Estimation of the Rate of a Doubly-Stochastic Time-Space Poisson Process

by

John Gubner and Prakash Narayan
<table>
<thead>
<tr>
<th>1. REPORT DATE</th>
<th>2. REPORT TYPE</th>
<th>3. DATES COVERED</th>
</tr>
</thead>
<tbody>
<tr>
<td>1985</td>
<td></td>
<td>00-00-1985 to 00-00-1985</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. TITLE AND SUBTITLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimation of the Rate of a Doubly-Stochastic Time-Space Poisson Process</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5a. CONTRACT NUMBER</th>
<th>5b. GRANT NUMBER</th>
<th>5c. PROGRAM ELEMENT NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5d. PROJECT NUMBER</th>
<th>5e. TASK NUMBER</th>
<th>5f. WORK UNIT NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>6. AUTHOR(S)</th>
<th>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>University of Maryland, Electrical Engineering Department, College Park, MD, 20742</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>8. PERFORMING ORGANIZATION REPORT NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</th>
<th>10. SPONSOR/MONITOR’S ACRONYM(S)</th>
<th>11. SPONSOR/MONITOR’S REPORT NUMBER(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>12. DISTRIBUTION/AVAILABILITY STATEMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approved for public release; distribution unlimited</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>13. SUPPLEMENTARY NOTES</th>
</tr>
</thead>
<tbody>
<tr>
<td>see report</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>14. ABSTRACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>see report</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>15. SUBJECT TERMS</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>16. SECURITY CLASSIFICATION OF:</th>
<th>17. LIMITATION OF ABSTRACT</th>
<th>18. NUMBER OF PAGES</th>
<th>19a. NAME OF RESPONSIBLE PERSON</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. REPORT</td>
<td>b. ABSTRACT</td>
<td>c. THIS PAGE</td>
<td></td>
</tr>
<tr>
<td>unclassified</td>
<td>unclassified</td>
<td>unclassified</td>
<td></td>
</tr>
</tbody>
</table>

| Standard Form 298 (Rev. 8-98) | Prescribed by ANSI Std Z39-18 |
ESTIMATION OF THE RATE OF A
DOUBLY-STOCHASTIC TIME-SPACE POISSON PROCESS

John Gubner and Prakash Narayan

Electrical Engineering Department
University of Maryland
College Park, Maryland 20742

Abstract

We consider the problem of estimating the rate of a doubly-stochastic, time-space Poisson process when the observations are restricted to a region $D \subseteq \mathbb{R}^2$. In the general case, we obtain a representation of the minimum mean-square-error (MMSE) estimate in terms of the conditional characteristic function of an underlying state process. In the case $D = \mathbb{R}^2$, we extend a known result to compute the MMSE estimate explicitly. For a special form of the rate process, a well-defined integral equation is presented which defines the linear MMSE estimate of the rate.

Key Words: doubly-stochastic, time-space Poisson process, MMSE estimate, linear MMSE estimate, likelihood ratio.

---

1This research was sponsored by the Office of Naval Research under grant no.N0001485-G-0102 and by the Minta Martin Fund for Aerospace Research from the University of Maryland at College Park.
I. Introduction

We consider a doubly-stochastic, time-space Poisson process $\mathbf{N}^0$ with intensity function $\lambda(t, r) = f(t, r - H(t)z_t)$, where $t > 0$ and $r \in \mathbb{R}^2$. Here, $f$ is a known, deterministic function; $z_t \in \mathbb{R}^n$ is the solution of an Ito stochastic differential equation, and $H(t)$ is a known, deterministic, $\mathbb{R}^{2 \times n}$-valued function. The process $\mathbf{N}^0$ under consideration counts events which occur in all of $\mathbb{R}^2$; however, suppose that only those events which occur within a region $D \subseteq \mathbb{R}^2$ can be observed. We wish to compute minimum mean-square-error (MMSE) estimates of $\lambda(t, r)$, given our limited observations. In the general case, $D \neq \mathbb{R}^2$, we obtain a representation of these estimates in terms of the conditional characteristic function of $z_t$.

