XI. DETECTION AND ESTIMATION THEORY*

Academic and Research Staff

Prof. H. L. Van Trees
Prof. A. B. Baggeroer
Prof. D. L. Snyder

Graduate Students

I. L. Bruyland
A. E. Eckberg
J. E. Evans
T. P. McGarty
J. R. Morrissey

A. PRELIMINARY RESULTS ON PERFORMANCE BOUNDS FOR THE DETECTION OF STOCHASTIC SIGNALS IN ADDITIVE WHITE GAUSSIAN NOISE

1. Introduction

The usefulness of tilted probability distributions and the semi-invariant moment-generating function $\mu(s)$ in bounding error performance for the detection of Gaussian signals in additive Gaussian noise has been considered in recent theses and publications.[1-4] The approach that has been used to obtain $\mu(s)$ requires that the signal process be Gaussian, and hence is rather restrictive. In this report, we present a new approach to computing $\mu(s)$ that may be quite useful for the class of problems for which the signal is a non-Gaussian Markov random process.

We consider the following canonical Bayes detection problem. Given observations $\{r(t), 0 \leq t \leq T\}$, determine which of the following hypotheses[1] is true:

$$H_1 : \dot{y}(t) = r(t) = z_1(t) + \dot{b}(t)$$

$$H_0 : \dot{y}(t) = r(t) = z_0(t) + \dot{b}(t) \qquad t \in I \triangleq [0, T], \tag{1}$$

where $\dot{b}(t)$ is a sample function of white Gaussian noise (with unit spectral density), and $\{z_k(\cdot), k=0, 1\}$ is a random process with known (not necessarily Gaussian) probability distribution such that

$$\int_t E[z_k(t) - Ez_k(t)]^2 < \infty \int_t E[z_k(t)]^2 \, dt < \infty. \tag{2}$$

Dots on b and y indicate that Eq. 1 represents a formal division by dt of the Itô equation[5, 6]

$$H_k : dy(t) = z_k(t) \, dt + db(t).$$

*This work was supported in part by the Joint Services Electronics Programs (U. S. Army, U. S. Navy, and U. S. Air Force) under Contract DA 28-043-AMC-02536(E).

Here, $b(t)$ is the Wiener, or Brownian motion, process. All stochastic differential equations are to be interpreted in the sense given by Itô.[7] A more recent discussion of his formulation has been given by Wonham.[8]

Furthermore, we assume that $\{z_k(t), t \in I\}$ and $\{\dot{b}(t), t \in I\}$ are statistically independent.

The likelihood ratio for this detection problem when $z_o(t) = 0$ has been shown[5,6] to be

$$\Lambda(T) = \exp\left[\int_I \hat{z}_1(u)\, dy(u) - \frac{1}{2}\int_I \hat{z}_1^2(u)\, du\right], \tag{3}$$

where

$$\hat{z}_1(t) = E[z_1(t)|\{r(\tau), \tau < t\}, H_1]$$

= the conditional mean of $z(t)$, given observations up to the instant t and under the assumption that hypothesis H_1 is true

= realizable least-squares estimate (not necessarily linear) of $z_1(t)$ when H_1 is true.

If H_1 is not the true hypothesis, then $\hat{z}_1(t)$ is not the least-squares estimate of $z_1(t)$, and might be called a "pseudo estimate" of $z_1(t)$. The bar in the first integral sign emphasizes that the integral is to be interpreted as an Itô stochastic integral. Kailath[6] gives an extensive discussion of (3), including a demonstration that (3) includes all previously known explicit formulas for signals in white Gaussian noise, as well as some discussion of the Itô definition of the stochastic integral.

Applying the "chain rule" for likelihood ratios to (3), we find that the likelihood ratio for the detection problem when $z_o(t) \neq 0$ is

$$\Lambda(T) = \exp\left\{\int [\hat{z}_1(u)-\hat{z}_o(u)]\, dy(u) - \frac{1}{2}\int \left[\hat{z}_1^2(u)-\hat{z}_o^2(u)\right] du, \tag{4}$$

where $\hat{z}_o(u)$ is the least-squares estimate of $z_o(u)$ when H_o is true.

We shall give some new results on the time evolution of the semi-invariant moment generating function:

$$\mu_o(s, T) = \ln E_{H_o}[e^{s\ell(T)}] = \ln E_{H_o}[\Lambda(T)^s], \tag{5}$$

where $\ell(T) = \ln \Lambda(T)$ is the log-likelihood ratio. The key ideas in our approach to the problem are the following.

