# The Asymptotic Capacity of the Direct Detection Photon Channel With a Bandwidth Constraint

David Brady
Department of Electrical and Computer Engineering
Northeastern University
Boston, MA 02115

Sergio Verdú
Department of Electrical Engineering
Princeton University
Princeton, NJ 08544

### 1. *Abstract*

In this paper we address the capacity of a direct detection photon counting channel with an average and symbol period constraint on the input. This channel model is ideally suited for, but not limited to, a communication system comprised of a semiconductor laser with electro-optic modulation, a fiber optic transmission medium, and an integrate-and-dump receiver. We find the limiting form of the capacity as the signal and noise energies increase proportionally and we show that the capacity grows logarithmically with the signal energy constraint.

### 2. *Introduction*

Consider a discrete-time memoryless channel with a nonnegative input $\mathcal{X}_i$ which is constrained in peak, $\mathcal{X}_i \leq M$, and sample average, $\bar{\mathcal{X}} \leq \sigma M$. The output of this channel is a conditionally Poisson count $\mathcal{Y}_i$ given the input $\mathcal{X}_i$. Under this conditioning $\mathcal{Y}_i$ has mean $\mathcal{X}_i + d$, where $d$ represents the average count due to additive, independent noise. This channel is called the pulsewidth-limited direct detection photon channel (PLPC) and is a suitable model for a fiber optic communication channel with photoelectric conversion at the receiver, especially when semiconductor laser sources and long-haul networks are used. With respect to this continuous-time application, the peak input constraint $M$ represents the product of the peak received optical power $A$ and the fixed symbol period $\Delta$, $\sigma$ is the ratio of the average laser power to the peak laser power, and $\lambda \stackrel{\Delta}{=} \frac{d}{\Delta}$ is the dark current intensity at the photodetector. In practical applications the fixed symbol period $\Delta$ is an important constraint, and represents the limitations on the rate of communication imposed by electro-optic modulation, fiber dispersion, and other system anomolies. The PLPC is a suitable model for the fiber optic channel provided that the temporal information about the photons (or photoelectrons) is ignored, and the output is restricted to the count in each subinterval. Since complete point process observations are difficult to achieve in practice, and subinterval counting may be implemented by integrating the photocurrent, this restriction represents a technological limitation. Note that the channel input is distorted in two ways: by the additive noise due to dark current, and by the inevitable quantum or self-noise due to the statistical nature of electromagnetic radiation.

It is of interest to find the capacity of this channel for arbitrary values of $A$, $\sigma$, $\lambda$ and $\Delta$. Unfortunately, the capacity expression does not have an explicit form, due in part to the lack of a closed form for the entropy of the Poisson distribution. Instead, previous work has addressed the asymptotic capacity, as some of the constraints vanish. By ignoring the modulation rate limitation, for example, it has been shown that the capacity of the PLPC (in nats per second) is given by [1], [2], [3], [4]

$$\lim_{\Delta\to 0}\frac{C}{\Delta}=\begin{cases}\sigma^*(A+\lambda)\log(1+\frac{\lambda}{A})+(1-\sigma^*)\lambda\log(\frac{\lambda}{A}), & \sigma^*\le\sigma\\ \sigma(A+\lambda)\log(1+\frac{\lambda}{A})+(1-\sigma)\lambda\log(\frac{\lambda}{A}), & \sigma^*>\sigma\end{cases}\tag{1}$$

where $\sigma^* A$ is the optimum average input for the simpler problem when $\sigma = 1$, i.e., the average input constraint is inactive, and $\sigma^*$ is given by

$$\sigma^*=\frac{\left(1+\frac{\lambda}{A}\right)^{1+\frac{\lambda}{A}}}{\left(\frac{\lambda}{A}\right)^{\frac{\lambda}{A}}\mathrm{e}}-\frac{\lambda}{A}.$$

In terms of the PLPC constraints, this asymptotic result is equivalent to $M\to 0$ or $\bar{X}\to 0$ with fixed $SNR \triangleq \bar{X}/d$. It is interesting to note that as $\Delta\to 0$, the PLPC channel capacity coincides with that of a channel whose output is the complete point process observation. This fact is due to the conditional orderliness, or "non-explosiveness" property of the double-stochastic point process. Since the PLPC restricts the information of each photoelectron arrival time to one of the $\Delta$-width subintervals, conditional orderliness assures that as $\Delta\to 0$, at most one photoelectron arrives in each subinterval, and the arrival time of each is known in this limit.