When $D = \mathbb{R}^2$, and $f(t, r) = e^{-\frac{1}{2}r^T R(t)^{-1} r}$, for some deterministic matrix $R(t)$, we extend a result of Rhodes and Snyder [1] to compute the MMSE estimate of $\lambda(t, r)$ explicitly. We also consider linear estimates of $\lambda(t, r)$ for the same choice of $f$ when $D \neq \mathbb{R}^2$. These filtering problems are frequently encountered in optical communication systems [2, 3], particularly in the context of hypothesis-testing; this issue is discussed in Section V.

II. Probabilistic Setting

Let $\mathcal{B}^2$ denote the Borel subsets of $\mathbb{R}^2$. Next, if $I$ is any interval of $\mathbb{R}$, let $\mathcal{B}(I)$ denote the Borel subsets of $I$. We define $\mathcal{B}(I) \otimes \mathcal{B}^2$ to be the smallest $\sigma$-field containing all sets of the form $E \times A$, such that $E \in \mathcal{B}(I)$ and $A \in \mathcal{B}^2$. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space on which we let

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

be a time-space point process. Sometimes, $\mathbf{N}^0$ is called a random point field or a random measure. Here, this means that with each $B \in \mathcal{B}(0, \infty) \otimes \mathcal{B}^2$, we associate a nonnegative, integer-valued random variable, $N(B) = N(\omega, B)$; in addition, for each $\omega \in \Omega$, $N(\omega, \cdot)$ is assumed to be an integer-valued measure on $\mathcal{B}(0, \infty) \otimes \mathcal{B}^2$. We let $F_t$ represent the times and locations at which points have occurred up to and including time $t$. More precisely, let
\( F_0 \) denote the trivial \( \sigma \)-field, and for \( t > 0 \), set

\[
F_t = \sigma \{ N(B) : B \in \mathcal{B}(0,t) \otimes \mathbb{B}^2 \}.
\]

Now, let \( D \) be a Borel subset of \( \mathbb{R}^2 \). We take \( \mathcal{G}_0 \) to be the trivial \( \sigma \)-field, and for \( t > 0 \), we set

\[
\mathcal{G}_t = \sigma \{ N(B \cap \{ (0,\infty) \times D \}) : B \in \mathcal{B}(0,t) \otimes \mathbb{B}^2 \}.
\]

Note that \( \mathcal{G}_t \) represents the history of the point process restricted to the region \( D \), up to time \( t \). We shall refer to \( \mathcal{G}_t \) as our "observations up to time \( t \)." On the same probability space, \( (\Omega, \mathcal{F}, \mathbb{P}) \), let \( X \) be an \( n \)-dimensional Gaussian random vector 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

\[
dx_t = F(t)x_t dt + V(t)dv_t; \quad x_0 = X.
\]

Here \( F \) and \( V \) are known matrices with appropriate dimensions. We also assume that \( F \) and \( V \) are piecewise-continuous so that a unique solution of (1) exists (see Davis [4], pp. 108-111). Let

\[
\mathbf{X}_0 \triangleq \sigma \{ z_t, 0 \leq s < \infty \}.
\]

For \( t > 0 \), let \( \mathbf{X}_t \) denote the smallest \( \sigma \)-field containing \( F_t \cup \mathbf{X}_0 \). We write this symbolically as

\[
\mathbf{X}_t \triangleq F_t \vee \mathbf{X}_0; \quad t > 0.
\]

