The Impact of Dark Current on the Wideband Poisson Channel

Ligong Wang and Gregory W. Wornell
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, MA, USA
{wlg,gww}@mit.edu

***Abstract*—We study the discrete-time Poisson channel under the constraint that its average input power (in photons per channel use) must not exceed some constant $\mathcal{E}$. We consider the wideband, high-photon-efficiency extreme where $\mathcal{E}$ approaches zero, and where the channel's "dark current" approaches zero proportionally with $\mathcal{E}$. Extending our previous work, we show that the influence of the dark current on channel capacity is mainly on the third-order term with respect to $\mathcal{E}$. We also show that pulse-position modulation with "soft-decision decoding" achieves data rates that accurately reflect such influence.**

I. Introduction

We consider the discrete-time memoryless Poisson channel whose input x is in the set $\mathbb{R}_0^+$ of nonnegative reals and whose output y is in the set $\mathbb{Z}_0^+$ of nonnegative integers. Conditional on the input $X = x$, the output Y has a Poisson distribution of mean $(\lambda + x)$, where $\lambda \geq 0$ is called the "dark current" and is a constant which does not depend on the input x. Hence the channel law $W(\cdot|\cdot)$ is

$$W(y|x) = e^{-(\lambda+x)} \frac{(\lambda + x)^y}{y!}, \quad x \in \mathbb{R}_0^+, y \in \mathbb{Z}_0^+. \tag{1}$$

This channel models pulse-amplitude modulated optical communication where the transmitter sends light signals in *coherent states* (which are usually produced using laser devices), and where the receiver employs *direct detection* (i.e., photon counting) [1]. The channel input x describes the expected number of *signal photons* (i.e., photons that come from the input light signal rather than noise) to be detected in the pulse duration, and is proportional to the light signal's intensity, the pulse duration, the channel's transmissivity, and the detector's efficiency; the channel output y is the actual number of photons that are detected in the pulse duration; and λ is the average number of extraneous counts that appear in y due to background radiation or to the detector's "dark clicks".

We impose an *average-power constraint*[1] on the input

$$\mathsf{E}[X] \leq \mathcal{E} \tag{2}$$

for some $\mathcal{E} > 0$.

This work was supported in part by the DARPA InPho program under Contract No. HR0011-10-C-0159, and by AFOSR under Grant No. FA9550-11-1-0183.

[1]Here "power" is in discrete time, means expected number of detected photons per channel use, and is proportional to the continuous-time physical power times the pulse duration.

In applications like free-space or outer-space optical communications, the cost of producing and successfully transmitting photons is high, hence high *photon efficiency* (information transmitted per photon) is desirable. As previously demonstrated [2], [3], this can be achieved in the *wideband* regime, where the pulse duration of the input approaches zero and, assuming that the *continuous-time* average input power is fixed, where $\mathcal{E}$ approaches zero proportionally with the pulse duration. Note that in this regime the average number of detected background photons or dark clicks also tends to zero proportionally with the pulse duration. Hence we have the linear relation

$$\lambda = c\mathcal{E}, \tag{3}$$

where c is some nonnegative constant that does not change with $\mathcal{E}$. Asymptotic results in this regime are relevant in scenarios where $\mathcal{E}$ is small and where λ is comparable to or much smaller than $\mathcal{E}$.

We denote the capacity (in nats) of the channel (1) under power constraint (2) with dark current (3) by $C(\mathcal{E}, c)$, then

$$C(\mathcal{E}, c) = \max_{\mathsf{E}[X] \leq \mathcal{E}} I(X;Y), \tag{4}$$

where the mutual information is computed from the channel law (1) and is maximized over input distributions satisfying (2), with dark current λ given by (3).

Various capacity results for the discrete-time Poisson channel have been obtained [2]–[8]. In particular, our earlier works [2], [3] considered the same scenario as the present paper and showed that [3, Theorem 1]

$$C(\mathcal{E}, c) = \mathcal{E} \log \frac{1}{\mathcal{E}} - \mathcal{E} \log\log \frac{1}{\mathcal{E}} + O(\mathcal{E}), \quad c \in [0, \infty), \tag{5}$$

and that this asymptotic expression can be achieved using pulse-position modulation (PPM).

