T-PPM: A Novel Modulation Scheme for Optical Communication Systems Impaired by Pulse-Width Inaccuracies

K. Kiasaleh[1] and T.-Y. Yan[2]

M-ary pulse-position modulation (PPM) has received considerable attention for direct-detection photon communications over unguided channels. The analysis generally assumes that the signaling set is orthogonal. However, the orthogonality of the signaling set will be destroyed by the finite area and bandwidth of optical detectors, resulting in severe intersymbol interference. This article presents the analysis of a trellis-based pulse-position modulation (T-PPM) scheme for photon communications with non-rectangular pulses. The novelty of the scheme includes the use of a set-partitioning methodology to increase the minimum distance using a simple convolutional encoder. The Viterbi algorithm is used at the receiver to separate the signaling set as part of the demodulation process. It has been shown that T-PPM will restore performance losses due to reduced peak intensity during the detection process. Furthermore, for a large range of background radiation levels, the average number of photons per information bit for T-PPM is smaller than that of the regular PPM. Numerical examples show that for a symbol-error rate of 10^{-3} when the received pulses extend over 4 PPM slots, the average laser energy per symbol for 256-ary T-PPM could be reduced by as much as 2 dB.

I. Introduction

Pulse-position modulation (PPM) has received considerable attention as the modulation of choice for direct-detection optical communications over unguided optical channels. There are several key aspects of this modulation that are critical to the deployment of this modulation scheme for deep-space communications. First, the presence of a pulse in the symbol frame regardless of the transmitted symbol benefits the clock recovery subsystem, whereas an on–off-keying (OOK) system may suffer a synchronization loss if a sequence of 0's is encountered. Also, if nonreturn-to-zero (NRZ) pulses are used, a sequence of 1's also can disrupt the synchronization subsystem of an OOK system. Unlike the OOK scenario, one does not require a priori knowledge of the signal or background noise radiation levels to implement an optimum PPM receiver. The other key requirement of systems considered for space applications is the peak laser power level that must be large enough to survive huge deep-space losses. For this reason, Q-switched lasers typically are employed for such applications. The current technology, however, does not support a scenario where a Q-switched laser can be toggled between the "on" and "off" states at a high rate,

[1] Erik Jonsson School of Engineering and Computer Science, The University of Texas at Dallas, Richardson, Texas.

[2] Communications Systems and Research Section.

limiting severely the data rate that can be supported using an OOK scheme. Hence, an M-ary PPM with large M is more suitable than its OOK counterpart for deep-space applications.[3] Finally, it can be shown that M-ary PPM for large M is more energy efficient than its OOK counterpart, which is of critical importance in a deep-space optical communication environment where energy consumption is a key constraint of the channel.

For many deep-space applications, due to large distances and pointing inaccuracies, the signal level received at the detector is significantly attenuated so that only a small number of photons per PPM slot interval are typically observed. Since thermal noise is present at the receiver, the detection of signal pulses in the above condition is significantly hampered. For this reason, avalanche photodetectors (APDs) are utilized to boost the signal level over the additive noise level present at the receiver. An APD, in principle, magnifies each incident photon to a large number of post-detection electrons. One major problem with today's APDs is the excess noise factor of the APD, which manifests itself as a random gain effect. That is, the ratio of the number of emitted electrons in response to an incident photon is a random variable with known statistics. Unfortunately, the complicated statistics of the APD response to incident photons do not lead to a closed-form solution for the error rate of optical PPM communication systems.

One key approximation for the statistics of the number of emitted electrons when the received number of photons obeys the Poisson statistics is due to Webb [1]. This approximation is shown to be fairly accurate over a wide range of conditions, and for this reason, this approximation is widely used to assess the error rates of various optical communication systems with APDs. When the number of incident photons in a slot interval is fairly large (as compared with the APD noise figure) and the APD possesses a large average gain, it has been shown that the APD's statistic could be approximated by a Gaussian probability density function (pdf) [2,3]. This approximation leads to a closed-form bit-error-rate expression for the PPM channels and usually is valid when large background radiation levels are present (daytime operation). However, for a large number of cases when the background radiation level is small, the Gaussian approximation fails to accurately predict the statistics of an APD detector. In fact, for a small number of incident photons, the probability mass function (PMF) of the number of released electrons significantly departs from a Gaussian form [2,3]. In that case, one has to resort to the Webb–McIntyre–Conradi (WMC) pdf, which does not lead to a closed-form expression for the probability of error [1]. More significantly, the resulting expression requires a prohibitively large computing time.

In some recent studies, via extensive simulation, it has been demonstrated that when the thermal noise level is non-negligible, the number of signal photons required to achieve a symbol-error rate in a range of 10^{-2} to 10^{-3} using the Gaussian-approximated model and that predicted by the WMC model are somewhat similar for background noise levels in excess of 1 photon per PPM slot. Since, for most applications of interest, one observes in excess of 1 background photon per PPM slot, the Gaussian model seems to be a reasonable model to exploit for the above scenario. This by no means implies that the Gaussian assumption is a valid approximation for all symbol-error-rate levels. It only underscores the negligible loss/gain in the signal photon count associated with the Gaussian assumption for symbol-error rates in the range of 10^{-2} to 10^{-3}. It also is imperative to note that, for most applications, an uncoded symbol-error rate of 10^{-2} to 10^{-3} is of interest. Obviously, with the inclusion of a forward error-correcting code (FEC), a symbol-error rate of 10^{-6}, which is required to support data communication, can be achieved.

Another generic property of optical sensors for detecting narrow laser pulses is the requirement of high-speed electronics to convert optical signals to electrical currents. In general, for deep-space direct-detection applications, larger detector areas potentially could collect more photons, thus increasing the

[3] In an M-ary PPM scenario with large M, the laser remains in an "on" state for $1/M$th of the symbol duration, allowing for $\log_2 M$ bits to be transmitted in a symbol interval. In an OOK scheme, given the same symbol duration (note that there exists an upper limit on how fast the laser can be toggled between the two states), only 1 bit can be transmitted over the same symbol interval.

number of photons per information bit. Unfortunately, larger detector sizes imply a lower bandwidth of the detector, which tends to smear the observed narrow pulses over several PPM slots. The proposed T-PPM excels in this area by allowing overlapping of the PPM signaling set. That is, as will be demonstrated in this article, the proposed T-PPM could circumvent this effect without degrading the system end-to-end performance. Furthermore, it will be shown in the ensuing sections that T-PPM can achieve a performance similar to that of its PPM counterpart at a reduced number of photons per information bit. In addition, T-PPM can accommodate lower peak laser power during slot detection by integrating energies over several slots. The lower peak power is of critical importance in energy-constrained, deep-space communication systems.[4]

In this article, we are interested in assessing the impact of T-PPM on enhancing the end-to-end performance of optical communication systems with pulse-width inaccuracies, which often are present when high-power Q-switched lasers are employed. The presence of such inaccuracies results in a significant degradation in the overall performance of conventional PPM systems.

II. System Model

For an M-ary PPM system, the transmitted laser power, $P(t)$, may be described as

$$P(t) = P_s \sum_{q=-\infty}^{\infty} h(t - C_q T_{\text{slot}} - NqT_s) \tag{1}$$

where P_s is the peak laser power in watts; $h(t)$ is the pulse shape, which is confined to a PPM symbol duration T_s s; T_{slot} is the PPM slot duration in seconds; C_q is the PPM symbol taking on the set $\{0, 1, \cdots, M-1\}$ with equal probability, with M denoting the PPM alphabet size; and $N \geq 1$ is a parameter that will be defined shortly. Note that $T_{\text{slot}} = T_s/M$ s. Ideally, the pulse shape $h(t)$ is confined to a PPM slot duration. In that case, one obtains an ideal PPM signaling scheme. In the scenario considered here, however, it is assumed that $h(t)$ extends beyond a slot interval, causing significant degradation in performance. Also, note that, in the formulation shown above, N is introduced to allow for one to generate a "silent" period in between symbol durations in which a PPM pulse is transmitted. This typically is a requirement for operating a high-power laser when the transmission of a symbol is interrupted by silent periods of several-symbol duration to allow for the recharging of the laser.

