Proceedings of 24th Conference
on Decision and Control
Ft. Lauderdale, FL • December 1985

TP3 - 3:00

OPTIMUM MULTI-USER DETECTION WITH POINT-PROCESS OBSERVATIONS

Sergio Verdú

Department of Electrical Engineering
Princeton University

1. Introduction

This paper is concerned with the derivation and analysis of optimum demodulators for optical digital communication systems where several users modulate coherent light of the same frequency. The information transfer is represented by a discrete-input multiple-access channel which outputs a scalar conditionally Poisson point-process modeling an ideal photodetector counting process. The rate of the counting process put out by the photodetector is a function of the instantaneous incident power of the sum of the electric fields modulated by several active users. If the users cooperate to maintain symbol synchronism among them and the light sources are modulated by nonoverlapping signal waveforms (time-division multiplexing), then the optimum demodulation of the information transmitted by each user is decoupled and can be carried out by a bank of optimum single-user detectors, each of which performs a weighted comparison of the product of the rate values at the arrival times under each hypothesis [1]. On the contrary, if the modulated signals overlap in the time-domain (e.g. because of lack of synchronism among the users) then single-user decision rules are no longer optimum in terms of error probability.

In this paper, we consider arbitrary signal constellations and both asynchronous and symbol-epoch synchronous Poisson multiple-access channels of the following types:

A) Additive-light Poisson Multiple-access Channel

Assuming that the incident light wave is planar and that the transmission medium is linear, the instantaneous complex value of the electric field at the photodetector screen depends on the transmitted sequence of symbols, $\mathbf{b}$, via

$$E_t(\mathbf{b}) = \sum_{k=1}^{K} e^{j\omega t + j\psi_k} \sum_{i=-M}^{M} s_k(t - iT - \tau_k;\, b_k(i)) \tag{1}$$

where K is the number of active users, $\tau_k \in [0, T)$, $\psi_k \in [0, 2\pi)$ and $b_k(i) \in A_k$ are the delay, phase and i^{th} symbol of the k^{th} user. The alphabet of each user A_k, is assumed to be finite and the waveforms of each user are such that $s_k(t; b) = 0$ for $t \notin [0, T]$, $b \in A_k$, $k = 1,...,K$. The ideal photodetector response to the impinging field is a Poisson counting process whose rate is given by

$$\rho_t(\mathbf{b}) = \alpha \mid E_t(\mathbf{b}) + \mid E \mid e^{j\phi_t} \mid^2 + \beta \tag{2}$$

where α, β, and $|E|$ are constants of the photodetector. The cases $|E| = 0$ and $|E| >> 1$ correspond to direct-detection and (homodyne, if $\phi_t = \omega t + \theta$) coherent-detection systems respectively ([2]). In the case of direct-detection this multiple-access model has been used by Chan [3] in the evaluation of asynchronous coded systems with suboptimum demodulation, and by Narayan and Snyder [4] in the derivation of optimal signal sets for bit-epoch synchronous systems.

The particular case where $A_i = A_j$, $s_i(t; b) = s_j(t; b)$, and $\psi_i = \psi_j$ for all $i \neq j$, $b \in A_i$, $t \in [0, T]$, and $\tau_{i+1} - \tau_i = T/K$ corresponds to the problem of single-user transmission through a dispersive medium where each symbol suffers suffers the interference of $K - 1$ signals. Maximum likelihood sequence detectors of point-processes whose rates are transmitted through finite-memory randomly-dispersive channels have been obtained by Morley and Snyder [5]. However, most optimum detectors proposed for intersymbol interference problems with point-process observations assume a particular case of the following model.

B) Additive-rate Poisson Multiple-access Channel

In this model, it is assumed that each user produces a Poisson point-process and that the receiver observes the sum of these processes plus, possibly, a dark current component. Therefore, the dependence of the rate on the transmitted sequence can be written as

$$\rho_t(\mathbf{b}) = \beta + \sum_{k=1}^{K} \sum_{i=-M}^{M} s_k(t - iT - \tau_k;\, b_k(i)) \tag{3}$$

While this model is the direct analogous to the Gaussian multiple-access channel, its scope of applicability in modeling optical multi-user systems is much narrower than the additive-light model because of the physical mechanism that converts the received optical energy into an electrical signal for subsequent processing. It can be used to model single-output photodetectors with independent screens for each user, and it has been employed in the analysis of single-user detection in multiple-access free-space applications [6]. In the particular case of intersymbol interference the additive-rate channel has been used extensively in modeling direct-detection dispersive optical fiber systems ([7, 8]). Even though these problems fall naturally into the additive-light channel model, this has been done previously both for analytical convenience (analogous schemes to those for additive Gaussian channels can be employed - especially in the case of linear modulation) and because in certain cases where the fiber modes are uncoupled or randomly coupled, Personick [9] showed that the superposition-of-power approximation may be accurate enough.

Optimum multi-user detectors according to the criteria of minimum probability of error and maximum likelihood sequence detection are derived in Section 2. In the asynchronous case, these detectors base every symbol decision on the whole observation interval and decompositions of the likelihood function suitable for the implementation of dynamic programming algorithms are found. In Section 3, lower and upper bounds on the minimum encoded bit error rate achievable by multi-user detectors are obtained. The particularization of these results to the intersymbol interference case solves an open problem: the error rate analysis of optimum single-user direct detection systems for dispersive optical fibers.

2. Optimum multi-user detectors

The multi-user detectors for point-process observations obtained in this section carry out the simultaneous demodulation of the information transmitted by all active users in the channel. The global and local optimality criteria of maximum likelihood sequence detection and minimum symbol error probability detection, respectively, are studied. The point-process observations are denoted by $\{dr_t,\ t_p \le t < t_f\}$ where, given the sequence $\mathbf{b} = \{b(i) \in A_1 \times \cdots \times A_K,\ -M \le i \le M\}$, r_t is a Poisson counting process with rate equal to $\rho_t(\mathbf{b})$ and $t_p = -MT + \tau_1$, $t_f = (M+1)T + \tau_K$. The conditional sample function density of the observations is given by (e.g. [10])

$$P[\{r_t,\ t_p \le t < t_f\} \mid \mathbf{b}] = exp\,[\int_{t_p}^{t_f} \ln \rho_t(\mathbf{b})\ dr_t - \int_{t_p}^{t_f} \rho_t(\mathbf{b})dt], \quad (4)$$

where the first integral in the right-hand side of (4) is a Riemann-Stieltjes integral:

