AD-A077 149 ILLINOIS UNIV AT URBANA-CHAMPAIGN COORDINATED SCIENCE LAB F/G 12/1
STATE ESTIMATION FOR LINEAR SYSTEMS DRIVEN SIMULTANEOUSLY BY WI--ETC(U)
DEC 78 S P AU DAAB07-72-C-0259
UNCLASSIFIED R-836 NL

1 OF 2
AD-
A077149

REPORT R-836 DECEMBER, 1978 UILU-ENG 78-2229

COORDINATED SCIENCE LABORATORY

LEVEL

AD A077149

STATE ESTIMATION FOR LINEAR SYSTEMS DRIVEN SIMULTANEOUSLY BY WIENER AND POISSON PROCESSES

SAMUEL PORIZA AU

UNIVERSITY OF ILLINOIS - URBANA, ILLINOIS

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle)	STATE ESTIMATION FOR LINEAR SYSTEMS DRIVEN SIMULTANEOUSLY BY WIENER AND POISSON PROCESSES
5. TYPE OF REPORT & PERIOD COVERED	Technical Rep[redacted]t
6. PERFORMING ORG. REPORT NUMBER	R-836; UILU-ENG-78-2229
7. AUTHOR(s)	Samuel Poriza Au
8. CONTRACT OR GRANT NUMBER(s)	DAAB-07-72-C-0259;DAAG-29-78-C-0016;DAAG-29-76-G-0154
9. PERFORMING ORGANIZATION NAME AND ADDRESS	Coordinated Science Laboratory University of Illinois at Urbana-Champaign Urbana, Illinois 61801
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	
11. CONTROLLING OFFICE NAME AND ADDRESS	Joint Services Electronics Program
12. REPORT DATE	Dec[redacted]78
13. NUMBER OF PAGES	111
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	
15. SECURITY CLASS. (of this report)	UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report)	Approved for public release; distribution unlimited
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)	DAAB07-72-C-0259, DAAG29-78-C-0016
18. SUPPLEMENTARY NOTES	
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)	White Processes Poisson Wiener Filtering Smoothing

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this thesis, the state estimation problem of linear stochastic systems driven simultaneously by Wiener and Poisson processes is considered. We are concerned with the case where the incident intensities of the Poisson processes are low and the system is observed in an additive white Gaussian noise.

The minimum mean-squared-error (MMSE) optimal filter is derived via the Doleans-Dade and Meyer differentiation rule for discontinuous semi-martingales and its corresponding basic filtering theorem for white Gaussian observation noise. The nonclosedness property and performance of the filter are

20. ABSTRACT (continued)

investigated. The results together with the performance of the linear optimal filtering schemes lead to the conclusion that causal filters and noncausal linear filters are inherently unsuitable for the state estimation for such class of systems.

A noncausal nonlinear suboptimal scheme is developed for the estimation problem based on a combined estimation and detection strategy. A first-order approximation scheme is included in the scheme to eliminate the error propagation effects that result from the sequential structure of the approach. The performance of the overall scheme is obtained analytically and simulated numerically. Both results agree closely indicating that there exists a λ^* such that if the Poisson intensity $\lambda \in (0,\lambda^*]$, the suboptimal sequential scheme performs better than the causal optimal filter and the noncausal linear filter.

lambda*

lambda an element of (0, lambda*>

UILU-ENG 78-2229

STATE ESTIMATION FOR LINEAR SYSTEMS DRIVEN SIMULTANEOUSLY BY WIENER AND POISSON PROCESSES

by

Samuel Poriza Au

This work was supported in part by the Joint Services Electronics Program (U.S. Army, U.S. Navy and U.S. Air Force) under Contract DAAB-07-72-C-0259 and DAAG-29-78-C-0016 and in part by the U.S. Army Research Office under Contract DAAG-29-76-G-0154.

STATE ESTIMATION FOR LINEAR SYSTEMS DRIVEN SIMULTANEOUSLY BY WIENER AND POISSON PROCESSES

BY

SAMUEL PORIZA AU

S.B.E.E., Massachusetts Institute of Technology, 1975
M.S., University of Illinois, 1977

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 1979

Thesis Advisers: Professor A. H. Haddad
Professor H. V. Poor

Urbana, Illinois

STATE ESTIMATION FOR LINEAR SYSTEMS DRIVEN
SIMULTANEOUSLY BY WIENER AND POISSON PROCESSES

Samuel Poriza Au, Ph.D.
Coordinated Science Laboratory and
Department of Electrical Engineering
University of Illinois at Urbana-Champaign, 1979

In this thesis, the state estimation problem of linear stochastic systems driven simultaneously by Wiener and Poisson processes is considered. We are concerned with the case where the incident intensities of the Poisson processes are low and the system is observed in an additive white Gaussian noise.

The minimum mean-squared-error (MMSE) optimal filter is derived via the Doleans-Dade and Meyer differentiation rule for discontinuous semimartingales and its corresponding basic filtering theorem for white Gaussian observation noise. The nonclosedness property and performance of the filter are investigated. The results together with the performance of the linear optimal filtering schemes lead to the conclusion that causal filters and noncausal linear filters are inherently unsuitable for the state estimation for such class of systems.

A noncausal nonlinear suboptimal scheme is developed for the estimation problem based on a combined estimation and detection strategy. A first order approximation scheme is included in the scheme to eliminate the error propagation effects that result from the sequential structure of the approach. The performance of the overall scheme is obtained analytically and simulated numerically. Both results agree closely indicating that there exists a λ^* such that if the Poisson intensity $\lambda \in (0,\lambda^*]$, the suboptimal sequential scheme performs better than the causal optimal filter and the noncausal linear filter.

ACKNOWLEDGMENTS

The author wishes to thank his advisors, Professor A. H. Haddad for his concerns and guidance during the preparation of this thesis, and Professor H. V. Poor for his encouragement and support during the entire course of the work. Most sincere thanks go to Professor J. Liu who first interested him in the aspects of random point processes that form the foundation of this thesis. He would also like to thank Professors M. B. Pursley and D. P. Bertsekas for serving on his dissertation committee.

It is a pleasure for the author to acknowledge his Mom for her love and inspirations. He also wishes to express his appreciation to all his friends for their support and encouragement during the time when he was at Urbana.

Thanks are also given to Mrs. Hazel Corray for her excellent typing.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 General

There are numerous physical phenomena that can be modeled as stochastic linear systems driven simultaneously by Wiener and Poisson processes. The Wiener driven part of the system is that part of the system that is continuous in the stochastic sense; and the Poisson driven part is that responsible for the discontinuities of the system process. An example is a man-maneuvered spacecraft. Its motion is composed of a nominal sum of responses due to the driver's discrete application of controls to the spacecraft and the response due to the continually varying turbulence and atmospheric perturbations. The superposition of the responses due to the random incidents of control applications can be modeled as a Poisson driven process if the control incidents arrive in a random fashion. The response of the spacecraft due to the continually varying turbulence and such alike can be modeled as a Wiener driven process.

As the indicent rate λ (the intensity) of the Poisson driving process tends to infinity, under fairly weak conditions, the linearly Poisson driven process tends to a Gaussian process [1]. Therefore the part of the linear system driven by Poisson processes of large intensities can be reasonably modeled as Gauss-Markov system, namely a Wiener driven model for the system process. Then the entire system consisting of a Poisson driven and a Wiener driven part reduces to a Wiener driven process in such case. When λ is small, as in the case of the man-maneuvered spacecraft, the Wiener driven model is no longer a good approximation to the Poisson driven part of the system. The system can only be appropriately modeled as a process driven both by Wiener and Poisson processes.

In this work, we consider the state estimation problem of such systems observed in additive white Gaussian noise. Compared to the problem for system driven only by Wiener processes, this has received little attention in the literatures, although such systems arise relatively often and naturally in practice.

Among the common examples, stochastic control systems that are driven in the environments of optimal control strategies typically exhibit discontinuities and "switchings" in their trajectories, e.g., the bang-bang controls under some time- and energy-optimality conditions [14-16,30]. Estimation of such stochastic systems often is necessary in pursuit-evasion games and in tracking of maneuvering targets [31,32].

To further illustrate the wide application of such a class of systems and the motivation for its estimation problem, we give the following illustrative example. In off-shore oil exploration, trains of sonic waves are injected into the earth's interior underneath the seabed; reflections of such waves are measured to estimate the physical density of the earth's interior as a function of the depth below the seafloor. Should there be no oil-layers within the depth of penetration of the wave, the density is a stochastically continuous function.

However, in case there is oil underneath, it is usually in layers of various thickness for example. Within the uncertainty of such aspects as thickness and quality etc., each of these layers which occur at various random spatial points, contributes a similar variational pattern (incident response) to the density function. Hence, the density function can be modeled as a system driven by a Poisson process with very low spatial intensity. The uncertainties of the layers can be modeled as a random parameter associated with each incident, called the mark of that incident.

Another example can be seen from the following environmental problem. A river is occasionally polluted by deposits of chemical products from various sources along the river. The amounts of a particular chemical that are deposited each time are independent stochastic variables with given distributions. The chemical is dispersed in the river by diffusion and transport and flows. The effect of such chemical to the various biological and environmental systems in the river can be studied by measuring continuously the concentration of the chemical and its arising effects at various points along the river. The problem hence relies heavily on the reconstruction or estimation of the times and amounts of chemical and such relevant factors from the measured data. Such a problem fits into our estimation problem naturally due to the Poisson nature of the modeled pollution deposits.

As seen from the above examples, the class of Wiener Poisson driven systems encompasses stochastic processes with discontinuous sample paths, unlike the Wiener driven systems that have continuous sample paths almost surely. In the area of fractures and defects studies, and earthquake analysis where the estimation problem arises naturally, due to their inherent discontinuities in the sample paths, the model becomes conveniently applicable.

Such systems for the case of finite dimensional state model have Markov property but in general are not Poisson, nor Gaussian except in some particular cases [2,3,4]. Due to the non-Gaussian property and the discontinuity of the sample paths, it takes an infinite number of its moments (cumulants) to characterize its stochastic properties completely. This is manifested later in the development of the optimal filter for such processes.

1.2. Problem Statement

We consider the state estimation problem of the following system driven simultaneously by a Wiener process and a Poisson process,

$$dx_t = \alpha_t x_t dt + \beta_t dW_t + b_t d\eta_t \qquad t \geq 0 \tag{1.1}$$

which is observed in white Gaussian noise, viz.

$$dy_t = h_t x_t d_t + d\nu_t \tag{1.2}$$

where x_t and y_t are the system state and observations respectively with x_0 assumed to be independent of $\{x_t, t > 0\}$. W_t, ν_t are independent generalized Wiener processes with $E(dW_t\ dW_t') = Q_t dt$ and $E(d\nu\ d\nu_t') = R_t dt$, Q_t, R_t being positive semidefinite and positive definite respectively. η_t is a Poisson process independent of W_t and ν_t with known rate parameter λ. Without loss of generality, we assume $\mathcal{A}_t$, $\mathcal{B}_t$, b_t and h_t to be all bounded known time-continuous functions for $t \geq 0$ with

$$b_t^2 \geq \zeta_1 > 0 \qquad t \geq 0 \tag{1.3}$$

for some constants ζ_1. The assumption is made to ensure the existence and uniqueness of a solution to (1.1) and (1.3). The problem under consideration is to construct for each $t \geq 0$, an estimate of the system state x_t on the basis of observations $\{y_s, s \leq u\}$ where $u > t$. When $u = t$, it is a filtering problem; $u > t$, a smoothing problem. The performance criterion to be used is the minimum mean-squared-error (MMSE). As λ is large, the estimation problem reduces to that of estimating a Wiener driven system in the presence of additive white Gaussian noise. In this work, we are mainly concerned with the case where λ is small, for which the linear estimate may not be as suitable. In order to avoid cumbersome algebraic notations we restrict ourselves to the scalar case. The extension of all the results in this work to the vector case is conceptually straightforward and can readily be obtained.

1.3 Summary of Past Results

The general problem of state estimation of linear systems driven by Wiener processes observed in additive white Gaussian noise has been treated extensively in the past [33-36] et al. Complete solutions to various modified filtering problems are available abundantly in the literature [33-

41]. However, the general problem of state estimation of linear systems driven by Wiener and Poisson processes in the presence of white Gaussian noise is relatively untreated.

Linear systems excited by Poisson processes was studied recently in the estimation problems [2,3,28]. The problem was solved by applying results developed most recently in martingale theory [4-8,20,42-47]. The MMSE optimal filters in a form of an infinite set of stochastic differential equations were derived. In the presentations of the results, the infinite dimensionality of the optimal filters were heuristically argued to be irreducible to finite form. Truncation approximations and the Ritz-Galerkin approximation methods [2,3,48] were suggested to solve the set of equations. Both in a numerical simulation study were proved to be disappointing in terms of improvement on performance over the Kalman filter. Indeed, the author has shown that the Kalman filter which utilizes only the first and the second order statistics, MMSE optimal for Markov-Gaussian systems, are extremely poor in performance for systems excited by Poisson marked processes [49,50].

In different formulations, random systems that evolve with stochastic jump processes has been extensively attended [51-57]. The basic theoretic properties of such stochastic systems were studied in great depth [7,8,12,27] 51] et al. Most of the works in area of estimation however centered mainly on the detection of jumps, sudden changes of stochastic properties of the systems and on the renewal processes associated with the point processes imbedded in the systems [58-67]. Under different criteria on performance, the approach of combined estimation detection schemes [68-72] were suggested for estimating and detecting the jump processes. A good collection of such combined schemes based on various criteria and structures of the problems are available [68-72].

Related estimation problems on stochastic systems driven by generalized Poisson processes have received extensive attentions [73-76] et al. In most cases, the estimation efforts were concentrated on the estimation of the average Poisson intensities of the driving point processes based on the noisy observations of the driven systems.

1.4 Chapters Review

We first derive in Chapter 2 the (MMSE) optimal filter for the problem. The approach will be closely parallel to that in [2,3], via the Dolean-Dade and Meyer decomposition rule for discontinuous semi-martingales and its corresponding filtering theorem [4,5,6]. The different aspects of the filter is discussed and investigated. Its performance is compared to that of the linear optimal filter.

In Chapter 3, we investigate the problem for the system driven only by the Poisson processes i.e., $\beta_t \equiv 0$. Since the estimation problem for the system driven only by Wiener processes has been widely studied and developed, this renders a better understanding and insights into the general problem. A suboptimal sequential scheme (SSS) based on a combined estimation-detection approach is developed for this particular problem. A compensation strategy is adopted in the SSS to eliminate the propagating error due to the sequential structure of the SSS. The asymptotic performance for the time-invariant case is derived and numerically simulated on a digital computer.

16

In Chapter 4, the general estimation problem is considered. The SSS is modified and extended as a suboptimal solution to the problem. The different aspects of the SSS and its extension for other systems will be presented.

In Chapter 5, we extend the estimation problem of the Poisson-Wiener driven linear systems to a wider class of stochastic systems, namely the conditional Gaussian systems. We derive the basic equations and functionals that are essential for the development of the SSS as a suboptimal approach to the estimation problem. While the derivations of the SSS will be similar as those in Chapter 4, we will not get into the detail derivations of the SSS here.

Chapter 6 concludes this thesis with a summary of the results and conclusions of the previous chapters.

CHAPTER 2

THE OPTIMAL FILTER

2.1. The Optimal Filter

For algebraic simplicity, we consider the scalar case of the system slightly modified from (1.1), given as follows,

$$dx_t = \mathcal{A}_t x_t dt + \mathcal{B}_t dW_t + b_t d\vartheta_t \qquad t \geq 0 \tag{2.1}$$

with

$$\vartheta_t = \eta_t - \lambda t \tag{2.2}$$

The principal effect is a deterministic change in the mean function of the system x_t, $t \geq 0$ which does not affect the essence of our problem.

The approach we use to obtain the optimum filter is via the semi-invariant generating function for x, $\ln E(e^{jux_t})$. First apply the differentiation rule of Dolean-Dade and Meyer (see Appendix A) for discontinuous semi-martingales to e^{jux_t}, putting it in the standard semi-martingale form. Then we apply the optimal filtering theorem for white Gaussian observation noise [2-6] (see Appendix B). The differentiation rule results in

$$de^{jux_t} = jue^{jux_{t-}}dx_{t-} + \frac{1}{2}(ju)^2 e^{jux_{t-}} d < x^c, x^c >_t + d \sum_{0<s<t} (e^{ju(x_s + b_s)} - e^{jux_s} - jue^{jux_{s-}} b_s) \tag{2.3}$$

where $< x^c, x^c >_t$ is the increasing process associated with the continuous part of the semi-martingale x_t, $t \geq 0$ [2,7,8]. The summation is carried out over those values of s where x jumps. Here $j = \sqrt{-1}$ and u is real. Considering the continuous part of x, we have

$$d < x^c, x^c >_t = \mathcal{B}_t^2 Q_t dt \tag{2.4}$$

and also discontinuous part, we have

$$d \sum_{0<s\leq t} \psi(X_{s-}) = d\int_0^t \psi(X_s)d\eta_s = \psi(X_t)d\eta_t \tag{2.5}$$

Substituting of (2.4), (2.5) into (2.3), one obtains

$$de^{jux_t} = [ju\alpha_t x_t e^{jux_t} + \frac{1}{2}(ju)^2\beta_t^2 Q_t e^{jux_t}]dt$$
$$+ jue^{jux_t}\beta_t dW_t + e^{jux_t}(e^{jub_t}-1)\,d\theta t + \lambda e^{jux_t}(e^{jub_t}-1)dt \tag{2.6}$$

which is in a form such that the optimal filtering theorem is readily applied. Let $\hat{\psi}_t$ denote the conditional expectation of the random variable ψ_t with respect to the growing σ-field generated by $\{y_s, 0 \leq s \leq t\}$. Then applying the filtering theorem, we have

$$d\widehat{e^{jux_t}} = [ju\alpha_t \widehat{x_t e^{jux_t}} + \frac{1}{2}(ju)^2\beta_t^2 Q_t \widehat{e^{iux_t}}]dt + \lambda\widehat{e^{iux_t}}(e^{jub_t}-1)dt$$
$$+ h_t[\widehat{x_t e^{jux_t}} - \hat{x}_t\widehat{e^{jux_t}}]v_t^{-1}[dy_t - h_t\hat{x}_t]\,dt \tag{2.7}$$

If we substitute the cumulant generating function, viz.

$$\widehat{e^{jux_t}} = \exp[\sum_{k=1}^{\infty} C_{k_t}\frac{(iu)^k}{k!}] \tag{2.8}$$

where C_{k_t} is the k-th conditional cumulant of x_t into (2.7)

we obtain the following set of equations,

$$dC_{1t} = \alpha_t C_{1t}dt = h_t v_t^{-1} C_{2t}(dy_t - h_t C_{1t}dt) \tag{2.9a}$$

$$dC_{2t} = 2\alpha_t C_{2t}dt + \beta_t^2 Q_t dt + \lambda b_t^2 dt - h_t^2 v_t^{-1} C_{2t}^2 dt \tag{2.9b}$$
$$+ h_t v_t^{-1} C_{3t}(dy_t - h_t C_{1t}dt)$$

$$dC_{k_t} = k\alpha_t C_{k_t}dt + \lambda b_t^k dt - \frac{1}{2}h_t^2 v_t^{-1}\sum_{j=1}^{k-1}\binom{k}{j} C_{j+1,t}C_{k-j+1}$$
$$+ h_t v_t^{-1} C_{k+1,t}(dy_t - h_t C_{1t}dt) \quad k = 3,4,5.... \tag{2.9c}$$

The conditional expectation of x_t and its variance denoted by P_{2t} are therefore given by

$$\hat{x}_t = C_{1t} \tag{2.10}$$

$$P_{2t} = \widehat{(x_t - \hat{x}_t)^2} = C_{2t} \tag{2.11}$$

Note that C_{k_t} $k = 2,3,4,\ldots$ are all invariant to any change in the mean function of the system x_t.

From Eqs. (2.9a) and (2.9b), we see that the optimal estimate has precisely the form of the Kalman filter, except that C_{2t} is observation-dependent. The differential equation for C_{k_t} involves, besides lower-order cumulants, also $C_{k+1,t}$ implying the optimal filter requires solving an infinite set of simultaneous stochastic differential equations.

There are many processes in which the optimal estimates exist in closed form. Among the common is the Gaussian case, where the additional relation

$$\widehat{x_t^3} = 3\hat{x}_t\widehat{x_t^2} - 2\hat{x}_t^3 \tag{2.12}$$

enables us to reduce the set of Eqs. (2.9) into a closed system of equations, the Kalman filter [7].

2.2 The Nonclosedness of the Optimal Filter

In order to study the property of closedness of the filter (2.8), we define the process I_t which drives the optimal filter as follows,

$$dI_t = dy_t - h_t\hat{x}_t dt = dy_t - h_t C_{1t} dt \tag{2.13}$$

If the process I_t exists in a well-defined manner then it is the innovation process with respect to the observation in our estimation problem [7-11]. In such case it is a Wiener process with respect to the σ-field generated by $\{y_s, 0 \leq s \leq t\}$. Furthermore its stochastic properties are identical to those of the Wiener process ν_t in the observations. It is called an innovation process because it carries and hence can generate the same "information" as the observations.

Note that solutions to a closed set of equations containing states driven by Wiener processes are all sample-continuous with probability one, provided they exist [7,8]. Hence if there exists an additional relationship in our case as Eq. (2.12) in the Gaussian case, which enables us to reduce the set of Eq. (2.9) to a closed set, the optimal estimate and its cumulants as solutions to the reduced set of equations will be all sample-continuous with probability one. Since the system we are estimating is discontinuous at various random times (termed as Markov times [7]) and the optimal estimate being instantaneous estimate of the system should exhibit a highly "discontinuous" sample path, we deduce that the set of Eqs. (2.9) is not closed, containing an infinite number of equations.

Then the question of existence of an optimal estimate as solution to such a non-closed set of equations naturally arises.

Note that to carry the same "information" as the observation y, the σ-field generated by the innovation process I has to be equal to that generated by the process y for all t, $0 \leq t < \infty$. That is, if we let $\mathcal{F}_t^I$ and $\mathcal{F}_t^y$ denote the σ-fields generated by I_s and y_s $0 \leq s \leq t < \infty$ respectively, by definition of the innovation process, the following must hold for all t

$$\mathcal{F}_t^I = \mathcal{F}_t^y \qquad 0 \leq t < \infty \tag{2.14}$$

Since $\mathcal{F}_t^I$ is generated by the innovation process which is a Wiener process, it is continuous in t, viz.

$$\lim_{s\uparrow t} \mathcal{F}_s^I = \lim_{s\downarrow t} \mathcal{F}_s^I = \mathcal{F}_t^I \qquad 0 \leq t < \infty \qquad (2.15)$$

However $\mathcal{F}_t^y$, generated by the observation process y which contains a discontinuous process, is only right-continuous. With non-zero probability, $\mathcal{F}_t^y$ is not continuous, i.e.

$$\lim_{s\uparrow t} \mathcal{F}_s^y \neq \mathcal{F}_t^y \qquad \text{for some } t, \quad 0 \leq t < \infty \qquad (2.16)$$

Therefore, from the above argument and Eq. (2.15), (2.16) we establish

$$\mathcal{F}_t^I \neq \mathcal{F}_t^y \qquad 0 \leq t < \infty \qquad (2.17)$$

which contradicts the assumption of the process I being the innovation process with respect to $\mathcal{F}_t^y$, viz. Eq. (2.14). And, by definition of the process I,

$$\mathcal{F}_t^I \subseteq \mathcal{F}_t^y \qquad (2.18)$$

through the construction of I.

Therefore, the process I is not an innovation process and it carries less information than the observation y. This implies the process I does not exist, and from Eq. (2.13), it follows that $\hat{x}_t$ does not exist, or at best exists in a certain stochastic limit sense, and not in the strong sense.

To summarize, the optimal estimate is given in terms of a set of infinite number of differential equations. Because of the non-closedness of the set of equations, the optimal estimate may not exist and at best it can be represented as a limit in some stochastic sense.

The investigation of how such a limit behaves is a difficult problem and does not bear any significance to our problem at hand. From a practical point of view, the non-closedness of the set of equations has made the implementation of such an estimate formidable. It should be apparent that simple truncation of the set results in a continuous estimate for x_t, generally of very poor performance because it is required to make instantaneous estimation of a discontinuous system. Such truncation may also lead to a possibly unstable filter.

2.3 The Optimal Filter and the Kalman Filter

We investigate the performances of the linear optimal filter, i.e., the Kalman filter, and the optimal filter. In particular, their asymptotic performances when λ is small are compared.

The Kalman filter which utilizes only the first and second order statistics can be readily obtained and is given by

$$d\check{X}_t = \mathcal{a}_t\check{X}_t dt + h_t v_t^{-1} P(t)(dy_t - h_t\check{X}_t dt) \tag{2.19a}$$

$$dP(t) = [2\mathcal{a}_t P(t) + \lambda b_t^2 + \mathcal{B}_t^2 Q_t - h_t^2 v_t^{-1} P^2(t)]dt \tag{2.19b}$$

where $\check{X}_t$ and $P(t)$ denote the Kalman estimate of the system and its variance respectively. Notice that the variance equation is realization-independent and can be solved ahead of time. Although the Kalman estimate always exists as solution to the Eq. (2.19a), $dy_t - h_t\check{X}_t dt$ is not an innovation process with respect to the observation since $\check{X}_t$ is not an optimal estimate. It should be apparent that any random process with only the same first and second moments as x_t will have the same Kalman filter.

Define now

$$\tilde{\phi} = \underset{\lambda \to 0}{\text{l.i.m.}}\ \phi \tag{2.20a}$$

$$\tilde{\phi}_\lambda = \underset{\lambda \to 0}{\text{l.i.m.}}\ \frac{\partial\phi}{\partial\lambda} \tag{2.20b}$$

where the derivative is taken in the quadratic mean sense.

