Throughput Maximization in Poisson Fading Channels with Limited Feedback

YuanYuan He, *Member, IEEE*, and Subhrakanti Dey, *Senior Member, IEEE*

***Abstract*—A shot-noise limited single-user single-input single-output (SISO) Poisson fading channel with partial channel state information (CSI) at the transmitter and perfect CSI at the receiver is considered. We address an optimal transmit power allocation problem that maximizes the ergodic capacity of the SISO Poisson fading channel subject to peak and average power constraints with only quantized CSI available at the transmitter, acquired via a no-delay and error-free feedback link with finite-rate from the receiver to the transmitter. Due to the non-convexity of the proposed optimization problem, a globally optimum solution is difficult to obtain. However, we manage to obtain a locally optimal quantized power allocation (QPA) scheme by solving its dual Lagrangian optimization problem. We develop two efficient optimal QPA algorithms for solving the dual optimization problem and show that both of these algorithms converge to the globally optimal solution of the dual problem. A low-complexity near-optimal QPA algorithm is also derived for the case of large number of feedback bits. The results are then extended to the high peak signal-to-noise ratio (SNR) regime and an explicit expression for the approximate asymptotic ergodic capacity behavior in the high SNR regime with high rate quantization (as the number of feedback bits goes to infinity) is also provided. It is seen via numerical simulations that this asymptotic capacity expression correctly approaches the capacity of the corresponding full CSI case as the number of feedback bits becomes large. Finally, the effectiveness of the derived algorithms is examined through numerical simulations.**

***Index Terms*—Free-space optical communication, wireless optical communications, Poisson fading channels, limited feedback, optimization.**

I. Introduction

FREE-SPACE OPTICS (FSO) has recently gained significant interest among a number of applications, e.g., metro network extensions, last-mile access, fibre-backup, back-haul for wireless networks [1]. Its remarkable advantages, such as flexibility, cost-effectiveness, high security, fast installation time, plentiful license-free spectrum and immunity from interference caused by external sources [2][6], have enabled it to become an increasingly attractive means for high data rate transmission. Despite so many benefits, a major drawback of the FSO channel is atmospheric turbulence-induced fading, which can result in significant degradation of the communication performance. This has hampered further widespread utilization of FSO. Hence, effective and powerful techniques to mitigate the undesirable effect of turbulence-induced fading are needed in FSO systems design.

Manuscript received May 8, 2012; revised August 17, 2012. The editor coordinating the review of this paper and approving it for publication was S. Hranilovic.

The authors are with the Department of Electrical and Electronic Engineering, University of Melbourne, Victoria 3010, Australia (e-mail: yuhe@unimelb.edu.au; sdey@ee.unimelb.edu.au).

Digital Object Identifier 10.1109/TCOMM.2013.082813.120307

One popular approach to combat the unfavorable impact of turbulence-induced fading so as to achieve higher system performance is to dynamically adapt transmit power or rate according to the instantaneous time-varying channel conditions, often referred to as channel state information (CSI), available at the transmitter and the receiver. In FSO communication, changes in the fading associated with atmospheric turbulence (in the order of milliseconds) occur on a sufficiently slow time scale with respect to the data rate (in the order of Gigabits per second)[4]. Therefore the block fading channel model, where the CSI changes independently from block to block but remains unchanged within each transmission block, can be employed to describe how the FSO channel varies with fading [5][4][2][3][6]. One obvious issue which can significantly affect the performance of such dynamic transmit power or rate adaptation algorithms is related to the degree of accuracy or resolution of the CSI available at the transmitter.

Many of the existing work on the resource optimization of FSO communication systems assume that both the transmitter and the receiver have perfect knowledge of full CSI, so that the transmitter can optimally and adaptively assign its transmit power or rate based on perfect CSI, achieving the optimal system performance. For example, for the case of shot-noise limited ideal photodetection FSO block fading channel with unlimited bandwidth, also referred to as the Poisson fading channel (where information is transmitted by modulating the intensity of an optical beam, and individual photon arrivals at the photodetector receiver follows a Poisson counting process), the authors of [3] studied the optimal power and rate control law for maximizing the ergodic capacity of a SISO Poisson fading channel subject to peak and average transmit power constraints and a service outage constraint. In [4], upper and lower bounds on ergodic and outage capacity, and approximate expressions for the capacity-versus-outage probability of MIMO Poisson fading channels with peak and average power constraints were investigated. In [5] and [6], the authors derived the exact expressions of ergodic and outage capacity of MIMO Poisson fading channel, respectively, and their corresponding optimal power allocation schemes. Other FSO fading channel models have been also considered, such as [7], where MIMO FSO fading channels with non-ideal photodetection (background noise modelled as additive white Gaussian noise (AWGN)) were investigated with pulse-position modulation (PPM) of the transmitted signals, and equal gain combining (EGC) at the receiver. In this paper, the authors analyzed the outage probability with full CSI and no CSI available at the transmitter.

However, the assumption on having perfect CSI at the transmitter is not achievable in a practical system due to inaccuracies of channel estimation at the receiver and the unreliability and data rate constraints of the feedback channel. To the best of our knowledge, imperfect CSI at the transmitter of the Poisson fading channel was first considered in [2]. The authors of [2] proposed a general class of ergodic capacity maximization problems of the SISO Poisson fading channel for different levels of CSI available at transmitter and with perfect CSI at the receiver. The characterization of optimal power control policies for two extreme cases of the general problem, i.e., perfect CSI and no CSI case, were fully derived. However, for partial CSI case, only a general form of power control law is provided. This motivated us to study the exact characterization of power allocation laws for the problem considered in [2] under a specific form of partial CSI available at the transmitter.

Our main contributions in this paper are as follows:

- We study the optimal power allocation schemes that maximize the ergodic capacity, as defined in [2], of a shot-noise limited SISO Poisson fading channel with quantized CSI at the transmitter, acquired via a no-delay and error-free feedback link with finite-rate from the receiver to the transmitter, and perfect CSI at the receiver. Note that the main difference between our work and [2] is in the consideration of quantized feedback of CSI to the transmitter.
- Although the considered optimization problem is non-convex, we develop a locally optimal QPA scheme for the original problem by solving its dual Lagrangian optimization problem with a 2-Step iterative design algorithm called 'QPA-DP'.
- Two simple and effective algorithms to implement the Step 1 of 'QPA-DP' are developed. The optimal algorithm 1 for Step 1 employs a simple iterative Lloyd-like algorithm similar to [9], [10]. However, unlike [9], [10], we are able to prove that at each iteration, a unique solution is found and thus the algorithm is guaranteed to converge to the optimal solution of Step 1. The optimal algorithm 2 for Step 1 is obtained by using some nice properties of the quantized power control law that we derive. Combining these algorithms with the algorithm for Step 2 of 'QPA-DP', we design two algorithms both of which converge to the globally optimal solution of the dual problem.
- Furthermore, in the case of a large number of feedback bits, we derive a fast and low-complexity suboptimal QPA algorithm.
- The results are then extended to the high peak signal-to-noise ratio (SNR) regime and an explicit expression for the approximate asymptotic ergodic capacity behavior in the high SNR regime for high rate quantization (as the number of feedback bits goes to infinity) is provided.
- Finally, the effectiveness of the derived algorithms are evaluated through numerical simulations.

This paper is organized as follows. Section II introduces the system model and the problem formulation. Section III presents a locally optimal QPA scheme for the proposed optimization problem obtained by solving its dual problem. Two effective optimal QPA algorithms are developed to find the global optimal solution of the dual problem. In Section IV, a low-complexity suboptimal QPA strategy is derived. Section V extends the results to the high SNR regime and the asymptotic capacity behavior of high SNR regime in high resolution quantization is also investigated. Simulation results are given in Section VI, followed by concluding remarks in Section VII.

II. System Model and Problem Formulation

We consider a shot-noise limited single-user single-input single-output (SISO) Poisson fading channel with unbounded bandwidth as described in [2][3]. It is assumed to be a block fading channel on account of the slowly varying nature of optical fade. Let $[0, T]$ be the time interval of signal transmission and reception. The channel input is a nonnegative real-valued signal $\{x(t), t \in [0, T]\}$, which is proportional to the transmitted optical power and is subjected to the peak and average power constraints below:

$$0 \le x(t) \le A, \quad t \in [0, T]; \qquad \frac{1}{T}\int_0^T x(t)dt \le \rho A, \quad (1)$$

where $A > 0$ stands for the peak power and $0 \le \rho \le 1$ denotes the average-to-peak power ratio. The corresponding channel output is a $\mathcal{Z}^+$-valued nondecreasing Poisson counting process $\{y(t), t \in [0, T]\}$, having an instantaneous rate (denoted as $R(t)$) equal to

$$R(t) = S\left[\left\lceil \frac{t}{T_c} \right\rceil\right] x(t) + \lambda_0, \quad t \in [0, T], \quad (2)$$

where $\mathcal{Z}^+$ denotes the space of positive integers, $\{S[k], k \in \mathcal{Z}^+\}$ is the independent and identically distributed (i.i.d.) nonnegative real-valued random channel fade with a known distribution in the k-th coherence interval$((k-1)T_c, kT_c]$ (T_c represents the channel coherence time and $\lceil x \rceil$ indicates the smallest integer greater than or equal to x), and $\lambda_0 \ge 0$ is the (constant) background noise (dark current) rate. The channel fade is assumed to follow a lognormal distribution as proposed in [4][2], i.e., $S \sim e^{\chi}$, where χ is normally distributed with mean $\overline{\mu}$ and variance σ^2. Its probability density function is given as $f_S(s) = \frac{1}{s\sigma\sqrt{2\pi}} e^{-\frac{(\log(s)-\overline{\mu})^2}{2\sigma^2}}$. We assume that the receiver has perfect or full knowledge of CSI, while various levels of CSI feedback can be made available to the transmitter. In general, the transmitter CSI can be specified in terms of the random variable $V = \phi(S)$, where the function $\phi(\cdot)$ denotes the mapping from the receiver CSI to the transmitter CSI. The maximum capacity $\mathcal{C}$ (in nats per second) of the SISO Poisson fading channel under the peak and average power constraints (1), achieved by i.i.d. $\{0, A\}$-valued channel inputs (corresponding to an ON-OFF keying (OOK) signaling scheme), is given by [2],

$$\mathcal{C} = \max_{\substack{p:V \to [0,1] \\ E[p(V)] \le \rho}} \lambda_0 E[p(V)(1 + Sr)\log(1 + Sr) - (1 + p(V)Sr)\log(1 + p(V)Sr)] \quad (3)$$

where $r = \frac{A}{\lambda_0}$ is the peak SNR. $p(V)$ is the probability of the channel input taking value A, conditioned on the available transmitter CSI, and can be interpreted as the duty cycle of the transmitted signal [2][3][4]. Since the transmitted power in OOK signaling scheme is directly proportional to the duty cycle, optimizing $p(.)$ in (3) can also be viewed as designing an optimal power allocation policy [3].

The two extreme cases of this general Problem (3): perfect CSI and no CSI available at the transmitter, where ϕ is an identity mapping (i.e., $V = S$) and a constant mapping (i.e., $V =$ a constant), respectively, have been considered in [2]. For the sake of completeness, the optimal power control scheme $p^*(.)$ for these two special cases are given below respectively as in [2],

(1) Full CSI at the transmitter case:

$$p^*(s) = [p_{\mu_f}(s)]^+, \tag{4}$$

where $[x]^+ = \max(x, 0)$,

$$p_{\mu_f}(s) = \frac{1}{sr}\left(e^{-\left(1+\frac{\mu_f}{As}\right)}(1+sr)^{\left(1+\frac{1}{sr}\right)} - 1\right) \tag{5}$$

and μ_f is the nonnegative Lagrange multiplier associated with the average power constraint and is determined by solving $\mu_f(E[[p_{\mu_f}(S)]^+] - \rho) = 0$. When $\mu_f > 0$ (i.e., average power constraint is active), $p^*(s) \in [0, \frac{1}{2}]$, while when $\mu_f = 0$ (i.e., average power constraint is inactive), $p^*(s) \in [e^{-1}, \frac{1}{2}]$.

(2) No CSI at the transmitter case:

$$p^* = \min\{\rho,\ p_0\}, \tag{6}$$

where p_0 is the solution of below equation:

$$E[Sr\log(1+p_0Sr)] = E[(1+Sr)\log(1+Sr) - Sr], \tag{7}$$

and $p^* \in [0, \frac{1}{2}]$, $p_0 \in [e^{-1}, \frac{1}{2}]$.

When only partial CSI available at the transmitter, in the next, we will design an optimum power allocation strategy for the capacity maximization Problem (3) based on quantized CSI, acquired via a no-delay and error-free feedback link with finite-rate from the receiver to the transmitter (also known as limited feedback technique).