We shall assume that \( \mathbb{N}^0 \) is an \( \{ \mathbf{X}_t \} \)-doubly-stochastic, time-space Poisson process, with \( \mathbf{X}_0 \)-measurable intensity (see Bremaud [5], pp. 21-23 and 233-238)

\[
\lambda(t, \tau) = \int f(t, \tau - H(t)x_t) dt,
\]

where \( t \in (0,\infty) \), \( \tau \in \mathbb{R}^2 \), and \( x_t \) is defined by (1). Assume that \( H : (0,\infty) \rightarrow \mathbb{R}^{2 \times n} \) and \( f : (0,\infty) \times \mathbb{R}^2 \rightarrow (0,\infty) \) are deterministic and known. We further assume that the function
\[ \mu(t) \triangleq \int_{\mathbb{R}^2} f(t, r) \, dr \]  

is finite for all \( t < \infty \). This means that for each \( t \geq 0 \), the process

\[ N(t) \triangleq \{ N(B) : B \in \mathcal{B}(t, \infty) \otimes \mathcal{B}^2 \} \]

is a Poisson random field under the measure \( \mathbb{P}(\cdot \mid \mathbf{X}_t) \), with rate \( \lambda(s, r) \), where \( s \in (t, \infty) \), and \( r \in \mathbb{R}^2 \). This implies the following. First, for \( B \in \mathcal{B}(0, \infty) \otimes \mathcal{B}^2 \), let

\[ \Lambda(B) \triangleq \int_{B} \lambda(s, r) \, ds \]  

and if \( B \in \mathcal{B}(t, \infty) \otimes \mathcal{B}^2 \) and \( n \) is an arbitrary, nonnegative integer,

\[ \mathbb{P}(N(B) = n \mid \mathbf{X}_t) = \frac{\Lambda(B)^n}{n!} e^{-\Lambda(B)}. \]  

and hence, for \( \theta \in \mathbb{R} \),

\[ \mathbb{E}[e^{i\theta N(B)} \mid \mathbf{X}_t] = \exp[(e^{i\theta} - 1) \Lambda(B)]. \]  

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

**Notation.** We let \( N_0 \equiv 0 \) and for \( t > 0 \), \( N_t \triangleq N((0, t] \times D) \).

**III. Nonlinear Filtering Results**

We first establish some notation in order to state our results more compactly. Let \( P_t(x) \), \( x \in \mathbb{R}^n \), denote the (regular) conditional probability of \( x_t \) given \( \mathcal{G}_t \). Let \( \psi_t(\eta) \), \( \eta \in \mathbb{R}^n \), denote the conditional characteristic function of \( x_t \) given \( \mathcal{G}_t \):

\[ \psi_t(\eta) \triangleq \mathbb{E}[e^{i\eta \cdot x_t} \mid \mathcal{G}_t] = \int_{\mathbb{R}^n} e^{i\eta \cdot x} \, dP_t(x); \quad \eta \in \mathbb{R}^n. \]

Next, let

\[ \lambda(t, r) \triangleq \mathbb{E}[\lambda(t, r) \mid \mathcal{G}_t] = \mathbb{E}[f(t, r - H(t)x_t) \mid \mathcal{G}_t], \]

and
\[ \hat{l}(t, \theta) \triangleq \int_{\mathbb{R}^2} \hat{\lambda}(t, r) e^{i \theta r} \, dr ; \quad \theta \in \mathbb{R}^2. \]

We also set
\[ F(t, \theta) \triangleq \int_{\mathbb{R}^2} f(t, r) e^{i \theta r} \, dr. \]

**Theorem 1.** Under the foregoing assumptions,

\[ \hat{l}(t, \theta) = F(t, \theta) \psi_t(H(t)^\theta). \]

**Proof.** Observe that
\[
\hat{l}(t, \theta) = \int_{\mathbb{R}^2} \mathbb{E} \left[ f(t, r - H(t)x) \mid G_t \right] e^{i \theta r} \, dr
\]
\[
= \int_{\mathbb{R}^2} \int_{\mathbb{R}^n} f(t, r - H(t)x) \, dP_t(x) \, e^{i \theta r} \, dr.
\]