1. The system that generates $\hat{z}_o$ and $\hat{z}_1$ is assumed to have a state-variable description of the type to be given in Eqs. 6 and 7.

2. That $\Lambda(T)^s$, $\hat{z}_o(t)$ and $\hat{z}_1(t)$ form a joint Markov process when $\hat{z}_o(t)$ is the realizable

least-squares estimate is used to obtain a partial differential equation related to the time evolution of $\mu_o(s, T)$.

3. Rather than work directly with the expectation of Eq. 5, we work with a certain conditional expectation called a quasi-transition function that has many of the properties of a transition density function. In particular, the time evolution of the quasi-transition function is given by a Fokker-Planck type of equation whose solution is closely related to $\mu_o(s, T)$.

We shall derive the basic equation for the time evolution of the quasi-transition function, and discuss some methods of solving the basic equation to give the time evolution of $\mu_o(s, T)$.

2. Derivation of the Basic Equation

We shall derive the basic equation for the time evolution of the quasi-transition function. The derivation proceeds as follows: First, we give the state-variable description of the dynamic system that generates the estimates $z_k(t)$ from the received data. From this state-variable representation, we can obtain the Fokker-Planck equation for the time evolution of the transition density function of the joint Markov process $\{\hat{z}_o(t), \hat{z}_1(t), \phi(t) = \Lambda(t)^s\}$ when H_o is true. From this Fokker-Planck equation we obtain the basic equation for the time evolution of the quasi-transition function.

We assume that the system that generates the least-squares estimate, $\hat{z}_k(t)$, is well modeled by the vector stochastic differential equation

$$\hat{z}_k(t) = \hat{h}_k(\hat{\underline{x}}_k(t), t) \tag{6}$$

$$d\hat{\underline{x}}_k(t) = \hat{\underline{f}}_k[t:\hat{x}_k(t)]\, dt + \hat{\underline{g}}_k[t:\hat{\underline{x}}_k(t)]\, [dy(t)-\hat{z}_k(t)], \tag{7}$$

where

$\hat{\underline{x}}_k(t)$ = estimator state vector with state $\hat{\underline{x}}_k(0)$ at $t = 0$

$\hat{\underline{f}}_k$, $\hat{\underline{g}}_k$ = memoryless functions of time that are possibly nonlinear functions of $\underline{x}_k(t)$

$\hat{h}_k$ = memoryless transformation of the $\hat{\underline{x}}_k(t)$.

Least-squares estimates have been discussed by several authors.[8-11] In many cases, an explicit formula for $\hat{z}(t)$ is not currently known, so we do not know how restrictive (6) and (7) are. In the only known cases for which explicit results exist,[8,9] the estimator is of the form given above. Approximate least-squares estimates of this form have been given by many people (see, for example, Snyder[11]).

Furthermore, we assume that $\hat{\underline{f}}_k$, $\hat{\underline{g}}_k$ and $\hat{h}_k$ satisfy certain conditions, so that when H_o is true, the $\hat{\underline{x}}_k$ and $\hat{\underline{z}}_k$ are continuous Markov processes whose transition density function exists and satisfies Kolmogorov's forward equation. Sufficient conditions have been given by Duncan.[12] The principal requirement is that various components of $\hat{\underline{f}}$ and $\hat{\underline{g}}$ and certain of their partial derivatives be globally Lipschitz continuous in $\hat{\underline{x}}_k$.

When hypothesis H_o is true, from Eq. 1 we have

$$dy(t) = z_o(t) + db(t). \tag{8}$$

Using the results of Frost[13] (see also Kailath[6]), we can reformulate (8) as

$$dy(t) = \hat{z}_o(t) + dv(t), \tag{9}$$

where the "innovation process" $\{v(t)\}$ is a Wiener process with the same intensity as b(t).

