XXVIII. PROCESSING AND TRANSMISSION OF INFORMATION*

Academic and Research Staff

Prof. P. Elias
Prof. R. G. Gallager
Prof. F. C. Hennie III
Prof. E. V. Hoversten
Prof. R. S. Kennedy
Prof. C. E. Shannon
Prof. R. N. Spann

Graduate Students

E. A. Bucher
D. L. Cohn
R. L. Greenspan
S. J. Halme
H. M. Heggestad
M. Khanna
Jane W-S. Liu
J. Max
J. C. Moldon
M. A. Sirbu
M. A. Tamny
W. C. Wilder
D. A. Wright
J. S. Zaborowski, Jr.

RESEARCH OBJECTIVES AND SUMMARY OF RESEARCH

The major goals of this research are to generate a deep understanding of communication channels and sources and to use this understanding in the development of reliable, efficient communication techniques.

1. Optical Communication

The fundamental limitations and efficient utilization of optical channels are the general concern of these investigations. Our interests now include the turbulent atmospheric channel, the cloud transmission channel, quantum-limited channels, and scatter channels; the investigations range from fundamental coding theorems through feasible near-optimum communication systems, to high-resolution astronomy.

During the past year we have shifted our emphasis from the general development of the turbulent atmospheric model to an investigation of its communication implications. Our principal conclusions, thus far, have been that the presence of turbulence does not reduce the channel capacity that would exist in its absence and, to be efficient, a receiver must exploit the spatial diversity that is contained within its aperture.[1] The investigation of such receivers continues (see Sec. XXVIII-B). Also, evaluation of some specific signaling schemes has been undertaken (polarization modulation and the transmitted reference system).

It has been apparent that, to be simple, a near-optimum receiver for the turbulent atmosphere must cleverly exploit the structure of the field phase front across the collecting aperture. Accordingly, a program to determine the characteristics of this structure has been initiated. This program will involve heterodyning phase experiments carried out in cooperation with the Electronics Research Center, NASA, Cambridge, and wavefront interference experiments, carried out in cooperation with the Smithsonian and Harvard College Observatories.

In other areas, three doctoral level investigations have developed beyond the proposal stage. One of these is concerned with the application of estimation theory to the problem of high-resolution astronomy (or surveillance) through the turbulent atmosphere. Preliminary results suggest that significant gains can be realized through the data-processing techniques suggested by estimation theory.[2] The second investigation pertains to the fundamental limitations upon the transmission of information by combined

*This work was supported principally by the National Aeronautics and Space Administration (Grant NsG-334), and in part by the Joint Services Electronics Program (U. S. Army, U. S. Navy, and U. S. Air Force) under Contract DA 28-043-AMC-02536(E).

temporal and spatial modulation, e.g., by a sequence of "images." Of particular concern is the interplay between time, bandwidth, aperture size, and background noise.[3] The third investigation is directed toward a fundamental examination of the role of quantum theory in communication theory.[4] The central issue in this investigation is the determination of the limitations imposed upon reliable communication by quantum effects and a determination of the receivers and waveforms which attain these limits. These three investigations will be completed during the coming year.

Two other lines of endeavor have been initiated. One of these, which is being attacked at both doctoral and master's levels, pertains to the limitations that clouds impose upon the reliability of optical communication. The present objective is to determine an appropriate statistical model for transmission through a cloud. The other endeavor is the establishment of a cw scatter link that will be used to investigate the feasibility of all-weather scatter communication.

R. S. Kennedy, E. V. Hoversten

References

1. R. S. Kennedy and E. V. Hoversten, "On the Reliability of the Atmosphere as an Optical Communication Channel," a paper presented at the IEEE International Symposium on Information Theory, San Remo, Italy, September 1967.

2. J. C. Moldon, "Optical Image Estimation in a Turbulent Atmosphere," Quarterly Progress Report No. 86, Research Laboratory of Electronics, M.I.T., July 15, 1967, pp. 221-229.

3. R. L. Greenspan, "Information Transmission via Spatial Modulation of Waveforms," Quarterly Progress Report No. 86, Research Laboratory of Electronics, M.I.T., July 15, 1967, pp. 247-255.

4. Jane W-S. Liu, "Quantum Communication Systems," Quarterly Progress Report No. 86, Research Laboratory of Electronics, M.I.T., July 15, 1967, pp. 230-239.

2. Channels and Coding

Upper and lower bounds have been found on the error probability that can be achieved with a systematic convolutional code.[1] If m is the number of check digits per information digit in the code, and N is the constraint length of the code in information symbols, then these bounds have the form $P_e \leq A_u \exp - (Nm+1)E_u$ and $P_e \geq A_L \exp - (Nm+1)E_L$. In these expressions A_u is independent of N, and A_L is slowly varying with N. The exponents E_u and E_L are equal if the capacity of the channel is close to the transmission rate. These results generalize earlier results for nonsystematic codes by Yudkin[2] and Viterbi,[3] in which Nm + 1 should be replaced by N(m+1). The reason for being interested in these new results is that systematic convolutional codes are far less sensitive to error propagation than nonsystematic codes. Work continues on the behavior of systematic convolutional codes.

Professor Kennedy has recently completed a monograph on Fading Dispersive Channels.[4] A mathematical model is developed for such channels and is shown to be equivalent to a diversity model. Bounds on minimum achievable error probability are found as a function of transmission rate and constraint time. It is shown that the exponential decay of error probability with increasing constraint time is much slower than for a nonfading additive Gaussian noise channel with the same signal-to-noise ratio, although the capacities are the same. This work was extended by Richters[5] in a recently completed Ph.D. thesis to the case of bandlimited signals. He found that for a fixed bandwidth, the infinite bandwidth exponent could be approached closely at small transmission rates, but the loss in exponent rapidly increased with increasing transmission rate.

A Ph.D. thesis has recently been completed by D. Chase[6] on the topic of unsynchronized noisy channels. He has shown that coding can be used to simultaneously correct transmission errors and to acquire synchronization. For transmission rates close to capacity, the lack of synchronization does not change the exponential dependence of error probability on code constraint length. For low transmission rates, the results are more complicated, and depend on the symmetry of the channel and whether the code can be changed from one block to the next. Some minimum distance bounds on unsynchronized code words are established which generalize earlier work on comma-free codes. Work continues on the topic of synchronization.

R. G. Gallager, R. S. Kennedy

References

1. E. Bucher, "Bounds on the Probability of Error for Systematic Convolutional Codes" (to be submitted to IEEE Trans. on Information Theory for publication).
2. H. Yudkin, "Channel State Testing in Information Decoding," Ph.D. Thesis, Department of Electrical Engineering, M.I.T., September 1964.
3. A. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm," IEEE Trans., Vol. IT-13, pp. 260-270, April 1967.
4. R. S. Kennedy, Dispersive Channels (John Wiley and Sons, Inc., New York, 1968).
5. J. S. Richters, "Communication over Fading Dispersive Channels," Technical Report 464, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Mass., November 10, 1967.
6. D. Chase, "Communication over Noisy Channels with no a priori Synchronization Information," Ph.D. Thesis, Department of Electrical Engineering, M.I.T., June 1967; also to appear as Technical Report 463 of the Research Laboratory of Electronics, M.I.T.