$$\int_{t_p}^{t_f} \ln \rho_t(\mathbf{b})\ dr_t = \sum_{i=1}^{N} \ln \rho_{a_i}(\mathbf{b}) \quad (5)$$

where $\{a_1, \cdots, a_N\}$ is the realization of the observed point-process in the interval $[t_p, t_f]$. Since the transmitted symbols are assumed equiprobable and independent a priori, the (globally optimum) maximum likelihood sequence detector chooses a sequence $b_1^*(-M) \in A_1, \cdots, b_K^*(M) \in A_K$ such that

$$\mathbf{b}^* \in \arg \max_{\mathbf{b} \in D} exp\,(-\int_{t_p}^{t_f} \rho_t(\mathbf{b})dt)\ \prod_{i=1}^{N} \rho_{a_i}(\mathbf{b}) \quad (6)$$

where D is the set of all possible sequences. The (locally optimum) minimum symbol error probability detector selects the sequence of symbols that maximize the posterior marginals

$$P_k(j, d) = \sum_{\substack{\mathbf{b} \in D \\ s.t. \\ b_k(j) = d \in A_k}} exp\,(-\int_{t_p}^{t_f} \rho_t(\mathbf{b})dt)\ \prod_{i=1}^{N} \rho_{a_i}(\mathbf{b});\ k = 1,...K; \quad (7)$$

In the asynchronous case, the symbols transmitted during nonoverlapping slots are no longer independent a posteriori, and the one-shot approach does not result in optimum decisions. Optimum demodulation is accomplished by dynamic programming decision algorithms of the forward type (Viterbi) in maximum likelihood sequence detection and of the backward-forward type in minimum error probability detection. Dynamic programming can be used in the maximization of (6) and (7) due to the sequential dependence of the conditional sample function density (4) on the components of the sequence $\mathbf{b}$.

The observation interval $[t_p, t_f)$ can be partitioned into subintervals of the form $[t_i, t_{i+1})$ where $t_i = \eta(i)T + \tau_\kappa(i)$, $i = -MK+1,...,(M+1)K$, and $t_f = t_{(M+1)K+1}$. (The integers $\eta(i)$ and $\kappa(i)$ are defined by the modulo-K decomposition of $i = \eta(i)K + \kappa(i)$, $\kappa(i) = 1,...,K$.) In each subinterval, the rate depends at most on K symbols, and consecutive subintervals depend on the same symbols but one; this is because the instants t_i correspond to the epoch end points of all transmitted symbols and because in the multiple-access models (2) and (3) the rate depends instantaneously on the transmitted signals. Hence we can define the following functions.

$$\sigma_t^i(x_i, u_i) = \rho_t(\mathbf{b}), \qquad t_i \le t < t_{i+1},\ i = -MK+1,...,(M+1)K \quad (8)$$

where $u_i = b_{\kappa(i)}(\eta(i)) \in A_{\kappa(i)}$ and

$$x_{i+1} = [x_{i+1}^1 \cdots x_{i+1}^{K-1}] = f(x_i, u_i) = [x_i^2 x_i^3 \cdots x_i^{K-1} u_i]^T .$$

This implies that the conditional sample function density (4) admits the following product decomposition

$$P[\{r_t,\ t_p \le t < t_f\} \mid \mathbf{b}] = \prod_{i=1-MK}^{(M+1)K-1} L_i(x_i, u_i) \quad (9)$$

where

$$L_i(x, u) = exp\,[\int_{t_i}^{t_{i+1}} \ln \sigma_t^i(x, u)\ dr_t - \int_{t_i}^{t_{i+1}} \sigma_t^i(x, u)dt] \quad (10)$$
$$= \prod_{t_i \le a_j < t_{i+1}} \sigma_{a_j}^i(x, u)\ \ exp\,[\int_{t_i}^{t_{i+1}} \sigma_t^i(x, u)dt].$$

It is easy to see that the maximization of (9) with respect to $\mathbf{b}$, or equivalently $\{u_i,\ i = 1-MK,...,(M+1)K-1\}$, can be carried out by forward dynamic programming. To that end one defines the cost-to-arrive function

$$J^{i+1}(x) = \max_{\substack{x_i \in A_{\kappa(i+1)} \times \cdots \times A_{\kappa(i-1)} \\ s.t.\ there\ exists \\ u \in A_{\kappa(i)},\ f(x_i, u) = x}} J^i(x_i)L_i(x_i, u);\ J^{1-MK} = 1, \quad (11)$$

and once $J^{(M+1)K}$ is obtained, the maximizing sequence of symbols is recovered by backtracking the optimum path of states generated by the recursion (11). As it is well known, in real time applications near-optimum performance can be achieved by the Viterbi algorithm which makes decisions after an adequately chosen fixed lag.

A related procedure can be used to obtain the marginals (7) in minimum error probability detection. Now instead of maximizing with respect to all sequences we have to sum the likelihood function over all sequences having a fixed symbol; this can be done by defining independent forward and backward recursions [11]:

$$F^{i+1}(x) = \sum_{\substack{x_i \in A_{\kappa(i+1)} \times \cdots \times A_{\kappa(i-1)} \\ s.t.\ there\ exists \\ u \in A_{\kappa(i)},\ f(x_i, u) = x}} F^i(x_i)L_i(x_i, u) \quad (12)$$

and

$$B^i(x) = \sum_{u_i \in A_{\kappa(i)}} B^{i+1}(f(x, u_i))L_i(x, u_i) \quad (13)$$

Then it is straightforward to show that the posterior marginal distributions of the transmitted symbols are given by

$$P_{\kappa(i)}(\eta(i), d) = \sum_{\substack{x_{i+1} \in A_{\kappa(i+2)} \times \cdots \times A_{\kappa(i)} \\ s.t.\ there\ exists \\ x_i,\ x_{i+1} = f(x_i, d)}} F^{i+1}(x_{i+1})\ B^{i+1}(x_{i+1}). \quad (14)$$

3. Probability of Error

In this section we analyze the uncoded minimum error probability achievable with additive-light and additive-rate Poisson multiple-access channels. The approach is to find upper and lower bounds on the finite-horizon error probability for each user and for arbitrary relative delays, i.e.,

$$P_\kappa(i) = P[b_\kappa(i) \ne \hat{b}_\kappa(i)], \quad (15)$$

where $b_\kappa(i) \in A_k$ is the transmitted symbol and $\hat{b}_\kappa(i)$ is the most likely symbol given the observation of the point-process realization on the interval, I^M, corresponding to the symbols $b_n(j)$, $n = 1,...,K$, $j = -M,...,M$; i.e.,

$$\hat{b}_\kappa(i) \in \arg \max_{b \in A_k} P[b_\kappa(i) = b \mid dr_t \sim \rho_t(\mathbf{b})]. \quad (16)$$

For the sake of notational convenience in presenting our results, we focus attention on the case where all users employ binary alphabets (and without loss of generality we label $A_1 = \cdots = A_K = \{-1, 1\}$). Bounds on the minimum bit error rate are obtained by direct passage to the limit $M \to \infty$ of the bounds on $P_\kappa(i)$. Except in particular cases, no explicit expressions are known for the error probability of binary one-shot single-user detection with point-process observations; in general, one must resort either to bounding or to numerical approximation (see Appendix).

