Optimal Power Allocation for Poisson Channels With Time-Varying Background Light

Ain-ul-Aisha, Lifeng Lai, *Member, IEEE*, and Yingbin Liang, *Member, IEEE*

***Abstract*—In this paper, we study Poisson fading channels with time-varying background light. Different from most of the existing work on fading Poisson channel that focus on the case with time-varying channel gain, our model is motivated by indoor optical wireless communication systems, in which the noise level is affected by the strength of the background light. We study both the single-input single-output and the multiple-input and multiple-output channels. For each channel, we consider scenarios with and without delay constraints. For the case without a delay constraint, we characterize the optimal power allocation scheme that maximizes the ergodic capacity. For the case with a strict delay constraint, we characterize the optimal power allocation scheme that minimizes the outage probability. We also provide several numerical examples to demonstrate the analytic results.**

***Index Terms*—Ergodic capacity, outage probability, Poisson channels, time-varying background light.**

I. Introduction

THE Poisson channel is a model suitable for free-space optical (FSO) communications and visible light communications (VLC) [2] with direct-detection receivers. In case of such receivers as the randomness in photon arrival is more than the thermal noise, Poisson model is considered as the appropriate model [3]. Compared to the Gaussian channel that has been extensively studied, the Poisson channel is less well understood due to several technical challenges. In particular, Poisson channels are non-linear, not scale-invariant and have continuous inputs and discrete outputs [4]. Consequently, Poisson channels are difficult to analyze.

Recently, there have been great interests in analyzing Poisson fading channels. There are two major types of fading models, that with time-varying channel gains and that with time-varying noise (e.g., background light) levels, for Poisson fading models. These two types of models are not equivalent and cannot be treated equivalently, because Poisson channels are not scale-invariant. This is different from Gaussian fading channels, in which channels with varying noise levels can be converted to channels with varying channel gains due to the scale-invariant property.

The first type of Poisson fading channels with time-varying channel gains have been studied in [5]–[8], which characterized the ergodic and outage capacities. These studies developed useful information theoretic tools that will also be used in our paper. Furthermore, [9] investigated this type of fading channels when the channel gains are log-normal random variables. The performance of the channel in both high and low signal to noise ratio regimes are studied based on lower and upper bounds on the channel capacity. [10] investigated the outage probabilities of several diversity schemes. [11] studied a single-input single-output (SISO) Poisson channel with channel state information (CSI) perfectly known at the receiver and partially known at the transmitter. The goal is to maximize the ergodic capacity of the channel, with partial information at the transmitter obtained by an error-free feedback link with a finite rate constraint from the receiver to the transmitter. [12] investigated the behavior of the outage capacity for the decode-and-forward multi-hop Poisson fading channel for FSO, where the atmospheric turbulence contributes to the fading in the channel. The paper has characterized the optimal power control function under different assumptions on the availability of the CSI at the transmitter.

The second type of Poisson fading channels with time-varying noise levels have been much less studied with only a few exceptions as we describe below. In fact, such models arise in many practical scenarios. For instance, in indoor optical wireless communications, the noise levels at the receiver are affected by the temperature and the strength of the background light, as the noise level increases when the temperature or the strength of the background light increases. In addition, the noise level is higher when other light sources are also on. [13] studied the optimal power allocation for 2-fold parallel poisson channel for constant dark current, which can be viewed as an equivalent SISO channel with time-varying noise levels. A recent work [14] also studied the channel with time-varying noise levels under an assumption that the transmitter knows the noise realization at the receiver.

In this paper, we study the second type of Poisson fading channels with time-varying noise levels. Our model is clearly different from the first type of Poisson fading models studied in [5]–[12], [15] due to the non-equivalence of the two types of models as we explain above. Our study is also different from [13], in which the model is equivalently SISO and the study

Manuscript received November 27, 2014; revised April 18, 2015, June 16, 2015, and August 22, 2015; accepted September 6, 2015. Date of publication September 14, 2015; date of current version November 13, 2015. The work of Ain-ul-Aisha and L. Lai was supported by the National Science Foundation under grant CCF-12-18541. The work of Y. Liang was supported by the National Science Foundation under grant CCF-12-18451. This paper was presented in part in the 48th Annual Conference on Information Sciences and Systems (CISS), 2014 [1]. The associate editor coordinating the review of this paper and approving it for publication was E. Agrell.

Ain-ul-Aisha and L. Lai are with the Department of Electrical and Computer Engineering, Worcester Polytechnic Institute, Worcester, MA 01609 USA (e-mail: aaisha@wpi.edu; llai@wpi.edu).

Y. Liang is with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13210 USA (e-mail: yliang06@syr.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2015.2478473

focused on the case with two channels (equivalently two noise levels), whereas here we study the more general SISO case with arbitrary number of noise levels and the multiple-input multiple-output (MIMO) channel. Our study also differentiates from [14] in that we make a mild assumption that only the noise level (a statistic quantity) rather than the realization of the noise is known at the transmitter.

More specifically, we study both SISO and MIMO channels with and without delay constraints. Our contributions lie in a comprehensive characterization of the optimal power allocation schemes to achieve the ergodic capacity (for the case with no delay constraints) and to minimize the outage probability (for the case with delay constraints). Here, the delay constraint is measured by the number of fading blocks after which the receiver decodes information (i.e., codewords). The delay can also be viewed as the number of fading blocks that a codeword is allowed to span. If there is no delay constraint, then the codeword length is allowed to be infinite, in which case we use the ergodic capacity as the performance metric. If there is a delay constraint that requires the receiver to decode after a finite number of blocks, then we focus on the outage probability that captures the performance of such a scenario. For both scenarios, we assume that the transmitter knows the noise level. As will be clear in the sequel, this is a reasonable assumption because the noise level here represents a statistic quantity but not the realization of the noise. Hence, the amount of feedback needed from the receiver to the transmitter is limited. In addition, as the noise level is affected by the background light, it can be effectively measured at the transmitter. This can further reduce the amount of feedback necessary for the transmitter to learn the value of the noise level.

For the case with no delay constraint, we establish the ergodic capacity, and characterize the corresponding optimal power allocation scheme as a function of the noise level that achieves the ergodic capacity. For the case with a strict delay constraint, our goal is to minimize the probability that the instantaneous achievable rate is less than a given threshold, i.e., the outage probability. Minimizing the outage probability directly is very challenging. In order to solve this problem, we apply the techniques developed in [16] to study a number of related optimization problems. From the solutions of these optimization problems, we then characterize the optimal power allocation scheme that minimizes the outage probability. Both problems are significantly more challenging than the corresponding problems in the Gaussian channels.

The remainder of the paper is organized as follows. In Section II, we introduce the system model. In Section III, we analyze the SISO channel. In Section IV, we extend our analysis to the MIMO case. We present several numerical examples in Section V. Finally, we provide concluding remarks in Section VI.

II. System Model

In this section, we introduce the model considered in this paper. As shown in Fig. 1, we consider a MIMO Poisson channel with N transmitter antennas and M receiver antennas. Let

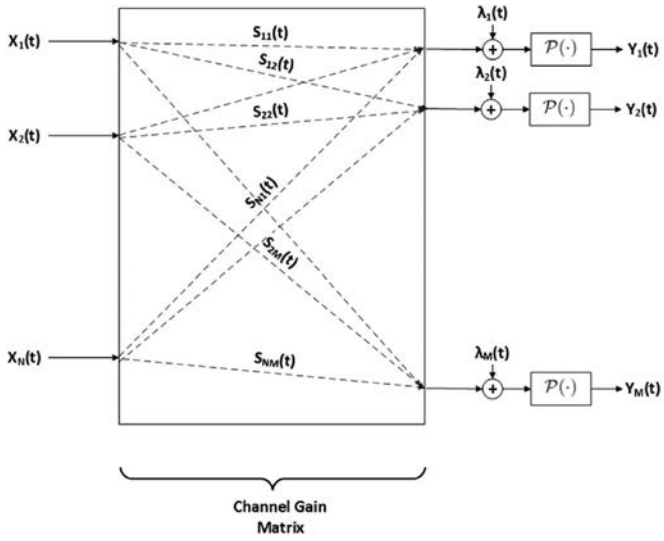


Fig. 1. MIMO Poisson channel with time-varying background light.

