The Discrete-Time Poisson Channel at Low Input Powers

Amos Lapidoth, *Fellow, IEEE*, Jeffrey H. Shapiro, *Fellow, IEEE*, Vinodh Venkatesan, *Student Member, IEEE*, and Ligong Wang, *Student Member, IEEE*

***Abstract*—The asymptotic capacity at low input powers of an average-power limited or an average- and peak-power limited discrete-time Poisson channel is considered. For a Poisson channel whose dark current is zero or decays to zero linearly with its average input power $\mathcal{E}$, capacity scales like $\mathcal{E}\log\frac{1}{\mathcal{E}}$ for small $\mathcal{E}$. For a Poisson channel whose dark current is a nonzero constant, capacity scales, to within a constant, like $\mathcal{E}\log\log\frac{1}{\mathcal{E}}$ for small $\mathcal{E}$.**

***Index Terms*—Asymptotic capacity, channel capacity, low signal-to-noise ratio (SNR), optical communication, Poisson channel.**

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*. We denote the Poisson distribution of mean ξ by $\mathcal{P}_\xi(\cdot)$ so

$$\mathcal{P}_\xi(y) = e^{-\xi}\frac{\xi^y}{y!}, \quad y \in \mathbb{Z}_0^+. \tag{1}$$

With this notation the channel law $W(\cdot|\cdot)$ is

$$W(y|x) = \mathcal{P}_{\lambda+x}(y), \quad x \in \mathbb{R}_0^+, y \in \mathbb{Z}_0^+. \tag{2}$$

This channel is often used to model pulse-amplitude modulated optical communication with a direct-detection receiver [1]. Here the input x is proportional to the product of the transmitted light intensity by the pulse duration; the output y models the number of photons arriving at the receiver during the pulse duration; and λ models the average number of extraneous counts that appear in y in addition to those associated with the illumination x.

The *average-power constraint*[1] is

$$\mathsf{E}[X] \leq \mathcal{E} \tag{3}$$

where $\mathcal{E} > 0$ is the maximum allowed average power.

The *peak-power constraint* is that with probability one

$$X \leq \mathcal{A}. \tag{4}$$

We assume throughout that $\mathcal{A} > 0$. In the absence of a peak-power constraint we write $\mathcal{A} = \infty$.

No analytic expression for the capacity of this channel is known. In [1] Shamai showed that the capacity-achieving input distribution is discrete with the number of mass points depending on $\mathcal{E}$ and $\mathcal{A}$. In [2], [3] Lapidoth and Moser derived the asymptotic capacity of the Poisson channel in the regime where both the allowed average power and allowed peak power tend to infinity with their ratio held fixed.

In the present paper we seek the asymptotic capacity of the Poisson channel when the allowed average input power tends to zero with the allowed peak-power—if finite—held fixed. We consider two different cases for the dark current λ. The first is when the dark current tends to zero proportionally to the average power. This corresponds to the wide-band regime where the pulse duration tends to zero.[2] The second case is when the dark current is constant. This corresponds to the regime where the transmitter is weak.

Our lower bounds on channel capacity in the various cases are all based on binary inputs. Our upper bounds are derived using duality (see [4] and references therein). In some cases our lower and upper bounds asymptotically coincide (Proposition 1). An efficient way to compute asymptotic capacities at low average input powers is to compute the capacity per unit cost [5]. However, we shall see that, apart from one case ((11)), the capacity per unit cost is infinite, i.e., the capacity tends to zero more slowly than linearly with the average power.

Manuscript received March 10, 2010; revised August 31, 2010; accepted November 10, 2010. Date of current version May 25, 2011. The material in this paper was presented in part at the 25th IEEE Convention of Electrical and Electronics Engineers in Israel, Eilat, Israel, 3-5 Dec. 2008. This work was conducted while V. Venkatesan was at the Signal and Information Processing Laboratory, ETH Zurich. J. H. Shapiro is supported by the W. M. Keck Foundation Center for Extreme Quantum Information Theory and the ONR Basic Research Challenge Program.

A. Lapidoth is with ETH Zurich, ETF E107, Sternwartstrasse 7, Zürich 8092, Switzerland (e-mail: amos.lapidoth@ethz.ch).

J. H. Shapiro is with the Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: jhs@mit.edu).

V. Venkatesan is with IBM Research, Säumerstrasse 4, Rüschlikon 8803, Switzerland (e-mail: ven@zurich.ibm.com).

L. Wang is with ETH Zurich, ETF E105, Sternwartstrasse 7, Zürich 8092, Switzerland (e-mail: wang@isi.ee.ethz.ch).

Communicated by H. Yamamoto, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2011.2134430

[1]The word "power" here has the meaning "average number of photons transmitted per channel use." If we denote by P the standard "power" in physics, namely, energy per unit time (in watts), then the notion of "power" in this paper is really $\eta P\Delta/\hbar\omega$, where η is the detector's quantum efficiency, Δ is the pulse duration (in sec), and $\hbar\omega$ is the photon energy (in joules) at the operating frequency ω (in rad/sec).

[2]Note that by "wide-band" we mean that the communication bandwidth, i.e., the reciprocal of the pulse duration Δ, is large enough so that $\eta P\Delta/\hbar\omega \ll 1$, but this bandwidth is still much smaller than the optical center frequency ω. Once the bandwidth becomes comparable to the optical center frequency, photon-flux is no longer proportional to input power, and therefore our channel model becomes inadequate.

Among the results in this paper, the special case of zero dark current was derived independently in [6], [7].

The rest of the paper is arranged as follows: in Section II we provide the results with some discussions; in Section III we prove the lower bounds; and in Section IV we prove the upper bounds.

II. Results

Let $C(\lambda, \mathcal{E}, \mathcal{A})$ denote the capacity of the Poisson channel with dark current λ under constraints (3) and (4):

$$C(\lambda, \mathcal{E}, \mathcal{A}) = \sup I(X;Y) \tag{5}$$

where the supremum is over all input distributions satisfying (3) and (4).

When λ is proportional to $\mathcal{E}$, the low-average-power asymptotic capacity of the Poisson channel is given in the following proposition. Note that this also includes the case where the dark current is zero.

Proposition 1 (Dark Current Proportional to $\mathcal{E}$): For any $c \geq 0$ and $\mathcal{A} \in (0, \infty]$,

$$\lim_{\mathcal{E} \downarrow 0} \frac{C(c\mathcal{E}, \mathcal{E}, \mathcal{A})}{\mathcal{E} \log \frac{1}{\mathcal{E}}} = 1. \tag{6}$$

Recall that, for any α, $\beta > 0$, the sum of two independent random variables with the Poisson distributions $\mathcal{P}_\alpha(\cdot)$ and $\mathcal{P}_\beta(\cdot)$ has the Poisson distribution $\mathcal{P}_{\alpha+\beta}(\cdot)$. Thus, we can produce any Poisson channel with nonzero dark current from a Poisson channel with zero dark current by having the receiver add an independent Poisson random variable to the channel output. Since this cannot increase capacity,

$$C(0, \mathcal{E}, \mathcal{A}) \geq C(c\mathcal{E}, \mathcal{E}, \mathcal{A}), \quad c, \mathcal{E}, \mathcal{A} > 0. \tag{7}$$

Consequently, to prove Proposition 1, we only need to show

$$\varliminf_{\mathcal{E} \downarrow 0} \frac{C(c\mathcal{E}, \mathcal{E}, \mathcal{A})}{\mathcal{E} \log \frac{1}{\mathcal{E}}} \geq 1, \quad c > 0, \mathcal{A} \in (0, \infty], \tag{8}$$

$$\varlimsup_{\mathcal{E} \downarrow 0} \frac{C(0, \mathcal{E}, \infty)}{\mathcal{E} \log \frac{1}{\mathcal{E}}} \leq 1. \tag{9}$$

We shall prove (8) in Section III-A and (9) in Section IV-A.

Remarks about Proposition 1:

- If we set $\mathcal{A} = \infty$, then the model considered in Proposition 1 can be used to describe pulse amplitude modulation on a *continuous-time* Poisson channel with constant dark current under an average input-power constraint and in the absence of a peak-power constraint. Proposition 1 shows that, as we let the pulse duration Δ tend to zero, capacity grows like $\frac{\mathcal{E}}{\Delta} \log \frac{1}{\Delta}$, where $\frac{\mathcal{E}}{\Delta}$ is the continuous-time average power[3] which remains constant as $\mathcal{E}$ tends to zero proportionally with Δ.
- Note that (6) does not depend on the peak input power $\mathcal{A}$. In fact, as the proof shows, (6) can be achieved using on-off signaling, where the "on" signal is chosen small but constant.[4] In the continuous-time picture, this choice corresponds to the peak power growing like the constant divided by Δ. As Δ tends to zero, the maximum continuous-time input power thus tends to infinity. (Note that, to achieve unbounded capacity—in our case $\frac{\mathcal{E}}{\Delta} \log \frac{1}{\Delta}$—on the continuous-time Poisson channel, it is necessary to use inputs that tend to infinity since peak-limited continuous-time Poisson channels have bounded capacities [8].)
- It is somewhat surprising that the RHS of (6) does not depend on the dark current. In particular, it does not depend on whether the dark current is zero or not. Intuitively this is because, when $\mathcal{E}$ is small, our "on" signal, which we hold constant, dominates the dark-current floor.
- The bound (9) can also be derived by noting that the capacity of the Poisson channel with zero dark current under only an average-power constraint is upper-bounded by the capacity of the pure-loss bosonic channel, and by using the explicit formula [9]

$$C_{\text{bosonic}}(\mathcal{E}) = (1+\mathcal{E}) \log(1+\mathcal{E}) - \mathcal{E} \log \mathcal{E} \tag{10}$$

of the latter.
- Because the pure-loss bosonic channel with coherent input states and direct detection reduces to a Poisson channel, the formula (6) and the achievability of its left-hand side using binary signaling combine with (10) to show that the asymptotic (quantum-receiver) capacity of the pure-loss bosonic channel is achievable with binary modulation (on-off signaling) and direct detection.
- To see how well capacity is approximated by its asymptotic expression, we compare this expression with nonasymptotic upper and lower bounds in Fig. 1. The upper bound is computed using (53) and (86); the lower bounds are computed using (17), (21) and (28). It can be seen that this approximation is useful for $c = 0.1$, and for $c = 1$ and $\mathcal{E} < 10^{-3}$. For $c = 1$ and $\mathcal{E} > 10^{-3}$ our choice of the input distribution becomes highly suboptimal, which is why the lower bound deviates significantly from the capacity asymptote.

