ISIT'2013 1569719473

Square Root Approximation to the Poisson Channel

Anagnostis Tsiatmas* †, Frans M. J. Willems*, Constant P. M. J. Baggen†
*Department of Electrical Engineering, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands
†Philips Research Eindhoven, High Tech Campus 34, 5656 AE Eindhoven, The Netherlands
Emails: a.tsiatmas@tue.nl, f.m.j.willems@tue.nl, stan.baggen@philips.com

***Abstract*—Starting from the Poisson model we present a channel model for optical communications, called the Square Root (SR) Channel, in which the noise is additive Gaussian with constant variance. Initially, we prove that for large peak or average power, the transmission rate of a Poisson Channel when coding and decoding methods for the SR Channel are used, converges to the capacity of the Poisson Channel. Then, we derive bounds and asymptotic expressions for the capacity of the SR Channel, and compare them to those of other optical channel models. Finally, signal-independent noise sources are discussed.**

I. Introduction

Many optical communication systems employ intensity modulation, in which non-negative input signals modulate the intensity of the transmitted wave. In case of direct-detection, a photodetector is used to convert linearly the optical intensity of the radiation incident on its surface into an output current signal. Such systems, called intensity-modulation/direct-detection (IM/DD) systems, are exploited, e.g., in optical fiber and free-space wireless communications when phase control of the optical wave is not possible, as it can be the case with LED sources.

Traditionally, IM/DD channels are modeled using Poisson processes, to account for the quantum nature of light. The channel output, i.e., the detected number of photons, is a r.v. which has a Poisson distribution with a parameter λ corresponding to the expected received intensity level. In this case, it is apparent that the uncertainty of the received signal depends on the signal level.

In recent years, free space optical intensity channels have been used to model short range infrared (IR) and Visible Light Communication (VLC) links [1]. In these models, noise is considered to be additive Gaussian and signal independent, since it mainly originates from light sources (daylight, fluorescent lamps, etc.) or from the electronics in the receiver. While this description is adequate in cases where the signal intensity is low compared to background radiation or thermal noise, as can be the case in infrared wireless links, this is not true when signal-dependent noise is dominant. This can be the case for VLC with high-brightness LEDs, where the randomness of the signal itself cannot be neglected [2].

In this paper, we introduce the Square Root (SR) Channel model which is an Additive White Gaussian Noise (AWGN) Channel and can account for the randomness of the signal, the ambient-light noise, and the thermal noise. The noise in this model is signal-independent. For the SR Channel model we investigate how well it fits the Poisson Channel model in capacity and mismatch capacity settings.

II. Channel Models

We consider four different models for the IM/DD channel. Without loss of generality, we assume that we signal in disjoint temporal intervals of unit duration, with the expected number of received photons $\lambda_i \geq 0$ in the i-th interval, $i \in \mathbb{N}$. The first and most accurate model is derived from the photon-counting and, thus, Poisson nature of the channel.

Definition 1: The memoryless discrete-time Poisson Channel has as input the r.v. $\Lambda \geq 0$, and as output the discrete r.v. X which is drawn from a Poisson distribution with a parameter $\Lambda + \lambda_0$ ($X \sim \mathcal{P}(\Lambda + \lambda_0)$). The non-negative term λ_0 is a constant related to the ambient light or to the random generation of electrons at the output of the photodetector [3]. The conditional output probability of this channel is:

$$w_{\mathrm{P}}(x|\lambda) = e^{-(\lambda+\lambda_0)} \frac{(\lambda+\lambda_0)^x}{x!}, \; x \in \mathbb{N}, \text{ and } \lambda \geq 0. \quad (1)$$

In this case, the variance of the received signal has a component that is linearly dependent on the input signal value λ and a signal-independent component that is modeled with the parameter λ_0. Using the central limit theorem, it can be shown that the distribution of an r.v. $X \sim \mathcal{P}(\lambda)$ for large values of λ approaches a Gaussian distribution $\mathcal{N}(\lambda, \lambda)$.

Definition 2: The memoryless discrete-time Optical Intensity (OI) Channel with input-dependent Gaussian noise has as input the r.v. $\Lambda \geq 0$, and as output the r.v. Y, which is the sum of the input r.v. Λ and two independent r.v.'s $N_1 \sim \mathcal{N}(0, \Lambda)$ and $N_2 \sim \mathcal{N}(\lambda_0, \lambda_0)$. Therefore, the conditional pdf of this channel is:

$$w_{\mathrm{G}}(y|\lambda) = \frac{1}{\sqrt{2\pi(\lambda+\lambda_0)}} e^{-\frac{(y-(\lambda+\lambda_0))^2}{2(\lambda+\lambda_0)}}, \; y \in \mathbb{R}, \text{ and } \lambda \geq 0. \quad (2)$$

While the OI Channel is closer to our understanding of AWGN channels than the Poisson Channel, the noise dependence on the signal level still complicates the signal design. A description with fixed noise variance, would allow us to use the widely available knowledge of coding and decoding methods for AWGN channels and enhance our understanding of the IM/DD channel behaviour. In such a description, e.g., the corresponding metrics can be shown to depend only on the distance between transmitted and received signals.

