for the case where the state process $x_t, t \geq 0$, is governed by the Ito equation (4.52), whereas the observed process y (discrete) is given by

$$y(t_k) = h(x(t_k)) + n(t_k), \quad (4.56)$$

where h is a vector in R^m satisfying our basic assumptions of section 4.2. The process $n(t_k); k = 1, 2, 3, \dots, N$, is assumed to be a zero mean white Gaussian noise sequence with values in R^m , so that

$$E\{n(t_k)n'(t_j)\} = \delta_{kj}R(t_k),$$

where R is $(m \times m)$ positive definite matrix and δ is the dirac function. As indicated earlier in chapter 1, the above filtering problem was also considered in [70] for the case where the signal process $x_t, t \geq 0$, is driven only by a continuous martingale. However, in this section, we are interested in obtaining the corresponding filter equation when the signal process is governed by the Ito equation (4.52). We shall use the result given in Remark 4.3 to obtain two differential equations for the estimate \hat{f}_t between and at observations.

Let $f \in C_0^2(R^n)$ and let $\mathcal{F}_{t_k}^y$ denote the σ -field generated by the process y up to time t_k . Since the process y is only available at the discrete points $t_k; 1 \leq k \leq N$, it is clear from equation (4.55) that the estimate \hat{f}_t is given by

$$\hat{f}_t = \begin{cases} \hat{f}(t|t_{k-1}), & \text{for } t \in (t_{k-1}, t_k), \\ \hat{f}(t|t_k), & \text{for } t = t_k, \end{cases} \quad (4.57)$$

where $\hat{f}(t|t_k) \equiv E\{f(x_t)|\mathcal{F}_{t_k}^y\}$ and $\hat{f}(0) = Ef(x_0)$. Utilizing equation (4.55), it follows that between observations, i.e, $t \in (t_{k-1}, t_k)$, the process $\hat{f}(t|t_{k-1})$ is governed by the following differential equation

$$\frac{d}{dt}\hat{f}(t|t_{k-1}) = (\widehat{A}f)(t|t_{k-1}), \quad t \in (t_{k-1}, t_k), \quad (4.58)$$

with initial condition $\hat{f}(t_{k-1}|t_{k-1})$, where the operator A is given by (4.12.a). On the other hand, at the observation time $t = t_k$, the estimate $\hat{f}(t_k|t_k)$ can be obtained by solving the following differential equation

$$\frac{d}{dt}\hat{f}(t_k|t) = [(\widehat{h}f)(t_k|t) - \widehat{h}(t_k|t)\hat{f}(t_k|t)] \cdot R^{-1}(t_k)[y(t_k) - \widehat{h}(t_k|t)], \quad (4.59)$$

for all $t \in (t_{k-1}, t_k]$, with the initial condition $\hat{f}(t_k|t_{k-1})$ given by the solution of (4.58). Note that equations (4.58) and (4.59) will be used in the next chapter to obtain the filter equations for navigation system.

Remark 4.4

The filter equations (4.58) and (4.59) are similar in nature to those obtained in [70] except that the generator A appearing in (4.58) is different.

In the next section, we present three examples to show that the results of Zakai [76], Pardoux [67], Dimasi and Ruggaldier [31], can be obtained from the result of Theorem 4.2.

4.6 Examples

Example 4.1 (Zakai case)

Let the processes $\{x_t, y_t; t \geq 0\}$, be governed by the following nonlinear Ito differential equations

$$\begin{aligned} dx_t &= a(x_t)dt + b(x_t)dW_t, \\ dy_t &= h(x_t)dt + dV_t, \end{aligned} \tag{4.60}$$

where a, b, W, h and V are as in section 4.2. Our problem is to obtain the corresponding filter equation from the result of Theorem 4.2.

Let μ_1 and μ_2 denote the measures induced, respectively, by the systems of equations (4.60) and

$$\begin{aligned} dx_t &= a(x_t)dt + b(x_t)dW_t, \\ dy_t &= dV_t. \end{aligned} \tag{4.61}$$

Comparing the systems (S1) and (S2), given in Lemma 4.2, with the systems (4.60) and (4.61), it is clear that $g \equiv 0, g_1 = g_2 \equiv 0$, and $\alpha(x, y) = -h(x)$. Therefore, it follows from assumption (iii) of Lemma 4.1 that $\varrho = \tilde{\varrho} = 1$ and hence from (4.15) we have $c(y, x) = -\alpha(x, y) = h(x)$. It is also clear from equation (4.12.a), for $g \equiv 0$, that the infinitesimal generator A of the process x is given by

$$A f \equiv \sum_{i=1}^n a_i \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^n (bb')_{ij} \frac{\partial^2 f}{\partial x_i \partial x_j}. \tag{4.62}$$

In this case, the filter equation (4.38) reduces to

$$d\varphi(t, x) = A^* \varphi(t, x)dt + \varphi(t, x)h(x) \cdot dy_t,$$

giving the usual Zakai equation, where A^* is the formal adjoint of the operator A , given by (4.62).

Example 4.2 (Pardoux's Case)

Suppose that the processes $\{x_t, y_t; t \geq 0\}$, are governed by the following stochastic differential equations

$$\begin{aligned} dx_t &= a(x_t)dt + b(x_t)dW_t, \\ dy_t &= \lambda(x_t)N(dt), \end{aligned} \tag{4.63}$$

where a, b, W are as in example 4.1, $\lambda > 0$ is scalar and $N((s, t))$, $s \leq t$, is a standard Poisson process (independent of W), with mean $(t - s)$, giving the number of jumps (of size one) during the time interval $(s, t]$. Again our problem is to obtain the filter equation corresponding to the above system.

Let μ_1 and μ_2 denote the measures induced, respectively, by the systems of equations (4.63) and

$$\begin{aligned} dx_t &= a(x_t)dt + b(x_t)dW_t, \\ dy_t &= N(dt). \end{aligned} \tag{4.64}$$

Then comparing the systems (S1) and (S2) of Lemma 4.1 with (4.63) and (4.64), it follows that $g \equiv 0, h \equiv 0, \alpha(x, y) = \lambda(x) - 1$,

$$g_1(x, y, u) = \lambda(x)I_{\{1\}}(u), \tag{4.65}$$

and

$$g_2(y, u) = I_{\{1\}}(u), \tag{4.66}$$

where

$$I_{\{1\}}(u) \equiv \begin{cases} 1, & \text{if } u \in \{1\}, \\ 0, & \text{otherwise.} \end{cases}$$

Further,

$$\begin{aligned} \int_{R_0^m} g_1(x, y, u)\gamma(du, dt) &= \int_{R_0^m} \lambda(x)I_{\{1\}}(u)\gamma(du, dt) \\ &= \lambda(x)\gamma(\{1\}, dt) \equiv \lambda(x)N(dt), \end{aligned} \tag{4.67}$$

and

$$\begin{aligned}
 \dot{q}(\{1\}, dt) &= \gamma(\{1\}, dt) - E(\gamma(\{1\}, dt)) \\
 &\equiv N(dt) - E(N(dt)) \\
 &= N(dt) - dt.
 \end{aligned} \tag{4.68}$$

Using assumption (iii) of Lemma 4.1 and equations (4.65) and (4.66), it follows that

$$\int_{\{u: I_{(1)}(u) \in A\}} \Pi(du) = \int_{\{u: \lambda(x) I_{(1)}(u) \in A\}} \varrho(x, y, u) \Pi(du), \quad A \in \mathcal{B}(R^n),$$

or equivalently

$$\int_A (I_{\{1\}}^{-1} \Pi)(d\eta) = \int_A \varrho(x, y, u) \lambda(x) (I_{\{1\}}^{-1} \Pi)(d\eta),$$

for any $A \in \mathcal{B}(R^n)$. Clearly, the above equality holds only if $\varrho(x, y, u) \lambda(x) \equiv 1$, which implies that

$$\tilde{\varrho}(x, y, u) = \frac{1}{\varrho(x, y, u)} = \lambda(x). \tag{4.69}$$

Using equations (4.65), (4.16) and (4.69), and removing the Wiener process V , it follows from equations (4.15) and (4.16) that

$$\int_0^t c(y_s, x_s) dM(s) = \int_0^t (\lambda(x) - 1)(N(ds) - ds),$$

where

$$M(t) = \int_0^t (N(ds) - ds).$$

Hence it is easy to verify that

$$c(y, x) = \lambda(x) - 1,$$

for any $x \in R^n$, and $y \in R$. In this case, the filter equation (4.38) reduces to

$$d\varphi(t, x) = \mathbf{A}^* \varphi(t, x) dt + \varphi(t, x) \{ \lambda(x) - 1 \} [dy_t - dt],$$

giving the result obtained in [67], where \mathbf{A}^* is the formal adjoint of the operator A given by (4.62).

Example 4.3 (Di Masi and Runggaldier Case)

Let the processes $\{x_t, y_t; t \geq 0\}$, be governed by the following stochastic differential equations

$$\begin{aligned} dx_t &= a(x_t)dt + b(x_t)dW_t, \\ dy_t &= h(x_t)dt + dV_t + N(dt), \end{aligned} \tag{4.70}$$

where a, b, W are as in Example 4.1; h is scalar, V is one-dimensional standard Wiener process and N (see Example 4.2) is a doubly stochastic Poisson process, independent of the processes W and V , with intensity $\lambda(x_t)$. Again we are interested in obtaining the corresponding filter equation using the result of Theorem 4.2.

Let μ_1 and μ_2 denote the measures induced, respectively, by the systems (4.70) and

$$\begin{aligned} dx_t &= a(x_t)dt + b(x_t)dW_t, \\ dy_t &= dV_t + N(dt), \end{aligned} \tag{4.71}$$

where under the measure μ_2 , the process $N(dt)$ is a standard Poisson process with intensity dt , independent of the process x_t . Comparing systems (S1) and (S2) of Lemma 4.1 with the systems (4.70) and (4.71) and using same arguments as those of Examples 4.1 and 4.2, it is easy to verify that

$$\begin{aligned} g_1(x, y, u) &= g_2(y, u) = 1, \\ \tilde{g}(x, y, u) &= \lambda(x), \\ \alpha(x, y) &= -h(x), \end{aligned}$$

and

$$c(y_t, x_t)[dy_t - dt] = (\lambda(x) - 1)[N(dt) - dt] + h(x_t)dV_t,$$

where, under the measure μ_2 , the process $(y_t - t); t \geq 0$, is a square integrable \mathcal{G}_t -martingale given by

$$y_t - t = V_t + \int_0^t [N(ds) - ds], \quad \mu_2 - a.s.$$

Using the above facts, it follows from equation (4.38) that

$$\begin{aligned} d\varphi(t, x) = & \mathbf{A}^* \varphi(t, x) dt + \varphi(t, x) h(x) [dy_t - N(dt)] \\ & + \varphi(t, x) (\lambda(x) - 1) [N(dt) - dt], \end{aligned} \quad (4.72)$$

giving the result obtained in [31], where \mathbf{A}^* is given by (4.62).

In the next chapter we consider the problem of modelling the errors that arise in navigation sensors. Then utilizing the result of section 4.5, we develop the corresponding linear filter equations from which the ship (or aircraft) position and velocity can be estimated.

4.7 Summary, Conclusion and Some Open Questions

In this chapter, we have considered the filtering problem for a class of systems governed by nonlinear Ito equations driven by Wiener and generalized Poisson processes. The observed process was also modelled as a nonlinear Ito differential equation driven by Wiener and Poisson processes. Utilizing the Radon-Nikodym derivative of measures induced by systems driven by right continuous martingales, we have obtained a stochastic integral equation for unnormalized density on the basis of which the filter equation was derived. Further, it is clear from equation (4.38) that in order to compute the estimate $\hat{f}_t, t \geq 0$, one needs only to know the history of the observed process up to time t as well as the mean of the Poisson noise γ appearing in (4.4). In contrast, the solutions of filter equations obtained by Di Masi [31], for example, (see equation (4.72)) requires observing the Poisson noise of the output process, which is practically difficult. Finally, three examples were presented showing that the results given in [31], [67] and [76], can be obtained from our main result given in Theorem 4.2.

In Lemma 4.1 we have assumed the existence of the functions g_2 and ϱ satisfying condition (iii) (see Lemma 4.1). We have been able to determine these functions explicitly for the given examples which represent many commonly used models in nonlinear filtering (see [31], [67] and [76]). However, the question of existence of such functions, given g_1 and the measure Π , has not been settled in this thesis and it remains as open problem.

In Theorem 4.1, we have also assumed the existence of a bounded \mathcal{G}_t -measurable function c satisfying the relation (4.15). This function has been explicitly determined for examples 4.1 and 4.2. In example 4.3 an implicit expression was given. It appears that when the observed process is only driven by a Poisson process, an explicit expression for c is possible. On the other hand, when the observed

process is driven by both Wiener and Poisson processes, the explicit expression is rather difficult. However, we observe that given g_2 and g (and hence \tilde{g}), the function c has the following probabilistic expression

$$c(y_t, x_t) = \lim_{\Delta t \rightarrow 0} \frac{\int_{R_0^m} [\tilde{g}(x_t, y_t, u) - 1] q(du, \Delta t) - \alpha(x_t, y_t) \cdot \Delta V_t}{\int_{R_0^m} g_2(y_t, u) q(du, \Delta t) + \Delta V_t},$$

μ_2 -a.s, where $\Delta V_t \equiv V_{t+\Delta t} - V_t$. The question of existence of this limit, remains also an open problem.

CHAPTER 5
MODELLING AND FILTERING OF ERRORS
ARISING IN NAVIGATION SYSTEMS

5.1 Introduction

Over the last few years considerable attention has been given towards developing a marine integrated navigation system that combines the navigation information, from a wide variety of sensors, to provide an accurate and reliable estimate of ship (or aircraft) position and velocity. As indicated earlier in chapter 1, this problem was investigated by several authors [1, 32, 59, 60, 64] and their work was mainly based on the assumption that the errors arise in navigation sensors, such as Dead Reckoning, Loran-C, Omega, Satellite, etc., can be modelled as linear stochastic differential equation (Ito differential equation) driven by Wiener process. Then based on these models and assuming that the observed process (discrete) is exited by white Gaussian noise, they developed a recursive discrete Kalman filter which consists of a set of difference equations for the estimator and covariance matrix. By solving these equations they obtained an estimate of these errors from which ship (or aircraft) position and velocity can be determined. In most circumstances, these error models and the corresponding Kalman filter give quite satisfactory results. However, there are some well known errors, particularly those arise in Loran-C and Omega radio aids, that do not conform to these standard models and therefore warrant further investigation.

In this chapter we are interested in the problem of modelling and filtering the errors that arise in Dead Reckoning and Loran-C navigation sensors. The Dead Reckoning error is modelled as a linear stochastic differential equation (Ito differential equation) driven by Wiener process, as already proposed in [60, 64]. On

the other hand, the Loran-C error is modelled as the sum of two independent random processes. The first process (continuous) is introduced to represent the signal propagation error (random) which is the difference between the actual transmitter-receiver signal travel time (through the atmosphere over land and sea boundary layers) and the ideal signal travel time predicted by assuming the vacuum speed of light. Since the propagation error is random and continuously varying, this process is modelled as an Ito (linear) differential equation driven by Wiener process [60,64]. In order to include the rapid behaviour of the Loran-C error, due to cycle selection error (or cycle hop), the second random process is modelled as a pure jump process driven by two independent Poisson random measures [8,9]. Using these models and assuming that the observed process (discrete) is excited by white Gaussian noise , we use the results of section 4.5 to develop a linear unbiased recursive filter from which ship (or aircraft) position and velocity can be easily determined. In fact our proposed filter equations are similar in nature to those of the conventional Kalman filter except that we have added extra terms in the estimator and covariance matrix equations to include the effect of the jump process presented in the state equation to explain the rapid behaviour of the Loran-C error.

In section 5.2 we present the Dead Reckoning error model, which was already proposed in [60, 64], along with our proposed model for Loran-C error as well as the observation model. In section 5.3, we utilize these models and the results of chapter 4, to develop a linear unbiased recursive filter. In section 5.4 we present a numerical algorithm for computing the required estimates. Finally, in section 5.5 we present a numerical example to illustrate the behaviour of the proposed filter compared to that of the conventional Kalman filter. In fact it is known that when the state (signal) process is driven by both Wiener and Poisson Processes, the result of Kalman does not hold and hence one expects that the corresponding

estimates will not be satisfactory. However, our comparison is mainly to illustrate the effectiveness of the proposed filter and to show how poor our estimates can be if one uses the wrong filter.

5.2 Dead Reckoning, Loran-C and Observation Models

5.2.1 Dead Reckoning Error Model

Since the Dead Reckoning sensors measure velocity relative to the water, a dominant source of error is the ocean current. Under the assumptions that the gyrocompass heading error can be neglected and the northern and eastern components of the velocity error are uncorrelated, the Dead Reckoning velocity errors together with the latitude and longitude errors $(\tilde{\lambda}, \tilde{L})$ are governed by the following linear stochastic differential equations [60, 64]

$$\begin{aligned}dV_N(t) &= -\frac{1}{T_c}V_N(t)dt + dW_1(t), \\dV_E(t) &= -\frac{1}{T_c}V_E(t)dt + dW_2(t), \\ \frac{d}{dt}\tilde{\lambda}(t) &= V_N(t), \\ \frac{d}{dt}\tilde{L}(t) &= V_E(t),\end{aligned}\tag{5.1}$$

where $T_c(> 0)$, denotes the ocean current correlation time and V_N and V_E are the northern and eastern components of the velocity error (ocean current and speed-log errors), respectively. Further, the process $W \equiv (W_1, W_2)'$ is a Wiener process with independent components.

5.2.2 Loran-C Error Model

Before presenting our proposed model for Loran-C receiver error, let us briefly describe the basic functions of this receiver and discuss the source of these errors. Loran-C is a hyperbolic radio positioning system developed by the U.S. navy for coastal navigation. The system basically works on the time difference between receiving pulses from two transmitters, a master and slave (secondary) transmitter.

This time difference defines a range difference, which defines a hyperbolic line of position on the earth surface. Measurement of another time difference between receiving another pulse from the master and a second slave transmitter, defines another hyperbolic line of position. The receiver location is then given by the point of intersection of these two lines. Further, these pulses are modulated at 100KHz and the receiver is designed to lock onto the third cycle on the rising edge of the pulse from each station (master or slave) in order to compute the corresponding time (range) differences. The third cycle of the ground-wave pulse is used to avoid the skywave contamination where the signal reflected off the ionosphere (skywave) interferes with the ground wave. In fact the ground-wave is usually far enough ahead of the skywave to keep its first few cycles uncontaminated.

The major errors that can affect the determination of these time differences are

(i) Land (sea) error (random)

This error is due to the difference between free space electromagnetic propagation and propagation through the atmosphere over a land (or sea) boundary condition which is assumed to be random.

(ii) Error due to receiver noise (random)

In practice, this error is usually modelled as a Gaussian random variable.

(iii) Cycle selection error (or cycle hop)

This error occurs when the receiver fails to select the third cycle of the pulse. For example, if the receiver selects the fourth cycle of the pulse this will cause an error of 10 micro-seconds in the time difference measurement and consequently the line of position will be in error by 3 Km or even more. The likelihood of the cycle selection error may increase due to the following reasons :

- a- local radio interference (frequently encountered in port) ,

- b- shielding of the antenna, for example due to dockside buildings ,
 - c- precipitation static, which occurs at the onset of snow flurries, rain showers or wet fog ,
 - d- icing or a coating of dirt on the antenna or antenna coupling ,
- and
- e- skywave interference by night, and particularly at dawn and dusk.

Although the cycle selection error does not occur frequently, neglecting its effect may lead to a very poor estimate of position.

Based on the above description, we assume that each Loran-C time difference error can be modelled as the sum of two independent random processes, which we denote by $e'(t)$ and $e''(t), t \geq 0$. The process $e'(t), t \geq 0$, is modelled as a diffusion process governed by the following linear Ito differential equation [60]

$$de'(t) = -\frac{1}{T}e'(t)dt + dW_3(t), \quad (5.2)$$

where T is a given positive parameter and W_3 is one dimensional Wiener process. The process $e''(t), t \geq 0$, is modelled as a pure jump process driven by two independent Poisson processes. We will adapt the procedure given in [8,9], which has been used to model the tie line on-off behaviour in the reliability dynamics of interconnected power systems, to develop a stochastic differential equation for the stochastic process $e''(t), t \geq 0$, which is introduced to explain the rapid changes of Loran-C time difference error due to cycle hop.

Since we will only consider the case where the Loran-C time difference error can have one cycle slip, it is clear from the state diagram presented in fig. 5.1 that $e''(t), t \geq 0$, is a random process with values in the set $\{0, \gamma\}$ where $\gamma > 0$ is a constant indicating the size of jump. Let $\{\tau_1, \tau_2, \dots\}$ denote the sequence of

random times at which the process e'' resides in the state '0' (i.e, no cycle selection error has occurred) and let $\{\sigma_1, \sigma_2, \dots\}$ denote the corresponding times at which e'' resides in the state γ . Clearly, with probability one, $\tau_m < \tau_{m+1}$, $\sigma_m < \sigma_{m+1}$ and $\tau_m < \sigma_m < \tau_{m+1}$ for all integers $m \geq 1$. Let

$$T_m(\gamma) \equiv \tau_{m+1} - \sigma_m, \quad (5.3.a)$$

and

$$T_m(0) \equiv \sigma_m - \tau_m, \quad (5.3.b)$$

denote the residence times, of the process e'' in the states ' γ ' and '0', respectively. We assume that the random variables $T_m(\gamma)$ and $T_m(0)$, $m \geq 1$, are exponentially distributed with parameters β and δ , respectively, independent of m . That is

$$\Pr\{T_m(\gamma) \leq T\} = \int_0^T \beta \exp(-\beta t) dt, \quad (5.4.i)$$

and

$$\Pr\{T_m(0) \leq T\} = \int_0^T \delta \exp(-\delta t) dt. \quad (5.4.ii)$$

Normally, the parameter β is very large compared to δ which implies that most of the time the process e'' is residing in the zero state. In other words, the cycle selection error does not occur very often.

Under the assumption that the random variables $T_m(\gamma)$ and $T_m(0)$, $m \geq 1$, are exponentially distributed with parameters independent of m , the process $e''(t)$, $t \geq 0$, is a homogeneous jump Markov process. Further, for $\Delta (> 0)$ sufficiently small, we have

$$\begin{aligned} \Pr\{e''(t) = \gamma | e''(t - \Delta) = 0\} &\cong \delta \Delta + o(\Delta), \\ \Pr\{e''(t) = 0 | e''(t - \Delta) = 0\} &\cong (1 - \delta \Delta) + o(\Delta), \\ \Pr\{e''(t) = 0 | e''(t - \Delta) = \gamma\} &\cong \beta \Delta + o(\Delta), \\ \Pr\{e''(t) = \gamma | e''(t - \Delta) = \gamma\} &\cong (1 - \beta \Delta) + o(\Delta). \end{aligned} \quad (5.5)$$

From the above relations, we arrive at the familiar state transition diagram:

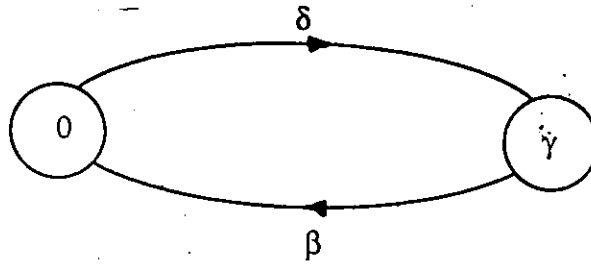


Fig. 5.1 State Diagram For The process e''

Thus the process $e''(t), t \geq 0$, is governed by a homogeneous jump Markov process with the state space $\{0, \gamma\}$. Introducing the random Poisson counting measures $N_0(dt)$ and $N_1(dt)$ on the Borel σ -field $\mathcal{B}(I)$ on $I \equiv [0, \infty)$ with

$$EN_0(\Delta) = \beta\Delta,$$

and

$$EN_1(\Delta) = \delta\Delta,$$

we can write a stochastic differential equation for the process $e''(t), t \geq 0$, as

$$de''(t) = 1(e''(t^-) = 0)N_1(dt) - 1(e''(t^-) = \gamma)N_0(dt), \quad (5.6)$$

where $1(x = \xi)$ denotes the indicator function given by

$$1(x = \xi) \equiv \begin{cases} 1, & \text{if } x = \xi, \\ 0, & \text{otherwise.} \end{cases}$$

It is easy to verify that equation (5.6) is mathematically equivalent to the model discussed above.

Using equations (5.2) and (5.6), the Loran-C time difference error, denoted e_c , is given by

$$e_c(t) = e'(t) + e''(t), \quad t \geq 0. \quad (5.7)$$

Combining equations (5.1), (5.2) and (5.6) and assuming that Loran-C receiver uses a master and two slave transmitters, the overall system model can be written as

$$\begin{aligned} dx(t) &= Fx(t)dt + BdW(t) + CN(dt), \\ x(0) &= x_0, \end{aligned} \tag{5.8}$$

where

$$x \equiv (V_N, V_E, \tilde{\lambda}, \tilde{L}, e_1', e_1'', e_2', e_2'')$$

and $W \equiv (W_1, W_2, \dots, W_4)'$ is 4-dimensional Wiener process with independent components such that

$$E\{(W(t) - W(s))(W(t) - W(s))'\} = (t - s)Q, \quad t \geq s. \tag{5.9}$$

Further, the process $N \equiv (N^1, N^2)'$ is 2-dimensional Poisson process, with independent components, given by

$$N^i(dt) = 1(e_i''(t^-) = 0)N_1^i(dt) + 1(e_i''(t^-) = \gamma)N_0^i(dt), \tag{5.10}$$

for $i = 1, 2$, where N_0^i and N_1^i are two independent Poisson processes with parameters β^i and δ^i , respectively. The time invariant matrices F, B and C are given by

$$F = \begin{pmatrix} -\frac{1}{T_c} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{T_c} & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{T} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{T} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{5.11}$$

$$B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{5.12}$$

and

$$C = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \gamma & 0 \\ 0 & 0 \\ 0 & \gamma \end{pmatrix} \quad (5.13)$$

Using equation (5.10), the (conditional) transition rate of the process N , denoted $\lambda = (\lambda^1, \lambda^2)'$, is given by

$$\begin{aligned} \int_t^{t+\Delta} \lambda^i(\theta) d\theta &\equiv E\left\{ \int_t^{t+\Delta} N^i(d\theta) | \sigma(e_i''(\theta^-)) \right\} \\ &\cong 1(e_i''(t) = 0) E\{N_1^i((t, t + \Delta)) | \sigma(e_i''(t))\} \\ &\quad - 1(e_i''(t) = \gamma) E\{N_0^i((t, t + \Delta)) | \sigma(e_i''(t))\} \\ &\cong 1(e_i''(t) = 0) \delta^i \Delta - 1(e_i''(t) = \gamma) \beta^i \Delta, \end{aligned} \quad (5.13.a)$$

for $i = 1, 2$, and $\Delta > 0$ sufficiently small, where $N((t, t + \Delta])$ denotes the number of jumps of size γ over the interval $(t, t + \Delta]$, and $\sigma(e_i''(t))$ denotes the σ -field generated by $\{e_i''(t)\}, t \geq 0$.

5.2.3 Observation Model

Under the assumption that each Loran-C receiver uses a master and two slave transmitters, the observed process $y(t), t \geq 0$, is related to the state $x(t), t \geq 0$, through the following linear relation [60,64]

$$y(t) = Hx(t) + V(t), \quad (5.14)$$

where $x(t), t \geq 0$, is the solution of equation (5.8), $y \in R^2$, H is (2×8) matrix and V is two dimensional Wiener process (with independent components) independent of the processes W and N and that

$$E\{(V(t) - V(s))(V(t) - V(s))'\} = (t - s)R, \quad t \geq s. \quad (5.15)$$

Further, the matrix H is given by

$$H = \begin{pmatrix} 0 & 0 & (\cos \Psi_M - \cos \Psi_{S_1}) & (\sin \Psi_M - \sin \Psi_{S_1}) & 1 & 1 & 0 & 0 \\ 0 & 0 & (\cos \Psi_M - \cos \Psi_{S_2}) & (\sin \Psi_M - \sin \Psi_{S_2}) & 0 & 0 & 1 & 1 \end{pmatrix}, \quad (5.16)$$

where Ψ_M denotes the bearing of the ship to the Loran-C master transmitter, and Ψ_{S_i} denotes the bearing to the slave transmitter i ; $i = 1, 2$. Note that the first term in the right hand side of (5.14) represents the difference between each Loran-C master-slave time (range) delay, as measured by Loran-C receiver, and the corresponding time (range) delay as calculated from Dead Reckoning position measurements and the known location and coding delay of each Loran-C transmitter. The second term represents the receiver noise with zero mean and constant covariance R .

Remark 5.1

Since it is clear from equations (5.8) and (5.14) that whenever a jump occurs in x it will be reflected in y ; therefore, observing y determines, automatically, the points of discontinuity of x . Hence the process $x(t), t \geq 0$, is a conditionally Gaussian process given the history $\{y_s; s \leq t\}$, provided that the initial state x_0 is a Gaussian random variable.

In the next section, we will use equations (5.8) and (5.14) and the results of section 4.5 to develop a recursive unbiased linear filter from which the state process $x(t), t \geq 0$, can be estimated given the observed path $\{y(s), s \leq t\}$.

5.3 Linear Unbiased Filter For Navigation System

Consider the navigation model (5.8) and suppose that the observed process $y(t), t \geq 0$, is given by (5.14). Let Ω denote the space of right continuous R^n ($n = 10$) valued functions having left limits. Let \mathcal{F} denote the σ -field generated by Borel subsets of Ω and let $(\Omega, \mathcal{F}, \mu)$ be a complete probability space on which all the random processes, described above, are defined. Let \mathcal{F}_t^y denote the σ -field generated by the process $\{y(s), s \leq t\}$ and $\sigma(x(t))$ the σ -field generated by $\{x(t)\}, t \geq 0$. In this section our main goal is to determine the conditional expectation $E\{x(t)|\mathcal{F}_t^y\}$ using the results of chapter 4. First let us determine the infinitesimal generator corresponding to the state process $x_t, t \geq 0$, with the help of the following Lemma.

Lemma 5.1

Consider the system (5.8) and let $\lambda = (\lambda^1, \lambda^2)'$ denote the conditional transition rate of the process N , given by

$$\begin{aligned} E\{N^i((t, t + \Delta t)) | e_i''(t) = \eta_i\} &\cong \lambda^i(t) \Delta t \\ &= 1(e_i''(t) = 0) \delta^i \Delta t - 1(e_i''(t) = \gamma) \beta^i \Delta t. \end{aligned} \quad (5.17)$$

for $i = 1, 2$. Let f be any twice continuously differentiable function on R^n ($n=10$).

Then

$$\lim_{\Delta t \rightarrow 0} E_{x, \eta} \frac{1}{\Delta t} \{f(x(t + \Delta t)) - f(x(t))\} = \mathbf{A}f, \quad (5.18)$$

where

$$\mathbf{A}f \equiv (f_x \cdot \{Fx + C\lambda(t)\}) + \frac{1}{2} \text{tr}(BQB' f_{xx}) + \frac{1}{2} \text{tr}(CAC' f_{xx}). \quad (5.19)$$

Here $E_{x, \eta}$ denotes the conditional expectation given $x(t) = x$ and $e(t) = \eta$; \cdot denotes the scalar product and $\text{tr}(B)$ denotes the trace of the matrix B . Further, \mathbf{A} is a diagonal matrix with elements $\mathbf{A}^{ii} = \lambda^i$, where λ^i is given by (5.17), $i = 1, 2$.

Proof

Using Taylor's series expansion, the function $f(x(t + \Delta t)), t \geq 0$, can be written as follows

$$\begin{aligned} f(x(t + \Delta t)) &= f(x(t)) + (f_x \cdot \Delta x(t)) \\ &\quad + \frac{1}{2}(f_{xx}\Delta x(t) \cdot \Delta x(t)) + o(|\Delta x|^2), \end{aligned} \quad (5.20)$$

where f_x and f_{xx} denote the first and second derivatives of f with respect to x and $\Delta x(t) \equiv x(t + \Delta t) - x(t)$

$$\cong Fx(t)\Delta t + B(W(t + \Delta t) - W(t)) + CN((t, t + \Delta t)). \quad (5.21)$$

Thus

$$\begin{aligned} E_{x,\eta}(f_x \cdot \Delta x(t)) &= (f_x \cdot Fx)\Delta t + B'f_x \cdot E_{x,\eta}(W(t + \Delta t) - W(t)) \\ &\quad + C'f_x \cdot E_{x,\eta}\{N((t, t + \Delta t))\} \\ &= (f_x \cdot \{Fx + C\lambda(t)\})\Delta t, \end{aligned} \quad (5.22.a)$$

and

$$\begin{aligned} \frac{1}{2}E_{x,\eta}\{f_{xx}\Delta x(t) \cdot \Delta x(t)\} &= \frac{1}{2}E_{x,\eta}B'f_{xx}B(W(t + \Delta t) - W(t)) \cdot (W(t + \Delta t) - W(t)) \\ &\quad + \frac{1}{2}E_{x,\eta}C'f_{xx}CN((t, t + \Delta t)) \cdot N((t, t + \Delta t)) \\ &= \frac{1}{2}tr(BQB'f_{xx})\Delta t + \frac{1}{2}tr(C\Lambda C'f_{xx})\Delta t, \end{aligned} \quad (5.22.b)$$

where Λ is a diagonal matrix with elements given by (5.17). Utilizing equations (5.20), (5.22.a) and (5.22.b) in (5.18) and the facts that $o(|\Delta x|^2) = o(\Delta t)$ and $\frac{o(\Delta t)}{\Delta t} \rightarrow 0$, as $\Delta t \rightarrow 0$, the result of the Lemma follows. ■

On the basis of the above Lemma, we now make use of the result of section 4.5 to obtain the filter equations corresponding to the navigation system (5.8) and (5.14).

First note that from equation (5.17) and the properties of conditional expectation, it is clear that

$$\begin{aligned}
 E\{N((t, t + \Delta t)) | \mathcal{F}_t^y\} &\equiv E\{E\{N^i((t, t + \Delta t)) | \mathcal{F}_t^y \vee \sigma(e_i''(t))\} | \mathcal{F}_t^y\} \\
 &= \int_t^{t+\Delta t} E\{\dot{\lambda}(\theta) | \mathcal{F}_t^y\} d\theta \\
 &= \int_t^{t+\Delta t} \Pr\{e_i''(\theta) = 0 | \mathcal{F}_t^y\} \delta^i d\theta \\
 &\quad - \int_t^{t+\Delta t} \Pr\{e_i''(\theta) = \gamma | \mathcal{F}_t^y\} \beta^i d\theta,
 \end{aligned} \tag{5.23}$$

for $i = 1, 2$ and $\Delta t > 0$ sufficiently small. Since, as indicated in Remark 5.1, the conditional density of the process x given y is Gaussian, the infinite dimensional filter equations (4.58) and (4.59) can be reduced to finite dimensional differential equations. Based on this fact, we now proceed to derive the filter equations for the navigation system given by equations (5.8) and (5.14).

Define

$$\hat{x}(t|t_{k-1}) \equiv E\{x(t) | \mathcal{F}_{t_{k-1}}^y\}, \quad t > t_{k-1}, \tag{5.24}$$

where $\mathcal{F}_{t_{k-1}}^y$ denotes the σ -field generated by the process y up to time t_{k-1} . Then setting $f \equiv x$ in equation (4.57), the estimate $\hat{x}(t), t \geq 0$, is given by

$$\hat{x}(t) = \begin{cases} \hat{x}(t|t_{k-1}) & \text{for } t \in (t_{k-1}, t_k), 1 \leq k \leq N, \\ \hat{x}(t|t_k) & \text{for } t = t_k, 1 \leq k \leq N. \end{cases} \tag{5.25}$$

Setting $f \equiv \hat{x}$ in (4.58) and (5.19) and using (5.23), it follows that between observations the process $\hat{x}(t_k|t_{k-1})$ can be determined by solving the following differential equation

$$\frac{d}{dt} \hat{x}(t|t_{k-1}) = F \hat{x}(t|t_{k-1}) + C \alpha(t), \tag{5.26}$$

for $i = 1, 2$. On the other hand, setting $f \equiv x$ and $h(x) = Hx$ in (4.59), one can easily verify that at the observation time t_k , the estimate $\hat{x}(t_k|t_k)$ is determined by solving the following differential equation

$$\begin{aligned} \frac{d}{dt} \hat{x}(t_k|t) &= [(\widehat{xx'})(t_k|t) - \hat{x}(t_k|t)\hat{x}'(t_k|t)] \times \\ &\times H'R^{-1}[y(t_k) - H\hat{x}(t_k|t)], \end{aligned} \quad (5.26.a)$$

for all $t \in (t_{k-1}, t_k]$, with the initial condition $\hat{x}(t_k|t_{k-1})$. Clearly, equations (5.26) and (5.26.a) describe the behaviour of the estimate $\hat{x}(t), t \geq 0$, between and at observations, respectively. In the remaining part of this section we shall follow similar arguments as those given in [45], to obtain the corresponding covariance matrix equations.

