

A Lower Bound for the Capacity of the Discrete-Time Poisson Channel

Alfonso Martinez
 Centrum Wiskunde & Informatica
 The Netherlands
 Email: alfonso.martinez@ieee.org

Abstract—A simple lower bound to the capacity of the discrete-time Poisson channel with average number of quanta of energy ε_s is derived. The rate $\frac{1}{2} \log(1 + \varepsilon_s)$ is shown to be the generalized mutual information of a modified minimum-distance decoder, when the input follows a gamma distribution of parameter $1/2$ and mean ε_s .

I. INTRODUCTION

Consider a memoryless discrete-time whose output Y is distributed according to a Poisson distribution of parameter X , the channel input. By construction, the output is a non-negative integer, and the input a non-negative real number. The channel transition probability $W(y|x)$ is thus given by

$$W(y|x) = e^{-x} \frac{x^y}{y!}. \quad (1)$$

This model, the discrete-time Poisson (DTP) channel, appears often in the analysis of optical communication channels. In this case, and without loss of generality, one can identify both input and output with an integer number of quanta of energy. In the following, we omit the words “quanta of”, and simply refer to energy.

Let $P_X(x)$ denote the probability density function of the channel input. We assume that the input energy is constrained, i. e. $E[X] \leq \varepsilon_s$, where $E[\cdot]$ denotes the expectation operator and ε_s is the average energy. Random variables are denoted by capital letters, and their realizations by small letters.

An exact formula for the capacity $C(\varepsilon_s)$ of the DTP channel is not known. Recently, Lapidoth and Moser [2], derived a lower bound for a Poisson channel with additive Poisson noise and a peak constraint, which particularized to our channel model gives

$$C(\varepsilon_s) \geq \log \left(\left(1 + \frac{1}{\varepsilon_s} \right)^{1+\varepsilon_s} \sqrt{\varepsilon_s} \right) - \left(1 + \sqrt{\frac{\pi}{24\varepsilon_s}} \right). \quad (2)$$

Observe that this bound becomes negative for vanishing ε_s . Capacity is given in nats and the logarithms are in base e .

A closed-form expression for the mutual information $I(X; Y)$ achieved by an input with a gamma distribution of

parameter ν was derived by Martinez in [3], namely

$$I(X; Y) = \int_0^1 \left(\varepsilon_s - \left(1 - \frac{\nu^\nu}{(\nu + \varepsilon_s(1-u))^\nu} \right) \frac{u^{\nu-1}}{1-u} \right) \frac{du}{\log u} + (\varepsilon_s + \nu) \log \frac{\varepsilon_s + \nu}{\nu} + \varepsilon_s (\psi(\nu + 1) - 1), \quad (3)$$

where $\psi(y)$ is Euler’s digamma function. For $\nu = 1/2$, numerical evaluation of the mutual information gives a rate which would seem to exceed $\frac{1}{2} \log(1 + \varepsilon_s)$ for all values of ε_s . In this paper, we prove that the rate $\frac{1}{2} \log(1 + \varepsilon_s)$ is indeed achievable by this input distribution. The analysis uses a suboptimum minimum-distance decoder, similar in spirit to Lapidoth’s analysis of nearest neighbor decoding [4].

II. MAIN RESULT

Let the input X follow a gamma distribution of parameter $1/2$ and mean ε_s , that is,

$$P_X(x) = \frac{1}{\sqrt{2\pi\varepsilon_s x}} e^{-\frac{x}{2\varepsilon_s}}. \quad (4)$$

This choice led to good lower and upper bounds in [2] and [3] respectively.

We consider a maximum-metric decoder; the codeword metric is given by the product of symbol metrics $q(x, y)$ over all channel uses. The optimum maximum-likelihood decoder, for which $q(x, y) = W(y|x)$, is somewhat unwieldy to analyze (Eq. (3) gives the exact mutual information). We consider instead a symbol decoding metric of the form

$$q(x, y) = e^{-ax - \frac{y^2}{x}}, \quad (5)$$

where $a = 1 + \frac{1}{\varepsilon_s}$. The reasons for this choice of a will be apparent later.

Clearly, the decoder is unchanged if we replace the symbol metric $q(x, y)$ by a symbol distance $d(x, y) = -\log q(x, y)$, and select the codeword with smallest total distance, summed over all channel uses. This alternative formulation is reminiscent of minimum-distance, or nearest-neighbor decoding. Indeed, the metric in Eq. (5) is equivalent to a minimum-distance decoder which uses the distance

$$d(x, y) = \frac{(y - \sqrt{ax})^2}{x} = \frac{y^2}{x} + ax - 2y\sqrt{a}. \quad (6)$$

The term $-2y\sqrt{a}$ is common to all symbols x and can be removed, since it does not affect the decision.

For $a = 1$, the distance in Eq. (6) naturally arises from a Gaussian approximation to the channel output, whereby the channel output is modeled as a Gaussian random variable of mean x and variance x . This approximation is suggested by the fact that a Poisson random variable of mean x approaches a Gaussian random variable of mean and variance x for large x .

