Figure of Merit for Direct-Detection Optical Channels

C.-C. Chen
Communications Systems Research Section

The capacity and sensitivity of a direct-detection optical channel are calculated and compared to those of a white Gaussian noise channel. Unlike Gaussian channels in which the receiver performance can be characterized using the noise temperature, the performance of the direct-detection channel depends on both signal and background noise, as well as the ratio of peak to average signal power. Because of the signal-power dependence of the optical channel, actual performance of the channel can be evaluated only by considering both transmit and receive ends of the systems. Given the background noise power and the modulation bandwidth, however, the theoretically optimum receiver sensitivity can be calculated. This optimum receiver sensitivity can be used to define the equivalent receiver noise temperature and calculate the corresponding G/T product. It should be pointed out, however, that the receiver sensitivity is a function of signal power, and care must be taken to avoid deriving erroneous projections of the direct-detection channel performance.

I. Introduction

Optical communication technology can offer potentially significant improvements in communication performance compared to current RF links. Much of the gain in optical receiver performance is due to the reduced operating wavelength and hence the increased antenna directivity. The small beam divergence resulting from the short operating wavelength implies that the received power needed for the communication link can be achieved with a much smaller transmit antenna aperture and a lower transmit power requirement. A smaller aperture also implies a lower weight communication package and hence an increased science payload capacity. Furthermore, a smaller beam divergence implies that the aperture size of future deep-space optical receiving terminals can be substantially smaller than the present-day Deep-Space Network (DSN) receivers.

For RF Earth receivers, an effective measure of the receiver performance is the ratio of the receiving antenna gain G_R to the system noise temperature T_{eq}. The received antenna gain G_R is proportional to the receiver aperture area and hence the received signal power. The channel efficiency, which is the maximum amount of information that can be relayed per unit of received energy, directly related to the receiver noise temperature. For RF receivers, this G_R/T_{eq} ratio, known as the Earth receiver figure of merit [1], is directly proportional to the receiver signal-to-noise ratio (SNR). Given the transmitter and the link distance, the performance of the ground terminal can be compared by calculating the figure of merit without performing a complete end-to-end link analysis.

Since the G_R/T_{eq} figure of merit is a convenient parameter to compare the receiver performance, it is desirable to extend the concept to optical frequencies and to derive an equivalent figure of merit for optical receivers. For coherent optical receivers, the extension is straightforward. The resulting receiver noise temperature is given by $T_{op} = hf/k_B$, where hf is the energy of the photon and k_B is Boltzmann's constant. For direct-detection receivers, however, the concept of noise temperature cannot be applied directly. Although the minimum detectable power is equal to the energy of a single photon, the receiver performance can be effectively improved by trading the receiver sensitivity for increasing bandwidth [2,3]. In the limit of infinite bandwidth expansion, the direct-detection receiver can achieve very high channel efficiency that is limited only by the background noise level. Furthermore, unlike RF links where the channel efficiency depends only on the receiver noise temperature, direct-detection optical communication link performance is a function of both the signal and noise powers, and the ratio of peak-to-average signal power.

Because the performance of a direct-detection optical channel depends on the peak and average signal power in addition to the background noise, the performance of the direct-detection channel can be evaluated only if both ends of the link are defined. For a given background power and bandwidth, however, there exists an optimum condition under which the maximum amount of information can be transmitted across the channel per received signal photon. This optimal channel efficiency is achievable only under a particular set of signal power and bandwidth constraints. However, it allows one to derive an upper bound on the direct-detection channel performance.

The purpose of this article is to present a simple calculation of the receiver sensitivity for an ideal direct-detection optical channel subjected to a system bandwidth constraint. This optimal sensitivity can then be used to define the equivalent receiver noise temperature of the direct-detection channel. The equivalent G_R/T_{eq} parameter can then be calculated. Caution should be exercised, however, when using this ratio as it behaves quite differently from the corresponding ratio for RF systems.

II. Figure of Merit for Communication Links

An important figure of merit of an RF communication link is the channel efficiency C_E. The efficiency can be defined as the maximum amount of information that can be relayed per unit of received energy. For a white Gaussian noise channel with noise power spectral density N_0, the ideal channel efficiency is simply

$$C_E = \frac{\log_2 e}{N_0} \text{ (bits/joule)} \quad (1)$$

Given the amount of received signal power P_S, the channel capacity, which is defined as the maximum data rate that can be transmitted across the channel, is

$$R_{max} = P_S C_E = \frac{P_S}{N_0 \ln 2} = \frac{P_S}{k_B T_{eq} \ln 2} \text{ (bits/sec)} \quad (2)$$

Note that in the last equality we have replaced the noise power spectral density with $k_B T_{eq}$, where k_B is Boltzmann's constant and T_{eq} is the equivalent receiver noise temperature. The channel efficiency and channel capacity shown in Eqs. (1) and (2) are derived for additive white Gaussian noise (AWGN) channels without bandwidth constraint. Practical systems, with a limited bandwidth expansion factor, will have lower channel efficiency and capacity.

The received signal power P_S is a function of the transmit power, the transmitter and receiver parameters, and the link distance. With a diffraction-limited transmitter operating with aperture area A_T at a distance D, the signal power collected at the receiving terminal can be written as

$$\begin{aligned} P_S &= P_T \eta_T G_T \left(\frac{1}{4\pi D^2}\right) L_T(\theta_T) \eta_R A_R \\ &= P_T \eta_T G_T \left(\frac{\lambda}{4\pi D}\right)^2 L_T(\theta_T) \eta_R G_R \end{aligned} \quad (3)$$

where η_T and η_R are the transmitter and receiver efficiencies, G_T and G_R are the transmit and receive antenna gains, $L_T(\theta_T)$ is the pointing loss for an angular pointing error of θ_T, A_R is the receiver aperture area, and P_T is the transmit power. The transmit and receive antenna gains are defined as the ratio of far-field intensity when a signal is transmitted from the antenna to that which is radiated from an isotropic radiator. For diffraction-limited apertures, the on-axis antenna gain is related to the aperture area and the operating wavelength λ by

$$G_T = 4\pi A_T/\lambda^2 \quad (4a)$$

$$G_R = 4\pi A_R/\lambda^2 \quad (4b)$$

Equation (3) shows that the total signal power received is proportional to the transmit and receive antenna gain and is inversely proportional to the link distance squared. The received power scales inversely with λ^2 because the far field intensity from a diffraction-limited transmitter is inversely proportional to λ^2. One should note that, although the received signal power is directly proportional to the receiver antenna gain in Eq. (4b), Eq. (3) does not imply that a diffraction-limited receiver must be used to collect the signal power. In fact, for a sufficiently large receiver field of view, the received power depends only on the total aperture area and not on the surface quality of the receiver. A nondiffraction-limited receiver, however, can admit much more noise power (Appendix A), or can degrade or even preclude the use of some forms of signal modulation.

