SRC TR 87-208

MS. 87-10

Filtering Results For A Joint Control-Decoding Problem In Optical Communications

by

J.A. Gubner

FILTERING RESULTS
FOR A
JOINT CONTROL-DECODING PROBLEM
IN OPTICAL COMMUNICATIONS

by

John A. Gubner

Thesis submitted to the faculty of the Graduate School

of the University of Maryland in partial fulfillment

of the requirements for the degree of

Master of Science

1985

APPROVAL SHEET

Title of Thesis: Filtering Results for a Joint Control-Decoding Problem in Optical Communications

Name of Candidate: John A. Gubner

Master of Science, 1985

Thesis and Abstract Approved: Prakash Narayan

Prakash Narayan

Assistant Professor

Electrical Engineering Department

Date Approved: August 12, 1985

CURRICULUM VITAE

Name: John A. Gubner

Permanent address: 8433 Carrollton Pkwy, New Carrollton, Maryland 20784.

Degree and date to be conferred: M.S. 1985.

Date of birth: June 11, 1961.

Place of birth: Washington, DC

Secondary education: Parkdale Senior High School, Riverdale, Maryland, 1979.

Collegiate institutions attended	Dates	Degree	Date of Degree
University of Maryland	1983-1985	M.S.	1985
University of Maryland	1979-1983	B.S.	1983

Major: Electrical Engineering.

Professional publications:

John A. Gubner, "A Joint Control-Decoding Problem in Optical Communications", in the *Procedings of the Nineteenth Annual Conference on Information Sciences and Systems,* The Johns Hopkins University, Baltimore, Maryland, 1985.

ABSTRACT

Title of Thesis: Filtering Results for a Joint Control-Decoding Problem in Optical Communications

John A. Gubner, Master of Science, 1985

Thesis directed by: Prakash Narayan

Assistant Professor

Electrical Engineering Department

We consider a doubly-stochastic time-space Poisson-process model for a direct-detection receiver in an optical communication system. Using a Bayesian decision approach to specify the design of the receiver, we encounter a likelihood ratio which, in general, is a function of a certain conditional expectation. We show how the design of the receiver leads to what we call the *Joint Control-Decoding Problem*. In a degenerate case, we completely solve the Joint Control-Decoding Problem and compute the conditional expectation mentioned. In the general case, we cannot compute the conditional expectation mentioned above, and hence, cannot proceed to solve the Joint Control-Decoding Problem; however, in order to gain insight into the general filtering problem given time-space point-process observations, we attempt to apply known filtering methods to the computation of a related conditional expectation. Finally, we consider linear estimates to substitute for the needed conditional expectation. In the case of a deterministic control, we reduce the linear estimation problem to the solution of a Fredholm integral equation. In the final chapter, we present a discrete-time version of our model which we hope will render the corresponding Discrete-Time Joint Control-Decoding Problem more tractable.

ACKNOWLEDGEMENTS

I wish to express my deep gratitude to my advisor, Professor Prakash Narayan, for introducing me to the *Joint Control-Decoding Problem.* Professor Narayan was always available to discuss my research and to offer suggestions and encouragement when I tackled formidable calculations.

I am also grateful to the other members of the faculty with whom I discussed some of the problems I encountered in the course of my research. In particular, I thank the other two members of my thesis committee, Professors A. M. Makowski and P. S. Krishnaprasad.

I thank also the Naval Research Laboratory which, through the University of Maryland, supported my graduate studies. In particular, I thank Mr. P. D. Stilwell of the Naval Research Laboratory, who introduced me to applications of optical systems in several different areas.

Finally, I thank my parents for their support, encouragement, and understanding throughout my education.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

I. An Optical Communication System

Consider the optical communication system shown in Figure 1. Here, an optical source transmits binary data by suitably modulating the intensity of a laser beam directed at the receiver. Let T be a fixed, positive constant. At time $t = 0$, $m(0)$ arrives at the optical source. During the interval $[0, T]$, the optical source modulates the intensity of the laser beam in accordance with whether $m(0)$ represents a "0" or a "1." Also during this time interval, the receiver observes the incident beam. At time $t = T$, the receiver produces its decision $\hat{m}(T)$ as to whether a "0" or a "1" was sent. This process is repeated on the intervals $[T, 2T]$, $[2T, 3T]$, and so on.

We assume that the receiver is a direct-detection device, by which we mean that the incident laser beam proceeds more or less unaltered to a photodetector, and that the output of that photodetector is then proportional to the intensity of the incident beam. This is in contrast to a heterodyne system in which the receiver generates a local laser beam to coincide physically with the incident laser beam. In this case the output of the photodetector oscillates at the frequency difference between the incident beam and the local beam (see Harger [1], pp. 1-6).

Refering again to Figure 1, if the optical source or the receiver is subject to vibrations, or if the laser beam passes through a turbulent atmosphere, the spot of laser light at the receiver will move randomly over the receiver's optical detection equipment [2]. If no action is taken, the receiver may eventually fail to intercept the laser beam completely, or even fail to intercept the beam at all. Hence, some type of tracking and control system must be incorporated into the receiver's structure in order to ensure continuous interception of the beam.

The major subsystems of a direct-detection receiver are shown in Figure 2. Figure 2 is adapted from Figure 1 in [2], and hence, so is the following discussion.

> The beam-acquisistion and fine-tracking equipment includes "... the collection optics, such as a telescope, and any spatial or temporal filtering used to reduce the effects of background radiation... [This equipment also includes] the elements used to effect tracking to maintain optical alignment. These might include servo-driven gimbal arrangements, bender bimorphs, and Risley prisms, all of which are electromechanical devices... These devices are assumed to be described by linear, stochastic differential-equations. This does not seem to be a restrictive assumption for the limited range of operation associated with the 'fine-tracking' mode of operation that we shall emphasize. Also, it is consistent with the linear, ordinary differential-equation models found in textbooks for most electromechanical devices." [2]
>
> Initially, the beam-acquisition controller is used to direct the beam-acquisition and fine-tracking equipment to acquire the laser beam. This ensures that when the fine-tracking controller-decoder is activated, laser light will be passing through the optical processor and photodetector subsystem.
>
> The purpose of the optical processor "is to convert variations in the angle of arrival of the incident beam of light into variations in the position of a spot of light on the active surface of the photodetector. This can be accomplished with a lens [system]." [2]
>
> We assume that the receiver has initially acquired the beam and that laser light is striking the photodetector surface at the receiver. We also assume that the fine-tracking controller-decoder (hereafter refered to as simply the controller-decoder) has been engaged and is directing the beam-acquisition and fine-tracking equipment (hereafter refered to as simply the fine-tracking equipment) to compensate for the random motions of the spot of laser light at the photodetector [2].

Our primary concern is the design of the controller-decoder. This subsystem generates both the estimate, $\hat{m}(T)$, of the message being sent, $m(0)$, as well as the control signal, $\{ u_t , 0 \leq t \leq T \}$, used to drive the fine-tracking equipment. Both controller subsystems are driven by the observations, G_t, which are produced by the photodetector. The observations G_t consist of the times and locations at which photoelectrons are generated in the receiver's photodetector up to time t. Based on the input $\{ G_t , 0 \leq t \leq T \}$, the controller-decoder must generate the two outputs, $\hat{m}(T)$ and $\{ u_t , 0 \leq t \leq T \}$, in an optimal way.

In [2] the controller-decoder is split into two sub-units which operate independently (see Figure 3), a fine-tracking controller subsystem and a communication subsystem. In [2, 3] the fine-tracking controller generates a signal $\{ u_t , 0 \leq t \leq T \}$ which will minimize a certain quadratic cost functional (see equation (2.29) and the *Remark* at the end of Chapter 2). In [2] it is assumed that the photodetector has an infinite photosensitive surface which is identified with the Euclidean plane, $\mathbb{R}^2$. Under this assumption, a formula for u_t is derived in [3]. Also in [2], a communication subsystem is considered in which maximum-likelihood estimates of message sequences are generated. Details are found in [4].

The goal in designing a receiver is to minimize the probability of a decoding error. It is not at all clear that in the system alluded to in the previous paragraph, generating $\{ u_t, 0 \leq t \leq T \}$ to satisfy the criterion mentioned above, will minimize the probability of a decoding error. Therefore, we shall consider the problem of designing a controller-decoder in which it is not assumed *a priori* that the fine-tracking controller and the communication subsystem will operate independently. This leads to the *Joint Control-Decoding Problem* we discuss in Section II. Further, we consider not only the probabilistic implications of the assumption that the photodetector has an infinite photosensitive surface, but we also examine the difficulties encountered when this assumption is dropped.

II. The Joint Control-Decoding Problem

As indicated in Section I, the design of the controller-decoder is related to the solution of an optimization problem. In Section IV of Chapter 2, we let $\{ u_t, 0 \leq t \leq T \}$ be any process generated from the observations $\{ G_t, 0 \leq t \leq T \}$ in a non-anticipative way. Using a Bayesian decision approach, we require a decoding rule which will minimize the probability of a decoding error, p_e. This implies decoding using the likelihood ratio test

$$L_T \underset{H_0}{\overset{H_1}{\gtrless}} 1,$$

where L_T is the likelihood ratio function, H_1 is the hypothesis that $m(0) = 1$, H_0 is the hypothesis that $m(0) = 0$, and 1 is the threshold (we use the minimum probability of error cost criterion, and assume equally likely hypotheses). Now, the probability of a decoding error is given by

$$p_e = \frac{1}{2} \left[\mathbf{P}(L_T < 1 \mid m(0) = 1) + \mathbf{P}(L_T > 1 \mid m(0) = 0) \right] .$$

We show that the probability law of the likelihood ratio function, L_T, and hence the probability of a decoding error, p_e, will depend, in general, on the manner in which the controller-decoder generates $\{ u_t , 0 \le t \le T \}$ from the observations $\{ G_t , 0 \le t \le T \}$. We use the term *Joint Control-Decoding Problem* to refer to the task of designing a controller-decoder to generate a process $\{ u_t , 0 \le t \le T \}$ which will minimize the probability of a decoding error.

Unfortunately, in the general case, it is not possible to write down explicitly all of the quantities appearing in the likelihood ratio. In particular, a certain conditional expectation which is needed to write down the likelihood ratio usually cannot be computed.

In this thesis we consider the Joint Control-Decoding Problem, and in particular, the computation of the conditional expectation which appears in the likelihood ratio. This conditional expectation, if computed, would be a functional of the photodetector output and of the control signal $\{ u_t , 0 \le t \le T \}$. The likelihood ratio could then be computed explicitly so that a decoder could be implemented, even for a suboptimal control signal. The long-range goal, of course, is to obtain the probability of error corresponding to the likelihood ratio test, as a function of the control, and then to choose a control which will minimize the probability of error.

III. Thesis Outline

In Chapter 2 we motivate and then explicitly describe a mathematical model for the Joint Control-Decoding Problem.

In Chapter 3 we consider the Joint Control-Decoding Problem under the assumption that the photodetector has an infinite photosensitive surface. We prove that "all controls are optimal" in the sense that the probability of a decoding error is not a function of the control signal. We show how this implies that the Joint Control-Decoding Problem degenerates into a simple detection problem with Poisson-process observations. We also use the work of Rhodes and Snyder [3], together with Theorem A1 of Appendix A, to compute the conditional expectation mentioned in Section II.

In Chapter 4 we drop the assumption that the photodetector has an infinite photosensitive surface. We consider a related time point process, associated with which, there is a conditional expectation similar to that appearing in the likelihood ratio. We attempt to apply the work of Boel and Benes [5] on time point processes to gain some insight into the general filtering problem given time-space point-process observations.

The conditional expectation for which we search is, of course, a functional of the observations. In Chapter 5 we restrict our attention to a certain class of *linear* functionals of the observations.

In Chapter 6 we discuss our results and point out directions for future research.

CHAPTER 2

MATHEMATICAL FORMULATION OF THE JOINT CONTROL-DECODING PROBLEM

I. Introduction

The receiver model we outline below in Sections II and III is essentially that given in [2], and we refer the reader to that paper for a detailed description and justification of the model. In Section IV we describe our model for the controller-decoder in which it is not assumed *a priori* that the fine-tracking will be carried out independently from the message decoding. Consequently, the criterion by which we optimize our control law is quite different from that discussed in Rhodes and Snyder [3], as we will indicate at the end of Section IV.

