XXIX. DETECTION AND ESTIMATION THEORY*

Academic and Research Staff

Prof. H. L. Van Trees
Prof. D. L. Snyder

Graduate Students

A. B. Baggeroer	L. D. Collins	M. Mohajeri
M. F. Cahn	T. J. Cruise	E. C. Wert
	R. R. Kurth	

RESEARCH OBJECTIVES AND SUMMARY OF RESEARCH

The work of our group can be divided into three major areas.

1. Sonar

The central problem is still the development of effective processing techniques for the output of an array with a large number of sensors. Specific topics of current interest are the following.

a. In Quarterly Progress Report No. 84 (pages 225-226) we described a state-variable formulation for the estimation and detection problem when the signal is a sample function from a nonstationary process that has passed through a dispersive medium. Work has continued in this area and we are attempting to find effective solution procedures for the equations that result. The state of current progress is described in A. B. Baggeroer's doctoral thesis.[1]

b. Iterative techniques to measure the interference and modify the array processor are still being studied. The current work includes both the study of fixed and adaptive arrays.

2. Communications

a. Digital Systems

The work on decision feedback systems that was described in Quarterly Progress Report No. 84 (pages 225-226) has been completed and is summarized in M. E. Austin's doctoral thesis.[2] An algorithm that used the past decisions in order to attempt to eliminate intersymbol interference caused by the dispersive nature of the channel was devised. Performance of this system was simulated and its probability of error was calculated. It was found to be a very effective procedure and, even in the presence of decision errors, operated reasonably well. Throughout this work perfect measurement of the channel was assumed. Further studies will include the effect of errors in channel measurement on the decision feedback system.

We have continued to work on the problem of evaluating the performance in the problem of detecting Gaussian signals in Gaussian noise. Some of the current results are summarized in Section XXIX-A. L. D. Collins has also formulated a signal design problem and is attempting to obtain an effective solution procedure for this. The results of this investigation will then be applied to a number of design problems in the radar and sonar field.

The study of digital and analog systems operating when there is a feedback channel

*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), and in part by U. S. Navy Purchasing Office Contract N00140-67-C-0210.

available from receiver to transmitter continues. The performance of several suboptimal systems has been computed, and work continues on the design of optimal systems and related problem design. Some of the current results are summarized in Section XXIX-B.

b. Analog Systems

(i) Work continues on the problem of estimating continuous waveforms in real time. We are using a state-variable approach based on Markov processes. Several specific problems have been studied experimentally. In order to investigate the accuracy obtainable and the complexity required in these systems by allowing the state equation to be nonlinear, we may also include interesting problems such as parameter estimation. This modification has been included and several parameter estimation problems are being studied experimentally.

(ii) One of the more significant results in linear estimation theory is the closed-form error expression developed originally by M. C. Yovits and J. L. Jackson.[3] We have been able to extend this result to a much larger class of problems. The current state of this work is described in M. Mohajeri's Master's thesis.[4] Research will continue along these lines.

3. Random Process Theory and Applications

a. State Variable and Continuous Markov Process Techniques

(i) In Quarterly Progress Report No. 84 (pages 226-227) we described an effective method for obtaining solutions to the Fredholm integral equation. As a part of this technique we found the Fredholm determinant. Subsequent research has shown that a similar determinant arises in a number of problems of interest. Specifically, we have been able to formulate several interesting design problems and carry through the solution. Baggeroer[5] has studied the problem of designing a signal under an energy and mean-square bandwidth constraint to combat the effects of non-white additive noise. The actual solution exhibited a number of interesting phenomena whose implications are now the subject of further research. Furthermore, the problems of peak-power and hard-bandwidth constraints have been formulated, but the solution has not yet been carried out. A related problem is the one of signal design for detection in a discrete resolution environment. E. C. Wert[6] is studying methods of solution for this problem.

A problem that is apparently dissimilar is that of estimating the parameters of a Gaussian random process. We have derived a bound on the minimum mean-square error obtainable by using any estimation scheme. Also, we have been able to express this bound in terms of the Fredholm determinant. This gives us a convenient computational procedure for actually evaluating the mean-square error bound.

(ii) Most of the channels of interest and practice are bandpass channels. The value of complex notation in dealing with problems of this type is well known. We have formulated a complex state-variable representation and have derived a number of the results needed in order to study problems of this type.[7]

b. System Identification Problem

The system identification problem is still an item of research. Applications of interest include measurement of spatial noise fields, random process statistics, and linear system functions.

c. Detection Techniques

Various extentions of the Gaussian detection problem are being studied. A particular topic of current interest is the detection of non-Gaussian Markov processes.

H. L. Van Trees

References

1. A. B. Baggeroer, "State Variables, Fredholm Theory, and Optimal Communication," Sc.D. Thesis, Department of Electrical Engineering, M.I.T., January 1968.
2. M. E. Austin, "Decision-Feedback Equalization for Digital Communication over Dispersive Channels" (to appear as Research Laboratory of Electronics Technical Report 461 and Lincoln Laboratory Technical Report 437).
3. M. C. Yovits and J. L. Jackson, "Linear Filter Optimization with Game Theory Considerations," IRE National Convention Record, Pt. 4, pp. 193-199, 1955.
4. M. Mohajeri, "Closed Form Error Expressions in Linear Filtering," S.M. Thesis, Department of Electrical Engineering, M.I.T., January 1968.
5. A. B. Baggeroer, "Optimal Signal Design for Additive Colored Noise Channels via State Variables," Internal Memorandum, Research Laboratory of Electronics, M.I.T., Cambridge, Massachusetts, 1967 (unpublished).
6. E. C. Wert, "A State Variable Approach to a Signal Design Problem," Sc.D. Thesis, Department of Electrical Engineering, M.I.T., January 1968.
7. A. B. Baggeroer, L. D. Collins, and H. L. Van Trees, "Complex State Variables: Theory and Applications" (unpublished).

A. ASYMPTOTIC APPROXIMATIONS TO THE ERROR PROBABILITY FOR DETECTING GAUSSIAN SIGNALS

In Quarterly Progress Report No. 85 (pages 253-265), we discussed the application of tilted probability distributions to the problem of evaluating the performance of optimum detectors for Gaussian signals received in additive Gaussian noise. The present report describes further work on this topic. We first present an asymptotic expansion for the error probabilities. The leading term in this expansion is asymptotically the same as the approximation postulated in the previous report. Second, we discuss the calculation of the semi-invariant moment-generating function $\mu(s)$, with particular emphasis on the case in which the random processes of interest can be modeled via state variables. This includes as a subclass all stationary processes with rational spectra. We conclude with a formulation of a signal design problem.

