Bounds on the Capacity of the Discrete-Time Poisson Channel

Amos Lapidoth   Stefan M. Moser*

Signal and Information Processing Laboratory
Swiss Federal Institute of Technology (ETH) Zurich
CH-8092 Zurich, Switzerland
e-mail: {lapidoth, moser}@isi.ee.ethz.ch

Abstract

The large-inputs asymptotic capacity of a peak and average power limited discrete-time Poisson channel is derived using a new firm (non-asymptotic) lower bound and an asymptotic upper bound. The upper bound is based on the dual expression for channel capacity and the recently introduced notion of *capacity-achieving input distributions that escape to infinity*. The lower bound is based on a lemma that lower bounds the entropy of a conditionally Poisson random variable in terms of the differential entropy of the conditional mean.

1   Introduction

We consider a memoryless discrete-time channel whose output Y takes value in the set of non-negative integers $\mathbb{Z}_0^+$ and whose input takes value in the set of non-negative reals $\mathbb{R}_0^+$. Conditional on the input $x \geq 0$, the output has a Poisson distribution of mean $x + \lambda_0$, where λ_0 is some non-negative constant. Thus

$$\Pr[Y = y | X = x] = e^{-(x+\lambda_0)} \frac{(x+\lambda_0)^y}{y!}, \quad y \in \mathbb{Z}_0^+, x \geq 0. \tag{1}$$

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; λ_0 similarly models the time-by-intensity product of the background radiation ("dark current"); and the output Y models the number of photons arriving at the receiver during the pulse duration. An average power constraint on the transmitter is accounted for by the average input constraint

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

and a peak power constraint by

$$0 \leq X \leq \mathsf{A}. \tag{3}$$

We use $0 < \alpha \leq 1$ to denote the average-to-peak ratio

$$\alpha = \mathsf{P}/\mathsf{A}. \tag{4}$$

The case $\alpha = 1$ corresponds to the absence of an average power constraint, whereas $\alpha \ll 1$ corresponds to a very weak peak power constraint.

No analytic expression for the capacity of the Poisson channel is known. In [1] Shamai showed that capacity-achieving input distributions are discrete with a finite number of mass points that increases to infinity as the constraints are relaxed.

In [2] Brady & Verdú considered the case of the channel with an average power constraint only. The following bounds were derived: Let $\mathsf{P}, \lambda_0 \uparrow \infty$ proportionally such that $\mathrm{SNR} \triangleq \frac{\mathsf{P}}{\lambda_0}$ is constant. Given $\epsilon > 0$ there exists an P_ϵ such that for all $\mathsf{P} > \mathsf{P}_\epsilon$ the capacity is bounded by

$$C(\mathsf{P}) \geq \frac{1}{2} \log \frac{\mathsf{P}}{2\pi} - \frac{1}{2} \log \left(1 + \frac{1}{\mathrm{SNR}} \right) - \epsilon; \tag{5}$$

$$C(\mathsf{P}) \leq \frac{1}{2} \log \frac{\mathsf{P}}{2\pi} + \log \left(\sqrt{\mathrm{SNR}} \left(1 + \frac{1}{\mathsf{P}_\epsilon} \right) + \frac{1}{\sqrt{\mathrm{SNR}}} \right) + 1 + \log \frac{3}{2} + \epsilon. \tag{6}$$

Note that the difference between the upper and lower bound is unbounded if the dark current is held constant while P tends to infinity.

Here we present results for the more general case of both peak and average power constraints. We improve the lower bound to channel capacity and introduce an asymptotic — peak and average powers tending to infinity with their ratio and the dark current held fixed — upper bound. The upper and lower bounds asymptotically coincide, thus yielding the asymptotic expansion for channel capacity.

2 Results

We begin with the case where only the average power constraint $\mathsf{E}[X] \leq \mathsf{P}$ is imposed.

Theorem 2.1. *The channel capacity $C(\mathsf{P})$ of a Poisson channel with dark current λ_0 under an average power constraint $\mathsf{E}[X] \leq \mathsf{P}$ is bounded as follows:*

$$C(\mathsf{P}) \geq \frac{1}{2} \log \mathsf{P} - \sqrt{\frac{\pi \left(\lambda_0 + \frac{1}{12} \right)}{2\mathsf{P}}} + (\mathsf{P} + 1) \log \left(1 + \frac{1}{\mathsf{P}} \right) - 1; \tag{7}$$

$$C(\mathsf{P}) \leq \frac{1}{2} \log \mathsf{P} + o(1). \tag{8}$$

Here, the error term $o(1)$ tends to zero as $\mathsf{P} \uparrow \infty$. Hence, the asymptotic expansion for the channel capacity is

$$\lim_{\mathsf{P} \uparrow \infty} \left\{ C(\mathsf{P}) - \frac{1}{2} \log \mathsf{P} \right\} = 0, \tag{9}$$

irrespective of λ_0, which is held fixed.