Definition 3: The memoryless discrete-time Square Root (SR) Channel (see Fig. 1) has as input the r.v. $\Lambda \geq 0$, and as output the r.v. Y, which is the sum of $\sqrt{\Lambda + \lambda_0}$ and a

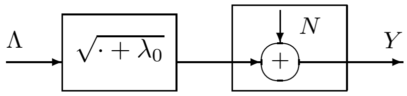

Fig. 1: The SR Channel.

r.v. $N \sim \mathcal{N}(0, 1/4)$. In this case, the conditional pdf of the channel is:

$$w_{\text{SR}}(y|\lambda) = \sqrt{\frac{2}{\pi}} e^{-2\left(y - \sqrt{\lambda + \lambda_0}\right)^2}, \; y \in \mathbb{R}, \text{ and } \lambda \geq 0. \quad (3)$$

The SR Channel is inspired by the square-root variance-stabilisation transform [4]. Curtiss showed that for a r.v. $X \sim \mathcal{P}(\lambda)$ and any constant $a \geq 0$ the distribution of the transformed r.v. $\sqrt{X + a}$ has variance that is asymptotically equal to $1/4$ as $\lambda \to \infty$.

The square-root transform used in the SR Channel, provides an additional advantage in the description of IM/DD channels. The photon-counting device acts as a squarer and produces a current signal proportional to the intensity of the optical wave. This squaring behaviour leads to a discrepancy between the optical and electrical power. The square-root transform reverses the behaviour of the detector, allowing the power in the optical and electrical domains to match.

To allow the comparison between the discrete output Poisson Channel and the continuous output SR Channel, we introduce the Modified Continuous Output Poisson Channel.

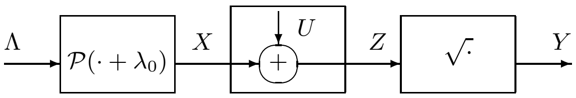

Fig. 2: The MP Channel.

Definition 4: The memoryless discrete-time Modified Continuous Output Poisson (MP) Channel (Fig. 2) has as input the r.v. $\Lambda \geq 0$ and as output the r.v. $Y = \sqrt{Z} = \sqrt{X + U}$, where the continuous r.v. U is drawn from a uniform distribution on $[0, 1)$. The conditional pdf of this channel is:

$$w_{\text{MP}}(y|\lambda) = 2ye^{-(\lambda+\lambda_0)} \frac{(\lambda + \lambda_0)^{\lfloor y^2 \rfloor}}{\lfloor y^2 \rfloor !}, \; y \geq 0, \text{ and } \lambda \geq 0. \quad (4)$$

Lemma 1: The mutual information of the Poisson Channel $I(\Lambda; X)$ is equal to the one of the MP Channel $I(\Lambda; Y)$.

Proof: We define $T = \lfloor Y^2 \rfloor = X$. Then, by applying the data processing inequality we have:

$$I(\Lambda; X) \geq I(\Lambda; Y) \geq I(\Lambda; T) = I(\Lambda; X). \quad (5)$$

■

Eq. (5) implies that the capacity of the Poisson Channel is equal to the capacity of the MP Channel and is achieved by the same input distribution.

III. Mismatch Capacity

Our aim is to examine the induced loss in capacity of the Poisson Channel when we decode the received messages using the optimum rule for the SR Channel. This problem falls in the field of mismatch capacity estimation [5].

Proposition 1: For the same input $\Lambda = \lambda$ and as $\lambda \to \infty$, the output distribution of the MP Channel will converge to the output distribution of the SR Channel.

Proof: We follow a method similar to [2]. Initially, we consider $\lambda_0 = 0$ and show that for $\lambda \to \infty$ the transformed variable $Y = \sqrt{Z}$ with pdf:

$$w_{\text{MP}}(y|\lambda) = 2ye^{-\lambda} \frac{\lambda^{\lfloor y^2 \rfloor}}{\lfloor y^2 \rfloor !}, \quad y \geq 0, \quad (6)$$

converges to the pdf of a Gaussian distribution with mean value equal to $\sqrt{\lambda}$ and variance equal to $1/4$:

$$w_{\text{SR}}(y|\lambda) = \sqrt{\frac{2}{\pi}} e^{-2\left(y - \sqrt{\lambda}\right)^2}, \quad y \in \mathbb{R}. \quad (7)$$