Several works have also addressed the limiting form of the PLPC capacity for a non-trivial modulation constraint $\Delta$. A lower bound to the PLPC capacity was investigated for an inactive peak constraint, $d = 0$, and $\bar{\mathcal{X}} \to \infty$ [5] . These restrictions represent the case in which the quantum model of electromagnetic propagation approaches the continuous model, and both the quantum and dark current noise may be neglected. It was shown that the capacity of the noiseless PLPC grows at least as fast as $\log\sqrt{\bar{\mathcal{X}}}$

$$\lim_{\substack{\bar{\mathcal{X}}\to\infty \\ SNR=\infty}} \frac{C}{\log\sqrt{\frac{\bar{\mathcal{X}}}{2\pi}}} \geq 1.$$

A lower bound to the PLPC capacity with an inactive average constraint was investigated by considering discrete input distributions having $K$ equally-spaced, equally-likely points of mass on $[0, M]$ [6] . For fixed $M$ and $d$, the mutual information was optimized numerically over $K$. A lower bound to the PLPC capacity was derived for $d = 0$ and large $M$

$$\lim_{\substack{M\to\infty \\ SNR=\infty}} \frac{C}{\log\sqrt{\frac{M}{2\pi}}} \geq 1. \tag{2}$$

In the absence of dark current, it was noted that the optimal spacing between the mass points was roughly $2.7\sqrt{M}$ for large $M$.

It has been shown recently that the capacity of the PLPC with an active peak constraint is achieved by a discrete input distribution having a finite set of mass points on $[0, M]$ [7]. This was demonstrated by arguments paralleling those of Smith [8] for the peak-constrained, additive Gaussian noise channel. It was proven that a binary input distribution maximizes the mutual information between channel input and output, not only as $M \to 0$, but also for $0 < M < \phi$, where $\phi$ is strictly positive. It was shown that equally-spaced, uniform, discrete input distributions do not maximize the mutual information between input and output.

In this paper we present the limiting form of the PLPC capacity for an inactive peak constraint, and for sufficiently large $\bar{\mathcal{X}}$. Since $M = \infty$, the results from [7] do not apply, and it is not known whether a discrete input distribution maximizes $\mathcal{I}(\mathcal{X}_i; \mathcal{Y}_i)$. Unlike previous works with an inactive peak constraint, we consider the practical case when the dark current cannot be neglected. In fact, we will be interested in the case when $d$ may dominate the signal energy, $\bar{\mathcal{X}}$. We develop bounds on the PLPC capacity for increasing $\bar{\mathcal{X}}$ and constant $SNR$, and show that these bounds tighten as $\bar{\mathcal{X}} \to \infty$. Since $d$ and $\bar{\mathcal{X}}$ are average counts, this represents a situation in which the signal and noise counts grow at the same rate, either by increasing the subinterval length or by lowering the carrier frequency of the system (and the energy per photon). Although this is a case in which the semi-classical model approaches the classical model, we should not expect the PLPC capacity to approach the (finite) capacity of the discrete time additive Gaussian noise channel with an energy constraint. In fact, it will be demonstrated that the PLPC capacity is bounded for sufficiently large $\bar{\mathcal{X}}$ by

$$K_l \leq \mathcal{C} - \log\sqrt{\frac{\bar{\mathcal{X}}}{2\pi}} \leq K_u$$

where $K_l, K_u$ are constants that depend only on $SNR$. This demonstrates that the lower bound on the asymptotic capacity with $SNR = \infty$ derived in [5] is also a lower bound to the (smaller) capacity of the PLPC with arbitrarily large dark current, $\lambda$. More importantly, we demonstrate that the lower bound is tight. The positive growth of the PLPC capacity for fixed $SNR$ is not intuitive, and suggests that the Poisson noise count is "less random" as the noise and signal energies increase proportionally. This result is an artifact of the Poisson distribution, and complements previous work on singular detection in additive Poisson noise [9].

