COMMUNICATION SCIENCES

AND

ENGINEERING

X. PROCESSING AND TRANSMISSION OF INFORMATION*

Academic and Research Staff

Prof. P. Elias
Prof. R. G. Gallager
Prof. M. E. Hellman
Prof. E. V. Hoversten
Prof. R. S. Kennedy
Prof. C. E. Shannon
Dr. Jane W. S. Liu
Dr. R. J. McAulay

Graduate Students

P. Alexander
J. R. Clark
S. J. Dolinar, Jr.
R. A. Flower
M. A. Freedman
J. G. Himes
J. E. Just
B. K. Levitt
G. Q. McDowell
L. S. Metzger
P. F. Moulton
A. J. R. Muller
B. Neumann
R. S. Orr
T. A. Welch
H. P. H. Yuen

A. RELATIONSHIPS BETWEEN RANDOM TREES AND BRANCHING PROCESSES

Let z be a non-negative, integer-valued random variable and let z_i, $1 \leq i < \infty$, be independent random variables, all having the same distribution as z. At t = 0 one particle is alive, and at t = 1 it gives birth to a random number z_1 of offspring. At t = 2 each of the z_1 first-generation particles gives birth to a random number (z_2, z_3, ..., z_{z_1+1}) of offspring, etc. If N(k) is the number of particles in the k^{th} generation, then N(k), k = 0, 1, 2, ..., is called a branching process.[1] Letting $\bar{z}$ denote the expected value of z, it can be shown[1] that

$$\overline{N}(k) = \bar{z}^k. \tag{1}$$

Thus if $\bar{z} > 1$, $\overline{N}(k) \to \infty$; if $\bar{z} < 1$, $\overline{N}(k) \to 0$; while if $\bar{z} = 1$, $\overline{N}(k) = 1$ for all k. It is also known[1] that if $\bar{z} \leq 1$, the process dies out (i.e., N(k) = 0 for some k) with probability one; while if $\bar{z} > 1$ the probability of ultimate extinction p is strictly less than 1, and is the smallest non-negative solution to

$$p = P_z(p), \tag{2}$$

where $P_z(s) = E[s^z]$ is the probability-generating function of z.

Here, we examine a process which is a generalization of a branching process and find many similarities. In this random-tree process one particle starts at the origin at t = 0. At t = 1 it gives birth to a random number z_1 of offspring, each of which moves a random distance $y_1, y_2, \ldots, y_{z_1}$ from the origin. At t = 2 each of the first-generation particles gives birth to a random number of offspring, each of which moves a random

*This work was supported by the National Aeronautics and Space Administration (Grant NGL 22-009-013).

distance $y_{z_1+1}, y_{z_1+2}, \ldots, y_{z_1+N(2)}$ from its starting position, etc. The random variables y_i, $1 \leq i < \infty$, are independent of one another and of the z_i, and have a common distribution. We shall let y denote the generic random variable.

Thus far the number of particles $N(k)$ in the random tree is itself a branching process. We now define a region of activity $(-\beta, a)$, such that if a particle ever leaves the region it produces no new offspring. This type of behavior is found in a nuclear reactor or a sequential decoding search, and other applications undoubtedly exist.

Let us consider the case where y is an integer-valued random variable (generalization to arbitrary distributions is being undertaken). Then the population distribution at time $k = 0, 1, 2, \ldots$, is specified by $\{N_i(k)\}_{i=-\infty}^{\infty}$, where $N_i(k)$ is the number of particles in position i, at time k. Since only particles in the region $(-\beta, a)$ play dynamic roles, we may restrict attention to $\{N_i(k)\}$, $-\beta < i < a$. Let us renumber the integers in the region from 1 to n (assuming there are n integers). We show that

$$\overline{\underset{\sim}{N}}(k+1) = \overline{\underset{\sim}{N}}(k)[\bar{z}P], \tag{3}$$

where $\underset{\sim}{N}(k) = (N_1(k), N_2(k) \ldots N_n(k))$ is the population distribution (row) vector, and

$$P = [p_{ij}] = [\Pr\{\text{transition from } i \text{ to } j | i\}], \qquad 1 \leq i, j \leq n \tag{4}$$

is the "reduced" state transition matrix. We then show that if $\bar{z} > (1/\lambda)$, where λ is the maximal eigenvalue of P, then $\overline{N}(k) \to \infty$; while $\bar{z} < (1/\lambda)$ implies $\overline{N}(k) \to 0$; and $\bar{z} = (1/\lambda)$ implies that there exists a $B < \infty$, such that $(1/B) \leq \overline{N}(k) < B$. The similarity to a branching process is striking.

It is further shown that as $n \to \infty$, $\lambda \to \min_s G_y(s)$, where $G_y(s) = E[e^{sy}]$ is the moment generating function of y.

The probability of ultimate extinction is one if $\bar{z} \leq (1/\lambda)$ and is strictly less than one if $\bar{z} > (1/\lambda)$. Further, a relation similar to (2) exists. Let

$$p_i(k) = \Pr\{\text{random tree dies out by time } k | \text{ initial state} = i\}. \tag{5}$$

Then

$$p_i(k+1) = P_z\left[\sum_j p_{ij}p_j(k)\right] \equiv P_z^{(i)}[\underset{\sim}{p}(k)], \tag{6}$$

where $\underset{\sim}{p}(k) = (p_1(k), \ldots, p_n(k))$. Then we have

$$\underset{\sim}{p}(k+1) = \underset{\sim}{P}_z[\underset{\sim}{p}(k)], \tag{7}$$

where $\underset{\sim}{P}_z = \left(P_z^{(1)}, P_z^{(2)}, \ldots, P_z^{(n)}\right)$. Since

$$\underset{\sim}{p}(1) = (q_o, q_o, \ldots q_o), \tag{8}$$

where $q_o = \Pr\{z=0\}$,

$$\underset{\sim}{p} = \lim_{k\to\infty} p(k) \tag{9}$$

is easy to find numerically and must be a solution of

$$\underset{\sim}{p} = \underset{\sim}{P}_z(\underset{\sim}{p}). \tag{10}$$

One last similarity exists. As $\alpha, \beta \to \infty$ the probability of extinction (starting from the origin) tends to the probability of extinction of the associated branching process, provided $\bar{z} > \left[1 \Big/ \left\{\min_s G_y(s)\right\}\right]$.

Thus it is seen that the random-tree process is amenable to analysis and has many similarities to a branching process. R. G. Gallager is also investigating this problem, particularly as it applies to sequential decoding, and has obtained similar results.

M. E. Hellman

References

1. E. Parzen, Stochastic Processes (Holden-Day, San Francisco, 1962).

B. FINITE MEMORY DECISION SCHEMES*

Let $X_1, X_2, \ldots, X_n, \ldots$ be a sequence of independent, identically distributed random variables with probability density p. There are two hypotheses H_o and H_1 with a priori probabilities π_o and π_1, respectively. Under $H_o: p = p_o$, while under $H_1: p = p_1$. One would like to determine the true state of nature with minimum probability of error. In general one is interested in a sequence of optimal decision rules $d_1(X_1)$, $d_2(X_1X_2)$, $\ldots, d_n(X_1, X_2, \ldots X_n), \ldots$. Since each decision may depend upon all preceding observations, the amount of data to be stored increases without bound.

Motivated by a desire to reduce the data to be stored, Hellman and Cover[1] have investigated the hypothesis testing problem under the restriction that the data be summarized by an updatable statistic T, where T takes on a finite number of values; $T \in \{1, 2, \ldots, m\} = S$. Motivated by a desire to bring the hypothesis testing problem further into the "real" world, we have considered the problem with the additional stipulation that the number of observations N be finite.

The statistic at time n is a time-independent function of the statistic at time n-1 and the observation X_n.