For our second case where the dark current is constant and does not scale with $\mathcal{E}$, the asymptotics depend critically on whether a peak-power constraint is present $(\mathcal{A} < \infty)$ or not $(\mathcal{A} = \infty)$:

Proposition 2 (Constant Nonzero Dark Current): For any $\lambda > 0$,

$$\lim_{\mathcal{E} \downarrow 0} \frac{C(\lambda, \mathcal{E}, \mathcal{A})}{\mathcal{E}} = \left(1 + \frac{\lambda}{\mathcal{A}}\right) \log\left(1 + \frac{\mathcal{A}}{\lambda}\right) - 1, \quad \mathcal{A} < \infty, \tag{11}$$

and

$$\frac{1}{2} \leq \varliminf_{\mathcal{E} \downarrow 0} \frac{C(\lambda, \mathcal{E}, \infty)}{\mathcal{E} \log \log \frac{1}{\mathcal{E}}} \leq \varlimsup_{\mathcal{E} \downarrow 0} \frac{C(\lambda, \mathcal{E}, \infty)}{\mathcal{E} \log \log \frac{1}{\mathcal{E}}} \leq 2. \tag{12}$$

[3]To be precise, $\frac{\mathcal{E}}{\Delta}$ is $\eta P / \hbar\omega$ where P is the "power" in physics. See Footnote 1.

[4]One can also approach asymptotic capacity with nonbinary inputs.

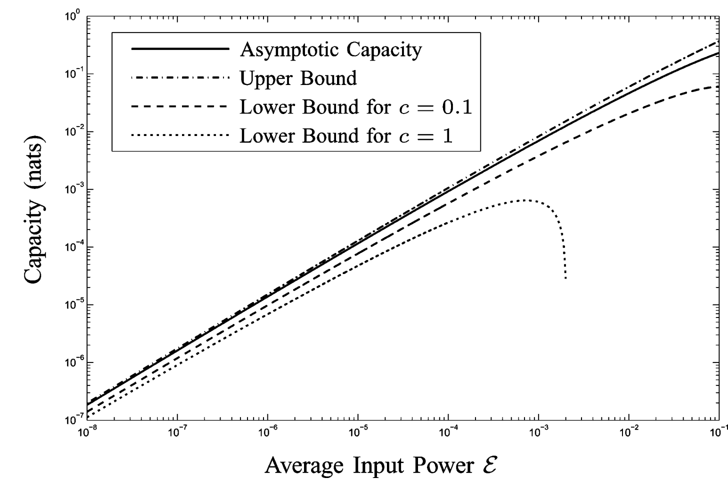

Fig. 1. Comparison of capacity asymptote $\mathcal{E} \log \frac{1}{\mathcal{E}}$ with nonasymptotic upper and lower bounds for dark current proportional to $\mathcal{E}$.

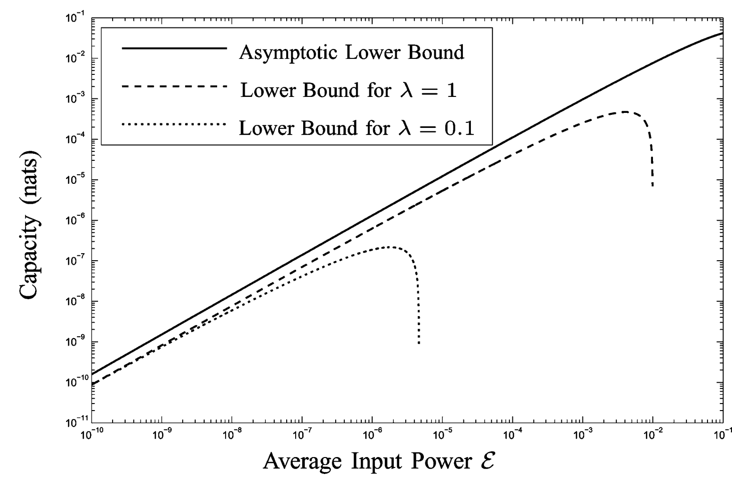

Fig. 2. Comparison of asymptotic lower bound $\frac{1}{2}\mathcal{E} \log\log \frac{1}{\mathcal{E}}$ with nonasymptotic lower bound for constant nonzero dark current.

The proof of (11) is a simple application of the formula for the capacity per unit cost [5, Theorem 2]. We shall prove the lower bound in (12) in Section III-B and the upper bound in Section IV-B.

Remarks about Proposition 2:

- In contrast to Proposition 1, here the capacity asymptote depends heavily on the peak input power $\mathcal{A}$. In particular, it is linear in $\mathcal{E}$ if $\mathcal{A}$ is finite, and it is proportional to $\mathcal{E} \log\log \frac{1}{\mathcal{E}}$ if $\mathcal{A}$ is infinite.
- As the proof shows, both (11) and (12) can be achieved with on-off signaling. In the case of (11), the "on" signal is equal to $\mathcal{A}$; while in the case of (12), the "on" signal tends to infinity as $\mathcal{E}$ tends to zero. These signaling schemes are in the same spirit as the one that achieves (6) in the sense that the "on" signal should be large compared to the dark-current floor.
- We compare the asymptotic and nonasymptotic lower bounds in Fig. 2. The nonasymptotic lower bounds are computed using (44). Interestingly, for most realistic values of $\mathcal{E}$, this nonasymptotic lower bound for $\lambda = 1$ is better than that for $\lambda = 0.1$. This is because, when $\lambda = 0.1$, our choice of the input distribution is good only for extremely small input powers. To get a sense of how good the asymptotic approximation is, we can always lower-bound the capacity when $\lambda = 0.1$ by the lower bound for $\lambda = 1$, which is rather close to the asymptotic lower bound whenever $\mathcal{E} < 10^{-3}$. Our nonasymptotic upper bounds are difficult to compute and are therefore not included in this figure.

III. Lower Bounds

The lower bounds in this section are obtained by choosing binary input distributions and then studying the corresponding mutual informations. We denote by Q^{b} the binary distribution

$$X = \begin{cases} 0, & \text{w.p. } (1-p) \\ \zeta, & \text{w.p. } p \end{cases} \tag{13}$$

where $\zeta > 0$ and $p \in (0, 1)$. If we choose the parameters ζ and p so that constraints (3) and (4) are satisfied, then

$$C(\lambda, \mathcal{E}, \mathcal{A}) \geq I(Q^{\mathrm{b}}, W). \tag{14}$$

A. Dark Current Proportional to $\mathcal{E}$

We next derive Inequality (8). To this end, we write out the mutual information $I(Q^{\mathrm{b}}, W)$ for the input distribution Q^{b} of (13) as

$$\begin{aligned} I(Q^{\mathrm{b}}, W) &= H(Y) - H(Y|X) && (15)\\ &= -\sum_{y=0}^{\infty}((1-p)\mathcal{P}_{\lambda}(y) + p\mathcal{P}_{\lambda+\zeta}(y)) \\ &\quad \cdot \log((1-p)\mathcal{P}_{\lambda}(y) + p\mathcal{P}_{\lambda+\zeta}(y)) \\ &\quad + (1-p)\sum_{y=0}^{\infty}\mathcal{P}_{\lambda}(y)\log\mathcal{P}_{\lambda}(y) \\ &\quad + p\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+\zeta}(y)\log\mathcal{P}_{\lambda+\zeta}(y) && (16)\\ &= I_0(\lambda,\zeta,p) + I_1(\lambda,\zeta,p) && (17) \end{aligned}$$

where in the last equality we defined

$$\begin{aligned} I_0(\lambda,\zeta,p) &\triangleq -\left((1-p)e^{-\lambda} + pe^{-(\lambda+\zeta)}\right) \\ &\quad \cdot \log\left((1-p)e^{-\lambda} + pe^{-(\lambda+\zeta)}\right) \\ &\quad - (1-p)\lambda e^{-\lambda} - p(\lambda+\zeta)e^{-(\lambda+\zeta)} && (18)\\ I_1(\lambda,\zeta,p) &\triangleq -\sum_{y=1}^{\infty}((1-p)\mathcal{P}_{\lambda}(y) + p\mathcal{P}_{\lambda+\zeta}(y)) \\ &\quad \cdot \log((1-p)\mathcal{P}_{\lambda}(y) + p\mathcal{P}_{\lambda+\zeta}(y)) \\ &\quad + (1-p)\sum_{y=1}^{\infty}\mathcal{P}_{\lambda}(y)\log\mathcal{P}_{\lambda}(y) \\ &\quad + p\sum_{y=1}^{\infty}\mathcal{P}_{\lambda+\zeta}(y)\log\mathcal{P}_{\lambda+\zeta}(y). && (19) \end{aligned}$$

Note that in the above decomposition we took out the terms corresponding to $y = 0$ in all three summations to form $I_0(\lambda,\zeta,p)$ and collected the remaining terms in $I_1(\lambda,\zeta,p)$.

We lower-bound $I_0(\lambda,\zeta,p)$ as

$$\begin{aligned} I_0(\lambda,\zeta,p) &\geq 0 - (1-p)\lambda e^{-\lambda} - p(\lambda+\zeta)e^{-(\lambda+\zeta)} && (20)\\ &\geq -\lambda - p(\lambda+\zeta). && (21) \end{aligned}$$