Next, for ease of motion, we define a new state vector $\hat{\underline{x}}$ to be the vector obtained by adjoining the vectors $\hat{\underline{x}}_o$ and $\hat{\underline{x}}_1$. More precisely, let

$$\hat{\underline{x}} = \begin{bmatrix} \hat{\underline{x}}_o \\ \hat{\underline{x}}_1 \end{bmatrix}. \tag{10}$$

Then, from (6), (7) and (9), we can reformulate our state equations as

$$\hat{z}_k(t) = \hat{z}_k(\hat{\underline{x}}(t)) = \hat{h}_k(\hat{\underline{x}}_k) \tag{11}$$

$$d\hat{\underline{x}}(t) = \hat{\underline{f}}[t:\underline{x}(t)]\, dt + \hat{\underline{g}}[t:\hat{\underline{x}}(t)]\, dv, \tag{12}$$

where

$$\hat{\underline{f}} = \begin{bmatrix} \hat{\underline{f}}_o(\hat{\underline{x}}_o) \\ \hline \hat{\underline{f}}_1(\hat{\underline{x}}_1) - \hat{\underline{g}}_1(\hat{\underline{x}}_1)\, \hat{h}_o(\hat{\underline{x}}_o) \end{bmatrix} \tag{13}$$

$$\hat{\underline{g}} = \begin{bmatrix} \hat{\underline{g}}_o(\hat{\underline{x}}_o) \\ \hline \hat{\underline{g}}_1(\hat{\underline{x}}_1) \end{bmatrix}. \tag{14}$$

The forward Kolmogorov equation for the transition density function, $p(\hat{\underline{x}}_t, t | \hat{\underline{x}}(0))$, of $\hat{\underline{x}}(t)$ (Dynkin[14]) is

$$\frac{\partial}{\partial t} p(\hat{\underline{x}}_t, t | \hat{\underline{x}}(0)) = L^+_{\hat{\underline{x}}}\, p(\hat{\underline{x}}_t, t | \hat{\underline{x}}(0)), \tag{15}$$

where

$$L^+_{\hat{\underline{x}}} = -\sum \frac{\partial}{\partial x_i} [\hat{f}_i(\hat{\underline{x}})(\ :\ )] + \frac{1}{2} \sum \sum_j \frac{\partial^2}{\partial x_i \partial x_j} [c_{ij}(t, \hat{\underline{x}}_t)(\ \cdot\ )] \tag{16}$$

$$x_i = (\hat{\underline{x}}_t)_i \tag{17}$$

$$\underline{c} = \hat{\underline{g}}^t \hat{\underline{g}}. \tag{18}$$

We shall now proceed to derive a partial differential equation satisfied by the quasi-transition function

$$r[\hat{\underline{x}}_t, t|\hat{\underline{x}}(0)] \triangleq E[\phi_t|\hat{\underline{x}}_t, t_1 \hat{\underline{x}}(0)]\ p(\hat{\underline{x}}_t, t|\hat{\underline{x}}(0)), \tag{19}$$

where

$$\phi_t = \exp\left\{s \oint [\hat{z}_1 - \hat{z}_o]\ dv - \frac{s}{2} \int [z_1 - z_o]^2\ du\right\} \tag{20}$$

and $E[\,.\,|\,.\,]$ is the conditional expectation of the multiplicative functional ϕ_t for all sample paths, with $\hat{\underline{x}}(t=0) = \hat{\underline{x}}(0)$ and $\hat{\underline{x}}(t) = \hat{\underline{x}}_t$. The form given for ϕ_t can be obtained by substituting (9) in (4). Note that $r[\,.\,|\,.\,]$ is related to $\mu_o(s,t)$ by

$$\mu_o(s,t) = \ln \int r[\hat{\underline{x}}_t, t|\hat{\underline{x}}(0)]\ d\hat{\underline{x}}_t d\hat{\underline{x}}(0). \tag{21}$$

Some of the properties of such a function where

$$\phi_t = \exp\left[\int_0^t V(\hat{z}_u)\ du\right],$$

with $V(\cdot)$ continuous and nonpositive, were first discussed by Kac[15] following their introduction by Feynman[13] in the form of the "Feynman integral." The use of such functions in quantum mechanics has been discussed by Gelfand and Yaglom,[16] Kac,[17] and Feynman and Hibbs.[18] Dynkin[19] discusses the properties of such functionals from the viewpoint of Markov process theory. In work more closely aligned to that at hand, Ray[20] has used the probabilistic properties of certain quasi-transition functions to obtain results regarding second-order linear differential operators, while Siegert,[21,22] in conjunction with Darling and Kac, has demonstrated the utility of such functions in computing the probability distribution of Gaussian processes after they have passed through certain nonlinear devices.