By combining Eqs. (1) and (2), it is seen that the channel capacity is related to the transmit and receive parameters by

$$R_{\max} = \left(\frac{1}{4\pi D^2}\right) P_T \eta_T G_T L_T(\theta_T) \eta_R A_R \left(\frac{\log_2 e}{k_B T_{\mathrm{eq}}}\right) \quad (5)$$

In general, it is desirable to derive a figure of merit which is independent of the link distance. For Gaussian channels, one such parameter is the product of channel capacity and link distance squared, $R_{\max} D^2$, which is given by

$$\begin{aligned} R_{\max} D^2 &= \left(\frac{1}{4\pi}\right) P_T \eta_T G_T L_T(\theta_T) \eta_R A_R \left(\frac{\log_2 e}{k_B T_{\mathrm{eq}}}\right) \\ &= \left(\frac{\lambda}{4\pi}\right)^2 P_T \eta_T G_T L_T(\theta_T) \eta_R G_R \left(\frac{\log_2 e}{k_B T_{\mathrm{eq}}}\right) \end{aligned} \quad (6)$$

The $R_{\max} D^2$ parameter defined in Eq. (6) is proportional to the transmit antenna gain, the receiver area, and is inversely proportional to the receiver noise temperature. For a given transmitter power and aperture area, this parameter is proportional to G_R/T_{eq}, which is commonly referred to as the Earth receiver figure of merit. The G_R/T_{eq} figure of merit is particularly useful in comparing different RF system performances. This is because for RF systems the transmitter aperture and available transmitter power usually do not vary much. As a result, the performance improvements are generally achieved by increasing aperture area, reducing operating wavelength, and lowering the receiver noise temperature; in other words, by improving the G_R/T_{eq} ratio.

III. Direct Detection Optical Channel

As was mentioned earlier, the channel efficiency of an ideal RF link, defined as the maximum amount of information that can be relayed across the channel per unit of received energy, is equal to $\log_2 e/k_B T_{\mathrm{eq}}$ (bit/J). Note that the channel efficiency depends only on the receiver noise temperature T_{eq} and is independent of the signal power and aperture size. This simple expression for channel efficiency emerges because the only noise source present in the RF link model is the AWGN. An important aspect of the AWGN channel is that the error rate performance of the link depends only on the ratio of signal and noise powers (signal-to-noise ratio). Consequently, the G_R/T_{eq} parameter, which is proportional to the SNR, is a good measure of the receiver performance.

In contrast, the dominant source of noise in a direct-detection channel is the shot noise inherent in the signal, which cannot be modeled as AWGN. When detected using a square-law detector such as a photodiode, the shot-noise fluctuation can result in a fluctuation of the received photocount [4]. This self-noise fluctuation implies that, even when the amount of background noise admitted by the receiver is negligible, the number of photons collected (received signal energy) over a period ΔT can still fluctuate. It should be noted that the quantum fluctuation is also present in the RF receiver. However, such a fluctuation is usually ignored since the mean field is much larger than the rms fluctuation.

It can be shown that, for reception of a coherent signal field and multimode thermal radiation, the photocounts follow Poisson statistics [5]. For detection of single-mode background radiation, the photocounts follow Bose-Einstein statistics. For most direct-detection receivers, however, the large mode mismatch at the receiver implies that more than one spatial and temporal mode is being received and, in the limit of large mode mismatch, the background photocount process also exhibits Poisson statistics [5].

Because of the Poisson nature of the counting process, the direct-detection optical channel is also known as a Poisson channel. The received photocount over a given interval, ΔT, which contains contributions from both signal and background noise, can be modeled as a Poisson random variable with mean and variance given by [6]

$$\begin{aligned} <N> &= (\lambda_S + \lambda_B)\Delta T \\ Var[N] &= <N> = (\lambda_S + \lambda_B)\Delta T \end{aligned} \tag{7}$$

where λ_S and λ_B are the detected signal and background photocount rates measured in photons per second. The photocount rates are related to the power input by [5]

$$\lambda_S = \frac{\eta_D}{hf} P_S \tag{8a}$$

$$\lambda_B = \frac{\eta_D}{hf} P_B \tag{8b}$$

where η_D is the detector quantum efficiency, and hf is the energy of the photon at the operating wavelength. Equation (7) showed that the variation in the received photocount depends on both signal and background power. Because of this signal-dependent fluctuation, the receiver performance will depend on both signal and background power instead of a single quantity (signal-to-noise ratio).

The channel capacity and energy efficiency of an ideal direct-detection channel has been evaluated by several authors [2,3,7–9]. It has been shown that, for a direct detection channel with peak signal power P_S and an average to peak signal power ratio of σ, the channel capacity is given by [8,9]

$$\begin{aligned} R_{\max} = \log_2 e\ \lambda_S \Bigg[&\bar{q}\left(\frac{1+\rho}{\rho}\right)\ln\left(\frac{1+\rho}{\rho}\right) \\ &-(1-\bar{q})\frac{1}{\rho}\ln\rho - \left(\frac{1+\bar{q}\rho}{\rho}\right)\ln\left(\frac{1+\bar{q}\rho}{\rho}\right)\Bigg] \end{aligned} \tag{9a}$$

where

$$\rho = \lambda_S/\lambda_B = P_S/P_B \tag{9b}$$

is the peak signal to background power ratio, and

$$\bar{q} = \min\left(\left(\frac{1+\rho}{\rho}\right)^{(1+\rho)/\rho} \rho^{1/\rho}/e - \rho, \sigma\right) \tag{9c}$$

The quantity $\bar{q}$ is the ratio of average to peak signal power that achieves the channel capacity. When there is no average power constraint, that is, when

$$\left[\left(\frac{1+\rho}{\rho}\right)^{(1+\rho)/\rho} \rho^{1/\rho}/e - \rho\right] < \sigma$$

the channel capacity is limited only by the available peak signal power at the receiver. In this case the channel capacity is $\log_2 e\bar{q}\lambda_S$. Note that $\bar{q}\lambda_S$ is simply the average signal count rate. In other words, given the peak signal count rate λ_S, the maximum rate for which the information can be relayed across the channel is simply equal to the number of average received signal photons per second times $\log_2 e$. Equivalently, the channel can transmit $\log_2 e \approx 1.44$ bits per photon received.