Given an unguided optical channel, one can describe the intensity of the received optical signal, $\lambda_r(t)$, as

$$\lambda_r(t) = \lambda_b + \lambda_s \sum_{q=-\infty}^{\infty} h(t - C_q T_{\text{slot}} - NqT_s + \epsilon_q T_{\text{slot}}) \tag{2}$$

where $\lambda_s = P_r \eta / h\nu$ denotes the peak intensity of the received optical signal in photons/s with η, P_r, h, and ν denoting the quantum efficiency of the detector, the received power collected by the optical assembly, Planck's constant, and the operating frequency of the laser in hertz, respectively. Moreover, λ_b in Eq. (2) denotes the intensity of the background noise in photons/s, and ϵ_q denotes the timing inaccuracy present due to imperfect time tracking at the receiver, which is assumed to be confined to $[-1/2, 1/2]$.

[4] For energy efficiency, one must resort to a large alphabet size PPM scenario. In that event, and for a relatively high data rate, one must employ fairly short laser pulses. Given the fixed required energy, this implies that the laser must generate a large peak power. The large peak power then becomes a major constraint of the system.

Since an APD is employed at the receiver, the collected primary electrons cause the generation of a large number of secondary electrons. In general, one can assume that for the mth detected primary electrons, the APD generates G_m secondary electrons. The cumulative generating function (CGF) of G_m is implicitly given by [4]

$$s = \mu_{G_m}(s) - b \ln\left[a + (1-a)\, e^{\mu_{G_m}(s)}\right] \tag{3}$$

where $\mu_{G_m}(s)$ denotes the CGF, $a = [1+\kappa(\bar{g}-1)]/\bar{g}$, and $b = 1/(1-\kappa)$, with $\bar{g}$ and κ denoting the average gain and the ionization factor of the APD, respectively. Let $F = [E\left(G_m^2\right)]/\bar{g}^2 = \kappa\bar{g} + (2 - [1/\bar{g}])(1-\kappa)$ denote the APD noise figure. Then, it can be shown that the moment-generating function of G_m may be expressed explicitly as [5]

$$M_{G_m}(s) = \frac{F}{(F-1)^2}\left[1 - \sqrt{1 - 2(F-1)\bar{g}s}\right] - \frac{\bar{g}}{F-1}s + 1 \tag{4}$$

Note that, by definition, $M_{G_m}(s) = e^{\mu_{G_m}(s)}$.

The response of the APD, in terms of the current in the load resistor of the APD, may be modeled as follows:

$$x(t) = \sum_{j=-\infty}^{\infty} G_j R(t - t_j) + n(t) \tag{5}$$

where t_j denotes the occurrence time of the jth primary electron, governed by a Poisson point process; $R(t)$ is the current response of the APD; and $n(t)$ is the thermal noise present at the receiver. The $n(t)$ commonly is modeled as a zero-mean, additive white Gaussian noise (AWGN) with a two-sided power spectrum density (PSD) level of $N_0/2 = (2kT_0)/R_L$ A^2/Hz with k, T_0, and R_L denoting Boltzmann's constant, the effective temperature of the receiver in kelvins, and the APD load resistance in ohms (Ω).

At the receiver, the PPM detector performs M distinct integrations of the observed current over the M slots of the PPM signal. The symbol associated with the slot with a maximum integrated current is declared as the demodulated symbol. This detection strategy is optimum when the integrated current is governed by Gaussian statistics or when a shot-noise-limited (Poisson-statistics) scenario is considered [2]. It also has been demonstrated that when the statistics of the APD are approximated by WMC distribution and the observed signal is further corrupted by AWGN, the optimum detection strategy is identical to that used for purely Gaussian or Poisson channels [6]. Therefore, we use the detection mechanism outlined above to render a decision regarding the transmitted symbol.[5]

Let us define $X_{l,n}$ as the outcome of integrating $x(t)$ over the interval $I_{l,n} = [lT_{\mathrm{slot}} + NnT_s, (l+1)T_{\mathrm{slot}} + NnT_s]$. Note that $I_{l,n}$ denotes the lth slot of the nth transmitted symbol.[6] We then have

$$X_{l,n} = \int_{lT_{\mathrm{slot}}+NnT_s}^{(l+1)T_{\mathrm{slot}}+NnT_s} x(\tau)\, d\tau; l = 0, 1, \cdots, M-1 \tag{6}$$

[5] In the presence of timing error, the above method is no longer optimal. Since in the ensuing analysis we consider negligible timing error as compared with the pulse-width inaccuracies, we resort to the aforementioned detection mechanism.

[6] In the ensuing analysis, and without loss of generality, we consider $N = 4$ (each transmitted symbol is followed by a silent period of $3T_s$ s). The selection of $N = 4$ is dictated by the requirements that currently are imposed on high-power Q-switched lasers that are being considered for deep-space communication.

Given that the APD's current response is such that it can be approximated with a delta function, we arrive at

$$X_{l,n} = eK_2(l,n) + \nu_N \tag{7}$$

where ν_N is a zero-mean Gaussian random variable with variance

$$\sigma_\nu^2 = \frac{2kT_0}{R_L} T_{\text{slot}} \tag{8}$$

e is the charge of an electron in C, and $K_2(l,n)$ is the number of secondary electrons observed over $I_{l,n}$. The PMF of $K_2(l,n)$ in its exact form originally was discovered by McIntyre [7] and was verified experimentally by Conradi [8]. This PMF is given by

$$\Pr\{K_2(l,n) = k_2 | K_1(l,n) = k_1\} = \frac{k_1 \Gamma\left(\frac{k_2}{1-\kappa} + 1\right)}{k_2 (k_2 - k_1)! \Gamma\left(\frac{\kappa k_2}{1-\kappa} + 1 + k_1\right)} \times \left[\frac{1 + \kappa(\bar{g} - 1)}{\bar{g}}\right]^{k_1 + \kappa k_2/(1-\kappa)} \left[\frac{(1-\kappa)(\bar{g} - 1)}{\bar{g}}\right]^{k_2 - k_1} \tag{9}$$

where $K_1(l,n)$ is the number of primary electrons observed over the interval $I_{l,n}$. Given that the primary electron statistics obey a Poisson model, Webb has demonstrated that the PMF of the secondary electron count may be approximated by

$$\Pr\left\{K_2(l,n) = k_2 | \bar{K}_1(l,n) = \bar{k}_1\right\} = \frac{1}{(2\pi C_1^2)^{1/2}} \left[\frac{1}{1 + \frac{(k_2 - \bar{k}_1 \bar{g})}{C_1 C_2}}\right]^{3/2} \exp\left\{\frac{-(k_2 - \bar{k}_1 \bar{g})^2}{2C_1^2 \left[1 + \frac{(k_2 - \bar{k}_1 \bar{g})}{C_1 C_2}\right]}\right\} \tag{10}$$

where $\bar{K}_1(l,n)$ is the average number of primary electrons observed over $I_{l,n}$, $C_1^2 = \bar{g}^2 \bar{k}_1 F$, and $C_2^2 = \bar{k}_1 F / (F-1)^2$. Given the Poisson statistics, then

$$\bar{K}_1(l,n) = \int_{lT_{\text{slot}} + NnT_s}^{(l+1)T_{\text{slot}} + NnT_s} \lambda_r(\tau)\, d\tau = K_b + K_s Y_{l,n} \tag{11}$$

where

$$Y_{l,n} = \frac{1}{T_{\text{slot}}} \int_{lT_{\text{slot}} + NnT_s}^{(l+1)T_{\text{slot}} + NnT_s} \sum_{q=-\infty}^{\infty} h(\tau - C_q T_{\text{slot}} - NqT_s + \epsilon_q T_{\text{slot}})\, d\tau$$

and $K_b = \lambda_b T_{\text{slot}}$ denotes the average number of primary electrons observed over a PPM slot interval due to background noise. Moreover, $K_s = \lambda_s T_{\text{slot}}$. Note that the above formulation implies that the average number of primary electrons due to signal power observed over the lth slot of the nth symbol, i.e., $K_s Y_{l,n}$, decreases as the pulses smear in time, maintaining the average number of primary electrons observed over a symbol interval (i.e., K_s) constant. That is, we assume that the energy observed over a symbol interval remains constant as pulses spread over the adjacent slots, leading to less observed energy per slot.

