The Poisson Optical Communication Channels: Capacity and Optimal Power Allocation

Samah A. M. Ghanem, *Member, IAENG*, Munnujahan Ara, *Member, IAENG*

***Abstract*— In this paper, the channel capacity for different models of Poisson optical communication channels has been derived. The closed form expression for the single input single output (SISO) Poisson channel -derived by Kabanov in 1978, and Davis in 1980- will be investigated. In addition, we derive closed form expressions for the capacity of the parallel Poisson channel and the capacity of the multiple access Poisson Channel (MAC) under the assumption of constant shot noise. The optimum power allocation is also derived for the different models; results have been analyzed in the context of information theory and optical communications. We then provide a set of simulation results establishing a comparison between Gaussian channels and Poisson Channels.**

***Index Terms*—Gaussian channels, MAC, Parallel Channels, Poisson Channels, Power Allocation, SISO.**

I. Introduction

INFORMATION THEORY provides one of its strongest developments via the notion of maximum bit rate or channel capacity. If the capacity can be found, then the goal of the engineer is to design an architecture which achieves that capacity. The seminal work of Shannon published in 1948 [1] gave birth to information theory. Shannon determined the capacity of memoryless channels, including channels impaired by additive white Gaussian noise (AWGN) for a given signal-to-noise ratio (SNR). Currently, the majority of worldwide data and voice traffic is transported using optical communication channels. As the demand for bandwidth continues to increase, it is of great importance to find closed form expressions of the information capacity for the optical communications applications at the backbone as well as the access networks. In particular, the capacity expressions of Poisson channels that model the application. However, applying concepts of information theory to the optical communications channels encounters major challenges. The most important difficulty is dealing with the simultaneous interaction of three phenomena in the optical channel: noise, filtering, and Kerr nonlinearity. These three phenomena are distributed along the propagation path, and influence each other leading to deterministic as well as stochastic impairments [2].

Therefore, in this paper, we accomplish an information-theoretic approach to derive the closed form expressions for: the SISO Poisson channel already found by Kabanov [3] and Davis [4], the parallel multiple input multiple output (MIMO) Poisson channels, as well as for the MAC Poisson channel using a direct detection or photon counting receiver and under constant noise; therefore, we simplify the framework of derivation. Several contributions have been done using information theoretic approaches to derive the capacity of Poisson channels under constant and time varying noise via martingale processes [3-7], or via approximations using Bernoulli processes [8], to define upper and lower bounds for the capacity and the rate regions of different models [9-10], to define relations between information measures and estimation measures [11], in addition to deriving optimum power allocation for such channels [6] [12]. However, this paper introduces a simple framework similar to [6] for deriving the capacity of Poisson channels for any model of consideration, with the assumption of constant stochastic martingale noise, i.e. for the sake of simplicity, we didn't model the noise as Gaussian within the stochastic intensity rate process. In addition, it builds upon derivations for the optimal power allocation for SISO, Parallel, and MAC models, or any other Poisson channel model of consideration.

In Poisson channels, the shot noise is the dominant noise whenever the power received at the photodetector is high; such noise is modeled as a Poisson random process. In fact, such framework has been investigated in many researches, see [3-7], [9-13]. Capitalizing on the expressions derived on [3-4], [6] and on the results by [6], [11], we re-investigate the derivation process in a simple step by step way; we then determine the optimal power allocation that maximizes the information rates.

The remainder of the paper is organized as follows; Section II introduces preliminary definitions and the communication framework. Section III introduces the SISO Poisson channel, as well as the optimal power allocation that maximizes the capacity. Section VI introduces the Parallel Poisson channel capacity expression as a normal generalization of the SISO setup, as well as the optimal power allocation. Section V introduces the MAC Poisson channel capacity as well as the optimal power allocation. Finally, we conclude the paper by some simulations and analytical results.

II. Preliminaries

A. The Communication Framework

In a communication framework, the information source inputs a message to a transmitter. The transmitter couples the message onto a transmission channel in the form of a signal which matches the transfer properties of the channel. The channel is the medium that bridges the distance between the transmitter and the receiver. This can be either a guided

Samah A. M. Ghanem is a researcher at the Institute of Telecommunications, Porto, Portugal. Samah is also a PhD student at the Faculty of Engineering (FEUP), University of Porto (UP), Portugal. (e-mail: samah.ghanem@gmail.com).

Munnujahan Ara is a researcher at the Institute of Telecommunications, Porto, Portugal. M. Ara is also a PhD student at FCUP, UP, Portugal. (e-mail: munnujahan@gmail.com).

(Advance online publication: 27 February 2012)

transmission such as a wire or wave guide, or it can be an unguided free space channel. A signal traverses the channel will suffer from attenuation and distortion. For example, electric power can be lost due to heat generation along a wire, and optical power can be attenuated due to scattering and absorption by air molecules in a free space.
Therefore, channels are characterized by a transfer function which models the input-output process. The input-output process statistics is dominated by the noise characteristics the modulated input experiences during its propagation along the communication medium, in addition to the detection procedure experienced at the channel output.
In particular, when the noise $n_G(t)$ is a zero-mean Gaussian process with double-sided power spectral density *N0/2*, the channel is called an Additive White Gaussian Channel (AWGN). However, when the electrical input is modulated by a light source, like a laser diode, the channel will be an optical channel with the dominant shot noise $n_d(t)$ arising from the statistical nature of the production and collection of photoelectrons when an optical signal is incident on a photodetector, such statistics characterized by a Poisson random process.
Fig. 1. illustrates both the AWGN and the Poisson optical channels. In this paper, we focus on the Poisson optical communication channel shown in Fig. 1. (b) and we derive capacity expressions for three different channel models, the SISO channel, the parallel MIMO, and the MAC Poisson channels.

