The Capacity of Avalanche Photodiode-Detected Pulse-Position Modulation

J. Hamkins[1]

The capacity is determined for an optical channel employing pulse-position modulation (PPM) and an avalanche photodiode (APD) detector. This channel is different from the usual optical channel in that the detector output is characterized by a Webb-plus-Gaussian distribution, not a Poisson distribution. The capacity is expressed as a function of the PPM order, slot width, laser dead time, average number of incident signal and background photons received, and APD parameters. The capacity also is examined for the ideal photon-counting (Poisson) channel. Based on a system using a laser and detector proposed for X2000 second delivery, numerical results provide upper bounds on the data rate, level of background noise, and code rate that the channel can support while operating at a given bit-error rate. For the particular case studied, the capacity-maximizing PPM order is near 2048 for nighttime reception and 16 for daytime reception. Reed–Solomon codes can handle background levels 2.3 to 7.6 dB below the ultimate level that can be handled by codes operating at the Shannon limit.

I. Introduction

The capacity of a channel is the highest data rate it can reliably support. Whenever the data rate is less than the capacity of the channel, there exists an error-correcting code for the channel that has an output probability of error as small as desired, and, conversely, whenever the data rate is more than the capacity, the probability of error is bounded away from zero.

The capacity of the optical channel depends on many factors, including the modulation scheme, laser, transmission medium, photodetector, and preamplifier. Unlike the band-limited additive white Gaussian noise (AWGN) channel in which all performance-influencing factors are relevant to the channel capacity only in how they affect the bandwidth and signal-to-noise ratio, there is not a method to simplify the formulation of the capacity of the optical channel to so few variables. For example, the capacity depends separately on the signal and background light levels, not simply on their ratio. In this article, the functional dependence of the capacity is distilled to the following six major parameters: (1) the pulse-position modulation (PPM) order, M, (2) the laser pulse width, T_s, (3) the necessary dead time between pulses, T_d, (4) the average number of signal photons per pulse incident on the detector, $\bar{n}_s$, (5) the average number of background photons per slot incident on the detector, $\bar{n}_b$, and (6) the detector itself. These parameters are represented by the vector $(M, \bar{n}_s, \bar{n}_b, T_s, T_d,$ detector), and we will write the capacity as $C = C(M, \bar{n}_s, \bar{n}_b, T_s, T_d,$ detector). For an avalanche photodiode (APD) detector, the parameters used are

[1] Communications Systems and Research Section.

the quantum efficiency, η; ionization ratio, k_{eff}; noise temperature, T; load resistance, R; noise equivalent one-sided bandwidth, B; bulk leakage current, I_b; and surface leakage current, I_s. Not explicitly included in the functional description of the capacity is the modulation extinction ratio, α_{er}, of the laser, which we fix at 10^6 throughout the article. A description of these parameters is contained in [7,18].

Numerical results in the article are based on a system using components currently available and suggested by X2000 second delivery for a Mars-type mission. This includes a 1064-nm pulsed Q-switched neodymium-doped yttrium aluminum garnet (Nd:YAG) laser; a super low K (SLiK) APD detector made by EG&G, where K is the ionization ratio; and a transimpedance preamplifier.

Future improvements made in lasers and detectors can be evaluated with the methods outlined in this article. The increase in capacity can be projected by re-evaluating the equations with new $(M, \bar{n}_s, \bar{n}_b, T_s, T_d,$ detector) parameters.

In the following section, the optical channel is described and the notation used in this article is given. We also discuss the various units in which capacity may be expressed. Section III gives the analytic capacity results, including derivations of the capacity of PPM, the probability of uncoded symbol error for the APD and ideal photon-counting detectors, and implications of the converse of Shannon's capacity theorem. In Section IV, we give the numerical capacity results, and, in Section V, we state conclusions and discuss future research needed in this area.

II. Preliminaries

A. Channel Description and Notation

1. Encoder, Laser Modulator, and Optical Channel. This article concerns the communications system shown in Fig. 1. The information bits $\mathbf{U} = (U_1, \cdots, U_k)$ are independent identically distributed (i.i.d.) binary random variables assumed to take on the values 0 and 1 with equal probability. The vector $\mathbf{U}$ is encoded to $\mathbf{X} = (X_1, \cdots, X_n)$, a vector of n M-PPM symbols. Each M-PPM symbol is a number in $\{0, \cdots, M-1\}$ (or equivalently, a block of $\log_2 M$ bits, if M is a power of two). There are one signaling slot and $M-1$ nonsignaling slots for each M-PPM symbol. The symbol indicates to the modulator in which of the M time slots of length T_s to pulse the transmitting laser. Between each M-PPM symbol, the laser requires dead time, T_d, to recharge and ready itself for sending another pulse. The laser is coupled to a telescope, and pulses are transmitted through the optical channel to the receiving telescope, where background light also enters the receiving telescope.

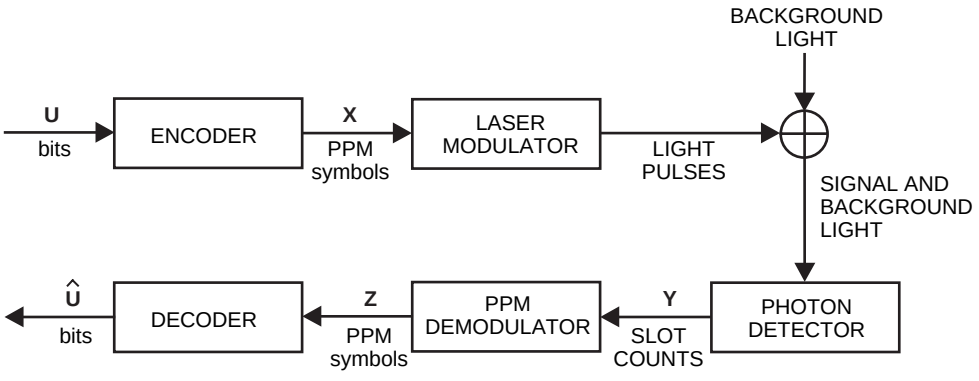

Fig. 1. An optical communications system.

2. Detector. At the receiver, light is focused on a photodetector, which we restrict to either an APD or an ideal photon counter. The detector integrates over slot times to produce $\mathbf{Y} = (\mathbf{Y}_1, \cdots, \mathbf{Y}_n)$, where $\mathbf{Y}_i = (y_{i,1}, \cdots, y_{i,M})$ are the M soft outputs for the ith M-PPM symbol, $1 \leq i \leq n$. The number of photons incident on a detector from an incident optical field of known intensity is a Poisson distributed random variable [7]. The number of photons absorbed by the detector is equal to the number of photons incident times the quantum efficiency, η, of the detector. The secondary electrons at the output of the detector may have a more complicated probability distribution [4,17,26]. In this article, for simplicity, we assume perfect timing synchronization and no interslot interference, which implies that the number of absorbed photons in each slot is independent of the number of photons absorbed in all other slots. Recent work has developed a method to combat interslot interference by using trellis-coded modulation [12,21].

3. PPM Demodulator and Decoder. Typically, the individual slot statistics at the output of the detector are not available to the decoder.[2] Instead, for $1 \leq i \leq n$, a PPM demodulator uses the M slot statistics of $\mathbf{Y}_i$ to make an M-PPM symbol decision, $Z_i \in \{0, \cdots, M-1\}$, by choosing the slot within each symbol that maximizes the number of detected photons, or, in case of a tie, by randomly choosing a slot among those with the maximum statistic. It has recently been proven that this is the maximum-likelihood rule for PPM detection when the statistics are governed by the sum of Webb and Gaussian deviates [25]. Perhaps surprisingly, the maximum-likelihood rule becomes more complicated than "pick the largest" when the detector output is approximated by a Gaussian distribution, in which a nonsignaling slot has mean μ_b and variance σ_b^2 and a signaling slot has mean $\mu_b + \mu_s$ and variance $\sigma_b^2 + \sigma_s^2$. We avoid this problem by not using the Gaussian approximations.