The partial differential equation for $r(\,.\,|\,.\,)$ will be obtained by making use of the fact that ϕ_o^t and $\hat{\underline{x}}_t$ are jointly Markovian. Applying Itô's differential rule[6] to (20), we have

$$d\phi = \frac{s^2 - s}{2} [\hat{z}_1(\hat{\underline{x}}) - \hat{z}_o(\hat{\underline{x}})]^2\ \phi\ dt + s[\hat{z}_1(\hat{\underline{x}}) - \hat{z}_o(\hat{\underline{x}})]\ \phi\ dv. \tag{22}$$

Equations 11, 12, and 22 define the vector Markov process $(\phi, \hat{\underline{x}}_t)$. We assume that the form of $\hat{z}_o(\hat{\underline{x}})$ and $\hat{z}_1(\hat{\underline{x}})$ is such that the transition density of $(\phi, \hat{\underline{x}})$ exists and satisfies the forward Kolmogorov equation

$$\frac{\partial p(\phi, \hat{\underline{x}}_t, t|\hat{\underline{x}}(0))}{\partial t} = L^+_{\hat{\underline{x}}}\, p(\phi, \hat{\underline{x}}_t, t|\hat{\underline{x}}(0)) - \frac{\partial}{\partial \phi}\left[\frac{s^2-s}{2}(\hat{z}_1-\hat{z}_o)^2\, p(\phi, \hat{\underline{x}}_t, t|\hat{\underline{x}}(0))\right]$$

$$+ \sum_i \frac{\partial^2}{\partial\phi\partial\hat{x}_i}\left[s(\hat{z}_1-\hat{z}_o)\, g_i(\hat{\underline{x}}_t, t)\, \phi p(\phi, \hat{\underline{x}}_t, t|\hat{\underline{x}}(0))\right]$$

$$+ \frac{1}{2}\frac{\partial^2}{\partial\phi^2}\left[s^2(\hat{z}_1-\hat{z}_o)^2\, \phi^2 p(\phi, \hat{\underline{x}}_t, t|\hat{\underline{x}}(0))\right]. \tag{23}$$

This can be obtained by noting that the forward operator for $p[\phi, \hat{\underline{x}}_t, t|\hat{\underline{x}}(0)]$ differs from the operator $L^+_{\hat{\underline{x}}}$ by the fact that there are additional additive terms involving partial derivatives with respect to ϕ. The particular form of these additional terms is specified by Eqs. 11, 12, and 22.

Equation 23 allows us to determine a partial differential equation satisfied by $r[\,.\,|\,.\,]$. First, we put Eq. 19 into a more convenient form:

$$r[\hat{\underline{x}}_t, t|\hat{\underline{x}}(0)] = \left\{\int \phi\, \frac{p(\phi, \hat{\underline{x}}_t, t|\hat{\underline{x}}(0))}{p(\hat{\underline{x}}_t, t|\hat{\underline{x}}(0))}\, d\phi\right\} p(\hat{\underline{x}}_t, t|\hat{\underline{x}}(0))$$

$$= \int \phi p(\phi, \hat{\underline{x}}_t, t|\hat{\underline{x}}(0))\, d\phi. \tag{24}$$

From (19), we see that

$$\frac{\partial r[\,.\,|\,.\,]}{\partial t} = \int \phi\, \frac{\partial p(\phi, \hat{\underline{x}}_t, t|\hat{\underline{x}}(0))}{\partial t}\, d\phi. \tag{25}$$

Substituting (23) in (25) and integrating by parts with appropriate boundary conditions (for example, (suppressing arguments): $p\,\Big|_{-\infty}^{\infty} = \frac{\partial p}{\partial \hat{x}_i}\Big|_{-\infty}^{\infty} = \hat{f}_k p\,\Big|_{-\infty}^{\infty} = \hat{g}_k p\,\Big|_{-\infty}^{\infty} = 0$), we obtain

$$\frac{\partial r[\hat{\underline{x}}_t, t|\hat{\underline{x}}(0)]}{\partial t} = L^+_{\hat{\underline{x}}}\, r(\hat{\underline{x}}_t, t|\hat{\underline{x}}(0)) + \frac{s^2-s}{2}\left[\hat{z}_1(\hat{\underline{x}}_t)-\hat{z}_o(\hat{\underline{x}}_t)\right]^2 r(\hat{\underline{x}}_t, t|\hat{\underline{x}}(0))$$

$$- s\sum_i \frac{\partial}{\partial\hat{x}_i}\left\{\left[\hat{z}_1(\hat{\underline{x}}_t)-\hat{z}_o(\hat{\underline{x}}_t)\right]\hat{g}_i(\hat{\underline{x}}_t, t)\, r(\hat{\underline{x}}_t, t|\hat{\underline{x}}(0))\right\}. \tag{26}$$

Equation 26 is the basic equation for the time evolution of the quasi-transition function $r[\,.\,|\,.\,]$. We shall discuss various methods of solving (26) to obtain $\mu_o(s, t)$.

