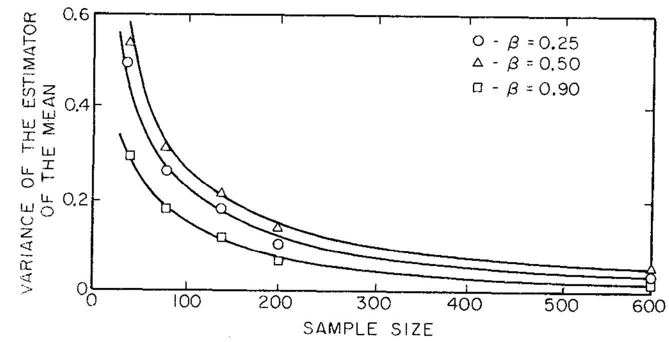

Fig. 2. Variance on the estimate of the mean of $p(x|\theta_1)$, ensemble averaged over 50 runs.

irrespective of β, the effective value of β is reflected in the variance associated with the estimators. A plot of the variance of the estimators (computed across the run ensemble) for three different values of β is shown in Fig. 2. This plot shows that as the effective value of β increases, the variance decreases.

V. Conclusions

We have looked at a parametric scheme for learning to recognize patterns with an imperfect teacher. The proposed scheme is computationally feasible since the reproducing properties of density functions are used in learning about the unknown parameters. All three different learning situations, namely, learning with a perfect teacher, learning with an imperfect teacher, and learning without a teacher, can be handled by the proposed scheme.

References

[1] A. K. Agrawala, "Learning with a probabilistic teacher," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 373–379, July 1970.
[2] R. O. Duda and R. C. Singleton, "Training a threshold logic unit with imperfectly classified patterns," in *IEEE WESCON Conv. Rec.*, pt. 3.2, pp. 1–6, Aug. 1964.
[3] A. W. Whitney and S. J. Dwyer, III, "Performance and implementation of the k-nearest neighbor decision rule with incorrectly identified training samples," in *Proc. 1970 IEEE Systems Science and Cybernetics Conf.*, Pittsburg, Pa., Oct. 14–16, pp. 32–38.
[4] K. Shanmugan and A. M. Breipohl, "Learning with an imperfect teacher," in *Proc. 1970 IEEE Systems Science and Cybernetics Conf.*, Pittsburg, Pa., Oct. 14–16, pp. 32–38.
[5] S. C. Fralick, "Learning to recognize patterns without a teacher," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 57–64, Jan. 1967.
[6] J. Spragins, "A note on the iterative application of Bayes' rule," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 544–549, Oct. 1965.
[7] H. J. Scudder, III, "Adaptive communication receivers," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 167–174, Apr. 1965.
[8] J. L. Doob, *Stochastic Processes.* Princeton, N.J.: Van Nostrand, 1962, pp. 319–334.

Channel Capacity and Maximum-Likelihood Detection for Atmospherically Disturbed Binary Photocounting Communications

R. Y. YEN, P. DIAMENT, AND M. C. TEICH

***Abstract*—A binary optical communication system, consisting of a laser transmitter, a log-normal atmospheric channel, a photocounting detector, and a likelihood-ratio decision box has been investigated. Background radiation and dark current have been taken into account. Curves are presented for the probability of error and the mutual information for varying levels of atmospheric turbulence. It was ascertained that optical preamplification could be useful in improving system performance.**

With the advent of the laser, there has been a considerable amount of effort involved in analyzing various optical communication systems utilizing a photocounting receiver. Photoelectron counting statistics [1]–[7], maximum-likelihood detection [8]–[12], and information-rate calculations [13], [14] have all been considered. In particular, Diament and Teich [5], [6], and Teich and Rosenberg [7] have obtained the photoelectron counting distributions for stochastic radiation after transmission through the turbulent atmosphere. Both Lachs and Jankowich [8] and Helstrom [9] have examined the decision threshold and error probability using maximum-likelihood detection for binary communication systems in which the laser signal is coherent and the background noise is thermal. Similar calculations have been carried out by many other authors using various radiation models [10], [11].