### 3. *Analysis*

Throughout this section $\Pi$ denotes a Poisson random variable with mean $m$. It will be demonstrated presently that the expectation of $\log \Pi!$ coincides with $\log m!$ for large $m$ by using Stirling's bounds to $k!$

$$\lfloor k \rceil ! \triangleq \sqrt{2\pi k}\left(\frac{k}{\mathrm{e}}\right)^k \leq k! \leq \lfloor k \rceil ! \left(1 + \frac{1}{12k-1}\right) \tag{3}$$

A reciprocal moment of the Poisson random variable will also be used in this demonstration. While reciprocal moments are difficult to compute for most random variables, $\mathrm{E}\left[\frac{1}{\Pi+1} \mathit{1}_{\{\Pi \geq k\}}\right]$ can be expressed in compact form.

**Lemma 1.** For any integer $k$,

$$\mathrm{E}\left[\frac{1}{\Pi+1} \mathit{1}_{\{\Pi \geq k\}}\right] = \frac{1}{m}\mathcal{P}\left[\Pi \geq k+1\right]$$

**Proof.** Immediate.

As shown in Lemma 2, this reciprocal moment is useful in finding tight bounds to the expectation of $\log \Pi$.

**Lemma 2.**

$$\log m - \log \frac{3}{2} + o(m^3 e^{-m}) \le \mathrm{E}\left[\log \Pi \mathbb{1}_{\{\Pi \ge 2\}}\right] \le \mathcal{P}[\Pi \ge 2] \log m - \mathcal{P}[\Pi \ge 2] \log \{\mathcal{P}[\Pi \ge 2]\}$$

**Proof.** The upper bound is immediate from Jensen's inequality, while the lower bound is most easily seen from the inequality

$$\frac{1}{k} \le \frac{3}{2}\frac{1}{1+k} \quad \forall k \ge 2$$

Over the range of integration, it follows that

$$\log \Pi \ge \log m - \log\left(\frac{3}{2}\frac{m}{1+\Pi}\right)$$

Taking the expectation over both sides for $\Pi \ge 2$, and applying Jensen's inequality to the second term, we have

$$\mathrm{E}\left[\log \Pi \mathbb{1}_{\{\Pi \ge 2\}}\right] \ge \mathcal{P}[\Pi \ge 2] \left\{ \log m - \log \left[ \frac{\frac{3}{2}\mathrm{E}\left[\frac{m}{\Pi+1}\mathbb{1}_{\{\Pi \ge 2\}}\right]}{\mathcal{P}[\Pi \ge 2]} \right] \right\}$$

Using Lemma 1 and the inequality $\log x \le x - 1$, we get

$$\mathrm{E}\left[\log \Pi \mathbb{1}_{\{\Pi \ge 2\}}\right] \ge \log m - \log \frac{3}{2} - \left\{ e^{-m} + m e^{-m} \right\} \log m + \frac{m^2}{2!} e^{-m}$$

As seen from (3), Lemma 2 will be used to establish bounds on $\mathrm{E}[\log \Pi!]$. In addition, bounds on $\mathrm{E}[\Pi \log \Pi]$ will also be required. In Lemma 3, we use Lemma 1 to bound this expectation.