First, we shall discuss some issues related to the least-squares estimates $\hat{z}_k(t)$. As indicated earlier, explicit formulas for the $\hat{z}_k(t)$ are, at present, unknown for most cases when $z_k(t)$ is non-Gaussian. There is great interest in determining the performance of detection systems when approximate least-squares estimates are used in the likelihood-ratio formula (4). Our development leading to (26) did not make use of the fact that $\hat{z}_1(t)$ was the exact least-squares estimate. Thus (26) can be used to bound the error performance of the (suboptimal) receiver if $\hat{z}_1(t)$ is an approximate least-squares estimate, with $\hat{z}_o(t)$ the exact least-squares estimate.

When $\hat{z}_o(t)$ is an approximate least-squares estimate, we cannot use the "innovation process" representation of dy(t) to eliminate $z_o(t)$ from Eq. 9 and to obtain Eq. 10. Thus we must assume that $z_o(t)$ has an explicit representation such that z_o, $\hat{z}_o$, $\hat{z}_1$, and $\Lambda(t)$ form a joint Markov process for which we can determine the Fokker-Planck equation of the joint density. In particular, if we assume that $z_o(t)$ is the dynamic response of a system that can be well modeled by

$$z_o(t) = h_o[t:\underline{x}_o(t)] \tag{27}$$

$$d\underline{x}_o(t) = f_o[t:\underline{x}_o(t)]\,dt + g_o[t:\underline{x}_o(t)]\,dw(t), \tag{28}$$

where $\{w(t)\}$ is a Wiener process independent of $\{b(t)\}$, then, by the same arguments used above, we can obtain a partial differential equation for the time evolution of $r[\,.\,|\,.\,]$ quite similar to (26), the only difference being that we must consider expectations conditional on $\underline{x}_o(t)$, as well as $\hat{\underline{x}}(t)$.

3. Methods of Solving the Basic Equation

We shall now present some methods of solving Eq. 26 so as to obtain $\mu_o(s,t)$. Two approaches will be discussed: (i) a Fourier transform technique used by Siegert,[22] and (ii) asymptotic solutions based on the spectrum of the "forward operator" of Eq. 26. Both approaches can be utilized when the $\hat{z}_k(t)$ are approximate estimates, as well as when $\hat{z}_o(t)$ is an exact least-squares estimate.

The Fourier transform approach is motivated by the observation that if we were able to solve (26) explicitly for $r[\,.\,|\,.\,]$, we would still have to evaluate the integral on the right-hand side of (22). Siegert[22] has demonstrated that in problems of this kind, it may be better to consider the Fourier transform of $r[\,.\,|\,.\,]$ with respect to $\hat{\underline{x}}(t)$. That is,

$$q(\underline{\eta}, t) = \int e^{j\underline{\eta}^T\hat{\underline{x}}_t}\, r[\hat{\underline{x}}_t, t\,|\,\hat{\underline{x}}(0)]\, d\hat{\underline{x}}_t. \tag{29}$$

Note that if we can obtain $q(\underline{\eta}, t)$, then $\mu_o(s,t)$ can be easily obtained from

$$\mu_o(s,t) = \ln q(0,t). \tag{30}$$

In some cases, including that of Gauss-Markov signals, it appears that the basic equation

$$\frac{\partial}{\partial t} r[\,.\,|\,.\,] = L_r^+ r[\,.\,|\,.\,], \tag{31}$$

after multiplying both sides of (26) by $\exp\left(j\underline{\eta}^t\hat{\underline{x}}_t\right)$ and integrating by parts, yields an equation of the form

$$\frac{\partial}{\partial t} q(\underline{\eta}, t) = L_q^+ q(\underline{\eta}, t), \tag{32}$$

where the operator L_q^+ involves various partial derivatives with respect to the components of $\underline{\eta}$.

