An Analysis of Heterodyne Pulse-Position Modulation Communication Systems Over Unguided, Turbulent Optical Channels

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

This article investigates the performance of the heterodyne detection scheme for optical pulse-position modulation (PPM) communications channels when nonideal laser transmitters are employed. Two aspects of this channel will be considered. First, the limitations of unguided, turbulent channels in accommodating coherent laser communications are addressed. This is followed by an investigation of the performance of a heterodyne optical PPM communication system. It is demonstrated here that, due to the presence of laser phase noise, noncoherent (or asynchronous) RF detection is a preferred mode of detection and that the overall performance is severely hampered in the presence of phase noise. The performance of the proposed receiver is compared with that of the avalanche photodiode detector (APD)-based direct-detection receiver. It is shown here that the APD-based detector outperforms its heterodyne counterpart across a wide range of background radiation levels.

I. Introduction

A preferred means of modulation for optical direct-detection channels is M-ary pulse-position modulation (PPM). In this article, we will examine the feasibility (from an error-rate standpoint) of using heterodyne detection when the modulation scheme is a 256-ary PPM. We consider the case when the local laser oscillator is strong enough to allow a shot-noise-limited scenario at the receiver and when phase noise due to laser instability is non-negligible. In this article, we make a distinction between "coherent" or "synchronous" detection and optically coherent detection. The former refers to a scenario wherein phase coherency can be established at the receiver (this may refer to the case when coherent detection is employed after optical detection), whereas the latter refers to a scenario wherein optical heterodyne or homodyne has been employed to convert the optical signal into its RF or baseband counterpart. In this case, phase coherency may be established at the receiver. For instance, it is quite possible that one can use optically coherent (heterodyne) detection for converting an M-ary frequency shift-keying (FSK)-modulated optical signal into its RF counterpart and use a noncoherent detection mechanism to recover the data symbol.

[1] Communications Systems and Research Section.

The other key factor in the design of unguided coherent optical channels is the atmospheric effect. In the following, we first describe the impact of atmosphere on the characteristics of a coherent optical beam. We follow this with a study of the performance of a heterodyne receiver for the detection of coherent PPM pulses.

II. Channel Effects (Weak Turbulence)

A turbulent channel has a profound impact on the quality of an optical beam. In the following, we consider a "weak" turbulent scenario wherein coherent communication can be considered. For the scenario when turbulence is considered to be strong, rather predictably the performance of the channel is substantially compromised, and, hence, a reliable link may not be established.

We begin by considering the spatial coherence function of an optical field. Let $\Gamma(\underline{r_1}, \underline{r_2})$ denote the mutual coherence function (MCF) of the field in the turbulent medium. That is,

$$\Gamma(\underline{r_1}, \underline{r_2}) = E\left\{f(\underline{r_1}) f(\underline{r_2})\right\}$$

where $E\{\}$ denotes the expected value of the enclosed and $f(\underline{r})$ denotes the electric field at $\underline{r}$ (a point in space is identified by a vector). In the case of weak turbulence, and using Tatarski's model [1], one can demonstrate that [1]

$$\Gamma(\underline{r_1}, \underline{r_2}) = A^2 \exp\left(-\left(\frac{\underline{\rho}}{\rho_0}\right)^{5/3}\right)$$

where $\rho_0 = [1.45k^2 \int_0^L C_n^2(x)dx]^{-3/5}$ for plane waves, $\rho_0 = [1.45k^2 \int_0^L C_n^2(x/L)dx]^{-3/5}$ for spherical waves, and $\underline{\rho} = \underline{r_1} - \underline{r_2}$. Furthermore, A denotes the amplitude of the electric field; $C_n^2(x)$ is the index of the refraction structure constant; L is the length of the turbulent medium; and $k = (2\pi)/\lambda$, with λ denoting the wavelength of the laser. For deep-space applications, one can assume that the wave impinging upon the receiver is a plane wave, and, hence, $\rho_0 = [1.45k^2 \int_0^L C_n^2(x)dx]^{-3/5}$ must be used in the above equation. This scenario is encountered in the downlink of a deep-space channel. Quite often, $r_0 = \rho_0 3.44^{5/3}$ [i.e., $\Gamma(\underline{r_1}, \underline{r_2}) = A^2 \exp(-3.44(\underline{\rho}/r_0)^{5/3})$] is used in signal-to-noise-ratio (SNR) calculation of optical heterodyne systems. This issue will be discussed later. This formulation allows for the computation of the MCF when the index of refraction structure constant is known or can be approximated. For instance, when the downlink channel (spacecraft to ground) is considered, the field may be considered as a plane wave. In that case, it has been shown that [1]

$$r_0 = 0.05\lambda^{6/5} \cos^{3/5}(\phi) \left[\Gamma\left(\frac{2}{3}\right) \Gamma\left(\frac{2}{3}\,\frac{H}{h_0}\right)\right]^{3/5}$$

where H and h_0 denote the height of the receiver and the height of the turbulent atmosphere, respectively, and $\Gamma(\alpha)$ and $\Gamma(\alpha, \beta)$ denote gamma and incomplete gamma functions, respectively. To arrive at the above, we have assumed that the index of refraction structure constant follows the Hufnagel and Stanley models [1].