We lower-bound $I_1(\lambda,\zeta,p)$ as

$$\begin{aligned} &I_1(\lambda,\zeta,p) \\ &= -\sum_{y=1}^{\infty}\big((1-p)\mathcal{P}_{\lambda}(y) + p\mathcal{P}_{\lambda+\zeta}(y)\big) \\ &\quad \cdot \log\left((1-p)\frac{\mathcal{P}_{\lambda}(y)}{\mathcal{P}_{\lambda+\zeta}(y)} + p\right) \\ &\quad + (1-p)\sum_{y=1}^{\infty}\mathcal{P}_{\lambda}(y)\log\frac{\mathcal{P}_{\lambda}(y)}{\mathcal{P}_{\lambda+\zeta}(y)} && (22)\\ &= -\sum_{y=1}^{\infty}\big((1-p)\mathcal{P}_{\lambda}(y) + p\mathcal{P}_{\lambda+\zeta}(y)\big) \\ &\quad \cdot \left(\log p + \log\left(1 + \frac{1-p}{p}\frac{\mathcal{P}_{\lambda}(y)}{\mathcal{P}_{\lambda+\zeta}(y)}\right)\right) \\ &\quad + (1-p)\sum_{y=1}^{\infty}\mathcal{P}_{\lambda}(y)\log\frac{e^{-\lambda}\frac{\lambda^y}{y!}}{e^{-(\lambda+\zeta)}\frac{(\lambda+\zeta)^y}{y!}} && (23)\\ &= -\sum_{y=1}^{\infty}\big((1-p)\mathcal{P}_{\lambda}(y) + p\mathcal{P}_{\lambda+\zeta}(y)\big) \\ &\quad \cdot \left(\log p + \underbrace{\log\left(1 + \frac{1-p}{p}\frac{\mathcal{P}_{\lambda}(y)}{\mathcal{P}_{\lambda+\zeta}(y)}\right)}_{\leq \frac{1-p}{p}\frac{\mathcal{P}_{\lambda}(y)}{\mathcal{P}_{\lambda+\zeta}(y)}}\right) \\ &\quad + (1-p)\zeta\underbrace{\sum_{y=1}^{\infty}\mathcal{P}_{\lambda}(y)}_{=1-e^{-\lambda}} \\ &\quad + (1-p)\log\frac{\lambda}{\lambda+\zeta}\underbrace{\sum_{y=1}^{\infty}\mathcal{P}_{\lambda}(y)y}_{=\lambda} && (24)\\ &\geq -\underbrace{\sum_{y=1}^{\infty}\big((1-p)\mathcal{P}_{\lambda}(y) + p\mathcal{P}_{\lambda+\zeta}(y)\big)}_{=(1-p)(1-e^{-\lambda})+p(1-e^{-(\lambda+\zeta)})}\log p \\ &\quad - \sum_{y=1}^{\infty}\big((1-p)\mathcal{P}_{\lambda}(y) + p\mathcal{P}_{\lambda+\zeta}(y)\big)\frac{1-p}{p}\frac{\mathcal{P}_{\lambda}(y)}{\mathcal{P}_{\lambda+\zeta}(y)} \\ &\quad + (1-p)(1-e^{-\lambda})\zeta - (1-p)\lambda\log\left(1+\frac{\zeta}{\lambda}\right) && (25)\\ &= \big((1-p)(1-e^{-\lambda}) + p(1-e^{-(\lambda+\zeta)})\big)\log\frac{1}{p} \\ &\quad - \frac{(1-p)^2}{p}\sum_{y=1}^{\infty}\underbrace{\frac{(\mathcal{P}_{\lambda}(y))^2}{\mathcal{P}_{\lambda+\zeta}(y)}}_{=e^{\frac{\zeta^2}{\lambda+\zeta}}\mathcal{P}_{\frac{\lambda^2}{\lambda+\zeta}}(y)} - (1-p)\underbrace{\sum_{y=1}^{\infty}\mathcal{P}_{\lambda}(y)}_{=1-e^{-\lambda}} \\ &\quad + (1-p)(1-e^{-\lambda})\zeta - (1-p)\lambda\log\left(1+\frac{\zeta}{\lambda}\right) && (26)\\ &= \big((1-p)(1-e^{-\lambda}) + p(1-\underbrace{e^{-(\lambda+\zeta)}}_{\leq e^{-\zeta}})\big)\log\frac{1}{p} \\ &\quad - \underbrace{\frac{(1-p)^2}{p}}_{\leq\frac{1}{p}}\underbrace{e^{\frac{\zeta^2}{\lambda+\zeta}}}_{\leq e^{\zeta}}\underbrace{\left(1-e^{-\frac{\lambda^2}{\lambda+\zeta}}\right)}_{\leq\frac{\lambda^2}{\lambda+\zeta}\leq\frac{\lambda^2}{\zeta}} - \underbrace{(1-p)}_{\leq 1}\underbrace{(1-e^{-\lambda})}_{\leq\lambda} \\ &\quad - \underbrace{(1-p)}_{\leq 1}\underbrace{(1-e^{-\lambda})}_{\leq\lambda}\zeta - (1-p)\lambda\log\left(1+\frac{\zeta}{\lambda}\right) && (27)\\ &\geq (1-p)(1-e^{-\lambda})\log\frac{1}{p} + p(1-e^{-\zeta})\log\frac{1}{p} \\ &\quad - \frac{1}{p}\frac{\lambda^2}{\zeta}e^{\zeta} - \lambda - \lambda\zeta - (1-p)\lambda\log\left(1+\frac{\zeta}{\lambda}\right). && (28) \end{aligned}$$

Choose any $\zeta \in (0, \mathcal{A}]$ and, for small enough $\mathcal{E}$, let $p = \mathcal{E}/\zeta$. Then the distribution (13) satisfies both constraints (3) and (4).

Let $\lambda = c\mathcal{E}$. Using (21) we can bound the asymptotic behavior of $I_0(\lambda,\zeta,p)$ as

$$\lim_{\mathcal{E}\downarrow 0} \frac{I_0\left(c\mathcal{E},\zeta,\frac{\mathcal{E}}{\zeta}\right)}{\mathcal{E}\log\frac{1}{\mathcal{E}}} \geq -\lim_{\mathcal{E}\downarrow 0}\frac{c\mathcal{E}}{\mathcal{E}\log\frac{1}{\mathcal{E}}} - \lim_{\mathcal{E}\downarrow 0}\frac{\frac{\mathcal{E}}{\zeta}(c\mathcal{E}+\zeta)}{\mathcal{E}\log\frac{1}{\mathcal{E}}} = 0. \tag{29}$$

Similarly, using (28) we can bound the asymptotic behavior of $I_1(\lambda,\zeta,p)$ as

$$\begin{aligned}
&\lim_{\mathcal{E}\downarrow 0} \frac{I_1\left(c\mathcal{E},\zeta,\frac{\mathcal{E}}{\zeta}\right)}{\mathcal{E}\log\frac{1}{\mathcal{E}}} \\
&\quad\geq \lim_{\mathcal{E}\downarrow 0}\frac{\left(1-\frac{\mathcal{E}}{\zeta}\right)\left(1-e^{c\mathcal{E}}\right)\log\frac{\zeta}{\mathcal{E}}}{\mathcal{E}\log\frac{1}{\mathcal{E}}} + \lim_{\mathcal{E}\downarrow 0}\frac{\frac{\mathcal{E}}{\zeta}(1-e^{-\zeta})\log\frac{\zeta}{\mathcal{E}}}{\mathcal{E}\log\frac{1}{\mathcal{E}}} \\
&\qquad - \lim_{\mathcal{E}\downarrow 0}\frac{\frac{\zeta}{\mathcal{E}}\frac{c^2\mathcal{E}^2}{\zeta}e^{\zeta}}{\mathcal{E}\log\frac{1}{\mathcal{E}}} - \lim_{\mathcal{E}\downarrow 0}\frac{c\mathcal{E}}{\mathcal{E}\log\frac{1}{\mathcal{E}}} - \lim_{\mathcal{E}\downarrow 0}\frac{c\mathcal{E}\zeta}{\mathcal{E}\log\frac{1}{\mathcal{E}}} \\
&\qquad - \lim_{\mathcal{E}\downarrow 0}\frac{\left(1-\frac{\mathcal{E}}{\zeta}\right)c\mathcal{E}\log\left(1+\frac{\zeta}{c\mathcal{E}}\right)}{\mathcal{E}\log\frac{1}{\mathcal{E}}} && (30)\\
&\quad = c + \frac{1-e^{-\zeta}}{\zeta} - 0 - 0 - 0 - c && (31)\\
&\quad = \frac{1-e^{-\zeta}}{\zeta}. && (32)
\end{aligned}$$

Combining (14), (17), (29), and (32) we obtain

$$\lim_{\mathcal{E}\downarrow 0}\frac{C(c\mathcal{E},\mathcal{E},\mathcal{A})}{\mathcal{E}\log\frac{1}{\mathcal{E}}} \geq \frac{1-e^{-\zeta}}{\zeta}, \quad \text{for all } \zeta \in (0,\mathcal{A}]. \tag{33}$$

We can make the right-hand side (RHS) of (33) arbitrarily close to 1 by choosing arbitrarily small ζ. Thus we obtain (8).

B. Constant Nonzero Dark Current

We next prove the lower bound in (12). To this end, we lower-bound the mutual information $I(Q^{\mathrm{b}},W)$ for the input distribution (13) as follows:

$$\begin{aligned}
&I(Q^{\mathrm{b}},W) \\
&\quad = H(Y) - H(Y|X) && (34)\\
&\quad = -\sum_{y=0}^{\infty}\left((1-p)\mathcal{P}_{\lambda}(y) + p\mathcal{P}_{\lambda+\zeta}(y)\right) \\
&\qquad\qquad \cdot \log\left((1-p)\mathcal{P}_{\lambda}(y) + p\mathcal{P}_{\lambda+\zeta}(y)\right) \\
&\qquad + (1-p)\sum_{y=0}^{\infty}\mathcal{P}_{\lambda}(y)\log\mathcal{P}_{\lambda}(y) \\
&\qquad + p\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+\zeta}(y)\log\mathcal{P}_{\lambda+\zeta}(y) && (35)\\
&\quad = -p\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+\zeta}(y)\log\left((1-p)\frac{\mathcal{P}_{\lambda}(y)}{\mathcal{P}_{\lambda+\zeta}(y)} + p\right) \\
&\qquad - (1-p)\sum_{y=0}^{\infty}\mathcal{P}_{\lambda}(y)\log\left((1-p) + p\frac{\mathcal{P}_{\lambda+\zeta}(y)}{\mathcal{P}_{\lambda}(y)}\right) && (36)\\
&\quad = -p\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+\zeta}(y) \\
&\qquad\qquad \cdot\left(\log\frac{\mathcal{P}_{\lambda}(y)}{\mathcal{P}_{\lambda+\zeta}(y)} + \log\left((1-p)+p\frac{\mathcal{P}_{\lambda+\zeta}(y)}{\mathcal{P}_{\lambda}(y)}\right)\right) \\
&\qquad - (1-p)\sum_{y=0}^{\infty}\mathcal{P}_{\lambda}(y)\log\left((1-p)+p\frac{\mathcal{P}_{\lambda+\zeta}(y)}{\mathcal{P}_{\lambda}(y)}\right) && (37)\\
&\quad = p\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+\zeta}(y)\log\frac{\mathcal{P}_{\lambda+\zeta}(y)}{\mathcal{P}_{\lambda}(y)} - \sum_{y=0}^{\infty}\Bigg(\underbrace{(1-p)\mathcal{P}_{\lambda}(y)}_{\geq 0} \\
&\qquad + \underbrace{p\mathcal{P}_{\lambda+\zeta}(y)}_{\geq 0}\Bigg)\underbrace{\log\left((1-p)+p\frac{\mathcal{P}_{\lambda+\zeta}(y)}{\mathcal{P}_{\lambda}(y)}\right)}_{\leq \log\left(1+p\frac{\mathcal{P}_{\lambda+\zeta}(y)}{\mathcal{P}_{\lambda}(y)}\right)\leq p\frac{\mathcal{P}_{\lambda+\zeta}(y)}{\mathcal{P}_{\lambda}(y)}} && (38)\\
&\quad \geq p\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+\zeta}(y)\log\frac{\mathcal{P}_{\lambda+\zeta}(y)}{\mathcal{P}_{\lambda}(y)} \\
&\qquad - \sum_{y=0}^{\infty}\Big((1-p)\mathcal{P}_{\lambda}(y) + p\mathcal{P}_{\lambda+\zeta}(y)\Big)p\frac{\mathcal{P}_{\lambda+\zeta}(y)}{\mathcal{P}_{\lambda}(y)} && (39)\\
&\quad = p\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+\zeta}(y)\log\left(\frac{e^{-(\zeta+\lambda)}\frac{(\zeta+\lambda)^y}{y!}}{e^{-\lambda}\frac{\lambda^y}{y!}}\right) \\
&\qquad - (1-p)p\sum_{y=0}^{\infty}\mathcal{P}_{\lambda}(y)\frac{\mathcal{P}_{\lambda+\zeta}(y)}{\mathcal{P}_{\lambda}(y)} \\
&\qquad - p^2\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+\zeta}(y)\frac{\mathcal{P}_{\lambda+\zeta}(y)}{\mathcal{P}_{\lambda}(y)} && (40)\\
&\quad = p\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+\zeta}(y)\log\left(e^{-\zeta}\left(1+\frac{\zeta}{\lambda}\right)^y\right) \\
&\qquad - (1-p)p\underbrace{\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+\zeta}(y)}_{=1} \\
&\qquad - p^2\sum_{y=0}^{\infty}\frac{\left(e^{-(\zeta+\lambda)}\frac{(\zeta+\lambda)^y}{y!}\right)^2}{e^{-\lambda}\frac{\lambda^y}{y!}} && (41)\\
&\quad = p\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+\zeta}(y)\left(-\zeta + y\log\left(1+\frac{\zeta}{\lambda}\right)\right) - (1-p)p \\
&\qquad - p^2\underbrace{\left(\sum_{y=0}^{\infty}e^{-(\lambda+2\zeta)}\frac{\left(\lambda+2\zeta+\frac{\zeta^2}{\lambda}\right)^y}{y!}e^{-\frac{\zeta^2}{\lambda}}\right)}_{=\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+2\zeta+\frac{\zeta^2}{\lambda}}(y)=1}e^{\frac{\zeta^2}{\lambda}} && (42)\\
&\quad = -p\zeta\underbrace{\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+\zeta}(y)}_{=1} + p\underbrace{\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+\zeta}(y)y}_{=(\zeta+\lambda)}\log\left(1+\frac{\zeta}{\lambda}\right) \\
&\qquad - p + p^2 - p^2e^{\frac{\zeta^2}{\lambda}} && (43)\\
&\quad = p(\zeta+\lambda)\log\left(1+\frac{\zeta}{\lambda}\right) - p\zeta - p - p^2\left(e^{\frac{\zeta^2}{\lambda}} - 1\right). && (44)
\end{aligned}$$