$X_n(t)$ be the input of the n^{th} transmitter antenna and $Y_m(t)$ be the doubly-stochastic Poisson process observed at the m^{th} receiver antenna. The relationship between them can be described as:

$$Y_m(t) = \mathcal{P}\left(\sum_{n=1}^{N} S_{nm}X_n(t) + \lambda_m(t)\right), \tag{1}$$

in which S_{nm} is the channel response between the n^{th} transmitter antenna and m^{th} receiver antenna, $\lambda_m(t)$ is the dark current at the m^{th} receiver antenna, which signifies the background light, and $\mathcal{P}(\cdot)$ is the non-linear transformation converting the light strength to the doubly-stochastic Poisson process that records the timing and number of photon's arrivals. In particular, for any time interval $[t, t+\tau]$, the probability that there are j photons arriving at receiver antenna m is

$$\Pr\{Y_m(t+\tau) - Y_m(t) = j\} = \frac{e^{-\Lambda_m}\Lambda_m^j}{j!}, \tag{2}$$

$$\Lambda_m = \int_t^{t+\tau} \left[\sum_{n=1}^{N} S_{nm}X_n(t') + \lambda_m(t')\right] dt'. \tag{3}$$

We consider the maximum and average sum power constraint, i.e. the transmitted signal $X_n(t)$ must satisfy the following constraints:

$$0 \le X_n(t) \le A_n, \tag{4}$$

$$\frac{1}{T}\int_0^T \sum_{n=1}^{N} X_n(t)dt \le \sigma \sum_{n=1}^{N} A_n, \tag{5}$$

in which A_n is the maximum power allowed for antenna n and σ is the average to peak power ratio. In our model, we assume that S_{nm} is constant while $\lambda_m(t)$ is time-varying. This model is motivated by the fact that the dark current is a physical parameter that depends on the temperature and the background lights in the environment, which naturally change throughout the day. We consider block fading model, in which $\lambda_m(t)$ is fixed for a block of symbols and then changes to another independent

value at the beginning of the next block. Furthermore, we assume that the transmitter knows λ_m. This can be justified for two reasons. Firstly, λ_m does not represent the realization of the noise, but is a statistical quantity that characterizes the average behavior of the noise process. Hence, it is reasonable that the receiver feeds back the information about λ_m albeit with a certain rate limit. Secondly, λ_m is affected by the background light and temperature, which can be measured at both the transmitter and the receiver. In this paper, we assume that the distribution of λ satisfies the following properties: (i) $\Pr\{\lambda(t) > 0\} = 1$, (ii) $\mathbb{E}[\lambda(t)] < \infty$, and (iii) $\mathbb{E}[\lambda(t)\log\lambda(t)] < \infty$.

We note that the model considered here is different from the one considered in [5], [6], [8], [11], [12], [15], in which the channel gain S_{nm} is time-varying while λ_m is fixed. This is because the case of varying dark currents can not be converted to the case of varying channel gain, due to the nonlinearity of the Poisson channel unlike the Gaussian channels. However, some techniques developed in these studies are useful for solving the problems studied in our paper.

III. SISO Channel Analysis

We first study a special case with $N = 1$ and $M = 1$, namely the SISO channel, to introduce main tools used in the MIMO case in Section IV. As $M = N = 1$, we drop the subscripts m and n in variables in this section for notational convenience.

A. Preliminary

In preparation for the further development, we first review existing results and techniques for the case when λ is fixed. In particular, Wyner [17] developed a binary approximation method that converts the complicated continuous time continuous input discrete output Poisson channel into a discrete time binary input binary output channel. It is much simpler to handle the binary channel, and it is shown in [17] that this binary approximation does not reduce the capacity.

In this binary approximation, the time is divided into intervals, each with duration Δ. In each time interval $(i-1)\Delta \le t \le i\Delta$, the input waveform $X(t)$ is set to be a constant, which is equal to A with probability μ and is equal to 0 with probability $1-\mu$. Hence, μ can be viewed as the duty-cycle. Therefore, to satisfy the average power constraint (5), we require $\mu \le \sigma$. Let X^Δ be a binary random variable with

$$X^\Delta = \begin{cases} 1 & \text{if } X(t) = A, \\ 0 & \text{if } X(t) = 0. \end{cases} \tag{6}$$

It is clear that $\Pr\{X^\Delta = 1\} = 1 - \Pr\{X^\Delta = 0\} = \mu$.

At the receiver side, the receiver records only whether or not there is exactly one photon arriving during each time interval $(i-1)\Delta \le t \le i\Delta$. Let Y^Δ be a binary random variable whose value is 1 if the receiver observes one photon in the small interval Δ, and is 0 otherwise. Using (2), one can easily compute the transition probabilities

$$\begin{aligned} \Pr\{Y^\Delta = 1 | X^\Delta = 0\} &= \lambda\Delta e^{-\lambda\Delta}, \\ \Pr\{Y^\Delta = 1 | X^\Delta = 1\} &= (SA+\lambda)\Delta e^{-(\lambda+SA)\Delta}. \end{aligned} \tag{7}$$

It is easy to see that the capacity of the binary channel defined by $X^\Delta \to Y^\Delta$ is $\max_{0\le\mu\le\sigma} I(X^\Delta; Y^\Delta)$, and the normalized value $\frac{1}{\Delta}\max_{0\le\mu\le\sigma} I(X^\Delta; Y^\Delta)$ is an achievable rate for the original Poisson channel. Remarkably, [17] showed that this simple scheme is capacity achieving, and the capacity of the SISO Poisson non fading channel is given by

$$C_{SISO} = \lim_{\Delta\to 0} \frac{1}{\Delta} \max_{0\le\mu\le\sigma} I(X^\Delta; Y^\Delta). \tag{8}$$

Using (7), it was shown in [17] that

$$\begin{aligned} C_{SISO} = \max_{0\le\mu\le\sigma} \big[&\mu(SA+\lambda)\log(SA+\lambda) \\ &+ (1-\mu)\lambda\log\lambda - (\mu SA+\lambda)\log(\mu SA+\lambda)\big]. \end{aligned} \tag{9}$$

Intuitively, the first term in (9) corresponds to the case when the transmitter is on (i.e., $X(t) = A$ and the Poisson arrival rate at the receiver is $SA+\lambda$), which happens with probability μ. The second term corresponds to the case when the transmitter is off (i.e., $X(t) = 0$ and the Poisson arrival rate at the receiver is λ), which happens with probability $1-\mu$. The third term corresponds to the average case (i.e., the average Poisson arrival rate at the receiver is $\mu SA+\lambda$). The optimal value of μ can be easily obtained by solving the optimization problem (9).

B. Ergodic Capacity

In this section, we characterize the ergodic capacity of the Poisson fading channel. Following [17], it can be shown that the input $X(t)$ can be limited to be two levels without loss of optimality: either $X(t) = A$ or $X(t) = 0$. However, the probability that $X(t) = A$ can be adjusted depending on the noise level $\lambda(t)$. Let $\mu(t)$ be the time-varying duty cycle of the two-level waveform, then the average power constraint can be written as

$$\mathbb{E}[\mu(t)] \le \sigma. \tag{10}$$

Using (9), the ergodic capacity can be characterized as

$$C_{SISO}^f = \max_{\mu(t)} \mathbb{E}[I(\mu(t), \lambda(t))], \tag{11}$$

$$\text{s.t.} \quad \mathbb{E}[\mu(t)] \le \sigma, \tag{12}$$

$$0 \le \mu(t) \le 1, \tag{13}$$

in which the expectation is over λ and

$$\begin{aligned} &I(\mu(t), \lambda(t)) \\ &\quad = \mu(t)(SA+\lambda(t))\log(SA+\lambda(t)) \\ &\qquad + (1-\mu(t))\lambda(t)\log\lambda(t) \\ &\qquad - (\mu(t)SA+\lambda(t))\log(\mu(t)SA+\lambda(t)) \\ &\quad \triangleq \mu(t)\zeta(SA, \lambda(t)) - \zeta(\mu(t)SA, \lambda(t)), \end{aligned} \tag{14}$$

where $\zeta(x, y) = (x+y)\log(x+y) - y\log y$.

In the following, for notational convenience, we write $\mu(t)$ as μ. We characterize the optimal power allocation μ^{opt} for the constrained optimization problem (11)–(13).

For $0 \le \mu \le 1$, we note that $\frac{\partial^2 I(\mu,\lambda(t))}{\partial \mu^2}$ is negative, which implies that $I(\mu, \lambda(t))$ is a strictly concave function of μ in the range of our interest.

To obtain the optimal power allocation solution μ^{opt} for (11), we first consider the unconstrained version of (11) with (12) and (13) ignored. In particular, for any fading block with a given value of $\lambda(t)$, we examine the maximal rate that the channel can support:

$$r_{max} = \max_{\mu} I(\mu, \lambda(t)). \tag{15}$$

Let μ^o be the corresponding maximizer. For this unconstrained problem, it is easy to obtain that

$$\mu^o = \frac{\left(1+\frac{\lambda(t)}{SA}\right)^{\left(1+\frac{\lambda(t)}{SA}\right)}}{\left(\frac{\lambda(t)}{SA}\right)^{\left(\frac{\lambda(t)}{SA}\right)}} e^{-1} - \frac{\lambda(t)}{SA}. \tag{16}$$

Now, we examine (16) in detail. We first have the following result, which can be proved easily.

Lemma 1:

$$0 \le \mu^o \le 1. \tag{17}$$

This result implies that μ^o satisfies the constraint (13).

In the following, we consider the constraint (12). We have two cases.