For the uplink (i.e., from the ground to the deep-space spacecraft), however, one is required to consider a spherical wave since the turbulence is in the vicinity of the transmitter.

III. Receiver SNR

The electric field at the aperture of a heterodyne optical receiver may be described as

$$f_a(\underline{r},t) = a_s e^{j\omega_c t+\phi} f_a(\underline{r})$$

where a_s and ϕ denote the amplitude and phase, respectively, of the modulated optical field and $f_a(\underline{r},t)$ denotes the time-dependent electric field at the aperture of the receiver in the absence of background noise. For most practical cases of interest, background noise is present at the receiver and, hence, the effect of such a noise process must be included in the analysis. The electric field then is processed by the receiver optics, which consequently produces the following field at the detector area prior to heterodyning:

$$f_d(\underline{r},t) = a_s e^{j\omega_c t} f_d(\underline{r}) + b(t) e^{j\omega_c t} f_b(\underline{r})$$

where $b(t)$ and $f_b(\underline{r})$ are the temporal and spatial components, respectively, of the background radiation at the detector area and $f_d(\underline{r})$ is the spatial component of the field at the detector area prior to heterodyning. Prior to detection, the optical field is mixed with the local laser field. Let

$$f_L(\underline{r},t) = a_L e^{j(\omega_c+\omega_{IF})t} f_L(\underline{r})$$

denote the local laser field in the detector field. It is assumed here that the center frequency of the local laser is at a frequency that is different from the received field. The difference frequency is known as the IF frequency, ω_{IF}. Assuming that the local laser possesses a power level substantially greater than those of the received signal and background noise and that the photodetector has a unity gain, the current at the output of the photodetector may be described as [1]

$$i_d(t) \approx e\alpha a_L^2 \int_{A_d} |f_L(\underline{r})|^2 \, dr + 2e\alpha a_L a_s \left[\int_{A_d} f_d(\underline{r}) f_L^*(\underline{r}) dr\right] \cos(\omega_{IF} t + \phi)$$

$$+ 2e\alpha a_L b(t) \left[\int_{A_d} f_b(\underline{r}) f_L^*(\underline{r}) dr\right] \cos(\omega_{IF} t) + n_{sn}(t) + n_c(t) \tag{1}$$

where e is the charge of an electron, $\alpha = \eta/(hf_c)$, with h, η, and f_c denoting Planck's constant, the quantum efficiency of the detector, and the frequency of the laser, respectively; $n_{sn}(t)$ denotes the shot-noise process at the output of the detector; $n_c(t)$ is the circuit noise due to the electronic processing that follows the optical detection; and A_c denotes the detector area. The circuit noise typically is suppressed by the strong local laser while the dc term at the output of the photodetector is removed by the bandpass filter that follows the photodetector. The shot-noise process has a power-spectrum level (one-sided) of

$$N_{sn} = e^2 \alpha a_L^2 \int_{A_d} |f_L(\underline{r})|^2 d\underline{r}$$

(this is obtained assuming that the local laser field is far stronger than any other field present and, hence, the effective power of the combined field is due to the local laser).

At this stage, assuming that the local field is matched to the capture field at the detector area [i.e., $f_L(\underline{r}) = f_s(\underline{r})$], one can define

$$A_r = \int_{A_d} |f_d(\underline{r})|^2 d\underline{r} = \int_{A_d} |f_L(\underline{r})|^2 d\underline{r}$$

Also, let

$$P_L = a_L^2 A_r$$

denote the portion of the local laser power that is collected by the detector. Then,

$$N_{sn} = \alpha e^2 P_L$$

Let us also define the effective area of the receiver (or the mean-square value of the spatial integral) as

$$A_c = E\left\{|I_{l,d}|^2\right\}$$

where

$$I_{l,d} = \int_{A_d} f_L^*(\underline{r}) f_d(\underline{r}) d\underline{r}$$

Due to the presence of atmospheric turbulence, we assume that the received field is random with its spatial coherence function satisfying the Rytov model described above. We also note that, since the local field is matched to that of the incoming signal, the contribution of background noise in the spatial domain is limited to the spatial domain of the source. That is, not all of the spatial modes of the background field are collected by the detector. Consequently, the power spectrum of the background-noise contribution at the output of the detector may be expressed as $2e^2\alpha^2 P_L N_{0b}$, where N_{0b} is the one-sided power spectrum of the random process $b(t)$. To arrive at this result, we have assumed that the background field is spatially incoherent, i.e., $\Gamma_b(\underline{r_1}, \underline{r_2}) = E\left\{f_b(\underline{r_1}) f_b^*(\underline{r_2})\right\} = \delta(\underline{r_1} - \underline{r_2})$ (see the term in Eq. (1) that accounts for the background-noise contribution).