By Fubini’s Theorem,
\[
\hat{l}(t, \theta) = \int_{\mathbb{R}^n} \int_{\mathbb{R}^2} f(t, r - H(t)x) e^{i \theta r} \, dr \, dP_t(x)
\]
\[
= F(t, \theta) \int_{\mathbb{R}^n} e^{i \theta H(t)x} \, dP_t(x)
\]
\[
= F(t, \theta) \int_{\mathbb{R}^n} e^{i H(t)^\theta y} \, dP_t(x)
\]
\[
= F(t, \theta) \psi_t(H(t)^\theta).
\]

QED

**Theorem 2.** If \( D = \mathbb{R}^2 \), and if

\[ f(t, r) = e^{-\frac{1}{2} r^T R(t)^{-1} r}, \quad (5) \]

for some deterministic, positive-definite matrix \( R(t) \), then
\[ \lambda(t, r) \triangleq \mathbb{E} \left[ \lambda(t, r) \mid \mathcal{G}_t \right] \]
\[ = \mathbb{E} \left[ f(t, r - H(t)x_t) \mid \mathcal{G}_t \right] \]
\[ = \frac{\sqrt{\det R(t)}}{\sqrt{\det Q_t}} \exp \left[ -\frac{1}{2} (r - H(t)\hat{x}_t)' Q_t^{-1} (r - H(t)\hat{x}_t) \right]. \]

where

\[ \hat{x}_t \triangleq \mathbb{E} \left[ x_t \mid \mathcal{G}_t \right], \]
\[ \hat{\Sigma}_t \triangleq \mathbb{E} \left[ (x_t - \hat{x}_t)(x_t - \hat{x}_t)' \mid \mathcal{G}_t \right] > 0, \quad \mathbb{P} - \text{a.s.}, \]
\[ Q_t \triangleq H(t)\hat{\Sigma}_t H(t)' + R(t), \]
and

\[ d\hat{x}_t = F(t)\hat{x}_t dt \]
\[ + \int_{\mathbb{R}^2} \hat{\Sigma}_t H(t - \gamma) Q_t^{-1} (r - H(t - \gamma) \hat{x}_t) N(dt \times dr); \quad \hat{x}_0 = m. \] (6)

\[ d\hat{\Sigma}_t = F(t)\hat{\Sigma}_t dt + \hat{\Sigma}_t F(t)' dt + V(t)V(t)' dt \]
\[ - \hat{\Sigma}_t H(t - \gamma) Q_t^{-1} H(t - \gamma)\hat{\Sigma}_t N(dt \times \mathbb{R}^2); \quad \hat{\Sigma}_0 = S. \] (7)

Proof. First, since \( D = \mathbb{R}^2, \mathcal{G}_t = F_t \). Next, in [1] 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 of the assumption that \( S \) is positive definite) satisfying (6) and (7) above. So,

\[ \psi_t(\eta) = e^{-\frac{1}{2} \eta' \hat{\Sigma}_t \eta}. \]

Next, from equation (5), it follows that

\[ F(t, \theta) = 2\pi \sqrt{\det R(t)} e^{-\frac{1}{2} \theta' R(t) \theta}. \]

Hence, by Theorem 1,

\[ l(t, \theta) = 2\pi \sqrt{\det R(t)} e^{\eta' H(t)\hat{x}_t - \eta' Q_t \eta}. \]

Taking inverse Fourier transforms, we see by inspection that
\[ \hat{\lambda}(t, r) = \frac{\sqrt{\det R(t)}}{\sqrt{\det Q_t}} \exp\left[-\frac{1}{2}(r - H(t) \hat{x}_t)' Q_t^{-1} (r - H(t) \hat{x}_t)\right]. \]

QED

When \( D \neq \mathbb{R}^2 \), or equation (5) does not hold, \( \psi_t(\eta) \) is, in general, not known. This has led us to consider linear estimates of \( \lambda(t, r) \). We discuss this in the next section.

