A. NONLINEAR MAP INTERVAL ESTIMATION

We shall consider two problems related to the equations specifying the nonlinear interval estimator. In the first part, we shall discuss a method for solving these equations which involves the optimum nonlinear realizable filter. In the second part, we shall present a technique for determining the optimum nonlinear realizable filter from the equations specifying the nonlinear interval estimator.

We shall assume that the following equations describe the system of interest:

\[ \frac{dx(t)}{dt} - f(x(t), t) + g(x(t), t) u(t) \text{ nonlinear state equation} \]  
\[ \mathbb{E}[u(t)u^T(\tau)] = Q_\delta(t-\tau) \]  
\[ \mathbb{E}[x(T_o)] = \bar{x}_o \text{ initial condition assumption} \]  
\[ \mathbb{E}[(x(T_o)-\bar{x}_o)(x(T_o)-\bar{x}_o)^T] = P_o \]  
\[ r(t) = s(x(t), t) + w(t) \text{ nonlinear observation equation} \]  
\[ \mathbb{E}[w(t)w^T(\tau)] = R_\delta(t-\tau) \text{ for } T_0 < t, \tau < T_f \]

For this system, the equations\(^1,\)\(^2\) describing the MAP interval estimate of the state vector \( \hat{x}(t) \) are

\[ \frac{d\hat{x}(t)}{dt} = f(\hat{x}(t), t) + g(\hat{x}(t), t) q^{-T}(\hat{x}(t), t) p(t) \]  
\[ \frac{d p(t)}{dt} = - \left[ \frac{\partial f}{\partial x} \right]_{|\dot{x} = \hat{x}(t)} \]  
\[ p(t) + C^T(\hat{x}(t), t) R^{-1}(r(t)-s(\hat{x}(t), t)), \]

where

*This work was supported by the Joint Services Electronics Programs (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract DA 36-039-AMC-03200(E)).
There is also a two-point boundary-value restriction on the equations. The boundary conditions are required to satisfy

$$-(T_0) - I_p P_o(T_0) (10)$$

$$P(T_f) = 0. (11)$$

The difficulty in solving these equations is the two-point boundary condition associated with the problem. We note that if we could obtain the value $x(T_f)$, that is, the state estimate at the end point of the interval, we would have a complete set of boundary conditions at this time. We could then solve these equations backwards in time from this end point as if it were an initial or, more precisely, final-value problem. The estimate at the end point of the interval, however, is identical to the realizable filter estimate, since it involves only an operation upon past data, that is, the data before the end point time, $T_f$. Therefore, one method that we propose for solving the nonlinear smoothing problem may be outlined as follows:

a. Determine a complete set of boundary conditions at the end point of the interval by obtaining the realizable filter estimate.

b. Solve the MAP interval estimation equations backwards in time from the end point of the observation interval, using this set of boundary conditions.

This method is not the only algorithm for solving the MAP estimation equations. Another technique is the method of quasi-linearization. With this technique, the estimation equations are linearized around some a priori estimate of the solution. Then these linear equations are solved exactly, by use of the transition matrix associated with this system. This new solution provides the next estimate around which the nonlinear equations are linearized. This technique is repeated until a satisfactory convergence criterion has been satisfied.

The equivalence of the two methods has not been shown. It is suspected that in the high signal-to-noise case they produce identical estimates. In the low signal-to-noise case, however, the quasi-linearization procedure is probably better. This is because the realizable filter estimate that we obtain is an approximate estimate. Probably, the best procedure is some combination of the two techniques.

We shall now present a derivation of the optimum nonlinear realizable filter by the use of the technique of invariant imbedding. Our derivation is a modified version of that given by Detchmendy and Shridar.3, 4

A fundamental difference between the interval estimator and the realizable filter is the time variable involved. In the interval estimator the important time variable is the time within the fixed observation interval, whereas in the realizable filter the important
time variable is the end-point time of the observation interval, which is not fixed but increases continually as the data are accumulated. For the realizable filter we want the estimate at the end point of the observation interval as a function of the length of the interval.