Case 1): If

$$\mathbb{E}[\mu^o] \le \sigma. \tag{18}$$

In this case, μ^o also satisfies condition (12). Since $I(\mu, \lambda(t))$ is a strictly concave function of μ in the range of our interest, it is clear that μ^o is the maximizer for the original problem with constraints. That is

$$\mu^{opt} = \mu^o \text{ and } C^f_{SISO} = \mathbb{E}[r_{max}]. \tag{19}$$

Case 2): If

$$\mathbb{E}[\mu^o] > \sigma. \tag{20}$$

In this case, μ^o does not satisfy the condition (12). Hence μ^o is not the maximizer for the problem (11) with the average power constraint. To obtain the optimal solution for (11), we consider the Lagrangian function:

$$\mathcal{L}(\mu, \eta) = \mathbb{E}[\psi(\mu)] \triangleq \mathbb{E}[I(\mu, \lambda(t)) - \eta\mu]. \tag{21}$$

Since $I(\mu, \lambda(t))$ is a strictly concave function of μ and $-\eta\mu$ is a linear function of μ, we know that $\psi(\mu)$ is a strictly concave function of μ. From the Euler-Lagrange equation:

$$\frac{\partial \psi(\mu)}{\partial \mu} = SA\left[-\left(\log\left(\mu+\frac{\lambda(t)}{SA}\right)+1\right) + \left(1+\frac{\lambda(t)}{SA}\right)\log\left(1+\frac{\lambda(t)}{SA}\right) - \frac{\lambda(t)}{SA}\log\frac{\lambda(t)}{SA}\right] - \eta = 0,$$

we have

$$\mu_\eta = \frac{\left(1+\frac{\lambda(t)}{SA}\right)^{\left(1+\frac{\lambda(t)}{SA}\right)}}{\left(\frac{\lambda(t)}{SA}\right)^{\left(\frac{\lambda(t)}{SA}\right)}} e^{-1} e^{\frac{-\eta}{SA}} - \frac{\lambda(t)}{SA}. \tag{22}$$

Now, we consider the constraint $0 \le \mu \le 1$. It is easy to see that $\mu_\eta \le \mu^o$ for any positive η. From Lemma 1, we know that $\mu_\eta \le 1$. However, μ_η might be smaller than 0. If this occurs, from the fact that $\psi(\mu)$ is a strictly concave function of μ, we know that $\psi(\mu)$ is a strictly decreasing function of μ in the range $\mu > 0$. Hence, if $\mu_\eta < 0$, the constraint $\mu \ge 0$ implies that $\mu^{opt} = 0$. If $\mu_\eta > 0$, we have $\mu^{opt} = \mu_\eta$. Hence, we can write $\mu^{opt} = \mu^+_{\eta^*}$, in which η^* should be chosen such that $\mathbb{E}[\mu^+_{\eta^*}] = \sigma$.

We summarize the above analysis with the following Theorem.

Theorem 2: The optimal power allocation scheme that solves (11) (i.e., achieves the ergodic capacity of the Poisson fading channel) is given by

$$\mu^{opt} = \begin{cases} \mu^+_{\eta^*} & \text{if } \mathbb{E}[\mu^o] > \sigma \\ \mu^o & \text{if } \mathbb{E}[\mu^o] \le \sigma. \end{cases} \tag{23}$$

C. Outage Probability

In this section, we study the scenario with a strict delay constraint. In particular, we assume that each codeword needs to be transmitted within a fading block. Let r_0 be the target rate. Then an outage event occurs if $I(\mu, \lambda(t)) < r_0$. Our goal is to find the optimal power allocation strategy that minimizes the outage probability. Hence, we solve the following optimization problem.

Problem-1 (P1):

$$\min \quad \Pr\{I(\mu, \lambda(t)) < r_0\}, \tag{24}$$

$$\text{s.t.} \quad \mathbb{E}[\mu] \le \sigma, \tag{25}$$

$$0 \le \mu \le 1. \tag{26}$$

Again, we use μ^{opt} to denote the solution for this optimization problem.

Directly solving **P1** is challenging. Following a similar strategy as used for the Gaussian channel [16], we first solve several related optimization problems, from which we obtain the optimal solution for (24).

In the first step, we examine the maximal rate r_{max}, obtained in (15), that the channel can support for any given value of $\lambda(t)$. Following the discussion on (15), the optimal duty cycle that achieves r_{max} is μ^o given in (16). From Lemma 1 it is clear that $0 \le \mu^o \le 1$. We now compare r_{max} with r_0. Intuitively, for a given block with $\lambda(t)$, if $r_{max} < r_0$, then we should not transmit at this fading block and save the power for the future use, because an outage event will occur no matter what the value of μ we choose. On the other hand, if $r_{max} > r_0$, there exist choices of μ such that the achievable rate is larger than r_0. These are the regions of interest in the second step below.

In the second step, we investigate for any given value of $\lambda(t)$, what is the minimal value of μ that achieves the target rate r_0 or higher.

Problem-2 (P2):

$$\min \quad \mu, \tag{27}$$
$$\text{s.t.} \quad I(\mu, \lambda(t)) \geq r_0. \tag{28}$$

Let $\hat{\mu}$ be the minimizer for **P2**. Clearly, for those values of $\lambda(t)$s such that r_{max} is less than r_0, **P2** does not have a solution. For other values of $\lambda(t)$s, using the solution μ^o specified in (16), we have the following two cases:

- If $r_{max} = r_0$, then $\hat{\mu}$ must be equal to μ^o.
- If $r_{max} > r_0$, then the optimal power $\hat{\mu}$ equals $\breve{\mu}$ that satisfies the following equation so as to reduce the power consumption

$$r_0 = I\left(\breve{\mu}(\lambda), \lambda(t)\right), \tag{29}$$

from which $\breve{\mu}$ can be solved easily.

With the solutions to **P2**, we can obtain the optimal solution for an unconstrained version of **P1**. In the unconstrained versionof **P1**, the average power constraint is ignored. It is easy to see that

$$\mu^*(\lambda) = \begin{cases} 0 & \text{if } r_{max} < r_0 \\ \mu^o & \text{if } r_{max} = r_0 \\ \breve{\mu} & \text{if } r_{max} > r_0. \end{cases} \tag{30}$$

is a solution for the unconstrained version of **P1**. It is also clear that $0 \leq \mu^*(\lambda) \leq 1$ due to Lemma 1.

Now, we include the average power constraint into consideration.

Theorem 3: The optimal power allocation μ^{opt} of **P1** that achieves the smallest outage probability is characterized as follows.

If $\mathbb{E}[\mu^*] \leq \sigma$, then $\mu^{opt} = \mu^*$.

If $\mathbb{E}[\mu^*] > \sigma$, then μ^{opt} is given by

$$\mu^{opt}(\lambda) = \begin{cases} \mu^*(\lambda) & \textit{with probability } \hat{w}, \\ 0 & \textit{with probability } (1-\hat{w}), \end{cases} \tag{31}$$

where $\hat{w}$ is given by

$$\hat{w} = \begin{cases} 1 & \textit{if } \mu^*(\lambda) < p^* \\ w^* & \textit{if } \mu^*(\lambda) = p^* \\ 0 & \textit{if } \mu^*(\lambda) > p^*, \end{cases} \tag{32}$$

with

$$w^* = \frac{\sigma - \Sigma(p^*)}{\overline{\Sigma}(p^*) - \Sigma(p^*)}, \tag{33}$$

and

$$p^* = sup\{p : \Sigma(p) < \sigma\}, \tag{34}$$

and

$$\Sigma(p) = \int_{R(p)} \mu dF(\lambda),$$
$$\overline{\Sigma}(p) = \int_{\overline{R}(p)} \mu dF(\lambda),$$

Here $F(\lambda)$ is the distribution of the dark current and the regions are defined by:

$$R(p) = \{\lambda \in \mathbb{R} \setminus r_0 : \mu < p\},$$
$$\overline{R}(p) = \{\lambda \in \mathbb{R} \setminus r_0 : \mu \leq p\}.$$

Proof: The proof follows similar steps as those in [16] and [5], and is omitted for brevity. □

Theorem 3 implies that if $\mu^*(\lambda)$ specified in (30) does not satisfy the average power constraint, the optimal solution can be obtained by setting it to be equal to μ^* with a probability $\hat{w}$, and we should choose $\hat{w}$ properly so that the average power constraint is satisfied.

IV. MIMO Channel Analysis

For the ease of the presentation, we will present the results for $M = 1$ and $N = 2$ in details. We will briefly discuss how to extend the results to the case with general values of M and N in Section IV-D.

A. With Constant λ

In this subsection, we will first focus on the case in which λ is a constant. Hence, the power allocation is only among transmitters. The solution to this problem will be used to study the general case with time-varying λ.

As discussed in Section III, without loss of optimality, the input of a single antenna Poisson channel can be limited to two levels, i.e., $X(t) = A$ or $X(t) = 0$. The same idea can be extended to the channel with two transmitter antennas and the input of each antenna can be restricted to two levels, i.e. $X_n(t) = A_n$ or $X_n(t) = 0$. Let μ_n be duty-cycle of antenna n, i.e., the probability of $X_n(t) = A_n$. Similar to Section III, the capacity of the Poisson channel with two transmitter antennas, one receiver antenna, and a fixed dark current level with a total power constraint is given by

$$C_{MIMO} = \max_{\substack{\mu_1 A_1 + \mu_2 A_2 \leq \sigma(A_1+A_2) \\ 0 \leq \mu_1 \leq 1 \\ 0 \leq \mu_2 \leq 1}} I(\mu_1, \mu_2, \lambda), \tag{35}$$

in which

$$\begin{aligned} &I(\mu_1, \mu_2, \lambda) \qquad (36) \\ &\triangleq (\mu_1 - \kappa)(S_1A_1 + \lambda)\log(S_1A_1 + \lambda) \\ &\quad + (\mu_2 - \kappa)(S_2A_2 + \lambda)\log(S_2A_2 + \lambda) \\ &\quad + \kappa(S_1A_1 + S_2A_2 + \lambda)\log(S_1A_1 + S_2A_2 + \lambda) \\ &\quad + (1 - (\mu_1 + \mu_2 - \kappa))\lambda\log\lambda \\ &\quad - (\mu_1S_1A_1 + \mu_2S_2A_2 + \lambda)\log(\mu_1S_1A_1 + \mu_2S_2A_2 + \lambda) \\ &= (\mu_1 - \kappa)\zeta(S_1A_1, \lambda) + (\mu_2 - \kappa)\zeta(S_2A_2, \lambda) \\ &\quad + \kappa\zeta(S_1A_1 + S_2A_2, \lambda) - \zeta(\mu_1S_1A_1 + \mu_2S_2A_2, \lambda), \qquad (37) \end{aligned}$$