The problem that we are considering is the general Gaussian binary detection problem.

$$\begin{aligned} H_1&: \; r(t) = s_1(t) + m_1(t) + w(t) \\ H_2&: \; r(t) = s_2(t) + m_2(t) + w(t) \end{aligned} \qquad T_i \leq t \leq T_f, \tag{1}$$

where $s_1(t)$ and $s_2(t)$ 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; and w(t) is a sample function of white Gaussian noise with spectral density $N_o/2$.

1. An Asymptotic Expansion for the Probability of Error

We shall now derive an asymptotic expansion for the error probabilities. We have previously shown that

$$\Pr[\epsilon|H_1] = \int_{\gamma}^{\infty} p_{\ell_s}(L) \exp[\mu(s)-sL]\, dL \tag{2}$$

$$\Pr[\epsilon|H_2] = \int_{-\infty}^{\gamma} p_{\ell_s}(L) \exp[\mu(s)+(1-s)L]\, dL. \tag{3}$$

For simplicity, we shall treat only $\Pr[\epsilon|H_1]$. The derivation for $\Pr[\epsilon|H_2]$ is very similar.

First, we introduce a normalized, zero-mean random variable

$$z_s = \frac{\ell_s - \dot{\mu}(s)}{\sqrt{\ddot{\mu}(s)}}. \tag{4}$$

Then

$$\Pr[\epsilon|H_1] = \exp[\mu(s)-s\dot{\mu}(s)] \int_0^{\infty} e^{-s\sqrt{\ddot{\mu}(s)}\,Z}\; p_{z_s}(Z)\, dZ. \tag{5}$$

Before proceeding from Eq. 5, let us point out the motivation for introducing the tilted random variable ℓ_s (and subsequently z_s).

One of the serious practical problems that we encounter in the straightforward evaluation of the error probability is that we are generally interested in the behavior far out on the tail of the probability density. Since the test statistic is made up of a large number of statistically independent components, we would like to apply the Central Limit theorem. This theorem, however, is of little use when the region of interest is the tail of the probability density.

But observe that in our alternative error expression, Eqs. 2 and 3, we no longer are integrating under the tail of a probability density, but rather we start integrating at the mean of the tilted variables ℓ_s. Furthermore, the integrand contains a decaying exponential factor that results in the value of the integral being determined primarily by the behavior of $p_{\ell_s}(L)$ <u>near the mean</u> rather than on the tail. Thus, we expect, at least heuristically, that the Central Limit theorem may be useful in approximating the error probabilities.

Unfortunately, in most cases of interest $p_{\ell_s}(L)$ does not tend to a Gaussian distribution in the limit as the number of independent components goes to infinity. This is a consequence of our covariance functions being positive definite, square integrable functions from which it immediately follows that the variance of ℓ_s remains finite as the number of independent components goes to infinity. In this case, a necessary and sufficient condition that $p_{\ell_s}(L)$ approaches the Gaussian density is that each component random variable in the sum be Gaussian,[1] which is not the case except in the known signal-detection problem. Experience has shown us, however, that the limiting distribution, while not converging to the Gaussian distribution, does not differ a great deal from the Gaussian. Therefore, it is fruitful to make an expansion of $p_{z_s}(Z)$ which is related to the Gaussian distribution. Such an expansion is the Edgeworth expansion, the first few terms of which are given below.[2]

$$\begin{aligned} p_{z_s}(Z) = \phi(Z) &- \left[\frac{\gamma_3}{6}\phi^{(3)}(Z)\right] \\ &+ \left[\frac{\gamma_4}{24}\phi^{(4)}(Z) + \frac{\gamma_3^2}{72}\phi^{(6)}(Z)\right] \\ &- \left[\frac{\gamma_5}{120}\phi^{(5)}(Z) + \frac{\gamma_3\gamma_4}{144}\phi^{(7)}(Z) + \frac{\gamma_3^2}{1296}\phi^{(4)}(Z)\right] \\ &+ \left[\frac{\gamma_6}{720}\phi^{(6)}(Z) + \frac{\gamma_4^2}{1152}\phi^{(8)}(Z) + \frac{\gamma_3\gamma_5}{720}\phi^{(8)}(Z) \right. \\ &\qquad \left. + \frac{\gamma_3^2\gamma_4}{1728}\phi^{(10)}(Z) + \frac{\gamma_3^4}{31104}\phi^{(12)}(Z)\right] \\ &- \dots \end{aligned} \tag{6}$$

where

$$\gamma_n = \frac{\mu^{(n)}(s)}{[\ddot{\mu}(s)]^{n/2}} \tag{7}$$

and

$$\phi^{(k)}(Z) = \frac{d^k}{dZ^k}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{Z^2}{2}\right) \qquad k = 0, 1, 2, \dots \tag{8}$$

This expansion may be obtained from a Hermite expansion of $p_{z_s}(Z)$ (also sometimes called a Gram-Charlier series) upon reordering the terms. It has the further advantage, for our purposes, that the coefficients are expressed in terms of the semi-invariants of the random variable z_s, which are readily computed from $\mu(s)$.

We now substitute the expansion (8) in the integral

$$\int_0^\infty p_{z_s}(Z) \exp[-s\sqrt{\ddot{\mu}(s)}\, Z]\, dZ,$$

and interchange orders of integration and summation. We are then faced with the task of evaluating integrals of the form

$$I_k(\alpha) = \int_0^\infty \phi^{(k)}(Z) \exp(-\alpha Z)\, dZ. \tag{9}$$

Repeated integrations by parts enable us to express these integrals in terms of the Gaussian error function

$$\Phi(X) = \int_{-\infty}^{X} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{Z^2}{2}\right) dz. \tag{10}$$