Let us now consider a more general class of solutions to the MAP equations. Instead of imposing the condition of Eq. 11, let us consider the class of solutions for the boundary condition

\[ p(T_f) = \eta. \] (12)

In general, the solution to Eq. 7 at the end point of the observation interval is now a function of the end-point time of the interval and the parameter \( \eta \). Let us denote the solution to Eq. 7 at the end point of the interval by \( \hat{x}(T_f, \eta) \), where we emphasize the dependence noted above. We also note that

\[ \hat{x}(T_f, \eta) = x(T_f) \] (13)

We now state that it can be shown that the function satisfies the following partial differential equation:

\[ \frac{\partial \hat{x}(T_f, \eta)}{\partial T_f} + \frac{\partial \hat{x}(T_f, \eta)}{\partial \eta} \pi(\hat{x}, \eta, T_f) = \Phi(\hat{x}, \eta, T_f). \] (14)

where

\[ \Phi(\hat{x}, \eta, T_f) = f(\hat{x}, T_f) + g(\hat{x}, T_f) Qg^T(\hat{x}, T_f) \eta \] (15)

\[ \pi(\hat{x}, \eta, T_f) = -\frac{\partial f}{\partial x}(x, T_f) \bigg|_{x=\hat{x}} \eta - C^T(\hat{x}, T_f) R^{-1}(r(T_f) - s(\hat{x}, T_f)). \] (16)

For convenience of notation, let us define the term

\[ K(x, r(T_f), T_f) = C^T(x, T_f) R^{-1}(r(T_f) - s(x, T_f)). \] (17)

We now try a solution to Eq. 14 of the form

\[ \hat{x}(T_f, \eta) = \hat{x}(T_f) + \hat{p}(T_f) \eta. \] (18)

where \( \hat{x}(T_f) \) and \( \hat{p}(T_f) \) are functions that are to be chosen appropriately. As indicated by the notation, the function is the desired estimate. We now substitute this trial solution in the invariant imbedding equation. Since we are interested in the solution when \( \eta \) is small, we expand the functions around \( \hat{x}(T_f) \). We want to consider terms to order \( |\eta| \). The expansions of (15) and (16) are
\[ \Phi(x, n, T_f) = f(\hat{x}(T_f), T_f) + \frac{\partial f}{\partial x} (x, T_f) \bigg|_{x=\hat{x}(T_f)} P(T_f) \eta + g(\hat{x}(T_f), T_f) Q g(\hat{x}(T_f), T_f) \]

\[ \pi(x, n, T_f) = -\frac{\partial f}{\partial x} (x, T_f)^T \bigg|_{x=\hat{x}(T_f)} \eta - \frac{\partial K}{\partial x} (x, \varepsilon, T_f) \bigg|_{x=\hat{x}(T_f)} P(T_f) \eta - K(\hat{x}(T_f), \varepsilon(T_f), T_f). \]

Substituting (19) and (20) in (14) yields

\[
\left\{ \frac{d\hat{x}(T_f)}{dT_f} = f(\hat{x}(T_f), T_f) + P(T_f) \kappa(\hat{x}(T_f), \varepsilon(T_f), T_f) \right. \\
+ \left. \left\{ \frac{dP(T_f)}{dT_f} = \frac{\partial f}{\partial x} (x, T_f) \bigg|_{x=\hat{x}(T_f)} P(T_f) - P(T_f) \frac{\partial f}{\partial x} (x, T_f) \bigg|_{x=\hat{x}(T_f)} \right\} \right. \\
- \left. P(T_f) \frac{\partial K}{\partial x} (\hat{x}, \varepsilon(T_f), T_f) \bigg|_{x=\hat{x}(T_f)} P(T_f) + g(\hat{x}(T_f), T_f) Q g(\hat{x}(T_f), T_f) \right\} + O(\eta^2) = 0. \]