B. The Units of Capacity

This article expresses the channel capacity in bits per second because ultimately the system designer wants to know how quickly data can be pumped through the channel using the given power available. The laser properties, optics efficiency, pointing accuracy, and space and atmospheric losses all affect C, but only through their influence on $\bar{n}_s$, $\bar{n}_b$, T_s, and T_d. Hence, we express the capacity as a function of the following parameters:

$$C = C(M, \bar{n}_s, \bar{n}_b, T_s, T_d, \text{ detector})$$

The units in which C is expressed affect the parameter values that maximize C. This fact, which might seem surprising at first, implies that work on maximizing photon efficiency (e.g., [2,13,16][3,4]) does not necessarily help determine the maximum data rate possible on the channel.

1. Bits Per Photon or Bits Per Channel Use. A channel capacity of C bits per channel use can be restated as $C/\bar{n}_s$ bits per signal photon, C/M bits per PPM slot (neglecting the dead time), and $C/(MT_s + T_d)$ bits per second. The capacity in bits per photon or bits per channel use is not bounded for noiseless PPM if perfect timing is assumed [19]. (Other practical constraints bound it [13,16].) Intuitively, the reason is that, by choosing increasing values of M and keeping the slot duration fixed, the statistics governing the number of photons detected in the signal slot remain the same, but the number of bits per symbol increases as $\log_2 M$. Thus, the capacity in bits per photon (or bits per channel use) increases as $\log_2 M$, an unbounded number as M increases.

---

[2] If individual slot statistics are available to the decoder, then the capacity will be higher. See Appendix A for a discussion of this case.

[3] J. Hamkins, "Lower Bounds on the Number of Required Photons for Reliable Optical Communication With PPM Signals," JPL Interoffice Memorandum 331.98.9.005 (internal document), Jet Propulsion Laboratory, Pasadena, California, November 1998.

[4] J. Hamkins, "More Numerical Capacity Results for the Photon Counting Channel," JPL Interoffice Memorandum 331.99.1.001 (internal document), Jet Propulsion Laboratory, Pasadena, California, January 1999.

This unbounded capacity in bits/photon is not particularly useful, however, because it necessitates a low data rate and wasted power. Lasers on a spacecraft can have power allocated to them on a continual basis, at least within the intervals of time set aside for transmission to Earth. This power is used primarily to charge the laser after it has fired a pulse. If the laser waits an extensive period of time between pulse firings, that power is being wasted. From an information theoretical standpoint, the waste can be quantified by the lost entropy of the signal. The information content of a set of signaling slots (ones) and nonsignaling slots (zeroes) decreases as their probabilities are made more disparate. An increasing M means that the information content per slot (or per unit time) is decreasing, because $M-1$ out of every M slots contain zeroes.

2. Bits Per Second. Instead of using an enormous value of M and transmitting one symbol, we would be better off transmitting two $(M/2)$-PPM symbols in the same amount of time (assuming $M >> T_d/T_s$), because there is a potential for $2\log_2(M/2)$ bits received, as opposed to only $\log_2 M$ bits. Neglecting dead time, the capacity of the errorless channel is $\log_2 M/M$ bits per slot, which is maximized when $M = 3$. (The noninteger maximum occurs when $M = e$.)

The optimum value of M may be much higher than three when the required dead time is taken into account. On an error-free channel using M-PPM, a slot time of T_s, and a laser dead time of T_d, the capacity in bits per second is

$$C = \frac{\log_2 M}{MT_s + T_d} \quad \text{bits/second}$$

M may be chosen to maximize this equation. For the laser used in this article, $T_s = 3.125 \times 10^{-8}$ seconds and $T_d = 4.32 \times 10^{-4}$ seconds, and an errorless channel capacity is optimized when $M = 2082$. For channels that produce errors, more complicated expressions of capacity result [shown later in Eqs. (1), (7), (11), and (A-1)] and a different optimal value of M emerges.

III. Analytic Results

In Section III.A, we derive the capacity of APD-detected PPM, in terms of the PPM order M and the probability of correct uncoded M-PPM symbol detection. A detailed summary of how to compute this probability then is given. In Section III.B, we discuss the capacity of the photon-counting channel for PPM signaling, as well as for the more general case of average and peak power-constrained signaling. In Section III.C, we use the converse to Shannon's capacity theorem to derive bounds on performance.

A. Capacity of APD-Detected PPM

1. Capacity as a Function of Correct PPM Symbol Detection. The capacity of the communications system in Fig. 1 is the maximum mutual information between the input and output,

$$C \triangleq \max_{p(\mathbf{X})} I(\mathbf{U};\hat{\mathbf{U}}) = \max_{p(\mathbf{X})} H(\hat{\mathbf{U}}) - H(\hat{\mathbf{U}}|\mathbf{U})$$

where $H(\hat{\mathbf{U}})$ is the entropy of $\hat{\mathbf{U}}$, $H(\hat{\mathbf{U}}|\mathbf{U})$ is the conditional entropy of $\hat{\mathbf{U}}$ given $\mathbf{U}$, and $I(\mathbf{U};\hat{\mathbf{U}})$ is the mutual information between $\mathbf{U}$ and $\hat{\mathbf{U}}$. Since the encoder and decoder are deterministic, invertible functions, the capacity of the system reduces in the usual way to

$$C = \max_{p(\mathbf{X})} I(\mathbf{X};\mathbf{Z}) = \max_{p(\mathbf{X})} H(\mathbf{Z}) - H(\mathbf{Z}|\mathbf{X})$$

The channel $\mathbf{X} \to \mathbf{Z}$ is an M-ary symmetric channel (repeated n times), whose capacity depends on the probability of correct uncoded symbol detection, $p \triangleq \Pr(X_i = Z_i)$. Under the assumptions of perfect timing and negligible interslot interference, the $M-1$ possible incorrect decisions are equally likely, and each incorrect M-PPM symbol has probability $q = (1-p)/(M-1)$. The capacity of the M-ary symmetric channel is given by [1]

$$C = \log_2 M + p \log_2 p + (M-1) q \log_2 q \quad \text{bits per channel use} \tag{1}$$

Thus, to compute the capacity of the optical channel, we need only determine p. Note that the analysis thus far has not depended on the particular type of detector used, only on the detector operating in a memoryless fashion.

2. The Probability of Correct Detection With an APD Detector. A low-noise APD enhances the detection of weak optical signals by amplifying the electrical current generated by absorbed photons. This is illustrated in Fig. 2, in which the diode symbol represents the more complicated solid-state components of the APD itself, and some of the APD parameters are shown in block diagram form. Unfortunately, in addition to amplifying the signal, the APD transforms the simple Poisson distribution of absorbed photons into a much more complicated probability density function (pdf) at the APD output. This pdf is known [4,17] but extremely complex to evaluate numerically. This Conradi–McIntyre distribution has been accurately approximated in a simpler formulation by Webb [26]. In particular, the probability that m secondary electrons are emitted from the APD in response to the absorption of, on average, $\bar{n}$ primary photons in a slot is approximately

$$\Pr_w(m|\bar{n}) = \frac{\exp\left[-\dfrac{(m-G\bar{n})^2}{2\bar{n}G^2F\left(1+\frac{m-G\bar{n}}{\bar{n}GF/(F-1)}\right)}\right]}{\sqrt{2\pi\bar{n}G^2F}\left[1+\dfrac{m-G\bar{n}}{\bar{n}GF/(F-1)}\right]^{3/2}} \tag{2}$$

where G is the average APD gain, F is the excess noise factor given by

$$F = k_{eff}G + \left(2 - \frac{1}{G}\right)(1 - k_{eff})$$

and k_{eff} is the ionization ratio. For values of m close to its mean, $G\bar{n}$, Eq. (2) can be approximated by a Gaussian pdf; however, $\Pr_w(m|\bar{n})$ departs greatly from a Gaussian pdf at both tails, which form the main contribution to error events in decoders [7].