Remark 1: In the low SNR regime for $\lambda_0 \gg 1$, it was shown in [2] that regardless of whether the transmitter is provided with CSI or not, the maximum capacity of the Poisson fading channel is $\mathcal{C}^{\mathrm{L}} = \frac{p^{\mathrm{L}}(1-p^{\mathrm{L}})E[S^2]A^2}{2\lambda_0} + O(\lambda_0^{-2})$, where $p^{\mathrm{L}} = \min\{\rho, \frac{1}{2}\}$, which means that in this scenario, transmitter CSI is not needed. Therefore, quantized CSI feedback is beneficial only when λ_0 is not too high (i.e., excluding the low SNR regime).

III. Optimum Quantized Power Allocation (QPA) with Limited Feedback

In this section, we consider the case where the receiver can obtain perfect information of CSI S and then forward some appropriately quantized information about S to the transmitter through a finite-rate feedback link. More specifically, given $\mathcal{B}$ bits of feedback, a power (duty cycle) codebook $\mathcal{P}= \{p_1, \ldots, p_{\mathcal{L}}\}$ of cardinality $\mathcal{L} = 2^{\mathcal{B}}$, known *a priori* by both the transmitter and the receiver, is designed off line purely on the basis of the statistics of S. Given a channel realization $S = s$, the receiver employs an index mapping ϕ from the current instantaneous s information to one of $\mathcal{L}$ integer indices (which partitions the scalar space of s into $\mathcal{L}$ disjoint regions $\mathcal{R}_1, \ldots, \mathcal{R}_{\mathcal{L}}$ and is defined as $\phi(s) = j$, if $s \in \mathcal{R}_j$, $j = 1, \ldots, \mathcal{L}$). The corresponding index $j = \phi(s)$ is then sent to the transmitter via the feedback link. The transmitter then uses the associated power codebook element (e.g., if the feedback signal is j, then p_j will be used as the duty cycle of the transmitted signal) to adapt its transmission strategy.

Let $\Pr(\mathcal{R}_j)$, $E[\bullet|\mathcal{R}_j]$ represent $\Pr(S \in \mathcal{R}_j)$ (the probability that S falls in the region $\mathcal{R}_j$) and $E[\bullet|S \in \mathcal{R}_j]$, respectively. Then Problem (3) with limited feedback can be expressed as

$$\begin{aligned}\mathcal{C}_{\mathcal{L}} = \max_{p_j \in [0,1], \mathcal{R}_j, \forall j} &\sum_{j=1}^{\mathcal{L}} \lambda_0 E[p_j(1+Sr)\log(1+Sr) \\ &\quad - (1+p_jSr)\log(1+p_jSr)|\mathcal{R}_j]\Pr(\mathcal{R}_j) \\ s.t. \quad &\sum_{j=1}^{\mathcal{L}} p_j \Pr(\mathcal{R}_j) \le \rho\end{aligned} \tag{8}$$

Our aim is thus to jointly optimize the channel partition regions and the power codebook, such that the capacity is maximized under the above constraints. It is easy to verify that (8) is a non-convex optimization problem. Thus it is very hard to find the globally optimal solution directly. However, in what follows, we are able to find a local optimum for Problem (8) by solving its Lagrange dual problem in an efficient manner.

A. A 2-Step Iterative Design Algorithm for Locally Optimum QPA

The Lagrangian of Problem (8) can be written as,

$$\begin{aligned}L(\{p_j\}, \{\mathcal{R}_j\}, \mu) = &\sum_{j=1}^{\mathcal{L}} \lambda_0 E\left[p_j\beta(Sr) - \beta(p_jSr)|\,\mathcal{R}_j\right]\Pr(\mathcal{R}_j) \\ &- \mu\left(\sum_{j=1}^{\mathcal{L}} p_j \Pr(\mathcal{R}_j) - \rho\right),\end{aligned} \tag{9}$$

where $\beta(x) = (1+x)\log(1+x)$ and μ is the nonnegative Lagrange multiplier associated with average power constraint. The dual problem of (8) is then defined as

$$\min_{\mu \ge 0} g(\mu) \tag{10}$$

where $g(\mu)$ is the Lagrange dual function, defined as,

$$\begin{aligned}g(\mu) &= \max_{p_j \in [0,1], \mathcal{R}_j, \forall j} L(\{p_j\}, \{\mathcal{R}_j\}, \mu) \\ &= \left(\max_{p_j \in [0,1], \mathcal{R}_j, \forall j} \mathcal{G}(\{p_j\}, \{\mathcal{R}_j\}, \mu)\right) + \mu\rho\end{aligned} \tag{11}$$

where

$$\mathcal{G}(\{p_j\}, \{\mathcal{R}_j\}, \mu) \triangleq \sum_{j=1}^{\mathcal{L}} \lambda_0 E\left[p_j\beta(Sr) - \beta(p_jSr) - \frac{\mu}{\lambda_0}p_j \middle| \mathcal{R}_j\right]\Pr(\mathcal{R}_j). \tag{12}$$

In order to solve the above dual problem we begin with an arbitrary initial value for μ. Then, Step 1 and Step 2 below are iteratively applied until a pre-specified convergence criterion is met:

Step 1: With a fixed value of μ, find a locally optimal solution (a locally optimal power codebook and the corresponding quantization regions) of the Lagrange dual function problem (11).

Step 2: With the resulting power codebook and channel partition, update the optimal value μ value by solving the dual problem (10)[1], i.e., $\mu\left(\sum_{j=1}^{\mathcal{L}} p_j \Pr(\mathcal{R}_j) - \rho\right) = 0$.

We call this 2-Step iterative design algorithm for solving the Lagrange dual problem (10) (i.e., a locally optimum QPA for Problem (8)), as the 'QPA-DP'algorithm. In the following, we will study how to implement **Step 1** and **Step 2**, and then combine these two steps to obtain the 'QPA-DP' algorithm.

B. An Inefficient Solution to Step 1

In **Step 1**, with a given μ, a general approach to find a solution for Problem (11) is to utilize Generalized Lloyd Algorithm (GLA) with a modified Lagrangian distortion measure $\sum_j E\left[d(S,j)|\mathcal{R}_j\right]\Pr(\mathcal{R}_j)$, where $d(s,j) = \lambda_0\Big(p_j\beta(sr) - \beta(p_j sr) - \frac{\mu}{\lambda_0}p_j\Big)$, as stated in [8]. This modified GLA is implemented using a sufficiently large number of training samples (channel realizations for S) and is designed based on two necessary optimality conditions: 1) with a given power codebook $\mathcal{P}= \{p_1,\ldots,p_{\mathcal{L}}\}$, the optimal channel CSI partition is determined by the nearest neighbor condition, namely, for $\forall j = 1,\ldots,\mathcal{L}$,

$$\mathcal{R}_j = \{s : d(s,j) \geq d(s,i), \forall i \in \{1,\ldots,\mathcal{L}\}, i \neq j\}. \quad (13)$$

Note that $\mathcal{R}_j, \forall j$ need not be unique since ties may be broken arbitrarily; 2) with a given channel partition $\mathcal{R}_1,\ldots,\mathcal{R}_{\mathcal{L}}$, the optimal power codebook is found by the generalized centroid condition, namely,

$$p_j^* = \arg\max_{p_j\in[0,1]} \; E[d(S,j)|\mathcal{R}_j]\Pr(\mathcal{R}_j), \quad \forall j = 1,\ldots,\mathcal{L}, \quad (14)$$

which is a concave optimization problem. Therefore, beginning with an arbitrary initial codebook, one can repeatedly apply the two optimality conditions described above until convergence to obtain a locally optimal power codebook and corresponding channel partitions for Problem (11).

However, the computational complexity of implementing GLA can be quite high and it can take a very long time to converge, due to the fact that GLA based quantizer design requires a large number of training samples drawn from empirical distributions. In order to avoid this computational burden, in the following, we will derive two alternative optimal QPA algorithms with lower complexity and faster convergence rates. Most importantly, we will show that they can find the unique optimal solution for the Problem (11) with a fixed μ.

C. Optimal Algorithm 1 for Step 1

Let a power codebook $\mathcal{P}= \{p_1,\ldots,p_{\mathcal{L}}\}$ and the corresponding channel partitioning $\mathcal{R}_1,\ldots,\mathcal{R}_{\mathcal{L}}$ denote an optimal solution to Problem (11). And let $\{s_1,\ldots,s_{\mathcal{L}}\}$ indicate the associated quantization thresholds on S axis, where $0 = s_1 < \cdots < s_{\mathcal{L}} < s_{\mathcal{L}+1} = \infty$. Then we have $\mathcal{R}_j = [s_j, s_{j+1})$ for $j = 1,\ldots,\mathcal{L}$. Our aim is to jointly optimize the power codebook $\{p_1,\ldots,p_{\mathcal{L}}\}$ and the channel thresholds $\{s_2,\ldots,s_{\mathcal{L}}\}$, so that the objective function in Problem (11) (or $\mathcal{G}(\{p_j\},\{\mathcal{R}_j\},\mu)$) is maximized. We employ a simple iterative algorithm, similar to [9][10]. Let k denote the iteration number. The algorithm starts from an initial guess of $\left\{s_2^{(0)},\ldots,s_{\mathcal{L}}^{(0)}\right\}$ satisfying $0 < s_2^{(0)} < \cdots < s_{\mathcal{L}}^{(0)} < \infty$, then produces a sequence of iterates $\left(\{p_j^{(k+1)}\},\{s_j^{(k+1)}\}\right) = \theta\left(\{p_j^{(k)}\},\{s_j^{(k)}\}\right)$, where the operator θ works as follows : for given channel thresholds $\left\{s_j^{(k)}\right\}$, a power codebook $\left\{p_j^{(k+1)}\right\}$ is chosen to maximize $\mathcal{G}\left(\{p_j^{(k+1)}\},\{\mathcal{R}_j^{(k)}\},\mu\right)$ and then for fixed $\left\{p_j^{(k+1)}\right\}$, $\left\{s_j^{(k+1)}\right\}$ is found to maximize $\mathcal{G}\left(\{p_j^{(k+1)}\},\{\mathcal{R}_j^{(k+1)}\},\mu\right)$, until the resulting value of objective function $\mathcal{G}(\{p_j\},\{\mathcal{R}_j\},\mu)$ converges within a pre-specified accuracy, namely, $\frac{\mathcal{G}^{(k+1)}-\mathcal{G}^{(k)}}{\mathcal{G}^{(k+1)}} < \epsilon$.

We carry out the following steps at each iteration (for simplicity, we will omit the iteration index):

1) Given a set of $\{s_2,\ldots,s_{\mathcal{L}}\}$ where $0 < s_2 < \cdots < s_{\mathcal{L}} < \infty$, the optimal $\{p_1,\ldots,p_{\mathcal{L}}\}$ is obtained by maximizing $\mathcal{G}(\{p_j\},\{\mathcal{R}_j\},\mu)$, i.e., $\forall j = 1,\ldots,\mathcal{L}$,

$$p_j = \arg\max_{p_j\in[0,1]} \lambda_0 E\left[p_j\beta(Sr) - \beta(p_j Sr) - \frac{\mu}{\lambda_0}p_j\middle|\mathcal{R}_j\right]\Pr(\mathcal{R}_j). \quad (15)$$

It is easy to verify that with a fixed channel partition, (15) is a concave optimization problem. Thus by using the KKT optimality conditions, the optimal power level is given as

$$p_j = \left[p_j^{\mu}\right]^+, \mu \geq 0, \quad (16)$$

where p_j^{μ} is the solution of the equation $\lambda_0 E\left[Sr\log\left(\frac{1+\varepsilon_\mu(Sr)Sr}{1+p_j^{\mu}Sr}\right)\middle|\mathcal{R}_j\right] = 0$, and

$$\varepsilon_\mu(sr) = \frac{1}{sr}\left(e^{-\left(1+\frac{\mu}{sr\lambda_0}\right)}(1+sr)^{\left(1+\frac{1}{sr}\right)} - 1\right). \quad (17)$$

2) Given $\{p_1,\ldots,p_{\mathcal{L}}\}$, the optimal $\{s_2,\ldots,s_{\mathcal{L}}\}$ is determined by maximizing $\mathcal{G}(\{p_j\},\{\mathcal{R}_j\},\mu)$, i.e., (18), Problem (18) is generally a non-convex optimization problem. However, we will prove in Lemma 3 below that there exists a unique maximum point for Problem (18), i.e., only one optimal solution. This optimal solution s_j is obtained by solving

$$\varphi(s_j) = 0, \quad \text{with} \quad \frac{\partial\varphi(s_j)}{\partial s_j} < 0. \quad (19)$$

where

$$\varphi(s_j) \triangleq \lambda_0\Big[p_{j-1}\beta(s_j r) - \beta(p_{j-1}s_j r) - \frac{\mu}{\lambda_0}p_{j-1} - p_j\beta(s_j r) + \beta(p_j s_j r) + \frac{\mu}{\lambda_0}p_j\Big]. \quad (20)$$

[1]When the average power constraint of Problem (8) is inactive, i.e., $\sum_{j=1}^{\mathcal{L}} p_j \Pr(\mathcal{R}_j) < \rho$, then we must have $\mu = 0$. Alternatively, when the average power constraint is active, i.e., $\sum_{j=1}^{\mathcal{L}} p_j \Pr(\mathcal{R}_j) = \rho$, we have $\mu \geq 0$ and can solve $\sum_{j=1}^{\mathcal{L}} p_j \Pr(\mathcal{R}_j) = \rho$ to obtain μ.

$$
\begin{aligned}
s_j = \arg\max_{s_j} \ \ & \lambda_0 E\left[p_{j-1}\beta(Sr) - \beta(p_{j-1}Sr) - \frac{\mu}{\lambda_0}p_{j-1} \middle| \mathcal{R}_{j-1} \right] \Pr(\mathcal{R}_{j-1}) \\
& + \lambda_0 E\left[p_j\beta(Sr) - \beta(p_jSr) - \frac{\mu}{\lambda_0}p_j \middle| \mathcal{R}_j \right] \Pr(\mathcal{R}_j), \qquad \forall j = 2, \ldots, \mathcal{L}
\end{aligned} \tag{18}
$$

In addition, we also show in Lemma 3 that the optimal $\{s_2, \ldots, s_{\mathcal{L}}\}$ obtained from (18) automatically satisfies $0 < s_2 < \cdots < s_{\mathcal{L}} < \infty$.