In Figure 1 these bounds are plotted together with the lower and upper bound (5) and (6) from Brady & Verdú [2].

Next consider the case where both average and peak power constraints are imposed. Holding the average-to-peak ratio α fixed, we distinguish between two cases: $\alpha \geq \frac{1}{3}$ (including the peak power constraint only case $\alpha = 1$) and $0 < \alpha < \frac{1}{3}$. We begin with the former:

Theorem 2.2. *The channel capacity $C(\mathsf{A}, \mathsf{P})$ of a Poisson channel with dark current λ_0 under a peak power constraint $0 \le x \le \mathsf{A}$ and an average power constraint $\mathsf{E}[X] \le \mathsf{P}$, where the ratio $\alpha = \frac{\mathsf{P}}{\mathsf{A}}$ lies in $\left[\frac{1}{3}, 1\right]$, is bounded as follows:*

$$C(\mathsf{A}, \mathsf{P}) \ge \frac{1}{2} \log \mathsf{A} - \frac{1}{2} \log \frac{\pi e}{2} + \left(\frac{\mathsf{A}}{3} + 1\right) \log\left(1 + \frac{3}{\mathsf{A}}\right) - 1$$

$$- \frac{1}{2} \log\left(1 + \frac{\lambda_0 + \frac{1}{12}}{\mathsf{A}}\right) + \sqrt{\frac{\lambda_0 + \frac{1}{12}}{\mathsf{A}}} \arctan\left(\sqrt{\frac{\mathsf{A}}{\lambda_0 + \frac{1}{12}}}\right); \qquad (10)$$

$$C(\mathsf{A}, \mathsf{P}) \le \frac{1}{2} \log \mathsf{A} - \frac{1}{2} \log \frac{\pi e}{2} + o(1). \qquad (11)$$

Here, the error term $o(1)$ tends to zero as the average and peak powers tend to infinity with their ratio held fixed at α, $\frac{1}{3} \le \alpha \le 1$. Hence, the asymptotic expansion for the channel capacity is

$$\lim_{\mathsf{A}\uparrow\infty} \left\{ C(\mathsf{A}, \alpha\mathsf{A}) - \frac{1}{2} \log \mathsf{A} \right\} = -\frac{1}{2} \log \frac{\pi e}{2}, \qquad \frac{1}{3} \le \alpha \le 1. \qquad (12)$$

In the second case $0 < \alpha < \frac{1}{3}$ we have:

Theorem 2.3. *The channel capacity $C(\mathsf{A}, \mathsf{P})$ of a Poisson channel with dark current λ_0 under a peak power constraint $0 \le x \le \mathsf{A}$ and an average power constraint $\mathsf{E}[X] \le \mathsf{P}$, where the ratio $\alpha = \frac{\mathsf{P}}{\mathsf{A}}$ lies in $\left(0, \frac{1}{3}\right)$, is bounded as follows:*

$$C(\mathsf{A}, \mathsf{P}) \ge \frac{1}{2} \log \mathsf{A} - \frac{1}{2} \log 2\pi e + (\alpha - 1)u - \log\left(\frac{1}{2} - \alpha u\right)$$

$$- e^u \left(\frac{1}{2} - \alpha u\right) \left(\log\left(1 + \frac{\lambda_0 + \frac{1}{12}}{\mathsf{A}}\right) - 2\sqrt{\frac{\lambda_0 + \frac{1}{12}}{\mathsf{A}}} \arctan\left(\sqrt{\frac{\mathsf{A}}{\lambda_0 + \frac{1}{12}}}\right) \right)$$

$$+ (\alpha\mathsf{A} + 1) \log\left(1 + \frac{1}{\alpha\mathsf{A}}\right) - 1; \qquad (13)$$

$$C(\mathsf{A}, \mathsf{P}) \le \frac{1}{2} \log \mathsf{A} + \alpha u - u - \log\left(\frac{1}{2} - \alpha u\right) - \frac{1}{2} \log 2\pi e + o(1). \qquad (14)$$

Here, $u \in \left(0, \frac{1}{2\alpha}\right)$ is the non-zero solution to

$$\sqrt{\pi} \mathrm{erf}\left(\sqrt{u}\right) \left(\frac{1}{2} - \alpha u\right) - \sqrt{u} e^{-u} = 0, \qquad (15)$$

and the error function $\mathrm{erf}\,(\cdot)$ is defined as

$$\mathrm{erf}\,(\xi) = \frac{2}{\sqrt{\pi}} \int_0^\xi e^{-t^2} \, dt. \qquad (16)$$

The error term $o(1)$ tends to zero as the average and peak powers tend to infinity with their ratio held fixed at α, $0 < \alpha < \frac{1}{3}$. Hence, the asymptotic expansion for the channel capacity is

$$\lim_{\mathsf{A}\uparrow\infty} \left\{ C(\mathsf{A}, \alpha\mathsf{A}) - \frac{1}{2} \log \mathsf{A} \right\} = -\frac{1}{2} \log 2\pi e + (\alpha - 1)u - \log\left(\frac{1}{2} - \alpha u\right), \quad 0 < \alpha < \frac{1}{3}, \quad (17)$$

where u is defined as above to be the non-zero solution to (15).