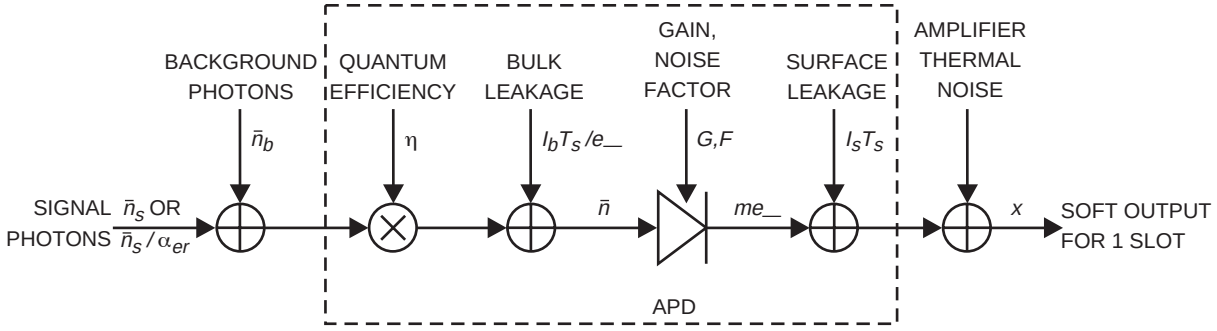

Fig. 2. The soft APD demodulator.

The detector output, x, is the sum of the charge due to the approximately Webb-distributed secondary electron emissions, a contribution from the APD surface leakage current, and Gaussian-distributed amplifier thermal noise, as shown in Fig. 2. Because of the thermal noise, the slot statistic x is not necessarily an integer, and may even be negative. The pdf of the sum charge is given by the convolution

$$p(x|\bar{n}) = \sum_{m=0}^{\infty} \phi\big(x, \mu_m, \sigma^2\big) \Pr_w(m|\bar{n}) \tag{3}$$

where $\phi(x, \mu_m, \sigma^2)$ is a Gaussian pdf with mean $\mu_m = me_- + I_s T_s$ and variance $\sigma^2 = (2e_- I_s + (4\kappa T/R))BT_s^2$, e_- is the electron charge, κ is Boltzmann's constant, T is the noise temperature, B is the single-sided noise bandwidth, and I_s is the APD surface leakage current. Note that $\Pr_w(m|\bar{n})$ and $p(x|\bar{n})$ are conditioned on the mean number of photons effectively *absorbed* by the detector, not *incident* on the detector. The relationship between incident and absorbed photons is governed by the quantum efficiency, η, of the detector, as shown in Fig. 2.

The average number of absorbed photons, $\bar{n}$, depends on whether the slot contains the signal. In a signaling slot, $\bar{n} = \eta\bar{n}_s + \eta\bar{n}_b + T_s I_b/e_-$; in a nonsignaling slot, $\bar{n} = \eta\bar{n}_s/\alpha_{er} + \eta\bar{n}_b + T_s I_b/e_-$. The $T_s I_b/e_-$ term represents the additional effective absorbed photons resulting from the APD bulk leakage current. The $\eta\bar{n}_s/\alpha_{er}$ term represents the laser-emitted photons absorbed when the laser is not sending a pulse. For practical purposes, the extinction ratio, α_{er}, is often inconsequential, being as high or higher than 10^6.

The probability of correct detection, p, is given by

$$p = \int_{-\infty}^{\infty} p\left(x|\eta\bar{n}_s + \eta\bar{n}_b + \frac{T_s I_b}{e_-}\right) \left[\int_{-\infty}^{x} p\left(y|\eta\bar{n}_b + \frac{\eta\bar{n}_s}{\alpha_{er}} + \frac{T_s I_b}{e_-}\right) dy\right]^{M-1} dx \tag{4}$$

where $p(x|\bar{n})$ is the conditional pdf of the detector slot statistic given that an average of $\bar{n}$ photons are absorbed by the detector, using Eq. (3). By plugging Eq. (4) into Eq. (1), the capacity is determined. In cases when Eq. (4) is too cumbersome to numerically evaluate, we may use a simpler expression as a bound and approximation. Using Jensen's inequality, p can be bounded by [22]

$$p \geq \left[1 - \int_{-\infty}^{\infty} p\left(x|\eta\bar{n}_s + \eta\bar{n}_b + \frac{I_b}{e_-}\right) \int_{x}^{\infty} p\left(y|\eta\bar{n}_b + \frac{\eta\bar{n}_s}{\alpha_{er}} + \frac{I_b}{e_-}\right) dy dx\right]^{M-1} \tag{5}$$

which will give a lower bound on capacity when plugged into Eq. (1). This bound always is tighter than the union bound [10], which implies that, as the probability of error gets small, the ratio of the bound to the true value tends to one.

B. Capacity of the Photon-Counting Channel

Both the incident and the absorbed photons of a photon counter have a Poisson distribution with mean values that differ by a factor of the quantum efficiency, η, of the photon counter. The photon-counting detector has been studied extensively in previous work on the optical channel (see, e.g., [2,3,5,6,9,11,13–16,19,20,23,27,28]). We may view the photon counter as a type of ideal photomultiplier tube (PMT), although a more proper statistical development of the PMT is developed in [24]. By using the ideal photon-counting model and the same quantum efficiency as that observed in real PMTs, we can obtain PMT performance approximations.

1. PPM Signaling on the Noisy Photon-Counting Channel. The capacity is determined by Eq. (1) and the probability of correct uncoded PPM detection, p, from the photon counter. Let $\text{Pois}(i,n) = n^i e^{-n}/i!$ denote the probability that a Poisson random variable with mean n takes on the value i, $i = 0, 1, \cdots$. The probability that an uncoded M-PPM symbol is detected correctly is given by

$$
\begin{aligned}
p = {} & \Pr\ (\text{uncoded PPM symbol is detected correctly}) \\
= {} & \frac{1}{M} \Pr \begin{pmatrix} \text{0 photons} \\ \text{detected in all} \\ \text{slots} \end{pmatrix} + \sum_{i=1}^{\infty} \Pr \begin{pmatrix} i \text{ photons} \\ \text{detected in} \\ \text{signal slot} \end{pmatrix} \times \Pr \begin{pmatrix} \text{less than } i \\ \text{photons} \\ \text{detected in all} \\ M-1 \\ \text{nonsignal slots} \end{pmatrix} \\
& + \sum_{i=1}^{\infty} \sum_{r=1}^{M-1} \frac{1}{r+1} \binom{M-1}{r} \Pr \begin{pmatrix} i \text{ photons} \\ \text{detected in} \\ \text{signal slot} \end{pmatrix} \times \Pr \begin{pmatrix} i \text{ photons} \\ \text{detected in} \\ \text{exactly } r \\ \text{nonsignal slots} \end{pmatrix} \times \Pr \begin{pmatrix} \text{less than } i \\ \text{photons} \\ \text{detected in} \\ M-1-r \\ \text{nonsignal slots} \end{pmatrix} \\
= {} & \frac{e^{-(\eta \bar{n}_s + M \eta \bar{n}_b)}}{M} + \sum_{i=1}^{\infty} \text{Pois}(i, \eta \bar{n}_s + \eta \bar{n}_b) \left[\sum_{m=0}^{i-1} \text{Pois}(m, \eta \bar{n}_b) \right]^{M-1} \\
& + \sum_{i=1}^{\infty} \sum_{r=1}^{M-1} \frac{1}{r+1} \binom{M-1}{r} \text{Pois}(i, \eta \bar{n}_s + \eta \bar{n}_b) \left(\text{Pois}(i, \eta \bar{n}_b)\right)^r \left[\sum_{m=0}^{i-1} \text{Pois}(m, \eta \bar{n}_b) \right]^{M-1-r}
\end{aligned}
$$

After some straightforward algebra, this simplifies to [3,8]

$$
p = \frac{1}{M} e^{-(\eta \bar{n}_s + M \eta \bar{n}_b)} + \sum_{k=1}^{\infty} \left[\text{Pois}(k, \eta \bar{n}_s + \eta \bar{n}_b) \left(\sum_{m=0}^{k-1} \text{Pois}(m, \eta \bar{n}_b) \right)^{M-1} \frac{((1+a)^M - 1)}{aM} \right] \tag{6}
$$

where $a = \text{Pois}(i, \eta \bar{n}_b) / \sum_{m=0}^{i-1} \text{Pois}(m, \eta \bar{n}_b)$. Equation (6) may be plugged into Eq. (1) to compute the capacity of PPM when an ideal photon counter is used.