3. Source Coding

A Ph.D. thesis has been completed by J. T. Pinkston,[1] clarifying the relationship of fixed-length to variable-length codes, subject to a distortion measure, particularly when the distortion is peak-limited, as well as average-limited. He has also shown that simple quantizers are strictly bounded away from the minimum distortion achievable with source codes.

Professor Gallager has extended Shannon's coding theorem for sources subject to a distortion measure to the case of arbitrary discrete ergodic sources with a broad class of distortion measures. According to this theorem,[2] any given source and distortion measure has a function $R(d^*)$ associated with it. If one transmits the source output, after appropriate coding, over a noisy channel, one can achieve an average distortion per source letter of d^* if and only if the capacity of the channel exceeds $R(d^*)$.

R. G. Gallager

References

1. J. T. Pinkston III, "Encoding Independent Sample Information Sources," Ph.D. Thesis, Department of Electrical Engineering, M.I.T., September 1967; also to appear as Technical Report 462 of the Research Laboratory of Electronics, M.I.T.
2. R. G. Gallager, Information Theory and Reliable Communication (John Wiley and Sons, Inc., New York, in press), see Chap. 9.

A. LOWER BOUND TO THE ERROR PROBABILITY FOR THE ATMOSPHERIC OPTICAL CHANNEL

An appropriate channel model for signaling through the turbulent atmosphere at optical frequencies is shown in Fig. XXVIII-1. This is a scalar channel model and thus it is assumed that neither polarization modulation nor spatial modulation is employed. The model is in terms of complex waveforms with the carrier frequency suppressed. Thus $x(t)$ is the complex envelope of the input signal as a function of time t and $y(t, \vec{r})$ is the complex envelope of the channel output as a function of time, t, and of $\vec{r}$, the position in the receiving aperture.

The complex process $n(t, \vec{r})$ represents the envelope of the relevant polarization component of the background light. Any front-end receiver noises are also included in $n(t, \vec{r})$. This noise is assumed to be a zero-mean complex Gaussian random process with independent components that are stationary in time and space. Further, it is assumed that $E[n(t, \vec{r}_1)n(\tau, \vec{r}_2)] = 0$ for all t, τ, $\vec{r}_1$, and $\vec{r}_2$ and $E[n(t, \vec{r}_1)n^*(\tau, \vec{r}_2)]$ is zero for $|\vec{r}_1 - \vec{r}_2|$ greater than a few wavelengths. (The background light does have a spatial correlation function satisfying this assumption.) Finally, $E[n(t, \vec{r})n^*(\tau, \vec{r}) = R_n(t-\tau)$ for all $\vec{r}$ and the Fourier transform of $R_n(t-\tau)$ can be assumed to be constant at the value $2N_o$ W/cps over any frequency range of interest. It is also assumed that any front-end receiver noises, such as the shot noise caused by heterodyne detection, can be modeled with a correlation function that is extremely narrow in the $r = |\vec{r}_1 - \vec{r}_2|$ variable. All that is needed now is that the spectrum of the integral of $n(t, \vec{r})$ over some area depend linearly on the area. This is reasonable for many front-end noises and consistent with the correlation-function assumption, as long as the area is large relative to the radiation wavelength.

The multiplicative process, $z(t, \vec{r})$, in Fig. XXVIII-1 represents the effects of the temporal and spatial fading caused by the turbulence; that is, the effects of the random refractive index variations in the atmosphere. The random process, $z(t, \vec{r})$, has the form

$$z(t, \vec{r}) = \exp \gamma(t, \vec{r}), \tag{1}$$

where $\gamma(t, \vec{r})$ is a complex Gaussian random process. For simplicity, the real and imaginary parts of $\gamma(t, \vec{r})$ are assumed to be statistically independent of each other and

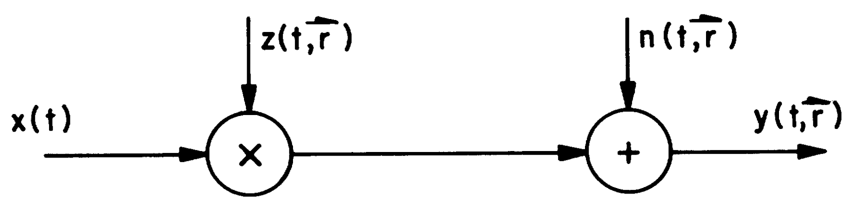

Fig. XXVIII-1. Model of the turbulent optical channel.

stationary in time.

The channel model of Fig. XXVIII-1 can be reduced with some approximation to that shown in Fig. XXVIII-2. The approximation involves the assumption that $\gamma(t, \bar{r})$ is completely correlated over those time intervals and spatial areas wherein it is correlated at all, and is completely uncorrelated from one such interval and area to another. The correlation time is assumed to be so large that the total decision interval falls within one coherence time. The lack of correlation from one decision interval to the next can be achieved by scrambling.

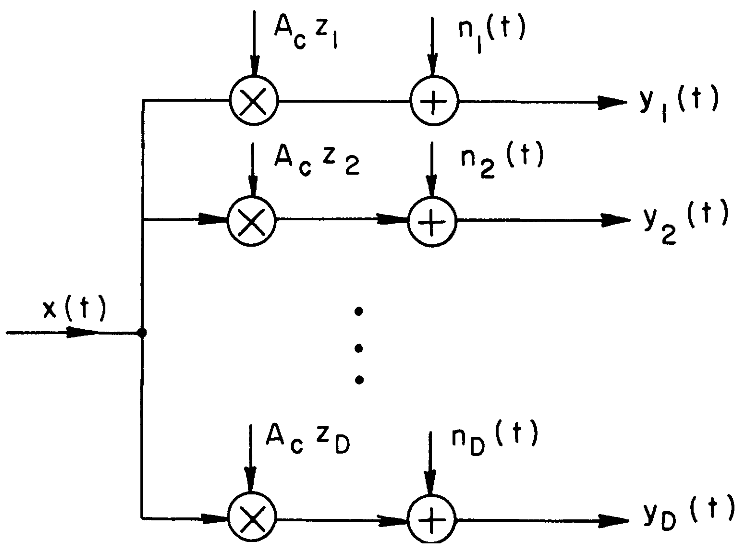

Fig. XXVIII-2. Diversity representation of the turbulent optical channel. Path quantities are independent and identically distributed. $n_i(t)$ are zero-mean complex random processes, whose real and imaginary parts are independent, with power density N_oA_c over the band of interest. $z_i = e^{\gamma_i} = e^{\chi_i + j\phi_i}$, where the γ_i are complex random variables.

The model of Fig. XXVIII-2 is obtained by integrating the channel output, $y(t, \bar{r})$, of Fig. XXVIII-1 over each coherence area. This involves no loss of optimality under the assumptions above. Each integration yields a (complex) temporal random process of the form

$$y_i(t) = A_c z_i x(t) + n_i(t), \tag{2a}$$

where x(t) is the channel input, z_i is the (constant) random value of $z(t, \bar{r})$ over the i^{th} aperture area and the time interval in question, A_c is the area of integration, and $n_i(t)$, the integrated noise, is a zero-mean complex Gaussian random process, whose real and imaginary parts are independent and possess a spectral density, N_oA_c, over all frequencies of interest. The receiving aperture is assumed to contain D coherence areas and, for simplicity, $z(t,\overline{r})$ is assumed to be spatially stationary over the aperture.