where κ is the probability with which both antenna 1 and 2 remain active. Each term in (36) has the same interpretation as the corresponding formula for the single antenna case (9). In particular, in (36), the first term corresponds to the case when only antenna 1 is active (i.e. $X_1(t) = A_1$ and the Poisson arrival rate at the receiver is $S_1A_1 + \lambda$), which happens with probability $\mu_1 - \kappa$. The second term corresponds to the case where only antenna 2 is active (i.e. $X_2(t) = A_2$ and the Poisson arrival rate at the receiver is $S_2A_2 + \lambda$), which happens with probability $\mu_2 - \kappa$. The third term corresponds to the case when both antennas 1 and 2 are active (i.e. $X_1(t) = A_1$ and $X_2(t) = A_2$ and the Poisson arrival rate at the receiver is $S_1A_1 + S_2A_2 + \lambda$), which happens with probability κ and the fourth term corresponds to the case when both of the transmitters are off and only the dark current is observed at the receiver. The last term corresponds to the average case (i.e. the average Poisson arrival rate at the receiver is $\mu_1S_1A_1 + \mu_2S_2A_2 + \lambda$).

Unlike the single antenna case in (9), it needs a bit of work to solve the optimization problem (35). First, using the property of $\zeta(x, y)$, it has been proved in [6], that if the antenna with the smaller duty cycle is on (i.e., the antenna with the smaller value of μ_n), the other antenna should also be active for the optimality. This implies that $\kappa = \min\{\mu_1, \mu_2\}$. Hence to calculate the optimal solution of (35), we consider the following two cases:

Case 1): $\mu_1 \geq \mu_2$. In this case, (37) can be simplified as

$$\begin{aligned} I(\mu_1 - \mu_2, \mu_2, \lambda) = &(\mu_1 - \mu_2)\zeta(S_1A_1, \lambda) \\ &+ \mu_2\zeta(S_1A_1 + S_2A_2, \lambda) \\ &- \zeta(\mu_1S_1A_1 + \mu_2S_2A_2, \lambda). \end{aligned} \tag{38}$$

Case 2): $\mu_1 \leq \mu_2$. In this case, (37) can be simplified as

$$\begin{aligned} I(\mu_2 - \mu_1, \mu_1, \lambda) = &(\mu_2 - \mu_1)\zeta(S_2A_2, \lambda) \\ &+ \mu_1\zeta(S_1A_1 + S_2A_2, \lambda) \\ &- \zeta(\mu_1S_1A_1 + \mu_2S_2A_2, \lambda). \end{aligned} \tag{39}$$

Hence, (35) can be written as

$$C_{MIMO} = \max\{C_{\mu_1\geq\mu_2}, C_{\mu_2\geq\mu_1}\}, \tag{40}$$

in which $C_{\mu_1\geq\mu_2}$ is given by (corresponds to case 1 above)

$$\begin{aligned} C_{\mu_1\geq\mu_2} \quad &= \quad \max I(\mu_1 - \mu_2, \mu_2, \lambda) && (41)\\ \text{s.t.} \quad &\quad \mu_1 - \mu_2 \geq 0, && (42)\\ &\quad \mu_1 \leq 1, && (43)\\ &\quad \mu_2 \geq 0, && (44)\\ &\quad \mu_1A_1 + \mu_2A_2 \leq \sigma(A_1 + A_2), && (45) \end{aligned}$$

and $C_{\mu_2\geq\mu_1}$ corresponds to case 2 above and can be written in a similar manner.

Hence, to solve (35), we need to find $C_{\mu_1\geq\mu_2}$ and $C_{\mu_2\geq\mu_1}$. Due to the symmetry, one can focus on case 1 and solve (41), as case 2 is similar. The solution to $C_{\mu_1\geq\mu_2}$ can be obtained using following steps. First, we will solve (41) by ignoring the total power constraint (45). Then, we check whether the obtained solution satisfies the total power constraint or not. If yes, the solution is optimal. If not, we then need to do further calculation.

Step 1: Following the strategy outlined above, we first consider the following optimization problem

$$\begin{aligned} C_{\mu_1\geq\mu_2} \quad &= \quad \max I(\mu_1 - \mu_2, \mu_2, \lambda) && (46)\\ \text{s.t.} \quad &\quad \mu_1 - \mu_2 \geq 0, && (47)\\ &\quad \mu_1 \leq 1, && (48)\\ &\quad \mu_2 \geq 0. && (49) \end{aligned}$$

To solve this problem, we further ignore (48) and (49). We will discuss at the end of step 1 that ignoring these conditions does not affect the solution.

Setting $x_1 = \mu_1 - \mu_2$ and $x_2 = \mu_2$, the Lagrangian for (46) with the constraint (47) only is given by:

$$\mathcal{L}(x_1, x_2, \eta) = I(x_1, x_2, \lambda) - \eta x_1.$$

Let (x_1^*, x_2^*, η^*) be the optimizer, and the corresponding KKT conditions are:

$$\begin{aligned} \left.\frac{\partial\mathcal{L}}{\partial x_1}\right|_{x_1^*,x_2^*} = \left.\frac{\partial I}{\partial x_1}\right|_{x_1^*,x_2^*} - \eta^* = 0,\\ \left.\frac{\partial\mathcal{L}}{\partial x_2}\right|_{x_1^*,x_2^*} = \left.\frac{\partial I}{\partial x_2}\right|_{x_1^*,x_2^*} = 0,\\ \eta^* x_1^* = 0. \end{aligned}$$

After solving the KKT conditions, we conclude that $x_1^* = 0$ and x_2^* can be obtained from

$$\left.\frac{\partial I}{\partial x_2}\right|_{x_1^*=0,x_2^*} = \lambda B \log\left(\frac{1 + B\alpha(B)}{1 + x_2^*B}\right) = 0,$$

where

$$\alpha(x) = \frac{1}{x}\left(e^{-1}(1 + x)^{\left(1+\frac{1}{x}\right)} - 1\right), \tag{50}$$

$$B = \frac{S_1A_1 + S_2A_2}{\lambda}. \tag{51}$$

Hence, $x_2^* = \alpha(B)$.

In summary, the solution to (46) with only the constraint (47) is $\mu_1^* = \mu_2^* = \alpha(B)$. It is easy to show that $0 \leq \alpha(B) \leq 1$. Hence, $\mu_1^* = \mu_2^* = \alpha(B)$ satisfy conditions (48) and (49). As the result, $(\alpha(B), \alpha(B))$ is the solution to (46) with constraints (47)–(49).

Step 2: In this step, we follow similar steps as those in [6] to check whether the solution $(\alpha(B), \alpha(B))$ obtained in Step 1 satisfies the sum power constraint (45) or not. If $\alpha(B) \leq \sigma$, then $(\alpha(B), \alpha(B))$ also satisfies the sum power constraint, hence is the optimal solution to (41). On the other hand, if $\alpha(B) > \sigma$, $(\alpha(B), \alpha(B))$ violates the sum power constraint, and hence can not be the solution. In this case, the sum power constraint is active in the optimal solution. To solve for optimality, we convert the given problem into a single variable optimization problem: $\max_{\mu_1^*} I(\mu_1^*) = I((1 + a)(\mu_1^* - \sigma), (1 + a)\sigma - a\mu_1^*, \lambda)$,

which is obtained by writing μ_2^* as function of μ_1^* using the average power constraint $\mu_1^* A_1 + \mu_2^* A_2 = \sigma(A_1 + A_2)$. Here, $a = A_1/A_2$ and μ_1^* is constrained to be between σ and $\left(1 + \frac{1}{a}\right)\sigma$. By simple calculations, the optimal solution for $\max_{\mu_1^*} I(\mu_1^*)$ is found out to be:

$$\omega = \frac{e^{-1}}{A_1(S_1 - S_2)} \times \left(\frac{(S_1 A_1 + \lambda)^{(S_1 A_1 + \lambda)(1+a)}}{\lambda^{\lambda}(S_1 A_1 + S_2 A_2 + \lambda)^{(S_1 A_1 + S_2 A_2 + \lambda)a}} \right)^{\frac{1}{A_1(S_1 - S_2)}} - \frac{(1+a)\sigma S_2 A_2 + \lambda}{A_1(S_1 - S_2)}. \quad (52)$$

Now, we check whether ω satisfies the corresponding constraint or not. If yes, ω is the optimal solution to $I(\mu_1^*)$ (therefore, it is also optimal for (41)). If not, we need to modify the solution. First, it is easy to check that $I(\mu_1^*)$ is a concave function of μ_1^*. Hence, if $\omega < \sigma$, we set $\mu_1^* = \sigma$, which implies $\mu_2^* = \sigma$. Similarly, if $\omega > (1 + 1/a)\sigma$, we set $\mu_1^* = (1 + 1/a)\sigma$, which implies $\mu_2^* = 0$.

We have the following Lemma regarding the optimal duty cycle that maximize $C_{\mu_1 \geq \mu_2}$.