**Lemma 3.**

$$m \log m \le \mathrm{E}\left[\Pi \log \Pi \mathbb{1}_{\{\Pi \ge 2\}}\right] \le m \log m + 1$$

**Proof.** The left hand side follows from Jensen's inequality, while the right hand side follows from

$$\Pi \log \Pi = \Pi \log m + \Pi \log \left\{1 + \frac{\Pi - m}{m}\right\}$$

$$\leq \Pi \log m + \Pi \frac{\Pi - m}{m}$$

Taking the expectation of the right hand side for all values of $\Pi$ produces the upper bound in the statement of the lemma.∎

As shown in (3) the tightness of the bounds of $\mathrm{E}\left[\log \Pi! \mathbb{1}_{\{\Pi \geq 2\}}\right]$ depends in part on the tightness of Stirling's bounds, which may be determined by $\mathrm{E}\left[\log\left(1 + \frac{1}{12\Pi - 1}\right) \mathbb{1}_{\{\Pi \geq 2\}}\right]$. In Lemma 4, we show that the bounds converge at least as fast as $\frac{1}{m}$.

**Lemma 4.**

$$\mathrm{E}\left[\log\left(1 + \frac{1}{12\Pi - 1}\right) \mathbb{1}_{\{\Pi \geq 2\}}\right] \leq \frac{1}{7m} \mathcal{P}\left[\Pi \geq 3\right]$$

**Proof.** For all $\Pi \geq 2$ we have

$$\frac{1}{12\Pi - 1} \leq \frac{1}{7(\Pi + 1)}$$

The lemma follows immediately from this inequality and Lemma 1.∎

From the Stirling bounds in (3) it is clear that bounds on $\mathrm{E}\left[\log \Pi!\right]$ follow from (3), as well as Lemmas 1 through 4. We combine and simplify these results in the following theorem.

**Theorem 1.** Let $\Pi$ be a Poisson random variable with mean $m$. Given $\epsilon > 0$, there exists an $m_\epsilon$ so that for all $m > m_\epsilon$

$$\log \lfloor m \rceil! - \log \frac{3}{2} - \epsilon \leq \mathrm{E}\left[\log \Pi!\right] \leq \log \lfloor m \rceil! + 1 + \epsilon$$

where $\lfloor m \rceil!$ is Stirling's lower bound to $m!$.∎

Theorem 1 allows us to find an asymptotic form for the Poisson entropy. Recalling that $\Pi$ is a Poisson random variable with mean $m$, the Poisson entropy may be written

$$H_P(m) \triangleq -\mathrm{E}\left[\log \frac{m^{\Pi} e^{-m}}{\Pi!}\right], \tag{4}$$

and our immediate aim is to show that

$$\lim_{m\to\infty} \frac{H_P(m)}{-\log \frac{m^m e^{-m}}{m!}} = 1.$$

The only quantity that is not known explicitly in (4) is $\mathrm{E}[\log \Pi!]$. As shown in Theorem 1, we can bound this expectation by quantities that differ by a constant as $m \to \infty$. More precisely, given $\epsilon > 0$, we may find an $m_\epsilon$ such that for all $m > m_\epsilon$,

$$\log \lfloor m \rceil ! - \log \frac{3}{2} - \epsilon \leq \mathrm{E}[\log \Pi!] \leq \log \lfloor m \rceil ! + 1 + \epsilon$$

where $\lfloor m \rceil !$ is Stirling's lower bound to $m!$. Using this result we tightly bound the Poisson entropy as the mean parameter $m$ grows.