Let $\varphi(t, \tau), 0 \leq \tau \leq t < \infty$, denote the transition matrix corresponding to the system (5.8). Then the solution of the differential equation (5.8) can be expressed as

$$x(t) = \varphi(t, t_{k-1})x(t_{k-1}) + \int_{t_{k-1}}^t \varphi(t, s)BdW(s) + \int_{t_{k-1}}^t \varphi(t, s)CN(ds). \quad (5.27)$$

Note that since the functions φB and φC are deterministic and bounded, the stochastic integrals appearing in the right hand of (5.27) are well defined. Since φ is also the transition matrix corresponding to the system (5.26), the solution $\hat{x}(t|t_{k-1}), t \in (t_{k-1}, t_k)$, is given by

$$\hat{x}(t|t_{k-1}) = \varphi(t, t_{k-1})\hat{x}(t_{k-1}|t_{k-1}) + \int_{t_{k-1}}^t \varphi(t, s)C\alpha(s)ds. \quad (5.28)$$

Define

$$P(t|t_{k-1}) \equiv E\{(x(t) - \hat{x}(t|t_{k-1}))(x(t) - \hat{x}(t|t_{k-1}))'\}, \quad (5.29)$$

for all $t \in (t_{k-1}, t_k)$. Then substituting (5.27) and (5.28) into (5.29) and using the facts that

$$E \int_{t_{k-1}}^t \varphi(t, s)BdW(s) \cdot \int_{t_{k-1}}^t \varphi(t, s)BdW(s) = \int_{t_{k-1}}^t \varphi(t, s)BQB'\varphi'(t, s)ds,$$

and

$$\begin{aligned}
 & E \int_{t_{k-1}}^t \varphi(t,s) CN(ds) \cdot \int_{t_{k-1}}^t \varphi(t,s) CN(ds) \\
 &= E \{ E \{ \int_{t_{k-1}}^t \varphi(t,s) CN(ds) \cdot \int_{t_{k-1}}^t \varphi(t,s) CN(ds) | \sigma(e''(s^-)) \} \} \\
 &= E \int_{t_{k-1}}^t \varphi(t,s) C \Lambda(s) C' \varphi'(t,s) ds,
 \end{aligned}$$

we obtain

$$\begin{aligned}
 P(t|t_{k-1}) &= \varphi(t, t_{k-1}) P(t_{k-1}|t_{k-1}) \varphi'(t, t_{k-1}) + \int_{t_{k-1}}^t \varphi(t,s) B Q B' \varphi'(t,s) ds \\
 &\quad + \int_{t_{k-1}}^t \varphi(t,s) C \Gamma(s) C' \varphi'(t,s) ds,
 \end{aligned} \tag{5.30}$$

for all $t \in (t_{k-1}, t_k)$, where

$$P(t_{k-1}|t_{k-1}) \equiv E \{ (x(t_{k-1}) - \hat{x}(t_{k-1}|t_{k-1})) (x(t_{k-1}) - \hat{x}(t_{k-1}|t_{k-1}))' \},$$

and Γ is a diagonal matrix with elements Γ^{ii} ; $i = 1, 2$, given by

$$\begin{aligned}
 \Gamma^{ii}(t) &= E \{ \left(\int_{t_{k-1}}^t N^i(ds) \right)^2 \} = E \{ E \{ \left(\int_{t_{k-1}}^t N^i(ds) \right)^2 | \sigma(e_i''(t_{k-1})) \} \} \\
 &\equiv E \int_{t_{k-1}}^t \lambda^i(s) ds.
 \end{aligned} \tag{5.31}$$

Defining

$$\Theta_0^i(t) \equiv \Pr \{ e_i''(t) = 0 \}, \tag{5.32.a}$$

and

$$\Theta_1^i(t) \equiv \Pr \{ e_i''(t) = \gamma \}, \tag{5.32.b}$$

for all $i = 1, 2$, it follows from (5.17), (5.23) and (5.31) that

$$\Gamma^{ii}(t) = \delta^i \int_{t_{k-1}}^t \Theta_0^i(s) ds + \beta^i \int_{t_{k-1}}^t \Theta_1^i(s) ds, \tag{5.33}$$

for $i = 1, 2$. From the transition diagram presented in fig. 5.1, it is easy to verify that for sufficiently small $\Delta t > 0$, the total probabilities Θ_0^i and Θ_1^i ; $i = 1, 2$, satisfy the following difference equations

$$\Theta_0^i(t + \Delta t) = (1 - \delta^i \Delta t) \Theta_0^i(t) + \beta^i \Theta_1^i(t) \Delta t,$$

and

$$\Theta_1^i(t + \Delta t) = \delta^i \Theta_0^i(t) \Delta t - (1 - \beta^i \Delta t) \Theta_1^i(t),$$

for all $i = 1, 2$. Clearly, as $\Delta t \rightarrow 0$, the above equations can be written in the following differential form

$$\begin{aligned} \frac{d}{dt} \Theta_0^i(t) &= -\delta^i \Theta_0^i(t) + \beta^i \Theta_1^i(t), \\ \frac{d}{dt} \Theta_1^i(t) &= \delta^i \Theta_0^i(t) - \beta^i \Theta_1^i(t), \end{aligned} \tag{5.34}$$

for all $t \geq 0$, and $i = 1, 2$. From equations (5.30) and (5.33), it follows that the matrices $P(t|t_{k-1})$ and $\Gamma(t)$, $t > t_{k-1}$, satisfy the following differential equations

$$\frac{d}{dt} P(t|t_{k-1}) = FP(t|t_{k-1}) + P(t|t_{k-1})F' + BQB' + C\Gamma(t)C', \tag{5.35}$$

with initial condition $P(t_{k-1}|t_{k-1})$, and

$$\frac{d}{dt} \Gamma^{ii} = \delta^i \Theta_0^i(t) + \beta^i \Theta_1^i(t), \tag{5.36}$$

for all $t > t_{k-1}$ and $i = 1, 2$, where $\Theta_0^i(t)$ and $\Theta_1^i(t)$ are the solutions of (5.34). Clearly, equations (5.26), (5.34), (5.35) and (5.36) represent the behaviour of the filter between observations. We now use similar arguments as those given above to obtain a differential equation for the covariance matrix $P(t_k|t_k)$, at the observation times t_k , $1 \leq k \leq N$.

Define

$$\begin{aligned} P(t_k|t) &\equiv E\{(x(t_k) - \hat{x}(t_k|t))(x(t_k) - \hat{x}(t_k|t))'\} \\ &= E\{(\widehat{xx'})^j(t_k|t) - \hat{x}(t_k|t)\widehat{x}^j(t_k|t)\}, \quad t \in (t_{k-1}, t_k]. \end{aligned} \tag{5.37}$$

Applying Ito differential rule to the above equation, we obtain

$$dP_{ij}(t_k|t) = E\{d(\widehat{x_i x_j})(t_k|t) - \widehat{x_i}(t_k|t)d\widehat{x_j}(t_k|t) - d\widehat{x_i}(t_k|t)\widehat{x_j}(t_k|t)\} - P_{ij}(t_k|t)\widetilde{H}_{ij}P_{ij}(t_k|t)dt \quad (5.38)$$

where P_{ij} ; $i, j = 1, 2, \dots, n$, denote the elements of the matrix P and

$$\widetilde{H} \equiv H'R^{-1}H. \quad (5.39)$$

Setting $f = x_i x_j$ in equation (4.59), we obtain

$$\begin{aligned} \frac{d}{dt}(\widehat{x_i x_j})(t_k|t) &= [(\widehat{x_i x_j x'})'(t_k|t) - (\widehat{x_i x_j})(t_k|t)\widehat{x'}(t_k|t)] \times \\ &\quad \times H'R^{-1}[y(t_k) - H\widehat{x}(t_k|t)], \end{aligned} \quad (5.40)$$

for all $t \in (t_{k-1}, t_k]$. Substituting (5.26) and (5.40) into (5.38), it is easy to verify that

$$\begin{aligned} \frac{d}{dt}P_{ij}(t_k|t) &= E\{\widehat{X}(t_k|t)H'R^{-1}[y(t_k) - H\widehat{x}(t_k|t)]\} \\ &\quad - P_{ij}(t_k|t)\widetilde{H}_{ij}P_{ij}(t_k|t) \end{aligned} \quad (5.41)$$

where

$$\begin{aligned} \widehat{X}(t_k|t) &\equiv E\{(x_i(t_k) - \widehat{x_i}(t_k|t))(x_j(t_k) - \widehat{x_j}(t_k|t)) \times \\ &\quad \times (x(t_k) - \widehat{x}(t_k|t))' | \mathcal{F}_t^y\}. \end{aligned} \quad (5.41.a)$$

Since as indicated in Remark 5.1 the process $x(t), t \geq 0$, is conditionally Gaussian, given \mathcal{F}_t^y , it follows that the third order moment appearing in (5.41.a) vanishes. Using this fact, it follows from (5.39) and (5.41) that the covariance matrix $P(t_k|t), t \in (t_{k-1}, t_k]$, is governed by the following differential equation

$$\frac{d}{dt}P(t_k|t) = -P(t_k|t)H'R^{-1}HP(t_k|t), \quad t \in (t_{k-1}, t_k], \quad (5.42)$$

with initial condition $P(t_k|t_{k-1})$ given by the solution of the differential equation (5.35). Using the definition (5.37), it follows from equation (5.26.a) that the estimate $\hat{x}(t_k|t_k)$ is given by the solution of the following differential equation

$$\frac{d}{dt}\hat{x}(t_k|t) = P(t_k|t)H'R^{-1}[y(t_k) - H\hat{x}(t_k|t)], \quad (5.43)$$

for all $t \in (t_{k-1}, t_k]$, with initial condition $\hat{x}(t_k|t_{k-1})$. It is easy to verify that the solutions of the differential equations (5.42) and (5.43) are given by (see [45] P.222)

$$P(t_k|t_k) = P(t_k|t_{k-1}) - G(t_k)HP(t_k|t_{k-1}), \quad (5.44)$$

and

$$\hat{x}(t_k|t_k) = \hat{x}(t_k|t_{k-1}) + G(t_k)[y(t_k) - H\hat{x}(t_k|t_{k-1})], \quad (5.45)$$

where

$$G(t_k) \equiv P(t_k|t_{k-1})H'[HP(t_k|t_{k-1})H' + R]^{-1}, \quad (5.46)$$

and $\hat{x}(t_k|t_{k-1})$ and $P(t_k|t_{k-1})$ are the solutions of the differential equations (5.26) and (5.35), respectively. We summarize the above results in the following Theorem.

Theorem 5.1

Let the (state) process $x(t), t \geq 0$, be governed by (5.8) and the observed process $y(t_k); k = 1, 2, \dots, N$, be given by (5.14). Then the corresponding linear unbiased (minimum variance) filter consists of the following set of equations :

(i) Between Observations

$$\frac{d}{dt}\hat{x}(t|t_{k-1}) = F\hat{x}(t|t_{k-1}) + C\alpha(t), \quad (5.47)$$

$$\frac{d}{dt}P(t|t_{k-1}) = FP(t|t_{k-1}) + P(t|t_{k-1})F' + BQB' + C\Gamma(t)C', \quad (5.48)$$

where

$$\alpha^i(t) = \Pr\{e_i''(t) = 0 | \mathcal{F}_{t_{k-1}}^y\} \delta^i - \Pr\{e_i''(t) = \gamma | \mathcal{F}_{t_{k-1}}^y\} \beta^i, \quad (5.49)$$

$$\frac{d}{dt} \Gamma^{ii}(t) = \delta^i \Theta_0^i(t) + \beta^i \Theta_1^i(t), \quad (5.50)$$

and

$$\frac{d}{dt} \Theta_0^i(t) = -\delta^i \Theta_0^i(t) + \beta^i \Theta_1^i(t), \quad (5.51.a)$$

$$\frac{d}{dt} \Theta_1^i(t) = \delta^i \Theta_0^i(t) - \beta^i \Theta_1^i(t), \quad (5.51.b)$$

(ii) At observation :

$$\hat{x}(t_k | t_k) = \hat{x}(t_k | t_{k-1}) + G(t_k) [y(t_k) - H \hat{x}(t_k | t_{k-1})], \quad (5.52)$$

$$P(t_k | t_k) = P(t_k | t_{k-1}) - G(t_k) H P(t_k | t_{k-1}), \quad (5.53)$$

where

$$G(t_k) \equiv P(t_k | t_{k-1}) H' [H P(t_k | t_{k-1}) H' + R]^{-1}. \quad (5.54)$$

Remark 5.2

Note that the usual Palman filter, for the case where $x(t), t \geq 0$, is only driven by the Wiener process W , can be obtained from the above filter equations by setting $C \equiv 0$.

In the next section we present a numerical algorithm on the basis of which the estimate $\hat{x}(t_k | t_k)$ and the corresponding covariance matrix $P(t_k | t_k)$ can be computed. Further, we present a numerical example to illustrate the expected behaviour of our proposed filter in comparison to the usual Palman filter.

5.4 Computational Algorithm and Examples

In this section we develop a numerical algorithm for generating the random processes $x(t)$ and $y(t), t \geq 0$, given by (5.8) and (5.14), respectively. Then on the basis of Theorem 5.1, we compute the estimate $\hat{x}(t_k|t_k)$ and the corresponding covariance matrix $P(t_k|t_k)$. In this algorithm, we have used the assumption that the random times during which the process $e''(t), t \geq 0$, (see equation (5.6)) resides in the states '0' or ' γ ' are exponentially distributed, since the transition rates δ and β are time invariant. Further, we assume also that the elements of the covariance matrices Q and R are very small compared to the jump size γ . With this assumption, the conditional probabilities appearing in equation (5.49), which are difficult to obtain, can be computed by observing the sudden changes in the process y .

The major steps in this algorithm may be summarized as follows :

Step 1

Given m, Q, R, β and δ generate the random processes W, N , and V , using the IMSL subroutines GGNML and GGEXT.

Step 2

Given k, T_c, T, γ and $x(t_{k-1})$, solve the stochastic differential equation (5.8) using Runge-Putta method to obtain $x(t_k)$.

Step 3

Given Ψ_M, Ψ_{S_1} and Ψ_{S_2} , compute $y(t_k)$ using equation (5.14).

Step 4

If $|y^i(t_{k-2}) - y^i(t_{k-1})| \geq r; i = 1, 2$, where $r > 0$ is a preassigned threshold, set $\alpha^i(t) = -\beta^i$ otherwise set $\alpha^i(t) = \delta^i, t_{k-1} \leq t < t_k; i = 1, 2$.

Step 5

Given $\hat{x}(t_{k-1}|t_{k-1})$ solve (5.47) to obtain $\hat{x}(t_k|t_{k-1})$.

Step 6

Given $P(t_{k-1}|t_{k-1})$, compute $P(t_k|t_{k-1})$ using (5.48), (5.50), (5.51.a) and (5.51.b).

Step 7

Compute the gain matrix $G(t_k)$ using equation (5.54).

Step 8

Compute $\hat{x}(t_k|t_k)$ and the corresponding covariance matrix $P(t_k|t_k)$, using equations (5.52) and (5.53), respectively.

Step 9

If $k \leq m$, set $k = k + 1$ and go to Step 2, otherwise stop .

Based on the above algorithm, we now present a numerical example to illustrate the effectiveness of the proposed filter given in Theorem 5.1.

Numerical Example

Let the parameters corresponding to the navigation system (5.8) and (5.9) be given by

$$\beta^1 = \beta^2 = 0.7$$

$$\delta^1 = \delta^2 = 0.05$$

$$Q_{11} = Q_{22} = 0.09 \text{ m}$$

$$Q_{33} = Q_{44} = 25.0 \text{ m}$$

$$R_{11} = R_{22} = 140. \text{ m}$$

$$\Psi_M = 60^\circ$$

$$\Psi_{S_1} = 180^\circ$$

$$\Psi_{S_2} = 270^\circ$$

$$\gamma = 3000 \text{ m}$$

and the matrix B is given by (5.12).

Using the above parameters and the proposed algorithm, we have generated different sample paths for the processes $x(t_k)$ and $y(t_k)$, $k = 1, 2, \dots, m$, and computed the corresponding estimate $\hat{x}(t_k|t_k)$ using the proposed filter (see Theorem 5.1) and the conventional Palman filter (see Remark 5.2). The numerical results given in figs. 5.2-5.34, present three sample paths for the actual state $x(t_k)$ and the observed process $y(t_k)$, $k = 1, 2, \dots, m$, as well as the corresponding estimate $\hat{x}(t_k|t_k)$ and r.m.s. for both proposed and Palman filters. From these results, one observes that the actual state and the estimated state obtained by the proposed

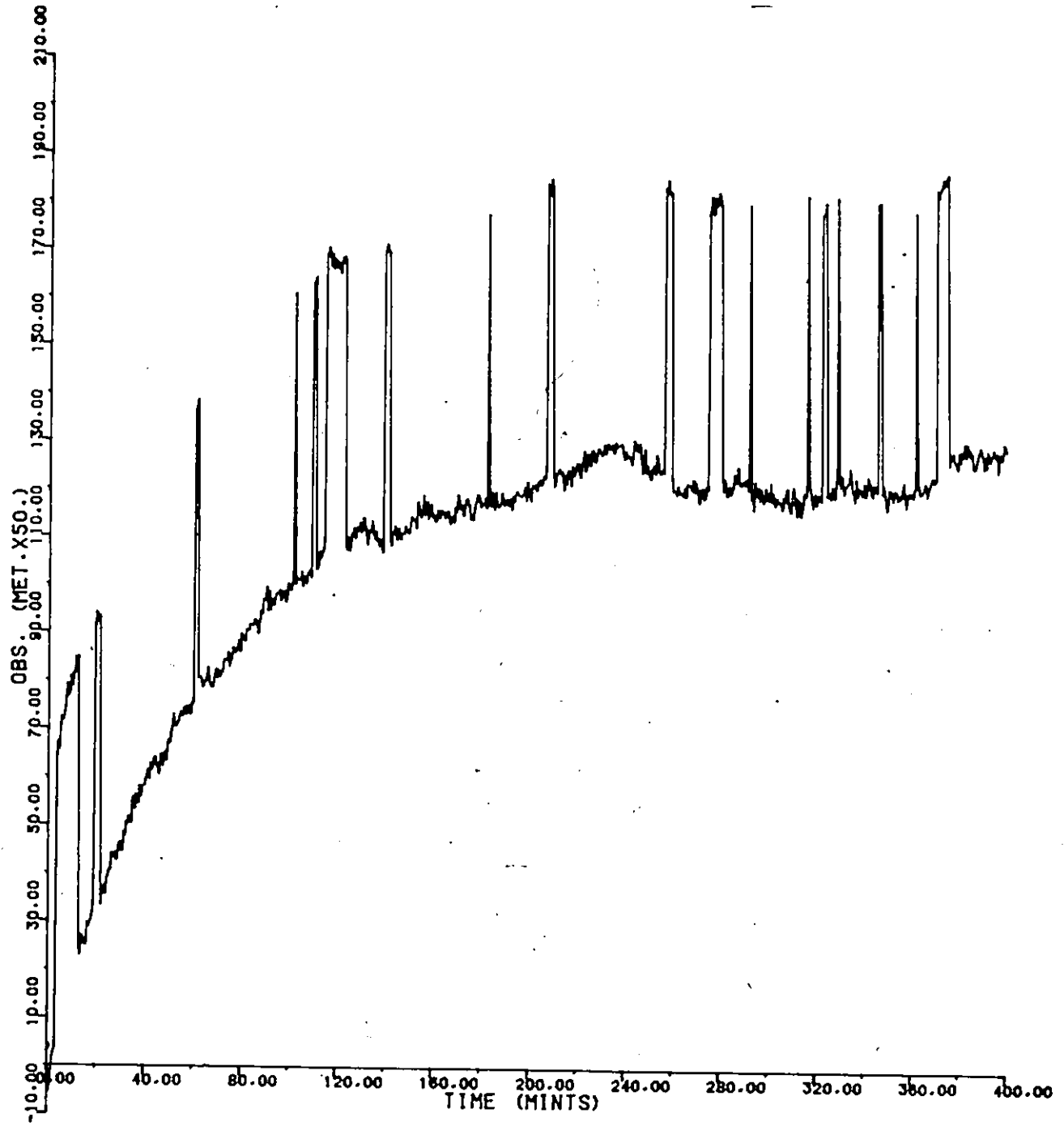


Fig. 5.2 Observed Process (sample path 1)

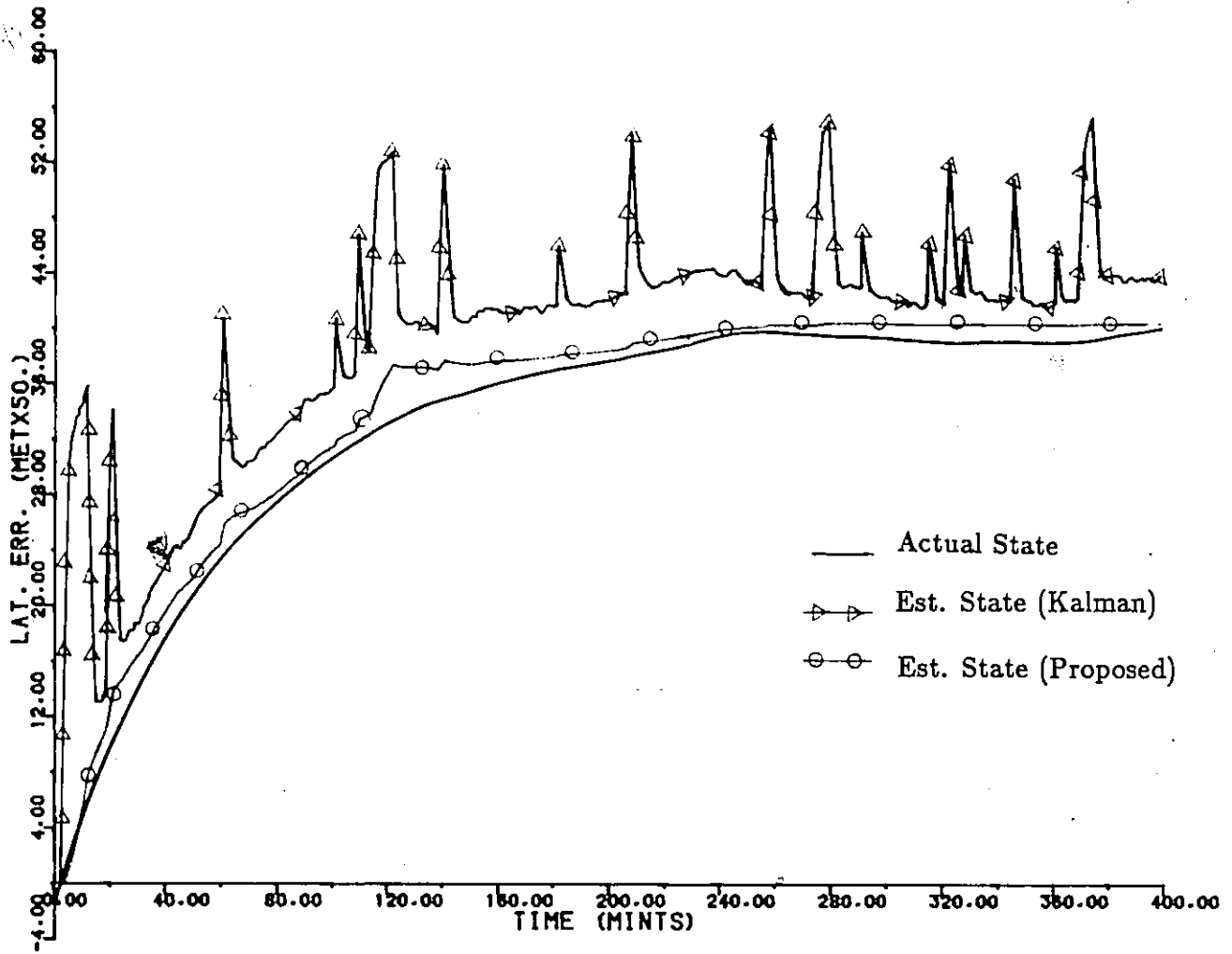


Fig. 5.3 Latitude Error (sample path 1)

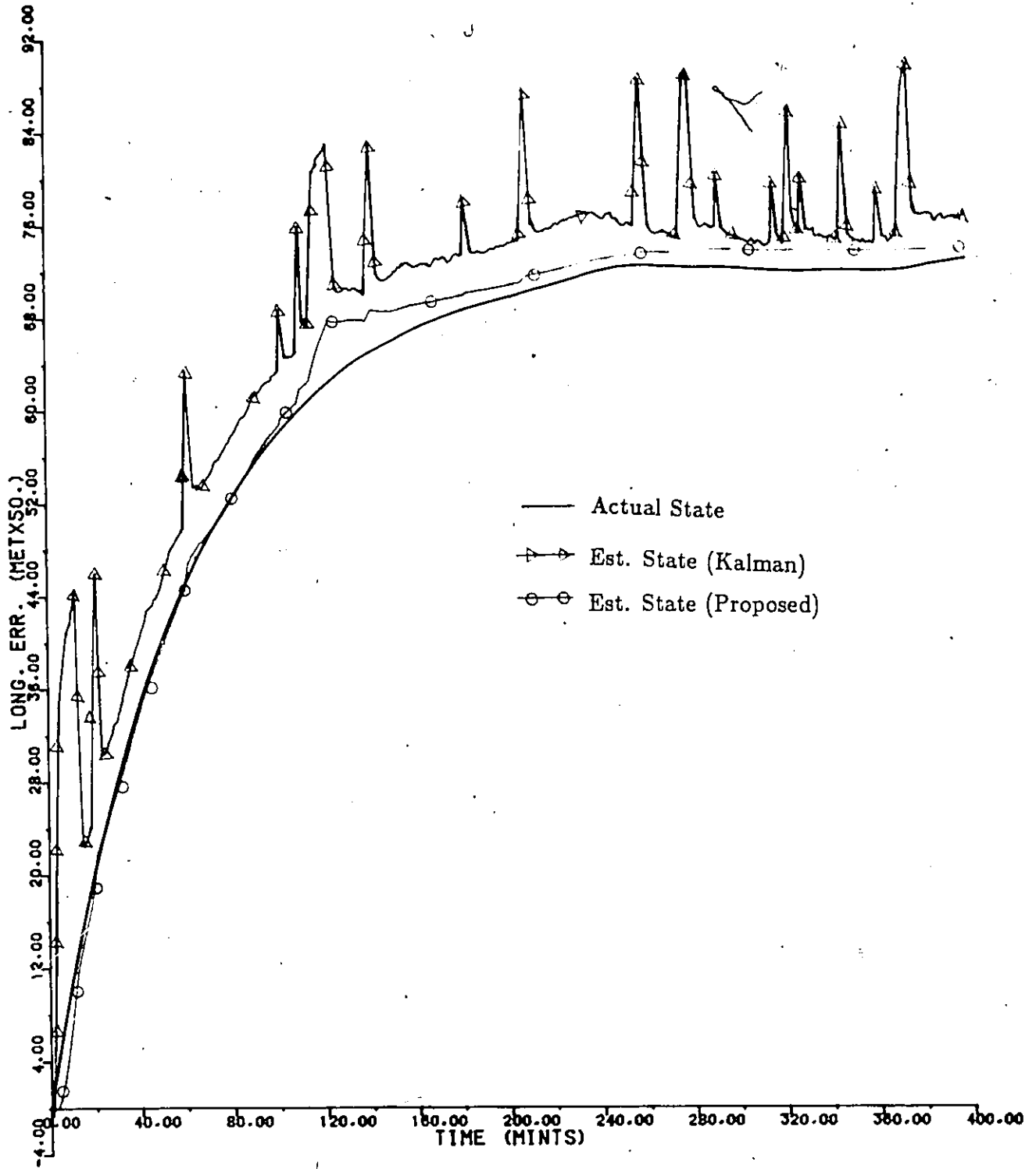


Fig. 5.4 Longitude Error (sample path 1)

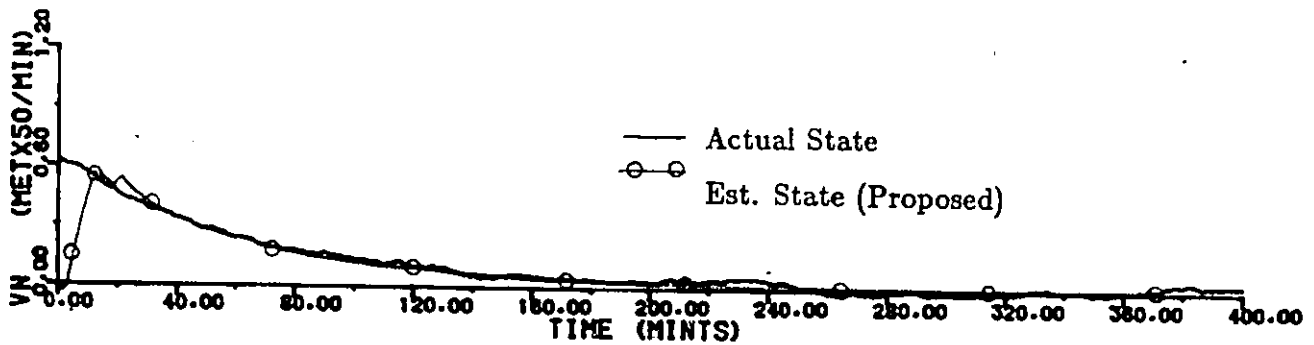
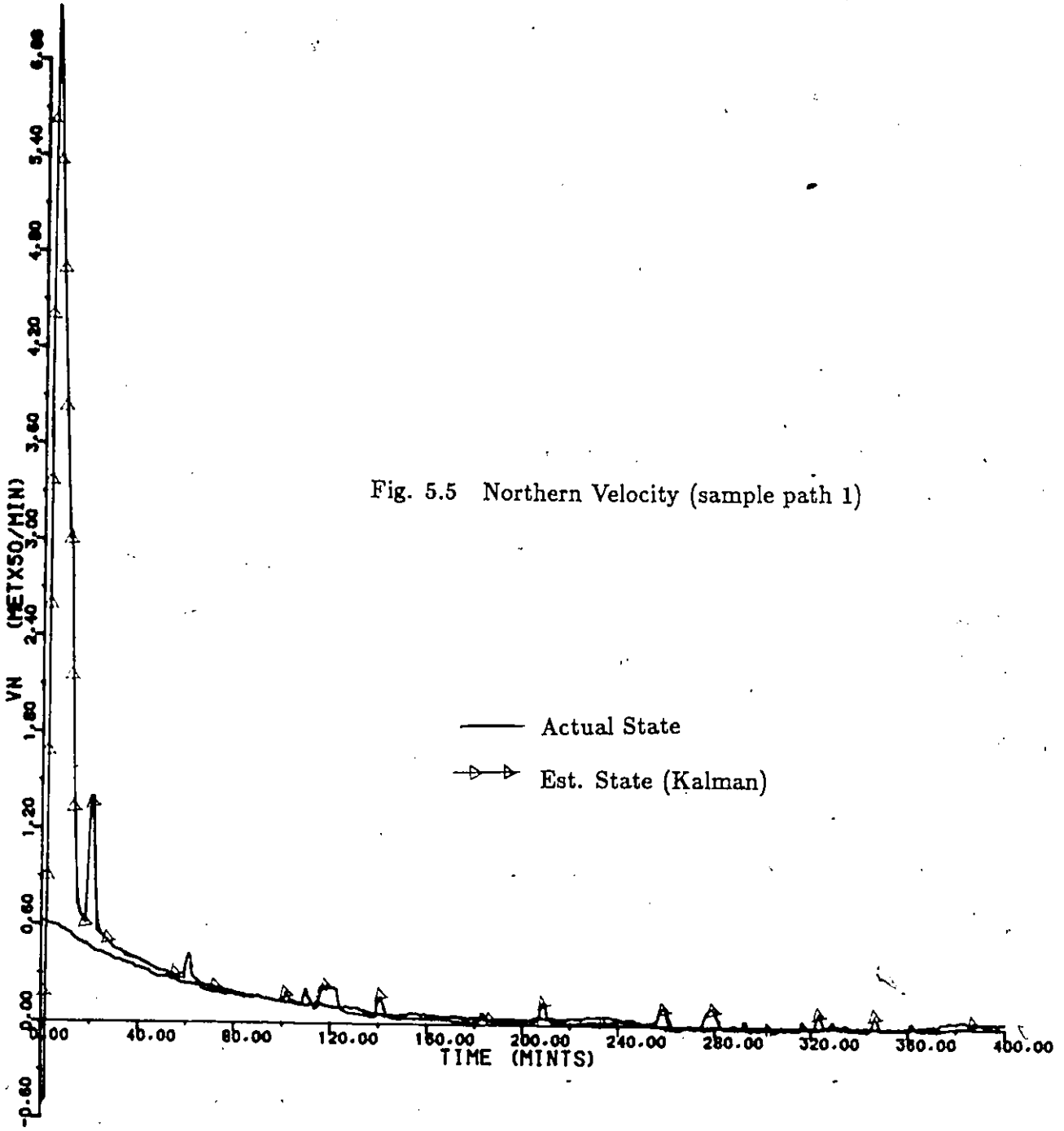
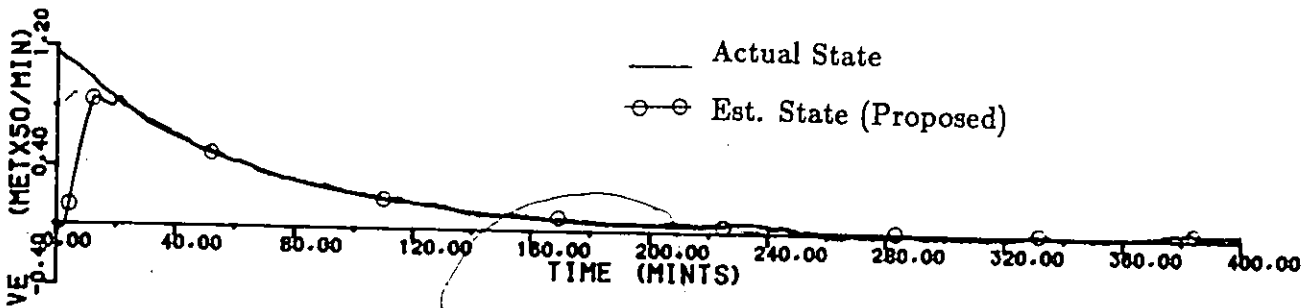
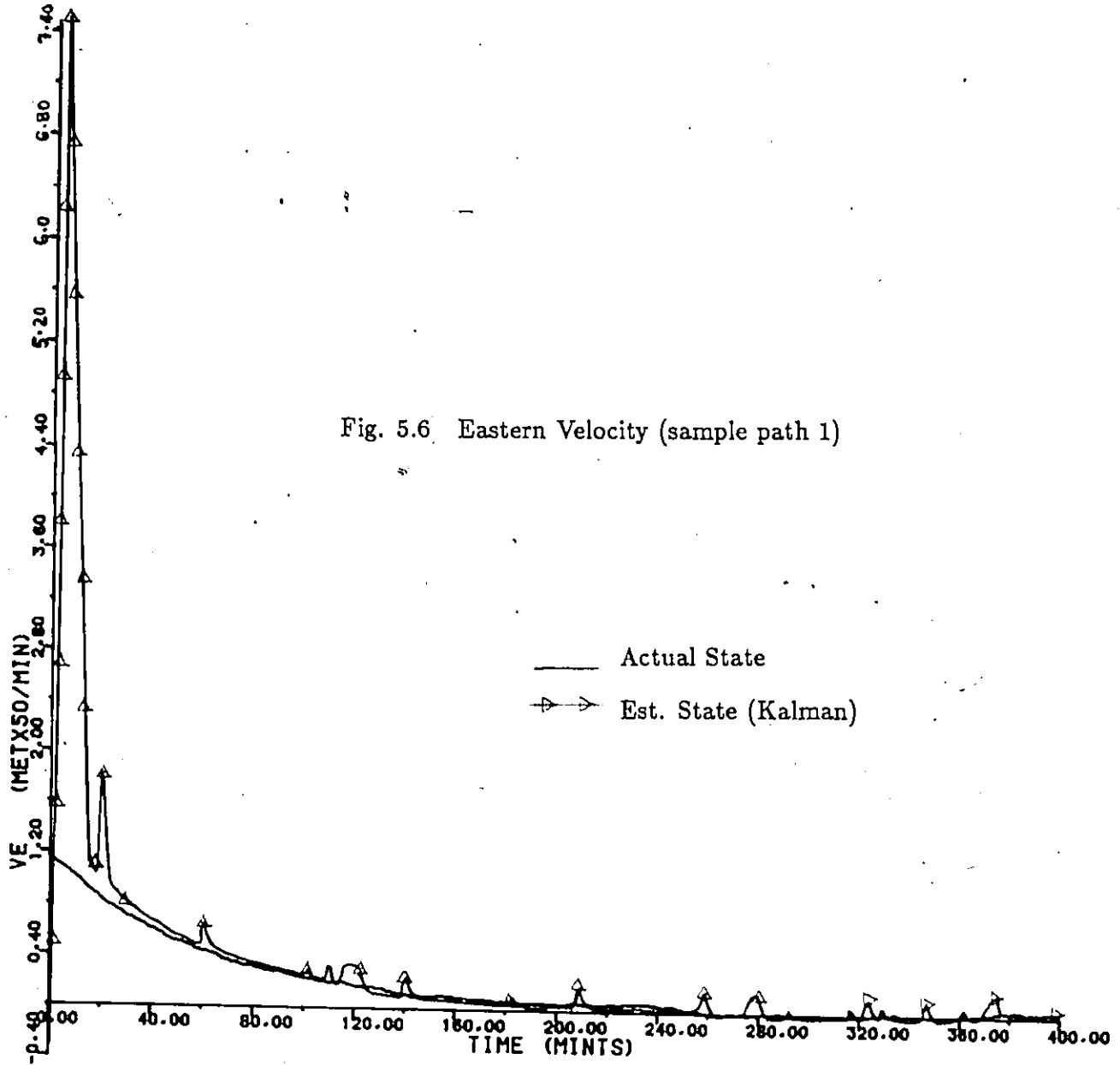