On the other hand, when

$$\left[\left(\frac{1+\rho}{\rho}\right)^{(1+\rho)/\rho} \rho^{1/\rho}/e - \rho\right] > \sigma$$

the channel is said to be average power limited. The channel capacity decreases with decreasing average signal count rate, $\sigma\lambda_S$. However, for a constant peak signal count rate, the amount of information carried per signal photon, i.e., the channel efficiency, increases with decreasing average to peak power ratio, σ. Given the peak signal to background power ratio, ρ, the channel efficiency is given by

$$\begin{aligned} C_E &= \frac{\log_2 e}{hf}\,\frac{R_{\max}}{\bar{q}\lambda_S} \\ &= \frac{\log_2 e}{hf}\,\Bigg[\left(\frac{1+\rho}{\rho}\right)\ln\left(\frac{1+\rho}{\rho}\right) \\ &\quad - \frac{1+\rho}{\rho}\ln\rho - \frac{1+\bar{q}\rho}{\bar{q}\rho}\ln(1+\bar{q}\rho)\Bigg] \end{aligned} \tag{10}$$

The photon efficiency of the channel, which is defined as the maximum amount of information relayed through the channel per photon, is related to the channel efficiency by

$$C_{\rm ph} = hf\,C_E$$

$$= \log_2 e \left[\left(\frac{1+\rho}{\rho}\right)\ln\left(\frac{1+\rho}{\rho}\right) - \frac{1+\rho}{\rho}\ln\rho - \frac{1+\bar{q}\rho}{\bar{q}\rho}\ln(1+\bar{q}\rho)\right] \qquad (11)$$

In the limit as the average to peak signal power ratio goes to zero, the photon efficiency is given by

$$\lim C_{\rm ph} = \log_2 e \times \left[\left(\frac{1+\rho}{\rho}\right)\ln\left(\frac{1+\rho}{\rho}\right) + \frac{1}{\rho}\ln\rho + \ln\rho - 1\right] \qquad (12)$$

Equations (9) through (12) illustrate an important aspect of the direct-detection channel; namely, the receiver sensitivity is a function of the average to peak signal power ratio (duty cycle), as well as the signal and noise power levels. Furthermore, one can trade bandwidth for receiver sensitivity by choosing modulation schemes with low average to peak power ratio. Plotted in Fig. 1 is the limiting receiver sensitivity as $\sigma \rightarrow 0$ versus the signal power to background noise ratio, ρ. As expected, the photon efficiency increases indefinitely as the amount of background noise decreases. Consequently, the limiting efficiency of a direct-detection channel is limited only by the amount of background noise the receiver collects.

Given the peak signal count rate λ_S, the channel capacity (data rate) can be optimized by letting the average count rate approach

$$\left[(1+\rho/\rho)^{(1+\rho)/\rho}\rho^{\frac{1}{\rho}}/e - \rho\right]\lambda_S$$

In this case, the limiting receiver sensitivity is $\log_2 e$ bits per photon. For near-Earth links where maximum data rate, and not maximum photon efficiency, is desired, the link should operate with average to peak signal power ratio equal to

$$\left[\left(\frac{1+\rho}{\rho}\right)^{(1+\rho)/\rho}\rho^{1/\rho}/e - \rho\right]$$

Since the channel capacity is achieved with a relatively high average to peak power ratio, semiconductor lasers, which can be modulated at a higher bandwidth, are more suitable for achieving high data rates. On the other hand, for a deep-space link where high power efficiency is desired, one should operate the link with a low average to peak power ratio, i.e., by choosing laser and modulation schemes such that $\sigma \rightarrow 0$, while holding the average power constant. Solid-state lasers, which can provide high peak power with a low duty cycle, are ideally suited for such applications.

IV. Bandwidth-Limited Direct-Detection Channel and Its Figure of Merit

In practice, the limiting performance given by Eqs. (9) through (12) cannot be achieved. Several factors contribute to the limitations on channel performance [10]. First, the timing resolution (bandwidth) of the receiver is limited by the response of the photodetector material, the timing resolution of the photodetector/preamplifier assembly, and the complexity of the decoding electronics. Furthermore, the bandwidth of the channel is affected by the modulation bandwidth of the laser. The maximum timing resolution at the receiver is bounded by the uncertainty principle to approximately 0.1 psec [10], and, in practice, the detector timing resolution is limited by the complexity of the electronics to approximately 1 nsec.

The amount of background noise admitted by the receiver is, in principle, limited only by the optical background noise. Even when the Sun (6000 K) is in the field of view, a diffraction-limited receiver with a 1-GHz-wide predetection filter will observe only 2.5×10^6 background counts/sec when operating at 532-nm wavelength. In practice, the nondiffraction-limited receiver admits much more background noise (Appendix A), and the predetection filter is usually much wider than the signal bandwidth. Furthermore, detector dark counts constitute an irreducible background level which is present even when no bright background object is absent. As a result, a practical receiver will admit much more background noise than the thermal background noise limit.

Another factor affecting channel performance is the average to peak signal power ratio. Although in principle an infinite peak to average power ratio is achievable, the maximum achievable peak to average power ratio for a solid state laser is limited by the laser parameters such as the pump power to threshold power ratio. A practical direct-detection channel will reach its complexity and bandwidth expansion limit long before it reaches the capacity limit.

As a result, a practical direct-detection channel will have performance that is much less than the theoretical thermal background noise limit as predicted by Pierce and Posner [3].

Shown in Table 1 is a list of projected parameters for direct-detection receivers. For short-term development support, a 1-m diffraction-limited telescope can be used to collect the downlink signal. The blur circle diameter of the 1-m diffraction-limited telescope is limited by atmospheric seeing to approximately 20 times its diffraction limit. Actual operational support for deep-space missions is planned by using several 10-m-class photon-bucket receivers to provide spatial diversity reception. The blur circle diameter of the 10-m photon-bucket is estimated to be approximately 2000 times more than an equivalent diffraction-limited aperture. The substantially worse surface quality implies that a much larger amount of background noise will be collected by the 10-m receiver. However, the larger collecting area can actually result in an improved system performance. A prototype research and development station, the Deep Space Optical Receiving Antenna (DSORA), is currently being studied by JPL, and a facility construction request for DSORA has been submitted to NASA for a projected 1997 start. The parameters shown in Table 1 are the projected parameters for DSORA.