The information divergence between these distributions is:

$$\begin{aligned} D(w_{\text{MP}} \| w_{\text{SR}}) =& E\left[\log \frac{w_{\text{MP}}(Y|\lambda)}{w_{\text{SR}}(Y|\lambda)}\right] \\ =& E\left[\frac{1}{2}\log(2\pi Z)\right] - \lambda + E\left[\lfloor Z \rfloor\right]\log(\lambda) \\ & - E\left[\log\left(\lfloor Z \rfloor !\right)\right] + 2E\left[\left(\sqrt{Z} - \sqrt{\lambda}\right)^2\right]. \quad (8) \end{aligned}$$

The three middle terms on the right-hand side of Eq. (8) are equal to $-H_{\text{P}}$, where H_{P} is the entropy of a Poisson distribution. From [6] we have that asymptotically for large values of λ:

$$H_{\text{P}} = \frac{1}{2}\log\left(2\pi e \lambda\right) - \frac{1}{12\lambda} + O\left(\frac{1}{\lambda^2}\right). \quad (9)$$

In addition, for the last term of Eq. (8) we can write:

$$2E\left[\left(\sqrt{Z} - \sqrt{\lambda}\right)^2\right] = 4\lambda + 1 - 4\sqrt{\lambda}E\left[\sqrt{Z}\right]. \quad (10)$$

If $V = 1/2 + Z - E[Z] = Z - \lambda$ and $\lambda > 0$, we obtain:

$$\begin{aligned} Y =& \sqrt{V + \lambda} = \sqrt{\lambda} \sum_{n=0}^{+\infty} \frac{(-1)^n (-\frac{1}{2})_n}{n!} \frac{(V)^n}{\lambda^n} \\ =& \sqrt{\lambda}\left(1 + \frac{1}{2}\frac{V}{\lambda} - \frac{1}{8}\frac{V^2}{\lambda^2} + \frac{1}{16}\frac{V^3}{\lambda^3} - \frac{5}{128}\frac{V^4}{\lambda^4} + \ldots\right), \quad (11) \end{aligned}$$

where $\left(-\frac{1}{2}\right)_n$ is the Pochhammer symbol. Eq. (11) holds for all $V \geq -\lambda$, therefore:

$$E\left[\sqrt{Z}\right] = \sqrt{\lambda}\left(1 + \frac{1}{2}\frac{\mu_1}{\lambda} - \frac{1}{8}\frac{\mu_2}{\lambda^2} + \frac{1}{16}\frac{\mu_3}{\lambda^3} - \frac{5}{128}\frac{\mu_4}{\lambda^4} + \ldots\right), \quad (12)$$

where μ_n, $n = 1, 2, 3, \ldots$ are the moments of V. It can be shown, by using the definition of the moments for V, that $\mu_n = \frac{1}{n+1}\sum_{i=1}^{n+1}\binom{n+1}{i}\tilde{\mu}_{n+1-i}$, with $\tilde{\mu}_n$ the n-th central moment of $X \sim \mathcal{P}(\lambda)$. For $\tilde{\mu}_n$ it is easy to show that $\tilde{\mu}_{n+1} = \lambda \sum_{k=0}^{n-1}\binom{n}{k}\tilde{\mu}_k$. Therefore, $\mu_1 = 1/2$, $\mu_2 = \lambda + 1/3$, $\mu_3 = 5\lambda/2 + 1/4$, and $\mu_4 = 3\lambda^2 + 5\lambda + 1/5$, while all terms μ_n/λ^n, for $n > 4$ are approaching zero faster than $1/\lambda^2$. Therefore:

$$E\left[\sqrt{Z}\right] = \sqrt{\lambda}\left(1 + \frac{1}{8\lambda} - \frac{1}{384\lambda^2} + O(\frac{1}{\lambda^3})\right), \quad (13)$$

and Eq. (10) can be written as:

$$2E\left[\left(\sqrt{Z}-\sqrt{\lambda}\right)^2\right]=\frac{1}{2}+\frac{1}{96\lambda}+O\left(\frac{1}{\lambda^2}\right). \quad (14)$$

Substituting Eq. (9) and Eq. (14) into Eq. (8), and using Jensen's inequality for the first term of the latter, we obtain:

$$\begin{aligned}D(w_{\mathrm{MP}}\|w_{\mathrm{SR}}) &\leq \frac{1}{2}\log\left(1+\frac{1}{2\lambda}\right)+\frac{3}{32\lambda}+O\left(\frac{1}{\lambda^2}\right)\\ &\leq \frac{1}{4\lambda}+\frac{3}{32\lambda}+O\left(\frac{1}{\lambda^2}\right)\\ &= \frac{11}{32\lambda}+O\left(\frac{1}{\lambda^2}\right). \end{aligned} \quad (15)$$

Since the divergence is always non-negative, as λ increases, $D(w_{\mathrm{MP}}\|w_{\mathrm{SR}})\to 0$ and the two distributions become identical. In case $\lambda_0>0$, the same derivation can be applied to show that the distribution of Y approaches $\mathcal{N}(\sqrt{\lambda+\lambda_0},1/4)$ as $\lambda\to\infty$, since the sum of two independent Poisson variables is a Poisson variable with parameter equal to the sum of their parameters. ■

Proposition 2: For all $\lambda\geq 0$ and $\lambda_0\geq 0$ we have $D(w_{\mathrm{MP}}\|w_{\mathrm{SR}})\leq 1/2\log(2\pi(\lambda+\lambda_0+1/2))+4(\lambda+\lambda_0)+1$.