To illustrate this technique, we now consider the case for which $z_o(t) = 0$ and $z_1(t)$ is the solution of the stochastic differential equation

$$\frac{d}{dt}[z_1(t)] = F(t)\, z_1(t)\, dt + u(t), \tag{33}$$

where $F(t)$ is a known function of time, and $u(t)$ is white Gaussian noise of unit spectral density. For this case, the (Kalman-Bucy) "estimator equation" is

$$\begin{aligned} d\hat{z}_1(t) = d\hat{x}_1(t) &= k_1(t)\,\hat{x}_1(t)\,dt + k_2(t)\,dy(t) \\ &= k_1(t)\,\hat{x}_1(t)\,dt + k_2(t)\,db(t), \end{aligned} \tag{34}$$

where

$$k_1(t) = F(t) - \xi_p(t) \tag{35}$$

$$k_2(t) = \xi_p(t), \tag{36}$$

and $\xi_p(t)$ is the estimation mean-square error; that is, $E\{[z_1(t)-\hat{z}_1(t)]\}$ is the solution to the "variance equation"

$$\frac{d\xi_p(t)}{dt} = 2F(t)\,\xi_p(t) - \xi_p^2(t) + 1. \tag{37}$$

Substituting in (26), we obtain

$$\begin{aligned} \frac{\partial r(\hat{x}_1, T)}{\partial T} &= -k_1(T)\frac{\partial}{\partial \hat{x}_1}[\hat{x}_1 r(\hat{x}_1, T)] + \frac{s^2 - s}{2}\hat{x}_1^2 r(\hat{x}_1, T) \\ &\quad + \frac{1}{2}k_2^2(T)\frac{\partial^2}{\partial \hat{x}_1^2} r(\hat{x}_1, T) - sk_2(T)\frac{\partial}{\partial \hat{x}_1}[\hat{x}_1 r(\hat{x}_1, T)]. \end{aligned} \tag{38}$$

It can be established that $q(\eta, T)$ satisfies the partial differential equation

$$\frac{\partial q(\eta, T)}{\partial T} = \eta[k_1(T)+sk_2(T)]\frac{\partial q(\eta, T)}{\partial \eta} - \frac{k_2^2(T)}{2}\eta^2 q(\eta, T) - \frac{s^2 - s}{2}\frac{\partial^2 q(\eta, T)}{\partial \eta^2}. \tag{39}$$

We now "guess" that $q(\eta, T)$ is of the form

$$q(\eta, T) = f(T) \exp\left[-\frac{1}{2}\sigma_1(T)\eta^2\right]. \tag{40}$$

Substituting (40) in (39), we obtain

$$\frac{\dot{f}}{f} - \frac{1}{2}\dot{\sigma}_1\eta^2 = \eta(k_1+sk_2)(-\sigma_1\eta) - \frac{1}{2}k_2^2\eta^2 - \frac{s^2-s}{2}\left[-\sigma_1+(-\sigma_1\eta)^2\right], \tag{41}$$

where dots indicate differentiation with respect to T. Equating powers of η, we obtain

$$\frac{\dot{f}}{f} = \frac{s^2-s}{2}\sigma_1 \tag{42}$$

$$\dot{\sigma}_1 = 2\left[k_1(t)+sk_2(t)\right]\sigma_1 + k_2^2(t) + (s^2-s)\sigma_1^2. \tag{43}$$

Setting $V(t) = \sigma_1(t)$ and noting from Eq. 30 that we are interested in the case $\eta = 0$, we obtain

$$\frac{\dot{f}}{f} = \frac{\partial\mu_o(s,T)}{\partial T} = \frac{s^2-s}{2}V(T) \tag{44}$$

$$\dot{V}(t) = 2\left[F(t)+(s-1)\xi_p(t)\right]V(t) + \xi_p^2(t) + (s^2-s)V^2(t). \tag{45}$$

We note that the differential equation for V(t) can be interpreted as the variance equation for the estimation of the Gauss-Markov process generated by

$$\dot{z}(t) = \left[F(t)+(s-1)\xi_p(t)\right]z(t) + \xi_p(t)\,u(t) \tag{46}$$

in white Gaussian noise of spectral density $(s-s^2)^{-1}$. From this, it can be shown that V(t) converges asymptotically to a strictly positive value, but we shall not go into the details here.

This particular case falls into a class of problems studied by Collins.[1] His result for this particular case is

$$\frac{\partial\mu(s,T)}{\partial T} = \frac{1-s}{2}\left[\xi_p(t|\,1) - \xi_p\left(t\,\middle|\,\frac{1}{1-s}\right)\right], \tag{47}$$

where

$\xi_p(t|N_o)$ = mean-square estimation error for the process $z_1(t)$ of Eq. 34 in white Gaussian noise of spectral density N_o.