It is seen from Eqs. (9) through (11) that the performance of direct-detection optical channels is determined by the average and peak signal powers, as well as the background noise power. Given a set of link parameters, the average signal and background powers can be calculated using Eq. (3) and Appendix A. For a constant average signal power, the channel performance can be improved by increasing the peak power to average signal power ratio. Since the receiver bandwidth limits the maximum rate at which arriving photons can be distinguished, the peak signal count rate, λ_S, of a practical direct-detection channel is simply equal to its receiver bandwidth limit. Based on the receiver bandwidth and average signal and background power, the channel efficiency and capacity of a band-limited direct-detection channel can be calculated. The equivalent noise temperature of the optical receiver can then be defined by equating the optimum channel efficiency with $\log_2 e/k_B T_{\text{eq}}$. That is,

$$C_{E_{\max}} \equiv \frac{C_{\text{ph}}}{hf} \equiv \frac{\log_2 e}{k_B T_{\text{eq}}} \text{ (bits/joule)} \qquad (13)$$

This equivalent noise temperature represents an upper bound of the direct-detection link performance. For a constant transmit power and aperture, the G/T figure of merit can be defined as:

$$(G/T)_{\text{optical}} = \eta_R \frac{(4\pi A_R/\lambda^2)}{T_{\text{eq}}} \qquad (14)$$

where T_{eq} is the equivalent noise temperature as given by Eq. (13), and η_R is the combined efficiency of the receiving optics and detector. The G/T ratio defined in Eq. (14) may also be interpreted as the equivalent G/T ratio needed to achieve similar performance for a coherent receiver. Since the direct detection receiver is not actually diffraction limited, we should refrain from calling the quantity in the numerator the receive antenna gain. One should note that the figure of merit as defined in Eq. (14) is calculated only under fixed signal and background power levels and receiver bandwidth. The figure of merit will vary as the link distance and power level change. Therefore, care should be taken when using the figure of merit to estimate the link performance at different power levels.

The projected performance of the direct-detection receiver is calculated and summarized in Table 2. The receiver figures of merit are calculated under night sky viewing and when Saturn is within the receiver field of view. The spectral irradiance of the night sky is assumed to be 5×10^{-6} W/m$^2 \cdot$sr$\cdot$mm. Given the background irradiance, the receiver bandwidth, and the optics efficiency, the total amount of background power collected by the receiver can be calculated. The maximum photon efficiency can then be calculated by assuming a modulation format with very small average to peak signal power ratio ($\sigma \to 0$). Shown in Table 2 are the maximum photon efficiencies achievable given the bandwidth and background noise constraints. It is seen that, when viewing night sky only, a photon efficiency of 15.2 bits/photon can be achieved with the 1-m receiver, whereas 19.8 bits/photon can be achieved using the 10-m photon bucket. When Saturn is in the field of view, however, the efficiency decreases to 9.3 bits/photon for the 1-m receiver and 7.7 bits/photon for the 10-m photon bucket. The better performance of the 1-m receiver when Saturn is in the FOV is due primarily to the reduced background power. The equivalent noise temperature of the receiver is then calculated using Eq. (13). The calculated equivalent noise temperatures are between 1970 and 25700 K when viewing night sky and between 4200 and 5070 K when Saturn is in the field of view.

Based on the equivalent noise temperature, a G/T figure of merit can then be calculated. The calculated receiver figure of merit is between 86.1 and 88.2 dB/K for the 1-m photon bucket and between 109.5 and 113.6 dB/K for the 10-m photon bucket. The variation is due to the reduction in receiver sensitivity due to the presence of background noise. One should note that the term G_R in Eq. (14) bears no relationship to the actual antenna gain, which is defined using its effective isotropic radiated power, since the photon bucket is not diffraction limited. Furthermore, the optimal receiver noise temperatures shown in Table 2 are calculated by assuming a low modulation duty cycle ($\sigma \rightarrow 0$) and a large bandwidth expansion factor. As a result, the limiting performance calculated in Table 2 is achievable only under a very low data rate. Practical direct-detection channels operate with a nonzero average to peak power ratio and hence have a higher noise temperature than that shown in Table 2.

V. Pulse Position Modulation for Direct-Detection Optical Channel

One practical modulation scheme which has low average to peak power ratio is pulse-position modulation (PPM), in which the information is conveyed through the channel by the time window in which the signal pulse is present [11,12]. In an M-ary PPM channel, each code word period is divided into M time slots. The transmit alphabet contains M symbols. Each symbol has a unique pulse location within the M time slots. At the receiving end, the decoder simply inspects the time windows and determines which time slot contains the signal pulse. Direct-detection PPM has been shown to be very effective in achieving high energy efficiency [13–15]. In the limit of infinite bandwidth expansion, i.e, at $M \rightarrow \infty$, PPM can achieve a photon efficiency limited only by the background noise level. Practical implementation constraints [10], however, limit the maximum order of PPM and hence impose an upper limit on the receiver sensitivity.

As expected, the performance of a PPM channel depends on the peak signal count rate, λ_S, the background count rate, λ_B, the modulation slot time, ΔT, and the order of modulation, M. Detailed derivation of the PPM channel capacity is given in Appendix B. Given the signal and background powers and slot time, ΔT, there exists a modulation order M that optimizes the channel efficiency. Shown in Fig. 2 are the maximum values of photon efficiency versus the average signal counts per slot for several background powers. Shown in Fig. 3 are the corresponding PPM alphabet sizes which achieve maximum photon efficiency. The maximum PPM order was artificially constrained to 256 to illustrate the effect of limited modulation bandwidth expansion. At high signal powers, the sensitivity of the receiver is limited by the modulation bandwidth, $1/\Delta T$, and the photon efficiency decreases with increasing average signal power. As the average power decreases, the photon efficiency can be improved by using a higher order PPM. If the maximum order of PPM is constrained, however, the photon efficiency will eventually reach a maximum. Further decrease in signal power will result in reducing photon efficiency. The sharp corners for the curves in Fig. 2 near 0.01 average signal counts/slot are due to the limiting PPM alphabet size.

It is seen from Fig. 2 that, in the limit of large signal power, the receiver sensitivity decreases. However, the channel capacity per slot, $R_{\max}\Delta T$, approaches a constant. This is because at high signal power, $R_{\max}$ is limited by the bandwidth. Shown in Fig. 4 is a plot of the capacity per slot versus the average signal count rate. Note that the limiting performance of approximately 0.53 bit/slot is achieved with average signal count per slot greater than 5.

The performance of the PPM modulated channel can be compared to that of a channel constrained only by the bandwidth. Shown in Table 3 are link design tables for the theoretically limited channel and two examples of PPM channels. It is seen that, for a Saturn return link using a 2-W transmitter and a 60-cm-diameter aperture, the maximum data rate sustainable is 52 Mbps. The maximum photon efficiency is 6.1 bits/photon and the equivalent noise temperature is 6,400 K. The capacity of an M-ary PPM channel, on the other hand, is somewhat lower. For a PPM channel with a 10-nsec slot time, the capacity is approximately 15 Mbps with an equivalent noise temperature of 22,170 K. For a PPM channel with a 1-nsec slot time, the capacity is 32 Mbps with an equivalent noise temperature of 10,240 K. Note that the effective photon efficiency and hence the receiver noise temperature change with modulation format. The $R \times D^2$ parameter can also be calculated based on realistic link parameters. Recall that the figure of merit defined in Eq. (14) is achievable only at the optimal signal level and modulation format. For a Saturn return link with the parameters shown in Table 3, the achievable photon efficiency is 6.1 bits/photon. This is very close to the 7.7-bits/photon limit based on the noise power. The resulting maximum $R \times D^2$ parameter is approximately 320 dB. For a PPM channel with a 1-nsec slot time, the $R \times D^2$ parameter is approximately 318.5 dB.