In the ensuing analysis, we assume that ϵ_q is negligible (since it is a fraction of a slot interval), whereas the intersymbol interference caused by the shape of the nonideal pulse shape, $h(t)$, is assumed to be quite significant. More precisely, we assume that $h(t)$ extends beyond a slot boundary and perhaps extends over 2 to 4 slot intervals. This causes severe degradation in performance in the absence of a strategy to circumvent intersymbol interference (ISI). In what follows, we assume that the laser pulse obeys a Gaussian shape. Namely,

$$h(t) = \frac{T_{\text{slot}}}{\sqrt{2\pi\sigma_h^2}} \exp\left[-\frac{\left(t - \dfrac{T_{\text{slot}}}{2}\right)^2}{2\sigma_h^2}\right] \tag{12}$$

where we have assumed that the laser pulse is centered around the midpoint of the slot interval and that the pulse has a standard deviation of σ_h s. Moreover, note that $\int_{-\infty}^{\infty} h(\tau)\, d\tau = T_{\text{slot}}$. This assumption is motivated by the fact that if $h(t)$ is replaced with an ideal nonreturn-to-zero (NRZ) pulse of unit amplitude, a similar result will be obtained. This, in turn, implies that regardless of the value of σ_h^2, the total laser energy over a symbol interval remains constant. That is, as the pulses are smeared in time, the peak laser power is reduced, maintaining the constant-energy assumption that is critical to space-borne optical communication systems. Considering that $h(t)$ extends over a few slot intervals and that we have a silent period of $3M$ slots, and assuming that $C_n = j$, we have

$$Z_{l,j} = Y_{l,n}\,|_{C_n=j} = \frac{1}{\sigma_h\sqrt{2\pi}} \int_{lT_{\text{slot}}}^{(l+1)T_{\text{slot}}} \exp\left[-\frac{\left(\tau - \dfrac{(2j+1)\,T_{\text{slot}}}{2}\right)^2}{2\sigma_h^2}\right] d\tau \tag{13}$$

Note that the integral on the right-hand side of the above equation is a function of l and j only.

III. Performance Analysis—PPM With Imperfect Pulse Shape

Given the above model, one can obtain the performance of a PPM system with imperfect pulses. A key obstacle, however, is the absence of a tractable model for the statistics of the APD signal when corrupted by an AWGN. In a recent simulation study, it was demonstrated that the K_s required for an APD-detected 256-ary PPM signal to achieve an error rate in the range of 10^{-2} to 10^{-3} and that predicted using a Gaussian model differ only slightly over a wide range for K_b ($1 \le K_b \le 1000$), allowing one to benefit from the Gaussian approximation for the above range of symbol-error rates. Furthermore, we benefit from a union bound to establish an upper bound on the error rate of the uncoded PPM system with imperfect pulse shapes. To that end,

$$P_{PPM}^{(u)} \le \frac{1}{M} \sum_{j=0}^{M-1} \sum_{l=0;j\neq l}^{M-1} \Pr\left(E_{l,j}\right) \tag{14}$$

where we have assumed equally likely PPM symbols. In the above equation, $\Pr(E_{l,j})$ is the pair-wise error probability of making a decision in favor of the lth symbol when the jth symbol actually is transmitted and $P_{PPM}^{(u)}$ is the symbol-error rate of the uncoded PPM system. This error event is possible only when the integrated current of the lth slot exceeds that of the jth slot. As noted earlier, we resort to Gaussian approximation for the range of error rates stated previously. If one allows for such an approximation, then $X_{l,n}$ is a Gaussian random variable when conditioned on C_q (assume $\epsilon_q = 0$). Hence, the mean and variance of this random variable will be of interest to establish performance. To that end,

$$m_{l,j} = E\{X_{l,n}|C_n = j\} = e\bar{g}(K_b + K_s Z_{l,j}) + I_{dc}T_{\text{slot}} \tag{15}$$

and

$$\sigma_{l,j}^2 = Var\{X_{l,n}|C_n = j\} = (e\bar{g})^2 F(K_b + K_s Z_{l,j}) + \sigma_\nu^2 + eI_{dc}T_{\text{slot}} \tag{16}$$

where $E\{.|C_n = j\}$ and $Var\{.|C_n = j\}$ are the expected value and the variance of the random variable $X_{l,n}$ conditioned on $\{C_n = j\}$ and I_{dc} denotes the surface dark current of the APD in amperes, which is assumed to be non-negligible here. In that event, and considering that an erroneous decision in favor of the lth symbol is rendered if the integrated APD current over the lth slot exceeds that of the jth slot (note that this is a comparison between $X_{l,n}$ and $X_{j,n}$, which are a pair of conditionally [when conditioned on $\{C_n = j\}$] independent Gaussian random variables), we have

$$\Pr(E_{l,j}) = \frac{1}{2}\,\text{erfc}\left(\frac{m_{j,j} - m_{l,j}}{\sqrt{2\left(\sigma_{l,j}^2 + \sigma_{j,j}^2\right)}}\right); l \neq j \tag{17}$$

where $\text{erfc}(x)$ is the complementary error function. Given that the pulse shape follows a Gaussian pattern, $Z_{l,j}$ and $Z_{j,j}$ may be obtained in terms of erfc(.) or erfc(.). However, we note that

$$m_{j,j} - m_{l,j} = e\bar{g}K_s(Z_{j,j} - Z_{l,j}) \tag{18}$$

and

$$\sigma_{l,j}^2 + \sigma_{j,j}^2 = (e\bar{g})^2 F[2K_b + K_s(Z_{l,j} + Z_{j,j})] + 2\left(\sigma_\nu^2 + eI_{dc}T_{\text{slot}}\right) \tag{19}$$

are functions of $Z_{j,j} - Z_{l,j}$ and $Z_{j,j} + Z_{l,j}$ only. Therefore, $Z_{j,j} - Z_{l,j}$ and $Z_{j,j} + Z_{l,j}$ are only of interest. With some effort, it may be shown that

$$Z_{j,j} - Z_{l,j} = 1 - \text{erfc}\left(\frac{T_{\text{slot}}}{2\sqrt{2}\sigma_h}\right) - \frac{1}{2}\text{erfc}\left(-\frac{T_{\text{slot}}\left(l - j + \frac{1}{2}\right)}{\sqrt{2}\sigma_h}\right) + \frac{1}{2}\text{erfc}\left(-\frac{T_{\text{slot}}\left(l - j - \frac{1}{2}\right)}{\sqrt{2}\sigma_h}\right) \tag{20}$$

and

$$Z_{j,j} + Z_{l,j} = 1 - \text{erfc}\left(\frac{T_{\text{slot}}}{2\sqrt{2}\sigma_h}\right) + \frac{1}{2}\text{erfc}\left(-\frac{T_{\text{slot}}\left(l-j+\frac{1}{2}\right)}{\sqrt{2}\sigma_h}\right) - \frac{1}{2}\text{erfc}\left(-\frac{T_{\text{slot}}\left(l-j-\frac{1}{2}\right)}{\sqrt{2}\sigma_h}\right) \quad (21)$$

Now, substituting Eqs. (20) and (21) in Eqs. (18) and (19), respectively, and then substituting the resulting expressions in Eq. (17), we arrive at a closed-form expression for the pair-wise error rate of a PPM system with pulse-width inaccuracies. The remaining task is to compute the upper bound using Eq. (14). We postpone a discussion on the performance of the system using Eq. (14) to Section VI of this article. Instead, we proceed to describe the proposed T-PPM system in the next section.