The above result provides a rather accurate approximation for the capacity $C(\mathcal{E}, c)$, as well as its gap to the capacity of the ideal optical channel equipped with a fully quantum receiver—which lies in the second-order term—where the latter can be found in [9]. Furthermore, extending previous results [4], [5], [10], it suggests PPM as a near-optimal modulation scheme, which greatly simplifies coding compared to on-off keying as suggested in [2], as the latter requires a highly skewed input distribution and is hence difficult to code.

The result (5) may be surprising in that it does not involve the constant c. In other words, in our regime of interest, the dark current influences neither the first- nor the second-order term in capacity. It is clear, however, that the dark current must affect capacity, and it is the aim of the current paper to find out how. Our main result, Theorem 1, shows that the first term in $C(\mathcal{E}, c)$ to be affected by c is the third-order term. It also shows that this term scales like $-\log c$ for large c.

In this paper we also address the coding question: how useful is PPM when c is large? This question has two parts. First, is PPM still near optimal in terms of capacity in this case? Second, does PPM still simplify coding? We answer the first part of the question in the affirmative, again in Theorem 1. However, unlike in [3] where the decoder needs only to record one pulse per PPM frame, here the decoder must make a "soft decision" by recording up to two pulses per frame to achieve the aforementioned $-\log c$ behavior of the third-order term. As for the second part of the question, we cannot fully answer it within this paper, but shall briefly discuss it in Section V.

The rest of this paper is arranged as follows. Section II introduces our notation and formally describes our setting for PPM. Section III presents our main result with some discussions. Section IV outlines the proof of the achievability part of our main result. Finally, Section V concludes the paper with some numerical results and remarks on future directions.

II. Notation and PPM

We usually use a lower-case letter like x to denote a constant, and an upper-case letter like X to denote a random variable.

We use natural logarithms, and measure information in nats.

We use the usual $o(\cdot)$ and $O(\cdot)$ notation to describe behaviors of functions of $\mathcal{E}$ in the limit where $\mathcal{E}$ approaches zero *with other parameters, if any, held fixed*. We emphasize that, in particular, we do *not* use $o(\cdot)$ and $O(\cdot)$ to describe how functions behave with respect to c.

We adopt the convention

$$0 \log 0 = 0. \tag{6}$$

We next formally describe what we mean by PPM. On the transmitter side:

- The channel uses are divided into frames of equal lengths;
- In each frame, there is only one channel input that is positive (the "pulse"), while all the other inputs are zeros;
- The pulses in all frames have the same amplitude.

On the receiver side, we distinguish between two cases, which we call *simple PPM* and *soft-decision PPM*, respectively. In simple PPM, the receiver records *at most one pulse* in each frame; if more than one pulse is detected in a frame, then that frame is recorded as an "erasure", as if no pulse were detected at all. In soft-decision PPM, the receiver records *up to two pulses* in each frame; frames containing no or more than two detected pulses are treated as erasures.

III. Main Result

Let $C_{\text{PE}}(\mathcal{E}, c)$ denote the *photon efficiency* of the channel:

$$C_{\text{PE}}(\mathcal{E}, c) \triangleq \frac{C(\mathcal{E}, c)}{\mathcal{E}}. \tag{7}$$

Clearly, it is equivalent to the capacity $C(\mathcal{E}, c)$ up to normalization with respect to $\mathcal{E}$. Henceforth we may choose to use $C(\mathcal{E}, c)$ or $C_{\text{PE}}(\mathcal{E}, c)$ as convenient. Note that (5) is equivalent to

$$C_{\text{PE}}(\mathcal{E}, c) = \log \frac{1}{\mathcal{E}} - \log\log \frac{1}{\mathcal{E}} + O(1). \tag{8}$$

The main result of this paper is a characterization of the $O(1)$ term in (8):

Theorem 1: The $O(1)$ term in (8) scales like $-\log c$ for large c:

$$\begin{aligned} &\lim_{c\to\infty} \varliminf_{\mathcal{E}\downarrow 0} \frac{C_{\text{PE}}(\mathcal{E}, c) - \log \frac{1}{\mathcal{E}} + \log\log \frac{1}{\mathcal{E}}}{\log c} \\ &= \lim_{c\to\infty} \varlimsup_{\mathcal{E}\downarrow 0} \frac{C_{\text{PE}}(\mathcal{E}, c) - \log \frac{1}{\mathcal{E}} + \log\log \frac{1}{\mathcal{E}}}{\log c} = -1. \end{aligned} \tag{9}$$

Furthermore, the limits in (9) are achievable with soft-decision PPM.