Given the above formulation, the SNR of a heterodyne receiver may now be expressed as

$$SNR = \frac{4e^2\alpha^2 \frac{a_s^2}{2} P_L A_c}{e^2\left(\alpha P_L + 2\alpha^2 P_L N_{0b} + N_{0c}\right) 2B_n}$$

where N_{0c} denotes the two-sided power-spectrum level of the circuit noise and B_n is the IF filter-noise bandwidth (one-sided).

Considering that the power of the local oscillator typically is much larger than other noise processes at the receiver, the SNR equation may be further approximated by

$$SNR \approx \frac{\alpha a_s^2 A_c}{(1 + 2\alpha N_{0b}) B_n} \tag{2}$$

The remaining task is to obtain A_c. This parameter may be interpreted as the effective aperture of the receiver (which is smaller than the actual aperture of the receiver due to turbulence). This parameter may be related to the MCF and the receiver optical transfer function (OTF). To elaborate,

$$A_c = \int_{A_d} \int_{A_d} \Gamma_d(\underline{r_1}, \underline{r_2}) f_L^*(\underline{r_1}) f_L(\underline{r_s}) d\underline{r_1} d\underline{r_2}$$

with

$$\Gamma_d(\underline{r_1}, \underline{r_2}) = E\left\{f_d(\underline{r_1}) f_d^*(\underline{r_2})\right\}$$

denoting the MCF of the received optical field at the detector area prior to mixing with the local laser. It is important to note that, when the received field remains coherent over the aperture area, the MCF at the detector area will satisfy $\Gamma_d(\underline{r_1}, \underline{r_2}) = \delta(\underline{r_1} - \underline{r_2})$. This is due to the fact that the MCFs in the detector and aperture planes are related via a two-dimensional Fourier transform relationship, and, hence, a constant MCF in the aperture (totally coherent field) will yield a delta function MCF in the detector area. Given this argument, $A_c = \int_{A_d} |f_L(\underline{r})|^2 d\underline{r} - A_r$. This implies that the "coherence area" of the field (the area on the detector available for collecting coherent power) will be identical to the effective detector area in the absence of turbulence. In the event that atmospheric turbulence is present, $A_c < A_r$, which implies that the effect of turbulence is to reduce the coherence area of the signal. This in turn reduces the effective power and, consequently, the SNR of the heterodyne detection. Considering that the relationship between the fields at the aperture and detector areas is that of the Fourier transform pair, assuming that the local laser field is matched to the received field at the detector area, and using the Parseval identity, we have

$$A_c = E\left| \int_{-\infty}^{\infty} f_a(\underline{r}) P(\underline{r}) d\underline{r} \right|^2$$

or

$$A_c = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Gamma_a(\underline{r_1}, \underline{r_2}) P(\underline{r_1}) P(\underline{r_2}) d\underline{r_1} d\underline{r_2}$$

where $P(\underline{r})$ denotes the pupil function of the aperture. Making the change of variables $\underline{r_1} = \underline{R} + (\underline{\rho}/2)$ and $\underline{r_2} = \underline{R} - (\underline{\rho}/2)$, and realizing that

$$\kappa(\underline{\rho}) = \int_{-\infty}^{\infty} P\left(\underline{R} + \frac{\underline{\rho}}{2}\right) P\left(\underline{R} - \frac{\underline{\rho}}{2}\right) d\underline{R}$$

is the optical transfer function (OTF) of the receiver aperture, we have

$$A_c = \int_{-\infty}^{\infty} \Gamma_a(\underline{\rho}) \kappa(\underline{\rho}) d\underline{\rho}$$

where it is assumed that the MCF in the above equation is a function of $\underline{r_1} - \underline{r_2} = \underline{\rho}$. Note that the OTF is merely the autocorrelation function of the pupil function. To go any further, an aperture type must be specified. If one assumes a circular aperture with diameter D, then it is rather easy to observe that the OTF is the common area of two overlapping circles whose centers are a distance of $|\underline{\rho}|$ apart. Note that the OTF is only a function of the magnitude of the vector $\underline{\rho}$. A similar observation is true for a turbulent field whose MCF obeys the Rytov model. That is, $\Gamma_a(\underline{\rho}) = \Gamma_a(|\rho|)$. Hence (assuming $\underline{\rho} = (r, \theta)$ is in the polar coordinate system),

$$A_c = \int_{-\pi}^{\pi} \int_0^D \Gamma_a(r)\kappa(r) r dr d\theta = 2\pi \int_0^D \Gamma_a(r)\kappa(r) r dr$$

The remaining task is to substitute for the OTF and MCF in the above equation. The OTF of a circular aperture has been previously obtained (see [1]) and can be shown to be

$$\kappa(r) = \begin{cases} \frac{2}{\pi}\left[\cos^{-1}\left(\frac{r}{D}\right) - \frac{r}{D}\sqrt{1-\left(\frac{r}{D}\right)^2}\right]; & 0 \le r \le D \\ 0; & \text{otherwise} \end{cases}$$