IV. T-PPM

It is important to note that the use of trellis-coded modulation (TCM) and overlapping PPM (OPPM) to enhance the capacity of PPM channels was originally discussed in [9]. There are several aspects of the present analysis that set it apart from its predecessor. First, the analysis in [9] assumes square shape pulses, whereas we will concern ourselves with Gaussian-type pulse shapes in this analysis. Second, the analysis in [9] assumes either quantum-limited or shot-noise-limited scenarios, whereas a more realistic situation (as is assumed in what follows) calls for the inclusion of additive Gaussian noise and APD excess noise factors.

In the previous section, we described the performance of a PPM system impaired by imperfect pulse shapes. Needless to say, the impact of an imperfect pulse shape on the performance of a PPM system may be quite severe for pulse shapes extending over several slot intervals. Hence, the introduction of TCM is necessary to overcome the impact of severe ISI caused by imperfect pulse shapes. Due to the regular shape of the pulse, and from Eq. (17), it can readily be seen that $\Pr(E_{l,j})$ is a decreasing function of $|j-l|$. That is, the pair-wise error rate is a decreasing function of the "distance" between symbols j and l, where distance is defined in terms of the number of PPM slots that exit between the positions (in time) of the jth and lth symbols.[7] In view of this observation, one can proceed with a set-partition strategy for the generation of T-PPM signals that is aimed at increasing the minimum distance of the PPM constellation.

We begin by introducing $A = \{1, 2, 3, \cdots, M\}$ as a set containing the uncoded PPM symbols. Note that the distance between a pair of symbols j and l selected from A is simply $|j-l|$. This set is further divided into sets $B_0 = \{1, 3, 5, 7, \cdots\}$ and $B_1 = \{2, 4, 6, \cdots\}$. Note that a pair of symbols selected from either one of these sets is at a minimum distance of 2, whereas the symbols selected from A are at a minimum distance of 1. Next, we proceed to subdivide set B_0 to generate sets $C_0 = \{1, 5, 9, \cdots\}$ and $C_2 = \{3, 7, 11, \cdots\}$. Set B_1 also can be subdivided to form $C_1 = \{2, 6, 10, \cdots\}$ and $C_3 = \{4, 8, 12, \cdots\}$. Finally, we perform one last partitioning of the previous sets to form $D_0 = \{1, 9, 17, \cdots\}$, $D_4 = \{5, 13, 21, \cdots\}$, $D_2 = \{3, 11, 19, \cdots\}$, $D_6 = \{7, 15, 23, \cdots\}$, $D_1 = \{2, 10, 18, \cdots\}$, $D_5 = \{6, 14, 22, \cdots\}$, $D_3 = \{4, 12, 20, \cdots\}$, and $D_7 = \{8, 16, 24\}$. The entire process is depicted in Fig. 1 for $M = 256$. Note that symbols selected from D_j for any j are at a minimum distance of 8. Furthermore, symbols selected from D_j and D_i are at a minimum distance of $|i-j|$ from each other. Although it appears that one can continue this process in hope of increasing the minimum distance beyond 8, the minimum distance between a pair of paths through the trellis (which ultimately dictates the overall performance) cannot be increased indefinitely with further set partitioning. Also, since we are interested only in eliminating the impact of ISI, a minimum distance of 4 is sufficient to establish orthogonality among the symbols in the set. That is, when the imperfect pulse $h(t)$ stretches over 2 to 4 PPM slots, a minimum distance of 4 slots among the symbols selected from C_j (for all j) ensures that there exists no overlap among the pulses in the set.

[7] Since the jth symbol is represented by a pulse in the jth slot, the distance between the jth and lth symbols is $|j-l|$.

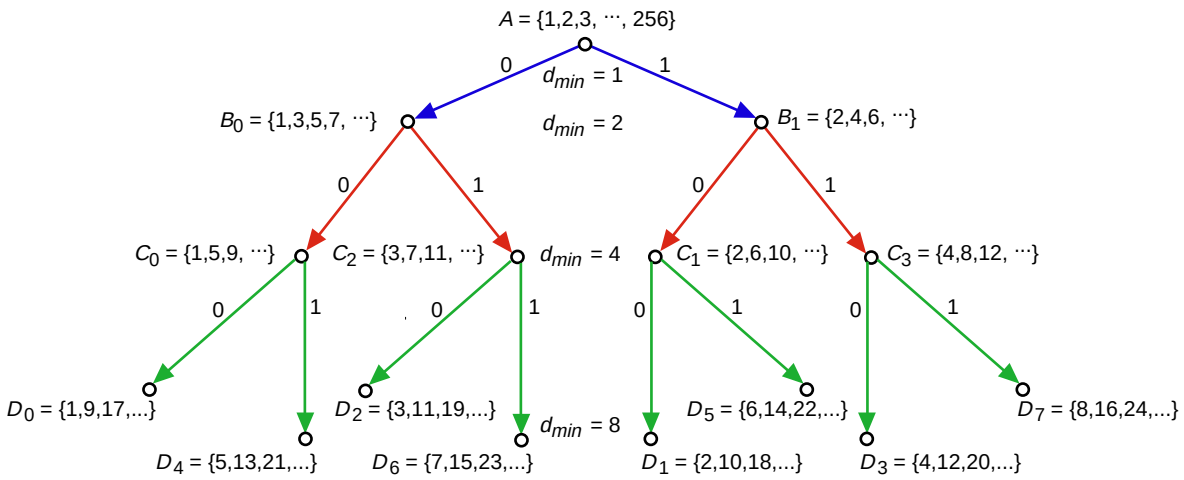

Fig. 1. Set partitions for the 256-ary PPM signal set.

Without loss of generality, we concern ourselves with rate 1/2 and 2/3 convolutional encoders (CEs) for the generation of T-PPM signals. As will be shown later, the rate 1/2 CE with a 4-state trellis is quite suitable for the problem at hand, although the rate 2/3 CE with an 8-state trellis can offer a larger minimum distance ($d_{\min}$). We note also that the alphabet size of the PPM is 256 (8 bits), which allows one to encode 7 data bits per T-PPM symbol (for both the rate 1/2 and rate 2/3 CEs, see Figs. 2 and 3). Given that for high-power lasers used in deep-space communications the frame rates usually are kept constant, this implies a loss in the data rate by a factor of 7/8. This, however, is a relatively small price to pay to circumvent the substantial ISI caused by imperfect pulses. The rate 1/2 and 2/3 convolutional encoders along with their respective trellises are depicted in Figs. 2 and 3, respectively. Note that, for the rate 1/2 CE case, a 4-state trellis is used and that the 2 bits produced by the encoder are used to select a set from the four possible C-type sets: C_0, C_1, C_2, and C_3. The remaining 6 bits are used to select a signal from the selected set (note that there are 2^6 signals in any of the four C-type sets). For the rate 2/3 code, which, as shown in Fig. 3, corresponds to an 8-state trellis, the 3 bits produced by the CE are used to select one of the eight possible D-type sets. Given that the code is a rate 2/3 CE, then the remaining 5 bits are used to select a signal from the selected set. Once again, note that there are 2^5 signals in any of the eight D-type sets. The above arrangement, then, leads to the existence of a number of parallel paths in the trellises depicted in Figs. 2 and 3. That is, for each path in the trellis shown in Fig. 2, there exist 64 parallel paths (the size of the C-type sets). This is a common feature of TCM systems that is due to the existence of uncoded bits used to select a signal from the constellation. The number of parallel paths for the trellis in Fig. 3 is 32, which is identical to the size of any of the D-type sets.