The channel model of Fig. XXVIII-2 is thus a D-fold uniform diversity system. Each of the D paths in the system are independent of each other and each suffers constant, or flat-flat, fading. The z_i associated with the i^{th} path can be written from Eq. 1 as

$$z_i = e^{\gamma_i} = e^{\chi_i + j\phi_i}. \tag{2b}$$

It is reasonable to assume that $\overline{\chi_i} = -\sigma^2$, where σ^2 is the variance of χ_i for all i.[1] This ensures that the expected value of $|z_i|^2$ is unity. The variance of ϕ_i is very large, but is not important in the following discussion.

We now suppose that the channel of Fig. XXVIII-2 is used to transmit one of M equiprobable (complex) waveforms, and the receiver is to decide which waveform was transmitted in such a fashion that the probability of error is minimized. If the transmitted waveforms are $S_j(t)$; $j = 1, \ldots, M$, it is well known that the optimum receiver is a maximum-likelihood receiver and that it need only evaluate the quantities[2]

$$y_{ij} = \int y_i(t)\, S_j^*(t)\, dt \qquad i = 1, \ldots, D; \qquad j = 1, \ldots, M. \tag{3}$$

It is easily shown that the likelihood functions that the receiver must evaluate are

$$L_k = \sum_{i=1}^{D} \ln \left\{ \iint du\, d\phi\, p(u, \phi) \exp - \left[\frac{|z|^2 E_k - 2\mathrm{Re}\,[y_{ik} z^*]}{2N_o} \right] \right\}; \qquad k = 1, \ldots, M, \tag{4}$$

where

$$z = u e^{j\phi}, \qquad E_k = A_c \int |S_k(t)|^2\, dt,$$

where $p(u, \phi)$ denotes the probability density of the amplitude and phase of the (complex) random variable z. The receiver decides that the transmitted waveform was that one, say n, for which $L_n \geq L_k$, for all k.

We now seek a lower bound to the error probability that is attainable when the channel of Fig. XXVIII-2 is used with any set of equiprobable complex envelopes $S_i(t)$; $i = 1, \ldots, M$, of average energy E. Clearly, this error probability will be as large as that which would occur if the z_i were known to the receiver. When they are known, the likelihood functions, L_k, of Eq. 4 become

$$L_k = \sum_{i=1}^{D} \frac{1}{2N_o} \left\{ 2\text{Re}\left[y_{ik} z_i^*\right] - |z_i|^2 E_k \right\}, \tag{5}$$

where the y_{ik} and E_k are defined by Eqs. 3 and 4.

Given the transmitted message, say n, and the z_i, the L_k are joint Gaussian random variables with means

$$L_k = \left[B_{nk} - \frac{1}{2} B_{kk}\right] \sum_{i=1}^{D} |z_i|^2 \tag{6a}$$

and covariances

$$(L_k - L_k)(L_j - L_j) = \frac{B_{kj}}{2} \sum_i |z_i|^2, \tag{6b}$$

where

$$B_{kj} = \text{Re}\left[\frac{A_c}{N_o} \int S_k(t)\, S_j^*(t)\, dt\right]. \tag{6c}$$

These statistics are precisely those that would result from the use of the (complex) waveforms $S_i(t)$ with an infinite-bandwidth additive white Gaussian noise channel of noise power density

$$\frac{N_o'}{2} = \frac{N_o}{A_c}\left[\sum_i |z_i|^2\right]^{-1}. \tag{7}$$

Consequently, for any specific values of the z_i, the error probability will be minimized when the $S_i(t)$ are chosen to form a simplex of the given average energy E.[3] Moreover, this choice does not depend on the values of the z_i.

Thus, if the z_i are known to the receiver, the minimum attainable error probability is just the average over the z_i of the error probability for a simplex system of energy-to-noise ratio

$$\frac{EA_c}{N_o} \sum_i |z_i|^2,$$

where E, the energy of the transmitted waveform, Re $[S_i(t) \exp j2\pi ft]$, is given by the expression

$$E = \frac{1}{2} \int |S_i(t)|^2 \, dt; \qquad i = 1, \ldots, M \tag{8}$$

Consequently,[4]

$$P_z[\epsilon] = \int_{-\infty}^{\infty} dx \, \overline{w(x-\sqrt{\beta})} \left\{ 1 - \left[\int_{-\infty}^{x} w(y) \, dy \right]^{M-1} \right\}, \tag{9a}$$

where

$$w(x) = \frac{1}{\sqrt{2\pi}} \exp -\frac{x^2}{2} \tag{9b}$$

$$\beta = \frac{2EA_cM}{N_o(M-1)} \sum_i |z_i|^2, \tag{9c}$$

and the bar denotes the average with respect to the z_i. The subscript z has been added to $P[\epsilon]$ to denote that this is the minimum attainable error probability when the z_i are known to the receiver. We next note that

$$P_z[\epsilon] = \overline{P[\epsilon|\beta]} \geq \int_0^{\bar{\beta}} P[\epsilon|\beta] \, p_\beta(\beta) \, d\beta \tag{10a}$$

or

$$P_z[\epsilon] \geq P[\epsilon|\beta=\bar{\beta}] \, P[\beta \leq \bar{\beta}], \tag{10b}$$

where $P[\epsilon|\beta]$ is the probability of error for a simplex of energy-to-noise ratio $\frac{\beta(M-1)}{M}$. Equation 10b follows from Eq. 10a because the probability of error for a simplex is monotone decreasing in the energy-to-noise ratio.

If $P[\beta \leq \bar{\beta}]$ exceeds a positive number, $K(\sigma)$, for each finite fixed σ and for all D the desired bound is obtained. This follows from Eq. 10b with

$$\bar{\beta} = \frac{2EA_cM}{N_o(M-1)} D \tag{11}$$

(recall that $\overline{|z_i|^2} = 1$), which yields

$$P_z[\epsilon] \geq K(\sigma) \, P[\epsilon|\bar{\beta}]$$

$$= K(\sigma) \int_{-\infty}^{\infty} dx \, w(x-\sqrt{\bar{\beta}}) \left\{ 1 - \left[\int_{-\infty}^{x} w(y) \, dy \right]^{M-1} \right\}. \tag{12}$$

The integral is, in fact, the probability of error for the simplex signal set of energy-to-noise ratio EA_cD/N_o. Thus the probability of error for the channel model of Fig. XXVIII-2 is greater than the product of a positive constant, $K(\sigma)$, and the minimum attainable probability of error for an infinite-bandwidth additive white Gaussian noise channel with energy-to-noise ratio EA_cD/N_o. The maximum rate at which the channel of Fig. XXVIII-2 can be used with arbitrarily low error probability is thus overbounded, for any equiprobable signal set, by

$$C_s = \frac{EA_cD}{TN_o} \log_2 e, \tag{13}$$

the capacity of an infinite-bandwidth additive white Gaussian noise channel with the same power-to-noise ratio. Thus Eq. 12 provides a lower bound to the error probability that can be obtained with the atmospheric optical channel (at least under the assumptions that led to Fig. XXVIII-2) for any set of equiprobable signals with average energy E Similarly, Eq. 13 provides an upper bound to the channel capacity of the atmospheric optical channel when turbulent effects are the dominant disturbance.