We call this simple iterative algorithm as "Iterative Power and Thresholds Optimization (IPTO)".

We will now present an important property of the function $\varepsilon_\mu(sr)$.

Lemma 1: With a finite r, when $\mu = 0$, the function $\varepsilon_\mu(sr)$ is monotonically decreasing in s, and $e^{-1} \le \varepsilon_\mu(sr) \le \frac{1}{2}$. While when $\mu > 0$, $\varepsilon_\mu(sr)$ has a unique maximum point, denoted as s^{M}, and is increasing when $s \in [0, s^{\mathrm{M}}]$ and then decreasing over $s \in (s^{\mathrm{M}}, \infty)$. In this case, $-\infty \le \varepsilon_\mu(sr) < \frac{1}{2}$.

Proof: See Appendix A. ■

Lemma 2: Based on Lemma 1, the optimal quantized power levels obtained by (16) of the IPTO algorithm satisfy: when $\mu = 0$, $\frac{1}{2} \ge p_1 > \cdots > p_{\mathcal{L}} \ge e^{-1}$; while, when $\mu > 0$, $0 \le p_j \le \frac{1}{2}$, $\forall j = 1, \ldots, \mathcal{L}$ and if the power levels of any two adjacent regions $\mathcal{R}_j$ and $\mathcal{R}_{j+1}$ satisfy $p_j > p_{j+1}$, we also have $\frac{1}{2} \ge p_j > p_{j+1} \ge e^{-1}$.

Proof: See Appendix B. ■

With Lemma 2, we can obtain the following Lemma:

Lemma 3: There exists a unique optimal solution $\{s_2, \ldots, s_{\mathcal{L}}\}$ for Problem (18), which can be obtained by solving (19). And the resulting $\{s_j\}_{j=2}^{\mathcal{L}}$ also satisfies $0 < s_2 < \cdots < s_{\mathcal{L}} < \infty$.

Proof: See Appendix C. ■

The IPTO algorithm can be formally summarized as below:

IPTO Algorithm:

Initialize $k = 0$, and choose arbitrary initial values of $\left\{s_2^{(0)}, \ldots, s_{\mathcal{L}}^{(0)}\right\}$ such that $0 < s_2^{(0)} < \cdots < s_{\mathcal{L}}^{(0)} < \infty$;

repeat

(1) Given $\left\{s_2^{(k)}, \ldots, s_{\mathcal{L}}^{(k)}\right\}$, optimal $\left\{p_1^{(k+1)}, \ldots, p_{\mathcal{L}}^{(k+1)}\right\}$ is given by (16), $\forall j = 1, \ldots, \mathcal{L}$;

(2) Given $\left\{p_1^{(k+1)}, \ldots, p_{\mathcal{L}}^{(k+1)}\right\}$, optimal $\left\{s_2^{(k+1)}, \ldots, s_{\mathcal{L}}^{(k+1)}\right\}$ is determined by solving (19), $\forall j = 2, \ldots, \mathcal{L}$;

(3) $k = k + 1$;

until Convergence;

Remark 2: In the IPTO algorithm, the value of the objective function of Problem (11) obtained at each step, i.e., $\left\{\mathcal{G}^{(k)} + \mu\rho\right\}$, is a monotonically increasing sequence and is upper bounded by the Lagrangian for the full CSI case. Meanwhile, at each iteration k, a unique value of $\left\{p_j^{(k+1)}\right\}, \left\{s_j^{(k+1)}\right\}$ is found. Thus the IPTO algorithm is guaranteed to converge to the optimal solution of Problem (11).

D. Optimal Algorithm 2 for Step 1

In order to obtain the optimal quantized power codebook $\{p_1, \ldots, p_{\mathcal{L}}\}$ and channel partition thresholds $\{s_2, \ldots, s_{\mathcal{L}}\}$ from the IPTO algorithm, one could equivalently solve the following system of simultaneous nonlinear equations:

$$
\begin{cases}
\varphi(s_j) = 0, \ \ \text{with } \dfrac{\partial \varphi(s_j)}{\partial s_j} < 0, \ \ j = 2, \ldots, \mathcal{L}; & \text{(21a)} \\
p_j = [p_j^\mu]^+, \qquad j = 1, \ldots, \mathcal{L}, & \\
\ \ \text{with } \ \lambda_0 E\left[Sr \log\left(\dfrac{1 + \varepsilon_\mu(Sr)Sr}{1 + p_j^\mu Sr} \right) \middle| \mathcal{R}_j \right] = 0 & \text{(21b)}
\end{cases}
$$

where $\varphi(s_j)$ and $\varepsilon_\mu(sr)$ are given in (20) and (17), respectively. From Remark 2, the IPTO algorithm converges to the unique optimal solution of Problem (11). Thus there must be only one $\{p_1, \ldots, p_{\mathcal{L}}\}$, $\{s_2, \ldots, s_{\mathcal{L}}\}$ satisfying the above nonlinear equations system (21a)(21b). This motivates us to develop an alternative optimal QPA algorithm to solve (21a)(21b) directly with the aid of the following Lemma.

Lemma 4: The optimal quantization thresholds must satisfy: $\forall j = 2, \ldots, \mathcal{L}$,

$$
s_j r \left(-1 + \left(1 + \frac{1}{s_j r} \right) \log(1 + s_j r) \right) - \frac{\mu}{\lambda_0} \ge 0. \tag{22}
$$

which also can be written as $s_j r \log(1 + \varepsilon_\mu(s_j r) s_j r) \ge 0$

Proof: The proof follows directly by applying the Mean Value Theorem and is omitted due to space restrictions. ■

Remark 3: It is easy to verify that the left hand side (LHS) of (22) in Lemma 4 is a monotonically increasing function, thus (22) can be rewritten as $s_j \ge s_{\mathrm{lb}}$, $\forall j = 2, \ldots, \mathcal{L}$, where s_{lb} satisfies $s_{\mathrm{lb}} r \left(-1 + \left(1 + \frac{1}{s_{\mathrm{lb}} r} \right) \log(1 + s_{\mathrm{lb}} r) \right) - \frac{\mu}{\lambda_0} = 0$ (or $s_{\mathrm{lb}} r \log(1 + \varepsilon_\mu(s_{\mathrm{lb}} r) s_{\mathrm{lb}} r) = 0$). When $\mu = 0$, $s_{\mathrm{lb}} = 0$ and in this case, $s_j \ge s_{\mathrm{lb}}$ is always met. Whereas, when $\mu > 0$, $s_{\mathrm{lb}} > 0$ and $s_{\mathrm{lb}} r \log(1 + \varepsilon_\mu(s_{\mathrm{lb}} r) s_{\mathrm{lb}} r) = 0$, clearly we have $\varepsilon_\mu(s_{\mathrm{lb}} r) = 0$. As a result, with $0 < s_2 < \cdots < s_{\mathcal{L}}$, as long as $s_2 \ge s_{\mathrm{lb}}$, all other partition thresholds will automatically satisfy $s_j \ge s_{\mathrm{lb}}, \forall j = 3, \ldots, \mathcal{L}$.

With Lemma 4, we can obtain the following Lemma:

Lemma 5: When $\mu = 0$, all the power levels are strictly positive, i.e., $p_j > 0, \forall j = 1, \ldots, \mathcal{L}$. When $\mu > 0$, the power level in the first region could be zero or positive, i.e., $p_1 \ge 0$, but the power levels in the last $\mathcal{L} - 1$ regions are always strictly positive, i.e., $p_j > 0, \forall j = 2, \ldots, \mathcal{L}$. Let $p_m, 1 \le m \le \mathcal{L}$, represent the largest power level when $\mu > 0$. Then the optimal power levels must satisfy, $0 \le p_1 < \cdots < p_{m-1} < p_m$ and $p_m > p_{m+1} > \cdots > p_{\mathcal{L}} \ge e^{-1}$.

Proof: See Appendix D. ■

Remark 4: From Lemma 5, we can see that Lemma 4 ensures at most one zero power level. Besides, when $\mu = 0$, there is no outage at all; while when $\mu > 0$, the first region could be in the outage.

From Lemma 5, the nonlinear equations system (21a)(21b) can be rewritten as,

$$\begin{cases} \varphi(s_j) = 0, \text{ with } \dfrac{\partial \varphi(s_j)}{\partial s_j} < 0, \quad j = 2, \ldots, \mathcal{L}; & \text{(23a)} \\ \lambda_0 E\left[Sr \log\left(\dfrac{1+\varepsilon_\mu(Sr)Sr}{1+p_j Sr} \right) \middle| \mathcal{R}_j \right] = 0, & \\ \qquad j = 1, \ldots, \mathcal{L}, \; p_1 = [p_1]^+. & \text{(23b)} \end{cases}$$

Instead of jointly solving the above $2\mathcal{L}-1$ nonlinear equations, we can show that all these equations actually involve only one variable, which can be solved in a straightforward manner. Given a specific value of s_2, we can solve (23b) with $j = 1$ for p_1 and set $p1 = [p1]^+$, and then solve (23a) with $j = 2$ for p_2. After that, by recursively solving (23b) with $j = n$ for s_{n+1} and (23a) with $j = n+1$ for p_{n+1}, where $n = 2, \ldots, \mathcal{L}-1$, we can successively obtain $s_3, p_3, \ldots, s_\mathcal{L}, p_\mathcal{L}$. As a result, equation (23b) with $j = \mathcal{L}$ thus has only one unknown variable s_2 and can be numerically solved for s_2. We call this optimal QPA algorithm as 'Power and Thresholds via Nonlinear Equations' (PTNE) method.

E. Optimum QPA: QPA-DP

So far we have presented two alternative optimal QPA algorithms called IPTO and PTNE for implementing **Step 1**. Now, we need to determine the optimal μ for **Step 2** by solving

$$\mu \left(\sum_{j=1}^{\mathcal{L}} p_j \Pr(\mathcal{R}_j) - \rho \right) = 0. \tag{24}$$

The algorithm for **Step 2** is straightforward as it is a convex problem. Since the average power $\sum_{j=1}^{\mathcal{L}} p_j \Pr(\mathcal{R}_j)$ can be shown to be a monotonically decreasing and continuous function of μ following a similar proof to that of Lemma 1 in [13], we can find the unique optimum for μ by solving (24) using a simple bisection method.