V. T-PPM Performance Analysis

In Figs. 2 and 3, we have depicted the trellises for the rate 1/2 and 2/3 T-PPM systems with 4 and 8 states, respectively. When the rate 1/2, 4-state trellis is used, one easily can conclude that the minimum distance between the all-zero path and any path that departs from and re-emerges with the all-zero path for the first time is 4. The path that leads to the minimum distance of 4 is due to a symbol selected from C_0. For the 8-state trellis, we also have depicted the path (dotted line) that is at a minimum distance of 5 from the all-zero path. Note that parallel transitions in this trellis are selected from the D-type sets, and, hence, a minimum distance of 8 exists between any pair of parallel paths in Fig. 3. As noted earlier, although we have been able to increase the minimum distance between parallel paths in the trellis via further set partitioning, the minimum distance between any path in the trellis and the all-zero path has not increased dramatically. Due to the complexity of the trellis in Fig. 3 and the fact that a

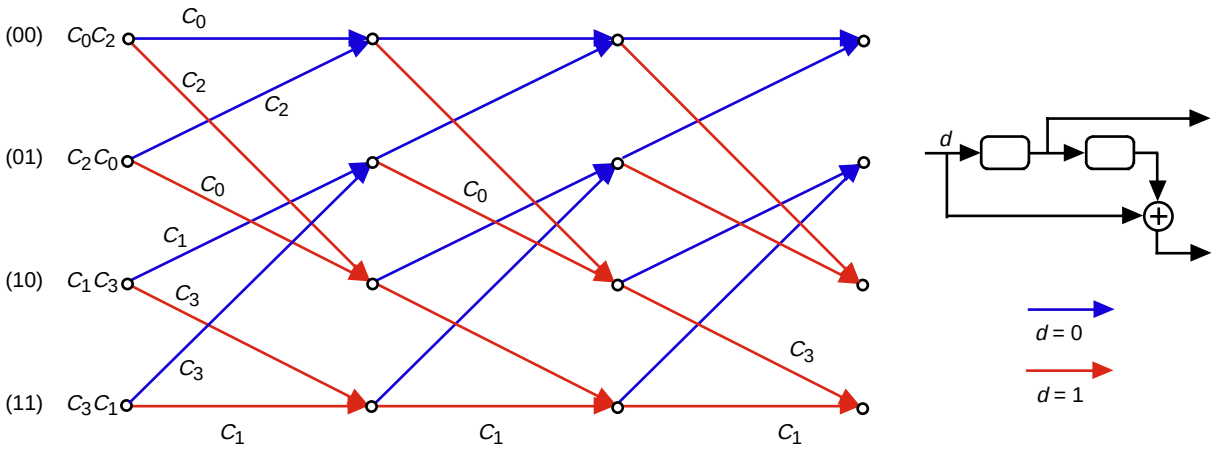

Fig. 2. The rate 1/2 convolutional encoder along with its 4-state trellis that uses the set partitioning of Fig. 1 for the generation of the rate 7/8 T-PPM signals.

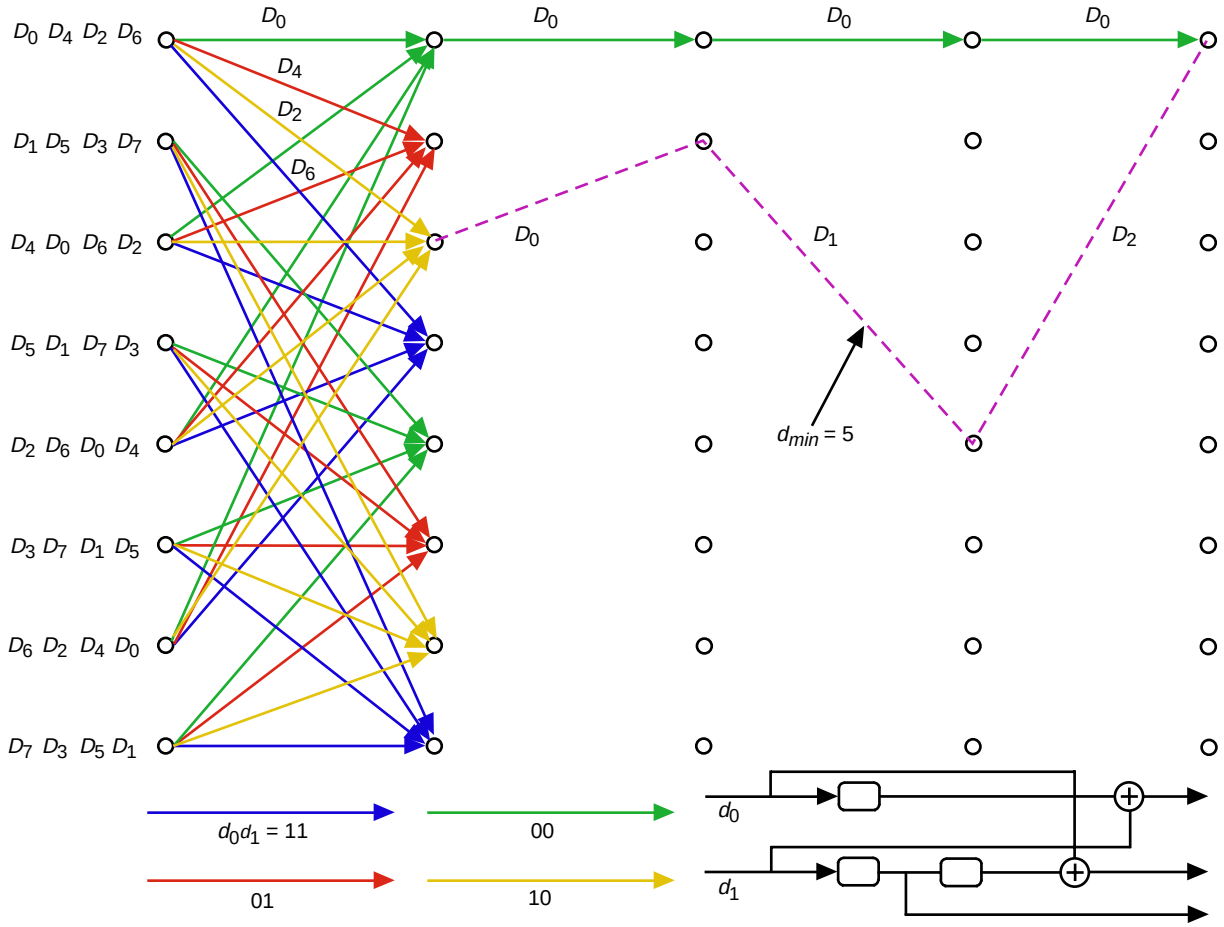

Fig. 3. The rate 2/3 convolutional encoder along with its 8-state trellis that uses the set partitioning of Fig. 1 for the generation of the rate 7/8 TCM–PPM signals.

minimum distance of 4 is sufficient to establish orthogonality among the symbols in the set, in what follows, we limit our discussion to the rate 1/2, 4-state trellis depicted in Fig. 2. To that end, let us consider the parallel paths in the trellis. Given that for any path there exist 63 other parallel paths for $M = 256$ in Fig. 2 (see Fig. 1), an upper bound on the error rate may be obtained with the aid of a union bound. This upper bound is given by

$$P_{PPM}^{(TCM)} \leq 63 \Pr\left(E_{l,l+4}\right)$$

where $\Pr\left(E_{l,l+4}\right)$ denotes the pair-wise error rate for a pair of PPM symbols with imperfect pulse shapes that are separated by 4 time slots. When $h\left(t\right)$ extends over 2 to 4 times the duration of a slot, then $\Pr\left(E_{l,l+4}\right)$ is not a function of l.