For small enough $\mathcal{E}$, we choose

$$\zeta = \sqrt{\lambda \log \frac{1}{\mathcal{E}}} \tag{45}$$

and

$$p = \frac{\mathcal{E}}{\zeta} = \frac{\mathcal{E}}{\sqrt{\lambda \log \frac{1}{\mathcal{E}}}}. \tag{46}$$

From (14) and (44) we then obtain

$$\begin{aligned}
&\varliminf_{\mathcal{E}\downarrow 0} \frac{C(\lambda,\mathcal{E},\infty)}{\mathcal{E}\log\log\frac{1}{\mathcal{E}}} \\
&\geq \varliminf_{\mathcal{E}\downarrow 0} \frac{\frac{\mathcal{E}}{\sqrt{\lambda\log\frac{1}{\mathcal{E}}}}\left(\sqrt{\lambda\log\frac{1}{\mathcal{E}}}+\lambda\right)\log\left(1+\frac{\sqrt{\lambda\log\frac{1}{\mathcal{E}}}}{\lambda}\right)}{\mathcal{E}\log\log\frac{1}{\mathcal{E}}} \\
&\quad + \varliminf_{\mathcal{E}\downarrow 0} \frac{-\mathcal{E}-\frac{\mathcal{E}}{\sqrt{\lambda\log\frac{1}{\mathcal{E}}}}-\left(\frac{\mathcal{E}}{\sqrt{\lambda\log\frac{1}{\mathcal{E}}}}\right)^2\left(e^{\frac{\lambda\log\frac{1}{\mathcal{E}}}{\lambda}}-1\right)}{\mathcal{E}\log\log\frac{1}{\mathcal{E}}} \qquad (47)\\
&= \frac{1}{2}+0 \qquad (48)\\
&= \frac{1}{2}. \qquad (49)
\end{aligned}$$

This establishes the lower bound in (12).

IV. Upper Bounds

In this section we prove the upper bounds on the asymptotic capacities of the Poisson channel. We shall use the duality bound [4] which states that, for any distribution $R(\cdot)$ on the output, the channel capacity satisfies

$$C \leq \sup \mathsf{E}\left[D\left(W(\cdot|X)\|R(\cdot)\right)\right], \tag{50}$$

where the supremum is taken over all allowed input distributions.

A. Zero Dark Current

We next prove (9). To this end, as in [3], we shall introduce in Section IV-A1 the Poisson channel with continuous output. This channel is equivalent to our channel but its output alphabet is not the nonnegative integers but the nonnegative reals. We shall then prove a lemma in Section IV-A2 before finally proving (9) in Section IV-A3.

1) Poisson Channel With Continuous Output: We introduce the Poisson channel with continuous output whose dark current is equal to zero. Its input x is the same as that of the original Poisson channel, and its output $\tilde{Y} \in \mathbb{R}_0^+$ is

$$\tilde{Y} \triangleq Y + U \tag{51}$$

where Y is the output of the original Poisson channel with zero dark current, and the random variable U is independent of (X, Y) and uniformly distributed on the interval $[0, 1)$. Then $\tilde{Y}$ is a continuous random variable whose conditional density $\tilde{w}(\tilde{y}|x)$ given $X = x$ is

$$\tilde{w}(\tilde{y}|x) = W(\lfloor \tilde{y} \rfloor | x) = \mathcal{P}_x(\lfloor \tilde{y} \rfloor) \tag{52}$$

where $\lfloor a \rfloor$ denotes the largest integer not exceeding a, and the second equality follows because $\lambda = 0$. Note that $W(\cdot|x)$ is a probability mass function on $\mathbb{Z}^+$ whereas $\tilde{w}(\cdot|x)$ is a density on $\mathbb{R}^+$.

Denoting the capacity of the channel $\tilde{w}(\cdot|\cdot)$ under constraints (3) and (4) by $\tilde{C}(0, \mathcal{E}, \mathcal{A})$,

$$C(0, \mathcal{E}, \mathcal{A}) = \tilde{C}(0, \mathcal{E}, \mathcal{A}), \quad \mathcal{E} > 0, \mathcal{A} \in (0, \infty] \tag{53}$$

because Y can be computed from $\tilde{Y}$ [3, Lemma 17].

2) A Lemma: The following lemma lower-bounds the differential entropy $h(\tilde{Y}|X = x)$ of $\tilde{Y}$ conditional on $X = x$.

Lemma 1: Let $\tilde{Y}$ be defined as in (51) and $\tilde{w}$ be given by (52), then

$$h(\tilde{Y}|X = x) \geq \frac{1 - e^{-x}}{2} \log 2\pi + \frac{1}{2} \sum_{y=1}^{\infty} \mathcal{P}_x(y) \log y. \tag{54}$$

Proof: By (51) we have as in [3, Lemma 17]

$$h(\tilde{Y}|X = x) = H(Y|X = x). \tag{55}$$

The RHS of (55) can be bounded as

$$\begin{aligned}
&H(Y|X=x) \\
&= -\sum_{y=0}^{\infty} W(y|x)\log W(y|x) \qquad (56)\\
&= -\sum_{y=0}^{\infty} \mathcal{P}_x(y)\log\left(e^{-x}\frac{x^y}{y!}\right) \qquad (57)\\
&= x\underbrace{\sum_{y=0}^{\infty}\mathcal{P}_x(y)}_{=1} - \log x\underbrace{\sum_{y=0}^{\infty}\mathcal{P}_x(y)\,y}_{=x} + \sum_{y=0}^{\infty}\mathcal{P}_x(y)\log(y!) \qquad (58)\\
&= x - x\log x + \sum_{y=1}^{\infty}\mathcal{P}_x(y)\log(y!) \qquad (59)\\
&\geq x - x\log x + \sum_{y=1}^{\infty}\mathcal{P}_x(y)\log\left(\sqrt{2\pi y}\left(\frac{y}{e}\right)^y\right) \qquad (60)\\
&= x - x\log x + \frac{1}{2}\log 2\pi \underbrace{\sum_{y=1}^{\infty}\mathcal{P}_x(y)}_{=1-\mathcal{P}_x(0)=1-e^{-x}} \\
&\quad + \frac{1}{2}\sum_{y=1}^{\infty}\mathcal{P}_x(y)\log y + \sum_{y=0}^{\infty}\mathcal{P}_x(y)y\log y - \underbrace{\sum_{y=0}^{\infty}\mathcal{P}_x(y)y}_{=x} \qquad (61)
\end{aligned}$$

$$= -x\log x + \frac{1-e^{-x}}{2}\log 2\pi + \frac{1}{2}\sum_{y=1}^{\infty}\mathcal{P}_x(y)\log y + \sum_{y=0}^{\infty}\mathcal{P}_x(y)\underbrace{y\log y}_{\geq x\log x+(1+\log x)(y-x)-x} \tag{62}$$

$$\geq -x\log x + \frac{1-e^{-x}}{2}\log 2\pi + \frac{1}{2}\sum_{y=1}^{\infty}\mathcal{P}_x(y)\log y + \sum_{y=0}^{\infty}\mathcal{P}_x(y)\left(x\log x + (1+\log x)(y-x)\right) \tag{63}$$

$$= -x\log x + \frac{1-e^{-x}}{2}\log 2\pi + \frac{1}{2}\sum_{y=1}^{\infty}\mathcal{P}_x(y)\log y + x\log x\underbrace{\sum_{y=0}^{\infty}\mathcal{P}_x(y)}_{=1} + (1+\log x)\underbrace{\sum_{y=0}^{\infty}\mathcal{P}_x(y)(y-x)}_{=0} \tag{64}$$

$$= \frac{1-e^{-x}}{2}\log 2\pi + \frac{1}{2}\sum_{y=1}^{\infty}\mathcal{P}_x(y)\log y. \tag{65}$$

Here, (60) follows by Stirling's Bound [10]

$$n! \geq \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n+1}} \geq \sqrt{2\pi n}\left(\frac{n}{e}\right)^n, \quad n\in\mathbb{Z}_0^+ \tag{66}$$

and (63) follows by bounding $y\log y$ by its Taylor expansion at $y=x$. Combining (55) and (65) proves the lemma. ∎

3) Proof of (9): According to (53), to prove (9), we need to prove

$$\overline{\lim_{\mathcal{E}\downarrow 0}}\frac{\tilde{C}(0,\mathcal{E},\infty)}{\mathcal{E}\log\frac{1}{\mathcal{E}}} \leq 1. \tag{67}$$