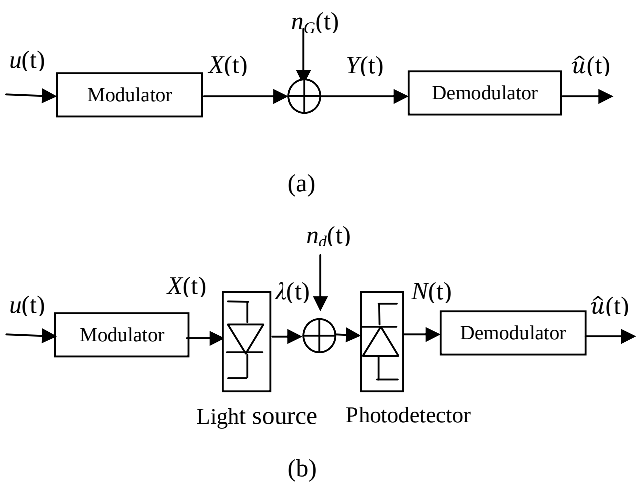

Fig.1. (a) The AWGN channel. (b) The Poisson optical channel.

B. Channel Capacity

Determining an ultimate limit to the rate at which we can reliably transmit information over a physical medium in a given environment is an earnest attempt of fundamental and practical consideration. Such a limit is referred to as the channel capacity and the process of evaluating this limit leads to an understanding of the technical solutions required to approach it. Capacity evaluations require information theory that must be adapted to the specific characteristics of the channel under study.
The capacity of a channel is the maximum mutual information, where the maximization is performed over all possible input distributions, i.e.,

$$C = \max_{p(X)} \; I(X;Y) \tag{1}$$

Where X corresponds to the channel input, and Y corresponds to the channel output. And the mutual information is defined by,

$$I(X;Y) = \iint p(X,Y) log\left(\frac{p(X,Y)}{p(X)p(Y)}\right) dxdy \tag{2}$$

With $p(X,Y)$ as the joint disctribution of X and Y, and the input distribution is $p(X)$ and the output distribution is $p(Y)$. Using Baye's rule, we can re-write (2) as follows,

$$I(X;Y) = \iint p(X,Y) log\left(\frac{p(Y|X)}{p(Y)}\right) dxdy \tag{3}$$

Where $p(Y|X)$ corresponds to the channel conditional probability which corresponds to the distribution of the noise, therefore, the channel will be considered as an AWGN if the conditional probability follows a Gaussian distribution, or cit will be considered as a Poisson channel if the conditional distribution follows a Poisson process.
We can simply write (3) in a more compact form, i.e.,

$$I(X;Y) = \mathbb{E}\left[log\left(\frac{p(Y|X)}{p(Y)}\right)\right] \tag{4}$$

With $\mathbb{E}[.]$ being the expectation operator over the joint distribution $P(X,Y)$ of the random variables X and Y.

C. Constrained Optimization Setup

To derive the optimal power allocation for different channel frameworks, it's worth to notice that different optimization criteria could be relevant. In particular, the optimization criteria could be the peak power, the average optical power, or the average electrical power. The average electrical power is the standard power measure in digital and wireless communications and it helps in assessing the power consumption in optical communications, while the average optical power is an important measure for safety considerations and helps in quantifying the impact of shot noise in wireless optical channels. In addition, the peak power, whether electrical or optical, gives a measure of tolerance against the nonlinearities in the system, for example the Kerr nonlinearity which is identified by a nonlinear phase delay in the optical intensity or in other words as the change in the refractive index of the medium as a function of the electric field intensity.
Therefore, in the context of AWGN channels, the constrained optimization setup relevant to single-user and multi-user systems usually takes the form,

$$\max \; I(X;Y) \tag{5}$$

Subject to the average electrical power constraint,

$$\mathbb{E}[|X|^2] \leq \; P. \tag{6}$$

However, in the context of Poisson optical channels the constrained optimization setup takes the form,

$$\max \; I(S_T;N_T) \tag{7}$$

Subject to the average optical power and peak power constraints,

$$\frac{1}{T}\,\mathbb{E}\left[\int_0^T \lambda(t)\,dt\right] \le \sigma P \quad (8)$$
$$0 \le \lambda(t) \le P. \quad (9)$$

P is the peak power and the ratio of average to peak power σ is used with $0 \le \sigma \le 1$.

III. THE SISO POISSON CHANNEL

Consider the SISO Poisson channel $\mathcal{P}$ shown in Fig.2. Let $N(t)$ represent the channel output, which is the number of photoelectrons counted by a direct detection device (photo-detector) in the time interval [0, T]. $N(t)$ has been shown to be a doubly stochastic Poisson process with instantaneous average rate $\lambda(t) + n$. The input $\lambda(t)$ is the rate at which photoelectrons are generated at time t in units of photons per second. And n is a constant representing the photo-detector dark current and background noise.