Maximum-likelihood detection through the turbulent atmosphere has recently been treated by Solimeno *et al.* [12]. These authors presented binary error probabilities for direct detection in the presence of log-normal fading. However, the case they consider is one in which the laser is ideal rather than noisy, and the fluctuations of the background radiation are discernable. That is, the counting time is assumed to be much less than the background coherence time, and thus the counting distribution in the absence of signal is Bose–Einstein, instead of Poisson, as considered here. In view of the fractional bandwidth of available optical interference filters, the coherence time of the background radiation is $\lesssim 10^{-12}$ s, and their assumption that the background statistics are resolved is unrealistic. Finally, Fillmore and Lachs [13] have calculated the information rates for a laser binary channel using a stable laser source with and without additive Gaussian noise. These authors also treat the noiseless case with a Gaussian signal. Most recently, Jodoin and Mandel [14] have evaluated the information rates for an optical communication channel in which the light beam is amplitude-modulated at the source by a filter of continuously variable transmittance.

In this correspondence, we consider the maximum-likelihood detection for a laser binary system in which the radiation, consisting of interfering coherent and chaotic components, is modulated by an on–off gate and passed through the log-normal turbulent atmosphere. The noise arising from background radiation and from dark current in the photodetector is considered to be independent and noninterfering and therefore leads to a Poisson distribution [5]. The basic counting distributions used in this work are obtained by the methods outlined by Diament and Teich [6]. In addition, we also calculate the mutual information and channel capacity for this atmospherically disturbed photocounting system, both with and without likelihood detection.

We consider a radiation source that emits an interfering superposition of an amplitude-stabilized beam and a chaotic (Gaussian) noise beam. This is a suitable model for laser radiation, and under certain conditions, for scattered radiation [6] as well. In the absence of atmospheric turbulence and with a detector of adequate bandwidth, the counting distribution $p_0(n,y,N)$ is then given by [1], [4], [6]

$$p_0(n,y,N) = (Hy)N^nH^{n+1} \exp(-yNH)L_n(-y[1+y]H), \quad (1)$$

where $H \equiv (N+1+y)^{-1}$, N is the overall mean count, y is the ratio of average coherent-to-chaotic irradiances, $L_n(x)$ is the Laguerre polynomial, and n is the number of counts.

After transmission through the turbulent atmosphere, as Diament and Teich have shown, the disturbed counting distribution $p_S(n,\sigma,y,N_S)$ is [5], [6]

$$p_S(n,\sigma,y,N_S) = \frac{p_0(n,y,M)\exp[-\frac{1}{2}\sigma^2 q_1^2(n,M)]}{[1-\sigma^2 q_2(n,M)]^{1/2}}, \quad (2)$$

where N_S is the overall mean count and σ is the standard deviation of the logarithmic irradiance, which varies between 0 (quiescent atmosphere) and about 1.5 (saturation value for the very turbulent atmosphere). The quantity q_j is given by

$$q_j(n,M) = \frac{\partial^j \ln p_0(n,y,M)}{\partial(\ln M)^j}, \quad (3)$$

the parameter M being determined implicitly for each count number n from the stationarity condition

$$\ln M = \ln N_S - \tfrac{1}{2}\sigma^2 + \sigma^2 q_1(n,M). \quad (4)$$

Equation (2) is generally valid for any reasonable single-peaked distribution.

Manuscript received March 30, 1971; revised August 18, 1971. This work was supported by NSF Grant NSF-GK-16649.
The authors are with the Department of Electrical Engineering and Computer Science, Columbia University, New York, N.Y. 10027.