II. Photodetectors, Point Processes, and Gaussian Beams

As in [2], we identify points on the receiver's photodetector surface with the Euclidean plane, $\mathbb{R}^2$. If a laser beam with light intensity-profile $I(t, r)$, where $t \in [0,\infty)$ and $r \in \mathbb{R}^2$, strikes the photodetector, then the occurrence of photoelectrons can be modeled as a time-space point process whose probabilistic intensity process $\lambda(t, r)$ is proportional to $I(t, r)$. (To be more accurate, we should add a deterministic term, say $d_0(t, r)$, to $\lambda(t, r)$ to account for the dark current. However, since $d_0(t, r)$ is deterministic, there is no loss of generality in our probabilistic calculations if we assume $d_0(t, r)$ to be identically zero).

Following [2], we assume that the intensity profile of the laser beam is Gaussian. Consequently, if the laser beam is centered at the origin,

$$I(t, r) = a(t)\exp[-\frac{1}{2}r' \; R(t)^{-1} r] , \tag{2.1}$$

where $a(t)$ modulates the intensity of the laser beam, $R(t)$ is a 2×2 positive definite matrix describing the shape of the spot of laser light, and $'$ denotes transpose [2]. The maximum value of $I(t, r)$ is, of course, $a(t)$. Later we shall assume that $a(t)$ is one of two known functions selected by the optical source in accordance with whether a "1" or a "0" is to be sent.

If the receiver or the optical source is subject to vibrations, or if the laser beam passes through a turbulent atmosphere, the spot of laser light will move randomly over the photodetector surface [2]. Let $y_t^m \in \mathbb{R}^2$ denote the random position of the center of the spot of laser light at time t. Hence, $I(t, r)$ is now given by

$$I(t, r) = a(t)\exp[-\frac{1}{2}(r - y_t^m)' \, R(t)^{-1} (r - y_t^m)] . \tag{2.2}$$

In general, $a(t)$ will also be a random process due to the effects mentioned above. Later, in order to simplify the analysis, we shall assume that $a(t)$ is strictly positive, deterministic, and known.

As shown in Figure 2, the receiver incorporates fine-tracking equipment to try to prevent the spot of laser light from drifting off the edge of the photodetector. We model the influence of the fine-tracking equipment on where the incident beam strikes the photodetector in the following way. Let $I(t, r)$ be given more generally by

$$I(t, r) = a(t)\exp[-\frac{1}{2}(r - [y_t^m - y_t^p])' \, R(t)^{-1} (r - [y_t^m - y_t^p])] . \tag{2.3}$$

Here, $-y_t^p$ denotes the position where the center of the spot of laser light would fall if y_t^m were identically zero. The vector $y_t^p \in \mathbb{R}^2$ is completely determined by the (stochastic) state of the receiver's fine-tracking equipment. From equation (2.3), we see that if the receiver knew y_t^m, and could direct the fine-tracking equipment to operate so as to cause y_t^p to be equal to y_t^m, the spot of laser light would be centered at the origin.

Continuing to follow the model in [2], we assume that

$$y_t^m = H^m(t) x_t^m , \tag{2.4}$$

and

$$y_t^p = H^p(t)x_t^p \ , \tag{2.5}$$

where

$$dx_t^m = F^m(t)x_t^m dt \ + \ V^m(t)dv_t^m \ ; \quad x_0^m = X^m \ , \tag{2.6}$$

and

$$dx_t^p = F^p(t)x_t^p dt \ + \ G^p(t)u_t\, dt \ + \ V^p(t)dv_t^p \ ; \quad x_0^p = X^p \ . \tag{2.7}$$

In the above equations, H^m, H^p, F^m, F^p, V^m, V^p, and G^p are known matrices with appropriate dimensions. The process $\{ u_t \, , \, t \geq 0 \}$ is the control law which drives the fine-tracking equipment. We take $\{ v_t^m, \, t \geq 0 \}$ and $\{ v_t^p, \, t \geq 0 \}$ to be standard Wiener processes. The initial conditions X^m and X^p are assumed to be normal with known means, m^m and m^p, and known covariances, S^m and S^p, respectively. The four quantities $\{ v_t^m, \, t \geq 0 \}$, $\{ v_t^p, \, t \geq 0 \}$, X^m, and X^p are assumed to be statistically independent. To simplify the notation, let

$$H(t) = [\, H^m(t) \quad -H^p(t) \,] \ , \tag{2.8}$$

$$x_t = \begin{bmatrix} x_t^m \\ x_t^p \end{bmatrix} , \ \text{and} \quad v_t = \begin{bmatrix} v_t^m \\ v_t^p \end{bmatrix} . \tag{2.9}$$

If we let F, V, G, m, and S be the obvious block matrices, then we can rewrite equations (2.3), (2.6), and (2.7) as

$$I(t\, , \, r) = a(t)\exp[\, -\frac{1}{2}(r - H(t)x_t)' \ R(t)^{-1}\,(r - H(t)x_t)\,] \ , \tag{2.10}$$

where

$$dx_t \ = F(t)x_t\, dt \ + \ G(t)u_t\, dt \ + \ V(t)dv_t \ ; \quad x_0 = X \ , \tag{2.11}$$

and X is normal with mean m and covariance S.

With the above motivation and equations in mind, we next make a precise statement of our probabilistic setting.

III. Probabilistic Setting

Let $(\Omega, F, \mathbf{P})$ be a probability space. On this probability space, let X be normal with known mean, m, and known, positive definite covariance, S. Let $\{ v_t, t \geq 0 \}$ be a standard Wiener process independent of X. We let the n-dimensional process $\{ x_t, t \geq 0 \}$ be the solution to the Ito stochastic differential equation

$$dx_t = F(t)x_t\, dt + G(t)u_t\, dt + V(t)dv_t \;; \quad x_0 = X \,. \tag{2.12}$$

Here F, G, and V are known matrices with appropriate dimensions. We also assume that F, G, and V are piecewise-continuous so that a unique solution of (2.12) exists (see Davis [6], pp. 108-111, especially Theorem 4.2.4 on p. 111), at least when $\{ u_t, t \geq 0 \}$ is deterministic. We shall further require that $\{ u_t, t \geq 0 \}$ be predictable with respect to $\{ G_t, t \geq 0 \}$, where G_t is defined below. (This is trivially true for any deterministic control; it is also true for any left-continuous process adapted to $\{ G_t, t \geq 0 \}$).

Let B^2 denote the Borel subsets of $\mathbb{R}^2$. Next, if I is any interval of $\mathbb{R}$, let $B(I)$ denote the Borel subsets of I. We define $B(I) \otimes B^2$ to be the smallest σ-field containing all sets of the form $E \times A$, such that $E \in B(I)$ and $A \in B^2$. The occurrence of photoelectrons at the photodetector is modeled as a time-space point process

$$\mathbf{N}^0 = \{ N(B) : B \in B(0,\infty) \otimes B^2 \} \,. \tag{2.13}$$

Sometimes, $\mathbf{N}^0$ is called a random point field or a random measure. Here, this means that to each $B \in B(0,\infty) \otimes B^2$, we associate a nonnegative, integer-valued random variable, $N(B) = N(\omega, B)$ (we will usually suppress the argument ω). In addition, for each $\omega \in \Omega$, $N(\omega, \bullet)$ is an integer-valued measure on $B(0,\infty) \otimes B^2$. The motivation for this abstract model is the following. Suppose C is a subset of $\mathbb{R}^2$, say a square centered at the origin. Let $0 < t_1 < t_2 < \infty$. Then $N((t_1, t_2] \times C)$ is interpreted as the number of photoelectrons observed in the region C during the time interval $(t_1, t_2]$.

We now place further restrictions on the random field $\mathbf{N}^0$. In order to do this, we first define the σ-fields $\{X_t\}$ by

$$X_t = \begin{cases} \sigma\{ x_s , 0 \le s < \infty \} ; & t = 0 \\ \\ \sigma\{ N(B), B \in B(0,t] \otimes B^2 ; \ x_s , 0 \le s < \infty \} ; & t > 0 \end{cases} \tag{2.14}$$

We shall assume, as in [2], that $\mathbf{N}^0$ is a $\{X_t\}$-doubly-stochastic, time-space Poisson process, with X_0-measurable intensity (see Bremaud [7], pp. 21-23 and 233-238)

$$\lambda_i(t, r, x_t) = \mu_i(t) \exp[-\frac{1}{2}(r - H(t)x_t)' \ R(t)^{-1} \ (r - H(t)x_t)] , \tag{2.15}$$

where

$t \in [0,\infty)$, $r \in \mathbb{R}^2$, and x_t is defined by equation (2.12),

$i = 0$ or 1 depending on the message,

μ_i, H, and R are deterministic and known, with

$\mu_i : [0,\infty) \to (0,\infty)$, $H : [0,\infty) \to \mathbb{R}^{2\times n}$, and

$R : [0,\infty) \to \mathbb{R}^{2\times 2}$ is positive definite.

This means that for each $t \ge 0$, the process

$$\mathbf{N}^t \ \triangleq \ \{ N(B) : B \in B(t, \infty) \otimes B^2 \} \tag{2.16}$$

is a Poisson random field under the measure $\mathbf{P}(\bullet \mid X_t)$, with rate $\lambda_i(s, r, x_s)$, where $s \in (t, \infty)$, and $r \in \mathbb{R}^2$. This implies the following. First, for $B \in B(0,\infty) \otimes B^2$, let $\Lambda(B) \triangleq \int_B \lambda_i(\tau, \rho, x_\tau)\, d\rho\, d\tau$. Then if $B \in B(t, \infty) \otimes B^2$ and n is an arbitrary, nonnegative integer,

$$\mathbf{P}(N(B) = n \mid X_t) = \frac{\Lambda(B)^n}{n!} e^{-\Lambda(B)} ,$$

and hence, for $\theta \in \mathbb{R}$,

$$\mathbf{E} [e^{j\theta N(B)} \mid X_t] = \exp[(e^{j\theta} - 1) \Lambda(B)] .$$

The second implication is that if B_1 and B_2 are disjoint sets in $B(t\,,\infty)\otimes B^2$, then the random variables $N(\,B_1\,)$ and $N(\,B_2\,)$ are independent under the measure $\mathbf{P}\,(\,\bullet\mid X_t\,)$.

At this point, the model appears to assume that the photodetector has a photosensitive surface equal to $\mathbb{R}^2$. We let F_t represent the history of photoevents in $\mathbb{R}^2$ up to and including time t. More precisely, let F_0 be the trivial σ-field, and for $t>0$, set

$$F_t = \sigma\{\, N(B) : B \in B(0,t]\otimes B^2 \,\}\,. \tag{2.17}$$

Note: We can now write $X_t = F_t \vee X_0$ for all $t \geq 0$.

To model the fact that photodetectors have a finite photosensitive area, we introduce the σ-fields $\{G_t\}$. Take G_0 to be the trivial σ-field, and for $t>0$, set

$$G_t = \sigma\{\, N(\,B \cap \{\,(0,\infty)\times D\,\}\,) : B \in B(0,t]\otimes B^2 \,\}\,. \tag{2.18}$$

In (2.18), D is a Borel subset of $\mathbb{R}^2$, typically a rectangle or a circle centered at the origin; D represents the actual photosensitive surface of the photodetector, and G_t represents the observations available from that region only, up to time t. In Chapter 3 we will set $D = \mathbb{R}^2$, in which case, $G_t = F_t$. The notion of the σ-fields $\{G_t\}$ is not found in [2, 3, 8]. However, conditioning on G_t seems to be more difficult than conditioning on F_t, and this difficulty is alluded to in [8]. The difficulty seems to arise from the fact that $\int_D \lambda_i(t\,,r\,,x_t)\,dr$ is, in general, a random process, while $\int_{\mathbb{R}^2} \lambda_i(t\,,r\,,x_t)\,dr = 2\pi\,\mu_i(t)\sqrt{\det R(t)}$ is a deterministic function.