2. PPM Signaling on the Noiseless Photon-Counting Channel. When no background photons are absorbed, $\bar{n}_b = 0$, and the capacity simplifies to [16]

$$
C = \frac{(1 - e^{-\eta \bar{n}_s}) \log_2 M}{(M T_s + T_d)} \quad \text{bits/second} \tag{7}
$$

This can provide an easy bound for the noisy photon-counting channel. As $\bar{n}_b \to 0$, the capacity will approach this noiseless case.

3. Average and Peak Power-Limited Signaling on the Noisy Photon-Counting Channel. An M-ary signaling scheme more general than PPM, which results in a higher capacity, was considered

in [27]. There are M waveforms $\lambda_m(t)$, $m \in \{1, \cdots, M\}$, where $\lambda_m(t)$ is the absorbed signal photons per second as a function of time. Each waveform is subject to a peak power constraint,

$$0 \leq \lambda_m(t) \leq A \ \text{ photons/second}, \quad \text{for } 0 \leq t \leq T \tag{8}$$

and to an average power constraint,

$$\frac{1}{T}\int_0^T \lambda_m(t)dt \leq \sigma A \ \text{ photons/second} \tag{9}$$

Background light adds a rate of λ_0 photons/second, and the probability that k photons are counted at the receiver in an interval of length τ is given by $\text{Pois}(k, \Lambda) = (e^{-\Lambda}\Lambda^k)/k!,\ k = 0, 1, 2, \cdots$, where

$$\Lambda = \int_t^{t+\tau} (\lambda(t') + \lambda_0)dt' \tag{10}$$

Let the probability of correct word detection be denoted p. For a fixed A, Λ_0, and σ, a rate $R \geq 0$ bits/photon is achievable if for each $\epsilon > 0$ there exists a system operating with parameters (M, T, σ, p) such that $p > 1 - \epsilon$ and $M \geq 2^{R\sigma A}$. The channel capacity, C, is the supremum of all achievable rates. Wyner found that the capacity of a photon-counting channel, when the signal is subject to the constraints in Eqs. (8) and (9) and the photon counter absorbs λ_0 background photons/second, is given by

$$C = A[q(1+s)\log_2(1+s) + (1-q)s\log_2 s - (q+s)\log_2(q+s)] \ \text{ bits/second} \tag{11}$$

where

$$s = \frac{\lambda_0}{A}$$

$$q = \min(\sigma, q_0(s))$$

$$q_0(s) = \frac{(1+s)^{1+s}}{s^s e} - s$$

C. Implications of the Converse of Shannon's Capacity Theorem

The converse of Shannon's channel-coding theorem applied to the communications system in Fig. 1 implies that any error-correcting code with code rate R_c information bits per transmitted bit satisfies

$$R_c(\log_2 M)(1 - \mathcal{H}_b(P_b)) \leq C(M, \bar{n}_s, \bar{n}_b, T_s, \text{ detector}) \ \text{ bits per channel use} \tag{12}$$

where $\mathcal{H}_b(x) \triangleq -x\log_2 x - (1-x)\log_2(1-x)$ is the binary entropy function and P_b is the coded bit-error rate (BER). Here, $R_c \log_2 M$ is the rate in bits per channel use. Note that capacity is expressed in bits per channel use, which removes its dependence on T_d. We may rewrite Eq. (12) as

$$P_b \geq \mathcal{H}_b^{-1}\left[1 - \frac{C(M, \bar{n}_s, \bar{n}_b, T_s, \text{ detector})}{R_c \log_2 M}\right] \tag{13}$$

For a given code rate R_c and fixed $(M, \bar{n}_s, \bar{n}_b, T_s, \text{ detector})$, Eq. (13) gives the minimum BER, P_b, that any rate-R_c code can achieve on the channel. Alternatively, we may write

$$R_c \leq \frac{C(M, \bar{n}_s, \bar{n}_b, T_s, \text{ detector})}{(\log_2 M)(1 - \mathcal{H}_b(P_b))} \tag{14}$$

For a given desired error rate, say $P_b = 10^{-6}$, Eq. (14) gives an upper bound on the code rate, i.e., the percentage of the transmission bits that carry information. Since the data rate $R_d = (R_c \log_2 M)/(MT_s + T_d)$, this translates directly into a bound on the data rate as well:

$$R_d \leq \frac{C(M, \bar{n}_s, \bar{n}_b, T_s, \text{ detector})}{(MT_s + T_d)(1 - \mathcal{H}_b(P_b))} \text{ bits/second} \tag{15}$$

IV. Numerical Capacity Results

All numerical evaluations were carried out on a 333-MHz Pentium II using programs written in C and Perl. We used parameters from a 1064-nm pulsed Q-switched Nd:YAG laser having slot width $T_s = 31.25$ ns, required dead time $T_d = 432,000$ ns, and modulation extinction ratio $\alpha_{er} = 10^6$. This laser was chosen based on its proposed use for X2000 second delivery.[5] The EG&G SLiK APD and follow-on electronics have the following parameters: $k_{eff} = 0.007$, $T = 300$ K, $R = 179,700\ \Omega$, $B = 1/2T_s$ Hz, $I_b = 4 \times 10^{-14}$ A, $I_s = 2 \times 10^{-9}$ A, and $\eta = 38$ percent. See Appendix B for a description of these parameters or [7] for a more detailed explanation. All numerical results reported in the article used an optimized APD gain. We discuss this optimization in Section IV.E; the optimal gain varied from 50 to 200, depending on the background level.

A. Bit-Error Rate Versus Background Level

We used Eq. (13) to determine the lowest bit-error rate theoretically possible for PPM signaling using the Nd:YAG laser and SLiK APD. The capacity was determined by numerically evaluating Eq. (5) and plugging into Eq. (1); substitution into Eq. (13) gives the bound on bit-error rate. Figures 3 through 5 indicate the bounds for $M = 256$, $M = 64$, and $M = 2$, respectively. As can be seen, when operating at a BER of 10^{-6}, the use of rate-7/8 codes promises the ability to withstand background levels over 40 dB stronger than an uncoded system. Rate-7/8 Reed–Solomon (RS) codes operate within 3.5 dB of the limit for rate-7/8 codes. In an uncoded system with $M = 256$, we must have $\bar{n}_b \leq 0.001$ in order to achieve a BER of 10^{-6}; with an RS(255,224) code, we required $\bar{n}_b \leq 7.1$; and capacity implies $\bar{n}_b \leq 16.0$. Note in Table 1, which provides the maximum background light that can be handled while operating with a coded BER of 10^{-6}, that, when $M = 64$, an RS code is further from capacity than when $M = 256$. Table 1 indicates that codes operating at the Shannon limit can withstand 2.3 to 7.6 dB higher levels of background light, as compared with RS codes. In the table, the parameters are $M = 256, 64$, and 2; $R_c = 7/8$ or $1/2$; $\bar{n}_s = 100$; $T_s = 31.25$ ns; and a SLiK detector.