The first few integrals are

$$I_o(\alpha) = \Phi(-\alpha) \exp\left(\frac{\alpha^2}{2}\right) \tag{11a}$$

$$I_1(\alpha) = \alpha\ I_o(\alpha) - \frac{1}{\sqrt{2\pi}} \tag{11b}$$

$$I_2(\alpha) = \alpha^2 I_o(\alpha) - \frac{2}{\sqrt{2\pi}} \tag{11c}$$

$$I_3(\alpha) = \alpha^3 I_o(\alpha) + \frac{1}{\sqrt{2\pi}} (1-\alpha^2) \tag{11d}$$

$$I_4(\alpha) = \alpha^4 I_o(\alpha) + \frac{1}{\sqrt{2\pi}} (\alpha-\alpha^3) \tag{11e}$$

$$I_5(\alpha) = \alpha^5 I_o(\alpha) + \frac{1}{\sqrt{2\pi}} (-3+\alpha^2-\alpha^4) \tag{11f}$$

$$I_6(\alpha) = \alpha^6 I_o(\alpha) + \frac{1}{\sqrt{2\pi}} (-3\alpha+\alpha^3-\alpha^5). \tag{11g}$$

Thus we have our approximation to the integral

$$\int_0^\infty p_{z_s}(Z) \exp[-s\sqrt{\ddot{\mu}(s)}\, Z]\, dZ,$$

and therefore also to Pr $[\epsilon | H_1]$. We simply evaluate $\mu(s)$ and its derivatives and then substitute. This procedure is far too complex to be used analytically, but if we use a digital computer to obtain $\mu(s)$ — as we must in many problems — then there is no disadvantage in having the computer evaluate our error approximation, too.

2. Calculation of $\mu(s)$ for Gaussian Random Processes

It should be pointed out that we have not used the Gaussian assumption up to this point. We now make use of this assumption in computing $\mu(s)$, in which case we have an explicit closed-form expression.[3]

$$\mu(s) = \frac{2}{N_o} \frac{s}{2} \left\{ \int_{T_i}^{T_f} \xi_{s_1}\left(t \middle| t: \frac{N_o}{2}\right) dt \right.$$

$$+ \frac{1-s}{2} \int_{T_i}^{T_f} \xi_{s_2}\left(t \middle| t: \frac{N_o}{2}\right) dt$$

$$- \frac{1}{2} \int_{T_i}^{T_f} \xi_{comp}\left(t \middle| t: \frac{N_o}{2}, s\right) dt$$

$$\left. - \frac{s(1-s)}{2} \int_{T_i}^{T_f} \left[(m_2(t)-m_1(t)) - \int_{T_i}^{T_f} h_{comp}(t,\tau: s)[m_2(\tau)-m_1(\tau)]\, d\tau \right]^2 dt \right\}. \tag{12}$$

Here, ξ_{s_1}, ξ_{s_2}, and ξ_{comp} denote the minimum mean-square linear estimation error for estimating $s_1(t)$, $s_2(t)$, and $s_{comp}(t) \triangleq \sqrt{s}\, s_1(t) + \sqrt{1-s}\, s_2(t)$, respectively, when observed over $[T_i, t]$ in additive white noise of spectral height $N_o/2$; and $h_{comp}(t,\tau: s)$ denotes the minimum mean-square linear estimator for estimating the composite process $s_{comp}(t)$. The last term in Eq. 12 can also be interpreted as $d^2_{comp}(s)$, the output signal-to-noise ratio for the problem of deciding which of the known signals $m_1(t)$ or $m_2(t)$ was sent when observed in colored Gaussian noise with covariance

$$K_n(t,\tau: s) = K_{comp}(t,\tau: s) + \frac{N_o}{2}\delta(t-\tau)$$

$$= s K_1(t,\tau) + (1-s) K_2(t,\tau) + \frac{N_o}{2}\delta(t-\tau). \tag{13}$$

An important subclass of problems in which we can readily compute the various terms that comprise $\mu(s)$ is composed of problems in which the random processes can be modeled as the output of linear state-variable systems that are driven by white Gaussian noise. This model includes all stationary processes with rational spectra as one important subclass of interesting problems.

The state-variable model allows us to use the results of the Kalman-Bucy formulation of the optimum linear filtering problem to determine the optimum receiver structure as well as to calculate $\mu(s)$.[4, 5] The straightforward way to compute the first three terms in Eq. 12 is to solve the appropriate matrix Ricatti equation and then integrate the result over the time interval $[T_i, T_f]$ as indicated.

We assume the following state-variable model for the random process generation.

$$\frac{d}{dt}\underline{x}(t) = \underline{F}(t)\,\underline{x}(t) + \underline{G}(t)\,\underline{u}(t) \tag{14}$$

$$y(t) = \underline{C}(t)\,\underline{x}(t) \tag{15}$$

$$E[\underline{u}(t)\,\underline{u}^T(\tau)] = \underline{Q}(t)\,\delta(t-\tau) \tag{16}$$

$$E[\underline{u}(t)] = \underline{0}. \tag{17}$$

Then

$$\xi_y\left(t \,\middle|\, t:\frac{N_o}{2}\right) = \underline{C}(t)\,\underline{\Sigma}(t)\,\underline{C}^T(t), \tag{18}$$

where

$$\frac{d}{dt}\underline{\Sigma}(t) = \underline{F}(t)\,\underline{\Sigma}(t) + \underline{\Sigma}(t)\,\underline{F}^T(t) + \underline{G}(t)\,\underline{Q}(t)\,\underline{G}^T(t) - \underline{\Sigma}(t)\,\underline{C}^T(t)\,\underline{R}^{-1}(t)\,\underline{C}(t)\,\underline{\Sigma}(t), \tag{19}$$

with

$$\underline{\Sigma}(T_i) = \underline{\Sigma}_o. \tag{20}$$

A useful computational alternative enables us to express the first three terms in $\mu(s)$ as the solution to sets of linear differential equations evaluated at a point. In particular, for the important case in which Eqs. 14 through 17 are time-invariant, which includes all stationary processes with rational spectra, the solution to this linear set of equations can be obtained in closed form via the appropriate transition matrix.