A fortiori, explicit expressions cannot be expected for the multi-user error probabilities. However, the bounds derived below are expressed in terms of binary one-shot error probabilities, reducing the analysis to one of single-user performance. All the results in this section are given for arbitrary relative delays; however, even the particular case of bit-epoch synchronous users evidences the role of all the basic ideas in our approach. Hence, on a first reading it may be helpful to assume $M = 0$, $\tau_k = 0$, $k = 1,...,K$, and to regard each sequence as a single K-dimensional vector.

The first lower bound is the single-user bound on the error probability of the additive-rate model.

Proposition 1: Assume that the additive-rate model (3) is in effect. Then, the minimum error probability of the k^{th} user is lower-bounded by

$$P_k(i) \geq \frac{1}{2} P\left[\int_0^T ln(d(t))\, dr_t - \int_0^T g(t)\, dt > 0 \mid dr_t \sim \beta + s_k(t;-1)\right] + \frac{1}{2} P\left[\int_0^T ln(d(t))\, dr_t - \int_0^T g(t)\, dt \leq 0 \mid dr_t \sim \beta + s_k(t;1)\right] \quad (17)$$

where $d(t) = [s_k(t;1) + \beta]/[s_k(t;-1) + \beta]$, and $g(t) = s_k(t;1) - s_k(t;-1)$

Proof: Suppose that $\{r_t^0, r_t^j, j = 1,...,K\}$ are independent point-processes generated with rates β and $\sum_{n=-M}^{M} s_j(t - nT - \tau_j; b_j(n))$ $j = 1,...,K$, respectively. Then $P_\kappa(i)$ is the minimum error probability of a detector of the i^{th} bit of the κ^{th} user which observes $\{r_t = \sum_{j=0}^{K} r_t^j, t \in I^M\}$. But, $P_\kappa(i)$ cannot be lower than the error probability, say $\hat{P}_\kappa(i)$, of a detector which in addition to observing $\{r_t, t \in I^M\}$, observes $\{r_t^j\, t \in I^M\}$, $j = 1,...,K$, $j \neq \kappa$. Now since $\{r_t^j, t \in I^M\}$ is independent of $\{r_t^\kappa, t \in I^M\}$, the optimum use of the side-information is to subtract it from $\{r_t, t \in I^M\}$. Hence, it follows that $\hat{P}_\kappa(i)$ is the single-user error probability, which is given by the right-hand side of (17) (e.g. [10]).

∇

The above bound simply says that the existence of other active users in the channel cannot improve probability of error. Somewhat surprisingly, this bound does not hold for the additive-light multiple-access channel. To see this, consider the following direct-detection example in which the single-user error probability is greater than the minimum error probability achievable by adding one extra user:

Example: Suppose that the amplification and dark current rate of the photodetector are $\alpha = 1$ and $\beta = \frac{1}{2}$, respectively, and the signal waveforms of user 1 are

$$s_1(t;b) = \begin{cases} 1.00 & b = +1 \\ & \qquad 0 \leq t < T = 1. \\ 0.50 & b = -1 \end{cases} \quad (18)$$

Therefore, the rate is equal to (via (1) and (2) with $|E| = 0$) $\rho_t(b = 1) = 1.5$ and $\rho_t(b = -1) = 0.75$, and it is straightforward to see that the optimum detector decides $\hat{b} = -1$ if the number of arrivals in $[0,1)$ is equal to either 0 or 1. Then, the single-user error probability is equal to

$$P_1[1 \text{ user}] = (2.5\, e^{-1.5} + 1 - 1.75 e^{-0.75})/2 = 0.3656. \quad (19)$$

Suppose now that we add an extra user whose bit-epoch coincides with that of user 1 ($\tau_1 = \tau_2$) and whose signal waveforms are

$$s_2(t;b) = \begin{cases} 1.00 & b = +1 \\ & \qquad 0 \leq t < T = 1. \\ 0.99 & b = -1 \end{cases} \quad (20)$$

Then, the optimum decision rule for user 1 is to choose $\hat{b}_1 = -1$ if the number of arrivals in $[0,1)$ is less than or equal to $N(\psi_1 - \psi_2)$, where

$$N(\phi) = \begin{cases} 3 & \text{if } 0.63 \leq \cos\phi \\ 2 & \text{if } -0.07 < \cos\phi < 0.63 \\ 1 & \text{if } -0.74 \leq \cos\phi < -0.07 \\ 0 & \text{if } \qquad \cos\phi < -0.74. \end{cases}$$

It can be checked that if the phases of both users are such that $cos(\psi_1 - \psi_2) \geq 0.35$, then, $0.3197 \leq P_1[2\ users] < P_1[1\ user]$.

∇

In the above counterexample, the reason for the error probability improvement in a large region of carrier phases can be traced back to the nonlinear dependency of the rates on each user's signal (2). For this reason, the presence of an extra user of adequate strength may actually improve the rate discrimination of the user of interest.