IV. The Controller-Decoder

Let T be as in Chapter 1. Let $\{\,u_t\,,\,0\leq t\leq T\,\}$ be any process predictable with respect to $\{\,G_t\,,\,0\leq t\leq T\,\}$, with otherwise arbitrary distributions. Suppose that our only task were to make a decision, based on our observations $\{\,G_t\,,\,0\leq t\leq T\,\}$, as to whether a "1" or a "0" was being sent. Then we would be faced with the standard binary-hypothesis testing problem. (Here we assume equally likely hypotheses and use the minimum probability of error cost criterion). The optimum decision rule for this problem is the likelihood ratio test (see Snyder [9], section 2.5), which we reproduce later in equations (2.21) and (2.22). Before writing down the likelihood ratio, we need to define the following quantity:

$$\hat{\lambda}_i(t, r) \triangleq \mathbf{E}[\lambda_i(t, r, x_t) \mid G_t]; \quad i = 0, 1. \tag{2.19}$$

Let N_t be the number of photoevents that have occurred up to and including time t in the region D. More precisely, let $N_0 \equiv 0$, and for $t > 0$, set

$$N_t = N((0,t] \times D). \tag{2.20}$$

If $N_t \geq 1$, let $(t_1, r_1), \ldots, (t_{N_t}, r_{N_t})$ be the times and locations of these events. Then the likelihood ratio, L_t, is given by (see Snyder [9], pp. 471-476)

$$L_t = \frac{\prod_{j=1}^{N_t} \hat{\lambda}_1(t_j, r_j) \ \exp[-\int_0^t \int_D \hat{\lambda}_1(s, r)\, dr\, ds]}{\prod_{j=1}^{N_t} \hat{\lambda}_0(t_j, r_j) \ \exp[-\int_0^t \int_D \hat{\lambda}_0(s, r)\, dr\, ds]}. \tag{2.21}$$

Convention. When $N_t = 0$, the factors preceeding *exp* in equation (2.21) are taken to be one.

With the above ideas in mind, we state the **Joint Control-Decoding Problem**.

Find a rule for generating a control signal, $\{u_t, 0 \leq t \leq T\}$, predictable with respect to $\{G_t, 0 \leq t \leq T\}$, such that the probability of error corresponding to the likelihood ratio test

$$L_T \underset{H_0}{\overset{H_1}{\gtrless}} 1 \tag{2.22}$$

is minimized.

In (2.22) H_1 is the hypothesis that a "1" is being sent, and H_0 is the hypothesis that a "0" is being sent.

Notation. If we let

$$f\ (t\ ,\ r\ ,\ x_t\ ) = \exp[-\frac{1}{2}(r - H(t)x_t)'\ R\,(t)^{-1}\,(r - H(t)x_t)\,]\ , \tag{2.23}$$

and

$$\hat{f}\ (t\ ,\ r\ ) = \mathbf{E}\ [\ f\ (t\ ,\ r\ ,\ x_t\ )\ \mid\ G_t\ ]\ , \tag{2.24}$$

then

$$\lambda_i(t\ ,\ r\ ,\ x_t\ ) = \mu_i(t)f\ (t\ ,\ r\ ,\ x_t\ )\ , \tag{2.25}$$

$$\hat{\lambda}_i(t\ ,\ r\ ) = \mu_i(t)\hat{f}\ (t\ ,\ r\ )\ , \tag{2.26}$$

and

$$L_t = \prod_{j=1}^{N_t} \frac{\mu_1(t_j)}{\mu_0(t_j)}\ \exp[-\int_0^t [\ \mu_1(s) - \mu_0(s)\ ]\int_D\ \hat{f}\ (s\ ,\ r\ )\ dr\ \ ds\ ]\ . \tag{2.27}$$

We see from equation (2.27) that our immediate goal is to compute

$$\int_D \hat{f}\ (s\ ,\ r\ )\ dr\ . \tag{2.28}$$

In Chapter 3 we consider the Joint Control-Decoding Problem under the assumption that the photodetector has an infinite photosensitive surface; mathematically, this amounts to setting $D = \mathbb{R}^2$. We prove that "all controls are optimal" in the sense that the probability of a decoding error is not a function of the control signal. We show how this implies that the Joint Control-Decoding Problem degenerates into a simple detection problem with Poisson-process observations. We also use the work of Rhodes and Snyder [3], together with Theorem A1 of Appendix A, to compute $\mathbf{E}\ [\ f\ (t\ ,\ r\ ,\ x_t\ )\ \mid\ F_t\ ]$ explicitly.

In Chapter 4 we drop the extra assumption that the photodetector has an infinite photosensitive surface. We consider the time point process $\{\ N_t\ ,\ t \geq 0\ \}$. Associated with this process, there is a conditional expectation similar to that in equation (2.19). We attempt to apply the work of Boel and Benes [5] on time point processes to gain some insight into the problem of finding $\hat{\lambda}_i(t, r)$ when $D \neq \mathbb{R}^2$.

Finally, in Chapter 5 we look at *linear* estimates of $\lambda_i(t, r, x_t)$ given G_t.

Remark. We are trying to find a rule for generating a control signal, $\{ u_t, 0 \leq t \leq T \}$, which will minimize the probability of a decoding error when the likelihood ratio detector described by equations (2.21) and (2.22) is used. This is quite different from the approach in Rhodes and Snyder [3]. In that paper, a control signal, $\{ u_t, 0 \leq t \leq T \}$, was sought which would minimize

$$\mathbf{E} \{ \int_0^T [u_t' M_1(t) u_t + x_t' M_2(t) x_t] dt + x_T' M_3 x_T \}, \tag{2.29}$$

where the matrix $M_1(t)$ is positive definite, and the matrices $M_2(t)$ and M_3 are nonnegative definite.

CHAPTER 3

RESULTS FOR THE INFINITE DETECTOR

I. Introduction

In this chapter we consider the Joint Control-Decoding Problem under the assumption that the photodetector has an infinite photosensitive surface; mathematically, this amounts to setting $D = \mathbb{R}^2$. Under this assumption, we prove that "all controls are optimal" in the sense that the probability of a decoding error is not a function of the control signal. We show how this implies that the Joint Control-Decoding Problem degenerates into a simple detection problem with Poisson-process observations. More precisely, assume $D = \mathbb{R}^2$. For each $t \geq 0$, define the random field

$$\mathbf{M}^t = \{ N(E \times \mathbb{R}^2) : E \in B(t, \infty) \} . \tag{3.1}$$

We establish that $\mathbf{M}^t$ is independent of the σ-field X_t defined in (2.14). In particular, this implies that $\mathbf{M}^0$ is independent of X_0. From this we conclude that the process $\{ N_t , t \geq 0 \}$ is independent of X_0 (note that $D = \mathbb{R}^2$ implies $N_t = N((0,t] \times \mathbb{R}^2))$. We next show that when $D = \mathbb{R}^2$, the quantity in (2.28) is deterministic. With these facts we examine (2.27) and find that the likelihood ratio, L_T, is independent of X_0. We conclude that when $D = \mathbb{R}^2$, the probability of a decoding error is not a function of the rule used to generate $\{ u_t , 0 \leq t \leq T \}$ from the observations.

Also, up to this point we have not needed the quantity $\hat{f}(t, r)$ in (2.24) in an explicit form. We have, however, been able to compute it and present it as Theorem 3.3.

II. The Photodetector with an Infinite Photosensitive Surface

We present the fact that $\mathbf{M}^t$ is independent of X_t and that $\{ N_t , t \geq 0 \}$ is independent of X_0 as Theorem 3.1. The proof depends primarily on the fact, mentioned in Chapter 2, that $\int_{\mathbb{R}^2} \lambda_i(t, r, x_t)\, dr = 2\pi\, \mu_i(t)\sqrt{\det R(t)}$.

Theorem 3.1. *For each $t \geq 0$, the random field $\mathbf{M}^t$ defined in equation (3.1) is independent of the σ-field X_t defined by equation (2.14). The process $\{ N_t, t \geq 0 \}$ is independent of X_0.*

Proof. To prove that $\mathbf{M}^t$ is independent of X_t, it is sufficient to show that the conditional characteristic function of $N(E \times \mathbb{R}^2)$ is deterministic for $E \in B(t, \infty)$. Now, it follows immediately from the assumption and comments following (2.14), that for $\theta \in \mathbb{R}$,

$$\mathbf{E}[\, e^{j\theta N(E\times\mathbb{R}^2)} \mid X_t \,] = \exp[\, (e^{j\theta} - 1) \int_E \int_{\mathbb{R}^2} \lambda_i(s, r, x_s)\, dr\, ds \,] \tag{3.2}$$

$$= \exp[\, (e^{j\theta} - 1) \int_E 2\pi\mu_i(s)\sqrt{\det R(s)}\, ds \,] .$$

Hence $\mathbf{M}^t$ is independent of X_t. Clearly, since $\{ N_t, t > 0 \} \subset \mathbf{M}^0$, $\{ N_t, t \geq 0 \}$ is independent of X_0.

QED

The following corollary is an immediate consequence of the assumption following (2.14) together with equation (3.2).

Corollary 3.2. *Under the (unconditional) measure $\mathbf{P}$, $\{ N_t, t \geq 0 \}$ is a Poisson process, and for each $t \geq 0$, $\mathbf{M}^t$ is a Poisson random field. Each process has the same deterministic intensity, $2\pi\, \mu_i(s)\sqrt{\det R(s)}$.*

To establish that (2.28) is deterministic is also quite straightforward. Observe that since $D = \mathbb{R}^2$,

$$\int_D \hat{f}(t,r)\,dr = \int_{\mathbf{R}^2} \mathbf{E}[\,f(t,r,x_t) \mid F_t\,]\,dr$$
$$= \mathbf{E}[\,\int_{\mathbf{R}^2} f(t,r,x_t)\,dr \mid F_t\,] \tag{3.3}$$
$$= \mathbf{E}[\,2\pi\sqrt{\det R(t)} \mid F_t\,]$$
$$= 2\pi\sqrt{\det R(t)}\,.$$

Substituting (3.3) into (2.27) and setting $t = T$ yields

$$L_T = \prod_{j=1}^{N_T} \frac{\mu_1(t_j)}{\mu_0(t_j)} \exp[-2\pi\int_0^T \sqrt{\det R(s)}\,[\,\mu_1(s)-\mu_0(s)\,]\,ds\,]\,. \tag{3.4}$$

Now, L_T depends only on N_T and its jump times, $\{ t_1, \ldots, t_{N_T} \}$. As a consequence of Corollary 3.2, under the measure $\mathbf{P}$, the probability law of N_T and its jump times is not affected by the manner in which we generate $\{ u_t, 0 \le t \le T \}$ from the observations $\{ G_t, 0 \le t \le T \}$. Hence, the probability of a decoding error is not a function of the control law used to drive the fine-tracking equipment. This means that under the assumption that the photodetector has an infinite photosensitive surface, the Joint Control-Decoding Problem degenerates into a simple detection problem with Poisson-process observations. The complete solution of this problem is given by (3.4) and (2.22).

Remark. Professor A. M. Makowski pointed out that the preceding comments hold even if we generate $\{ u_t, 0 \le t \le T \}$ from $\{ X_t, 0 \le t \le T \}$. This means that perfect knowledge of $F_t \vee X_0$ cannot improve our ability to decode.

We next present Theorem 3.3 in which we compute $\mathbf{E}[\,f(t,r,x_t) \mid G_t\,]$, which, since $D = \mathbf{R}^2$, is just $\mathbf{E}[\,f(t,r,x_t) \mid F_t\,]$.

Theorem 3.3. *Under the assumptions of Chapter 2,*

$$\mathbf{E}\,[\,\exp[-\frac{1}{2}(r-H(t)x_t)'\;R(t)^{-1}\,(r-H(t)x_t)\,]\;\mid\;F_t\,] \tag{3.5}$$

$$=\frac{\sqrt{\det R(t)}}{\sqrt{\det Q_t}}\exp[-\frac{1}{2}(r-H(t)\hat{x}_t)'\;Q_t^{-1}\,(r-H(t)\hat{x}_t)\,]\,,$$

where

$$\hat{x}_t \triangleq \mathbf{E}\,[\,x_t\;\mid\;F_t\,]\,, \tag{3.6}$$

$$\hat{\Sigma}_t \triangleq \mathbf{E}\,[\,(x_t-\hat{x}_t)(x_t-\hat{x}_t)'\;\mid\;F_t\,]>0\,,\quad \mathbf{P}\text{ -a.s.}, \tag{3.7}$$

$$Q_t \triangleq H(t)\hat{\Sigma}_t H(t)' + R(t)\,, \tag{3.8}$$

and

$$d\hat{x}_t = F(t)\hat{x}_t\,dt + G(t)u_t\,dt + \int_{\mathbb{R}^2}\hat{\Sigma}_{t-}H(t-)'\;Q_{t-}^{-1}\;(r-H(t-)\hat{x}_{t-})\,N(dt\times dr)\,;\quad \hat{x}_0=m\;, \tag{3.9}$$

$$d\hat{\Sigma}_t = F(t)\hat{\Sigma}_t\,dt + \hat{\Sigma}_t F(t)'\;dt + V(t)V(t)'\;dt - \hat{\Sigma}_{t-}H(t-)'\;Q_{t-}^{-1}\;H(t-)\hat{\Sigma}_{t-}N(dt\times\mathbb{R}^2)\,;\quad \hat{\Sigma}_0=S. \tag{3.10}$$

Proof. In [3] it is proved that the conditional density of x_t given F_t is *Gaussian* with conditional mean $\hat{x}_t$ and conditional covariance $\hat{\Sigma}_t$ (which is positive definite almost surely because we are assuming that S is positive definite) satisfying (3.9) and (3.10) above. Our result, equation (3.5), follows immediately from Theorem A1 of Appendix A.