[5] G. Ortiz, personal communication, Communications Systems and Research Section, Jet Propulsion Laboratory, Pasadena, California, March 1999.

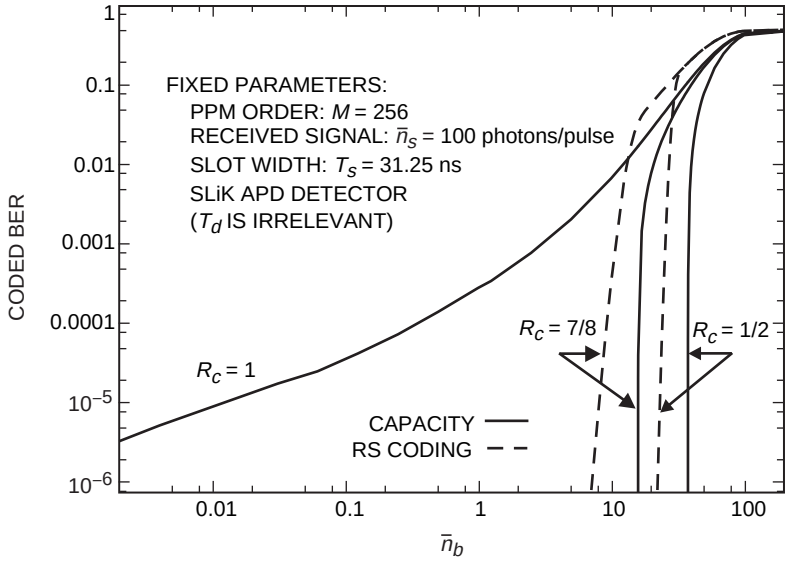

Fig. 3. A comparison of RS performance with the Shannon limit.

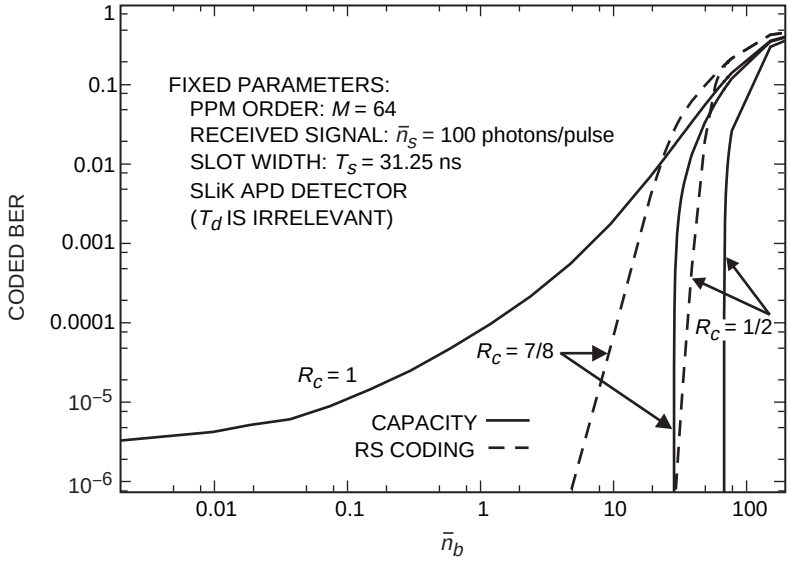

Fig. 4. A comparison of RS performance with the Shannon limit.

B. Code Rate Versus Background Level

Using Eq. (14), a bound on the highest coding rate possible while operating at a given BER and $(M, \bar{n}_s, \bar{n}_b, T_s,$ detector) was calculated. This code rate is the percentage of transmission symbols that carry information. The remainder of the transmissions carry redundancy used for coding. It also can be viewed as the fraction of the maximum data rate, $\log_2 M/(MT_s + T_d)$, that it is possible to achieve on the channel. Both the maximum code rate and the code rate needed by RS coding are shown in Fig. 6.

C. Data Rate Versus Background Level

Using Eq. (15), a bound on the highest data rate possible while operating at a given BER and $(M, \bar{n}_s, \bar{n}_b, T_s, T_d,$ detector) was calculated. As $\bar{n}_b \to 0$, the data rate tends to the maximum dictated by

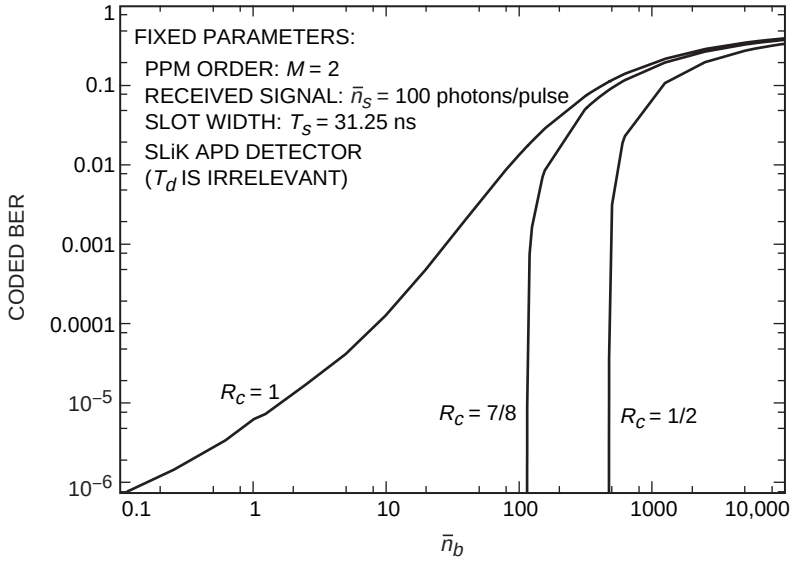

Fig. 5. The Shannon limit on the BER of 2-PPM.

Table 1. Maximum background light while operating with a coded BER of 10^{-6}.

M	R_c	$\bar{n}_b$, maximum	$\bar{n}_b$, RS coding	Difference, dB
256	7/8	16.0	7.1	3.5
64	7/8	29.3	5.1	7.6
2	7/8	115	—	—
256	1/2	37.8	22.5	2.3
64	1/2	69.9	30.5	3.6
2	1/2	475	—	—

M, T_s, and T_d: $\log_2 M/(MT_s + T_d)$. Figure 7 shows the maximum attainable data rate for various M, a range of $\bar{n}_b$, and with fixed $\bar{n}_s$, T_s, T_d, and detector. Also shown is the RS coding performance when $M = 256$.

D. Optimization of PPM Order

Figure 7 begs the question of what PPM order optimizes the data rate. For nighttime reception in which $\bar{n}_b \ll 1$, the optimal PPM order is near $M = 2048$. This closely follows the discussion in Section II.B.2 regarding the errorless channel. For daytime reception in which $\bar{n}_b \approx 100$, we can see from Fig. 7 that the optimal PPM order is under 256. To be more precise, the order of PPM that maximizes capacity in bits per second can be seen directly from a plot of capacity versus M. This is shown in Fig. 8, and the optimal PPM orders for various values of $\bar{n}_b$ are summarized in Table 2. In Table 2, the parameters are $P_b = 10^{-6}$, $\bar{n}_s = 100$, $\bar{n}_b \in \{0.1, 1, 10, 50, 100\}$, $T_s = 31.25$ ns, $T_d = 432,000$ ns, and the SLiK APD detector.

This suggests the use of a multiple PPM order communications system. During nighttime reception, it should use M on the order of thousands, and during daytime reception, it should use M on the order of dozens. Unoptimized PPM orders can be costly. As can be seen from Fig. 8, using $M = 2036$ during the day would be disastrous for the data rate. Using $M = 18$ at night reduces capacity by over half.