We turn to the derivation of a pair of lower bounds that hold for both multiple-access channels. The first step is to define the average error probability, $P^{av}(\epsilon)$, of an error sequence $\epsilon \in E = \left\{\epsilon = \{\epsilon(i) \in \{-1,0,1\}^K, i = -M,...,M; \epsilon(j) \neq 0 \text{ for some } j\}\right\}$. Note that in contrast to the Gaussian case, the error probability of the binary test between two sequences - **b** and **d** - is not completely specified by the error sequence $\epsilon = \frac{1}{2}(\mathbf{b} - \mathbf{d})$, i.e., in general

$$P_e[\mathbf{b}, \mathbf{b} - 2\epsilon] \neq P_e[\mathbf{d}, \mathbf{d} - 2\epsilon] \triangleq \frac{1}{2} P[\Lambda(\omega \mid \mathbf{d} - 2\epsilon; \mathbf{d}) \geq 1 \mid \mathbf{d} \text{ transmitted}] + \frac{1}{2} P[\Lambda(\omega \mid \mathbf{d}; \mathbf{d} - 2\epsilon) > 1 \mid \mathbf{d} - 2\epsilon \text{ transmitted}], \quad (21);$$

where the likelihood ratio is denoted by

$$\Lambda(\{dr_t, t \in I^M\} \mid \mathbf{b}; \mathbf{d}) = \frac{P[\{dr_t, t \in I^M\} \mid \mathbf{b}]}{P[\{dr_t, t \in I^M\} \mid \mathbf{d}]}.$$

Hence, we define the average error probability of ϵ as

$$P^{av}(\epsilon) = E[P_e[\mathbf{b}, \mathbf{b} - 2\epsilon]] \quad (22)$$

where the expectation is over $D(\epsilon) \triangleq \{\mathbf{d} \in D, \mathbf{d} - 2\epsilon \in D\}$, i.e., the ensemble of transmitted sequences congruent with the error sequence ϵ (if $\epsilon_n(j) \neq 0$ then $b_n(j) = \epsilon_n(j)$, $n = 1,...,K$; $j = -M,...,M$). Note that $P_e[\mathbf{b}, \mathbf{b} - 2\epsilon]$ is independent of the symbols in **b** which do not overlap with nonzero components of ϵ. Therefore, we need not take into account all the transmitted sequences in order to evaluate (22), and we can write

$$P^{av}(\epsilon) = \frac{1}{|\Upsilon(\epsilon)|} \sum_{b \in \Upsilon(\epsilon)} P_e[\mathbf{b}, \mathbf{b} - 2\epsilon] \quad (23)$$

where

$$\Upsilon(\epsilon) = \{\mathbf{b} \in D, \epsilon \in A(\mathbf{b}) \text{ and } b_j(n) = 1 \text{ if } nK + j < p(\epsilon) - K + 1 \text{ or } nK + j > f(\epsilon) + K - 1\}$$

and $p(\epsilon)$ and $f(\epsilon)$ are the first and last nonzero components of ϵ; i.e., if $u_m = \epsilon_{\kappa(m)}(\eta(m))$ then $u_{p(\epsilon)} \neq 0$, $u_{f(\epsilon)} \neq 0$ and $u_m = 0$ for $m < p(\epsilon)$ or $m > f(\epsilon)$. Consequently, only $|\Upsilon(\epsilon)| = 2^{(2K + f(\epsilon) - p(\epsilon) - 1 - w(\epsilon))}$ sequences need to be considered for the computation of (22) where $w(\epsilon)$ is the number of nonzero components in ϵ.

Proposition 2: For each sequence $\mathbf{b} \in D$, denote the worst-case pairwise error probability with any other sequence differing in the i^{th} bit of the κ^{th} user by

$$R_\kappa(\mathbf{b}, i) = \max_{\substack{\mathbf{d} \in D \\ b_\kappa(i) \neq d_\kappa(i)}} P_e[\mathbf{b}, \mathbf{d}]. \tag{24}$$

Let $Z_\kappa(i) = \{\epsilon \in E, \epsilon_\kappa(i) \neq 0\}$ and denote the largest error probability between two sequences whose difference belongs to $Z_\kappa(i)$ by

$$\overline{R}_\kappa(i) = \max_{\mathbf{b} \in D} R_\kappa(\mathbf{b}, i). \tag{25}$$

Then, the minimum error probability of the i^{th} bit of the κ^{th} user is lower bounded by

$$P_\kappa(i) \geq \max_{\epsilon \in Z_\kappa(i)} 2^{1-w(\epsilon)} P^{av}(\epsilon) \tag{26}$$

and by

$$P_\kappa(i) \geq P[R_\kappa(\mathbf{b}, i) = \overline{R}_\kappa(i)]\, \overline{R}_\kappa(i) \tag{27}$$

where $P[R_\kappa(\mathbf{b}, i) = \overline{R}_\kappa(i)]$ is the a priori probability that the transmitted sequence attains the maximum in (25).

Proof: Let $p(\omega \mid F)$ denote the conditional sample function density of the observables $\omega = \{dr_t, t \in I^M\}$ with respect to the measure υ induced by the unit-rate Poisson point-process. The minimum error probability of the i^{th} bit of the κ^{th} user attainable by a detector which observes $\{dr_t, t \in I^M\}$, can be expressed as

$$P_\kappa(i) = \frac{1}{2} \int_\Omega \min\{p(\omega \mid b_\kappa(i) = 1),\, p(\omega \mid b_\kappa(i) = -1)\} d\upsilon \tag{28}$$

Let $D^\pm = \{\mathbf{b} \in D, b_\kappa(i) = \pm 1\}$; then for every $\omega \in \Omega$, and every one-to-one function $h: D^+ \rightarrow D^-$, we have

$$\begin{aligned} &\frac{1}{2} \min\{p[\omega \mid b_\kappa(i) = 1],\, p[\omega \mid b_\kappa(i) = -1]\} \\ &= \min\{\sum_{b \in D^+} p[\omega \mid \mathbf{b}] \cdot P[\mathbf{b}],\, \sum_{b \in D^-} p[\omega \mid \mathbf{b}] \cdot P[\mathbf{b}]\} \\ &\geq C \sum_{b \in D^+} \min\{p[\omega \mid \mathbf{b}],\, p[\omega \mid h(\mathbf{b})]\}, \end{aligned} \tag{29}$$

where $C = P[\mathbf{b}], \mathbf{b} \in D$, and the last inequality follows because every term in the summation of the right-hand side of (29) is dominated by a corresponding term in both $\sum_{b \in D^+} p[\omega \mid \mathbf{b}]$ and $\sum_{b \in D^-} p[\omega \mid \mathbf{b}]$. Uniting (28) and (29), we obtain a bound on $P_\kappa(i)$ in terms of the error probabilities of binary sequence-selection tests. This bound depends on the choice of the function $h: D^+ \rightarrow D^-$. For every $\epsilon \in E$ such that $\epsilon_k(i) = 1$, we can define $h(\mathbf{b}) = \mathbf{b} - 2\epsilon$ over the subset of D^+ that contains all the sequences $\mathbf{b}$ that are congruent with ϵ, i.e., $D^+ \cap D(\epsilon)$. Hence, it follows that, for every $\epsilon \in E$ such that $\epsilon_\kappa(i) = 1$, we have

$$P_\kappa(i) \geq C \sum_{\mathbf{b} \in D^+ \cap D(\epsilon)} P_e[\mathbf{b}, \mathbf{b} - 2\epsilon] = 2^{1-w(\epsilon)} P^{av}(\epsilon), \tag{30}$$

and noticing that $2^{-w(\epsilon)} P^{av}(\epsilon) = 2^{-w(-\epsilon)} P_{av}(-\epsilon)$, the bound (26) follows.