Fig. 5.6 Eastern Velocity (sample path 1)



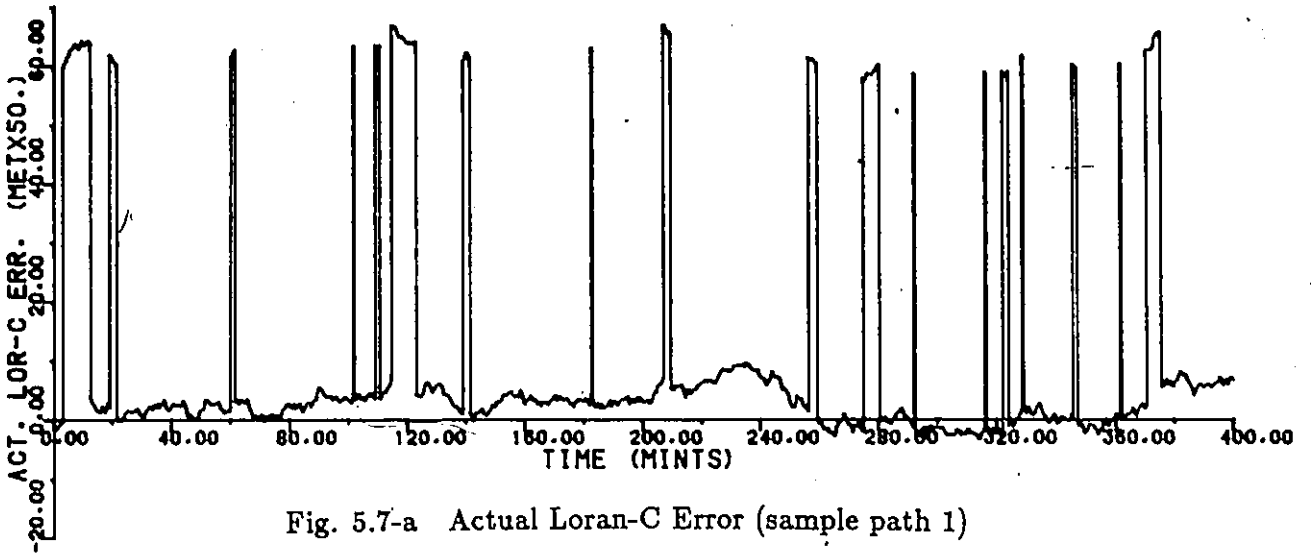


Fig. 5.7-a Actual Loran-C Error (sample path 1)

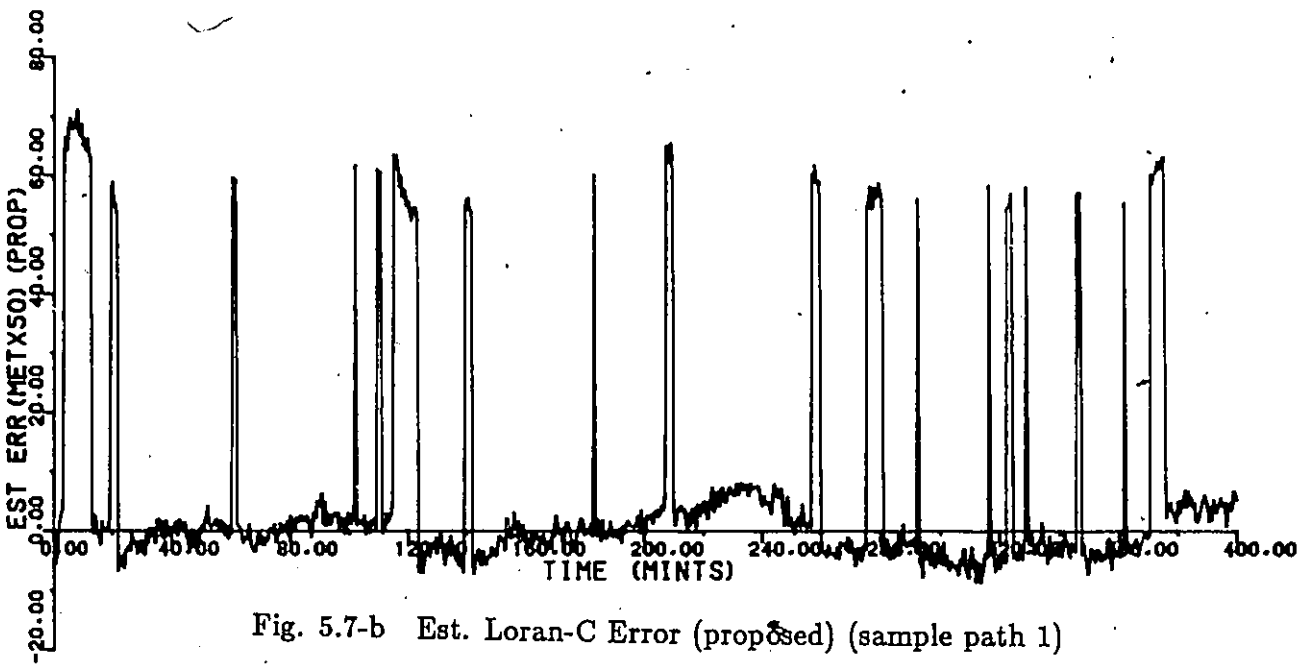


Fig. 5.7-b Est. Loran-C Error (proposed) (sample path 1)

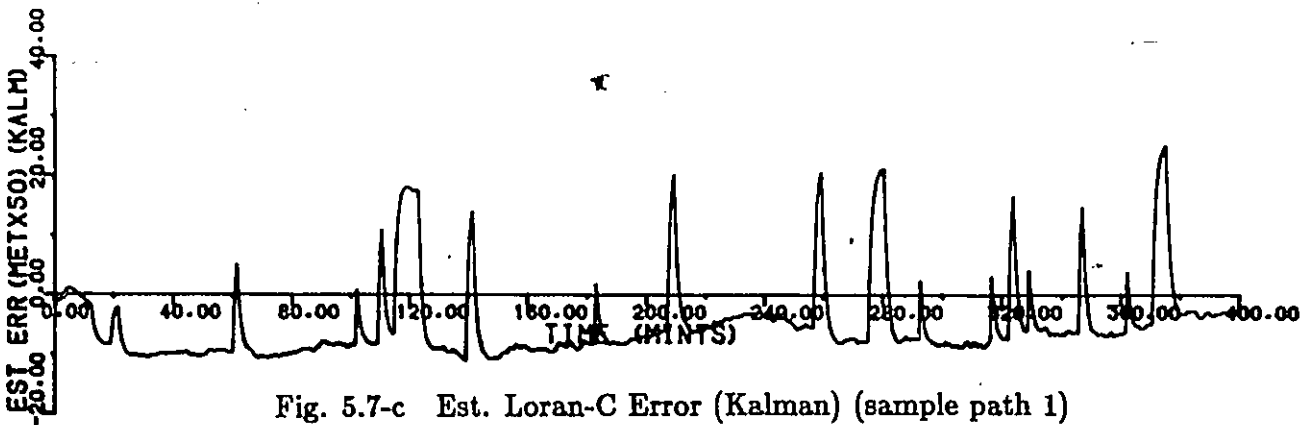


Fig. 5.7-c Est. Loran-C Error (Kalman) (sample path 1)

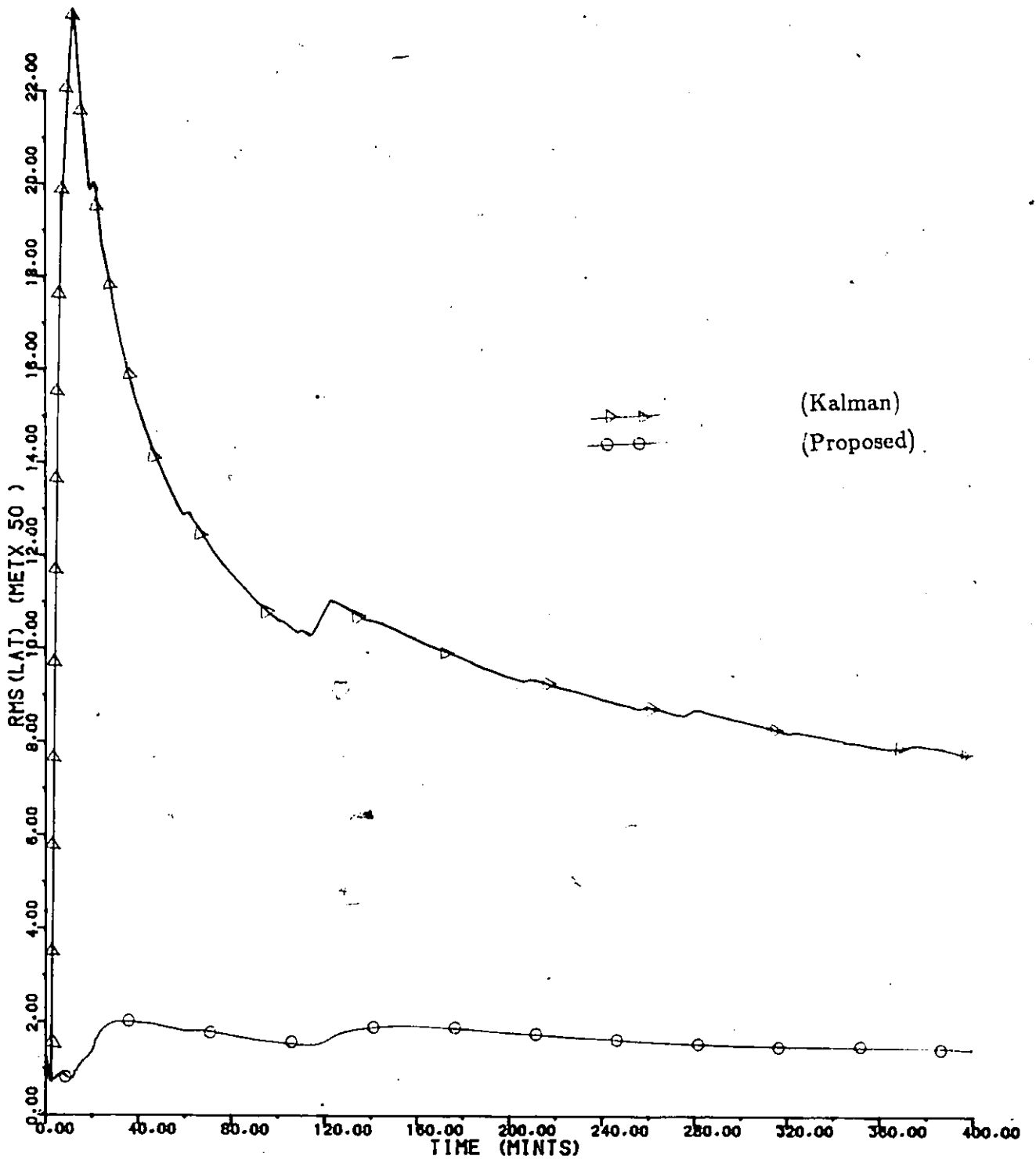


Fig. 5.8 R.M.S. for Latitude (sample path 1)

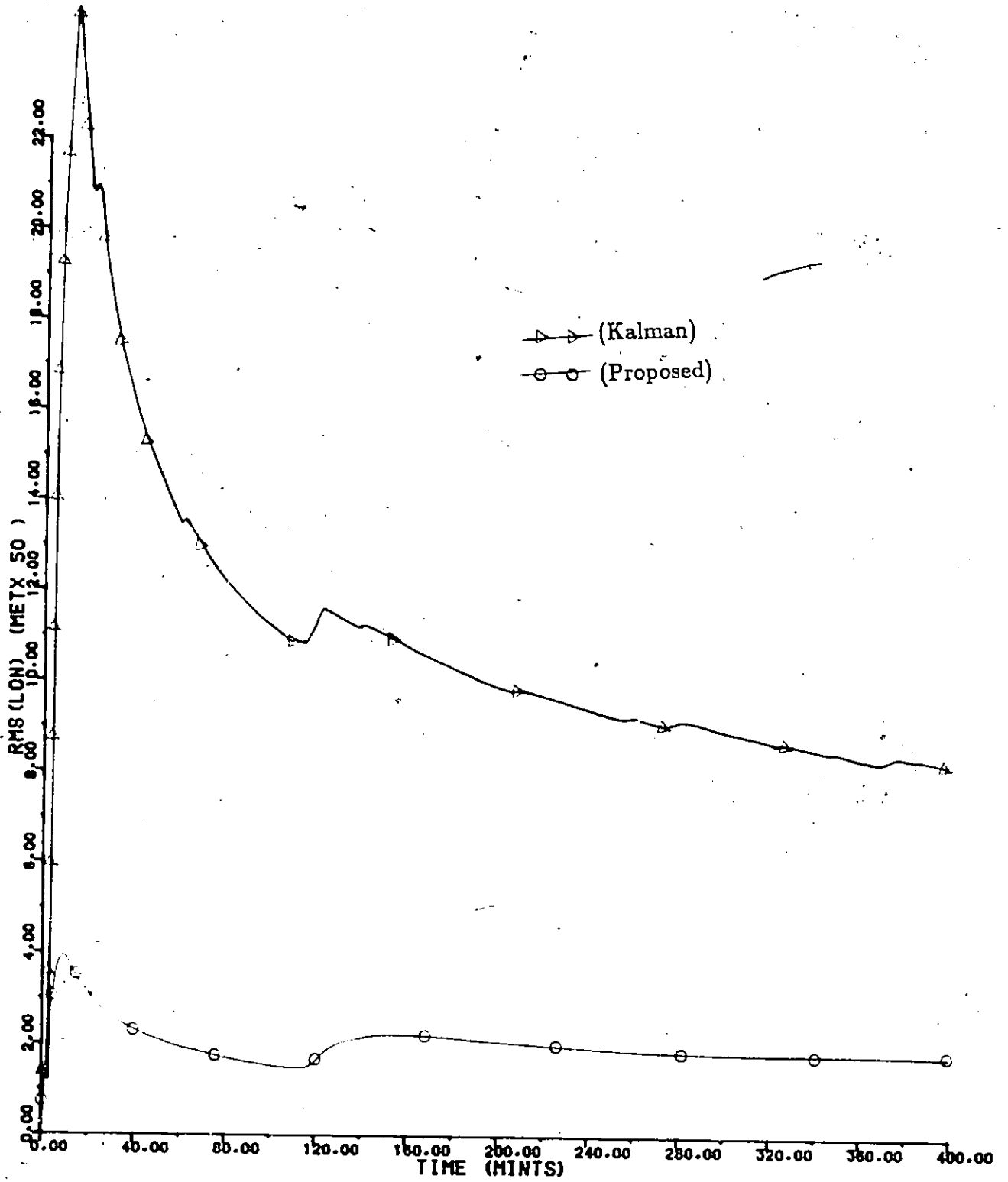


Fig. 5.9 R.M.S. for Longitude (sample path 1)

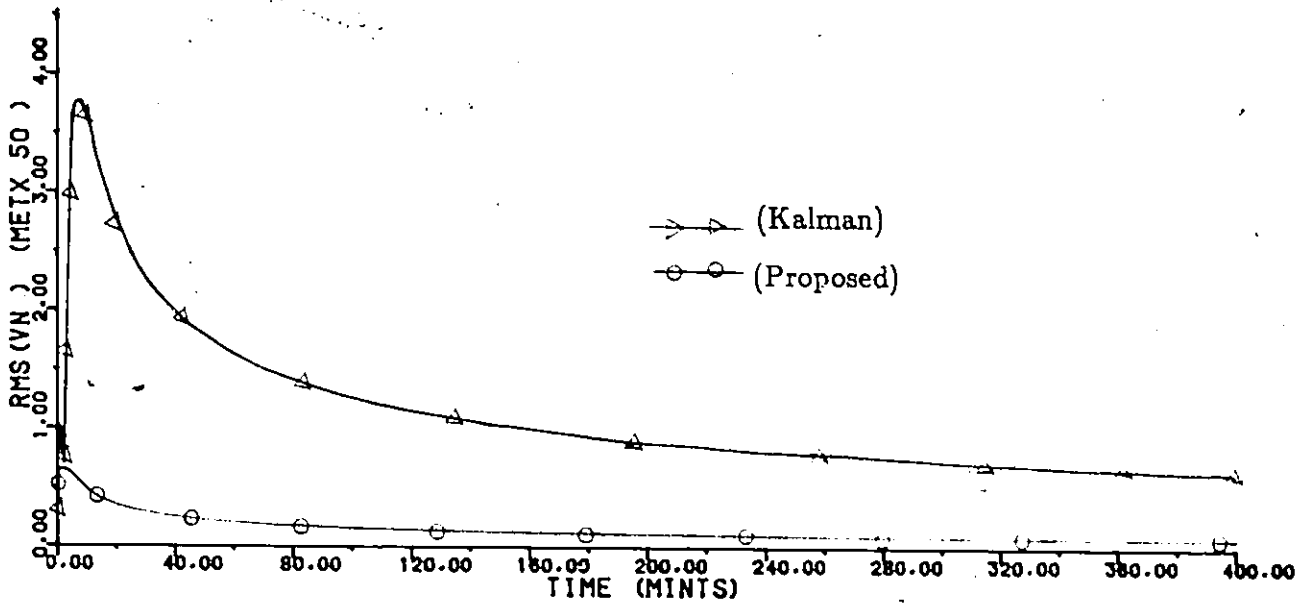


Fig. 5.10 R.M.S. for Northern Velocity (sample path 1)

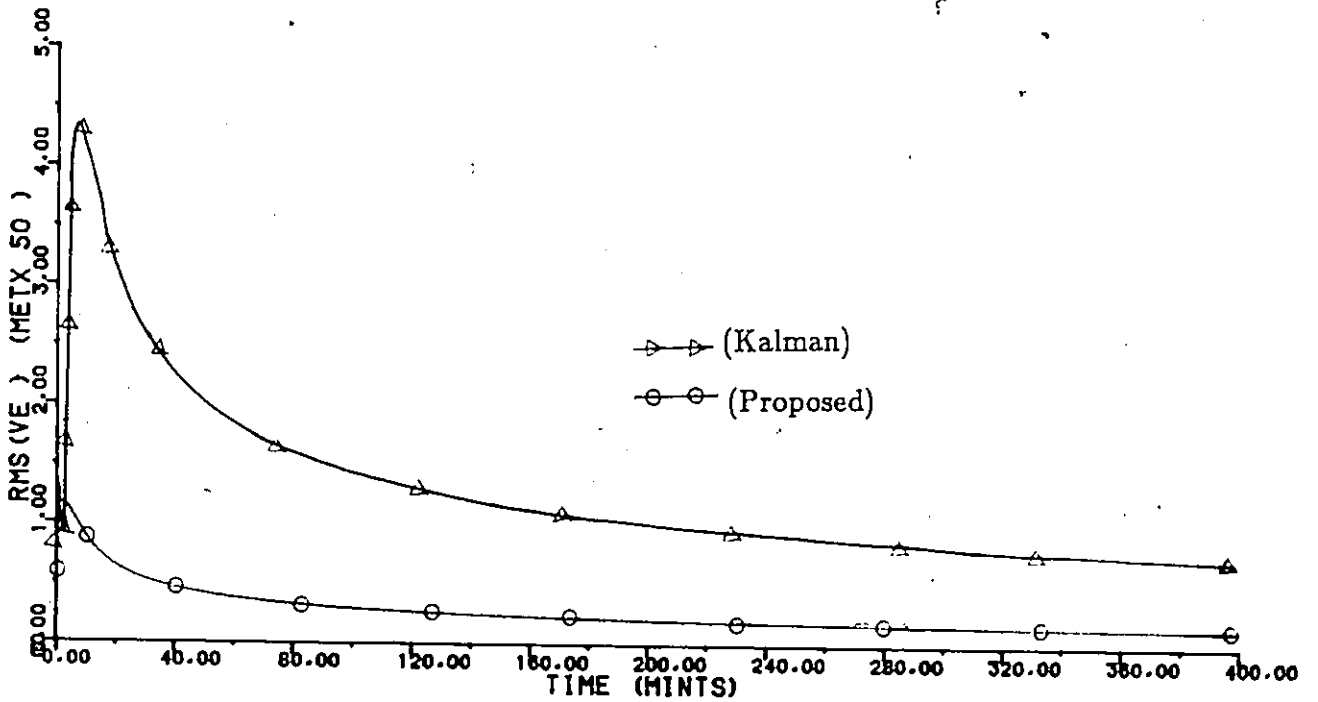
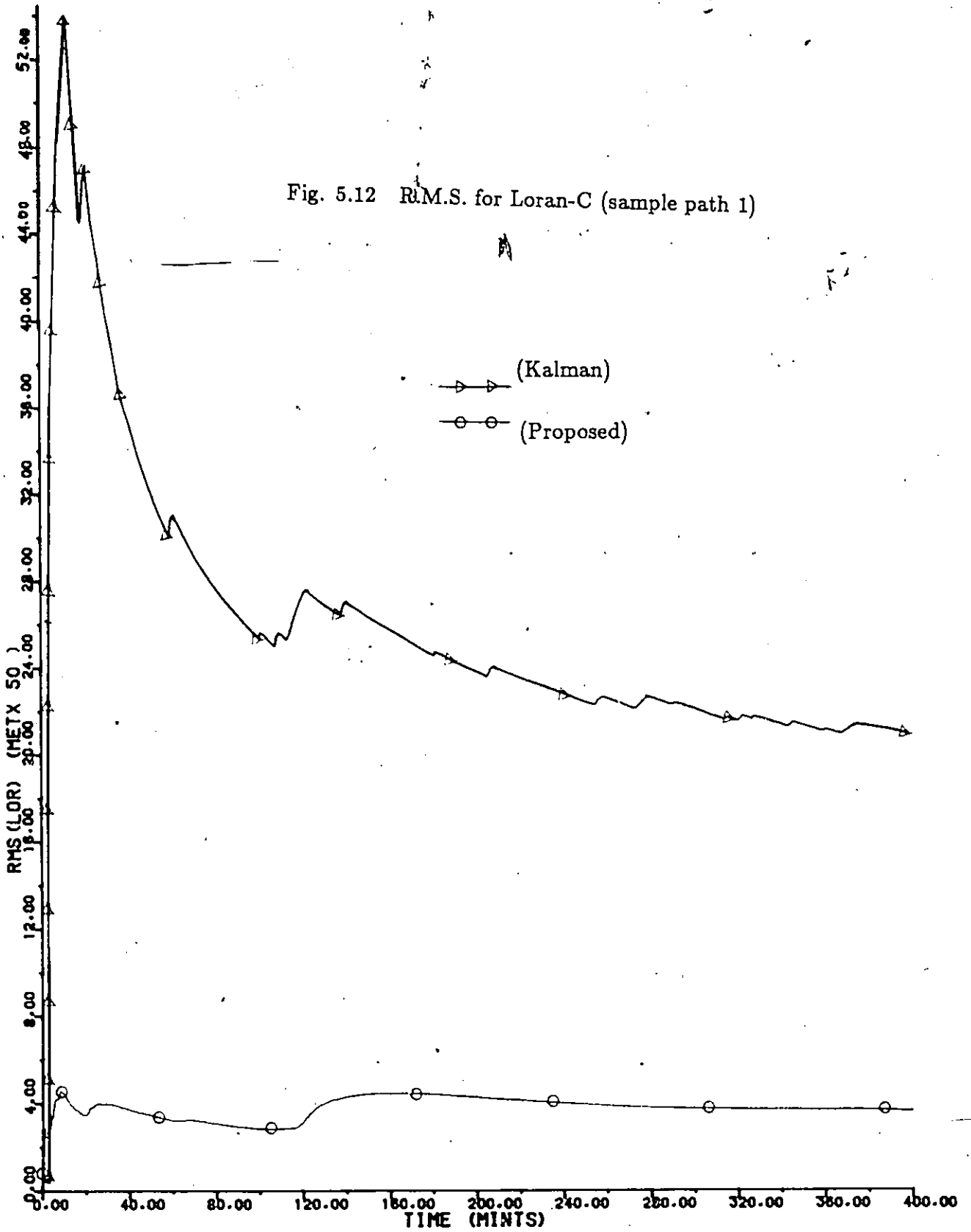


Fig. 5.11 R.M.S. for Eastern Velocity (sample path 1)



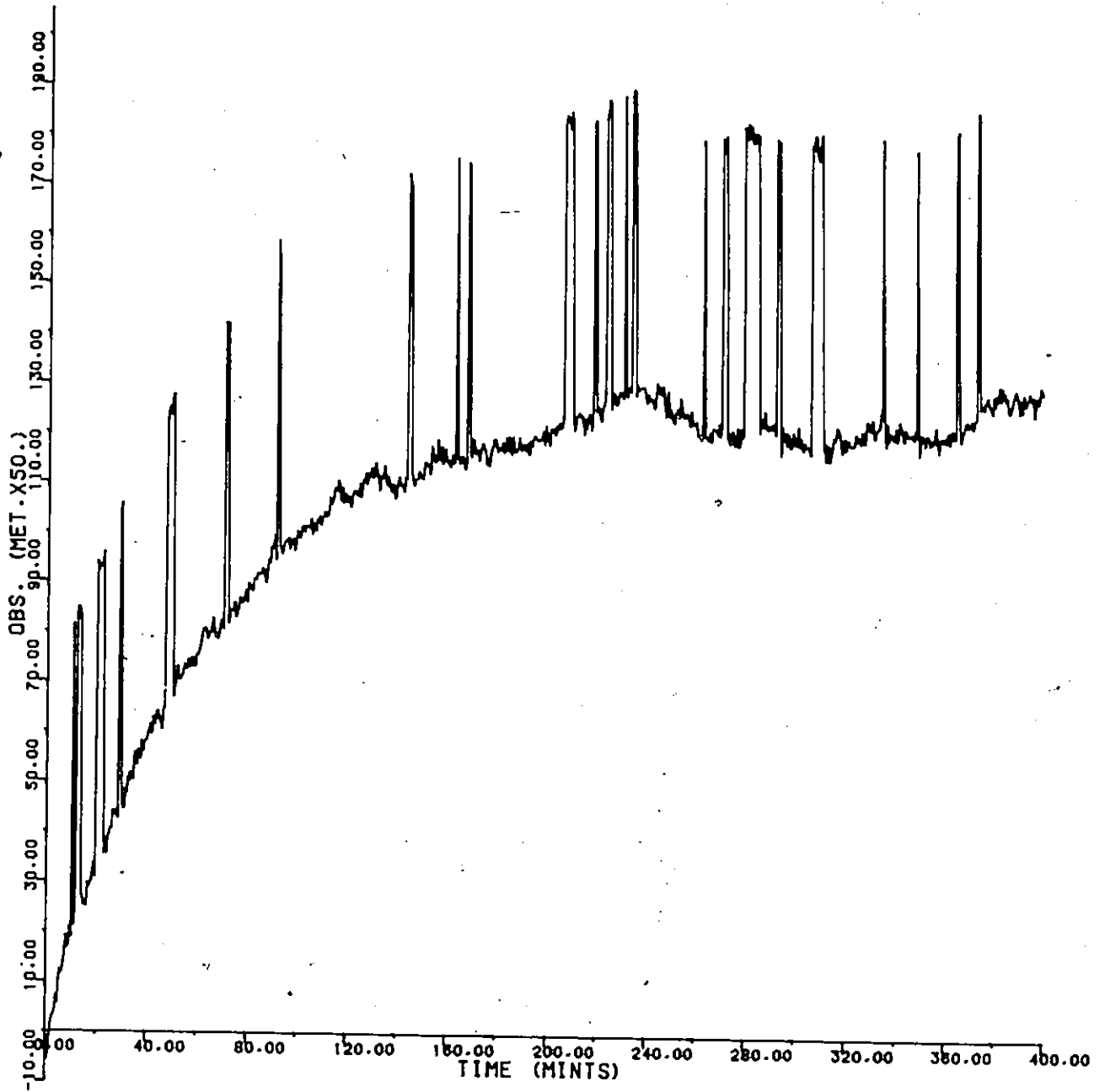


Fig. 5.13 Observed_Process (sample path 2)

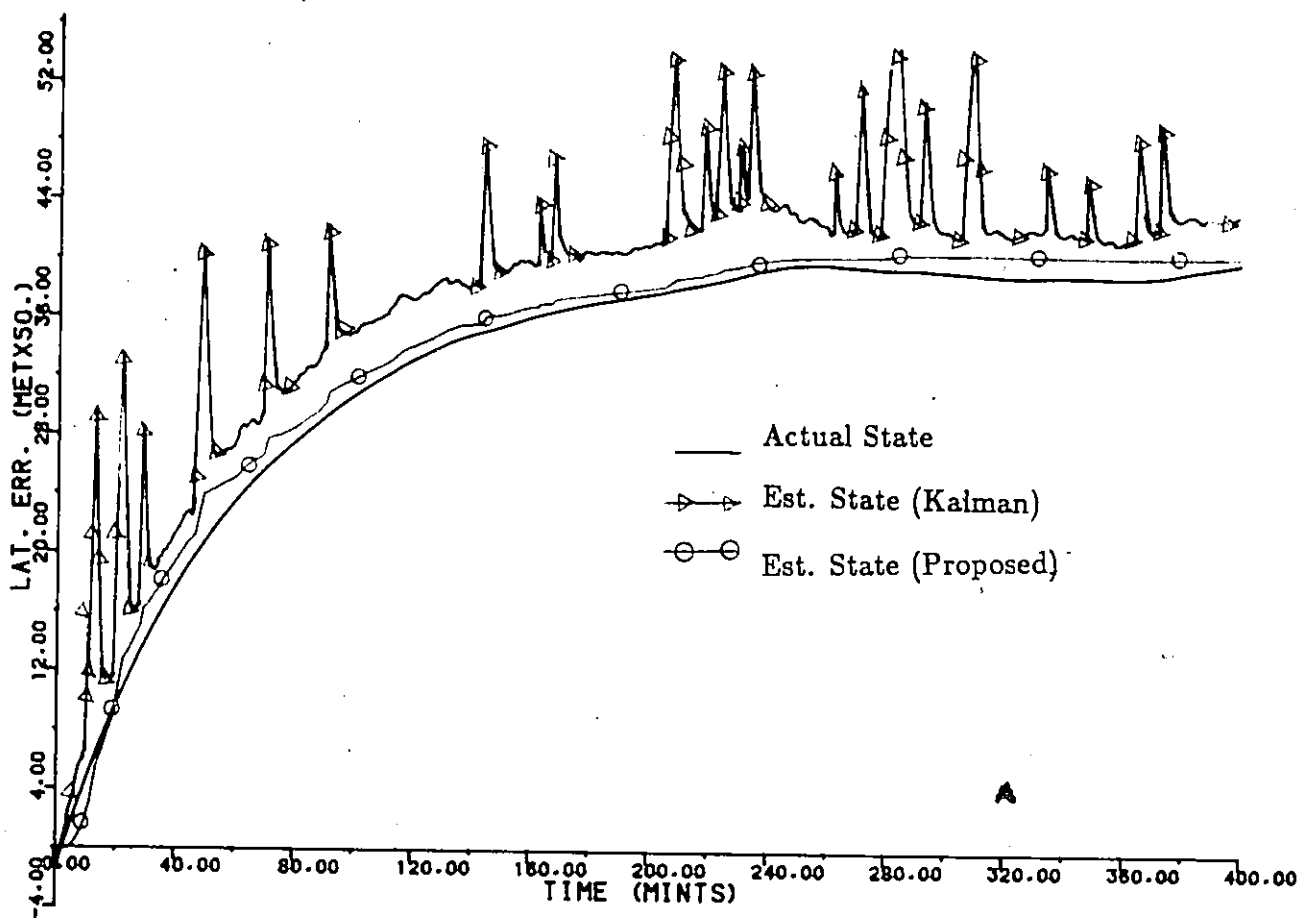


Fig. 5.14 Latitude Error (sample path 2)

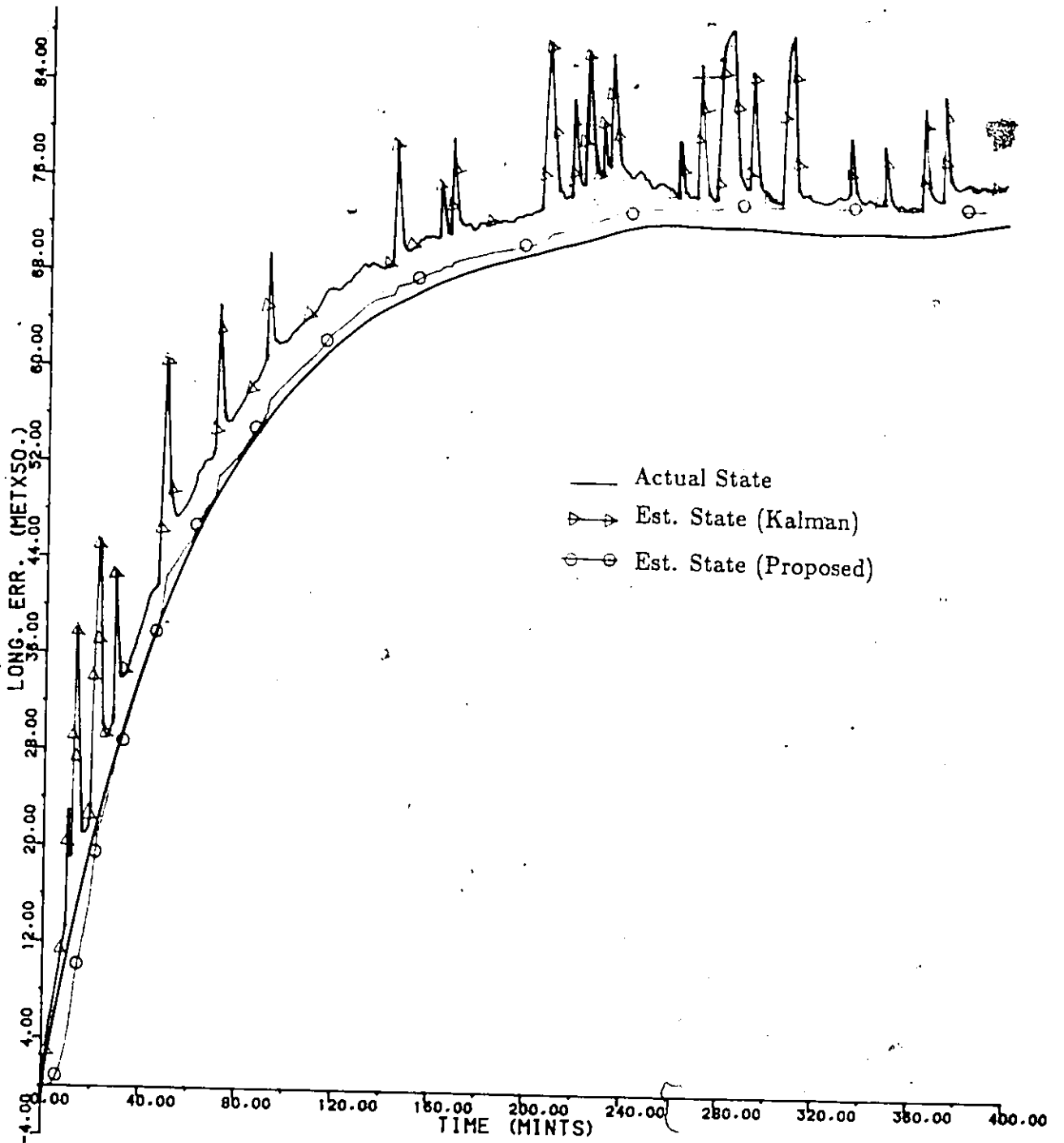


Fig. 5.15 Longitude Error (sample path 2)