*This work was supported in part by the National Science Foundation (Grant GK-5800).

$$T_n = f(T_{n-1}, X_n).$$

The independence of the observations and the action of the updating rule induce a Markov process on the memory space S. This process may be viewed as the operation of an m-state automaton acting on input X_i and yielding output $d(T_i) \in \{H_o, H_1\}$.

As a specific example, consider the symmetric coin-toss problem. Here, X_i takes on one of two possible values, heads (H) or tails (T). The hypotheses are

$$H_o: \quad \Pr\{H\} = p > \frac{1}{2}; \Pr\{T\} = 1 - p \equiv q \qquad \pi_o = \frac{1}{2}$$

$$H_1: \quad \Pr\{H\} = q; \Pr\{T\} = p \qquad \pi_1 = \frac{1}{2}$$

It is seen[1] that the infinite-time solution to the symmetric coin-toss problem embodies all the important aspects of optimal memory decision schemes. Thus, let us restrict attention to this problem in our search to learn about optimal, finite-time decision rules.

A greatest lower bound[1] to the infinite-time probability of error is

$$P_\infty^* = 1/[1+(p/q)^{m-1}].$$

No automaton can actually have a probability of error equal to P^*. There do exist sequences of ϵ-optimal automata. (A sequence of automata $A_1, A_2, \dots A_n, \dots$ is said to be ϵ-optimal if the associated sequence of probabilities of error approaches P_∞^* as a limit.)

An intuitively pleasing decision rule is the saturable counter.[1] This rule moves from state i to state i+1 if H is observed, and from state i+1 to i if T is observed ($1 \leq i \leq m-1$). Since H is more likely under H_o than under H_1, decide H_o in state i if $i > m/2$, and H_1 otherwise.

A slight modification greatly improves the saturable counter. If in state 1 an H is observed, move to state 2 only with small probability δ; if in state m a T is observed, move to state m - 1 with small probability $k\delta$. Such a machine is called a saturable counter with δ traps. From the symmetry of the problem k = 1 seems an optimal choice for any δ, and indeed it is. As shown,[1] if A_n is a saturable counter with $\delta = 1/n$, then $\{A_n\}$ is ϵ-optimal.

Unfortunately, for the infinite-time problem there exist other ϵ-optimal sequences which are structurally very different from the saturable counter. It was felt, however, that these were "less natural" solutions, and that the sequence of optimal finite-time solutions, $\left\{A_N^*\right\}$ would approach a structure similar to the saturable counter. In particular, one might expect that the optimal solution A_N^* for N observations would have the following properties:

0. Upward transitions occur only on H; downward transitions occur only on T.
1. The machine is symmetric.
2. Transitions only occur between adjacent states.
3. As $N \to \infty$, the probability of a transition from state 1 to state 2 (or from m to m-1) on H (or on T) tends monotonically to 0.
4. Randomization is needed not only for transitions from state 1 to 2, and from m to m - 1, but also for transitions between interior states. Moreover, it is felt that randomization will only be needed for transitions from a better to a worse state (i.e., from a state with a low probability of error to one with a higher probability of error).

To test these conjectures, one could do a computer search over the space of all possible automata and find A_N^*. However, there are $2m(m-1)$ parameters over which to optimize so that even for moderate values of m this is infeasible in practice.

Therefore we assumed that conjecture 0 is in fact true and tested the validity of conjectures 1, 2, 3, and 4 under this assumption, thereby halving the number of parameters. We found that conjectures 2, 3, and 4 are correct, but conjecture 1 is not true for m odd.

We believe that this anomaly is due to the fact that state $(m+1)/2$ provides no information either in favor of H_0 or H_1, so that a slightly asymmetric machine is better. However, it can be predicted that as $N \to \infty$ the asymmetry should vanish and, in fact, such behavior was observed.

For m = 3 or 4 the only randomization occurs in transitions from state 1 to 2, and from m to m - 1. Thus the A_N^* are saturable counters with δ-traps in the end states only. For m = 4, k = 1 the machine is symmetric. Although for m = 3, $k \neq 1$ it is true that $k \to 1$ as $N \to \infty$. Therefore assuming k = 1 does not affect the asymptotic behavior in this case.

For m > 4, randomization is needed in the interior as well, but much can be learned by considering the saturable counter with δ-traps in the end states only, and with k = 1. Let us consider such machines and optimize over δ.

First, let $P_\infty(\delta)$ be the infinite-time probability of error for a saturable counter with δ-traps in the end states only. For small δ it can be shown that

$$P_\infty(\delta) \approx P_\infty^* + k_1\delta.$$

It is shown[1] that $1/\delta$ has significance as a kind of time constant for the machine. Let us assume, therefore, that $P_N(\delta)$, the probability of error after N observations, is

$$P_N(\delta) \approx P_\infty^* + k_1\delta + k_2 \exp(-k_3\delta N).$$

Taking the partial with respect to δ and setting it equal to zero yields

$$\delta_N^* = \frac{\ln N}{k_3 N} + \frac{\ln (k_2k_3/k_1)}{k_3 N}$$

as the optimal value of δ, and

$$P_N^* = P_N\left(\delta_N^*\right)$$

$$= P_\infty^* + (k_1/k_3N)[1+\ln N + \ln (k_2k_3/k_1)].$$

Thus asymptotically δ_N^* and P_N^* both approach their limits $\left(0 \text{ and } P_\infty^*\right)$ as $\ln N/N$.

As has been mentioned, for $m > 4$ the type of machine assumed is not optimal. However, we felt that even then the results gave an indication of the true behavior. Experimental evidence (i.e., numerical calculations) bears this out.

R. A. Flower, M. E. Hellman

References

1. M. Hellman and T. Cover, "Learning with Finite Memory," Ann. of Math. Statist. 41, 765-782 (June 1970).

C. DETECTOR STATISTICS FOR OPTICAL COMMUNICATION THROUGH THE TURBULENT ATMOSPHERE

1. Introduction

It has been established, both theoretically[1] and experimentally,[2,3] that optical signals transmitted through the turbulent atmosphere are subjected to log-normal fading. Specifically, the spatial fading may be represented by a random process of the form

$$\exp \gamma(\vec{r}) = \exp[\chi(\vec{r}) + j\phi(\vec{r})] \tag{1}$$

multiplying the field that would otherwise be received in the absence of turbulence, where $\vec{r}$ is the two-dimensional spatial coordinate in the plane of the receiving aperture R, $\gamma(\vec{r})$ is a complex Gaussian random process whose real and imaginary parts, $\chi(\vec{r})$ and $\phi(\vec{r})$, are assumed to be statistically independent and stationary, and the variance of the phase term $\phi(\vec{r})$ is so large that we can assume it is uniformly distributed over $[0, 2\pi)$.[4]

In order to examine the performance of an optical communication link in the atmosphere, we must determine the fading statistics of the detector output in the receiver.

In a direct detection system, the fading affects the Poisson rate parameter μ of the photodetector in the receiving aperture, where μ is proportional to the total field intensity in the aperture.[5] Under the assumption that the received signal field would have no spatial variation over the aperture in the absence of fading, and ignoring the additive noise field, for an ideal photodetector,

$$\mu \propto \int_R d\vec{r}\ |e^{\gamma(\vec{r})}|^2 = \int_R d\vec{r}\ e^{2\chi(\vec{r})}. \tag{2}$$

In a heterodyne detection system, the receiver extracts a single spatial mode of the field incident on the receiving aperture.[6] The fading parameter of interest is the magnitude u of this spatial mode; under the same assumptions which led to Eq. 2, we can write

$$u^2 \propto \left|\int_R d\vec{r}\ e^{\gamma(\vec{r})}\ e^{-jk\vec{\theta}\cdot\vec{r}}\right|^2, \tag{3}$$

where the spatial mode has direction cosines $\vec{\theta}$.[7] We can also regard u^2 as the field intensity at some point in the focal plane of the receiving lens.