The proof of Theorem 1 has two parts. The achievability part asserts that soft-decision PPM can achieve a rate that satisfies

$$\varliminf_{c\to\infty} \varliminf_{\mathcal{E}\downarrow 0} \frac{C_{\text{PE-PPM(SD)}}(\mathcal{E}, c) - \log \frac{1}{\mathcal{E}} + \log\log \frac{1}{\mathcal{E}}}{\log c} \geq -1 \tag{10}$$

and is sketched in Section IV. The converse part is omitted due to space limitation; it uses the same method as the converse proof in [3] (the *duality bound* [11]) but involves more careful analyses. For the complete proof, we refer to [12].

To better understand our main result, Theorem 1, we make the following remarks.

- Theorem 1 shows that the first term in $C_{\text{PE}}(\mathcal{E}, c)$ that c—or the dark current—affects is the third, constant term. Indeed, though we have not given an exact expression for the constant term, (9) shows that c affects the constant term in such a way that, for large c, the constant term is approximately $-\log c$. Equivalently, the first term in $C(\mathcal{E}, c)$ that is affected by c is the third, $O(\mathcal{E})$ term.
- Theorem 1 suggests the approximation

$$C_{\text{PE}}(\mathcal{E}, c) \approx \log \frac{1}{\mathcal{E}} - \log\log \frac{1}{\mathcal{E}} - \log c. \tag{11}$$

The correction terms, roughly speaking, are either vanishing for small $\mathcal{E}$, or small compared to $\log c$ for large c. Note that if we fix the dark current, i.e., if we fix the product $c\mathcal{E}$, then the first and third terms on the right-hand side of (11) cancel. This is intuitively in agreement with [2, Proposition 2], which states that, for fixed dark current, photon efficiency scales like some constant times $\log\log(1/\mathcal{E})$ and hence *not* like $\log(1/\mathcal{E})$. We note, however, that (9) cannot be derived directly using (8) together with [2, Proposition 2], because we cannot change the order of the limits. In (9) we do not let $\mathcal{E}$ tend to zero and c tend to infinity simultaneously. Instead, we first let $\mathcal{E}$ tend to zero to close down onto the constant term in $C_{\text{PE}}(\mathcal{E}, c)$ with respect to $\mathcal{E}$, and then let c tend to infinity to study the asymptotic behavior of this constant term with respect to c.

- The approximation (11) is good for large c, but diverges as c tends to zero. We hence need a better approximation for the constant term, which behaves like $-\log c$ for large c, but which does not diverge for small c. As both the nonasymptotic bounds and the numerical simulations we show later will suggest, $-\log(1+c)$ is a good approximation:

$$C_{\text{PE}}(\mathcal{E}, c) \approx \log \frac{1}{\mathcal{E}} - \log\log \frac{1}{\mathcal{E}} - \log(1+c). \qquad (12)$$

- In [3] we showed that PPM is optimal up to the second-order term in photon efficiency (equivalently, in capacity). Now we see that, even when the third-order term is taken into account, (soft-decision) PPM is still not far from optimal, in the sense that it achieves the optimal asymptotic behavior of this term with respect to c.
- In Section IV we show that

$$\begin{aligned} C_{\text{PE-PPM}}(\mathcal{E}, c) \geq \log \frac{1}{\mathcal{E}} - \log\log \frac{1}{\mathcal{E}} - c - \frac{3}{2} \\ - \log(1+c) + o(1). \end{aligned} \qquad (13)$$

A careful analysis will confirm that the bound (13) is tight in the regime of interest, in the sense that simple PPM cannot achieve a constant term that is better than linear in c (while being second-order optimal). This is in contrast to soft-decision PPM, which can achieve a constant term that is logarithmic in c. In particular, simple PPM is clearly *not* third-order optimal.[2]

IV. Proof of Achievability (10)

In this section we derive the achievability bound (10) for soft-decision PPM. To this end, we first derive a lower bound on the photon efficiency achievable with *simple* PPM that is slightly tighter than the one in [3], and then base our soft-decision PPM bound on it. Due to space limitation, we only provide the key steps in our calculations, but we try to highlight the difference between simple and soft-decision PPM. For complete derivations of these bounds, see [12].