Using the fact that $\xi_p(t|N_o)$ is the solution of the "variance equation"

$$\frac{d}{dt}\xi_p(t|N_o) = 2F(t)\,\xi_p(t|N_o) - \frac{1}{N_o}\xi_p^2(t|N_o) + 1, \tag{48}$$

it can be shown by substitution that (45) and (47) are equivalent.

Collin's approach to this problem involved the following steps.

1. Obtaining an expression for $\mu(s, T)$ in terms of a Fredholm determinant by making a Karhunen-Loève series expansion of the signal process and using the fact that for a Gaussian process, the coefficients in the expansion are statistically independent Gaussian random variables to obtain an explicit answer in terms of the eigenvalues and s.

2. Using some resolvent identities, which apparently first appeared in Siegert,[22] to relate the Fredholm determinant to some expressions for the realizable mean-square estimation error.

The principal advantage of the approach presented here over that used by Collins and Kurth is that their approach can only be used when the received process is Gaussian, whereas our approach would appear to be applicable to a larger class of signal processes, since it exploits the properties of Markov rather than of Gaussian processes.

Finally, we discuss an asymptotic method based on the spectrum of the operator L_r^+ of Eq. 31. A standard way of solving equations such as (26) is to assume that $r[\,.\,|\,.\,]$ is of the form

$$r[\underline{\hat{x}}_t, t\,|\,\underline{\hat{x}}(0)] = d[\underline{\hat{x}}_t, t\,|\,\underline{\hat{x}}(0)]\; e(t). \tag{49}$$

Substituting (49) in (31), we find that (31) can be solved for all t and $\underline{\hat{x}}_t$ only if

$$\frac{1}{e}\,\frac{de}{dt} = \lambda = \frac{1}{d[\cdot]}\, L_r^+\, d[\cdot]. \tag{50}$$

From (50), we see that if

$$\left(L_r^+ - \lambda\right)\; d[\cdot] = 0, \tag{51}$$

then

$$e(t) = k\, e^{\lambda t} \tag{52}$$

and (49) is indeed a solution to (31).

When the set of eigenvalues $\{\lambda_i\}$ forms a discrete set (that is, L_r^+ has a discrete spectrum), then as $t \to \infty$ the general solution for $r[\,.\,|\,.\,]$ is dominated by the lowest order eigenvalue

$$r[\,.\,|\,.\,] = \sum_{i=1}^{\infty} c_i d_i[\underline{\hat{x}}_t, \underline{\hat{x}}(0)]\; e^{\lambda_i t} \longrightarrow c_1 d_1[\underline{\hat{x}}_t, \underline{\hat{x}}(0)]\; e^{\lambda_1 t}, \tag{53}$$

where $d_i[\cdot]$ is the eigenfunction associated with the i^{th} eigenvalue and $\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4$ ...... Thus for large t, we see that

$$\mu_o(s, t) \longrightarrow \lambda_1 t. \tag{54}$$

Computing the lowest order eigenvalue of a differential operator has been the subject of much study in mathematics, physics, and chemistry (for example, in physics the lowest order eigenvalue often corresponds to the lowest energy level of the system). In particular, approximate solutions that can be obtained by variational methods (see Gould[23]) would appear to be quite useful in practice.

We shall mention one such method, the Rayleigh-Ritz method. Suppose any solution to (31) can be represented by a series of the form given in (53), and that $u[\hat{\underline{x}}_t, \hat{\underline{x}}(0)]$ is the class of functions satisfying the boundary conditions for $r[\,.\,|\,.\,]$. Then

$$\lambda_1 = \max_u \frac{\int L_r^+ u[\hat{\underline{x}}_t, \hat{\underline{x}}(0)]\, d\hat{\underline{x}}_t d\hat{\underline{x}}(0)}{\int u[\hat{\underline{x}}_t, \hat{\underline{x}}(0)]\, d\hat{\underline{x}}_t d\hat{\underline{x}}(0)}. \tag{55}$$

Typically, we would use a finite sum of orthogonal functions for $u[\cdot]$ and keep on increasing the number of terms until the value of λ_1 stabilizes (see Slater[24]).

4. Concluding Comments

We have presented a new method of computing $\mu(s,t)$ for binary detection problems when the signal process is a non-Gaussian Markov process. We have shown that the time evolution of $\mu(s,t)$ is closely related to the solution of a certain nonlinear differential equation and discussed some methods of solving such an equation.