VI. Conclusion

The discussion of the direct-detection channel above illustrates the difficulty in defining a figure of merit of the direct-detection receiver. The Poisson statistics imply that the system cannot be characterized using only the signal-to-noise ratio. Instead, receiver sensitivity depends on both signal and noise power. Furthermore, sensitivity depends on the modulation format. One can trade system bandwidth (via reducing the average to peak power ratio) for photon efficiency. As a result, the performance of the channel can be characterized only by defining both ends of the system contrary to the microwave case.

Although the actual performance cannot be defined without knowing the signal and background powers, one can derive the limiting channel capacity based on the background noise power and bandwidth constraints and, from the limiting channel performance, define an equivalent receiver noise temperature. The system performance calculated using the G/T parameters, therefore, can be used as an indicator of the limiting channel performance given the bandwidth and noise power constraints. Based on the projected receiver parameters shown in Table 1, it is seen that the maximum G/T ratio for an optical receiver in 10 to 15 years is 113.6 dB/K when operating under a night-sky background. When Saturn or Jupiter is within the field of view, the G/T ratio decreases to approximately 109.5 dB/K. The limiting G/T performance calculated above is achieved at low average to peak power ratio and, consequently, at a very low data rate. For the practical link shown in Table 3, the G/T ratio is between 103.0 dB/K for a 16-ary PPM channel with a 10-nsec slot time, and 108.4 dB/K for a 64-ary PPM with a 1-nsec slot time.

It should be emphasized that the G/T ratio calculated in this article is an indicator of the limiting channel performance given the bandwidth and noise power constraints. Practical system performance can be substantially improved by increasing the receiver bandwidth. Furthermore, unlike RF channels where the noise temperature can be easily measured using a known background source, the equivalent noise temperature of a direct-detection PPM channel depends on both peak signal to background power ratio, ρ, as well as the average to peak signal power ratio, σ. Changes in link parameters such as aperture size, link distance, and transmitter power can result in changes in signal and noise powers admitted by the receiver and hence in the equivalent noise temperature. As a result, care must be taken in applying the equivalent noise temperature and hence the G/T ratio in projecting direct-detection link performance.

Acknowledgments

The author would like to acknowledge the following individuals for their helpful discussions: C. Stelzried, S. Butman, J. Lesh, T. Yan, and R. Clauss.

References

[1] W. L. Morgan and G. D. Gordon, *Communications Satellite Handbook*, New York: John Wiley & Sons, 1989.

[2] J. R. Pierce, "Optical Channels: Practical Limits With Photon Counting," *IEEE Trans. Commm.*, vol. COM-26, no. 12, pp. 1819–1821, 1978.

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

[4] P. L. Kelley and W. H. Kleiner, "Theory of Electromagnetic Field Measurement and Photoelectron Counting," *Phys. Rev.*, vol. 136, pp. A316–A334, 1964.

[5] S. Karp and J. R. Clark, "Photon Counting: A Problem in Classical Noise Theory," *IEEE Trans. Info. Theory*, vol. IT-16, no. 6, pp. 672–680, 1970.

[6] A. Papoulis, *Probability, Random Variables, and Stochastic Processes,* New York: McGraw-Hill, 1965.

[7] J. L. Massey, "Capacity, Cutoff Rate, and Coding for a Direct-Detection Optical Channel," *IEEE Trans. Commun.,* vol. COM-39, no. 11, pp. 1615–1621, November 1981.

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

[9] A. D. Wyner, "Capacity and Error Exponent for the Direct Detection Photon Channel—Part I," *IEEE Trans. Info. Theory,* vol. 34, no. 6, pp. 1449–1961, November 1988.

[10] S. A. Butman, J. Katz, and J. R. Lesh, "Practical Limitations on Noiseless Optical Channel Capacity," *DSN Progress Report,* vol. 42-55, Jet Propulsion Laboratory, Pasadena, California, pp. 12–14, November 1979.

[11] R. Gagliardi and S. Karp, "M-ary Poisson Detection and Optical Communications," *IEEE Trans. Comm. Theory,* vol. CT-17, no. 2, p. 208, 1969.

[12] R. J. McEliece, "Practical Codes for Photon Communications," *IEEE Trans. Info. Theory*, vol. IT-27, no. 4, pp. 393–397, 1981.

[13] J. R. Lesh, "Capacity Limit of the Noiseless Energy-Efficient Optical PPM Channel," *TDA Progress Report 42-67*, vol. July–September, Jet Propulsion Laboratory, Pasadena, California, pp. 54–58, November 15, 1981.

[14] R. G. Lipes, "Pulse-Position Modulation Coding as Near-Optimum Utilization of Photon Counting Channels with Bandwidth and Power Constraints," *DSN Progress Report*, vol. 42-56, Jet Propulsion Laboratory, Pasadena, California, pp. 108–113, January 1980.

[15] D. L. Zwillinger, "Maximizing Throughput Over an Average-Power Limited and Band-Limited Optical Pulse Position Modulation Channel," *TDA Progress Report 42-62*, vol. April–June 1981, Jet Propulsion Laboratory, Pasadena, California, pp. 81–86, January 1981.

Table 1. Projected receiver parameters for a direct detection optical receiver.

Parameters	Near term	Projected in 10–15 years
Detector quantum efficiency, μm	0.5 at 0.532	0.8 at 0.532
Receiver optics efficiency	0.4	>0.5
Filter transmissivity	0.6	>0.8
Filter bandwidth, nm	0.1	<0.001
Receiver diameter, m	≈1 diffraction limited	10, photon bucket
Receiver spatial mode mismatch	≈400	$\approx 4 \times 10^6$
Temporal bandwidth mismatch	$\approx 5 \times 10^3$	≈12
Receiver bandwidth, MHz	200	1000
Maximum data rate, Mbps	106	530
Detector dark count rate	2,000	200
Operating wavelength, nm	532	532

Table 2. Loss factors for a direct detection ground receiver.