As $\lambda \to 0$, it can be seen easily that the optimal filter given by Eq. (2.9) reduces to the closed form of the Kalman filter given by (2.19) with the assumption of Gaussian initial condition of x_t. We have

$$d\tilde{C}_{1t} = \mathcal{a}_t\tilde{C}_{1t} + h_t v_t^{-1}\tilde{C}_{2t}dI_t \tag{2.21a}$$

$$d\tilde{C}_{2t} = 2\mathcal{a}_t\tilde{C}_{2t} = \mathcal{B}_t^2 Q_t - h_t^2 v_t^{-1}\tilde{C}_{2t}^2 \tag{2.21b}$$

in which case

$$\tilde{C}_{1t} = \tilde{\check{X}}_t \tag{2.22a}$$

$$\tilde{C}_{2t} = \tilde{P}(t) \tag{2.22b}$$

and
$$\tilde{C}_{k_t} = 0 \qquad t \geq 0 \quad \text{for} \quad k = 3,4,5,\ldots. \tag{2.22c}$$

Taking derivatives of Eq. (2.9) and using Eq. (2.22c), we have

$$d\tilde{C}_{1\lambda t} = \mathcal{a}_t\tilde{C}_{1\lambda t}dt + h_t v_2^{-1}\tilde{C}_{2\lambda t}dI_t \tag{2.23a}$$

$$d\tilde{C}_{k\lambda t} = k[\mathcal{a}_t - h_t^2 v_t^{-1}\tilde{P}(t)]\tilde{C}_{k\lambda t}dt + b_t^k dt$$
$$+ h_t v_t^{-1}\tilde{C}_{k+1,\lambda t}dI_t \qquad k = 2,3,4,\ldots \tag{2.23b}$$

Suppose we write

$$C_{1t} = \tilde{C}_{1t} + \lambda\tilde{C}_{1\lambda t} + \frac{1}{2}\lambda^2\tilde{C}_{1\lambda\lambda t} + o(\lambda^2) \tag{2.24a}$$

$$\check{X}_t = \tilde{\check{X}}_t + \lambda\tilde{\check{X}}_{\lambda t} + \frac{1}{2}\lambda^2\tilde{\check{X}}_{\lambda\lambda t} + o(\lambda^2) \tag{2.24b}$$

Hence, from Eq. (2.5), we have for the difference between the two estimates

$$(\check{X}_t - C_{1t}) = \lambda\,(\tilde{\check{X}}_{\lambda t} - \tilde{C}_{1\lambda t}) + o(\lambda) \tag{2.25}$$

Combining Eq. (2.19) and (2.23), we first have

$$d(\check{\tilde{X}}_{\lambda t} - \tilde{C}_{1\lambda t}) = [\mathcal{a}_t - h_t^2 v_t^{-1} \tilde{P}_\lambda(t)][\check{\tilde{X}}_{\lambda t} - \tilde{C}_{1\lambda t}]dt$$

$$+ h_t v_t^{-1}[\tilde{P}_\lambda(t) - \tilde{C}_{2\lambda t}]dI_t \tag{2.26}$$

Let us confine ourselves to investigate the asymptotic values, for $t \to \infty$. Assume there exists $t' < \infty$, such that

$$\mathcal{a}_t - h_t^2 v_t^{-1} \tilde{P}_\lambda(t) \leq K_1 < 0 \qquad \text{for } t > t' \tag{2.27}$$

$$\mathcal{a}_t - h_t^2 v_t^{-1} \tilde{P}(t) \leq K_2 < 0 \qquad \text{for } t > t' \tag{2.28}$$

which is readily satisfied by any stable systems.

Therefore from (2.25) and (2.26), we have

$$\lim_{t\to\infty} E(\check{X}_t - C_{1t}) = \lim_{t\to\infty} \lambda\, E(\check{\tilde{X}}_{\lambda t} - \tilde{C}_{1\lambda t}) + o(\lambda) = o(\lambda) \tag{2.29}$$

Similarly, we have

$$E(\check{X}_t - C_{1t})^2 = \lambda^2 E(\check{\tilde{X}}_{\lambda t} - \tilde{C}_{1\lambda t})^2 + o(\lambda^2) \tag{2.30}$$

Applying the differentiation rule to $(\check{\tilde{X}}_{\lambda t} - \tilde{C}_{1\lambda t})^2$, we obtain

$$d(\check{\tilde{X}}_{\lambda t} - \tilde{C}_{1\lambda t})^2 = 2[\mathcal{a}_t - h_t^2 v_t^{-1} \tilde{P}_\lambda(t)](\check{\tilde{X}}_{\lambda t} - \tilde{C}_{1\lambda t})^2 dt$$

$$+ h_t^2 v_t^{-1}[P_\lambda^2(t) - C_{2\lambda t}^2]dt$$

$$+ 2(\check{\tilde{X}}_{\lambda t} - \tilde{C}_{1\lambda t}) h_t v_t^{-1}[\tilde{P}_\lambda(t) - \tilde{C}_{2\lambda t}]dI_t \tag{2.31}$$

Hence if the system (2.31) is asymptotically stable

$$\lim_{t\to\infty} E(\check{X}_t - C_{1t})^2 = \lim_{t\to\infty} \lambda^2 E(\check{\tilde{X}}_{\lambda t} - \tilde{C}_{1\lambda t})^2 + o(\lambda^2)$$

$$= \lim_{t\to\infty} \frac{-h_t^2 v_t^{-1}}{2(\mathcal{a}_t - h_t^2 v_t^{-1} \tilde{P}_\lambda(t))} E[\tilde{P}_\lambda^2(t) - \tilde{C}_{2\lambda t}^2] + o(\lambda^2) \tag{2.32}$$

However, we also have

$$d(\tilde{P}_\lambda^2(t) - \tilde{C}_{2\lambda t}^2) = 4[\mathcal{A}_t - h_t^2 v_t^{-1}\tilde{P}(t)][\tilde{P}_\lambda^2(t) - \tilde{C}_{2\lambda t}^2]dt$$

$$+ 2b_t^2[\tilde{P}_\lambda(t) - \tilde{C}_{2\lambda t}]dt - h_t^2 v_t^{-1}\tilde{C}_{3\lambda t}^2$$

$$- 2h_t v_t^{-1}\tilde{C}_{2\lambda t}\tilde{C}_{3\lambda t}dI_t$$

$$\Rightarrow \lim_{t\to\infty} E[P_\lambda^2(t) - C_{2\lambda t}^2] = \lim_{t\to\infty} \frac{-2b_t^2 E(\tilde{P}_\lambda(t) - \tilde{C}_{2\lambda t}) + h_t^2 v_t^{-1} E(\tilde{C}_{3\lambda t}^2)}{4(\mathcal{A}_t - h_t^2 v_t^{-1}\tilde{P}(t))}$$

$$= \lim_{t\to\infty} \frac{h_t^2 v_t^{-1} E(\tilde{C}_{3\lambda t}^2)}{4(\mathcal{A}_t - h_t^2 v_t^{-1}\tilde{P}(t))} \tag{2.33}$$

Similarly,

$$\lim_{t\to\infty} E(\tilde{C}_{3\lambda t}^2) = \lim_{t\to\infty} \frac{-1}{6[\mathcal{A}_t - h_t^2 v_t^{-1}\tilde{P}(t)]}\{h^t v^{-1} E(\tilde{C}_{4\lambda t}^2)$$

$$+ \frac{(b_t^2)^3}{3(\mathcal{A}_t - h^2 v^{-1}\tilde{P}(t))}\} \tag{2.34}$$

Continuing the process, we deduce that as t is large

$$E(\tilde{X}_{\lambda t} - \tilde{C}_{1\lambda t})^2 \sim [\mathcal{A}_t - h_t^2 v_t^{-1}\tilde{P}_\lambda(t)]^{-1}\{\sum_{k=3}^{\infty} \frac{1}{k \cdot k!}(\frac{h_t^2 v_t^{-1} b_t^2}{2\alpha_t})\} \tag{2.35}$$

where $\alpha_t = \mathcal{A}_t - h_t^2 v_t^{-1}\tilde{P}(t)$

Observe that the sum given by Eq. (2.35) is a convergent series, because the ratio of the $(k+1)$st term to the k-th term is

$$\frac{k}{(k+1)^2}\left|\frac{h_t^2 v_t^{-1} b_t^2}{2[\mathcal{A}_t - h_t^2 v_t^{-1}\tilde{P}(t)]}\right| \to 0 \tag{2.36}$$

uniformly as $k \to \infty$ for $t > t'$. The behavior of the limit of the mean squared difference (2.32) therefore depends on the behavior of b_t^2.

First for small b_t^2 as $t \to \infty$, as expected

$$\lim_{t\to\infty} E(\check{X}_t - C_{1t})^2 = 0 + o(\lambda^2) \qquad (2.37)$$

In such case, the optimal filter reduces to the Kalman filter asymptotically as $\lambda \to 0$.

However for large b_t^2, it is easily deduced that

$$\lim_{t\to\infty} E(\tilde{\check{X}}_{\lambda t} - \tilde{C}_{1\lambda t})^2 \sim \lim_{t\to\infty} - [\mathcal{C}_t - h_t^2 v_t^{-1} \tilde{P}_\lambda(t)](\frac{h_t^2 v_t^{-1} b_t^2}{2\alpha t})^2 \qquad (2.38)$$

Therefore for small λ but large b_t^2

$$E(\check{X}_t - C_{1t})^2 \propto (\lambda b_t^2)^2 \qquad \text{large } t \qquad (2.39)$$

which is the square of the power of the Poisson driven part of the system. Now, let us study the performance of the optimal filter in terms of its error distribution over the Wiener driven part and the Poisson driven part. Assume further that the system is stable and has a steady state after some time $t' > 0$.

From Eq. (2.23b), we have

$$\lim_{t\to\infty} E(\tilde{C}_{2\lambda t}) = \frac{-b_t^2}{2[\mathcal{C}_t - h_t^2 v_t^{-1} \tilde{P}(t)]} \geq 0 \qquad (2.40)$$

Writing

$$\begin{aligned} C_{2t} &= \tilde{C}_{2t} + \lambda \tilde{C}_{2\lambda t} + O(\lambda) \\ &= \tilde{P}(t) + \lambda \tilde{C}_{2\lambda} + O(\lambda) \end{aligned} \qquad (2.41)$$

For large t, we have

$$E(C_{2t}) = \bar{P}(t) + \frac{-\lambda b_t^2}{2[\alpha_t - h_t^2 v_t^{-1} \bar{P}(t)]} + o(\lambda) \tag{2.42}$$

The first quantity on the right side of the equation is the variance in filtering the Wiener driven part i.e., the continuous part of the system. The second quantity is that due to the filtering of the Poisson driven part. Notice that it depends on $\bar{P}(t)$.

Comparing the error variance in filtering the Poisson driven part of the system to its variance, we have the ratio, denoted by ϵ,

$$\begin{aligned} \epsilon &= \lim_{t\to\infty} \frac{-\lambda b_t^2}{2[\alpha_t - h_t^2 v_t^{-1} \bar{P}(t)]} \Big/ \frac{-\lambda b_t^2}{2\alpha_t} \\ &= \lim_{t\to\infty} \frac{\alpha_t}{\alpha_t - h_t^2 v_t^{-1} \bar{P}(t)} \end{aligned} \tag{2.43}$$

Hence when there is no Wiener driven part in the system,

$$\bar{P}(t) = 0$$

making $\epsilon = 1$ (2.44)

In fact it is shown by the author that for system driven only by Poisson inputs, the steady-state error variance of the linear optimal filter is at least as large as the variance of the system [12]. This manifests the fact that the optimal filter and the Kalman filter, both being instantaneous filters fail to estimate the jumps. This can also be seen by noting that if the system is a fast decaying system i.e. $|\alpha_t|$ is large, the filter fails to estimate the jumps, giving $\epsilon \to 1$. And the quantity in (2.43)

$h_t^2 v_t^{-1}$ denotes the signal-to-noise ratio in the observations. As it increases, the error decreases. From Eq. (2.43), the ratio ϵ decreases as $\tilde{P}(t)$ increases for fixed power of the Poisson driven part. This appears at first to be contradictary. This is due to the fact that $\tilde{P}(t)$ only increases with the power of the Wiener driven part of the system. If the energy of the Wiener part in the system dominates that of the Poisson driven part, the error made by the filter will be dominated by the estimate error of the Wiener part, i.e. $\tilde{P}(t)$. This is a direct consequence of the ratio of the power of the Wiener part and that of the Poisson part in the system itself. Hence in systems where the Wiener part dominates over the Poisson part, a reasonable approximate filter can be constructed by simply neglecting the Poisson part in the system.

In case, the opposite is true, the power of the Poisson part dominates, both of the Kalman filter and the optimal filter performs very badly as an instantaneous estimate of the system.

2.4 Conclusions

In this chapter, we derive the optimal filter for the problem. We deduce that the set of equations defining the optimal filter is infinite (non-closed) by arguing that there is no additional relationship that would reduce the set to a finite set. If such a relationship exists, then we are guaranteed to have a solution in the strong sense for C_{1t} because the finite set of equations in our case is guaranteed to have a strong solution, In such case, the process I defined in (2.13) is the innovation process, and hence a Wiener process with respect to the σ-field generated by the observations. For a finite dimensional vector differential equations as the reduced filter equations driven by a Wiener process, the solution is always continuous with probability one. Hence we would have a continuous estimator for a discontinuous system, contradicting the optimality assumption of the estimator.

We further argued that the solution for C_{1t} does not exist in the strong sence by showing that I in fact is not an innovation process in our problem. The optimal filter is unrealizable and unimplementable.

For the performance of the optimal filter, we compare it with that of the linear optimal filter. At least for the case of stable systems, the improvement in performance of the optimal filter over the linear optimal filter is extremely small when λ is small.

We shall show that the linear optimal filters (causal and non-causal) which utilize the first and second order statistics only, perform extremely poorly in the estimation problem of the Poisson driven systems. Hence the optimal filter being a causal estimator is not suitable for the state estimation problem for Poisson driven systems.

CHAPTER 3

ESTIMATION OF THE POISSON DRIVEN PROCESSES

3.1 General

In this chapter the optimal filter and suboptimal approximation estimation scheme for the Poisson driven case will be considered. It can be easily seen that by setting $\lambda = 0$ in the optimal filter Eq. (2.9), the problem reduces to the standard linear quadratic Gaussian problem and the filter reduces to the standard Kalman filter which is finite dimensional (closed). The nonclosedness of the optimal filter (2.9) arises from the Poisson driven part of the system due to its non-Gaussian property and hence an infinite number of moments are required to characterize completely the filter.

The optimal filter for the purely Poisson driven process may be obtained from Eq. (2.9) by setting $\beta_t \equiv 0$ and is given as follows [2,3,28]

$$dC_{1t} = \alpha_t C_{1t} dt + h_t v_t^{-1} C_{2t} dI_t \qquad (3.1a)$$

$$dC_{kt} = k\alpha_t C_{kt} dt + \lambda b_t^k dt - \frac{1}{2} h_t^2 v_t^{-1} \sum_{j=1}^{k-1} \binom{k}{j} C_{j+1,t} C_{k-j+1,t} dt$$
$$+ h_t v_t^{-1} C_{k+1,t} dI_t \qquad k = 2,3,4\ldots \qquad (3.1b)$$

Note that the filter as in Eq. (2.9) is still infinite dimensional. The Wiener part indeed only comes into play in the variance equation, i.e. the C_{2t} equation.

The problem may be generalized by including a mark process generated by the Poisson driving input. The problem in this case is a modification of (2.1) and the process is defined by

$$dx_t = a_t x_t dt + b_t(U) d\eta_t \qquad t \geq 0 \qquad (3.2a)$$

where U denotes the mark of the Poisson process η_t and is assumed to have a known probability density f(U). By definition of the mark, the random variable U will be generated independently according to f(U) at every incident of the Poisson process η_t. The mark can be considered as a random parameter generated by each incident, e.g., the marks represent the random magnitudes of the Poisson impulses.

Let the i-th incident time and mark of the Poisson process be denoted by τ_i, u_i respectively. Also denote the number of incidents over the semi-closed interval [s,t) by N[s,t) or if s = 0, N(t) for simplicity.

The objective is to obtain a near-optimal estimation scheme for x_t from noisy observations viz.

$$dy_t = h_t x_t dt + dv_t \qquad (3.2b)$$

3.2 The Optimal Linear Filter

As indicated in Chapter 2, the improvement of the optimal filter (which is unimplementable) over the linear optimal filter is relatively small in terms of error variance. We establish in the section that the linear optimal filter for the Poisson driven case with low intensity is extremely poor. Consequently both of the Kalman filter and the optimal filter are unacceptable in their performances. This is basically due to the fact that the system process to be estimated is only right-continuous, making instantaneous estimation or detection of the jumps in its sample paths formidable.

From Eq. (2.44), we have seen that the normalized error variance of the optimal filter tends to unity as $\lambda \to 0$. Now we show that in the <u>non-causal</u> linear optimal smoothing, the steady state normalized error variance for the time invariant case also tends to unity <u>uniformly</u> as $\lambda \to 0$. This result indicates that an improvement over the optimal linear noncasual filter may be possible only if a time-delay is allowed in a nonlinear filtering scheme. In order to present the derivation, the system (3.2a) will first be reformulated and expressed as

$$dx_t = \mathcal{a}_t x_t dt + \int_U b_t(U)\ M(dt,dU) \tag{3.3}$$

where M(·) denotes the measure of the underlying Poisson marked process. Let $\Phi(t,s)$ denote the state transition function of the homogeneous part of the system, then

$$x_t = \Phi(t,0)x_0 + \int_0^t \int_U \Phi(t,s)\ b_s(U)M(ds,dU) \tag{3.4}$$

Hence from Eq. (3.2b), the observations become

$$y_t = v_t + h_t\Phi(t,0)x_0 + \int_0^t \int_U h_t\Phi(t,s)b_s(U)M(ds,dU) \tag{3.5}$$

By definition of the stochastic counting integral whose existence and uniqueness only requires the point process M to be conditional orderly and with probability one have a finite number of points in a finite interval [17, 18, 29] which are all satisfied by our Poisson assumption of η_t, we obtain the following

$$x_t = \Phi(t,0)x_0 + \sum_{i=1}^{N(t)} g(t)\ \tau_i;u_i) \qquad t \geq 0 \tag{3.6}$$

where

$$g(t,\tau;u) = \Phi(t,\tau)b_t(u) \tag{3.7}$$

which is the integrand in the integral in (3.4).

If we assume x_0 is known (this assumption may be removed without difficulties), a new observation process z_t may now be defined by

$$z_t \stackrel{\Delta}{=} y_t - h_t\Phi(t,0)x_0 \tag{3.8}$$

$$= v_t + \sum_{i=1}^{N(t)} h(t,\tau_i;u_i) \qquad t \geq 0$$

$$= v_t + \sum_{i=1}^{N(T)} h(t,\tau_i;u_i) \qquad 0 \leq t < T \tag{3.8b}$$

where

$$h(t,\tau;u) = h_t\Phi(t,\tau)b_t(u) \tag{3.9}$$

Equation (3.8b) follows from the causality of g and h in the integral, viz

$$g(t,\tau;u) = 0 = h(t,\tau;u) \qquad \text{for} \qquad t < \tau. \tag{3.10}$$

We now examine the performance of linear optimal smoothing. The case of *finite interval and infinite interval linear smoothing are both considered.*

For the fixed interval smoothing over an interval of length T, the steady state error variance for the optimal linear smoother, denoted by ϵ_T is given from [19, pp.256], as follows

$$\epsilon_T = \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi_x(\omega)d\omega - \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{|\phi_{zx}(\omega)|^2}{\phi_z(\omega)}d\omega + \int_{-\infty}^{-T}|G(t)|^2dt$$

where ϕ_x, ϕ_z are the spectral densities of x(t) and z(t) respectively, ϕ_{zx}, the cross spectral density of x(t) and z(t). G(t) is an appropriate time-function derived from the spectral densities. Therefore

$$\epsilon_T \geq \epsilon_\infty = \frac{1}{2\pi}\int_{-\infty}^{\infty}\phi_x(\omega)d\omega - \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{|\phi_{xz}(\omega)|^2}{\phi_z(\omega)}d\omega .$$

The substitution of the spectral densities with the appropriate expressions in terms of the transforms of the incident response [1], and using the notation

$$\overline{h^2(t,\tau;u)} = \underset{U}{E}\, h^2(t,\tau;U)$$

results in

$$\epsilon_T \geq \frac{1}{2\pi}\int_{-\infty}^{\infty}\lambda|H(j\omega;\bar{u}|^2 d\omega - \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\lambda^2|H(j\omega;\bar{u})|^4 d\omega}{\sigma^2+\lambda|H(j\omega;\bar{u})|^2}$$

Normalization with the signal variance which is given by

$$\mathrm{Var}(x(t)) = \lambda\int_{-\infty}^{\infty}h^2(t-\tau,\bar{u})d\tau \tag{3.11}$$

finally yields

$$\frac{\epsilon_T}{\mathrm{Var}(x_t)} \geq \left[\frac{1}{2\pi}\int_{-\infty}^{\infty}|H(j\omega;\bar{u})|^2 d\omega - \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{|H(j\omega;\bar{u})|^4 d\omega}{(\frac{\sigma^2}{\lambda})+|H(j\omega;\bar{u})|^2}\right] \Big/ \int_{-\infty}^{\infty}h^2(t-\tau,\bar{u})d\tau \tag{3.12}$$

The behavior of the normalized error variance as $\lambda \to 0$ becomes therefore

$$\lim_{\lambda\to 0}\frac{\epsilon_T}{\mathrm{Var}(x_t)} \geq \frac{\frac{1}{2\pi}\int_{-\infty}^{\infty}|H(j\omega;\bar{u})|^2 d\omega}{\int_{-\infty}^{\infty}h^2(t;\tau,\bar{u})d\tau} = 1 \tag{3.13}$$

where the Pareseval's Theorem has been used. From Eq. (3.13) we can deduce that the performance of the linear optimal filter or smoother is extremely poor regardless of the signal-to-noise ratio and behavior of the incident responses when λ is small, since the normalized variance tends to unity <u>uniformly</u> as $\lambda \to 0$. This result is not too surprising since linear techniques only utilize first and second order statistics, and the Poisson driven process which is highly "discontinuous" contains considerable energy in the higher order statistics.

Since causal estimation for such systems is inherently of poor quality when λ is small, the problem leads naturally to the alternative of allowing some delay in the estimation process i.e., smoothing. In view of the poor performance of the linear optimal smoothing and the nonclosedness of the optimal estimates, a suboptimal non-linear smoothing scheme will be considered for the estimation problem. While the scheme may not perform as well as ultimately possible, it is expected to perform better than the optimal linear infinite-interval smoother for low enough incident rate λ.

3.3 The Sequential Smoothing Scheme (SSS)

In this section, we derive a sequential smoothing strategy to the problem given by (3.2). The system to be estimated is a Poisson marked process; hence the knowledge of the realization of the Poisson driving process η_t and its associated mark process is sufficient to determine the realization of x_t.

The approach to the estimation problem considered here is to first solve the particular problem of estimating the incident process η_t and U, via a combined sequential estimation and detection scheme based on the criterion of maximum aposteriori probability (MAP). Then the estimate of any function $\phi(t,\eta_t,U)$ is reconstructed suboptimally through the estimate of η_t and U, namely $\phi(t,\hat{\eta}_t,\hat{U})$. Although this approach is by no means optimal in the sense of MMSE, this is nevertheless a robust alternative when the optimal scheme is intractable. In the proposed sequential scheme, the observations are processed in subintervals whose length, Δ, is chosen such that the probability that each component of the driving process having two or more than two incidents within the interval is negligibly small.

Each subinterval of observations is processed to detect and estimate, if it exists, the incident in the interval as well as to update the estimates of past incidents.

In order to reduce the memory requirement and computational complexity of the scheme, a finite memory scheme is selected whose memory size depends on the marginal improvements in the performance and the desired performance relative to an infinite memory scheme. Consequently we choose not to update (re-estimate) the estimates of the past incidents occurring earlier than L subintervals away from the new intervals as shown in Fig. 1.

Note that there is a maximum delay $L\Delta$ inherent in the structure of the scheme. Any estimates with delay longer than $L\Delta$ are the same as that with delay $L\Delta$ since they have been finalized. L usually is large and is chosen as a tradeoff between performance and computational complexity. Also note that the scheme structure is actually independent of the dealy imposed by the original problem. The delay reconstruction process is carried out independently as the scheme sequentially updates the estimates of the incidents.

3.3.1 Preliminary Requirements for the SSS

To be able to apply the SSS described above, there are a few basic conditions we have to satisfy due to the combined nature of the detection estimating scheme.

Let $(\Omega,\mathcal{F},P)$ be the underlying probability space for our estimation problem. Then the random processes defined in the problem statement are all measurable random processes in continuous time $t \in [0,\infty)$. For example, if B is a Borel set $\in \mathcal{B}$ of the real line of R^1, η_t satisfies

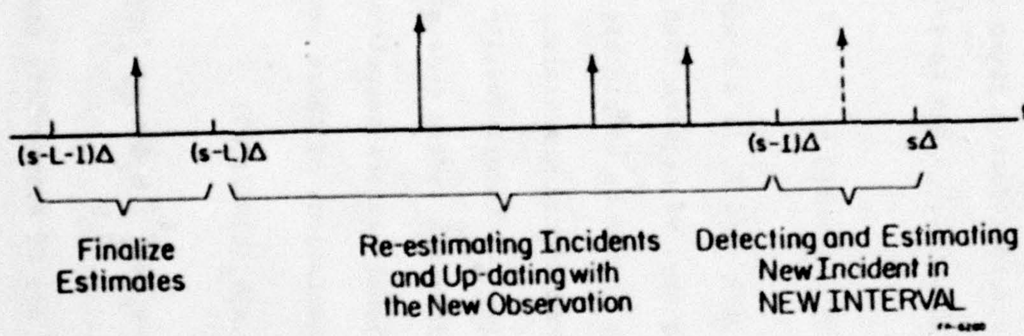

Figure 1. Schematic representation of the sequential combined scheme: the additional observations over $[(s-1)\Delta, s\Delta)$ is used to update the estimates over $[(s-L)\Delta, (s-1)\Delta)$ and finalized that in $[(s-L-1)\Delta, (s-L)\Delta)$.

$$\{(\omega,t) : \eta_t(\omega) \in B\} \in \mathcal{F} \times \mathcal{B}[0,\infty) \tag{3.14}$$

where $\mathcal{B}[0,\infty)$ is a σ-algebra of Borel sets on interval $[0,\infty)$. That simply says the random process η_t besides having well-defined trajectories can always be characterized in probablistic terms.

In the SSS, in order to reconstruct a given functional $G(t,\eta_t(U))$ based on the estimate of $\eta_t(U)$, $G(t,\eta_t)$ has to satisfy the following condition,

$$\{(\omega,t): \; G(t,\eta_t(\omega)) \in B\} \in \mathcal{F}_t^{\eta} \times \mathcal{B}[0,t) \qquad \forall \; t \in [0,\infty) \tag{3.15}$$

where B is a Borel set $\in \mathcal{B}$ of R^1 and $\mathcal{F}_t^{\eta}$ is a σ-algebra generated by $\{\eta_s, \; 0 \leq s < t\}$. That is to say, we require that $G(t,\eta_t)$ be independent of the future of η_t and that it can be completely and deterministically determined by the knowledge of $\{\eta_s : 0 \leq s < t\}$.