We use the linear system of equations which is equivalent to Eq. 19.[4]

$$\frac{d}{dt}\underline{\phi}_1(t) = \underline{F}(t)\,\underline{\phi}_1(t) + \underline{G}(t)\,\underline{Q}(t)\,\underline{G}^T(t)\,\underline{\phi}_2(t) \tag{21}$$

$$\frac{d}{dt}\underline{\phi}_2(t) = \underline{C}^T(t)\,\frac{2}{N_o}\,\underline{C}(t)\,\underline{\phi}_1(t) - \underline{F}^T(t)\,\underline{\phi}_2(t) \tag{22}$$

$$\underline{\phi}_1(T_i) = \underline{\Sigma}_o \tag{23}$$

$$\underline{\phi}_2(T_i) = \underline{I}. \tag{24}$$

Then

$$\frac{2}{N_o}\int_{T_i}^{T_f} \xi_y\left(t\,|\,t: \frac{N_o}{2}\right) dt$$

$$= \frac{2}{N_o}\int_{T_i}^{T_f} \underline{C}(t)\, \underline{\phi}_1(t)\, \underline{\phi}_2^{-1}(t)\, \underline{C}^T(t)\, dt$$

$$= \int_{T_i}^{T_f} \text{Tr}\left[\left(\underline{C}^T(t)\frac{2}{N_o}\underline{C}(t)\, \underline{\phi}_1(t)\right) \underline{\phi}_2^{-1}(t)\right] dt$$

$$= \int_{T_i}^{T_f} \text{Tr}\left[\underline{\phi}_2^{-1}(t)\, d\phi_2(t)\right] + \int_{T_i}^{T_f} \text{Tr}\,[\underline{F}(t)]\, dt$$

$$= \int_{T_i}^{T_f} d[\ell n \det \underline{\phi}_2(t)] + \int_{T_i}^{T_f} \text{Tr}\,[\underline{F}(t)]\, dt$$

$$= \ell n \det \underline{\phi}_2(T_f) + \int_{T_i}^{T_f} \text{Tr}\,[\underline{F}(t)]\, dt. \tag{25}$$

In the derivation above we have made use of several properties of the trace operation.[6]

The first term in Eq. 25 is readily computed in terms of the transition matrix of the canonical system

$$\frac{d}{dt}\,\underline{\theta}(t, T_i) = \begin{bmatrix} \underline{F}(t) & \underline{G}(t)\, \underline{Q}(t)\, \underline{G}^T(t) \\ \underline{C}^T(t)\frac{2}{N_o}\underline{C}(t) & -\underline{F}^T(t) \end{bmatrix} \underline{\theta}(t, T_i) \tag{26}$$

$$\underline{\theta}(T_i, T_i) = \underline{I}. \tag{27}$$

If we partition the transition matrix

$$\underline{\theta}(t, T_i) = \begin{bmatrix} \underline{\theta}_{11}(t, T_i) & \underline{\theta}_{12}(t, T_i) \\ \underline{\theta}_{21}(t, T_i) & \underline{\theta}_{22}(t, T_i) \end{bmatrix}, \tag{28}$$

then

$$\underline{\phi}_2(T_f) = \underline{\theta}_{21}(T_f, T_i)\, \underline{\Sigma}_o + \underline{\theta}_{22}(T_f, T_i). \tag{29}$$

In the case of constant-parameter systems, it is particularly easy to compute the transition matrix in terms of a matrix exponential.[6]

It is interesting and useful to observe that $\mu(s:t)$, where we have explicitly indicated the dependence on the time t, may be regarded as a state variable of a (realizable) dynamic system. Straightforward differentiation of Eq. 12 yields

$$\frac{\partial\mu(s:t)}{\partial t} = \frac{s}{N_o}\xi_{s_1}\left(t\,\middle|\,t:\frac{N_o}{2}\right) + \frac{1-s}{N_o}\xi_{s_2}\left(t\,\middle|\,t:\frac{N_o}{2}\right) - \frac{1}{N_o}\xi_{comp}\left(t\,\middle|\,t:\frac{N_o}{2}, s\right)$$

$$-\frac{s(1-s)}{N_o}\left[(m_2(t)-m_1(t)) - \int_{T_i}^{t} h_{comp}(t,\tau:s)(m_2(\tau)-m_1(\tau))\,d\tau\right]^2, \tag{30}$$

with the initial condition

$$\mu(s:T_i) = 0. \tag{31}$$

Since we can compute each term on the right-hand side of Eq. 30 in real time, that is, as the outputs of realizable dynamic systems, $\mu(s:t)$ can also be computed as the output of such a system. This naturally leads us to consider optimization problems in which it is desired to use $\mu(s:T_f)$ as a performance index that is to be minimized as a function of various system parameters, subject to some constraints. We have at our disposal the techniques of optimal control theory (Pontryagin's minimum principle) which provide a set of necessary conditions for the minimum.[7] One such problem is formulated below.

3. A Signal Design Problem

We shall formulate a typical problem in optimal signal design. By using Pontryagin's maximum principle, we obtain a set of differential equations and boundary conditions which the optimum signal must satisfy. The (numerical) solution of these equations and the classification of the various solutions is a topic of current research.

The communication problem of interest is diagrammed in Fig. XXIX-1. This problem is closely related to the more realistic problem of communicating over a Rayleigh fading channel. We wish to pick s(t), subject to constraints on its energy and bandwidth, to maximize the performance of the optimum detector. The performance measure that we would like to minimize is the error probability, but the complexity of the error expressions is not very encouraging. Instead we minimize $\mu(s)$ which appears as the dominant term in our error approximations.

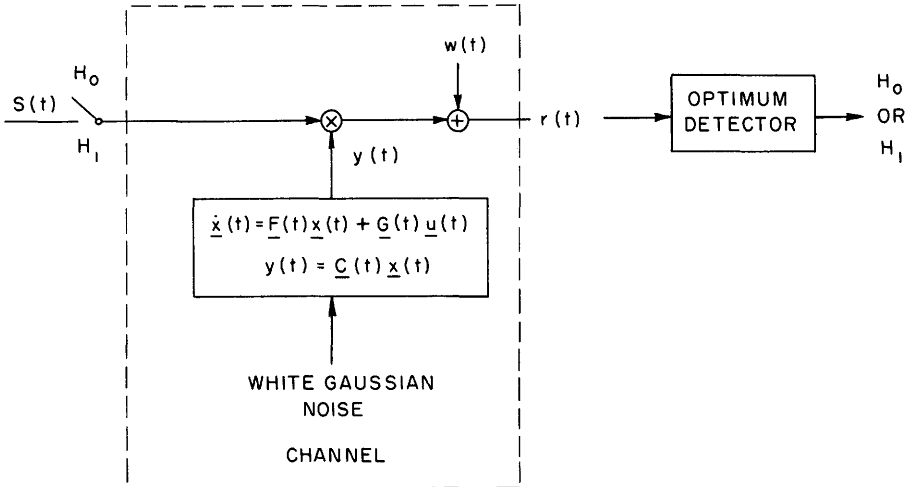

Fig. XXIX-1. Communication system model.