To prove (67) using the duality bound (50), we choose the distribution $\tilde{R}(\cdot)$ on $\tilde{Y}$ to be of density

$$f_{\tilde{R}}(\tilde{y}) = \begin{cases} 1-p, & 0\leq\tilde{y}<1 \\ p\cdot\frac{\tilde{y}^{\nu-1}e^{-\frac{\tilde{y}}{\beta}}}{\beta^{\nu}\Gamma\left(\nu,\frac{1}{\beta}\right)}, & \tilde{y}\geq 1 \end{cases} \tag{68}$$

where $\beta>0$ is arbitrary (e.g., $\beta=1$), whereas $\nu\in(0,1]$ and $p\in(0,1)$ will be specified later, and $\Gamma(\cdot,\cdot)$ denotes the Incomplete Gamma Function

$$\Gamma(a,\xi) = \int_{\xi}^{\infty} t^{a-1}e^{-t}\,\mathrm{d}t, \qquad a,\xi\geq 0. \tag{69}$$

We next apply the duality bound to upper-bound the capacity of the channel $\tilde{w}(\cdot|\cdot)$ using the above output distribution:

$$\tilde{C}(0,\lambda,\infty) \leq \sup_{\mathsf{E}[X]\leq\mathcal{E}} D\left(\tilde{w}(\cdot|X)\|f_{\tilde{R}}\right). \tag{70}$$

We bound $D\left(\tilde{w}(\cdot|x)\|f_{\tilde{R}}(\cdot)\right)$ as follows:

$$D\left(\tilde{w}(\cdot|x)\|f_{\tilde{R}}(\cdot)\right) = \int_0^{\infty}\tilde{w}(\tilde{y}|x)\log\frac{\tilde{w}(\tilde{y}|x)}{f_{\tilde{R}}(\tilde{y})}\,\mathrm{d}\tilde{y} \tag{71}$$

$$= -h(\tilde{Y}|X=x) + \int_0^1\tilde{w}(\tilde{y}|x)\log\frac{1}{f_{\tilde{R}}(\tilde{y})}\,\mathrm{d}\tilde{y} + \int_1^{\infty}\tilde{w}(\tilde{y}|x)\log\frac{1}{f_{\tilde{R}}(\tilde{y})}\,\mathrm{d}\tilde{y} \tag{72}$$

$$= -h(\tilde{Y}|X=x) + \int_0^1\tilde{w}(\tilde{y}|x)\log\frac{1}{1-p}\,\mathrm{d}\tilde{y} + \int_1^{\infty}\tilde{w}(\tilde{y}|x)\log\frac{\beta^{\nu}\Gamma\left(\nu,\frac{1}{\beta}\right)}{p\tilde{y}^{\nu-1}e^{-\frac{\tilde{y}}{\beta}}}\,\mathrm{d}\tilde{y} \tag{73}$$

$$= -h(\tilde{Y}|X=x) + \log\frac{1}{1-p}\underbrace{\int_0^1\tilde{w}(\tilde{y}|x)\,\mathrm{d}\tilde{y}}_{=e^{-x}} + \log\frac{\beta^{\nu}\Gamma\left(\nu,\frac{1}{\beta}\right)}{p}\underbrace{\int_1^{\infty}\tilde{w}(\tilde{y}|x)\,\mathrm{d}\tilde{y}}_{=1-e^{-x}} + (1-\nu)\int_1^{\infty}\tilde{w}(\tilde{y}|x)\log\tilde{y}\,\mathrm{d}\tilde{y} + \frac{1}{\beta}\underbrace{\int_1^{\infty}\tilde{w}(\tilde{y}|x)\tilde{y}\,\mathrm{d}\tilde{y}}_{=\mathsf{E}[\tilde{Y}|X=x]-\int_0^1\tilde{w}(\tilde{y}|x)\tilde{y}\,\mathrm{d}\tilde{y}} \tag{74}$$

$$= -h(\tilde{Y}|X=x) + e^{-x}\log\frac{1}{1-p} + (1-e^{-x})\log\frac{\beta^{\nu}\Gamma\left(\nu,\frac{1}{\beta}\right)}{p} + (1-\nu)\int_1^{\infty}\tilde{w}(\tilde{y}|x)\log\tilde{y}\,\mathrm{d}\tilde{y} + \frac{1}{\beta}\left(\underbrace{\mathsf{E}\left[\tilde{Y}|X=x\right]}_{=x+\frac{1}{2}} - \int_0^1\underbrace{\tilde{w}(\tilde{y}|x)}_{=e^{-x}}\tilde{y}\,\mathrm{d}\tilde{y}\right) \tag{75}$$

$$= -h(\tilde{Y}|X=x) + e^{-x}\log\frac{1}{1-p} + (1-e^{-x})\log\frac{\beta^{\nu}\Gamma\left(\nu,\frac{1}{\beta}\right)}{p} + \underbrace{(1-\nu)}_{\geq 0}\underbrace{\int_1^{\infty}\tilde{w}(\tilde{y}|x)\log\tilde{y}\,\mathrm{d}\tilde{y}}_{\leq\sum_{y=1}^{\infty}W(y|x)\log(y+1)} + \frac{1}{\beta}\left(x+\frac{1}{2}-\frac{1}{2}e^{-x}\right) \tag{76}$$

$$\begin{aligned}
&\leq -h(\tilde{Y}|X=x)+e^{-x}\log\frac{1}{1-p}\\
&\quad+(1-e^{-x})\log\frac{\beta^{\nu}\Gamma\left(\nu,\frac{1}{\beta}\right)}{p}\\
&\quad+(1-\nu)\sum_{y=1}^{\infty}\underbrace{W(y|x)}_{=\mathcal{P}_x(y)}\log(y+1)\\
&\quad+\frac{x}{\beta}+\frac{(1-e^{-x})}{2\beta} &(77)\\
&\leq -\frac{(1-e^{-x})}{2}\log 2\pi-\frac{1}{2}\sum_{y=1}^{\infty}\mathcal{P}_x(y)\log y\\
&\quad+e^{-x}\log\frac{1}{1-p}+(1-e^{-x})\log\frac{\beta^{\nu}\Gamma\left(\nu,\frac{1}{\beta}\right)}{p}\\
&\quad+(1-\nu)\sum_{y=1}^{\infty}\mathcal{P}_x(y)\log(y+1)\\
&\quad+\frac{x}{\beta}+\frac{(1-e^{-x})}{2\beta} &(78)
\end{aligned}$$

where the last inequality follows from Lemma 1.

Substituting $\nu=\frac{1}{2}$ in (78) yields

$$\begin{aligned}
&D\left(\tilde{w}(\cdot|x)||f_{\tilde{R}}(\cdot)\right)\\
&\leq -\frac{(1-e^{-x})}{2}\log 2\pi-\frac{1}{2}\sum_{y=1}^{\infty}\mathcal{P}_x(y)\log y\\
&\quad+e^{-x}\log\frac{1}{1-p}+(1-e^{-x})\log\frac{\beta^{\frac{1}{2}}\Gamma\left(\frac{1}{2},\frac{1}{\beta}\right)}{p}\\
&\quad+\frac{1}{2}\sum_{y=1}^{\infty}\mathcal{P}_x(y)\log(y+1)\\
&\quad+\frac{x}{\beta}+\frac{(1-e^{-x})}{2\beta} &(79)\\
&= -\frac{(1-e^{-x})}{2}\log 2\pi+\frac{1}{2}\sum_{y=1}^{\infty}\mathcal{P}_x(y)\underbrace{\log\left(1+\frac{1}{y}\right)}_{\leq\log 2}\\
&\quad+e^{-x}\log\frac{1}{1-p}+(1-e^{-x})\log\frac{\beta^{\frac{1}{2}}\Gamma\left(\frac{1}{2},\frac{1}{\beta}\right)}{p}\\
&\quad+\frac{x}{\beta}+\frac{(1-e^{-x})}{2\beta} &(80)\\
&\leq -\frac{(1-e^{-x})}{2}\log 2\pi+\frac{1}{2}\underbrace{\sum_{y=1}^{\infty}\mathcal{P}_x(y)}_{=(1-e^{-x})}\log 2\\
&\quad+e^{-x}\log\frac{1}{1-p}+(1-e^{-x})\log\frac{\beta^{\frac{1}{2}}\Gamma\left(\frac{1}{2},\frac{1}{\beta}\right)}{p}\\
&\quad+\frac{x}{\beta}+\frac{(1-e^{-x})}{2\beta} &(81)\\
&= -\frac{(1-e^{-x})}{2}\log\pi+e^{-x}\log\frac{1}{1-p}\\
&\quad+(1-e^{-x})\log\left(\beta^{\frac{1}{2}}\Gamma\left(\frac{1}{2},\frac{1}{\beta}\right)\right)+(1-e^{-x})\log\frac{1}{p}\\
&\quad+\frac{x}{\beta}+\frac{(1-e^{-x})}{2\beta} &(82)\\
&= \underbrace{(1-e^{-x})}_{\leq x}\log\frac{1}{p}+\underbrace{e^{-x}}_{\leq 1}\log\frac{1}{1-p}+\frac{x}{\beta}\\
&\quad+\underbrace{(1-e^{-x})}_{\leq x}\left(\frac{1}{2}\log\beta+\log\frac{\Gamma\left(\frac{1}{2},\frac{1}{\beta}\right)}{\sqrt{\pi}}+\frac{1}{2\beta}\right) &(83)\\
&\leq x\log\frac{1}{p}+\log\frac{1}{1-p}+\frac{x}{\beta}\\
&\quad+x\max\left\{0,\left(\frac{1}{2}\log\beta+\log\frac{\Gamma\left(\frac{1}{2},\frac{1}{\beta}\right)}{\sqrt{\pi}}+\frac{1}{2\beta}\right)\right\}. &(84)
\end{aligned}$$

From (84) and (70), we get

$$\begin{aligned}
&\tilde{C}(0,\mathcal{E},\infty)\\
&\leq \sup_{\mathsf{E}[X]\leq\mathcal{E}}\mathsf{E}\Bigg[X\log\frac{1}{p}+\log\frac{1}{1-p}+\frac{X}{\beta}\\
&\quad+X\max\left\{0,\left(\frac{1}{2}\log\beta+\log\frac{\Gamma\left(\frac{1}{2},\frac{1}{\beta}\right)}{\sqrt{\pi}}+\frac{1}{2\beta}\right)\right\}\Bigg] &(85)\\
&\leq \mathcal{E}\log\frac{1}{p}+\log\frac{1}{1-p}+\frac{\mathcal{E}}{\beta}\\
&\quad+\mathcal{E}\max\left\{0,\left(\frac{1}{2}\log\beta+\log\frac{\Gamma\left(\frac{1}{2},\frac{1}{\beta}\right)}{\sqrt{\pi}}+\frac{1}{2\beta}\right)\right\}. &(86)
\end{aligned}$$

Note that (86) holds for all $p\in(0,1)$ and $\beta>0$. Choosing $p=\frac{\mathcal{E}}{1+\mathcal{E}}$ in (86) and letting $\mathcal{E}$ tend to zero yields (67) and hence concludes the proof of (9).