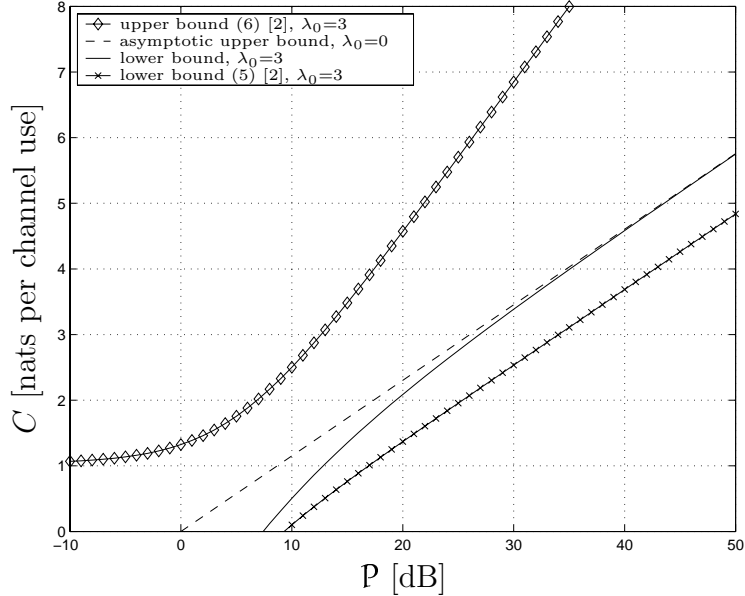See Figure 2 for plots of these bounds for different values of α.

Figure 1: Firm lower and asymptotic upper bound on the capacity for a Poisson channel with average power constraint $\mathsf{E}[X] \leq \mathsf{P}$. Additionally the lower and upper bound (5) and (6) by Brady & Verdú [2] are plotted. Apart from the asymptotic upper bound which does not depend on the dark current all bounds assume a dark current $\lambda_0 = 3$. In (5) and (6) P_ϵ has been set to infinity and ϵ to 0.
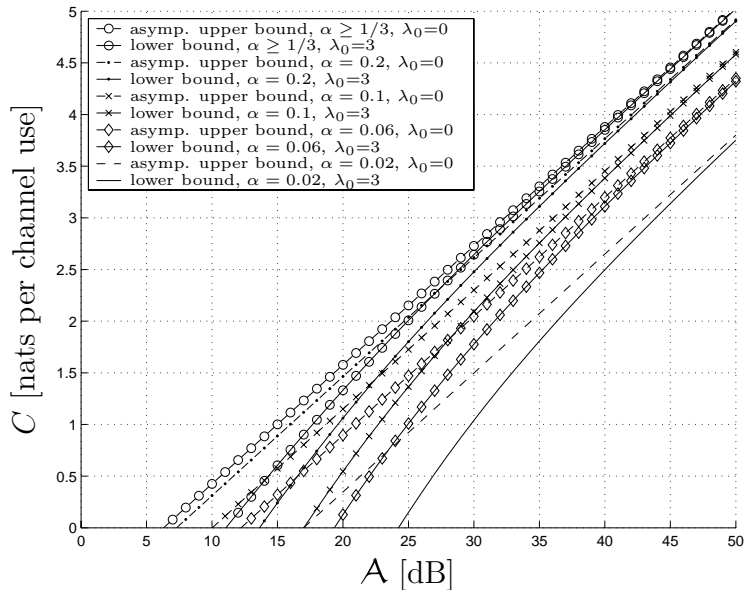

Figure 2: Firm lower and asymptotic upper bounds on the capacity for a Poisson channel under peak and average power constraint. The parameter α denotes the ratio of average to peak power constraint $\frac{\mathsf{P}}{\mathsf{A}}$. For $\alpha \geq \frac{1}{3}$ (including the peak power constraint only case $\alpha = 1$) the bounds do not depend on α anymore.