IV. Linear Filtering Results

We call \( \hat{\lambda}_L(t, r) \) a linear estimate of \( \lambda(t, r) \) given \( \mathcal{G}_t \), if \( \hat{\lambda}_L \) can be written in the form

\[ \hat{\lambda}_L(t, r) = \int_0^t \int_D h(t, r; \tau, \rho) \left[ N(d\tau \times d\rho) - \hat{\lambda}(\tau, \rho) d\tau d\rho \right] + h_0(t, r), \quad (8) \]

where \( h \) and \( h_0 \) are deterministic, and \( \hat{\lambda}(t, r) \triangleq \mathbb{E}[\lambda(t, r)] \). We wish to choose \( h \) and \( h_0 \) to minimize

\[ \mathbb{E}\left[ |\lambda(t, r) - \hat{\lambda}_L(t, r)|^2 \right]. \quad (9) \]

Lemma 1. (Grandell [6]). Let \( \hat{\lambda}_L(t, r) \) be given by (8). Under the conditions outlined in Section II, the quantity in (9) will be minimized if \( h_0(t, r) = \hat{\lambda}(t, r) \), and if \( h \) satisfies

\[ \Gamma(t, r; \tau, \rho) = \int_0^t \int_D h(t, r; \sigma, \varsigma) \Gamma(\sigma, \varsigma; \tau, \rho) d\varsigma d\sigma + h(t, r; \tau, \rho) \hat{\lambda}(\tau, \rho), \quad (10) \]

where

\[ \Gamma(t, r; \tau, \rho) \triangleq \text{cov}\left[\lambda(t, r), \lambda(\tau, \rho)\right]. \]

With Lemma 1 in mind, we state our Theorem 3.

Theorem 3. If \( f(t, r) \) is given by (5), and the conditions outlined in Section II hold, then
\[
\bar{x}(t, r) = \frac{\sqrt{\text{det} R(t)}}{\sqrt{\text{det} Q(t)}} \exp\left[ -\frac{1}{2} (r - H(t) \bar{x}(t))^\top Q(t)^{-1} (r - H(t) \bar{x}(t)) \right],
\]

where

\[
\bar{x}(t) \triangleq \mathbb{E} [x_t],
\]

\[
\Sigma(t) \triangleq \text{cov} [x_t],
\]

\[
Q(t) \triangleq H(t) \Sigma(t) H(t)^\top + R(t).
\]

Furthermore,

\[
\Gamma(t, r; \tau, \rho) + \bar{x}(t, r) \bar{x}(\tau, \rho) = \sqrt{\frac{\text{det} R(t)}{\text{det} Q(t, r)}} \times
\exp\left[ -\frac{1}{2} \begin{pmatrix} \tau \\ \rho \end{pmatrix} - \begin{bmatrix} H(t) & 0 \\ 0 & H(\tau) \end{bmatrix} \begin{bmatrix} \bar{x}(t) \\ \bar{x}(\tau) \end{bmatrix} \right] Q(t, r)^{-1} \left( \begin{pmatrix} \tau \\ \rho \end{pmatrix} - \begin{bmatrix} H(t) & 0 \\ 0 & H(\tau) \end{bmatrix} \begin{bmatrix} \bar{x}(t) \\ \bar{x}(\tau) \end{bmatrix} \right),
\]

where

\[
\Sigma(t, \tau) \triangleq \text{cov} [x_t, z_r],
\]

and

\[
Q(t, \tau) \triangleq \begin{bmatrix} Q(t) & H(t) \Sigma(t, \tau) H(\tau)^\top \\ H(\tau) \Sigma(t, \tau) H(\tau)^\top & Q(\tau) \end{bmatrix}.
\]

**Proof.** For completeness, we make the following observations. Recall that

\[
dx_t = F(t) x_t \, dt + V(t) \, d\eta_t; \quad x_0 = X.
\]

Let \( \Phi(t_2, t_1) \) be the transition matrix corresponding to \( F(t) \). Then

\[
\bar{x}(t) = \Phi(t, 0) m,
\]

and
\[ \Sigma(t, \tau) = \Phi(t, 0) S \Phi(t, 0)^t + \int_0^{\min(t, \tau)} \Phi(t, s) V(s) V(s)^t \Phi(t, s)^t \, ds . \]

Note that \( \Sigma(t) = \Sigma(t, t) \).