Time, years	Near term		In 10–15 years	
	Night sky only	Saturn in FOV	Night sky only	Saturn in FOV
Antenna gain, dB[a]	134.5	134.5	154.5	154.5
Efficiency, dB				
Optics efficiency	−6.2	−6.2	−4.0	−4.0
Detector quantum efficiency	−3.0	−3.0	−1.0	−1.0
Atmospheric attenuation[b]	−3.0	−3.0	−3.0	−3.0
Receiver bandwidth, MHz	200	200	1000	1000
Total background count rate	2000	1.2×10^5	400	1.8×10^6
Maximum photon efficiency, bits/photon	15.2	9.3	19.8	7.7
Equivalent noise temperature, K	2570	4200	1970	5070
Figure of merit, dB/K	88.2	86.1	113.6	109.5

[a] Antenna gains are calculated by assuming the telescopes have 18.5-percent central obscuration area.

[b] Assume single site, 65-percent weather confidence at 30-deg elevation.

Table 3. Projected link performance for a Saturn return link using a 1-GHz bandwidth receiver.

Parameters	Theoretical limit	MPPM (10-nsec slot time)	MPPM (1-nsec slot time)
Transmitter power, W	2	2	2
Transmitter aperture, cm	60	60	60
Transmitter optics efficiency	>0.5	>0.5	>0.5
Link distance, AU	10	10	10
Atmospheric attenuation	0.5	0.5	0.5
Receiver aperture, m	10, photon bucket	10, photon bucket	10, photon bucket
Receiver optics efficiency	>0.5	>0.5	>0.5
Narrow-band filter transmission	>0.8	>0.8	>0.8
Filter bandwidth, nm	<0.001	<0.001	<0.001
Detector quantum efficiency	0.5	0.5	0.5
Receiver spatial mode mismatch	4×10^6	$\approx 4 \times 10^6$	$\approx 4 \times 10^6$
Receiver bandwidth, MHz	1000	1000	1000
Modulation	N/A		
Slot width, nsec		10	1
PPM order		$M = 16$	$M = 64$
Background count rate	1.8×10^6 (Saturn in FOV)	1.8×10^6 (Saturn in FOV)	1.8×10^6 (Saturn in FOV)
Signal photocount rate	8.6×10^6	8.6×10^6	8.6×10^6
Effective photon efficiency, bits/photon	6.1	1.76	3.81
Effective noise temperature, K	6,400	22,170	10,240
Channel capacity, Mbps	52	15.1	32.7
G/T value, dB/K	108.4	103.0	106.4
$R_{max} D^2$, dB(m^2/sec)	320.7	315.3	318.7

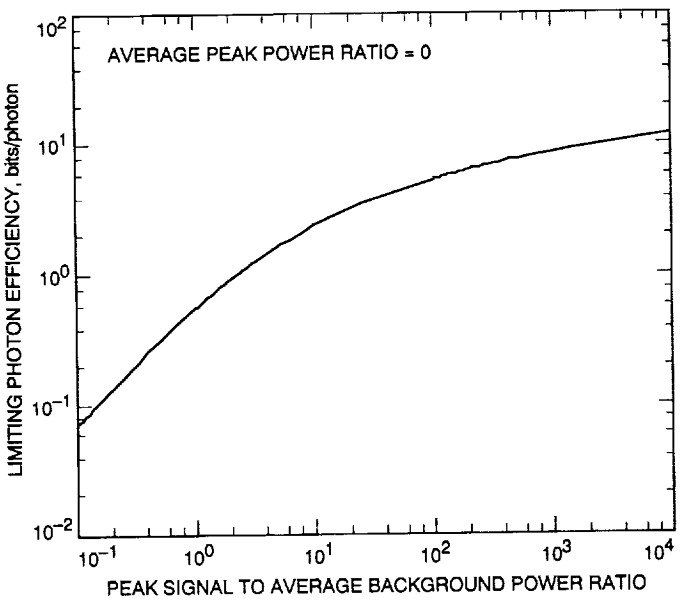

Fig. 1. Photon efficiency versus peak signal to average background power ratio in the limit of average power to peak power ratio equals 0 (i.e., with extremely low duty cycle).

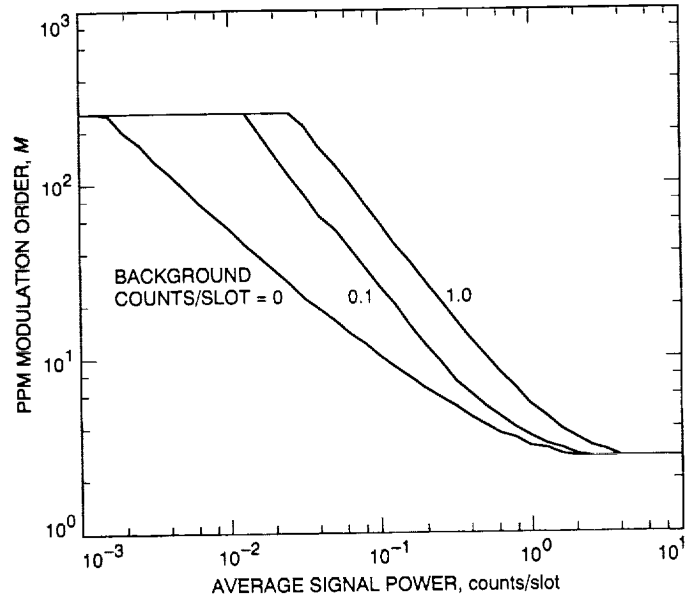

Fig. 3. PPM alphabet size that achieves maximum photon efficiency versus the average signal power for an *M*-ary PPM channel with constant slot width. Maximum PPM alphabet size is constrained to 256.

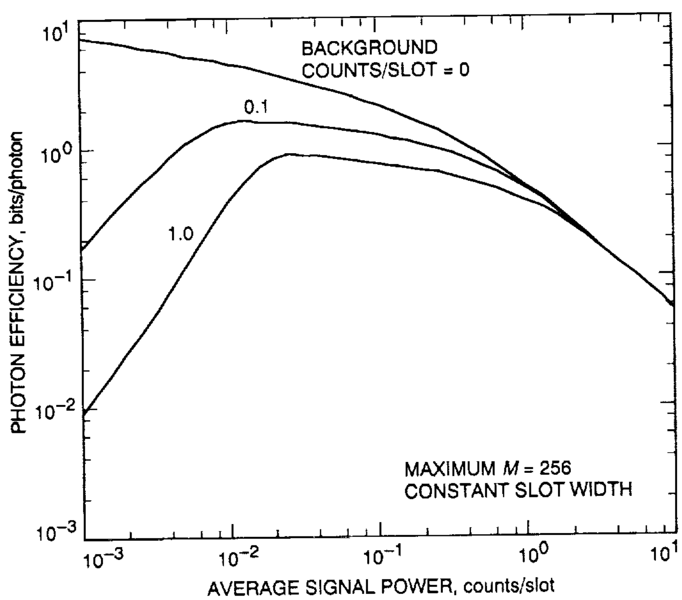

Fig. 2. Optimum photon efficiency versus average signal power for a PPM channel with constant slot width and maximum *M* constrained to 256.