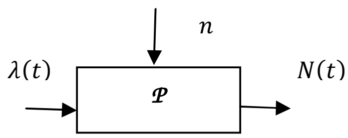

Fig. 2. The SISO Poisson channel model

A. Derivation of the Capacity of SISO Poisson Channels

Let $p(N_T)$ be the sample function density of the compound regular point process $N(t)$ and $p(N_T|S_T)$ be the conditional sample function of $N(t)$ given the message signal process $S(t)$ in the time interval [0,T]. Then we have,

$$p(N_T|S_T) = e^{-\int_0^T (\lambda(t)+n)\,dt + \int_0^T \log(\lambda(t)+n)\,dN(t)} \quad (10)$$
$$p(N_T) = e^{-\int_0^T (\widehat{\lambda(t)}+n)\,dt + \int_0^T \log(\widehat{\lambda(t)}+n)\,dN(t)} \quad (11)$$

We use the following consistent notation in the paper, $\widehat{\lambda(t)}$ is the estimate of the input $\lambda(t)$. $\mathbb{E}[.]$ is the expectation operation over time. Therefore, the mutual information is defined as follows,

$$I(S_T; N_T) = \mathbb{E}\left[\log\left(\frac{p(N_T|S_T)}{p(N_T)}\right)\right] \quad (12)$$

Theorem1 (Kabanov'78[3]-Davis'80 [4]):

The capacity of the SISO Poisson channel is given by:

$$C = \frac{K}{P}(P+n)\log(P+n) + \left(1-\frac{K}{P}\right)n\log(n) - (K+n)\log(K+n) \quad (13)$$

Proof:

Substitute (10) and (11) in (12), we have,

$$I(S_T; N_T) = \mathbb{E}\left[-\int_0^T (\lambda(t) - \widehat{\lambda(t)})\,dt + \int_0^T \log\left(\frac{\lambda(t)+n}{\widehat{\lambda(t)}+n}\right) dN(t)\right]$$

Since $\mathbb{E}\big[\widehat{\lambda(t)}\big] = \mathbb{E}\big[\mathbb{E}[\lambda(t)|N_T]\big] = \mathbb{E}[\lambda(t)]$, it follows that,

$$I(S_T; N_T) = \mathbb{E}\left[\int_0^T \log\left(\frac{\lambda(t)+n}{\widehat{\lambda(t)}+n}\right) dN(t)\right]$$

And $N(t) - \int_0^T \log(\lambda(t)+n)$ is a martingale[1] from theorems of stochastic integrals, see [6], [11], and [15] therefore,

$$\begin{aligned}
I(S_T; N_T) &= \mathbb{E}\left[\int_0^T (\lambda(t)+n)\log\left(\frac{\lambda(t)+n}{\widehat{\lambda(t)}+n}\right) dt\right] \\
&= \int_0^T \mathbb{E}[(\lambda(t)+n)\log(\lambda(t)+n)] \\
&\qquad - \mathbb{E}\big[(\lambda(t)+n)\log\big(\widehat{\lambda(t)}+n\big)dt\big] \\
&= \int_0^T \mathbb{E}[(\lambda(t)+n)\log(\lambda(t)+n)] \\
&\qquad - \mathbb{E}\big[\mathbb{E}[(\lambda(t)+n)]\log\big(\widehat{\lambda(t)}+n\big)|N_T\big]dt \\
&= \int_0^T \mathbb{E}[(\lambda(t)+n)\log(\lambda(t)+n)] \\
&\qquad - \mathbb{E}[\mathbb{E}[(\lambda(t)+n)|N_T]]\log\big(\widehat{\lambda(t)}+n\big)\,dt \\
&= \int_0^T \mathbb{E}[(\lambda(t)+n)\log(\lambda(t)+n)] \\
&\quad - \mathbb{E}\big[\big(\widehat{\lambda(t)}+n\big)\log\big(\widehat{\lambda(t)}+n\big)\big]\,dt
\end{aligned} \quad (14)$$

See [6] for similar steps. In [11], it has been shown that the derivative of the input–output mutual information of a Poisson channel with respect to the intensity of the dark current is equal to the expected error between the logarithm of the actual input and the logarithm of its conditional mean estimate, it follows that,

$$\frac{dI(S_T;N_T)}{d\lambda(t)} = \mathbb{E}\left[\log\left(\frac{\lambda(t)+n}{\widehat{\lambda(t)}+n}\right)\right] \quad (15)$$

The right hand side term of (15) is the derivative of the mutual information corresponding to the integration of the estimation errors. This plays as a counter part to the well known relation between the mutual information and the minimum mean square error (MMSE) in Gaussian channels in [14].

The capacity of the SISO Poisson channel given in Theorem1 (13) is defined as the maximum of (14) solving the following optimization problem,

$$\max\ I(S_T; N_T) \quad (16)$$

Subject to average and peak power constraints,

$$\frac{1}{T}\,\mathbb{E}\left[\int_0^T \lambda(t)\,dt\right] \le \sigma P$$
$$0 \le \lambda(t) \le P \quad (17)$$

With P is the maximum power and the ratio of average to peak power σ is used with $0 \le \sigma \le 1$.

We can easily check that the mutual information is strictly convex via its second derivative with respect to $\lambda(t)$ as follows,

$$\frac{d^2 I(S_T;N_T)}{d\lambda^2(t)} = \log\left(\frac{\lambda(t)+n}{\widehat{\lambda(t)}+n}\right) > 0.$$

Therefore, the mutual information is convex with respect to $\lambda(t)$. Now solving:

[1]In probability theory, a martingale is a stochastic process such that the conditional expected value of an observation at some time t, given all observations up to some earlier time s, is equal to the observation at that earlier time s, therefore, $\mathbb{E}\big[\widehat{\lambda(t)}\big] = \mathbb{E}\big[\mathbb{E}[\lambda(t)|N_T]\big] = \mathbb{E}[\lambda(t)]$, see [15].