3 Derivation

3.1 Lower Bounds

The lower bounds are based on two lemmas yielding a lower bound to $H(Y)$ and an upper bound to $H(Y|X)$ in terms of the input X. The latter is well known and is based

on the variance of Y given $X = x$:

Lemma 3.1. *The entropy $H(Y|X = x)$ is upper bounded by*

$$H(Y|X = x) \leq \frac{1}{2}\log 2\pi e \left(x + \lambda_0 + \frac{1}{12}\right). \tag{18}$$

Proof. See [3, Theorem 16.3.3]. □

As to the former we have the following

Lemma 3.2. *Let Y be the output of a Poisson channel with input X and dark current λ_0. Assume that X has a finite positive expectation $\mathsf{E}[X] = \mathsf{P}$. Then*

$$H(Y) \geq h(X) + (1 + \mathsf{P})\log\left(1 + \frac{1}{\mathsf{P}}\right) - 1 \tag{19}$$

$$> h(X). \tag{20}$$

Proof. See Appendix A. □

The lower bound (13) can now be derived using these lemmas in the following way:

$$I(X;Y) = H(Y) - H(Y|X) \tag{21}$$

$$\geq h(X) + (\alpha\mathsf{A} + 1)\log\left(1 + \frac{1}{\alpha\mathsf{A}}\right) - 1 - \frac{1}{2}\log 2\pi e$$

$$- \frac{1}{2}\mathsf{E}\left[\log\left(X + \lambda_0 + \frac{1}{12}\right)\right] \tag{22}$$

$$= h(X) + (\alpha\mathsf{A} + 1)\log\left(1 + \frac{1}{\alpha\mathsf{A}}\right) - 1 - \frac{1}{2}\log 2\pi e$$

$$- \frac{1}{2}\mathsf{E}[\log X] - \frac{1}{2}\mathsf{E}\left[\log\left(1 + \frac{\lambda_0 + \frac{1}{12}}{X}\right)\right]. \tag{23}$$

We choose now an input distribution of the form

$$Q(x) = \frac{1}{\sqrt{\beta\pi x} \cdot \mathrm{erf}\left(\sqrt{\frac{\mathsf{A}}{\beta}}\right)} \cdot e^{-\frac{x}{\beta}}, \qquad 0 \leq x \leq \mathsf{A}, \tag{24}$$

where $\beta > 0$ is a parameter chosen such that $\mathsf{E}[X] = \alpha\mathsf{A}$, *i.e.*,

$$\frac{\beta}{2} - \frac{\beta\sqrt{\frac{\mathsf{A}}{\beta}}e^{-\frac{\mathsf{A}}{\beta}}}{\sqrt{\pi}\mathrm{erf}\left(\sqrt{\frac{\mathsf{A}}{\beta}}\right)} = \alpha\mathsf{A}. \tag{25}$$

For this distribution:

$$h(X) = \frac{1}{2}\log\beta + \log\left(\sqrt{\pi}\mathrm{erf}\left(\sqrt{\frac{\mathsf{A}}{\beta}}\right)\right) + \frac{\alpha\mathsf{A}}{\beta} + \frac{1}{2}\mathsf{E}[\log X]; \tag{26}$$

$$\mathsf{E}\left[\log\left(1 + \frac{a}{X}\right)\right] \leq \int_0^{\mathsf{A}} \log\left(1 + \frac{a}{x}\right) \cdot \frac{1}{\sqrt{\beta\pi x} \cdot \mathrm{erf}\left(\sqrt{\frac{\mathsf{A}}{\beta}}\right)}\,\mathrm{d}x \tag{27}$$

$$= \frac{2\pi\sqrt{a} + 2\sqrt{\mathsf{A}}\log\left(1 + \frac{a}{\mathsf{A}}\right) - 4\sqrt{a}\arctan\left(\sqrt{\frac{a}{\mathsf{A}}}\right)}{\sqrt{\beta\pi} \cdot \mathrm{erf}\left(\sqrt{\frac{\mathsf{A}}{\beta}}\right)}. \tag{28}$$

We substitute β with $\frac{A}{u}$, where $u \in \left(0, \frac{1}{2\alpha}\right)$ is defined by (25) as non-zero solution of

$$\sqrt{\pi}\,\mathrm{erf}\left(\sqrt{u}\right) = \frac{\sqrt{u}\,e^{-u}}{\frac{1}{2} - \alpha u}. \tag{29}$$

Using (26), (28), and (29) in (23) we get

$$I(X;Y) \geq \frac{1}{2}\log\beta + \log\sqrt{\pi}\,\mathrm{erf}\left(\sqrt{u}\right) + \alpha u + (\alpha A + 1)\log\left(1 + \frac{1}{\alpha A}\right) - 1 - \frac{1}{2}\log 2\pi e$$

$$- \frac{\pi\sqrt{\lambda_0 + \frac{1}{12}} + \sqrt{A}\log\left(1 + \frac{\lambda_0 + \frac{1}{12}}{A}\right) - 2\sqrt{\lambda_0 + \frac{1}{12}}\arctan\left(\sqrt{\frac{\lambda_0 + \frac{1}{12}}{A}}\right)}{\sqrt{\beta\pi}\cdot\mathrm{erf}\left(\sqrt{u}\right)} \tag{30}$$

$$= \frac{1}{2}\log A + \alpha u - u - \log\left(\frac{1}{2} - \alpha u\right) + (\alpha A + 1)\log\left(1 + \frac{1}{\alpha A}\right) - 1$$