Combining the algorithms for **Step 1** and **Step 2**, the QPA-DP algorithm for solving the dual problem (10), can then be summarized by the following steps:

QPA-DP algorithm:

Given $\mu = 0$, obtain the optimal $\{p_j\}$, $\{s_j\}$ by IPTO or PTNE algorithm and calculate $\rho_0 = \sum_{j=1}^{\mathcal{L}} p_j \Pr(\mathcal{R}_j)$;
If $\rho > \rho_0$
  $\mu = 0$;
Else
  Initialize $\mu_{\min}$ and $\mu_{\max}$;
  repeat
  (1) $\mu = (\mu_{\min} + \mu_{\max})/2$;
  (2) Given μ, the optimal $\{p_j\}$, $\{s_j\}$ are obtained by IPTO or PTNE algorithm;
  (3) if $\left(\sum_{j=1}^{\mathcal{L}} p_j \Pr(\mathcal{R}_j) - \rho\right) > 0$, then $\mu_{\min} = \mu$,
  else if $\left(\sum_{j=1}^{\mathcal{L}} p_j \Pr(\mathcal{R}_j) - \rho\right) < 0$, then $\mu_{\max} = \mu$;
  until Convergence : $|\mu_{\min} - \mu_{\max}| < \epsilon$

Remark 5: Obviously, the QPA-DP algorithm converges to the global optimum solution of the dual problem (10). However, due to the non-convexity of original problem (8), the solution obtained by the QPA-DP algorithm may not be the global optimum of the original problem (8). But the optimal μ, $\{p_j\}$, $\{s_j\}$ obtained by the QPA-DP algorithm satisfy the KKT conditions of the original problem (8), thus it must be a locally optimum solution of Problem (8).

IV. Suboptimal QPA-DP Algorithm

In this section, we will develop a suboptimal, but computationally efficient algorithm for implementing **Step 1** of the QPA-DP algorithm (i.e., we will replace IPTO or PTNE by a more efficient algorithm). We aim this algorithm to be close to optimum in the high rate quantization regime, i.e., for a large number of feedback bits $\mathcal{B}$.

According to [11], for sufficiently large $\mathcal{B}$ (or $\mathcal{L}$), the design of optimal channel partitions is of less significance. Thus, in high rate quantization, the criterion of equal probability per region (EPrPR) can be utilized to design a simple suboptimal (asymptotically optimum) channel partition policy, as stated in [11][12]. Let $F_S(s)$ denote the cumulative distribution function of the lognormally distributed channel fade, given as, $F_S(s) = \frac{1}{2} + \frac{1}{2}\operatorname{erf}\left[\frac{\log(s)-\overline{\mu}}{\sqrt{2\sigma^2}}\right]$. Applying EPrPR directly to our problem, we can obtain $\forall j = 2, \ldots, \mathcal{L}, F_S(s_j) = \frac{j-1}{\mathcal{L}}$, which implies $\forall j = 2, \ldots, \mathcal{L}$

$$s_j = \exp\left\{ \overline{\mu} - \sqrt{2}\sigma \operatorname{erf}^{-1}\left(1 - 2\frac{j-1}{\mathcal{L}}\right) \right\}. \tag{25}$$

Then with given $\{s_j\}$, we can easily and immediately solve (23b) for $p_1, \ldots, p_\mathcal{L}$. However, the problem is that, when $\mu > 0$, there may exist some quantization thresholds obtained by (25) that do not satisfy $s_j \geq s_{\text{lb}}$, as required by Lemma 4. This can lead to more than one region having zero power levels when solving (23b) for $\{p_j\}$, thus resulting in a reduced throughput. Next we will propose a modified EPrPR scheme to obtain a set of suboptimal quantization thresholds for large $\mathcal{L}$.

Lemma 6: When $\mu > 0$, p_1 must be zero for a sufficiently large $\mathcal{L}$ and as $\mathcal{L} \to \infty$, $s_2 \to s_{\text{lb}}$.

Proof: See Appendix E. ■

Remark 6: Lemma 6 shows that when $\mu > 0$, with high rate quantization, the power level for the first region $\mathcal{R}_1$ approaches zero, i.e., $\mathcal{R}_1$ becomes an outage region. In addition, as $\mathcal{L} \to \infty$, the boundary between the outage and the non-outage regions approaches to s_{lb}, which is consistent with the perfect CSI case.

Remark 3 implies that in order to meet $s_j \geq s_{\text{lb}}, \forall j = 2, \ldots, \mathcal{L}$ we just require $s_2 \geq s_{\text{lb}}$. Then by Lemma 6, we have for large $\mathcal{L}$, $s_2 \approx s_{\text{lb}}$. Therefore, if s_2 obtained by (25) meets $s_2 \geq s_{\text{lb}}$, i.e., $\exp\left\{\overline{\mu} - \sqrt{2}\sigma \operatorname{erf}^{-1}\left(1 - 2\frac{1}{\mathcal{L}}\right)\right\} \geq s_{\text{lb}}$, we can directly apply the EPrPR algorithm to obtain the other suboptimal quantization thresholds, and then solve (23b) for $p_1, \ldots, p_\mathcal{L}$. Otherwise, if $s_2 = \exp\left\{\overline{\mu} - \sqrt{2}\sigma \operatorname{erf}^{-1}\left(1 - 2\frac{1}{\mathcal{L}}\right)\right\} < s_{\text{lb}}$, we first use s_{lb} as an approximation for s_2, which leads to only the first region being in outage. We can then apply EPrPR to all other non-outage regions to obtain $F_S(s_j) = F_S(s_{\text{lb}}) + \frac{j-2}{\mathcal{L}-1}(1 - F_S(s_{\text{lb}})), \forall j =$

$3, \ldots, \mathcal{L}$. Thus, in this case, we have $\forall j = 3, \ldots, \mathcal{L}$,

$$s_j = \exp\left\{\overline{\mu} - \sqrt{2}\sigma \operatorname{erf}^{-1}\left(1 - 2\left(F_S(s_{\text{lb}}) + \frac{j-2}{\mathcal{L}-1}\left(1 - F_S(s_{\text{lb}})\right)\right)\right)\right\}. \tag{26}$$

After obtaining $\{s_j\}$, we can solve (23b) for $p_1, \ldots, p_{\mathcal{L}}$. We name this suboptimal algorithm for **Step 1** of QPA-DP as the modified EPrPR (MEPrPR).

V. QPA-DP Algorithm in the High SNR Regime

In the high SNR regime (in the limit as $\lambda_0 \to 0$ or $r = \frac{A}{\lambda_0} \to \infty$), the system of nonlinear equations (21a)(21b) becomes

$$\begin{cases} s_j = \dfrac{\mu}{A} \dfrac{p_{j-1} - p_j}{p_j \log(p_j) - p_{j-1}\log(p_{j-1})}, \quad \text{with} \\ p_j \log\left(\dfrac{1}{p_j}\right) > p_{j-1} \log\left(\dfrac{1}{p_{j-1}}\right), \; j = 2, \ldots, \mathcal{L}; \quad (27a) \\ p_j = \exp\left\{-1 - \dfrac{\mu}{AE[S|\mathcal{R}_j]}\right\}, \quad j = 1, \ldots, \mathcal{L}, \quad (27b) \end{cases}$$

and the capacity with limited feedback, denoted as $\mathcal{C_L}^{\text{H}}$, becomes,

$$\mathcal{C_L}^{\text{H}} = \sum_{j=1}^{\mathcal{L}} Ap_j \log\left(\frac{1}{p_j}\right) E[S|\mathcal{R}_j] \Pr(\mathcal{R}_j) \tag{28}$$

where $\{p_j\}$ and $\{s_j\}$ can be obtained from solving (27a)(27b) by using IPTO or PTNE algorithm. (1) When $\mu = 0$, i.e., the average power constraint is not active, from (27b), we have $p_1 = \cdots = p_{\mathcal{L}} = e^{-1}$, which implies $\rho \geq e^{-1}$ (due to $\sum_{j=1}^{\mathcal{L}} p_j \Pr(\mathcal{R}_j) < \rho$) and only one power level is needed to fully obtain the maximum capacity. In other words, no feedback is required. Thus in this case, the maximum capacity is always

$$\mathcal{C_L}^{\text{H}} = Ae^{-1}E[S]. \tag{29}$$

(2) When $\mu > 0$, i.e., the average power constraint is active (implying $\sum_{j=1}^{\mathcal{L}} p_j \Pr(\mathcal{R}_j) = \rho < e^{-1}$), from (27b), it is easy to obtain $0 < p_1 < \cdots < p_{\mathcal{L}} < e^{-1}$, implying there is no outage in the high SNR regime. Since the function $x \log\left(\frac{1}{x}\right)$ is increasing when $x \in \left(0, e^{-1}\right)$, the inequality $p_j \log\left(\frac{1}{p_j}\right) > p_{j-1} \log\left(\frac{1}{p_{j-1}}\right)$ in (27a) is always satisfied. Thus, applying (27b) into (28), the maximum capacity (28) with $\mu > 0$ becomes,

$$\begin{aligned} \mathcal{C_L}^{\text{H}} &= \sum_{j=1}^{\mathcal{L}} Ap_j E[S|\mathcal{R}_j] \Pr(\mathcal{R}_j) + \sum_{j=1}^{\mathcal{L}} \mu p_j \Pr(\mathcal{R}_j) \\ &= Ae^{-1} \sum_{j=1}^{\mathcal{L}} e^{-\frac{\mu}{AE[S|\mathcal{R}_j]}} E[S|\mathcal{R}_j] \Pr(\mathcal{R}_j) + \mu\rho. \end{aligned} \tag{30}$$

When $\mathcal{L}$ is large, due to there being no outage in the high SNR regime, we can directly apply EPrPR (25), to obtain a suboptimal channel partition policy, as stated in Section IV. Then, by applying (25), we have, $\forall j = 1, \ldots, \mathcal{L}$,

$$\begin{aligned} E[s|\mathcal{R}_j] \approx \frac{\mathcal{L}}{2} e^{\frac{\sigma^2}{2} + \overline{\mu}} \Bigg[&\operatorname{erf}\left(\frac{\sigma}{\sqrt{2}} + \operatorname{erf}^{-1}\left(1 - 2\frac{j-1}{\mathcal{L}}\right)\right) \\ &- \operatorname{erf}\left(\frac{\sigma}{\sqrt{2}} + \operatorname{erf}^{-1}\left(1 - 2\frac{j}{\mathcal{L}}\right)\right)\Bigg] \\ \triangleq a_j, & \end{aligned} \tag{31}$$

(where $\triangleq$ stands for definition). Thus for high rate quantization, the maximum capacity in the high SNR regime with $\mu > 0$ can be approximated as

$$\mathcal{C_L}^{\text{H}} \approx \frac{Ae^{-1}}{\mathcal{L}} \sum_{j=1}^{\mathcal{L}} e^{-\frac{\mu}{Aa_j}} a_j + \mu\rho. \tag{32}$$

VI. Numerical Results

In this section, we will evaluate the capacity performance of a shot-noise limited SISO Poisson fading channel with the proposed QPA strategies via numerical simulations. The channel fade with a lognormal distribution is considered to be normalized, such that $E[S] = e^{\frac{\sigma^2}{2} + \overline{\mu}} = 1$ by setting $\overline{\mu} = -\frac{\sigma^2}{2}$, as advocated in [2], [4]. According to [2], [4], the value of the variance of $\log(S)$, σ^2 can be chosen anywhere between 0 (negligible fading) to 2 (severe turbulence). Here we consider the moderate fading scenario with $\sigma^2 = 0.4$ and fix the peak power to be 1, i.e., $A = 1$.

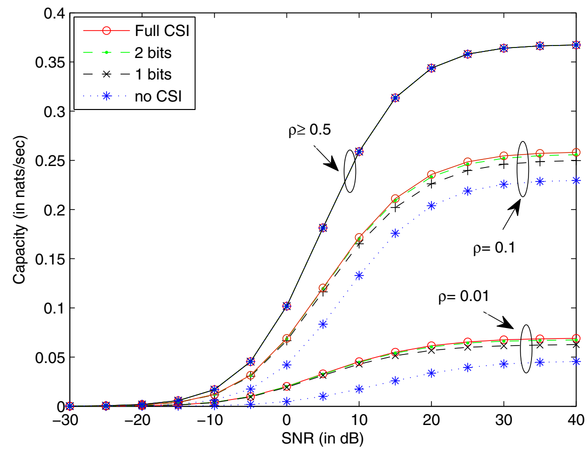


Fig. 1. Effect of increasing feedback bits on the capacity performance of the SISO Poisson fading channel.

Fig. 1 illustrates the ergodic capacity performance of the SISO Poisson fading channel obtained by QPA-DP strategy with feedback bits $\mathcal{B} = \{1, 2\}$ versus the peak SNR ($r = \frac{A}{\lambda_0}$), for three different average power constraint situations, i.e., $\rho \geq 0.5$, $\rho = 0.1$ and $\rho = 0.01$, respectively, in order to study the effect of increasing the number of feedback bits on the capacity performance. For comparison, we also plot the corresponding capacity performance of the two extreme cases, i.e., full CSI case and no CSI case. First, it can be easily observed that when $\rho \geq 0.5$, which implies the average power constraint is always inactive ($\mu = 0$) regardless of

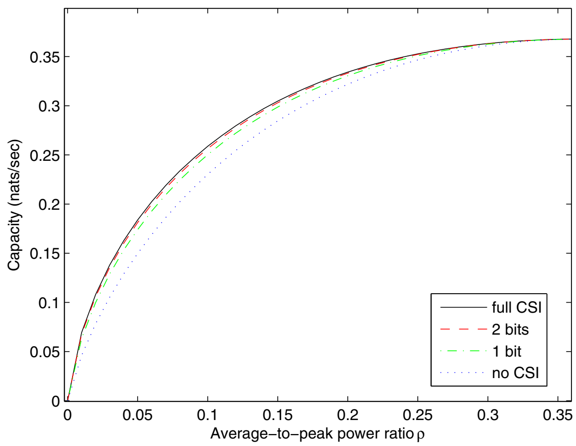


Fig. 2. The capacity performance of the Poisson fading channel in the high SNR regime with limited feedback.