All that remains is to show that $P[\beta\leq\bar{\beta}]$ can be bounded away from zero for any finite fixed σ and every positive D. First, consider any finite value for D. The probability of interest can be written

$$P[\beta\leq\bar{\beta}] = P\left[\frac{\Sigma\ |z_i|^2}{D} \leq 1\right]$$

$$= 1 - P\left[\frac{\Sigma|z_i|^2}{D} > 1\right]. \tag{14}$$

From Markov's inequality

$$P[\beta > \bar{\beta}] < \frac{\overline{|\beta|}}{\bar{\beta}} = 1, \tag{15}$$

as β is a non-negative random variable. Combining Eqs. 14 and 15 yields the desired result

$$P[\beta\leq\bar{\beta}] > 0 \tag{16}$$

for any finite D.

The behavior of $P[\beta\leq\bar{\beta}]$ as D approaches infinity can be determined by use of the Central Limit theorem. To see this, consider

$$P[\beta \leq \bar{\beta}] = P\left[\frac{\sum_{i=1}^{D} |z_i|^2}{D} < 1\right]$$

$$= P\left[\frac{\sum_{i=1}^{D} |z_i|^2 - D}{\left[D\left(e^{4\sigma^2} - 1\right)\right]^{1/2}} \leq 0\right]$$

$$= P[s_D \leq 0], \tag{17a}$$

where

$$s_D = \frac{\sum_{i=1}^{D}\left[|z_i|^2 - 1\right]}{\left[D\left(e^{4\sigma^2} - 1\right)\right]^{1/2}}. \tag{17b}$$

The sequence of random variables $\{s_D\}$ has mean zero and unit variance. By an application of the Central Limit theorem, s_D converges in distribution to a zero-mean, unit-variance, Gaussian random variable as D approaches infinity and

$$P[\beta \leq \bar{\beta}] \xrightarrow[D \to \infty]{} \Phi(0) = \frac{1}{2}. \tag{18}$$

The results of Eqs. 16 and 18 guarantee the existence of the desired bound on $P[\beta \leq \bar{\beta}]$ for all D.

E. V. Hoversten, R. S. Kennedy

References

1. E. V. Hoversten, "The Atmosphere as an Optical Communication Channel," 1967 IEEE International Convention Record.
2. J. M. Wozencraft and I. M. Jacobs, Principles of Communication Engineering (John Wiley and Sons, Inc., New York, 1965), Chap. 7.
3. H. J. Landau and D. Slepian, "On the Optimality of the Regular Simplex Code," Bell System Tech. J., Vol. XLV, No. 8, pp. 1247-1272, October 1966.
4. J. M. Wozencraft and I. M. Jacobs, op. cit., Chap. 4.

B. ON OPTIMUM RECEPTION THROUGH A TURBULENT ATMOSPHERE

1. Introduction

It is well known that an optical signal propagating through a turbulent atmosphere suffers from strong log-normal fading.[1] Besides, the field at the receiving aperture is not spatially coherent. The optimum receiver for some such channels has been found and the error probability has been bounded for orthogonal waveforms.[2] This report evaluates the behavior of the likelihood function. The results can be used to construct an optimum receiver and to find its performance bounds. The structure of the simple binary case with no diversity is also discussed.

2. Solution of the Detection Problem

Suppose x(t) is the complex input envelope carrying the message, and $y(t, \vec{r})$ is the corresponding output as a function of time and position in the receiving aperture. Consider a sufficiently small area A_c in the aperture, on which the signal is coherent. The signal integrated over that area has following form

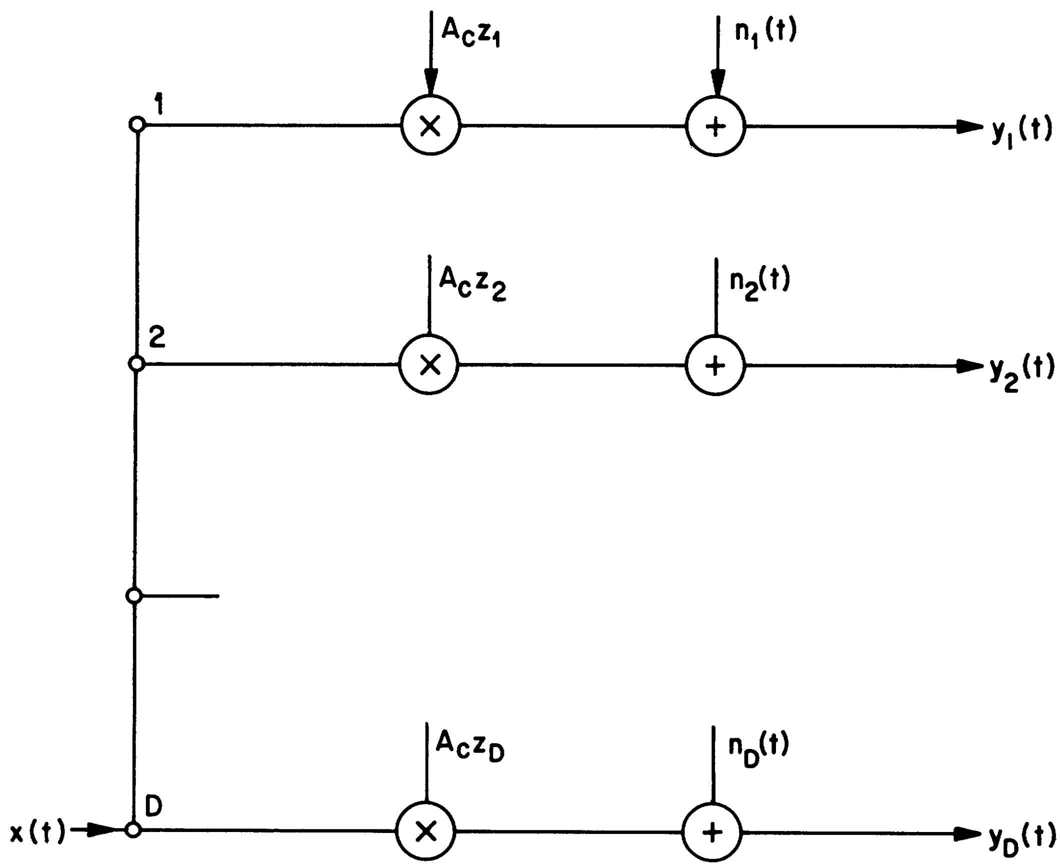

Fig. XXVIII-3. Diversity representation of the turbulent optical channel.

$$y(t) = \int_{A_c} y(t,\vec{r})\, dA = zA_c x(t) + n(t). \tag{1}$$

Here, $z = \exp \gamma(t)$ denotes the complex fading resulting from turbulence. It is known that $\gamma(t)$ is a relatively slowly changing (time constant 100-10 msec) complex Gaussian process.[1] In this report only intervals within which z stays effectively constant are discussed. The other component of Eq. 1, $n(t)$, stands for the noise generated by the surrounding space and the receiver front end. It is assumed to be a zero-mean complex Gaussian noise, whose real and imaginary parts are independent, and of power density N_oA_c over the band of interest. All of the illuminated receiver aperture can be divided into smaller coherent areas, for each of which z and n are statistically independent of corresponding quantities of the other areas. In this way, a diversity representation of the turbulent optical channel is obtained (Fig. XXVIII-3).