In order to obtain (27) it is possible to use an argument similar to that of Forney [12]. If the transmitted sequence, $\mathbf{b}$, is such that $R_\kappa(\mathbf{b}, i) = \overline{R}_\kappa(i)$, then the detector is told by a genie that the true sequence is either $\mathbf{b}$ or

$$\mathbf{d}^*(\mathbf{b}) \in \arg \max_{\substack{\mathbf{d} \in D \\ b_\kappa(i) \neq d_\kappa(i)}} P_e[\mathbf{b}, \mathbf{d}].$$

In that case, the detector must perform a binary test between $\mathbf{b}$ and $\mathbf{d}^*(\mathbf{b})$ whose error probability is $P_e[\mathbf{b}, \mathbf{d}^*(\mathbf{b})] = \overline{R}_\kappa(i)$. Note that

$$R_\kappa(d^*(\mathbf{b}), i) = \max_{\substack{\mathbf{c} \in D \\ c_\kappa(i) = b_\kappa(i)}} P_e[\mathbf{c}, d^*(\mathbf{b})] \geq \overline{R}_\kappa(i) \geq R_\kappa(d^*(\mathbf{b}), i), \tag{31}$$

so $R_\kappa(d^*(\mathbf{b}), i) = \overline{R}_\kappa(i)$. (If this were not true, then the detector could decide without error, because the transmitted sequence would have lower worst-case error probability.)

∇

The upper bound on error probability of the κ^{th} user is equal to the sum, over all *indecomposable* sequences that affect the κ^{th} user, of $2^{-w(\epsilon)} P^{av}(\epsilon)$ (cf. the lower bound (26) which is equal to the maximum of such quantities - and coincides with the upper bound in the single-user case). An error sequence $\epsilon \in E$ is *decomposable* into $\epsilon^1 \in E$ and $\epsilon^2 \in E$ if

i) $\epsilon = \epsilon^1 + \epsilon^2$

ii) if $\epsilon_n(j) = 0$ then $\epsilon_n^1(j) = \epsilon_n^2(j) = 0$,

iii) there exists a covering $I_1^M \cup I_2^M = I^M$ such that for $t \in I_l^M, l = 1,2$

$$\Delta_l(\epsilon^l) \triangleq \sum_{k=1}^{K} \sum_{i=-M}^{M} [s_k(t - \tau_k - iT; \epsilon_k^l(i)) - s_k(t - \tau_k - iT; -\epsilon_k^l(i))]\, e^{j\psi_k} =$$

Proposition 3: The minimum error probability of the i^{th} bit of the κ^{th} user is upper bounded by

$$P_\kappa(i) \leq \sum_{\epsilon \in Q_\kappa(i)} 2^{-w(\epsilon)} P^{av}(\epsilon) \tag{32}$$

where $Q_\kappa(i) = \{\epsilon \in Z_\kappa(i);\ \epsilon \text{ is indecomposable}\}$.

Proof: The right-hand side of (32) is an upper bound on the error probability achieved by the multi-user maximum likelihood sequence detector. In order to prove this bound we need not assume any specific decision algorithm; however, it will be convenient to place an assumption on the tie-breaking rule in the maximization of the posterior likelihood (depending on the rate, ties may occur with nonzero probability): the detector outputs one of the most likely sequences whose i^{th} symbol of the κ^{th} user is equal to -1 (if such a sequence exists).

Consider the following inclusions between events in the probability space where the transmitted $\mathbf{b} \in D$ and the observed point-process realization $\omega = \{dr_t, t \in I^M\} \in \Omega$ are defined:

$$\begin{aligned} &\{(\omega, \mathbf{b}) \in \Omega \times D:\ b_\kappa(i) \neq b_\kappa^*(i)\} \\ &= \bigcup_{\epsilon \in Z_\kappa^+(i)} \{(\omega, \mathbf{b}) \in \Omega \times D(\epsilon):\ \mathbf{b} - 2\epsilon \in \arg\max_{\mathbf{d} \in D} P[\omega \mid \mathbf{d}]\} \\ &\quad \cup \{(\omega, \mathbf{b}) \in \Omega \times D^-:\ \arg\max_{\mathbf{d} \in D} P[\omega \mid \mathbf{d}] \subset D^+\} \\ &\subset \bigcup_{\epsilon \in Z_\kappa^+(i)} \{(\omega, \mathbf{b}) \in \Omega \times D(\epsilon):\ \text{if } d_\kappa(i) = 1 \text{ then } \Lambda(\omega \mid \mathbf{b} - 2\epsilon; \mathbf{d}) \geq 1\} \\ &\quad \bigcup_{\epsilon \in Z_\kappa^-(i)} \{(\omega, \mathbf{b}) \in \Omega \times D(\epsilon):\ \text{if } d_\kappa(i) = -1 \text{ then } \Lambda(\omega \mid \mathbf{b} - 2\epsilon; \mathbf{d}) > 1\} \\ &\subset \bigcup_{\epsilon \in Q_\kappa^+(i)} \{(\omega, \mathbf{b}) \in \Omega \times D(\epsilon):\ \Lambda(\omega \mid \mathbf{b} - 2\epsilon; \mathbf{b}) \geq 1\} \\ &\quad \bigcup_{\epsilon \in Q_\kappa^-(i)} \{(\omega, \mathbf{b}) \in \Omega \times D(\epsilon):\ \Lambda(\omega \mid \mathbf{b} - 2\epsilon; \mathbf{b}) > 1\}, \end{aligned} \tag{33}$$