We require that the functions \( x(T_f) \) and \( P(T_f) \) be so chosen that the terms within the first and second braces of Eq. 21 vanish identically. This specifies the solution to the invariant imbedding equation to \( O(\eta^2) \). The term within the first brace specifies the estimation equation. This is given by

\[ \frac{d\hat{x}(T_f)}{dT_f} = f(\hat{x}(T_f), T_f) - P(T_f) C^T(\hat{x}(T_f), T_f) R^{-1}(\varepsilon(T_f) - \varepsilon(\hat{x}(T_f), T_f)). \]

The term within the second brace specifies part of the dynamics of the estimation equation. This is given by

\[ \frac{dP(T_f)}{dT_f} = \frac{\partial f}{\partial x} (x, T_f) \bigg|_{x=\hat{x}(T_f)} P(T_f) + P(T_f) \frac{\partial f}{\partial x} (x, T_f) \bigg|_{x=\hat{x}(T_f)} \]

\[ -P(T_f) \frac{\partial K}{\partial x} (x, \varepsilon(T_f), T_f) \bigg|_{x=\hat{x}(T_f)} P(T_f) + g(\hat{x}(T_f), T_f) Q g(\hat{x}(T_f), T_f), \]

where (see Eq. 17)

\[ K(x; \varepsilon, T_f) = C^T(x, T_f) R^{-1}(\varepsilon(T_f) - \varepsilon(x, T_f)). \]

The initial conditions to these differential equations are found by setting \( T_f \) equal to \( T_0 \) and using Eq. 10. This gives
Equations 22 and 23 are identical to those derived by Snyder by approximating the solution to the Fokker-Planck equation describing the a posteriori density. We note that we obtain the same result—that the optimum nonlinear realizable filter may not be equivalent to successively linearizing the system under consideration and then applying the Kalman-Bucy filtering equations. We also note that our solution is an approximate one because we only approximated the solution to the invariant imbedding equation.

A. B. Baggeroer

References


B. PERFORMANCE BOUNDS FOR OPTIMUM DETECTION FOR GAUSSIAN SIGNALS

In this report, we shall apply the technique of tilted probability distributions to the problem of evaluating the performance of optimum detector for Gaussian signals received in additive Gaussian noise.

Stated in mathematical terms, we shall consider the following binary detection problem.

\[
\begin{align*}
H_1: \quad r(t) &= s_{r_1}(t) + m_1(t) + w(t) \\
H_2: \quad r(t) &= s_{r_2}(t) + m_2(t) + w(t)
\end{align*}
\]

where \( s_{r_1}(t) \) and \( s_{r_2}(t) \) are sample functions from zero-mean Gaussian random...
processes with known covariance functions $K_1(t, \tau)$ and $K_2(t, \tau)$, respectively; $m_1(t)$ and $m_2(t)$ are known waveforms, $w(t)$, in a sample function of white Gaussian noise of spectral density $\frac{\sigma^2}{2}$.

Since the optimum detector for such problems is well known,\textsuperscript{1-3} we shall not dwell upon it here. It suffices to indicate that for a large class of performance criteria, the detector bases its decision on the likelihood ratio, or some monotone function of the likelihood ratio. This is the class of detectors with which we shall be concerned.

One structure for the optimum detector is shown in Fig. XIX-1. A direct evaluation of the error probabilities is conceptually possible, but practically it is extremely difficult, for we are faced with the problem of computing the probability distribution at the output of a time-variant nonlinear filter. It is this motivation that has led us to consider alternative methods of performance evaluation.

One alternative which is widely used is the output signal-to-noise ratio,\textsuperscript{4} which is valid in the so-called low-energy coherence, or threshold, case. One can show by a simple example that this performance measure can give incorrect results in some problems of interest, and therefore it must be applied with caution.