QED

Remark. From "Proof (2)" of Theorem A1, we see that the following more general result is also true. Let $\psi_t(\eta)$, $\eta\in\mathbb{R}^n$, be the conditional characteristic function of x_t given G_t, even if $D\neq\mathbb{R}^2$. Then the Fourier transform of $\mathbf{E}\,[\,f\,(t,r,x_t)\;\mid\;G_t\,]$ as a function of r, is given by

$$\int_{\mathbb{R}^2}\mathbf{E}\,[\,f\,(t,r,x_t)\;\mid\;G_t\,]e^{j\theta' r}\,dr = 2\pi\sqrt{\det R(t)}\,e^{-\frac{1}{2}\theta' R(t)\theta}\psi_t(H'\theta)\,;\quad \theta\in\mathbb{R}^2.$$

We have an immediate corollary concerning the l-th conditional moments of $f(t, r, x_t)$.

Corollary 3.4. *For* $l \geq 1$,

$$\mathbf{E}\,[f(t, r, x_t)^l \mid F_t] = \frac{\sqrt{\det R(t)}}{l^2 \sqrt{\det Q_t(l)}} \exp[-\frac{1}{2}(r - H(t)\hat{x}_t)' \; Q_t(l)^{-1} \, (r - H(t)\hat{x}_t)], \tag{3.11}$$

where

$$Q_t(l) \triangleq H(t)\hat{\Sigma}_t H(t)' + l^{-1} R(t) .$$

Remark. Observe that (3.9) is coupled to (3.10). However, (3.10) evolves independently of (3.9). Recalling that $\{ (t_n, r_n) \}$ is the family of times and locations of the photoelectrons we have observed, we interpret (3.10) as

$$\frac{d}{dt}\hat{\Sigma}_t = F(t)\hat{\Sigma}_t + \hat{\Sigma}_t F(t)' + V(t)V(t)' , \tag{3.12}$$

for $t \in [0, t_1)$ or $t \in [t_n, t_{n+1})$, the initial conditions being

$$\begin{aligned} &\hat{\Sigma}_0 = S , \\ &\hat{\Sigma}_{t_n} = \hat{\Sigma}_{t_n -} - \hat{\Sigma}_{t_n -} H(t_n -)' \; Q_{t_n -}^{-1} \; H(t_n -)\hat{\Sigma}_{t_n -} . \end{aligned} \tag{3.13}$$

After solving (3.12) and (3.13), we interpret (3.9) as

$$\frac{d}{dt}\hat{x}_t = F(t)\hat{x}_t + G(t)u_t , \tag{3.14}$$

for $t \in [0, t_1)$ or $t \in [t_n, t_{n+1})$, with initial conditions

$$\begin{aligned} &\hat{x}_0 = m , \\ &\hat{x}_{t_n} = \hat{x}_{t_n -} + \hat{\Sigma}_{t_n -} H(t_n -)' \; Q_{t_n -}^{-1} \; (r_n - H(t_n -)\hat{x}_{t_n -}) . \end{aligned} \tag{3.15}$$

In the next chapter we will give an indication of the difficulties encountered if instead of computing $\mathbf{E}\,[\lambda_i(t, r, x_t) \mid F_t]$, one tries to compute $\mathbf{E}\,[\lambda_i(t, r, s_t) \mid G_t]$.

CHAPTER 4

A LOOK AT THE GENERAL FILTERING PROBLEM

I. Introduction

In this chapter we no longer assume that the photodetector has an infinite photosensitive surface. We consider the time point process $\{ N_t , t \geq 0 \}$. Associated with this process, there is a conditional expectation similar to that in equation (2.19). We attempt to apply the work of Boel and Benes [5] on time point processes to gain some insight into the general filtering problem given time-space point-process observations.

Recall from Chapter 2 that our immediate goal is the calculation of (2.28), which can be rewritten as

$$\mathbf{E} [\int_D \lambda_i(t , r , x_t) \, dr \mid G_t] , \tag{4.1}$$

where

$$\lambda_i(t , r , x_t) = \mu_i(t) \exp[-\frac{1}{2}(r - H(t)x_t)' \, R(t)^{-1} (r - H(t)x_t)] . \tag{4.2}$$

In this chapter we are going to demonstrate some of the difficulties in computing

$$\mathbf{E} [\int_D \lambda_i(t , r , x_t) \, dr \mid H_t] , \tag{4.3}$$

when $D \neq \mathbb{R}^2$ and where

$$H_t \triangleq \sigma\{ N_s , 0 \leq s \leq t \} . \tag{4.4}$$

Observe that since $N_t \triangleq N((0,t] \times D) = N(\{ (0,t] \times \mathbb{R}^2 \} \bigcap \{ (0,\infty) \times D \})$, we have $H_t \subset \sigma\{ N(B \bigcap \{ (0,\infty) \times D \}) : B \in B(0,t] \otimes B^2 \} \triangleq G_t$. Now, define

$$\overline{X}_t = H_t \vee X_0 . \tag{4.5}$$

We claim that $\{ N_t , t \geq 0 \}$ is an $\{\overline{X}_t\}$-doubly-stochastic Poisson process with stochastic intensity

$$\lambda_t = \int_D \lambda_i(t, r, x_t)\, dr \ . \tag{4.6}$$

This means, of course, that under the measure $\mathbf{P}(\bullet \mid \overline{X}_t)$,

$$\overline{\mathbf{N}}^t \triangleq \{ N_s - N_t, s \geq t \} \tag{4.7}$$

is a Poisson process with $\overline{X}_0$-measurable intensity λ_t. This follows because for $t \geq 0$, $\overline{\mathbf{N}}^t \subset \mathbf{N}^t$ and $\overline{X}_t \subset X_t$ and because $\overline{X}_0 = X_0$ (recall the discussion following (2.14)).

In Section II we will outline an approach due to Boel and Benes [5] for computing $\mathbf{E}[\lambda_t \mid H_t]$. We will show that in the general case, the method leads from the problem of computing $\mathbf{E}[\lambda_t \mid H_t]$ directly, to a new problem which is no less tractable; however, there is a special case (generalized very slightly from [5]) where this method yields a well-defined family of partial differential equations. If one can solve these PDE's, then the conditional moment generating function of λ_t given H_t can be recovered, and from this, of course, $\mathbf{E}[\lambda_t \mid H_t]$ can be found. In Section III we demonstrate that for our particular λ_t, this method does not lead to a tractable approach for finding $\mathbf{E}[\lambda_t \mid H_t]$.

II. Outline of Method

Following [5], we begin with our processes $\{ N_t, t \geq 0 \}$ and $\{ \lambda_t, t \geq 0 \}$ as described in Section I; our goal is to compute $\mathbf{E}[\lambda_t \mid H_t]$ for $t \geq 0$. To simplify the calculations, we assume that $\mu_i(t) \equiv 1$ and that $u_t \equiv 0$ in the remainder of this chapter. Then

$$\lambda_t = \int_D \exp[-\frac{1}{2}(r - H(t)x_t)' \, R(t)^{-1} \, (r - H(t)x_t)] \, dr \ , \tag{4.8}$$

where $\{ x_t, t \geq 0 \}$ satisfies

$$dx_t = F(t)x_t\, dt + V(t)dv_t \ ; \quad x_0 = X. \tag{4.9}$$

The general method proceeds as follows. (Note that it will make no use of the functional form of λ_t, other than the fact that it is a twice-continuously-differentiable function of x_t, where x_t is the solution to an arbitrary Ito stochastic differential equation). In view of (4.8), (4.9), and Ito's rule, we may write

$$d\lambda_t = \alpha_t \, dt + \beta_t \, dv_t \ . \tag{4.10}$$

Here α_t and β_t are determined by Ito's formula; α_t is a scalar and β_t is a row vector. Now, let f be an arbitrary, twice-continuously-differentiable function defined on $[0,\infty)$. Applying Ito's rule to f and (4.10),

$$d(f(\lambda_t)) = \gamma_t \, dt + \delta_t \, dv_t \ . \tag{4.11}$$

Again, γ_t and δ_t are determined by Ito's rule. For any such f, define

$$\hat{f}_t = \mathbf{E}[\, f(\lambda_t) \mid H_t \,] \, . \tag{4.12}$$

According to [5],

$$d\hat{f}_t = \mathbf{E}[\, \gamma_t \mid H_t \,]dt + k_{t-}(dN_t - \hat{\lambda}_t \, dt) \, , \tag{4.13}$$

where

$$\hat{\lambda}_t \triangleq \mathbf{E}[\, \lambda_t \mid H_t \,] \, , \tag{4.14}$$

and

$$\begin{aligned} k_t &= \hat{\lambda}_t^{-1} \, \mathbf{E}[\, (f(\lambda_t) - \hat{f}_t)(\lambda_t - \hat{\lambda}_t) \mid H_t \,] \\ &= \hat{\lambda}_t^{-1} \, \mathbf{E}[\, f(\lambda_t)\lambda_t \mid H_t \,] - \hat{f}_t \ . \end{aligned} \tag{4.15}$$

At this point, we should say a few words about stochastic differential equations of the form

$$dz_t = a_t \, dt + b_t \, dN_t \ ; \quad z_0 = Z.$$

Let $\{ t_n \}$ be the sequence of jump times of $\{ N_t , t \geq 0 \}$. We interpret the above equation as the sequence of ordinary sample-path differential equations

$$\frac{dz_t}{dt} = a_t \ ; \quad 0 \le t < t_1 , \quad z_0 = Z ,$$

$$\frac{dz_t}{dt} = a_t \ ; \quad t_n < t < t_{n+1} , \quad z_{t_n +} = z_{t_n -} + b_{t_n} .$$

In light of this, (4.13) implies that $\hat{f}_t$ is differentiable, and therefore continuous, on $(0, t_1)$ and (t_n , t_{n+1}) for $n \geq 1$. Now consider the quantity $f(\lambda_t)\lambda_t$. This is also a twice-continuously-differentiable function of λ_t. By the preceding argument, $\mathbf{E}[f(\lambda_t)\lambda_t \mid H_t]$ is differentiable (as a function of t) and therefore continuous between the jump times, $\{ t_n \}$. As a consequence, k_t is continuous and differentiable between the jump times.