To compute \( \bar{\lambda}(t, \tau) = E[\lambda(t, \tau)] \), observe that \( z_t \) is Gaussian with mean \( \bar{x}(t) \) and covariance \( \Sigma(t) \). By considering the proofs of Theorem 1 and Theorem 2, equation (11) is immediate.

The computation of (12) is similar, but requires some judicious preliminary arithmetic.

First, observe that \( \Gamma(t, \tau; \tau, \rho) + \bar{\lambda}(t, \tau) \bar{\lambda}(\tau, \rho) \) is just another way of writing \( E[\lambda(t, \tau) \lambda(\tau, \rho)] \). Next, rewrite \( \lambda(t, \tau) \lambda(\tau, \rho) \) as

\[
\exp\left[-\frac{1}{2} \begin{bmatrix} r \\ \rho \end{bmatrix} \begin{bmatrix} H(t) & 0 \\ 0 & H(\tau) \end{bmatrix} \begin{bmatrix} x_t \\ x_\tau \end{bmatrix} \begin{bmatrix} R(t)^{-1} & 0 \\ 0 & R(\tau)^{-1} \end{bmatrix} \begin{bmatrix} r \\ \rho \end{bmatrix} - \begin{bmatrix} H(t) & 0 \\ 0 & H(\tau) \end{bmatrix} \begin{bmatrix} x_t \\ x_\tau \end{bmatrix} \right],
\]

which is equal to

\[
\exp\left[-\frac{1}{2} \begin{bmatrix} r \\ \rho \end{bmatrix} \begin{bmatrix} H(t) & 0 \\ 0 & H(\tau) \end{bmatrix} \begin{bmatrix} x_t \\ x_\tau \end{bmatrix} \begin{bmatrix} R(t) & 0 \\ 0 & R(\tau) \end{bmatrix}^{-1} \begin{bmatrix} r \\ \rho \end{bmatrix} - \begin{bmatrix} H(t) & 0 \\ 0 & H(\tau) \end{bmatrix} \begin{bmatrix} x_t \\ x_\tau \end{bmatrix} \right]. \tag{15}
\]

Because \( \{ z_t, t \geq 0 \} \) 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 \( \begin{bmatrix} \Sigma(t) & \Sigma(t, \tau) \\ \Sigma(\tau, t) & \Sigma(\tau) \end{bmatrix} \). By the same reasoning used to deduce (11), (12) also follows.

**QED**

**Remark.** In equation (10), if we regard \( t \) and \( \tau \) 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 \( h \).
V. Discussion

The filtering problems considered above often arise in the design and implementation of receivers for optical communication systems. Typically, a binary message source is used by a transmitter to select the modulation of the intensity of a laser beam in accordance with whether a "0" or a "1" is to be sent. The laser beam travels to a receiver and strikes its photodetector. We assume that the laser beam has an intensity profile of the form

$$\nu_i(t) f(t, r) ; \quad i = 0, 1.$$ 

Here, $\nu_i(t)$ is a known, deterministic function, where $i = 0$ or $1$ has been selected by the transmitter.