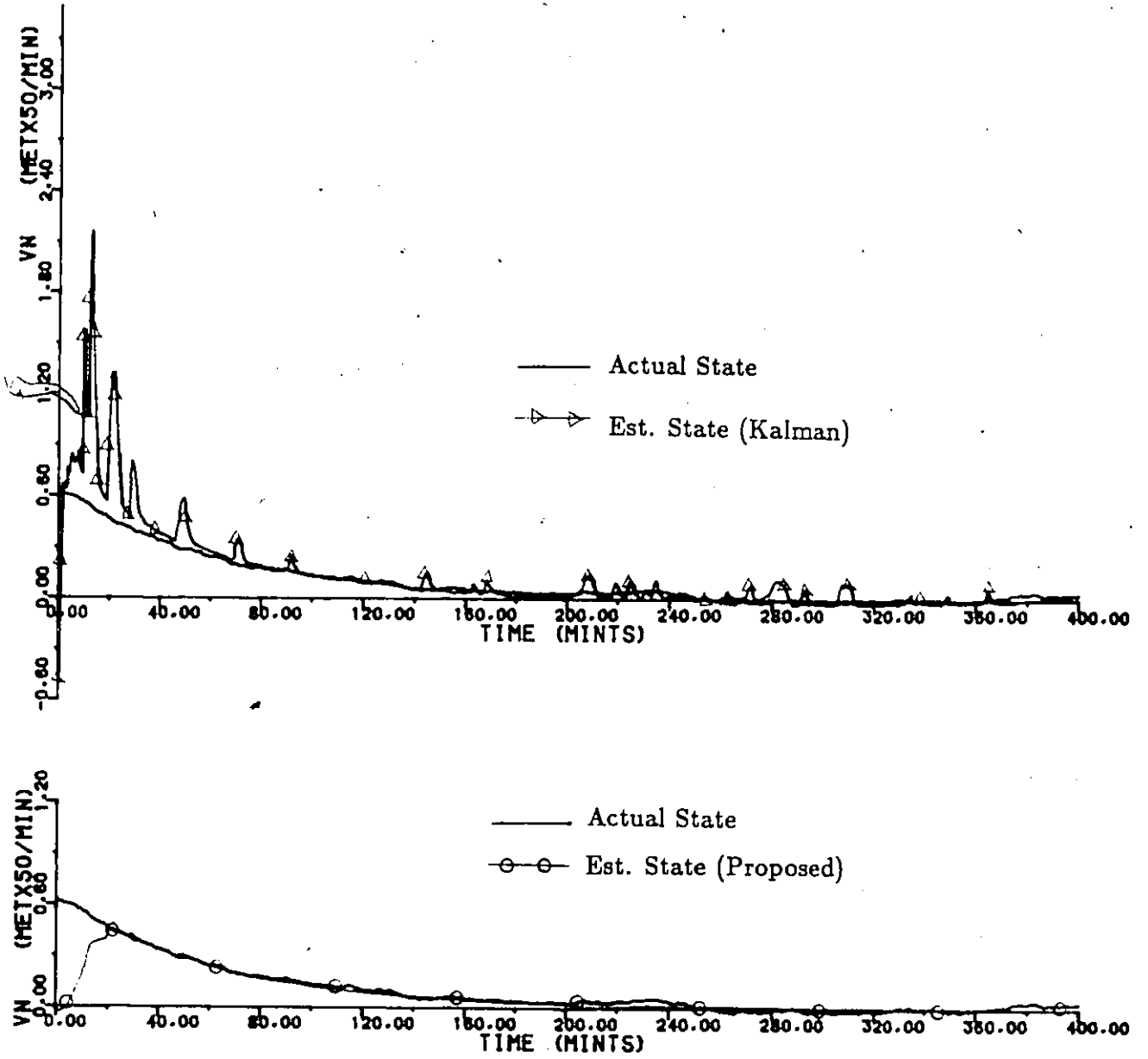


Fig. 5.16 Northern Velocity (sample path 2)

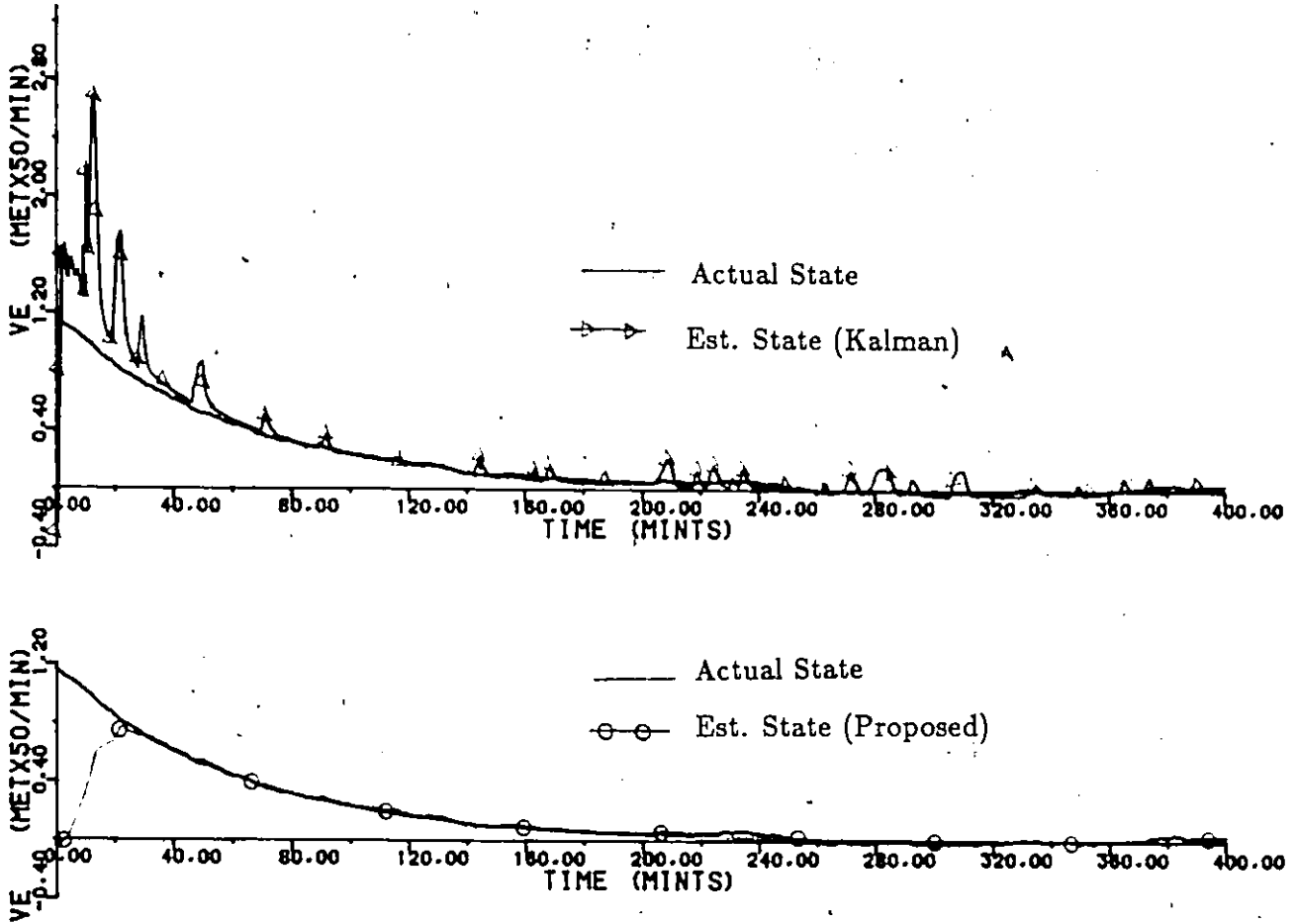


Fig. 5.17 Eastern Velocity (sample path 2)

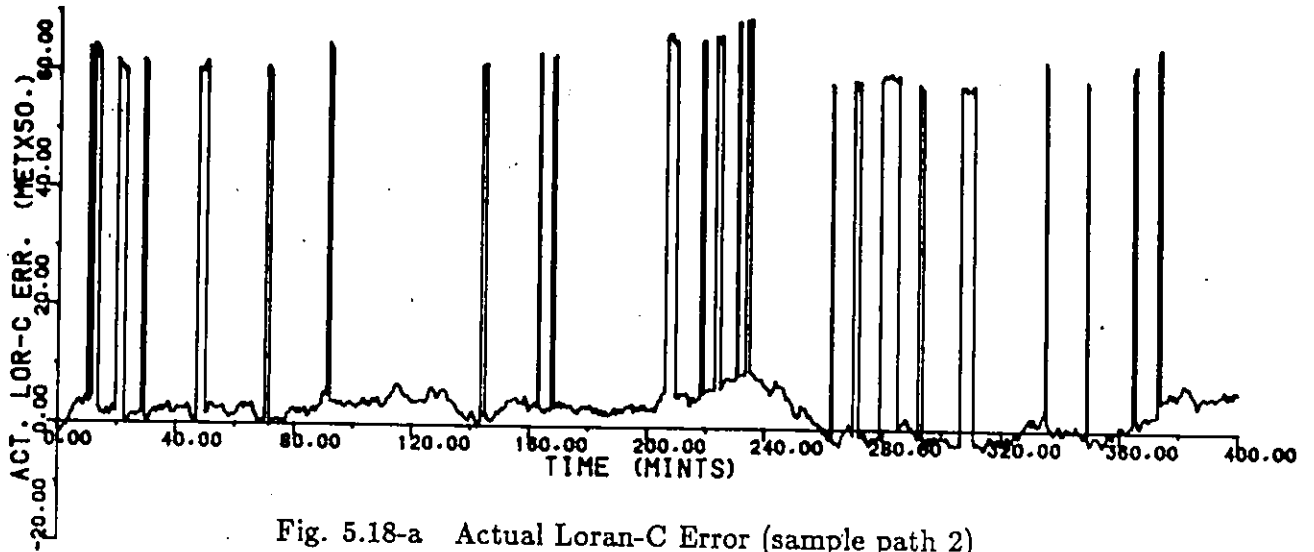


Fig. 5.18-a Actual Loran-C Error (sample path 2)

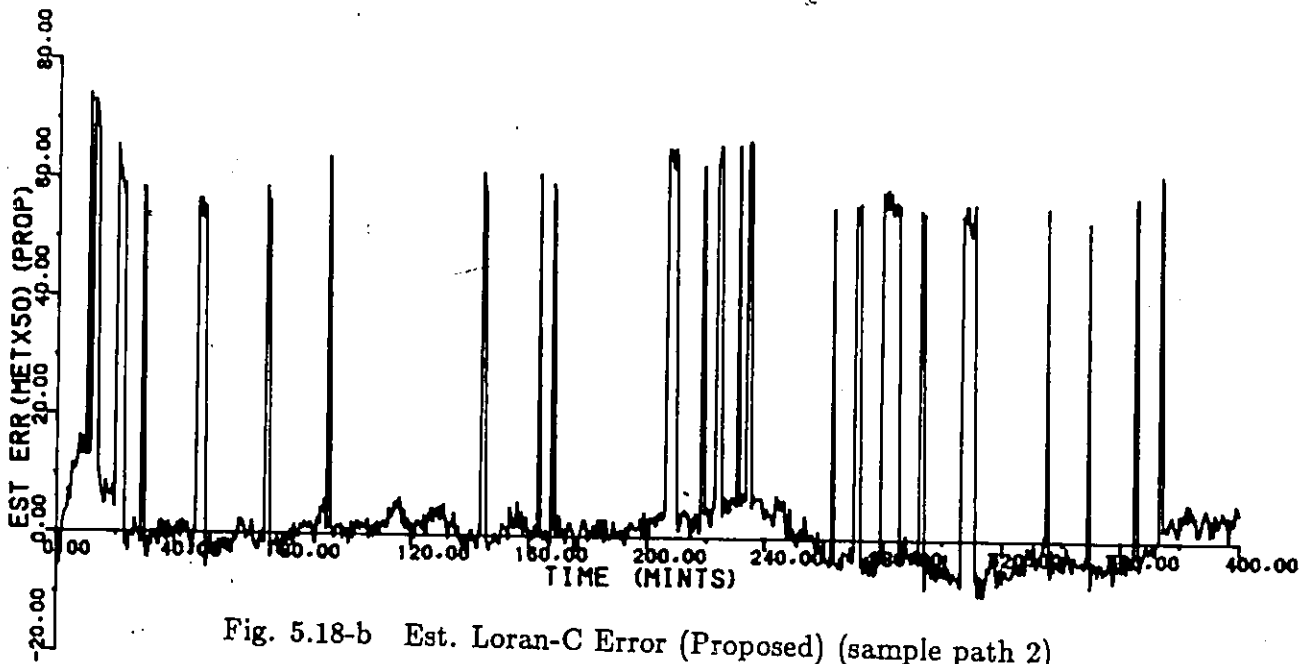


Fig. 5.18-b Est. Loran-C Error (Proposed) (sample path 2)

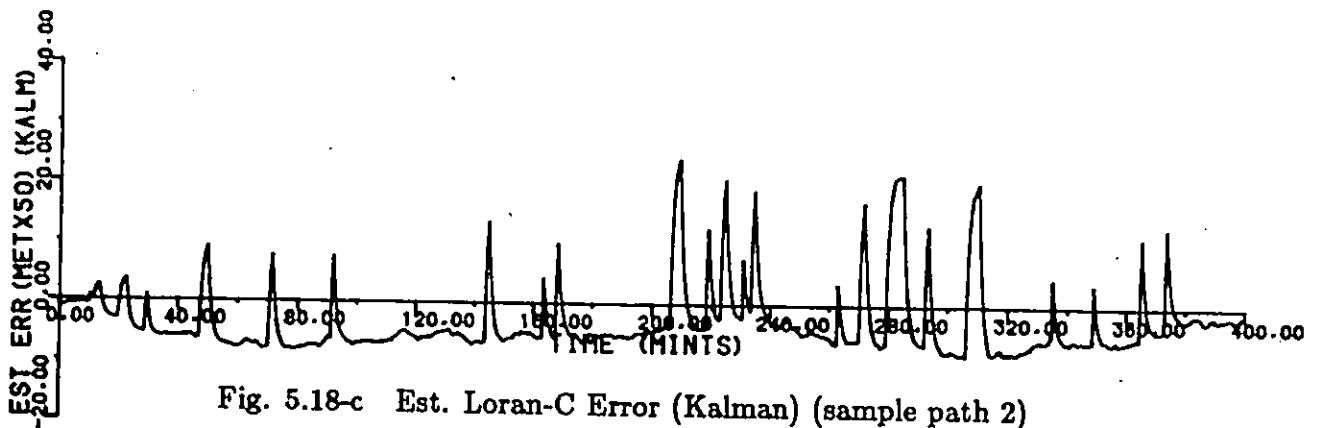


Fig. 5.18-c Est. Loran-C Error (Kalman) (sample path 2)

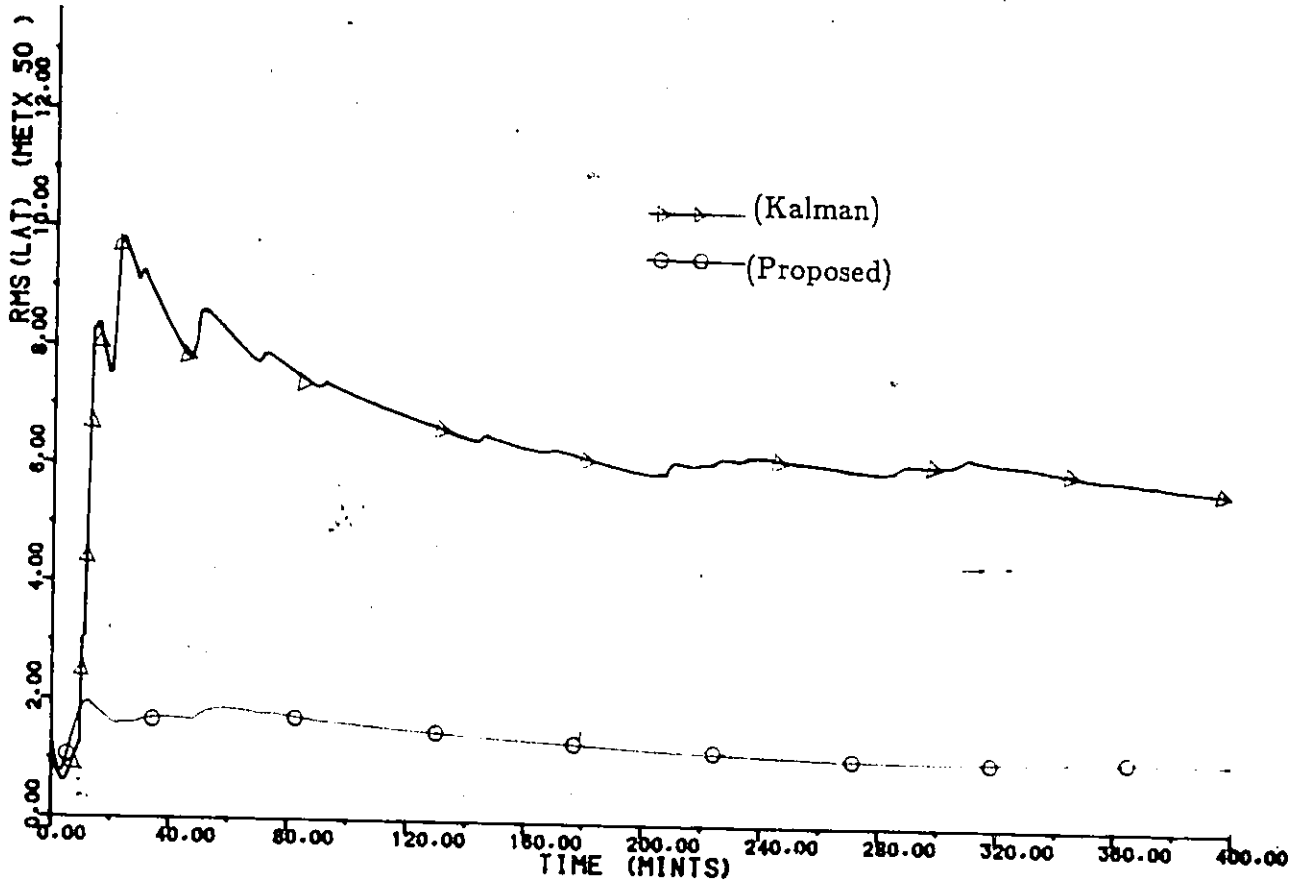


Fig. 5.19 R.M.S. for Latitude (sample path 2)

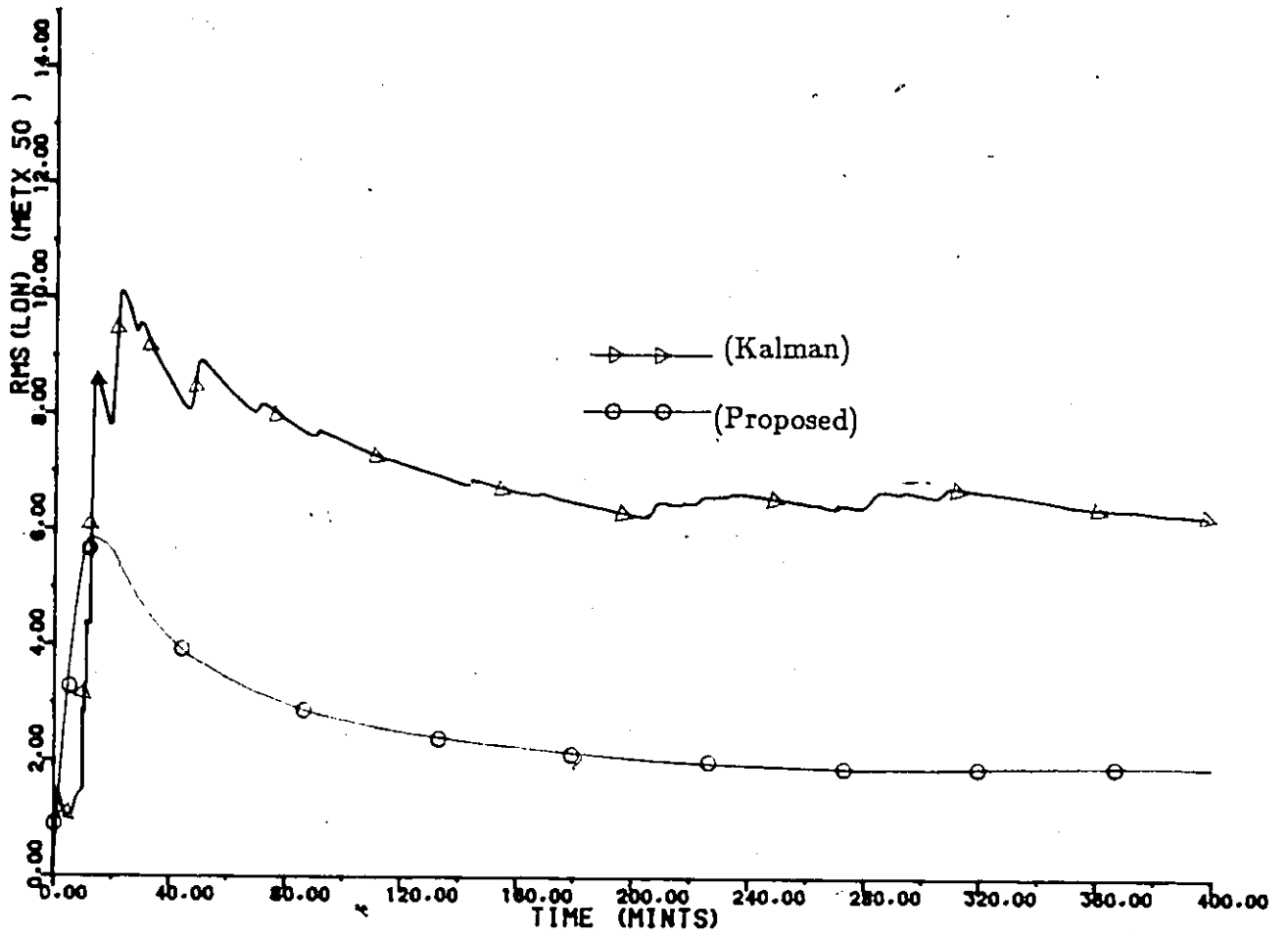


Fig. 5.20 R.M.S. for Longitude (sample path 2)

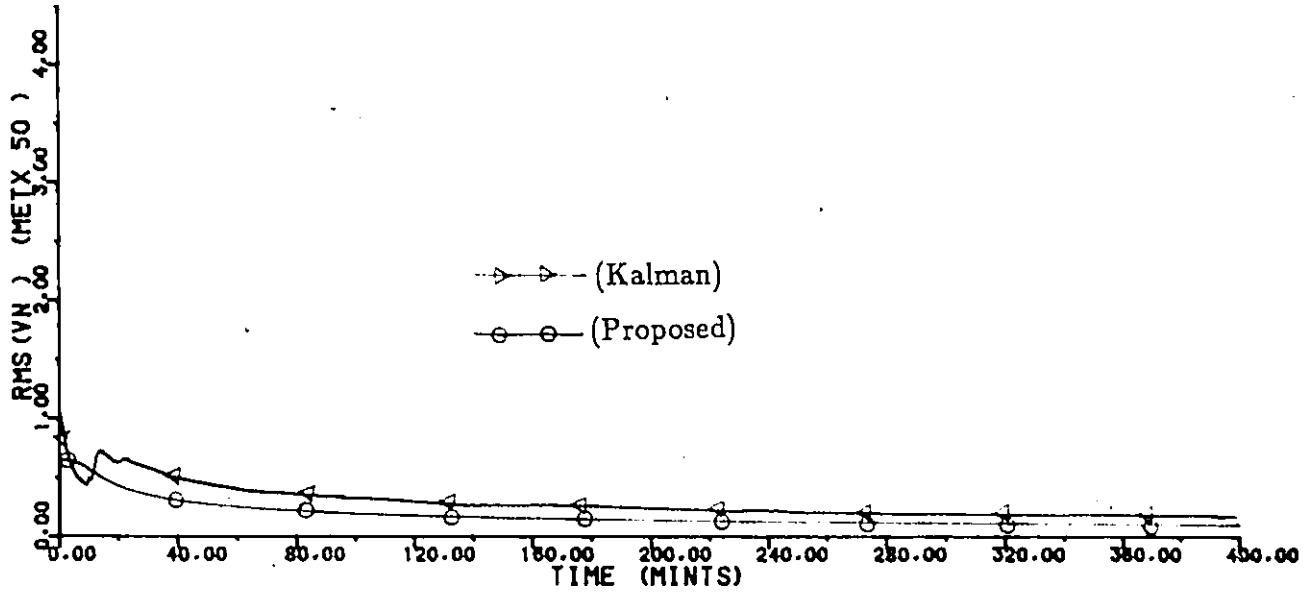


Fig. 5.21 R.M.S. for Northern Velocity (sample path 2)

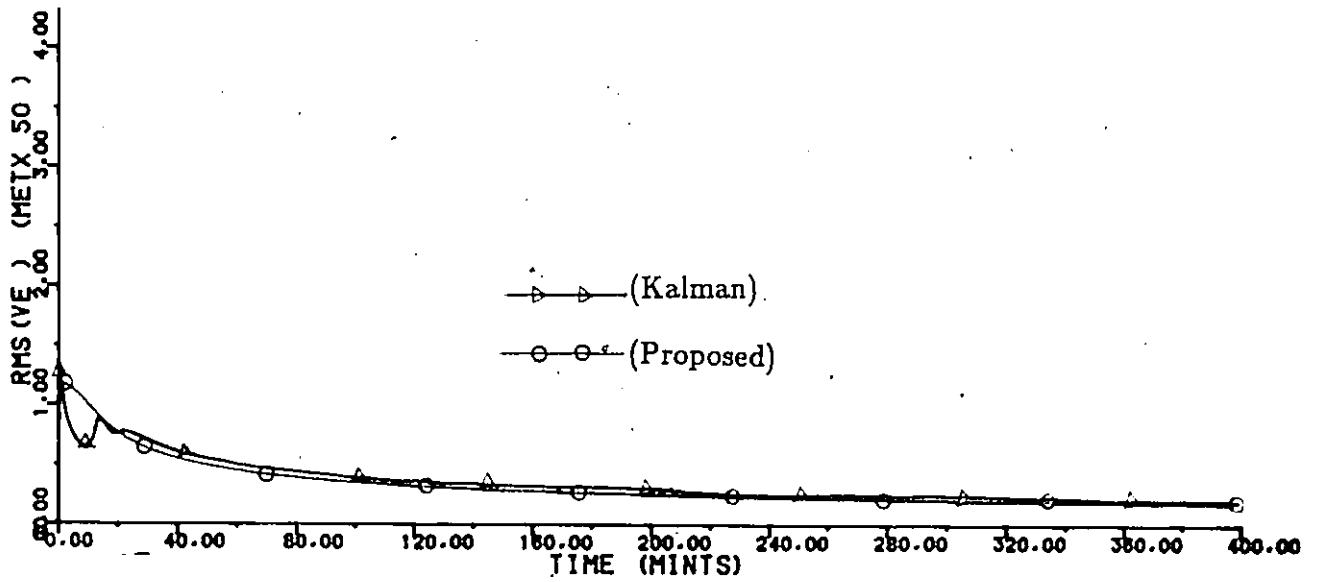


Fig. 5.22 R.M.S. for Eastern Velocity (sample path 2)

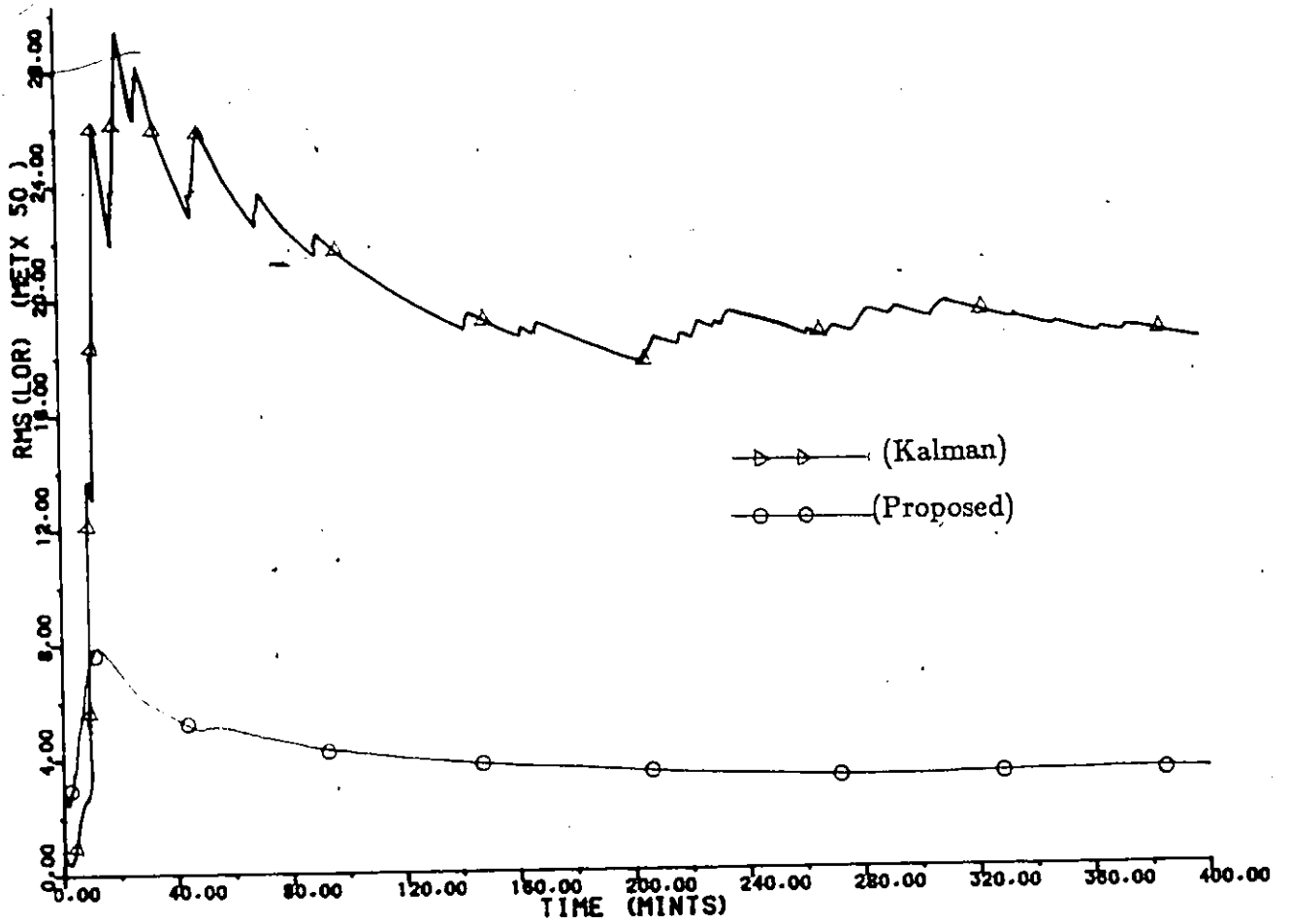


Fig. 5.23 R.M.S. for Loran-C (sample path 2)

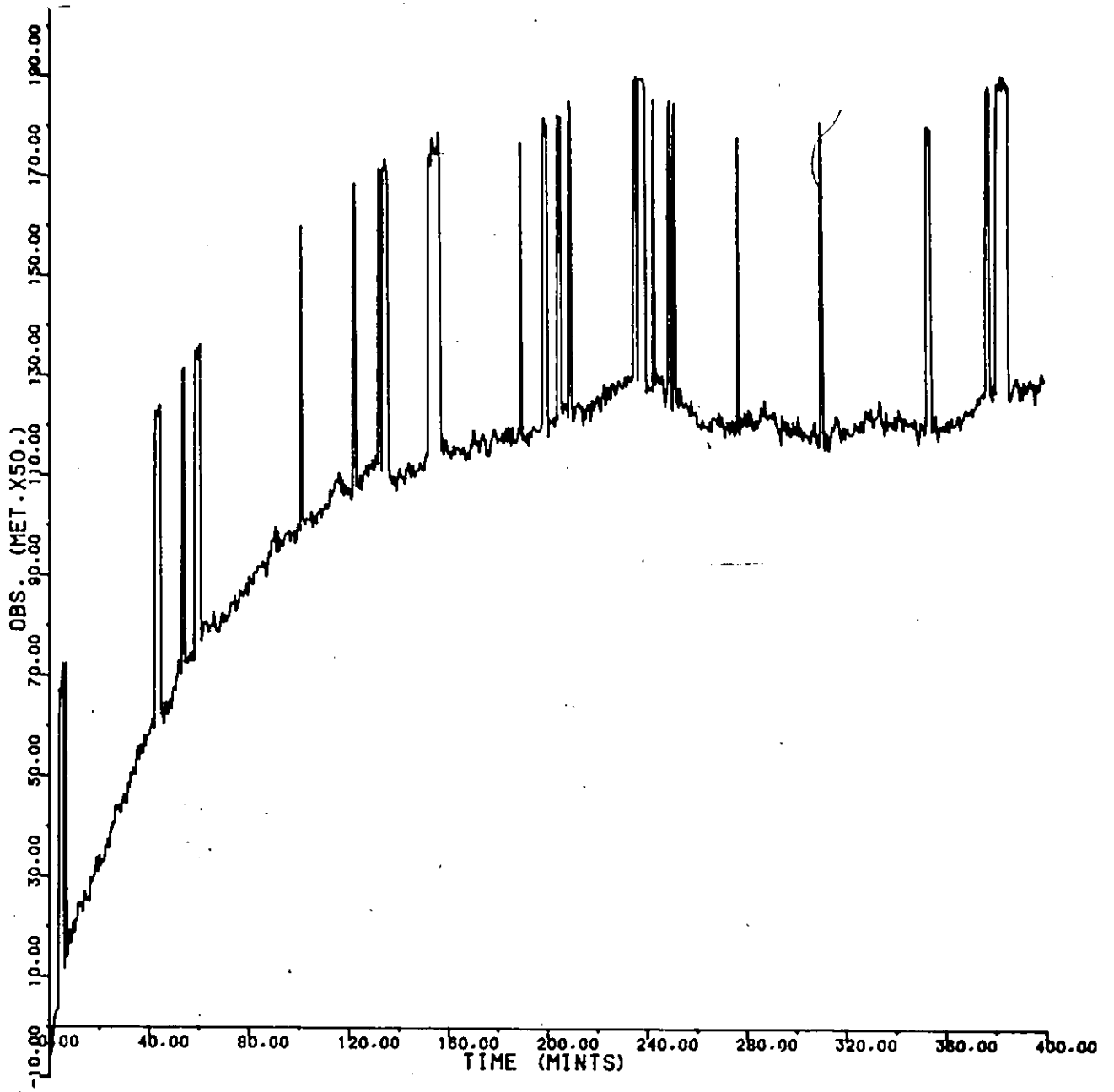


Fig. 5.24 Observed Process (sample, path 3)

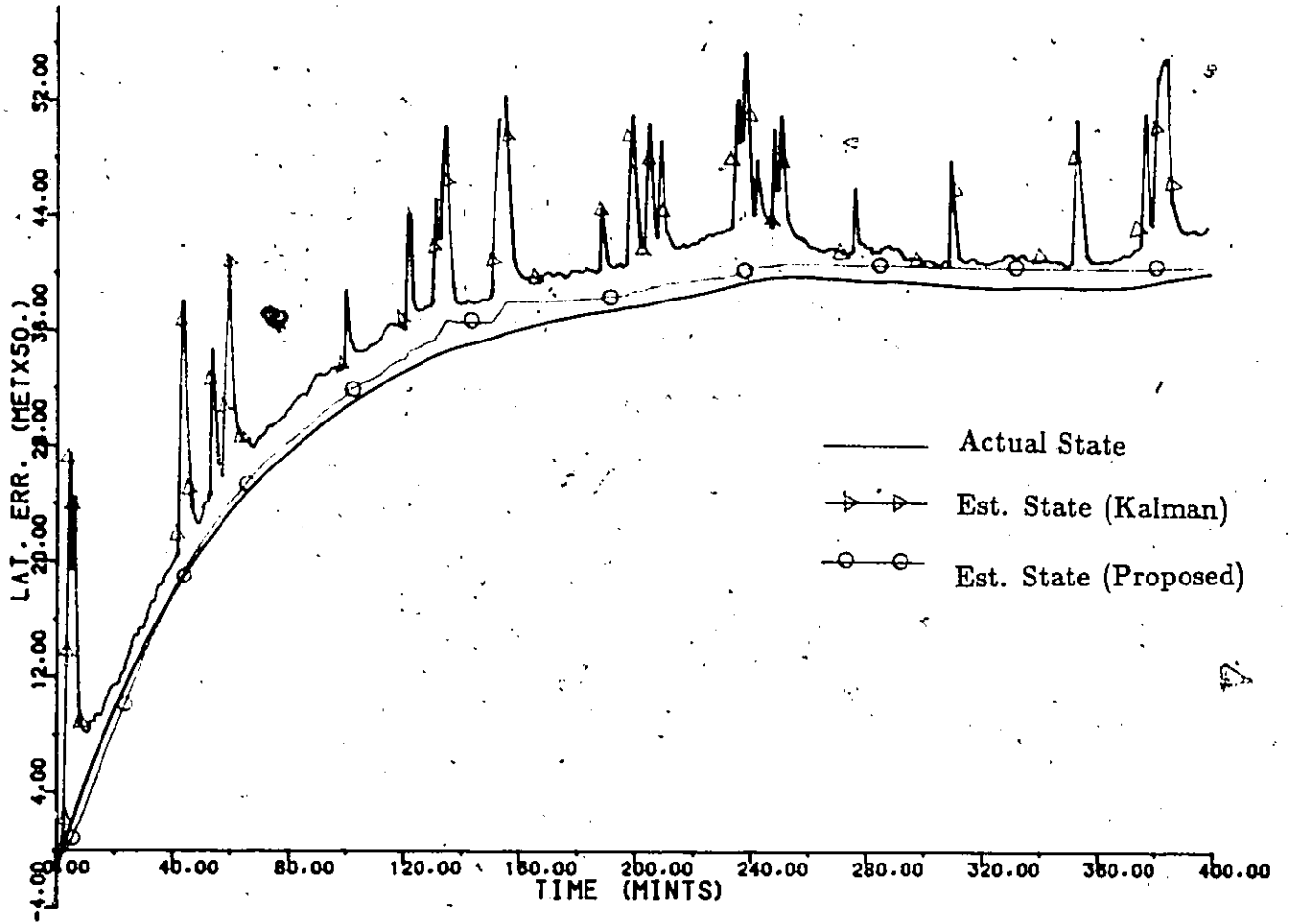


Fig. 5.25 Latitude Error (sample path 3)