Assuming that the background radiation and the circuit noise are rather negligible, the SNR of the receiver may be approximated as

$$SNR \approx \frac{\alpha a_s^2 A_c}{B_n} = \frac{4\alpha a_s^2}{B_n} \int_0^D r\Gamma_a(r) \left[\cos^{-1}\left(\frac{r}{D}\right) - \frac{r}{D}\sqrt{1-\left(\frac{r}{D}\right)^2}\right] dr$$

$$= \frac{4\alpha a_s^2}{B_n} \int_0^D r \exp\left(-3.44\left(\frac{r}{r_0}\right)^{5/3}\right) \left[\cos^{-1}\left(\frac{r}{D}\right) - \frac{r}{D}\sqrt{1-\left(\frac{r}{D}\right)^2}\right] dr$$

$$= \frac{\pi r_0^2 \eta a_s^2}{4hf_c B_n} \times \left\{\frac{16\beta^2}{\pi} \int_0^1 u \exp\left(-3.44(\beta u)^{5/3}\right) \left[\cos^{-1}(u) - u\sqrt{1-u^2}\,\right] du\right\} \quad (3)$$

where $\beta = D/r_0$. As can be seen, the SNR at the receiver is a function of β, the ratio of the receiver diameter and the coherence radius of the atmosphere. It has been shown [2] that, as β is increased, the SNR can be improved. However, as one continues to increase β, the SNR approaches a bound, underscoring the fact that any further increase in the aperture diameter to combat the impact of turbulence will not result in an improved SNR. Assuming that $\beta > 10$ can be used, then the integral inside the brackets in Eq. (3), which is always less than 1, approaches unity. Hence,

$$SNR \le \frac{\pi r_0^2 a_s^2 \eta}{4hf_c B_n}$$

Note that the numerator of the equation may be interpreted as an effective receiving area that is identical to the area of the coherence region of the received field. Also, as one considers background noise and circuit noise, the above upper bound is reduced.

IV. PPM Modulation

In the analysis that follows, we are interested in PPM modulation, which is suitable for energy-efficient deep-space optical communication. In this form of modulation, a single pulse is placed in one of the slots in the symbol interval. Given that the background radiation is present, one can describe the intensity of the received optical M-ary PPM signal as

$$\lambda_r(t) = \lambda_b + \lambda_s \sum_{i=0}^{\infty} P(t - C_i T_c - iT_s)$$

where λ_b and λ_s are the background and signal intensities, respectively, in photons/s; $P(t)$ is the pulse shape describing the laser pulse; C_i is the PPM symbol taking on an integer value from the set $\{0, 1, \cdots, M-1\}$ with equal probability (M is the PPM symbol alphabet); T_s is the PPM symbol duration in seconds; and $T_c = T_s/M$ is the PPM slot duration in seconds. In this analysis, it is assumed that the pulse shape, $P(t)$, is Gaussian in shape and is confined to a slot duration. In direct-detection-type receivers, the optical signal intensity, such as the one shown above, will determine the statistics of the number of photons in a given symbol interval.

In the analysis that follows, however, we are interested in assessing the performance of the PPM modulation scheme when the received signal is mixed with a local laser prior to photodetection. A major difference between the analysis presented here and that in the literature is the use of a pulsed laser for heterodyne detection. Traditionally, coherent detection is associated with continuous-wave (CW) lasers whose phase instabilities are well approximated using a Lorentzian distribution. In the analysis presented here, we consider heterodyne detection using high-power, pulsed lasers, whose phase stability, to the best of the authors' knowledge, has never been the subject of any extensive study. Naturally, such lasers are used for noncoherent (or direct-detection) applications, and, thus, the phase stability is of little concern. For the problem at hand, however, the statistics of the phase of the laser are of significant importance to the detection process. Hence, in the absence of an analytical model, we carry out our analysis making a number of assumptions. First, we assume that slot and symbol synchronization have been established. Second, it is assumed that the local laser is phase locked with the incoming signal to the extent that the phase difference between the phase of the incoming signal and that of the local laser is assumed to be constant over at least one PPM slot. However, we make no assumption with regard to the phase difference between the two signals from symbol to symbol. This is mainly due to the fact that the phase of the optical signal generated by high-power, pulsed lasers is quite unpredictable and may vary substantially from pulse to pulse (which in this case implies symbol to symbol). However, given that the slot interval typically is quite small, one can assume negligible drift in laser phase over a given PPM slot. This assumption has not been substantiated experimentally, although some recent experiments have shown promising results. This assumption, nonetheless, is quite reasonable assuming that, for high-speed, deep-space exploration, one requires large PPM alphabet sizes and small PPM symbol durations. As this trend continues, one can expect a fairly small PPM slot duration for future systems.

Finally, we assume that the local laser power is substantially greater than those of the incoming signal and the thermal noise due to electronic devices that are present at the receiver. Furthermore, we assume a shot-noise-limited detection due to the presence of a strong local laser.