Proof: This follows immediately from Eq. (8) and the inequalities $\log(X!)\geq X\log(X)-X$ and $E\left[\sqrt{Z}\right]\geq 0$. ■

We now investigate the mismatch capacity C_{M} of the MP Channel when the SR Channel decoding rule is applied. While C_{M} is not known in general, we consider its lower bounds [5]:

$$I_{\mathrm{GMI}}=\int_{\Lambda\times Y}p(\lambda)w_{\mathrm{MP}}(y|\lambda)\log\frac{w^s_{\mathrm{SR}}(y|\lambda)}{\int_\Lambda p(\lambda')w^s_{\mathrm{SR}}(y|\lambda')\mathrm{d}\lambda'}\mathrm{d}y\mathrm{d}\lambda, \quad (16)$$

where $p(\lambda)$ is the pdf of the input distribution for the MP Channel, $w_{\mathrm{MP}}(y|\lambda)$ is the conditional pdf of the MP Channel, $w_{\mathrm{SR}}(y|\lambda)$ is induced by the decoding rule for the SR Channel, and $s\geq 0$ is a free parameter. For $s=1$ we obtain I_{LB}, which is:

$$\begin{aligned}I_{\mathrm{LB}}=&I(\Lambda;Y)-\int_{\Lambda\times Y}p(\lambda)w_{\mathrm{MP}}(y|\lambda)\log\frac{w_{\mathrm{MP}}(y|\lambda)}{w_{\mathrm{SR}}(y|\lambda)}\mathrm{d}y\mathrm{d}\lambda\\ &+\int_{\Lambda\times Y}p(\lambda)w_{\mathrm{MP}}(y|\lambda)\log\frac{\int_\Lambda p(\lambda')w_{\mathrm{MP}}(y|\lambda')\mathrm{d}\lambda'}{\int_\Lambda p(\lambda')w_{\mathrm{SR}}(y|\lambda')\mathrm{d}\lambda'}\mathrm{d}y\mathrm{d}\lambda,\end{aligned} \quad (17)$$

with $I(\Lambda;Y)$ the mutual information of the MP channel. Since the third term on the right-hand side of Eq. (17) is an information divergence, and, thus, is always non-negative, we can write:

$$I(\Lambda;Y)-I_{\mathrm{LB}}\leq\int_\Lambda p(\lambda)D\left(w_{\mathrm{MP}}(Y|\lambda)\|w_{\mathrm{SR}}(Y|\lambda)\right)\mathrm{d}\lambda. \quad (18)$$

From [7] it is known that for the Poisson Channel with a peak power constraint ($\Lambda\leq A$) an input distribution with pdf $p(\lambda)=\frac{1}{2\sqrt{A\lambda}}$ for $\lambda\in[0,A]$ is capacity-achieving as $A\to\infty$. Similarly, in case of an average power constraint ($E[\Lambda]\leq\mathcal{E}$), a distribution with pdf $p(\lambda)=\frac{1}{\sqrt{2\pi\mathcal{E}\lambda}}\exp\left(-\frac{\lambda}{2\mathcal{E}}\right)$ for $\lambda\geq 0$ is capacity-achieving as $\mathcal{E}\to\infty$. These results hold $\forall\lambda_0\geq 0$.

Theorem 1. *As A or $\mathcal{E}\to\infty$ the mismatch capacity C_{M} of the Poisson Channel using the SR Channel decoding rule is converging to the actual capacity C_{P} of the Poisson Channel.*

Proof: From *Lemma 1* we know that the capacity of the MP Channel is equal to the capacity of the Poisson Channel and achieved by the same input distribution. We first consider the capacity achieving input distribution under a peak power constraint for $A\to\infty$. For this input distribution, there is a $\lambda^*\geq 0$ such that for $\lambda\geq\lambda^*$ the information divergence $D(w_{\mathrm{MP}}\|w_{\mathrm{SR}})\leq 1/(\lambda+\lambda_0)$, as implied by Eq. (15) of *Proposition 1*. Therefore, using Eq. (18), we have:

$$\begin{aligned}0\leq C_{\mathrm{P}}-I_{\mathrm{LB}}\leq&\int_0^{\lambda^*}p(\lambda)D\left(w_{\mathrm{MP}}(Y|\lambda)\|w_{\mathrm{SR}}(Y|\lambda)\right)\mathrm{d}\lambda\\ &+\int_{\lambda^*}^{A}p(\lambda)\frac{1}{\lambda+\lambda_0}\mathrm{d}\lambda\triangleq C_1+C_2.\end{aligned} \quad (19)$$