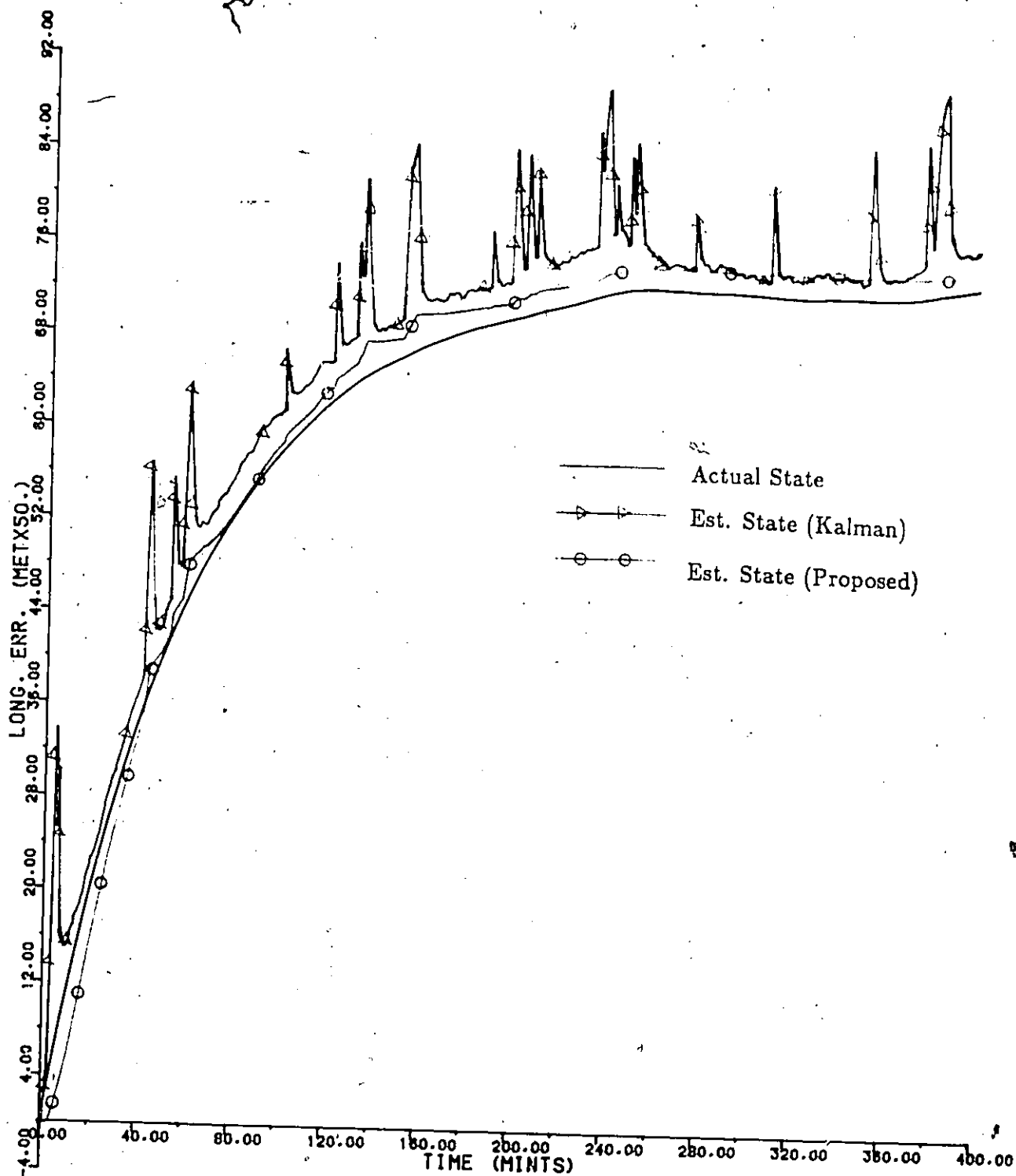
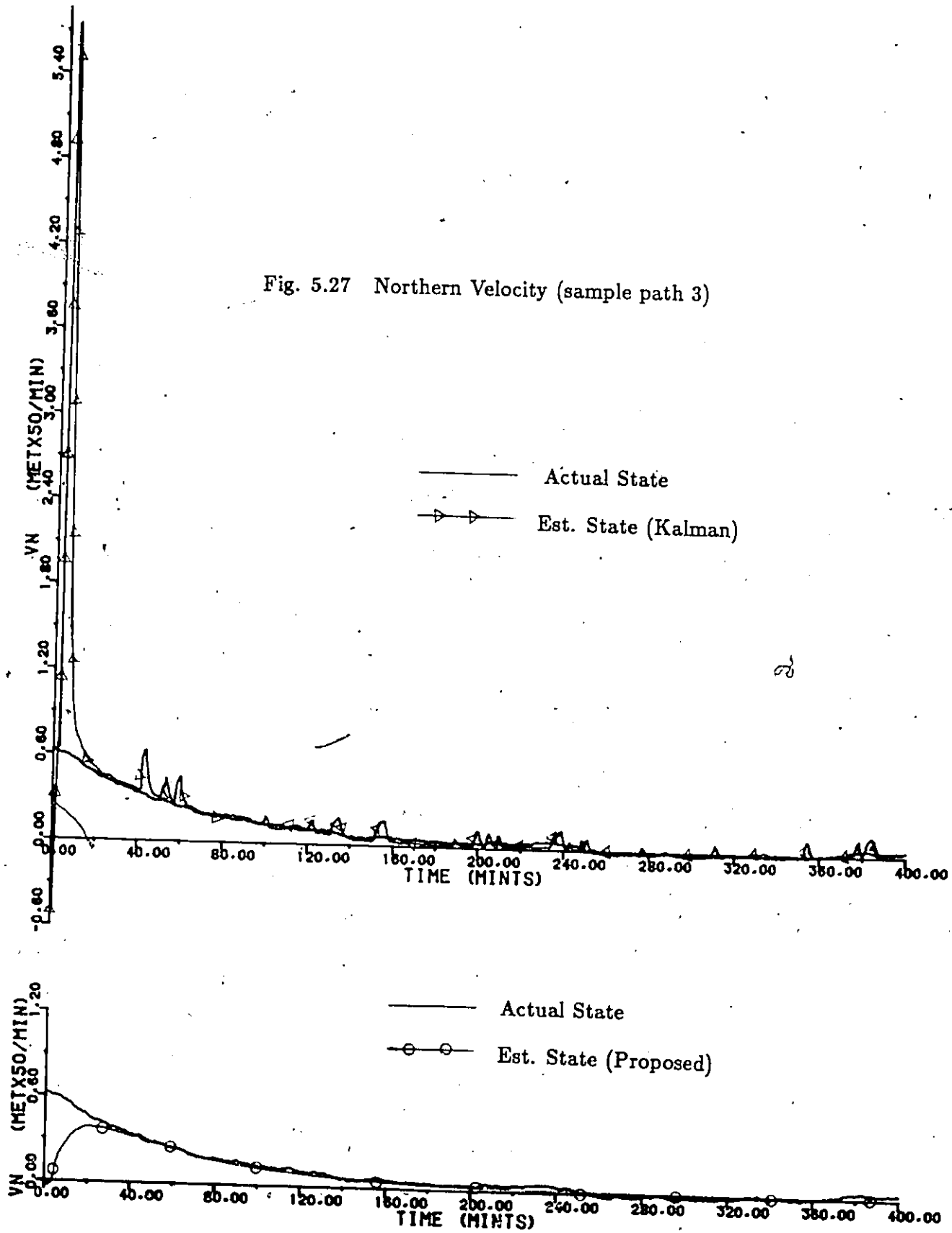


Fig. 5.26 Longitude Error (sample path 3)

Fig. 5.27 Northern Velocity (sample path 3)



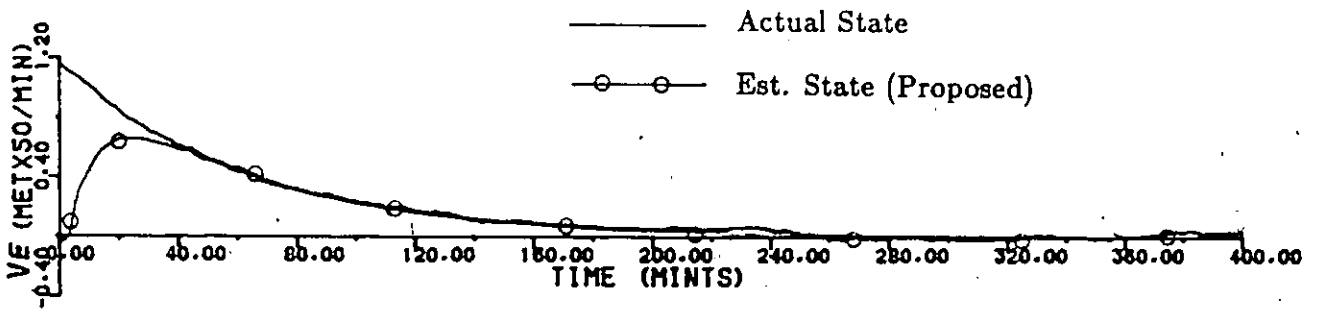
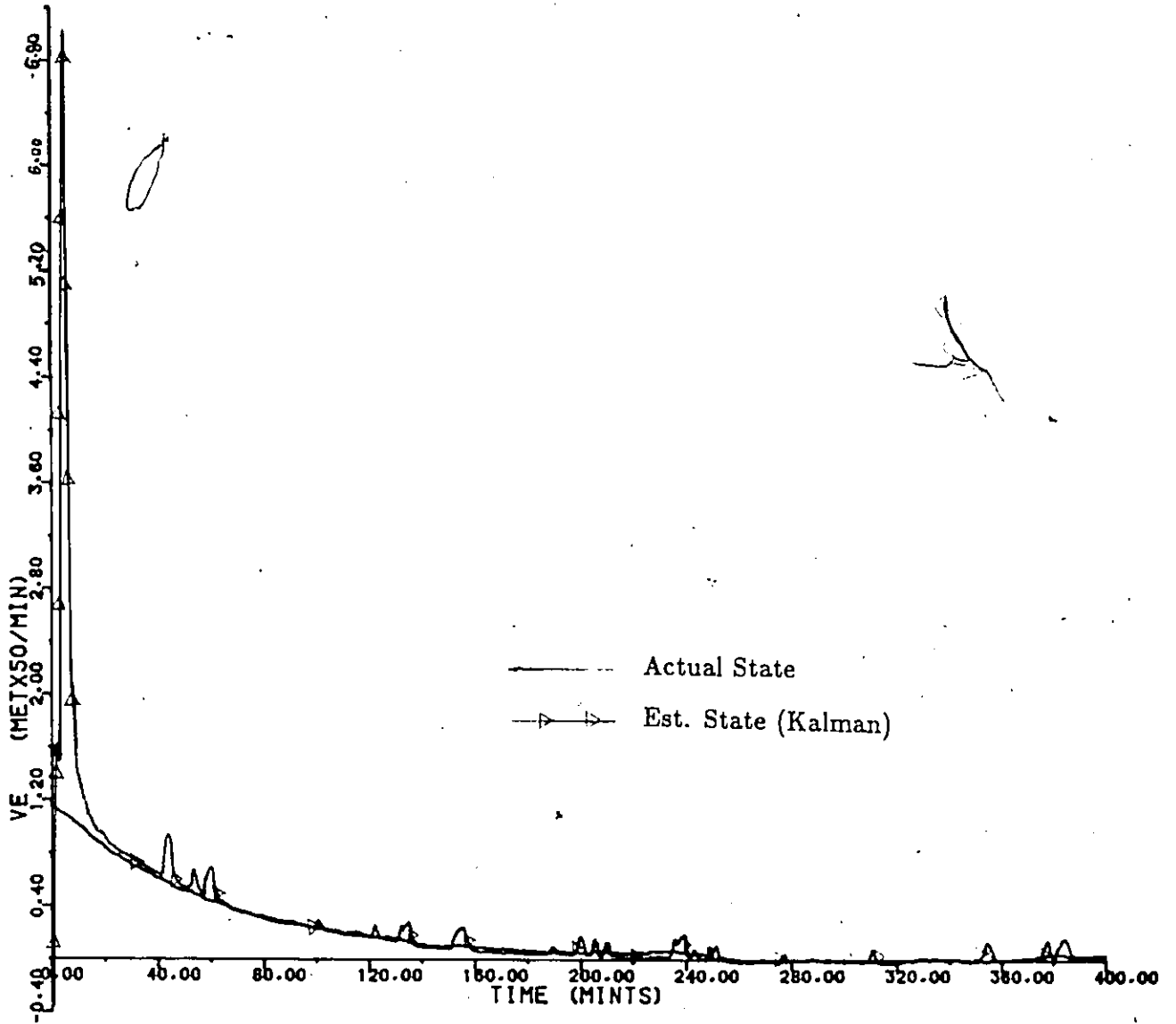


Fig. 5.28 Eastern Velocity (sample path 3)

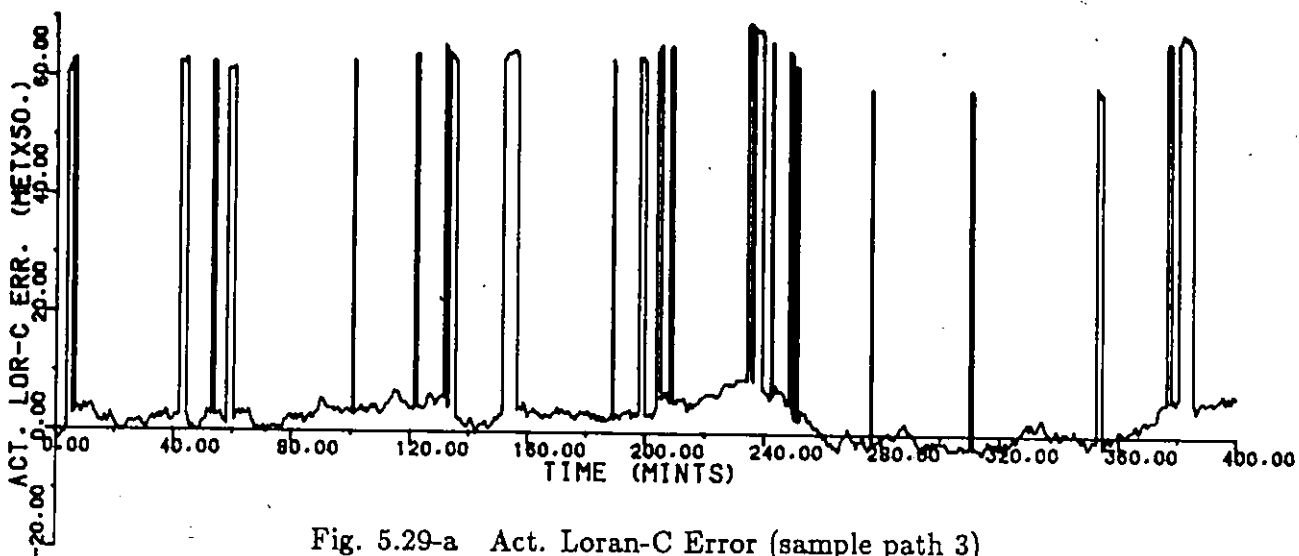


Fig. 5.29-a Act. Loran-C Error (sample path 3)

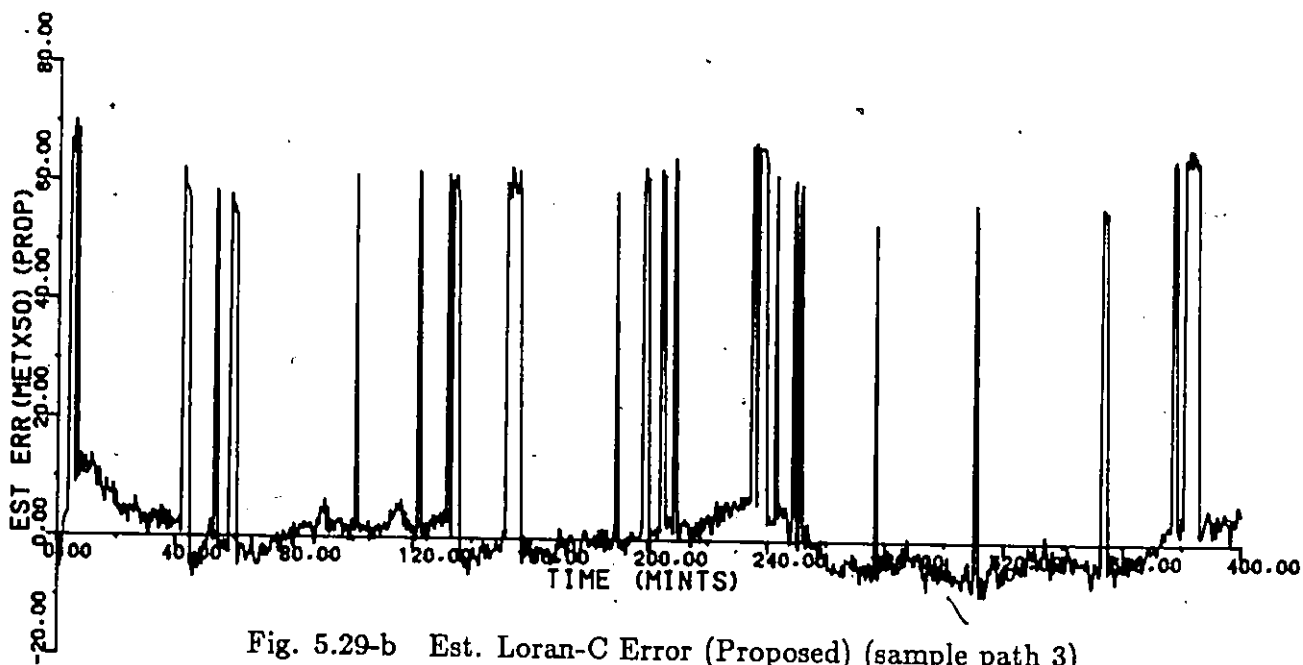


Fig. 5.29-b Est. Loran-C Error (Proposed) (sample path 3)

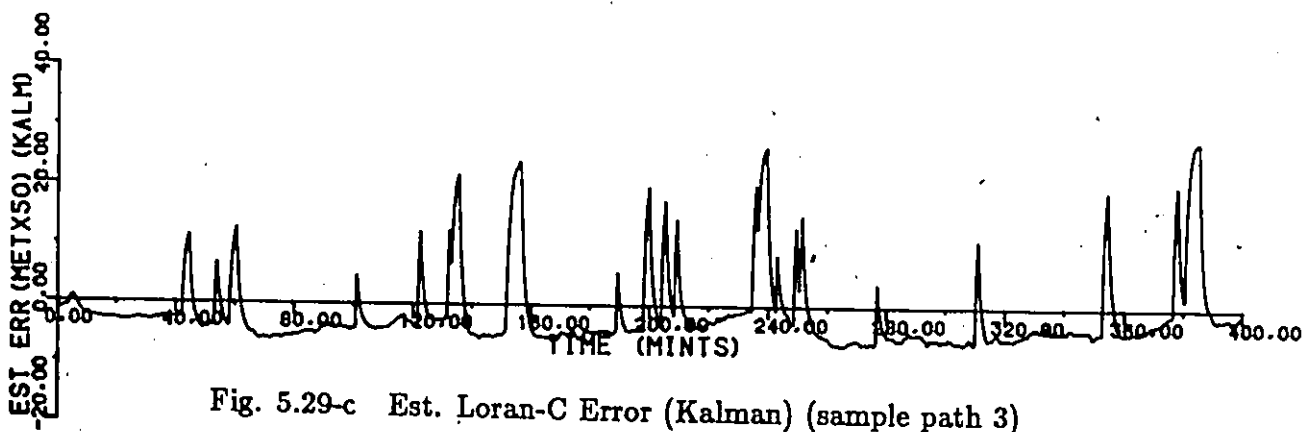


Fig. 5.29-c Est. Loran-C Error (Kalman) (sample path 3)

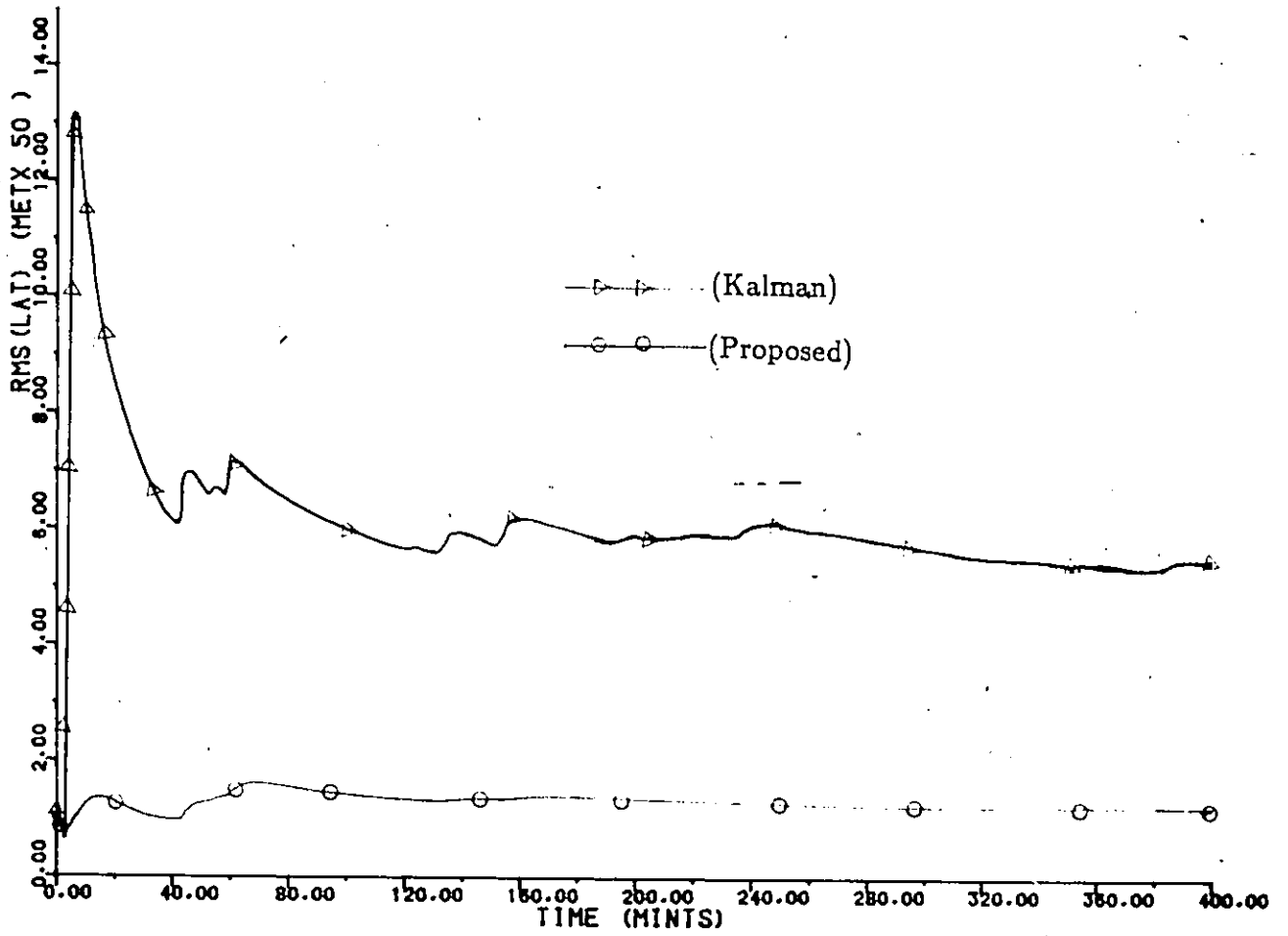


Fig. 5.30 R.M.S. for Latitude (sample path 3)

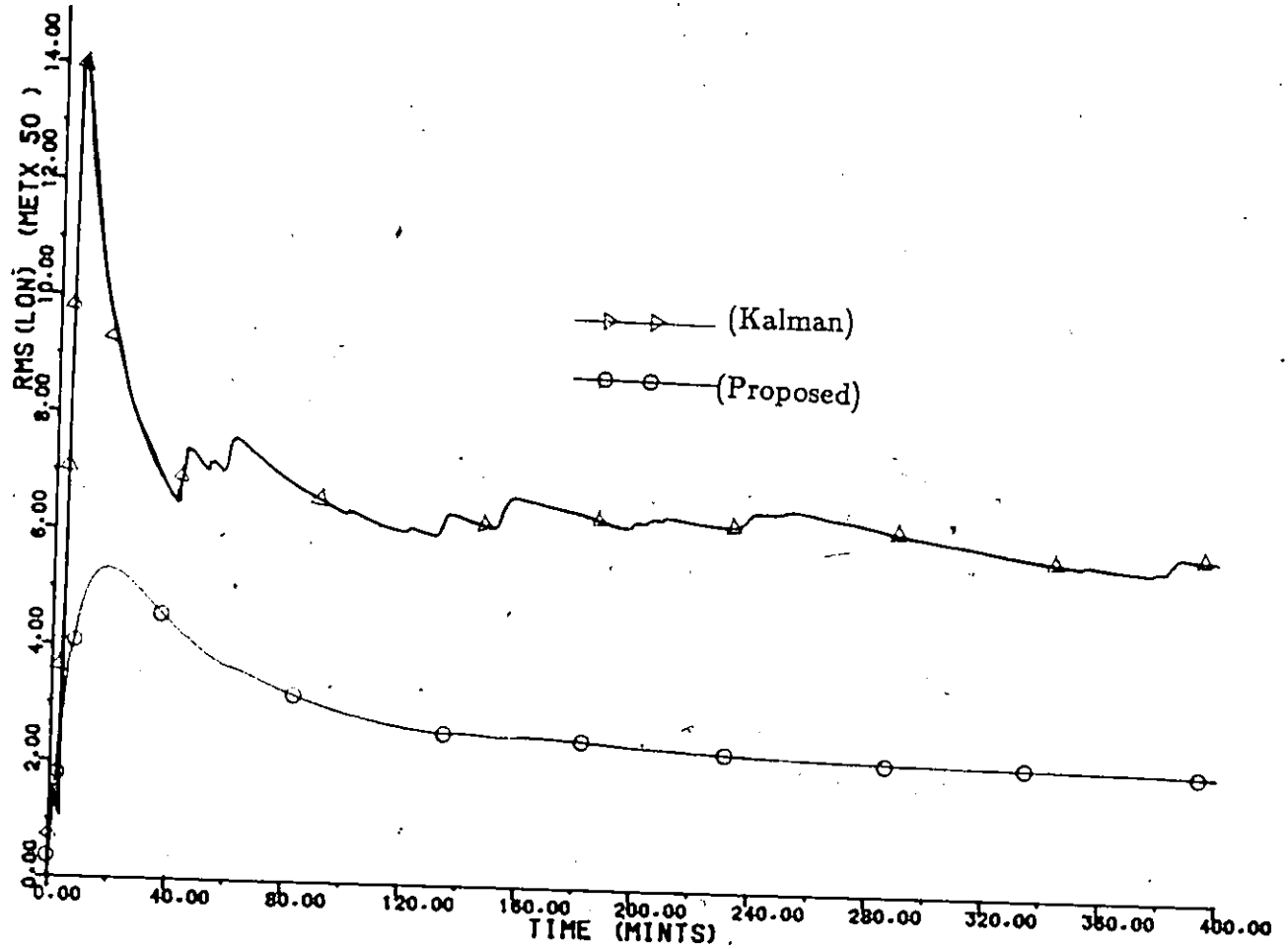


Fig. 5.31 R.M.S. for Longitude (sample path 3) .

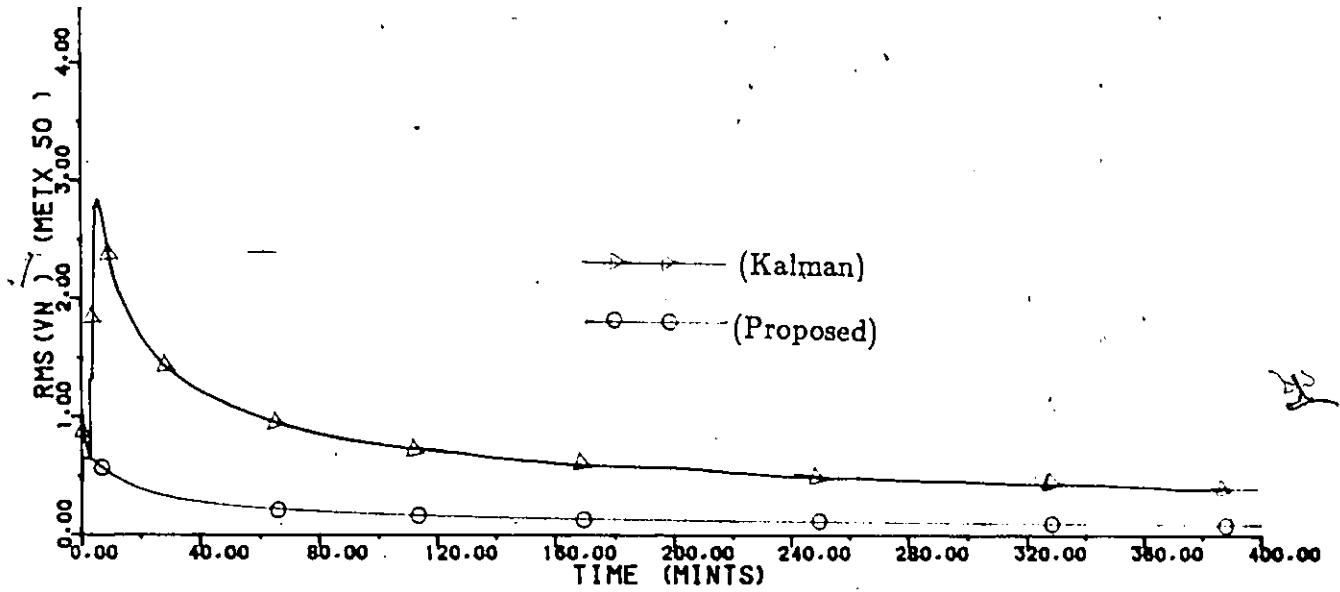


Fig. 5.32 R.M.S. for Northern Velocity (sample path 3)

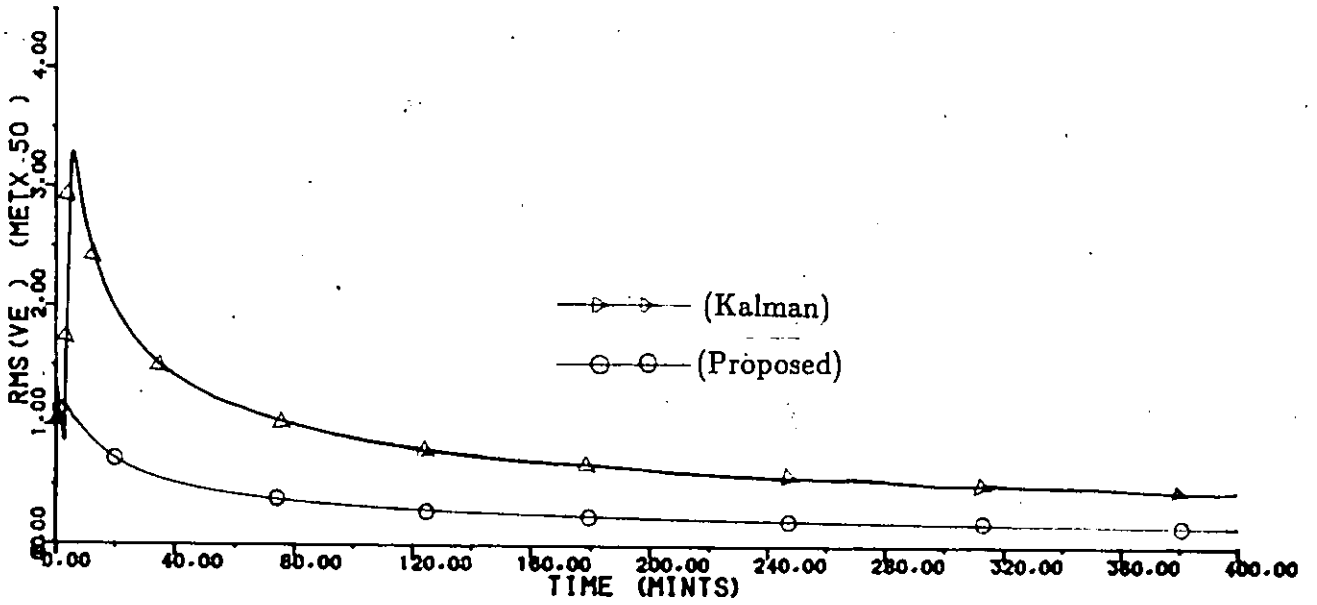


Fig. 5.33 R.M.S. for Eastern Velocity (sample path 3)

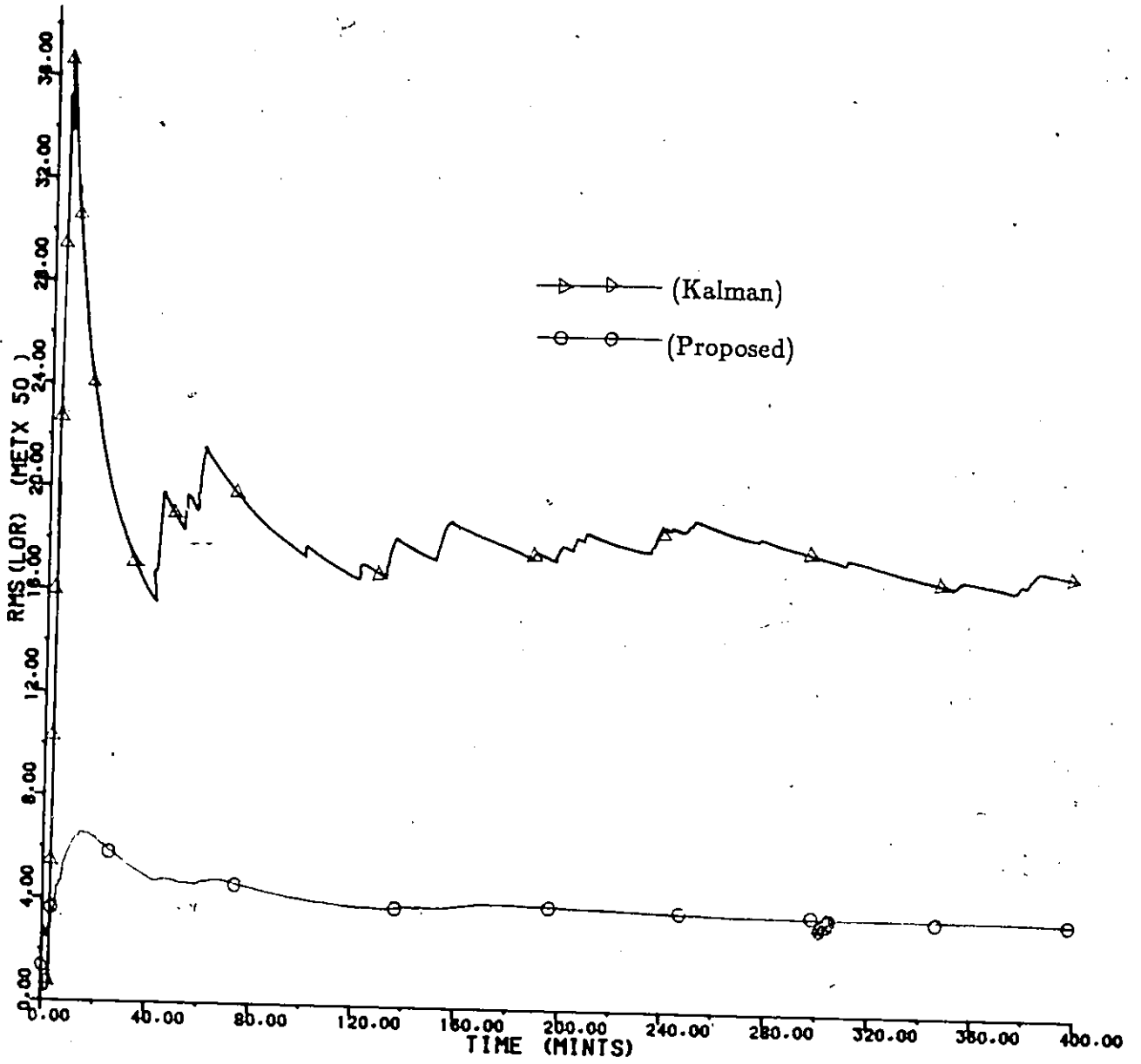


Fig. 5.34 R.M.S. for Loran-C (sample path 3)

filter have similar behaviour and their values are fairly close (see figs. 5.3-5.7, 5.14-5.18 and 5.25-5.29). It is also clear from figs. 5.7, 5.18 and 5.29 that whenever the Loran-C time (range) difference error jumps from one state to another, the filter tends to follow these jumps with reasonable accuracy. In contrast, the Palman filter does not properly respond to these jumps and hence the behaviour of the estimated and actual states are quite different. For further illustration of the behaviour of the proposed filter in comparison to Palman filter, we have also computed the r.m.s. values corresponding to the process $\tilde{x}(t_k) \equiv x(t_k) - \hat{x}(t_k|t_k)$ and the results are shown in figs. 5.8-5.12, 5.19-5.23 and 5.30-5.34 for the three sample paths. From these figures it is clear that the r.m.s. values obtained by Palman filter are almost 2-7 times those obtained by the proposed filter. Indeed from these results it is clear that the response of the proposed filter to the rapid changes of Loran-C error is much better compared to that of the Palman filter, and hence one expects better estimates of position and velocity.

Finally, it should be noted that our model (5.8) and the corresponding filter proposed in this chapter, allows only for a single jump in the Loran-C time (range) difference error. However, this model, and hence the corresponding filter, can be extended to include the effect of multiple jumps (in both directions) by the appropriate modelling of the process N (see equation (5.10)). This work is presently under investigation.