Defining

$$\phi(s , t) = e^{s \lambda_t} , \tag{4.16}$$

we have by Ito's rule that

$$\partial_t \phi(s , t) = \phi(s , t)[s \alpha_t + \frac{1}{2} s^2 | \beta_t |^2]dt + s \phi(s , t)\beta_t \, dv_t . \tag{4.17}$$

Set

$$\rho(s , t) = \phi(s , t)[s \alpha_t + \frac{1}{2} s^2 | \beta_t |^2] , \tag{4.18}$$

and let

$$\hat{\rho}(s , t) = \mathbf{E}[\rho(s , t) \mid H_t] . \tag{4.19}$$

Next, let

$$\hat{\phi}(s , t) = \mathbf{E}[\phi(s , t) \mid H_t] = \mathbf{E}[e^{s \lambda_t} \mid H_t] . \tag{4.20}$$

Thus, $\hat{\phi}(s , t)$ is the conditional moment generating function of λ_t given H_t. If we can determine $\hat{\phi}(s , t)$, then

$$\hat{\lambda}_t = \frac{\partial \hat{\phi}}{\partial s} \Big|_{s=0} .$$

Following the procedure described at the beginning of this section, we see that

$$\partial_t \hat{\phi}(s\,,\,t\,) = \hat{\rho}(s\,,\,t\,)dt \;+\; k_{t-}(dN_t \;-\; \hat{\lambda}_t\,dt\,)\,, \tag{4.21}$$

where now

$$k_t \;= \hat{\lambda}_t^{-1}\,\mathbf{E}\,[\;\phi(s\,,\,t\,)\lambda_t \;\;|\;\; H_t\;] - \;\hat{\phi}(s\,,\,t\,)\,. \tag{4.22}$$

At this point, observe that since $\phi(s\,,\,t\,) = e^{s\lambda_t}$, we have

$$\phi(s\,,\,t\,)\lambda_t \;= \frac{\partial}{\partial s}\,\phi(s\,,\,t\,)\,. \tag{4.23}$$

Thus,

$$k_t \;= \hat{\lambda}_t^{-1}\,\frac{\partial}{\partial s}\,\hat{\phi}(s\,,\,t\,) - \;\hat{\phi}(s\,,\,t\,)\,, \tag{4.24}$$

since

$$\mathbf{E}\,[\;\frac{\partial}{\partial s}\,\phi(s\,,\,t\,)\;\;|\;\;H_t\;] = \frac{\partial}{\partial s}\,\mathbf{E}\,[\;\phi(s\,,\,t\,)\;\;|\;\;H_t\;] = \frac{\partial}{\partial s}\,\hat{\phi}(s\,,\,t\,)\,. \tag{4.25}$$

We now define

$$g_0(t\,) = \exp[\,-\int_0^{\,t}\,(\hat{\lambda}_\tau \;-\; \mu_0)\;d\,\tau\,]\,\prod_{t_n \le t}\left(\frac{\hat{\lambda}_{t_n -}}{\mu_0}\right)\,. \tag{4.26}$$

The quantity $g_0(t\,)$ defined in (4.26) can be interpreted as the Radon-Nikodym derivative of $\mathbf{P}$ with respect to new measure $\mathbf{P}_0$. More precisely, for each set $C \in H_t$,

$$\mathbf{P}(C\,) = \int_C\; g_0(t\,)\;d\,\mathbf{P}_0\,.$$

For more information on this "change of measure" interpretation, see ([5], section II and the comments following equation (19)). Clearly, $g_0(t\,)$ is continuous between the jump times, $\{\;t_n\;\}$, and, $g_0(t_n\,) = \dfrac{\hat{\lambda}_{t_n -}}{\mu_0}\,g_0(t_n\,-\,)$. Also, note that since $\hat{\lambda}_t$ is continuous between the jump times, $g_0(t\,)$ is continuously differentiable between the jump times. Therefore, we can write

$$dg_0(t) = (\mu_0 - \hat{\lambda}_t) g_0(t)dt + (\frac{\hat{\lambda}_{t-}}{\mu_0} - 1) g_0(t-)dN_t \; ; \quad g_0(0) = 1 \; . \tag{4.27}$$

The next step is to define the function $g(s, t)$ by

$$g(s, t) = g_0(t)\,\hat{\phi}(s, t) \; . \tag{4.28}$$

Since $\hat{\phi}(0, t) = 1$, we see that $\hat{\phi}(s, t) = g(s, t)/g(0, t)$. This means that finding $g(s, t)$ is equivalent to finding $\hat{\phi}(s, t)$. The plan is to use (4.21) and (4.27) to obtain a partial differential equation for $g(s, t)$. So, for t between the jump times, $\{ t_n \}$, we may use the product rule with (4.21) and (4.27). Thus,

$$\frac{\partial g}{\partial t} = g_0(t)[\,\hat{\rho}(s, t) - k_t \hat{\lambda}_t\,] + \hat{\phi}(s, t)[\,(\mu_0 - \hat{\lambda}_t) g_0(t)\,] \tag{4.29}$$

$$= g_0(t)\hat{\rho}(s, t) - \frac{\partial g}{\partial s} + \hat{\lambda}_t\, g(s, t) + \mu_0\, g(s, t) - \hat{\lambda}_t\, g(s, t) \; .$$

Simplifying and rearranging terms,

$$\frac{\partial g}{\partial t} + \frac{\partial g}{\partial s} = \mu_0 g(s, t) + g_0(t)\hat{\rho}(s, t) \; . \tag{4.30}$$

To determine the behavior of $g(s, t)$ at the jump times, observe that if $t = t_n$ for any $n \geq 1$, then (4.21) and (4.24) imply

$$\hat{\phi}(s, t) - \hat{\phi}(s, t-) = k_{t-} \tag{4.31}$$

$$= \hat{\lambda}_{t-}^{-1} \frac{\partial}{\partial s} \hat{\phi}(s, t-) - \hat{\phi}(s, t-) \; .$$

So,

$$\hat{\phi}(s, t) = \hat{\lambda}_{t-}^{-1} \frac{\partial}{\partial s} \hat{\phi}(s, t-) \; . \tag{4.32}$$

Multiplying both sides of (4.32) by $g_0(t)$, we find that

$$g(s, t) = g_0(t)\, \hat{\lambda}_{t-}^{-1} \frac{\partial}{\partial s} \hat{\phi}(s, t-) \; . \tag{4.33}$$

Substituting $g_0(t) = \frac{\hat{\lambda}_{t-}}{\mu_0} g_0(t-)$ into (4.33), we have

$$g(s, t) = \mu_0^{-1} \frac{\partial}{\partial s} g(s, t-) . \tag{4.34}$$

The final piece of information needed to start solving (4.30) is the initial data $\{ g(s, 0), \text{all } s \}$. However, since $g(0, 0) = 1$, we have $g(s, 0) = \hat{\phi}(s, 0) \triangleq \mathbf{E}[e^{s\lambda_0}]$. We assume that the unconditional moment-generating function is known. Then the course of action is to solve

$$\frac{\partial g}{\partial t} + \frac{\partial g}{\partial s} = \mu_0 g(s, t) + g_0(t)\hat{\rho}(s, t) , \tag{4.35}$$

for t between the jump times, using the boundary conditions

$$\begin{aligned} g(s, 0) &= \mathbf{E}[e^{s\lambda_0}] \\ g(s, t_n) &= \mu_0^{-1} \frac{\partial}{\partial s} g(s, t_n-) ; \quad n \geq 1 . \end{aligned} \tag{4.36}$$

The problem in solving (4.35) and (4.36) is that usually we don't know $g_0(t)\hat{\rho}(s, t)$. Recall that

$$\rho(s, t) = \phi(s, t)[s\,\alpha_t + \frac{1}{2} s^2 | \beta_t |^2] ,$$

where

$$d\lambda_t = \alpha_t\, dt + \beta_t\, dv_t .$$

Now, just suppose that $\alpha_t = p(\lambda_t)$ and $| \beta_t |^2 = q(\lambda_t)$ where p and q are *polynomials*. Recall that $\phi(s, t) = e^{s\lambda_t}$. With the obvious meaning of the symbols $p(\frac{\partial}{\partial s})$ and $q(\frac{\partial}{\partial s})$,

$$\rho(s, t) = s\ p(\frac{\partial}{\partial s})\phi(s, t) + \frac{1}{2} s^2 q(\frac{\partial}{\partial s})\phi(s, t) , \tag{4.37}$$

and hence

$$\hat{\rho}(s, t) = s\ p(\frac{\partial}{\partial s})\hat{\phi}(s, t) + \frac{1}{2} s^2 q(\frac{\partial}{\partial s})\hat{\phi}(s, t) . \tag{4.38}$$

We now see that

$$g_0(t)\hat{\rho}(s, t) = s\ p(\frac{\partial}{\partial s})g(s, t) + \frac{1}{2} s^2 q(\frac{\partial}{\partial s})g(s, t) . \tag{4.39}$$

Under these very special conditions, (4.35) becomes

$$\frac{\partial g}{\partial t} + \frac{\partial g}{\partial s} = \mu_0 g + s\ p(\frac{\partial}{\partial s})g + \frac{1}{2} s^2 q(\frac{\partial}{\partial s})g . \tag{4.40}$$

In [5], examples in which α_t and $| \beta_t |^2$ *are* polynomials in λ_t are worked out in detail; equations of the form of (4.40) are derived and solved.

III. An Application of the General Method

In this section, we apply the above procedures to the following problem. Similar to (4.8) and (4.9), let

$$\lambda_t = \int_D e^{-\frac{1}{2}(r - Hx_t)' R^{-1} (r - Hx_t)}\, dr , \tag{4.41}$$

and

$$dx_t = Fx_t\, dt + V dv_t , \tag{4.42}$$

where we have assumed that H, R, F, and V are no longer functions of time in order to simplify the calculations. We introduce the following notation. Let

$$u(r, x) = e^{-\frac{1}{2}(r - Hx)' R^{-1} (r - Hx)} , \tag{4.43}$$

and set

$$\lambda(x) = \int_D u(r, x)\, dr . \tag{4.44}$$

Then (4.41) becomes

$$\lambda_t = \lambda(x_t) = \int_D u(r, x_t)\, dr \ . \tag{4.45}$$

The first step is to apply Ito's rule to (4.45) and (4.42). Before doing so, note that

$$\frac{\partial u}{\partial x} = u(r, x)[\, r' R^{-1} H - x' H' R^{-1} H \,] , \tag{4.46}$$

and that

$$\frac{\partial^2 u}{\partial x^2} = u(r, x)(H' R^{-1} [\, (r-Hx)(r-Hx)' - R \,] R^{-1} H \,) . \tag{4.47}$$

Hence,

$$\frac{\partial \lambda}{\partial x} = (\int_D u(r, x) r' \, dr - \lambda(x)\, x' H' \,) R^{-1} H , \tag{4.48}$$

and

$$\frac{\partial^2 \lambda}{\partial x^2} = \int_D u(r, x) H' R^{-1} (r-Hx)(r-Hx)' R^{-1} H \, dr - \lambda(x) H' R^{-1} H . \tag{4.49}$$

Applying Ito's rule to (4.45) and (4.42) yields

$$d\lambda_t = \alpha(x_t) dt + \beta(x_t) dv_t , \tag{4.50}$$

where

$$\begin{aligned} \alpha(x) = {} & (\int_D u(r, x) r' \, dr - \lambda(x)\, x' H' \,) R^{-1} H F x \\ & + \frac{1}{2} \int_D u(r, x) \, | V' H' R^{-1} (r-Hx) |^2 \, dr \\ & - \frac{1}{2} \lambda(x) \ \mathrm{tr}\,(V' H' R^{-1} H V) , \end{aligned} \tag{4.51}$$

and

$$\beta(x) = (\int_D u(r, x) r' \, dr - \lambda(x)\, x' H' \,) R^{-1} H V . \tag{4.52}$$

Clearly, $\alpha(x_t)$ and $|\beta(x_t)|^2$ do not take a form which will give us a partial differential equation only in terms of the unknown function $g(s, t)$.

Remark. When $D = \mathbb{R}^2$, (4.48) and (4.49) become

$$\frac{\partial \lambda}{\partial x} = (2\pi \sqrt{\det R} - \lambda(x)) \, x' \, H' \, R^{-1} H , \tag{4.53}$$

and

$$\frac{\partial^2 \lambda}{\partial x^2} = (2\pi \sqrt{\det R} - \lambda(x)) \, H' \, R^{-1} H . \tag{4.54}$$

Now, (4.51) and (4.52) become

$$\alpha(x) = (2\pi \sqrt{\det R} - \lambda(x)) \, x' \, H' \, R^{-1} H \, Fx + \frac{1}{2} (2\pi \sqrt{\det R} - \lambda(x)) \, \mathrm{tr} \, (V' \, H' \, R^{-1} H \, V) , \tag{4.55}$$

and

$$\beta(x) = (2\pi \sqrt{\det R} - \lambda(x)) \, x' \, H' \, R^{-1} H \, V . \tag{4.56}$$

Note that even in this case, $\alpha(x)$ and $| \beta(x) |^2$ are not polynomials in $\lambda(x)$, even though the second term of (4.55) *is* a first degree polynomial in $\lambda(x)$.

IV. Remarks

(i) We have seen the difficulties in computing $\mathbf{E} [\lambda_t \mid H_t]$. Clearly, any similar attempt to compute $\mathbf{E} [\lambda_t \mid G_t]$ would require a "time-space version" of (4.13) (see, for example, Bremaud [7], Theorem T9 on pp. 240-241).

(ii) In the next chapter we meet with some success when we look at *linear* estimates of $\lambda_i (t , r , x_t)$ given G_t.