Now suppose that equiprobable complex orthogonal waveforms $S_j(t)$; $j = 1, \dots, M$ are transmitted, and the objective is to minimize the total error probability. Statistically, this means the testing of M hypotheses when there are D unwanted complex parameters z_i. In terms of vector notation[4] Eq. 1 may be rewritten

$$\underline{y} = \underline{v} + \underline{n}, \tag{2}$$

where the components of the vectors $\underline{y}$, $\underline{v}$, and $\underline{n}$ are defined by

$$y_{ij} = \int y_i(t)\, S_j^*(t)\, dt, \qquad v_{ij} = A_c z_i \int x_i(t)\, S_j^*(t)\, dt$$

$$n_{ij} = \int n_i(t)\, S_j^*(t)\, dt, \qquad i = 1, \dots, D; \qquad j = 1, \dots, M. \tag{3}$$

The vectors are double-indexed for convenience. The real and imaginary parts of n_{ij} are independent normal random variables with mean zero and variance N_oE_j. If the message k is sent, $v_{ij} = z_iE_k\delta_{jk}$, where $\delta_{jk} = 1$ for $j = k$, and zero otherwise. Also

$$E_k = A_c \int |S_k(t)|^2\, dt. \tag{4}$$

Given the z_i and the k^{th} message, the probability density of $\underline{y}$ is

$$p_{\underline{y}|\underline{z},k}(y|z,k) = \prod_j (2N_oE_j)^{M/2} \exp - \frac{1}{2N_oE_j} \sum_i |y_{ij} - z_iE_k\delta_{jk}|^2. \tag{5}$$

To form the likelihood function L_k the dummy hypothesis (no signal sent) is used, so that the likelihood ratio becomes

$$\Lambda_{\underline{y}|\underline{z},k} = \frac{p_{\underline{y}|\underline{z},k}(\underline{y}|\underline{z},k)}{p_{\underline{y}|\underline{z},0}(\underline{y}|\underline{z},0)} = \sum_{i=1}^{D} \exp - \frac{|y_{ik}-z_iE_k|^2 - |y_{ik}|^2}{2N_oE_k}. \tag{6}$$

Next, this ratio is averaged over $\underline{z}$. (Division by dummy hypothesis before averaging is correct because the probability density in question does not contain $\underline{z}$.) Assuming that the z_i are independent and identically distributed, and taking a logarithm of the result, we obtain the likelihood function L_k.

$$L_k = \ln \overline{\Lambda_{\underline{y}|\underline{z},k}} = \sum_{i=1}^{D} \ln \left\{ \iint dud\phi\ p(u,\phi) \exp\left[- \frac{|y_{ik}-zE_k|^2 - |y_{ik}|^2}{2N_oE_k} \right] \right\}$$

$$= \sum_{i=1}^{D} \ln \left\{ \iint dud\phi\ p(u,\phi) \exp\left[- \frac{|z|^2 E_k - 2\mathrm{Re}\,[y_{ik}z]}{2N_o} \right] \right\}, \tag{7}$$

with $z = u \exp j\phi$. The decision rule is to pick up that message k for which L_k is largest.

The probability density of z is taken to be such that the angle ϕ is uniformly distributed, and the amplitude distribution u is taken to be log-normal, normalized in such a way that $\overline{|z|^2} = 1$. Hence

$$p(u) = \frac{1}{\sqrt{2\pi}\,\sigma u} \exp - \frac{(\sigma^2 + \ln u)^2}{2\sigma^2} \tag{8}$$

and

$$L_k = \sum_{i=1}^{D} \ln \left\{ \int_0^\infty \frac{du}{\sqrt{2\pi}\,\sigma u} I_o\left(u \frac{y_{ik}}{N_o}\right) \exp\left[- \frac{u^2E_k}{2N_o} + \frac{(\sigma^2 + \ln y)^2}{2\sigma^2} \right] \right\}, \tag{9}$$

where the integral with respect to ϕ has been carried out, and $I_o(\cdot)$ is the modified Bessel function of order zero. Figure XXVIII-4 shows that the optimum receiver first correlates each of the D received signals by each of the M signal candidates. Because the signal is a complex envelope, the processing blocks indicated in the figure are more complicated than they are for just real signals. Equivalently, the correlation receiver can be realized by a bank of matched filters. The envelopes of the correlation products are then processing in a nonlinear memoryless device, and finally combined as indicated. A logic circuit then gives an indication of which inputs $L_1, \ldots, L_M$ is largest.

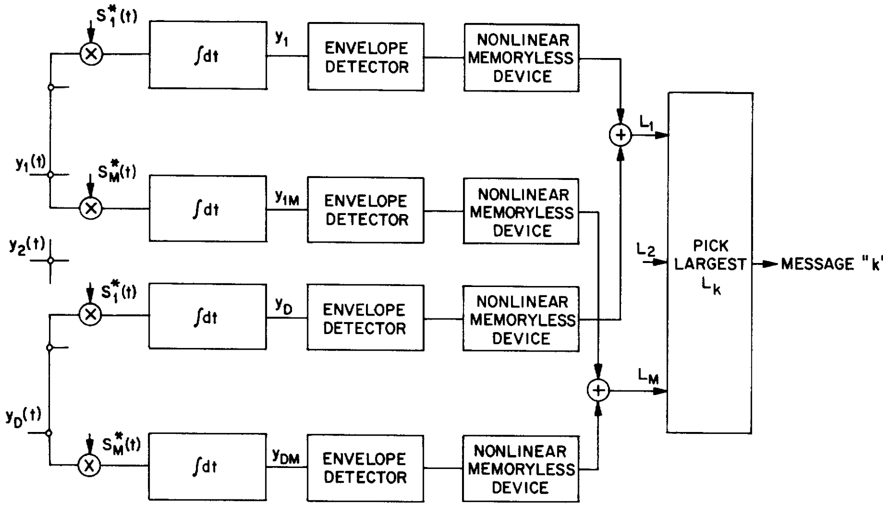

Fig. XXVIII-4. Structure of the optimum diversity receiver.

3. Nonlinear Detection Law

The question that we now consider is, What is the detection law used in the nonlinear memoryless device of Fig. XXVIII-4? This involves the evaluation of the integral in Eq. 9. To get some idea of the nature of the law, one can look at the integral for $\sigma^2 \ll 1$, that is, for small or medium turbulence. Then p(u) behaves almost as an impulse at u = 1. Then set 1 + x = u, and ln u ≈ x, since x « 1 at the domain of x, thereby giving the main contribution to the integral. Furthermore, suppose that $y_{ik}/N_o \gg 1$, so that the large-argument approximation of $I_o(\cdot)$ can be used. Then setting $1 + x \approx p$, we obtain the following results:

$$L = \ln\left\{\int_0^\infty \frac{1}{\sqrt{2\pi}\,\sigma u} I_o\left(u\frac{|y|}{N_o}\right)\exp\left[-\frac{u^2E}{2N_o}+\frac{(\sigma^2+\ln u)^2}{2\sigma^2}\right]\right\}$$

$$\approx \ln\left\{\int_{-1}^\infty \frac{1}{2\pi\sigma(1+x)^{3/2}\sqrt{\frac{|y|}{N_o}}}\exp\left[(1+x)\frac{|y|}{N_o}-\frac{(1+x)^2E}{2N_o}-\frac{(\sigma^2+x)^2}{2\sigma^2}\right]\right\}$$