Thus we have the cost functional

$$J = \frac{1-s}{N_o} \int_{T_i}^{T_f} \xi_1\left(t \middle| t: \frac{N_o}{2}\right) dt - \frac{1}{N_o} \int_{T_i}^{T_f} \xi_2\left(t \middle| t: \frac{N_o}{2}\right) dt, \tag{32}$$

where

$$\xi_1\left(t \middle| t: \frac{N_o}{2}\right) = s^2(t)\, \underline{C}(t)\, \underline{\Sigma}_1(t)\, \underline{C}^T(t) \tag{33a}$$

$$\xi_2\left(t \middle| t: \frac{N_o}{2}\right) = s^2(t)\, \underline{C}(t)\, \underline{\Sigma}_2(t)\, \underline{C}^T(t) \tag{33b}$$

$$\frac{d}{dt}\underline{\Sigma}_1(t) = \underline{F}(t)\, \underline{\Sigma}_1(t) + \underline{\Sigma}_1(t)\, \underline{F}^T(t) + \underline{G}(t)\, \underline{Q}(t)\, \underline{G}^T(t) - s^2(t) \frac{2}{N_o} \underline{\Sigma}_1(t)\, \underline{C}(t)\, \underline{C}^T(t)\, \underline{\Sigma}_1(t) \tag{34a}$$

$$\frac{d}{dt}\underline{\Sigma}_2(t) = \underline{F}(t)\, \underline{\Sigma}_2(t) + \underline{\Sigma}_2(t)\, \underline{F}^T(t) + (1-s)\, \underline{G}(t)\, \underline{Q}(t)\, \underline{G}^T(t) - s^2(t) \frac{2}{N_o} \underline{\Sigma}_2(t)\, \underline{C}(t)\, \underline{C}^T(t)\, \underline{\Sigma}_2(t) \tag{34b}$$

$$\underline{\Sigma}_1(T_i) = \underline{\Sigma}_o \tag{35a}$$

$$\underline{\Sigma}_2(T_i) = (1-s)\underline{\Sigma}_o. \tag{35b}$$

By the results of the previous section, this can be rewritten in terms of the solution to a linear system of equations.

$$J = \frac{1-s}{2}\left[\ln \det \underline{X}_2(T_f) + \int_{T_i}^{T_f} \mathrm{Tr}\,[\underline{F}(t)]\,dt\right]$$

$$-\frac{1}{2}\left[\ln \det \underline{X}_4(T_f) + \int_{T_i}^{T_f} \mathrm{Tr}\,[\underline{F}(t)]\,dt\right]$$

$$= \frac{1-s}{2} \ln \det \underline{X}_2(T_f) - \frac{1}{2} \ln \det \underline{X}_4(T_f) - \frac{s}{2}\int_{T_i}^{T_f} \mathrm{Tr}\,[\underline{F}(t)\,dt]. \tag{36}$$

The last term is independent of the modulation and may be ignored. The first two terms in Eq. 36 are obtained from the solution of the following system of differential equations:

$$\frac{d}{dt}\underline{X}_1(t) = \underline{F}(t)\,\underline{X}_1(t) + \underline{G}(t)\,\underline{Q}(t)\,\underline{G}^T(t)\,\underline{X}_2(t) \tag{37a}$$

$$\frac{d}{dt}\underline{X}_2(t) = s^2(t)\,\underline{C}^T(t)\,\frac{2}{N_o}\,\underline{C}(t)\,\underline{X}_1(t) - \underline{F}^T(t)\,\underline{X}_2(t) \tag{37b}$$

$$\frac{d}{dt}\underline{X}_3(t) = \underline{F}(t)\,\underline{X}_3(t) + (1-s)\,\underline{G}(t)\,\underline{Q}(t)\,\underline{G}^T(t)\,\underline{X}_4(t) \tag{37c}$$

$$\frac{d}{dt}\underline{X}_4(t) = s^2(t)\,\underline{C}^T(t)\,\frac{2}{N_o}\,\underline{C}(t)\,\underline{X}_3(t) - \underline{F}^T(t)\,\underline{X}_4(t), \tag{37d}$$

with the boundary conditions

$$\underline{X}_1(T_i) = \underline{\Sigma}_o \tag{38a}$$

$$\underline{X}_2(T_i) = \underline{I} \tag{38b}$$

$$\underline{X}_3(T_i) = (1-s)\underline{\Sigma}_o \tag{38c}$$

$$\underline{X}_4(T_i) = \underline{I}. \tag{38d}$$

Thus we have a fixed-terminal time problem with a terminal cost.

Also, we have the energy and "bandwidth" constraints

$$E = \int_{T_i}^{T_f} s^2(t)\,dt \tag{39}$$

$$B^2 = \frac{1}{E}\int_{T_i}^{T_f} (\dot{s}(t))^2\,dt. \tag{40}$$

We introduce additional state variables to incorporate these constraints. It also is convenient to let the control variable u(t) be the derivative of s(t), and to let s(t) be an additional state variable.

$$x_5(t) = s(t) \tag{41}$$

$$\dot{x}_5(t) = \dot{s}(t) = u(t) \tag{42}$$

$$\dot{x}_6(t) = s^2(t) = x_5^2(t) \tag{43}$$

$$\dot{x}_7(t) = (\dot{s}(t))^2 = u^2(t). \tag{44}$$

Thus the constraints become target values (boundary conditions) for $x_6(t)$ and $x_7(t)$. Furthermore, we impose the additional boundary conditions $s(T_i) = s(T_f) = 0$, in order that the signal have no jump discontinuities at the ends of the interval. Thus

$$x_5(T_i) = x_5(T_f) = 0 \tag{45}$$

$$x_6(T_f) = E \tag{46}$$

$$x_7(T_f) = E B^2. \tag{47}$$

Now we write down the Hamiltonian.