2. Probability Density Expansion

We will examine the statistical nature of μ and u^2 by using a mathematical technique described by Cramer,[8] which demonstrates whether an indeterminate probability density function $p_y(a)$ converges to an arbitrary probability density $p_x(a)$, provided only that lower order moments for both x and y can be determined. Notationally, we can expand $p_y(a)$ in terms of any $p_x(a)$ as follows:

$$p_y(a) = p_x(a)\left[1 + \sum_{i=1}^{\infty} r_i(a)\right], \tag{4}$$

where

$$r_i(a) \equiv \frac{D_i(a)}{D_iD_{i-1}} \sum_{k=0}^{i} L_{i,k}(\ell_k - m_k); \qquad i \geq 1,$$

$$\ell_k \equiv \overline{y^k}, \qquad m_k \equiv \overline{x^k},$$

$$D_i(a) \equiv \begin{vmatrix} m_o & m_1 & \cdots & m_i \\ m_1 & m_2 & \cdots & m_{i+1} \\ \vdots & \vdots & & \\ m_{i-1} & m_i & \cdots & m_{2i-1} \\ a^o & a^1 & \cdots & a^i \end{vmatrix} \qquad D_i \equiv \begin{vmatrix} m_o & m_1 & \cdots & m_i \\ m_1 & m_2 & \cdots & m_{i+1} \\ \vdots & \vdots & & \vdots \\ m_{i-1} & m_i & \cdots & m_{2i-1} \\ m_i & m_{i+1} & \cdots & m_{2i} \end{vmatrix} = L_{i+1,i+1}$$

$L_{i,k}$ is the cofactor of a^k in $D_i(a)$.

We can always select a $p_x(a)$ for which $m_1 = \ell_1$, and since $m_o = \ell_o = 1$, $r_2(a)$ is the first nonvanishing correction term in the expansion:

$$p_y(a) = p_x(a) \left\{ 1 + \underbrace{(\ell_2 - m_2) \frac{D_2(a)}{D_2}}_{r_2(a)} + \underbrace{\left[L_{3,2}(\ell_2 - m_2) \frac{D_3(a)}{D_3 D_2} + (\ell_3 - m_3) \frac{D_3(a)}{D_3} \right]}_{r_3(a)} + \dots \right\}. \tag{5}$$

Furthermore, if $p_x(a)$ is specified by two independent parameters of our choosing, we can arrange to have $m_1 = \ell_1$ and $m_2 = \ell_2$, so that $r_3(a)$ is the first nonvanishing correction term in the expansion:

$$p_y(a) = p_x(a) \left\{ 1 + \underbrace{(\ell_3 - m_3) \frac{D_3(a)}{D_3}}_{r_3(a)} + \underbrace{\left[L_{4,3}(\ell_4 - m_4) \frac{D_4(a)}{D_4 D_3} + (\ell_4 - m_4) \frac{D_4(a)}{D_4} \right]}_{r_4(a)} + \dots \right\}. \tag{6}$$

3. Direct Detection Fading Statistics

We cannot determine a sufficient number of moments from the integral expression for μ in Eq. 2 needed to use the expansion technique above. Therefore, assume that the integral of the log-normal random process $e^{2\chi(\vec{r})}$ can be approximated by a sum of n identically distributed, statistically independent, log-normal random variables $e^{2\chi_\ell}$, where n is the number of degrees of freedom of $e^{2\chi(\vec{r})}$ over the aperture R:

$$\mu \propto \sum_{\ell=1}^{n} e^{2\chi_\ell} \equiv I_n. \tag{7}$$

In some sense, n can be thought of as the number of coherence areas of $e^{2\chi(\vec{r})}$ contained in R.

For convenience, we shall normalize I_n, so that

$$y \equiv \frac{I_n - m_{I_n}}{\sigma_{I_n}} = \frac{I_n - n\, e^{2m_\chi + 2\sigma_\chi^2}}{n\, e^{2m_\chi + 2\sigma_\chi^2} \sqrt{\dfrac{e^{4\sigma_\chi^2} - 1}{n}}}, \tag{8}$$

where m_{I_n}, $\sigma_{I_n}^2$, m_χ, and σ_χ^2 are the means and variances of I_n and χ_ℓ. The first four moments of y are then given by

$$\ell_1 = 0, \qquad \ell_3 = \left(e^{4\sigma_\chi^2}+2\right)\sqrt{\frac{e^{4\sigma_\chi^2}-1}{n}},$$

$$\ell_2 = 1, \qquad \ell_4 = 3 + \frac{\left(e^{4\sigma_\chi^2}-1\right)\left(e^{12\sigma_\chi^2}+3\,e^{8\sigma_\chi^2}+6\,e^{4\sigma_\chi^2}+6\right)}{n}. \tag{9}$$

According to the Central Limit theorem, $p_y(\alpha)$ should converge to a normalized Gaussian probability density as n gets large. We shall therefore first expand $p_y(\alpha)$ in terms of

$$p_x(\alpha) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{\alpha^2}{2}}, \tag{10}$$

which has moments

$$\begin{array}{llll} m_1 = 0 & m_3 = 0 & m_5 = 0 & m_7 = 0 \\ m_2 = 1 & m_4 = 3 & m_6 = 15 & m_8 = 105. \end{array} \tag{11}$$

As one indication of how closely $p_y(\alpha)$ is approximated by $p_x(\alpha)$ in Eq. 10, a comparison of Eqs. 9 and 11 suggests that

$$\ell_k \xrightarrow[n\to\infty]{} m_k.$$

Combining Eqs. 6, 9, and 11, we find after some algebraic manipulation that

$$r_3(\alpha) = \frac{e^{4\sigma_\chi^2}+2}{6}\sqrt{\frac{e^{4\sigma_\chi^2}-1}{n}}\,(\alpha^3-3\alpha) \sim n^{-1/2} \tag{12}$$

$$r_4(\alpha) = \frac{\left(e^{4\sigma_\chi^2}-1\right)\left(e^{12\sigma_\chi^2}+3\,e^{8\sigma_\chi^2}+6\,e^{4\sigma_\chi^2}+6\right)}{24n}\,(\alpha^4-6\alpha^2+3) \sim n^{-1}. \tag{13}$$

The orders of $r_3(\alpha)$ and $r_4(\alpha)$ in n suggest that all of the correction terms in the expansion converge to zero so that $p_y(\alpha)$ converges to $p_x(\alpha)$ in Eq. 10 as n gets large.

However, Mitchell has concluded that I_n converges in distribution to a log-normal random variable more rapidly than to a Gaussian random variable.[9] To show this, he uses the expansion technique above modified to compare cumulative probability distributions instead of densities; he uses Eq. 4 in the form

$$P_y(\alpha) = P_x(\alpha) + \sum_{i=1}^{\infty} \underbrace{\int_{-\infty}^{\alpha} d\beta \; p_x(\beta) \; r_i(\beta)}_{R_i(\alpha)}. \tag{14}$$

Let us check his contention by expanding $p_y(\alpha)$ in terms of the probability density function of the normalized log-normal random variable

$$x = \frac{e^{\psi} - m_{I_n}}{\sigma_{I_n}}; \qquad \text{where } \psi \text{ is } N(\eta, \rho^2), \tag{15}$$

and η and ρ are chosen to set $m_1 = 0$, $m_2 = 1$:

$$e^{\eta} = n\, e^{2m_\chi + 2\sigma_\chi^2 - \frac{1}{2}\rho^2}$$

$$e^{\rho^2} = 1 + \epsilon; \qquad \epsilon \equiv \frac{e^{4\sigma_\chi^2} - 1}{n}. \tag{16}$$

Then we can show that

$$m_3 = \sqrt{\epsilon}\,(3+\epsilon) \xrightarrow[n\to\infty]{} 0$$

$$m_4 = 3 + 16\epsilon + 15\epsilon^2 + 6\epsilon^3 + \epsilon^4 \xrightarrow[n\to\infty]{} 3, \tag{17}$$

so we suspect that corresponding moments of x and y converge to the same values as n gets large, just as in the Gaussian expansion.