$$\approx \frac{|y|}{N_o}-\frac{E}{2N_o}-\frac{\sigma^2}{2}+\frac{\sigma^2}{2}\frac{\left(|y|/N_o-E/N_o-1\right)^2}{1+\sigma^2E/N_o}-\frac{1}{2}\ln\left(2\frac{|y|}{N_o}\left(1+\sigma^2\frac{E}{N_o}\right)\right), \qquad (10)$$

in which subscripts have been dropped out. Neglecting the logarithmic part, we see that the nonlinear memoryless device is, in fact, somewhere between a linear and a quadratic detector. Using the small-argument approximation for $I_o(\cdot)$, we see that

$$L \approx \left[\frac{|y|}{2N_o}\right]^2 - \frac{E}{2N_o} - \frac{\sigma^2}{2}\,\frac{\left(\frac{1}{2}\,|y|/N_o^2 - E/N_o - 1\right)^2}{1+\sigma^2 E/N_o} - \frac{1}{2}\ln\left(1+\sigma^2\frac{E}{N_o}\right) \tag{11}$$

for small y/N_o.

To obtain more accurate results and to examine the behavior of L for larger σ, we decided to evaluate the left side of Eq. 10 numerically. Figure XXVIII-5 displays some of the results. The IBM OS-360 computer was used with the FORTRAN compiler. Each detection curve has an initial square-law portion as predicted by (11), and for small σ there is also a linear portion agreeing with (10). For very large inputs,

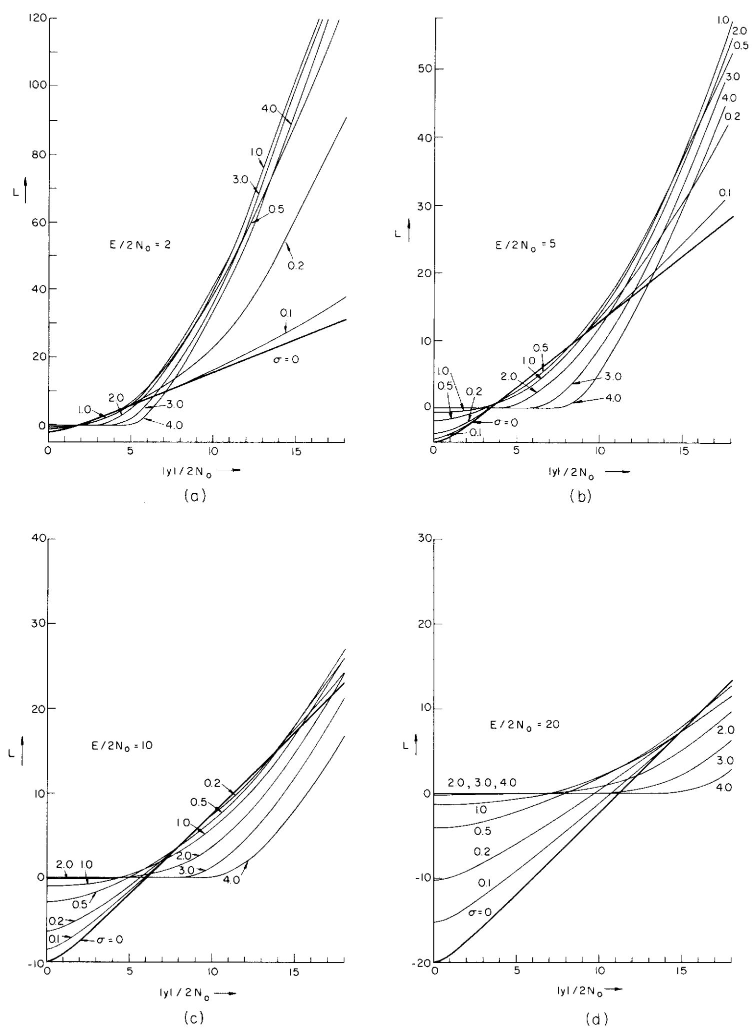

Fig. XXVIII-5. Optimum detection law characteristics as function of normalized input signal $|y|/2N_o$, and average signal-to-envelope-noise ratio, $E/2N_o$, for various values of the log-normal fading variance, σ.

$|y|/N_o$, the behavior is again close to that of a square-law detector. For $\sigma > 1$, the detector has a clear threshold. With each family of curves the asymptotic behavior is given by

$$\lim_{\sigma\to 0} L = \ln I_o\left(\frac{y}{N_o}\right) - \frac{E}{2N_o}, \tag{12}$$

as indicated in the figure.

4. Probability of Error

The probability of error of the optimum receiver of Fig. XXVIII-4 can be bounded[2] as follows.

$$P[\epsilon] \leq 2 \cdot 2^{-TC_g E(R)}, \tag{13}$$

where $C_g = DE \log_2 e/(2N_oT)$, $R = \log_2 M/T$, with T being the duration of a message, and E defined by (4), under the assumption that all signal energies are equal. The error exponent is given by

$$E(R) = \max_{0\leq\rho\leq 1} [E_o(\rho)-\rho R/C_g] \tag{14}$$

$$E_o(\rho) = -\frac{1+\rho}{a_p} \ln \left\{\int_0^\infty dy\, e^{y} \left[\int_0^\infty du\, p(u)\, I_o(2u\sqrt{y a_p})\, e^{u^2 a_p}\right]^{\frac{1}{1+\rho}}\right\}, \tag{15}$$

where $a_p = E/2N_o$, and $p(u)$ is given by (8). The physical meaning of a_p is "energy-to-envelope noise ratio per diversity path." The inner integral has already been evaluated numerically. For small σ the approximations (10) and (11) are helpful. For $a_p \gg 1$, a very simple result follows roughly:

$$E_o(\rho) \approx \frac{\rho}{1+\rho\left(1+2\sigma^2 a_p\right)}. \tag{16}$$

Therefore

$$E(R) \approx \frac{(1-R/C_g)^2}{1+2\sigma^2 a_p} \qquad C_g \geq R \geq R_{crit} \tag{17}$$

$$E(r) \approx \frac{1}{2\left(1+\sigma^2 a_p\right)} - \frac{R}{C_g} \qquad 0 \leq R \leq R_{crit} \tag{18}$$

where $R_{crit} \approx C_g/\left(4\left(1+\sigma^2\alpha_p\right)^2\right)$. By adding some logarithmic terms to Eq. 16, we can express the behavior of the exponent more accurately. The more accurate and complicated expression shows that the zero-rate error exponent has a maximum at a certain value of α_p. This result agrees with the findings of Kennedy and Hoversten.[2] Figure XXVIII-6 illustrates the behavior of the reliability of the channel, as compared with nonfading Gaussian and Rayleigh channel reliabilities.[2]

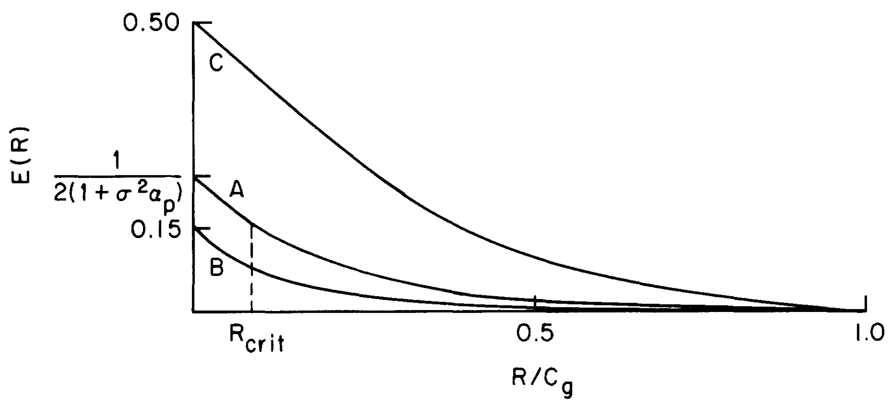

Fig. XXVIII-6. Reliability curves for infinite-bandwidth channels.
(a) Log-normal, diversity below optimum.
(b) Rayleigh fading, optimum diversity.
(c) Gaussian constant channel.