VI. Numerical Results

To underscore the impact of imperfect pulse shapes, a pair of adjacent Gaussian-type pulses are plotted in Figs. 4 through 6. The laser energy (the area covered by the pulse) within one slot time is computed to be 98, 78.8, and 59.5 percent, respectively, in Figs. 4 through 6. Note that, for a slot duration of 20 ns, for instance, the case depicted in Fig. 6 corresponds to an 80-ns pulse scenario. It is immediately obvious that ISI increases in an exponential fashion with σ_h. Since we are interested in a scenario where $h\left(t\right)$ is limited to 2 to 4 times the duration of a time slot, we limit our analysis to the cases where $0 \leq \sigma_h \leq 0.6T_{\text{slot}}$ s.

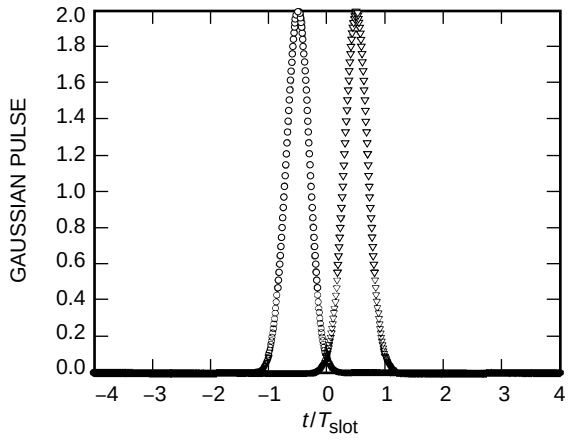

Fig. 4. A pair of adjacent Gaussian-type pulses with $\sigma_h = 0.2T_{\text{slot}}$.

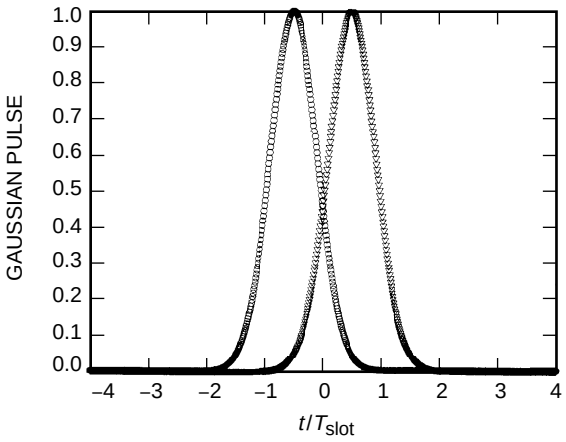

Fig. 5. A pair of adjacent Gaussian-type pulses with $\sigma_h = 0.4T_{\text{slot}}$.

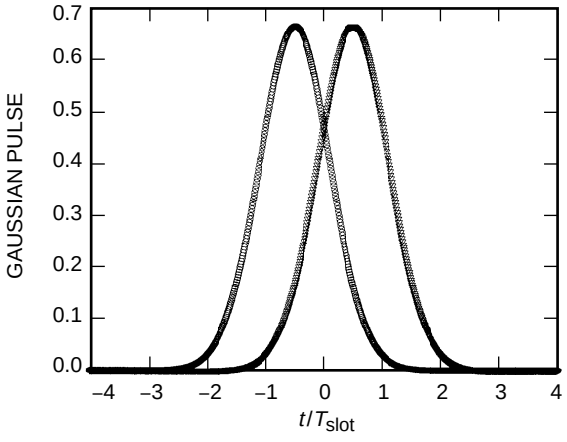

Fig. 6. A pair of adjacent Gaussian-type pulses with $\sigma_h = 0.6T_{\text{slot}}$.

We first consider the case where a symbol-error rate of 10^{-2} is of interest, and subsequently we obtain the required K_s to achieve the desired error rate for a given K_b. Without loss of generality, we limit our discussion to the following system parameters: $M = 256$, $T_{\text{slot}} = 20$ ns, $\bar{g} = 40$, $\kappa = 0.007$, $T_0 = 100$ K (cooled receiver), $R_L = 146.65$ kΩ, and $I_{dc} = 2$ nA. In Figs. 7 and 8, the K_s required to achieve a symbol-error rate of 10^{-2} is plotted versus K_b when the above set of parameters is used. We note that the Gaussian assumption considered here does not lead to performance measures that agree with the simulation results using WMC statistics for the entire range of K_b when an error rate of 10^{-2} is of interest. In particular, when $K_b < 50$ (Fig. 7), the Gaussian assumption is less reliable than for the case of $K_b > 50$ (Fig. 8). Nonetheless, the loss in K_s using the Gaussian approximation is not significant, and hence we proceed with our analysis using the results shown above. Before doing so, however, it is imperative to note that the simulation results for the required K_s using WMC statistics (for the above set of parameters) are shown to be in close agreement with those predicted using the Gaussian assumption for a wide range of K_b when a symbol-error rate of 10^{-3} is considered [10]. Since an error rate of 10^{-3} is typically of interest, we consider the numerical results shown below for a 10^{-3} error rate to be a good approximation of the results obtained using the more realistic WMC statistics for modeling the APD output statistics.

The first observation that can be made from Figs. 7 and 8 is that, when a severe ISI is present (see the curves associated with 80- and 60-ns pulses), the performance degrades substantially as compared with the perfect pulse-shape case. The second significant observation is that, in the event of having perfect pulses, the performance improves using T-PPM. This may be attributed to the coding gain. Note that 1 of the 7 information bits almost always will be detected correctly, since the major error is due to parallel transitions in the set.

We also note that the gain in performance as a result of using T-PPM increases almost exponentially with an increase in σ_h. In Figs. 9 and 10, a similar set of results is depicted when an error rate of 10^{-3} is of interest. We focus on these results, since an error rate of 10^{-3} typically is needed to ensure a coded performance[8] of 10^{-6}.

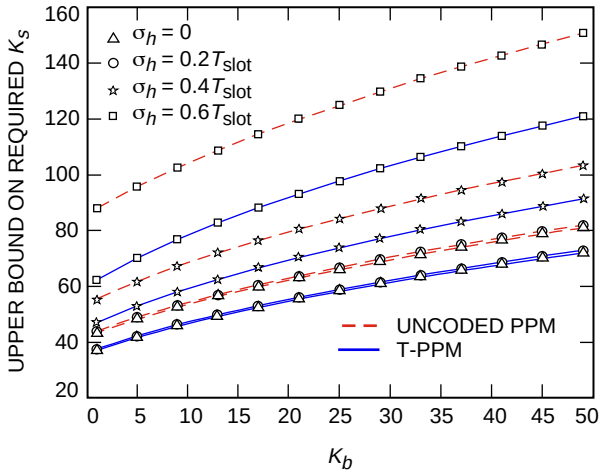

Fig. 7. The K_s required to achieve a symbol-error rate of 10^{-2} for 256-ary PPM with imperfect pulses when K_b is less than 50.