To examine the conditions for recoverability of η_t based on the observation y_t, let us consider the vector case of the problem. Let x_t, η_t and y_t be n-,p-,q-dimensional vectors respectively, then Eqs. (3.4) and (3.5) are still valid representations of these processes, viz.

$$x_t = \Phi(t,0)x_0 + \int_0^t \int_U \Phi(t,s)b_s(U)M(ds,dU) \tag{3.16}$$

$$z_t \triangleq y_t - h_t\Phi(t,0)x_0 = v_t = \int_0^t \int_U h_t\Phi(t,s)b_s(U)M(ds,dU) \tag{3.17}$$

where now h_t, $B_t(U)$, $\Phi(t,s)$ are of appropriate dimensions. To estimate η_t, it is essential that the observation z_t is a non-trivial function of all the p components of η_t and their corresponding mark processes. From Eq. (3.17), it follows quite immediately that z_t satisfies this requirement if the following holds

(1) the $q \times p$ matrix $h_t \Phi(t,s) b_s(U)$ contains no identitically zero columns almost everywhere in $t \geq s \geq 0$ and U.

(2) the system input $n \times p$ matrix $b_s(U)$ contains no columns completely independent of U.

Note that condition (2) can be easily satisfied with a modification of the system equations. Since our main concern is to reconstruct X_t, it is already satisfied. Examining condition (1), we find that it is a considerably relaxed condition compared to the usual requirements of controllability and observability of the system in estimating a Wiener driven system.

3.3.2 The Estimation and Detection Equations

To obtain the equations for the joint estimation and detection, we first consider the simple MAP interval estimation and detection problem over an interval of length T. Therefore, the estimates of the incidents and marks, denoted respectively by $\hat{\underline{\tau}}, \hat{\underline{u}}$, have to satisfy the following

$$\ln f_{N(T),\underline{\tau},\underline{u}}(\hat{N},\hat{\underline{\tau}},\hat{\underline{u}}|\mathcal{F}_T) = \max_{N^*,\underline{\tau}^*,\underline{u}^*} \ln f_{N(T),\underline{\tau},\underline{u}}(N^*,\underline{\tau}^*,\underline{u}^*|\mathcal{F}_T^z) \tag{3.18}$$

where $f(\cdot)$ is the conditional density of N(T) and τ,u, the vectors containing all the incident times and marks over $[0,T\Delta)$. $\mathcal{F}_T^z$ is the σ-field generated by the observations $\{z(t);0 \leq t < T\}$. Since the dimensions of $\underline{\tau},\underline{u}$ depends on N(T), the maximization can be carried out first by obtaining the MAP estimates of $\underline{\tau},\underline{u}$ assuming the value of N(T) is fixed as follows:

$$\max_{N^*,\underline{\tau}^*,\underline{u}^*} \ln f_{N(T)\underline{\tau},\underline{u}}(N^*,\underline{\tau}^*,\underline{u}^*|\mathcal{F}_T^z) = \max\{\max_{N^*,\underline{\tau}^*,\underline{u}^*} \ln f_{N(T),\underline{\tau},\underline{u}}(N^*,\underline{\tau}^*,\underline{u}^*|\mathcal{F}_T^z)\} \tag{3.19}$$

$$= \max_{N^*} \ln f_{N(T)\tau,u}(N^*,\tau(N^*),u(N^*)|\mathcal{F}_T^z) \tag{3.20}$$

where $\hat{\underline{\tau}}(N^*)$, $\hat{\underline{u}}(N^*)$ stand for the MAP estimates of $\underline{\tau},\underline{u}$ given $N(T) = N^*$.

The maximization of $\underline{\tau}^*,\underline{u}^*$ in Eq. (3.19) gives a set of equations for $\hat{\underline{\tau}}(N^*)$ and $\hat{\underline{u}}(N^*)$, the estimation equations. The maximization over N* in Eq. (3.20) gives the detection equations. By obtaining an expression for ℓnf, which is given as

$$\ell nf_{N(T),\underline{\tau},\underline{u}}(N^*,\underline{\tau}^*,\underline{u}^*|\mathcal{F}^Z_T) = \ell nP(\{z_t;0\leq t<T\}|\underline{\tau}^*,\underline{u}^*,N^*)$$

$$+\ \ell nf(\tau^*|N^*) + \ell nf(u^*/N^*) + \ell nP(N^*) \qquad (3.21)$$

and recognizing that the first quantity on the right-hand side of the equation has an equivalent expression, namely the log-likelihood function of $\underline{\tau}^*,\underline{u}^*$ which is given as

$$\frac{1}{\sigma^2}\int_0^T dt[z_t - \sum_{i=1}^{N^*} h(t,\tau_i^*;u_i^*)][\sum_{j=1}^{N^*} h(t,\tau_i^*;u_i^*)]$$

we obtain the following set of equations by optimizing over $\underline{\tau}^*,\underline{u}^*$,

$$\frac{1}{\sigma^2}\int_0^T dt[z(t) - \sum_{i=1}^{N(T)} h(t;\hat{\tau}_i,\hat{u}_i)][\frac{\partial h(t;\hat{\tau}_j,\hat{u}_j)}{\partial \hat{u}_j}] + \frac{\partial}{\partial \hat{u}_j}\ell nf(\hat{u}_j) = 0 \qquad (3.22a)$$

$$\frac{1}{\sigma^2}\int_0^T dt[z(t) - \sum_{i=1}^{N(T)} h(t;\hat{\tau}_i,\hat{u}_i)][\frac{\partial h(t;\hat{\tau}_j,\hat{u}_j)}{\partial \hat{\tau}_j}] = 0 \qquad (3.22b)$$

for $1\leq j\leq N(T)$ and $0<\tau_1,<\tau_2\ldots<\tau_{N(T)}<T$. Equation (3.22) actually is the necessary condition for the MAP estimates of $\underline{\tau},\underline{u}$ given N(T) and always guarantees a set of solution. The uniqueness of the solution set does not play a role in the scheme since we have to maximize over N(T) subsequently. The expression to be maximized with respect to N is obtained by substituting the distributions associated with the Poisson process into Eq. (3.21) which reduces to

$$J(N) \overset{\Delta}{=} N \ln\lambda + \sum_{i=1}^{N} \ln f_U(\hat{u}_i)$$
$$+ \frac{1}{\sigma^2}\int_0^T dt[z(t) - \frac{1}{2}\sum_{j=1}^{N} h(t;\hat{\tau}_j,\hat{u}_j)][\sum_{i=1}^{N} h(t;\hat{\tau}_i,\hat{u}_i)] \tag{3.23}$$

3.3.3 Sequential Approximation Approach

The resulting MAP scheme defined by the solution of (3.18) and the maximization of (3.19) will now be approximated by a sequential scheme. In the scheme, we sequentially process the observations in subintervals each of length Δ, which is chosen such that the probability of having two or more incidents in each interval is negligibly small. Hence, we choose Δ such that

$$\lambda\Delta \overset{\Delta}{=} \alpha \ll 1 \tag{3.24}$$

The observation over the k^{th} subinterval $[(k-1)\Delta, k\Delta)$ is used to finalize the estimates of the incidents prior to $(k-L)\Delta$. We define the new information carrying observation $\tilde{y}(t,k)$ with the removal of the finalized estimates by

$$\tilde{y}(t,k) = z(t) - \sum_{i=1}^{N((k-L)\Delta)} h(t;\hat{\tau}_i,\hat{u}_i), \qquad t \in [(k-L)\Delta, k\Delta) \tag{3.25}$$

This new observation in (3.25) will be used in the MAP equations (3.18) and (3.19) to yield the approximate sequential estimation-detection scheme. The equations to be satisfied by the unfinalized MAP estimates $\hat{\underline{\tau}}, \hat{\underline{u}}$ in the interval $[(k-L)\Delta, k\Delta)$ when the k^{th} subinterval is observed, are obtained from (3.22) and (3.25) and may be written as

$$\frac{1}{\sigma^2}\int_{(k-L)\Delta}^{k\Delta} dt[\tilde{y}(t,k) - \sum_{i=N((k-L)\Delta)+1}^{N(k\Delta)} h(t;\hat{\tau}_i,\hat{u}_i)][\frac{\partial h(t;\hat{\tau}_j,\hat{u}_j)}{\partial \hat{u}_j}]$$
$$+ \frac{\partial}{\partial \hat{u}_j} \ln f(\hat{u}_j) = 0 \tag{3.26a}$$

$$\frac{1}{\sigma^2}\int_{(k-L)\Delta}^{k\Delta} dt[\tilde{y}(t,k) - \sum_{N((k-L)\Delta)+1}^{N(k\Delta)} h(t;\hat{\tau}_i,\hat{u}_i)][\frac{\partial h(t;\hat{\tau}_j,\hat{u}_j)}{\partial \hat{u}_j}] = 0 \qquad (3.26b)$$

for $N((k-L)\Delta)+1 \leq j \leq N(k\Delta)$

$$(k-L)\Delta \leq \tau_{N((k-L)\Delta)+1} < \ldots < \tau_j < \tau_{j+1} < \ldots \tau_{N(k\Delta)} < k\Delta.$$

These equations are to be solved for fixed $N(k\Delta)$ (whose values will be discussed in the sequel) and then the proper choice of $\hat{N}(k\Delta)$, the estimate of $N(k\Delta)$ should be made. The estimate $\hat{N}(k\Delta)$ of $N(k\Delta)$ is made by maximizing the following expression which is obtained directly from (3.23) over the value of N,

$$J_1(N) = N\ell n\lambda + \sum_{i=N((k-L)\Delta)+1}^{N} \ell nf_U(\hat{u}_i) + \frac{1}{\sigma^2}\int_{(k-L)\Delta}^{k\Delta} dt[\tilde{y}(t,k)$$

$$- \frac{1}{2} \sum_{i=N((k-L)\Delta)+1}^{N} h(t;\hat{\tau}_i,\hat{u}_i)][\sum_{j=N((k-L)\Delta)+1}^{N} h(t;\hat{\tau}_j,\hat{u}_j)] \qquad (3.27)$$

Note that the detection is carried out sequentially with the assumption of Eq. (3.24), which limits the values of $N(k\Delta)$ for the maximization to a set of 3 values as follows:

(1) If an incident was detected in the previous interval $[(k-2)\Delta,(k-1)\Delta)$

$$N(k\Delta) \in \{N((k-1)\Delta),\ \ n((k-1)\Delta)+1,\ \ n((k-1)\Delta)+1\} \qquad (3.28a)$$

The third value is included in case that the previously detected incident has been a false detection.

(2) Similarly, if no incident was detected in the previous interval $[(k-2)\Delta,(k-1)\Delta)$, then $N(k\Delta)$ is limited to the set.

$$N(k\Delta) \in \{N((k-1)\Delta)\ N((k-1)\Delta+1,\ \ N((k-1)\Delta)+2\} \qquad (3.28b)$$

where the last value accounts for the possibility of a miss in the earlier subinterval. Note also that for each value of $N(k\Delta)$, it is required to solve a set of equations (3.26) for the estimates $\hat{\underline{\tau}}(N)$, $\hat{\underline{u}}(N)$. However, these sets of equations are uncoupled to each other and can be solved in parallel. Since all of the incidents except the new incident (if it exists) have all been previously evaluated, iterative perturbation methods may be employed to solve them.

The set of values for $N(k\Delta)$ can obviously be increased to include more values to cover the possibility of multiple wrong detections in previous subintervals.

3.4 Performance of the SSS

3.4.1 Asymptotic Performance of the SSS

The analysis of the performance of the sequential approximation scheme is quite involved; exact analytic results are rather complex to obtain. We consider only the scalar time-invariant case and derive the asymptotic performance of the scheme as the intensity tends to zero. For comparison purpose, we express the performance in terms of the error variance of the estimate of the system state normalized by the variance of the state (average power).

When the intensity is small, the estimation error resulting from the scheme can be modeled as Poisson filtered process, with each error incident response as a result of the detection and estimation over each subinterval. Since we are making estimation and detection for every subinterval, the intensity of the driving process is $1/\Delta$. Therefore the estimation error, denoted by $W(t)$ can be represented as

$$W(t) = \int_V \int_0^t \epsilon(t;\sigma,V)\, \mathcal{M}(d\sigma,dV) \tag{3.29}$$

where $\epsilon(t;\sigma,V)$ is the incident response of the average error made according to the result of the detection indicated by the mark V, and $\mathcal{M}$ is the measure of the underlying error driving provess with intensity $1/\Delta$.

Define the mark V as follos

$$\begin{aligned} V &= d_{10} \quad \text{indicating a miss in the detection} \\ V &= d_{11} \quad \text{indicating a correct detection of an incident} \\ V &= d_{01} \quad \text{indicating a false alarm} \\ V &= d_{00} \quad \text{indicating a correct detection of no incident} \end{aligned} \tag{3.30}$$

Hence, the distribution of the mark is given by

$$\begin{aligned} p(V = d_{10}) &= \alpha p_{10}(\Delta) \\ p(V = d_{11}) &= \alpha(1 - p_{10}(\Delta)) \\ p(V = d_{01}) &= (1-\alpha)p_{01}(\Delta) \\ p(V = d_{00}) &= (1-\alpha)(1-p_{01}(\Delta)) \end{aligned} \tag{3.31}$$

where $p_{01}(\Delta)$, $p_{10}(\Delta)$ denote the probability of a false alarm and a miss in the detection over an interval of length Δ.

Note that with this model,

$$\epsilon^2(t\ \tau;d_{00}) = 0 \tag{3.32}$$

$$\epsilon^2(t\ \tau;d_{01}) = \underset{U}{E}\ h^2(t;\tau,U) \tag{3.33}$$

The variance of W(t) is therefore

$$\mathrm{Var}(W(t)) = \frac{1}{\Delta}\ \underset{V}{E} \int_0^\infty \epsilon^2(t\ 0;V)dt \tag{3.34}$$

$$= \frac{\lambda}{\alpha}\ \underset{V}{E} \int_0^\infty \epsilon^2(t\ 0;V)dt \tag{3.35}$$

The substitution of Eq. (3.31) into (3.34) yields

$$\mathrm{Var}(W(t)) = \frac{\lambda}{\alpha}\{p_{01}(\Delta)\int_0^\infty \epsilon^2(t,0;d_{01})dt$$

$$+ \lambda\Delta[1 - p_{10}(\Delta)]\int_0^\infty \epsilon^2(t,0;d_{11})dt$$

$$+ \lambda\Delta\int_0^\infty dt[p_{10}(\Delta)\epsilon^2(t,0;d_{10}) - p_{01}(\Delta)\epsilon^2(t,0;d_{01})]\} \tag{3.36}$$

which after normalization by Var(X(t)) given in (3.11), becomes in the limit as $\lambda \to 0$,

$$\lim_{\lambda\to 0}\frac{\mathrm{Var}(W(t))}{\mathrm{Var}(X(t))} = \lim_{\lambda\to 0}\frac{p_{01}(\frac{\alpha}{\lambda})}{\alpha} = \frac{1}{\alpha}p_{01}(\infty) \tag{3.37}$$

where $p_{01}(\infty)$ is now the probability of wrong detection in a semi-infinite interval. The probability of false alarm, $p_{01}(\Delta)$ is given by (see [12])

$$p_{01}(\Delta) = \mathrm{erf}\left(\frac{\ln(1/\alpha)}{\sqrt{e_\Delta}} + \frac{\sqrt{E_\Delta}}{2}\right) \tag{3.38}$$

where

$$E_\Delta \geq \frac{1}{\sigma^2}\underset{U}{E}\int_0^\Delta h^2(t;0,U)dt \tag{3.39}$$

In fact when λ is small, we have

$$\frac{p_{01}(\Delta)}{\alpha} \leq \frac{\mathrm{Var}(W(t))}{\mathrm{Var}(X(t))} \leq \frac{p_{01}(\Delta)}{\alpha} + \lambda\left[\frac{\int_0^\infty \epsilon^2(t;0,d_{11})dt}{\underset{U}{E}\int_0^\infty h^2(t;0,U)dt}\right] \tag{3.40}$$

Hence, we expect the normalized error variance to approach $p_{01}(\infty)/\alpha$ as $\lambda \to 0$. Furthermore it increases linearly as λ with slope

$$\left[\frac{\int_0^\infty \epsilon^2(t;0,d_{11})dt}{\underset{U}{E}\int_0^\infty h^2(t;0,U)dt}\right]$$

which is expected to be very small when the signal-to-noise ratio is reasonably high, i.e. when

$$\int_0^\infty \epsilon^2(t;0,d_{11})dt \ll \underset{U}{E} \int_0^\infty h^2(t;0,U)dt \tag{3.41}$$

Consequently the normalized variance is approximately $p_{01}(\Delta)/\alpha$ when λ is small.

$$\frac{Var(W(t))}{Var(X(t))} \approx \frac{p_{01}(\Delta)}{\alpha} \tag{3.42}$$

In view of the fact that the normalized variance of the error resulted from the linear optimal filter tends to unity regardless of the signal-to-noise ratio and behavior of the incident responses, the scheme performs extremely well over the linear optimal filter, since under normal conditions of signal-to-noise ratio

$$\frac{p_{01}(\Delta)}{\alpha} \ll 1 \tag{3.43}$$

when λ is small. The results imply that there exists a rate $\lambda^* > 0$ such that for $\lambda \in (0,\lambda^*]$ the suboptimal nonlinear scheme performs better than the optimal linear noncausal scheme.

3.4.2 Performance Simulation of the SSS

The sequential scheme was simulated on a digital computer; in each case, it was done over a total length of time to include 20 incidents. The performance in terms of normalized error variance was plotted against the intensity ranging from 0 to 2. L was chosen to be 4 and $\alpha = 0.15$. The system state was reconstructed with delay of 2.

The computed performance, Eq. (3.42) was plotted as a continuous graph in each case and the simulated performance for different values of λ were plotted as points with appropriate notations. For comparison purposes, the computed normalized steady state error variance of the linear optimal filter (LOF) was larger than 0.90 in all the cases over the range $\lambda = 0$ to 2. Several examples were considered as follows:

Example I: $\dot{x}(t) = -2x(t) + U\eta(t)$ $\qquad x(0) = 0$

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

Hence $h(t;\tau,U) = Ue^{-2(t-\tau)}$

$$y(t) = \sum_{i=1}^{N(t)} U_i e^{-2(t-\tau)} \qquad t \geq 0$$

The mark U is assumed to be Gaussian with variance 5 and three possible cases for the mean: (i) 5 (ii) 7 (iii) 10 (see Fig. 2).

Example II: $\dot{x}(t) = U\eta(t)$ $\qquad x(0) = 0$

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

The incident response is a step function indicated by $\mu(\cdot)$,

$$h(t;\tau,U) = U\mu(t-\tau)$$

Again, U is Gaussian with variance 5 and mean taking one of three values: (i) 5 (ii) 7 (iii) 10 (see Fig. 3).

Example III: The incident response is a rectangular function,

$$h(t,\tau;U) = U[\mu(t-\tau) - \mu(t-\tau-L)]$$

We include two cases $L = 0.2$ and 0.5 with same Gaussian U as above (see Fig. 4).

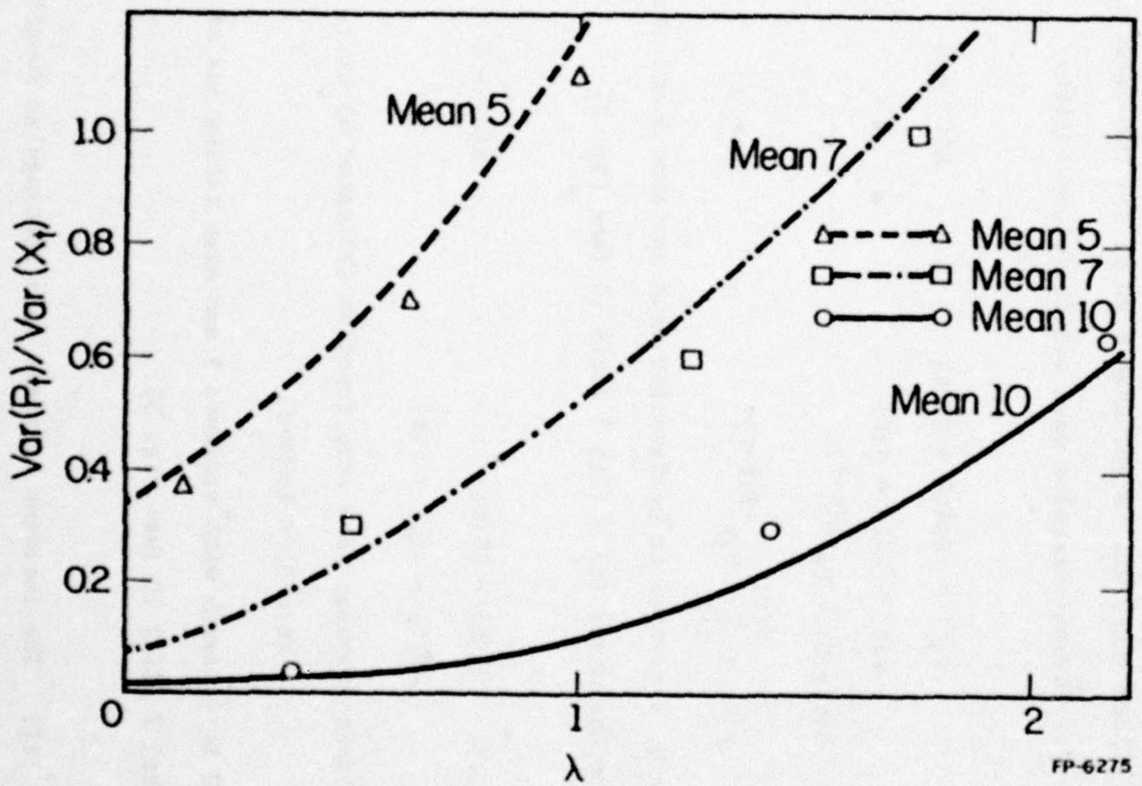

Figure 2. Simulated and computer performance, $Var(\epsilon_t)$ vs. λ, for Example I.

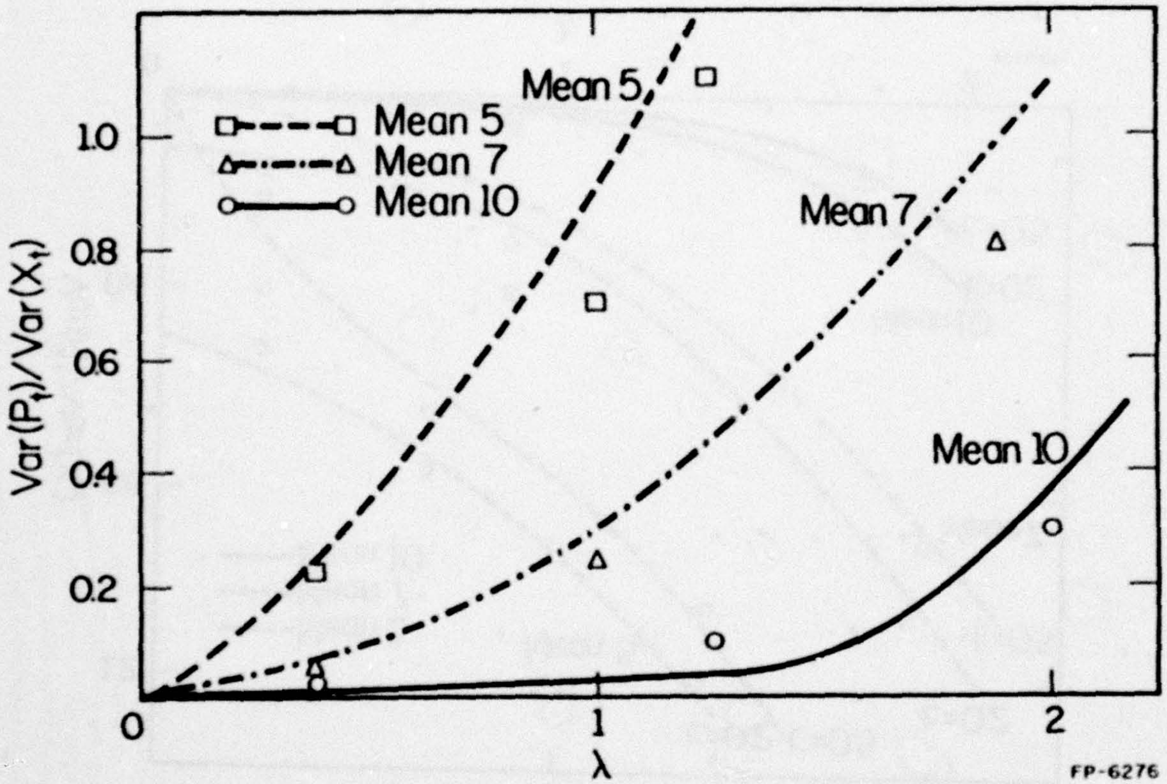

Figure 3. Simulated and computed performance, $Var(\epsilon_t)$ vs. λ, for Example II.

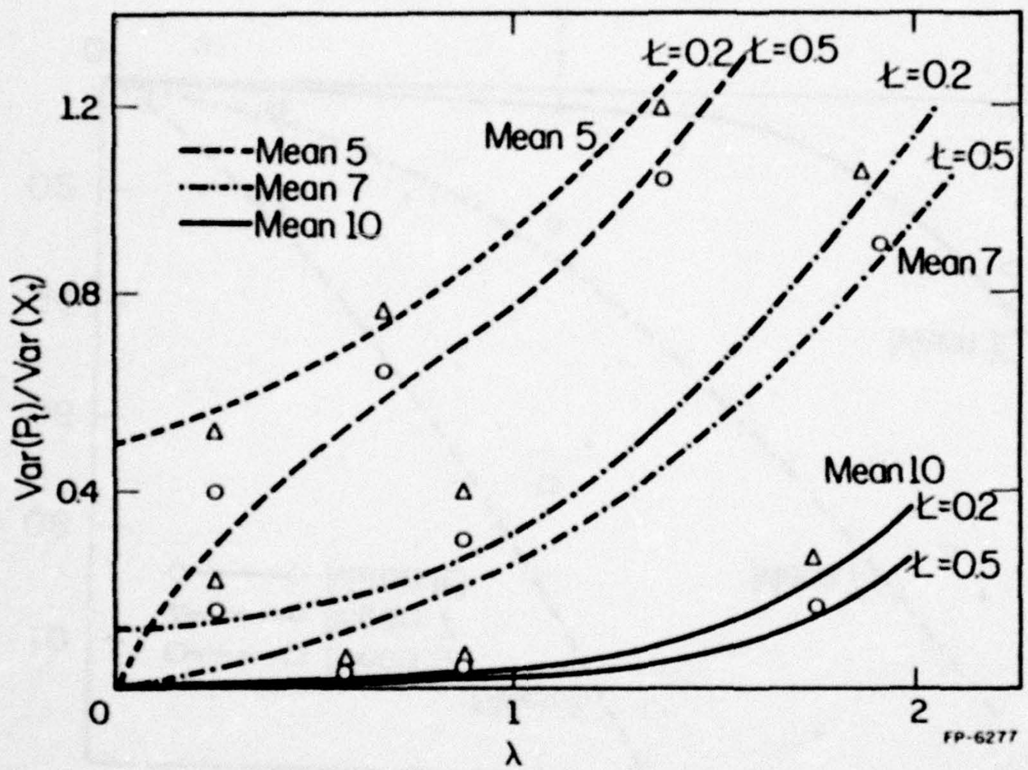

Figure 4. Simulated and computed performance, $Var(\epsilon_t)$ vs. λ, for Example III.