We model the surface of the receiver's photodetector as $\mathbb{R}^2$. If the receiver, for example, is subject to vibrations, the center of the spot of laser light may wander randomly over the photodetector surface [2]. We assume, as in [2], that the center of the spot of laser light is given by $H(t)x_t \in \mathbb{R}^2$. The output of photoelectrons from the photodetector is modeled by the process $N^0$, with stochastic intensity now given by

$$\lambda_i(t, r) = \nu_i(t) f(t, r - H(t)x_t). \quad (16)$$

Of course, an actual photodetector does not have an infinite photosensitive surface. We account for this fact by assuming that only those photoelectrons which occur in a region $D \subseteq \mathbb{R}^2$ are observed. For example, in this setting, $D$ might be a square or a circle centered at the origin. After observing photoelectrons occurring in $D$ during some time interval $[0, T]$, a decision as to whether a "0" or a "1" was sent has to be made based on one of the estimates $\hat{\lambda}_i(t, r)$ or $\hat{\lambda}_{i, L}(t, r)$. As an example of a decoding scheme, we could use the likelihood ratio test

$$\frac{H_1}{H_0}$$

$$L_T > 1,$$
to make the decision, using the minimum probability of error cost criterion and assuming equiprobable hypotheses (see Snyder [3], section 2.5). The likelihood ratio, \( L_T \), is given by (see Snyder [3], pp. 471-476)

\[
L_T = \frac{\prod_{j=1}^{N_T} \lambda_i(t_j, r_j) \exp[-\int_0^T \int_D \hat{\lambda}_i(s, r) \, dr \, ds]}{\prod_{j=1}^{N_T} \lambda_0(t_j, r_j) \exp[-\int_0^T \int_D \hat{\lambda}_0(s, r) \, dr \, ds]},
\]

where \( t_j \) and \( r_j \) are respectively the time and the location of the \( j \)th photoevent in the region \( D \), and we adopt the convention that when \( N_T = 0 \), the factors preceding \( \exp \) in equation (17) are taken to be unity. Here, of course,

\[
\hat{\lambda}_i(t, r) \triangleq \mathbb{E} \left[ \lambda_i(t, r) \mid \mathcal{G}_t \right]; \quad i = 0, 1.
\]

Now, using (16), (17) simplifies to

\[
L_T = \prod_{j=1}^{N_T} \frac{\nu_i(t_j)}{\nu_0(t_j)} \exp[-\int_0^T \int_D \hat{\lambda}_i(s, r) - \hat{\lambda}_0(s, r) \, dr \, ds].
\]

In the general case, \( D \neq \mathbb{R}^2 \), \( \hat{\lambda}_i(t, r) \) is not known, and hence, \( L_T \) cannot be computed. However, when \( D = \mathbb{R}^2 \), it turns out that we do not need to know \( \hat{\lambda}_i(t, r) \) in order to compute \( L_T \). Observe that if \( D = \mathbb{R}^2 \), then

\[
\int_D \hat{\lambda}_i(s, r) - \hat{\lambda}_0(s, r) \, dr = \mathbb{E} \left[ \int_{\mathbb{R}^2} \lambda_i(s, r) - \lambda_0(s, r) \, dr \mid \mathcal{G}_s \right]
\]

\[
= \mathbb{E} \left[ (\nu_i(s) - \nu_0(s)) \int_{\mathbb{R}^2} f(s, r - H(s)z_s) \, dr \mid \mathcal{G}_s \right]
\]

\[
= \mathbb{E} \left[ (\nu_i(s) - \nu_0(s)) \mu(s) \mid \mathcal{G}_s \right]
\]

\[
= \mu(s) \left[ \nu_i(s) - \nu_0(s) \right].
\]

In equation (20) we used the fact that for all \( r_0 \in \mathbb{R}^2 \),

\[
\mu(s) \triangleq \int_{\mathbb{R}^2} f(s, r) \, dr = \int_{\mathbb{R}^2} f(s, r - r_0) \, dr.
\]

Thus, when \( D = \mathbb{R}^2 \), (19) becomes
\[ L_T = \prod_{i=1}^{N_T} \frac{\nu_i(t_i)}{\nu_0(t_i)} \exp\left[-\int_0^T \mu(s)\left(\nu_1(s) - \nu_0(s)\right) ds\right]. \] (21)

With (21) in mind, consider the following theorem.