This difference is lower-bounded by 0, since the mismatch capacity cannot be larger than the actual capacity of the channel. Furthermore, we upper-bound C_1 using *Proposition 2* and C_2 by extending the upper-bound to infinity. After a few calculations, it can be shown that $C_1\leq\frac{g_1(\lambda^*,\lambda_0)}{\sqrt{A}}$, and $C_2\leq\frac{g_2(\lambda^*,\lambda_0)}{\sqrt{A}}$, where g_1, g_2 are functions only of λ^* and λ_0. Therefore $I_{\mathrm{LB}}\to C_{\mathrm{P}}$, or $C_{\mathrm{M}}\to C_{\mathrm{P}}$ as $A\to+\infty$. Similar results emerge in the case of an average power constraint. In this case, for the asymptotic capacity achieving input distribution as $\mathcal{E}\to\infty$ we have $p(\lambda)=\frac{1}{\sqrt{2\pi\mathcal{E}\lambda}}\exp\left(\frac{-\lambda}{2\mathcal{E}}\right)\leq\frac{1}{\sqrt{2\pi\mathcal{E}\lambda}}$, $\forall\lambda>0$. Therefore, the upper-bounds of C_1 and C_2 are analogous to the derived ones for the peak power constraint case, and $C_{\mathrm{M}}\to C_{\mathrm{P}}$ for $\mathcal{E}\to+\infty$. ■

Similar results appear in the context of cutoff-rate [8] and asymptotic bit error rate behaviour in binary transmission [9].

Corollary 1: The result of *Theorem 1* implies that asymptotically for large values of A or $\mathcal{E}$ we can use decoding techniques developed for the SR Channel, which is just an AWGN Channel, for a Poisson Channel without inducing any loss in information transmission rate of the latter. Note that decoding techniques for AWGN channels are well understood.

IV. Capacity Results for the SR Channel

We define $S\triangleq\sqrt{\Lambda+\lambda_0}$ and we can then rewrite the SR Channel equation as follows:

$$Y=S+N, \quad (20)$$

where $N\sim\mathcal{N}(0,1/4)$ and S is restricted such that $S\geq\sqrt{\lambda_0}$.

The capacity of the SR Channel with the restriction that $S\geq\sqrt{\lambda_0}$ and additionally (a) a peak power constraint ($S^2\leq A+\lambda_0$) or (b) an average power constraint ($E\left[S^2\right]\leq\mathcal{E}+\lambda_0$) is the supremum among all possible input distributions of the mutual information between channel input S and output Y,

$$C=\sup_{P_S(s)}I(S;Y), \quad (21)$$

where the optimisation is subject to the stated constraints.

Since a closed form expression for the capacity of the SR Channel in the cases of a peak or an average power constraint is not known at present, we derive, here, analytic upper and lower bounds on the channel's capacity for each of these cases. The derived bounds for the SR Channel are valid both for low and high power regions and are compared with bounds for the OI and Poisson Channels.

A. Peak Power Constraint

In the case of a peak power constraint $S^2 \leq A + \lambda_0$ in addition to the $S \geq \sqrt{\lambda_0}$ constraint, we know from [10] that the capacity achieving distribution is discrete with a finite number of points. Nevertheless, this has no influence on our derivations.

In order to derive a lower bound, we consider an input distribution uniform on the interval $L = [\sqrt{\lambda_0}, \sqrt{A+\lambda_0}] = [S_{\min}, S_{\max}]$. We define l as the length of L. Therefore, by using the entropy power inequality (EPI) [11], we have:

$$\begin{aligned} C(A) &\geq I(S;Y) = h(Y) - h(Y|S) = h(S+N) - h(N) \\ &\geq \frac{1}{2}\log\left(e^{2h(S)} + e^{2h(N)}\right) - \frac{1}{2}\log\left(e^{2h(N)}\right) \\ &= \frac{1}{2}\log\left(1 + \frac{2l^2}{\pi e}\right) = C_{\mathrm{LB}}(A), \end{aligned} \tag{22}$$

since the differential entropy of the input distribution is $h(S) = \log(l)$ and the power of the noise $\sigma^2 = 1/4$.