At the receiver, the independent, additive, and noninterfering noise generated by dark current and background radiation yields a Poisson counting distribution $p_H(n,N_H)$, which, in the absence of signal, is given by

$$p_H(n,N_H) = \frac{N_H{}^n}{n!} e^{-N_H}. \tag{5}$$

Here N_H is the total noise mean count. The overall counting distribution p_{SH} arising from both the atmospherically disturbed laser signal and the Poisson noise is then obtained from the convolution of p_S and p_H:

$$p_{SH}(n,\sigma,y,z,N) = \sum_{m=0}^{n} p_S(m,\sigma,y,Nz/[1+z])p_H(n-m, N/[1+z]), \tag{6}$$

where the total mean count is $N = N_S + N_H$, and the final signal-to-noise ratio is defined by $z = N_S/N_H$.

Now, using the maximum-likelihood criterion [10], the decision threshold n_D is determined by the minimum n satisfying the condition $p_{SH}(n)/p_H(n) \geq (1-Q)/Q$, where Q is the *a priori* probability that the laser signal is transmitted. In terms of the parameter n_D, the probability of error P_e is given by

$$P_e = Q\left[1 - \sum_{n=n_D}^{\infty} p_{SH}(n)\right] + (1-Q)\sum_{n=n_D}^{\infty} p_H(n). \tag{7}$$

In Fig. 1(a) the probability of error P_e is plotted as a function of the final signal-to-noise ratio z for $\sigma = 1.0$, with N_H and y as parameters. The quantity Q is taken to be 0.5. It can be seen that for fixed y, increasing the value of signal N_S and noise N_H, while holding the ratio constant, reduces P_e and therefore improves receiver performance. Furthermore, it can also be seen that increasing the coherent-to-chaotic ratio further improves performance. From Fig. 1(b), with N_H, y, and z held constant, it is apparent that increasing turbulence serves to degrade performance by giving a larger probability of error, as expected. Thus, optimum performance is achieved by maximizing N_S, z, and y, and by operating in channels for which σ is minimum. This result is intuitively meaningful and implies that optical preamplification may be useful in improving the system performance, provided that it does not introduce excessive noise of its own.

Simple Photocounting Receiver

Consider a binary system in which the input to the channel U is taken to be equal to 1 when the laser signal is present and 0 when the laser signal is absent. The output of the system V_1 is equal to the number of photoelectrons emitted; thus V_1 can take on all nonnegative integer values $0,1,2,\cdots,\infty$. In this receiver configuration we exclude the likelihood-ratio test box.

Letting $p(u=1) = Q$, $p(u=0) = 1-Q$, and $p(v=n) = r(n)$, the mutual information $I(U;V)$ can be expressed as [15]–[18]

$$I(U;V) = \sum_{n=0}^{\infty}\left[Qr(n \mid 1)\log\frac{r(n \mid 1)}{r(n)} + (1-Q)r(n \mid 0)\log\frac{r(n \mid 0)}{r(n)}\right], \tag{8}$$

with

$$r(n) = Qr(n \mid 1) + (1-Q)r(n \mid 0). \tag{9}$$

If the likelihood-ratio test is excluded, we have $r(n \mid 0) = p_H(n,N_H)$ and $r(n \mid 1) = p_{SH}(n,\sigma,y,z,N)$, as given by (5) and (6), respectively. With $r(n \mid 0)$ and $r(n \mid 1)$, we can calculate $I(U;V)$ for any given value of Q, and thus find the channel capacity C, which is given by

$$C = \max_{p(u)} I(U;V). \tag{10}$$

For the system considered here, this quantity is a function of the radiation statistics, the strength of atmospheric turbulence, and the various mean levels of signal and noise radiation.