**Theorem 2.** Let $\Pi$ be a Poisson random variable with mean $m$. Given $\epsilon > 0$, there exists an $m_\epsilon$ so that for all $m > m_\epsilon$ the Poisson entropy is bounded in the following way:

$$-\log \frac{m^m e^{-m}}{\lfloor m \rceil !} - \log \frac{3}{2} - \epsilon \leq H_P(m) \leq -\log \frac{m^m e^{-m}}{\lfloor m \rceil !} + 1 + \epsilon$$

where $\lfloor m \rceil !$ is Stirling's lower bound to $m!$.∎

This result is applicable to the PLPC channel, since the mutual information between the $\mathcal{X}_i$ and $\mathcal{Y}_i$ may be described in term of the Poisson entropy,

$$\mathcal{I}(\mathcal{X}_i; \mathcal{Y}_i) = H(\mathcal{Y}_i) - \mathrm{E}[H_P(\mathcal{X}_i + d)].$$

Conditioned on the input, the mean of $\mathcal{Y}_i$ is bounded from below by $d$, which grows arbitrarily large in the limit. This justifies the application of Theorem 2 to bound $\mathcal{I}(\mathcal{X}_i; \mathcal{Y}_i)$ for sufficiently large $d$.

**Theorem 3.** Given $\epsilon > 0$ there exists a $d_\epsilon$ such that for all average noise energies $d > d_\epsilon$ and all distributions on $\mathcal{X}$ such that $\mathcal{X} \geq 0$ *a.s.*,

$$-1 - \epsilon \leq \mathcal{I}(\mathcal{X}_i; \mathcal{Y}_i) - H(\mathcal{Y}_i) - \mathrm{E}\left[\log \frac{(\mathcal{X}_i + d)^{(\mathcal{X}_i + d)} e^{-(\mathcal{X}_i + d)}}{\lfloor (\mathcal{X}_i + d) \rceil !}\right] \leq \log \frac{3}{2} + \epsilon.$$

∎

By substituting the definition of Stirling's lower bound, these bounds take a form which is more amenable to bounding

$$-1-\epsilon \leq \mathcal{I}(\mathcal{X}_i;\mathcal{Y}_i) - H(\mathcal{Y}_i) + \mathrm{E}\left[\frac{1}{2}\log[2\pi(\mathcal{X}_i+d)]\right] \leq \log\frac{3}{2}+\epsilon. \tag{5}$$

We shall develop bounds on the capacity of the PLPC by bounding the output entropy $H(\mathcal{Y}_i)$ in (5). As in [5] we achieve a lower bound on $\mathcal{I}(\mathcal{X}_i;\mathcal{Y}_i)$ by assuming an exponential distribution on the input.

**Theorem 4.** Let $\bar{\mathcal{X}}, d \to \infty$ proportionally. Given $\epsilon > 0$ there exists an $\bar{\mathcal{X}}_\epsilon$ such that for all $\bar{\mathcal{X}} > \bar{\mathcal{X}}_\epsilon$ the capacity of the PLPC is lower bound by

$$\mathcal{C} \geq \frac{1}{2}\log\frac{\bar{\mathcal{X}}}{2\pi} - \frac{1}{2}\log\left(1+\frac{1}{SNR}\right) - \epsilon$$

where $SNR = \bar{\mathcal{X}}/d$.

**Proof.** When the input to the channel is exponential with mean $\bar{\mathcal{X}}$, it can be shown that the output distribution has the form

$$\begin{aligned}\mathcal{P}[\mathcal{Y}=y] &= (1-\rho)\rho^y e^{\frac{1}{SNR}}\mathcal{P}[\Pi \leq y]\\ &\leq (1-\rho)\rho^y e^{\frac{1}{SNR}}\end{aligned} \tag{6}$$

where $\rho \triangleq \frac{\bar{\mathcal{X}}}{1+\bar{\mathcal{X}}}$, and $\Pi$ is a Poisson random variable with mean $m = d + \frac{1}{SNR}$. A lower bound on the output entropy follows from (6),

$$H(\mathcal{Y}_i) \geq \log[1+\bar{\mathcal{X}}] + (d+\bar{\mathcal{X}})\log\left[1+\frac{1}{\bar{\mathcal{X}}}\right] - \frac{d}{\bar{\mathcal{X}}} \tag{7}$$

The theorem follows by substituting (7) in (5), and by applying Jensen's inequality to the term $\mathrm{E}\left[\frac{1}{2}\log[2\pi(\mathcal{X}_i+d)]\right]$ in (5). ∎