Minimum-distance decoders were considered by Lapidoth [4] in his analysis of additive non-Gaussian-noise channels. For our channel model, even though noise is neither additive (it is signal-dependent), nor Gaussian, similar techniques to the ones used in [4] can be applied. More specifically, since we have a *mismatched decoder*, we determine the generalized mutual information [5]. For a given decoding metric $q(x, y)$ and a nonnegative number s , it can be proved [5] that the following rate—the generalized mutual information—is achievable

$$I_{\text{GMI}}(s) = \mathbb{E} \left[\log \frac{q(X, Y)^s}{\mathbb{E}[q(X', Y)^s]} \right]. \quad (7)$$

The expectation is carried out according to $P_X(x)W(y|x)$. This quantity is obviously a lower bound to the channel capacity.

Our main result is

Theorem 1. *In the discrete-time Poisson channel with average signal energy ε_s , the rate $\frac{1}{2} \log(1 + \varepsilon_s)$ is achievable.*

This rate is reminiscent of the capacity of a real-value Gaussian channel with average signal-to-noise ratio ε_s . Similarly to the situation in this channel, the rate is achieved by a form of minimum-distance decoding. Differently, the input follows a gamma distribution, rather than a Gaussian.

Proof: We evaluate the generalized mutual information $I_{\text{GMI}}(s)$ for an input distributed according to the gamma density, in Eq. (4). First, we evaluate the expectation in the denominator [6, Eq. 3.471-15]

$$\int_0^\infty \frac{e^{-\frac{x'}{2\varepsilon_s} - a s x' - \frac{s y^2}{x'}}}{\sqrt{2\pi\varepsilon_s x'}} dx' = \frac{e^{-|y|\sqrt{\frac{2s(1+2a\varepsilon_s s)}{\varepsilon_s}}}}{\sqrt{1+2a\varepsilon_s s}}. \quad (8)$$

Further, using the expression of the first two moments of the Poisson distribution, namely¹

$$\sum_y W(y|x)y = x, \quad \sum_y W(y|x)y^2 = x^2 + x, \quad (9)$$

together with the input constraint $\mathbb{E}[X] = \varepsilon_s$, we can explicitly

¹The moment generating function of a Poisson random variable of mean x is readily computed to be $e^{x(e^t-1)}$. The first two moments are the first two derivatives, evaluated at $t = 0$.

carry out the expectation in Eq. (7),

$$I_{\text{GMI}}(s) = \mathbb{E}[\log q(X, Y)^s] - \mathbb{E}[\log \mathbb{E}[q(X', Y)^s]] \quad (10)$$

$$= s \mathbb{E} \left[-aX - \frac{Y^2}{X} \right] + \mathbb{E} \left[Y \sqrt{\frac{2s(1+2a\varepsilon_s s)}{\varepsilon_s}} + \log \sqrt{1+2a\varepsilon_s s} \right] \quad (11)$$

$$= -s((a+1)\varepsilon_s + 1) + \sqrt{2\varepsilon_s s(1+2a\varepsilon_s s)} + \frac{1}{2} \log(1+2a\varepsilon_s s). \quad (12)$$

Choosing $\hat{s} = \frac{2\varepsilon_s}{(a-1)^2\varepsilon_s^2 + 2\varepsilon_s(a+1)+1}$, the first two summands cancel out. And for $a = 1 + \frac{1}{\varepsilon_s}$ we have that $2a\hat{s} = 1$, and therefore

$$I_{\text{GMI}}(\hat{s}) = \frac{1}{2} \log(1 + \varepsilon_s). \quad (13)$$

The same rate, $\frac{1}{2} \log(1 + \varepsilon_s)$, is also achievable by a decoder with $a = 1$. In this case, we have to replace the generalized mutual information by the alternative expression I_{LM} [5], given by

$$I_{\text{LM}} = \mathbb{E} \left[\log \frac{a(X)q(X, Y)^s}{\mathbb{E}[a(X')q(X', Y)^s]} \right]. \quad (14)$$

As for I_{GMI} , s is a non-negative number; $a(x)$ is a weighting function. Setting $a(x) = e^{-\frac{x}{\varepsilon_s}}$ we have that I_{LM} is given by Eq. (11), thus proving the achievability.

The bound provided in this paper is simpler and tighter than Eq. (2). It would be interesting to extend Theorem 1 to channel models $Y = S(X) + Z$, where $S(X)$ corresponds to the case considered here and Z is some additive noise Z , with a Poisson or a geometric distribution. A different input distribution and another modified decoding metric are likely required for either case.

REFERENCES

- [1] D. Brady and S. Verdú, "The asymptotic capacity of the direct detection photon channel with a bandwidth constraint," in *Proc. on the 28th Annual Allerton Conf. on Communication, Control, and Computing*, Allerton House, Monticello (USA), September 1990, pp. 691–700.
- [2] A. Lapidoth and S. M. Moser, "Bounds on the capacity of the discrete-time Poisson channel," in *Proceedings of the 41st Allerton Conf. on Communication, Control, and Computing*, October 2003.
- [3] A. Martinez, "Spectral efficiency of optical direct detection," *J. Opt. Soc. Am. B*, vol. 24, no. 4, pp. 739–749, April 2007.
- [4] A. Lapidoth, "Nearest neighbor decoding for additive non-gaussian noise channels," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1520–1529, September 1996.
- [5] A. Ganti, A. Lapidoth, and I. E. Telatar, "Mismatched decoding revisited: general alphabets, channels with memory, and the wide-band limit," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2315–2328, November 2000.
- [6] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, 6th ed., A. Jeffrey, Ed. Academic Press, 2000.