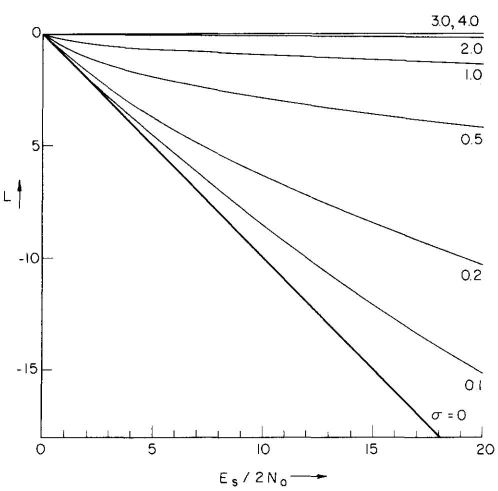

Fig. XXVIII-7.

Error exponent for a log-normal channel with no diversity and two orthogonal signals as a function of average signal-to-envelope-noise ratio $E_s/2N_o$, and the fading variance, σ.

When there is no diversity (D=1) the results will be simpler. In this case, the optimum receiver is just a correlator-square-law envelope detector. Hence the result in Wozencraft and Jacobs[5] can be used:

$$P[\epsilon] = \frac{1}{2}\, \overline{e^{-u^2 E_s/(2N_o)}} = \frac{1}{2} \int p(u)\, e^{-u^2 E_s/2N_o}\, du$$

$$= \frac{1}{2} \exp L, \tag{19}$$

where E_s is the signal energy, N_o is the noise power density per unit area, and L is to be computed for y = 0. The exponent for this case is displayed in Fig. XXVIII-7. The error does not decrease exponentially as the signal-to-noise ratio increases. For small σ, Eq. 11 can be applied by setting y = 0. Thus

$$L \approx -\frac{(E_s/N_o)(1-\sigma^2)^2}{2\left(1+\sigma^2 E_s/N_o\right)} - \frac{1}{2} \ln \left(1+\sigma^2 E_s/N_o\right). \tag{20}$$

When $E_s/N_o \to \infty$, the error probability goes to zero slowly, in fact slower than in the case of Rayleigh fading.

Equation 19 can be generalized to the case of M orthogonal signals (no diversity) by using the union bound.[3]

S. J. Halme

References

1. V. I. Tatarski, Wave Propagation in a Turbulent Medium (McGraw-Hill Book Co., Inc., New York, 1961), pp. 206-293.
2. R. S. Kennedy and E. V. Hoversten, "On the Reliability of the Atmosphere as an Optical Communications Channel," a paper presented at the IEEE International Symposium on Information Theory, San Remo, Italy, September 11-15, 1967 (to be published in IEEE Transactions on Information Theory).
3. J. M. Wozencraft and I. M. Jacobs, Principles of Communication Engineering (John Wiley and Sons, Inc., New York, 1965).
4. Ibid., see Chap. 4.
5. Ibid., p. 533.

C. COMPARISON OF THE PERFORMANCE OF THE PHOTON COUNTER AND THE CLASSICAL OPTIMUM RECEIVER

In this report, the minimum probabilities of error for the reception of binary signals attainable by a photon counter and by the classical optimum receiver are compared. We shall show that for the type of binary on-off signals considered, the photon counter yields a smaller error probability in the limit of low signal-to-noise ratio and low noise level. On the other hand, it is known that in the classical limit (high signal and noise levels), the classical optimum receiver that yields the smallest probability of error is not a photon counter. We shall also show that for binary orthogonal signals, the classical optimum receiver yields a smaller error probability at all signal and noise levels.

For the purpose of comparing the performances of the two types of receivers for the reception of binary on-off signals, it suffices to consider the simplest equally likely binary signal set defined by the correspondence

$$m = m_0 \longleftrightarrow \vec{J}(\underline{r}, t) = 0$$

$$m = m_1 \longleftrightarrow \vec{J}(\underline{r}, t) = I \cos (\omega t+\phi)\ \delta^{(3)}(\underline{r})[u_{-1}(t)-u_{-1}(t-T)]\ \vec{e}_r.$$

In these equations, $\vec{J}(\underline{r}, t)$ is a classical current distribution in the transmitting aperture of infinitesimal dimensions.

It has been shown,[1] that the optimum receiving system for this signal set consists of a resonant cavity, the natural frequency of the only dominant mode of which is ω. The cavity is initially empty and is exposed to the signal source for the time interval $(0, T)$ when its aperture is open. At thermal equilibrium, the state of the electromagnetic field inside of the cavity after the exposure is given by the density operator

$$\rho_0 = \frac{1}{\pi \langle n \rangle} \int \exp\left(-\frac{|\alpha|^2}{\langle n \rangle}\right) |\alpha\rangle\langle\alpha|\ d^2\alpha \tag{1a}$$

if m_0 is transmitted. On the other hand, the state of the field is given by the density operator

$$\rho_1 = \frac{1}{\pi \langle n \rangle} \int \exp\left(-\frac{|\alpha-\sigma|^2}{\langle n \rangle}\right) |\alpha\rangle\langle\alpha|\ d^2\alpha \tag{1b}$$

if m_1 is transmitted. In Eqs. 1a and 1b, $\langle n \rangle$ is the average number of photons in the chaotic thermal noise field. $|\sigma|^2$ is a function of the parameters I, T, ω, etc. (see the previous report[2]) and is proportional to the average received signal energy. Without loss of generality, we shall assume that σ is real.