Example IV: The incident response is assumed to be a ramp function of slop U.

$$h(t,\tau;U) = U(t-\tau)$$

Again, U is Gaussian with variance 5 and mean (i) 40 (ii) 100 (iii) 200 (see Fig. 5).

To investigate the effect of increasing the number of values in the set $\{N(k\Delta)\}$ for maximization i.e. allowing delay detection, the case in Example IV was reconsidered and compared to the case when one more value of $N(k\Delta)$ in the detection was allowed. The results are shown in Fig. 6. The simulation agrees extremely closely with the approximate expression for the performance when λ is small. And as Fig. 6 indicates, the performance of the scheme improves considerably if we increase the set of values for $N(k\Delta)$ in the detections.

3.5 Conclusion

In this chapter, we have developed and examined a sequential estimation-detection scheme as an approximate solution to the state estimation problem. Its asymptotic performance has been derived and shown by digital simulations to be an extremely good indicator of the true performance when the intensity is small. Hence, under the situation of reasonable signal-to-noise ratio (SNR) in the observations, the scheme performs considerably better than in the MMSE sense, the linear optimal filters when λ is small. Indeed the asymptotic performance of the linear optimal filter is extremely disappointing, regardless of the SNR or how the incident responses behave.

Notice that the SSS performs reasonably well for both stable and unstable systems under normal signal-to-noise environments as indicated

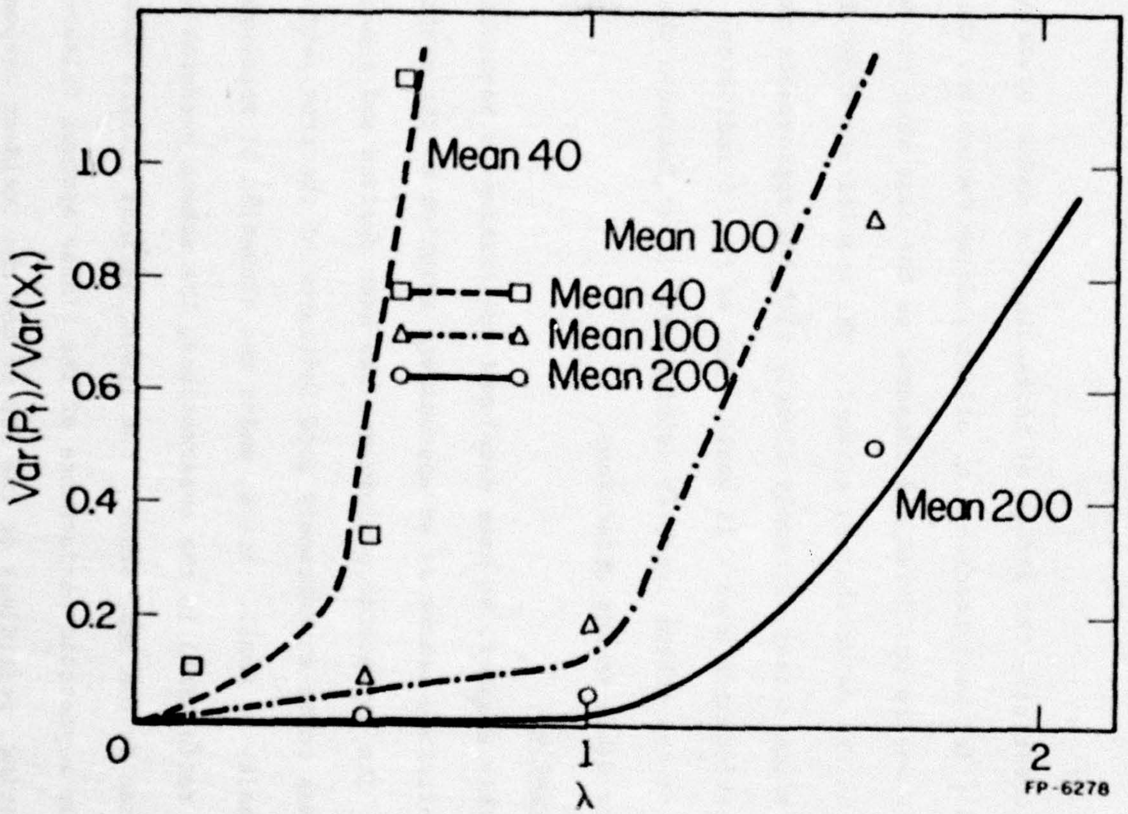

Figure 5. Simulated and computed performance, $Var(\epsilon_t)$ vs. λ, for Example IV.

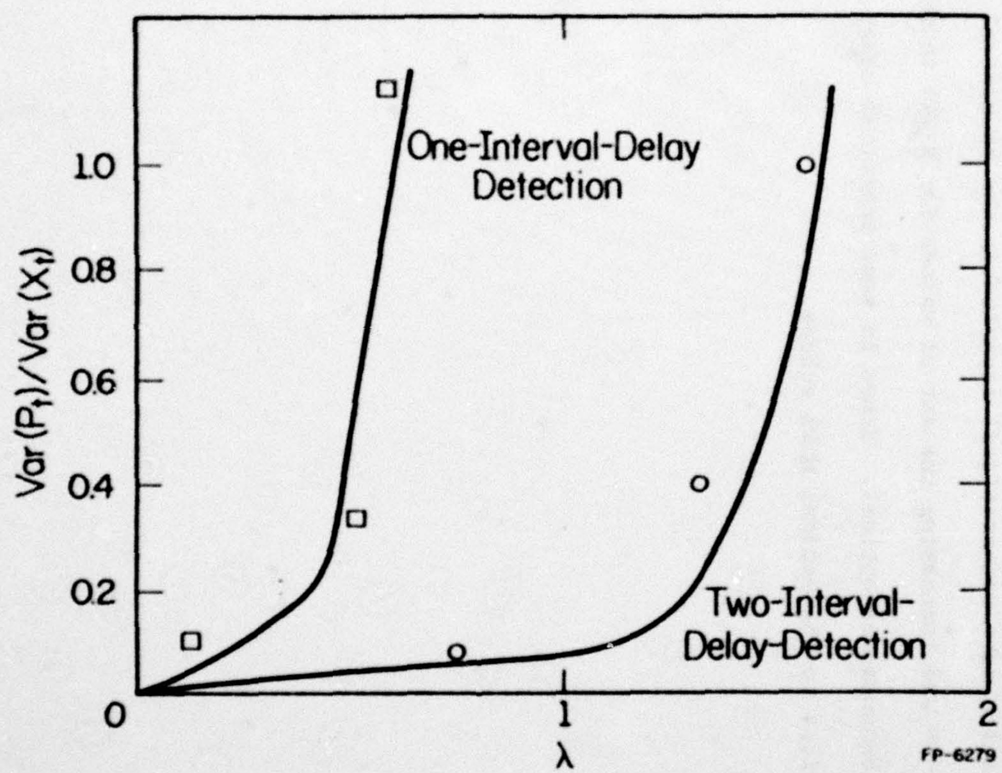

Figure 6. Simulated and computed performance, $Var(\epsilon_t)$ vs. λ, for Example IV with one additional value for $N(k\ell)$.

in all the examples (see Figs. 2-6). In fact is some cases, it works better for unstable systems than for the stable case as example I and II show. As indicated in example IV and Fig. 6, the performance of the SSS can be further improved by an additional delay in the reconstruction process.

As we allow more computational capacity in the sequential scheme, choosing $L \to \infty$ and increasing the set of values for $N(k\Delta)$ in the detection, the scheme becomes MAP optimal. Hence in many practical cases, it may indeed be close to the optimal MMSE scheme.

CHAPTER 4

STATE ESTIMATION OF THE WIENER-POISSON DRIVEN LINEAR SYSTEM

4.1 General

We now address ourselves to the general estimation problem stated in Chapter 1. As in Chapter 3, we allow an additional uncertainty in the system by introducing a mark process excited by the Poisson process as follows

$$dX_t = \mathcal{A}_t X_t dt + \mathcal{B}_t dW_t + b_t(U) d\eta_t \quad (4.1a)$$

where U is the mark process associated with the Poisson process η_t and is assumed to have known statistics, i.e. probability density $f_U(u)$. Again the objective is to develop an estimation scheme with delay (smoothing) for X_t from observations, viz.,

$$dy_t = h_t X_t dt + dv_t \quad (4.1b)$$

In this chapter, we develop a suboptimal sequential scheme (SSS) as the one in Chapter 3 for Poisson driven systems. The scheme is a modification of the one considered earlier, and consists of performing detection and estimation of the Poisson incidents in a sequence of small subintervals. The reconstruction of the system state X_t will be obtained from the estimates of the Poisson incident process followed by a smoothing filter. The estimation and detection of the Poisson incidents are performed using the MAP criterion as in Chapter 3, while the smoothing filter employs the MMSE criterion. Basically the only difference in the SSS for the process (4.1) from that in Chapter 3 is that in the reconstruction process we apply the MMSE smoothing to the Wiener driven part of the system whereas in

Chapter 3 the reconstruction is purely deterministic once the estimation and detection of the Poisson process is finalized.

Due to the sequential structure of the SSS, the errors arising from any deviations in the estimation and detection of η_t will tend to propagate and accumulate. A compensation approach will be included in the SSS to eliminate the error propagation effect. The numerical algorithms in the optimization procedures and the solution of the equations in the SSS are discussed and presented in an algorithmic form. The asymptotic performance for the time-invariant case for small intensity is derived analytically and simulated numerically.

4.2 Single Interval Suboptimal Smoothing

The system given by Eq. (4.1) is by assumption a measurable process of the Wiener input process and the Poisson input process together with the initial (random) conditions. *If the Poisson input is given, and the initial* conditions are Gaussian, the system process can be easily shown to be conditionally Gaussian [7]. This conditional Gaussian property enables us to obtain an optimal estimate in closed form if the Poisson process is known. Since the system is driven linearly and independently by the two input processes, that optimal estimate is in fact the optimal MMSE estimate of the Wiener driven part of the system, and in our case is just the Kalman filter.

Let the conditional filtered and smoothed estimates of the system X_t over the interval $[0,T)$ be denoted by m_t and $m_{t/T}$ respectively, i.e.

$$m_t \triangleq E(X_t | \mathcal{F}_t^y, \mathcal{F}_T^\eta) \qquad 0 \le t < T \qquad (4.2)$$

$$m_{t|T} \triangleq E(X_t | \mathcal{F}_T^y, \mathcal{F}_T^\eta) \qquad 0 \le t < T \qquad (4.3)$$

where in our notation, $\mathcal{F}_s$ denotes the σ-field generated by $\{\xi_t;\ 0 \leq t < s\}$. Note that m_t and $m_{t/T}$ are not the filtered and smoothed estimates of the system but rather, the pseudo-estimates based on the given knowledge of the Poisson input η_t. The Poisson input can be specified over [0,T) by giving the number of incidents over the interval and the incident times and marks.

To indicate the dependence on N[0,T) (N(T) for short) and $(\underline{\tau},\underline{u})$ of m_t and $m_{t/T}$, we use the explicit notation

$$m_t = m_t(N(T),\ \underline{\tau},\underline{u}) \qquad 0 \leq t < T \qquad (4.4)$$

$$m_{t/T} = m_{t/T}(N(T),\ \underline{\tau},\underline{u})\ . \qquad 0 \leq t < T \qquad (4.5)$$

With these notations, a likelihood function of m_t observed in white Gaussian noise, denoted by $LR[m_t]$, can be constructed using the concept of estimator-correlator receiver [10,11] et al.

$$LR[m_t(N(T),\tau,u)] = \exp[\int^T m_t \dot{y}_t dt - \frac{1}{2}\int^T m_t^2 dt\} \qquad (4.6)$$

Hence for the MAP estimate of N(T), u and $\underline{\tau}$, we have

$$J[N,\underline{\tau},\underline{u}] = \ell n\ LR[m_t(N,\underline{\tau},\underline{u}] + \ell n f_{N(T),\underline{\tau},\underline{u}}(N,\underline{\tau},\underline{u}) \qquad (4.7)$$

where $f_{N(T)\underline{\tau},\underline{u}}(\cdot,\cdot,\cdot)$ is the apriori probability of N(T) and $\underline{\tau},\underline{u}$.

Therefore the MAP estimate of N(T) and $\underline{\tau},\underline{u}$ denoted by $\hat{N}$ and $\hat{\underline{\tau}},\hat{\underline{u}}$ satisfy the following,

$$J[\hat{N},\hat{\underline{\tau}},\hat{\underline{u}}] = \max_{N^*,\underline{\tau}^*,\underline{u}^*} J[N^*,\underline{\tau}^*,\underline{u}^*] \qquad (4.8)$$

Notice that the dimension of $\underline{\tau},\underline{u}$ is N. The maximization can be done by assuming the value of N and maximizing with respect to $\underline{\tau},\underline{u}$ first. Therefore,

$$J[\hat{N},\hat{\underline{\tau}},\hat{\underline{u}}] = \max_{N^*} \max_{\underline{\tau}^*,\underline{u}^*} \{J[N^*,\underline{\tau}^*,\underline{u}^*]\} \tag{4.9a}$$

$$= \max_{N^*} J[N^*, \hat{\tau}(N^*), \hat{u}(N^*)] \tag{4.9b}$$

Although N* can take up any positive integer value, large values of N* can be neglected in the maximization by an appropriate choice of the length of the interval T. The aposteriori probability of N(T) is uniformly continuous with respect to its apriori probability, which can be made arbitrarily small with the choice of the time interval. Hence the maximization can be done over a set of finite elements, N* = 0,1.2,...K neglecting values higher than K.

After the MAP estimates of N and $\underline{\tau},\underline{u}$ are determined, the estimates of the system can be readily determined in a suboptimal fashion by substituting the estimates of τ,u,N in the expressions for m_t and $m_{t/T}$ (4.4) and (4.5) (see Fig. 7).

From Eq. (4.1), we have

$$X_t = \Phi(t,0)X_0 + \int_0^t \Phi(t,s)B_s dW_s + \int_0^t \Phi(t,s)b_s(U)d\eta_s \tag{4.10}$$

If $\{\eta_s(U),\ 0 \le s \le t\}$ is known, the system is, in effect, a Wiener driven process. The last term of the equation just gives the system a different mean function. Hence we can write m_t at follows,

$$m_t = \hat{X}_t^c + \int^t \Phi(t,s)b_s(U)\ d\eta_s \tag{4.11}$$

where $\hat{X}_t^c$ is the estimate of the Wiener driven part, and is given by its Kalman filter

$$\dot{\hat{X}}_t^c = \mathcal{A}_t\hat{X}_t^c + P_t^c h_t v_t^{-1}[\dot{y}_t - h_t \int_0^t \Phi(t,s)b_s(U)d\eta_s - h_t\hat{X}_t^c] \quad t \geq 0 \tag{4.12a}$$

$$\dot{P}_t^c = 2\mathcal{A}_t P_t^c + \mathcal{B}_t^2 Q_t + h_t^2 v_t^{-1} P_t^{c^2} \quad t \geq 0 \tag{4.12b}$$

with $\hat{X}_0^c = \hat{X}_0$, $\quad P_0^c = P_0$

After some algebraic manipulations, we have equivalently,

$$m_t = X_t^0 + X_t^1 + \int_0^t \Phi(t,s)b_s(U)d\eta_s \tag{4.13}$$

and

$$\dot{\hat{X}}_t^0 = \mathcal{A}_t\hat{X}_t^0 + P_t^c h_t v_t^{-1}[\dot{y}_y - h_t\hat{X}_t^0] \tag{4.14}$$

$$\hat{X}_t^1 = \int_0^t \Phi(t,s)b_s(U)\ [\int_s^t \varphi(t,u)h_u^2 v_u^{-1} P_u^c du]\ d\eta_s \tag{4.15a}$$

$$\triangleq \sum_{i=1}^{N(t)} g_t^1(\tau_i;u_i) \tag{4.15b}$$

where $\varphi(t,u)$ is the state transition function of the homogeneous part of the filter in Eq. (4.12) and $g_t^1(\tau_i;u_i)\ \forall\ i$ satisfy the causality condition by definition of the integral (4.15), i.e.

$$g_t^1(\tau_i;u_i) = 0 \qquad \forall\ \ t < \tau_i \qquad \forall i \tag{4.16}$$

Similarly, we have

$$\int_0^t \Phi(t,s)b_s(U)d\eta_s \triangleq \sum_{i=1}^{N(t)} g_t^0(\tau_i;u_i) \tag{4.17}$$

Note that in Eq. (4.13), $\hat{X}_t^0$ is the only term on the right side of the equation that does not depend on the realization of η_t; it is the only term that depends on the observation $\dot{y}_t$ in fact. The remaining two terms depend only on the realization of η_t and not $\dot{y}_t$, they are completely determined by the knowledge of N(t), $\underline{\tau}$,$\underline{u}$.

60

The smoothed pseudo-estimate $m_{t/T}$ can be readily constructed from $\hat{X}_t^c$ and is given as follows [7,9,12] (see also Fig. 7)

$$m_{t/T} = \hat{X}^c_{t/T} + \sum_{i=1}^{N(T)} g_t^0(\tau_i, u_i) \tag{4.18}$$

where

$$\dot{\hat{X}}^c_{t/T} = \mathcal{A}_t \hat{X}_{t/T} + \mathcal{B}_t^2 Q_t (P_t^c)^{-1} [\hat{X}^c_{t/T} - \hat{X}^c_t] \tag{4.19a}$$

$$\dot{P}^c_{t/T} = 2[\mathcal{A}_t + \mathcal{B}_t^2 Q_t (P_t^c)^{-1}]\ P^c_{t/T} - \mathcal{B}_t^2 Q_t \tag{4.19b}$$

with $\quad \hat{X}^c_{T/T} = \hat{X}^c_T, \qquad P^c_{T/T} = P^c_T \quad$ (4.19c)

The substitution of Eqs. (4.13-4.17) into Eq. (4.7) yields an equivalent expression for the MAP maximization which can be expressed as

$$\begin{aligned} J_T[N, \underline{\tau}, \underline{u}] &= \int_0^T dt [\hat{X}_t^1 + \sum_{i=1}^N g_t^0(\tau_i, u_i)][-\hat{X}_t^0 - \dot{y}_t] \\ &\quad - \frac{1}{2}\int_0^T dt [X_t^1 + \sum_{i=1}^N g_t^0(\tau_i, u_i)]^2 + N\ell n\lambda \\ &= \int_0^T dt \left\{ \sum_{i=1}^N [g_t^t(\tau_i, u_i) + g_t^0(\tau_i, u_i)] \right\} \\ &\quad \cdot \left\{ \dot{y}_t - \hat{X}_t^0 - \frac{1}{2} \sum_{i=1}^N [g_t^1(\tau_i, u_i) + g_t^0(\tau_i, u_i)] \right\} + N\ell n\lambda \end{aligned} \tag{4.20}$$

This expression can be further simplified into

$$J_T(N, \underline{\tau}, \underline{u}) = \int_0^T dt \left\{ [\sum_{i=1}^N g_t(\tau_i, u_i)][z_t - \frac{1}{2}\sum_{i=1}^N g_t(\tau_i, u_i)] \right\} + N\ell n\lambda \tag{4.21}$$

where

$$g_t(\tau, u) \triangleq g_t^0(\tau, u) + g_t^1(\tau, u) \tag{4.22}$$

$$z_t \triangleq \dot{y}_t + \hat{X}_t^0 \tag{4.23}$$

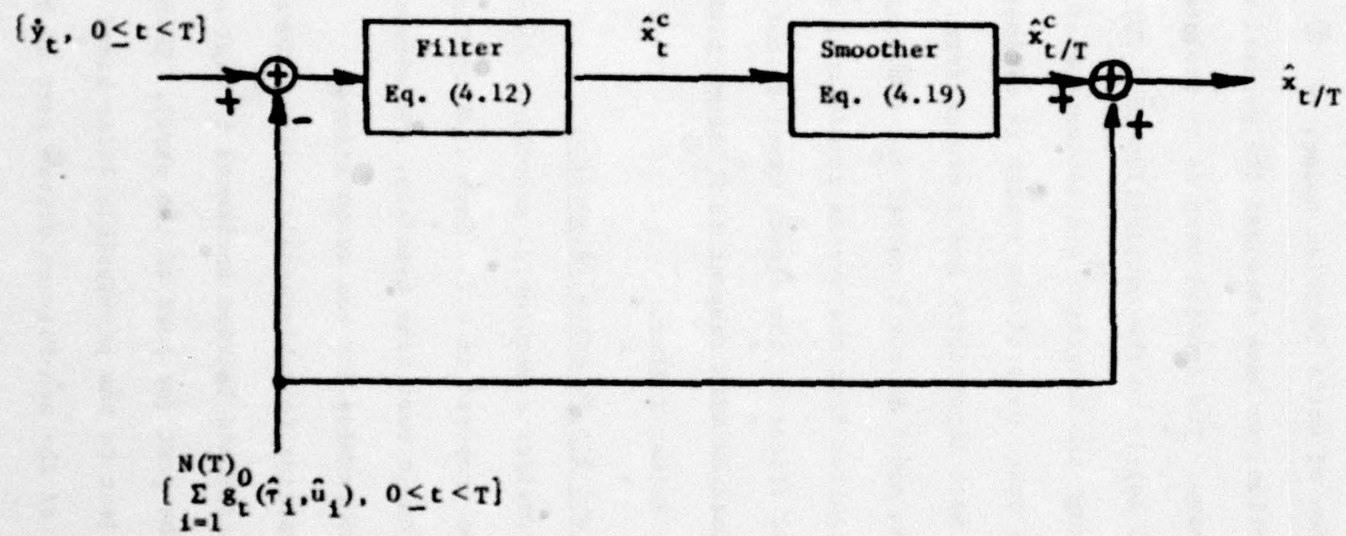

Figure 7. Schematic representation of the reconstruction of the system state after the Poisson incidents and marks are estimated.

Note that J = 0 when N = 0. Also (4.21) is in the conventional form of the likelihood function where z_t is the observation of the signal $\sum_{i=1}^{N} g_t(\tau_i, u_i)$ in the presence of white Gaussian noise.

To summarize this section, we have obtained the general equations for the proposed smoothing scheme. The crucial step is to isolate the parts of the system which depend solely on the realization of $\Pi_t(U)$. They are finally reconstructed through the detection and estimation of the incidents of the process $\Pi_t(U)$. The other part of the system is estimated using a MMSE optimal filter which most importantly has a closed form. Notice that the linearity of the system model is not required for obtaining the MMSE closed filter. It only requires that the system conditional on the σ-field $\mathcal{F}_t^{\Pi}$ gives rise to an optimal filter of the closed type; in our case the system is conditionally Gaussian with respect to $\mathcal{F}_t^{\Pi}$ hence gives rise to a closed optimal filter, the Kalman filter.

4.3 Sequential Smoothing and Its Numerical Algorithm

In this section, we consider a sequential smoothing scheme approximately implement the scheme proposed in 4.2. Such a sequential scheme will make the implementation in real time feasible, and offers the advantages of sequential iterative programming for the computations.

In the scheme, observations in subintervals, each of duration Δ are processed sequentially to estimate Poisson incidents in that new subinterval and also update those in the past (or part of the past). Then the system state is reconstructed subject to the permissible delay based on these estimates and the estimate of the non-Poisson driven part of the system.

The approach for obtaining the estimates of the incidents is similar to that in the previous chapter. Basically, upon receiving a new sub-interval of observations, the problem of a single interval smoothing as in the last section has to be solved. Although there are some nice properties arising from the sequential structure of the scheme the major component of the scheme lies in the maximization of the expression in Eq. (4.21).

Note that the expression to be maximized does not have a second derivative with respect to $\underline{\tau},\underline{u}$, since it involves the derivatives of z_t which contains white-noise components. Hence, second order iterative numerical techniques for optimization do not apply.

Another point to note in the maximization is that the number of incidents, i.e. the dimension of $\underline{\tau},\underline{u}$ increases as the length of smoothing interval increases by Δ each time. Hence a robust approximation approach which requires up-dating only part of the previous estimates is considered.

In order to develop the algorithm, we shall assume that we have the estimates of the number of incidents and their times and marks to time $t = n\Delta$, denoted by $N(n\Delta)$, $\underline{\tau}(n\Delta)$, $\underline{u}(n\Delta)$ respectively. Then we add as new subinterval of observations of length Δ, where Δ is to be so chosen, as to satisfy

$$\lambda\Delta = \alpha << 1$$

i.e., the probability of having more than one incident in the new sub-interval is negligibly small. Hence the number of incidents up to $t = n\Delta$ can take on two values, namely

$$N(n\Delta) \quad \text{and} \quad N(n\Delta) + 1 \tag{4.24}$$

Therefore expression (4.21) only needs to be maximized for two sets of $\underline{\tau}((n+1)\Delta)$, which can be done independently. Note that the entrees of the vectors $\underline{T}$,$\underline{u}$ have been all or all-but-one estimated previously using the data up to $t = n\Delta$. Although they have to be updated since more data has been considered, their estimates can serve as an approximate point in an iterative algorithm for the new estimates. There are quite a few iterative techniques that are applicable to the maximization. Among the available, we choose an approach that does not require taking any derivatives and converges at a relatively fast speed.