In an attempt to overcome the computational difficulties associated with an exact computation of error probabilities, while at the same time having a performance measure that is of wider applicability than the output signal-to-noise ratio, we have been investigating bounds on error probabilities, and, in particular, we have been seeking bounds that become asymptotically exact as the transmitted signal energy becomes large.

1. Tilted Probability Distribution

The technique that we employ is usually called "tilting" of probability distributions. It was introduced into information and coding theory by Shannon,\textsuperscript{5, 6} and has been employed with great success. Earlier applications in the field of mathematical statistics are due to Chernoff\textsuperscript{7} and Esscher.\textsuperscript{8} We only summarize the notation and results
here, before proceeding to our specific problem.

All of our results center around the semi-invariant moment-generating function $\mu(s)$, which is merely the logarithm of the conditional characteristic function $M_{\mathbb{H}|H_1}(s)$

$$\mu(s) = \ln M_{\mathbb{H}|H_1}(s)$$

$$= \ln \int p_{\mathbb{H}|H_1}(R) p_{\mathbb{H}|H_2}(R) \, dR,$$

where $p_{\mathbb{H}|H_1}(R)$ denotes the probability density for the received vector $R$, conditioned on the hypothesis $H_i$, $i = 1, 2$. It can be readily shown that

$$\mu'(s) = \int p_{\mathbb{H}|H_1}(R) \ell(R) \, dR$$

$$\mu''(s) = \int p_{\mathbb{H}|H_1}(R) \ell^2(R) \, dR - [\mu'(s)]^2,$$

which are the mean and variance of $\ell(R)$ with respect to the probability density

$$p_{\mathbb{H}|H_1}(R) = \exp[-\mu(s)] Pr[H_1(R) Pr[H_2(R)].$$

We shall refer to $p_{\mathbb{H}|H_1}(R)$ as the "tilted" density. The amount of "tilting" depends on the value of the parameter $s$. It follows from our definition that

$$p_{\mathbb{H}|H_1}(R) = p_{\mathbb{H}|H_1}(R) \exp[\mu(s) - s\ell(R)]$$

$$p_{\mathbb{H}|H_2}(R) = p_{\mathbb{H}|H_2}(R) \exp[\mu(s) + (1-s)\ell(R)].$$

Hence the error probabilities may be expressed in terms of $\mu(s)$ and the tilted density $p_{\mathbb{H}|H_1}(R)$

$$Pr[\mathbb{H}|H_1] = \int_{\{R: \ell(R) > \gamma\}} p_{\mathbb{H}|H_1}(R) \, dR$$

$$= \int_0^\gamma p_{\mathbb{H}|H_1} L \exp[\mu(s)-sL] \, dL.$$
\[ 
\Pr \left[ \mathcal{E} \mid H_2 \right] = \int_{\{R : \ell(R) < \gamma\}} \Pr \left[ \mathcal{E} \mid H_2(R) \right] \, dR \\
= \int_{-\infty}^{\gamma} p_L(L) \exp[\mu(s) + (1-s)L] \, dL, 
\]

where \( \gamma \) denotes the threshold level, and \( p_L(L) \) is the tilted probability density for the log-likelihood ratio corresponding to the nonlinear transformation \( \ell_s = L(r_s) \).

A simple, well-known upper bound that follows immediately from Eqs. 8 and 9 is the Chernoff bound. For example, if in (7) we bound \( e^{-sL} \) by \( e^{-s\gamma} \), then bound \( \int_{\gamma}^{\infty} p_L(L) \, dL \) by unity, we obtain the following upper bounds:

\[ 
\Pr \left[ \mathcal{E} \mid H_1 \right] \leq \exp[\mu(s) - s\gamma] \\
\Pr \left[ \mathcal{E} \mid H_2 \right] \leq \exp[\mu(s) + (1-s)\gamma].
\]

Then we can minimize the bound by proper choice of the parameter \( s \). This choice is the solution to

\[ 
\mu'(s) = \gamma. 
\]