A. Simple PPM

Consider the following simple PPM scheme:

Scheme 1:

- The channel uses are divided into frames of length b, so each frame contains b input symbols $x_1, \ldots, x_b$ and b corresponding output symbols $y_1, \ldots, y_b$. We set

$$b = \left\lfloor \frac{1}{\mathcal{E} \log \frac{1}{\mathcal{E}}} \right\rfloor. \qquad (14)$$

- Within each length-b frame, there is always one input that equals η, and all the other $(b-1)$ inputs are zeros. Each frame is then fully specified by the position of its unique nonzero symbol, i.e., its pulse position. We consider each frame as a "super input symbol" $\tilde{x}$ that takes value in $\{1, \ldots, b\}$. Here $\tilde{x} = i$ means

$$x_i = \eta \qquad (15)$$

$$x_j = 0, \quad j \neq i. \qquad (16)$$

To meet the average-power constraint (2) with equality, we require

$$\eta = b\mathcal{E}. \qquad (17)$$

- The b output symbols $y_1, \ldots, y_b$ are mapped to one "super output symbol" $\tilde{y}$ that takes value in $\{1, \ldots, b, ?\}$ in the following way: $\tilde{y} = i$, $i \in \{1, \ldots, b\}$, if y_i is the *unique* nonzero term in $\{y_1, \ldots, y_b\}$; and $\tilde{y} = ?$ if there is more than one or no nonzero term in $\{y_1, \ldots, y_b\}$.

We have the following lower bound on the photon efficiency achieved by the above scheme.

Proposition 1: For small enough $\mathcal{E}$, Scheme 1 achieves photon efficiency

$$\begin{aligned} &C_{\text{PE-PPM}}(\mathcal{E}, c) \\ &\geq \left(1 - \frac{\eta}{2}\right) \log b - c\eta \log b - \left(1 + \frac{c\mathcal{E}}{\eta}\right) \log(1+c) \\ &\quad - \left(1 + \frac{c\mathcal{E}}{\eta}\right) \left\{\log(1 - c\eta) + \log\left(1 - \frac{\eta}{2}\right)\right\} - 1 - \frac{c\mathcal{E}}{\eta} \end{aligned} \qquad (18)$$

where b and η are given in (14) and (17), respectively.

It is easy to check that Proposition 1 implies (13).

Proof Sketch of Proposition 1: The transition matrix of the PPM "super channel" is found to be:

$$\tilde{W}(i|i) = e^{-(b-1)c\mathcal{E}} - e^{-\eta - bc\mathcal{E}} \triangleq p_0; \qquad (19)$$

$$\tilde{W}(j|i) = e^{-\eta-(b-1)c\mathcal{E}} - e^{-\eta - bc\mathcal{E}} \triangleq p_1, \quad i \neq j; \qquad (20)$$

$$\tilde{W}(?|i) = 1 - p_0 - (b-1)p_1. \qquad (21)$$

Denote the capacity of this super channel by $\tilde{C}(\mathcal{E}, c, b, \eta)$, then

$$\tilde{C}(\mathcal{E}, c, b, \eta) = \max_{P_{\tilde{X}}} I(\tilde{X}; \tilde{Y}). \qquad (22)$$

Note that the total input power (i.e., expected number of detected signal photons) in each frame equals η. Therefore

$$C_{\text{PE-PPM}}(\mathcal{E}, c) \geq \frac{\tilde{C}(\mathcal{E}, c, b, \eta)}{\eta}. \qquad (23)$$

It can be easily verified that the optimal input distribution for (22) is the uniform distribution, which induces the following marginal distribution on $\tilde{Y}$:

$$P_{\tilde{Y}}(i) = \frac{p_0 + (b-1)p_1}{b}, \quad i \in \{1, \ldots, b\}; \qquad (24)$$

$$P_{\tilde{Y}}(?) = 1 - p_0 - (b-1)p_1. \qquad (25)$$

[2] A scheme between simple and soft-decision PPM is the following. When detecting multiple pulses in a frame, the receiver randomly selects and records one of the positions (possibly together with a "quality" flag). This scheme outperforms simple PPM in photon efficiency, but its third-order term is still linear in c: it scales like $-c/2$ instead of $-c$.