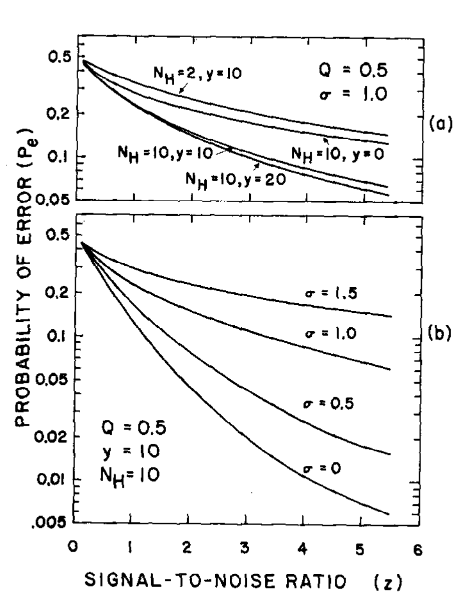

Fig. 1. Probability of error P_e versus final signal-to-noise ratio z. (a) Q and σ are fixed and N_H and y are varied as parameters. (b) Q, y, and N_H are fixed and σ is varied as a parameter.

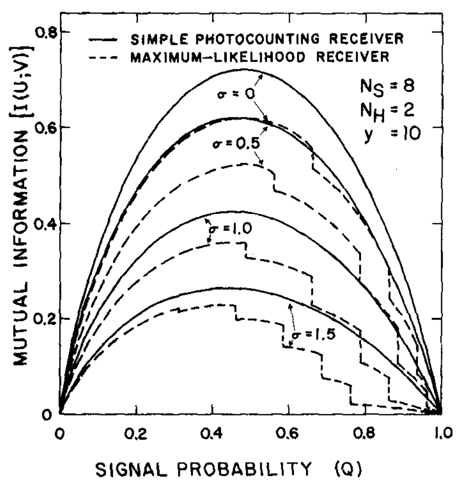

Fig. 2. Mutual information $I(U;V)$ versus signal probability Q with turbulence σ as a parameter. N_S, N_H, and y are fixed.

Maximum-Likelihood Photocounting Receiver

In this case, the output V_2 may take on only two values, 0 and 1, as determined by the likelihood-ratio test. Thus,

$$r(0 \mid 0) = \sum_{n=0}^{n_D-1} p_H(n,N_H), \quad r(1 \mid 0) = 1 - r(0 \mid 0),$$

$$r(0 \mid 1) = \sum_{n=0}^{n_D-1} p_{SH}(n,\sigma,y,z,N), \quad r(1 \mid 1) = 1 - r(0 \mid 1)$$

and the mutual information is given by the well-known formula

$$I(U;V) = \sum_{m=0}^{1}\left[Qr(m \mid 1)\log\frac{r(m \mid 1)}{r(m)} + (1-Q)r(m \mid 0)\log\frac{r(m \mid 0)}{r(m)}\right]. \tag{11}$$

In this case, $r(v \mid u)$ is a function of the *a priori* probability Q, since $r(v \mid u)$ depends on n_D, which is a function of Q.

In Fig. 2 we present the results for $I(U;V)$ versus Q for both types of receiver. The solid curves refer to the simple photocounting receiver without the likelihood-ratio test, while the dashed curves represent the maximum-likelihood photocounting receiver. We have arbitrarily set $N_S = 8$, $N_H = 2$, and $y = 10$, and have plotted curves for various values of σ from 0 to 1.5. The channel capacity, represented by the peak of the curve, occurs near $Q = 0.5$ and decreases with increasing turbulence. Similar curves are obtained by fixing N_H, y, and σ, while varying N_S, and by fixing N_S, N_H, and σ, while varying y. In these cases the channel capacity increases with increasing N_S and y, respectively. In all cases, the mutual information achieves its maximum value near $Q = 0.5$.

An obvious distinction between the two receivers is that the maximum-likelihood receiver exhibits discontinuities in the $I(U; V)$ versus Q curves. This is because of the discrete nature of the decision threshold n_D, which jumps from one integer to the next at certain values of Q, as Q varies. The likelihood receiver yields a lower capacity than the simple photocounting receiver in all cases. Since both the input and output take on values of 0 and 1 only with the binary photocounting likelihood receiver, it may be considered as an asymmetrical binary channel [18] with varying error transition probability, i.e., $r(1 \mid 0)$ and $r(0 \mid 1)$. Thus it is expected that the channel capacity will not exceed 1.