A solution exists and is unique, provided the threshold \( \gamma \) lies between the means of the conditional densities \( p_L|H_1(L) \) and \( p_L|H_2(L) \). This condition is usually satisfied in the applications. An important step in our work consists in finding tighter bounds than Eqs. 10 and 11. After all, the arguments leading to these bounds were rather crude, and while the exponential behavior exhibited in Eqs. 10 and 11 is adequate for many applications, particularly in information and coding theory, we would like to retain the algebraic dependence for our applications to radar, sonar, and uncoded communication systems. A bound of a particularly attractive form is

\[ 
\Pr \left[ \mathcal{E} \mid H_1 \right] \leq \frac{\exp[\mu(s) - s\mu'(s)]}{\sqrt{2\pi s^2 \mu''(s)}} \\
\Pr \left[ \mathcal{E} \mid H_2 \right] \leq \frac{\exp[\mu(s) + (1-s)\mu'(s)]}{\sqrt{2\pi(1-s)^2 \mu''(s)}}
\]

where \( \mu'(s) = \gamma \).

For large values of transmitted energy, this bound is considerably tighter than the Chernoff bound, although at the other extreme the Chernoff bound is tighter. Clearly, the bound that should be used in a given situation is the one that is the tightest.

We have not yet obtained a rigorous derivation of Eqs. 13 and 14 that is of sufficient
generality to take care of all problems of interest to us. In the sequel, when we discuss
the application of these bounding techniques to the Gaussian problem, we shall discuss
some preliminary results in this direction.

2. Application to the Gauss-in-Gauss Problem

We shall specialize the bounding techniques to the case in which the received vectors
have Gaussian conditional probability densities with statistically independent components.
Then, by interpreting the components of these vectors as coordinates in a Karhunen-
Loève expansion, we shall generalize our results to the case in which the received sig-
nals are sample functions from random processes; that is, they are infinite dimen-
sional.

The present problem is the evaluation of \( \mu(s) \). We assume that on hypothesis
\( H_j \), \( j = 1, 2, \ldots \), \( \mathbf{r} \) is an \( N \) component Gaussian vector with

\[
E[r_i] = m_{ij}
\]

\[
V_0[r_i] = \lambda_{ij} + \frac{N_o}{2}
\]

\[
E[r_i r_k] = m_{ij} m_{kj}, \quad k \neq i
\]

for \( i, k = 1, 2, \ldots, N \)

\[
j = 1, 2
\]

Substituting the \( N \)-dimensional Gaussian densities in the definition of \( \mu(s) \), we find

\[
\mu(s) = \frac{1}{2} \sum_{i=1}^{N} \left\{ \frac{1}{2} \ln \left( 1 + \frac{2\lambda_{ii}}{N_o} \right) + \ln \left( 1 + \frac{2\lambda_{ij}}{N_o} \right) - \ln \left[ \frac{2}{N_o} \left( s \lambda_{ii} + (1-s) \lambda_{ij} \right)^2 + 1 \right] \right\} \left( m_{ii} - m_{ij} \right)^2
\]

\[
- \frac{(m_{ij} - m_{ij})^2}{\frac{\lambda_{ii}}{1 - s} + \frac{\lambda_{ij}}{s} + \frac{N_o}{2}}
\]

\[
0 \leq s \leq 1.
\]

Equation 18 is now in a very convenient form to let \( N \rightarrow \infty \). This limiting operation is
frequently used in detection and estimation theory, so we need not worry about its jus-
tification here. Furthermore, the infinite series which results is convergent in all cases
of interest to us. It then remains to interpret the various terms as closed-form expres-
sions involving the (conditional) mean and covariance functions of the random process
\( r(t) \), for (17) is not at all convenient from a computational point of view for the case
\( N \rightarrow \infty \). To illustrate the manner in which we find a closed-form expression for Eq. 18,
consider the first term in the series,

\[
\frac{s}{2} \sum_{i=1}^{\infty} \ln \left( 1 + \frac{2\lambda_{i1}}{N_o} \right).
\]

Now, for the moment, we wish to focus on a related problem; namely that of estimating the zero-mean Gaussian random process \( s_r(t) \) when observed in additive white Gaussian noise of spectral height \( N_o/2 \).