5.5 Summary and Conclusion

In this chapter, the problem of modelling and filtering the errors that arise in dead Reckoning and Loran-C navigation sensors has been considered. The Dead Reckoning error was modelled as a diffusion process (linear Ito equation driven by Wiener process) as in [60,64]. In order to include the effect of propagation as well as cycle selection errors that arise in Loran-C radio signal, we have modelled each master-slave time (range) difference as sum of two independent random processes. The first process was modelled as a diffusion Markov process [60]. The second as a pure jump Markov process driven by two independent Poisson processes. Using the fact that the signal process $x(t), t \geq 0$, is conditionally Gaussian, given \mathcal{F}_t^y , (since the process y carries all the information about the jump term appearing in equation (5.8)), the third order moment of x vanishes and in this case the filtering problem reduces to solving a set of finite number of differential equations from which the required estimate can be determined. In fact, as indicated in section 5.3, the filter equations (5.45)-(5.52) are similar in nature to those of the conventional Palman filter except that we have included extra terms to compensate the effect of the jump term appearing in equation in equation (5.8). Further, using these filter equations, we have developed a numerical algorithm for computing the estimate $\hat{x}(t_k|t_k)$ and the corresponding covariance matrix $P(t_k|t_k)$. Based on this algorithm, we have presented a numerical example to illustrate the behaviour of the proposed filter. From this example, we have observed that under the rapid changes of Loran-C time (range) difference error, the proposed filter has the tendency to follow closely these rapid changes and hence one expects better estimates for ship (aircraft) position and velocity. Finally, it should be noted that the navigation model, and the corresponding filter, that have been proposed in this chapter represent the case where

Loran-C error has only a single jump with constant magnitude (see fig. 5.1). However, this model and the corresponding filter can be extended to include multiple jumps (in both directions) for Loran-C cycle selection error by the proper choice of the process N (see equation (5.10)). This work is currently under investigation.

CHAPTER 6

APPLICATION OF NONLINEAR FILTERING TO PARAMETER IDENTIFICATION PROBLEM

6.1 Introduction

An important and essential aspect of modelling any physical system is the identification of parameters in the model equation. These model equations are usually inferred on the basis of fundamental physical laws and some idealizing assumptions but contains certain parameters which are completely unknown because of lack of precise understanding of the system, or partly known because of poor measurement data. The analyst must determine these parameters on the basis of the available field data.

In this chapter we consider the identification problem for the drift coefficients of systems governed by Ito equations with and without Markov jump Parameters. Following similar arguments as those given in [36] and utilizing the results of chapters 2 and 3, we formulate these identification problems as control problems where the unnormalized densities act as the states, the unknown parameters (vectors or matrices) as controls and the likelihood ratio as the objective functional. The question of existence of elements, in the parameter set, that maximize the likelihood ratio is discussed. The method used in the proof of existence is mainly based on similar arguments as those given in [3,6,7,12] for the control and identification problems of distributed parameter systems. Further, the necessary conditions for optimal identification, using variational method, are also presented. This requires the Gateaux differentiability of the unnormalized densities on the parameter set, which is also proved using similar arguments as those in [3,6,7,12].

In section 6.2, we formulate the identification problems and present the necessary assumptions and notations which have been used in the sequel. In section 6.3, we formulate the identification problems as control problems and show that these problems have solutions. In section 6.4, we present the necessary conditions for optimal identification. Based on these necessary conditions, we develop in section 6.5 an iterative procedure for computing the maximum likelihood estimate of the unknown parameters along with two numerical examples to illustrate some of the results of this chapter.

6.2 Formulation of Identification Problems, Assumptions and Notations

In this chapter, we consider the following identification problems

Problem (I1)

Let the (state) process x_t and the output process $y_t, t \in [0, T]$, be governed by the following Ito equations

$$\begin{aligned} dx_t &= a(t, x_t, \alpha)dt + b(t, x_t)dW_t, \\ x(0) &= x_0, \end{aligned} \tag{6.1}$$

and

$$\begin{aligned} dy_t &= h(x_t)dt + \sigma_0(y_t)d\widetilde{W}_t, \\ y(0) &= 0, \end{aligned} \tag{6.2}$$

where x, a are vectors in R^n ; b is $(n \times n)$ matrix ; y, h are vectors in R^m , σ_0 is $(m \times m)$ matrix and α is unknown parameter taking values from a compact convex set Σ in R^m . The processes W_t and $\widetilde{W}_t, t \geq 0$, are independent n and m dimensional standard Wiener processes, independent of the initial state x_0 . Assuming that all the above random processes and vectors are defined on a complete probability space $(\Omega_0, \mathcal{B}_0, P_0)$, our problem is to determine the unknown parameter vector α , given the history of the observed (output) process y up to time t , that is \mathcal{F}_t^y .

Problem (I2)

Let the processes $x_t, y_t, t \in [0, T]$, be governed by the following Ito equations

$$\begin{aligned} dx_t &= f(t, x_t)dt + b(t, x_t)dW_t, \\ x(0) &= x_0, \end{aligned} \tag{6.3}$$

and

$$\begin{aligned} dy_t &= g(x_t, \beta)dt + \sigma_0(y_t)d\widetilde{W}_t, \\ y(0) &= 0, \end{aligned} \tag{6.4}$$

where $x, f \in R^n; y, g \in R^m, \sigma_0 \in R^{(m \times m)}$, and W and \widetilde{W} are as in the problem (I1). The unknown parameter vector β is assumed to take values in a compact convex set Σ in R^m . The problem is to determine the unknown parameter β given \mathcal{F}_t^y .

Problem (I3)

Let the processes $x_t, y_t, t \in [0, T]$, be governed by the following Ito equations

$$\begin{aligned} dx_t &= \check{f}(t, x_t, \xi_t)dt + b(t, x_t)dW_t, \\ x(0) &= x_0, \end{aligned} \tag{6.5}$$

and

$$\begin{aligned} dy_t &= h(x_t)dt + d\widetilde{W}_t, \\ y(0) &= 0, \end{aligned} \tag{6.6}$$

where $x, \check{f} \in R^n; b, W, y, h$ and \widetilde{W} are as in problems (I1) and (I2). The process $\xi_t, t \geq 0$, is a temporally homogeneous Markov chain (see chapter 3) taking values in a finite set $E \equiv \{e_1, e_2, \dots, e_N\}$, with transition probability matrix $S(t), t \geq 0$, satisfying the following (matrix) differential equation

$$\begin{aligned} \frac{d}{dt}S(t) &= \Lambda S(t), \quad t \in [0, T], \\ S(0) &= I, \end{aligned} \tag{6.7}$$

where I denotes the identity matrix and Λ is unknown. The matrix Λ is assumed to be constant with values in some compact convex set \mathcal{P} . Further, the elements $\{\lambda_{ij}; i, j = 1, 2, \dots, N\}$ of the matrix Λ are given by equation (3.3) and satisfy the

properties (P1) and (P2) of section 3.2. Again our problem is to determine the unknown matrix Λ given \mathcal{F}_t^y .

Using the maximum likelihood ratio approach [17,39,63] and the filter equations obtained in chapters 2 and 3, we formulate in the next section the above identification problems as control problems, where the unnormalized densities (solutions of (2.41) or (3.24)) act as the states, the unknown parameters (constant) as controls, and the likelihood ratio (see section 2.3) as the objective functional. Then following similar arguments as those given in [3,6,7,12], we show that these identification problems have solutions. For this we need the following assumptions and notations.

Assumptions

- (A1) The function a is continuous in its arguments and for all $\alpha \in \Sigma, t \in [0, T]$, $a(t, \cdot, \alpha)$ satisfies the assumptions of chapter 2. Further, the mapping $\alpha \rightarrow a(\cdot, \cdot, \alpha)$ is continuous and once differentiable.
- (A2) The (matrix) function b is continuous in its arguments and for all $t \in [0, T]$, $b(t, \cdot)$ satisfies the assumptions of chapter 2.
- (A3) The function $h \in C_b^2(R^n)$, where $C_b^2(R^n)$ denotes the space of bounded, twice continuously differentiable functions on R^n .
- (A4) The (matrix) function σ_0 satisfies assumption (A5) of chapter 2.
- (A5) The function f is continuous in its arguments and for all $t \in [0, T]$, $f(t, \cdot)$ satisfies the assumptions of chapter 2 (with a being replaced by f).
- (A6) For each $\beta \in \Sigma$, the function $g \in C_b^2(R^n)$; the mapping $\beta \rightarrow g(\cdot, \beta)$ is continuous and once differentiable, and $\beta \rightarrow \frac{\partial}{\partial x} g(x, \beta), x \in R^n$ is continuous.
- (A7) For all $e \in E$ and $t \in [0, T]$, the function $\check{f}(t, x, e)$ is continuous in t and x . Further, for all $e \in E$ and $t \in [0, T]$, $\check{f}(t, \cdot, e)$ satisfies the assumptions of chapter 3 (with a being replaced by \check{f}).

Notations

Let $\eta_t, t \geq 0$, be any random process and let \mathcal{F}_t^η denote the σ -field generated by the process η up to time t . Let $L_2(I; R^n)$ denote the equivalence classes of measurable functions f on I with values in R^n such that $\int_I |f(t)|^2 dt < \infty$, where $|\cdot|$ denotes the usual Euclidean norm.

For any Banach space E we shall use $L_\infty(I; E)$ to denote the space of strongly measurable E valued functions on I with the norm

$$\|f\|_\infty = \text{ess. sup}\{\|f(t)\|_E, t \in I\}.$$

Let $C(I; E)$ denote the space of strongly continuous E valued functions on I furnished with uniform topology $\|f\|_C = \sup\{\|f(t)\|_E, t \in I\}$. For any pair of Banach spaces E and F , we use $\mathcal{L}(E, F)$ to denote the space of bounded linear operators from E to F . For notational convenience, we use $\varphi(t)$ to denote $\varphi(t, x)$ and h to denote $h(x), x \in R^n$. For any vectors $a, b \in R^n$, we use $(a \cdot b)$ to denote the scalar product in R^n . Further notations will be introduced in the sequel as required.

6.3 Existence Theory For Identification Problems

In this section, we use the results of chapters 2 and 3 and the the maximum likelihood approach to formulate the identification problems (I1), (I2) and (I3) as control problems. Then utilizing optimal control Theory, we show that these problems have solutions.

6.3.1 Identification Problem (I1)

Let the state process x_t and the output process $y_t, t \in [0, T]$, be governed by the Ito equations (6.1) and (6.2), respectively, with α being the unknown parameter. Using the results of section 2.3, it is clear that for every $\alpha \in \Sigma$, the unnormalized density $\varphi^\alpha(t) \equiv \varphi^\alpha(t, \cdot)$, satisfies the following Zakai equation

$$\begin{aligned} d\varphi^\alpha(t) &= \mathbf{A}^*(t, \alpha)\varphi^\alpha(t)dt + \varphi^\alpha(t)h \cdot \Gamma_0^{-1}(y_t)dy_t, \\ \varphi^\alpha(0) &= p_0, \end{aligned} \tag{6.8}$$

for all $\alpha \in \Sigma, t \in [0, T]$, where $h \equiv h(\cdot), p_0$ is the initial density of the random variable x_0 , and the operator \mathbf{A}^* is given by

$$\mathbf{A}^*(t, \alpha)f \equiv - \sum_{i=1}^n \frac{\partial}{\partial x_i} (a_i(t, x, \alpha)f) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} (\sigma_{ij}(t, x)f), \tag{6.9}$$

with

$$\sigma(t, x) \equiv (bb')(t, x).$$

Further, as indicated in section 2.4, the likelihood ratio $E_2\{L_t | \mathcal{F}_t^y\}$, is given by (see equation (2.67))

$$\begin{aligned} E_2\{L_t^\alpha | \mathcal{F}_t^y\} &\equiv \langle \varphi^\alpha(t), 1 \rangle \\ &= 1 + \int_0^t \langle \varphi^\alpha(s), h \rangle \cdot \Gamma_0^{-1}(y_s) dy_s, \end{aligned} \tag{6.10}$$

where E_2 denotes the conditional expectation with respect to the measure μ_2 induced by the systems (6.1) and (6.2) with $h \equiv 0$, and $\varphi^\alpha(t)$, $\alpha \in \Sigma$, $t \in [0, T]$, is the solution of (6.8).

Let $L_2(E)$ denote the equivalence classes of measurable functions on E such that $\int_E |f(x)|^2 dx < \infty$. Consider the Sobolev space

$$H^1 \equiv \{f \in L_2(R^n) : \frac{\partial f}{\partial x_i} \in L_2(R^n); 1 \leq i \leq n\},$$

with $H^{-1} \equiv (H^1)'$ being the dual space. As indicated in section 2.4, under our assumptions (A1) and (A2) of section 6.2, the operator A^* (see equation (6.9)) gives rise to the following (divergence) form

$$\begin{aligned} \langle A^*u, v \rangle &= \langle u, Av \rangle \\ &= -\frac{1}{2} \int_{R^n} \sum_{i,j=1}^n \sigma_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx + \int_{R^n} \sum_{i=1}^n \tilde{a}_i \frac{\partial u}{\partial x_i} v dx, \end{aligned} \tag{6.11}$$

for all $\alpha \in \Sigma$ and $u, v \in H^1$, where $\langle \cdot, \cdot \rangle$ denotes the pairing of H^1 and H^{-1} , and

$$\begin{aligned} \tilde{a}_i(t, x, \alpha) &\equiv a_i(t, x, \alpha) - \frac{1}{2} \sum_{j=1}^n \frac{\partial \sigma_{ij}}{\partial x_j}, \\ \sigma(t, x) &\equiv (bb')(t, x). \end{aligned} \tag{6.12}$$

Clearly, both A and $A^* \in \mathcal{L}(H^1, H^{-1})$, where $\mathcal{L}(H^1, H^{-1})$ denotes the class of bounded linear operators from H^1 to H^{-1} .

Defining

$$\varphi^\alpha(t) \equiv \exp(h \cdot Z_t) V^\alpha(t), \quad t \in [0, T], \alpha \in \Sigma, \tag{6.13}$$

where \cdot denotes the scalar product and

$$Z_t \equiv \int_0^t \Gamma_0^{-1}(y_s) dy_s, \quad t \in [0, T], \tag{6.14}$$

then for each $\alpha \in \Sigma$, the process $V^\alpha(t), t \in [0, T]$, is governed by the following linear parabolic partial differential equation (see section 2.5)

$$\begin{aligned} \frac{d}{dt} V^\alpha(t) &= F^*(t, \alpha) V^\alpha(t), \quad t \in [0, T], \\ V^\alpha(0) &= p_0. \end{aligned} \quad (6.15)$$

Here the operator F^* is given by

$$\begin{aligned} \langle F^*(t, \alpha)v, u \rangle &= \langle v, F(t, \alpha)u \rangle \\ &= \frac{1}{2} Z_t \cdot \sum_{i,j=1}^n \int_{R^n} \sigma_{ij} \frac{\partial h}{\partial x_i} \frac{\partial u}{\partial x_j} v dx - \frac{1}{2} \sum_{i,j=1}^n \int_{R^n} \sigma_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx \\ &\quad + \sum_{i=1}^n \int_{R^n} \tilde{a}_i \frac{\partial v}{\partial x_i} u dx + \sum_{i=1}^n \int_{R^n} \check{a}_i u v dx, \end{aligned} \quad (6.16)$$

where

$$\tilde{a}_i(t, x, \alpha) \equiv a_i(t, x, \alpha) - \frac{1}{2} \sum_{j=1}^n \frac{\partial \sigma_{ij}}{\partial x_j}(t, x) - \frac{1}{2} Z_t \cdot \sum_{j=1}^n \sigma_{ij}(t, x) \frac{\partial h}{\partial x_j}, \quad (6.17)$$

and

$$\begin{aligned} \check{a}_i(t, x, \alpha) &\equiv \left(\frac{1}{2} \sum_{j=1}^n \frac{\partial \sigma_{ij}}{\partial x_j}(t, x) - a_i(t, x, \alpha) \right) \frac{\partial h}{\partial x_i} \cdot Z_t \\ &\quad + \frac{1}{2} \sum_{j=1}^n \sigma_{ij}(t, x) \left(\frac{\partial h}{\partial x_i} \cdot Z_t \right) \left(\frac{\partial h}{\partial x_j} \cdot Z_t \right) - \frac{1}{2} (\Gamma_0^{-1} h \cdot h). \end{aligned} \quad (6.18)$$

Under the assumptions (A1), (A2), (A3) and (A4) of section 6.2 and the fact that $\Pr\{\sup_{t \in [0, T]} |Z_t| < \infty\} = 1$, one can verify that the operator $F^* \in \mathcal{L}(H^1, H^{-1})$ and the following properties hold

(P1) For all $\alpha \in \Sigma$ and $u, v \in H^1$, the mapping $t \rightarrow \langle F^*(t, \alpha)u, v \rangle$ is measurable and there exists a constant $c > 0$ such that

$$|\langle F^*(t, \alpha)u, v \rangle| \leq c \|u\|_{H^1} \|v\|_{H^1}, \quad \text{for all } t \in [0, T].$$

(P2) There exist constants $\gamma > 0$ and $\eta \geq 0$ such that

$$-\langle F^*u, u \rangle + \eta \|u\|_{L_2}^2 \geq \gamma \|u\|_{H^1}^2,$$

for all $t \in [0, T]$, $\alpha \in \Sigma$, and $u \in H^1$.

(P3) For any sequence of parameters $\{\alpha^n\}$ that converges to α^0 in Σ ,

$$F^*(t, \alpha^n) \rightarrow F^*(t, \alpha^0),$$

in the strong operator topology of $\mathcal{L}(H^1, H^{-1})$.

(P4) The mapping $\alpha \rightarrow F^*(t, \alpha)$ is once Gateaux differentiable in the strong operator topology of $\mathcal{L}(H^1, H^{-1})$ in the sense that

$$\lim_{\epsilon \rightarrow 0} \left\| \frac{F^*(t, \alpha^0 + \epsilon\alpha)u - F^*(t, \alpha^0)u}{\epsilon} - \bar{F}^*(t, \alpha^0; \alpha)u \right\|_{H^{-1}} = 0,$$

for all $t \in [0, T]$; $\alpha^0, \alpha \in \Sigma$; $0 \leq \epsilon \leq 1$ and $u \in H^1$, where $\bar{F}^*(t, \alpha^0; \alpha)$ denotes the Gateaux differential of the operator F^* at the point α^0 in the direction α .

Utilizing the properties (P1) and (P2), it follows from Lemma 2.4 that for every $\alpha \in \Sigma$, the initial value problem (6.15) has a unique (weak) solution $V^\alpha \in L_2((0, T); H^1) \cap C([0, T]; L_2(R^n))$.

Using the definitions (6.13) and (6.14) and assuming that the function $h \in C_b^2(R^n) \cap H^1(R^n)$, the identification problem (II) can be reformulated as follows

Problem (II)'

Given $\mathcal{F}_t^y, t \in [0, T]$, find an $\alpha^0 \in \Sigma$ such that $J(\alpha^0) \geq J(\alpha)$ for all $\alpha \in \Sigma$, where

$$E_2\{L_t^\alpha | \mathcal{F}_t^y\} \equiv J(\alpha) \equiv \int_0^T dZ_t \cdot \langle V^\alpha(t), h \exp(h \cdot Z_t) \rangle_{H^1 - H^{-1}}, \quad (6.19)$$

and $Z_t, t \in [0, T]$, is given by (6.14) and $V^\alpha(t), t \in [0, T]$, is the solution of the initial value problem (6.15). Note that equation (6.19) is obtained from equation (6.10) by using equations (6.13) and (6.14).

Remark 6.1

Throughout this chapter we shall call α^0 the maximum likelihood estimate of the unknown parameter α .

We now present the following result which claims that the identification problem (I1)' has a solution. In the proof we shall follow the same arguments as those given by Ahmed [12].

Theorem 6.1

Consider the identification problem (I1)' and suppose that the operator F^* satisfies the properties (P1), (P2) and (P3). Let the function $h \in C_b^2(\mathbb{R}^n) \cap H^1(\mathbb{R}^n)$ and suppose that the process $Z_t, t \in [0, T]$, as given by equation (6.14), satisfy the following property

$$\Pr\left\{ \sup_{t \in [0, T]} |Z_t| < \infty \right\} = 1. \quad (6.19.1)$$

Then the mapping $\alpha \rightarrow J(\alpha)$, where J is given by equation (6.19), is continuous on Σ and the problem (I1)' has a solution.

Proof

Let $\{\alpha^n\} \in \Sigma$ and suppose that $\{\alpha^n\} \rightarrow \alpha^0$. Clearly, (by the compactness of the set Σ) $\alpha^0 \in \Sigma$ and hence, under the hypotheses of the Theorem, it follows from Lemma 2.4 that the system

$$\begin{aligned} \frac{d}{dt}V(t) &= F^*(t, \alpha^0)V(t), & t \in [0, T], \\ V(0) &= p_0, \end{aligned} \quad (6.20)$$

has a unique solution $V^0 \in L_2((0, T); H^1) \cap C([0, T]; L_2(R^n))$. Similarly, corresponding to each $\alpha^n \in \Sigma$, the system

$$\begin{aligned} \frac{d}{dt}V(t) &= F^*(t, \alpha^n)V(t), \quad t \in [0, T], \\ V(0) &= p_0, \end{aligned} \tag{6.21}$$

has a unique solution $V^n \in L_2((0, T); H^1) \cap C([0, T]; L_2(R^n))$.

Define

$$z^n(t) \equiv V^n(t) - V^0(t), \quad t \in [0, T], \tag{6.22}$$

where V^n and V^0 are the solutions of (6.20) and (6.21). Then it is clear that $z^n(t), t \in [0, T]$, is the solution of the following initial value problem

$$\begin{aligned} \frac{d}{dt}z^n(t) &= F^*(t, \alpha^n)z^n(t) + (F^*(t, \alpha^n) - F^*(t, \alpha^0))V^0(t), \\ z^n(0) &= 0, \end{aligned} \tag{6.23}$$

for all $t \in [0, T]$, where $V^0(t), t \geq 0$, is the solution of (6.20). Scalar multiplying both sides of the above equation by z^n and integrating with respect to $t \in [0, T]$, we have

$$\begin{aligned} \frac{1}{2}|z^n(t)|_{L_2}^2 - \int_0^t \langle F^*(\theta, \alpha^n)z^n(\theta), z^n(\theta) \rangle d\theta &= \int_0^t \langle (F^*(\theta, \alpha^n) \\ &\quad - F^*(\theta, \alpha^0))V^0(\theta), z^n(\theta) \rangle d\theta. \end{aligned} \tag{6.24}$$

Using the property (P2), it follows from (6.24) that

$$\begin{aligned} |z^n(t)|_{L_2}^2 - 2\eta \int_0^t |z^n(\theta)|_{L_2}^2 d\theta + 2\gamma \int_0^t |z^n(\theta)|_{H^1}^2 d\theta \\ \leq 2 \int_0^t \langle (F^*(\theta, \alpha^n) - F^*(\theta, \alpha^0))V^0(\theta), z^n(\theta) \rangle d\theta. \end{aligned}$$

Using Schwartz inequality, it follows from the above equation that

$$\begin{aligned} |z^n(t)|_{L_2}^2 - 2\eta \int_0^t |z^n(\theta)|_{L_2}^2 d\theta + 2\gamma \int_0^t |z^n(\theta)|_{H^1}^2 d\theta \\ \leq 2 \left(\int_0^t \|(F^*(\theta, \alpha^n) - F^*(\theta, \alpha^0))V^0(\theta)\|_{H^{-1}}^2 d\theta \right)^{1/2} \left(\int_0^t |z^n(\theta)|_{H^1}^2 d\theta \right)^{1/2}. \end{aligned} \quad (6.25)$$

Using the elementary inequality $ab \leq \frac{1}{2\epsilon}a^2 + \frac{\epsilon}{2}b^2$; $a, b \in R$ and $\epsilon > 0$, and taking $\epsilon = \gamma$, it follows from (6.25) that

$$\begin{aligned} |z^n(t)|_{L_2}^2 - 2\eta \int_0^t |z^n(\theta)|_{L_2}^2 d\theta + \gamma \int_0^t |z^n(\theta)|_{H^1}^2 d\theta \\ \leq \frac{1}{\gamma} \int_0^t \|(F^*(\theta, \alpha^n) - F^*(\theta, \alpha^0))V^0(\theta)\|_{H^{-1}}^2 d\theta, \end{aligned} \quad (6.26)$$

for all $t \in [0, T]$, and $\alpha^n, \alpha^0 \in \Sigma$. From equation (6.26), it is clear that

$$\begin{aligned} |z^n(t)|_{L_2}^2 + \gamma \int_0^t |z^n(\theta)|_{H^1}^2 d\theta \leq \frac{1}{\gamma} \int_0^t \|(F^*(\theta, \alpha^n) - F^*(\theta, \alpha^0))V^0(\theta)\|_{H^{-1}}^2 d\theta \\ + 2\eta \int_0^t \left\{ |z^n(\theta)|_{L_2}^2 + \gamma \int_0^\theta |z^n(s)|_{H^1}^2 ds \right\} d\theta, \end{aligned}$$

and hence by Gronwall's Lemma, one obtains

$$\begin{aligned} |z^n(t)|_{L_2}^2 + \gamma \int_0^t |z^n(\theta)|_{H^1}^2 d\theta \leq \left\{ \frac{1}{\gamma} \int_0^T \|(F^*(\theta, \alpha^n) - F^*(\theta, \alpha^0))V^0(\theta)\|_{H^{-1}}^2 d\theta \right\} \exp(2\eta T). \end{aligned} \quad (6.27)$$

This implies that the sequence $\{z^n\}$ is contained in a bounded subset of $L_2((0, T); H^1) \cap L_\infty([0, T]; L_2(R^n))$. Using the properties (P1) and (P3), it is clear that

$$\lim_{n \rightarrow \infty} \sup_{t \in [0, T]} |z^n(t)|_{L_2}^2 = 0,$$

and hence $V^n(t) \rightarrow V^0(t)$ in $L_2(R^n)$ as $n \rightarrow \infty$, uniformly in $t \in [0, T]$.

Define

$$J(\alpha^n) \equiv \int_0^T dZ_t \cdot \langle V^n(t), h \exp(h \cdot Z_t) \rangle_{H^1 - H^{-1}},$$

and

$$J(\alpha^0) \equiv \int_0^T dZ_t \cdot \langle V^0(t), h \exp(h \cdot Z_t) \rangle_{H^1-H^{-1}},$$

where $V^n(t)$ and $V^0(t), t \in [0, T]$, are the solutions of (6.20) and (6.21), respectively.

Then

$$\lim_n (J(\alpha^n) - J(\alpha^0)) = \lim_n \int_0^T dZ_t \cdot \langle V^n(t) - V^0(t), h \exp(h \cdot Z_t) \rangle_{H^1-H^{-1}}. \quad (6.28)$$

Since under our assumptions, given \mathcal{F}_t^y , the function $h \exp(h \cdot Z_t), t \in [0, T]$, is in $H^{-1}(R^n)$ and $V^n \rightarrow V^0$ in $L_2((0, T); H^1)$, it follows by the properties of the stochastic integral [68] that the right hand side of (6.28) converges in the mean in the limit and hence there exists a subsequence $\{n\}_k \in \{n\}$ for which

$$J(\alpha^{n_k}) \rightarrow J(\alpha^0), \quad \mathcal{F}_T^y \text{ almost surely.}$$

Since the parameter set Σ is compact and $\alpha \rightarrow J(\alpha)$ is continuous on Σ , $J(\alpha)$ attains its maximum (or minimum) on Σ . This completes the proof. ■

6.3.2 Identification Problem (I2)

In this section, we use similar arguments as those given above to show that the identification problem (I2) has a solution. Consider the systems (6.3) and (6.4), with β being the unknown parameter vector and the parameter set Σ is compact and convex. Then it follows from section 2.3 that for each $\beta \in \Sigma$, the unnormalized density $\varphi^\beta(t), t \in [0, T]$, satisfies the following Zakai equation

$$\begin{aligned} d\varphi^\beta(t) &= A^*(t)\varphi^\beta(t)dt + \varphi^\beta(t)g(\beta) \cdot \Gamma_0^{-1}(y_t)dy_t, \\ \varphi^\beta(0) &= p_0, \end{aligned} \quad (6.29)$$

for all $t \in [0, T]$, where p_0 denotes the initial density, and the operator A^* is given by

$$A^*\varphi = - \sum_{i=1}^n \frac{\partial}{\partial x_i} (f_i(t, x)\varphi) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} (\sigma_{ij}(t, x)\varphi). \quad (6.30)$$

Further, the likelihood ratio, is given by

$$\begin{aligned} E_2\{L_t^\beta | \mathcal{F}_t^y\} &\equiv \langle \varphi^\beta(t), 1 \rangle \\ &= 1 + \int_0^t \langle \varphi^\beta(s), g(\beta) \rangle \cdot \Gamma_0^{-1}(y_s) dy_s, \end{aligned} \quad (6.31)$$

where $\varphi^\beta(t), \beta \in \Sigma, t \in [0, T]$, is the solution of (6.29). Define

$$\varphi^\beta(t) \equiv \exp(g(\beta) \cdot Z_t) V^\beta(t), \quad t \in [0, T], \quad \beta \in \Sigma, \quad (6.32)$$

and

$$\begin{aligned} \langle D^*(t, \beta)v, u \rangle &= \langle v, D(t, \beta)u \rangle \\ &= \frac{1}{2} Z_t \cdot \sum_{i,j=1}^n \int_{R^n} \sigma_{ij} \frac{\partial g(\beta)}{\partial x_i} \frac{\partial u}{\partial x_j} v dx - \frac{1}{2} \sum_{i,j=1}^n \int_{R^n} \sigma_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} dx \\ &\quad + \sum_{i=1}^n \int_{R^n} \tilde{f}_i \frac{\partial v}{\partial x_i} u dx + \sum_{i=1}^n \int_{R^n} \check{f}_i u v dx \end{aligned} \quad (6.33)$$

where

$$\tilde{f}_i(t, x, \beta) \equiv f_i(t, x) - \frac{1}{2} \sum_{i=1}^n \frac{\partial \sigma_{ij}}{\partial x_j}(t, x) - \frac{1}{2} Z_t \cdot \sum_{j=1}^n \sigma_{ij}(t, x) \frac{\partial g}{\partial x_j}(x, \beta), \quad (6.34)$$

$$\begin{aligned} \check{f}_i(t, x, \beta) &\equiv \left(\frac{1}{2} \sum_{j=1}^n \frac{\partial \sigma_{ij}}{\partial x_j}(t, x) - f_i(t, x) \right) \frac{\partial g}{\partial x_i}(x, \beta) \cdot Z_t \\ &\quad + \frac{1}{2} \sum_{j=1}^n \sigma_{ij}(t, x) \left(\frac{\partial g}{\partial x_i} \cdot Z_t \right) \left(\frac{\partial g}{\partial x_j} \cdot Z_t \right)(x, \beta) \\ &\quad - \frac{1}{2} (\Gamma_0^{-1}(y_t) g(x, \beta) \cdot g(x, \beta)), \end{aligned} \quad (6.35)$$

$Z_t, t \in [0, T]$, is given by (6.14) and $u, v \in H^1$. Then for each $\beta \in \Sigma$, $V^\beta(t), t \geq 0$, is governed by the following linear (parabolic) partial differential equation

$$\begin{aligned} \frac{d}{dt} V^\beta(t) &= D^*(t, \beta) V^\beta(t), \quad t \in [0, T], \\ V^\beta(0) &= p_0. \end{aligned} \quad (6.36)$$

Again under the assumptions (A4), (A5), (A6) and the fact that

$$\Pr\left\{ \sup_{t \in [0, T]} |Z_t| < \infty \right\} = 1, \quad (6.37)$$

it is clear that the operator D^* , as given by (6.33), satisfies the properties (P1)-(P4). Hence by Lemma 2.4, it follows that for each $\beta \in \Sigma$, the initial value problem (6.36) has a unique (weak) solution $V^\beta \in L_2((0, T); H^1) \cap C([0, T]; L_2(R^n))$.

Using the definition (6.32) and assuming that for all $\beta \in \Sigma$, the function $g(\beta) \in C_b^2(R^n) \cap H^1(R^n)$, the identification problem (I2) can be reformulated as follows.

Problem (I2)'

Given $\mathcal{F}_t^y, t \in [0, T]$, find a $\beta^0 \in \Sigma$ such that $J(\beta^0) \geq J(\beta)$ for all $\beta \in \Sigma$, where

$$J(\beta) \equiv \int_0^T dZ_t \cdot \langle V^\beta(t), g(\beta) \exp(g(\beta) \cdot Z_t) \rangle_{H^1-H^{-1}}, \quad (6.38)$$

and the \mathcal{F}_t^y -measurable process $Z_t, t \in [0, T]$, (see (6.14)) satisfies (6.37) and $V^\beta(t), t \in [0, T]$, is the solution of (6.36). We now present the following existence result for the identification problem (I2)'.

Theorem 6.2

Consider the identification problem (I2)' and suppose that the operator D^* satisfy the properties (F1), (F2), and (F3) and for all $\beta \in \Sigma, g(\beta) \in C_b^2(R^n) \cap H^1(R^n)$. Let $Z_t, t \in [0, T]$, be given by (6.14) and suppose (6.37) hold. Then the mapping $\beta \rightarrow J(\beta)$, where J is given by (6.38), is continuous on the parameter set Σ and the problem (I2)' has a solution.

Proof

Let $V^n(t)$ and $V^0(t), t \in [0, T]$, be the solutions of the initial value problem (6.36), corresponding to the parameters β^n and β^0 , respectively. Then following similar arguments as those of Theorem 6.1, it is easy to verify that $V^n(t) \rightarrow V^0(t)$ in $L_2(R^n)$, uniformly in $t \in [0, T]$, whenever $\beta^n \rightarrow \beta^0$ in Σ .

Define

$$J(\beta^n) \equiv \int_0^T dZ_t \cdot \langle V^n(t), e^n(t) \rangle_{H^1-H^{-1}},$$

and

$$J(\beta^0) \equiv \int_0^T dZ_t \cdot \langle V^0(t), e^0(t) \rangle_{H^1-H^{-1}},$$

where

$$e^n(t) \equiv g(\beta^n) \exp(g(\beta^n) \cdot Z_t). \quad (6.39)$$

Then

$$\begin{aligned} \lim_n (J(\beta^n) - J(\beta^0)) &= \lim_n \int_0^T dZ_t \cdot \langle V^n(t) - V^0(t), e^n(t) \rangle_{H^1-H^{-1}} \\ &\quad + \lim_n \int_0^T dZ_t \cdot \langle V^0(t), e^n(t) - e^0(t) \rangle_{H^1-H^{-1}}. \end{aligned} \quad (6.40)$$

Since under our assumptions, the function $e^n(t)$, given \mathcal{F}_t^y , is in $L_2((0, T); H^{-1})$ for all n , $e^n \rightarrow e^0$ in $L_2((0, T); H^{-1})$ and $V^n \rightarrow V^0$ in $L_2((0, T); H^1)$, it follows from standard properties of stochastic integral that the right hand side of equation (6.40) converges in the limit in the mean and hence by similar arguments as those of Theorem 6.1, the continuity of J follows. The proof of the Theorem now follows by the compactness of the parameter set Σ . ■

6.3.3 Identification Problem (I3)

Consider the systems (6.3) and (6.6) and suppose that assumptions (A2), (A3), and (A7) hold. Let the process $\xi_t, t \in [0, T]$, be a temporally homogeneous Markov chain taking values from a finite set $E \equiv \{e_1, e_2, \dots, e_N\}$. Let the transition probability matrix $S(t), t \in [0, T]$, corresponding to the process ξ , satisfy the differential equation (6.7) with the infinitesimal transition matrix A being unknown. We assume that A is a constant matrix taking values from a compact convex set \mathcal{P} , with

elements $\{\lambda_{ij}; i, j = 1, 2, \dots, N\}$ satisfying the properties given in section 3.2. Using similar arguments as those given above, we wish to show that the identification problem of the unknown matrix Λ has a solution.

First we note that the solution of the differential equation (6.7), denoted $S(t, \Lambda), t \in [0, T]$, can be expressed as

$$S(t, \Lambda) = \exp(t\Lambda), \quad t \in [0, T]. \quad (6.41)$$

Then defining

$$\Theta_i(t, \Lambda) \equiv \Pr\{\xi_t = e_i; \quad 1 \leq i \leq N\}, \quad (6.42)$$

it is clear that for each $\Lambda \in \mathcal{P}$, the total probability

$$\Theta \equiv (\Theta_1, \Theta_2, \dots, \Theta_N)'$$

of the Markov chain ξ is related to the transition probability matrix $S(t, \Lambda)$ through the following relation

$$\Theta(t, \Lambda) = S(t, \Lambda)\Theta_0 = \exp(t\Lambda)\Theta_0, \quad (6.43)$$

for all $t \in [0, T]$, where $\Theta_0 \equiv \Pr\{\xi(0) = \xi_0\}$.