1. Performance of the Classical Optimum Receiver

It is known that in the classical limit, the (classical) optimum receiver measures the dynamical variable $\frac{a^{+}+a}{2}$, of the field inside of the receiving cavity. This variable is just the amplitude of the component of the electric field which is in phase with the transmitted signal field, that is, with the electromagnetic field that would exist in the receiving cavity in the absence of noise. The probability of error, $P_c(\epsilon)$, for this classically optimum receiver has been given previously.[3] It is

$$P_c(\epsilon) = Q\left(\frac{\sigma}{\sqrt{2<n>+1}}\right) \tag{2}$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \exp\left(-\frac{1}{2} y^2\right) dy.$$

2. Performance of the Photon Counter

The photon counter measures the energy of the electromagnetic field inside of the receiving cavity. When the density operator of the field in P-representation has a weight function $p(\alpha)$, the probability distribution function of the number of photons detected by an ideal photon counter [1] is given by

$$p(n) = \int p(\alpha) \frac{|\alpha|^{2n}}{n!} \exp(-|\alpha|^2)\, d^2\alpha.$$

Therefore, the conditional probability distributions of the photon count are

$$p(n/m_0) = \frac{1}{1+<n>} \left(\frac{<n>}{1+<n>}\right)^n \tag{3a}$$

$$p(n/m_1) = \frac{1}{1+<n>} \left(\frac{<n>}{1+<n>}\right)^n \sum_{r=0}^{n} \binom{n}{r} \frac{1}{r!} \left[\frac{\sigma^2}{<n>(1+<n>)}\right]^r \exp\left(-\frac{\sigma^2}{1+<n>}\right). \tag{3b}$$

To minimize the probability of error, $P_p(\epsilon)$, the receiver sets the estimate m to

$$\begin{aligned} &m_0 \quad \text{if} \quad p(n/m_0) \geq p(n/m_1) \\ &m_1 \quad \text{if} \quad p(n/m_1) > p(n/m_0), \end{aligned} \tag{4}$$

where n is the observed value of the number of photons. It follows that

$$P_p(\epsilon) = \frac{1}{2}\left\{\sum_{n \epsilon A} p(n/m_1) + \sum_{n \epsilon A} p(n/m_0)\right\},$$

where A is the set of integers n that are such that

$$\left\{\sum_{r=0}^{n} \binom{n}{r} \frac{1}{r!} \left[\frac{\sigma^2}{<n>(1+<n>)}\right]^r\right\} \cdot \exp\left(-\frac{\sigma^2}{1+<n>}\right) \leq 1.$$

Unfortunately, one cannot find an analytic expression for $P_p(\epsilon)$ for all values of σ and $<n>$. An upper bound on the error probability is

$$P_p(\epsilon) \leq \frac{1}{2}\left\{\sum_{n=1}^{\infty} p(n/m_0) + p(n=0/m_1)\right\}.$$

That is,

$$P_p(\epsilon) \leq \frac{1}{2}\left[\frac{<n>}{1+<n>} + \frac{1}{1+<n>} \exp\left(\frac{-\sigma^2}{1+<n>}\right)\right], \tag{5}$$

where the equality sign holds only in the case $\frac{\sigma^2}{<n>} \ll 1$, and $<n> < 1$.

3. Comparison of the Two Types of Receivers

In the classical limit, that is, at high signal and noise levels, we have

$$P_c(\epsilon) \leq P_p(\epsilon).$$

This inequality follows from the fact that the classical optimum receiver is the receiver that yields the minimum attainable probability of error.

In the case of low signal-to-noise ratio $\left(\frac{\sigma^2}{<n>} \ll 1\right)$, the error probability of the classical optimum receiver in Eq. 2 can be approximated by

$$P_c(\epsilon) \approx \frac{1}{2} \exp\left(-\frac{\sigma^2}{4<n>+2}\right) \approx \frac{1}{2}\left(1 - \frac{\sigma^2}{4<n>+2}\right), \tag{6}$$

and Eq. 5 becomes

$$P_p(\epsilon) \leq \frac{1}{2}\left\{\frac{<n>}{1+<n>} + \frac{1}{1+<n>} \exp\left(-\frac{\sigma^2}{1+<n>}\right)\right\}$$
$$\approx \frac{1}{2}\left\{1 - \frac{\sigma^2}{(1+<n>)^2}\right\}.$$

Comparing this expression with that in Eq. 6, we see that, for $\frac{\sigma^2}{<n>} \ll 1$ and $<n> \leq 1 + \sqrt{2}$,

$$P_p(\epsilon) \leq P_c(\epsilon).$$

That is, in the limit of low signal-to-noise ratio and low noise level, the classical optimum receiver no longer yields the minimum error probability.

For binary orthogonal signals, however, it can be shown that the classical optimum receiver yields a lower probability of error for all σ and $<n>$, when compared with the photon counter. The binary orthogonal signals are defined by the correspondence:

$$m = m_0 \longleftrightarrow \rho_0 = \frac{1}{\pi^2 <n>^2} \int \exp\left[-\frac{|\alpha_1-\sigma|^2 + |\alpha_2|^2}{<n>}\right] |\alpha_1, \alpha_2\rangle \langle \alpha_1, \alpha_2| \, d^2\alpha_1 d^2\alpha_2$$

$$m = m_1 \longleftrightarrow \rho_1 = \frac{1}{\pi^2 <n>^2} \int \exp\left[-\frac{|\alpha_1|^2 + |\alpha_2-\sigma|^2}{<n>}\right] |\alpha_1 \alpha_2\rangle \langle \alpha_1 \alpha_2| \, d^2\alpha_1 d^2\alpha_2,$$

where ρ_0 and ρ_1 are density operators specifying the state of the electromagnetic field in the receiving cavities.[1] When the number of photons in the two relevant modes are measured separately, the joint conditional probability of the photon count is

$$p(n_1, n_2/m_0) = \left(\frac{1}{1+<n>}\right)^2 \left(\frac{<n>}{1+<n>}\right)^{n_1+n_2} \left\{\sum_{r=0}^{n_1} \binom{n_1}{r} \frac{1}{r!} \left[\frac{\sigma^2}{<n>(1+<n>)}\right]^r\right\} \exp\left(-\frac{\sigma^2}{1+<n>}\right)$$

$$p(n_1, n_2/m_1) = \left(\frac{1}{1+<n>}\right)^2 \left(\frac{<n>}{1+<n>}\right)^{n_1+n_2} \left\{\sum_{r=0}^{n_2} \binom{n_2}{r} \frac{1}{r!} \left[\frac{\sigma^2}{<n>(1+<n>)}\right]^r\right\} \exp\left(-\frac{\sigma^2}{1+<n>}\right).$$

The receiver sets the estimate m to

$$m_0 \quad \text{if} \quad n_1 \geq n_2$$

$$m_1 \quad \text{if} \quad n_1 \leq n_2.$$

Hence the probability of error is

$$P_p(\epsilon) = \sum_{n_1=0}^{\infty} \sum_{n_2=n_1}^{\infty} \left(\frac{1}{1+<n>}\right)^2 \left(\frac{<n>}{1+<n>}\right)^{n_1+n_2} \left\{\sum_{r=0}^{n_1} \binom{n_1}{r} \frac{1}{r!} \left[\frac{\sigma^2}{<n>(1+<n>)}\right]^r\right\} \exp\left[-\frac{\sigma^2}{1+<n>}\right]$$

$$= \frac{1+<n>}{1+2<n>} \exp\left[-\frac{\sigma^2}{1+2<n>}\right].$$

This value of error probability is always larger than that of the classical optimum receiver given in the previous report.[1]

$$P_c(\epsilon) = Q\left(\sqrt{\frac{2\sigma^2}{1+2\langle n\rangle}}\right) \leq \frac{1}{2}\exp\left(-\frac{\sigma^2}{1+2\langle n\rangle}\right).$$

Jane W-S. Liu

References

1. Jane W-S. Liu, "Quantum Communication Systems," Quarterly Progress Report No. 86, Research Laboratory of Electronics, M.I.T., July 15, 1967, pp. 230-239.

2. *Ibid.*, see Equation 8, p. 233.

3. *Ibid.*, see Equation 18, p. 238.