We achieve the upper bound to the PLPC capacity by lower bounding the term $\mathrm{E}\left[\frac{1}{2}\log[2\pi(\mathcal{X}_i+d)]\right]$ in (5), and by assuming that there is some distribution on $\mathcal{X}_i$ such that $\mathcal{Y}_i$ is geometric with mean $\bar{\mathcal{X}}+d$. This follows from the fact that the geometric random variable maximizes the entropy functional over all distributions on the integers subject to a constrained mean. While it can be shown that there is no distribution on $\mathcal{X}_i$ such that $\mathcal{Y}_i$ is geometrically distributed for $d>0$, it is clear that the entropy of a geometric random variable with mean $\bar{\mathcal{X}}+d$ is an upper bound to $H(\mathcal{Y}_i)$.

**Theorem 5.** Let $d, \bar{\mathcal{X}} \to \infty$ proportionally. Given $\epsilon > 0$ there exists a $\bar{\mathcal{X}}_\epsilon$ such that for all $\bar{\mathcal{X}} > \bar{\mathcal{X}}_\epsilon$ the capacity of the PLPC is upper bounded by

$$\mathcal{C} \leq \frac{1}{2}\log\frac{\bar{\mathcal{X}}}{2\pi} + \log\left(\sqrt{SNR}(1+\frac{1}{\bar{\mathcal{X}}_\epsilon}) + \frac{1}{\sqrt{SNR}}\right) + 1 + \log\frac{3}{2} + \epsilon. \tag{8}$$

**Proof.** The upper bound may be achieved by replacing the expectation in (5) by the minimum value of the integrand, $\frac{1}{2}\log 2\pi d$, and by using the entropy of the geometric random variable in place of $H(\mathcal{Y}_i)$. In this way we find

$$H(\mathcal{Y}_i) \leq \log(1+\bar{\mathcal{X}}+d) + (\bar{\mathcal{X}}+d)\log\left(1+\frac{1}{\bar{\mathcal{X}}+d}\right)$$

and the bound follows immediately.∎

## 4. References


[1] Y.M. Kabanov, "*The Capacity of a Channel of the Poisson Type*", *Theory of Probability and Applications,* **Vol. 23,** *January, 1978, pp. 143-147*

[2] A.A. Lazar, "*On the Capacity of the Poisson Type Channel*", *Proceedings of the Fourteenth Annual Conference on Information Sciences and Systems, Princeton, NJ, March 26-28, 1980, pp.285-289*

[3] M.H.A. Davis, "*Capacity and Cutoff Rate for Poisson-Type Channels*", *IEEE Transactions on Information Theory,* **Vol. IT-26, No. 6,** *November, 1980, pp. 710-715*

[4] A.D. Wyner, "*Capacity and Error Exponent for the Direct Detection Photon Channel-Part I*", *IEEE Transactions on Information Theory,* **Vol. IT-34, No. 6,** *November, 1988, pp. 1449-1461*

[5] J.P. Gordon, "*Quantum Effects in Communications Systems*", *Proceedings of the IRE,* **Vol. 50,** *September, 1962, pp. 1898-1908*

[6] E. Hisdal, "*Information in a Photon Beam vs. Modulation-Level Spacing,*" *Journal of the Optical Society of America,* **Vol. 61, No.3,** *March 1971, pp.328-332*

[7] S. Shamai (Shitz), "*The Capacity of a Pulse Amplitude Modulated Direct Detection Photon Channel,*" *preprint*

[8] J.G. Smith, "*The Information Capacity of Amplitude- and Variance-Constrained Scalar Gaussian Channels,*" *Information and Control,* **Vol. 18,** *pp.203-219, 1971*

[9] I. Bar-David, "*Error Probabilities in Random-Phase Binary Optical Signalling*", *IEEE Transactions on Information Theory, May, 1974*