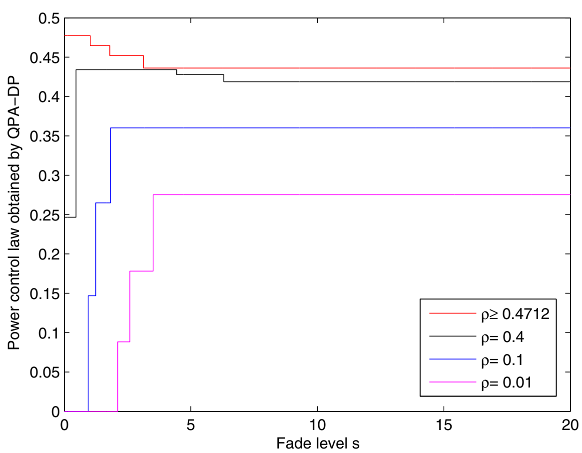


Fig. 3. Locally optimal power control law with 2 bits of CSI feedback at the transmitter (SNR $= 0$ dB).

the SNR value, the capacity performance of the full CSI case is almost identical to that of no CSI case. Thus in this case any knowledge of CSI at the transmitter does not improve capacity substantially. However, the improvement becomes more pronounced when ρ is small. When $\rho = 0.1$ or $\rho = 0.01$, introducing one extra bit of quantized CSI feedback substantially shrinks the capacity gap between the no CSI and the full CSI case. Even only one bit feedback can dramatically eliminate most of the capacity gap between the no CSI and full CSI case and as few as 2 bits of feedback are required to achieve a capacity performance very close to that of full CSI case. To be specific, for $\rho = 0.1$ and SNR$= 20$ dB, with no CSI, 1 bit feedback and 2 bits of feedback, the capacity losses due to imperfect CSI are approximately 13.38%, 3.95% and 0.93% respectively. A similar behavior can be also observed from Fig. 2 for the capacity performance obtained by QPA-DP in high SNR regime ($r \to \infty$).

Fig. 3 shows the behaviour of the locally optimal power control law obtained by the QPA-DP versus the channel fade for different values of average power ρ, with 2 bits of feedback and SNR $= 0$ dB. It can be observed from Fig. 3 that each power control law is a stepwise curve, since with 2 bits of feedback, the entire channel fade space is quantized into 4 regions and within each region, the power level is constant. More specifically, Fig. 4 diagrammatically provides an example of how the state space is partitioned for $\rho = 0.01$ (for other values of ρ, similar graphs can be easily obtained), where the channel state space is quantized into 4 regions by the quantization thresholds s_2, s_3, s_4. The first region has zero power and thus is in outage. In Fig. 3, with SNR $= 0$ dB, the QPA-DP algorithm yields $\rho_0 = 0.4712$. Then when $\rho > \rho_0$, the average power constraint of Problem (8) is inactive and $\mu = 0$. In this case, the power control law is a decreasing stepwise curve as shown in Fig. 3. Whereas, when $\rho \leq \rho_0$, the average power constraint is satisfied with equality at optimality, and in this case, if $\rho \ll \rho_0$ (such as $\rho = 0.01$ and $\rho = 0.1$ as shown in Fig. 3), the power control law is an increasing stepwise curve. If $\rho \lesssim \rho_0$ (such as $\rho = 0.4$ in Fig. 3), the power control law is a first increasing and then decreasing stepwise curve. A similar observation for the full CSI case can be found in [2]. In addition, Fig. 5 illustrates the power control law comparison between full CSI, optimal algorithm (QPA-DP) and suboptimal strategy (MEPrPR), where in order to clearly show the comparison, only results for $\rho = 0.01$ and $\rho = 0.4$ are plotted.

Fig. 6 depicts the capacity performance of the suboptimal QPA-DP algorithm (MEPrPR) versus peak SNR with feedback bits $\mathcal{B} = \{2, 3\}$, for $\rho = 0.1$ case and $\rho = 0.01$ case respectively, and compares these results with the corresponding throughput performance obtained by the optimal QPA-DP algorithm. As observed from Fig. 6(a), for any $\rho = 0.1$ or $\rho = 0.01$, even with only 2 bits feedback, the capacity performances of MEPrPR and corresponding optimal QPA-DP are almost indistinguishable. For clearer visualization, in Fig. 6(b), we zoom into the detail of the area A in Fig. 6(a), which shows that with the same number of feedback bits, the performance of optimal QPA-DP is only slightly better than that of MEPrPR, and with increasing number of feedback bits, the capacity loss due to the suboptimal quantization policy gradually vanishes. For example, at $\rho = 0.1$ and SNR$= 25$ dB, with 2 bits and 3 bits of feedback, the capacity gap between MEPrPR and optimal QPA-DP is about 0.0006 and 0.0002, respectively. Furthermore, we compared the computational speed of MEPrPR and the optimal QPA-DP algorithm. QPA-DP and MEPrPR were implemented in MATLAB (version 7.11.0.584 (R2010b)) on a Dual-Core processor (CPU AMD Athlon$^{\text{TM}}$II P360 with a clock speed of 2.30 GHz and a memory of 2 GB). It was seen that for a fixed μ with 3 bits of feedback, QPA-DP took approximately 67.14 seconds whereas MEPrPR took only 0.33 seconds to achieve comparable levels of accuracy. These results confirm that MEPrPR is a computationally highly efficient near-optimal algorithm, especially for a moderate to large number of feedback bits.

In addition, Fig. 7 shows the asymptotic capacity behavior of suboptimal QPA-DP in the high SNR regime (derived in Section V) versus the number of quantization levels $\mathcal{L}$ with $\rho = 0.1$, and compares the result with the full CSI performance of the high SNR regime given by (29) of [2]. It can be seen from Fig. 7 that the capacity increases significantly as the number of quantization level $\mathcal{L}$ increases, however, as $\mathcal{L}$

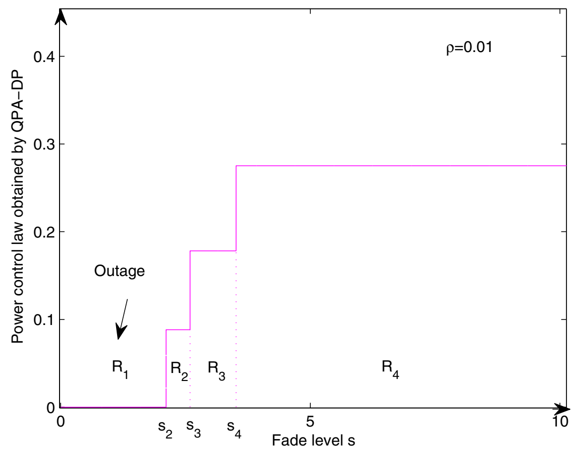


Fig. 4. Illustration of channel state partitions for 2 bits of feedback with average power $\rho = 0.01$. The channel state space is quantized into 4 regions by the quantization thresholds s_2, s_3, s_4, and the first region is in outage.

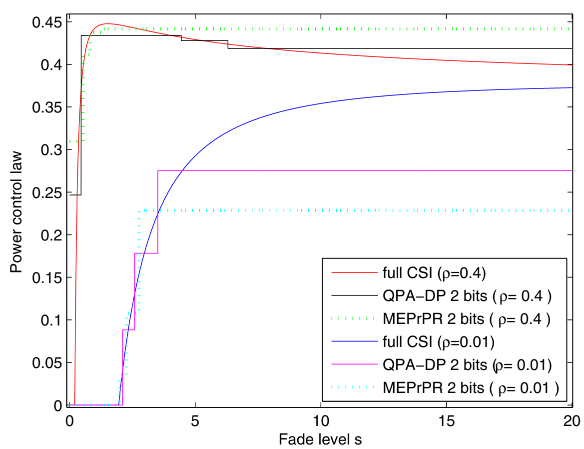
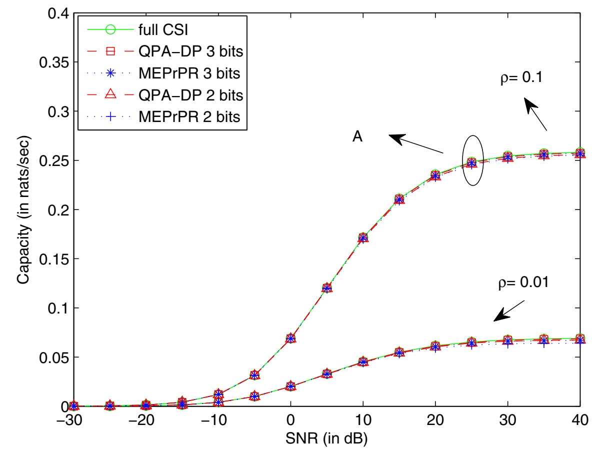
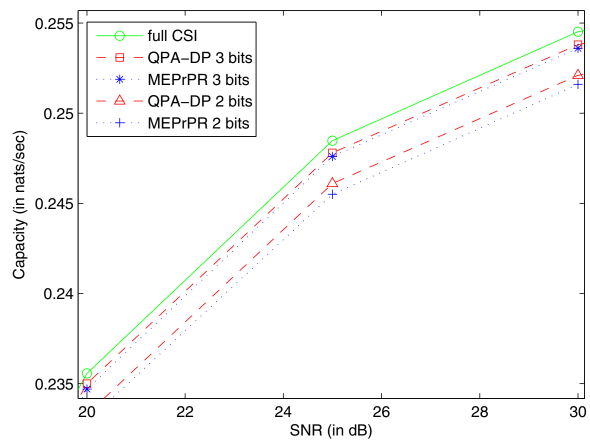


Fig. 5. Power control law comparison between full CSI, optimal algorithm (QPA-DP) and suboptimal strategy (MEPrPR).



(a)



(b) Zooming in the area of A in (a)

Fig. 6. Capacity performance comparison between suboptimal QPA-DP (MEPrPR) and optimal QPA-DP.

increases beyond a certain number ($\mathcal{L} \geq 2^3$, i.e., $\mathcal{B} \geq 3$ bits), the capacity curve starts to saturate and gradually approaches the corresponding full CSI performance. This further confirms that only a small number of feedback bits is enough to obtain a capacity close to the perfect CSI-based performance.

VII. Conclusions and Extensions

In this paper, we have considered the ergodic capacity maximization problem of a shot-noise limited SISO Poisson fading channel subject to peak and average power constraints, with only quantized CSI available at the transmitter acquired via limited feedback, and perfect CSI at the receiver. Two optimal QPA algorithms for solving the dual problem of the original optimization problem are developed and it is shown that both can converge to the globally optimal solution of the dual problem, which is also a locally optimal QPA solution for the original optimization problem. Furthermore, a fast and low-complexity suboptimal QPA algorithm is derived for high rate quantization. The results are also extended to the high SNR regime and an explicit expression for the approximate asymptotic ergodic capacity behavior in the high SNR regime for a large number of feedback bits is also provided. Although the presented optimal QPA design methods result in locally optimal solutions (due to the non-convexity of the original problem), numerical results illustrate that only 2 bits of feedback result in capacity performance almost identical to that of the full CSI case at the transmitter. Future work will involve extending the results to MIMO Poisson fading channels along with consideration of other types of communication performance measures, such as outage probability.