B. Constant Nonzero Dark Current

In this section we shall prove the upper bound in (12), namely

$$\varlimsup_{\mathcal{E}\downarrow 0}\frac{C(\lambda,\mathcal{E},\infty)}{\mathcal{E}\log\log\frac{1}{\mathcal{E}}}\leq 2. \qquad (87)$$

To this end, we shall prove two lemmas in Section IV-B1, then derive a general upper bound on $C(\lambda,\mathcal{E},\infty)$ in Section IV-B2, and finally prove (87) in Section IV-B3.

1) Lemmas: We next present two lemmas. The first lemma shows that the tail of a Poisson distribution of mean ξ behaves like $\left(\frac{\xi}{n}\right)^n$ for large n.

Lemma 2: If Y is a mean-ξ Poisson random variable, then, for any $n\in\mathbb{Z}^+$,

$$L_{\xi}(n)\leq\Pr[Y\geq n]\leq U_{\xi}(n), \qquad (88)$$

where

$$L_\xi(n) \triangleq \exp\left(n - \xi + n\log\xi - n\log n - \frac{1}{12n} - \frac{1}{2}\log(2\pi n)\right) \tag{89}$$

$$U_\xi(n) \triangleq \begin{cases} \exp(n - \xi + n\log\xi - n\log n), & n > \xi \\ 1, & n \leq \xi. \end{cases} \tag{90}$$

Proof: To prove the lower bound, we observe that for every $n \in \mathbb{Z}^+$

$$\Pr[Y \geq n] \geq \Pr[Y = n] = e^{-\xi}\frac{\xi^n}{n!}. \tag{91}$$

Using Stirling's Bound

$$n! \leq \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n}} \tag{92}$$

we obtain from (91) that

$$\Pr[Y \geq n] \geq e^{-\xi}\xi^n\left(\sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n}}\right)^{-1} = L_\xi(n). \tag{93}$$

This establishes the lower bound in (88).

To prove the upper bound, we recall that the moment generating function of the Poisson distribution of mean ξ is

$$\mathsf{E}[e^{\theta Y}] = e^{\xi(e^\theta - 1)}, \quad \theta \in \mathbb{R}. \tag{94}$$

Consequently, by the Chernoff Bound,

$$\Pr[Y \geq n] \leq \frac{\mathsf{E}[e^{\theta Y}]}{e^{\theta n}} = \frac{e^{\xi(e^\theta - 1)}}{e^{\theta n}}, \quad \theta > 0. \tag{95}$$

When $n > \xi$, letting $\theta = \log\left(\frac{n}{\xi}\right)$ in (95) yields the upper bound in (88)

$$\Pr[Y \geq n] \leq \frac{e^{n-\xi}}{e^{n\log\frac{n}{\xi}}} = U_\xi(n), \quad n > \xi. \tag{96}$$

When $n \leq \xi$, the upper bound in (88) is trivial. ∎

The second lemma is a simple property of convex functions.

Lemma 3: For any convex function $f\colon [a,b] \to \mathbb{R}$

$$\frac{f(b) - f(a)}{b - a} \geq \frac{f(c) - f(a)}{c - a}, \quad c \in (a,b). \tag{97}$$

Proof: Since $c \in (a,b)$, we can write $c = \alpha a + (1-\alpha)b$ for some $\alpha \in (0,1)$. The convexity of $f(\cdot)$ implies

$$f(c) \leq \alpha f(a) + (1-\alpha)f(b). \tag{98}$$

Thus,

$$\begin{aligned} \frac{f(c) - f(a)}{c - a} &\leq \frac{\alpha f(a) + (1-\alpha)f(b) - f(a)}{(\alpha a + (1-\alpha)b) - a} && (99) \\ &= \frac{(1-\alpha)(f(b) - f(a))}{(1-\alpha)(b-a)} && (100) \\ &= \frac{f(b) - f(a)}{b - a}. && (101) \end{aligned}$$

∎

2) An Upper Bound on $C(\lambda, \mathcal{E}, \infty)$: We shall next apply (50) to upper-bound $C(\lambda, \mathcal{E}, \infty)$. To this end, we choose $R(\cdot)$ to be

$$R(y) = \begin{cases} \mathcal{P}_\lambda(y), & y \in \{0, 1, \ldots, \eta - 1\} \\ \delta(1-p)p^{y-\eta}, & y \in \{\eta, \eta+1, \ldots\} \end{cases} \tag{102}$$

where $\eta \in \mathbb{Z}^+$ and $p \in (0,1)$ are constants that will be specified in Section IV-B-III, and δ is the normalizing factor

$$\delta \triangleq \Pr[Y \geq \eta | X = 0] = \sum_{y=\eta}^{\infty} \mathcal{P}_\lambda(y). \tag{103}$$

In the following calculations we shall assume that η is large compared to λ.

To upper-bound the capacity using (50), we write $D(W(\cdot|x) \| R(\cdot))$ as

$$D\left(W(\cdot|x)\|R(\cdot)\right) = \underbrace{\sum_{y=0}^{\eta-1} W(y|x)\log\frac{W(y|x)}{R(y)}}_{\triangleq D_1(x)} + \underbrace{\sum_{y=\eta}^{\infty} W(y|x)\log\frac{W(y|x)}{R(y)}}_{\triangleq D_2(x)} \tag{104}$$

and study $D_1(x)$ and $D_2(x)$ separately. Substituting (102) and the channel law (2) in $D_1(x)$ yields

$$\begin{aligned} D_1(x) &= \sum_{y=0}^{\eta-1} \mathcal{P}_{\lambda+x}(y)\log\frac{e^{-(\lambda+x)}\frac{(\lambda+x)^y}{y!}}{e^{-\lambda}\frac{\lambda^y}{y!}} \\ &= -x\sum_{y=0}^{\eta-1}\mathcal{P}_{\lambda+x}(y) + \log\left(1 + \frac{x}{\lambda}\right)\sum_{y=0}^{\eta-1}\mathcal{P}_{\lambda+x}(y)y. \end{aligned} \tag{105}$$

The second term $D_2(x)$ can be upper-bounded as

$$D_2(x) = \sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y)\log\frac{e^{-(\lambda+x)}\frac{(\lambda+x)^y}{y!}}{\delta(1-p)p^{y-\eta}} \tag{106}$$

$$= \left(-(\lambda+x) + \log\frac{1}{\delta} + \log\frac{1}{1-p} - \eta\log\frac{1}{p}\right)\sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y) + \log\frac{\lambda+x}{p}\sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y)y - \sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y)\log(y!) \quad (107)$$

$$\le \left(-(\lambda+x) + \log\frac{1}{\delta} + \log\frac{1}{1-p} - \eta\log\frac{1}{p}\right)\sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y) + \log\frac{\lambda+x}{p}\sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y)y - \sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y)(y\log y - y) \quad (108)$$

where the inequality follows from Stirling's Bound

$$n! \ge \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n+1}} \ge \left(\frac{n}{e}\right)^n, \quad n\in\mathbb{Z}^+. \quad (109)$$

Substituting (105) and (108) in (104) we obtain

$$\begin{aligned} &D\big(W(\cdot|x)\|R(\cdot)\big) \\ &\le -x\sum_{y=0}^{\eta-1}\mathcal{P}_{\lambda+x}(y) + \log\left(1+\frac{x}{\lambda}\right)\sum_{y=0}^{\eta-1}\mathcal{P}_{\lambda+x}(y)y \\ &\quad + \left(-(\lambda+x) + \log\frac{1}{\delta} + \log\frac{1}{1-p} - \eta\log\frac{1}{p}\right)\sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y) \\ &\quad + \log\frac{\lambda+x}{p}\sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y)y - \sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y)(y\log y - y) \quad (110)\\ &= \underbrace{-x\left(\sum_{y=0}^{\eta-1}\mathcal{P}_{\lambda+x}(y) + \sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y)\right)}_{\le 0} \\ &\quad + \left(\underbrace{-\lambda}_{\le 0} + \log\frac{1}{\delta} + \log\frac{1}{1-p} \underbrace{-\eta\log\frac{1}{p}}_{\le 0}\right)\sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y) \\ &\quad + \log(\lambda+x)\underbrace{\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+x}(y)y}_{=\lambda+x} - \log\lambda\underbrace{\sum_{y=0}^{\eta-1}\mathcal{P}_{\lambda+x}(y)y}_{=(\lambda+x)-\sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y)y} \\ &\quad + \left(1+\log\frac{1}{p}\right)\sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y)y - \sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y)y\log y \quad (111)\end{aligned}$$

$$\begin{aligned} &\le \underbrace{\left(\log\frac{1}{\delta} + \log\frac{1}{1-p}\right)\Pr[Y\ge\eta|X=x]}_{\triangleq D_3(x)} \\ &\quad + \underbrace{\left(1+\log\frac{1}{p}+\log\lambda\right)\sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y)y}_{\triangleq D_4(x)} \\ &\quad + \underbrace{(\lambda+x)\log\left(1+\frac{x}{\lambda}\right) - \sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y)y\log y}_{\triangleq D_5(x)}. \quad (112)\end{aligned}$$

It follows from (50) and (112) that $C(\lambda,\mathcal{E},\infty)$ is upper-bounded by

$$\begin{aligned} C(\lambda,\mathcal{E},\infty) &\le \sup_{\mathsf{E}[X]\le\mathcal{E}} \mathsf{E}[D_3(X)+D_4(X)+D_5(X)] \quad (113)\\ &\le \sup_{\mathsf{E}[X]\le\mathcal{E}} \mathsf{E}[D_3(X)] + \sup_{\mathsf{E}[X]\le\mathcal{E}} \mathsf{E}[D_4(X)] + \sup_{\mathsf{E}[X]\le\mathcal{E}} \mathsf{E}[D_5(X)]. \quad (114)\end{aligned}$$

To find an upper bound on the capacity, we shall next upper-bound the three terms on the RHS of (114) separately.