Let \( \hat{s}_{r_1}(t) \) denote the minimum mean-square-error point estimate of \( s_{r_1}(t) \). The resulting mean-square error is

\[
\xi_{r1}(t; \frac{N_o}{2}) = E \left[ \left( s_{r_1}(t) - s_{r_1}(t) \right)^2 \right],
\]

where we have explicitly included the noise level as a parameter. Furthermore, we explicitly indicate the dependence of the Karhunen-Loève eigenvalues and eigenfunctions on the length of the interval:

\[
\lambda_{i1}(\tau) \phi_i(t; \tau) = \int_0^T K_1(t, \tau) d_1(\tau) d\tau.
\]

Then,

\[
\sum_{i=1}^{\infty} \ln \left( 1 + \frac{2\lambda_{i1}}{N_o} \right)
\]

\[
= \sum_{i=1}^{\infty} \ln \left( 1 + \frac{2\lambda_{i1}(\tau)}{N_o} \right)
\]

\[
= \int_0^T d\tau \frac{\partial}{\partial t} \sum_{i=1}^{\infty} \ln \left( 1 + \frac{2\lambda_{i1}(t)}{N_o} \right)
\]

\[
= \int_0^T d\tau \sum_{i=1}^{\infty} \frac{\lambda_{i1}(t)}{N_o + \lambda_{i1}(t)} \phi_i^2(t; \tau). (21)
\]

In Eq. 21, we have used the result of Huang and Johnson

\[
\frac{\partial \lambda_{i1}(t)}{\partial t} = \lambda_{i1}(t) \phi_i^2(t; t).
\]

Then we have
\[
\frac{s}{2} \sum_{i=1}^{\infty} \ln \left( 1 + \frac{2 \lambda_{1}}{N_0} \right) = \frac{s}{2} \int_{0}^{\tau} \xi_1 \left( t: \frac{N_0}{2} \right) dt.
\]

(23)

Similar results hold for the other terms in Eq. 18. The final closed-form expression for \( \mu(s) \) is

\[
\mu(s) = \frac{2}{N_0} \left\{ \begin{array}{c}
\frac{s}{2} \int_{0}^{T} \xi_1 \left( t: \frac{N_0}{2} \right) dt + \frac{1-s}{2} \int_{0}^{T} \xi_2 \left( t: \frac{N_0}{2} \right) dt \\
- \frac{1}{2} \int_{0}^{T} \xi_{1+2} \left( t: \frac{N_0}{2}, s \right) dt \\
- \frac{s(1-s)}{2} \int_{0}^{T} \left[ m_1(t) - m_2(t) - \int_{0}^{T} h_{1+2}(t, \tau:s)[m_1(\tau) - m_2(\tau)] h_1 \right]^2 dt
\end{array} \right\}
\]

(24)

The terms in Eq. 24 that have not been previously defined are the following. \( \xi_{1+2} \left( t: \frac{N_0}{2}, s \right) \) is the minimum mean-square point estimation error for estimating the fictitious random process \( s_{1+2}(t), \) which is defined by

\[
s_{1+2}(t) = s_{r_1}(t) + (1-s) s_{r_2}(t) \quad 0 \leq s \leq 1.
\]

(25)

It is easy to see that this fictitious random process plays the same role as the tilted random vector \( r_s \) in the finite-dimensional case. \( h_{1+2}(t, \tau:s) \) denotes the minimum mean-square point estimator of \( s_{1+2}(t). \) The important computational advantage of Eq. 24, as compared with Eq. 18, stems from the availability of techniques for efficiently solving linear estimation problems. In particular, the Kalman-Bucy formulation \(^{10}\) of the estimation problem provides us with a direct approach for calculating the pertinent estimation errors. Furthermore, this approach to the problem is readily implemented on a digital computer, which is an important practical advantage.