Lemma 4: The optimal solution to the optimization problem (41) is given by

$$(\mu_1^*, \mu_2^*) = \begin{cases} (\alpha(B), \alpha(B)) & \text{If } \alpha(B) \leq \sigma, \\ (\omega, (1+a)\sigma - a\omega) & \text{If } \alpha(B) > \sigma \text{ and } \sigma \leq \omega \leq (1 + 1/a)\sigma, \\ (\sigma, \sigma) & \text{If } \alpha(B) > \sigma \text{ and } \omega < \sigma, \\ ((1 + 1/a)\sigma, 0) & \text{If } \alpha(B) > \sigma \text{ and } \omega > (1 + 1/a)\sigma. \end{cases}$$

We can obtain $C_{\mu_2 \geq \mu_1}$ in a similar manner and therefore finally obtain C_{MIMO} via (40).

B. Ergodic Capacity

We now study the case in which $\lambda(t)$ is a time varying random variable and characterize the ergodic capacity. We derive the optimal power allocation strategy using the results developed Section IV-A. To maximize the ergodic capacity in the presence of the time-varying dark current, we have:

$$C_{MIMO}^f = \max_{\substack{\mathbb{E}[\mu_1(t)A_1 + \mu_2(t)A_2] \leq \sigma(A_1 + A_2) \\ 0 \leq \mu_1(t) \leq 1 \\ 0 \leq \mu_2(t) \leq 1}} \mathbb{E}[\Phi(t)], \quad (53)$$

where the expectation is over random variable $\lambda(t)$ and $\Phi(t) = I(\mu_1(t), \mu_2(t), \lambda(t))$ and we use $(\mu_1^{opt}, \mu_2^{opt})$ to denote the maximizer for (53). We follow the same strategy developed in the SISO case, and use a two-step approach to solve this problem.

Step 1: First, we ignore the average power constraint, and find the maximal rate that the channel can support for any given value of $\lambda(t)$.

$$r_{max} = \max I(\mu_1, \mu_2, \lambda(t)), \quad (54)$$
$$\text{s.t.} \quad 0 \leq \mu_1 \leq 1, \quad (55)$$
$$0 \leq \mu_2 \leq 1. \quad (56)$$

We use (μ_1^o, μ_2^o) to denote the maximizer of this problem.

As shown in Section IV-A, for any given value of $\lambda(t)$,

$$\max_{\mu_1, \mu_2} I(\mu_1, \mu_2, \lambda(t)) = \max\{\Theta_1(t), \Theta_2(t)\},$$

where $\Theta_1(t) = \max_{\mu_1, \mu_2} I(\mu_1 - \mu_2, \mu_2, \lambda(t))$ and $\Theta_2(t) = \max_{\mu_1, \mu_2} I(\mu_2 - \mu_1, \mu_1, \lambda(t))$. Therefore, problem (54) can be split into two separate problems. Due to the symmetry, we focus on $\Theta_1(t)$ (the other case being similar), which can be written as:

$$\max \quad I(\mu_1 - \mu_2, \mu_2, \lambda(t)), \quad (57)$$
$$\text{s.t.} \quad \mu_1 \leq 1, \quad (58)$$
$$\mu_2 \geq 0, \quad (59)$$
$$\mu_1 \geq \mu_2. \quad (60)$$

For a given value of $\lambda(t)$, problem (57) is the same as problem (46). Following the analysis of (46), it is clear that $\mu_1 = \mu_2 = \alpha(B(t))$ is the solution to (57), where $\alpha(\cdot)$ is defined in (50) and $B(t)$ is defined in (51) with λ replaced by $\lambda(t)$. Similarly, $\mu_1 = \mu_2 = \alpha(B(t))$ is also the solution to problem $\Theta_2(t)$. As the result, the solution to problem (54) is $(\mu_1^o, \mu_2^o) = (\alpha(B(t)), \alpha(B(t)))$.

Step 2: Now we check whether the solution (μ_1^o, μ_2^o) obtained in Step 1 satisfies the average power constraint or not, i.e., we check whether the following inequality holds or not:

$$\mathbb{E}\left[\mu_1^o A_1 + \mu_2^o A_2\right] \leq \sigma(A_1 + A_2). \quad (61)$$

If (61) holds, the optimal solution to (53) is $(\mu_1^{opt}, \mu_2^{opt}) = (\mu_1^o, \mu_2^o)$, and $C_{MIMO}^f = \mathbb{E}[r_{max}]$.

If (61) does not hold, then (μ_1^o, μ_2^o) can not be the optimal solution to (53), as it violates the average power constraint. Therefore, in the optimal solution, the average power constraint is active, and hence problem (53) is equivalent to:

$$C_{MIMO}^f = \max_{\substack{\mathbb{E}[\sigma'(t)] = \sigma \\ 0 \leq \sigma'(t) \leq 1}} \mathbb{E}[\Psi(t)], \quad (62)$$

where $\Psi(t) = \max_{\substack{\mu_1(t)A_1 + \mu_2(t)A_2 = \sigma'(t)(A_1 + A_2) \\ 0 \leq \mu_1(t) \leq 1 \\ 0 \leq \mu_2(t) \leq 1}} \Phi(t)$. We note that, in the optimization problem inside $\mathbb{E}[\cdot]$, the value of $\lambda(t)$ is fixed, and hence the results in Section IV-A are applicable. More specifically, we use results in step 2 of Section IV-A, because the sum power constraint is active here. Using the results in Section IV-A, we can write the mutual information term inside $\mathbb{E}[\cdot]$ as a function of $\sigma'(t)$, and hence the problem is converted into an optimization problem of $\sigma'(t)$. Due to its complex form (as shown in Lemma 4), it is difficult to obtain an analytical form of the optimal solution. However, one can set $\mu_1(t) =$

$\mu_2(t)$ and obtain a computable solution, which is optimal when σ is large.

In summary, we have the following proposition regarding the optimal power allocation scheme for the ergodic capacity.

Proposition 5: The optimal power allocation strategy that achieves the ergodic capacity (namely the optimization problem in (53)) is given by

$$\left(\mu_1^{opt}, \mu_2^{opt}\right) = \begin{cases} \left(\alpha\left(B(t)\right), \alpha\left(B(t)\right)\right) & \text{If } \mathbb{E}\left[\alpha\left(B(t)\right)\right] \le \sigma, \\ \text{solution of (62)} & \text{If } \mathbb{E}\left[\alpha\left(B(t)\right)\right] > \sigma. \end{cases} \tag{63}$$

As we can see here, when the average power constraint σ is large, we obtain the closed form expression for the optimal allocation scheme. When σ is small, we do not have the closed form solution as the form of the function is too complicated. Alternatively, suboptimal solutions with good properties can be numerically computed easily as discussed in the paragraph after (62).

C. Outage Probability

In this section, we consider the case with a strict delay constraint, for which the strategy developed in Section III-C for the SISO case is useful. Let r_0 be the target rate, hence an outage event occurs if $I(\mu_1, \mu_2, \lambda(t)) < r_0$. The goal is to minimize the outage probability.

Problem-1-*MIMO* (P1-M):

$$\min \quad \Pr\left\{I\left(\mu_1, \mu_2, \lambda(t)\right) < r_0\right\}, \tag{64}$$

$$\text{s.t.} \quad \mathbb{E}[\mu_1 A_1 + \mu_2 A_2] \le \sigma(A_1 + A_2). \tag{65}$$

We use $(\mu_1^{opt}, \mu_2^{opt})$ to denote the minimizer of this problem.

Similar to the SISO case studied in Section III, we first solve several related optimization problems, which help to obtain the optimal solution for **P1-M**.

We first examine the maximal rate r_{max} that the channel can support for any given value of $\lambda(t)$, namely, the optimization problem (54). Following Step 1 in Section IV-B, $(\mu_1^o, \mu_2^o) = (\alpha(B(t)), \alpha(B(t)))$ is the power allocation strategy that achieves r_{max} for each block.

Similar to Section III, we then compare r_{max} with r_0. If $r_{max} < r_0$, then the transmitter should not transmit anything and save the power for future use. If $r_{max} = r_0$, then the only choice of power control that avoids outage is the power that achieves r_{max}. The interesting case is when $r_{max} > r_0$. In this case, there are multiple (in fact infinitely many) power control choices that can avoid outage. In the following, we find the minimal sum power that avoids the outage.

Problem-2-*MIMO* (P2-M):

$$\min \quad \mu_1 A_1 + \mu_2 A_2, \tag{66}$$

$$\text{s.t.} \quad I\left(\mu_1, \mu_2, \lambda(t)\right) \ge r_0. \tag{67}$$

We use $(\hat{\mu}_1, \hat{\mu}_2)$ to denote the optimal solution to **P2-M**.

For those $\lambda(t)$s such that r_{max} is less than r_0, **P2-M** does not have a solution. For other values of $\lambda(t)$s, we consider the following two cases:

- If $r_{max} = r_0$ then $\hat{\mu}_1 = \hat{\mu}_2 = \alpha(B(t))$.
- If $r_{max} > r_0$ then $(\hat{\mu}_1, \hat{\mu}_2) = (\check{\mu}_1, \check{\mu}_2)$, such that $(\check{\mu}_1, \check{\mu}_2)$ is the solution of the following problem **P2a-M**.