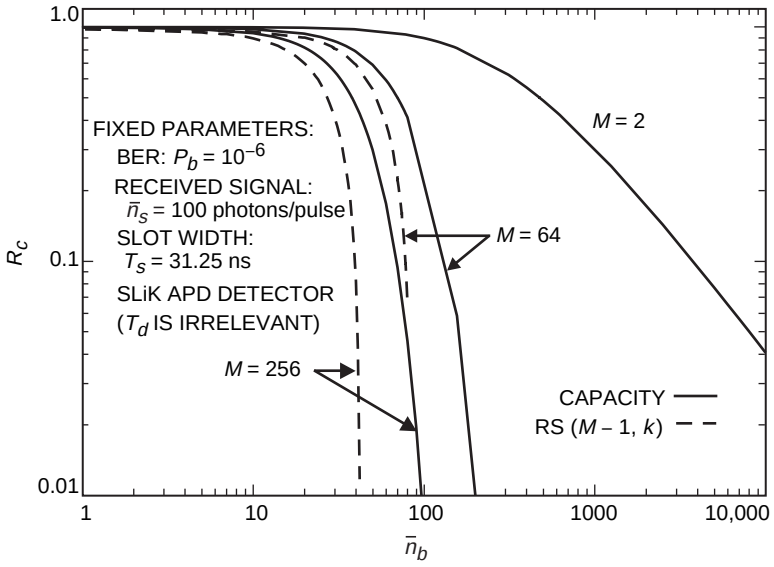

Fig. 6. A comparison of the required RS code rate with the maximum code rate implied by capacity while operating with $P_b = 10^{-6}$.

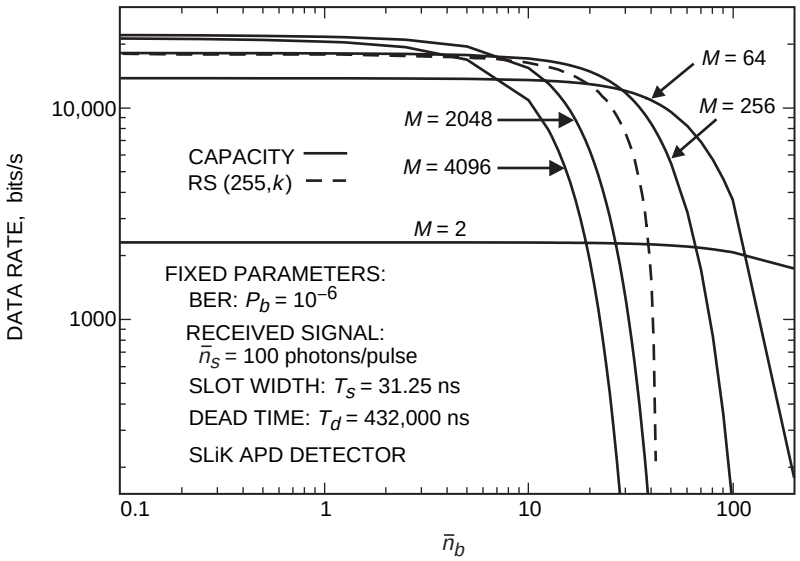

Fig. 7. The capacity of *M*-PPM on an optical channel, with M∈ {2,64,256,2048,4096}.

E. APD Gain Optimization

The APD gain is a parameter required to evaluate performance. For example, Eq. (2) depends on the gain. All numerical results in this article use an optimized gain. For each value of $\bar{n}_b$, the numerical capacity or other needed quantity was computed over a range of gains and the largest one chosen. In the interest of time, the gain was restricted to multiples of five. In all cases considered, a gain difference of five (and typically much more than five) from the optimal value made little difference in the numerical results. Shown in Fig. 9 are the optimal gain values. Optimal APD gains are also reported in [22].

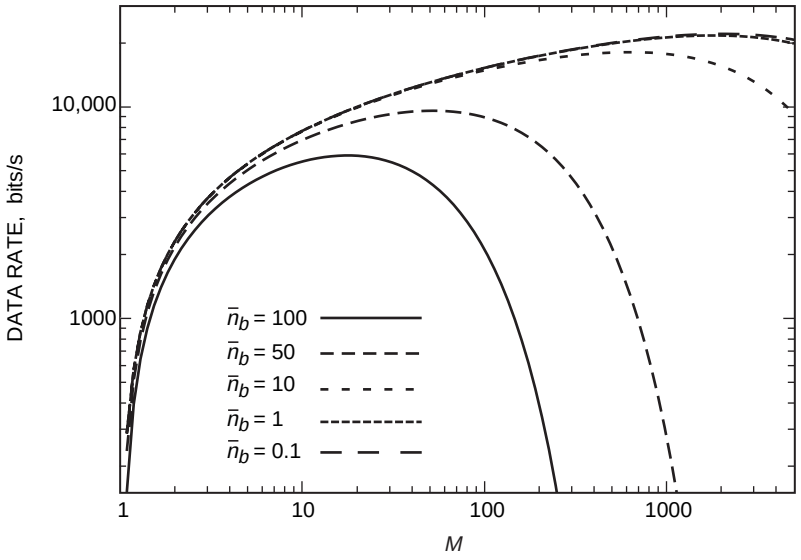

Fig. 8. The capacity of *M*-PPM on an optical channel, with $P_b = 10^{-6}$, $\bar{n}_s = 100$, $\bar{n}_b \in \{0.1,1,10,50,100\}$, $T_s = 31.25$ ns, $T_d = 432{,}000$ ns, and the SLiK APD detector.

Table 2. Optimal PPM orders *M*.

$\bar{n}_b$	Optimal M
0.1	2036
1	1815
10	634
50	52
100	18

F. Comparison of Simulation With the Upper Bound of Uncoded APD-Detected PPM

Most numerical results in this article required determination of the probability of uncoded PPM symbol-detection error. Two approaches were taken—simulation and bounding. Using the method given in [7] to simulate the statistical properties of the APD, a channel was simulated for 100,000 256-PPM symbols. The probability of uncoded symbol error is shown in Fig. 10 and is compared with the upper bound used in Eq. (5) that was used to derive the remainder of the numerical results in the article. Since the upper bound is tighter than the union bound, it necessarily converges to the true value. We see this happening, if slowly, in Fig. 10.

V. Conclusions

This article considered an X2000 second delivery laser and detector, representing the current technology available. Capacity was reported in terms of BER versus background level, code rate versus background level, and data rate versus background level. Optimization of the PPM order and APD gain also were discussed.

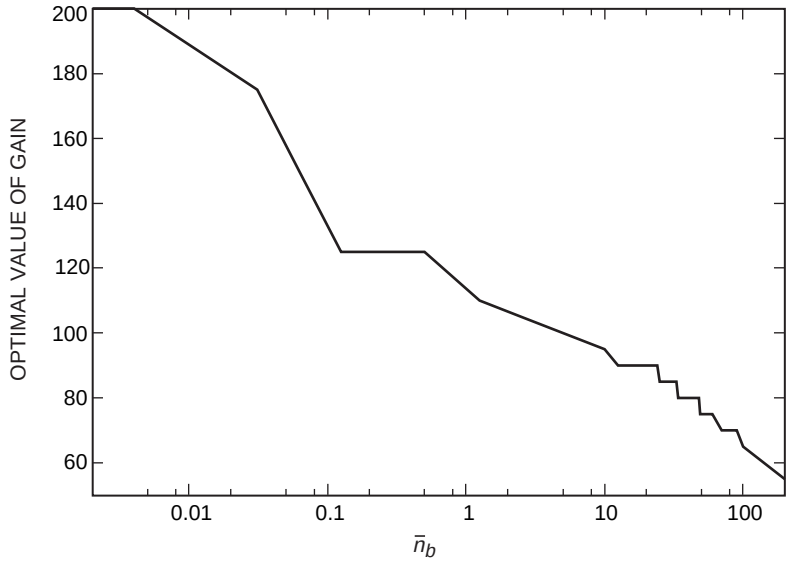

Fig. 9. The optimal gain, as a function of $\bar{n}_b$, for $\bar{n}_s = 100$, $T_s = 31.25$ ns, $T_d = 432{,}000$ ns, and the SLiK APD detector.