To solve for the correction terms in the expansion is an algebraically complicated task in general; however, if ϵ is small, we can show that to dominant order in n

$$r_3(\alpha) \approx \frac{\left(e^{4\sigma_\chi^2} - 1\right)^{3/2}}{6\sqrt{n}} (\alpha^3 - 3\alpha) \sim n^{-1/2}; \qquad n \gg e^{4\sigma_\chi^2} - 1 \tag{18}$$

$$r_4(\alpha) \approx \frac{\left(e^{4\sigma_\chi^2} - 1\right)^2 \left(e^{8\sigma_\chi^2} + 4\,e^{4\sigma_\chi^2} - 8\right)}{24n} (\alpha^4 - 6\alpha^2 + 3) \sim n^{-1}; \qquad n \gg e^{4\sigma_\chi^2} - 1. \tag{19}$$

Extrapolating the results of Eqs. 18 and 19, it appears that $p_y(\alpha)$ converges to the probability density function of x in Eq. 15. Thus, I_n converges to both a Gaussian and a log-normal random variable as n gets large. For large n, $r_3(\alpha)$ is the dominant correction term in both the Gaussian and log-normal expansions, and hence $|r_3(\alpha)|$ is a measure of the accuracy of approximating $p_y(\alpha)$ by $p_x(\alpha)$ in each case. But comparing Eqs. 12 and 18, it is evident that

$$\left|r_3(\alpha)\right|_{\text{log-normal}} \approx \left(\frac{e^{4\sigma_\chi^2}-1}{e^{4\sigma_\chi^2}+2}\right) \left|r_3(\alpha)\right|_{\text{Gaussian}} < \left|r_3(\alpha)\right|_{\text{Gaussian}}; \quad \forall\ \alpha, n \gg e^{4\sigma_\chi^2}-1. \tag{20}$$

This reinforces Mitchell's conclusion: <u>I_n does in fact converge more rapidly to a log-normal than to a Gaussian random variable over its entire sample space as n gets large, and this is especially true when σ_χ^2 is small.</u>

Since $r_3(\alpha)$ is the dominant correction term when convergence occurs, we can conclude that when $|r_3(\alpha)| \ll 1$, $p_y(\alpha)$ is accurately approximated by $p_x(\alpha)$. Note, however, that if $|r_3(\alpha)|$ is not small, it is nonetheless possible that $\left|\sum_{i=3}^{\infty} r_i(\alpha)\right|$ is still small, so that convergence is uncertain in this case. As an example, for $\sigma_\chi^2 = 10^{-1}$, in Fig. X-1 we have graphically illustrated the regions of convergence in (n, α) space for the Gaussian and log-normal approximations, by arbitrarily assuming convergence when $|r_3(\alpha)| < 10^{-1}$. This clearly gives us some quantitative information about the convergence of $p_y(\alpha)$ near the tail of the density function.

Contrast this with Mitchell's analysis: he assumes convergence of $P_y(\alpha)$ to $P_x(\alpha)$ when

$$|R_3(\alpha)| \equiv \left|\int_{-\infty}^{\alpha} d\beta\ p_x(\beta)\ r_3(\beta)\right| \ll 1; \tag{21}$$

however, this does not reveal very much about the behavior of $p_y(\alpha)$ near its tail. For example, the tail of $p_y(\alpha)$ may differ drastically from $p_x(\alpha)$, and this may result in a large correction term $r_3(\alpha)$ for large α. But if $p_x(\alpha)$ decays fast enough as α gets large, the product $p_x(\alpha)\ r_3(\alpha)$ may remain sufficiently small that $R_3(\alpha)$ is still small, even for very large α.

4. Heterodyne Detection Fading Statistics

As in the previous section, in order to calculate the moments needed in the expansion, we shall approximate the integral expression for u^2 in Eq. 3 by the finite sum

$$u^2 \propto \left|\sum_{\ell=1}^{n} e^{\chi_\ell + \overbrace{j(\phi_\ell - k\theta_\ell)}^{\mu_\ell}}\right|^2 \equiv J_n. \tag{22}$$

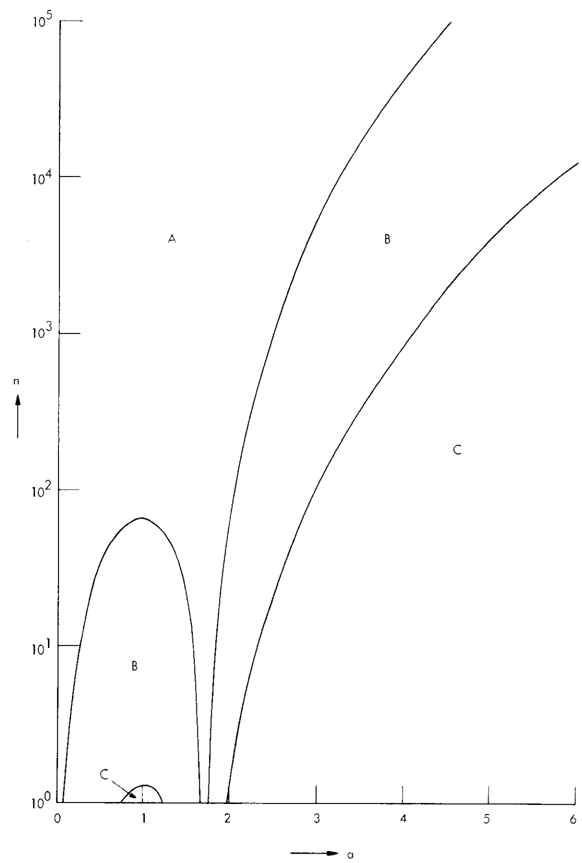

Fig. X-1. Regions of convergence of the probability density function for $I_n \equiv \sum_{\ell=1}^{n} e^{2\chi_\ell}$ to the log-normal and Gaussian densities, where each χ_ℓ is $N\left(m_\chi, \sigma_\chi^2\right)$, the χ_ℓ's are independent, and a is the distance in standard deviations σ_{I_n} from the mean m_{I_n} of I_n. For the specific case of $\sigma_\chi^2 = 10^{-1}$, A is the region where convergence occurs for both the log-normal and Gaussian approximations, B is the region of convergence for the log-normal approximation only, and convergence is uncertain in region C.

In Eq. 22, the χ_ℓ's are each $N\left(m_\chi, \sigma_\chi^2\right)$, the ϕ_ℓ's, and consequently the μ_ℓ's, are each uniformly distributed over $[0, 2\pi)$, and the χ_ℓ's and μ_ℓ's are statistically independent. Again we shall examine a normalized version of J_n:

$$y \equiv \frac{J_n - m_{J_n}}{\sigma_{J_n}} = \frac{J_n - n\, e^{2m_\chi + 2\sigma_\chi^2}}{n\, e^{2m_\chi + 2\sigma_\chi^2}\sqrt{1+\delta}}; \qquad \delta \equiv \frac{e^{4\sigma_\chi^2} - 2}{n}. \tag{23}$$

The first four moments of y are given by

$$\ell_1 = 0, \qquad \ell_2 = 0,$$

$$\ell_3 = \frac{2n^2 + 6n\left(e^{4\sigma_\chi^2} - 2\right) + \left(e^{12\sigma_\chi^2} - 9\, e^{4\sigma_\chi^2} + 12\right)}{n^2(1+\delta)^{3/2}} \xrightarrow[n\to\infty]{} 2,$$

$$\ell_4 = \frac{9n^3 + 42n^2\left(e^{4\sigma_\chi^2} - 2\right) + 6n\left(2e^{12\sigma_\chi^2} + 3e^{8\sigma_\chi^2} - 30e^{4\sigma_\chi^2} + 36\right) + \left(e^{24\sigma_\chi^2} - 16e^{12\sigma_\chi^2} - 18e^{8\sigma_\chi^2} + 144e^{4\sigma_\chi^2} - 144\right)}{n^3(1+\delta)^2} \xrightarrow[n\to\infty]{} 9. \tag{24}$$