Problem-2a-*MIMO* (P2a-M):

$$\min \quad \mu_1 A_1 + \mu_2 A_2, \tag{68}$$

$$\text{s.t.} \quad I\left(\mu_1, \mu_2, \lambda(t)\right) = r_0. \tag{69}$$

To solve this problem, we consider two subproblems.

The first subproblem is:

P2a-M-sub1:

$$\min \quad \mu_1 A_1 + \mu_2 A_2, \tag{70}$$

$$\text{s.t.} \quad I\left(\mu_1 - \mu_2, \mu_2, \lambda(t)\right) = r_0, \tag{71}$$

$$\mu_1 \ge \mu_2. \tag{72}$$

Let $(\tilde{\mu}_1, \tilde{\mu}_2)$ be the solution.

The second subproblem of **P2a-M** is

P2a-M-sub2:

$$\min \quad \mu_1 A_1 + \mu_2 A_2, \tag{73}$$

$$\text{s.t.} \quad I\left(\mu_2 - \mu_1, \mu_1, \lambda(t)\right) = r_0, \tag{74}$$

$$\mu_2 \ge \mu_1. \tag{75}$$

Let $(\bar{\mu}_1, \bar{\mu}_2)$ be the solution. Following the similar approach in Corollary 1 of [5], the solutions to these sub problems can be found using KKT conditions.

As $I(\mu_1, \mu_2, \lambda) = \max\{I(\mu_1 - \mu_2, \mu_2, \lambda), I(\mu_2 - \mu_1, \mu_1, \lambda)\}$, then from the solutions to the two subproblems **P2a-M-sub1** and **P2a-M-sub2**, the optimal solution of **P2a-M** is given by

$$(\check{\mu}_1, \check{\mu}_2) = \begin{cases} (\tilde{\mu}_1, \tilde{\mu}_2) & \text{if } \tilde{\mu}_1 A_1 + \tilde{\mu}_2 A_2 \le \bar{\mu}_1 A_1 + \bar{\mu}_2 A_2, \\ (\bar{\mu}_1, \bar{\mu}_2) & \text{otherwise.} \end{cases}$$

Therefore, the solution to problem **P-1M** with the absence of the average power constraint can be written as:

$$\left(\mu_1^*(\lambda), \mu_2^*(\lambda)\right) = \begin{cases} (0, 0) & \text{if } r_{max}(\lambda) < r_0 \\ \left(\mu_1^o, \mu_2^o\right) & \text{if } r_{max}(\lambda) = r_0 \\ (\check{\mu}_1, \check{\mu}_2) & \text{if } r_{max}(\lambda) > r_0. \end{cases} \tag{76}$$

Now, we check the average power constraint. Depending on whether (μ_1^*, μ_2^*) satisfies $\mathbb{E}[\mu_1^* A_1 + \mu_2^* A_2] \le \sigma(A_1 + A_2)$ or not, we have the following two cases.

- **Case 1**): $\mathbb{E}[\mu_1^* A_1 + \mu_2^* A_2] \le \sigma(A_1 + A_2)$. In this case, we have $(\mu_1^{opt}, \mu_2^{opt}) = (\mu_1^*, \mu_2^*)$.
- **Case 2**): $\mathbb{E}[\mu_1^* A_1 + \mu_2^* A_2] > \sigma(A_1 + A_2)$. In this case, we need to modify (μ_1^*, μ_2^*) to obtain $(\mu_1^{opt}, \mu_2^{opt})$. Following similar arguments as in Section III, we can conclude that, if $\mathbb{E}[\mu_1^* A_1 + \mu_2^* A_2] > \sigma(A_1 + A_2)$, then the optimal solution $(\mu_1^{opt}, \mu_2^{opt})$ of **P1-M** is given by

$$\left(\mu_1^{opt}, \mu_2^{opt}\right) = \begin{cases} (\mu_1^*, \mu_2^*) & \text{with probability } \hat{w} \\ (0, 0) & \text{with probability } (1 - \hat{w}), \end{cases} \tag{77}$$

in which $\hat{w}$ has the following form

$$\hat{w} = \begin{cases} 1 & \text{if } \bar{\sigma}(t) < p^* \\ w^* & \text{if } \bar{\sigma}(t) = p^* \\ 0 & \text{if } \bar{\sigma}(t) > p^*, \end{cases} \tag{78}$$

where w^* is given by (33), p^* is given by (34) and $\bar{\sigma} = \frac{\mu_1^* A_1 + \mu_2^* A_2}{A_1 + A_2}$.

In summary, we have the following proposition regarding the optimal power allocation strategy that minimize the outage probability.

Proposition 6: The optimal power allocation strategy that minimizes the outage probability is given by

$$\left(\mu_1^{opt}, \mu_2^{opt}\right) = \begin{cases} \textit{specified in (76)} & \textit{If } \mathbb{E}\left[\mu_1^* A_1 + \mu_2^* A_2\right] \\ & \quad \leq \sigma(A_1 + A_2), \\ \textit{specified in (77)} & \textit{If } \mathbb{E}\left[\mu_1^* A_1 + \mu_2^* A_2\right] \\ & \quad > \sigma(A_1 + A_2). \end{cases}$$

D. Extension to General MIMO Case

For the case with an arbitrary number of transmit and receive antennas, i.e., general values of N and M, we can follow the similar steps as in the previous sections to obtain the optimal power control policy that maximizes the ergodic capacity and the optimal power control policy that minimizes the outage probability. In particular, for any N, if the transmitter antenna with the smallest duty cycle is active, then the other antennas should also be active. Hence, there are $N+1$ states. As the result, the mutual information $I(\mu^N(t), \lambda(t))$ has $N+1$ terms, each corresponding to one state. Thus in order to obtain the ergodic capacity, we solve

$$C_{MIMO}^f = \max_{\mathbb{E}[\sum_i \mu_i(t) A_i] \leq \sigma \sum_i A_i} \mathbb{E}\left[I\left(\mu^N(t), \lambda(t)\right)\right], \tag{79}$$

where

$$I\left(\mu^N(t), \lambda(t)\right) = \sum_{m=1}^{M} \left[\sum_{n=1}^{N} \nu_n \zeta \left(\sum_{k=1}^{n} S_{km} A_k, \lambda_m(t) \right) - \zeta \left(\sum_{n=1}^{N} \nu_n \sum_{k=1}^{n} S_{km} A_k, \lambda_m(t) \right) \right], \tag{80}$$

and when $\mu_n > \mu_{n+1}$, $n = 1, \ldots, N-1$,

$$\nu_n = \begin{cases} \mu_n - \mu_{n+1} & n = 1, \ldots, N-1 \\ \mu_N & n = N. \end{cases} \tag{81}$$

The solution of (79) can be obtained following the same approach as in Section IV-B by examining cases with different ordering of duty cycles.

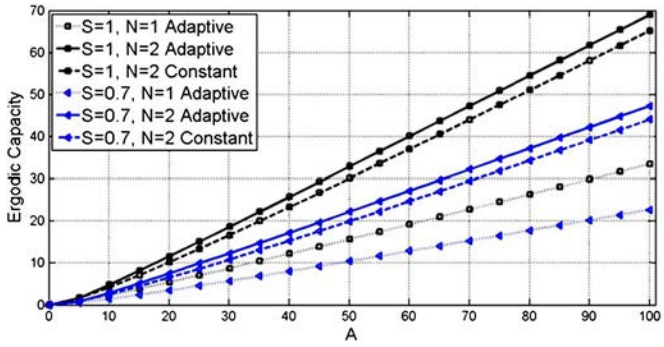


Fig. 2. The ergodic capacity vs. A.

To minimize the outage probability, we consider

$$\min \quad Pr\{I(\mu^N(t), \lambda(t)) < r_0\}, \tag{82}$$

$$s.t. \quad \mathbb{E}\left[\sum_i \mu_i(t) A_i\right] \leq \sigma \sum_i A_i, \tag{83}$$

which can be solved following the similar steps as explained in Section IV-C by examining cases with different ordering of duty cycles.

V. Numerical Results

In this section, we present several numerical examples to illustrate the results obtained in the previous two sections. In the simulations, we set $M = 1$, and consider two cases with $N = 1$ and $N = 2$, respectively. Furthermore, $\lambda(t)$ is chosen as a uniform random variable such that $\lambda(t) \sim \mathcal{U}[1, 2]$. For fair comparison between $N = 1$ and $N = 2$ case, we have ensured that both the total power constraints and average power constraints are equal by setting $A = 30$ for $N = 1$ and $A_1 = A_2 = 15$ for $N = 2$.

A. Ergodic Capacity

Fig. 2 illustrates the ergodic capacity as a function of the maximum amplitude for both $N = 1$ and $N = 2$, for $\lambda(t) \sim \mathcal{U}[1, 2]$. It is evident from the figure that as A increases, the ergodic capacity increases monotonically. If the number of antennas at the transmitter is larger, the increment in the ergodic capacity is more evident. For comparison, in the figure, we also plot curves for the ergodic capacity when the constant power allocation strategy is employed. The figure shows that the adaptive power allocation for multiple transmitting antennas observe significant improvement in the ergodic capacity when A is increased, as compared to the case with the constant power control for multiple transmitting antennas and adaptive power control for a single transmitting antenna.