4.3.1 Numerical Iterative Algorithm

The iterative algorithm consists of first determining for any given iterated point a direction along which the maximization of J can be done, then the scalar maximization is done by some efficient search procedures. More precisely, if we let $\underline{\tilde{\tau}}^k$ be the k-th approximation point to the optimal $\underline{\tau}^*$ which lies in some N-dimensional Euclidean space E^N, the new approximation point $\underline{\tilde{\tau}}^{k+1}$ will be given as

$$\underline{\tilde{\tau}}^{k+1} = \underline{\tilde{\tau}}^k + \tilde{\alpha}\, \underline{d} \qquad \underline{\tilde{\tau}}^{k+1}, \underline{d} \in E^N \tag{4.25}$$

where $\underline{d}$ is the chosen direction for maximization of J and $\tilde{\alpha}$ is the maximization constant that gives the maximum J subject to the constraint that the ordered incident times must satisfy

$$0 \le \tilde{\tau}_{i-1}^{k+1} < \tilde{\tau}_i^{k+1} < n\Delta \qquad i = 2,3\text{-}N \tag{4.26}$$

Actually constraint (4.26) gives the range for $\tilde{\alpha}$ in the optimization, since

$$\tilde{\tau}_{i-1}^{k+1} < \tilde{\tau}_i^{k+1} \qquad i = 2\text{---}N$$

we have

$$\tilde{\tau}^k_{i-1} + \tilde{\alpha} d_{i-1} < \tilde{\tau}^k_i + \tilde{\alpha} d_i$$

which results in

$$\tilde{\alpha} = \begin{cases} \dfrac{\tau_i - \tau_{i-1}}{d_i - d_{i-1}}, & d_{i-1} > d_i \\[2ex] \dfrac{\tau_i - \tau_{i-1}}{d_i - d_{i-1}}, & d_{i-1} < d_i \end{cases} \tag{4.27}$$

$i = 1, 2, \ldots, N+1$.

We take $\tilde{\tau}_0^{k+1} = 0$, $d_0 = d_{N+1} = 0$ and $\tilde{\tau}^{k+1}_{N+1} = n\Delta$, and without loss of generality we assume $d_1 > 0$, we have

$$-\frac{\tau_1^{k+1}}{d_1} < \max_{\substack{i \\ d_{i-1} < d_i}} \left(\frac{\tau_i - \tau_{i-1}}{d_{i-1} - d_i} \right) < \tilde{\alpha} < \min_{\substack{i \\ d_{i-1} > d_i}} \left(\frac{\tau_i - \tau_{i-1}}{d_{i-1} - d_i} \right) < \frac{n\Delta - \tau_1^{k+1}}{d_1} \tag{4.28}$$

which is the feasible range of $\tilde{\alpha}$ for the maximization search procedure.

Mathematically, if the iterative maximization algorithm is denoted by A, it can be represented as

$$A = M \bullet D \tag{4.29}$$

where $D = E^N \rightarrow E^{2N}$ is a point-to-set mapping that determines a direction for maximization for the objective function J; and $M = E^{2N} \rightarrow E^N$ is a point-to-set mapping that gives a better approximation point for the maximization given the direction and the iterative point.

We choose the Golden-Section Search for M and the Conjugate Directions Technique for D. The convergence of algorithm A solely depends on the closedness of M and D. For our choice of M and D, it can be shown easily both closed and hence convergent procedures [13,14].

Due to the unimodal property of J in the mean sense, the Golden Section Search procedure is a very efficient, convergent search technique over the feasible range of $\tilde{\alpha}$. It does not require any derivatives of the objective function J, avoiding the analytic difficulty of taking derivatives of white noises.

The conjugate directions usually defined with respect to a quadratic function are a set of directions such that maximizing in each of the directions will give the maximum of the quadratic function. In E^N, there are N such directions in the set. The generation of such directions for a given quadratic form does not require taking derivatives and for maximization, the order of the directions maximized is immaterial. Hence the algorithm converges within N steps in case of maximizing a quadratic function.

Note that our objective function J to be maximized is in general non-quadratic. Generation of such a set of conjugate directions is impossible. Nevertheless, we exploit a quadratic approximation to J without explicitly calculating the Hessian matrix, which in our case is not well defined due to the white processes involved. We iteratively generate a set of conjugate directions which gives a better optimization to J. With several iterations, the requisite information about the Hessian is developed. The overall procedure can be thought of as generating conjugate directions to a quadratic that approximates J. Since the algorithm for the procedure is closed [13], the overall scheme is convergent.

For the details of the complete maximization scheme see Appendix C.

For our choice of Δ, the length of the subinterval, at most there is only one incident time that has never been estimated, the rest are all close to the optimal value maximizing J. The numerical procedures should converge extremely fast within a few iterations.

4.3.2 Sequential Approximations

The computational complexity and hence the efficiency and rate of convergence in the algorithm depend heavily on the number of incident times to be estimated. To overcome the computational difficulties, the following approximations can be made to enhance the speed of the overall scheme:

(1) Updating or re-estimating the recent incident times only. In the sequential scheme, upon receiving the n^{th} subinterval of observations, the smoothing is performed such that the estimates of the incidents prior to the (n-k+1)-th subinterval are all finalized, i.e. they are not re-estimated. More precisely, the expression for J Eq. (4.21) is approximated in the following manner,

$$J_{n\Delta}[N,\underline{\tau},\underline{u}] = \int_0^{n\Delta} dt\{\sum_{i=1}^{N} g_t(\tau_i,u_i)\}\{z_t - \frac{1}{2}\sum_{i=1}^{N} g_t(\tau_i,u_i)\} + N\ell n\lambda$$

$$= \int_0^{n\Delta} dt\{\sum_{i=1}^{N((n-K)\Delta)} g_t(\tau_i,u_i) + \sum_{i=N((n-K)\Delta)+1}^{N} g_t(\tau_i,u_i)\}$$

$$\cdot\{z_t - \frac{1}{2}\sum_{i=1}^{N((n-K)\Delta)} g_t(\tau_i,u_i) - \frac{1}{2}\sum_{i=N((n-K)\Delta)+1}^{N} g_t(\tau_i,u_i)\}$$

$$+ [N((n-K)\Delta) + N - N((n-K)\Delta)]\,\ell n\lambda \qquad (4.30)$$

If τ_i, u_i for $i = 1,2---N((n-K)\Delta)$ are all finalized and not estimated again, than an approximate expression $\tilde{J}$ can be formed

$$\tilde{J}_{n\Delta}[\tilde{N},\tilde{\underline{\tau}},\tilde{\underline{u}}] = \int_{(n-K)\Delta}^{n\Delta} dt\{\sum_{i=1}^{\tilde{N}} g_t(\tilde{\tau}_i,\tilde{u}_i)\}\{\tilde{z}_t - \frac{1}{2}\sum_{i=1}^{\tilde{N}} g_t(\tilde{\tau}_i,\tilde{u}_i)\} + \tilde{N}\ell n\lambda \qquad (4.31)$$

where $\tilde{N} = N(n\Delta) - N((n-K)\Delta)$ (4.31a)

$\tilde{\tau}_i = \tilde{\tau}_{i+N((n-K)\Delta)}$ (4.31b)

$\tilde{u}_i = \tilde{u}_{i+N((n-K)\Delta)}$ (4.31c)

and $\tilde{z}_t = z_t - \frac{1}{2} \sum_{i=1}^{N((n-K)\Delta)} g_t(\tau_i, u_i)$ (4.31d)

Notice that expression $\tilde{J}$ in (4.31) is in the same form as J in (4.21). The difference is that the number of unknowns $\tilde{\tau}_i$'s, $\tilde{u}_i$'s now is less than τ_i's, u_i's by exactly $2N((n-K)\Delta)$. Indeed the expected number of unknowns in $\tilde{J}$ is $2K(\lambda\Delta)$ which is small since $(\lambda\Delta)$ is small.

(II) Shortening the interval of integral in J Eq. (4.21) by increasing the lower limit of the integral, placing all the weight of the observation in the shortened interval for the estimation. If the interval for the integration in the expression of J is $[0,n\Delta)$, it can be shortened to $[M\Delta,n\Delta)$ where M is a parameter to be determined. Combining approximation I and II, we can write the expression for $\tilde{J}$ as

$$\tilde{J}_{n\Delta}[\tilde{N},\underline{\tilde{\tau}},\underline{\tilde{u}}] = \int_{(n-K+M)\Delta}^{n\Delta} dt\{\sum_{i=1}^{\tilde{N}} g_t(\tilde{\tau}_i,\tilde{u}_i)\}\{\tilde{z}_t - \frac{1}{2}\sum_{i=1}^{N} g_t(\tilde{\tau}_i,\tilde{u}_i)\} \quad (4.32)$$

Note that in (4.31), all the $\tilde{\tau}_i$'s are within the interval of the integral, and in (4.32) some of them are not but their incident responses are.

While both approximation I and II reduce the computational complexity and enhance the speed of the process, they are different in nature. The application of one or both and subsequently the choice of K and M depend on the tradeoff of various aspects as availability of computational facilities, numerical operations, speed and desired performance of the overall scheme.

Basically, for stable systems especially those of large negative eigenvalues, Approximation I is a robust choice since new observations are unlikely to improve very much the estimates of thes incidents occurring in the remote past. For unstable systems, Approximation II nevertheless is a good alternative in reducing the computational burden. In that case the incidents in the past, are still being re-estimated, but based primarily on their effect on the recent observations. Such a weighting of latter observations is consistent with the unstable nature of the incident responses.

4.4 Asymptotic Performance of the SSS

Since the general performance of the SSS is difficult to derive, we consider only the time-invariant case and derive the asymptotic performance of the scheme as the intensity tends to zero.

From Eq. (4.19) which gives the estimates of the system, the estimation error $\mathcal{E}_t$ resulting from the overall scheme can be expressed as a sum of two errors: one from the estimates of the Poisson driven part and one from the smoothing of the conditional Wiener driven part of the system. The smoother $\hat{X}_{t/T}$ given in Eq. (4.12) is derived conditionally on the assumed knowledge of $\eta_t(U)$, i,e, $\mathcal{F}_t^{\eta}$; due to the linearity of the smoother, its error can be expressed as a sum of an η_t-free term and an η_t-dependent term. Therefore we have

$$\mathcal{E}_t = \epsilon_t^0 + \epsilon_t^n + e_t^n \qquad 0 \leq t < T \qquad (4.33)$$

where ϵ_t^0 and ϵ_t^n denote the η_t-free and η_t-dependent term of the smoothing error of the Wiener driven part of the system; e_t^n denotes the error of the estimate of the Poisson driven part. Since the conditional estimate $\hat{X}_{t/T}^c$

70

with respect to $\mathcal{F}_t^{\eta} \times \mathcal{F}_t^{y}$, it can be shown easily from Eq. (4.12) that

$$\epsilon_t^0 = P_{t/T}^c \tag{4.34}$$

which is the variance of the conditional smoother given in Eq. (4.18). In fact it is that variance one would obtain in smoothing the Wiener driven part of the system using an MMSE optimal smoother. Note that the estimate $\hat{X}_{t/T}$ is conditional on $\mathcal{F}_t^{\eta}$, but the variance $P_{t/T}^c$ is independent of η_t.

When the intensity λ is small, the estimation error resulting from the scheme can be modeled as a Poisson filtered process [1,50], with each incident response as a result of the detection and estimation over each subinterval of length Δ. Hence the Poisson filtered process driving the error process of the scheme has an intensity of $1/\Delta$.

Combining Eqs. (4.33) and (4.34), we have

$$\mathcal{E}_t = P_{t/T}^c + \int_V \int_0^t [\epsilon(t;\sigma,V) + e(t;\sigma,V)]\mathcal{M}(d\sigma,dV) \qquad t < T \tag{4.35}$$

Here $\epsilon(t,\sigma;V)$ and $e(t,\sigma;V)$ are the average error response made according to the result of the detection indicated by the mark V, and $\mathcal{M}$ is the measure of the underlying error driving process with intensity $1/\Delta$.

Define the mark V in a similar fashion as in Chapter 3,

$V = d_{10}$ indicating a miss in the detection

$V = d_{11}$ indicating a correct detection of an incident

$V = d_{01}$ indicating a false alarm

$V = d_{00}$ indicating a correct detection of no incident

The distribution of the mark is given by

$$
\begin{aligned}
P(V = d_{10}) &= \alpha p_{10}(\Delta) \\
P(V = d_{11}) &= \alpha(1-p_{10}(\Delta)) \\
P(V = d_{01}) &= (1-\alpha) p_{01}(\Delta) \\
P(V = d_{00}) &= (1-\alpha)(1-p_{01}(\Delta))
\end{aligned}
\tag{4.36}
$$

where $p_{01}(\Delta)$ and $p_{10}(\Delta)$ denote the probability of a false alarm and a miss in the detection. Hence the variance of $\mathcal{E}_t$, denoted by $\mathrm{Var}(\mathcal{E}_t)$ is given as follows

$$
\begin{aligned}
\mathrm{Var}(\mathcal{E}_t) &= P^c_{t/T} + \frac{1}{\Delta} \underset{V}{E} \int_0^T [\mathcal{E}(t;0,V) + e(t;0,V)]^2 dt \\
&= P^c_{t/T} + \frac{\lambda}{\alpha} \underset{V}{E} \int_0^T [\mathcal{E}(t;0,V) + e(t;0,V)]^2 \, dt \quad \text{for } t < T
\end{aligned}
\tag{4.37}
$$

The substitution of Eq. (4.36) into (4.37) yields

$$
\begin{aligned}
\mathrm{Var}(\mathcal{E}_t) - P^c_{t/T} = \frac{\lambda}{\alpha} \{ & \alpha p_{10}(\Delta) \int_0^T [\mathcal{E}(t;0,d_{10}) + e(t;,0,d_{10})]^2 \, dt \\
& + \alpha(1-p_{10}(\Delta)) \int_0^T [\mathcal{E}(t;0,d_{11}) + e(t;0,d_{11})]^2 dt \\
& + (1-\alpha) p_{01}(\Delta) \int_0^T [\mathcal{E}(t;0,d_{01}) + e(t;0,d_{01})]^2 dt
\end{aligned}
\tag{4.38}
$$

Note that we have only three terms in Eq. (4.38) because in case of a correct detection of no incident ($V = d_{00}$), there is no η_t-dependent error, i.e.

$$
\mathcal{E}(t;0,d_{00}) = e(t;0,d_{00}) \equiv 0 \tag{4.39}
$$

Assuming that the average error made by a correct detection is relatively small, since $\alpha \ll 1$,

$$\mathrm{Var}(\mathcal{E}_t) - P^c_{t/T} \cong \lambda\left(\frac{P_{01}(\Delta)}{\alpha}\right) \int_0^T [\mathcal{E}(t,0;d_{01}) + e(t,0,d_{01})]^2 dt \quad \text{for } t < T \tag{4.40}$$

For the time-invariant case i.e. the system function and observations are not time dependent, from the definition of $\mathcal{E}$ and e, we deduce easily from Eqs. (4.15), (4.16) and (4.17) that

$$e(t;0,d_{01}) = \underset{U}{E}\, g^0_t(0,U) = ube^{at} \tag{4.41a}$$

$$\mathcal{E}(t;0,d_{01}) = \underset{U}{E}\{B^2Q(P^c)^{-1} \int_0^t ds \cdot \exp\{[a + B^2Q(P^c)^{-1}][t-s]\} \cdot g^1_s(0,U)\}$$

$$= \underset{U}{E}\,[- B^2Q(P^c)^{-1}] \int_0^t du \, \exp\{[a + B^2Q(P^c)^{-1}][t-u]\}$$

$$\cdot \int_0^u ds \, \exp\{(a-P^ch^2v^{-1})(u-s)\} g^0_s(0,U) \tag{4.41b}$$

$$= (\bar{u}be^{at})\left[\frac{(P^c)^{-1}}{h^2v^{-1}}\right]\left\{1 - \frac{B^2Q(P^c)^{-1}e^{-P^ch^2v^{-1}t} + P^ch^2v^{-1}e^{+B^2QP^{-1}t}}{(B^2QP^{c^{-1}} + P^ch^2v^{-1})}\right\} \tag{4.42}$$

where $\bar{u}$ denotes the mean of U. For comparison purposes, let us normalize both sides of Eq. (4.40) with the variance of the Poisson driven part of the system which is given as [1]

$$\mathrm{Var}(X^n_t) = \lambda \int_0^T [g^0_t(0)]^2 dt$$

$$= \lambda \int_0^T [e(t;0,d_{01})]^2 dt \tag{4.43}$$

Therefore from Eqs. (4.40) and (4.43), the error variance in estimating the Poisson driven part of the system normalized to the variance of that part is

$$\mathrm{Var}(\mathcal{E}_t^n) \triangleq \frac{\mathrm{Var}(\mathcal{E}_t) - P^c_{t/T}}{\mathrm{Var}(X_t^n)}$$

$$= [\frac{P_{01}(\Delta)}{\alpha}] \frac{\int_0^T [\mathcal{E}(t;0,d_{01}) + e(t;0,d_{01})]^2 dt}{\int_0^T [\mathcal{E}(t;0,d_{01})]^2 dt} \tag{4.44}$$

As λ tends to zero, $\Delta \to \infty$ and $T \to \infty$ therefore

$$\lim_{\lambda \to 0} \mathrm{Var}(\mathcal{E}_t^n) = \lim_{T,\Delta \to \infty} \frac{P_{01}(\Delta)}{\alpha} \frac{\int_0^T [\mathcal{E}(t;0,d_{01}) + e(t;0,d_{01})]^2 dt}{\int_0^T [\mathcal{E}(t;0,d_{01})]^2 dt}$$

$$= \lim_{T \to \infty} \frac{P_{01}(T)}{\alpha} \frac{\int_0^T (\mathcal{E}_t + e_t)^2 dt}{\int_0^T \mathcal{E}_t^2 \, dt} \tag{4.45}$$

The probability of wrong detection in an interval of length $T, P_{01}(T)$, is given by (see [10,12])

$$P_{01}(T) = \mathrm{erfc}[\frac{\ln(1/\alpha)}{\sqrt{E_T}} + \frac{\sqrt{E_T}}{2}] \tag{4.46}$$

where $\sqrt{E_T}$ is the normalized energy of the signal for detection, and is easily bounded for our sequential scheme by

$$E_T \geq v^{-1} \int_0^T [g_t^0(0) + g_t^1(0)]^2 dt$$

$$= v^{-1} \int_0^T (\mathcal{E}_t + g_t^1)^2 dt \tag{4.47}$$

Therefore, the asymptotic performance of our scheme in terms of the normalized error variance is, from Eqs. (4.34) to (4.47)

$$\lim_{\lambda \to 0} \mathrm{Var}(\mathcal{E}_t^n) = \frac{1}{\alpha} \lim_{T \to \infty} \mathrm{erfc}\left[\frac{\ln(1/\alpha)}{\sqrt{E_T}} + \frac{\sqrt{E_T}}{2}\right] \frac{\int_0^T (\mathcal{E}_t^2 + e_t)^2 dt}{\int_0^T \mathcal{E}_t^2 \, dt} \tag{4.48}$$

Note that in arriving at Eq. (4.48), we have invoked no assumption on the stability of the system. For stable system where the system matrix has all the poles on the left-half plane, or in the scalar case $a < 0$, all the quantities on the right side of Eq. (4.48) can be shown to be bounded and hence the asymptotic limit as $T \to \infty$ always exist. Since if $a < 0$,

$$\int_0^\infty \mathcal{E}_t^2 \, dt < \infty \; ; \quad \mathrm{erfc}\left[\frac{\ln(1/\alpha)}{\sqrt{E_\infty}} + \frac{\sqrt{E_\infty}}{2}\right] < \infty \; ; \qquad \sqrt{E_\infty} < \infty \tag{4.49}$$

Now, at first glance the term

$$\int_0^\infty (\mathcal{E}_t + e_t)^2 dt$$

appears to be possibly unbounded, since e_t (Eq. (4.42) contains an exponential term $\exp[(a + B^2QP^{c-1})t]$, which can still be unstable, i.e. $a + B^2QP^{c-1} > 0$. However, this is the eigenvalue of the non-causal smoother through which e_t is generated. Here, even for the case $a + B^2QP^{c-1} > 0$, e_t is still always bounded due to the end-point condition, $\hat{x}^c_{T/T} = \hat{x}^c_T$, of the smoother. For unstable system, $a > 0$, and the limit of the error variance becomes

$$\lim_{\lambda \to 0} \mathrm{Var}(\mathcal{E}_t^n) = \frac{1}{\alpha} \lim_{T \to \infty} \mathrm{erfc}\left[\frac{\ln(1/\alpha)}{\sqrt{E_T}} + \frac{\sqrt{E_T}}{2}\right] \frac{\int_0^T (\mathcal{E}_t + e_t)^2 dt}{\int_0^T \mathcal{E}_t^2 \, dt} \tag{4.50}$$

Using an exponential bound for erfc(.) and substituting Eq. (4.47) into (4.50), we have

$$\lim_{\lambda \to 0} \mathrm{Var}(\mathcal{E}_t^n) \le \frac{1}{\alpha} \lim_{T\to\infty} \exp\left(-\frac{v^{-1}}{8}\int_0^T (\mathcal{E}_t + g_t^1)^2 dt\right) \frac{\int_0^T (\mathcal{E}_t + e_t)^2 dt}{\int_0^T \mathcal{E}_t^2 dt}$$

Application of the L'apital's rule for the limit consecutively yields

$$\lim_{\lambda \to 0} \mathrm{Var}(\mathcal{E}_t^n) \le \frac{16av}{\alpha} \lim_{T\to\infty} \frac{\left(1 + \frac{e_T}{\mathcal{E}_T}\right)^2}{\mathcal{E}_T^2 \exp\left(\frac{v^{-1}}{8}\int_0^T (\mathcal{E}_t + g_t^1)^2 dt\right)} = 0 \qquad (4.51)$$

The result in Eq. (4.51) is not too surprising because in detecting a signal with increasing power over a semi-infinite interval, the signal-to-noise ratio goes to infinity rendering correct detection and subsequently precise estimation with probability one.

Hence in our sequential approximation scheme, for unstable systems, it is better to keep estimating all the previous incidents. It should be noted that in Approximation II, the error resulting from the finalized estimates of the incidents will propagate thus degrading the performance of the scheme.

An equalizing approach to eliminate the growing error propagation is developed in the next section. With this additional procedure, the instability of the system in many cases while not giving rise to a unstable error propagation from the finalized estimates can in effect improve the estimation and detection performance if the system state contains sufficiently large energy in the estimation and detection interval.

4.5 The Error Propagation and Compensation Approach

In this section, we examine the effect of the error due to the deviation of the finalized estimates of the incidents on subsequent estimations and detections in the sequential scheme. Due to the sequential structure of the scheme, the error arising from the finalized estimates will exhibit a propagation effect on latter estimates and detections, degrading the overall performance of the scheme. This is especially the case when the approximations (I) and (II) of Section 4.2 are adopted.

For an unstable system, the errors individually grow with time and tend to accumulate and propagate through the subsequent intervals. For stable systems, a single error tends to fade out with time. But the propagation effect due to it is quite nominal to continue for a relatively longer duration because of its generating phenomenon. If the time for the resultant propagation effect to fade is long enough, the errors will accumulate and reach a perpetual propagation effect.

Although our approximate expression for the asymptotic performance for $\lambda \to 0^+$ derived in section 4.4 is quite acceptable, it is intuitively clear that the error propagation effect will drastically degrade the performance as λ increases from 0.

Here we propose a rather simple and robust approach to overcome the effect sequentially by a compensating scheme. More precisely, it will estimate the error due to the previous finalized estimate over the interval over which the estimation and detection of the incidents are being processed. By taking away the errors due to the finalized estimates sequentially, the propagation effect is eliminated, and sequentially we have a new smoothing problem over a new single interval of length $K\Delta$ each time.

4.5.1 Error Pattern

To investigate the error propagating pattern in the SSS, let us assume that we have just applied the scheme over the interval $[0,K\Delta)$ and obtained the estimates $\underline{\tau},\underline{u}$ and the estimate of the system state based on the observations over $[0,K\Delta)$. Now we finalize the incident over $[0,\Delta)$ (assumed exists) and take in a new subinterval of observation of length Δ; apply the scheme and obtain another set of estimates $\tilde{\underline{\tau}},\tilde{\underline{u}}$ over the new interval $[\Delta,(K+1)\Delta)$.

It should be clear that even if we consider the entire interval $[0,(K+1)\Delta)$ for the SSS, the estimates of the incidents and their marks should be more or less as $\underline{\tau},\underline{u}$, especially the ones occurred at earlier time. In a lot of cases, the estimates of the early incidents and marks are identically the same, since the additional subinterval of observation of length Δ is unlikely to improve those estimates. Hence it is quite safe to assume for all practical purpose that over the subinterval $[\Delta,2\Delta)$,

$$(\tilde{\tau},\tilde{u}) \cong (\tau,u) \tag{4.52}$$

Indeed from Eq. (4.21), we see that (τ,u) has to satisfy the following

$$\int_0^{K\Delta} dt[z_t - \sum_{i=1}^{N(k\Delta)} g_t(\tau_i,u_i)] \frac{\partial g_t(\tau_2,u_2)}{\partial \tau_2} = 0 \tag{4.53a}$$

$$\int_0^{K\Delta} dt[z_t - \sum_{i=1}^{N(k\Delta)} g_t(\tau_i,u_i)] \frac{\partial g_t(\tau_2,u_2)}{\partial \tau_2} + \frac{\partial}{\partial u_2} \ell n f_U(u_2) = 0 \tag{4.53b}$$

In particular, we have from causality of g,

$$\int_{\Delta}^{2\Delta} dt\{z_t - g_t(\tau_1,u_1) - g_t(\tau_2,u_2)\} \frac{\partial g_t(\tau_2,u_2)}{\partial \tau_2} = \xi_\tau \simeq 0 \tag{4.54a}$$

$$\int_{\Delta}^{2\Delta} dt\{z_t - g_t(\tau_1,u_1) - g_t(\tau_2,u_2)\} \frac{\partial g_t(\tau_2,u_2)}{\partial u_2} + \frac{\partial}{\partial u_2} \ln f_U(u_2) = \xi_u \simeq 0 \tag{4.54b}$$

where ξ_τ and ξ_u are two constants and should be closed to zero. Similarly for the estimates $\tilde{\tau}_2,\tilde{u}_2$, denoting the estimates of incident in $[\Delta,2\Delta)$, we have

$$\int_{\Delta}^{2\Delta} dt\{\tilde{z}_t - g_t(\tilde{\tau}_2,\tilde{u}_2)\} \frac{\partial g_t(\tilde{\tau}_2,\tilde{u}_2)}{\partial \tilde{\tau}_2} = \tilde{\xi}_\tau \sim 0 \tag{4.55a}$$

$$\int_{\Delta}^{2\Delta} dt\{\tilde{z}_t - g_t(\tilde{\tau}_2,\tilde{u}_2)\} \frac{\partial g_t(\tilde{\tau}_2,\tilde{u}_2)}{\partial \tilde{u}_2} + \frac{\partial}{\partial \tilde{u}_2} \ln f_U(\tilde{u}_2) = \tilde{\xi}_u \sim 0 \tag{4.55b}$$

where $\tilde{\xi}_\tau$ and $\tilde{\xi}_u$ are two constants close to zero. But $\tilde{z}_t = z_t - g_t(\tau_1,u_1)$ from Eq. (4.31d). Hence we have

$$\int_{\Delta}^{2\Delta} dt\{z_t - g_t(\tau_1,u_1) - g_t(\tilde{\tau}_2,\tilde{u}_2)\} \frac{\partial g_t(\tilde{\tau}_2,\tilde{u}_2)}{\partial \tilde{\tau}_2} = \tilde{\xi}_\tau \sim 0 \tag{4.56a}$$

$$\int_{\Delta}^{2\Delta} dt\{z_t - g_t(\tau_1,u_1) - g_t(\tilde{\tau}_2,\tilde{u}_2)\} \frac{\partial g_t(\tau_2,u_2)}{\partial \tilde{\tau}_2} + \frac{\partial}{\partial \tilde{u}_2} \ln f_U(\tilde{u}_2) = \tilde{\xi}_u \sim 0 \tag{4.56b}$$

Comparing Eq. (4.54) and Eq. (4.56), we assume that $\tilde{\tau}_2,\tilde{u}_2$ are close to τ_2,u_2 for all practical purposes.