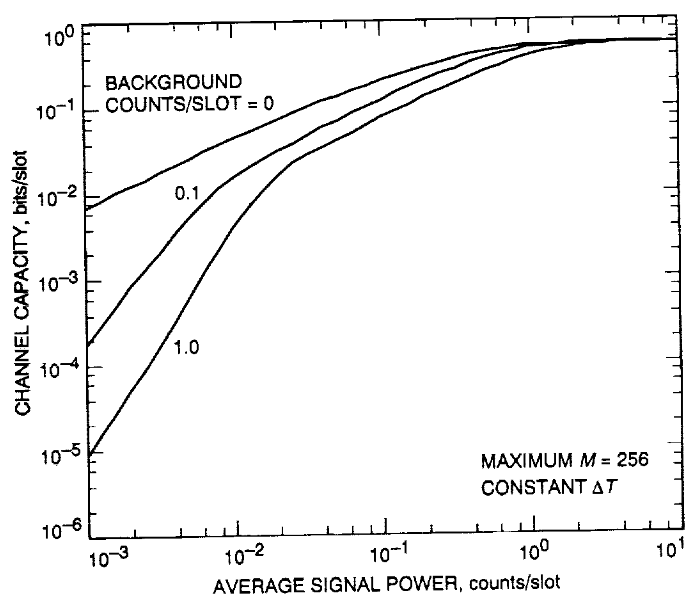

Fig. 4. Optimum channel capacity versus average signal power for an *M*-ary PPM channel for which the slot time is kept constant and the maximum *M* is constrained to 256. Note that in the limit of high signal power, the channel capacity approaches 0.53 bits/slot.

Appendix A

Received Background Power Calculation

The intensity pattern at the receiver focal plane can be related to the incoming signal amplitude and the wavefront quality of the primary aperture. For simplicity, it will be assumed that the incoming signal is a plane wave and that all the distortion due to the optical system can be summarized by the wavefront distortion of the primary aperture, $A(\mathbf{x})$, where $\mathbf{x}$ is the coordinate in the aperture plane. If it is further assumed that the incident light beam has unit intensity and is incident from an angle θ_S from the surface normal, the focal plane intensity pattern can be expressed as

$$I_f(\mathbf{r}) = \left(\frac{1}{\lambda f}\right)^2 \left|\mathcal{F}\left[A(\mathbf{x})\, e^{i\mathbf{k_S}\cdot\mathbf{x}}\right]_{\mathbf{r}/\lambda f}\right|^2 \cos\theta_S \qquad \text{(A-1)}$$

where $\mathbf{r}$ is the coordinate in the receiver focal plane, λ is the wavelength, f is the focal length of the optical system, and $\mathbf{k_S}$ is the incident wave vector. The factor $\cos\theta_S$ in Eq. (A-1) accounts for the reduction in signal intensity when the aperture is illuminated off angle. Equation (A-1) can be simplified by noting that the Fourier transform of a tilted wavefront results in a translation on the receiver focal plane:

$$I_f(\mathbf{r}) = \left(\frac{1}{\lambda f}\right)^2 \left|\mathcal{A}\left[\frac{\mathbf{r}-\mathbf{u_S}}{\lambda f}\right]\right|^2 \cos\theta_S \qquad \text{(A-2)}$$

where $\mathcal{A}(\mathbf{r})$ is the Fourier transform of the aperture distortion function $A(\mathbf{x})$. The translation on the focal plane, $\mathbf{u_S}$, is related to the angle of incidence by

$$|\mathbf{u_S}| = f\cos\theta_S \qquad \text{(A-3)}$$

The amount of power received by the detector is simply the integral of Eq. (A-1) over the detector area:

$$P_R = \int_r w_f(r) I_f(r) d^2 r \qquad \text{(A-4)}$$

where $I_f(r)$ is the intensity distribution over the detector focal plane, and $w_f(r)$ is the detector aperture function; $w_f(r) = 1$ if $r < r_D$, and $w_f(r) = 0$ if $r > r_D$. The integration of Eq. (A-4) can be equally carried out in angular space. The detector's angular field of view, Ω_D, defines the boundary of integration, and Eq. (A-4) can be equivalently written as

$$P_R = \int_{\Omega_D} \frac{1}{\lambda^2}\left|\mathcal{A}\left[\left(\frac{\theta-\theta_S}{\lambda}\right)\right]\right|^2 \cos\Omega_S d\Omega \qquad \text{(A-5)}$$

The receiver aperture, in principle, can be used for transmitting a signal. The far-field amplitude can again be related to the wavefront at the aperture using a Fourier transform relationship. The antenna gain $G(\Omega)$ is defined as the intensity at far field versus that of an isotropic radiator. That is,

$$G(\Omega) = \frac{4\pi z^2}{P_{in}} I(\Omega) \qquad \text{(A-6)}$$

where P_{in} is the input optical power, and z is the propagation distance. The far field intensity pattern resulting from an aperture pattern of $A(\mathbf{x})$ can be given by

$$I(\theta) = \left(\frac{1}{\lambda z}\right)^2 \left|\mathcal{A}\left[\frac{\theta}{\lambda}\right]\right|^2 \qquad \text{(A-7)}$$

By substituting the far field intensity pattern into the expression for antenna gain, it is seen that the antenna gain can be related to the Fourier transform of the aperture by

$$G(\theta) = \frac{16}{d_R^2\lambda^2}\left|\mathcal{A}\left[\frac{\theta}{\lambda}\right]\right|^2 \qquad \text{(A-8)}$$

Given the relationship between antenna gain and signal intensity, the total amount of background power received can then be calculated. The amount of power radiated by a patch of sky with solid angle $d\Omega_S$ can be characterized by Planck's radiation law. The intensity of light as seen on the receiver aperture can be written as

$$I_{d\Omega_S} = \left(\frac{2hf}{\lambda^2}\right)\frac{1}{e^{hf/k_B T}-1}\eta B \Delta f d\Omega_S \quad (\mathrm{W/cm^2 \mu m}) \qquad \text{(A-9)}$$

where the factor of 2 came from the two orthogonal polarization modes, and the parameter η is the emissivity of the blackbody. By combining Eqs. (A-1) through (A-9), the total power collected on the receiver from a patch of background source that subtends the solid angle $d\Omega_S$ is given by

$$P_R(\theta_S)d\Omega_S = \int_{\Omega_D} \frac{1}{\lambda^4}\left|\mathcal{A}\left[\frac{\mathbf{r}-\mathbf{u}_S}{\lambda f}\right]\right|^2 \times \cos\theta_S \frac{2\eta h f B}{e^{hf/k_B T}-1} B\Delta f d\Omega_S d\Omega$$

$$= \int_{\Omega_D}\left(\frac{d_R^2}{16\lambda^2}\right) G(\theta-\theta_S) \times \cos\theta_S \frac{2\eta h f B}{e^{hf/k_B T(\Omega_S)}-1}\Delta f d\Omega d\Omega_S \qquad \text{(A-10)}$$