Since $I(S;Y) = h(Y) - h(Y|S) = h(Y) - h(N) = h(Y) - \frac{1}{2}\log\left(\frac{\pi e}{2}\right)$, for deriving an upper bound we have to determine the maximum-entropy output distribution. For $y \leq S_{\min}$ and $y \geq S_{\max}$ the output distribution has average power equal to half the power of the noise at most, while there is no constraint for $y \in L$. It can be shown that the maximum-entropy output distribution [11] under these constraints is of the form:

$$f(y) = \begin{cases} \frac{\epsilon}{L} & \text{if } y \in L, \\ \sqrt{\frac{2(1-\epsilon)^3}{\pi}}\exp\left(\frac{-(y-S_{\max})^2}{1/[2(1-\epsilon)]}\right) & \text{if } y > S_{\max}, \\ \sqrt{\frac{2(1-\epsilon)^3}{\pi}}\exp\left(\frac{-(y-S_{\min})^2}{1/[2(1-\epsilon)]}\right) & \text{if } y < S_{\min}, \end{cases} \tag{23}$$

with ϵ the portion of output probability in L. The output distribution achieves the maximum entropy when $(1-\epsilon)^{3/2} = \sqrt{\pi/2}\epsilon/l$, for which f turns out to be continuous. In this case, the derived upper bound on the capacity is: $C_{\mathrm{UB}}(A) = \epsilon\log(l) + \frac{1-\epsilon}{2}\log\left(\frac{\pi e}{2(1-\epsilon)}\right) + H_2(\epsilon) - \frac{1}{2}\log\left(\frac{\pi e}{2}\right)$.

Asymptotic Expression for the Capacity: As $A \to \infty$ we have that $l \to \sqrt{A} \to \infty$ and $\epsilon \to 1$. In this case, both $C_{\mathrm{LB}}(A)$ and $C_{\mathrm{UB}}(A)$ approach $\frac{1}{2}\log(A) - \frac{1}{2}\log\left(\frac{\pi e}{2}\right)$, which is the asymptotic expression for the channel capacity achieved when the distribution of the channel input S tends towards a uniform distribution on L.

Comment: For $A \to \infty$, the capacity-achieving uniform input distribution in the S-domain for the SR Channel asymptotically matches the capacity-achieving input distribution for the Poisson Channel in the Λ-domain. This can be shown by using the r.v. transformation: $\Lambda = S^2 - \lambda_0$.

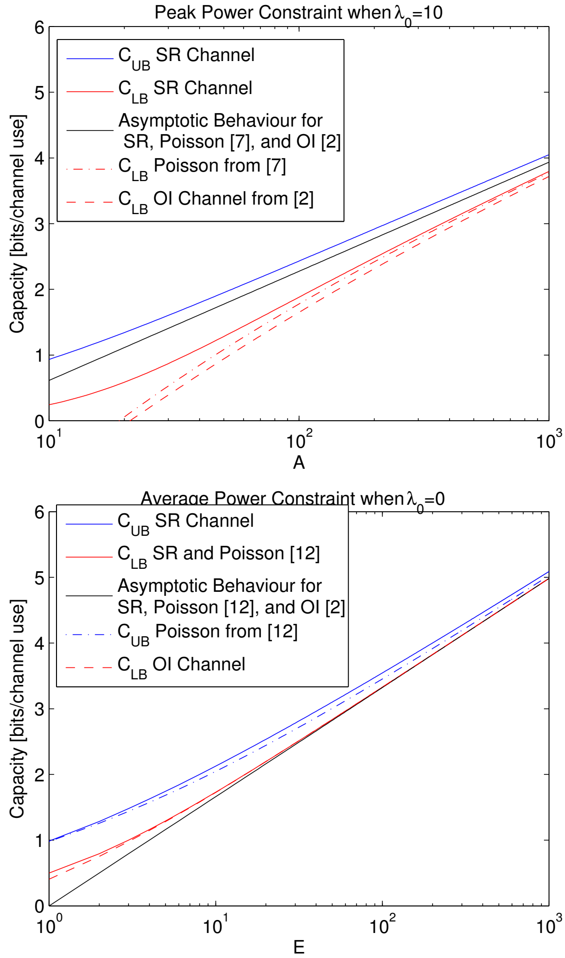

Fig. 3: Bounds for the capacity of the SR, OI and Poisson Channels under (a) a peak and (b) an average power constraint.

B. Average Power Constraint

We now study the case of an average power constraint $E\left[S^2\right] \leq \mathcal{E} + \lambda_0$ in addition to the $S \geq \sqrt{\lambda_0}$ constraint, when $\lambda_0 = 0$.

For the lower bound we consider an input that corresponds to a single-sided Gaussian distribution of the form $p(s) = \sqrt{\frac{2}{\pi\mathcal{E}}}\exp\left(-\frac{s^2}{2\mathcal{E}}\right)$, $s \geq 0$. In this case, by using EPI, we derive the lower bound: $C_{\mathrm{LB}}(\mathcal{E}) = \frac{1}{2}\log(1+\mathcal{E})$. Note that the differential entropy of a single-sided Gaussian is $h(S) = \frac{1}{2}\log\left(\frac{\pi e\mathcal{E}}{2}\right)$. Interestingly, $C_{\mathrm{LB}}(\mathcal{E})$ is also the best known lower bound on the actual capacity of the Poisson Channel [12].