$$- \frac{1}{2}\log 2\pi e - \left(\frac{1}{2} - \alpha u\right)e^u\left\{\pi\sqrt{\frac{\lambda_0 + \frac{1}{12}}{A}} + \log\left(1 + \frac{\lambda_0 + \frac{1}{12}}{A}\right)\right.$$

$$\left. - 2\sqrt{\frac{\lambda_0 + \frac{1}{12}}{A}}\arctan\sqrt{\frac{\lambda_0 + \frac{1}{12}}{A}}\right\} \tag{31}$$

$$= \frac{1}{2}\log A - \frac{1}{2}\log 2\pi e + \alpha u - u - \log\left(\frac{1}{2} - \alpha u\right)$$

$$- \left(\frac{1}{2} - \alpha u\right)e^u\left(\log\left(1 + \frac{\lambda_0 + \frac{1}{12}}{A}\right) - 2\sqrt{\frac{\lambda_0 + \frac{1}{12}}{A}}\arctan\sqrt{\frac{A}{\lambda_0 + \frac{1}{12}}}\right)$$

$$+ (\alpha A + 1)\log\left(1 + \frac{1}{\alpha A}\right) - 1, \tag{32}$$

where we have used

$$\arctan\left(\frac{1}{\xi}\right) = \frac{\pi}{2} - \arctan\xi, \qquad \xi \in \mathbb{R}. \tag{33}$$

This concludes our proof.

The derivations of (10) and (7) are very similar and therefore omitted. Note that (10) corresponds to (13) with $u \downarrow 0$.

3.2 Asymptotic Upper Bounds

For the derivation of the asymptotic upper bounds we will assume that the dark current is zero, $i.e.$, $\lambda_0 = 0$. This is no real restriction because firstly any upper bound to the capacity of a Poisson channel without dark current is also an upper bound to the case with non-zero dark current. And secondly we will show that asymptotically the dark current has no impact on the capacity at all. The derivation of the upper bounds relies on the following:

- In order to avoid the problems arising from the fact that the output of a Poisson channel is discrete, we introduce a new channel defined as $\tilde{Y} = Y + U$, where U is uniformly distributed between 0 and 1. Then $\tilde{Y} \geq 0$ is continuous with a distribution

$$\tilde{W}(\tilde{y}|x) = e^{-x}\frac{x^{\lfloor\tilde{y}\rfloor}}{\lfloor\tilde{y}\rfloor!}, \quad \tilde{y} \geq 0, x \geq 0, \tag{34}$$

where $\lfloor \cdot \rfloor$ denotes the operation that truncates any real-number to the next smaller (or equal) integer. Note that $I(X; \tilde{Y}) = I(X; Y)$ and $h(\tilde{Y}|X = x) = H(Y|X = x)$. Further, note that

$$\lim_{x \uparrow \infty} \left\{ \mathsf{E}\left[\log \tilde{Y} \middle| X = x \right] - \log x \right\} = 0; \tag{35}$$

$$\lim_{x \uparrow \infty} \left\{ h(\tilde{Y}|X = x) - \frac{1}{2} \log x \right\} = \lim_{x \uparrow \infty} \left\{ H(Y|X = x) - \frac{1}{2} \log x \right\} \geq \frac{1}{2} \log 2\pi e. \tag{36}$$

- The upper bounds are based on the following inequality [4]:

$$C \leq \mathsf{E}_{Q^*}\left[D\big(\tilde{W}(\cdot|X) \| R(\cdot) \big) \right] \tag{37}$$

$$= \mathsf{E}_{Q^*}\left[-\int_0^\infty \tilde{W}(\tilde{y}|X) \log R(\tilde{y}) \, \mathrm{d}\tilde{y} - h(\tilde{Y}|X = x) \right], \tag{38}$$

where Q^* denotes the capacity-achieving input distribution to the Poisson channel fulfilling all input constraints, and where $R(\cdot)$ denotes any distribution over $\mathbb{R}_0^+$ that can be chosen freely.

- In [4, Definition 4.11, p. 8] the notion of *capacity achieving input distributions that escape to infinity* is defined. Using [4, Theorem 4.13, p. 9] one can show that the capacity of a Poisson channel can be achieved by such distributions, *i.e.*, for any P_0 there exists input distributions $\{Q_\mathsf{P}^*\}_{\mathsf{P} \geq 0}$ satisfying the peak and average power constraints such that

$$\lim_{\mathsf{P} \uparrow \infty} \left\{ C(\mathsf{P}/\alpha, \mathsf{P}) - I(Q_\mathsf{P}^*; W) \right\} = 0 \tag{39}$$

and

$$\lim_{\mathsf{P} \uparrow \infty} Q_\mathsf{P}^*(X \leq \mathsf{P}_0) = 0. \tag{40}$$