Fig. 3 further illustrates the comparison between the ergodic rate achieved with and without adaptive power control. In this figure, we plot the rate as a function of σ for $A_1 = A_2 = 15$, $N = 2$ and $M = 1$ when $\lambda(t) \sim \mathcal{U}[0.5, 8.0]$. It can be observed from the figure that initially as σ increases, the ergodic capacity increases. But after reaching a certain threshold point,

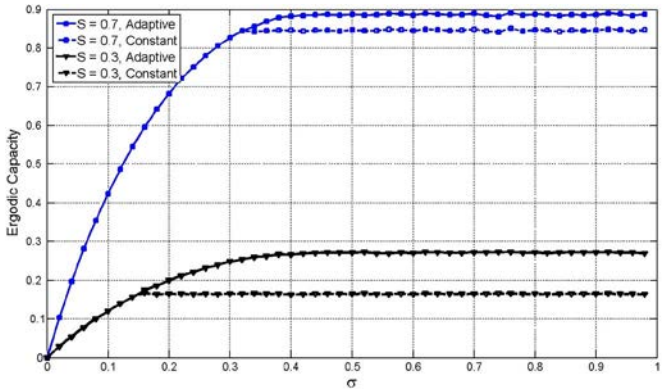


Fig. 3. The ergodic capacity vs. σ for $N = 2$, comparing the adaptive and the constant power allocation.

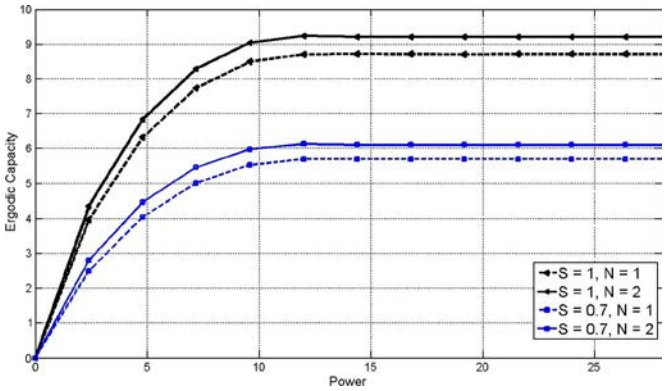


Fig. 4. The ergodic capacity vs. power for both multi-antenna and single-antenna cases.

increasing σ does not affect the ergodic capacity. This is due to the fact that once σ is large enough, as discussed in (63), $(\alpha(B(t)), \alpha(B(t)))$ satisfy the total power constraint $\mathbb{E}[\mu_1 A_1 + \mu_2 A_2] \leq \sigma(A_1 + A_2)$, and then the optimal power allocation strategy and the ergodic capacity do not depend on σ. It is shown in the figure that the gain achieved by adaptive power allocation is more obvious when $\mathbb{E}[\mu_1 A_1 + \mu_2 A_2] \leq \sigma(A_1 + A_2)$ is satisfied and for the smaller values of S. From the figure, it is clear that when σ is small, both the adaptive and constant power control have the same rate. This can be explained as follows. For the adaptive power control, solving (52) shows that for the given parameters, $\omega < 0$. Therefore, from Lemma 4, we know that the optimal power allocation strategy is $(\mu_1, \mu_2) = (\sigma, \sigma)$, which is the same as the constant power control case.

Fig. 4 illustrates the ergodic capacity as a function of the total power for both $N = 1$ and $N = 2$ with $A = 30$ and $A_1 = A_2 = 15$. As the total power (i.e. σA for $N = 1$ and $\sigma(A_1 + A_2)$ for $N = 2$) increases, the ergodic capacity also increases. As discussed above, after reaching a certain threshold, the increase in power does not affect the ergodic capacity.

Fig. 5 compares the ergodic capacity of two transmitter antennas and one receiver with the one transmitter and two receiver antennas while keeping the power constraints (i.e., total power and average power constraint) to be the same when $S = 0.5$. For both cases we assume that for each receiver, the corresponding dark current is uniformly distributed in the interval [1, 2]. It is clear from the figure that the

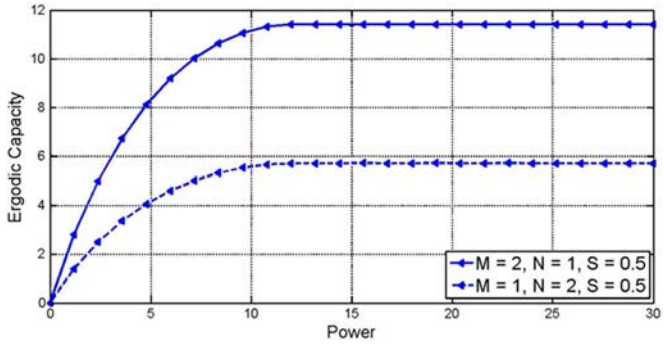


Fig. 5. The ergodic capacity vs. power for $M = 1$, $N = 2$ and $M = 2$, $N = 1$.

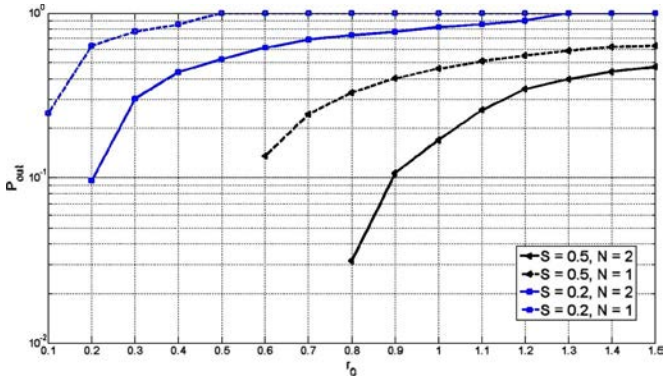


Fig. 6. P_{out} vs. r_0, comparing $N = 1$ with $N = 2$.

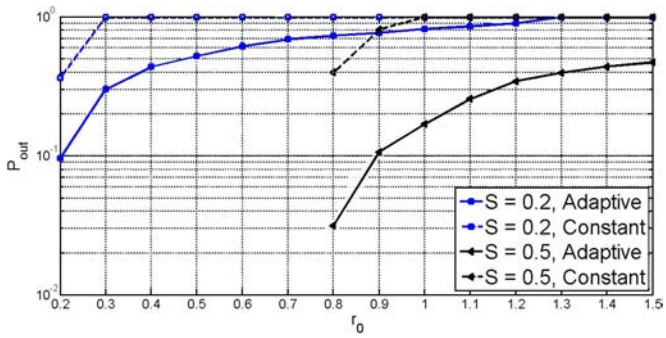


Fig. 7. P_{out} vs. r_0, comparing the adaptive power allocation with the constant power allocation.

two receiver antennas case has a better performance. This can be explained by the mutual information formulas. For $(M = 1,\ N = 2)$, when $\mu_1 \geq \mu_2$, the mutual information is given by $(\mu_1 - \mu_2)\zeta(S_1 A_1, \lambda) + \mu_2 \zeta(S_1 A_1 + S_2 A_2, \lambda) - \zeta(\mu_1 S_1 A_1 + \mu_2 S_2 A_2, \lambda)$, while for $(M = 2, N = 1)$, it is given by: $\mu\zeta(SA, \lambda_1) - \zeta(\mu SA, \lambda_1) + \mu\zeta(SA, \lambda_2) - \zeta(\mu SA, \lambda_2)$.

B. Outage Probability

Fig. 6 plots the outage probability as a function of the target rate r_0. For this simulation, $A = 30$, $A_1 = A_2 = 15$ and $\sigma = 0.03725$ for both $N = 1$ and $N = 2$, for the fair comparison between $N = 1$ and $N = 2$. It is evident from the figure that as r_0 increases, P_{out} also increases. Furthermore, the outage probability for $N = 1$ increases more rapidly than that of $N = 2$. As the figure is in logarithmic scale and log 0 is not defined, hence

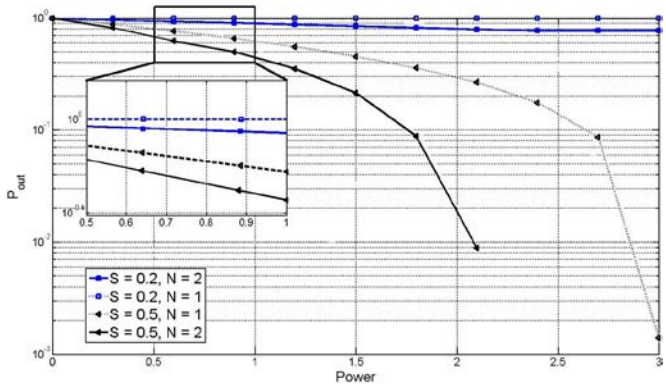


Fig. 8. P_{out} vs. power, comparing $N = 1$ with $N = 2$.