We next explicitly compute $\tilde{C}(\mathcal{E}, c, b, \eta)$ to be

$$\tilde{C}(\mathcal{E}, c, b, \eta) = p_0 \log \frac{bp_0}{p_0 + (b-1)p_1} + (b-1)p_1 \log \frac{bp_1}{p_0 + (b-1)p_1}. \quad (26)$$

We then lower-bound it as

$$\tilde{C}(\mathcal{E}, c, b, \eta) \geq p_0 \log b - p_0 \log\left(1 + \frac{bp_1}{p_0}\right) - p_0. \quad (27)$$

Next note that p_0 can be lower- and upper-bounded as

$$(1 - bc\mathcal{E})\left(\eta - \frac{1}{2}\eta^2\right) \leq p_0 \leq \eta + c\mathcal{E}, \quad (28)$$

while p_1 can be upper-bounded as

$$p_1 \leq c\mathcal{E}. \quad (29)$$

Plugging (28) and (29) into (27) and making some simplifications yield (18). ■

B. Soft-decision PPM

Consider the following soft-decision PPM scheme:

Scheme 2: The transmitter performs the same PPM as in Scheme 1. The receiver maps the b output symbols to a "super symbol" $\hat{y}$ that takes value in $\{1, \ldots, b, ?\} \cup \big\{\{i, j\} \colon i, j \in \{1, \ldots, b\}, i \neq j\big\}$. The mapping rule is as follows. Take $\hat{y} = i$ if y_i is the *unique* nonzero term in $\{y_1, \ldots, y_b\}$; take $\hat{y} = \{i, j\}$ if y_i and y_j are the *only two* nonzero terms in $\{y_1, \ldots, y_b\}$; if there are more than two or no nonzero term in $\{y_1, \ldots, y_b\}$, take $\hat{y} = ?$.

Proposition 2: For small enough $\mathcal{E}$, Scheme 2 achieves photon efficiency

$$\begin{aligned} &C_{\text{PE-PPM(SD)}}(\mathcal{E}, c) \\ &\geq \left(1 - \frac{\eta}{2}\right)\log b - c\eta \log b - \left(1 + \frac{c\mathcal{E}}{\eta}\right)\log(1 + c) \\ &\quad - \left(1 + \frac{c\mathcal{E}}{\eta}\right)\left\{\log(1 - c\eta) + \log\left(1 - \frac{\eta}{2}\right)\right\} - 1 - \frac{c\mathcal{E}}{\eta} \\ &\quad + \left(1 - \frac{\eta}{2}\right)(b-1)\left(c\mathcal{E} - \frac{c^2\mathcal{E}^2}{2}\right)(1 - c\eta)\log\frac{b}{2} \\ &\quad - \frac{c^2}{2}\eta - c(\eta + c\mathcal{E}) \end{aligned} \quad (30)$$

where b and η are given in (14) and (17), respectively.

It is easy to check that Proposition 2 implies the desired achievability bound (10).

Proof Sketch of Proposition 2: We compute the transition matrix of the super channel that results from Scheme 2. We first note

$$\hat{W}(i|i) = p_0, \quad (31)$$

$$\hat{W}(j|i) = p_1, \quad i \neq j, \quad (32)$$

where p_0 and p_1 are given in (19) and (20), respectively. For the remaining elements of the transition matrix we have

$$\hat{W}\big(\{i, j\}\big|i\big) = (1 - e^{-\eta - c\mathcal{E}})(1 - e^{-c\mathcal{E}})\, e^{-(b-2)c\mathcal{E}} \triangleq p_2 \quad (33)$$

$$\hat{W}\big(\{j, k\}\big|i\big) = e^{-\eta - c\mathcal{E}}(1 - e^{-c\mathcal{E}})^2\, e^{-(b-3)c\mathcal{E}} \triangleq p_3 \quad (34)$$

$$\hat{W}(?|i) = 1 - p_0 - (b-1)p_1 - (b-1)p_2 - \frac{(b-1)(b-2)}{2}p_3 \triangleq p_4 \quad (35)$$

for all $\{i, j, k\} \subseteq \{1, \ldots, b\}$. Choosing a uniform $\tilde{X}$ (which is again optimal) yields

$$\begin{aligned} I(\tilde{X}; \hat{Y}) &= p_0 \log \frac{bp_0}{p_0 + (b-1)p_1} \\ &\quad + (b-1)p_1 \log \frac{bp_1}{p_0 + (b-1)p_1} \\ &\quad + (b-1)p_2 \log \frac{bp_2}{2p_2 + (b-2)p_3} \\ &\quad + \frac{(b-1)(b-2)}{2}p_3 \log \frac{bp_3}{2p_2 + (b-2)p_3}. \end{aligned} \quad (36)$$