An upper bound can be determined again similar to the case of a peak power constraint. When $y < 0$, the average power of the output distribution is equal to half the power of the noise at most (i.e., $E\left[Y^2\right] \leq 1/8$), while in the case that $y \geq 0$ we have $E\left[Y^2\right] \leq \mathcal{E} + 1/4$. Therefore, the maximum-entropy distribution for the output Y is of the form:

$$f(y) = \begin{cases} \sqrt{\frac{2\epsilon^3}{\pi(\mathcal{E}+1/4)}}\exp\left(-\frac{y^2}{2[\mathcal{E}+1/4]/\epsilon}\right) & \text{if } y \geq 0, \\ \sqrt{\frac{16(1-\epsilon)^3}{\pi}}\exp\left(-\frac{y^2}{1/[4(1-\epsilon)]}\right) & \text{if } y < 0. \end{cases} \tag{24}$$

In this case, ϵ is the portion of the probability in the non-negative y region and the maximum entropy is achieved when $\epsilon = 1/\left(1+(\frac{1}{8\mathcal{E}+2})^{\frac{1}{3}}\right)$, which makes f continuous. This value of ϵ leads to the upper bound: $C_{\text{UB}}(\mathcal{E}) = \frac{\epsilon}{2}\log\left(\frac{\pi e(\mathcal{E}+1/4)}{2\epsilon}\right) + \frac{1-\epsilon}{2}\log\left(\frac{\pi e}{16(1-\epsilon)}\right) + H_2(\epsilon) - \frac{1}{2}\log\left(\frac{\pi e}{2}\right)$.

Asymptotic Expression for the Capacity: As the value of $\mathcal{E}\to\infty$ we have $\epsilon\to 1$, and $C_{\text{UB}}(\mathcal{E}) \to \frac{1}{2}\log\left(\frac{\pi e(\mathcal{E}+1/4)}{2}\right) - \frac{1}{2}\log\left(\frac{\pi e}{2}\right) \to \frac{1}{2}\log(\mathcal{E})$. Simultaneously, as $\mathcal{E}\to\infty$, we get $C_{\text{LB}}(\mathcal{E}) \to \frac{1}{2}\log(\mathcal{E})$, which means that this is asymptotically the capacity of the SR Channel under an average power constraint which is achieved by a single-sided Gaussian distribution.

Comment: For $\mathcal{E}\to\infty$ the capacity-achieving single-sided Gaussian input distribution matches the capacity-achieving input distribution for the Poisson Channel. This can be shown by applying the r.v. transformation: $\Lambda = S^2$.

In the examined cases, the matching between the capacity-achieving input distributions of the SR and the Poisson Channels as A or $\mathcal{E}\to\infty$, proves that, in this region, coding methods for the SR Channel could be used for the Poisson Channel without inducing any loss in capacity.

Proposition 3: As $\mathcal{E}\to\infty$, the OI Channel with $\lambda_0 = 0$ has a lower bound $C_{\text{LB}} = \frac{1}{2}\log(\mathcal{E}) + \frac{1}{2}\log\left(1+\frac{2}{\mathcal{E}}\right) - (\mathcal{E}+1) + \sqrt{\mathcal{E}(\mathcal{E}+2)}$ and an asymptotic upper-bound $C_{\text{UB}} = \frac{1}{2}\log(\mathcal{E})$.

Proof: Eq. (36) in [2] holds also without signal-independent noise, as is proved by Eq. (138) and Appendix A there. Following the derivation of Eq. (23)–(24), without signal-independent noise, and taking $\sigma^2 = 1/\varsigma^2$, we can immediately derive the proposed expressions for C_{LB} and C_{UB}. ■

Fig. 3 illustrates the derived bounds under (a) a peak power constraint with $\lambda_0 = 10$ and (b) an average power constraint with $\lambda_0 = 0$, together with bounds for the Poisson Channel [7], [12], and for the OI Channel [2]. The upper bounds from [7] and [2] are asymptotic only. The difference between the upper and lower bounds on the capacity for the SR Channel is small even for low values of A and $\mathcal{E}$. Note that in most practical systems A or $\mathcal{E}$ values are much higher than those illustrated.