(Advance online publication: 27 February 2012)

$\max\left(\int_0^T \mathbb{E}[(\lambda(t)+n)\log(\lambda(t)+n)] - \mathbb{E}\left[\left(\widehat{\lambda(t)}+n\right)\log\left(\widehat{\lambda(t)}+n\right)\right] - \frac{\xi}{T}\mathbb{E}[\lambda(t)]\right)$,
with ξ as the lagrangian multiplier.

The possible values of $\mathbb{E}[(\lambda(t)+n)\log(\lambda(t)+n)]$ must lie in the set of all y-coordinates of the closed convex hull of the graph y=$(x+n)\log(x+n)$. Hence, the maximum mutual information achieved using the distribution $p(\lambda=P)=1-p(\lambda=0)=\alpha$. Where $0\le\alpha\le1$, so that $\mathbb{E}[\lambda(t)]=K$.
So, we must have $[\lambda(t)]=\sum\lambda(t)p(\lambda)$. It follows that, $K=Pp(\lambda=P)=P\alpha$. Then, $\alpha=\frac{k}{P}$. And then the capacity in (13) is proved.

B. Optimum Power Allocation for SISO Poisson channels

We need to solve the following optimization problem,

$$\max\left(\frac{K}{P}(P+n)\log(P+n)+\left(1-\frac{K}{P}\right)n\log(n)-(K+n)\log(K+n)-\frac{\xi}{T}K\right) \quad (18)$$

Since (18) is concave with respect to K, i.e. the second derivative of (18) with respect to K is negative. Using the Lagrangian corresponding to the derivative of the objective with respect to K, and the Karush–Kuhn–Tucker (KKT) conditions, the optimal power allocation is the following,

$$K^*=(P+n)e^{-\left(1+\frac{\xi}{T}\right)+\frac{n}{P}\log\left(1+\frac{P}{n}\right)}-n \quad (19)$$

IV. THE PARALLEL POISSON CHANNEL

Consider the Parallel Poisson channel shown in Fig.3.

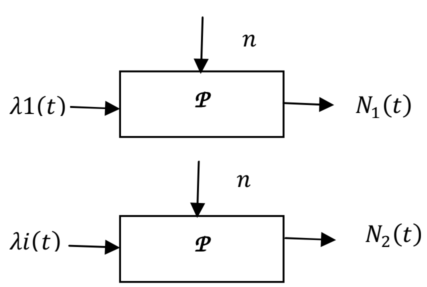

Fig. 3. The Parallel Poisson channel model

Consider a 2-fold parallel Poisson channel, then, $N_1(t)$ and $N_2(t)$ are doubly stochastic Poisson processes with instantaneous average rates $\lambda1(t)+n$ and $\lambda2(t)+n$ respectively.

A. Derivation of the Capacity of Parallel Poisson Channels

Let $p(N_1,N_2)$ and $p(N_1,N_2|S_1,S_2)$ be the joint density and conditional sample function of the compound regular point processes $N_1(t)$ and $N_2(t)$ respectively, given the message signal processes $S_1(t)$ and $S_2(t)$ in the time interval [0, T]. Then we have,

$$p(N_1,N_2|S_1,S_2)=p(N_1|S_1)p(N_2|S_2) \quad (20)$$
$$p(N_1,N_2)=p(N_1)p(N_2) \quad (21)$$

$p(N_1|S_1), p(N_2|2), p(N_1)$, and $p(N_2)$ are given by (10) and (11) respectively for each input $\lambda i(t)$. Therefore, the mutual information is defined as follows,

$$I(S_T;N_T)=\mathbb{E}\left[log\left(\frac{p(N_1|S_1)p(N_2|S_2)}{p(N_1)p(N_2)}\right)\right] \quad (22)$$

Theorem2:
The capacity of the 2-input parallel Poisson channel is given by the sum capacity of the independent SISO Poisson channels as follows:

$$C=\frac{K1}{P1}(P1+n)\log(P1+n)+\left(1-\frac{K1}{P1}\right)n\log(n)-(K1+n)\log(K1+n)+\frac{K2}{P2}(P2+n)\log(P2+n)+\left(1-\frac{K2}{P2}\right)n\log(n)-(K2+n)\log(K2+n) \quad (23)$$

Proof:
Substitute (20) and (21) in (22), we have,

$$I(S_T;N_T)=\mathbb{E}\left[\int_0^T\log\left(\frac{\lambda1(t)+n}{\widehat{\lambda1(t)}+n}\right)dN1(t)\right]+\mathbb{E}\left[\int_0^T\log\left(\frac{\lambda2(t)+n}{\widehat{\lambda2(t)}+n}\right)dN2(t)\right]$$

Following the same steps of the proof of theorem1, we can easily find that,

$$I(S_T;N_T)=\int_0^T\mathbb{E}[(\lambda1(t)+n)\log(\lambda1(t)+n)]-\mathbb{E}\left[\left(\widehat{\lambda1(t)}+n\right)\log\left(\widehat{\lambda1(t)}+n\right)\right]+\mathbb{E}[(\lambda2(t)+n)\log(\lambda2(t)+n)]-\mathbb{E}\left[\left(\widehat{\lambda2(t)}+n\right)\log\left(\widehat{\lambda2(t)}+n\right)\right]dt \quad (24)$$