We first consider $\sup_{\mathsf{E}[X]\le\mathcal{E}} \mathsf{E}[D_3(X)]$. By Lemma 2,

$$\mathsf{E}[D_3(X)] \le \left(\log\frac{1}{\delta} + \log\frac{1}{1-p}\right)\mathsf{E}\left[U_{\lambda+X}(\eta)\right]. \quad (115)$$

Further, when $\mathsf{E}[X]\le\mathcal{E}$, we can bound $\mathsf{E}\left[U_{\lambda+X}(\eta)\right]$ as

$$\begin{aligned} &\mathsf{E}[U_{\lambda+X}(\eta)] \\ &= U_\lambda(\eta) + \mathsf{E}\left[X\cdot\frac{U_{\lambda+X}(\eta)-U_\lambda(\eta)}{X}\right] \quad (116)\\ &\le U_\lambda(\eta) + \mathsf{E}[X]\cdot\sup_{x\in\mathbb{R}^+}\left(\frac{U_{\lambda+x}(\eta)-U_\lambda(\eta)}{x}\right) \quad (117)\\ &\le U_\lambda(\eta) + \mathcal{E}\cdot\sup_{x\in\mathbb{R}^+}\left(\frac{U_{\lambda+x}(\eta)-U_\lambda(\eta)}{x}\right) \quad (118)\end{aligned}$$

where in the first step we adopt the convention $\frac{0}{0}=0$. To upper-bound the supremum in (118), we first observe that $U_{\lambda+x}(\eta)$ is convex in x for $x\in[0,\eta-\sqrt{\eta}-\lambda]$. Then, by Lemma 3, for $x<\eta-\sqrt{\eta}-\lambda$ we have

$$\begin{aligned} &\frac{U_{\lambda+x}(\eta)-U_\lambda(\eta)}{x} \\ &\le \frac{U_{\eta-\sqrt{\eta}}(\eta)-U_\lambda(\eta)}{\eta-\sqrt{\eta}-\lambda} \quad (119)\\ &\le \frac{1}{\eta-\sqrt{\eta}-\lambda}, \quad x<\eta-\sqrt{\eta}-\lambda. \quad (120)\end{aligned}$$

On the other hand, when $x \geq \eta - \sqrt{\eta} - \lambda$ we have

$$\frac{U_{\lambda+x}(\eta) - U_\lambda(\eta)}{x} \leq \frac{1}{x} \leq \frac{1}{\eta - \sqrt{\eta} - \lambda}, \quad x \geq \eta - \sqrt{\eta} - \lambda. \quad (121)$$

Thus we obtain

$$\sup_{x \in \mathbb{R}^+} \left(\frac{U_{\lambda+x}(\eta) - U_\lambda(\eta)}{x} \right) \leq \frac{1}{\eta - \sqrt{\eta} - \lambda}. \quad (122)$$

Combining (118) and (122) yields that for η larger than some $\eta(\lambda)$,

$$\begin{aligned} \sup_{\mathsf{E}[X] \leq \mathcal{E}} \mathsf{E}[U_{\lambda+X}(\eta)] &\leq U_\lambda(\eta) + \frac{\mathcal{E}}{\eta - \sqrt{\eta} - \lambda} && (123) \\ &= \exp(\eta - \lambda + \eta \log \lambda - \eta \log \eta) + \frac{\mathcal{E}}{\eta - \sqrt{\eta} - \lambda} && (124) \\ &\leq \exp(\eta + \eta \log \lambda - \eta \log \eta) + \frac{\mathcal{E}}{\eta - \sqrt{\eta} - \lambda}. && (125) \end{aligned}$$

We next use Lemma 2 to bound δ as

$$\delta = \Pr[Y \geq \eta | X = 0] \geq L_\lambda(\eta). \quad (126)$$

Therefore,

$$\begin{aligned} \log \frac{1}{\delta} &\leq \log \frac{1}{L_\lambda(\eta)} \\ &\leq \eta \log \eta + \frac{1}{12\eta} + \frac{1}{2} \log(2\pi\eta) + \lambda - \eta \log \lambda \quad (127) \end{aligned}$$

where the second inequality is obtained by omitting nonpositive terms. Using (115), (125) and (127) we obtain that for η larger than some $\eta(\lambda)$,

$$\begin{aligned} \sup_{\mathsf{E}[X] \leq \mathcal{E}} \mathsf{E}[D_3(X)] &\leq \left(\eta \log \eta + \frac{1}{12\eta} + \frac{1}{2} \log(2\pi\eta) + \lambda - \eta \log \lambda + \log \frac{1}{1-p} \right) \\ &\quad \cdot \left(\exp\left(\eta + \eta \log \lambda - \eta \log \eta + \frac{\mathcal{E}}{\eta - \sqrt{\eta} - \lambda} \right) \right). \quad (128) \end{aligned}$$

We next derive an upper bound on $\sup_{\mathsf{E}[X] \leq \mathcal{E}} \mathsf{E}[D_4(X)]$. To this end, we observe

$$\begin{aligned} \sum_{y=\eta}^{\infty} \mathcal{P}_{\lambda+x}(y) y &= \sum_{y=\eta}^{\infty} e^{-(\lambda+x)} \frac{(\lambda+x)^y}{y!} \cdot y && (129) \\ &= (\lambda+x) \sum_{y=\eta}^{\infty} e^{-(\lambda+x)} \frac{(\lambda+x)^{y-1}}{(y-1)!} && (130) \\ &= (\lambda+x) \sum_{y=\eta-1}^{\infty} \mathcal{P}_{\lambda+x}(y) && (131) \\ &= (\lambda+x) \Pr[Y \geq \eta - 1 | X = x] && (132) \\ &\leq (\lambda+x) U_{\lambda+x}(\eta - 1) && (133) \end{aligned}$$

where the inequality follows by Lemma 2. Using the above inequality we obtain

$$\mathsf{E}[D_4(X)] \leq \max\left\{0, \left(1 + \log\frac{1}{p} + \log \lambda\right) \cdot \mathsf{E}[(\lambda + X) U_{\lambda+X}(\eta - 1)]\right\}. \quad (134)$$

Similarly to (118), when $\mathsf{E}[X] \leq \mathcal{E}$ we can bound the expectation on the RHS of (134) as

$$\begin{aligned} \mathsf{E}[(\lambda + X) U_{\lambda+X}(\eta - 1)] &\leq \lambda U_\lambda(\eta - 1) \\ &\quad + \mathcal{E} \cdot \sup_{x \in \mathbb{R}^+} \left(\frac{(\lambda+x) U_{\lambda+x}(\eta - 1) - \lambda U_\lambda(\eta - 1)}{x} \right). \quad (135) \end{aligned}$$

To bound the supremum on the RHS of (135), we observe that $(\lambda + x) U_{\lambda+x}(\eta - 1)$ is convex in x on $[0, \eta - \sqrt{\eta} - \lambda]$. Thus, by Lemma 3 we have, for $x < \eta - \sqrt{\eta} - \lambda$,

$$\begin{aligned} \frac{(\lambda+x) U_{\lambda+x}(\eta - 1) - \lambda U_\lambda(\eta - 1)}{x} &\leq \frac{(\eta - \sqrt{\eta}) U_{\eta - \sqrt{\eta}}(\eta - 1) - \lambda U_\lambda(\eta - 1)}{\eta - \sqrt{\eta} - \lambda} && (136) \\ &\leq \frac{\lambda}{\eta - \sqrt{\eta} - \lambda} + 1, \quad x < \eta - \sqrt{\eta} - \lambda. && (137) \end{aligned}$$

When $x \geq \eta - \sqrt{\eta} - \lambda$, we have

$$\begin{aligned} \frac{(\lambda+x) U_{\lambda+x}(\eta - 1) - \lambda U_\lambda(\eta - 1)}{x} &\leq \frac{\lambda + x}{x} && (138) \\ &\leq \frac{\lambda}{\eta - \sqrt{\eta} - \lambda} + 1. && (139) \end{aligned}$$

Thus the supremum on the RHS of (135) can be bounded as

$$\sup_{x \in \mathbb{R}^+} \left(\frac{(\lambda+x) U_{\lambda+x}(\eta - 1) - \lambda U_\lambda(\eta - 1)}{x} \right) \leq \frac{\lambda}{\eta - \sqrt{\eta} - \lambda} + 1. \quad (140)$$

Combining (134), (135), (140) and the definition of $U_\lambda(\cdot)$ we have that for η larger than some $\eta(\lambda)$,

$$\begin{aligned} \sup_{\mathsf{E}[X] \leq \mathcal{E}} \mathsf{E}[D_4(X)] &\leq \max\left\{0, \left(1 + \log\frac{1}{p} + \log \lambda\right) \cdot \left(\mathcal{E} + \frac{\lambda \mathcal{E}}{\eta - \sqrt{\eta} - \lambda} \right.\right. \\ &\quad \left.\left. + \lambda \cdot e^{\eta - 1 - \lambda + (\eta - 1) \log \lambda - (\eta - 1) \log(\eta - 1)} \right)\right\}. \quad (141) \end{aligned}$$

We now consider $\sup_{\mathsf{E}[X]\leq\mathcal{E}} \mathsf{E}[D_5(X)]$. For $x \leq \eta - \lambda$, we simply bound $D_5(x)$ by

$$D_5(x) \leq (\lambda + x)\log\left(1 + \frac{x}{\lambda}\right), \quad x \leq \eta - \lambda. \tag{142}$$

When $x > \eta - \lambda$, we use the inequality

$$y \log y \geq (\lambda + x)\log(\lambda + x) + (1 + \log(\lambda + x))(y - (\lambda + x)) \tag{143}$$

to obtain

$$\begin{aligned}
\sum_{y=\eta}^{\infty} \mathcal{P}_{\lambda+x}(y) y \log y
&\geq (\lambda + x)\log(\lambda + x)\sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y) \\
&\quad + \big(1 + \log(\lambda + x)\big)\sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y)\big(y - (\lambda + x)\big) &(144)\\
&\geq (\lambda + x)\log(\lambda + x)\sum_{y=\eta}^{\infty}\mathcal{P}_{\lambda+x}(y) \\
&\quad + \big(1 + \log(\lambda + x)\big)\underbrace{\sum_{y=0}^{\infty}\mathcal{P}_{\lambda+x}(y)\big(y - (\lambda + x)\big)}_{=0} &(145)\\
&= (\lambda + x)\log(\lambda + x)\Pr[Y \geq \eta | X = x] &(146)\\
&= (\lambda + x)\log(\lambda + x)\left(1 - \Pr[Y \leq \eta - 1 | X = x]\right), &(147)
\end{aligned}$$

where the second inequality is obtained by adding nonpositive terms. Thus we may bound $D_5(x)$, when $x > \eta - \lambda$, by

$$D_5(x) \leq (\lambda + x)\log\left(1 + \frac{x}{\lambda}\right)\Pr[Y \leq \eta - 1 | X = x] + (\lambda + x)\max\left\{\log\frac{1}{\lambda}, 0\right\}. \tag{148}$$

By combining (142) and (148) and by adding nonnegative terms, we can upper-bound $\mathsf{E}[D_5(X)]$, for all input distributions satisfying $\mathsf{E}[X] \leq \mathcal{E}$, by

$$\begin{aligned}
\mathsf{E}[D_5(X)]
&\leq \mathsf{E}[X]\max\left\{\log\frac{1}{\lambda}, 0\right\} \\
&\quad + \lambda \max\left\{\log\frac{1}{\lambda}, 0\right\}\cdot\Pr[X \geq \eta - \lambda] \\
&\quad + \mathsf{E}[D_6(X)] &(149)\\
&\leq \mathcal{E}\max\left\{\log\frac{1}{\lambda}, 0\right\} + \frac{\lambda\mathcal{E}}{\eta - \lambda}\max\left\{\log\frac{1}{\lambda}, 0\right\} \\
&\quad + \mathsf{E}[D_6(X)] &(150)
\end{aligned}$$

where in the second step we applied Markov's inequality to $\Pr[X \geq \eta - \lambda]$, and where

$$D_6(x) = (\lambda + x)\log\left(1 + \frac{x}{\lambda}\right) \cdot \begin{cases} 1, & x \leq \eta - \lambda \\ \Pr[Y \leq \eta - 1 | X = x], & x > \eta - \lambda. \end{cases} \tag{151}$$