V. Discussion

We now discuss the effect of signal-independent noise contributions in the SR Channel description. Apart from the ambient light noise, there is a thermal noise contribution $N_{\text{th}} \sim \mathcal{N}(0,\sigma^2_{\text{th}})$ due to the electronics in the receiver. It can be shown that in this case the channel can be approximated as $Y = \sqrt{\Lambda+\lambda_0+\sigma^2_{\text{th}}} + N$, where $N\sim\mathcal{N}(0,1/4)$. Adding a bias term of σ^2_{th} and observing that for large values of σ^2_{th} the $\mathcal{N}(\sigma^2_{\text{th}},\sigma^2_{\text{th}})$ converges to $\mathcal{P}(\sigma^2_{\text{th}})$ establishes this result. In this context, the impairment to the signal by the thermal noise power can be seen as the result of a virtual amount of photons of ambient light received by the detector. This allows a unified description of signal-independent noise in the additive parameter λ_0 under the square root. In the S-domain, increasing λ_0 results in a down-scaling of the signal space due to the square root behaviour. The square-root transforms the signal-dependent Poisson noise to an an additive Gaussian noise term of constant variance. Additional ambient-light and thermal noise do not change the variance of the AWGN term, but reduce the distances between the signals, which leads the effective signal to noise ratio to decrease. For instance, it can be shown for the signals $s_1 = \sqrt{\lambda_1+\lambda_0}$ and $s_2 = \sqrt{\lambda_2+\lambda_0}$, when $\lambda_0 \gg \lambda_1$, λ_2 with $\lambda_1 > \lambda_2$, that their distance d_{12} satisfies:

$$d_{12} = s_1 - s_2 \leq \sqrt{\lambda_0}\left(\frac{\lambda_1-\lambda_2}{2\lambda_0}\right) = d^0_{12}\frac{\sqrt{\lambda_1}+\sqrt{\lambda_2}}{2\sqrt{\lambda_0}}, \quad (25)$$

where d^0_{12} is the distance of these signals for $\lambda_0 = 0$. Observe that d_{12} decreases when λ_0 increases.

VI. Conclusion

In this paper, we have introduced the Square Root Channel, which is an Additive White Gaussian Noise Channel, to model optical IM/DD links. We have shown that for large peak or average powers, coding and decoding methods for the SR Channel when used for the Poisson Channel reach capacity. We have also estimated upper and lower bounds on the capacity of the SR Channel under a peak or average power constraint. For practical values of the optical power these upper and lower bounds are actually close. Finally, we investigated the influence of signal-independent noise sources.

Acknowledgment

The authors would like to thank Ronald Rietman, Ludo Tolhuizen and the reviewers for their helpful remarks.

References

[1] J. M. Kahn, and J. R. Barry, "Wireless Infrared Communications," *Proc. IEEE*, vol. 85, no. 2, pp. 265–298, Feb. 1997.
[2] S. M. Moser, "Capacity Results of an Optical Intensity Channel with Input-Dependent Gaussian Noise," *IEEE Trans. Inform. Theory*, vol. 58, no. 1, pp. 207–223, Jan. 2012.
[3] I. Bar-David, "Communication under the Poisson regime," *IEEE Trans. Inform. Theory*, vol. 15, no. 1, pp. 31–37, Jan. 1969.
[4] J. H. Curtiss, "On Transformations Used in the Analysis of Variance," *Annal. Math. Statist.*, vol. 14, no. 2, pp. 107–122, Jun. 1943.
[5] A. Ganti, A. Lapidoth, and I. E. Telatar, "Mismatched Decoding Revisited: General Alphabets, Channels with Memory, and the Wide-Band Limit," *IEEE Trans. Inform. Theory*, vol. 46, no. 7, pp. 2315–2328, Nov. 2000.
[6] R. J. Evans, J. Boersma, N. M. Blachman, and A. A. Jagers, "The entropy of a Poisson distribution: Problem 87-6," *SIAM Rev.*, vol. 30, no. 2, pp. 314–317, Jun. 1988.
[7] A. Lapidoth, and S. M. Moser, "On the Capacity of the Discrete-Time Poisson Channel," *IEEE Trans. Inform. Theory*, vol. 55, no. 1, pp. 303–322, Jan. 2009.
[8] D. L. Snyder, and I. B. Rhodes, "Some Implications of the Cutoff-Rate Criterion for Coded Direct-Detection Optical Communication Systems," *IEEE Trans. Inform. Theory*, vol.it-46, no. 3, pp. 327–338, May 1980.
[9] S. Verdú, "Asymptotic error probability of binary hypothesis testing for Poisson point-process observations," *IEEE Trans. Inform. Theory*, vol. 32, no. 1, pp. 113–115, Jan. 1986.
[10] J. G. Smith, "The Information Capacity of Amplitude- and Variance-Constrained Scalar Gaussian Channels," *Information and Control*, vol. 18, pp. 203–219, Feb. 1971.
[11] T. M. Cover, and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.
[12] A. Martinez, "Spectral efficiency of optical direct detection," *J. Opt. Soc. Am. B* vol. 24, no. 4, pp. 739–749, Apr. 2007.