CHAPTER 5

RESULTS FOR THE LINEAR FILTERING PROBLEM

I. Introduction

As shown in Chapter 4, the computation of $\hat{\lambda}_i(t, r) = \mu_i(t)\hat{f}(t, r)$ has proven to be intractable when $D \neq \mathbb{R}^2$. This has motivated us to look at *linear* estimates of $\lambda_i(t, r, x_t)$ given G_t. In this chapter we show that if $u_t = u(t)$ for some deterministic control signal $u(t)$, then we can compute the quantities necessary to write down the integral equation which defines the best linear estimate of $\lambda_i(t, r, x_t)$ given G_t.

II. Linear Estimators

To simplify the notation, we suppress the subscript i in the remainder of this chapter. We call $\hat{\lambda}_L(t, r)$ a *linear* estimate of $\lambda(t, r, x_t)$ given G_t, if $\hat{\lambda}_L$ can be written in the form

$$\hat{\lambda}_L(t, r) = \int_0^t \int_D h(t, r; \tau, \rho)\,[\,N(d\tau \times d\rho) - \overline{\lambda}(\tau, \rho)\,d\tau\,d\rho\,] + h_0(t, r), \quad (5.1)$$

where h and h_0 are deterministic, $\overline{\lambda}(t, r) \triangleq \mathbf{E}[\,\lambda(t, r, x_t)\,]$, and $N(d\tau \times d\rho)$ represents the number of photoelectrons generated by the photodetector in an infinitesimal area $d\rho$ during a time interval $d\tau$. We wish to choose h and h_0 to minimize

$$\mathbf{E}[\,|\lambda(t, r, x_t) - \hat{\lambda}_L(t, r)|^2\,]. \quad (5.2)$$

Lemma 5.1. (Grandell [10]). *Let $\hat{\lambda}_L(t, r)$ be given by (5.1). Under the assumption following (2.14), (5.2) will be minimized if $h_0(t, r) = \overline{\lambda}(t, r)$, and if h satisfies*

$$\Gamma(t, r; \tau, \rho) = \int_0^t \int_D h(t, r; \sigma, \varsigma)\Gamma(\sigma, \varsigma; \tau, \rho)\,d\varsigma\,d\sigma + h(t, r; \tau, \rho)\overline{\lambda}(\tau, \rho), \quad (5.3)$$

where

$$\Gamma(t\,,\,r\,;\,\tau,\,\rho)\;\triangleq\;\mathbf{cov}\;\,[\,\lambda(t\,,\,r\,,\,x_t\,),\,\lambda(\tau,\,\rho,\,x_\tau)\,]\;.$$

With reference to Lemma 5.1, we state our Theorem 5.2.

Theorem 5.2. *If the assumptions of Chapter 2 hold, and if $u(t)$ is a deterministic control signal, then*

$$\overline{\lambda}(t\,,\,r\,) = \mu(t\,)\frac{\sqrt{\det R\,(t\,)}}{\sqrt{\det\,Q\,(t\,)}}\exp[\,-\,\frac{1}{2}(r\,-H\,(t\,)\overline{x}\,(t\,))'\;\;Q\,(t\,)^{-1}\,(r\,-H\,(t\,)\overline{x}\,(t\,))\,]\;, \qquad (5.4)$$

where

$$\overline{x}\,(t\,)\;\triangleq\;\mathbf{E}\;[x_t\,]\;,$$

$$\Sigma(t\,)\;\triangleq\;\mathbf{cov}\;[x_t\,]\;,$$

$$Q\,(t\,)\;\triangleq\;H(t\,)\Sigma(t\,)H(t\,)'\;\;+\;R\,(t\,).$$

Furthermore,

$$\Gamma(t\,,\,r\,;\,\tau,\,\rho)\,+\,\overline{\lambda}(t\,,\,r\,)\overline{\lambda}(\tau,\,\rho) = \mu(t\,)\mu(\tau)\sqrt{\frac{\det R\,(t\,)\;\;\det R\,(\tau)}{\det\,Q\,(t\,,\,\tau)}}\;\;\times$$

$$\exp[\,-\,\frac{1}{2}(\begin{bmatrix} r \\ \rho \end{bmatrix} - \begin{bmatrix} H(t\,) & 0 \\ 0 & H(\tau) \end{bmatrix}\begin{bmatrix} \overline{x}\,(t\,) \\ \overline{x}\,(\tau) \end{bmatrix})'\;\;Q\,(t\,,\,\tau)^{-1}\,(\begin{bmatrix} r \\ \rho \end{bmatrix} - \begin{bmatrix} H(t\,) & 0 \\ 0 & H(\tau) \end{bmatrix}\begin{bmatrix} \overline{x}\,(t\,) \\ \overline{x}\,(\tau) \end{bmatrix})\,], \qquad (5.5)$$

where

$$\Sigma(t\,,\,\tau)\;\triangleq\;\mathbf{cov}\;\;[\,x_t\,,\,x_\tau\,]\;,$$

and

$$Q\,(t\,,\,\tau)\;\triangleq\;\begin{bmatrix} Q\,(t\,) & H(t\,)\Sigma(t\,,\,\tau)H(\tau)' \\ H(\tau)\Sigma(\tau,\,t\,)H(t\,)' & Q\,(\tau) \end{bmatrix}\;.$$

Proof. Let $u_t = u(t)$ for some deterministic control $u(t)$. For completeness, we make the following observations. Equation (2.12) becomes

$$dx_t = F(t)x_t\,dt + G(t)u(t)dt + V(t)dv_t \;;\quad x_0 = X.$$

Let $\Phi(t, t_0)$ be the transition matrix corresponding to $F(t)$. Then

$$\bar{x}(t) = \Phi(t, 0)m + \int_0^t \Phi(t, s)G(s)u(s)\,ds \;,$$

and

$$\Sigma(t, \tau) = \Phi(t, 0)S\,\Phi(\tau, 0)' + \int_0^{\min(t, \tau)} \Phi(t, s)\;V(s)V(s)'\;\Phi(\tau, s)'\;ds$$

Note that $\Sigma(t) = \Sigma(t, t)$. Because $u(t)$ is deterministic, $\{\, x_t\,,\, t \geq 0 \,\}$ is a Gaussian process.

To compute $\bar{\lambda}(t, r) = \mathbf{E}\,[\,\lambda(t, r, x_t)\,]$, observe that x_t is *Gaussian* with mean $\bar{x}(t)$ and covariance $\Sigma(t)$. Now apply Theorem A1, and (5.4) is immediate.

The computation of (5.5) is similar, but requires some judicious preliminary arithmetic. First, observe that $\Gamma(t, r; \tau, \rho) + \bar{\lambda}(t, r)\bar{\lambda}(\tau, \rho)$ is just another way of writing $\mathbf{E}\,[\,\lambda(t, r, x_t)\lambda(\tau, \rho, x_\tau)\,]$. Next, rewrite $\lambda(t, r, x_t)\lambda(\tau, \rho, x_\tau)$ as

$$\mu(t)\mu(\tau)\exp[-\frac{1}{2}(\begin{bmatrix} r \\ \rho \end{bmatrix} - \begin{bmatrix} H(t) & 0 \\ 0 & H(\tau) \end{bmatrix}\begin{bmatrix} x_t \\ x_\tau \end{bmatrix})' \begin{bmatrix} R(t)^{-1} & 0 \\ 0 & R(\tau)^{-1} \end{bmatrix} (\begin{bmatrix} r \\ \rho \end{bmatrix} - \begin{bmatrix} H(t) & 0 \\ 0 & H(\tau) \end{bmatrix}\begin{bmatrix} x_t \\ x_\tau \end{bmatrix})],$$

which is equal to

$$\mu(t)\mu(\tau) \tag{5.6}$$
$$\times \exp[-\frac{1}{2}(\begin{bmatrix} r \\ \rho \end{bmatrix} - \begin{bmatrix} H(t) & 0 \\ 0 & H(\tau) \end{bmatrix}\begin{bmatrix} x_t \\ x_\tau \end{bmatrix})' \begin{bmatrix} R(t) & 0 \\ 0 & R(\tau) \end{bmatrix}^{-1} (\begin{bmatrix} r \\ \rho \end{bmatrix} - \begin{bmatrix} H(t) & 0 \\ 0 & H(\tau) \end{bmatrix}\begin{bmatrix} x_t \\ x_\tau \end{bmatrix})].$$

Because $\{\, x_t\,,\, t \geq 0 \,\}$ is a Gaussian process, $\begin{bmatrix} x_t \\ x_\tau \end{bmatrix}$ is a Gaussian random vector with mean,

$\begin{bmatrix} \bar{x}(t) \\ \bar{x}(\tau) \end{bmatrix}$, and covariance $\begin{bmatrix} \Sigma(t) & \Sigma(t, \tau) \\ \Sigma(\tau, t) & \Sigma(\tau) \end{bmatrix}$. Clearly, Theorem A1 now applies to the expression in (5.6), and (5.5) readily follows.

QED

Remark. In equation (5.3), if we regard t and r as fixed, and divide through by $\bar{\lambda}(\tau, \rho)$, then the result has the form of the Fredholm equation

$$g = Bh + h ,$$

for known function g, known operator B, and unknown function h.

CHAPTER 6

CONCLUSIONS AND FURTHER RESEARCH

I. Conclusions

In this thesis we introduced the Joint Control-Decoding Problem as a part of the design of a direct-detection receiver for an optical communication system. This led us to investigate the computation of the conditional expectation $\mathbf{E}\,[\,\lambda_i(t\,,\,r\,,\,x_t)\;\mid\;G_t\,]$. We considered three "versions" of this problem: the case $D = \mathbb{R}^2$, the case $D \neq \mathbb{R}^2$, but conditioning on H_t, and finally, the case of linear estimators. In this section, we make a few comments about our results. In the next section we introduce a discrete-time version of the Joint Control-Decoding Problem as a possible starting point for future research.

In Chapter 3 we assumed that the photodetector had an infinitely large photosensitive surface ($D = \mathbb{R}^2$). Our results suggest that while the Joint Control-Decoding Problem is well-posed, the model is not realistic when $D = \mathbb{R}^2$, since no control was necessary for optimal decoding. Heuristicly, suppose the optical source sends messages by modulating the intensity of a laser beam. The position of the beam on an infinitely large photodetector carries no information about the message being sent, while the intensity of the beam determines how many photoelectrons will be observed at the detector. Intuitively then, only the total number of photoelectrons at each time $t \in (0,T]$ carries information about the incoming message. Because we assume $D = \mathbb{R}^2$, no matter where the beam falls on the photodetector, we count all photoelectrons; hence, use of the fine-tracking equipment to control where the beam falls cannot improve our ability to decode.

In Chapter 4 we dropped the assumption that $D = \mathbb{R}^2$. We tried to gain some insight into the general filtering problem for time-space point-process observations by examining an analgous problem for time point processes.

The kind of result we were looking for in Chapter 4 was at least something along the lines of Theorem 3.3, from which we obtained

$$\mathbf{E}\,[\,\lambda_i(t\,,\,r\,,\,x_t)\;\mid\;F_t\,] = \mu_i(t)\,\frac{\sqrt{\det R(t)}}{\sqrt{\det Q_t}}\,e^{-\frac{1}{2}(r-H(t)\hat{x}_t)'\,Q_t^{-1}\,(r-H(t)\hat{x}_t)}\,,$$

together with equations for computing $\hat{x}_t$ and Q_t. With a result like this for $D \neq \mathbb{R}^2$, at least a decoder could be implemented for any given control, even a suboptimal one. We note, however, that the long-term goal is to choose $\{\,u_t\,,\,0 \leq t \leq T\,\}$ so as to minimize the probability of error corresponding to the likelihood ratio test in (2.22).

In Chapter 5, we looked at *linear* estimators given doubly-stochastic, time-space, Poisson-process observations. Here, one would like to substitute $\hat{\lambda}_L$ into L_T and to choose a deterministic control $\{\,u(t),\,0 \leq t \leq T\,\}$ to minimize the probability of error corresponding to the likelihood ratio test in (2.22). Even if this could not be done, (2.22) could still be used with a suboptimal, deterministic control. In terms of the Joint Control-Decoding Problem, Theorem 5.2 is of limited usefulness since it considers only deterministic controls. However, if we consider a similar situation in which decoding is the only concern (say $u(t) \equiv 0$) but it is not reasonable to assume $D = \mathbb{R}^2$, and if the integral equation (5.3) can be solved, then $\hat{\lambda}_L(t\,,\,r)$ might be used in the likelihood ratio, L_T, instead of $\mathbf{E}\,[\,\lambda(t\,,\,r\,,\,x_t)\;\mid\;G_t\,]$, for $t \in [0,\,T]$.