Note that the two sets of estimates of incidents and marks over the subinterval $[\Delta,2\Delta)$ are very close does not imply that we do not have errors in the estimates. It simply says that over the intervals $[0,K\Delta)$ and $[\Delta,(K+1)\Delta)$, the MAP estimates of the incidents and marks over $[\Delta,2\Delta)$ are very close. Actually, the uncertainty of the incidents and marks is distributed in an MAP fashion over the estimates of the incidents in the entire interval. Finalizing the first estimate or not does not affect very much the estimate of the incident right after it since they have been estimated together and shared the uncertainty accordingly.

Hence finalization of estimates of incidents and marks does not affect much the subsequent estimation and detection of incident right after it provided K is reasonably large. Nevertheless, the finalization of estimates of an incident does introduce an error function into the subsequent subintervals. With the reconstruction scheme of using a filter (4.12) and a smoother (4.19), the resultant error can be sufficiently large that should not be neglected with the consideration that it would propagate and perpetuate into the subsequent intervals.

Unlike the case that the finalized estimates does not affect too much of the detection and estimation of its immediately subsequent incident, the error arising from a finalized estimate of an incident affects most the reconstruction of the system state over its immediately subsequent subinterval in the SSS. In the reconstruction process, the smoother (4.19) can be regarded as a filter driven backward in time over $[\Delta,K\Delta)$ by the filter (4.12) output, which in turn is driven by the entire error function over $[\Delta,K\Delta)$. Hence, the error over $[\Delta,2\Delta)$ is a result of a cascaded system of filters, both of which in effect are driven by the error function arising from the finalized estimate through a time-interval of $[\Delta,K\Delta)$.

Let us denote the error from the finalized estimate by $r_t(\tau_0,u_0)$ where τ_0,u_0 are the finalized estimates.

Hence $\quad \tau_0 < \Delta$

and $\quad r_t(\tau_0,u_0) = 0 \qquad \tau_0 > t \qquad (4.57)$

Recall that $\varphi(t,s)$ is the state transition matrix of the filter (4.12); and let us denote the same for smoother (4.19) by $\varphi_b(t,s)$.

From Eqs.(4.12), (4.19), the resultant error w_t^1 due to $r_t(\tau_0,u_0)$ can be easily derived and is given as follows

$$
\begin{aligned}
w_t^1(\tau_0,u_0) &= \varphi(t,(K+1)\Delta)\varphi_b(k+1)\Delta,\Delta)r_\Delta(\tau_0,u_0) \\
&+ \int_0^{(K+1)\Delta} \varphi(t,(K+1)\Delta)\varphi_b((K+1)\Delta,\sigma)r_\sigma(\tau_0,u_0)d\sigma \\
&- \int_t^{(K+1)\Delta} \varphi(t,\sigma)\varphi_b(\sigma,\Delta)r_\Delta(\tau_0,u_0)d\sigma \\
&- \int_t^{(K+1)\Delta} \varphi(t,\sigma)d\sigma \int_\Delta^\sigma \varphi_b(\sigma,u)r_u(\tau_0,u_0)du
\end{aligned}
$$

$$\Delta \leq t < (K+1)\Delta \qquad (4.58)$$

However, before the estimates τ_0,u_0 are finalized, a similar error due to the deviations of these estimates also exists over $[0,K\Delta)$ through the reconstruction process. Denoting this error function by $w_t^0(\tau_0,u_0)$, we have a similar expression for it as follows,

$$
\begin{aligned}
w_t^0(\tau_0,u_0) &= \varphi(t,K\Delta)\varphi_b(K\Delta,\tau_0)\ r_{\tau_0}(\tau_0,u_0) \\
&+ \int_0^{K\Delta} \varphi(t,K\Delta)\varphi_b(K\Delta,\sigma)\ r_\sigma(\tau_0,u_0)d\sigma \\
&- \int_t^{K\Delta} \varphi(t,\sigma)\varphi_b(\sigma,\tau_0)\ r_{\tau_0}(\tau_0,u_0)d\sigma \\
&- \int_t^{K\Delta} \varphi(t,\sigma)d\sigma \int^\sigma \varphi_b(\sigma,u)\ r_u(\tau_0,u_0)du \qquad 0 \leq t < K\Delta \qquad (4.59)
\end{aligned}
$$

Notice that $w_t^0(\tau_0,u_0)$ and $w_t^1(\tau_0,u_0)$ are both zero identically if $r_t(\tau_0,u_0)$ is zero.

4.5.2 Compensation Approach to Eliminate Error Propagation

In this section we develop a robust compensation scheme to eliminate the error propagation effect in the SSS. Based on the results in 4.41 about the error patterns, we like to assess the error due to the finalized estimate of the incident in a sequential manner.

Let us denote the reconstructed estimates of the system state $\overset{\star}{X}_t$ and $\tilde{X}_t$ over $[\Delta,2\Delta)$ respectively from the two sets of incident estimates $(\underline{\hat{\tau}},\underline{\hat{u}})$ and $(\underline{\tilde{\tau}},\underline{\tilde{u}})$. Hence from previous discussions and results (4.52) (4.58) and (4.59),

$$(\overset{\star}{X}_t - \tilde{X}_t) \cong w_t^0(\tau_0,u_0) - w_t^1(\tau_0,u_0) \qquad \Delta \leq t \leq 2\Delta \qquad (4.60)$$

Note that $\overset{\star}{X}_t$, $\tilde{X}_t$ are just the sequential reconstructed system states, and are most readily available in the SSS process. Equation (4.60) simply says by comparing the reconstructed system state over a subinterval sequentially, we can approximately compute the error due to the finalized estimates sequentially.

It should be noted that in arriving Eq. (4.60), we have only assumed that an additional subinterval observation of length Δ cannot improve much about the estimates in remotely early subintervals. The information of the error arising from the finalized estimate is gained by the sequential reconstruction process; it is over the same additional subinterval of length Δ that the information is gained, but that subinterval is about $2K\Delta$ afar. For unstable systems in which case error tends to grow with time, a small interval of observations will give considerable information when is far away.

For our compensation purposes, let us consider the first order effect compensation only (higher order compensation can easily be extended). Since by definition, the error function

$$r_t(\tau_0,u_0) \triangleq g_t^0(\tau,u) - g_t^0(\tau_0,u_0) \tag{4.61}$$

which can be expressed as follows

$$r_t(\tau_0,u_0) = [g_t^0(\tau_0,u_0) - g_t^0(\tau_0,u_0)]$$

$$+ [\frac{\partial}{\partial\tau_0} g_t^0(\tau_0,u_0)](\Delta\tau_0) + O(\Delta\tau_0)$$

$$+ [\frac{\partial}{\partial u_0} g_t^0(\tau_0,u_0)](\Delta u_0) + O(\Delta u_0) \tag{4.62}$$

Putting Eq. (4.62) into Eq. (4.58), (4.59) and (4.60) we have a relationship which can be algebraically represented as

$$(\overset{*}{X}_t - X_t) \cong \overline{\overline{w_t}}(\tau_0,u_0)(\Delta\tau) + \overline{w_t}(\tau_0,u_0)(\Delta u_0) \tag{4.63}$$

where $\overline{\overline{w_t}}$ and $\overline{w_t}$ are the appropriate expressions. A simple computation method to evaluate $(\Delta\tau)$ and (Δu) is

$$\int_\Delta^{2\Delta} (\overset{*}{X}_t - \bar{X}_t)dt = (\Delta\tau)\int_\Delta^{2\Delta} \overline{\overline{w_t}}(\tau_0,u_0)dt + (\Delta u)\int_\Delta^{2\Delta} \overline{w_t}(\tau_0,u_0)dt \tag{4.64a}$$

and similarly we have

$$\int_\Delta^{2\Delta} (\overset{*}{X}_t - \bar{X}_t)^2 dt = (\Delta\tau)^2 \int_\Delta^{2\Delta} \overline{\overline{w_t}}^2(\tau_0,u_0)dt$$

$$+ (\Delta\tau)(\Delta u)\int_\Delta^{2\Delta} \overline{w_t}(\tau_0,u_0)\,\overline{\overline{w}}(\tau_0,u_0)dt$$

$$+ (\Delta u)^2 \int^{2\Delta} \overline{w_t}^2(\tau_0,u_0)dt \tag{4.64b}$$

The two Eqs. (4.64a) and (4.64b) can determine the approximate values of ($\Delta\tau$) and (Δu) and hence the error function from the finalized estimate (τ_0, u_0).

To summarize the compensation approach, we have developed a scheme to learn about the error functions from the finalized estimates of incidents sequentially. The learning is done through the structure of the reconstruction process which utilizes in effect an interval of $2K\Delta$ long. By compensating the error function approximately in a sequential fashion in the SSS, the error propagating effect can be eliminated or at least dimenished considerably. For unstable systems, a length of $2K\Delta$ is sufficient to give a lot of information about the error functions, one would deduce that the compensation approach works even better than it would in a stable system.

4.6 Performance Simulations of the SSS

The sequential scheme was simulated on a digital computer to obtain samples of its performance in the state estimation problem. In each case, the simulation was done by first allowing the scheme to run for a time interval to include 20 incidents. Then the performance was computed over the subsequent interval of duration $30/\lambda$, which on the average would allow about 30 Poisson incidents. For comparison purposes, the time scale were chosen so that in each case λ was equal to one. We chose $\alpha = 0.15$, and the length for the integral $(K+M) = 4$. The system state was reconstructed with delay of 2.

In each example, the sample performance will be presented in form of three graphs plotter over a time interval of length 30. The first graph will indicate the normalized error variance derived analytically i.e. $\mathrm{Var}(\epsilon_t^n)$. The second graph will give the simulated performance in terms of the normalized error variance of the SSS without the compensation strategy. The third will give that simulated performance of the SSS with the compensation strategy.

Several examples were considered as follows:

Example I: $\dot{x}(t) = -2x(t) + 2\dot{W}_t + U\dot{\eta}(t);$ $\quad x(0) = 0$

$y(t) = x(t) + v(t)$

The mark U is assumed to be Gaussian with variance 5 and mean = (a) 5 (b) 7. (See Figs. 8a and 8b.)

Example II: $\dot{x}(t) = 2x(t) + 2\dot{W}_t + U\dot{\eta}_{(t)}$; $\quad x(0) = 0$

$y(t) = x(t) + v(t)$

The mark U is assumed to be Gaussian with variance 5 and mean = (a) 5 (b) 7. (See Figs. 9a and 9b.)

Example III: $\dot{x}(t) = 3x(t) + \dot{W}(t) + \dot{\eta}(t);$ $\quad x(0) = 0$

$y(t) = x(t) + v(t)$

(For results, see Fig. 10.)

From the simulation study, the results indicated that the sequential approximation scheme did exhibit an error propagation phenomenon both for stable and unstable systems. In cases of unstable systems as Example II and III had indicated, the propagating errors could build up to great magnitudes severely degrading the overall performance of the SSS. (See Figs. 9a, 9b and 10.) The results in all the examples confirmed the ability of eliminating most of the propagating errors of the compensation strategy. In the case of an unstable system, the adoption of this compensation approach

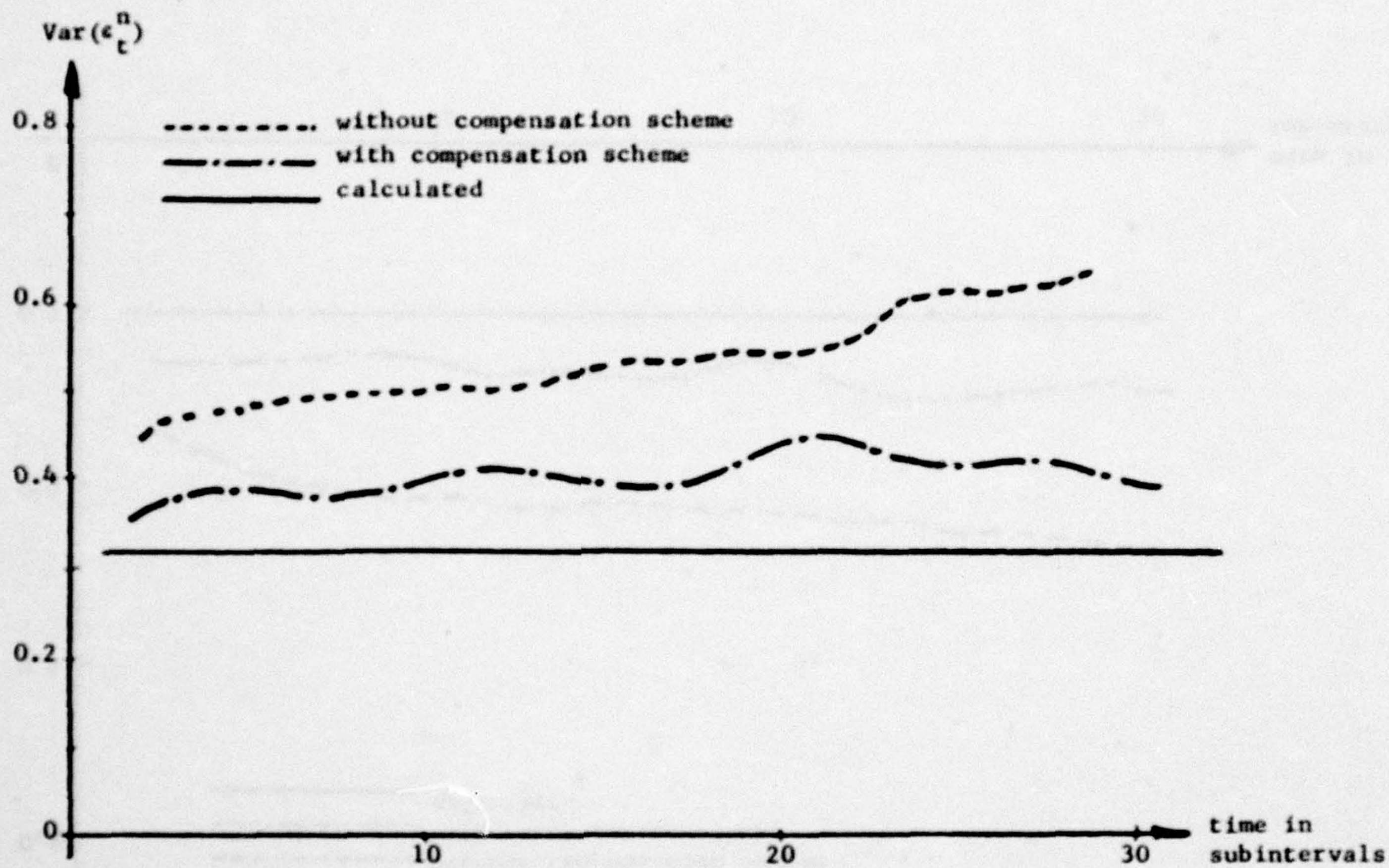

Figure 8a. Simulated performance of the SSS with and without the compensation scheme in comparison to the analytically derived performance (Example 4-Ia).

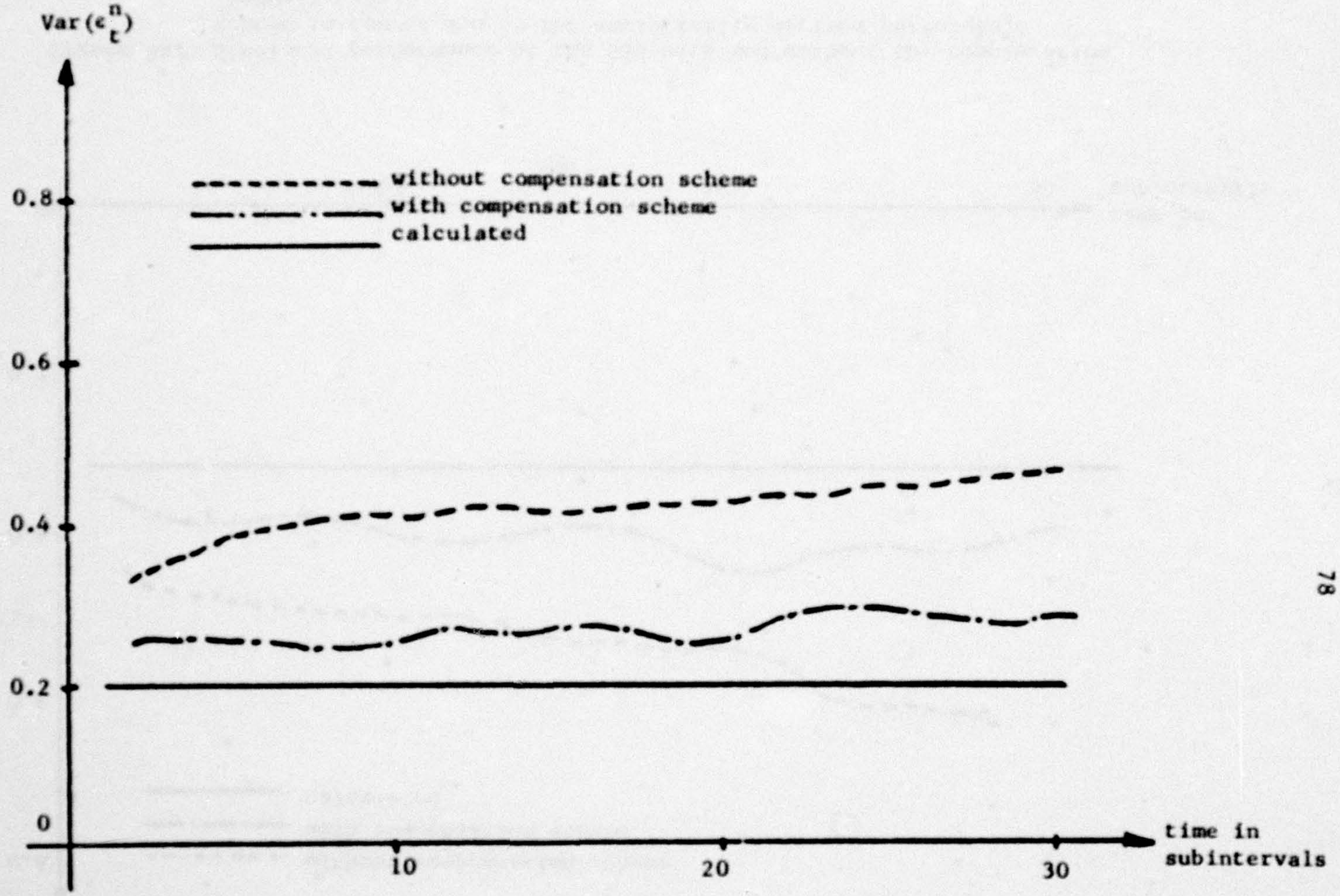

Figure 8b. Simulated performance of the SSS with and without the compensation scheme in comparison to the analytically derived performance (Example 4-Ib).

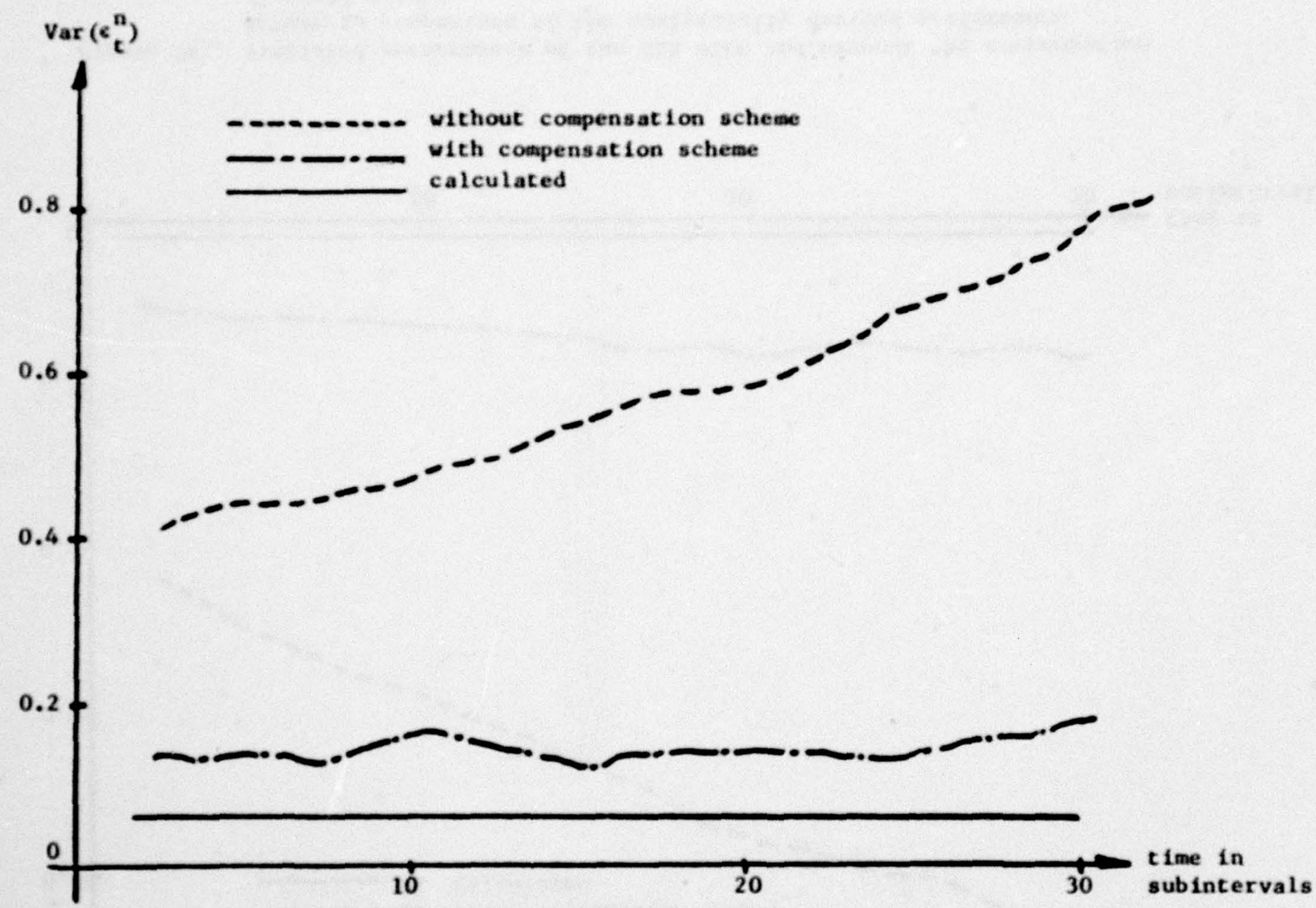

Figure 9a. Simulated performance of the SSS with and without the compensation scheme in comparison to the analytically derived performance (Example 4-IIa).

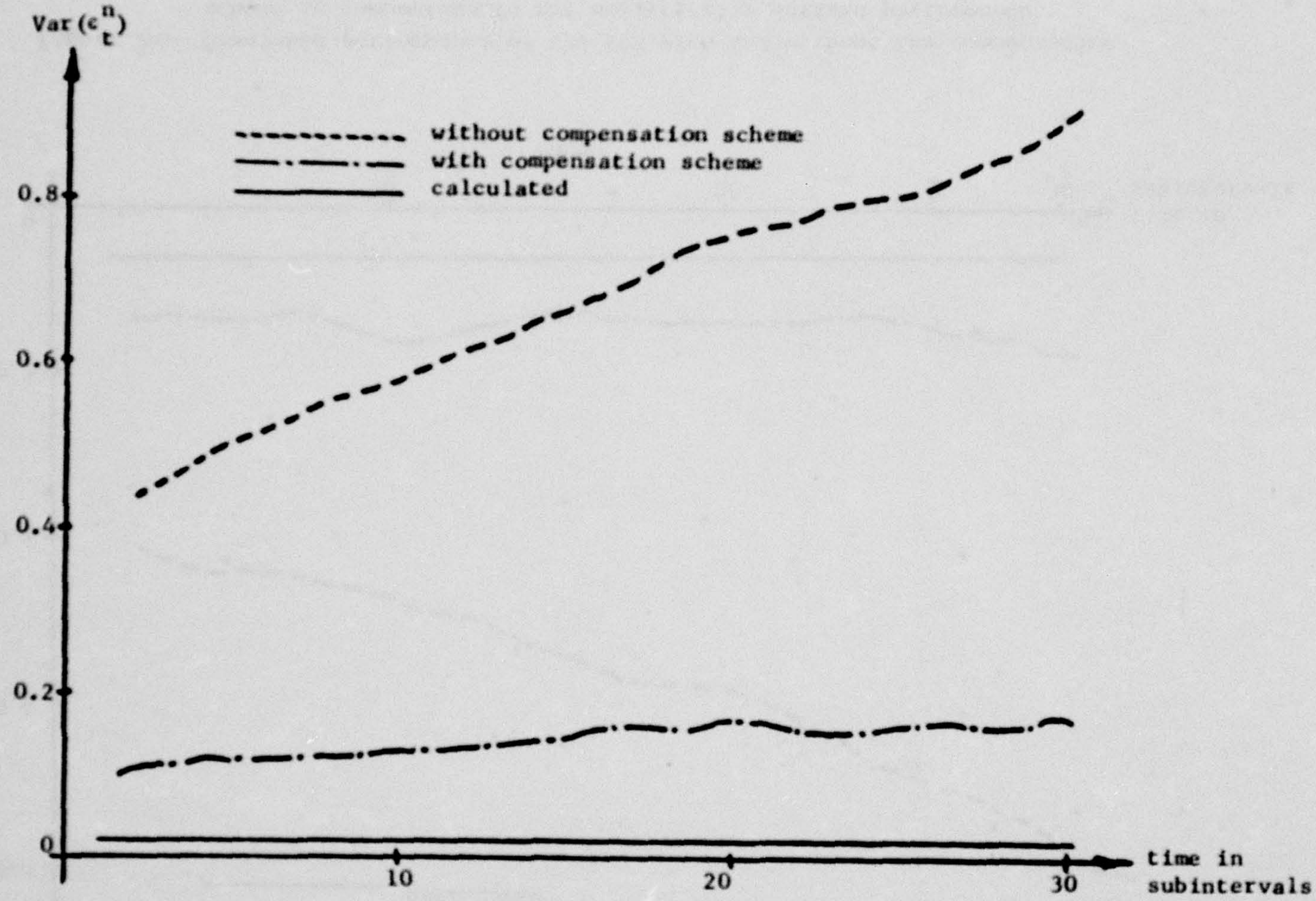

Figure 9b. Simulated performance of the SSS with and without the compensation scheme in comparison to the analytically derived performance (Example 4-IIb).