The derivation of (14) is based on (38) with the choice

$$R(\tilde{y}) = \begin{cases} pR_1(\tilde{y}) = p \cdot \dfrac{\tilde{y}^{\mu-1} e^{-\tilde{y}/\beta}}{\beta^\mu \gamma(\mu, \mathsf{A}(1+\delta)/\beta)} & \forall\, 0 \leq \tilde{y} \leq \mathsf{A}(1+\delta) \\ (1-p)R_2(\tilde{y}) = (1-p) \cdot e^{-(\tilde{y} - \mathsf{A}(1+\delta))} & \forall\, \tilde{y} > \mathsf{A}(1+\delta) \end{cases}, \tag{41}$$

where $\mu, \beta, \delta > 0$,

$$p = \Pr\left[\tilde{Y} \leq \mathsf{A}(1+\delta) \middle| X = \mathsf{A} \right], \tag{42}$$

and where $\gamma(\cdot, \cdot)$ denotes the incomplete gamma function

$$\gamma(\mu, \xi) = \int_0^\xi e^{-t} t^{\mu-1} \, \mathrm{d}t \qquad \mu > 0. \tag{43}$$

Plugging $R(\tilde{y})$ into the first term of (38) yields

$$-\int_0^\infty \tilde{W}(\tilde{y}|x) \log R(\tilde{y}) \, \mathrm{d}\tilde{y}$$

$$= \underbrace{-\log p \cdot \Pr\left[\tilde{Y} \leq \mathsf{A}(1+\delta) \middle| X = x \right]}_{I_\mathrm{a}} \underbrace{- \int_0^{\mathsf{A}(1+\delta)} \tilde{W}(\tilde{y}|x) \log R_1(\tilde{y}) \, \mathrm{d}\tilde{y}}_{I_\mathrm{b}}$$

$$\underbrace{- \log(1-p) \cdot \Pr\left[\tilde{Y} > \mathsf{A}(1+\delta) \middle| X = x \right]}_{I_\mathrm{c}} + \underbrace{\int_{\mathsf{A}(1+\delta)}^\infty \tilde{W}(\tilde{y}|x)\big(\tilde{y} - \mathsf{A}(1+\delta)\big) \, \mathrm{d}\tilde{y}}_{I_\mathrm{d}}. \tag{44}$$

We will now look at each term individually:

$$I_{\mathrm{a}} = \Pr\left[\tilde{Y} \le \mathsf{A}(1+\delta) \Big| X = x\right] \log \frac{1}{p} \le \log \frac{1}{p}; \tag{45}$$

$$I_{\mathrm{b}} = \int_0^{\mathsf{A}(1+\delta)} \tilde{W}(\tilde{y}|x) \log \frac{\beta^\mu \gamma\left(\mu, \frac{\mathsf{A}(1+\delta)}{\beta}\right)}{\tilde{y}^{\mu-1} e^{-\tilde{y}/\beta}} \, \mathrm{d}\tilde{y} \tag{46}$$

$$= \left\{ \mu \log \beta + \log \gamma\left(\mu, \frac{\mathsf{A}(1+\delta)}{\beta}\right) \right\} \Pr\left[\tilde{Y} \le \mathsf{A}(1+\delta) \Big| X = x\right]$$

$$+ (1-\mu) \int_0^{\mathsf{A}(1+\delta)} \tilde{W}(\tilde{y}|x) \log \tilde{y} \, \mathrm{d}\tilde{y} + \frac{1}{\beta} \int_0^{\mathsf{A}(1+\delta)} \tilde{y} \tilde{W}(\tilde{y}|x) \, \mathrm{d}\tilde{y} \tag{47}$$

$$= \left\{ \mu \log \beta + \log \gamma\left(\mu, \frac{\mathsf{A}(1+\delta)}{\beta}\right) \right\} \Pr\left[\tilde{Y} \le \mathsf{A}(1+\delta) \Big| X = x\right]$$

$$+ (1-\mu)\left\{ \mathsf{E}\left[\log \tilde{Y} \;\Big|\; X = x\right] - \underbrace{\int_{\mathsf{A}(1+\delta)}^\infty \tilde{W}(\tilde{y}|x) \log \tilde{y} \, \mathrm{d}\tilde{y}}_{\ge 0 \text{ for } \mathsf{A} > 1} \right\}$$

$$+ \frac{1}{\beta}\left\{ \mathsf{E}\left[\tilde{Y} \;\Big|\; X = x\right] - \underbrace{\int_{\mathsf{A}(1+\delta)}^\infty \tilde{y} \tilde{W}(\tilde{y}|x) \, \mathrm{d}\tilde{y}}_{\ge 0} \right\} \tag{48}$$

$$\le \left\{ \mu \log \beta + \log \gamma\left(\mu, \frac{\mathsf{A}(1+\delta)}{\beta}\right) \right\} \Pr\left[\tilde{Y} \le \mathsf{A}(1+\delta) \Big| X = x\right]$$