We have briefly mentioned the need for a bound that would be tighter than the Chernoff bound. We shall now indicate some results in this direction, when we make use of our Gaussian assumption.

A straightforward substitution of the appropriate Gaussian densities in Eq. 5 shows that the tilted density \( p_{r_s}(R) \) is also a Gaussian probability density. Therefore, the tilted random variable \( f_{s} \) is a "generalized Chi-square" random variable, that is, it is the sum of the square of Gaussian random variables which have nonzero means and unequal variances.
In the special case for which all variances are the same, the means are all zero, and the number of components, N, is finite, \( \ell_s \) has a Chi-squared distribution, and we have an analytic expression for the probability density \( p_{\ell_s} (L) \). This enables us to rigorously derive the following bounds by using a very simple argument.

\[
P_r[\theta | H_1] < \left( 2\pi s^2 \mu''(s) \left( 1 - \frac{2}{N} \right) \right)^{-1/2} \exp[\mu(s) - s\mu'(s)] \quad (26)
\]

\[
P_r[\theta | H_2] < \left( 2\pi (1-s)^2 \mu''(s) \left( 1 - \frac{2}{N} \right) \right)^{-1/2} \exp[\mu(s) + (1-s)\mu'(s)] \quad \text{for } N > 2 \quad (27)
\]

Observe that these are in the form of (12) and (13) and, in fact, for moderate values of \( N \) the factor \( \left( 1 - \frac{2}{N} \right)^{-1/2} \approx 1 \).

Another special case of more interest to us occurs when the variances of the various components are all distinct, but occur in pairs, as is always the case for a signal transmitted at RF through some physical channel. In this case, we can obtain an expression for the probability density \( p_{\ell_s} (L) \). We have shown that this density has only one maximum. Furthermore, in each of a large number of cases treated numerically, the maximum value was not found to differ appreciably from that for a Gaussian density having the same variance, namely \( (2\pi \mu''(s))^{-1/2} \). Furthermore, various methods of asymptotically approximating the integrals in Eqs. 8 and 9 lead to this same dependence for the dominant behavior as \( \mu''(s) \) becomes large.

Therefore, we feel confident that for the Gaussian detection problem that we have been discussing, the expressions in Eqs. 13 and 14 provide good approximations for the error probabilities.

3. Examples

We shall give a few simple examples illustrating the application of the techniques just discussed. We shall choose one example from each of the three levels of detection problems: known signal, signal with a finite number of random parameters, and random signal.

**EXAMPLE 1: Known Signal**

\( H_1: \ r(t) = m_1(t) + n_1(t) + w(t) \)

\( 0 \leq t \leq \tau \) \quad (28)

\( H_2: \ r(t) = m_2(t) + n_1(t) + w(t) \)

Here the random "signal" components \( n_1(t) \) are identical and represent the colored component of the noise. Substituting in Eq. 24 (or 18), we find

\[
\mu(s) = -\frac{s(1-s)}{2} d^2
\]
where

\[
d^2 \triangleq \frac{(E(\ell | H_1) - E(\ell | H_0))^2}{\sqrt{\text{Var}(\ell | H_1) \text{Var}(\ell | H_0)}}
\]

\[
= \int_0^T \left[ m_1(t) - m_2(t) - \int_0^T h_{1+2}(t, \tau)[m_1(\tau) - m_2(\tau)] \, d\tau \right]^2
\]  

(30)

Here, \( h_{1+2}(t, \tau) \) is independent of \( s \), since

\[
s_{1+2}(t) = s_{r_1}(t) = s_{r_2}(t) = n_c(t).
\]  

(31)

And, as a matter of fact,

\[
\delta(t-\tau) - h_{1+2}(t, \tau) = h_w(t, \tau),
\]  