The Central Limit theorem for complex random variables tells us that the probability density of J_n should converge to that of a central chi-square random variable z with two degrees of freedom:

$$p_z(a) = \frac{1}{\lambda}\, e^{-\frac{a}{\lambda}}\, u_{-1}(a); \qquad \overline{z^k} = \lambda^k k! \tag{25}$$

We shall therefore expand $p_y(a)$ in terms of $p_x(a)$, where

$$x \equiv \frac{z - n\, e^{2m_\chi + 2\sigma_\chi^2}}{n\, e^{2m_\chi + 2\sigma_\chi^2}\sqrt{1+\delta}}. \tag{26}$$

In order to have $m_1 = 0$, we must set $\lambda = n\, e^{2m_\chi + 2\sigma_\chi^2}$; then we can show that

$$m_2 = \frac{1}{1+\delta} \xrightarrow[n\to\infty]{} 1, \qquad m_3 = \frac{2}{(1+\delta)^{3/2}} \xrightarrow[n\to\infty]{} 2, \qquad m_4 = \frac{9}{(1+\delta)^2} \xrightarrow[n\to\infty]{} 9, \tag{27}$$

so, apparently corresponding moments of x and y converge to identical values as n gets large.

Evaluating the first two nonzero correction terms, we find that

$$r_2(a) = \frac{\delta}{4}\left[(1+\delta)a^2-2(1+\delta)^{1/2}a-1\right] \sim n^{-1}, \tag{28}$$

$$r_3(a) = \frac{\left(e^{12\sigma_\chi^2}-9e^{4\sigma_\chi^2}+12\right)}{36n^2}\left[(1+\delta)^{3/2}a^3-6(1+\delta)a+3(1+\delta)^{1/2}a+4\right] \sim n^{-2}, \tag{29}$$

from which we can conclude that J_n does behave like the chi-square random variable z for large n.

Finally, since we have shown that the sum of n real log-normal random variables tends to a log-normal random variable for large n, it is conceivable that the sum of n complex log-normal random variables looks log-normal for large n. Therefore, we should try to expand $p_y(a)$ in terms of $p_x(a)$, where χ is given in Eq. 15. To set $m_1 = 0$ and $m_2 = 1$, η and ρ must satisfy

$$e^\eta = n\,e^{2m_\chi+2\sigma_\chi^2-\frac{1}{2}\rho^2}, \qquad e^{\rho^2} = 2+\delta. \tag{30}$$

Then we can show that

$$m_3 = \frac{4+9\delta+6\delta^2+\delta^3}{(1+\delta)^{3/2}} \xrightarrow[n\to\infty]{} 4,$$

$$m_4 = \frac{41+150\,\delta+\dots}{(1+\delta)^2} \xrightarrow[n\to\infty]{} 41. \tag{31}$$

So, ℓ_k and m_k do not converge to the same values for $k \geq 3$ as n gets large. This is an indication that J_n does not behave like a log-normal random variable for large n. If the nonzero correction terms are evaluated, we find that for $i \geq 3$, $r_i(a) \sim n^0$, which corroborates our previous conclusion.

5. Conclusions

With appropriate reservations for the approximations of the integrals in Eqs. 2 and 3 by finite sums of independent random variables in Eqs. 7 and 22, we have used the Cramer expansion technique to demonstrate that in an atmospheric medium, if the collecting aperture of the receiver is large relative to a spatial coherence area of the atmospheric turbulence, and ignoring any background noise,

(i) the direct detection fading parameter μ is approximately log-normal, and

(ii) the heterodyne detection fading parameter u^2 is approximately chi-square, or equivalently, u is approximately Rayleigh.

B. K. Levitt

References

1. V. I. Tatarski, Wave Propagation in a Turbulent Medium (McGraw-Hill Book Company, New York, 1961).
2. D. L. Fried, G. E. Mevers, and M. P. Keister, Jr., "Measurements of Laser-Beam Scintillation in the Atmosphere," J. Opt. Soc. Am., Vol. 57, pp. 787-797, June 1962.
3. J. E. Ehrenberg, "A Study of the Effects of Atmospheric Turbulence on Intensity Properties at 6328 Å and 1.15 μ," S.M. Thesis, Department of Electrical Engineering, M.I.T., February 1968.
4. R. S. Kennedy and E. V. Hoversten, "On the Atmosphere as an Optical Communication Channel," IEEE Trans., Vol. IT-14, No. 5, pp. 717-719, September 1968.
5. S. Karp and R. S. Kennedy (eds.), NASA SP-217, "Optical Space Communication," p. 12, 1969.
6. Ibid., p. 15.
7. B. K. Levitt, "Variable-Rate Optical Communication through the Turbulent Atmosphere," Quarterly Progress Report No. 98, Research Laboratory of Electronics, M.I.T., pp. 88-91, July 15, 1970.
8. H. Cramer, Mathematical Methods of Statistics (Princeton University Press, New Jersey, 1946), pp. 131-133.
9. R. L. Mitchell, "Permanence of the Lognormal Distribution," J. Opt. Soc. Am., Vol. 58, pp. 1267-1272, September 1968.

D. DIRECT DETECTION ERROR PROBABILITIES FOR OPTICAL COMMUNICATION OVER A RAYLEIGH FADING CHANNEL

Under certain conditions, usually satisfied in practice, optical direct detection statistics in the presence of a temporally and spatially white background noise field can be accurately modelled by a Poisson process.[1] In this report we shall use this model to determine error probabilities for the case of binary PPM signalling over a Rayleigh fading channel.

If message m_1 is sent, the time-varying Poisson rate parameter during a single baud of duration T can be represented by

$$\mu(t)\big|_{m_1} = \begin{cases} u^2\mu_s + \mu_n; & 0 < t < \frac{T}{2} \\ \mu_n; & \frac{T}{2} < t < T \end{cases}$$

where μ_s and μ_n are signal and noise rate parameters, and the random variable u accounts for the fading. Denote the number of counts registered by the direct detector in the first- and second-half bauds by n_1 and n_2; under the Poisson model n_1 and n_2 are statistically independent with distributions

$$\Pr\left[n_1 \mid m_1, u\right] = \frac{1}{n_1!}\left[\frac{L}{2}\,(1+u^2\alpha)\right]^{n_1} e^{-\frac{L}{2}(1+u^2\alpha)}$$

$$\Pr\left[n_2 \mid m_1, u\right] = \frac{1}{n_2!}\left[\frac{L}{2}\right]^{n_2} e^{-\frac{L}{2}} \tag{1}$$

conditioned on the channel fading parameter u. For convenience, we have introduced the following parameters in Eq. 1:

$\alpha \equiv \mu_s/\mu_n$ = a priori signal-to-noise ratio,

$L \equiv \mu_n T$ = average number of noise counts per baud.