The capacity of the Parallel Poisson channel given in Theorem2 (23) is defined as the maximum of (24) solving the following optimization problem,

$$\max\ I(S_T;N_T) \quad (25)$$

Subject to average and peak power constraints,

$$\frac{1}{T}\mathbb{E}\left[\int_0^T\lambda1(t)\right]\le\sigma P1$$
$$\frac{1}{T}\mathbb{E}\left[\int_0^T\lambda2(t)\right]\le\sigma P2$$
$$0\le\lambda1(t)\le P1,\ 0\le\lambda2(t)\le P2 \quad (26)$$

With $P1$ and $P2$ are the maximum power and the ratio of average to peak power σ is used with $0\le\sigma\le1$.
Hence, the maximum mutual information achieved using the distribution of any input i such that $p(\lambda i=Pi)=1-p(\lambda i=0)=\alpha i$. Where $0\le\alpha i\le1$ so that $\mathbb{E}[\lambda i(t)]=Ki$.
So, we must have $[\lambda i(t)]=\sum\lambda i(t)p(\lambda i)$. It follows that, $Ki=Pip(\lambda i=Pi)=P\alpha i$. Then, $\alpha i=\frac{ki}{Pi}$. Therefore, the capacity in (23) is proved.

B. Optimum Power Allocation of Parallel Poisson Channels

We need to solve the following optimization problem,

$$\max\left(\frac{K1}{P1}(P1+n)\log(P1+n)+\left(1-\frac{K1}{P1}\right)n\log(n)-(K1+n)\log(K1+n)+\frac{K2}{P2}(P2+n)\log(P2+n)+\left(1-\frac{K2}{P2}\right)n\log(n)-(K2+n)\log(K2+n)-\frac{\xi}{T}(K1+K2)\right) \quad (27)$$

Using the Lagrangian corresponding to the derivative of the objective with respect to K, and applying the Karush–Kuhn–Tucker (KKT) conditions, the optimal power allocation follows the optimal power allocation for the SISO setup in (19). Therefore, for an *i*-fold parallel Poisson channel, the optimal power allocation is the following,

$$Ki^* = (Pi+n)e^{-\left(1+\frac{\xi}{T}\right)+\frac{n}{Pi}\log\left(1+\frac{Pi}{n}\right)} - n \qquad (28)$$

See [12] for similar results related to optimum power allocation for a 2-fold Parallel Poisson channel where the power constraint was the sum of both average input powers.

V. THE MAC POISSON CHANNEL

Consider the MAC Poisson channel shown in Fig. 4.

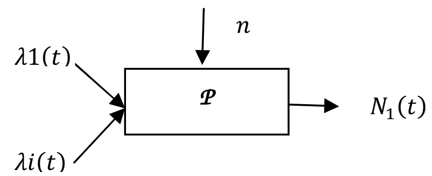

Fig. 4. The MAC Poisson channel model

Consider a 2-input MAC Poisson channel, then, $N_1(t)$ is a doubly stochastic Poisson processes with instantaneous average rates $\lambda 1(t) + \lambda 2(t) + n$.

A. *Derivation of the Capacity of MAC Poisson Channels*

Let $p(N_1)$ and $p(N_1|S_1,S_2)$ be the joint density and conditional sample function of the compound regular point process $N_1(t)$ given the message signal processes $S_1(t)$ in the time interval [0, T]. Then we have,

$$p(N_1|S_1,S_2) = e^{-\int_0^T(\lambda 1(t)+\lambda 2(t)+n)\,dt+\int_0^T\log(\lambda 1(t)+\lambda 2(t)+n)dN(t)} \qquad (29)$$

$$p(N_1) = e^{-\int_0^T(\widehat{\lambda 1(t)}+\widehat{\lambda 2(t)}+n)\,dt+\int_0^T\log(\widehat{\lambda 1(t)}+\widehat{\lambda 2(t)}+n)dN(t)} \qquad (30)$$

Therefore, the mutual information is defined as follows,

$$I(S_T;N_T) = \mathbb{E}\left[log\left(\frac{p(N_1|S_1,S_2)}{p(N_1)}\right)\right] \qquad (31)$$

Theorem3:

The capacity of the 2-input MAC Poisson channel is given by:

$$C = \left(\frac{K1}{P}+\frac{K2}{P}\right)(P+n)\log(P+n) + \left(1-\left(\frac{K1}{P}+\frac{K2}{P}\right)\right)nlog(n) -(K1+K2+n)\log(K1+K2+n) \qquad (32)$$

Proof:

Substitute (29) and (30) in (31), we have,

$$I(S_T;N_T) = \mathbb{E}\left[-\int_0^T(\lambda 1(t)-\widehat{\lambda 1(t)})\,dt - \int_0^T(\lambda 2(t)-\widehat{\lambda 2(t)})\,dt + \int_0^T\log\left(\frac{\lambda 1(t)+\lambda 2(t)+n}{\widehat{\lambda 1(t)}+\widehat{\lambda 2(t)}+n}\right)dN(t)\right]$$

Since $\mathbb{E}[\widehat{\lambda 1(t)}+\widehat{\lambda 2(t)}] = \mathbb{E}[\mathbb{E}[\lambda 1(t)+\lambda 2(t)|N_T]] = \mathbb{E}[\lambda 1(t)+\lambda 2(t)]$, it follows that,

$$I(S_T;N_T) = \mathbb{E}\left[\int_0^T\log\left(\frac{\lambda 1(t)+\lambda 2(t)+n}{\widehat{\lambda 1(t)}+\widehat{\lambda 2(t)}+n}\right)dN(t)\right]$$