The results presented here can be reduced to several special cases, including pure coherent radiation ($y = \infty$), pure chaotic radiation ($y = 0$), the vacuum channel ($\sigma = 0$), the saturated turbulent channel ($\sigma \simeq 1.5$), and the noiseless system ($z = 0$).

References

[1] J. A. Armstrong and A. W. Smith, "Experimental studies of intensity fluctuation in lasers," in *Progress in Optics*, E. Wolf, Ed. Amsterdam: North-Holland, 1967, pp. 213–257.

[2] C. Freed and H. A. Haus, "Photoelectron statistics produced by a laser operating below and above the threshold of oscillation," *IEEE J. Quantum Electron.*, vol. QE-2, pp. 190–195, Aug. 1966.

[3] G. Lachs, "Theoretical aspects of mixtures of thermal and coherent radiation," *Phys. Rev.*, vol. 138, pp. B1012-B1016, May 1965; also G. Lachs, "Approximate photocount statistics for coherent and chaotic radiation of arbitrary spectral shape," *J. Appl. Phys.*, vol. 42, pp. 602–609, Feb. 1971.

[4] P. Diament and M. C. Teich, "Photoelectron-counting distributions for irradiance-modulated radiation," *J. Opt. Soc. Amer.*, vol. 60, pp. 682–689, May 1970.

[5] ——, "Photodetection of low-level radiation through the turbulent atmosphere," *J. Opt. Soc. Amer.*, vol. 60, pp. 1489–1494, Nov. 1970.

[6] ——, "Optical detection of laser or scattered radiation transmitted through the turbulent atmosphere," *Appl. Opt.*, vol. 10, pp. 1664–1667, July 1971.

[7] M. C. Teich and S. Rosenberg, "*N*-fold joint photocounting distribution for modulated laser radiation: Transmission through the turbulent atmosphere," *J. Optoelectron.*, vol. 3, pp. 63–76, May 1971.

[8] G. Lachs and E. Jankowich, "Some results on error rates for a laser binary communication system," *Proc. IEEE* (Lett.), vol. 56, pp. 744–745, Apr. 1968.

[9] C. W. Helstrom, "Performance of an ideal quantum receiver of a coherent signal of random phase," *IEEE Trans. Aerosp. Electron. Syst.* (Corresp.), vol. AES-5, pp. 562–564, May 1969; "Photoelectric detection of coherent light in filtered background light," *IEEE Trans. Aerosp. Electron. Syst.* (Corresp.), vol. AES-7, pp. 210–213, Jan. 1971.

[10] W. K. Pratt, *Laser Communications Systems*. New York: Wiley, 1969, pp. 167–168, 196–199.

[11] I. Bar-David, "Communication under the Poisson regime," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 31–37, Jan. 1969.

[12] S. Solimeno, E. Corti, and B. Nicoletti, "Optical communication in turbulent media," *J. Opt. Soc. Amer.*, vol. 60, pp. 1245–1251, Sept. 1970.

[13] G. L. Fillmore and G. Lachs, "Information rates for photocount detection systems," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 465–468, July 1969.

[14] R. Jodoin and L. Mandel, "Information rate in an optical communication channel," *J. Opt. Soc. Amer.*, vol. 61, pp. 191–198, Feb. 1971.

[15] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 623–656, 1948.

[16] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968, p. 18.

[17] R. B. Ash, *Information Theory*. New York: Interscience, 1965, pp. 1–26, 46–50.

[18] C. W. McMullen, *Communication Theory Principles*. New York: Macmillan, 1968, p. 268.

The Systematic Selection of Cyclically Equivalent Codes

I. S. REED AND C. T. WOLVERTON

***Abstract*—A systematic procedure for constructing one member of each cyclic equivalence class of an $(n, k + 1)$ Reed–Solomon code is presented. It is shown that if the procedure is modified to exclude those codewords that do not have maximum period, the resulting set of codewords constitutes a synchronizable code having comma freedom of degree $n - 2k$.**

Manuscript received February 21, 1971; revised May 12, 1971. This work was supported by the U.S. Air Force under Contract F 19628-71-C-0002.