For equally-likely signals, the probability of a communication error on a single transmission, conditioned on the fading parameter u, can be shown to have the following forms[2]:

$$P(\epsilon \mid u) = \Pr\left[n_2 \geq n_1 \mid m_1, u\right] - \frac{1}{2}\Pr\left[n_2 = n_1 \mid m_1, u\right]$$

$$= e^{-\frac{L}{2}(2+u^2\alpha)}\left[\sum_{j=0}^{\infty}(1+u^2\alpha)^{-j/2}\, I_j\left(L\sqrt{1+u^2\alpha}\right) - \frac{1}{2}\, I_o\left(L\sqrt{1+u^2\alpha}\right)\right] \tag{2}$$

$$= Q_m\left[\sqrt{L},\ \sqrt{L(1+u^2\alpha)}\right] - \frac{1}{2}\, e^{-\frac{L}{2}(2+u^2\alpha)}\, I_o\left(L\sqrt{1+u^2\alpha}\right) \tag{3}$$

$$= \frac{1}{2}\left\{1 + Q_m\left[\sqrt{L},\ \sqrt{L(1+u^2\alpha)}\right] - Q_m\left[\sqrt{L(1+u^2\alpha)},\ \sqrt{L}\right]\right\} \tag{4}$$

$$= \frac{1}{2}\left(\frac{u^2\alpha}{2+u^2\alpha}\right)\int_{\frac{L}{2}(2+u^2\alpha)}^{\infty} dx\, e^{-x}\, I_o\left(\frac{2\sqrt{1+u^2\alpha}}{2+u^2\alpha}\, x\right) \tag{5}$$

where $I_j(\cdot)$ is the j^{th}-order modified Bessel function of the first kind, and

$$Q_m[a, b] \equiv \int_b^{\infty} dx\, x\, e^{-\left(\frac{a^2+x^2}{2}\right)}\, I_o(ax) \tag{6}$$

or, alternately,

$$Q_m[a,b] = e^{-\left(\frac{a^2+b^2}{2}\right)} \sum_{j=0}^{\infty} \left(\frac{a}{b}\right)^j I_j(ab) \tag{7}$$

are expressions for Marcum's Q function of radar theory[3,4] which has been tabulated.[5]

Professor E. V. Hoversten has suggested the following useful approximation for the conditional error probability above. In Eq. 5, the argument of the Bessel function satisfies the inequality

$$\frac{2\sqrt{1+u^2\alpha}}{2+u^2\alpha}\, x \geq L\sqrt{1+u^2\alpha}$$

over the range of integration. But the large argument asymptotic approximation for the 0^{th} order Bessel function is

$$I_o(z) \cong \frac{e^z}{\sqrt{2\pi z}}; \qquad z \gg 1. \tag{8}$$

Therefore, Eq. 5 can be approximated by

$$P(\epsilon|u) \cong \frac{u^2\alpha}{4\sqrt{\pi(2+u^2\alpha)}\,(1+u^2\alpha)^{1/4}} \int_{\frac{L}{2}(2+u^2\alpha)}^{\infty} dx \frac{1}{\sqrt{x}}\, e^{-\frac{\left(1-\sqrt{1+u^2\alpha}\right)^2}{2+u^2\alpha}x};$$

$$L\sqrt{1+u^2\alpha} \gg 1.$$

Making the substitution

$$y = \left(\sqrt{1+u^2\alpha} - 1\right)\sqrt{\frac{2x}{2+u^2\alpha}}$$

leads to the form

$$P(\epsilon|u) \cong \frac{u^2\alpha}{2\left(\sqrt{1+u^2\alpha}-1\right)(1+u^2\alpha)^{1/4}}\, Q\left[\sqrt{L}\left(\sqrt{1+u^2\alpha}-1\right)\right]; \quad L\sqrt{1+u^2\alpha} \gg 1, \tag{9}$$

where

$$Q(z) \equiv \frac{1}{\sqrt{2\pi}} \int_z^{\infty} dy\, e^{-\frac{y^2}{2}}$$

$$\cong \frac{1}{z\sqrt{2\pi}}\, e^{-\frac{z^2}{2}}; \qquad z \gtrsim 3. \tag{10}$$

is the Gaussian error function.[6] Using the approximation of Eq. 10 in Eq. 9, we have

$$P(\epsilon|u) \cong \frac{u^2 a}{\sqrt{8\pi L}\left(1-\sqrt{1+u^2 a}\right)^2 (1+u^2 a)^{1/4}} e^{-\frac{L}{2}\left(1-\sqrt{1+u^2 a}\right)^2};$$

$$L\sqrt{1+u^2 a} \gg 1, \quad \sqrt{L}\left(\sqrt{1+u^2 a}-1\right) \gtrsim 3. \tag{11}$$

This result is particularly satisfying since it agrees exponentially with the Chernoff bound for the conditional error probability:

$$P(\epsilon|u) \leq e^{-\frac{L}{2}\left(1-\sqrt{1+u^2 a}\right)^2} \tag{12}$$

Equation 11 can be further simplified if $u^2 a$ is large:

$$P(\epsilon|u) \cong \frac{e^{-\frac{Lau^2}{2}}}{a^{1/4}\sqrt{8\pi Lu}}; \qquad L\sqrt{1+u^2 a} \gg 1, \quad \sqrt{L}\left(\sqrt{1+u^2 a}-1\right) \gtrsim 3, \quad u^2 a \gg 1. \tag{13}$$

We will now examine $P(\epsilon)$, the probability of a communication error on a single transmission averaged over the channel fading, for the specific case where u is Rayleigh:

$$p(u) = \frac{2u}{\overline{u^2}} e^{-\frac{u^2}{\overline{u^2}}} u_{-1}(u), \tag{14}$$

where $\overline{u^2}$ is the second moment of u, and $u_{-1}(\cdot)$ is the unit step function. From Eqs. 3, 6, and 14, we can write

$$P(\epsilon) = \overline{P(\epsilon|u)}^{u}$$

$$= \underbrace{\frac{L}{2K} e^{\frac{L}{2K}} \int_0^\infty du\,(2au)\, e^{-\frac{L}{2K}(1+u^2 a)} \int_{\sqrt{L(1+u^2 a)}}^\infty dx\, x\, e^{-\left(\frac{L+x^2}{2}\right)} I_o(x\sqrt{L})}_{I}$$

$$- \underbrace{\frac{L}{2K} e^{\frac{L}{2K}} \int_0^\infty du\,(au)\, e^{-\frac{L}{2K}(1+u^2 a)}\, e^{-\frac{L}{2}(2+u^2 a)} I_o\left(L\sqrt{1+u^2 a}\right)}_{II} \tag{15}$$

where we have defined the parameter

$$K \equiv \frac{\overline{u^2}\mu_s T}{2} = \frac{\overline{u^2} L\alpha}{2} = \text{average number of signal counts per baud.}$$

To evaluate the first integral in Eq. 15, let

$$y = 1 + u^2\alpha$$

and then interchange the order of integration:

$$I = \frac{L}{2K} e^{\frac{L}{2K}} \int_1^\infty dy\, e^{-\frac{L}{2K}y} \int_{\sqrt{Ly}}^\infty dx\, x\, e^{-\left(\frac{L+x^2}{2}\right)} I_o(x\sqrt{L})$$

$$= e^{\frac{L}{2K}} \int_{\sqrt{L}}^\infty dx\, x\, e^{-\left(\frac{L+x^2}{2}\right)} I_o(x\sqrt{L}) \underbrace{\int_1^{x^2/L} dy\, \frac{L}{K} e^{-\frac{L}{2K}y}}_{e^{-\frac{L}{2K}} - e^{-\frac{x^2}{2K}}}$$

$$= Q_m[\sqrt{L}, \sqrt{L}] - e^{-\frac{L(K-1)}{2K}} \int_{\sqrt{L}}^\infty dx\, x\, e^{-\frac{x^2(K+1)}{2K}} I_o(x\sqrt{L}).$$

By making the substitution

$$z = x\sqrt{\frac{K+1}{K}}$$

in the integral above, we find that

$$I = Q_m[\sqrt{L}, \sqrt{L}] - \left(\frac{K}{K+1}\right) e^{\frac{L}{2K(K+1)}} Q_m\left[\sqrt{\frac{LK}{K+1}}, \sqrt{\frac{L(K+1)}{K}}\right].$$

Finally, we can evaluate the second integral in Eq. 15 by introducing the change of variable

$$x = \sqrt{\frac{L(K+1)(1+u^2\alpha)}{K}};$$

we then find that

$$II = \frac{1}{2(K+1)} e^{\frac{L}{2K(K+1)}} Q_m\left[\sqrt{\frac{LK}{K+1}}, \sqrt{\frac{L(K+1)}{K}}\right].$$