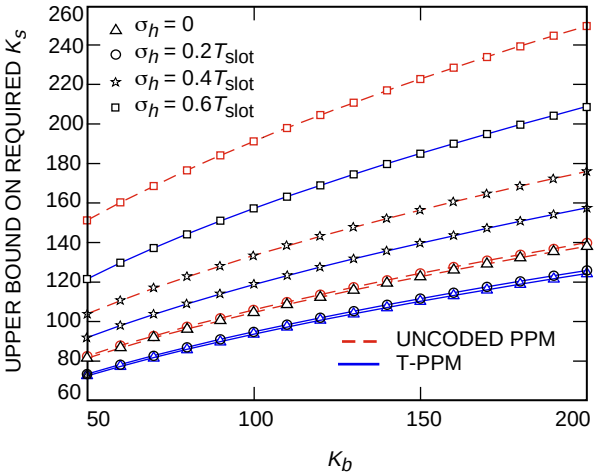

Fig. 8. The K_s required to achieve a symbol-error rate of 10^{-2} for 256-ary PPM with imperfect pulses when K_b is more than 50.

[8] It is anticipated that T-PPM will be used in conjunction with other more powerful coding schemes. That is, the information bits provided to T-PPM may be viewed as coded channel symbols provided by an FEC encoder (such as a Reed–Solomon encoder).

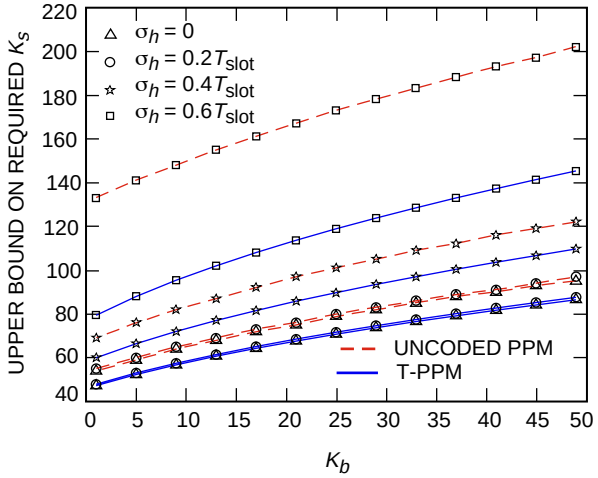

Fig. 9. The K_s required to achieve a symbol-error rate of 10^{-3} for 256-ary PPM with imperfect pulses when K_b is less than 50.

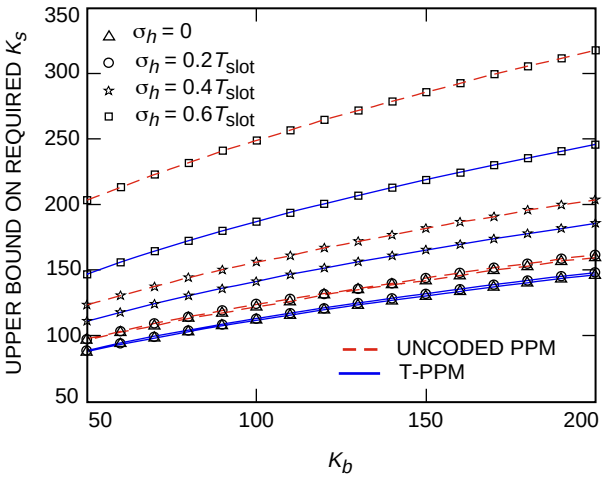

Fig. 10. The K_s required to achieve a symbol-error rate of 10^{-3} for 256-ary PPM with imperfect pulses when K_b is more than 50.

Before discussing such results, it is imperative to note that the proposed T-PPM system transmits 7 information bits, as compared with the 8 bits that are transmitted over each symbol duration for the uncoded system. Hence, it is appropriate to consider the required number of photons per information bit in comparing the T-PPM scenario with the uncoded PPM case. We then proceed to use the average number of required photons per information bit as a measure of efficiency of the modulation scheme. Hence, in Figs. 11 through 14, we depict the average number of photons per information bit that is required to achieve the desired error rate for a wide range of background radiation photon counts. We note, however, that the loss of 1 bit in order to achieve the desired performance, given that there exists a limit on laser energy, is a reasonable consideration. Hence, we discuss the performance of T-PPM using both the average number of photons as well as the average number of photons per information bit as a function of background radiation level in what follows.

From Figs. 9 and 10, it can be concluded that when $K_b = 1$ (night operation), one requires $K_s = 133$ to achieve an error rate of 10^{-3} when 80-ns pulses are used (20-ns PPM slots). This number is reduced to $K_s = 79$ with T-PPM, a reduction of about 2.2 dB in the required laser energy. However, as noted earlier, T-PPM conveys only 7 bits of information. If one uses the number of photons per bit, in that case the required number of photons per bit reduces from $133/8 = 16.6$ to $79/7 = 11.2$ using T-PPM (see Fig. 13). This is a substantial gain in the overall system efficiency. For $K_b = 100$, the K_s required to achieve an error rate of 10^{-3} when 80-ns pulses are used reduces from 249 to 187 using T-PPM, a reduction of nearly 1.2 dB in the required laser energy. The number of photons per bit improves from 31.3 to 26.7 using T-PPM (see Fig. 14). It is important to note that for small pulse spreading (60-ns or smaller pulses), the gain in performance is not substantial and, hence, the gain achieved using T-PPM is noticeable only when substantial pulse-width inaccuracies are present. In fact, for all cases considered, the system efficiency in terms of the number of photons per bit increases or remains the same when the pulse width is equal to or less than 60 ns. However, when the pulse width is increased to 80 ns, a substantial improvement in system efficiency is observed (see Figs. 11 through 14).

When an error rate of 10^{-2} is of interest and $K_b = 200$, from Fig. 8 one can conclude that $K_s = 250$ photons is needed[9] when 80-ns pulses are used. This number is reduced to 209 photons for a T-PPM system, a savings of 0.77 dB in the required laser power. The number of photons per bit improves from 31.2 to 29.8 (see Fig. 12), a gain smaller than that observed for the 10^{-3} error-rate scenario.

[9] Note that, for this large background radiation level, the Gaussian assumption is fairly accurate.

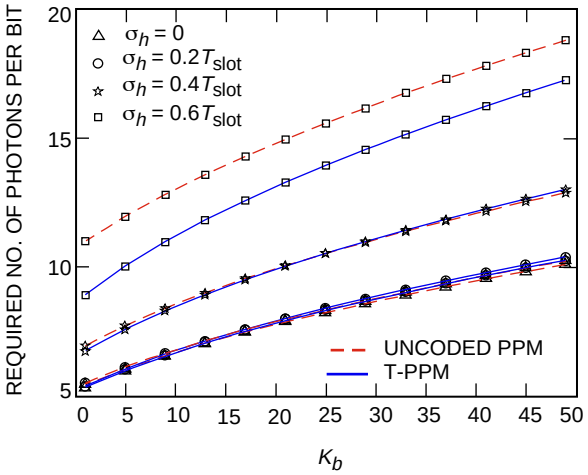

Fig. 11. The average number of photons per information bit required to achieve a symbol-error rate of 10^{-2} for 256-ary PPM with imperfect pulses when K_b is less than 50.

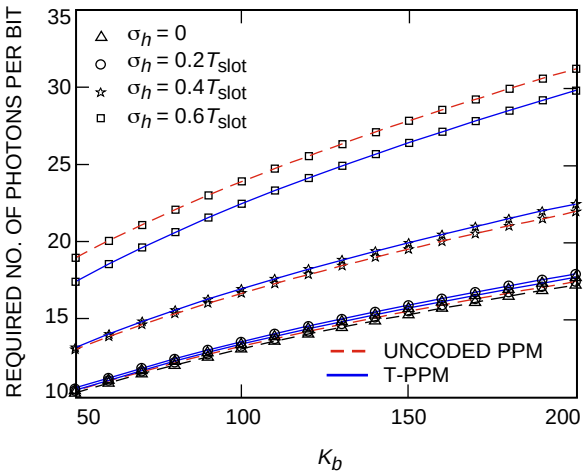

Fig. 12. The average number of photons per information bit required to achieve a symbol-error rate of 10^{-2} for 256-ary PPM with imperfect pulses when K_b is more than 50.