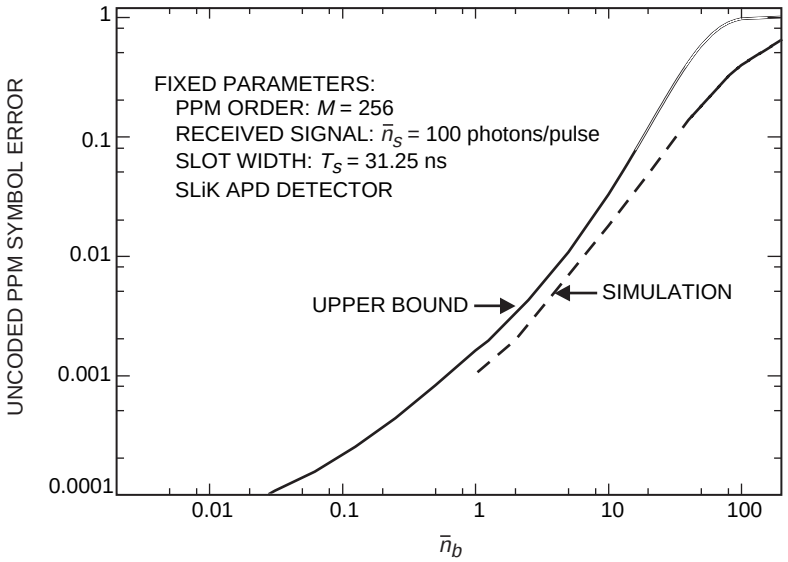

Fig. 10. The probability of uncoded 256-PPM symbol error on an optical channel.

Results indicate that, for 256-PPM and rate-7/8 coding, RS codes can handle all but the last 3.5 dB of the background levels that capacity promises can be handled while operating at a BER of 10^{-6}.

The optimal value of PPM order depends greatly on the background light. For nighttime reception, the optimal PPM order was found to be $M = 2036$, while for daytime reception, $M = 18$. With mismatched PPM order, the capacity reduces by more than a factor of two, which suggests that multiple-order PPM systems should be used if feasible.

Future advances in lasers and detectors have not been considered in this article. Evaluating capacity for these advancements would provide very useful information regarding the limits at which the optical channel can operate. This work is straightforward but as of yet undone.

This article also gives a framework that can be used for evaluating the sensitivity of the capacity to each parameter. Holding all parameters fixed but one, it is possible to show the sensitivity of capacity to each parameter. This would provide valuable feedback to laser and detector developers and to system designers, who then could expend effort in the areas leading to the biggest system gains. For the APD, this would include a study of the effects of the quantum efficiency, thermal noise levels, dark currents, and so forth; for the lasers, this would include the repetition rate and the pulse power. Also, note that in this article we mostly kept $\bar{n}_s$ fixed at 100 photons per pulse. It is important to know how the capacity changes for varying $\bar{n}_s$.

Also unknown is the capacity loss due to the hard PPM symbol demodulator. Removing it and providing soft slot statistics directly to the decoder would improve capacity, and a study to quantify this gain would be an important advancement in our understanding of the optical channel.

Acknowledgments

The author thanks Meera Srinivasan for providing the C program to bound the symbol error using the APD, Juan Ceniceros for providing many FOCAS link tables, Gerry Ortiz for providing X2000 second delivery laser and detector parameters, and Bob McEliece for helpful discussions regarding the units of capacity and optimization of the PPM order.

References

[1] R. B. Ash, *Information Theory*, New York: Dover, 1965.

[2] S. A. Butman, J. Katz, and J. R. Lesh, "Bandwidth Limitations on Noiseless Optical Channel Capacity," *IEEE Trans. Commun.*, vol. COM-30, no. 5, pp. 1262–1264, May 1982.

[3] C.-C. Chen, "Figure of Merit for Direct-Detection Optical Channels," *The Telecommunications and Data Acquisition Progress Report 42-109, January–March 1992*, Jet Propulsion Laboratory, Pasadena, California, pp. 136–151, May 15, 1992.
http://tmo.jpl.nasa.gov/tmo/progress_report/42-109/109L.PDF

[4] J. Conradi, "The Distribution of Gains in Uniformly Multiplying Avalanche Photodiodes: Experimental," *IEEE Trans. on Electron Devices*, vol. ED-19, no. 6, pp. 713–718, June 1972.

[5] M. H. A. Davis, "Capacity and Cutoff Rate for Poisson-Type Channels," *IEEE Trans. Inform. Theory*, vol. IT-26, no. 6, pp. 710–715, November 1980.

[6] D. Divsalar, R. M. Gagliardi, and J. H. Yuen, "PPM Performance for Reed–Solomon Decoding Over an Optical–RF Relay Link," *IEEE Trans. Commun.*, vol. COM-32, no. 3, pp. 302–305, March 1984.

[7] F. M. Davidson and X. Sun, "Gaussian Approximation Versus Nearly Exact Performance Analysis of Optical Communication Systems With PPM Signaling and APD Receivers," *IEEE Trans. Commun.*, vol. 36, no. 11, pp. 1185–1192, November 1988.

[8] R. M. Gagliardi, *Introduction to Communications Engineering*, New York: John Wiley & Sons, 1995.

[9] C. W. Helstrom, "Comments on 'The Capacity of the Photon Counting Channel,'" *IEEE Trans. Inform. Theory*, vol. IT-28, no. 3, p. 556, May 1982.

[10] L. W. Hughes, "A Simple Upper Bound on the Error Probability for Orthogonal Signals in White Noise," *IEEE Trans. Commun.*, vol. 40, no. 4, p. 670, April 1992.

[11] Y. M. Kabanov, "The Capacity of a Channel of the Poisson Type," *Theory Probab. Appl.*, vol. 23, pp. 143–147, 1978.

[12] K. Kiasaleh and T.-Y. Yan, "T-PPM: A Novel Modulation Scheme for Optical Communication Systems Impaired by Pulse-Width Inaccuracies," *The Telecommunications and Mission Operations Progress Report 42-135, July–September 1998*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–16, November 15, 1998.
http://tmo.jpl.nasa.gov/tmo/progress_report/42-135/135G.pdf

[13] J. R. Lesh, "Capacity Limit of the Noiseless, Energy-Efficient Optical PPM Channel," *IEEE Trans. Commun.*, vol. COM-31, no. 4, pp. 546–548, April 1983.

[14] R. G. Lipes, "Pulse-Position-Modulation Coding as Near-Optimum Utilization of Photon Counting Channel With Bandwidth and Power Constraints," *The Deep Space Network Progress Report 42-56, January and February 1980*, Jet Propulsion Laboratory, Pasadena, California, pp. 108–113, April 15, 1980.

[15] J. L. Massey, "Capacity, Cutoff Rate, and Coding for a Direct-Detection Optical Channel," *IEEE Trans. Commun.*, vol. COM-29, no. 11, pp. 1615–1621, November 1981.

[16] R. J. McEliece, "Practical Codes for Photon Communication," *IEEE Trans. Inform. Theory*, vol. IT-27, no. 4, pp. 393–398, July 1981.

[17] R. J. McIntyre, "The Distribution of Gains in Uniformly Multiplying Avalanche Photodiodes: Theory," *IEEE Transactions on Electron Devices*, vol. ED-19, no. 6, pp. 703–713, June 1972.