Given the above assumptions, we can proceed with the problem formulation. A notation similar to that of the preceding section is used here. However, the notation is modified to account for the pulsed nature of the optical signal at the transmitter as well as at the receiver. To that end, let the field generated by the local laser over the kth slot of the nth PPM symbol be

$$f_{LO}(\underline{r}, t) = a_L e^{j(\omega_{LO} t + \phi_{LO}(t))} P(t - kT_c - nT_s) f_{LO}(\underline{r})$$

where a_L, ω_{LO}, and ϕ_{LO} are the peak field amplitude, the frequency of laser in Hz, and the phase of the laser due to laser instabilities, respectively, and $f_{LO}(\underline{r})$ denotes the spatial component of the local field. Note that the laser phase noise can severely hamper the performance of a phase-modulated system. Furthermore, $P(t)$ denotes a nonreturn-to-zero (NRZ) pulse of duration T_c. For the analysis that follows, however, we are concerned with the performance of PPM modulation, which can be classified as an M-ary orthogonal signaling. Such signaling schemes are not as severely impacted by phase instability as their phase-modulation counterparts. The received signal field, similar to that presented in the preceding section, may now be represented as

$$f_d(\underline{r},t) = a_s e^{j(\omega_s t+\phi_s(t))} \sum_{i=-\infty}^{\infty} P(t - C_i T_c - iT_s) f_d(\underline{r}) + b(t) e^{j\omega_s t} f_b(\underline{r})$$

Given that the local laser power is large and that a shot-noise-limited operation (i.e., the receiver thermal noise can be ignored) is achievable here, the current at the photodetector load resistance can be approximated as

$$i_d(t) \approx 2e\alpha a_s I_{l,d} \sqrt{P_L} \cos(\omega_{IF} t + \phi_e(t)) \sum_{i=-\infty}^{\infty} P(t - C_i T_c - iT_s) + n_{het}(t)$$

where $n_{het}(t)$ denotes a zero-mean, white Gaussian noise process that accounts for shot noise and background radiation with a one-sided power spectrum

$$N_{het} = \alpha e^2 P_L (1 + 2\alpha N_{0b})$$

Given the above formulation, one can suggest an optimal receiver in the absence of phase noise. It is imperative to note that the phase noise is a major impairment, and, thus, must be taken into account in assessing the performance of the receiver. However, to this date, an optimal receiver for the detection of phase-noisy optical receivers has not been found. This complication is due mainly to the complexity of the statistics of the phase-noise process.

For the sake of simplicity and without loss of generality, we limit our analysis to the 0th ($j = 0$) transmitted symbol. In the absence of phase noise and turbulence, an optimal receiver will yield the following decision variables [3]:

$$\Gamma_q = \int_{(q-1)T_c+\hat{\tau}}^{qT_c+\hat{\tau}} i_d(t) \cos(\omega_{IF} t) dt; \quad q = 1, 2, \cdots, M-1$$

and the decision rule will be

$$C_0 = \arg \max_q \{\Gamma_q; q = 1, 2, \cdots, M-1\}$$

Obviously, the above decision rule relies on a perfect knowledge of the IF signal phase. This further implies that one must establish phase synchronization (coherent or synchronous IF detection) at this stage. However, as noted earlier, the phase of the laser may vary from pulse to pulse, making it quite difficult, if not impossible, to successfully track the phase of the IF signal using a conventional tracking loop. For these reasons, we believe that a realistic approach here is to perform a noncoherent detection (in IF) of the signal shown above. That is,

$$C_0 = \arg \max_s \left\{\Gamma_s^{NC}; s = 1, 2, \cdots, M-1\right\}$$

where

$$\Gamma_s^{NC} = \sqrt{\Gamma_{s,I}^2 + \Gamma_{s,Q}^2}$$

with $\Gamma_{s,I} = \int_{(s-1)T_c+\hat{\tau}}^{sT_c+\hat{\tau}} i_d(t)\cos(\omega_{IF}t)dt$ and $\Gamma_{s,Q} = \int_{(s-1)T_c+\hat{\tau}}^{sT_c+\hat{\tau}} i_d(t)\sin(\omega_{IF}t)dt$. This detection mechanism is optimal when turbulence is absent and the phase of the IF signal is assumed to be constant and uniformly distributed on $[-\pi, \pi]$. In reality, the phase of the IF signal will be time varying and nonuniform in distribution. Nonetheless, as the duration of the laser pulse is reduced in time, the phase drift in both transmitter and receiver lasers can be considered to be negligible over a PPM slot. It is important to note that a receiver implementation that follows the above decision rule may be classified as an asynchronous receiver without post-detection processing (see [4, pp. 276–284]). Finally, it is imperative to note that the presence of turbulence causes a time-dependent SNR. Hence, the decision rule described above leads to a sub-optimal architecture in the presence of time-dependent turbulence. For the problem at hand, we assume that the PPM frames are repeated at a rate of 50 kHz. That is, the PPM pulses are spaced in time by 20 μs. Since the atmospheric conditions change at a much slower rate (5 to 50 Hz), one may assume that the SNR (or more specifically, the field amplitude) remains constant for many PPM symbols. This further implies that one may obtain the performance of the above receiver architecture by conditioning the performance on the statistics of the amplitude. The performance then may be obtained by averaging the resulting expression over the random variations of the amplitude.