And $N(t)-\int_0^T\log(\lambda 1(t)+\lambda 2(t)+n)$ is a martingale from theorems of stochastic integrals, see [6], [11], and [15] therefore,

$$\begin{aligned}
I(S_T;N_T) &= \mathbb{E}\left[\int_0^T(\lambda 1(t)+\lambda 2(t)+n)\log\left(\frac{\lambda 1(t)+\lambda 2(t)+n}{\widehat{\lambda 1(t)}+\widehat{\lambda 2(t)}+n}\right)dt\right]\\
&= \int_0^T\mathbb{E}[(\lambda 1(t)+\lambda 2(t)+n)\log(\lambda 1(t)+\lambda 2(t)+n)] - \mathbb{E}\left[(\lambda 1(t)+\lambda 2(t)+n)\log\left(\widehat{\lambda 1(t)}+\widehat{\lambda 2(t)}+n\right)dt\right]\\
&= \int_0^T\mathbb{E}[(\lambda 1(t)+\lambda 2(t)+n)\log(\lambda 1(t)+\lambda 2(t)+n)] - \mathbb{E}\left[\mathbb{E}[(\lambda 1(t)+\lambda 2(t)+n)]\log\left(\widehat{\lambda 1(t)}+\widehat{\lambda 2(t)}+n\right)|N_T\right]dt\\
&= \int_0^T\mathbb{E}[(\lambda 1(t)+\lambda 2(t)+n)\log(\lambda 1(t)+\lambda 2(t)+n)] - \mathbb{E}[\mathbb{E}[(\lambda 1(t)+\lambda 2(t)+n)|N_T]]\log\left(\widehat{\lambda 1(t)}+\widehat{\lambda 2(t)}+n\right)\,dt\\
&= \int_0^T\mathbb{E}[(\lambda 1(t)+\lambda 2(t)+n)\log(\lambda 1(t)+\lambda 2(t)+n)] -\mathbb{E}\left[\left(\widehat{\lambda 1(t)}+\widehat{\lambda 2(t)}+n\right)\log\left(\widehat{\lambda 1(t)}+\widehat{\lambda 2(t)}+n\right)\right]dt \qquad (33)
\end{aligned}$$

The capacity of the MAC Poisson channel given in Theorem3 (32) is defined as the maximum of (33) solving the following optimization problem subject to average power and peak power constraints,

$$\max\ I(S_T;N_T) \qquad (34)$$

Subject to average and peak power constraints,

$$\frac{1}{T}\mathbb{E}\left[\int_0^T\lambda 1(t)+\lambda 2(t)\,dt\right] \le \sigma P$$
$$0 \le \lambda 1(t) \le P1$$
$$0 \le \lambda 2(t) \le P2 \qquad (35)$$

With $P1$ and $P2$ are the maximum power and the ratio of average to peak power σ is used with $0 \le \sigma \le 1$.

Now, solving:

$$\max\left(\int_0^T\mathbb{E}[(\lambda 1(t)+\lambda 2(t)+n)\log(\lambda 1(t)+\lambda 2(t)+n)] - \mathbb{E}\left[\left(\widehat{\widehat{\lambda 1(t)}+\widehat{\lambda 2(t)}}+n\right)\log\left(\widehat{\widehat{\lambda 1(t)}}+\widehat{\lambda 2(t)}+n\right)\right] - \frac{\xi}{T}\mathbb{E}[\lambda 1(t)+\lambda 2(t)]\right),$$

with ξ as the lagrangian multiplier. The possible values of $\mathbb{E}[(\lambda 1(t)+\lambda 2(t)+n)\log(\lambda 1(t)+\lambda 2(t)+n)]$ must lie in the set of all y-coordinates of the closed convex hull of the graph y=$(x1+x2+n)\log(x1+x2+n)$.

Suppose that the maximum power for both inputs is $P1+P2=\sigma P$. Hence, the maximum mutual information achieved using the distribution $p(\lambda = P) = 1-p(\lambda=0) = \alpha$. Where $0 \le \alpha \le 1$ so that $\mathbb{E}[\lambda 1(t)] = K1$, $\mathbb{E}[\lambda 2(t)] = K2$. So, we have $\mathbb{E}[\lambda 1(t)+\lambda 2(t)] = \sum(\lambda 1(t)p(\lambda 1)+$

$\lambda 2(t))p(\lambda 2)$. It follows that, $K1 = Pp(\lambda 1 = P) = P\alpha$. $K2 = Pp(\lambda 2 = P) = P(1-\alpha)$. Then, $\alpha = \frac{k1}{P}$ and $1-\alpha = \frac{k2}{P}$ and then the capacity in Theorem3 (32) is proved.

It's worth to note that we also have $K3 = P1p(0 \le \lambda 1(t) \le \sigma P) + P2p(0 \le \lambda 2(t) \le \sigma P) = P1\alpha + P2(1-\alpha)$, however, $K3$ is not considered in the capacity equations since we only need the maximum and the minimum powers for both $\lambda 1(t)$ and $\lambda 2(t)$ to get the maximum expected value.

B. *Optimum Power Allocation of MAC Poisson Channels*

We need to solve the following optimization problem,

$$\max\left(\left(\frac{K1}{P}+\frac{K2}{P}\right)(P+n)\log(P+n) + \left(1-\frac{K1}{P}\right)nlog(n) + \left(1-\frac{K2}{P}\right)nlog(n) - (K1+K2+n)\log(K1+K2+n) - \frac{\xi}{T}(K1+K2)\right) \quad (36)$$

Using the Lagrangian corresponding to the derivative of the objective with respect to K, and the Karush–Kuhn–Tucker (KKT) conditions, the optimal power allocation is the solution of the following equation,

$$K1^* + K2^* = (P+n)e^{-\left(1+\frac{\xi}{T}\right)+\frac{n}{P}\log\left(1+\frac{P}{n}\right)} - n \quad (37)$$

The optimum power allocation solution introduces the fact that orthogonalizing the inputs via time or frequency sharing will achieve the capacity; therefore it follows the importance for interface solutions to aggregate different inputs to the Poisson channel.