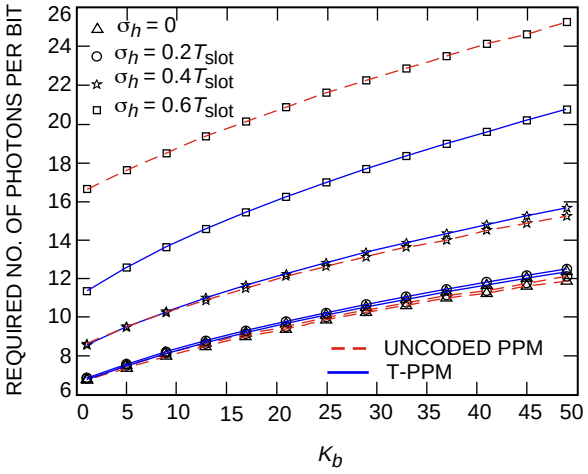

Fig. 13. The average number of photons per information bit required to achieve a symbol-error rate of 10^{-3} for 256-ary PPM with imperfect pulses when K_b is less than 50.

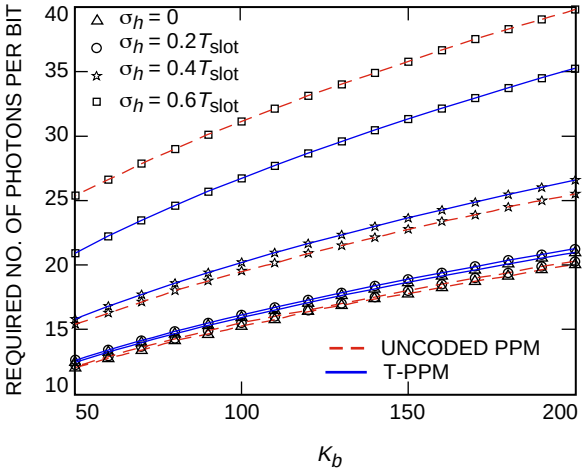

Fig. 14. The average number of photons per information bit required to achieve a symbol-error rate of 10^{-3} for 256-ary PPM with imperfect pulses when K_b is more than 50.

Finally, in Figs. 15 and 16, we examine the performance of T-PPM when the slot duration is reduced with the remaining parameters kept fixed. Given that the contribution of thermal noise and APD dark current to the output of the APD have variances that are directly proportional to the integration interval (slot duration), as one decreases the slot duration, a gain in performance (in terms of a reduction in the number of photons per information bit required to achieve a given error rate, 10^{-3}) is observed. More significantly, as one decreases the slot duration, for a fixed-background-radiation intensity level, the average number of received photons per slot decreases accordingly. As seen from these figures, for the 2-ns slot scenario, a significant improvement in performance is observed. This result is not surprising. To elaborate, as one decreases the slot duration with the total number of signal photons per slot kept constant (i.e., when a higher-power laser is utilized), the performance approaches that of a quantum-limited system, leading to the considerable gain in performance observed in Figs. 15 and 16.

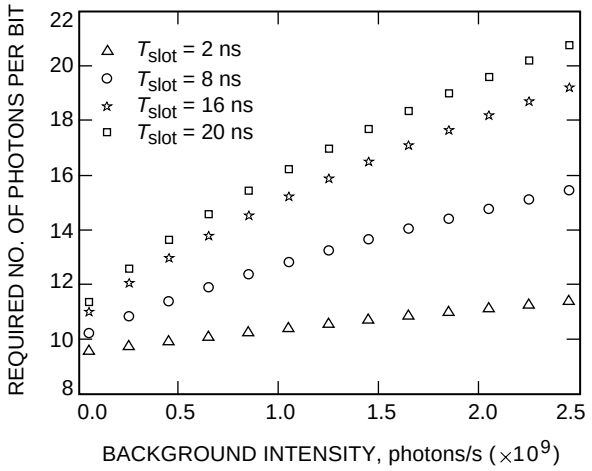

Fig. 15. The average number of photons per information bit required to achieve a symbol-error rate of 10^{-3} for T-PPM with imperfect pulses (background intensity less than 2.5×10^9 photons/s). For all cases, $\sigma_h = 0.6T_{slot}$.

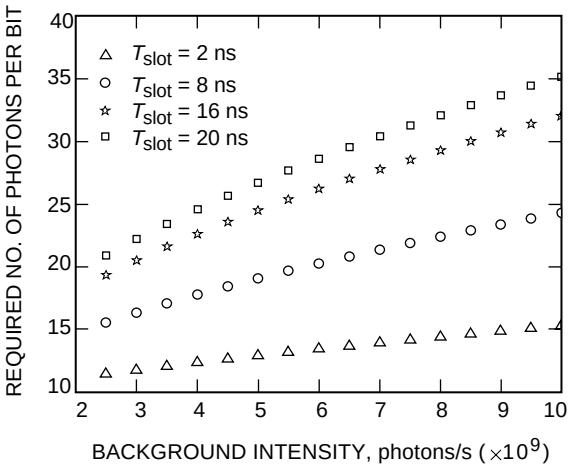

Fig. 16. The average number of photons per information bit required to achieve a symbol-error rate of 10^{-3} for T-PPM with imperfect pulses (background intensity more than 2.5×10^9 photons/s). For all cases, $\sigma_h = 0.6T_{slot}$.

VII. Conclusions

This article introduced a robust trellis-based pulse-position modulation (T-PPM) as a technique for deep-space optical communication and analyzed its performance using union bounds. The analysis assumes the use of a maximal-likelihood receiver for demodulation while the signaling pulses are allowed to extend over several PPM slots. It has been shown that, using a simple convolutional encoder at the transmitter and a Viterbi algorithm at the receiver, T-PPM restores the performance losses due to reduced intensity during the detection process.

Furthermore, using the average number of photons per information bit as a performance measure, T-PPM requires less energy than its regular PPM counterpart by affording a smaller PPM slot width. Numerical examples show that, for a symbol error of 10^{-3} when the received pulses extend over 4 PPM slots, the average laser energy per symbol for 255-ary T-PPM could be reduced by as much as 2 dB. In addition, the increase in the transmitter efficiency could be more profound if the pulse-width duration became narrower.

References

[1] P. P Webb, "Properties of Avalanche Photodiodes," *RCA Rev.*, vol. 35, pp. 234–278, June 1974.

[2] R. M. Gagliardi and S. Karp, *Optical Communications*, New York: John Wiley and Sons, Inc., 1976.

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

[4] S. D. Personick, "Statistics of a General Class of Avalanche Detectors With Applications to Optical Communication Systems," *Bell System Technical Journal*, vol. 50, no. 10, pp. 3075–3095, December 1971.

[5] J. T. K. Tang and K. B. Letaief, "The Use of WMC Distribution for Performance Evaluation of APD Optical Communication Systems," *IEEE Transactions on Communications*, vol. 46, no. 2, pp. 279–285, February 1998.

[6] 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.

[7] R. J. McIntyre, "The Distribution of Gains in Uniformly Multiplying avalanche Photodiodes: Theory," *IEEE Transactions on Electron. Devices*, vol. ED-19, pp. 703–713, June 1972.

[8] J. Conradi, "The Distribution of Gains in Uniformly Multiplying Avalanche Photodiodes: Experimental," *IEEE Transactions on Electron. Devices*, vol. ED-19, pp. 714–718, June 1972.

[9] C. N. Georghiades, "Some Implications of TCM for Optical Direct-Detection Channels," *IEEE Transactions on Communications*, vol. 37, no. 5, pp. 481–487, May 1989.

[10] 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