II. Future Research

The filtering problems we have discussed have proved to be quite formidable. We present here a modification of our model which we hope will lead to more tractable filtering problems, and perhaps to a solution of the Joint Control-Decoding Problem.

The idea here is to present a discrete-time, discrete-space version of the probabilistic setting introduced in Section III of Chapter 2. As before, let D represent the photosensitive surface of the photodetector. Let $D_1,\,.\,.\,.\,,\,D_K$ be a partition of D into disjoint subrectangles. With each D_k, $1 \le k \le K$, we associate a process $\{ N_t(k);\ t = 0, 1, 2, \ldots \}$ such that $N_0(k) \equiv 0$, and $n_t(k) \triangleq N_t(k) - N_{t-1}(k)$ for $t \ge 1$, takes only the values 0 and 1. We also assume that the events $\{n_t(k) = 1\}$, $1 \le k \le K$, are disjoint, so that simultaneous jumps do not occur. We call $\{ n_t(k);\ t \ge 1,\ 1 \le k \le K \}$ a discrete-time, multivariate point process. Next, let

$$G_t^* = \sigma\{ N_s(k);\ 0 \le s \le t,\ 1 \le k \le K \}.$$

Note that G_0^* is the trivial σ-field, and that for $t \ge 1$, we have $G_t^* = \sigma\{ n_s(k);\ 1 \le s \le t,\ 1 \le k \le K \}$. Let X be a Gaussian random vector with mean m and covariance S. Let $\{v_t,\ t \ge 0\}$ be a sequence of independent, identically distributed, Gaussian random vectors independent of X. Let $\{ u_t,\ t \ge 0 \}$ denote the fine-tracking control signal. We require that u_t be G_{t-1}^*-measurable for $t \ge 1$, and that u_0 be a constant. Let F, G, and V to be known matrices with appropriate dimensions; then let

$$x_{t+1} = F(t)x_t + G(t)u_t + V(t)v_t\ ;\quad x_0 = X.$$

Now define

$$X_t^* = \begin{cases} \sigma\{ X \}\ ; & t = -1 \\ \sigma\{ x_s,\ 0 \le s \le t+1\ ;\ \ N_s(k);\ 0 \le s \le t,\ 1 \le k \le K \}\ ; & t \ge 0 \end{cases}$$

The notable features of the model so far are that $G_t^* \subset X_t^*$, that x_t is X_{t-1}^*-measurable, and that u_t is G_{t-1}^*-measurable. Let

$$\lambda_t^i(k) \triangleq \mathbf{E}\,[\, n_t(k) \mid X_{t-1}^* \,]\ ;\quad t \ge 1\ ,\ \ i = 0, 1,$$

and assume that

$$\lambda_t^i(k) = \mu_i(t) \int_{D_k} \exp[-\tfrac{1}{2}(r - H(t)x_t)'\, R(t)^{-1}\,(r - H(t)x_t)\,]\, dr\ . \tag{6.1}$$

Set

$$\hat{\lambda}_t^i(k) = \mathbf{E}\,[\, \lambda_t^i(k) \mid G_{t-1}^* \,]\ ;\quad t \ge 1. \tag{6.2}$$

We can show that the likelihood ratio, L_t^*, is given by

$$L_t^* = \prod_{s=1}^{t} [\, 1 + \sum_{k=1}^{K} \beta_s(k)(\, n_s(k) - \hat{\lambda}_s^0(k)\,)\,] \, ; \quad t \geq 1, \tag{6.3}$$

where

$$\beta_s(k) \triangleq \frac{\hat{\lambda}_s^1(k)}{\hat{\lambda}_s^0(k)} - \frac{1 - \sum_{j=1}^{K} \hat{\lambda}_s^1(j)}{1 - \sum_{j=1}^{K} \hat{\lambda}_s^0(j)} \, . \tag{6.4}$$

Note that $L_0^* \equiv 1$. Also, since we are assuming that simultaneous jumps do not occur,

$$\sum_{j=1}^{K} \hat{\lambda}_s^i(j) = \mathbf{P}_i \left(\bigcup_{j=1}^{K} \{\, n_t(j) = 1 \,\} \;\middle|\; G_{s-1}^* \right) .$$

Here, $\mathbf{P}_i$ is the probability $\mathbf{P}$ under hypothesis H_i. Clearly, we assume that $\sum_{j=1}^{K} \hat{\lambda}_s^0(j) < 1$ and that $\hat{\lambda}_s^0(k) > 0$ for each $1 \leq k \leq K$. We see from equations (6.3) and (6.4) that we must compute $\hat{\lambda}_t^i(k)$ in order to implement the likelihood ratio test

$$L_t^* \underset{H_0}{\overset{H_1}{\gtrless}} 1 \, , \tag{6.5}$$

even for a suboptimal control. We now state the Discrete-Time Joint Control-Decoding Problem. (Let T be a fixed, positive integer).

Find a rule for generating a control signal, $\{\, u_t \,,\, t = 0, 1, \ldots, T \,\}$, predictable with respect to $\{\, G_t^*, t = 0, 1, \ldots, T \,\}$, such that the probability of error corresponding to the likelihood ratio test in (6.5) is minimized.

Our initial investigation into the computation of (6.2) has led us to the following preliminary results. We have been able to extend a representation result of Bremaud (Exercise E12, equation (3.19) on page 70 of [7], the solution of which is given on page 80 there) to the case of discrete-time, multivariate point processes. Based on this we have extended some of the filtering results of Segall [11] to the case of discrete-time, multivariate point processes. We hope that future research, perhaps along these lines, may yield an explicit formula or a set of recursive equations for computing $\hat{\lambda}_t^i(k)$ in terms of the observations. We hope that eventually this discrete-time version of the Joint Control-Decoding Problem can be rendered more tractable, at least under some conditions yet to be determined.

APPENDIX A

In this appendix we present two proofs of a theorem which we have used throughout the text in order to compute various integrals. The second proof is the simpler one.

Theorem A1. *Suppose that*

$$p(x) = \frac{\exp[-\frac{1}{2}(x-m)' S^{-1}(x-m)]}{(2\pi)^{n/2}\sqrt{\det S}}, \tag{A1}$$

where x and m belong to $\mathbb{R}^n$, and S is a positive definite $n \times n$ matrix. Next, let

$$g(r, x) = \exp[-\frac{1}{2}(r - Hx)' R^{-1}(r - Hx)], \tag{A2}$$

where $r \in \mathbb{R}^k$, H is a $k \times n$ matrix, and R is a positive definite $k \times k$ matrix. Then

$$\int_{\mathbb{R}^n} g(r, x)p(x)\, dx = \frac{\sqrt{\det R}}{\sqrt{\det Q}} \exp[-\frac{1}{2}(r - Hm)' Q^{-1}(r - Hm)], \tag{A3}$$

where

$$Q = HSH' + R .$$

Proof (1). Let

$$f(r) = \int_{\mathbb{R}^n} g(r, x)p(x)\, dx . \tag{A4}$$

Define $B = H' R^{-1} H + S^{-1}$ and set $b = H' R^{-1} r + S^{-1}m$. Note that as a consequence of the hypotheses on R and S, Q and B are positive definite. Now observe that

$$f(r) = \frac{\exp[-\frac{1}{2}(r' R^{-1} r + m' S^{-1} m - b' B^{-1} b)]}{\sqrt{\det S}}$$

$$\times \int_{\mathbb{R}^n} (2\pi)^{-n/2} \exp[-\frac{1}{2}(Bx - b)' B^{-1} (Bx - b)] \, dx$$

$$= \frac{\exp[-\frac{1}{2}(r' R^{-1} r + m' S^{-1} m - b' B^{-1} b)]}{\sqrt{\det S}\sqrt{\det B}} .$$

We now evaluate $b' B^{-1} b$. It is not hard to verify that $B(S - SH' Q^{-1} HS) = I$ and hence that $B^{-1} = S - SH' Q^{-1} HS$ (see [12], p. 656). From the definition of Q, $Q - R = HSH'$, and so

$$R^{-1} - Q^{-1} = Q^{-1} HSH' R^{-1} . \tag{A5}$$

Taking the transpose of (A5) yields

$$R^{-1} - Q^{-1} = R^{-1} HSH' Q^{-1} . \tag{A6}$$

It follows that

$$B^{-1} H' R^{-1} = SH' R^{-1} - SH' Q^{-1} HSH' R^{-1}$$

$$= SH' (R^{-1} - Q^{-1} HSH' R^{-1})$$

using (A5),

$$= SH' Q^{-1} .$$

From these equations,

$$\begin{aligned} B^{-1} b &= B^{-1}(H' R^{-1} r + S^{-1} m) \\ &= B^{-1} H' R^{-1} r + (S - SH' Q^{-1} HS) S^{-1} m \\ &= SH' Q^{-1} r + m - SH' Q^{-1} Hm \\ &= m + SH' Q^{-1} (r - Hm) . \end{aligned}$$

After a little more computation using (A6) we arrive at

$$b' B^{-1} b = r' R^{-1} r + m' S^{-1} m - (r - Hm)' Q^{-1} (r - Hm) .$$

Hence,

$$f(r) = \frac{\exp[-\frac{1}{2}(r - Hm)' Q^{-1} (r - Hm)]}{\sqrt{\det S}\sqrt{\det B}} . \tag{A7}$$

Clearly,

$$\int_{\mathbf{R}^k} f(r)\, dr = \frac{(2\pi)^{k/2}\sqrt{\det Q}}{\sqrt{\det S}\sqrt{\det B}} . \tag{A8}$$

However, by Fubini's Theorem,

$$\begin{aligned} \int_{\mathbf{R}^k} f(r)\, dr &= \int_{\mathbf{R}^n} \int_{\mathbf{R}^k} g(r, x) p(x)\, dr\, dx \\ &= \int_{\mathbf{R}^n} (2\pi)^{k/2} p(x) \sqrt{\det R}\, dx \\ &= (2\pi)^{k/2}\sqrt{\det R} . \end{aligned} \tag{A9}$$

From equations (A7), (A8), and (A9) we see that

$$f(r) = \frac{\sqrt{\det R}}{\sqrt{\det Q}} \exp[-\frac{1}{2}(r - Hm)' Q^{-1} (r - Hm)] .$$

QED

Remark. We concluded from (A8) and (A9) that $\frac{\sqrt{\det Q}}{\sqrt{\det S}\sqrt{\det B}} = \sqrt{\det R}$. This is equivalent to the statement that

$$\det (I_k + R^{-1} H S H') = \det (I_n + S H' R^{-1} H) , \tag{A10}$$

where I_k and I_n are the $k \times k$ and the $n \times n$ identity matrices respectively. It was pointed out to me by a fellow student, D. C. MacEnany, that (A10) follows immediately from the identity (see [12], p. 651) $\det (I_k + \alpha\beta) = \det (I_n + \beta\alpha)$, if we set $\alpha = R^{-1}H$ and $\beta = SH'$.

We now present a second, simpler proof of Theorem A1.