J. E. Evans

References

1. L. D. Collins, "Asymptotic Approximations to the Error Probability for Detecting Gaussian Signals," Ph.D. Thesis, Department of Electrical Engineering, M.I.T., June 1968.
2. R. L. Kurth, "Distributed-Parameter State-Variable Techniques Applied to Communication over Dispersive Channels," Ph.D. Thesis, Department of Electrical Engineering, M.I.T., June 1969.
3. L. D. Collins, "Asymptotic Approximations to the Error Probability for Detecting Gaussian Signals," Technical Paper 6-1, WESCON, Los Angeles, California, August 1968.
4. L. D. Collins and R. L. Kurth, "Asymptotic Approximations to the Error Probability for Square-Law Detection of Gaussian Signals," Quarterly Progress Report No. 90, Research Laboratory of Electronics, M.I.T., July 15, 1968, pp. 191-201.
5. T. E. Duncan, "Evaluation of Likelihood Functions," Inform. Contr. 13, 62-74 (1968).
6. T. Kailath, "A General Likelihood-Ratio Formula for Random Signals in Gaussian Noise," IEEE Trans., Vol. IT-15, No. 3, pp. 350-361, May 1969.
7. K. Itô, "On Stochastic Differential Equations," Mem. Am. Math. Soc. 4, 1-51 (1951).
8. W. M. Wonham, "Some Applications of Stochastic Differential Equations to Optimum Nonlinear Filtering," J. SIAM Control 2, 347-369 (1965).

9. R. Kalman and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," ASME Trans. J. Basic Engineering 83D, 95-107 (December 1961).

10. H. J. Kushner, "Dynamic Equations for Optimum Nonlinear Filtering," J. Differential Equations 3, 171-190 (1967).

11. D. L. Snyder, The State-Variable Approach to Continuous Estimation with Applications to Analog Communication Theory (The M. I. T. Press, Cambridge, Mass., 1968).

12. T. E. Duncan, "Probability Densities for Diffusion Processes with Applications to Nonlinear Filtering Theory and Detection Theory," Ph.D. Thesis, Department of Electrical Engineering, Stanford University, May 1967.

13. R. Feynman, "Space-Time Approach to Non-relativistic Quantum Mechanics," Rev. Mod. Phys. 20, 367-387 (1948).

14. E. Dynkin, Markov Processes, Vol. I (Springer-Verlag, Berlin, translation, 1965).

15. M. Kac, "On Some Connections between Probability Theory and Differential and Integral Equations," Proc. Second Berkeley Symposium on Mathematical Statistics and Probability (University of California Press, Berkeley, 1951), pp. 189-215.

16. I. M. Gelfand and A. M. Yaglom, "Integration in Functional Spaces and Its Applications in Quantum Physics," Usp. Mat. Nauk 11, 77 (January 1956). (English translation in J. Mathemat. Phys. Vol. 1, No. 1, pp. 48-69, January 1960.)

17. M. Kac, Probability and Related Topics in Physical Sciences (Interscience Publishers, Inc., New York, 1959).

18. R. P. Feynman and A. R. Hibbs, Quantum Mechanics and Path Integrals (McGraw-Hill Book Company, New York, 1965).

19. E. B. Dynkin, "Transformations of Markov Processes Connected with Additive Functionals," Proc. Fourth Berkeley Symposium on Mathematical Statistics and Probability (University of California Press, Berkeley, 1961), pp. 117-142.

20. D. Ray, "On Spectra of Second-Order Differential Operators," Am. Math. Soc. Trans. 77, 299-321 (1954).

21. D. A. Darling and A. J. F. Siegert, "A Systematic Approach to a Class of Problems in the Theory of Noise and Other Random Phenomena: Part I," IRE Trans., Vol. IT-3, pp. 32-37, March 1957.

22. A. J. F. Siegert, "A Systematic Approach to a Class of Problems in the Theory of Noise and Other Random Phenomena: Part II. Examples," IRE Trans., Vol. IT-3, pp. 38-43, March 1957.

23. S. H. Gould, Variational Methods for Eigenvalue Problems (University of Toronto Press, Toronto, 1957).

24. J. C. Slater, Quantum Theory of Atomic Structure, Vol. I (McGraw-Hill Book Company, New York, 1960), pp. 113-119.