The performance of the above decision rule can readily be obtained [3], assuming that the laser pulses are confined to the PPM slot duration and that the slot and symbol synchronization subsystems have yielded a negligible timing error. Furthermore, to arrive at the desired expression for symbol and bit-error rates, one must condition the performance on the amplitude fluctuations due to turbulence. Considering the above approximations, and assuming that the Lth symbol has been transmitted over the PPM symbol duration of interest, we have

$$\Gamma_{s,I} = \begin{cases} e\alpha a_s I_{l,d}\sqrt{P_L}\displaystyle\int_{(L-1)T_c+\hat{\tau}}^{LT_c+\hat{\tau}} \cos(\phi_e(s))ds + N_I; & s = L \\ N_I; & \text{otherwise} \end{cases}$$

and

$$\Gamma_{s,Q} = \begin{cases} e\alpha a_s I_{l,d}\sqrt{P_L}\displaystyle\int_{(L-1)T_c+\hat{\tau}}^{LT_c+\hat{\tau}} \sin(\phi_e(s))ds + N_Q; & s = L \\ N_Q; & \text{otherwise} \end{cases}$$

where N_I and N_Q are a pair of zero mean, independent Gaussian random variables with identical variances given by

$$\sigma_{sn}^2 = \alpha e^2 P_L \frac{T_c}{2}(1 + 2\alpha N_{0b})$$

Note that the impact of phase noise is the reduction in SNR. To go any further, one must consider the statistics of the phase noise in the above equation. As noted earlier, in this analysis, we consider the case when the laser phase noise may be considered to be "slow varying" or constant over a PPM slot duration. Given this approximation, then

$$\Gamma_{s,I} = \begin{cases} e\alpha a_s I_{l,d}\sqrt{P_L}\cos(\phi_e)T_c + N_I; & s = L \\ N_I; & \text{otherwise} \end{cases}$$

and

$$\Gamma_{s,Q} = \begin{cases} e\alpha a_s I_{l,d}\sqrt{P_L}\sin(\phi_e)T_c + N_Q; & s = L \\ N_Q; & \text{otherwise} \end{cases}$$

This, in turn, leads to

$$\Gamma_s^{NC} = \begin{cases} \sqrt{\left(e\alpha a_s |I_{l,d}|\sqrt{P_L}\sin(\phi_e + \theta_{l,d})T_c + N_I\right)^2 + \left(e\alpha a_s |I_{l,d}|\sqrt{P_L}\cos(\phi_e + \theta_{l,d})T_c + N_Q\right)^2}; & s = L \\ \sqrt{N_I^2 + N_Q^2}; & \text{otherwise} \end{cases}$$

where $\theta_{l,d}$ is the phase of $I_{l,d}$. If one assumes that ϕ_e is a uniformly distributed random variable[2] (this leads to a worst-case scenario), the probability density function of Γ_s^{NC} when conditioned on $I_{l,d}$ is given by

$$f_{\Gamma_s|I_{l,d}}(x|I_{l,d}) = \begin{cases} \dfrac{x}{\sigma_{sn}^2}\exp\left(-\dfrac{x^2 + A^2}{2\sigma_{sn}^2}\right) I_0\left(\dfrac{Ax}{\sigma_{sn}^2}\right) U(x); & s = L \\ \dfrac{x}{\sigma_{sn}^2}\exp\left(-\dfrac{x^2}{2\sigma_{sn}^2}\right) U(x); & \text{otherwise} \end{cases}$$

where $A = e\alpha a_s |I_{l,d}|\sqrt{P_L}T_c$ and $I_0()$ is the 0th-order modified Bessel function of the first kind.

In the absence of timing error and when the pulse shape is confined to a chip duration, PPM signaling constitutes an orthogonal form of signaling at the receiver. Given this assumption, we have

$$P_b(e|I_{l,d}) = \frac{M/2}{(M-1)}\sum_{n=1}^{m-1}(-1)^{n+1}\binom{M-1}{n}\frac{1}{n+1}\exp\left[-\frac{n}{(n+1)}SNR(I_{l,d})\right]$$

where

$$SNR(I_{l,d}) = \frac{\left(e\alpha a_s |I_{l,d}|\sqrt{P_L}T_c\right)^2}{2\alpha e^2 P_L \dfrac{T_c}{2}(1 + 2\alpha N_{0b})} = \frac{\alpha a_s^2 |I_{l,d}|^2 T_c}{1 + 2\alpha N_{0b}}$$

and $P_b(e|I_{l,d})$ denotes the system bit-error rate when conditioned on amplitude variations due to turbulence. Note that this definition of the SNR is identical to Eq. (2) when the one-sided bandwidth of the IF filter is considered to be $B_n = 1/T_c$ (note that an integrate-and-dump (I&D) filter possesses a one-sided bandwidth of $1/T_c$).