Proof (2). Consider equations (A2) and (A4). The function $f(r)$ is essentially a convolution integral. Let

$$F(\theta) = \int_{\mathbb{R}^k} f(r) e^{j\theta' r} \, dr \ ; \quad \theta \in \mathbb{R}^k .$$

Using Fubini's Theorem,

$$\begin{aligned} F(\theta) &= \int_{\mathbb{R}^n} p(x) \int_{\mathbb{R}^k} g(r, x) e^{j\theta' r} \, dr \, dx \\ &= \int_{\mathbb{R}^n} p(x) (2\pi)^{k/2} \sqrt{\det R} \;\; e^{j\theta' Hx - \frac{1}{2}\theta' R \theta} \, dx \\ &= (2\pi)^{k/2} \sqrt{\det R} \;\; e^{-\frac{1}{2}\theta' R \theta} \int_{\mathbb{R}^n} p(x) e^{j(H'\theta)' x} \, dx \\ &= (2\pi)^{k/2} \sqrt{\det R} \;\; e^{-\frac{1}{2}\theta' R \theta} \, e^{j\theta' Hm - \frac{1}{2}\theta' HSH' \theta} \\ &= (2\pi)^{k/2} \sqrt{\det R} \;\; e^{j\theta' Hm - \frac{1}{2}\theta' Q \theta} . \end{aligned}$$

Taking inverse Fourier transforms, we see by inspection that

$$f(r) = \frac{\sqrt{\det R}}{\sqrt{\det Q}} \exp[-\frac{1}{2}(r - Hm)' \, Q^{-1} (r - Hm)] .$$

QED

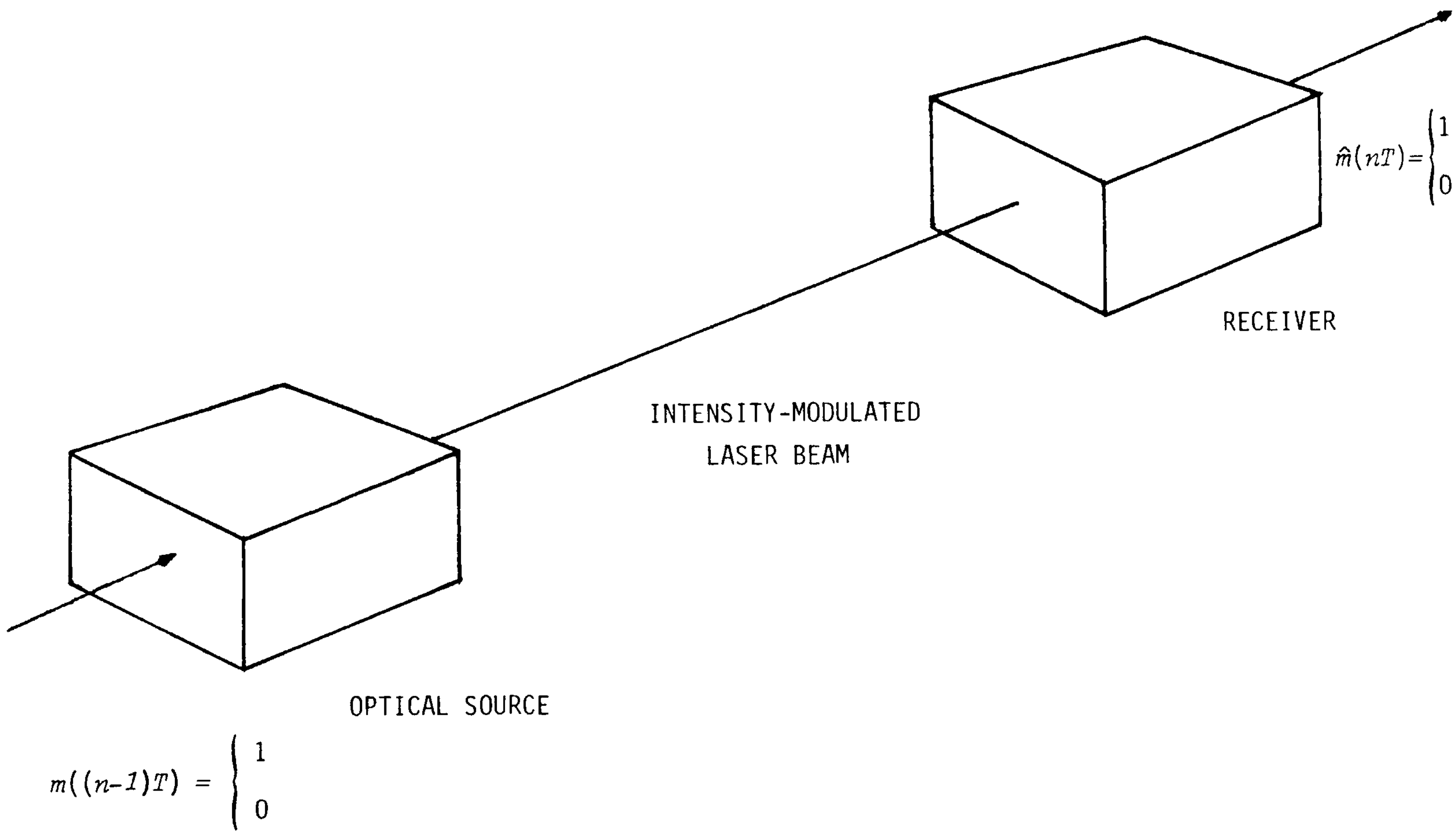

Figure 1. Optical Communication System.

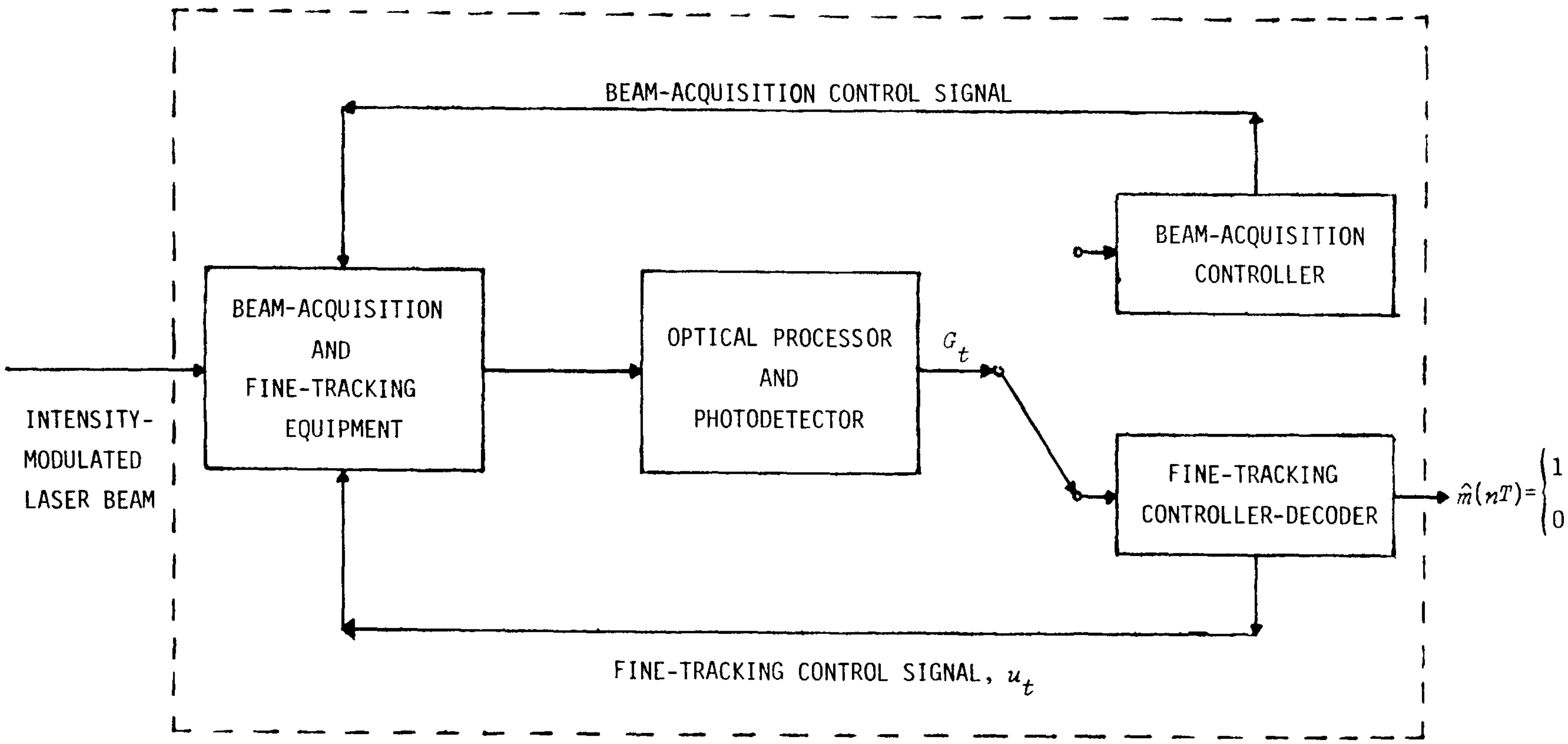

Figure 2. Receiver Subsystems.

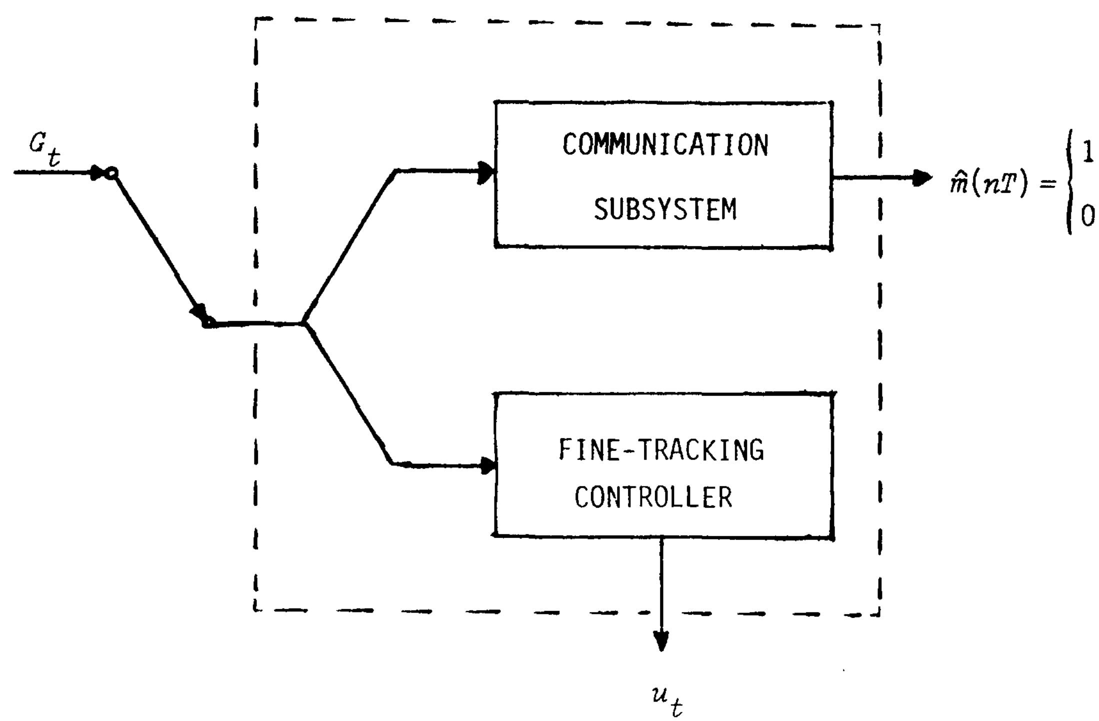

Figure 3. Uncoupled Fine-Tracking Controller-Decoder.

REFERENCES

[1] R. O. Harger, *Optical Communication Theory*, Dowden, Hutchingson, & Ross, Stroudsburg, PA (1977).

[2] D. L. Snyder, "Applications of Stochastic Calculus for Point Process Models Arising in Optical Communication," pp. 789-804 in *Communication Systems and Random Process Theory*, ed. J. K. Skwirzynski, Sijthoff and Noordhoff, Alphen aan der Rijn, The Netherlands (1978).

[3] I. B. Rhodes and D. L. Snyder, "Estimation and Control Performance for Space-Time Point-Process Observations," *IEEE Transactions on Automatic Control*, vol. AC-22, No. 3, pp. 338-346 (June 1977).

[4] R. E. Morley, Jr. and D. L. Snyder, "Maximum Likelihood Sequence Estimation for Randomly Dispersive Channels," *IEEE Transactions on Communications*, vol. COM-27, No. 6, pp. 833-839 (June 1979).

[5] R. K. Boel and V. E. Benes, "Recursive Nonlinear Estimation of a Diffusion Acting as the Rate of an Observed Poisson Process," *IEEE Transactions on Information Theory*, vol. IT-26, No. 5, pp. 561-575 (Sept. 1980).

[6] M. H. A. Davis, *Linear Estimation and Stochastic Control*, Chapman and Hall, London (1977).

[7] P. Bremaud, *Point Process and Queues, Martingale Dynamics*, Springer-Verlag, New York (1981).

[8] D. L. Snyder and P. M. Fishman, "How to Track a Swarm of Fireflies by Observing Their Flashes," *IEEE Transactions on Information Theory*, vol. IT-21, pp. 692-695 (Nov. 1975).

[9] D. L. Snyder, *Random Point Processes*, John Wiley & Sons, New York (1975).

[10] J. Grandell, "A Note on Linear Estimation of the Intensity of a Doubly Stochastic Poisson Field," *Journal of Applied Probability*, vol. 8, pp. 612-614 (1971).

[11] A. Segall, "Recursive Estimation from Discrete-Time Point Processes," *IEEE Transactions on Information Theory*, vol. IT-22, No. 4, pp. 422-431 (July 1976).

[12] T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ (1980).