At this point, note that the first two summands on the right-hand side of (36) constitute $I(\tilde{X}; \tilde{Y})$, which we analyzed in Section IV-A. The last two summands together can be lower-bounded by the expression

$$(b-1)p_2 \log \frac{b}{2} - \frac{b^2}{2}p_3 - bp_2. \quad (37)$$

By upper- and lower-bounding p_2 and p_3 similarly to (28) and (29) we obtain the following lower bound on the additional mutual information that is gained by considering output frames with two detection positions:

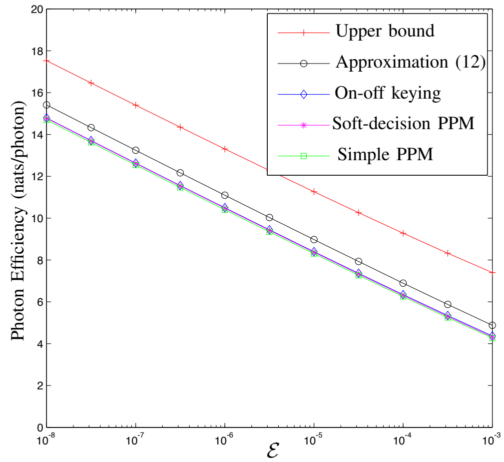
$$\begin{aligned} &I(\tilde{X}; \hat{Y}) - I(\tilde{X}; \tilde{Y}) \\ &\geq (b-1)\left(\eta - \frac{\eta^2}{2}\right)\left(c\mathcal{E} - \frac{c^2\mathcal{E}^2}{2}\right)(1 - bc\mathcal{E})\log\frac{b}{2} \\ &\quad - \frac{b^2}{2}c^2\mathcal{E}^2 - b(\eta + c\mathcal{E})c\mathcal{E}. \end{aligned} \quad (38)$$

Dividing the above by η and adding it to the right-hand side of (18) yield (30). ■

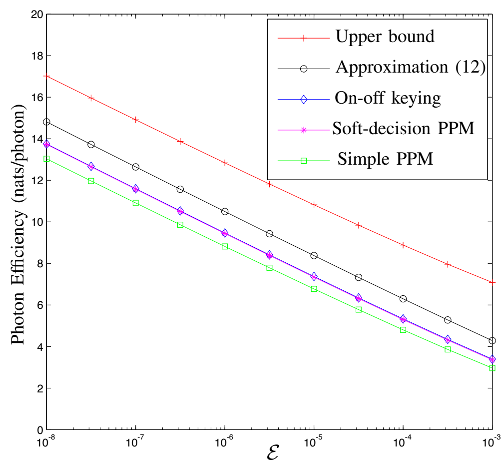
V. Numerical Comparison and Concluding Remarks

We numerically compare the approximation (12) with nonasymptotic upper and lower bounds on photon efficiency. Specifically, the plotted on-off-keying lower bound is obtained by computing the mutual information of the channel with "on" signal equaling $1/\big(\log(1/\mathcal{E})\big)$, and with the receiver ignoring multiple detected photons. The simple-PPM lower bound is computed using (26). The soft-decision-PPM lower bound is computed using (36). Expression for the plotted upper bound can be found in [12]. We plot these bounds for $c = 0.1$, $c = 1$, and $c = 10$ in Figure 1. The figures show the following.

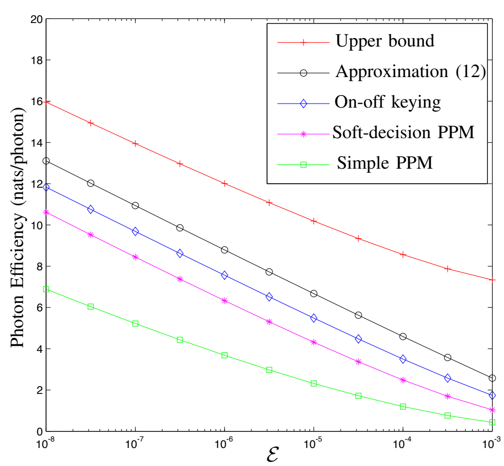
- The approximation (12) is reasonably accurate for small enough $\mathcal{E}$.
- The on-off-keying and the soft-decision-PPM bounds are consistently close to the approximation (12).