For large M, the above expression is difficult to compute, and, hence, one can exploit the union bound to arrive at an upper bound on performance. That is, the overall bit-error rate may be upper bounded as [3]

$$P_b(e|I_{l,d}) \leq \frac{M}{2}P_{b,2}(e|I_{l,d}) \tag{4}$$

[2] This assumption is only necessary for the case of $s = L$.

where $P_{b,2}(e|I_{l,d})$ denotes the performance of a binary PPM receiver operating under a set of conditions similar to that of its M-ary counterpart. This probability is given by

$$P_{b,2}(e|I_{l,d}) = \frac{1}{2}\exp\left(-\frac{SNR(I_{l,d})}{2}\right)$$

Finally, to arrive at the average error rate, we resort to an approximation. Namely,

$$P_{b,2}(e) \approx \frac{1}{2}\exp\left(-\frac{SNR}{2}\right)$$

where $SNR = E\{SNR(I_{l,d})\}$ and is given by Eq. (3). Similarly, the error rate also may be approximated as

$$P_b(e) \approx \frac{M/2}{(M-1)}\sum_{n=1}^{M-1}(-1)^{n+1}\binom{M-1}{n}\frac{1}{n+1}\exp\left[-\frac{n}{(n+1)}SNR\right] \tag{5}$$

To arrive at the above approximations, we have used $E\{g(x)\} \approx g(E(x))$. This approximation can be used when $g(x)$ is a well-behaved function in the vicinity of $E(x)$ and $\mathrm{var}(x)$ can be considered to be negligible as compared with the mean value. The first assumption is easily satisfied since the conditional error rates that are given above take on an exponential form. The second assumption is somewhat more difficult to satisfy. However, it can be argued that the variance of SNR is substantially smaller than the square of the mean SNR, justifying the above approximation.

V. Numerical Results

In this section, we will provide some numerical results to shed light on the performance of a heterodyne PPM receiver. For the case of interest, we use $M = 256$. Also, to gain insight into the achievable performance in the absence of turbulence, we first consider the case when $A_c = A_d$ (that is, the received field remains coherent over the entire receiver aperture). In this analysis, we are interested in symbol-error rates in the range of 10^{-3} to 10^{-2}, which is shown in Fig. 1. In this figure, the performance is depicted as a function of

$$SNR = \alpha a_s^2 E\left\{|I_{l,d}|^2\right\}T_c = \frac{\alpha a_s^2 A_c T_c}{1 + 2\alpha N_{0b}}$$

which in the absence of turbulence reduces to

$$SNR = \frac{\alpha a_s^2 A_d T_c}{1 + 2\alpha N_{0b}} = \frac{\alpha P_s T_c}{1 + 2\alpha N_{0b}} = \frac{\alpha E_s}{1 + 2\alpha N_{0b}}$$

Therefore, the average SNR in heterodyne systems may be viewed as the average number of recovered photons per PPM slot divided by the number of photons collected in the same time slot due to background radiation. Notice that the effect of atmospheric turbulence is to reduce the useful number of photons per slot, and, hence, to compromise the performance of the receiver. For this reason, the above may be viewed as the lower bound on the number of received photons required to achieve a given performance level. From Fig. 1, it can be seen that an SNR of 12.4 dB is needed to achieve a symbol-error rate of 10^{-2} in the absence of background radiation. This number increases to 13.6 dB for a symbol-error

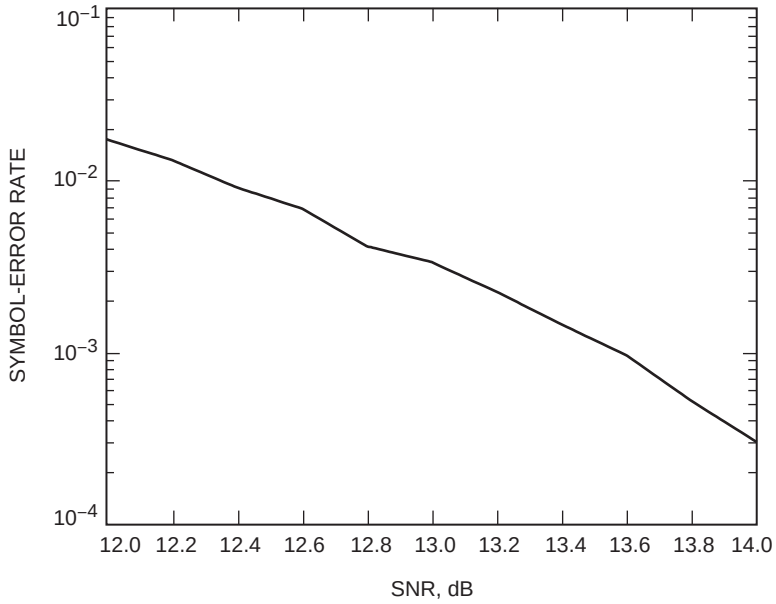

Fig. 1. Performance of the heterodyne 256-ary PPM receiver.