Therefore, Eq. 15 becomes

$$P(\epsilon) = Q_m[\sqrt{L}, \sqrt{L}] - \left(\frac{K+\frac{1}{2}}{K+1}\right) e^{\frac{L}{2K(K+1)}} Q_m\left[\sqrt{\frac{LK}{K+1}}, \sqrt{\frac{L(K+1)}{K}}\right]. \tag{16}$$

Using Eq. 7 in Eq. 16, we can write

$$P(\epsilon) = e^{-L} \sum_{j=0}^{\infty} \left[1-\left(\frac{K+\frac{1}{2}}{K+1}\right)\left(\frac{K}{K+1}\right)^{2j}\right] I_j(L). \tag{17}$$

But

$$\left(\frac{K}{K+1}\right)^{2j} \leqslant 1, \qquad \forall\ K, j \geqslant 0,$$

so that Eq. 17 leads to the lower bound

$$P(\epsilon) \geqslant \frac{e^{-L}}{2(K+1)} \sum_{j=0}^{\infty} I_j(L) = \frac{1}{2(K+1)} Q_m[\sqrt{L}, \sqrt{L}]. \tag{18}$$

Since tables of the Marcum Q function are not always readily available, it is convenient to use a property of this function[7] to rewrite the lower bound of Eq. 18 in the form

$$P(\epsilon) \geqslant \frac{1}{4(K+1)} \left[1+e^{-L} I_o(L)\right]. \tag{19}$$

We can also upper bound $P(\epsilon)$ by averaging the Chernoff bound for $P(\epsilon|u)$ in Eq. 12 over u. For Rayleigh u, we have, in our previous notation

$$P(\epsilon) \leqslant \overline{e^{-\frac{L}{2}\left(1-\sqrt{1+u^2 a}\right)^2}}^{\,u}$$

$$= \frac{L}{K} e^{\frac{L}{2K}} \int_0^{\infty} du\ (au)\ e^{-\frac{L(1+u^2 a)}{2K}}\ e^{-\frac{L}{2}\left(1-\sqrt{1+u^2 a}\right)^2}.$$

Making the substitution

$$x = \sqrt{1+u^2 a} - \frac{K}{K+1}$$

in the integral above, we find that

$$P(\epsilon) \leq \frac{L}{K} e^{\frac{L}{2K(K+1)}} \int_{1/K+1}^{\infty} dx\, x\, e^{-\frac{L(K+1)}{2K} x^2} + \frac{L}{K+1} e^{\frac{L}{2K(K+1)}} \int_{1/K+1}^{\infty} dx\, e^{-\frac{L(K+1)}{2K} x^2}$$

$$= \frac{1}{K+1} \left[1 + \sqrt{\frac{2\pi LK}{K+1}}\, e^{\frac{L}{2K(K+1)}}\, Q\left(\sqrt{\frac{L}{K(K+1)}}\right)\right]. \tag{20}$$

We also have the following useful approximation to the upper bound above:

$$P(\epsilon) \lesssim \frac{1}{K}\left[1 + \sqrt{\frac{\pi L}{2}}\right]; \qquad \frac{L}{K(K+1)} \ll 1, \quad K \gg 1. \tag{21}$$

This is interesting because Eqs. 19 and 21 suggest that $P(\epsilon)$ is inversely proportional to the parameter K under certain conditions, which is further corroborated below.

As a final exercise, we can use the approximation for Marcum's Q function derived in Appendix A to calculate a simple approximation for $P(\epsilon)$. Thus, using Eq. A. 1 together with Eq. 16, we find that

$$P(\epsilon) \approx \frac{1}{2} + \frac{1}{4(K+1)\sqrt{2\pi L}} - \sqrt{\frac{K+1}{K}} \left(\frac{K+\frac{1}{2}}{K}\right)^2 e^{\frac{L}{2K(K+1)}}\, Q\left[\sqrt{\frac{L}{K(K+1)}}\right]; \qquad \sqrt{L} \gg 1. \tag{22}$$

Now, if $K \gg 1$ and $\sqrt{\frac{K^2}{L}} \gg 1$ as well, using Eq. A. 2 we can show that Eq. 22 reduces to

$$P(\epsilon) \approx \frac{1}{K}\sqrt{\frac{L}{2\pi}}; \qquad \sqrt{L} \gg 1, \quad K \gg 1, \quad \frac{K^2}{L} \gg 1. \tag{23}$$

If the average number of signal counts equals or exceeds the average number of noise counts, that is $K \geq L$, then the restriction on Eq. 23 can be rewritten in the form

$$P(\epsilon) \approx \frac{1}{K}\sqrt{\frac{L}{2\pi}}; \qquad P(\epsilon) \ll \frac{L}{K\sqrt{2\pi}}, \quad K \geq L, \tag{24}$$

which is a useful approximation for small $P(\epsilon)$.

Appendix A

We will now derive a useful approximation for Marcum's Q function. Combining Eqs. 6 and 8, we have

$$Q_m(a, b) \approx \frac{1}{\sqrt{2\pi a}} \int_{b-a}^{\infty} dx\, \sqrt{x+a}\, e^{-\frac{x^2}{2}}; \qquad ab \gg 1.$$

Using the first two terms in the Taylor series expansion for $\sqrt{x+a}$ about x = b−a, we find that

$$Q_m(a,b) \approx \frac{1}{\sqrt{2\pi a}} \int_{b-a}^{\infty} dx \sqrt{b} \left\{1 + \frac{1}{2b}[x-(b-a)]\right\} e^{-\frac{x^2}{2}}$$

$$= \frac{1}{2}\left\{\sqrt{\frac{b}{a}}\left(1+\frac{a}{b}\right) Q(b-a) + \frac{1}{\sqrt{2\pi ab}} e^{-\frac{(b-a)^2}{2}}\right\};$$

provided $ab \gg 1$, $b \gg 1$, $b \geq a$. (A. 1)

Also, note that the Taylor series expansion for the Gaussian error function Q(z) about z = 0 is

$$Q(z) = \frac{1}{2} - \frac{z}{\sqrt{2\pi}} + \dots . \tag{A. 2}$$

B. K. Levitt

References

1. J. R. Clark and E. V. Hoversten, "Poisson Process as a Statistical Model for Photodetectors Excited by Gaussian Light," Quarterly Progress Report No. 98, Research Laboratory of Electronics, M.I.T., July 15, 1970, pp. 95-101.
2. W. K. Pratt, Laser Communication Systems (John Wiley and Sons, Inc., New York, 1969), pp. 207-208.
3. J. I. Marcum, "A Statistical Theory of Target Detection by Pulsed Radar," IRE Trans. on Information Theory, Vol. IT-6, pp. 159-160, April 1960.
4. M. Schwartz, W. R. Bennett, and S. Stein, Communication Systems and Techniques (McGraw-Hill Book Company, New York, 1966), pp. 585-589.
5. J. I. Marcum, "Table of Q Functions," Rand Corp. Research Memo RM 339, January 1, 1950.
6. J. M. Wozencraft and I. M. Jacobs, Principles of Communication Engineering (John Wiley and Sons, Inc., New York, 1965), pp. 82-83.
7. M. Schwartz, W. R. Bennett, and S. Stein, op. cit., p. 586.

E. WEAKLY COHERENT QUANTUM CHANNELS I

We derive a bound to the error probability of a weakly coherent quantum channel used with a photon counting receiver and orthogonal, equal energy, equi-probable, signals. The nature of these channels has been discussed elsewhere[1,2] and the bound has been presented previously.[2-5] We here limit our attention to the derivation of the bound. For simplicity we consider a single polarization component of the field at the aperture.