[18] G. S. Mecherle, *Maximized Data Rate Capability for Optical Communication Using Semiconductor Devices With Pulse Position Modulation*, Ph.D. Thesis, University of Southern California, Los Angeles, California, May 1986.

[19] J. R. Pierce, E. C. Posner, and E. R. Rodemich, "The Capacity of the Photon Counting Channel," *IEEE Trans. Inform. Theory*, vol. IT-27, no. 1, pp. 61–77, January 1981.

[20] D. L. Snyder and I. B. Rhoades, "Some Implications of the Cutoff-Rate Criterion for Coded Direct-Detection Optical Communication Systems," *IEEE Trans. Inform. Theory*, vol. IT-26, no. 3, pp. 327–338, May 1980.

[21] M. Srinivasan and V. Vilnrotter, "Receiver Structure and Performance for Trellis-Coded Pulse-Position Modulation in Optical Communication Systems," *The Telecommunications and Mission Operations Progress Report 42–135, July–September 1998*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–11, November 15, 1998.
http://tmo.jpl.nasa.gov/tmo/progress_report/42-135/135H.pdf

[22] M. Srinivasan and V. Vilnrotter, "Symbol-Error Probabilities for Pulse-Position Modulation Signaling With an Avalanche Photodiode Receiver and Gaussian Thermal Noise," *The Telecommunications and Mission Operations Progress Report 42-134, April–June 1998*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–11, August 15, 1998.
http://tmo.jpl.nasa.gov/tmo/progress_report/42-134/134E.pdf

[23] H. H. Tan, "Capacity of a Multimode Direct Detection Optical Communication Channel," *The Telecommunications and Data Acquisition Progress Report 42-63, March and April 1981*, Jet Propulsion Laboratory, Pasadena, California, pp. 51–70, June 15, 1981.

[24] H. H. Tan, "A Statistical Model of the Photomultiplier Gain Process With Applications to Optical Pulse Detection," *The Telecommunications and Data Acquisition Progress Report 42-68, January and February 1982*, Jet Propulsion Laboratory, Pasadena, California, pp. 55–67, April 15, 1982.
http://tmo.jpl.nasa.gov/tmo/progress_report/42-68/68H.PDF

[25] V. Vilnrotter, M. Simon, and M. Srinivasan, *Maximum Likelihood Detection of PPM Signals Governed by an Arbitrary Point Process Plus Additive Gaussian Noise*, JPL Publication 98-7, Jet Propulsion Laboratory, Pasadena, California, April 1998.

[26] P. P. Webb, R. J. McIntyre, and J. Conradi, "Properties of Avalanche Photodiodes," *RCA Review*, vol. 35, pp. 234–278, June 1974.

[27] A. D. Wyner, "Capacity and Error Exponent for the Direct Detection Photon Channel–Part I," *IEEE Trans. Inform. Theory*, vol. 34, no. 6, pp. 1449–1461, November 1988.

[28] A. D. Wyner, "Capacity and Error Exponent for the Direct Detection Photon Channel–Part II," *IEEE Trans. Inform. Theory*, vol. 34, no. 6, pp. 1462–1471, November 1988.

[29] J. Hamkins, "Performance of Binary Turbo-Coded 256-ary Pulse-Position Modulation," *The Telecommunications and Mission Operations Progress Report 42-138, April–June 1999*, Jet Propulsion Laboratory, Pasadena, California, pp. 1–15, August 15, 1999.
http://tmo.jpl.nasa.gov/tmo/progress_report/42-138/138B.PDF

Appendix A
Capacity of the Soft-Decision Optical Channel

This article considered the capacity of the "hard-decision" PPM optical channel, in which the decoder operates on PPM symbol decisions. However, there are practical ways to provide the decoder with additional information that can improve performance. When additional information is available, the capacity of the channel will increase (or, at least, will not decrease). Preliminary work on this has been started.[6]

There are several forms the additional information could take. It could be a reliability metric associated with each PPM symbol decision, indicating the conditional probability that the symbol decision is correct given the values of all the slot statistics. Or, the information could consist of the l most likely PPM symbols and each of their reliabilities. Decoders can incorporate the additional information into a decoding algorithm that performs better as a result. Ultimately, the symbol detector could be removed entirely, and the decoder could operate on all M soft statistics directly. This approach has been taken [29] and has shown improvement over Reed–Solomon coding in the cases considered there. This option is often within practical limits. In situations when full slot statistics are impractical, it still is useful to quantify the capacity one is giving up by not being able to use such an approach.

The capacity of the communications system when the symbol demodulator (see Fig. 1) is removed is at least as high as the channel that contains the symbol demodulator. This is a simple consequence of the data-processing theorem. Using Fig. 1 with the demodulator removed, let $X \in \{0, \cdots, M-1\}$ denote the M-PPM symbol sent, and let $\mathbf{Y} = (y_1, \cdots, y_M)$ be the vector of slot statistics, $y_i \in \Re$. The capacity of the modified communications system is

$$C = \max_{p(X)} \sum_{j=0}^{M-1} \int_{\mathbf{Y}} p(X=j) p(\mathbf{Y}|X=j) \log_2 \left(\frac{p(\mathbf{Y}|X=j)}{\sum_{l=0}^{M-1} p(X=l) p(\mathbf{Y}|X=l)} \right) d\mathbf{Y}$$

$$= \int_{\mathbf{Y}} p(\mathbf{Y}|X=j) \log_2 \left(\frac{p(\mathbf{Y}|X=j)}{\frac{1}{M} \sum_{l=0}^{M-1} p(\mathbf{Y}|X=l)} \right) d\mathbf{Y} \qquad \text{(A-1)}$$

This can be difficult to compute for APD statistics and typical PPM orders such as $M = 256$.

As a practical matter, codes are not yet available that can take advantage of the additional soft information within the individual slot statistics. RS codes cannot use soft information, except to the extent that they can define the erasure probabilities. Recent work on turbo-coded PPM has shown promise [29], but code rates needed for the optical channel have not been studied yet. Numerical analysis could indicate whether the difference in capacity is worth the extra effort needed to retain the soft slot statistics.

[6] S. Dolinar, D. Divsalar, and F. Pollara, "Capacity of PPM on Gaussian and Webb Channels," JPL Interoffice Memorandum Draft (internal document), Jet Propulsion Laboratory, Pasadena, California, November 1998.

Appendix B
Parameters and Notation

The following is a list of parameters and notation used in this article.

Laser and Modulator

M	256, 64, 2	PPM order
T_s	3.125×10^{-8}	Width of the PPM slot required by the laser, in seconds
T_d	4.32×10^{-4}	Dead time between PPM symbols required by the laser, in seconds
α_{er}	10^6	Modulation extinction ratio

Received Light

$\bar{n}_s$	100	Average number of signal photons incident on the photodetector, per pulse
$\bar{n}_b$	0.001–10,000	Average number of background photons incident on the photodetector, per slot

APD Detector and Follow-On Electronics

η	38%	Quantum efficiency
k_{eff}	0.007	Ionization ratio
T	300	Noise temperature, in kelvins
G	50–200	Gain
R	179,700	Load resistance implied by transimpedance model, $5.75 \times 10^{12} \times T_s$, in ohms
B	$\frac{1}{2T_s}$	Noise equivalent one-sided bandwidth, in hertz
I_b	4×10^{-14}	Bulk leakage current, in amperes
I_s	2×10^{-9}	Surface leakage current, in amperes

Constants

κ	1.38×10^{-23}	Boltzmann's constant, in joules/kelvin
e_-	1.6×10^{-19}	Electron charge, in coulombs

Error probabilities

p	Probability of correct uncoded PPM detection
q	Probability uncoded PPM symbol i is detected as symbol j, $j \neq i$