We can also differentiate (32) with respect to the maximum power P at which the capacity of the 2-input MAC Poisson channel is achieved with the optimal peak power P^* with respect to the average power of both inputs is the solution of,

$$(K1+K2)P^{*2} + (K1+K2)nP^* + (K1+K2)n(P^*+n)log\left(\frac{n}{P^*+n}\right) = 0 \quad (38)$$

VI. DISCUSSION

A. *Mathematical Analysis*

The solutions provided in the paper show that the capacity of Poisson channels is a function of the average and peak power of the input. It can be easily seen that similar to the Gaussian Parallel channels; Poisson parallel channels have the characteristic that their throughput is the sum of their independent SISO channels. In [9] the authors studied the capacity regions and the maximum achievable mutual information, or we can call it the upper bound of the information rate of the two-user MAC Poisson channel with equal average input powers. However, they pointed out an interesting observation that we can also see here via Theorem3; that is; in contrary to the Gaussian MAC, in the Poisson MAC the maximum throughput is bounded in the number of inputs, and similar to the Gaussian MAC in terms of achieving the capacity via orthogonalizing the inputs. We can also see that the maximum power is a function of the average power through which both can be optimized to maximize the capacity.

B. *Simulation Analysis*

Fig.5. shows the capacity of the SISO, parallel, and MAC Poisson channels with respect to the average power and under a maximum power $P = 10$, and dark current $n = 0.1$, it can be easily noticed through the mathematical results as well as the simulations that the capacity of parallel Poisson channels is exactly double the capacity of the SISO Poisson channels if we consider the average power $K1 = K2 = K$ and the maximum power constraint is met and equal for both channels, i.e. $P1 = P2 = P$. However, on the one hand, it is clear that at the low average power regime, the MAC Poisson channel capacity under same conditions lie between both channels. While it decays as the average power increases if inputs are not orthogonal.

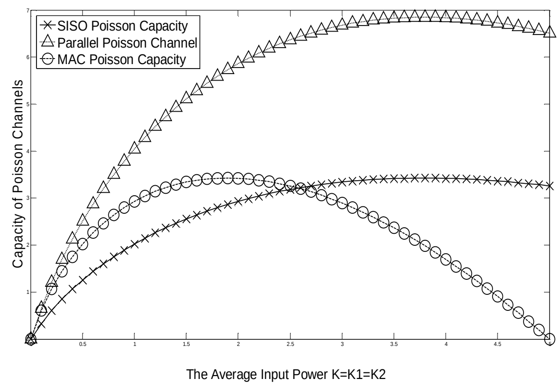

Fig.5. Capacity of Poisson Channels (photons/sec) versus the average power K.

On the other hand, for a different setup where one input average power is lowest and the other input power is maximum, i.e. a time or frequency shared inputs, it turns out that the MAC capacity is higher than that of Parallel channel, this is due to the fact that the dark noise is much more influencing the Parallel setup than that in the MAC setup. It can be easily verified that the MAC capacity can be maximized when $\frac{K1}{P} = \frac{K2}{P} = \frac{K}{P}$. However, when equal input maximum powers $K1 = K2 = P$ are used the capacity decays to zero. Similarly, for the Gaussian MAC, at equal input powers or more precisely, when the arbitrary inputs to the Gaussian MAC lie in the null space or the Voronoi region of the channel matrix, the capacity faces a decay to zero in the total achievable rate of the MAC, while when they differ i.e. inputs are orthogonal, the capacity moves into maximum.

Fig.6 shows the capacity of SISO, Parallel, and MAC Poisson channels with respect to the detector dark current, it shows that the capacity is a decreasing function with respect to n; however, for the MAC the capacity increases after a certain point with respect to n. We can also see via Fig. 5 and Fig. 6 that the two main factors in the MAC capacity is the orthognalization and the maximum power, while increasing the average power for one or the two inputs will not add positively to the capacity.

Fig.7 shows the optimum power allocation results, it can be deduced via the mathematical formulas as well as the simulations that the power allocation is a decreasing value with respect to the dark current for all Poisson channels. It means that the power allocation for the Poisson channels in

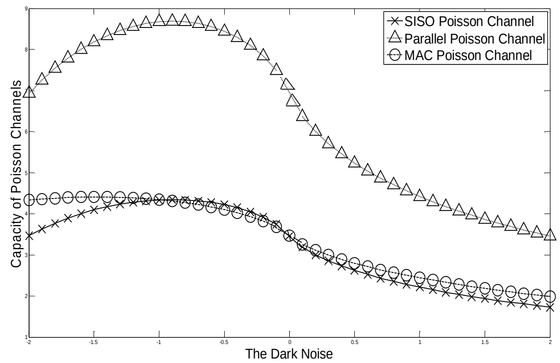

Fig.6. Capacity of Poisson Channels versus the dark current n, (photons/sec).