where $\mathbf{b}^*$ is the sequence selected by the detector, $Z_\kappa^\pm(i) = \{\epsilon \in E:\ \epsilon_\kappa(i) = \pm 1\}$ and $Q_\kappa^\pm(i) = Q_\kappa(i) \cap Z_\kappa^\pm(i)$. The equality in (33) follows because according to the above assumption on the tie-breaking rule, if $b_\kappa^*(i)$ is erroneous then the transmitted sequence is such that either $b_\kappa(i) = -1$ and it differs in that position from all most likely sequences or $b_\kappa(i) = +1$ and at least one most likely sequence differs in that position. The first inclusion in (33) follows because if $\mathbf{b} - 2\epsilon \in \arg\max_{\mathbf{d} \in D} P[\omega \mid \mathbf{d}]$ then $\Lambda(\omega \mid \mathbf{b} - 2\epsilon; \mathbf{d}) \geq 1$

for all $\mathbf{d} \in D$ and if $\arg\max_{\mathbf{d} \in D} P[\omega \mid \mathbf{d}] \subset D^+$ then each of the most likely sequences has strictly greater likelihood than the sequences in D^-. The main part of the present proof is to show the last inclusion in (33). First we show that for every $\epsilon \in Z_\kappa^+(i)$ there exists $\epsilon' \in Q_\kappa^+(i)$ such that

$$\{(\omega, \mathbf{b}) \in \Omega \times D(\epsilon):\ \text{if } \mathbf{d} \in D^+ \text{ then } \Lambda(\omega \mid \mathbf{b} - 2\epsilon;\ \mathbf{d}) \geq 1\} \subset \{(\omega, \mathbf{b}) \in \Omega \times D(\epsilon'):\ \Lambda(\omega \mid \mathbf{b} - 2\epsilon';\ \mathbf{b}) \geq 1\}. \tag{34}$$

Fix $\epsilon \in Z_\kappa^+(i)$. If $\epsilon \in Q_\kappa^+(i)$ then it is clear that $\epsilon' = \epsilon$ satisfies (34) because $D(\epsilon) \subset D^+$. If $\epsilon \notin Q_\kappa^+(i)$ then ϵ is decomposable into $\epsilon' + \epsilon''$ such that $\epsilon' \in Q_\kappa^+(i)$. This follows from the associativity of the operation of sequence decomposition, i.e., if ϵ^* is decomposable into $\epsilon^a + \epsilon^b$ and ϵ^a is decomposable into $\epsilon^c + \epsilon^d$, then ϵ^* is decomposable into $\epsilon^c + (\epsilon^d + \epsilon^b)$. In order to show this property (which does not hold for the decomposition of sequences used in the Gaussian multiple-access problem [13]) suppose that $I_a \cup I_b$ and $I_c \cup I_d$ are the respective coverings of I^M required by the definition of decomposability. Then, for $t \in I_a \cap I_d$ we have $\Delta_t(\epsilon^a) = \Delta_t(\epsilon^c) + \Delta_t(\epsilon^d)$ and $\Delta_t(\epsilon^a) = \Delta_t(\epsilon^d) = 0$. Hence if $t \in I_c \cup (I_a \cap I_d)$ then $\Delta_t(\epsilon^c) = 0$. On the other hand, if $t \in I_b \cap I_d$ then $\Delta_t(\epsilon^b + \epsilon^d) = \Delta_t(\epsilon^b) + \Delta_t(\epsilon^d) = 0$. Therefore, $I^M = (I_c \cup (I_a \cap I_d)) \cup (I_b \cap I_d)$ is a valid covering for the decomposition $\epsilon^* = \epsilon^c + (\epsilon^d + \epsilon^b)$. So in order to decompose ϵ into $\epsilon' + \epsilon''$ such that $\epsilon' \in Q_\kappa^+(i)$, it is possible to take advantage of the associative property to decompose successively the subsequence that includes $\epsilon_\kappa(i) = 1$ until an indecomposable sequence is obtained.

Now if ϵ is decomposable into $\epsilon' + \epsilon''$, then it is straightforward to check both for the additive-light and additive-rate multiple-access channels that for all $t \in I^M$ and $\mathbf{b} \in D(\epsilon)$ the following equations hold.

$$\rho_t(\mathbf{b} - 2\epsilon) - \rho_t(\mathbf{b} - 2\epsilon'') = \rho_t(\mathbf{b} - 2\epsilon') - \rho_t(\mathbf{b}) \tag{35}$$

and

$$\rho_t(\mathbf{b} - 2\epsilon)\, \rho_t(\mathbf{b}) = \rho_t(\mathbf{b} - 2\epsilon')\, \rho_t(\mathbf{b} - 2\epsilon'') \tag{36}$$

Thus, for every $\omega \in \Omega$.

$$\begin{aligned}&\Lambda(\omega \mid \mathbf{b} - 2\epsilon';\ \mathbf{b}) = \\ &exp\left[\int_{I^M} \ln \frac{\rho_t(\mathbf{b} - 2\epsilon')}{\rho_t(\mathbf{b})}\, dr_t - \int_{I^M} (\rho_t(\mathbf{b} - 2\epsilon') - \rho_t(\mathbf{b}))dt\right] = \\ &exp\left[\int_{I^M} \ln \frac{\rho_t(\mathbf{b} - 2\epsilon)}{\rho_t(\mathbf{b} - 2\epsilon'')}\, dr_t - \int_{I^M} (\rho_t(\mathbf{b} - 2\epsilon) - \rho_t(\mathbf{b} - 2\epsilon''))dt\right] = \\ &\Lambda(\omega \mid \mathbf{b} - 2\epsilon;\ \mathbf{b} - 2\epsilon'').\end{aligned} \tag{37}$$

Since $\epsilon_\kappa''(i) = 0$ if $(\omega, \mathbf{b}) \in \Omega \times D(\epsilon)$ belongs to the left-hand side of (34), then $\Lambda(\omega \mid \mathbf{b} - 2\epsilon;\ \mathbf{b} - 2\epsilon'') \geq 1$. Hence, using (37) and $D(\epsilon) \subset D(\epsilon')$, the inclusion in (34) follows. In order to complete the proof of (33) it suffices to show that for each $\epsilon \in Z_\kappa^-(i)$ there exists $\epsilon' \in Q_\kappa^-(i)$ such that

$$\{(\omega, \mathbf{b}) \in \Omega \times D(\epsilon):\ \text{if } \mathbf{d} \in D^- \text{ then } \Lambda(\omega \mid \mathbf{b} - 2\epsilon;\ \mathbf{d}) > 1\} \subset \{(\omega, \mathbf{b}) \in \Omega \times D(\epsilon'):\ \Lambda(\omega \mid \mathbf{b} - 2\epsilon';\ \mathbf{b}) > 1\}; \tag{38}$$

but this can be shown by the argument that led to (34).