Consider the system for which the receiver aperture field contains a zero mean Gaussian noise component that is white in space and time and a random signal component. The noise power density, N, of the noise is assumed to be much less than hf_o. The signal component corresponds to the transmission of one of M messages. The components resulting from the transmission of different messages are orthogonal to each other in that the sample functions of the random signal fields corresponding to different messages are orthogonal.

Analytically, the above assumptions may be restated as follows. When the j^{th} message is transmitted, the complex envelope $U(\vec{r}, t)$ of the receiver's aperture field can be expressed as

$$U(\vec{r}, t) = \sum_i S_{ij}\phi_i(\vec{r}, t) + N(\vec{r}, t)$$

where $N(\vec{r}, t)$ is the white Gaussian noise component of the field and the $\phi_i(\vec{r}, t)$ are a set of orthonormal functions over the receiving aperture. The S_{ij} $i = 1, 2, \ldots$ are random variables that characterize the signal component of the field when message j is transmitted. The assumption that the signal components of the field are orthogonal for different messages is equivalent to the statement that, for all i,

$$S_{ij} S_{ik} = 0 \qquad \text{for all } j \neq k; \quad j = 1, \ldots M, \quad k = 1, \ldots M. \tag{1}$$

The random variables S_{ij} can be viewed as the modal field amplitudes associated with the signal component of the received field. For weakly coherent quantum channels they behave as though they were statistically independent zero mean complex Gaussian random variables with variances that are small relative to the energy of a carrier photon. The following analysis is directed toward such channels with the additional assumption that the noise power density N is much less than hf_o.

Suppose that the receiver decomposes the received "signal plus noise" field into M orthogonal components, indexed 1, 2, ... M, such that each component contains all of the energy associated with one, and only one, of the signals. Suppose further that the average noise energy in each of these orthogonal components are equal. The energy, or photon count, in each such component is then measured and the i^{th} message is assumed to have been transmitted if the energy in the i^{th} component exceeds the energies in the remaining M-1 components. If none of the energies exceed all of the other ones, the choice among those with the maximum energy is made by choosing each with equal probability.

The error probability of the receiver just described will satisfy the inequality

$$Pe \leq \frac{1}{M} \sum_j P[n_j \leq n_i \text{ one or more } i \neq j \mid \text{message } j]. \tag{2}$$

Here the n_i, i = 1, ... M, are the photon counts in the M components into which the receiver divides the received field. Our objective is to bound the right member of this equation by a simpler (but weaker) expression. As a first step in that direction we employ some known results[6] to obtain

$$Pe \leq \sum_{n_j} P[n_j \mid \text{message } j] \{\sum_{j} P[n_i \geq n_j \mid \text{message } j;\, n_j]\}^{\rho}. \tag{3}$$

In this equation ρ is a free parameter whose value is constrained to be in the interval (0, 1). The validity of the equation rests upon the assumption that the receiver energy measurements are on orthogonal components of the received field and hence are conditionally independent.

To proceed, we invoke the assumption that N and the variances of the S_{ij} are small relative to hf_o. This implies that the n_i and n_j are approximately Poisson distributed.[7,8] We further invoke the assumptions that the received energies associated with the different messages are equal and that the average noise energies in each of the M receiver field components are equal. This implies that the photon counts in each of the components that do not contain signal energy are identically distributed and that the probability distribution of the count in the component that does contain signal energy is independent of which message was transmitted. Specifically

$$P[n_i \mid \text{message } j] = \frac{\lambda_o^{n_i}}{n_i!} \exp -\lambda_o \qquad i \neq j,\; n_i = 0, 1, \ldots \tag{4a}$$

$$P[n_j \mid \text{message } j] = \frac{(\lambda_o+\lambda_1)^{n_j}}{n_j!} \exp -(\lambda_o+\lambda_1) \qquad n_j = 0, 1, \ldots . \tag{4b}$$

Here λ_o is the average number of detected noise photons in each one of the M receiver components and λ_1 is the average number of signal photons detected in that component that contains signal energy.

Upon introducing Eq. 4 into Eq. 3 and using standard Chernov bounding techniques[6] one obtains

$$Pe \leq M^{\rho} \exp -\{\rho\lambda_o(1-e^{t}) - (\lambda_o+\lambda_1)(1-e^{-t\rho})\}. \tag{5}$$

Minimizing the bound with respect to the free parameters t and ρ yields

$$Pe \leq \exp -K\epsilon_d(\mu, \beta) \tag{6a}$$

where

$$K = \log_2 M \tag{6b}$$

$$\mu = \lambda_1/\lambda_o \tag{6c}$$

$$\beta = \lambda_1/K \tag{6d}$$

and

$$\epsilon_d(\mu,\beta) = \max_{0\leq\rho\leq 1} \left\{\frac{(1+\rho)\beta}{\mu} + \beta - \rho \ln 2 - \frac{\beta(1+\rho)}{\mu}(1+\mu)^{1/(1+\rho)}\right\}. \tag{6e}$$

The exponent $\epsilon_d(\mu,\beta)$ can also be expressed parametrically as

$$\text{if} \quad \beta \geq \beta_{crit} = \frac{\mu \ln 2}{1-\sqrt{1+\mu}\,[1-\ln(1+\mu)]} \tag{7a}$$

$$\epsilon_d(\mu,\beta) = \frac{\beta}{\mu}\left[\sqrt{1+\mu}-1\right]^2 - \ln 2 \tag{7b}$$

$$\text{if} \quad \beta_{crit} \geq \beta \geq \beta_{cap} = \frac{\mu \ln 2}{-\mu + (1+\mu)\ln(1+\mu)} \tag{7c}$$

$$\epsilon_d(\mu,\beta) = \frac{\beta}{\mu}\left[1+\mu+\rho-(1+\rho)(1+\mu)^{1/(1+\rho)}\right] - \rho \ln 2 \tag{7d}$$

where ρ is determined from the relationship

$$\beta = \mu\left\{1-(1+\mu)^{1/(1+\rho)}\left[1-\frac{\ln(1+\mu)}{1+\rho}\right]\right\}^{-1} \ln 2. \tag{7e}$$

Finally,

$$\text{if} \quad \beta \leq \beta_{cap}, \quad \epsilon_d(\mu,\beta) = 0. \tag{7f}$$

The bound of Eqs. 6 and 7 has been presented previously.[2-5] We note in passing that the derivation depends upon the statistics of the incident field only through the assumption that the n_i are Poisson distributed. Thus the bound applies to the detection of known signals as well as to weakly coherent random signals – provided that the contribution of the background noise to the count is itself Poisson distributed.

R. S. Kennedy

References

1. R. S. Kennedy, "Weakly Coherent Photon Channels," 1970 International Symposium on Information Theory, June 15-19, 1970, Noordwijk, The Netherlands.

2. R. S. Kennedy, "Communication through Optical Scattering Channels," to appear in Proc. IEEE (Special Issue on Optical Communication).

3. R. S. Kennedy, Internal Memorandum No. 62L-0106 (Lincoln Laboratory, M. I. T.), "Performance of Optical Scatter Communication Systems," January 29, 1968.

4. R. S. Kennedy, "On the Performance of Direct Detection Optical Communication Systems Using Orthogonal Alphabets," IEEE International Conference on Communications, Boulder, Colorado, June 9-11, 1969.

5. E. V. Hoversten and R. S. Kennedy, "Efficient Optical Communication within the Earth's Atmosphere," in Opto-Electronics Signal Processing Techniques, AGARD Conference Proceeding No. 50, February 1970.

6. R. G. Gallager, Information Theory and Reliable Communication (John Wiley and Sons, New York, 1968).

7. L. Mandel, "Fluctuations of Light Beams," in Progress in Optics, Vol. II, Ed. E. Wolf (North-Holland Publishing Co., Amsterdam, 1963), pp. 181-248.

8. S. Karp, E. O'Neill, and R. Gagliardi, "Communication Theory for the Free Space Channel," to appear in Proc. IEEE (Special Issue on Optical Communication).