To upper-bound $\Pr[Y \leq \eta - 1 | X = x]$ when $x > \eta - \lambda$, we use the Chernoff Bound and (94) to write

$$\begin{aligned}
\Pr[Y \leq \eta - 1 | X = x] &\leq \frac{\mathsf{E}[e^{\theta Y}]}{e^{\theta(\eta - 1)}} \\
&= \frac{e^{(\lambda + x)(e^{\theta} - 1)}}{e^{\theta(\eta - 1)}}, \quad \theta < 0. &(152)
\end{aligned}$$

Letting $\theta = \log\left(\frac{\eta - 1}{\lambda + x}\right)$ in the above inequality yields

$$\begin{aligned}
&\Pr[Y \leq \eta - 1 | = x] \\
&\quad \leq e^{\eta - 1 - (\lambda + x) + (\eta - 1)\log(\lambda + x)(-(\eta - 1)\log(\eta - 1)}, \\
&\qquad\qquad x > \eta - \lambda. &(153)
\end{aligned}$$

Substituting (153) into the definition of $D_6(x)$ we obtain

$$D_6(x) \leq D_7(x), \quad x \in \mathbb{R}_0^+ \tag{154}$$

where

$$D_7(x) \triangleq (\lambda + x)\log\left(1 + \frac{x}{\lambda}\right) \cdot \begin{cases} 1, & x \leq \eta - \lambda \\ e^{\eta - 1 - (\lambda + x) + (\eta - 1)\log(\lambda + x) - (\eta - 1)\log(\eta - 1)}, & x > \eta - \lambda. \end{cases} \tag{155}$$

Thus, when $\mathsf{E}[X] \leq \mathcal{E}$,

$$\begin{aligned}
\mathsf{E}[D_6(X)] &\leq \mathsf{E}[D_7(X)] &(156)\\
&= \mathsf{E}\left[X \cdot \frac{D_7(X)}{X}\right] &(157)\\
&\leq \mathcal{E} \cdot \sup_{x \in \mathbb{R}^+}\left(\frac{D_7(x)}{x}\right) &(158)
\end{aligned}$$

where we adopt the convention $\frac{0}{0} = 0$. It can be checked by computing the derivative of $\frac{D_7(x)}{x}$ with respect to x that, for η larger than some $\eta(\lambda)$, $\frac{D_7(x)}{x}$ is monotonically decreasing in x for $x > \eta - \lambda$. Using this observation, Lemma 3, and the fact that $(\lambda + x)\log\left(1 + \frac{x}{\lambda}\right)$ is convex in x on $\mathbb{R}^+$, we have that the supremum on the RHS of (158) is achieved when $x = \eta - \lambda$. Therefore, when $\mathsf{E}[X] \leq \mathcal{E}$,

$$\mathsf{E}[D_6(X)] \leq \mathcal{E} \cdot \frac{\eta\log\frac{\eta}{\lambda}}{\eta - \lambda}. \tag{159}$$

Combining (150) and (159) we obtain

$$\sup_{\mathsf{E}[X]\leq\mathcal{E}} \mathsf{E}[D_5(X)] \leq \mathcal{E}\cdot\left(1 + \frac{\lambda}{\eta - \lambda}\right)\cdot\max\left\{\log\frac{1}{\lambda}, 0\right\} + \mathcal{E}\cdot\frac{\eta\log\left(\frac{\eta}{\lambda}\right)}{\eta - \lambda}. \tag{160}$$

Thus, for any $\lambda > 0$, $\mathcal{E} > 0$, $p \in (0, 1)$, and η larger than some $\eta(\lambda)$, we can combine (114), (128), (141) and (160) to obtain an upper bound on $C(\lambda, \mathcal{E}, \infty)$.

3) Proof of (87): For small enough $\mathcal{E}$, we choose

$$\eta = \left\lfloor \log \frac{1}{\mathcal{E}} \right\rfloor, \tag{161}$$

and let $p \in (0,1)$ have any fixed value that does not depend on $\mathcal{E}$. In this case, it follows by (114) that

$$\begin{aligned}\varlimsup_{\mathcal{E}\downarrow 0} \frac{C(\lambda,\mathcal{E},\infty)}{\mathcal{E}\log\log\frac{1}{\mathcal{E}}} \leq \varlimsup_{\mathcal{E}\downarrow 0} \frac{\sup_{\mathsf{E}[X]\leq\mathcal{E}} \mathsf{E}[D_3(X)]}{\mathcal{E}\log\log\frac{1}{\mathcal{E}}} \\ + \varlimsup_{\mathcal{E}\downarrow 0} \frac{\sup_{\mathsf{E}[X]\leq\mathcal{E}} \mathsf{E}[D_4(X)]}{\mathcal{E}\log\log\frac{1}{\mathcal{E}}} \\ + \varlimsup_{\mathcal{E}\downarrow 0} \frac{\sup_{\mathsf{E}[X]\leq\mathcal{E}} \mathsf{E}[D_5(X)]}{\mathcal{E}\log\log\frac{1}{\mathcal{E}}}.\end{aligned} \tag{162}$$

Substituting (161) into (128) yields

$$\varlimsup_{\mathcal{E}\downarrow 0} \frac{\sup_{\mathsf{E}[X]\leq\mathcal{E}} \mathsf{E}[D_3(X)]}{\mathcal{E}\log\log\frac{1}{\mathcal{E}}} \leq 1. \tag{163}$$

Substituting (161) into (141) yields

$$\varlimsup_{\mathcal{E}\downarrow 0} \frac{\sup_{\mathsf{E}[X]\leq\mathcal{E}} \mathsf{E}[D_4(X)]}{\mathcal{E}\log\log\frac{1}{\mathcal{E}}} \leq 0. \tag{164}$$

Finally, substituting (161) into (160) yields

$$\varlimsup_{\mathcal{E}\downarrow 0} \frac{\sup_{\mathsf{E}[X]\leq\mathcal{E}} \mathsf{E}[D_5(X)]}{\mathcal{E}\log\log\frac{1}{\mathcal{E}}} \leq 1. \tag{165}$$

Combining (162), (163), (164) and (165) proves (87).

References

[1] S. Shamai (Shitz), "Capacity of a pulse amplitude modulated direct detection photon channel," *Proc. IEE*, vol. 137, no. 6, pp. 424–430, Dec. 1990.

[2] A. Lapidoth and S. M. Moser, "Bounds on the capacity of the discrete-time Poisson Channel," in *Proc. 41st Allerton Conf. Commun., Control, Comput.*, Monticello, IL, Oct. 1–3, 2003.

[3] A. Lapidoth and S. M. Moser, "On the capacity of the discrete-time Poisson Channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 303–322, Jan. 2009.

[4] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multiple-antenna systems on flat fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2426–2467, Oct. 2003.

[5] S. Verdú, "On channel capacity per unit cost," *IEEE Trans. Inf. Theory*, vol. 36, pp. 1019–1030, Sep. 1990.

[6] V. Venkatesan, "On low power capacity of the Poisson channel," Master's thesis, Signal Inf. Proc. Lab., ETH Zurich, Switzerland, 2008.

[7] A. Martinez, "Low-signal-energy asymptotics of capacity and mutual information for the discrete-time Poisson channel," *IEEE Trans. Inf. Theory*, Aug. 2008, submitted for publication.

[8] A. D. Wyner, "Capacity and error exponent for the direct detection photon channel—Part I," *IEEE Trans. Inf. Theory*, vol. 34, no. 6, pp. 1449–1461, Nov. 1988.

[9] V. Giovanetti, 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] W. Feller, *An Introduction to Probability Theory and its Applications*, 3rd ed. New York: Wiley, 1957/1968, vol. 1.

Amos Lapidoth (S'89–M'95–SM'00–F'04) received the B.A. degree in mathematics (summa cum laude, 1986), the B.Sc. degree in electrical engineering (summa cum laude, 1986), and the M.Sc. degree in electrical engineering (1990) all from the Technion—Israel Institute of Technology. He received the Ph.D. degree in electrical engineering from Stanford University in 1995.

In the years 1995–1999 he was an Assistant and Associate Professor at the Department of Electrical Engineering and Computer Science at the Massachusetts Institute of Technology, and was the KDD Career Development Associate Professor in Communications and Technology. He is now Professor of Information Theory at ETH Zurich in Switzerland. He is the author of the book *A Foundation in Digital Communication*, published by Cambridge University Press in 2009. His research interests are in digital communications and information theory.

Dr. Lapidoth served in the years 2003–2004 and 2009 as Associate Editor for Shannon Theory for the IEEE TRANSACTIONS ON INFORMATION THEORY.

Jeffrey H. Shapiro (S'67–M'70–SM'84–F'95) received the S.B., S.M. E.E., and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology in 1967, 1968, 1969, and 1970, respectively. From 1970 to 1973 he was an Assistant Professor of electrical sciences and applied physics at Case Western Reserve University, Cleveland, OH. Since 1973 he has been on the faculty at MIT, where he is now the Julius A. Stratton Professor of Electrical Engineering and the Director of the Research Laboratory of Electronics. His current research interests center on the application of communication theory to optical systems. He is best know for his work on the generation, detection, and application of squeezed-state light beams, but he has also published extensively in the areas of atmospheric optical communication, coherent laser radar, and quantum information science. Professor Shapiro is a Fellow of the Optical Society of America, the American Physical Society and the Institute of Physics. In 2008 he was co-recipient of the Quantum Electronics Award from the IEEE Photonics Society (formerly known as the IEEE Lasers and Electro-Optics Society), and he also received the Quantum Communication Award for Theoretical Research from Tamagawa University, Tokyo, Japan.

Vinodh Venkatesan (S'08) received the Bachelor of Technology degree in electrical engineering from the Indian Institute of Technology Madras in 2006 and the Master of Science degree in information technology and electrical engineering from ETH Zurich, Switzerland, in 2008. He is currently a Pre-doc at IBM Research in Zurich while working toward the Ph.D. degree at EPFL, Switzerland. His research interests are in data storage systems, network coding, and information theory.

Ligong Wang (S'08) received the B.E. degree in electronic engineering from Tsinghua University, Beijing, China, in 2004 and the M.Sc. degree in electrical engineering from ETH Zurich, Switzerland, in 2006. He is currently working toward the Ph.D. degree at the ETH Zurich, Switzerland. His research interests are in classical and quantum information theory.