$$\begin{aligned} H = {} & \mathrm{Tr}\left[\underline{X}_1(t)\, \underline{P}_1^T(t)\right] + \mathrm{Tr}\left[\underline{X}_2(t)\underline{P}_2^T(t)\right] \\ & + \mathrm{Tr}\left[\underline{X}_3(t)\underline{P}_3^T(t)\right] + \mathrm{Tr}\left[\underline{X}_4(t)\underline{P}_4^T(t)\right] \\ & + x_5(t)\, p_5(t) + x_6(t)\, p_6(t) + x_7(t)\, p_7(t) \\ = {} & \mathrm{Tr}\left[\left(\underline{F}(t)\underline{X}_1(t) + \underline{G}(t)\underline{Q}(t)\underline{G}^T(t)\underline{X}_2(t)\right)\underline{P}_1^T(t)\right] \\ & + \mathrm{Tr}\left[\left(x_5^2(t)\underline{C}^T(t)\,\frac{2}{N_o}\,\underline{C}(t)\underline{X}_1(t) - \underline{F}^T(t)\underline{X}_2(t)\right)\underline{P}_2^T(t)\right] \\ & + \mathrm{Tr}\left[\left(\underline{F}(t)\underline{X}_3(t) + (1-s)\underline{G}(t)\underline{Q}(t)\underline{G}^T(t)\underline{X}_4(t)\right)\underline{P}_3^T(t)\right] \\ & + \mathrm{Tr}\left[\left(x_5^2(t)\underline{C}^T(t)\,\frac{2}{N_o}\,\underline{C}(t)\underline{X}_3(t) - \underline{F}^T(t)\underline{X}_4(t)\right)\underline{P}_4^T(t)\right] \\ & + u(t)\, p_5(t) + x_5^2(t)\, p_6(t) + u^2(t)\, p_7(t), \end{aligned} \tag{48}$$

where

$\underline{P}_1(t)$, $\underline{P}_2(t)$, $\underline{P}_3(t)$, $\underline{P}_4(t)$, $p_5(t)$, $p_6(t)$, and $p_7(t)$

are the co-state (Lagrange multiplier) functions. We now obtain the differential equations for the co-states.

$$\frac{d}{dt}\underline{P}_1(t) = -\frac{\partial H}{\partial \underline{X}_1} = -\underline{F}^T(t)\,\underline{P}_1(t) - x_5^2(t)\,\underline{C}^T(t)\frac{2}{N_o}\underline{C}(t)\,\underline{P}_2(t) \tag{49a}$$

$$\frac{d}{dt}\underline{P}_2(t) = -\frac{\partial H}{\partial \underline{X}_2} = -\underline{G}(t)\,\underline{Q}(t)\,\underline{G}^T(t)\,\underline{P}_1(t) + \underline{F}(t)\,\underline{P}_2(t) \tag{49b}$$

$$\frac{d}{dt}\underline{P}_3(t) = -\frac{\partial H}{\partial \underline{X}_3} = -\underline{F}^T(t)\,\underline{P}_3(t) - x_5^2(t)\,\underline{C}^T(t)\frac{2}{N_o}\underline{C}(t)\,\underline{P}_4(t) \tag{49c}$$

$$\frac{d}{dt}\underline{P}_4(t) = -\frac{\partial H}{\partial \underline{X}_4} = -(1-s)\,\underline{G}(t)\,\underline{Q}(t)\,\underline{G}^T(t)\,\underline{P}_3(t) + \underline{F}(t) + \underline{P}_4(t) \tag{49d}$$

$$\frac{d}{dt}p_5(t) = -\frac{\partial H}{\partial x_5} = -\frac{2}{N_o}x_5(t)\left\{2\mathrm{Tr}\left[\underline{C}^T(t)\underline{C}(t)\left(\underline{X}_1(t)\underline{P}_2^T(t)+(1-s)\underline{X}_3(t)\underline{P}_4^T(t)\right)\right]+p_6(t)\right\} \tag{49e}$$

$$\frac{d}{dt}p_6(t) = -\frac{\partial H}{\partial x_6} = 0 \rightarrow p_6(t) = \text{constant} \tag{49f}$$

$$\frac{d}{dt}p_7(t) = -\frac{\partial H}{\partial x_7} = 0 \rightarrow p_7(t) = \text{constant.} \tag{49g}$$

Furthermore, we require that the Hamiltonian be minimized.

$$\frac{\partial H}{\partial u} = p_5(t) + 2u(t)\,p_7 = 0$$

$$\rightarrow u(t) = -\frac{p_5(t)}{2p_7} \tag{50}$$

$$\frac{\partial^2 H}{\partial u^2} = 2p_7 > 0$$

$$\rightarrow p_7 > 0 \tag{51}$$

Thus we can eliminate u(t) from our equations (in particular, from Eq. 42), to obtain

$$\dot{x}_5(t) = -\frac{p_5(t)}{2p_7}, \tag{52}$$

Finally, there are the boundary conditions on the co-states. Since we have a terminal cost,

$$\underline{P}_1(T_f) = \frac{\partial J}{\partial \underline{X}_1} = \underline{0} \tag{53a}$$

$$\underline{P}_2(T_f) = \frac{\partial J}{\partial \underline{X}_2} (1-s)\left[\underline{X}_2^{-1}(T_f)\right]^T \tag{53b}$$

$$\underline{P}_3(T_f) = \frac{\partial J}{\partial \underline{X}_3} = \underline{0} \tag{53c}$$

$$\underline{P}_4(T_f) = \frac{\partial J}{\partial \underline{X}_4} = -\left[\underline{X}_4^{-1}(T_f)\right]^T. \tag{53d}$$

Therefore, we have obtained a system of 10 coupled nonlinear differential equations (8 are matrix equations) with 10 boundary conditions. These are Eqs. 37a through 37d, 52, and 49a through 49c with the boundary conditions (38a) through (38d), (45) and (53a) through (53d). In addition, there are two undetermined constants (Lagrange multipliers), p_6 and p_7, associated with the energy and bandwidth constraints.

Two factors contribute to the difficulty of solving these equations: one is the fact that they are nonlinear, and the second is the split boundary condition. At present, we are investigating possible solution techniques.

4. Summary

We have obtained an asymptotic approximation for the error probabilities for optimum detection of random signals in terms of the semi-invariant moment-generating function $\mu(s)$. We then obtained closed-form expressions for $\mu(s)$, by assuming that the signals are sample functions from Gaussian random processes. When these processes can be modeled as the output of linear systems that can be modeled via state variables, we were able to obtain $\mu(s)$ in a particularly simple form from the solution to a set of linear differential equations. We then applied these results to the problem of designing an optimal signal for communicating over a singly-spread fading channel. We employed the minimum principle of Pontryagin to obtain a system of differential equations and boundary conditions which the optimum signal must satisfy.