Clearly, the mapping $\Lambda \rightarrow \Theta(t, \Lambda)$ is continuous, bounded and differentiable on the set \mathcal{P} for all $t \in [0, T]$. Using equation (3.29), it follows that the conditional expectation $E_2\{L_t | \mathcal{F}_t^y; \xi_t = e_i\}, 1 \leq i \leq N$, (which denotes the likelihood ratio given $\xi_t = e_i$), is given by

$$E_2\{L_t | \mathcal{F}_t^y; \xi_t = e_i\} \equiv \langle \varphi_i(t), 1 \rangle = 1 + \int_0^t \langle \varphi_i(s), h \rangle \cdot dy_s, \quad (6.44)$$

for all $1 \leq i \leq N$, where $\varphi_i(t), t \in [0, T]$, satisfies the following linear system of coupled stochastic partial differential equations (see section 3.4)

$$\begin{aligned} d\varphi_i(t) &= \{(A_i^* \varphi_i)(t) + \sum_{k=1}^N \lambda_{ki} \varphi_k(t)\} dt + \varphi_i(t) h \cdot dy_t, \\ \varphi_i(0) &= p_0, \end{aligned} \quad (6.45)$$

for all $1 \leq i \leq N$ and $t \in [0, T]$, where p_0 is the initial density and the operator A^* is given by

$$\begin{aligned} (A_k^* \varphi)(t, x) \equiv & - \sum_{i=1}^n \frac{\partial}{\partial x_i} (f_i(t, x, e_k) \varphi) \\ & + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j} (\sigma_{ij}(t, x) \varphi), \end{aligned} \quad (6.46)$$

for all $1 \leq k \leq N$.

Defining

$$\Gamma^* \Phi \equiv (A^* + \Lambda) \Phi, \quad (6.47)$$

where $\Phi \equiv (\varphi_1, \dots, \varphi_N)'$ and

$$A^* \equiv \begin{pmatrix} A_1^* & 0 & \dots & 0 \\ 0 & A_2^* & \dots & 0 \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ 0 & 0 & \dots & A_N^* \end{pmatrix}, \quad (6.48)$$

with Λ as given above, one can rewrite equation (6.45) in the following vector form

$$\begin{aligned} d\Phi(t) &= \Gamma^*(t, \Lambda) \Phi(t) dt + \Phi(t) (h \cdot dy_t), \\ \Phi(0) &= \Phi_0. \end{aligned} \quad (6.49)$$

Hence it follows from equation (6.44) that

$$\langle \Phi(t), 1 \rangle = r + \int_0^t \langle \Phi(s), h \cdot dy_s \rangle, \quad (6.50)$$

where $r \equiv (1, 1, \dots, 1)'$. Scalar multiplying equation (6.50) by Θ , where Θ is given by equation (6.43), we obtain

$$\Theta(t, \Lambda) \cdot \langle \Phi(t), 1 \rangle = \Theta(t, \Lambda) \cdot \left(r + \int_0^t \langle \Phi(s), h \cdot dy_s \rangle \right), \quad (6.51)$$

which we denote the weighted likelihood ratio.

Now we proceed as section 3.5 to show that equation (6.49) has a solution. For this we need the following definitions.

Let $(L_2(R^n))^N$ denote N -copies of the space $L_2(R^n)$ and consider the Sobolev space

$$H^1 \equiv \{f \in (L_2(R^n))^N : \frac{\partial f}{\partial x_i} \in (L_2(R^n))^N; \quad 1 \leq i \leq n\}, \quad (6.51.1)$$

with H^{-1} being the dual space. Define

$$\Phi(t) \equiv \exp(h \cdot y_t) V(t), \quad t \in [0, T], \quad (6.52)$$

and

$$\langle \tilde{\Gamma}^*(t, \Lambda) u, v \rangle \equiv \langle \Gamma^*(t, \Lambda) u \exp(h \cdot y_t), v \exp(-h \cdot y_t) \rangle - \frac{1}{2} \langle |h|^2 u, v \rangle, \quad (6.53)$$

for any $u, v \in H^1$, where $\langle \cdot, \cdot \rangle$ denotes the pairing of H^1 and H^{-1} , and $\Phi(t), t \in [0, T]$, satisfies equation (6.49). Then following similar arguments as those of section 3.5, one can easily verify that $V(t), t \in [0, T]$, is governed by the following partial differential equation

$$\begin{aligned} \frac{d}{dt} V(t) &= \tilde{\Gamma}^*(t, \Lambda) V(t), \quad t \in [0, T], \\ V(0) &= \Phi_0 \end{aligned} \quad (6.54)$$

With the help of equations (6.46), (6.47), and (6.48), the operator $\tilde{\Gamma}^*$, as given by (6.53), can also be written explicitly as

$$\begin{aligned} \sum_{m,t=1}^N \langle \tilde{\Gamma}_{mt}^* u_t, v_m \rangle &= \frac{1}{2} y_t \cdot \sum_{m=1}^N \sum_{i,j=1}^n \int_{R^n} \sigma_{ij} \frac{\partial h}{\partial x_i} \frac{\partial u_m}{\partial x_j} v_m dx \\ &\quad - \frac{1}{2} \sum_{m=1}^N \sum_{i,j=1}^n \int_{R^n} \sigma_{ij} \frac{\partial u_m}{\partial x_i} \frac{\partial v_m}{\partial x_j} dx \end{aligned}$$

$$\begin{aligned}
 & + \sum_{m=1}^N \sum_{i=1}^n \int_{R^n} \tilde{a}_i^m \frac{\partial v_m}{\partial x_i} u_m dx \\
 & + \sum_{m=1}^N \sum_{i=1}^n \int_{R^n} \check{a}_i^m u_m v_m dx, \tag{6.55}
 \end{aligned}$$

where

$$\tilde{a}_i^m \equiv \tilde{a}_i(t, x, e_m) \equiv \check{f}_i^*(t, x, e_m) - \frac{1}{2} \sum_{j=1}^n \frac{\partial \sigma_{ij}}{\partial x_j} - \frac{1}{2} y_t \cdot \sum_{j=1}^n \sigma_{ij} \frac{\partial h}{\partial x_j}, \tag{6.56}$$

and

$$\begin{aligned}
 \check{a}_i^m \equiv \check{a}_i(t, x, e_m) & \equiv \left(\frac{1}{2} \sum_{j=1}^n \frac{\partial \sigma_{ij}}{\partial x_j} - \check{f}_i(t, x, e_m) \right) \left(\frac{\partial h}{\partial x_i} \cdot y_t \right) \\
 & + \frac{1}{2} \sum_{j=1}^n \sigma_{ij} \left(\frac{\partial h}{\partial x_i} \cdot y_t \right) \left(\frac{\partial h}{\partial x_j} \cdot y_t \right) \\
 & + \sum_{l=1}^N \lambda_{ml} - \frac{1}{2} |h|^2. \tag{6.57}
 \end{aligned}$$

Clearly, under assumptions (A2), (A3), and (A7) of section 6.2, and the fact that

$$\Pr\{ \sup_{t \in [0, T]} |y_t| < \infty \} = 1, \tag{6.58}$$

the operator $\tilde{\Gamma}^* \in \mathcal{L}(H^1, H^{-1})$ and satisfies the following properties

(Pi) For all $\Lambda \in \mathcal{P}$, and $u, v \in H^1$, the mapping $t \rightarrow \langle \tilde{\Gamma}^*(t, \Lambda)u, v \rangle$ is measurable and there exists a constant $c > 0$ such that

$$| \langle \tilde{\Gamma}^*(t, \Lambda)u, v \rangle | \leq c \|u\|_{H^1} \|v\|_{H^1}, \quad \text{for all } t \in [0, T].$$

(Pii) There exist constants $\gamma > 0$ and $\eta \geq 0$ such that

$$- \langle \tilde{\Gamma}^*(t, \Lambda)u, u \rangle + \eta |u|_{L_2}^2 \geq \gamma \|u\|_{H^1}^2,$$

for all $t \in [0, T]$, $\Lambda \in \mathcal{P}$, and $u \in H^1$.

(Piii) For any sequence of matrices $\{\Lambda^n\}$ that converges to Λ^0 in \mathcal{P} ,

$$\tilde{\Gamma}^*(t, \Lambda^n) \rightarrow \tilde{\Gamma}^*(t, \Lambda^0),$$

in the strong operator topology of $\mathcal{L}(H^1, H^{-1})$.

(Piv) The mapping $\Lambda \rightarrow \tilde{\Gamma}^*(t, \Lambda)$ is once Gateaux differential in the strong operator topology of $\mathcal{L}(H^1, H^{-1})$, and the Gateaux differential is given by

$$\lim_{\epsilon \rightarrow 0} \frac{\tilde{\Gamma}^*(t, \Lambda^0 + \epsilon \Lambda)u - \tilde{\Gamma}^*(t, \Lambda^0)u}{\epsilon} = \Lambda u, \quad (6.59)$$

for all $t \in [0, T]$; $\Lambda^0, \Lambda \in \mathcal{P}$; $0 \leq \epsilon \leq 1$ and $u \in H^1$.

Again with the help of the properties (Pi) and (Pii), given above, it follows from Lemma 2.4 that for each $\Lambda \in \mathcal{P}$, the initial value problem (6.54) has a unique (weak) solution $V(t, \Lambda) \in L_2((0, T); H^1) \cap C([0, T]; (L_2(R^n))^N)$.

Using equations (6.51), (6.52) and (6.54) and assuming that $h \in C_b^2(R^n) \cap H^1(R^n)$ one can state the identification problem (I3) as follows.

Problem (I3)'

Given $\mathcal{F}_t^y, t \in [0, T]$, find a $\Lambda^0 \in \mathcal{P}$, such that $J(\Lambda^0) \geq J(\Lambda)$ for all $\Lambda \in \mathcal{P}$, where

$$J(\Lambda) \equiv \Theta(T, \Lambda) \cdot \int_0^T \langle V(t, \Lambda), \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}}, \quad (6.60)$$

and $\Theta(T, \Lambda)$ is given by (6.43) and $V(t), t \in [0, T]$, is the solution of (6.54).

We now present the following result which shows that the identification problem (I3)' has a solution.

Theorem 6.3

Consider the identification problem (I3)' and suppose that the operator $\tilde{\Gamma}^*$ satisfy the properties (Pi)-(Piii) given above. Let the process $y_t, t \in [0, T]$, satisfy the property (6.58) and the function $h \in C_b^2(R^n) \cap H^1(R^n)$. Then the mapping $\Lambda \rightarrow J(\Lambda)$, where J is given by (6.60), is continuous on \mathcal{P} and the problem (I3)' has a solution.

Proof

Let $V(t, \Lambda^n) \equiv V^n(t)$ and $V(t, \Lambda^0) \equiv V^0(t), t \in [0, T]$, be the solutions of the initial value problem (6.54) corresponding to Λ^n and Λ^0 , respectively. Then following similar arguments as those of Theorem 6.1, it is easy to verify that $V^n(t) \rightarrow V^0(t)$ in $(L_2(R^n))^N$, uniformly in $t \in [0, T]$, whenever $\Lambda^n \rightarrow \Lambda^0$ in \mathcal{P} .

Define

$$J(\Lambda^n) \equiv \Theta(T, \Lambda^n) \cdot \int_0^T \langle V^n(t), \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}},$$

and

$$J(\Lambda^0) \equiv \Theta(T, \Lambda^0) \cdot \int_0^T \langle V^0(t), \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}}.$$

Then

$$\begin{aligned} & \lim_n (J(\Lambda^n) - J(\Lambda^0)) \\ &= \lim_n \Theta(T, \Lambda^n) \cdot \int_0^T \langle V^n(t) - V^0(t), \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}} \\ & \quad + \lim_n (\Theta(T, \Lambda^n) - \Theta(T, \Lambda^0)) \cdot \int_0^T \langle V^0(t), \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}}. \end{aligned}$$

Since by assumption, given \mathcal{F}_t^y , the function $h \exp(h \cdot y_t), t \in [0, T]$, is in

$H^{-1}(R^n)$; $\Theta(T, \Lambda^n)$ is bounded for all n ; $\Theta(T, \Lambda^n) \rightarrow \Theta(T, \Lambda^0)$ and $V^n \rightarrow V^0$ in $L_2((0, T); H^1)$, it follows, by standard properties of stochastic integrals, that the right hand side of the above equation converges in the mean in the limit. Hence by similar arguments as those of Theorem 6.1, the continuity of J follows. The proof of the Theorem now follows by the compactness of the set \mathcal{P} . ■

In the next section we shall use the variational method and the Gateaux differentiability of the process $V(t), t \in [0, T]$, (see equations (6.15), (6.36) and (6.54)) to derive the necessary conditions of optimality for the identification problems (I1)', (I2)' and (I3)'.

6.4 Necessary Conditions For Optimal Identification

In this section we present the necessary conditions of optimality for the identification problems (I1)', (I2)' and (I3)' as stated in the preceding section. Our derivation of these necessary conditions is mainly based on the variational method and the Gateaux differentiability of the process $V(t), t \geq 0$, (see equations (6.15), (6.36) and (6.54)) which are the standard tools in optimal control problems [3,5,6,7,12].

6.4.1 Identification Problem (I1)'

Consider the identification problem (I1)' and suppose that the unknown parameter vector α takes values in a compact convex set Σ . In this section we shall make use of the Gateaux differential of V^α , with respect to α , (see (6.15)), to derive the corresponding necessary conditions of optimality from which the optimal parameter α^0 can be determined. Indeed we show that the Gateaux differential of V at α^0 in the direction α , defined as

$$\bar{V}(t, \alpha^0; \alpha) \equiv \lim_{\varepsilon \rightarrow 0} \frac{V^{\alpha^\varepsilon} - V^{\alpha^0}}{\varepsilon}, \quad \alpha^\varepsilon \equiv \alpha^0 + \varepsilon \alpha,$$

exists and that it is the solution of a related differential equation. In this regard, we have the following result.

Lemma 6.1

Let $V^\alpha, \alpha \in \Sigma$, denote the solution of the equation (6.15). Then at each point $\alpha^0 \in \Sigma$, the mapping $\alpha \rightarrow V^\alpha$ has a Gateaux differential in the direction $\alpha - \alpha^0$, denoted $\bar{V}(t, \alpha^0; \alpha - \alpha^0)$, and it is the solution of the following differential equation

$$\begin{aligned} \frac{d}{dt} \tilde{e}(t) &= F^*(t, \alpha^0) \tilde{e}(t) + \bar{F}^*(t, \alpha^0; \alpha - \alpha^0) V^0(t), \quad t \in [0, T], \\ \tilde{e}(0) &= 0, \end{aligned} \tag{6.61}$$

where $\bar{F}^*(t, \alpha^0; \alpha - \alpha^0)$ denotes the Gateaux differential of the operator F^* (see equation (6.16)) at α^0 in the direction $\alpha - \alpha^0$, and $V^0(t), t \in [0, T]$, is the solution of equation (6.15) corresponding to α^0 .

Proof

Let $\alpha, \alpha^0 \in \Sigma$. Since Σ is convex, $\alpha^\varepsilon \equiv \alpha^0 + \varepsilon(\alpha - \alpha^0) \in \Sigma$, for all $\varepsilon \in [0, 1]$. For each $u \in H^1$ and $t \in [0, T]$, define

$$\bar{F}^*(t, \alpha^0; \alpha - \alpha^0)u \equiv \lim_{\varepsilon \rightarrow 0} \frac{F^*(t, \alpha^\varepsilon)u - F^*(t, \alpha^0)u}{\varepsilon}, \quad (6.62)$$

and

$$\tilde{V}^\varepsilon(t) \equiv \frac{1}{\varepsilon}(V^{\alpha^\varepsilon}(t) - V^{\alpha^0}(t)) \equiv \frac{1}{\varepsilon}(V^\varepsilon(t) - V^0(t)). \quad (6.63)$$

Then using equations (6.15) and (6.63), we obtain

$$\frac{d}{dt} \tilde{V}^\varepsilon(t) - F^*(t, \alpha^\varepsilon) \tilde{V}^\varepsilon(t) = \frac{1}{\varepsilon}(F^*(t, \alpha^\varepsilon) - F^*(t, \alpha^0))V^0(t), \quad (6.64)$$

for all $t \in [0, T]$, where $\tilde{V}^\varepsilon(0) = 0$. We shall show that the Gateaux differential, given by the (weak) limit in $L_2((0, T); H^1)$, of \tilde{V}^ε , or a subsequence thereof, exists and it is the solution of equation (6.61).

Clearly, the right hand side of equation (6.64) is in $L_2((0, T); H^{-1})$ for all $\varepsilon \in (0, 1]$ and also as $\varepsilon \rightarrow 0$, the limit is well defined and equals

$$\bar{F}^*(t, \alpha^0; \alpha - \alpha^0)V^0(t) \in L_2((0, T); H^{-1}) \quad (\text{see property (P4)}).$$

Then using the properties (P1), (P2) and (P3) of the operator F^* and following similar arguments as those of Theorem 6.1, we arrive at the following estimate

$$\begin{aligned} & |\tilde{V}^\varepsilon(t)|_{L_2}^2 + \gamma \int_0^t |\tilde{V}^\varepsilon(\theta)|_{H^1}^2 d\theta \\ & \leq \exp(2\eta T) \frac{1}{\gamma} \int_0^T \left\| \frac{F^*(s, \alpha^\varepsilon) - F^*(s, \alpha^0)}{\varepsilon} V^0(s) \right\|_{H^{-1}}^2 ds, \end{aligned}$$

for all $t \in [0, T]$. Hence it follows from the above inequality and the property (P4) that the set $\{\tilde{V}^\varepsilon, \varepsilon \in [0, 1]\}$ is contained in a bounded subset of $L_2((0, T); H^1) \cap L_\infty([0, T]; L_2(R^n))$. Thus from every sequence $\tilde{V}^n \equiv \tilde{V}^{\varepsilon_n}$, with $\varepsilon_n \in [0, 1]$, one can extract a subsequence relabeled as \tilde{V}^n and there exists a $\tilde{V}^0 \in L_2((0, T); H^1)$ such that $\tilde{V}^n \rightarrow \tilde{V}^0$ (weakly) in $L_2((0, T); H^1)$. Hence the Gateaux differential of V exists and it is given by

$$\bar{V}(t, \alpha^0; \alpha - \alpha^0) = \tilde{V}^0(t).$$

It remains only to show that \tilde{V}^0 is the solution of the differential equation (6.61).

Indeed since for all $t \in [0, T]$, $F^*(t, \alpha^{\varepsilon_n}) \rightarrow F^*(t, \alpha^0)$ in the strong operator topology of $\mathcal{L}(H^1, H^{-1})$, and

$$\tilde{V}^n \rightarrow \tilde{V}^0 \quad (\text{weakly}) \text{ in } L_2((0, T); H^1),$$

it follows that

$$F^*(t, \alpha^{\varepsilon_n})\tilde{V}^n \rightarrow F^*(t, \alpha^0)\tilde{V}^0 \quad (\text{weakly}) \text{ in } L_2((0, T); H^{-1}).$$

Since by the property (F4)

$$\frac{1}{\varepsilon}(F^*(t, \alpha^{\varepsilon_n}) - F^*(t, \alpha^0))V^0(t) \rightarrow \bar{F}^*(t, \alpha^0; \alpha - \alpha^0)V^0(t),$$

in $L_2((0, T); H^{-1})$, it follows from (6.64) that $\frac{d}{dt}\tilde{V}^n(t) \in L_2((0, T); H^{-1})$ for all n and $\frac{d}{dt}\tilde{V}^n \rightarrow \Psi$ in $L_2((0, T); H^{-1})$ for suitable Ψ and Ψ is the distributional derivative of \tilde{V}^0 . Hence \tilde{V}^0 satisfies

$$\frac{d}{dt}\tilde{V}^0(t) = F^*(t, \alpha^0)\tilde{V}^0(t) + \bar{F}^*(t, \alpha^0; \alpha - \alpha^0)V^0(t),$$

in the sense of distribution in H^{-1} .

Since $\tilde{V}^0 \in L_2((0, T); H^1)$, $\frac{d}{dt}\tilde{V}^0 \in L_2((0, T); H^{-1})$, and $\tilde{V}^0 \in C([0, T]; L_2(R^n))$ it is clear that $\tilde{V}^0(0)$ is well defined and equals $\tilde{V}^n(0) = 0$, for all $\varepsilon_n \in [0, 1]$. Hence

\tilde{V}^0 satisfies the differential equation (6.61) (in the sense of distribution) and one may identify \tilde{V}^0 as \tilde{z} . This completes the proof. ■

With the help of the above Lemma, we now derive the necessary conditions of optimality for the identification problem (I1)'.

Theorem 6.4

Consider the identification problem (I1)' and suppose that the hypotheses of Theorem 6.1 hold. Then the maximum likelihood estimate α^0 of the unknown parameter vector α is determined by the simultaneous solution of the system equation

$$\begin{aligned} \frac{d}{dt}V(t) &= F^*(t, \alpha^0)V(t), & t \in [0, T], \\ V(0) &= p_0, \end{aligned} \tag{6.65}$$

the adjoint equation

$$\begin{aligned} -\frac{d}{dt}\Psi(t) &= F(t, \alpha^0)(\Psi(t) - \exp(h \cdot Z_t)), & t \in [0, T], \\ \Psi(T) &= \exp(h \cdot Z_T), \end{aligned} \tag{6.66}$$

and the inequality

$$J'_0(\alpha^0 - \alpha) = \int_0^T \langle \bar{F}^*(t, \alpha^0; \alpha - \alpha^0) V^0(t), \Psi^0(t) - \exp(h \cdot Z_t) \rangle_{H^{-1}-H^1} dt \leq 0, \quad (6.67)$$

for all $\alpha \in \Sigma$, where F denotes the formal adjoint of the operator F^* and \bar{F}^* is the Gateaux differential. Further, V^0 and Ψ^0 are the solutions of equations (6.65) and (6.66), respectively.

Proof

Since by Lemma 6.1, the mapping $\alpha \rightarrow V^\alpha$ is Gateaux differentiable on Σ , it follows that J , as defined by equation (6.19), has also a Gateaux differential. Then in order that J attains its maximum at $\alpha^0 \in \Sigma$, it is necessary that

$$J'_0(\alpha - \alpha^0) \equiv \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \{J(\alpha^\varepsilon) - J(\alpha^0)\} \leq 0, \quad (6.68)$$

where

$$\alpha^\varepsilon \equiv \alpha^0 + \varepsilon(\alpha - \alpha^0).$$

Defining $V^{\alpha^\varepsilon}(t) \equiv V^\varepsilon(t)$ and $V^{\alpha^0}(t) \equiv V^0(t)$, it follows from equations (6.19) and (6.68) that

$$0 \geq J'_0(\alpha - \alpha^0) \equiv \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^T dZ_t \cdot \langle V^\varepsilon(t) - V^0(t), h \exp(h \cdot Z_t) \rangle_{H^{-1}-H^1}.$$

Using the facts that given \mathcal{F}_t^y , the function $h \exp(h \cdot Z_t), t \in [0, T]$, is in $H^{-1}(R^n)$,

$$V^\varepsilon - V^0 \in L_2((0, T); H^1) \cap L_\infty([0, T]; L_2(R^n))$$

and

$$\frac{1}{\varepsilon} (V^\varepsilon(t) - V^0(t)) \rightarrow \bar{V}(t, \alpha^0; \alpha - \alpha^0), \quad (\text{see Lemma 6.1}),$$

it follows by the properties of stochastic integral that

$$0 \geq J'_0(\alpha - \alpha^0) = \int_0^T dZ_t \cdot \langle \bar{V}(t, \alpha^0; \alpha - \alpha^0), h \exp(h \cdot Z_t) \rangle_{H^1 - H^{-1}}, \quad (6.69)$$

\mathcal{F}_T^y -a.s., for all $\alpha \in \Sigma$. The inequality (6.69) can be further simplified by introducing the so called adjoint variable Ψ , which is the solution of the following differential equation

$$\begin{aligned} -\frac{d}{dt}\Psi(t) &= F(t, \alpha^0)(\Psi(t) - \exp(h \cdot Z_t)), \quad t \in [0, T], \\ \Psi(T) &= \exp(h \cdot Z_T), \end{aligned} \quad (6.70)$$

where F is the formal adjoint of the operator F^* given by (6.16). Reversing the time $t \rightarrow T - t$, and noting that the operator F satisfies the properties (P1) and (P2), it follows from Lemma 2.4 that the equation (6.70) has a unique (weak) solution $\Psi \in L_2((0, T); H^1) \cap C([0, T]; L_2(R^n))$.

Since

$$\begin{aligned} 0 &= \int_0^T d \langle \bar{V}(\alpha^0; \alpha - \alpha^0), \Psi(t) - \exp(h \cdot Z_t) \rangle \\ &= \int_0^T \langle \bar{V}(\alpha^0; \alpha - \alpha^0), d\Psi(t) \rangle - \int_0^T dZ_t \cdot \langle \bar{V}(\alpha^0; \alpha - \alpha^0), h \exp(h \cdot Z_t) \rangle \\ &\quad + \int_0^T \langle d\bar{V}(\alpha^0; \alpha - \alpha^0), \Psi(t) - \exp(h \cdot Z_t) \rangle, \end{aligned} \quad (6.71)$$

it follows from (6.69) that

$$\begin{aligned} J'_0(\alpha - \alpha^0) &= \int_0^T \langle d\bar{V}(\alpha^0; \alpha - \alpha^0), \Psi(t) - \exp(h \cdot Z_t) \rangle \\ &\quad + \int_0^T \langle \bar{V}(\alpha^0; \alpha - \alpha^0), d\Psi(t) \rangle. \end{aligned} \quad (6.72)$$

Utilizing the result of Lemma 6.1 and using equation (6.70), it is clear that the inequality (6.67) can be obtained from equation (6.72). This completes the proof.

■

In the next two subsections we shall use similar arguments as those of Lemma 6.1 and Theorem 6.1 to obtain the necessary conditions of optimality for the identification problems (I2)' and (I3)'. Some of the proofs of these results are similar to those given above and hence omitted.

6.4.2 Identification Problem (I2)'

Lemma 6.2

Let V^β denote the solution of the initial value problem (6.36) corresponding to $\beta \in \Sigma$. Then at each point $\beta^0 \in \Sigma$, the function $\beta \rightarrow V^\beta$ has a Gateaux differential in the direction $\beta - \beta^0$, denoted $\bar{V}(t, \beta^0; \beta - \beta^0)$, and it is the solution of the following differential equation

$$\begin{aligned} \frac{d}{dt} \tilde{z}(t) &= D^*(t; \beta^0) \tilde{z}(t) + \bar{D}^*(t, \beta^0; \beta - \beta^0) V^0(t), & t \in [0, T], \\ \tilde{z}(0) &= 0, \end{aligned} \tag{6.73}$$

where \bar{D}^* denotes the Gateaux differential of the operator D^* (see (6.33)) and $V^0(t), t \in [0, T]$, is the solution of (6.36) corresponding to β^0 .

Proof

The proof essentially follows from similar arguments as those of Lemma 6.1

■

Theorem 6.5

Consider the identification problem (I2)' and suppose that the hypotheses of Theorem 6.2 hold. Then the maximum likelihood estimate β^0 of the unknown parameter vector β is determined by the simultaneous solution of the system equation

$$\begin{aligned} \frac{d}{dt}V(t) &= D^*(t, \beta^0)V(t), \\ V(0) &= p_0, \end{aligned} \quad (6.74)$$

the adjoint equation

$$\begin{aligned} -\frac{d}{dt}\Psi(t) &= D(t, \beta^0)(\Psi(t) - e(t, \beta^0)), \\ \Psi(T) &= e(T, \beta^0), \end{aligned} \quad (6.75)$$

and the inequality

$$\begin{aligned} J'_0(\beta - \beta^0) &= \int_0^T \{ \langle \bar{D}^*(t, \beta^0; \beta - \beta^0)V^0(t), \Psi^0(t) - e(t, \beta^0) \rangle_{H^{-1}-H^1} dt \\ &\quad + dZ_t \cdot \langle V^0(t), \bar{g}(\beta^0; \beta - \beta^0)e(t, \beta^0) \rangle_{H^1-H^{-1}} \\ &\quad + dZ_t \cdot \langle V^0(t), g(\beta^0)\bar{e}(t, \beta^0; \beta - \beta^0) \rangle_{H^1-H^{-1}} \} \leq 0, \end{aligned} \quad (6.76)$$

where

$$e(t, \beta^0) = \exp(g(\beta^0) \cdot Z_t), \quad (6.76.1)$$

and D is the formal adjoint of the operator D^* . Further, \bar{D}^* , \bar{g} , \bar{e} are the Gateaux differentials of D^* , g , e , respectively and $V^0(t)$ and $\Psi^0(t)$, $t \in [0, T]$, are the solutions of equations (6.74) and (6.75).

Proof

Since by Lemma 6.2, the function $\beta \rightarrow V^\beta$ is Gateaux differentiable on Σ , it follows that J , as defined by (6.38), has also a Gateaux differential. Then in order that J attains its maximum at β^0 , it is necessary that

$$J'_0 \equiv \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \{ J(\beta^0 + \epsilon(\beta - \beta^0)) - J(\beta^0) \} \leq 0, \quad (6.77)$$

for all $\beta \in \Sigma$. Defining

$$\beta^\varepsilon \equiv \beta^0 + \varepsilon(\beta - \beta^0),$$

$$V^\varepsilon(t) \equiv V^{\beta^\varepsilon}(t),$$

$$V^0(t) \equiv V^{\beta^0}(t),$$

$$e(t, \beta^\varepsilon) \equiv \exp(g(\beta^\varepsilon) \cdot Z_t),$$

and using (6.38), it follows from (6.77) that

$$\begin{aligned} 0 \geq J'_0(\beta - \beta^0) &= \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^T dZ_t \cdot \langle V^\varepsilon(t) - V^0(t), g(\beta^\varepsilon)e(t, \beta^\varepsilon) \rangle_{H^1-H^{-1}} \\ &+ \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^T dZ_t \cdot \langle V^0(t), (g(\beta^\varepsilon) - g(\beta^0))e(t, \beta^\varepsilon) \rangle_{H^1-H^{-1}} \\ &+ \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^T dZ_t \cdot \langle V^0(t), g(\beta^0)(e(t, \beta^\varepsilon) - e(t, \beta^0)) \rangle_{H^1-H^{-1}}. \end{aligned} \quad (6.78)$$

Since by assumption, for all $\beta \in \Sigma, t \in [0, T]$, the function $\beta \rightarrow g(\beta)e(t, \beta)$ is in $H^{-1}(R^n)$, given \mathcal{F}_t^y , the mapping $\beta \rightarrow g(\beta)$ is continuous and once differentiable on Σ , and

$$\frac{1}{\varepsilon}(V^\varepsilon(t) - V^0(t)) \rightarrow \bar{V}(t, \beta^0; \beta - \beta^0), \quad (\text{weakly in } L_2((0, T); H^1),$$

it follows by standard properties of stochastic integral and similar arguments as those of Theorem 6.1 that

$$\begin{aligned} 0 \geq J'_0(\beta - \beta^0) &= \int_0^T dZ_t \cdot \langle \bar{V}(t, \beta^0; \beta - \beta^0), g(\beta^0)e(t, \beta^0) \rangle_{H^1-H^{-1}} \\ &+ \int_0^T dZ_t \cdot \langle V^0(t), \bar{g}(\beta^0; \beta - \beta^0)e(t, \beta^0) \rangle_{H^1-H^{-1}} \\ &+ \int_0^T dZ_t \cdot \langle V^0(t), g(\beta^0)\bar{e}(t, \beta^0; \beta - \beta^0) \rangle_{H^1-H^{-1}}, \end{aligned} \quad (6.79)$$

\mathcal{F}_T^y almost surely, where

$$\bar{g}(\beta^0; \beta - \beta^0) \equiv \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (g(\beta^\varepsilon) - g(\beta^0)),$$

and

$$\bar{e}(t, \beta^0; \beta - \beta^0) \equiv \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} (e(t, \beta^\varepsilon) - e(t, \beta^0)).$$

The inequality (6.79) can be further simplified by introducing the adjoint variable Ψ , which is the solution of the following differential equation

$$\begin{aligned} -\frac{d}{dt} \Psi(t) &= D(t, \beta^0)(\Psi(t) - e(t, \beta^0)), \\ \Psi(T) &= e(T, \beta^0), \end{aligned} \tag{6.80}$$

where D is the formal adjoint of the operator D^* (see (6.33)) and $e(t, \beta), t \in [0, T], \beta \in \Sigma$, as defined above. Again, by Lemma 2.4, the above equation has a unique (weak) solution $\Psi \in L_2((0, T); H^1) \cap C([0, T]; L_2(R^n))$.

Using similar arguments as those of Theorem 6.4, equation (6.79) can be written as

$$\begin{aligned} 0 \geq J'_0(\beta - \beta^0) &= \int_0^T \langle d\bar{V}(t, \beta^0; \beta - \beta^0), \Psi(t) - e(t, \beta^0) \rangle_{H^{-1}-H^1} \\ &+ \int_0^T \langle \bar{V}(t, \beta^0; \beta - \beta^0), d\Psi(t) \rangle_{H^1-H^{-1}} \\ &+ \int_0^T dZ_t \langle V^0(t), \bar{V}(\beta^0; \beta - \beta^0) e(t, \beta^0) \rangle_{H^1-H^{-1}} \\ &+ \int_0^T dZ_t \langle V^0(t), g(\beta^0) \bar{e}(t, \beta^0; \beta - \beta^0) \rangle_{H^1-H^{-1}}. \end{aligned} \tag{6.81}$$

Utilizing the result of Lemma 6.2 and using equation (6.80), it is easy to verify that equation (6.76) follows from (6.81). This completes the proof. ■

6.4.3 Identification Problem (I3)'

Lemma 6.3

Let $V(t, \Lambda), t \in [0, T], \Lambda \in \mathcal{P}$, be the solution of the equation (6.54). Then at each point $\Lambda^0 \in \mathcal{P}$, the function $\Lambda \rightarrow V(t, \Lambda), t \in [0, T]$, has a Gateaux differential in the direction $\Lambda - \Lambda^0$, denoted $\bar{V}(t, \Lambda^0; \Lambda - \Lambda^0)$, and it is the solution of the following differential equation

$$\begin{aligned} \frac{d}{dt} \tilde{z}(t) &= \tilde{\Gamma}^*(t, \Lambda^0) \tilde{z}(t) + (\Lambda - \Lambda^0) V^0(t), \\ \tilde{z}(0) &= 0, \end{aligned} \tag{6.82}$$

for all $t \in [0, T]$, where $V^0(t) \equiv V(t, \Lambda^0)$, denotes the solution of (6.54) corresponding to Λ^0 .

Proof

The proof essentially follows from similar arguments as those of Lemmas 6.1 and 6.2 and the fact that (see equation (6.59))

$$\lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \{ \tilde{\Gamma}^*(t, \Lambda^0 + \varepsilon(\Lambda - \Lambda^0)) - \tilde{\Gamma}^*(t, \Lambda^0) \} u = (\Lambda - \Lambda^0) u,$$

for all $u \in H^1$, where H^1 is defined by (6.51.1). ■

Theorem 6.6

Consider the identification problem (I3)' and suppose that the hypotheses of Theorem 6.3 hold. Let the process ξ_t and the corresponding transition probability matrix $S(t), t \in [0, T]$, satisfy the properties given in sections 6.2 and 6.3. Then the maximum likelihood estimate Λ^0 of the unknown matrix Λ is determined by the

simultaneous solution of the system equation

$$\begin{aligned} \frac{d}{dt} V(t) &= \tilde{\Gamma}^*(t, \Lambda^0) V(t), \quad t \in [0, T], \\ V(0) &= \Phi_0, \end{aligned} \quad (6.83)$$

the adjoint equation

$$\begin{aligned} -\frac{d}{dt} \Psi(t) &= \tilde{\Gamma}(t, \Lambda^0) (\Psi(t) - q(t)), \quad t \in [0, T], \\ \Psi(T) &= q(T), \end{aligned} \quad (6.84)$$

and the inequality

$$\begin{aligned} J'_0(\Lambda - \Lambda^0) &= \int_0^T \langle (\Lambda - \Lambda^0) V^0(t), \Psi^0(t) - \hat{q}(t) \rangle_{H^{-1}-H^1} dt \\ &+ \int_0^T \langle V^0(t), \bar{\Theta}(T, \Lambda^0; \Lambda - \Lambda^0) \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}} \leq 0, \end{aligned} \quad (6.85)$$

where

$$q(t) \equiv \Theta(T, \Lambda^0) \exp(h \cdot y_t), \quad (6.86)$$

$$\bar{\Theta}(T, \Lambda^0; \Lambda - \Lambda^0) \equiv T \exp(T \Lambda^0) (\Lambda - \Lambda^0) \Theta_0, \quad (6.87)$$

and $\Theta_0 \equiv \Pr\{\xi(0) = \xi_0\}$. Further, $V^0(t)$ and $\Psi^0(t)$, $t \in [0, T]$, are the solutions of equations (6.83) and (6.84), respectively.