Integration of Eq. (A-10) over the entire sky then gives the expression of total background power collected.

$$P_N = \int\int_{\Omega_D}\left(\frac{d_R^2}{16\lambda^2}\right) G(\theta-\theta_S) \times \cos\theta_S \frac{2\eta h f B}{e^{hf/k_B T}-1} d\Omega d\Omega_S \qquad \text{(A-11)}$$

where the integration is performed first over the detector field of view, Ω_D, then over the extent of the background source distribution. For a nondiffraction-limited receiver, the integral given by Eq. (A-11) can be substantially larger than the power received by a single mode receiver. Consequently, a nondiffraction-limited receiver can admit much more noise than a diffraction-limited one.

A simple figure of merit of the optical receiver is the number of background noise modes it collects. This spatial mode mismatch factor, F_B, is given by

$$F_B = \int\int_{\Omega_D}\left(\frac{d_R^2}{16\lambda^2}\right) G(\theta-\theta_S)\cos\theta_S d\Omega d\Omega_S \qquad \text{(A-12)}$$

The number of background modes, F_B, varies from 1 for a diffraction-limited receiver to greater than 2×10^6 for the proposed 10-m photon-bucket receiver.

Appendix B

Capacity of Optical PPM Channel

The capacity per channel use for a general M-ary channel is defined as

$$C = I(\mathbf{x};\mathbf{y}) = H(\mathbf{y}) - H(\mathbf{y}|\mathbf{x}) \qquad \text{(B-1)}$$

where $\mathbf{x}$ and $\mathbf{y}$ represent the transmit and receive symbol sets, respectively, $I(\mathbf{x};\mathbf{y})$ is the mutual information function, $H(\mathbf{y})$ is the entropy of $\mathbf{y}$, and $H(\mathbf{y}|\mathbf{x})$ is the conditional entropy function. The capacity per transmission is given as the entropy of the received symbol set minus the amount of conditional entropy between the transmission and reception signal set. If the transmission is perfect, the conditional entropy between $\mathbf{x}$ and $\mathbf{y}$ is zero, and the capacity per transmission is simply equal to the entropy of the signal set.

For PPM systems, the demodulator can be implemented as follows: The receiver simply counts the number of photons that are received during the M time slots, and then chooses the slot with the largest photocount. If there is more than one slot having the largest count, the demodulator can either randomly choose between these slots, or it can assign a special erasure symbol to the output. In the first case, the PPM channel can be modeled as an M-ary symmetric channel with probability of correct reception p, and probability of erroneous decision q, for each of the remaining $(M-1)$ symbols. The probabilities p and q are given by

$$p = \frac{1}{M}e^{-K_S+MK_B} + \sum_{k=1}^{\infty} \frac{(K_S+K_B)^k}{k!}e^{-(K_S+K_B)}$$

$$\times \left[\sum_{m=0}^{k-1} \frac{K_B^m}{m!}e^{-K_B}\right]^{M-1} \frac{1}{aM}\left[(1+a)^M - 1\right]$$

$$a \equiv \frac{K_B^k}{k!} \bigg/ \sum_{m=0}^{k-1} \frac{K_B^m}{m!} \qquad \text{(B-2a)}$$

$$q = (1-p)/(M-1) \qquad \text{(B-2b)}$$

where $K_S = \lambda_S \Delta T$ and $K_B = \lambda_B \Delta T$ are the average signal and background counts received during one time slot. By assuming that the M transmission symbols have equal probability, the capacity per word (channel use) for this M-ary symmetric channel is given by

$$C_{\mathrm{MSC}} = \log M + p\log p + (M-1)q\log q \qquad \text{(B-3)}$$

Similarly, the PPM channel can also be modeled as an M-ary erasure channel with probability of correct transmission α, and probability of erasure γ. The probability of erroneous decoding to the remaining $(M-1)$ symbol is given by $\beta = (1-\alpha-\gamma)/(M-1)$. The probabilities can be given by

$$\alpha = \sum_{k=1}^{\infty} \frac{(K_S+K_B)^k}{k!}e^{-(K_S+K_B)}$$

$$\times \left[\sum_{m=0}^{k-1}\left(\frac{K_B^m}{m!}e^{-K_B}\right)\right]^{M-1} \qquad \text{(B-4a)}$$

$$\beta = \sum_{k=1}^{\infty} \frac{K_B^k}{k!}e^{-K_B}\left[\sum_{m=0}^{k-1}\frac{K_B^m}{m!}e^{-K_B}\right]^{M-2}$$

$$\times \left[\sum_{m=0}^{k-1}\frac{(K_S+K_B)^m}{m!}e^{-(K_S+K_B)}\right] \qquad \text{(B-4b)}$$

$$\gamma = 1 - \alpha - (M-1)\beta \qquad \text{(B-4c)}$$

Again, if it is assumed that the transmission symbols have equal probabilities, the capacity per channel use is given by

$$C_{\mathrm{MEC}} = (1-\gamma)\log\frac{M}{1-\gamma}$$

$$+ \alpha\log\alpha + (M-1)\beta\log\beta \qquad \text{(B-5)}$$

In the limit where the background noise is zero, the probability of error β goes to zero, and the capacity (per word) of the PPM channel reduces to the familiar form for the M-ary erasure channel studied by Pierce:

$$C_{\mathrm{MEC}} = \alpha \log M = (1 - e^{K_S}) \log M \qquad \text{(B-6)}$$

The capacity per channel use is slightly different for the two demodulation schemes. For this calculation, the M-ary PPM channel is modeled as an erasure channel.

Given the capacity per channel use and the expected signal count, the photon efficiency, C_{ph}, is given by

$$C_{\mathrm{ph}} = C_{\mathrm{MEC}}/K_S \qquad \text{(B-7)}$$

The capacity per second, $R_{\max}$, is given by

$$R_{\max} = \frac{C_{\mathrm{MEC}}}{M\Delta T} \qquad \text{(B-8)}$$

and the capacity per slot, $C_{\Delta T}$, is simply

$$C_{\Delta T} = \frac{C_{\mathrm{MEC}}}{M} \qquad \text{(B-9)}$$