L. D. Collins

References

1. H. Cramèr, Mathematical Methods of Statistics (Princeton University Press, Princeton, N.J., 1956).
2. M. Abramowitz and I. A. Stegun, Handbook of Mathematical Functions (National Bureau of Standards, Washington, D.C., 1964), Sec. 26.2.47.
3. L. D. Collins "Closed-form Expressions for the Fredholm Determinant for State-variable Covariance Functions (submitted to Correspondence Section, IEEE Proc.).

4. R. E. Kalman and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," Trans. ASME J. Basic Engrg., pp. 95-108, March 1961.
5. F. C. Schweppe, "Evaluation of Likelihood Functions for Gaussian Signals," IEEE Trans., Vol. IT-11, No. 1, pp. 61-70, January 1965.
6. M. Athans and F. C. Schweppe, "Gradient Matrices and Matrix Calculations," Technical Note 1965-53, Lincoln Laboratory, M.I.T., November 17, 1965.
7. M. Athans and P. L. Falb, Optimal Control (McGraw-Hill Publishing Co., New York, 1966).

B. SCHEME TO UTILIZE NOISELESS FEEDBACK TO REDUCE CHANNEL NOISE

1. Introduction

A much-studied communication problem is that of communicating over an additive white Gaussian noise channel at the ultimate error-free rate given by channel capacity. A great deal of effort has been expended toward developing systems that approach the channel capacity rate arbitrarily closely. Such systems involve block coding, sequential decoding, orthogonal signaling (see Wozencraft and Jacobs[1]), or other equally complex algorithms for approaching channel capacity. Recently, many authors[2-5] have demonstrated that simpler systems will achieve channel capacity if a noiseless feedback link is available from the receiver to the transmitter. The availability of a noiseless feedback link does not change the forward channel capacity.

In this report noiseless feedback will be applied to reduce the additive channel noise in a manner independent of the particular type of channel signaling. If this feedback technique is applied to the digital example mentioned above, channel capacity will be achieved without coding. If this technique is applied to a process-estimation system, the rate distortion bound on mean-square error can be obtained.

A noiseless feedback channel is shown to be effective in reducing the channel noise and leaving the transmitted signal uncorrupted. In other words, the over-all feedback system behaves exactly as the no-feedback system, but with a reduced "effective" additive noise.

2. No-Feedback System

Figure XXIX-2 shows a communication system without feedback. m(t) is the transmitted signal that depends on the information that is being conveyed through the channel. It might be the transmitted signal in a digital problem, a parameter-estimation problem or a process-estimation problem. In the example below a parameter estimation problem is studied; for the present, the exact form that m(t) takes is unimportant.

The receiver in Fig. XXIX-2 depends on the structure of the modulation and operates to recover the information conveyed through the channel. The receiver would be a

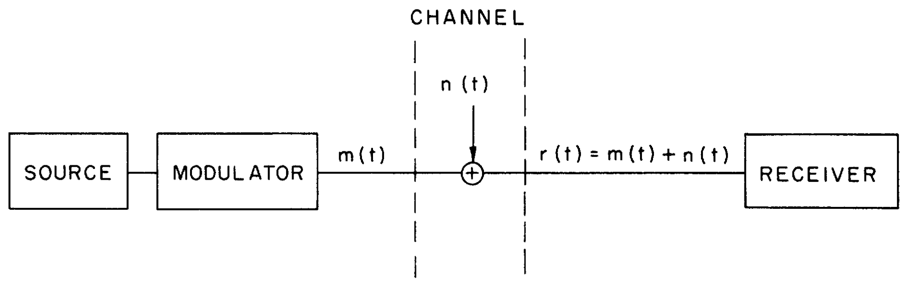

Fig. XXIX-2. No-feedback communication system.

decoder if block coding were used.

3. Feedback System

Figure XXIX-3 shows the system of Fig. XXIX-2 modified to include a feedback channel. Since the feedback channel is noiseless, a waveform, $\hat{m}(t)$, can be subtracted at the transmitter and added at the receiver without affecting the message waveform m(t). The gain K(t) is necessary in order to adjust the transmitter to meet power (energy) constraints that are assumed identical to the no-feedback transmitter requirements.

The "ESTIMATOR" in Fig. XXIX-3 represents a filter that generates the minimum mean-square estimate of m(t), given $K(\tau)\, m(\tau) + n(\tau)$ (ESTIMATOR input) for $\tau \leqslant t$. If K(t) is chosen so that the average transmitted power in the feedback system is identical to that of the no-feedback system, then

$$K^2(t)\; E\{(m(t)-\hat{m}(t))^2\} = E\{m^2(t)\},$$

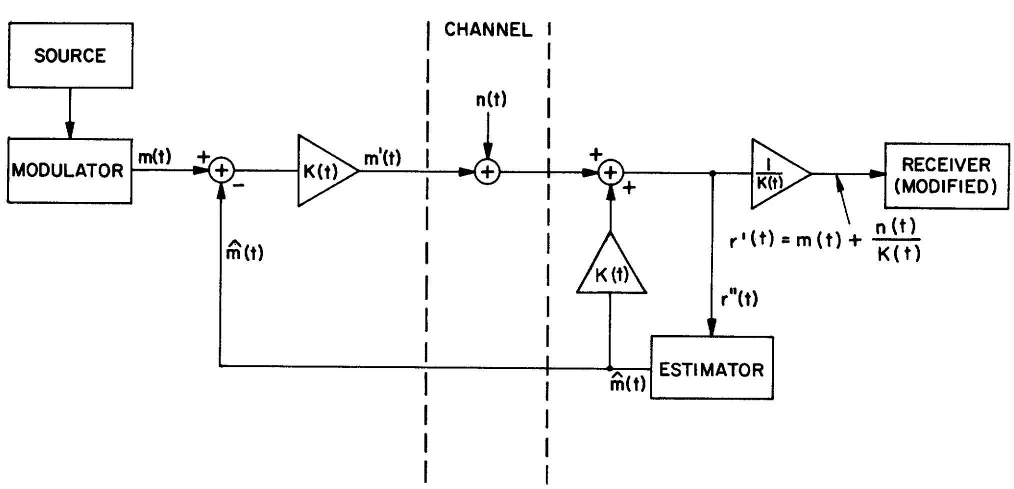

Fig. XXIX-3. Feedback communication system.

and $K(t) \leq 1$ for all t. This fact implies that the effective noise $n(t)/K(t)$ into the feedback receiver is less in magnitude than the noise n(t) input to the no-feedback receiver for all t; hence, the feedback system performs better.