Appendix

A. *Proof for Lemma 1*

Note that

$$\frac{\partial \varepsilon_\mu(sr)}{\partial s} = \frac{r}{(sr)^2}(\eta_\mu(sr)+1), \tag{33}$$

where $\eta_\mu(sr) = \frac{1}{sr}e^{-(1+\frac{\mu}{sr\lambda_0})}(1+sr)^{(1+\frac{1}{sr})}(\frac{\mu}{\lambda_0} - \log(1+sr))$. Then we have

$$\frac{\partial \eta_\mu(sr)}{\partial s} = \frac{re^{-(1+\frac{\mu}{sr\lambda_0})}(1+sr)^{\frac{1}{sr}}}{(sr)^3}\left[\hat{f}(sr) + \left(\frac{\mu}{\lambda_0}\right)^2(1+sr)\right.$$

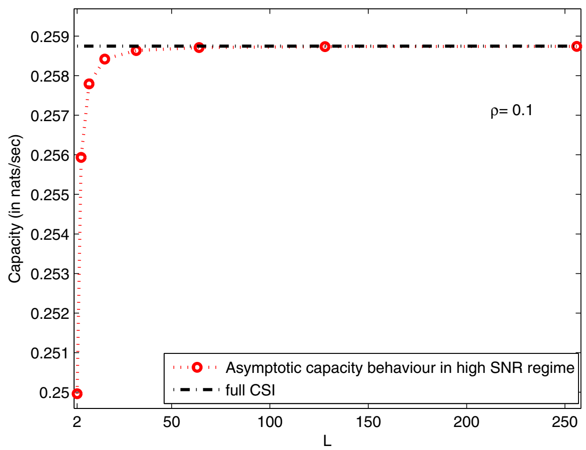


Fig. 7. Asymptotic capacity behavior in the high SNR regime versus the number of quantization levels $\mathcal{L}$.

$$-2\frac{\mu}{\lambda_0}(1+sr)\log(1+sr)\Bigg], \tag{34}$$

where $\hat{f}(sr) = -(sr)^2 + (1+sr)\log^2(1+sr)$. Since $s \geq 0$, we have $\frac{\partial^2 \hat{f}(sr)}{\partial s^2} = \frac{2r^2(-sr+\log(1+sr))}{1+sr} \leq 0$ and $\frac{\partial \hat{f}(sr)}{\partial s}|_{s=0} = 0$, which gives $\frac{\partial \hat{f}(sr)}{\partial s} \leq 0$. Due to $\hat{f}(sr)|_{s=0} = 0$, we can get $\hat{f}(sr) \leq 0$.

(1). When $\mu = 0$, (34) becomes,

$$\frac{\partial \eta_0(sr)}{\partial s} = \frac{re^{-1}(1+sr)^{\frac{1}{sr}}}{(sr)^3}\hat{f}(sr) \leq 0, \qquad s \geq 0. \tag{35}$$

Thus, $\eta_0(sr) \leq \lim_{s\to 0}\eta_0(sr) = -1$, which, combining with $\lim_{s\to 0}\frac{\partial \varepsilon_0(sr)}{\partial s} = -\frac{r}{24} < 0$, gives, $\frac{\partial \varepsilon_0(sr)}{\partial s} < 0$. Therefore $\varepsilon_0(sr)$ is a monotonic decreasing function, and $e^{-1} = \lim_{s\to\infty}\varepsilon_0(sr) \leq \varepsilon_0(sr) \leq \lim_{s\to 0}\varepsilon_0(sr) = \frac{1}{2}$.

(2). When $\mu > 0$, $\lim_{s\to 0}\frac{\partial \varepsilon_\mu(sr)}{\partial s} = \infty$ and $\lim_{s\to\infty}\frac{\partial \varepsilon_\mu(sr)}{\partial s} = 0$, thus when $s \in (0,\infty)$, $\frac{\partial \varepsilon_\mu(sr)}{\partial s} = 0$ implies $\eta(s) = -1$. From the definition of $\eta_\mu(sr)$, only when $\log(1+sr) \geq \frac{\mu}{\lambda_0}$, i.e., $s \in \left[e^{\frac{\mu}{\lambda_0 r}} - 1, \infty\right)$, $\eta_\mu(sr) \leq 0$. Thus in this case, from (34), we have,

$$\begin{aligned}\frac{\partial \eta_\mu(sr)}{\partial s} &\overset{(a)}{\leq} \frac{re^{-(1+\frac{\mu}{sr\lambda_0})}(1+sr)^{\frac{1}{sr}}}{(sr)^3}\left[\hat{f}(sr) - \left(\frac{\mu}{\lambda_0}\right)^2(1+sr)\right]\\ &\overset{(b)}{<} -\frac{re^{-(1+\frac{\mu}{sr\lambda_0})}(1+sr)^{\frac{1}{sr}}}{(sr)^3}\left(\frac{\mu}{\lambda_0}\right)^2(1+sr)\\ &< 0\end{aligned} \tag{36}$$

where (a) is obtained by applying $\log(1+sr) \geq \frac{\mu}{\lambda_0}$ and (b) is due to the fact that when $s > 0$, function $\hat{f}(sr) < 0$. Thus when $s \in \left[e^{\frac{\mu}{\lambda_0 r}} - 1, \infty\right)$, $\eta_\mu(sr)$ is monotonically decreasing over s and

$$-\infty = \lim_{s\to\infty}\eta_\mu(sr) < \eta_\mu(sr) \leq \eta_\mu(sr)\Big|_{s=e^{\frac{\mu}{\lambda_0 r}}-1} = 0, \tag{37}$$

which implies that equation $\eta_\mu(sr) = -1$ (i.e., $\frac{\partial \varepsilon_\mu(sr)}{\partial s} = 0, s \in (0,\infty)$) only has one solution, denoted as s^{M}, and $s^{\mathrm{M}} \in \left(e^{\frac{\mu}{\lambda_0 r}} - 1, \infty\right)$. Therefore, we can obtain that when $s \in \left[0, s^{\mathrm{M}}\right]$, $\frac{\partial \varepsilon_\mu(sr)}{\partial s} \geq 0$, and when $s \in \left(s^{\mathrm{M}}, \infty\right)$, $\frac{\partial \varepsilon_\mu(sr)}{\partial s} < 0$, which indicates that $\varepsilon_\mu(sr)$ is increasing when $s < s^{\mathrm{M}}$, until it achieves its unique maximum value at s^{M} point, and then decreasing after $s > s^{\mathrm{M}}$. And due to $\varepsilon_\mu(sr) < \varepsilon_0(sr)$, $\lim_{s\to\infty}\varepsilon_\mu(sr) = e^{-1}$ and with finite r, $\lim_{s\to 0}\varepsilon_\mu(sr) = -\infty$, we have $-\infty < \varepsilon_\mu(sr) < \frac{1}{2}$.

B. Proof for Lemma 2

From (16), we have that each power level $p_j = [p_j^\mu]^+, j = 1, \ldots, \mathcal{L}$, where p_j^μ satisfies,

$$\lambda_0 E\left[Sr\log\left(\frac{1+\varepsilon_\mu(Sr)Sr}{1+p_j^\mu Sr}\right)\middle|\mathcal{R}_j\right] = 0, \quad \mu \geq 0. \tag{38}$$

In order to meet (38), we must have $\min(\varepsilon_\mu(Sr)|\mathcal{R}_j) \leq p_j^\mu \leq \max(\varepsilon_\mu(Sr)|\mathcal{R}_j)$ (otherwise, (38) will become $\lambda_0 E\left[Sr\log\left(\frac{1+\varepsilon_\mu(Sr)Sr}{1+p_j^\mu Sr}\right)\middle|\mathcal{R}_j\right] \neq 0$), which leads to,

$$[\min(\varepsilon_\mu(Sr)|\mathcal{R}_j)]^+ \leq p_j \leq [\max(\varepsilon_\mu(Sr)|\mathcal{R}_j)]^+. \tag{39}$$

(1) When $\mu = 0$, according to Lemma 1, we have $\frac{1}{2} = \varepsilon_0(s_1 r) > \varepsilon_0(s_2 r) > \cdots > \varepsilon_0(s_{\mathcal{L}+1}) = e^{-1}$, due to $0 = s_1 < s_2 < \cdots < s_{\mathcal{L}} < s_{\mathcal{L}+1} = \infty$. Thus in each region $\mathcal{R}_j, \forall j = 1, \ldots, \mathcal{L}$, $\varepsilon_0(s_{j+1}r) < \varepsilon_0(sr) \leq \varepsilon_0(s_j r)$, and then from (39), we can obtain $\varepsilon_0(s_{j+1}r) < p_j \leq \varepsilon_0(s_j r)$, which gives $\frac{1}{2} \geq p_1 > \cdots > p_{\mathcal{L}} \geq e^{-1}$.

(2) When $\mu > 0$, from Lemma 1, $-\infty \leq \varepsilon_\mu(sr) \leq \frac{1}{2}$. Then from (39), we have $0 \leq p_j \leq \frac{1}{2}, \forall j = 1, \ldots, \mathcal{L}$. Now we will show that if the power levels of two adjacent regions $\mathcal{R}_j$ and $\mathcal{R}_{j+1}$ $(j = 1, \ldots, \mathcal{L}-1)$ satisfy $p_j > p_{j+1}$, then $p_j > p_{j+1} \geq e^{-1}$. Assume $p_j > p_{j+1} \ngeq e^{-1}$, then we must have $e^{-1} > p_j > p_{j+1}$ or $p_j \geq e^{-1} > p_{j+1}$. For any of these two cases, due to $p_{j+1} < e^{-1}$, $p_{j+1}^\mu < e^{-1}$, and then from (38), we must have $\min(\varepsilon_\mu(Sr)|\mathcal{R}_{j+1}) < e^{-1}$, otherwise, it will lead to $\lambda_0 E\left[Sr\log\left(\frac{1+\varepsilon_\mu(Sr)Sr}{1+p_{j+1}^\mu Sr}\right)\middle|\mathcal{R}_{j+1}\right] > 0$. From Lemma 1, $\varepsilon_\mu(sr)$ is a first increasing and then decreasing function over s with an unique maximum point and $\lim_{s\to\infty}\varepsilon_\mu(sr) = e^{-1}$, thus if $\exists s^t$ such that $\varepsilon_\mu(s^t r) < e^{-1}$, then $\varepsilon_\mu(sr)$ must be monotonically increasing over $[0, s^t]$. Therefore $\min(\varepsilon_\mu(Sr)|\mathcal{R}_{j+1}) < e^{-1}$ and $\forall s \in \mathcal{R}_j < \forall s \in \mathcal{R}_{j+1}$, gives $\forall s \in \mathcal{R}_j, \varepsilon_\mu(sr) < \min(\varepsilon_\mu(Sr)|\mathcal{R}_{j+1})$. Thus from (39), we can get $p_j < [\min(\varepsilon_\mu(Sr)|\mathcal{R}_{j+1})]^+$. Due to $[\min(\varepsilon_\mu(Sr)|\mathcal{R}_{j+1})]^+ < p_{j+1}$, we can conclude that $p_j < p_{j+1}$, which is in contradiction with $p_j > p_{j+1}$. Therefore we must have $p_j > p_{j+1} \geq e^{-1}$.

C. Proof of Lemma 3

Let $\omega(s_j)$ denote the objective function of Problem (18), i.e.,

$$\begin{aligned}\omega(s_j) = \lambda_0 E\left[p_{j-1}\beta(Sr) - \beta(p_{j-1}Sr) - \frac{\mu}{\lambda_0}p_{j-1}\middle|\mathcal{R}_{j-1}\right]\\ \Pr(\mathcal{R}_{j-1}) + \lambda_0 E\left[p_j\beta(Sr) - \beta(p_jSr) - \frac{\mu}{\lambda_0}p_j\middle|\mathcal{R}_j\right]\Pr(\mathcal{R}_j).\end{aligned} \tag{40}$$

According to the Fermat's theorem, every local extremum of the differentiable function $\omega(s_j)$ on the open set $(0,\infty)$, satisfies $\frac{\partial \omega(s_j)}{\partial s_j} = 0$, i.e., by using Leibniz integral rule, we have,

$$\frac{\partial \omega(s_j)}{\partial s_j} = \varphi(s_j) f_S(s_j) = 0, \tag{41}$$

where $f_S(s)$ is the pdf of S, and

$$\begin{aligned}\varphi(s_j) =& \lambda_0 \bigg[p_{j-1}\beta(s_j r) - \beta(p_{j-1}s_j r) - \frac{\mu}{\lambda_0} p_{j-1} \\ & - p_j \beta(s_j r) + \beta(p_j s_j r) + \frac{\mu}{\lambda_0} p_j \bigg]. \end{aligned} \tag{42}$$

Since $f_S(s) > 0$, when $s \in (0, \infty)$, we have $\frac{\partial \omega(s_j)}{\partial s_j} = 0$ equivalent to $\varphi(s_j) = 0$.