$$+ (1-\mu)\mathsf{E}\left[\log \tilde{Y} \;\Big|\; X = x\right] + \frac{1}{\beta}\left(x + \frac{1}{2}\right), \tag{49}$$

where for the inequality we assumed that $\mu < 1$ and $\mathsf{A} > 1$. Further,

$$I_{\mathrm{c}} \le \Pr\left[\tilde{Y} > \mathsf{A}(1+\delta) \Big| X = \mathsf{A}\right] \log \frac{1}{1-p} = (1-p) \log \frac{1}{1-p}, \tag{50}$$

where we have used the monotonicity of the Poisson distribution and the fact that $X \le \mathsf{A}$. Finally,

$$I_{\mathrm{d}} \le \sum_{y=\lfloor \mathsf{A}(1+\delta) \rfloor}^\infty y e^{-x} \frac{1}{y!} x^y - \underbrace{\mathsf{A}(1+\delta) \Pr\left[\tilde{Y} > \mathsf{A}(1+\delta) \Big| X = x\right]}_{\ge 0} \tag{51}$$

$$\le x \sum_{y=\lfloor \mathsf{A}(1+\delta) \rfloor}^\infty e^{-x} \frac{1}{(y-1)!} x^{y-1} \tag{52}$$

$$= x \sum_{y=\lfloor \mathsf{A}(1+\delta) \rfloor - 1}^\infty e^{-x} \frac{1}{y!} x^y \tag{53}$$

$$= x \int_{\lfloor \mathsf{A}(1+\delta) \rfloor - 1}^\infty \tilde{W}(\tilde{y}|x) \, \mathrm{d}\tilde{y} \tag{54}$$

$$= x \Pr\left[\tilde{Y} > \lfloor \mathsf{A}(1+\delta) - 1 \rfloor \Big| X = x\right] \tag{55}$$

$$\le \mathsf{A} \Pr\left[\tilde{Y} > \lfloor \mathsf{A}(1+\delta) - 1 \rfloor \Big| X = x\right] \tag{56}$$

$$\le \mathsf{A} \Pr\left[\tilde{Y} > \mathsf{A}(1+\delta_2) \Big| X = x\right] \tag{57}$$

$$\leq \mathsf{A}\Pr\Big[\tilde{Y} > \mathsf{A}(1+\delta_2)\Big|X = \mathsf{A}\Big] \tag{58}$$

$$\leq \mathsf{A}e^r e^{\mathsf{A}(e^r - 1 - r - r\delta_2)}, \tag{59}$$

where in (57) we assume that $\mathsf{A} \gg 1$ and we have chosen $0 < \delta_2 < \delta$. The second last inequality follows from the monotonicity of the Poisson distribution. And for the last inequality we used Chernoff's bound $\Pr[W \geq w] \leq e^{-rw}\mathsf{E}\big[e^{rW}\big]$ with the choice $r = \log(1 + \delta_2)$. Note that $(1 + \delta_2)\log(1 + \delta_2) - \delta_2 > 0$ for $\delta_2 > 0$.

Combining all these results together with (35), (36), and

$$\tilde{p}(\mathsf{A}) := \mathsf{E}_{Q^*}\Big[\Pr\Big[\tilde{Y} \leq \mathsf{A}(1+\delta)\Big|X = x\Big]\Big] \to 1 \quad \text{as } \mathsf{A} \uparrow \infty \tag{60}$$

yields

$$C \leq \tilde{p}(\mathsf{A})\mu \log \beta + \tilde{p}(\mathsf{A}) \log \gamma\left(\mu, \frac{\mathsf{A}(1+\delta)}{\beta}\right) + \left(\frac{1}{2} - \mu\right) \mathsf{E}_{Q^*}[\log X]$$
$$+ \frac{1}{\beta}\left(\alpha\mathsf{A} + \frac{1}{2}\right) - \frac{1}{2}\log(2\pi e) + o(1), \tag{61}$$

where $o(1)$ tends to zero as $\mathsf{A} \uparrow \infty$. To further simplify we choose $\mu = 1/2$ and substitute β with $\mathsf{A}(1+\delta)/u$, $u > 0$:

$$C \leq \tilde{p}(\mathsf{A})\frac{1}{2}\log \mathsf{A}(1+\delta) - \tilde{p}(\mathsf{A})\frac{1}{2}\log u + \tilde{p}(\mathsf{A}) \log \gamma\left(\frac{1}{2}, u\right)$$
$$+ \frac{u}{\mathsf{A}(1+\delta)}\left(\alpha\mathsf{A} + \frac{1}{2}\right) - \frac{1}{2}\log(2\pi e) + o(1) \tag{62}$$

$$= \frac{1}{2}\tilde{p}(\mathsf{A}) \log \mathsf{A} - \frac{1}{2}\tilde{p}(\mathsf{A}) \log u + \tilde{p}(\mathsf{A}) \log \sqrt{\pi}\mathrm{erf}\left(\sqrt{u}\right)$$
$$+ \frac{\alpha u}{1+\delta} - \frac{1}{2}\log(2\pi e) + \frac{1}{2}\tilde{p}(\mathsf{A})\log(1+\delta) + o(1) \tag{63}$$

$$= \tilde{p}(\mathsf{A})\left(\frac{1}{2}\log \mathsf{A} + \frac{1}{2}\log(1+\delta) - u - \log\left(\frac{1}{2} - \alpha u\right)\right)$$
$$- \frac{1}{2}\log 2\pi e + \frac{\alpha u}{1+\delta} + o(1). \tag{64}$$