rate of 10^{-3}. To go any further, one must specify the operating laser frequency and the quantum efficiency of the photodetector. We consider an operating wavelength of 532 nm, which leads to $f_c = (3 \times 10^8)/(532 \times 10^{-9}) = 5.6 \times 10^{14}$ Hz. Furthermore, the quantum efficiency of the detector can take on a value of 0.3 for germanium detectors operating at a wavelength of 532 nm. This number is reduced to about 0.15 for silicon detectors. Assuming, then, that a germanium detector is used and that background noise is absent, one requires 58 received photons per PPM slot to achieve a symbol-error rate of 10^{-2} (this number takes into account the loss due to quantum efficiency of less than 1). For a symbol-error rate of 10^{-3} and when background noise is absent, one requires 76 photons. Note that these numbers have been obtained assuming a negligible background radiation and when the atmospheric turbulence is considered negligible. If background radiation is present, it impacts the required number of signal photons adversely. Note that $\alpha N_{0b} = \eta(N_{0b}/hf_c) = \eta N_{bn}$ may be regarded as the average number of photons that are detected over a PPM slot due to background radiation, with N_{bn} denoting the average number of received background photons. The factor of 2 in the denominator of the SNR described above is due to the squaring operation of the photodetector.

For a symbol-error rate of 10^{-2}, then, we can compute the required number of received photons for various levels of background radiation. The resulting numbers may be compared with those of an APD-based, direct-detection receiver for PPM. In a recent study ([5], Figs. 9 and 10), an upper bound on the number of received photons required to achieve a given performance was computed as a function of the background radiation level. In Table 1, the required number of signal photons is given for various levels of background radiation for a germanium-type detector when an error rate of 10^{-2} is of interest. From Table 1, it immediately becomes obvious that, when background noise cannot be ignored, the heterodyne detection performs poorly as compared with its APD-based detection mechanism. It also is important to note that the numbers for the APD-based receiver have been obtained when thermal noise is present (the receiver is operating at room temperature). A similar calculation also can be made for an error rate of 10^{-3} and is provided in Table 2.

Once again, our earlier observation is reaffirmed. First, it is quite obvious that the APD-based detection is superior in performance across the board to that of the heterodyne detection considered here. In fact, in the absence of background radiation, the APD-based detector maintains its superiority.

Table 1. The upper bound on the required number of photons per PPM slot for APD-based receivers and the required number of photons per PPM slot for heterodyne receivers with a symbol-error rate of 10^{-2}.

N_{bn}	Required number of signal photons per PPM slot (heterodyne)	Upper bound on the required number of signal photons per PPM slot (direct detection)
0	58	43
20	406	60
40	754	75
100	1798	115
200	3538	120

Table 2. The upper bound on the required number of photons per PPM slot for APD-based receivers and the required number of photons per PPM slot for heterodyne receivers with a symbol-error rate of 10^{-3}.

N_{bn}	Required number of signal photons per PPM slot (heterodyne)	Upper bound on the required number of signal photons per PPM slot (direct detection)
0	76	52
20	532	77
40	988	90
100	2356	125
200	4636	150

Furthermore, the phase of the local laser, as well as that of the transmitter laser, is assumed to be relatively constant over a PPM slot. As noted earlier, if this condition is violated, there will be further degradation in performance. Moreover, in the event that atmospheric turbulence is present, one can expect a larger number of photons required to achieve the above performance.

In view of the above observations, one can conclude readily that when synchronous detection is not feasible in a heterodyne-type optical receiver and when PPM modulation is used, an APD-based direct-detection receiver yields a performance superior to that of its heterodyne counterpart. It also is imperative to note that a heterodyne receiver requires a number of additional circuitry to function properly. These include a (stable) local laser and additional circuitry to perform polarization matching and other necessary functions that are unique to heterodyne detection. Provided that the proposed system is intended for deployment in space, it is highly unlikely that one can achieve the necessary phase stability to perform the asynchronous detection described above when high-power lasers are employed. In the event that such stability can be achieved, and background noise can be ignored, the performance of the asynchronous heterodyne receiver remains inferior, as seen above, to that of its ADP-based counterpart.

References

[1] S. Karp, R. M. Gagliardi, S. E. Moran, and L. B. Stotts, *Optical Channels*, New York: Plenum Press, 1988.

[2] D. Fried, "Optical Heterodyne Detection of an Atmospherically Distorted Wavefront," *Proc. IEEE*, vol. 55, no. 1, pp. 57–67, January 1967.

[3] J. G. Proakis, *Digital Communications*, vol. 3, New York: McGraw-Hill, Inc., 1995.

[4] L. Kazovsky, S. Benedetto, and A. Willner, *Optical Fiber Communication System*, Boston: Artech House, 1996.

[5] K. Kiasaleh and T.-Y. Yan, "T-PPM: A Novel Modulation Scheme for Optical Communications 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