The first and second derivative of $\varphi(s_j)$ with respect to s_j are

$$\begin{aligned}\frac{\partial \varphi(s_j)}{\partial s_j} &= p_{j-1} A \log\left(\frac{1+s_j r}{1+p_{j-1}s_j r}\right) - p_j A \log\left(\frac{1+s_j r}{1+p_j s_j r}\right), \\ \frac{\partial^2 \varphi(s_j)}{\partial s_j^2} &= \frac{\lambda_0 r^2 (p_{j-1}-p_j)(1-p_{j-1}-p_j-p_{j-1}p_j s_j r)}{(1+s_j r)(1+p_{j-1}s_j r)(1+p_j s_j r)}, \end{aligned} \tag{43}$$

respectively. Let $s^0 = \frac{1-p_{j-1}-p_j}{p_{j-1}p_j r}$. From (43), we have,

(1) $\boxed{\text{Case } p_{j-1} < p_j}$: If $s_j < s^0$, then $\frac{\partial^2 \varphi(s_j)}{\partial s_j^2} < 0$; while if $s_j \geq s^0$, then $\frac{\partial^2 \varphi(s_j)}{\partial s_j^2} \geq 0$. Thus we can obtain that $\frac{\partial \varphi(s_j)}{\partial s_j}$ is first decreasing over $[0, s^0)$ and then increasing over $[s^0, \infty)$. And we also have $\left.\frac{\partial \varphi(s_j)}{\partial s_j}\right|_{s_j=0} = 0$, then,

(a) If $\lim_{s_j \to \infty} \frac{\partial \varphi(s_j)}{\partial s_j} < 0$, i.e., $p_{j-1}\log(\frac{1}{p_{j-1}}) < p_j \log(\frac{1}{p_j})$, $\frac{\partial \varphi(s_j)}{\partial s_j}$ will have only one zero-crossing point, i.e., $s_j = 0$. In this case, we have $\frac{\partial \varphi(s_j)}{\partial s_j} \leq 0$, which implies that $\varphi(s_j)$ is a monotonically decreasing function of s_j. Since $\varphi(s_j)|_{s_j=0} = -\mu(p_{j-1}-p_j) > 0$ (note that $\varphi(s_j)|_{s_j=0} \neq 0$, since according to Lemma 2, if $\mu = 0$, then we must have $p_{j-1} > p_j$), and

$$\begin{aligned}\lim_{s_j \to \infty} \varphi(s_j) &= \lim_{s_j \to \infty} s_j A \left(p_{j-1}\log\left(\frac{1}{p_{j-1}}\right) - p_j \log\left(\frac{1}{p_j}\right)\right) \\ & \quad - \mu(p_{j-1}-p_j) \\ &= -\infty, \end{aligned} \tag{44}$$

we can obtain that $\varphi(s_j) = 0$ has only one solution, thus the function $\omega(s_j)$ has a unique maximum point.

(b) Otherwise, if $p_{j-1}\log\left(\frac{1}{p_{j-1}}\right) > p_j \log\left(\frac{1}{p_j}\right)$, i.e., $\lim_{s_j \to \infty} \frac{\partial \varphi(s_j)}{\partial s_j} > 0$, then $\frac{\partial \varphi(s_j)}{\partial s_j}$ will have two zero-crossing points, i.e., $s_j = 0$ and $s_j = s_\triangle > 0$, where $s_\triangle$ satisfies $\left.\frac{\partial \varphi(s_j)}{\partial s_j}\right|_{s_j = s_\triangle} = 0$, namely,

$$p_{j-1} A \log\left(\frac{1+s_\triangle r}{1+p_{j-1}s_\triangle r}\right) - p_j A \log\left(\frac{1+s_\triangle r}{1+p_j s_\triangle r}\right) = 0. \tag{45}$$

Therefore, we have $\frac{\partial \varphi(s_j)}{\partial s_j} \leq 0$ over the range $[0, s_\triangle]$, and $\frac{\partial \varphi(s_j)}{\partial s_j} > 0$ over the range $(s_\triangle, \infty)$, which gives that the function $\varphi(s_j)$ is first decreasing over $[0, s_\triangle]$ and then increasing over $(s_\triangle, \infty)$. By applying (45), we have

$$\begin{aligned}&\varphi(s_\triangle) = \\ &\lambda_0 p_{j-1}\left(\log(1+s_\triangle r) - \frac{1}{p_{j-1}}\log(1+p_{j-1}s_\triangle r) - \frac{\mu}{\lambda_0}\right) \\ &\quad - \lambda_0 p_j \left(\log(1+s_\triangle r) - \frac{1}{p_j}\log(1+p_j s_\triangle r) - \frac{\mu}{\lambda_0}\right) \end{aligned} \tag{46}$$

Since $-\frac{1}{x}\log(1+x)$ is monotonically increasing over $x > 0$, with $p_{j-1} < p_j$, we have $-\frac{1}{p_{j-1}}\log(1+p_{j-1}s_\triangle r) < -\frac{1}{p_j}\log(1+p_j s_\triangle r)$. Applying this result into (46), we have $\varphi(s_\triangle) < 0$. Then due to $\varphi(0) = -\mu(p_{j-1}-p_j) > 0$ and

$$\begin{aligned}\lim_{s_j \to \infty} \varphi(s_j) &= \lim_{s_j \to \infty} s_j A \left(p_{j-1}\log\left(\frac{1}{p_{j-1}}\right) - p_j \log\left(\frac{1}{p_j}\right)\right) \\ & \quad - \mu\left(p_{j-1}-p_j\right) \\ &= \infty, \end{aligned} \tag{47}$$

we can get that $\varphi(s_j) = 0$ has two solutions, which implies that the function $\omega(s_j)$ in this case has one maximum and one minimum.

(2) $\boxed{\text{Case } p_{j-1} > p_j}$: When $s_j < s^0$, we have $\frac{\partial^2 \varphi(s_j)}{\partial s_j^2} > 0$; while when $s_j \geq s^0$, we have $\frac{\partial^2 \varphi(s_j)}{\partial s_j^2} \leq 0$. Thus we can obtain that the function $\frac{\partial \varphi(s_j)}{\partial s_j}$ is first increasing over $\left[0, s^0\right)$ and then decreasing on $\left[s^0, \infty\right)$. From Lemma 2, we get $\frac{1}{2} \geq p_{j-1} > p_j \geq e^{-1}$ for any $\mu \geq 0$, which gives $p_{j-1}\log\left(\frac{1}{p_{j-1}}\right) < p_j \log\left(\frac{1}{p_j}\right)$, thus $\left.\frac{\partial \varphi(s_j)}{\partial s_j}\right|_{s_j=\infty} < 0$. With $\left.\frac{\partial \varphi(s_j)}{\partial s_j}\right|_{s_j=0} = 0$, we have that $\frac{\partial \varphi(s_j)}{\partial s_j}$ has two zero-crossing points, i.e., $s_j = 0$ and $s_j = s_\triangle > 0$. Therefore, we have $\frac{\partial \varphi(s_j)}{\partial s_j} \geq 0$ over range $[0, s_\triangle]$, and then $\frac{\partial \varphi(s_j)}{\partial s_j} < 0$ over the range $(s_\triangle, \infty)$, which implies that the function $\varphi(s_j)$ is first increasing over $[0, s_\triangle]$ and then decreasing over $(s_\triangle, \infty)$. In this case, since $p_{j-1} > p_j$, we have $-\frac{1}{p_{j-1}}\log(1+p_{j-1}s_\triangle r) > -\frac{1}{p_j}\log(1+p_j s_\triangle r)$. Applying this result into (46), we have $\varphi(s_\triangle) > 0$. Then since $\varphi(0) = -\mu(p_{j-1}-p_j) \leq 0$ and

$$\begin{aligned}\lim_{s_j \to \infty} \varphi(s_j) &= \lim_{s_j \to \infty} s_j A \left(p_{j-1}\log\left(\frac{1}{p_{j-1}}\right) - p_j \log\left(\frac{1}{p_j}\right)\right) \\ & \quad - \mu(p_{j-1}-p_j) \\ &= -\infty, \end{aligned} \tag{48}$$

we obtain that when $\mu = 0$, $\varphi(s_j) = 0$ has only one solution at $s \in (0, \infty)$, which is the only maximum of the function $\omega(s_j)$; while when $\mu > 0$, $\varphi(s_j) = 0$ has two solutions, which indicates that function $\omega(s_j)$ has one minimum and one maximum.

Therefore, based on the above analysis, we can conclude that there exists only one maximum point in Problem (18) and thus only one optimal solution to Problem (18). Obviously, this optimal solution s_j is obtained by solving:

$$\varphi(s_j) = 0, \qquad \text{with} \quad \frac{\partial \varphi(s_j)}{\partial s_j} < 0. \tag{49}$$

Now we will prove that the optimal quantization thresholds $s_j, \forall j = 1, \ldots, \mathcal{L}$ always meet $0 < s_2 < \cdots < s_{\mathcal{L}} < \infty$. As we mentioned before, for each iteration k, given $0 < s_2^{(k)} <$

$\cdots < s_{\mathcal{L}}^{(k)} < \infty$, the optimal $\left\{p_j^{(k+1)}\right\}$ is obtained by (16), i.e., $p_j^{(k+1)} = \left[p_j^{\mu}\right]^+$, where p_j^{μ} is the solution of the equation, $\lambda_0 E\left[Sr\log\left(\frac{1+\varepsilon_\mu(Sr)Sr}{1+p_j^{\mu}Sr}\right)\middle|\mathcal{R}_j^{(k)}\right] = 0$. By applying the mean value theorem (MVT) for integration, above equation can be rewritten as,

$$\lambda_0 \log\left(\frac{1+\varepsilon_\mu(s^q r)s^q r}{1+p_j^{\mu}s^q r}\right) E\left[Sr\middle|\mathcal{R}_j^{(k)}\right] = 0, \tag{50}$$

which gives [2], $p_j^{\mu} = \varepsilon_\mu\left(s_j^q r\right)$, where $s_j^q \in \left(s_j^{(k)}, s_{j+1}^{(k)}\right)$ and obviously, $0 < s_1^q < \cdots < s_{\mathcal{L}}^q < \infty$. Thus,

$$p_j^{(k+1)} = \left[p_j^{\mu}\right]^+ = \left[\varepsilon_\mu\left(s_j^q r\right)\right]^+, \forall j = 1, \ldots, \mathcal{L}. \tag{51}$$

Then, with given $\left\{p_j^{(k+1)}\right\}$, we have proved that the optimal $\left\{s_j^{(k+1)}\right\}$ can be found by solving (19). $\varphi\left(s_j^{(k+1)}\right) = 0$ can be rewritten as,

$$\frac{\mu}{\lambda_0} = \beta\left(s_j^{(k+1)}r\right) - \frac{\beta\left(p_j^{(k+1)}s_j^{(k+1)}r\right) - \beta\left(p_{j-1}^{(k+1)}s_j^{(k+1)}r\right)}{p_j^{(k+1)} - p_{j-1}^{(k+1)}} \tag{52}$$

According to the MVT, we have

$$\begin{aligned}&\frac{\beta\left(p_j^{(k+1)}s_j^{(k+1)}r\right) - \beta\left(p_{j-1}^{(k+1)}s_j^{(k+1)}r\right)}{p_j^{(k+1)} - p_{j-1}^{(k+1)}}\\ &= s_j^{(k+1)}r\left(1+\log\left(1+\triangle p_j s_j^{(k+1)}r\right)\right)\end{aligned} \tag{53}$$

where $\triangle p_j \in \left(\min\left(p_{j-1}^{(k+1)}, p_j^{(k+1)}\right), \max\left(p_{j-1}^{(k+1)}, p_j^{(k+1)}\right)\right)$. Applying (53) into (52), we can obtain,

$$\triangle p_j = \varepsilon_\mu\left(s_j^{(k+1)}r\right). \tag{54}$$

From (51), we have,

$$\begin{aligned}\varepsilon_\mu\left(s_j^{(k+1)}r\right) \in &\left(\min\left(\left[\varepsilon_\mu\left(s_{j-1}^q r\right)\right]^+, \left[\varepsilon_\mu\left(s_j^q r\right)\right]^+\right),\right.\\ &\left.\max\left(\left[\varepsilon_\mu\left(s_{j-1}^q r\right)\right]^+, \left[\varepsilon_\mu\left(s_j^q r\right)\right]^+\right)\right)\end{aligned} \tag{55}$$

When $\mu = 0$, from Lemma 1, $\varepsilon_0(sr)$ is monotonically decreasing in s, thus, from (55), we must have $s_j^{(k+1)} \in \left(s_{j-1}^q, s_j^q\right)$. While when $\mu > 0$, from Lemma 1, (i.e., $\varphi\left(s_j^{(k+1)}\right) = 0,$) (54) may have one or two solutions but must have only one solution in the set $\left(s_{j-1}^q, s_j^q\right)$. Hence, we have proved that only one solution of (54) satisfies (19). Next, we will show that when $\mu > 0$, this unique solution must belong to the set $\left(s_{j-1}^q, s_j^q\right)$.