It remains to take probabilities of the events at both sides of (33). Using the union bound and $P[\mathbf{b} \in D(\epsilon)] = 2^{-w(\epsilon)}$ (since all sequences in D are equally likely a priori) we obtain

$$\begin{aligned}P_\kappa(i) \leq &\sum_{\epsilon \in Q_\kappa^+(i)} P[(\omega, \mathbf{b}) \in \Omega \times D(\epsilon), \Lambda(\omega \mid \mathbf{b} - 2\epsilon;\ \mathbf{b}) \geq 1] \\ &+ \sum_{\epsilon \in Q_\kappa^-(i)} P[(\omega, \mathbf{b}) \in \Omega \times D(\epsilon), \Lambda(\omega \mid \mathbf{b} - 2\epsilon;\ \mathbf{b}) > 1] \\ = &\sum_{\epsilon \in Q_\kappa^+(i)} E[P[\Lambda(\omega \mid \mathbf{b} - 2\epsilon;\ \mathbf{b}) \geq 1 \mid \mathbf{b} \text{ transmitted}] P[\mathbf{b} \in D(\epsilon)]] \\ &+ \sum_{\epsilon \in Q_\kappa^-(i)} E[P[\Lambda(\omega \mid \mathbf{b} - 2\epsilon;\ \mathbf{b}) > 1 \mid \mathbf{b} \text{ transmitted}] P[\mathbf{b} \in D(\epsilon)]] \\ = &\sum_{\epsilon \in Q_\kappa^+(i)} 2^{-w(\epsilon)-1} \{E[P[\Lambda(\omega \mid \mathbf{b} - 2\epsilon;\ \mathbf{b}) \geq 1 \mid \mathbf{b} \text{ transmitted}] + \\ &E[P[\Lambda(\omega \mid \mathbf{b};\ \mathbf{b} - 2\epsilon) > 1 \mid \mathbf{b} - 2\epsilon \text{ transmitted}]\} \\ &+ \sum_{\epsilon \in Q_\kappa^-(i)} 2^{-w(\epsilon)-1} \{E[P[\Lambda(\omega \mid \mathbf{b} - 2\epsilon;\ \mathbf{b}) > 1 \mid \mathbf{b} \text{ transmitted}] + \\ &E[P[\Lambda(\omega \mid \mathbf{b};\ \mathbf{b} - 2\epsilon) \geq 1 \mid \mathbf{b} - 2\epsilon \text{ transmitted}]\} \\ = &\sum_{\epsilon \in Q_\kappa(i)} 2^{-w(\epsilon)}\, P^{av}(\epsilon).\end{aligned} \tag{39}$$

∇

The idea behind the decomposability of an error sequence, $\epsilon = \epsilon' + \epsilon''$, is that only one component is active at one time, i.e., for every $t \in I^M$ the rate $\rho_t(\mathbf{b} - 2\epsilon)$ is equal to either $\rho_t(\mathbf{b} - 2\epsilon')$ or $\rho_t(\mathbf{b} - 2\epsilon'')$. For example, if the signals corresponding to the nonzero components of ϵ' and ϵ'' do not overlap, (e.g. for synchronous PPM signals, or in the asynchronous case for sequences with $K - 1$ consecutive null components) then the sequence $\epsilon' + \epsilon''$ is decomposable. However, as the following example illustrates decomposability is still possible with overlapping components. Consider a bit-synchronous three-user case where the signals are such that if $0 \leq t \leq T'$ then $s_3(t;1) = s_3(t;-1)$ and if $T' < t \leq T$ then $s_1(t;1) = a \leq b = s_1(t;-1)$ and $s_2(t;1) = b - a$, $s_2(t;-1) = 0$; it is easy to check that the error vector $[\,1\ 1\ 1\,]$ is decomposable into $[\,1\ 1\ 0\,] + [\,0\ 0\ 1\,]$.

The reduction of the above minimum error probability analysis to the case $s_i(t;b) = s_j(t;b)$, $\psi_i = \psi_j$ and $\tau_i = (i-1)T/K$ corresponds to the single-user point-process intersymbol interference problem, which has received considerable attention [8, 5, 7, 14] due to its importance in the analysis of optical direct-detection digital communication systems via dispersive media. Except for the tighter constant achieved in Proposition 2, the particularization of the lower bound in (27) yields the bound by Mazo and Salz [8] which in turn is akin to the bounds derived in [15] and [12] for the Gaussian problem. However, no counterpart was found in the Poisson intersymbol interference problem to the upper bound by Forney [16] on the error probability of maximum likelihood sequence detection for white Gaussian channels, and even with the analytically simpler additive-rate model no success has been reported in the search for an upper bound for optimum direct-detection systems. An answer to this question is given by Proposition 3 for both the additive-light and the additive-rate models. The computation of the bounding series in (32) via the symbolic transfer function technique (e.g. [17]) or the branch-and-bound approach [18] can be accomplished by using the Bhattacharyya bound (A.2 particularized to $s = 1/2$) in the case of direct-detection and the exact expression (A.1) for the binary-test error probability in the case of coherent-detection (see Appendix).

Appendix

The results of the multi-user error probability analysis of Section 3 were given in terms of the probability of error of binary hypothesis tests with Poisson point-process observations. Closed form analysis of binary tests is only possible in certain cases, such as when one of the rate functions dominates the other and their ratio is piecewise constant, or in the problem of coherent detection, in which a strong field is added to the received electromagnetic field prior to photodetection (cf. (2)). In that case the rates of the photoelectron stream put out by the photodetector are uniformly large and the minimum error probability has been shown [19] to be equal to that of a $\sigma = \frac{1}{2}$ white Gaussian noise channel whose signals are equal to the square root of the point-process rates, i.e.,

$$P_e = Q\left(\left(\int [\lambda_1^{1/2}(t) - \lambda_0^{1/2}(t)]^2 dt\right)^{1/2}\right). \qquad (A.1)$$

On the contrary, in the case of direct detection with arbitrary rate functions no closed or semi-closed expressions are known and one must resort to approximation or bounding using the following family of upper and lower bounds.