**Theorem 4.** The random field

\[ \mathbf{M}^t \sim \{ N(E \times \mathbb{R}^2) : E \in \mathcal{B}(t, \infty) \}, \]

is independent of the \( \sigma \)-field \( \mathbf{X}_t \).

**Proof.** To prove that \( \mathbf{M}^t \) is independent of \( \mathbf{X}_t \), it is sufficient to show that the conditional characteristic function of \( N(E \times \mathbb{R}^2) \) is deterministic for \( E \in \mathcal{B}(t, \infty) \). Now, it follows immediately from the assumption that \( \mathbf{N}^0 \) is an \( \{\mathbf{X}_t\} \)-doubly-stochastic, time-space Poisson process, that for \( \theta \in \mathbb{R} \),

\[ \mathbb{E} \left[ e^{i\theta N(E \times \mathbb{R}^2) | \mathbf{X}_t} \right] = \exp\left[(e^{i\theta} - 1) \int_E \int_{\mathbb{R}^2} \lambda_i(s, r) dr ds \right] \]

\[ = \exp\left[(e^{i\theta} - 1) \int_E \nu_i(s) \int_{\mathbb{R}^2} f(s, r - H(s)z_s) dr ds \right] \]

\[ = \exp\left[(e^{i\theta} - 1) \int_E \nu_i(s) \mu(s) ds \right]. \]

Hence \( \mathbf{M}^t \) is independent of \( \mathbf{X}_t \).

QED

It follows from equation (21) and Theorem 4 that for all \( t \geq 0 \), the random variable \( L_t \) is independent of the \( \sigma \)-field \( \mathbf{X}_t \).

If we replace equation (1) by

\[ dx_t = F(t)z_t dt + G(t)u_t dt + V(t)du_t; \quad x_0 = X, \] (22)

where \( \{ u_t, t \geq 0 \} \) is predictable with respect to \( \{ \mathbf{E}_t, t \geq 0 \} \) and \( G(t) \) is a known matrix with appropriate dimensions, then most of the above results hold with only minor
modifications. The term $G(t)u_t$ in (22) is interpreted as a control signal driven by the output of the photodetector. Since $H(t)x_t$ represents the center of the spot of laser light striking the receiver, one might try to use $G(t)u_t$ to drive $z_t$ to the origin. This problem is addressed in [1]. If (1) is replaced by (22), Theorem 1 is unchanged. Theorem 2 still holds except that equation (6) must be replaced by

$$\dot{x}_t = F(t)x_t dt + G(t)u_t dt$$

$$+ \int_{\mathbb{R}^2} \sum_l H(t-l)p_{l}^{-1}(r-H(t-l)x_t)N(dt \times dr); \quad x_0 = m.$$  

Lemma 1 is unchanged, and if $u_t = u(t)$ for some deterministic control $\{u(t), t \geq 0\}$, then Theorem 3 holds; of course, (13) becomes (22) and (14) is replaced by

$$\mathcal{F}(t) = \Phi(t,0)m + \int_0^t \Phi(t,s)G(s)u(s)ds.$$  

In addition, the results of the preceding paragraphs of Section V, including Theorem 4, are unchanged by substituting equation (22) for equation (1). Note also that since $G_t \subseteq X_t$, and $L_t$ is independent of $X_t$ when $D = \mathbb{R}^2$, it follows that $L_T$ is independent of the control law $\{u_t, 0 \leq t \leq T\}$ when $D = \mathbb{R}^2$. This implies that the probability of a decoding error corresponding to the likelihood ratio test preceding equation (17) is not a function of the control law $\{u_t, 0 \leq t \leq T\}$ when $D = \mathbb{R}^2$. In this sense, all controls are optimal, when $D = \mathbb{R}^2$. In general, when $D \neq \mathbb{R}^2$, this is not to be expected.

REFERENCES