Improvement will also be obtained for other choices of K(t) (satisfying the transmitted power constraint), although this fact is not as obvious as the choice of K(t) above. In some applications the estimate $\hat{m}(t)$ may be too complicated or impossible to compute. In such cases a suboptimum ESTIMATOR could be inserted, and thereby the feedback channel would be utilized to improve the system performance. The following example is in such a situation.

If $K(t) \neq$ constant, the receiver must be modified from the no-feedback receiver to incorporate time-variant noise.

4. Example

This brief example demonstrates some of the ideas presented for the conversion of a no-feedback system to one employing feedback.

Without feedback assume that the value of a random variable θ (uniform in $[-.5,.5]$) is to be conveyed across an additive white noise channel (noise density, $N_0/2$). The transmitter is limited in energy to E_0 in the time interval $[0,T]$. This problem is analogous to the feedback schemes of Omura[2] and of Schalkwijk and Kailath.[4]

Assume that the transmitter uses pulse amplitude modulation, with the height of the pulse being proportional to θ. Then

$$m(t) = \sqrt{\frac{12E_0}{T}}\,\theta \tag{1}$$

in order to maintain the transmitted energy constraint. The receiver is assumed to be the minimum mean-square linear estimator of θ; therefore, the estimate at t = T is

$$\hat{\theta}(T) = \int_0^T dt\, \frac{\sqrt{\frac{12E_0}{T}}\,\frac{1}{12}}{\frac{N_0}{2}+E_0}\, r(t). \tag{2}$$

The normalized (or fractional) variance of $\hat{\theta}(T)$ for this no-feedback scheme is

$$\xi_{\substack{\text{no}\\ \text{feedback}}} = \frac{1}{1+\frac{2E_0}{N_0}}. \tag{3}$$

This performance is optimal for the constraints of simple linear modulation and a linear receiver. A linear receiver is not optimal for this problem because of the non-Gaussian probability density of the random variable θ. Obviously, a more complex modulation scheme will work better.

In order to add feedback, the "ESTIMATOR" block in Fig. XXIX-3 needs to be determined. Assume that a Kalman filter is used to form the estimate $\hat{m}(t)$. The input to the filter is

$$r''(t) = K(t)\, m(t) + n(t) = K(t) \sqrt{\frac{12E_0}{T}}\, \theta + n(t), \tag{4}$$

and the estimate $\hat{m}(t)$ is

$$\hat{m}(t) = \sqrt{\frac{12E_0}{T}}\, \hat{\theta}(t), \tag{5}$$

with $\hat{\theta}(t)$ the output of the Kalman filter

$$\frac{d}{dt}\hat{\theta}(t) = -K^2(t)\, P(t)\, \frac{24E_0}{N_0T}\, \hat{\theta}(t) + \frac{2}{N_0} \sqrt{\frac{12E_0}{T}}\, K(t)\, P(t)\, r''(t) \tag{6}$$

$$\hat{\theta}(0) = 0.$$

The covariance

$$P(t) = \underline{E}\{(\hat{\theta}(t)-\theta)^2\} \tag{7}$$

satisfies

$$\dot{P}(t) = -\frac{24E_0}{N_0T} K^2(t)\, P^2(t)$$

$$P(0) = \frac{1}{12}. \tag{8}$$

The Kalman filter is not the minimum variance estimator of θ, because of the probability density of θ; hence, the ESTIMATOR is suboptimal and could be improved by a more complex filter.

With the further constraint that the mean power of the feedback system agree with the no-feedback system, K(t) is evaluated as

$$K^2(t)\, E[(m(t)-\hat{m}(t))^2] = \frac{12E_0}{T} K^2(t)\, P(t) = \frac{E_0}{T}. \tag{9}$$

The performance P(t) (from Eqs. 8 and 9) satisfies

$$P(t) = -\frac{2E_0}{N_0}\frac{P(t)}{T} \qquad P(0) = \frac{1}{12} \tag{10}$$

Integrating yields the performance of the feedback system:

$$\frac{P(T)}{\frac{1}{12}} = \xi_{\text{feedback}} = e^{-2E_0/N_0}.$$

This fractional error is less than the no-feedback error in Eq. 3. For large values of $2E_0/N_0$ the difference is considerable; the effective signal-to-noise ratio of the feedback system is approximately the no-feedback ratio exponentiated.

5. Conclusions

A technique for the utilization of a noiseless feedback channel has been proposed for use in arbitrary communication systems over white noise channels. The system is not designed to utilize the feedback channel optimally, but the technique presented turns out to be optimal for many communication systems.

This technique achieves improvement by inserting some feedback devices between the no-feedback transmitter and the no-feedback receiver. Virtually no changes are made in the basic signaling structure in converting from a no-feedback system to a feedback system. The feedback channel operates to reduce the transmitted power, but does not alter the type of message modulation/demodulation employed. Obviously, the actual transmitted signal is slightly different in the two systems.

T. J. Cruise

References

1. J. M. Wozencraft and I. M. Jacobs, Principles of Communication Engineering (John Wiley and Sons, Inc., New York, 1965).

2. J. K. Omura, "Signal Optimization for Channels with Feedback," Technical Report No. 7001-1, Stanford Electronics Laboratories, Stanford, California, August 1966.

3. G. L. Turin, "Comparison of Sequential and Nonsequential Detection Systems with Uncertainty Feedback," IEEE Trans., Vol. IT-12, No. 1, pp. 5-8, January 1966.

4. J. Schalkwijk and T. Kailath, "A Coding Scheme for Additive Noise Channels with Feedback – Part I: No Bandwidth Constraint," IEEE Trans., Vol. IT-12, No. 2, pp. 172-182, April 1966.

5. P. Elias, "Networks of Gaussian Channels with Applications to Feedback Systems," IEEE Trans., Vol. IT-13, No. 3, pp. 493-501, July 1967.