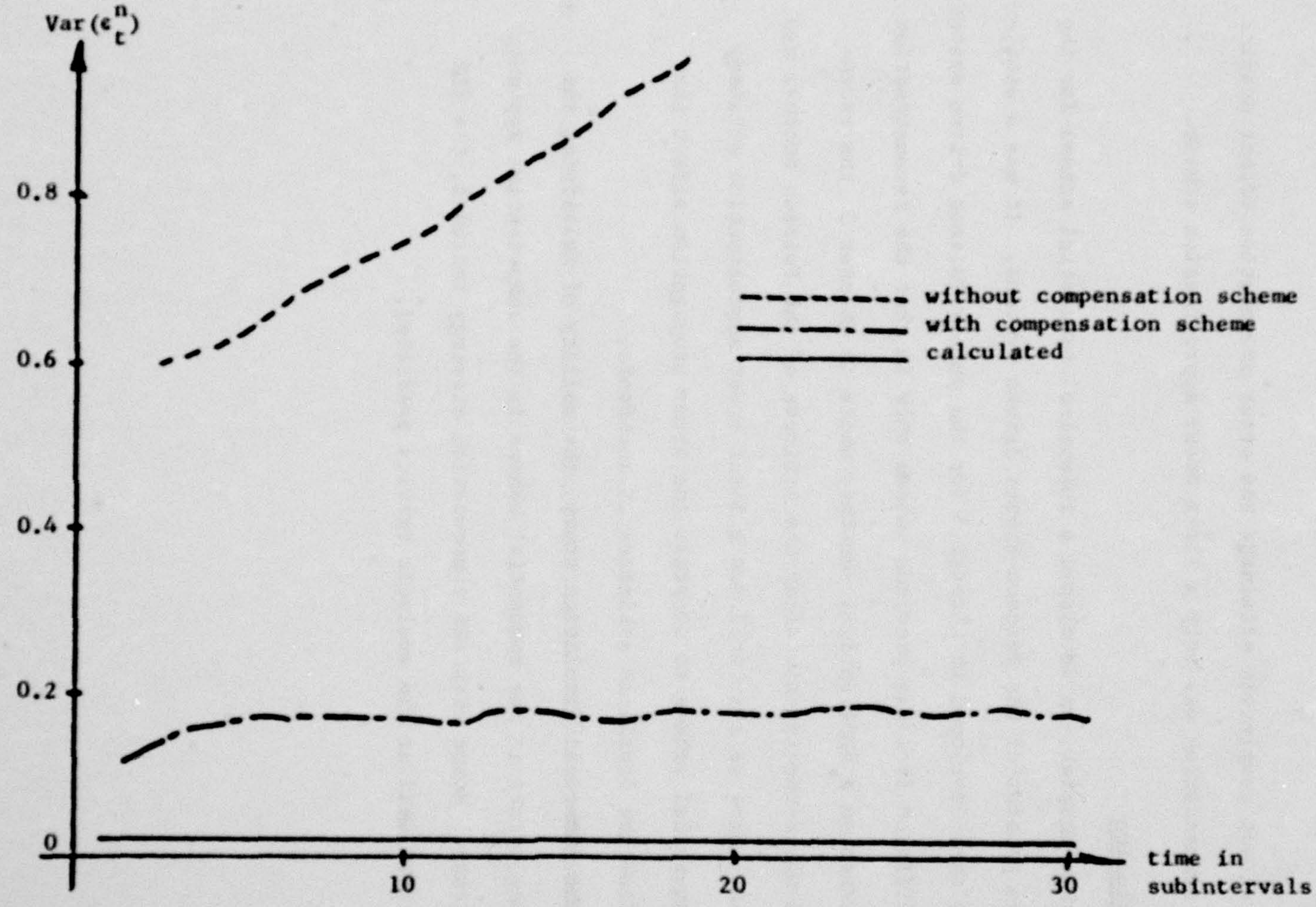

Figure 10. Simulated performance of the SSS with and without the compensation scheme in comparison to the analytically derived performance (Example 4-III).

could in effect completely eliminate the error propagation effect despite that the compensation was only a first order approximation scheme.

4.7 Conclusion

In this chapter, we developed a suboptimal sequential scheme for the estimation problem of the Poisson-Wiener driven systems. It was a similar scheme to that developed in Chapter 3 for the purely Poisson driven system. The SSS differed from the previous scheme only in that the reconstruction was performed on a MMSE optimal smoother while in Chapter 3 the reconstruction was deterministic after the estimate of the Poisson process was determined. Here we also included a first order approximation strategy in the sequential scheme to suppress the error propagation effect that stemmed from the finalized estimates of incidents.

In the numerical simulation study, the ability of nullifying the propagating errors in the sequential scheme in the compensation approach was confirmed. Hence with the compensation strategy included, the SSS would perform well as the analytic results predicted.

CHAPTER 5

EXTENSION OF THE SUBOPTIMAL SEQUENTIAL SCHEME (SSS)

As it has been indicated in the derivations of the SSS in previous chapters, the key requirement for applicability of the scheme is the existence of a closed (finite dimensional) MMSE optimal filter for the pseudo-estimates of the system state conditional on the σ-field generated by the Poisson-marked process. Furthermore, the corresponding likelihood function for the Poisson-marked process must exist, so that the detections and estimations of the process can be performed using the MAP criterion.

In this chapter, we shall consider the state estimation problem for a class of systems, namely the class of conditional Gaussian systems, which readily satisfy the two requirements above. As we have developed the SSS to quite an extent in previous two chapters, here we derive only the closed filters for the pseudo-estimates together with an expression for the likelihood function for this class of systems. It should be obvious that with these results, the SSS for the estimation problem can be readily obtained using the similar approach as before.

In order to state the system in precise terms, the notation $x = (x_t, \mathcal{F}_t)$ will be used to denote the random process $x_t(\omega)$, measurable with respect to $\mathcal{F}_t$. Let $(\Omega, \mathcal{F}, P)$ be the (complete) probability space as given, with a nondecreasing right continuous family of sub-σ-algebras $(\mathcal{F}_t)$, $0 \leq t < T$.

The system considered in this chapter can be represented by

$$dx_t = [\mathcal{A}_1(t,\Pi,y) + \mathcal{A}_2(t,y)x_t]dt + \sum_{i=1}^{2} \mathcal{B}_i(t,y)dW_i(t) \tag{5.1}$$

$$+ \mathcal{B}_0(t,U)d\Pi_t$$

and is observed through the following process,

$$dy_t = [\mathcal{C}_1(t,y) + \mathcal{C}_2(t,y)x_t]dt + \sum_{i=1}^{2} \mathcal{D}_i(t,y)dW_i(t) \qquad (5.2)$$

Here $\mathcal{A}_i(t,\xi)$, $\mathcal{B}_i(t,\xi)$, $\mathcal{C}_i(t,\xi)$, $\mathcal{D}_i(t,\xi)$, $i = 1,2$ are (measurable) functionals, assumed to be non-anticipative (i.e. $\mathcal{F}_t$- measurable where $\mathcal{F}_t$ is the σ-algebra in the space C_T of the continuous functions $\xi = \{\xi_s, s<T\}$, generated by the functions ξ_s, $s \leq t$). And $\mathcal{B}_0(t,U)$ is assumed to be a measurable function in time and the mark U driven by the Poisson-marked process $\eta = (\eta_t(U), \mathcal{F}_t)$. It should be clear that the random process η is characterized by the Poisson incident times $\underline{\tau}$ and the corresponding mark $\underline{u}$. Note that the system (5.1) to be estimated can depend on the entire history of the observed process.

In order for the system given in (5.1) to be conditional Gaussian, the following relations have to be satisfied [7,8,22,23],

(1) for each $\xi \in C_T$, with probability one

$$\int_0^T \{ \sum_{i=1}^{2} [|\mathcal{A}_i(t,\xi)| + |\mathcal{C}_i(t,\xi)| + \mathcal{B}_i^2(t,\xi) + \mathcal{D}_i^2(t,\xi)] + \mathcal{B}_0^2(t,U)\}dt < \infty \qquad (5.3)$$

$$(2) \quad \int^T [\sum_{i=1}^{2} \mathcal{C}_i^2(t,\xi)]\, dt < \infty \qquad (5.4)$$

$$\inf_{\xi \in C_T} \mathcal{D}_i(t,\xi) \geq c_i > 0 \qquad i = 1,2, \quad 0 \leq t \leq T \qquad (5.5)$$

(3) For any $\xi, \zeta \in C_T$, $i = 1,2$

$$|\mathcal{D}_i(t,\xi) - \mathcal{D}_i(t,\zeta)|^2 \leq L_1 \int_0^t |\xi_s - \zeta_s|^2 dK(s) + L_2|\xi_t - \zeta_t|^2 \qquad (5.6)$$

$$\mathcal{B}_i^2(t,\xi) \le L_1 \int_0^t (1+\xi_s^2)dK(s) + L_2(1+\xi_t^2) \tag{5.7}$$

where K(s) is a nondecreasing right continuous function, $0 \le K(s) \le 1$.

(4)
$$\int_0^T E|\mathcal{C}_2(t,\xi)x_t|dt < \infty \tag{5.8}$$

$$E|x_t| < \infty \qquad 0 \le t \le T \tag{5.9}$$

$$P\{\int_0^T \mathcal{C}_2^2(t,\xi)m_t^2 dt < \infty\} = 1 \tag{5.10a}$$

where
$$m_t = E(x_t|\mathcal{F}_t^y,\mathcal{F}_t^\eta) \tag{5.10b}$$

In Eqs. (5.1) and (5.2) we also assume that $W_1 = (W_1(t),\mathcal{F}_t)$, $W_2 = (W_2(t),\mathcal{F}_t)$ are independent Wiener processes and the random initial conditions x_0 and y_0 are independent of W_1 and W_2.

The system with the observations satisfying (5.1) and (5.2) is conditionally Gaussian if with probability one the conditional distribution $F_{y_0}(a) = P(x_0 \le a|y_0)$ is Gaussian. That is,

$$F_{y,\eta}(\theta_0,\theta_1 \cdots \theta_n) = P(x_{t_0} \le \theta_0, \cdots x_{t_n} \le \theta_n|\mathcal{F}_t^y,\mathcal{F}_t^\eta) \tag{5.11}$$

is Gaussian almost surely, [7,8,22,23] for any t and $0 \le t_0 < t_1 \ldots t_n < t$.

The class of conditional Gaussian processes similar to that given in (5.1) without the presence of the Poisson process $\eta = (\eta_t,\mathcal{F}_t)$ was first formulated and investigated by Liptser [22] and subsequently was used for the optimal nonlinear filtering problems [7,8,22]. For this class of processes, the MMSE optimal filters were closed due to the conditional

Gaussianness of the processes. (See Appendix D.) It should be emphasized that the Kalman filter is an optimal filter of a particular case of (5.1) when η is absent. When the system processes are Gaussian, the optimal filter as a result is linear, however in the conditionally Gaussian case the optimal filter is, in general, nonlinear [23-26].

Using the optimal filters derived for this class of conditional Gaussian processes, the pseudo-estimate for $x = (x_t, \mathcal{F}_t)$, $0 \leq t \leq T$ and its conditional variance denoted respectively by the m_t and P_t can be easily shown to satisfy the following, (Appendix D)

$$m_t = m_t^0 + m_t^1 \tag{5.12}$$

such that

$$dm_t^0 = [\mathcal{A}_1(t,\eta,y) + \mathcal{A}_2(t,y)m_t^0]dt + [\frac{\mathcal{B}_1(t,y)\mathcal{D}_1(t,y) + \mathcal{B}_2(t,y)\mathcal{D}_2(t,y) + P_t\mathcal{C}_2(t,y)}{\mathcal{D}_1^2(t,y) + \mathcal{D}_2^2(t,y)}]$$

$$\cdot\ [dy_t - (\mathcal{C}_1(t,\eta,y) + \mathcal{C}_2(t,\eta)m_t^0) - \mathcal{C}_1(t,y)\mathcal{B}_0(t,U)d\eta_t] \tag{5.13a}$$

$$\dot{P}_t = 2\mathcal{A}_2(t,y)P_t + \mathcal{B}_1^2(t,y) + \mathcal{B}_2^2(t,y) - [\frac{\mathcal{B}_1(t,y)\mathcal{D}_2(t,y) + P_t\mathcal{C}_1(t,y)}{\mathcal{D}_2(t,y)}]^2 \tag{5.13b}$$

with initial conditions

$$m_0 = E(x_0 | y_0) \tag{5.14a}$$

$$P_0 = E[(x_0 - m_0)^2 | y_0] \tag{5.14b}$$

and

$$m_t^1 = \int_0^t \Phi_y(t,s)\, \mathcal{B}_0(s,U)d\eta_s \tag{5.15}$$

where $\Phi_y(t,s)$ is the state transition function of the system (5.1) given the observations $\{y_v;\ v \le t\}$. Φ in this case is a functional of $y = (y_t, \mathcal{F}_t)$ since from Eq. (5.1), the system function $\mathcal{A}_2$ is a functional of y.

Hence we have a finite dimensional MMSE optimal filter for the pseudo-estimate of $x = (x_t, \mathcal{F}_t)$ conditional on the σ-algebra generated by $\eta = (\eta_t, \mathcal{F}_t)$. In order to derive the equations for the detection and estimation of the Poisson incidents and their corresponding marks, we shall obtain the likelihood function for $\{\eta_t;\ 0 \le t < T\}$. It should be obvious that once this is obtained, the incidents and their marks can be substituted into the function in the place of $\{\eta_t;\ 0 \le t < T\}$, so that an expression for the aposteriori probability of the incidents and marks can be obtained. Hence from now on, the equations containing η are understood to be functions of the Poisson incidents and marks.

Applying the estimator-correlator principle, the likelihood function for $\{\eta_t;\ 0 \le t < T\}$, denoted by $\Lambda_T[\eta]$ can be obtained and is given by [22,23], (see Appendix D also)

$$\Lambda_T[\eta] = \exp\Big\{ - \int_0^T \Big[\frac{\mathcal{C}_1(s,y) + \mathcal{C}_2(s,y) m_s(y,\eta)}{\mathcal{S}_1^2(s,y) + \mathcal{S}_2^2(s,y)}\Big] dy_s$$

$$+ \frac{1}{2} \int_0^T \frac{[\mathcal{C}_1(s,y) + \mathcal{C}_2(s,y) m_s(y,\eta)]^2}{\mathcal{S}_1^2(s,y) + \mathcal{S}_2^2(s,y)} ds \Big\} \qquad (5.16)$$

where we denote the dependence of m_t on y,η by $m_t(y,\eta)$. Hence an equivalent expression for the aposteriori probability density for $\{\eta_t;\ 0 \le t < T\}$ can be obtained and is given by

AD-A077 149 ILLINOIS UNIV AT URBANA-CHAMPAIGN COORDINATED SCIENCE LAB F/G 12/1
STATE ESTIMATION FOR LINEAR SYSTEMS DRIVEN SIMULTANEOUSLY BY WI--ETC(U)
DEC 78 S P AU DAAB07-72-C-0259

UNCLASSIFIED R-836 NL

2 OF 2
AD-A077149

END
DATE
FILMED
12-79
DDC

$$J_T[\eta] = \int_0^T - \frac{\mathcal{C}_2(s,y)\ m_s(y,\eta)}{\mathcal{S}_1^2(s,y) + \mathcal{S}_2^2(s,y)}\ dy_s$$

$$+ \int_0^T \{ \frac{\mathcal{C}_1(s,y)\mathcal{C}_2(s,y)m_s(y,\eta)}{\mathcal{S}_1^2(s,y) + \mathcal{S}_2^2(s,y)} + \frac{[\mathcal{C}_2(s,y)m_s(y,\eta)]^2}{\mathcal{S}_1^2(s,y) + \mathcal{S}_2^2(s,y)} \}\ ds$$

$$+ \ell n\ f_T(\eta) \tag{5.17}$$

where we denote the apriori probability density of $\{\eta_t;\ 0 \leq t < T\}$ by $f_T(\eta)$. Note that for our case $f_T(\eta)$ is a well defined probability density, since η is a separable process [27]. In fact, when $\{\eta_t;\ 0 \leq t < T\}$ is explicitly expressed in terms of the incidents and marks, $f_T(\eta)$ is just a function of the joint probability density of the marks and the joint probability density of the incidents over $[0,T)$.

Similarly as in Chapters 3 and 4, the maximization of $J_T[\eta]$ will give the MAP estimates of the incidents and their marks. For the reconstruction of the system, the optimal MMSE smoother will be used and can be represented as follows [22,23] (see also Appendix D)

$$d_t m^0_{t/T} = [\mathcal{A}_1(t,y) + \mathcal{A}_2(t,y)m^0_{t/T}]dt$$
$$+ \underline{\mathcal{B}}(t,y)P_t^{-1}[m_t - m_{t/T}]dt$$
$$+ (\mathcal{B} * \mathcal{S})(t,y)(\mathcal{S} * \mathcal{S})^{-1}(t,y)[dy_t - (\mathcal{C}_1(t,y) + \mathcal{C}_1(t,y)m_{t/T})dt] \tag{5.18a}$$

$$\dot{P}_{t/T} = 2[\underline{\mathcal{A}}(t,y) + \underline{\mathcal{B}}(t,y)P_t^{-1}]P_{t/T} - \underline{\mathcal{B}}(t,y) \tag{5.18b}$$

with initial conditions

$$P_{T/T} = P_T \tag{5.19a}$$

and

$$m_{T/T} = m_T \tag{5.19b}$$

Here the notations we use are as follows

$$\mathcal{B} * \mathcal{S}(t,y) = \mathcal{B}_1(t,y)\mathcal{S}_1(t,y) + \mathcal{B}_2(t,y)\mathcal{S}_2(t,y) \tag{5.20}$$

and

$$\underline{\mathcal{C}}(t,y) = \mathcal{C}_1(t,y) - (\mathcal{B} * \mathcal{S})(t,y)(\mathcal{S} * \mathcal{S})^{-1}(t,y)\mathcal{C}_1(t,y) \tag{5.21}$$

$$\underline{\mathcal{B}}(t,y) = \mathcal{B} * \mathcal{B}(t,y) - (\mathcal{B} * \mathcal{S})^2(t,y)(\mathcal{S} * \mathcal{S})^{-1}(t,y) \tag{5.22}$$

The reconstruction of the system state $x = (x_t, \mathcal{F}_t)$ can be achieved by

$$\hat{x}_{t/T} = m^0_{t/T} + m^1_t \tag{5.23}$$

where m^1_t is given in Eq. (5.15).

To summarize this chapter, we have extended our original estimation problem to a general class of conditional Gaussian processes. As a result, the SSS can be similarly developed as a near-optimal scheme to the problem. Here we have derived all the basic equations, expressions for the combined estimation detection scheme. The sequential suboptimal scheme (SSS) can be easily derived based on these equations and expressions. While this is conceptually straightforward and highly parallel to the algorithm in Chapter 4, the detail development of the sequential scheme is omitted here. The procedures and equations in the SSS are similar to those in Chapter 4. In fact the estimation problems in Chapter 3 and 4 are particular cases of the one considered here.

CHAPTER 6

SUMMARY AND CONCLUSIONS

In this thesis, the problem of state estimation for the linear Poisson-Wiener driven systems with low Poisson intensity was considered. The MMSE optimal filter was first obtained via the Dolean-Dade and Meyer differentiation rule for discontinuous semi-martingales, and the basic filtering theorem for white Gaussian observation noise. The optimal filter obtained was not closed in the sense that it was the solution of a set of an infinite number of stochastic differential equations, unlike the Gaussian case where an additional relationship of its moments reduced the set of equations to a finite dimensional set, resulting in the classical Kalman filter.

The nonclosedness of the optimal filter for our problem was proved with the following two arguments:

(1) If a similar relationship existed to reduce the infinite set of equations to a finite set, then the solutions to the set of equations were guaranteed to exist and furthermore were unique, hence the process $I = (I_t, \mathcal{F}_t^y)$ defined by $dI_t = dy_t - h_t C_{1t} dt$ was an innovation process and therefore a Wiener process. Therefore the estimator of the system state being the solution to a finite set of equations driven by a Wiener process would have to be continuous with probability one. Since the system being estimated was discontinuous with non-zero probability, the optimality of the estimator, was contradicted. We hence concluded that there was no such relationship that would reduce the infinite set of equations to a finite set, so that the optimal filter was not finite dimensional.

(2) By the definition of an innovation process, the information generated by it had to be equal to that generated by the observations. That is to say, the two σ-algebras generated by the innovation process and the observations individually had to be equal for all t, $0 \leq t < \infty$. However $y = (y_t, \mathcal{F}_t)$ being an observation process of a discontinuous system in white Gaussian noise was discontinuous with non-zero probability, implying that $\mathcal{F}_t^y$ was only right-continuous with non-zero probability. In order for a $\mathcal{F}_t^y$-measurable process to be an innovation process, its σ-algebra was necessarily equal to $\mathcal{F}_t^y$ and hence left discontinuous with non-zero probability. This contradicted the intrinsic property of an innovation process being a Wiener process having all the stochastic properties as the white Gaussian observation noise and its σ-algebra was continuous with probability one. Therefore, for our estimation problem, there was no $\mathcal{F}_t^y$-measurable innovation process, which explained why the optimal filter driven by the process I performed so poorly. The definition of I implies that it is not an innovation process only if the estimator C_{1t} does not exist in the strong sense as a solution to the set of equations. Hence the set of equations must be infinite, otherwise C_{1t} would exist as a strong solution.

Therefore, the optimal filter was physically unimplementable. Nevertheless its performance was investigated and compared to that of the linear optimal filter, the Kalman filter. When the Poisson intensity was low, the improvement over the Kalman filter was negligibly small. Due to the fact that the linear optimal schemes only utilized the first and second order statistics its performance was quite poor. With the result that the optimal filter performance was still unacceptable, we therefore concluded

100

that causal filtering was inherently unsuitable for the state estimation of the Wiener-Poisson driven systems. Since the linear optimal non-causal filter also exhibited poor performance, and the optimal non-causal schemes were physically unimplementable, we investigated nonlinear, noncausal suboptimal schemes.

In this work, a sequential suboptimal scheme (SSS) to the estimation problem was developed. Basically it detected and estimated the Poisson driving process using an MAP criterion, and then reconstructed the entire system state using a MMSE optimal noncausal filter. In the numerical procedures for optimization, the SSS utilized numerical algorithms that did not require differentiation of functions containing white processes and hence avoided the problem of ill-defined derivatives.

In Chapter 3, the estimation problem for the purely Poisson driven system was considered. The SSS was developed as a sequential suboptimal scheme. The asymptotic performance of the SSS as the Poisson intensity goes to zero was analytically derived and numerically simulated on a digital computer. Both results agreed extremely closely indicating that the SSS performs very well and that there exists a $\lambda^* > 0$, such that for the Poisson intensity $\lambda \in (0,\lambda^*]$ the SSS performs better than any causal filtering schemes and any non-causal linear schemes.

In Chapter 4, the SSS was developed for the Poisson-Wiener driven system. A sequential compensation scheme was included in the scheme to eliminate the error propagation effects which arose due to the sequential structure of the approach. It was found that the propagating errors which stemmed from the finalized estimates of the Poisson driving process did not strongly affect the immediate subsequent detection and estimation of incidents, but affected to a great extent the performance of the reconstruction of the entire system state. The sequential compensation strategy consisted of sequentially comparing a subinterval of the reconstructed state of the system to assess the propagating error and then compensating it with a first order approximation scheme to eliminate the propagating errors. In a numerical simulation study, the SSS was shown to perform basically as predicted by the analytical results. With the compensation strategy added, the overall performance improved in general very little, only that the error propagation effect was mostly eliminated.

In Chapter 5, we extended the state estimation problem of the original systems to a larger class of systems, namely a class of conditional Gaussian systems. This class included a wide variety of physical systems and random phenomena. However each member of this class gave rise to a MMSE optimal filter of finite dimensions for its pseudo-estimates, i.e. the estimates of the system state conditional on the information of the underlying Poisson driving process. This property enabled us to apply the principle of the SSS to the estimation problem. For the Poisson-Wiener driven linear system, which was a particular case of the conditional Gaussian systems, the MMSE optimal filters for its pseudo-estimates were linear. However, in general the MMSE optimal filters for such class of systems were non-linear.

For future considerations, the estimation problem is to extend to larger-class of stochastic systems. Suboptimal schemes then may have to be modified and extended to solve such problems and further improve performances. Finally, the control aspects of the estimation problem have to be examined to great extent since estimation and control dually arise naturally and interact in a coupling fashion that the two problems have to be investigated together.

Another aspect of the estimation problem that is important and practical in nature is when the stochastic system is excited by Wiener and Poisson white processes nonlinearly. Analysis of these systems have to be carried through to great depth. Approximation scheme as suboptimal strategy to the problem may have to be developed both analytically and numerically.

APPENDIX A: Doleans-Dade and Meyer Differential Rule for Discontinuous Semi-Martingales [2,19 et al.]

Let X_t be an n-dimensional vector-valued, discontinuous semi-martingale, and let ϕ be a twice differentiable function defined on R^n. Then the process $Q_t \triangleq \phi(X_t)$ is also a semi-martingale, such that

$$dQ_t = \phi_x'(X_{t-})dX_t + \frac{1}{2}\,tr[\phi_{xx}(X_{t-})d < X^c, X^c >_t]$$

$$+ d\{ \sum_{t_0 \le s \le t} [\phi(X_s) - \phi(X_{s-}) - \phi_x'(X_{s-})(X_s - X_{s-})]\}$$

Here ϕ_x is the gradient of the function ϕ, ϕ_{xx} its Jacobian, while X^c is the continuous part of X, and the summation is carried out over those values of s where X jumps. $< N,N >$ in the notation of Wong [20] is the increasing process associated with the continuous martingale or semi-martingale N. Note that this expression reduces to the ordinary Ito differential rule when the process is continuous.