The first equality follows from the identity $\gamma\left(\frac{1}{2}, u\right) = \sqrt{\pi}\mathrm{erf}\left(\sqrt{u}\right)$ where $\gamma(\cdot, \cdot)$ is defined in (43) and $\mathrm{erf}(\cdot)$ in (16). In the second equality we have chosen $u \in (0, 1/2\alpha]$ to be the non-zero solution to (29). Note that such a non-zero solution always exists as long as $0 < \alpha < \frac{1}{3}$. This concludes our proof.

The derivations of (11) and (8) are very similar and therefore omitted.

A   Proof of Lemma 3.2

Proof. Given $X = x$, Y can be written as $Y = Y_1 + Y_2$, where $Y_1 \sim \mathcal{P}o(x)$ and $Y_2 \sim \mathcal{P}o(\lambda_0)$, $Y_1 \perp\!\!\!\perp Y_2$. But

$$H(Y) = H(Y_1 + Y_2) \geq H(Y_1 + Y_2|Y_2) = H(Y_1|Y_2) = H(Y_1), \tag{65}$$

and we can restrict ourselves to the case where $\lambda_0 = 0$.

The proof of this case is based on the data processing inequality for relative entropy [5, Chapter 1, Lemma 3.11(ii)]. Let $P_X(\cdot)$ denote the distribution of X and let $P_{\mathrm{E}}(\cdot)$

denote the mean-P exponential distribution on $\mathbb{R}_0^+$. Let $P_Y(\cdot)$ be the distribution of Y when Y is conditionally Poisson given X, for $X \sim P_X$, and let $P_G(\cdot)$ be the distribution of Y when Y is conditionally Poisson given X and $X \sim P_E$. It can be shown that $P_G(\cdot)$ is a mean-P geometric distribution on $\mathbb{Z}_0^+$:

$$P_G(y) = (1 - p) \cdot p^y, \qquad y \in \mathbb{Z}_0^+ \tag{66}$$

where

$$p = \frac{\mathsf{P}}{1 + \mathsf{P}}. \tag{67}$$

By the data processing theorem we obtain:

$$D\big(P_X(\cdot)\big\|P_E(\cdot)\big) \geq D\big(P_Y(\cdot)\big\|P_G(\cdot)\big), \tag{68}$$

where $D(\cdot\|\cdot)$ denotes relative entropy [3, Chapter 2]. The first inequality in the lemma's statement now follows by evaluating the left hand side of (68)

$$D\big(P_X(\cdot)\big\|P_E(\cdot)\big) = \int_0^\infty P_X(x) \log \frac{P_X(x)}{\frac{1}{\mathsf{P}} e^{-x/\mathsf{P}}} \, \mathrm{d}x = -h(X) + \log \mathsf{P} + 1, \tag{69}$$

evaluating the right hand side of (68)

$$D\big(P_Y(\cdot)\big\|P_G(\cdot)\big) = \sum_{y=0}^\infty P_Y(y) \log \frac{P_Y(y)}{(1 - p)p^y} = -H(Y) - \mathsf{P} \log p + \log \frac{1}{1 - p}, \tag{70}$$

and using (67).

The second inequality follows by noting that $(1 + \mathsf{P}) \log(1 + \mathsf{P}^{-1}) - 1$ is monotonically decreasing in P and approaches zero, as $\mathsf{P} \to \infty$. $\qquad\qquad\square$

References

[1] S. Shamai(Shitz), "On the capacity of a pulse amplitude modulated direct detection photon channel," in *Proc. IEE*, vol. 137, pt. I, Dec. 1990, pp. 424–430.

[2] D. Brady and S. Verdú, "The asymptotic capacity of the direct detection photon channel with a bandwidth constraint," in *Proc. 28th Allerton Conf. Comm., Contr. and Comp.*, Oct. 3–5, 1990, pp. 691–700.

[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* John Wiley & Sons, 1991.

[4] A. Lapidoth and S. M. Moser, "Capacity bounds via duality with applications to multi-antenna systems on flat fading channels," 2003, to app. in *IEEE Trans. Inform. Theory,* av. at <http://www.isi.ee.ethz.ch/~moser>.

[5] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* Academic Press, 1981.