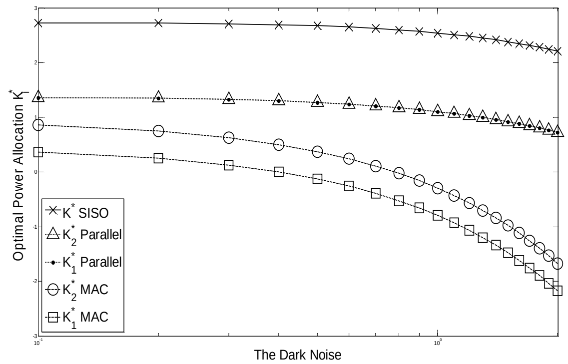

Fig.7. Optimum Power Allocation versus the dark current n, (photons/sec).

some way or another follows a waterfilling alike interpretation in the Gaussian setup where less power is allotted to the more noisy channels [16]. However, it is well known that the optimum power allocation is an increasing function in terms of the maximum power.

A. General Analysis

Here, we will introduce some important points about the capacity of Poisson channels in comparison to Gaussian channels within the context of this paper. Firstly, in comparison to the Gaussian capacity, the channel capacity of the Poisson channel is maximized with binary inputs, i.e. [0, 1], while the distribution that achieves the Gaussian capacity is a Gaussian input distribution. Secondly, the maximum achievable rates for the Poisson channel is a function of its maximum and average powers due to the nature of the Poisson processes that follows a stochastic random process with martingale characteristics, while in Gaussian channels, the processes are random and modeled by the normal distribution. Thirdly, the optimum power allocation for the Poisson channels is very similar for different models depending on the defined power constraints, and in comparison to the Gaussian optimum power allocation; it follows a similar interpretation to the waterfilling, at which more power is allocated to stronger channels, i.e. power allocation is inversely proportional to the more noisy channel. However, although the optimal inputs distribution for the Poisson channel is a binary input distribution, the optimal power allocation is a waterfilling alike, i.e. unlike the Gaussian channels with arbitrary inputs where it follows a mercury-waterfilling interpretation to compensate for the non-Gaussianess in the binary input [17].

VII. CONCLUSION

In this paper, we show via information theoretic approach that the capacity of optical Poisson channels is a function of the average and maximum power of the inputs, the capacity expressions have been derived as well as the optimal power allocation for different channel models. It is shown -through the limitation on users within the capacity of the Poisson MAC- that the interface solutions for the aggregation of multiple users/channels over a single Poisson channel are of great importance. However, a technology like orthogonal frequency division multiplexing (OFDM) for optical communications stands as one interface solution. While it introduces attenuation via narrow filtering, etc. it therefore follows the importance of optimum power allocation which can mitigate such effects, hence, we build upon optimum power allocation derivations.

REFERENCES

[1] C. E. Shannon , "A mathematical theory of communication," The Bell System Technical Journal, vol. 27, pp. 379-423 and 623-656, 1948.

[2] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel,“Capacity limits of optical fiber networks,” Journal of Lightwave Technology,vol. 28, no. 4, pp. 662–701, Feb. 2010.

[3] Y. M. Kabanov, “The Capacity of a Channel of the Poisson Type”, Theory of Probability and Its Applications, vol. 23, No. 1, 1978.

[4] M. H. A. Davis, “ Capacity and Cutoff Rate for Poisson-Type Channels”, IEEE Transactions on Information Theory, Vol. 26, No. 6, 1980.

[5] M. Fray, “Information Capacity of the Poisson Channel”, IEEE Transactions on Information Theory, Vol. 37, No. 2, 1991.

[6] H. H. Han, “Capacity of Multimode Direct Detection Optical Communication Channel”, TDA Progress Report 42-63, 1981.

[7] Samah A. M. Ghanem, Munnujahan Ara, “Capacity and Optimal Power Allocation of Poisson Optical Communication Channels”, Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering and Computer Science 2011, WCECS 2011, 19-21 October, 2011, San Francisco, USA, pp817-821.

[8] D. H. Johnson, I. N. Goodman, “Inferring the capacity of the vector Poisson channel with a Bernoulli model”, Network-computation in Neural Systems,Vol. 19, No. 1, 2008.

[9] A. Lapidoth, S. Shamai, “The Poisson Multiple Access Channel”, IEEE Transactions on Information Theory, Vol. 44, No. 1, 1998.

[10] L. Lifeng, L. Yingbin, S. Shamai, “On the capacity region of the Poisson interference channels”, IEEE International Symposium on Information Theory, Austin, TX, 2010.

[11] G. Dongning, S. Shamai, S. Verdu, “Mutual information and conditional mean estimation in Poisson channels”, IEEE Transactions on Information Theory, Vol. 54, No. 5, 2008.

[12] M. M. Alem-Karladani, L. Sepahi, M. Jazayerifar, K. Kalbasi, “Optimum power allocation in parallel Poisson optical channel”, 10th International Conference on Telecommunications, Zagreb, 2009.

[13] A. Segall, T. Kailath, “The Modeling of Randomly Modulated Jump Processes”, IEEE Transactions on Information Theory, Vol. IT-21, No. 1, 1975.

[14] G. Dongning, S. Shamai, S. Verdu, “Mutual information and minimum mean-square error in Gaussian channels”, IEEE Transactions on Information Theory, Vol. 51, 2005.

[15] L. C. G. Rogers and David Williams. Diffusions, Markov processes, and martingales. Vol. 2.Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000.

[16] T. M. *Cover*, Joy A. *Thomas*, “Elements of Information Theory”, Vol. 1, 2nd Edition, by J. Wiley & Sons, ltd.. ©2006.

[17] A. Lozano, A. Tulino and S. Verdú, “Mercury/Waterfilling: Optimum Power Allocation with Arbitrary Input Constellations”, IEEE International Symposium on Information Theory, Australia, 2005.