(a) Case $c = 0.1$. Note that the lowest three curves almost overlap.



(b) Case $c = 1$. The on-off keying and soft-decision PPM curves overlap.



(c) Case $c = 10$.

Fig. 1. Comparing the approximation (12) to nonasymptotic upper bound, and to lower bounds for simple PPM, soft-decision PPM, and on-off keying for three cases: $c = 0.1$, $c = 1$, and $c = 10$.

- As c increases, the simple-PPM bound deviates significantly from the other lower bounds, and hence also from the actual value of $C_{\text{PE}}(\mathcal{E}, c)$.

The capacity bounds as well as the asymptotic results in this paper show how dark current influences the communication rates and code design in optical channels in the wideband regime. As we have demonstrated, PPM is nearly optimal in this regime. When c is small, the simple PPM super channel has high erasure probability but low "error" (by which we mean the receiver detects a single pulse at a position that is different from the transmitted signal) probability. In this case, Reed-Solomon codes can perform rather close to the theoretical limit. However, when $c > 1$, Reed-Solomon codes can no longer achieve any positive rates on this channel. Nevertheless, we believe that, for $c > 1$, PPM still has its advantages over on-off keying in terms of code design. This is because the optimal input distribution for (both simple and soft-decision) PPM is uniform, whereas the optimal input distribution for on-off keying for this channel is highly skewed. The uniformity of PPM inputs allows the usage of more structured codes, in particular *linear codes*. One possible direction, for instance, is to employ the idea of *multilevel codes* [13], [14] on this channel.

References

[1] J. H. Shapiro, "The quantum theory of optical communications," *IEEE J. Select. Top. Quantum Electron.*, vol. 15, no. 6, pp. 1547–1569, 2009.

[2] A. Lapidoth, J. H. Shapiro, V. Venkatesan, and L. Wang, "The discrete-time Poisson channel at low input powers," *IEEE Trans. Inform. Theory*, vol. 57, pp. 3260–3272, June 2011.

[3] L. Wang and G. W. Wornell, "A refined analysis of the Poisson channel in the high-photon-efficiency regime," in *Proc. Inform. Theory Workshop (ITW)*, (Lausanne, Switzerland), pp. 587–591, Sept. 3–7, 2012.

[4] J. Pierce, "Optical channels: practical limits with photon counting," *IEEE Trans. Commun.*, vol. 26, pp. 1819–1821, Dec. 1978.

[5] J. L. Massey, "Capacity, cutoff rate, and coding for direct-detection optical channel," *IEEE Trans. Commun.*, vol. 29, pp. 1615–1621, Nov. 1981.

[6] S. Shamai (Shitz), "Capacity of a pulse amplitude modulated direct detection photon channel," in *Proc. IEE*, vol. 137, pt. I (Communications, Speech and Vision), pp. 424–430, Dec. 1990.

[7] A. Lapidoth and S. M. Moser, "On the capacity of the discrete-time Poisson channel," *IEEE Trans. Inform. Theory*, vol. 55, pp. 303–322, Jan. 2009.

[8] H. W. Chung, S. Guha, and L. Zheng, "On capacity of optical channels with coherent detection," in *Proc. IEEE Int. Symp. Inform. Theory*, (Saint Petersburg, Russia), July 31–August 5 2011.

[9] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen, "Classical capacity of the lossy bosonic channel: the exact solution," *Phys. Rev. Lett.*, vol. 92, no. 2, p. 027902(4), 2004.

[10] Y. Kochman and G. W. Wornell, "On high-efficiency optical communication and key distribution," in *Proc. Inf. Theory and Appl. Workshop*, (San Diego, CA, USA), Feb. 5–10 2012.

[11] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat fading channels," *IEEE Trans. Inform. Theory*, vol. 49, pp. 2426–2467, Oct. 2003.

[12] L. Wang and G. W. Wornell, "A refined analysis of the Poisson channel in the high-photon-efficiency regime." Subm. to *IEEE Trans. Inform. Theory*, 2013, arXiv:1401.5767.

[13] H. Imai and S. Hirakawa, "A new multilevel coding method using error correcting codes," *IEEE Trans. Inform. Theory*, vol. 23, pp. 371–377, 1977.

[14] U. Wachsmann, R. F. H. Fischer, and J. B. Huber, "Multilevel codes: Theoretical concepts and practical design rules," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1361–1391, July 1999.