I. S. Reed is with the University of Southern California, Los Angeles, Calif.
C. T. Wolverton is with the MITRE Corporation, Bedford, Mass.

I. Introduction

In [1] kth-order near-orthogonal codes were developed. A code E' in this class is constructed by choosing one member from each cyclic equivalence class of an $(n, k + 1)$ Reed–Solomon (RS) code E. (The cyclic equivalence class of a codeword consists of all distinct cyclic shifts of that codeword; equivalence under cyclic shifts, subsequently referred to as ρ equivalence, divides E into equivalence classes.)

To actually construct a code E' requires that a procedure be developed for selecting one codeword from each ρ-equivalence class. In principle one can use the following procedure. Choose a codeword from E and see if any of its cyclic shifts are equal to previously selected codewords of E'. If so, reject the codeword; otherwise select it as a member of E'. However, if the number of codewords in E is large, the computation time for this procedure may be prohibitive. Therefore, one would like to find a systematic procedure for selecting the members of E'. Such a procedure is presented in this correspondence. A simple modification of the procedure allows one to construct codes that have comma freedom[1] of degree $n - 2k$ and are therefore synchronizable [3].

Solomon [2] suggested a different technique for modifying RS codes to obtain synchronizable codes. His technique yielded codes with comma freedom of degree $n - 2m$, where m is the smallest integer greater than $k + 1$ and relatively prime to n. His codes are easier to implement than those developed in [1]. However, for some values of n and k the degree of comma freedom achievable with those codes may be significantly higher than that achievable with Solomon's codes.

II. Notations and Definitions

In this section we introduce the notation and definitions necessary for the exposition in the following section.

Denote by (a,b) the greatest common divisor of the integers a and b. Then $(a,b) = 1$ if and only if a and b are relatively prime. The notation $a \mid b$ means a divides b.

For any finite set of integers $\{I\}$ denote the least common multiple of the members of $\{I\}$ by lcm $\{I\}$.

Let α be a primitive element of $GF(q)$, the finite field of q elements, and let p_i be the period of α^i, where

$$p_i = \frac{q - 1}{(q - 1, i)}.$$

For each i, $1 \le i \le q - 2$, $\{(\alpha^i)^m, m = 0, 1, \cdots, p_i - 1\}$ is a subgroup of the group H of nonzero elements of $GF(q)$. Denote this subgroup by H_i and denote the distinct cosets of H_i in H by $H_i, \alpha H_i, \cdots, \alpha^{c_i} H_i$, where $c_i = (q - 1, i) - 1$.

Let k be any integer such that $1 \le k < q - 1$, and define the sets

$$P = \{i: 1 \le i \le k \text{ and } (q - 1, i) = 1\}$$

and

$$\bar{P} = \{i: 1 \le i \le k \text{ and } i \notin P\}.$$

In particular, if the number of integers in P is N, then let

$$P = \{i_1 = 1, i_2, \cdots, i_N\}$$

and let

$$\bar{P} = \{i_{N+1}, i_{N+2}, \cdots, i_K, i_{K+1} = 0\}$$

For any vector $x = (x_1, \cdots, x_n)$ with components in an arbitrary field, define the index set of x, denoted by $s(x)$, as

$$s(x) = \{l: x_l \neq 0\}.$$

III. Construction Procedure

A kth-order near-orthogonal code E' can be obtained from an $(n, k + 1)$ RS code E as follows. Let α be a primitive element of

[1] Comma freedom of degree r means that if $(a_1, \cdots a_n)$ and $(b_1 \cdots b_n)$ are two codewords, any subsequence of n consecutive digits of the sequence $(a_2 \cdots a_n, b_1 \cdots b_{n-1})$ is at distance r or more from any codeword.