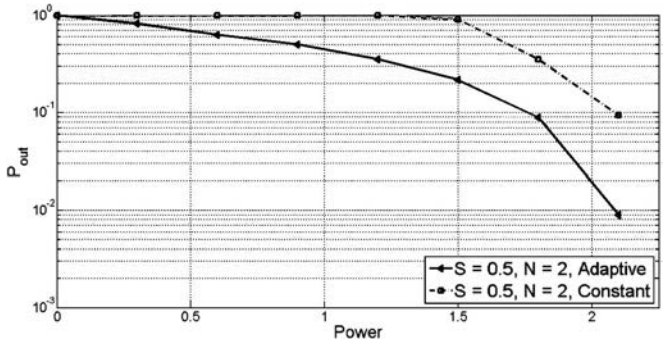


Fig. 9. P_{out} vs. power, comparing the adaptive power allocation with the constant power allocation.

the figure only shows the values when the outage probability is larger than 0. Fig. 7 compares the relationship between the target rate r_0 and the probability for the adaptive power allocation and the constant power allocation when $A_1 = A_2 = 15$ and $\sigma = 0.03725$. It shows that the outage probability of the adaptive power allocation responds gradually to the increase in the target rate as compared to the constant power allocation where the outage probability abruptly increases when r_0 increases. This figure also shows the values of outage probability when $P_{out} > 0$. Fig. 8 shows that as the power increases, P_{out} decreases and $r_0 = 1.2$. Furthermore, we can observe that, after a certain point, the increase in power does not lead to further decrease in P_{out}. Similar to the ergodic case, the reason for this phenomena is that once the available power is large enough, the optimal power allocation strategy and hence the outage probability does not depend on the available power anymore. From the figure, we also see that when $S = 0.2$, the outage probability for $N = 1$ is always 1. The value of $P_{out} = 0$ for $S = 0.5$, $N = 2$ when power is larger than 2.45 and $P_{out} = 0$ for $S = 0.5$, $N = 1$, when power is larger than 3.

Fig. 9 shows the improvement achieved by the multi-antenna dynamic power allocation as compared to multi-antenna constant power allocation and shows that when the available power increases, the adaptive power allocation scheme reduces the outage probability more significantly as compared to the constant power allocation for $A_1 = A_2 = 15$ and $r_0 = 1.2$. $P_{out} = 0$ when power is larger than 2.45.

Fig. 10 compares the outage probabilities for the cases with ($M = 1$, $N = 2$) and ($M = 2$, $N = 1$) with $S = 0.2$. From the

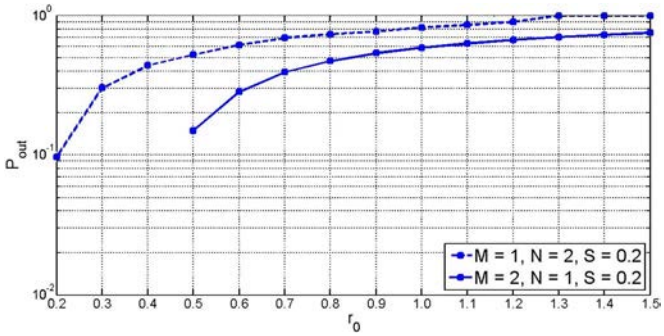


Fig. 10. P_{out} vs. r_0, Comparing $M = 1$, $N = 2$ with $M = 2$, $N = 1$ with adaptive power allocation.

figure, it is clear that P_{out} is higher for $M = 1$, $N = 2$ case than for the $M = 2$, $N = 1$, which is consistent with the mutual information formula and the comparison in Fig. 5.

VI. Conclusion

We have studied Poisson fading channels with varying noise levels. We have considered cases with and without strict delay constraints. For the case without a strict delay constraint, we have characterized the optimal power allocation scheme that achieves the ergodic capacity. For the case with a strict delay constraint, we have characterized the optimal power allocation strategy that minimizes the outage probability. We have also provided numerical results to illustrate the analytical results obtained in this paper.

References

[1] Ain-ul-Aisha, L. Lai, and Y. Liang, "On fading Poisson channels with varying noise levels," in *Proc. Conf. Inf. Sci. Syst.*, Princeton, NJ, USA, Mar. 2014, pp. 1–5.

[2] S. Hranilovic, *Wireless Optical Communication Systems*. New York: Springer, 2005.

[3] S. M. Haas, "Capacity of and coding for multiple-aperture, wireless, optical communications," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Mass. Inst. Technol., Cambridge, MA, USA, 2003.

[4] D. Guo, S. Shamai, and S. Verdú, "Mutual information and conditional mean estimation in Poisson channels," *IEEE Trans. Inf. Theory*, vol. 54, pp. 1837–1849, May 2008.

[5] K. Chakraborty and S. Dey, "Outage capacity of MIMO Poisson fading channels," *IEEE Trans. Inform. Theory*, vol. 54, pp. 4887–4907, Nov. 2008.

[6] K. Chakraborty, "Relaiable communication over optical fading channels," Ph.D. dissertation, Dept. Elect. Eng., Univ. Maryland, College Park, MD, USA, 2005.

[7] K. Chakraborty, "Capacity of the mimo optical fading channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Adelaide, Australia, Sep. 2005, pp. 530–534.

[8] K. Chakraborty and P. Narayan, "The Poisson fading channel," *IEEE Trans. Inf. Theory*, vol. 35, no. 3, pp. 2349–2364, Jul. 2007.

[9] S. M. Haas and J. H. Shapiro, "Capacity of wireless optical communications," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 8, pp. 1346–1357, Oct. 2003.

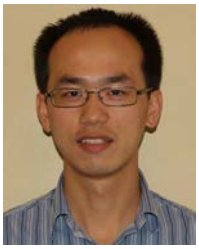
[10] E. J. Lee and V. W. S. Chan, "Part I: Optical communication over the clear turbulent atmospheric channel using diversity," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 9, pp. 1896–1906, Nov. 2004.

[11] Y. He and S. Dey, "Throughput maximization in Poisson fading channels with limited feedback," *IEEE Trans. Commun.*, vol. 61, no. 10, pp. 4343–4355, Oct. 2013.

[12] M. Sapari, M. M. Rad, and M. Uysal, "Multi-hop relaying over the atmospheric Poisson channel: Outage analysis & optimization," *IEEE Trans. Commun.*, vol. 60, no. 3, pp. 817–829, Mar. 2012.

[13] M. Alem-Karladani, L. Sepahi, M. Jazayerifar, and K. Kalbasi, "Optimum power allocation in parallel Poisson optical channel," in *Proc. Int. Conf. Telecommun.*, Zagreb, Croatia, Jun. 2009, pp. 285–288.
[14] S. Bross, A. Lapidoth, and L. Wang, "The Poisson channel with side information," in *Proc. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Oct. 2009, pp. 574–578.
[15] M. A. Kashani and M. Uysal, "Outage performance and diversity gain analysis of free-space optical multi-hop parallel relaying," *IEEE Trans. Commun.*, vol. 5, no. 8, pp. 901–909, Aug. 2013.
[16] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1468–1489, Jul. 1999.
[17] A. D. Wyner, "Capacity and error exponent for the direct detection photon channel—Parts I and II," *IEEE Trans. Inf. Theory*, vol. 34, pp. 1449–1471, Nov. 1988.

Ain-ul-Aisha received the B.E and M.E degrees from National University of Sciences and Technology, Islamabad, Pakistan, in 2008 and 2012, respectively. She is now pursuing the Ph.D. degree at Worcester Polytechnic Institute, Worcester, MA, USA. Her research interests include information theory, intelligent networks, and routing protocols for cognitive radio networks.



Lifeng Lai (M'07) received the B.E. and M.E. degrees from Zhejiang University, Hangzhou, China, in 2001 and 2004, respectively, and the Ph.D. degree from The Ohio State University at Columbus, Columbus, OH, USA, in 2007. He was a Postdoctoral Research Associate at Princeton University, Princeton, NJ, USA, from 2007 to 2009, and was an Assistant Professor at University of Arkansas, Little Rock, AR, USA, from 2009 to 2012. Since August 2012, he has been an Assistant Professor at Worcester Polytechnic Institute, Worcester, MA, USA. His research interests include information theory, stochastic signal processing and their applications in wireless communications, security, and other related areas.

Dr. Lai was a Distinguished University Fellow of The Ohio State University from 2004 to 2007. He is a co-recipient of the Best Paper Award from IEEE Global Communications Conference (GLOBECOM) in 2008, the Best Paper Award from IEEE Conference on Communications (ICC) in 2011, and the Best Paper Award from IEEE Smart Grid Communications (SmartGridComm) in 2012. He received the National Science Foundation CAREER Award in 2011, and Northrop Young Researcher Award in 2012. He served as a Guest Editor for IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, Special Issue on Signal Processing Techniques for Wireless Physical Layer Security. He is currently serving as an Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and an Associate Editor for IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.

Yingbin Liang (S'01–M'05) received the Ph.D. degree in electrical engineering from the University of Illinois at Urbana-Champaign, Urbana-Champaign, IL, USA, in 2005. In 2005–2007, she was a Postdoctoral Research Associate at Princeton University, Princeton, NJ, USA. In 2008–2009, she was an Assistant Professor at the Department of Electrical Engineering at the University of Hawaii, Honolulu, HI, USA. Since December 2009, she has been on the faculty at Syracuse University, Syracuse, NY, USA, where she is an Associate Professor. Her research interests include information theory, wireless communications and networks, and machine learning.

Dr. Liang was a Vodafone Fellow at the University of Illinois at Urbana-Champaign during 2003–2005, and received the Vodafone-U.S. Foundation Fellows Initiative Research Merit Award in 2005. She also received the M. E. Van Valkenburg Graduate Research Award from the ECE Department, University of Illinois at Urbana-Champaign, in 2005. In 2009, she received the National Science Foundation CAREER Award, and the State of Hawaii Governor Innovation Award. More recently, her paper received the 2014 EURASIP Best Paper Award for the *EURASIP Journal on Wireless Communications and Networking*. She is currently serving as an Associate Editor for the Shannon Theory of the IEEE TRANSACTIONS ON INFORMATION THEORY.