Proof

Since by Lemma 6.3, the mapping $\Lambda \rightarrow V(t, \Lambda)$ is Gateaux differentiable, it follows that J , as given by equation (6.60), has also a Gateaux differential on \mathcal{P} . Then in order that J attains its maximum at Λ^0 , it is necessary that

$$J'_0(\Lambda - \Lambda^0) \equiv \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \{J(\Lambda^\varepsilon) - J(\Lambda^0)\} \leq 0, \quad (6.88)$$

for all $\Lambda \in \mathcal{P}$, where $\Lambda^\varepsilon \equiv \Lambda^0 + \varepsilon(\Lambda - \Lambda^0)$. Defining

$$V^\epsilon(t) \equiv V(t, \Lambda^\epsilon)(t),$$

and

$$V^0(t) \equiv V(t, \Lambda^0), \quad t \in [0, T],$$

and using equation (6.60), it follows from equation (6.88) that

$$0 \geq J'_0(\Lambda - \Lambda^0)$$

$$= \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_0^T \langle V^\epsilon(t) - V^0(t), \Theta(T, \Lambda^\epsilon) \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}} \quad (6.89)$$

$$+ \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_0^T \langle V^0(t), (\Theta(T, \Lambda^\epsilon) - \Theta(T, \Lambda^0)) \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}}.$$

Since $h \in C_b^2(\mathbb{R}^n) \cap H^1(\mathbb{R}^n)$, $\Theta(T, \Lambda^\epsilon) \rightarrow \Theta(T, \Lambda^0)$ on \mathcal{P} , and

$$\frac{1}{\epsilon} (V^\epsilon(t) - V^0(t)) \rightarrow \bar{V}(t, \Lambda^0; \Lambda - \Lambda^0), \quad (\text{weakly in } L_2([0, T]; H^1),$$

following similar arguments as those given in Theorems 6.4 and 6.5, it follows that

the first term in the right hand side of the equation (6.89) is given by

$$\begin{aligned} & \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_0^T \langle V^\epsilon(t) - V^0(t), \Theta(T, \Lambda^\epsilon) \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}} \\ &= \int_0^T \langle \bar{V}(t, \Lambda^0; \Lambda - \Lambda^0), \Theta(T, \Lambda^0) \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}}, \end{aligned} \quad (6.90)$$

\mathcal{F}_T^y almost surely. As indicated earlier, the probability measure $\Theta(t, \Lambda)$, corresponding to the Markov chain $\xi_t, t \in [0, T]$, is given by

$$\Theta(t, \Lambda) = \exp(t\Lambda)\Theta_0, \quad t \in [0, T]. \quad (6.91)$$

Then the Gateaux differential of Θ is given by

$$\begin{aligned} \bar{\Theta}(T, \Lambda^0; \Lambda - \Lambda^0) &\equiv \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \{ \exp(T\Lambda^\epsilon) - \exp(T\Lambda^0) \} \Theta_0 \\ &= \exp(T\Lambda^0) \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \{ \exp(T\epsilon(\Lambda - \Lambda^0)) - I \} \Theta_0 \\ &= T \exp(T\Lambda^0) (\Lambda - \Lambda^0) \lim_{\epsilon \rightarrow 0} \exp(T\epsilon(\Lambda - \Lambda^0)) \Theta_0 \\ &= T \exp(T\Lambda^0) (\Lambda - \Lambda^0) \Theta_0. \end{aligned} \quad (6.92)$$

Hence following similar arguments as before, it is easy to verify that the second term in the right hand side of equation (6.89), is given by

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \int_0^T \langle V^0(t), (\Theta(T, \Lambda^\varepsilon) - \Theta(T, \Lambda^0)) \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}} \\ = \int_0^T \langle V^0(t), \bar{\Theta}(T, \Lambda^0; \Lambda - \Lambda^0) \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}}, \end{aligned} \quad (6.93)$$

\mathcal{F}_T^y almost surely, where $\bar{\Theta}$ is given by equation. Using equations (6.90) and (6.93); it follows from equation (6.89) that

$$\begin{aligned} 0 \geq J'_0(\Lambda - \Lambda^0) = \int_0^T \langle \bar{V}(t, \Lambda^0; \Lambda - \Lambda^0), \Theta(T, \Lambda^0) \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}} \\ + \int_0^T \langle V^0(t), \bar{\Theta}(T, \Lambda^0; \Lambda - \Lambda^0) \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}}. \end{aligned} \quad (6.94)$$

The equation (6.94) can be further simplified by introducing the following adjoint equation

$$\begin{aligned} -\frac{d}{dt} \Psi(t) &= \tilde{\Gamma}(t, \Lambda^0)(\Psi(t) - q(t)), \quad t \in [0, T], \\ \Psi(T) &= q(T), \end{aligned} \quad (6.95)$$

where

$$q(t) \equiv \Theta(T, \Lambda^0) \exp(h \cdot y_t). \quad (6.96)$$

Since the operator $\tilde{\Gamma}$ satisfies the properties (i) and (ii) of section 6.3.3, it follows by Lemma 2.4 that the system (6.95) has a unique (weak) solution $\Psi \in L_2((0, T); H^1) \cap C([0, T]; (L_2(R^n))^N)$. Following similar arguments as those of Theorems 6.4 and 6.5, equation (6.94) can be written as

$$\begin{aligned}
 0 \geq J'_0(\Lambda - \Lambda^0) &= \int_0^T \langle d\bar{V}(t, \Lambda^0; \Lambda - \Lambda^0), \Psi(t) - q(t) \rangle_{H^{-1}-H^1} \\
 &+ \int_0^T \langle \bar{V}(t, \Lambda^0, \Lambda - \Lambda^0), d\Psi(t) \rangle_{H^1-H^{-1}} \\
 &+ \int_0^T \langle V^0(t), \bar{\Theta}(T, \Lambda^0; \Lambda - \Lambda^0) \exp(h \cdot y_t) h \cdot dy_t \rangle_{H^1-H^{-1}}.
 \end{aligned}
 \tag{6.97}$$

Utilizing the result of Lemma 6.3 and using equation (6.95), it is easy to verify that equation (6.85) follows from (6.97). This completes the proof. ■

Based on the above necessary conditions, we develop in the next section a numerical algorithm for computing the maximum likelihood estimates of the unknowns α, β , and Λ corresponding to the problems (I1)', (I2)' and (I3)'. Then we present two examples to illustrate some of the results of this chapter.

6.5 Computational Algorithm and Examples

Based on the necessary conditions presented in the preceding section, an iterative procedure can be developed for determining the maximum likelihood estimates of the unknowns. For simplicity of presentation, we discuss the contents of this algorithm with reference to Theorem 6.4 ; but the same algorithm holds for Theorems 6.5 and 6.6 if references are made to the appropriate equations. In this algorithm, we have used the IMSL subroutine GGNML for generating the Wiener processes W and \tilde{W} and with the help of Runge-Kutta method we have obtained the state process x_t and the output process y_t , for all $t \in [0, T]$. Then using finite difference scheme [65] and Gradient method [23,73], we have used the necessary conditions obtained in the previous section to compute the required estimates of the unknowns.

Algorithm

1. Set $i = 1$, and guess $\alpha^i \in \Sigma$ for α^0 .
2. Solve the differential equations (6.1) and (6.2), with $\alpha^0 = \alpha^i$, using Runge-Kutta method and obtain the observed path $\{y_s; 0 \leq s \leq T < \infty\}$.
3. Solve the system equation (6.65) with $\alpha^0 = \alpha^i$.
4. Compute the objective function $J(\alpha^i)$ using (6.19).
5. Solve the adjoint equation (6.66) with $\alpha^0 = \alpha^i$.
6. Compute the gradient, denoted J'_{α^i} , with the help of inequality (6.67) (see example 6.1).
7. Update the parameter by setting

$$\alpha^{i+1} = \alpha^i - \epsilon J'_{\alpha^i},$$

where $\epsilon > 0$ is chosen sufficiently small so that

$$J(\alpha^{i+1}) \geq J(\alpha^i), \quad \text{and } \alpha^{i+1} \in \Sigma.$$

8. Set $i = i + 1$ and repeat from step 2 until the following convergence criterion is satisfied

$$|J(\alpha^{i+1}) - J(\alpha^i)| \leq \rho,$$

where $\rho > 0$ is chosen sufficiently small.

Remark 6.2

As in any gradient method, the iterated parameter would tend to seek a local minimum. This could be avoided, however, by repeating the above procedure for different initial guesses.

Example 6.1

Let the state process x_t and the output process y_t , $t \in [0, T]$, be governed by the following (scalar) stochastic differential equations

$$dx_t = (\alpha \cdot a(t, x))dt + b(t, x)dW_t, \quad (6.98)$$

and

$$dy_t = h(x_t)dt + dV_t, \quad (6.99)$$

where x, b, y, h are scalars ; $a \in R^m$ and the unknown parameter vector α takes values from a compact convex subset of R^m . The processes W and V are independent one dimensional standard Wiener processes. Assuming that the functions a, b and h satisfy our basic assumptions of section 6.2, our problem is to determine the unknown parameter vector α , given the σ -field \mathcal{F}_t^y .

Comparing equations (6.2) and (6.99) and using (6.14), it is clear that $Z_t \equiv y_t$. Integrating (6.16) by parts and using (6.17) and (6.18), one can easily verify that the operator F^* is given by

$$\begin{aligned}
 F^*u &= \frac{1}{2}\sigma u_{xx} + u_x\{\sigma h_x y_t + \sigma_x - \alpha \cdot a\} \\
 &\quad + u\left\{\frac{1}{2}\sigma_{xx} - (\alpha \cdot a)h_x y_t + \frac{1}{2}\sigma_x h_x y_t - \alpha \cdot a_x\right. \\
 &\quad \left. + \frac{1}{2}\sigma(h_x y_t)^2 + \frac{1}{2}\sigma h_{xx} y_t - \frac{1}{2}h^2\right\}, \tag{6.100}
 \end{aligned}$$

where a_x denotes the partial derivative of a with respect to x and $\sigma(t, x) \equiv b^2(t, x)$. Using equation (6.100), it is clear that the Gateaux differential \bar{F}^* of the operator F^* , is given by

$$\bar{F}^*(t, \alpha^0; \alpha - \alpha^0)u = (\alpha - \alpha^0) \cdot B^*(t)u, \tag{6.101}$$

where

$$B^*(t)u \equiv -u_x a - u\{a h_x y_t + a_x\}. \tag{6.102}$$

Substituting (6.101) in the inequality (6.67), we obtain

$$\begin{aligned}
 0 &\geq J'_0(\alpha - \alpha^0) \\
 &= (\alpha - \alpha^0) \cdot \int_0^T \langle B^*(t)V^0(t), \Psi^0(t) - \exp(hy_t) \rangle_{H^{-1}-H^1} dt, \tag{6.103}
 \end{aligned}$$

from which one can identify the gradient J'_{α^0} as

$$J'_{\alpha^0} = \int_0^T \langle B^*(t)V^0(t), \Psi^0(t) - \exp(hy_t) \rangle_{H^{-1}-H^1} dt, \tag{6.104}$$

where B^* is given by (6.102), V^0 and Ψ^0 are the solutions of the differential equations (6.65) and (6.66), respectively. Note that the gradient J'_{α^0} , as given by (6.104), is used in step 7 of the above algorithm to update the iterated parameter (see steps 6 and 7 in the algorithm).

For the numerical simulations, we have assumed that the functions a, b and h are given by

$$a(t, x) = (x^2, x)',$$

$$b(t, x) = 0.5x,$$

$$h(x) = 2x^2,$$

and $\alpha = (\alpha_1, \alpha_2)'$. Further, the true values of the parameters α_1 and α_2 were taken as 1.0 and 1.0, respectively, and the final time $T = 1.0$. Starting with initial guesses $\alpha_1 = 2.0$ and $\alpha_2 = 0.5$, we have used the above algorithm for computing the estimated parameters, and the results are summarized in Table I given below.

Table I

Iteration No.	α_1	α_2	$J(\alpha_1, \alpha_2)$
0	2.0	0.5	0.244632
10	1.818024	0.562947	3.095176
20	1.626672	0.628594	6.097868
30	1.429875	0.695021	12.144006
40	1.232492	0.760402	19.546665
50	1.040175	0.823527	30.403365
78	0.981176	0.998050	115.095427
79	0.983153	0.998741	115.122290
true	1.0	1.0	—

Example 6.2

Let the processes $\{(x_t, y_t); t \in [0, T]\}$ be governed by the following (scalar) stochastic differential equations

$$dx_t = f(t, x_t)dt + b(t, x_t)dW_t, \quad (6.105)$$

and

$$dy_t = (\beta \cdot \gamma(x_t))dt + dV_t, \quad (6.106)$$

where x, f, b, y are scalars ; $\gamma \in R^m$ and the unknown parameter vector β takes values from some compact convex subset of R^m . Assuming that the processes W and V are as in example 6.1, our problem is to determine the unknown parameter vector β given the σ -field \mathcal{F}_t^y .

— Comparing equations (6.4) and (6.106), it is clear that

$$g(\beta, x_t) = \beta \cdot \gamma(x_t),$$

and

$$Z_t \equiv y_t.$$

Integrating equation (6.33) by parts and using equations (6.34) and (6.35), one can easily verify that the operator D^* is given by

$$\begin{aligned} D^*(t, \beta)u &= \frac{1}{2}\sigma u_{xx} + u_x\{\sigma\gamma_x y_t \cdot \beta + \sigma_x - f\} \\ &+ u\left\{\frac{1}{2}\sigma_{xx} - f\gamma_x y_t \cdot \beta + \frac{1}{2}\sigma_x \gamma_x y_t \cdot \beta - f_x\right. \\ &\left. + \frac{1}{2}\sigma(\gamma_x y_t \cdot \beta)^2 + \frac{1}{2}\sigma\gamma_{xx} y_t \cdot \beta - \frac{1}{2}(\gamma \cdot \beta)^2\right\}, \end{aligned} \quad (6.107)$$

where $\sigma(t, x) \equiv b^2(t, x)$.

Using equation (6.107), it follows that the Gateaux differential D^* of the operator D^* , is given by

$$D^*(t, \beta^0; \beta - \beta^0)u = (\beta - \beta^0) \cdot G^*(t, \beta^0)u, \quad (6.108)$$

where

$$\begin{aligned} G^*(t, \beta^0)u \equiv & u_x \sigma \gamma_x y_t + u \left\{ -f \gamma_x y_t + \frac{1}{2} \sigma_x \gamma_x y_t \right. \\ & + \frac{1}{2} \sigma (\gamma_x y_t \cdot \beta^0) \gamma_x y_t + \frac{1}{2} \sigma \gamma_{xx} y_t \\ & \left. + \frac{1}{2} (\gamma \cdot \beta^0) \gamma \right\}. \end{aligned} \quad (6.109)$$

Setting $g(\beta) = \gamma \cdot \beta$ and $Z_t = y_t$ in (6.76.1), we obtain

$$e(t, \beta) = \exp(\gamma \cdot \beta y_t). \quad (6.110)$$

Hence the Gateaux differentials \bar{g} and \bar{e} of the functions g and e are given by

$$\begin{aligned} \bar{g}(\beta^0; \beta - \beta^0) & \equiv \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \{ (\beta^0 + \epsilon(\beta - \beta^0)) \cdot \gamma - \beta^0 \cdot \gamma \} \\ & = (\beta - \beta^0) \cdot \gamma, \end{aligned} \quad (6.111)$$

and

$$\begin{aligned} \bar{e}(t, \beta^0; \beta - \beta^0) & \equiv \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \{ \exp \{ \beta^0 + \epsilon(\beta - \beta^0) \cdot \gamma y_t \} - \exp(\beta^0 \cdot \gamma y_t) \} \\ & \equiv \exp(\beta^0 \cdot \gamma y_t) \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \{ \exp \{ \epsilon(\beta - \beta^0) \cdot \gamma y_t \} - 1 \} \\ & = \exp(\beta^0 \cdot \gamma y_t) \gamma y_t \cdot (\beta - \beta^0). \end{aligned} \quad (6.112)$$

Thus using (6.110), (6.111) and (6.112), the inequality (6.76) reduces to

$$J'_0(\beta - \beta^0) = (\beta - \beta^0) \cdot J'_{\beta^0}, \quad (6.113)$$

where the gradient J'_{β^0} is given by

$$\begin{aligned} J'_{\beta^0} \equiv & \int_0^T \{ \langle G^*(t, \beta^0) V^0(t), \Psi^0(t) - \exp(\beta^0 \cdot \gamma y_t) \rangle dt \\ & + \langle V^0(t), \gamma \exp(\beta^0 \cdot \gamma y_t) \rangle dy_t \\ & + \langle V^0(t), (\beta^0 \cdot \gamma) \exp(\beta^0 \cdot \gamma y_t) \gamma y_t \rangle dy_t \}. \end{aligned} \quad (6.114)$$

where V^0 and Ψ^0 are the solutions of the differential equations (6.74) and (6.75), respectively.

For the numerical simulations, we have taken the functions f, b and γ as

$$f(t, x) = x^2,$$

$$b(t, x) = 0.5x,$$

$$\gamma(x) = (x^2, x)',$$

and $\beta = (\beta_1, \beta_2)'$. Further, the true values of the parameters β_1 and β_2 were taken as 1.0 and 0.5, respectively. Again starting with initial guesses $\beta_1 = 1.5$ and $\beta_2 = 1.0$, we have used the above algorithm to compute the estimated parameters and the results are summarized in Table II given below.

Table II

Iteration No.	β_1	β_2	$J(\beta_1, \beta_2)$
0	1.5	1.0	1.941166
10	1.329837	0.847493	23.790562
20	1.183540	0.711856	51.170798
25	1.118244	0.649336	67.598066
30	1.057736	0.589891	86.355767
35	1.001739	0.533229	107.968160
37	1.001570	0.511287	117.557222
38	0.991206	0.500347	122.582724
39	1.001402	0.489505	122.594950
true	1.0	0.5	

6.6 Summary and Conclusion

In this chapter, we have considered the identification problem of the drift coefficients of partially observed diffusions with and without Markov jump parameters. Using the filter equations obtained in chapters 2 and 3 and following similar procedure as in [36], we have formulated the identification problems as control problems in which the unknown parameters (vector or matrix) act as the controls, the unnormalized densities as the states, and the likelihood ratios as the objective functionals. Utilizing the techniques of optimal control Theory, we have shown these identification problems have solutions. Further, using the variational method, we have derived the necessary conditions for optimal identification. This requires the Gateaux differentiability of the unnormalized densities on the parameter set, which was also proved using similar arguments as those given in [3,6,7,12]. Finally, using gradient method, we have presented an iterative procedure for computing the estimated parameters along with two numerical examples to illustrate some of the results of this chapter.

CHAPTER 7

CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

7.1 Conclusions

In this thesis filtering and identification problems of nonlinear stochastic dynamical systems, have been considered. A complete derivation of the filter equation (Zakai equation) for the case where the state as well as the observed processes are of diffusion type, was presented. Using Ito differential rule and utilizing the properties of the likelihood ratio, we have shown that Kushner equation can be obtained from that of Zakai and vice versa. Further, using the standard result on the existence of solutions of partial differential equations, the question of existence and uniqueness of solutions of Zakai equation, was discussed.

Using Girsanov transformation and utilizing semigroup Theory, the result of Zakai was extended to the case where the state process is governed by nonlinear Ito differential equation whose coefficients are perturbed by :

- (i) unobservable temporally homogeneous Markov chain, or,
- (ii) deterministic Markov chain, or,
- (iii) periodically observable Markov chain.

We have presented a complete derivation for the filter equation corresponding to case (i) from which those of cases (ii) and (iii) were obtained. Further, using the standard result on existence of solutions of partial differential equations, we have discussed the question of existence and uniqueness of solutions of the filter equation corresponding to case (i).

The result of Skorokhod for Radon-Nikodym derivatives of measures induced by right continuous processes was utilized to extend the result of Zakai to the

case where both the state as well as the observed processes are driven by Wiener processes and Poisson random measures. In fact this technique is much simpler than martingale theoretic approach adapted in the literature. Further, we have noticed that by using the filter equation (4.38), the required estimate can be obtained if the observed process and the mean of the Poisson process γ (see equation (4.4)) are known. In contrast, the solutions of the filter equations obtained in [31] (see example 4.3) or in [72], require observing the noise of the output process, which is practically difficult.

Modelling the errors that arise in Dead Reckoning and Loran-C navigation sensors, has been also investigated. It has been shown that due to the discontinuous behaviour of the time (range) difference error of Loran-C radio signal, this error can be modelled as the sum of two independent stochastic processes. The first process was modelled as a linear Ito differential equation driven by Wiener process and the second as a pure jump process driven by two independent Poisson processes. Utilizing the filter equation obtained for the jump processes and using the fact that the state process is conditionally Gaussian, given the output process, we have obtained the corresponding filter equations for the navigation system. Further, a numerical example was presented to illustrate the effectiveness of the proposed filter. From this example, we have observed that the proposed filter tends to follow closely the rapid changes of the Loran-C error. Indeed our simulation results indicate that the actual and estimated states have similar behaviour and hence one expects better estimates for position and velocity of the craft in question.

Techniques of optimal control Theory and nonlinear filter Theory, have been utilized in identifying the drift parameters for partially observed diffusions with and without jump coefficients. Using variational method and the Gateaux differentiability of the unnormalized density, it has been shown that the optimal parameter

(vector or matrix) maximizing the likelihood ratio, is determined by the simultaneous solution of the system equation (linear partial differential equation for unnormalized density), the corresponding adjoint equation and an associated maximality condition. Finally, an algorithm, which utilizes these optimality conditions, has been presented in order to compute the unknown parameters.

7.2 Suggestions For Further Research

As a continuation of this thesis, further research could be conducted along several directions.

The numerical aspects of Kushner and Zakai equations along with a detailed numerical comparison between these two equations, is an interesting problem to investigate.

Another interesting area is the extension of the filter equation obtained in section 3.4 to the case where the drift and diffusion parameters are allowed to be unbounded. It is interesting also to study the question of existence and uniqueness of solutions of the corresponding filter equation.

The extension of the filter equation obtained in section 4.4, for discontinuous processes, under weaker assumptions along with the proof of the existence of the functions g_2 , ϱ , and c (see sections 4.2 and 4.3) would be of interest.

The extension of the results obtained for the identification problems of nonlinear partially observable stochastic systems, under weaker assumptions, is also of importance. Further, it is also of importance to study the consistency of the estimated parameter, as the length of the observation period goes to infinity. Finally, a computer software for parameter identification of nonlinear partially observable stochastic systems, would be also very useful.

REFERENCES

- [1] Abott , J.C. , and Gent , G.R., "Self Adaptive Filter For The Integration of Navigation Data" , J. Navigation , P. 1-26 , 1970.
- [2] Ahmed , N.U. and Wong , H.W., "A Minimum Principle For Systems Governed By Ito Differential Equations With Markov Jump Parameters , " Differential Games and Control Theory II , Lecture Notes in Pure and Applied Mathematics Vol. 30 , P. 265-294 , New York , 1977.
- [3] Ahmed , N.U., "Optimal Control of a Class of Strongly Nonlinear Parabolic Systems," J. Math. Analys. Appl. , Vol. 61 , P. 188-207 , 1977.
- [4] Ahmed , N.U., "Optimal Control of Stochastic Systems , " Probabilistic Analys. and Related Topics , (A.T. Bharucha Ried ed.) , P. 2-68 , Academic Press , New York , 1979.
- [5] Ahmed , N.U., "Stochastic Control on Hilbert Space for Linear Evolution Equations With Random Operator Valued Coefficients , " SIAM J. Contr. Optim. , Vol. 19 , No. 3 , P. 401-430 , 1981.
- [6] Ahmed , N.U. and Teo , K.L. , "Optimal Control of Distributed Parameter Systems" , Elsevier North Holland , New York , Oxford , 1981 .
- [7] Ahmed , N.U., "Necessary Conditions of Optimality for a Class of Second Order Hyperbolic Systems With Spatially Dependent Controls in the Coefficients , " J. Optim. Theory Appl. , Vol. 38 , No. 3 , P. 423-446 , 1982.
- [8] Ahmed , N.U., Dabbous , T.E., and Biswas , S.K. , "Modelling of Reliability Dynamics of Large Scale Interconnected Power Systems , " Proc. 4th International Symposium on Large Eng. Systems , P. 173-178 , Calgary , 1982.

- [9] Ahmed , N.U. and Dabbous , T.E. , "Modelling and on Line Control of Reliability Dynamics of Large Scale Interconnected Power Systems , " *Int. J. System Sci.* , Vol. 14 , No. 12 , P. 1321-1353 , 1983.
- [10] Ahmed , N.U. and Dabbous , T.E. , "Nonlinear Filtering of Systems Governed by Ito Differential Equations With Jump Parameters , " *J. Math. Anal. Appl.*, Vol. 115, No. 1, P. 76-92, 1986.
- [11] Ahmed , N.U. and Dabbous , T.E. , "Study of Discontinuous Jump Processes in Omega and Loran-C Radio Signals , " Report Submitted to Defence Research and Establishment , Ottawa , Canada , Aug. 1985.
- [12] Ahmed , N.U. , "Identification of Linear Operators in Differential Equations on Banach Space , " Abstract , American Math. Society 92nd Annual Meeting , Jan. 7-11 , P. 150 , 1986. To appear also in *Operator Methods of Optimal Control Problems*, Lecture Notes in Pure and Applied Math., Marcel Dekker.
- [13] Ahmed , N.U. and Dabbous , T.E. , "Parameter Identification of Diffusion Processes Using Nonlinear Filtering , " Submitted.
- [14] Ahmed, N.U., Dabbous, T.E., Liang, D.F. and McMillan, J.C., "Filtering of Discontinuous Processes Arising in Marine Integrated Navigation Systems," Submitted.
- [15] Allinger , D.F. and Mitter , S.K. , "New Results on the Innovations Problem For Nonlinear Filtering , " *Stochastics* , Vol. 4 , P. 339-348 , 1981.
- [16] Banks , H.T. , "A Survey of Some Problems and Recent Results For Parameter Estimation and Optimal Control in Delay and Distributed Parameter Systems,"

- Proc. Conf. on Volterra and Functional Differential Equations , Blacksburg ,
June , 1981.
- [17] Balakrishnan, A.V. , "Identification-Inverse Problems For Partial Differential Equations, A Stochastic Formulation , " 6th IFIP Conf. on Optim. , Novosibirsk , 1974 , Proc. in Lecture Notes in Computer Science , Springer-Verlag , Heidelberg , Berlin , New York , 1975.
- [18] Balakrishnan, A.V. , "Radon Nikodym Derivatives of a Class of Weak Distributions on Hilbert Spaces , " J. Appl. Math. Optim. , Vol. 3 , No. 2/3 , P. 209-225 , 1977.
- [19] Baras, J.S. , Blankenship , G.L. and Mitter , S.K. , "Nonlinear Filtering of Diffusion Processes , " Proc. IFAC Congr. , Koyoto , Japan , Aug. 1981.
- [20] Baras, J.S. , Blankenship , G.L. and Hopkins , W.E. , "Existence, Uniqueness and Asymptotic Behaviour of Solutions to a Class of Zakai Equations With Unbounded Coefficients , " IEEE Trans. on Aut. Contr. , Vol. AC-28 , No. 2 P. 203-214 , 1983.
- [21] Benes, V. and Karatzas , I. , "On the Relation of Zakai's and Mortensen's Equations , " SIAM J. Contr. Optim. , Vol. 21 , No. 3 , P. 472-489 , 1983.
- [22] Bensoussan, A. and Lions, J.L. , "Application des Inequations Variationnelles en Controle Stochastique , " Dunod , Paris , 1978.
- [23] Biswas, S.K. , "Modelling and Stabilization of Flexible Space Craft , " Ph.D. Thesis , University of Ottawa , Canada , 1985.
- [24] Blankenship, G.L. and Menaldi, J.L. , "Optimal Stochastic Scheduling of Power

- Generation Systems With Scheduling Delays and Large Cost Differentials , " SIAM J. Contr. Optim. , Vol. 22 , No. 1 , P. 121-132 , 1984.
- [25] Boel, R. , Varaiya, P. and Wong, E. , "Martingales on Jump Processes I , " SIAM J. Contr. Optim. , Vol. 13 , No. 5 , P. 999-1021 , 1975.
- [26] Boel, R., Varaiya, P. and Wong, E. , "Martingales on Jump Processes II , " SIAM J. Contr. Optim. , Vol. 13 , No. 5 , P.1022-1060 , 1975.
- [27] Bucy, R.S., "Nonlinear Filtering Theory , " IEEE Trans. Aut. Contr. , 10 , P. 198-212 , 1965.
- [28] Dabbous, T.E. and Ahmed, N.U. , "Nonlinear Filtering of Diffusion Processes With Discontinuous Observations , " Stoch. Analys. Appl. , Vol. 2 , No. 1 , P. 87-106 , 1984.
- [29] Davis, M.H.A. , "Pathwise Nonlinear Filtering , " Stochastic Systems : The Mathematics of Filtering and Identification and Applications , (NATO Advanced Study Institute Series) , Dordrecht , Reidel , P. 505-528 , 1980.
- [30] Davis, M.H and Marcus, S.I., "An Introduction to Nonlinear Filtering," Stochastic Systems : The Mathematics of Filtering and Identification and Applications (NATO Advanced Study Institute Series) , Dordrecht , Reidel , P. 53-75 , 1981.
- [31] Di Masi, G.B. and Runggaldier, W.J. , "An Approximation to Optimal Nonlinear Filtering With Discontinuous Observations , " Stochastic Systems : The Mathematics of Filtering and Identification and Applications (NATO Advanced Study Institute Series) , Dordrecht , Reidel , P. 583-590 , 1980.
- [32] Duckenfield, M.J., Marsh, J. and Milbourn, A.J. , "Studies in Integrated Navigation Systems , " Dep. of Elec. Eng. and Electronics , Brunel University ,

Internal Report , 1979.

- [33] Elliot, R.J. , "Stochastic Calculus and Applications , " Springer - Verlag , Heidelberg Berlin , New York , 1982.
- [34] Fleming, W.H. , "Measure Valued Processes in Control of Partially Observable Stochastic Systems , " Appl. Math. Optim. , 6 , P. 271-285 , 1980.
- [35] Fleming, W.H. and Mitter, S.K. , "Optimal Control and Pathwise Nonlinear Filtering of Nondegenerate Diffusions , " Presented at the 20th IEEE Conf. Decision Contr. , San Diego , CA , Dec. 1981.
- [36] Fleming, W.J. and Pardoux, E. , "Optimal Control for Partially Observed Diffusions , " SIAM J. Contr. Optim. , Vol. 20 , No. 2 , P. 261-285 , 1982.
- [37] Gihman, I.I. and Skorokhod, A.V., "The Theory of Stochastic Processes III," Springer - Verlag , Heidelberg Berlin , New York , 1979.
- [38] Girsanov, I.V. , "On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution of Measures , " Theory of Prob. Appl. , Vol. 5 , No. 3 , P. 285-301 , 1960.
- [39] Gland, L.F. , "Nonlinear Filtering Equation and Problem of Parameter Estimation , " Stochastic Systems : The Mathematics of Filtering and Identification and Applications (NATO Advanced Studies Institute Series) , Dordrecht , Reidel , P. 613-620 , 1980.
- [40] Grigelionis, B., "Stochastic Nonlinear Filtering Equations and Semimartingales," Lecture Notes in Math. , Vol. 972 , P. 62-99 , Springer - Verlag , Heidelberg Berlin , New York , 1983.

- [41] Ho, Y.C. and Lee, R.C.K. , "A Bayesian Approach to Problem in Stochastic Estimation and Control , " IEEE Trans. Aut. Contr. , 9 , P. 333-339 , 1964.
- [42] Jazwinski, A.H. , "Nonlinear Filtering With Discrete Observations , " AIAA 3rd Aerospace Science Meeting , New York , Paper No. 66-39, 1966.
- [43] Jazwinski, A.H. , "Filtering For Nonlinear Dynamical Systems , " IEEE Trans. Aut. Contr. , 11 , P. 765-766 , 1966.
- [44] Jazwinski, A.H. , "Adaptive Filtering , " Automatica , Vol. 5 , P. 475-485 , 1969.
- [45] Jazwinski, A.H. , "Stochastic Processes and Filtering Theory , " Academic Press , New York , London , 1970.
- [46] Kallianpur, G. , "Stochastic Filter Theory , " Springer - Verlag , Heidelberg Berlin , New York , 1980.
- [47] Kallianpur, G. and Karandikar, R.L. , "A Finitely Additive White Noise Approach to Nonlinear Filtering , " Appl. Math. Optim. , 10 , P. 159-185 , 1983.
- [48] Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems," Trans. ASME , Ser. D : J. Basic Eng. , 82 , P. 35-45 , 1960.
- [49] Kalman, R.E. and Bucy, R.S. , "New Results in Linear Filtering and Prediction Theory , " Trans. ASME , Ser. D : J. Basic Eng. , 83 , P. 95-108 , 1961.
- [50] Kubrusly, C.S. , "Distributed Parameter System Identification : A Survey , " Int. J. Contr. , Vol. 26 , P. 509-535 , 1977.

- [51] Kunita, H. , "Asymptotic Behaviour Of the Nonlinear Filtering Errors of Markov Processes , " J. Multivariate Analys. , 1 , P. 365-393 , 1971.
- [52] Kunita, H. , "Stochastic Partial Differential Equations Connected With Non-linear Filtering , " Lecture Notes in Math. , Vol. 972 , P.101-169 , Springer - Verlag , Heidelberg Berlin , New York , 1983.
- [53] Kushner, H.J. , "On the Dynamical Equations of Conditional Probability Density Functions With Application to Optimal Stochastic Control Theory , " J. Math. Analys. Appl. , 8 , P.332-344 , 1964.
- [54] Kushner, H.J. , "On the Differential Equations Satisfied by Conditional Probability Densities of Markov Processes , " SIAM J. Contr. , 2 , P. 106-119 , 1964.
- [55] Kushner, H.J. , "Dynamical Equations For Optimal Nonlinear Filtering , " J. Diff. Equations , 3 , P. 179-190 , 1967.
- [56] Kushner, H.J. , "Nonlinear Filtering : The Exact Dynamical Equations Satisfied by the Conditional Modes , " IEEE Trans. Aut. Contr. , 12 , P. 262-267 , 1967.
- [57] Kushner, H.J. , "Stochastic Stability and Control , " Academic Press, New York, London , 1967.
- [58] Kushner, H.J. , "A Robust Discrete State Approximation to the Optimal Non-linear Filtering For a Diffusion , " Stochastics , Vol. 3 , P. 75-83 , 1979.
- [59] Liang, D.F. , McMillan, J.C. and Maskell, C.A. , "Design and Test of a Marine Integrated Navigation Systems , " Report , Defence Research and Establishment Ottawa , Canada .

- [60] Liang, D.F. and McMillan, J.C. , "Development Analysis of Marine Integrated Navigation System , " IEEE PLANS - Position Location Navigation Symposium Record , P. 233-245 ; 1982.
- [61] Liang, D.F. , "Exact and Approximation State Estimation Techniques For Non-linear Dynamical Systems , " Advances in Contr. and Dynamic Systems , Academic Press , New York , London , Vol. 19 , P. 1-71 , 1983.
- [62] Liang, D.F. , "Comparisons of Nonlinear Recursive Filters For Systems With Non Negligible Nonlinearities , " Advances in Contr. and Dynamic Systems , Academic Press , New York , London , Vol. 20 , P. 341-400 , 1984.
- [63] Liptser, R.S. and Shirayayev, A.N. , "Studies of Random Processes I and II , " Springer - Verlag , Heidelberg Berlin , New York , 1978.
- [64] McMillan, J.C. , "A Kalman Filter For Marine Navigation , " MPhil. Thesis , University of Waterloo , Canada , 1980.
- [65] Mitchell, A.R. and Griffiths, D.F. , "The finite Difference Methods in Partial Differential Equations , " John Wiley , Chinchester , New York , 1980.
- [66] Pardoux, E. , "Stochastic Partial Differential Equations and Filtering of Diffusion Processes , " Stochastics, Vol. 3 , No. 2 , P. 127-167 , 1979.
- [67] Pardoux, E. , "Nonlinear Filtering, Prediction and Smoothing , " Stochastic Systems : The Mathematics of Filtering and Identification and Applications (NATO Advanced Study Institute Series) , Dordrecht , Reidel , P. 529-577 , 1980.
- [68] Skorokhod, A.V. , "Studies in The Theory of Random Processes , " Addison - Wesley , New York , 1965.

- [69] Stratonovich, R.L., "Conditional Markov Processes," Theory of Prob. Appl., 5, P. 156-178, 1960.
- [70] Takeuchi, Y. and Akashi, H., "Nonlinear Filtering Formulas For Discrete - Time Observations," SIAM J. Contr. Optim., Vol. 19, No. 2, P. 244-261, 1981.
- [71] Vaca, M.V. and Snyder, D.L., "Estimation and Decision For Observations Derived From Martingales : Part I," IEEE Trans. on Inform. Theory, Vol. IT-22, No. 6, P. 691-707, 1976.
- [72] Vaca, M.V. and Snyder, D.L., "Estimation and Decision For Observations Derived From Martingales : Part II," IEEE Trans. on Inform. Theory, Vol. IT-24, No. 1, P. 32-45, 1978.
- [73] Wong, H.W., "Optimal Control of Stochastic Differential Systems With Applications," Ph.D. Thesis, University of Ottawa, Ottawa, Canada, 1980.
- [74] Wonham, W.M., "Statistic Problem in Optimal Control," IEEE Intern. Conv. Record, 11, P. 114-124, 1963.
- [75] Yavin, Y., "An Alternative Approach to Nonlinear Filtering : Jump Observations," Technical Report, National Research Institute For Mathematical Sciences.
- [76] Zakai, M., "On the Optimal Filtering of Diffusion Processes," Z. Wahrsch., Verw. Geb. 11, 230-243, 1969.
- [77] Zakai, M. and Jacob, Z., "Lower and Upper Bounds on the Optimal Filtering Error of Certain Diffusion Processes," IEEE Trans. on Inform. Theory, Vol. IT-18, P. 325-331, 1972.

VITA

NAME: TAYEL ESSAWY DABBOUS

BORN: December 20, 1948, CAIRO, EGYPT.

EDUCATION:

University : Ain Shams University, CAIRO, EGYPT.
B.Sc. Engg. (1971); in Electrical Engineering.

University of Ottawa, OTTAWA, CANADA.
M.A.Sc. (1981); in Electrical Engineering.