APPENDIX B: Fundamental Filtering Theorem for White Gaussian Observation Noise [2-7]

Let Q_t, $t \geq t_0$ be the vector-valued semi-martingale defined by

$$dQ_t = R_t dt + dM_t \qquad t \geq t_0$$

where M is a vector-valued martingale with respect to a growing family of σ-fields $\mathcal{F}_t$, $t \geq t_0$ and R is a process that is adapted to $\mathcal{F}$. Let furthermore the vector valued process Y_t, $t \geq t_0$ be the semi-martingale defined by

$$dY_t = H_t dt + dW_t \qquad t \geq t_0$$

where H is another vector-valued process that is adapted to $\mathcal{F}$, and W is vector-valued Brownian motion and a martingale with respect to $\mathcal{F}$, such that $E(dW_t \, dW_t') = V(t)dt$. The prime denotes the transpose, and $V(t) > 0$, $t \geq t_0$. Define $\mathcal{F}_t^y$, $t \geq t_0$, as the growing family of σ-fields generated by Y, and assume the notation

$$\hat{\phi}_t \triangleq E(\phi_t | \mathcal{F}_t^y)$$

Then, the process $\hat{Q}_t$, $t \geq t_0$ satisfies the following equation,

$$d\hat{Q}_t = R_t dt + [\widehat{Q_t H_t'} - \hat{Q}_t \hat{H}_t' + \hat{C}_t] \, V^{-1}(t) \, [dY_t - \hat{H}_t dt], \qquad t \geq t_0$$

Here

$$C_t \triangleq \frac{d}{dt} < M^c, W >_t$$

M^c is the continuous part of the martingale M. If M_1 and M_2 are two vector valued continuous martingales or semi-martingales, $< M_1, M_2 >$ is a matrix stochastic process, the (i,j)-th element of which is given by $< M_{1i}, M_{2j} >$ which is given as follows

$$< M_{1i}, M_{2j} > \triangleq \frac{1}{2} \{< M_{1i} + M_{2j}, M_{1i} + M_{2j} > - < M_{1i}, M_{1i} > - < M_{2j}, M_{2j} >\}$$

in the notation of Wong.

APPENDIX C: Numerical Algorithm Map A = M *D [13,14 et al.]

For the algorithms in the form of an algorithmic map A as

$$A = M * D$$

roughly, after the map D determines a direction, the map M maximizes the objective function J in that direction. We choose M to be the Golden-Section Search D, the conjugate-direction method. The algorithmic convergence of A depends on the closedness of the point-to-set map A which follows from the closedness of M and D.

(1) Golden-Section Search (M)

The procedures start with a feasible interval $I = [\beta_0, \beta_1]$ containing the optimal point $\hat{\tau}$. Then, given an interval

$$I^k = [\beta_0^k, \beta_1^k]$$

containing $\hat{\tau}$, they determine $I^{k+1}[\beta_0^{k+1}, \beta_1^{k+1}]$ also containing $\hat{\tau}$, such that $I^{k+1} \subset I^k$. The Golden-Section search is for J concave. It requires use of the Fibonacci fractions

$$F_1 = \frac{3 - \sqrt{5}}{2} \approx 0.38$$

and

$$F_2 = \frac{\sqrt{5} - 1}{2} = 1 - F_1 \approx 0.62$$

Observe that $F_1 = (F_2)^2$. Given an interval $I = [\beta_0, \beta_1]$ of length $(\beta_1 - \beta_0)$ the next interval is selected as follows. Let

$$\beta_0' = \beta_0 + F_1(\beta_1 - \beta_0)$$

and

$$\beta_1' = \beta_0 + F_2(\beta_1 - \beta_0) = \beta_1 - F_1(\beta_1 - \beta_0)$$

be points on the interval

If $J(\beta_0') > J(\beta_1')$, the new interval is $[\beta_0, \beta_1']$

If $J(\beta_0') < J(\beta_1')$, the new interval is $[\beta_0', \beta_1]$

If $J(\beta_0') = J(\beta_1')$, the new interval is either $[\beta_0, \beta_1']$ or $[\beta_0', \beta_1]$.

We are assured by the concavity or (local) concavity that the optimal point $\hat{\tau}$ in on the resulting interval. It should be clear that the algorithm is closed. Before we state the conjugate-direction method (D), let us introduce a procedure of simple maximization, a spacer step.

(2) Spacer Steps

In optimizing $J(\underline{\tau})$, $\underline{\tau} \in E^N$, it is important to have a step such as that used above which, given a nonoptimal point $\underline{\tau}^k$, generates a point $\underline{\tau}^{k+1}$ for which

$$J(\underline{\tau}^{k+1}) > J(\underline{\tau}^k)$$

We call this step a space step. The spacer step must also possess one other property, should $\underline{\tau}^k$ be a solution, then the spacer step much indicate this fact. Thus the spacer step must determine whether $\underline{\tau}^k$ is optimal, and if $\underline{\tau}^k$ is not optimal, it must calculate a better point. For our purpose we choose the cyclic-coordinate-ascent method, as it avoids calculation of derivatives, and is quite simple to implement.

The cyclic-coordinate-ascent method optimizes cyclically in each of the coordinate directions. For example in E^3, the coordinate directions are

$$c_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad c_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ and } c_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

(3) Conjugate-Direction Method (D)

Basically, the conjugate-directions of a quadratic function p in E^N are a set of N directions in E^N such that optimization in each of these directions once only will either find the maximum or reveal p to be unbounded above. Also the order of the directions optimized is immaterial.

Since our objective function J is not quadratic in general, we shall nevertheless using this method for optimization of J exploiting a quadratic approximation to J.

(i) Generating conjugate directions for quadratic $p \in E^N$:

Let $d_1, \ldots d_r$, $r < N$ be conjugate. Suppose τ' and τ'' maximize p in two different manifolds containing $d_1, \ldots d_r$ and $p(\tau') < p(\tau'')$.

$$d_1, d_2 \ldots d_r, d_{r+1}$$

where $d_{r+1} = \tau'' - \tau'$

are conjugate. Thus, given r conjugate directions $r < N$, another conjugate direction is generated.

(ii) Conjugate-directions for $J \in E^N$: we shall develop an algorithm which attempts to calculate conjugate directions for J. The procedure can be thought of as calculating conjugate directions to a quadratic that approximates J. The quadratic approximation notion is thus being utilized.

N directions will be always used. Initially these start out being N arbitrary directions $d_1^1, d_2^1 \ldots d_N^1$. At the end of iteration one, the direction d_1^1 is deleted, the directions d_i^2, $i = 1,2 \ldots N-1$ are determined by $d_{i-1}^2 = d_i^1$ for $i = 2, \ldots N$, and by making d_N^2 a new direction conjugate to d_{N-1}^2. The procedure continues in the manner.

<u>Initial Step</u> An initial point $\underline{\tau}_1$ and N directions d_i^1, $i = 1,2 \ldots N$ are given with $d_N^1 \neq 0$. Fir $i = 1,2 \ldots N$ calculate $\underline{\tau}_{i+1}^0 \in M(\underline{\tau}_i^0, d_i^1)$, i.e. by the Golden-section search.
Define $\underline{\tau}_{N+2}^0 = \underline{\tau}_{N+1}^0$ and set $k = 1$.

<u>Iteration k</u> $= \underline{\tau}_{N+2}^{k-1}$ and d_i^k, $i = 1, \ldots N$ are given

(a) Spacer step on $\underline{\tau}_{N+2}^{k-1}$ yielding $\underline{\tau}_1^k$

(b) For $i = 1,2, \ldots N$ calculate $\underline{\tau}_{i+1}^k \in M(\tau_i^k, d_i^k)$

(c) Set

$$d_{N+1}^k = \underline{\tau}_{N+1}^k - \tau_{N+2}^{k-1}$$

and calculate

$$\underline{\tau}_{N+2}^k \in M(\underline{\tau}_{N+1}^k, d_{N+1}^k)$$

(d) Set $d_i^{k+1} = d_{i+1}^k$ $i = 1,2 \ldots N$

Go to iteration k with k+1 replacing k. Stop the procedure whenever the spacer step on $\underline{\tau}_{N+2}^{k-1}$ determines that $\underline{\tau}_{N+2}^{k-1}$ is a solution.

APPENDIX D: Conditional Gaussian Processes

Let an unobservable process $x = (x_t, \mathcal{F}_t)$ and an observable process $y = (y_t, \mathcal{F}_t)$ be processes of the diffusion type with

$$dx_t = [\mathcal{a}_1(t,y) + \mathcal{a}_2(t,y)y]dt + \mathcal{B}_1(t,y)dW_1(t) + \mathcal{B}_2(t,y)dW_2(t) \tag{D.1}$$

$$dy_t = [\mathcal{C}_1(t,y) + \mathcal{C}_2(t,y)x_t]dt + \mathcal{B}_1(t,y)dW_1(t) + \mathcal{B}_2(t,y)dW_2(t) \tag{D.2}$$

where same assumptions and definitions of the processes involved in the equations are made as Eqs. (5.3) to (5.10).

We have the following results [7,8,21-25]

(1) If the conditional distribution $F_{y_0}(a) = \underline{P}(x_0 \leq a | y_0)dy_s$ is Gaussian, then the random process $x = (x_t, \mathcal{F}_t)$ is conditional Gaussian almost surely. That is, for any t and $0 \leq t_0 < \underline{t}_1 < \cdots t_n \leq t$, the conditional distributions

$$F_{y_0 t}(\theta_0 \cdots \theta_n) = P(x_{t_0} \leq \theta_0 \cdots x_{t_n} \leq \theta_n | \mathcal{F}_t^y) \tag{D.3}$$

is Gaussian.

(2) Let $\tilde{y} = (\tilde{y}_t, \mathcal{F}_t)$ be another process satisfies

$$d\tilde{y}_t = \mathcal{B}_1(t,\tilde{y})dW_1(t) + \mathcal{B}_2(t,\tilde{y})dW_2(t) \tag{D.4}$$

and let μ_y and $\mu_{\tilde{y}}$ be two measures corresponding to the processes y and $\tilde{y}$ defined by Eqs. (D.2) and (D.4).

Then μ_y is equivalent to $\mu_{\tilde{y}}$ i.e.

$$\mu_y \sim \mu_{\tilde{y}} \tag{D.5}$$

and the density

$$\Psi_t(y) = \frac{d\mu_{\tilde{y}}}{d\mu_y}(t,y) \tag{D.6}$$

is given by

$$\Psi_t(y) = \exp\{ -\int_0^t [\frac{\mathcal{A}_1(s,y) + \mathcal{A}_2(s,y)\, m_s(y)}{\mathcal{B}_1^2(s,y) + \mathcal{B}_2^2(s,y)}]\, dy_s$$

$$+ \frac{1}{2}\int_0^t \frac{[\mathcal{A}_1(s,y) + \mathcal{A}_2(s,y)\, m_s(y)]^2}{\mathcal{B}_1^2(s,y) + \mathcal{B}_2^2(s,y)}\, ds\} \qquad (D.7)$$

where $m_s(y) = E\{x_s | \mathcal{F}_s^y\}$ (D.8)

(3) The estimate of x and its variance denoted by m_t and P_t satisfy

$$dm_t = [\mathcal{A}_1(t,y) + \mathcal{A}_2(t,y)m_t]\, dt$$

$$+ \frac{\mathcal{B} * \mathcal{B}(t,y) + P_t\mathcal{C}_2(t,y)}{\mathcal{B} * \mathcal{B}(t,y)}\{dy_t - [\mathcal{A}_1(t,y) - \mathcal{A}_2(t,y)m_t]\, dt\} \qquad (D.9)$$

$$\dot{P}_t = 2\mathcal{A}_2(t,y)P_t + \mathcal{B} * \mathcal{B}(t,y) - \frac{[\mathcal{B} * \mathcal{B}(t,y) + P_t\mathcal{A}_2(t,y)]^2}{\mathcal{B} * \mathcal{B}(t,y)} \qquad (D.10)$$

subject to the conditions

$$m_0 = E(x_0 | y_0) \qquad (D.11)$$

$$P_0 = E[(x_0 - m_0)^2 | y_0] \qquad (D.12)$$

(4) And the smoothed estimate

$$m_{t/T} = E(x_t | \mathcal{F}_T^y)$$

and its variance denoted by $P_{t/T}$ satisfy

$$d_t m_{t/T} = [\mathcal{A}_1(t,y) + \mathcal{A}_2(t,y)m_{t/T}]dt + \underline{\mathcal{B}}(t,y)P_t^{-1}[m_t - m_{t/T}]dt$$

$$+ \mathcal{B} * \mathcal{S}(t,y)(\mathcal{S} * \mathcal{S})^{-1}(t,y)\{dy_t - [\mathcal{A}_1(t,y) + \mathcal{A}_2(t,y)m_{t/T}]dt\} \quad \text{(D.13)}$$

$$\dot{P}_{t/T} = 2[\underline{\mathcal{A}}9t,y) + \underline{\mathcal{B}}(t,y)P_t^{-1}]\ P_{t/T} - \underline{\mathcal{B}}(t,y) \quad \text{(D.14)}$$

where

$$\mathcal{B} * \mathcal{S}(t,y) = \mathcal{B}_1(t,y)\mathcal{S}_1(t,y) + \mathcal{B}_2(t,y)\mathcal{S}_2(t,y)$$

and

$$\underline{\mathcal{A}}(t,y) = \mathcal{A}_1(t,y) - (\mathcal{B} * \mathcal{S})(t,y)(\mathcal{S} * \mathcal{S})^{-1}(t,y)\mathcal{A}_2(t,y)$$

$$\underline{\mathcal{B}}(t,y) = \mathcal{B} * \mathcal{B}(t,y) - (\mathcal{B} * \mathcal{S})^2(t,y)(\mathcal{S} * \mathcal{S})^{-1}(t,y)$$

REFERENCES

1. D. Snyder, Random Point Processes, Wiley, 1975.

2. H. Kwakernaak, "Filter for Systems Excited by Poisson White Noise," p. 463-492 in A. Bensonssan and J. L. Lions, Ed., Control Theory, Numerical Methods and Computer Systems Modelling, Lecture Notes in Economics and Math. Systems, Vol. 107, Springer-Verlag, Berlin, 1975.

3. H. Kwakernaak, "An Estimate of the Variance Reduction in Filtering for Linear Systems Excited by Poisson White Noise," Memorandum No. 56, Department of Appl. Math., Trente Univ. of Tech., Oct. 1974.

4. J. Clark, "Two Recent Results in Nonlinear Filtering Theory," IN: Recent Mathematical Developments in Control, edited by D. J. Bell, Academic Press, London, 1973.

5. M. Fujisaka, G. Kallianpur, and H. Kunita, "Stochastic Differential Equations for the Nonlinear Filtering Problem," Osaka J. Math., 9, 1, 19-40, 1972.

6. J. H. van Schuppen, "Estimation Theory for Continuous Time Processes, a Martingale Approach," Electronics Research Laboratory Memorandum No. ERL-M405, College of Engineering, Univ. of California, Berkeley, Calif., 1973.

7. R. S. Liptser and A. N. Shiryayev, Statistics of Random Processes I General Theory), Springer-Verlag, 1978.

8. R. S. Liptser and A. N. Shiryayev, Statistics of Random Processes II (Applications), Springer-Verlag, 1978.

9. T. Kailath, "An Innovations Approach to Least-Squares Estimation," IEEE Trans. on Automatic Control, Vol. AC-13, No. 6, Dec. 1968.

10. T. Kailath, "A General Likelihood-Ratio Formula for Random Signals in Gaussian Noise," IEEE Trans. on Info. Theory, Vol. IT-15, No. 3, May 1969.

11. Van Trees, Detection, Estimation and Modulation Theory, Part I, Wiley, 1968.

12. D. Snyder, "Information Processing for Observed Jump Processes," Information and Control, 22, No. 1, 69-78, Feb. 1973.

13. W. I. Zangwill, Nonlinear Programming: A Unified Approach, Prentice-Hall, N.Y., 1969.

14. D. G. Luenberger, Optimization by Vector Space Methods, Wiley, 1969.

15. A. Gelb, Applied Optimal Estimation, M.I.T. Press, 1974.

16. J. S. Meditch, Stochastic Optimal Linear Estimation and Control, McGraw-Hill, Inc., 1969.

17. I. I. Gihman and A. V. Skorokhod, Stochastic Differential Equations, Springer, New York, 1972.

18. A. V. Skorokhod, Studies in the Theory of Random Processes, Addison-Wesley, Reading, Mass., 1964.

19. W. B. Davenport and W. L. Root, Random Signals and Noise, McGraw-Hill Inc., 1958.

20. C. Doleans-Dade and P. Meyer, "Integrales Stochastiques par Rapport aux Martingales Locales," In: Seminaire de Probabilities IV, Lecture Notes in Math., Vol. 124, Springer, Berlin, 1970.

21. E. Wong, "Recent Progress in Stochastic Processes - a Survey," IEEE Trans. Information Theory, Vol. IT-19, No. 3, 1973.

22. R. S. Liptser, "On Filtering and Extrapolation of the Components of Diffusion Type Markov Processes," Teoria Verojatn. i Primenen, XII, No. 4, 1967.

23. R. S. Liptser and A. N. Shiryayev, "Nonlinear Filtering of Diffusion Type Markov Processes," Trudy matem. in-ta im. V. A. Steklova AN SSSR 104, 1968.

24. ____________,"Extrapolation of Multivariate Markov Processes with Incomplete Data," Teoria Verojatn, i Primenen. XIII, No. 1, 1968.

25. ____________, "On the Cases of Effective Solving the Problems of Optimal Nonlinear Filtering, Integrability, and Extrapolation," Teoria Verojatn. i Primenen. XIII, No. 3, 1968.

26. ____________, "Nonlinear Interpolation of the Components of Diffusion Type Markov Processes (Forward Equations, Effective Formulas)," Teoria verojatn. i Primenen. XIII, No. 4, 1968.

27. E. Wong, Stochastic Processes in Information and Dynamical Systems, McGraw-Hill Inc., 1971.

28. M. A. Nauta, "Recursive Estimation of Impulse Driven Process," Master Thesis in Department of Appl. Math., Trente Univ. of Tech., Netherlands, June 1977.

29. J. Doob, Stochastic Processes, Wiley, New York, 1953.

30. M. Athans and P. L. Falb, Optimal Control, McGraw-Hill Inc., 1966.

31. R. Isaacs, Differential Games, Wiley Inc., 1965.

32. A. Blaquiere, F. Gerard and G. Leitmann, Quantitative and Qualitative Games, Academic Press, Inc., New York, 1969

33. R. Kalman and R. Bucy, "New Results in Linear Filtering and Prediction Theory," ASME J. Basic Engr., March 1961.

34. D. Snyder, "Optimum Linear Filtering of an Integrated Signal in White Noise," IEEE Trans. Aerospace Electron. Systems, AES-2, No. 2, 231-232, March 1966.

35. A. B. Baggeroer, "A State Variable Technique for the Solution of Fredholm Integral Equations," Proc. 1967 Intl. Conf. Communications, Minneapolis, Minn., June 12-14, 1967.

36. A. E. Bryson and M. Frazier, "Smoothing for Linear and Nonlinear Dynamic Systems," USAF Tech. Rpt. ASD-TDR-63-117, Feb. 1963.

37. A. Bensonssan, Filtrage Optimal des Systemes Lineaires, Dunod, Paris, 1971.

38. R. Bucy, "Nonlinear Filtering Theory," IEEE Trans. Automatic Control AC-10, 1965.

39. I. A. Bolshakov and V. G. Repin, "Problems of Nonlinear Filtering," Automatika i telemekhanika XXII, 4, 1961.

40. R. Bucy and P. D. Joseph, Filtering for Stochastic Processes with Application to Guidance, Interscience, N.Y., 1968.

41. N. Wiener, Extrapolation, Interpolation and Smoothing of Stationary Time Series, Wiley Inc., N.Y., 1949.

42. J. Clark, "The Representation of Functionals of Brownian Motion by Stochastic Integrals," Ann. Math. Statist., Vol. 41, pp. 1282-1295, 1970.

43. C. Doleans, "Variation Quadratique des Martingales Continues a Droite," Ann. Math. Statist., Vol. 40, pp. 284-289, 1969.

44. C. Doleans-Dade, "Quelques Application de la Formula de Changement de Variables pour les Semimartingales," Z. Wahrscheinlichkeits, verw. Geb., Vol. 16, pp. 181-194, 1970.

45. P. A. Meyer, "Square Integrable Martingales, a survey," in Martingales (Lecture Notes in Mathematics), Vol. 190, Berlin-Springer, 1971, pp. 32-37.

46. __________, "Non-Square Integrable Martingales, etc.," in Martingales, Vol. 190, Berlin-Springer, 1971, pp. 38-43.

47. E. Wong, "Representation of Martingales, Quadratic Variation and Applications," SIAM J. Contr., Vol. 9, 1971.

48. J. L. Zeman, "Approximate Analysis of Stochastic Processes in Mechanics," Udine Courses and Lectures No. 95, Springer, Berlin, 1971.

49. S. Au and A. H. Hadded, "Near-Optimal Estimation Detection Scheme for Poisson Driven Processes," Intl. Symposium in Info. Theory, Oct. 1977, Ithaca, N.Y.

50. S. Au and A. H. Haddad, "Suboptimal Sequential Estimation-Detection Scheme for Poisson Driven Linear Systems," Information Science (to appear).

51. I. Rubin, "Regular Point Processes and Their Detection," IEEE Trans. on Info. Theory, Vol. 17-18, No. 5, Sept. 1972.

52. A. J. Lawrence, "Selective Interaction of a Stationary Point Processes and a Renewal Process," J. Appl. Prob., F, pp. 483-489, 1973.

53. ______________, "Selective Interaction of a Poisson and Renewal Process," J. Appl. Prob. 8, pp. 170-183, March 1971.

54. A. Baggeroer, "Maximum Aposteriori Interval Estimation," WESCON, Paper No. 7/8, Aug. 1966.

55. V. Ya. Lumel'Skii, "Algorithm for Detecting the Time of Changes of Properties of a Random Process," Translated from Automatika i Telemekhanika, No. 10, Oct. 1972.

56. L. A. Tel'kanis, "Application of the Optimal Bayesian Learning Algorithm to Determine the Time Instants of Property Changes of Random Signals," Translated from Automatika i Telemekhanika, No. 6, June 1969.

57. L. I. Borodkin and V. V. Mottl, "Algorithm for Finding the Jump Times of Random Process Equation Parameters," Translated from Automatika i Telemakhanika, No. 6, 1976.

58. I. Bar-David, "Communication Under the Poisson Regime," IEEE Trans. Inform. Theory, Vol. IT-15, pp. 31-37, Jan. 1969.

59. I. Bar-David and A. Nemirovsky, "Level Crossings of Nondifferentiable Shot Processes," IEEE Trans. Inform. Theory, Vol. IT-18, pp. 27-34, Jan. 1972.

60. F. J. Beutler and O. A. Z. Leneman, "The Spectral Analysis of Impulse Processes," Inform. Contr., Vol. 12, pp. 236-258, 1968.

61. __________, "On the Statistics of Random Pulse Processes," Inform. Contr., Vol. 18, pp. 326-341, 1971.

62. M. Brown, "Discrimination of Poisson Processes," Ann. Math. Statist., Vol. 42, pp. 773-776, 1971.

63. S. Karp and R. M. Gagliardi, "On the Representation of a Continuous Stochastic Intensity by Poisson Shot Noise," IEEE Trans. Inform. Theory, Vol. IT-16, pp. 142-147, Mar. 1970.

64. S. Karp and J. R. Clark, "Photon Counting: A Problem in Classical Noise Theory," IEEE Trans. Inform. Theory, Vol. IT-16, pp. 672-680, Nov. 1970.

65. O. Macchi, "Stochastic Point Processes and Multicoincidences," IEEE Trans. Inform. Theory, Vol. IT-17, pp. 2-7, Jan. 1971.

66. D. L. Snyder, "Filtering and Detection for Doubly Stochastic Poisson Processes," IEEE Trans. Inform. Theory, Vol. IT-18, pp. 91-102, Jan. 1972.

67. ___________, "Smoothing for Doubly Stochastic Poisson Processes," IEEE Trans. Inform. Theory, Vol. IT-18, pp. 558-562, Sept. 1972.

68. D. Middleton and R. Esposito, "Simultaneous Optimum Detection and Estimation of Signals in Noise," IEEE Trans. on Inf. Theory, IT-14, pp. 434-444, May 1968.

69. A. Fredriksen, D. Middleton and D. Vandelind, "Simultaneous Signal Detection and Estimation Under Multiple Hypothesis," IEEE Trans. on Inf. Theory, 17-18, pp. 607-614, Sept. 1972.

70. D. W. Kelsey and A. H. Haddad, "A Note on Detection for Point Minimax Detection Estimation Schemes," IEEE Trans. on Automatic Control, AL-18, pp. 558-559, Oct. 1973.

71. R. A. Padilla and A. H. Haddad, "On the Estimation of Uncertain Signals Using an Estimation-Detection Scheme," IEEE Trans. on Automatic Control, AC-21, Aug. 1976.

72. D. W. Kelsey, "On Joint Prediction-Detection Systems," Rep. R-581, CSL, University of Illinois at Urbana-Champaign, Urbana, IL, Aug. 1972.

73. J. Clark and E. V. Hoversten, "Poisson Process as a Statistical Model for Photodetectors Excited by Gaussian Light," Quarterly Progress Report No. 98, Research Laboratory of Electronics, M.I.T., Camb., Mass., July 1970.

74. J. Clark, "Estimation for Poisson Processes with Applications in Optical Communication," Ph.D. Thesis, Dept. of Electrical Engineering, M.I.T., Camb., Mass., Sept. 1971.

75. K. P. Dunn and I. B. Rhodes, "A Generalized Representation Theorem with Applications to Estimation and Hypothesis Testing," Proceedings of the 11th Annual Allerton Conf. on Circuit and System Theory, Univ. of Ill., Urbana, Oct. 1973.

76. O. Macchi and B. Picinbono, "Estimation and Detection of Weak Optical Signals," IEEE Trans. Info. Theory, IT-18, No. 5, Sept. 1972.

VITA

Samuel Poriza Au was born in Kowloon, Hong Kong on November 27, 1953. He received his Science Bachelor degree in Electrical Engineering from the Massachusetts Institute of Technology, Cambridge in August 1975, and his Master of Science degree in Electrical Engineering from the University of Illinois, Urbana-Champaign in August 1976.

From August 1975 to May 1976, he was a research assistant in the Ultrasonic Laboratory of the University of Illinois. During part of that period, Mr. Au has also been a graduate teaching assistant with the Department of Electrical Engineering. Since August 1976, he has been a research assistant at the Coordinated Science Laboratory, University of Illinois. His interests are in the estimation and detection theory in control systems and communication systems.