Proposition 4: Let $\{r_t, t \in I\}$ be a Poisson counting process with rates $\lambda_1(t)$ and $\lambda_0(t)$, $t \in I$ under hypotheses one and zero respectively. If both hypotheses are equiprobable the minimum probability of error is bounded by

$$P_e \le exp(\mu(s)), \qquad (A.2)$$

and

$$P_e \ge \frac{1}{2}(1 - \frac{1}{h})\, exp(\mu(s) - s\mu'(s) + \min\{-s\sqrt{h\mu''(s)}, \mu'(s) - (1-s)\sqrt{h\mu''(s)}\}) \qquad (A.3)$$

for $0 \le s \le 1$ and $0 < h$, and

$$\mu(s) = \int_I [\lambda_1^s(t)\lambda_0^{1-s}(t) - s\lambda_1(t) - (1-s)\lambda_0(t)]dt \qquad (A.4)$$

Proof: The proof of the Chernoff upper bound (A.2) can be found in [10]. The lower bound (A.3) is reminiscent of the bounds by Shannon, Gallager and Berlekamp [20] for discrete memoryless channels. Let $p_i(\omega)$ $i = 0,1$ be the sample function density of the observables $\omega = \{dr_t, t \in I\}$ with respect to the measure v induced by the unit-rate Poisson point-process on the interval I, and define for every $0 \le s \le 1$ the probability density function

$$q(\omega) = p_1^{1-s}(\omega) p_0^{s}(\omega)\, exp(-\mu(s)) \qquad (A.5)$$

Since both hypotheses are equiprobable, we can write using (A.5)

$$P_e = \frac{1}{2}\int_\Omega \min\{p_1(\omega), p_0(\omega)\} d v =$$
$$= \frac{1}{2} exp(\mu(s)) \int_\Omega exp(\min\{-s \ln\Lambda(\omega), (1-s)\ln\Lambda(\omega)\})\, q(\omega) d v \qquad (A.6)$$

where $\Lambda(\omega) = p_1(\omega)/p_0(\omega)$. Now, it is easy to check that for every $h > 0$ and every x,

$$exp(\min\{-sx, (1-s)x\}) \ge exp(\min\{-s(\mu'(s) + \sqrt{h\mu''(s)}), (1-s)(\mu'(s) - \sqrt{h\mu''(s)})\}) \cdot I\{|x - \mu'(s)| \le \sqrt{h\mu''(s)}\} \qquad (A.7)$$

i.e., a double-sided exponential function dominates a pulse of adequate height centered at $\mu'(s)$. Hence, (A.6) and (A.7) imply

$$P_e \ge \frac{1}{2} \exp\left(\mu(s) - s\mu'(s) + \min\{-s\sqrt{h\mu''(s)}, \mu'(s) - (1-s)\sqrt{h\mu''(}\right.$$
$$\int_\Omega I\{|\ln\Lambda(\omega) - \mu'(s)| \le \sqrt{h\mu''(s)}\} q(\omega) d v \qquad (A.8)$$

It remains to bound the integral in the right-hand side of (A.8). To that end, it is straightforward to show that $\mu'(s)$ and $\mu''(s)$ are the mean and variance, respectively, of the log likelihood ratio under $q(\omega)dv$ (see[20]). Therefore, the Chebychev inequality implies that the integral in (A.8) is lower bounded by $(1 - \frac{1}{h})$ and (A.3) follows.

References

1. I. Bar-David, "Communication Under the Poisson Regime," *IEEE Trans. Inform. Theory,* vol. IT-15, no. 1, pp. 31-37, January 1969.
2. R. M. Gagliardi and S. Karp, *Optical Communications,* Wiley, New York, 1976.
3. V. W. S. Chan, "A Multiple-user Random-access Optical Communication System," *Proc. 1979 IEEE International Communications Conf.*, pp. 1.4.1-1.4.5.
4. P. Narayan and D.L. Snyder, "Optimal signal sets for multiple-access channels," *Proc. 1982 Conf. Decision and Control,* pp. 378-379, Dec. 1982.
5. R.E. Morley, Jr. and D.L. Snyder, "Maximum Likelihood Sequence Estimation for Randomly Dispersive Channels," *IEEE Trans. Commun.*, vol. COM-27, no. 6, pp. 833-839, June 1979.
6. G.D. Peterson and C.S. Gardner, "Cross-correlation Interference Effects in Multiaccess Optical Communications," *IEEE Trans. Aeros. and Electro. Syst.*, vol. AES-17, no. 2, pp. 199-207, March 1981.
7. G.J. Foschini, R.D. Gitlin, and J. Salz, "Optimum Direct Detection for Digital Fiber Optic Communication Systems," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1389-1430, October 1975.
8. J.E. Mazo and J. Salz, "On Optical Data Communication Via Direct Detection of Light Pulses," *Bell Syst. Tech. J.*, vol. 55, no. 3, pp. 347-369, March 1976.
9. S.D. Personick, "Baseband Linearity and Equalization in Fiber Optic Digital Communication Systems," *Bell Syst. Tech. J.*, vol. 52, no. 7, pp. 1175-1194, September 1973.
10. D.L. Snyder, *Random Point Processes,* Wiley, New York:, 1975.
11. S. Verdu, "Optimum multi-user signal detection," Ph. D. Dissertation, Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign. Report T-151 Coordinated Science Laboratory, Urbana, IL, Aug. 1984.
12. G.D. Forney, "Lower Bounds on Error Probability in the Presence of Large Intersymbol Interference," *IEEE Trans. Commun.*, vol. COM-20, pp. 76-77, February 1972.
13. S. Verdu, "Minimum Probability of Error for Asynchronous Gaussian Multiple-Access Channels," *IEEE Trans. on Information Theory,* vol. IT-32, Jan. 1986.
14. R. Dogliotti, A. Luvison, and G. Pirani, "Error Probability in Optical Fiber Transmission Systems," *IEEE Trans. Information Theory,* vol. IT-25, pp. 170-178, March 1979.
15. J.E. Mazo, "Faster-Than-Nyquist Signaling," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1451-1462, October 1975.
16. G.D. Forney, "Maximum Likelihood Sequence Estimation of Digital Sequences in the Presence of Intersymbol Interference," *IEEE Trans. Inform. Theory,* vol. IT-18, no. 3, pp. 363-378, May 1972.
17. A.J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding,* McGraw-Hill, New York:, 1979.
18. S. Verdu, "New Bound on the error probability of maximum likelihood sequence detection of signals subject to intersymbol interference," *Proc. 1985 Conf. Information Sciences and Systems,* pp. 413-418, March 1985.
19. S. Verdu, "Asymptotic Error Probability of Binary Hypothesis Testing for Poisson Point-Process Observations," *IEEE Trans. Information Theory,* vol. IT-32, Jan. 1986.
20. C. Shannon, R. Gallager, and E. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels, I," *Information and Control,* vol. 10, pp. 65-103, 1967.