(32)

which is a realizable whitening filter for this problem.\(^\text{13}\) Thus Eq. 30 is equivalent to the more familiar form

\[
\theta^2 = \int_0^T \int_0^T [m_1(t) - m_2(t)] Q_n(t, \tau)[m_1(\tau) - m_2(\tau)] \, dt \, d\tau
\]  

(33)

\[
\mu'(s) = d^2 \left(s - \frac{1}{2}\right) = \gamma.
\]

(34)

Hence we can solve explicitly for \( s \) in this case, and we obtain the bounds

\[
\Pr[\delta | H_1] \leq \frac{1}{\sqrt{2\pi d^2 \left(\frac{1}{2} + \frac{\gamma}{d^2}\right)}} \exp \left[-\frac{d^2 \left(\frac{1}{2} + \frac{\gamma}{d^2}\right)}{2}\right]
\]  

(35)

\[
\Pr[\delta | H_1] \leq \frac{1}{\sqrt{2\pi d^2 \left(\frac{1}{2} - \frac{\gamma}{d^2}\right)}} \exp \left[-\frac{d^2 \left(\frac{1}{2} - \frac{\gamma}{d^2}\right)}{2}\right].
\]  

(36)

These are exactly the same results that we would obtain using the familiar bound

\[
\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \, dy \leq \frac{1}{\sqrt{2\pi}} e^{-x^2/2}
\]  

(37)

in the exact error expression for this problem.
EXAMPLE 2: Slow Rayleigh Fading with Diversity

This is another case in which analytical results are available for comparison purposes.\textsuperscript{11}

One of two orthogonal signals, $s_1(t)$ or $s_2(t)$, is received in additive white Gaussian noise after transmission through a fading channel in which the signal undergoes a random attenuation and phase shift. The attenuation is a Rayleigh random variable, and the phase shift is uniform $0 \leq \theta < 2\pi$. We assume that there are $N$ such channels operating in space, time, or frequency diversity. We also assume that the average received energy is the same in each channel on either hypothesis. Substitution in (17) yields

Fig. XIX-2. Error probabilities for slow Rayleigh fading with diversity.
Fig. XIX-3. Upper bounds on error probability for single-pole Rayleigh fading.
\[
\mu(s) = \frac{1}{N} \left\{ \ln \left(1 + \frac{E}{N_0}\right) - \ln \left(1 + \frac{sE}{N_0}\right) - \ln \left[1 + (1-s) \frac{E}{N_0}\right] \right\}
\]

where \(E\) is the average received signal energy in each diversity channel.

Figure XIX-2 shows the bound

\[
\Pr [\delta^c] \leq \frac{1}{\sqrt{\pi \frac{4}{\mu''(1/2)}}} \exp \left[\mu\left(\frac{1}{2}\right)\right],
\]

as well as the exact results for several values of \(N\).

EXAMPLE 3: Rayleigh Fading, Single-Pole Spectrum

This simple example illustrates the application of our results to a problem for which it would be extremely difficult to obtain error probabilities in any other manner. The model is the same as that of the previous case, except that, now the channel attenuation is a sample function from a random process.

For purposes of illustration, we assume only one Rayleigh channel (that is, no explicit diversity) and a single-pole spectrum for the fading.

The upper bound on the probability of error is shown in Fig. XIX-3. To our knowledge, no other calculations are available for comparison. It is interesting to observe that this optimum receiver does not exhibit the irreducible error probability of one sub-optimum receiver that has been analyzed.\(^{12}\)

4. Summary

We have discussed the application of tilted probability distributions to the evaluation of the performance of optimum detectors for Gaussian signals in Gaussian noise. Two main points were covered: (i) obtaining bounds that are tighter than the Chernoff bound and ones that are asymptotically exact; (ii) obtaining closed-form expressions for \(\mu(s)\) in terms of minimum mean-square estimation errors, which may be readily computed by using Kalman-Bucy filtering techniques. We concluded with three simple examples.

L. D. Collins

References