(1) $\boxed{\text{Case } p_{j-1}^{(k+1)} < p_j^{(k+1)}}$:

a) If $p_{j-1}^{(k+1)}\log\left(\frac{1}{p_{j-1}^{(k+1)}}\right) < p_j^{(k+1)}\log\left(\frac{1}{p_j^{(k+1)}}\right)$, we have shown that in this case, $\frac{\partial\varphi(s_j^{(k+1)})}{\partial s_j^{(k+1)}} < 0$, $s_j^{(k+1)} > 0$ and $\varphi\left(s_j^{(k+1)}\right) = 0$ only has one solution. Then we must have the solution $s_j^{(k+1)} \in \left(s_{j-1}^q, s_j^q\right)$.

b) If $p_{j-1}^{(k+1)}\log\left(\frac{1}{p_{j-1}^{(k+1)}}\right) < p_j^{(k+1)}\log\left(\frac{1}{p_j^{(k+1)}}\right)$, in this case, we have proved that $\varphi\left(s_j^{(k+1)}\right) = 0$ has two solutions, denoted as s_j^1 and s_j^2 and $s_j^1 < s_j^2$, and only $\left.\frac{\partial\varphi(s_j^{(k+1)})}{\partial s_j^{(k+1)}}\right|_{s_j^{(k+1)}=s_j^1} < 0$, namely, only s_j^1 is the maximum point of $\omega\left(s_j^{(k+1)}\right)$. Let s^{M} denote the maximum point of the function $\varepsilon_\mu(sr)$, then according to Lemma 1 and (54), we have $s_j^1 < s^{\mathrm{M}} < s_j^2$. Since $p_{j-1}^{(k+1)} < p_j^{(k+1)}$, we must have $s_{j-1}^q < s_j^1 < s^{\mathrm{M}}$ (otherwise if $s_{j-1}^q \geq s_j^1$, then s_j^2 must belong to the set $\left(s_{j-1}^q, s_j^q\right)$, thus we have $s_j^1 \leq s_{j-1}^q < s_j^2 < s_j^q$, which implies $p_{j-1}^{(k+1)} > p_j^{(k+1)}$), and $s_j^q > s_j^1$ (otherwise, none of s_j^1 and s_j^2 would exist in the set $\left(s_{j-1}^q, s_j^q\right)$), which gives $s_j^1 \in \left(s_{j-1}^q, s_j^q\right)$.

(2) $\boxed{\text{Case } p_{j-1}^{(k+1)} > p_j^{(k+1)}}$: The proof follows along similar lines and is omitted to save space. Therefore, the optimal $\left\{s_j^{(k+1)}\right\}$ obtained by (19) always satisfies $s_j^{(k+1)} \in \left(s_{j-1}^q, s_j^q\right), \forall j = 2, \ldots, \mathcal{L}$, and then since $0 < s_1^q < \cdots < s_{\mathcal{L}}^q < \infty$, we have $0 < s_2^{(k+1)} < \cdots < s_{\mathcal{L}}^{(k+1)} < \infty$. This completes the proof for Lemma 3.

D. Proof for Lemma 5

When $\mu = 0$, from Lemma 2, it is obvious that $p_j > 0, j = 1, \ldots, \mathcal{L}$. When $\mu > 0$, from (21b), we have that the optimal quantized power for region $\mathcal{R}_j, j = 1, \ldots, \mathcal{L}$, is $p_j = \left[p_j^{\mu}\right]^+$, where p_j^{μ} is determined by solving the equation

$$\lambda_0 E\left[Sr\log\left(\frac{1+\varepsilon_\mu(Sr)Sr}{1+p_j^{\mu}Sr}\right)\middle|\mathcal{R}_j\right] = 0, \tag{56}$$

which also can be expressed as,

$$\begin{aligned}&E\left[Sr\log\left(1+p_j^{\mu}Sr\right)\middle|\mathcal{R}_j\right]\\ &= E\left[Sr\left(-1+\left(1+\frac{1}{Sr}\right)\log(1+Sr)\right) - \frac{\mu}{\lambda_0}\middle|\mathcal{R}_j\right].\end{aligned} \tag{57}$$

If $E\left[Sr\left(-1+\left(1+\frac{1}{Sr}\right)\log(1+Sr)\right) - \frac{\mu}{\lambda_0}\middle|\mathcal{R}_j\right] \leq 0$, then to satisfy the equation (57), we must have $p_j^{\mu} \leq 0$, which implies $p_j = \left[p_j^{\mu}\right]^+ = 0$. On the other hand, if $E\left[Sr\left(-1+\left(1+\frac{1}{Sr}\right)\log(1+Sr)\right) - \frac{\mu}{\lambda_0}\middle|\mathcal{R}_j\right] > 0$, p_j^{μ} has to be strictly positive in order to meet the equation (57), which gives $p_j = \left[p_j^{\mu}\right]^+ > 0$.

From Lemma 4, we have $\forall j = 2, \ldots, \mathcal{L}$,

$$s_j r\left(-1+\left(1+\frac{1}{s_j r}\right)\log(1+s_j r)\right) - \frac{\mu}{\lambda_0} \geq 0. \tag{58}$$

Let $y(s) = sr\left(-1+\left(1+\frac{1}{sr}\right)\log(1+sr)\right) - \frac{\mu}{\lambda_0}$. It is easy to get that $\frac{\partial y(s)}{\partial s} = r\log(1+sr) > 0$ for $s > 0$, thus function $y(s)$ is monotonically increasing over variable $s \in (0, \infty)$. Then with (58), we obtain that $y(s_2) \geq 0$ and for $s \in (s_2, \infty)$,

$$sr\left(-1+\left(1+\frac{1}{sr}\right)\log(1+sr)\right) - \frac{\mu}{\lambda_0} > 0, \tag{59}$$

[2]Note that there exists at most one p_j^{μ} for which there may exist two different s_j^q in the set $\left(s_j^{(k)}, s_{j+1}^{(k)}\right)$ which meets $p_j^{\mu} = \varepsilon_\mu\left(s_j^q r\right)$, but it does not affect the following analysis.

which implies, $\forall j = 2, \ldots, \mathcal{L}$,

$$E\left[Sr\left(-1+\left(1+\frac{1}{Sr}\right)\log(1+Sr)\right)-\frac{\mu}{\lambda_0}\middle|\mathcal{R}_j\right] > 0. \tag{60}$$

Therefore, we have $p_j > 0, \forall j = 2, \ldots, \mathcal{L}$, i.e., the optimal quantized powers in the last $\mathcal{L}-1$ regions are strictly positive. However due to $y(0) = -\frac{\mu}{\lambda_0} < 0$ and $y(s_2) \geq 0$, $E\left[y(S)|\mathcal{R}_1\right]$ could be positive, negative or zero, implying the power level in the first region could be positive or zero, namely, $p_1 \geq 0$.

Now we will show that the optimal power levels of case $\mu > 0$ always satisfy $0 \leq p_1 < \cdots < p_{m-1} < p_m$ and $p_m > p_{m+1} > \cdots > p_{\mathcal{L}} \geq e^{-1}$, where p_m is the largest power level. By applying the MVT for integration, (56) becomes, $\lambda_0 \log\left(\frac{1+\varepsilon_\mu(s^q r)s^q r}{1+p_j^\mu s^q r}\right) E\left[Sr|\mathcal{R}_j\right] = 0$, which gives $p_j^\mu = \varepsilon_\mu\left(s_j^q r\right), \forall j = 2, \ldots, \mathcal{L}$, with $s_j^q \in (s_j, s_{j+1})$. Obviously, $0 < s_1^q < \cdots < s_{\mathcal{L}}^q < \infty$. We have proved above that all the power levels are always strictly positive except p_1, which could be zero or positive (i.e., $p_1 \geq 0$), thus, $p_1 = \left[\varepsilon_\mu\left(s_1^q r\right)\right]^+$, $p_j = \varepsilon_\mu\left(s_j^q r\right), \forall j = 2, \ldots, \mathcal{L}$. According to Lemma 1, when $\mu > 0$, $\varepsilon_\mu(sr)$ is first increasing in $s \in \left[0, s^{\mathrm{M}}\right]$ (where s^{M} is the unique maximum point of $\varepsilon_\mu(sr)$), and then decreasing over $s \in \left(s^{\mathrm{M}}, \infty\right)$ with $\lim_{s\to\infty} \varepsilon_\mu(sr) = e^{-1}$. Without loss of generality, assuming $s^{\mathrm{M}} \in [s_n, s_{n+1}), n \in [1, \mathcal{L}]$, we must have $0 < s_1^q < \cdots < s_{n-1}^q < s_n \leq s^{\mathrm{M}}$ and $s^{\mathrm{M}} < s_{n+1} < s_{n+1}^q \cdots < s_{\mathcal{L}}^q < \infty$, which gives $0 \leq p_1 < \cdots < p_{n-1}$ and $p_{n+1} > \cdots > p_{\mathcal{L}} \geq e^{-1}$, respectively. If $p_{n-1} < p_n$ and $p_n > p_{n+1}$, then $0 \leq p_1 < \cdots < p_{n-1} < p_n$ and $p_n > p_{n+1} > \cdots > p_{\mathcal{L}} \geq e^{-1}$. Whereas, if $p_n < p_{n-1}$ (or $p_n < p_{n+1}$), then we must have $s_M < s_n^q < s_{n+1}$ (or $s_n < s_n^q < s_M$) , which implies $p_n > p_{n+1}$ (or $p_n > p_{n-1}$), thus, we obtain $0 \leq p_1 < \ldots p_{n-2} < p_{n-1}$ and $p_{n-1} > p_n > \cdots > p_{\mathcal{L}} \geq e^{-1}$ (or $0 \leq p_1 < \cdots < p_n < p_{n+1}$ and $p_{n+1} > p_{n+2} > \cdots > p_{\mathcal{L}} \geq e^{-1}$). Any of these cases can be written in the form of $0 \leq p_1 < \cdots < p_{m-1} < p_m$ and $p_m > p_{m+1} > \cdots > p_{\mathcal{L}} \geq e^{-1}$, where $p_m, m \in [1, \mathcal{L}]$ is the largest power level.

E. Proof for Lemma 6

First of all, we will prove that as $\mathcal{L} \to \infty$, $p_1 \to \varepsilon_\mu(s_t r)$ where $s_t = \lim_{\mathcal{L}\to\infty} s_2 \geq s_{\mathrm{lb}}$. From Remark 3, $s_{\mathcal{L}} > \cdots > s_2 \geq s_{\mathrm{lb}} > 0$ and $\varepsilon_\mu(s_{\mathrm{lb}}) = 0$. When $\mathcal{L} \to \infty$, according to the Monotonic Sequence Theorem, this strictly monotonic and lower bounded sequence $\{s_j\}_{j=2}^{\mathcal{L}}$, must converge to its greatest lower bound $s_t = \lim_{\mathcal{L}\to\infty} s_2$. Consequently, given an arbitrarily small $\epsilon > 0$, we can always find a sufficiently large $\mathcal{L}$ such that $s_3 - s_2 < \epsilon$. Therefore, when $\mathcal{L} \to \infty$, $s_3 - s_2 \to 0$. From (23b) with $j = 2$, we can obtain $p_2 = \varepsilon_\mu\left(s_2^q r\right)$, where $s_2^q \in (s_2, s_3)$, then as $\mathcal{L} \to \infty$, $s_3 - s_2 \to 0$ implies $s_2^q \to s_t$. Therefore, $p_2 \to \varepsilon_\mu(s_t r)$, as $\mathcal{L} \to \infty$. By applying the MVT, the (23a) with $j = 2$ can be rewritten as $\triangle p_2 = \varepsilon_\mu(s_2 r)$, where $\triangle p_2$ satisfies

$$\frac{\beta(p_2 s_2 r) - \beta(p_1 s_2 r)}{p_2 - p_1} = \left.\frac{\partial \beta(p s_2 r)}{\partial p}\right|_{p=\triangle p_2}, \tag{61}$$

and $\triangle p_2 \in (\min(p_1, p_2), \max(p_1, p_2))$. Obviously, $\lim_{\mathcal{L}\to\infty} \triangle p_2 = \varepsilon_\mu(s_t r) = \lim_{\mathcal{L}\to\infty} p_2$. Let $h = p_2 - p_1$, and then as $\mathcal{L} \to \infty$, (61) becomes,

$$\lim_{\mathcal{L}\to\infty} \frac{\beta(p_2 s_2 r) - \beta((p_2 - h)s_2 r)}{h} = \left.\frac{\partial \beta(p s_2 r)}{\partial p}\right|_{p=p_2}, \tag{62}$$

which gives $\lim_{\mathcal{L}\to\infty} h = 0$, i.e., $\lim_{\mathcal{L}\to\infty} p_1 = \lim_{\mathcal{L}\to\infty} p_2 = \varepsilon_\mu(s_t r)$.

Now suppose $p_1 > 0$ for any given arbitrarily large $\mathcal{L}$, which implies $s_t > s_{\mathrm{lb}}$. Then p_1 must satisfy the optimality condition $\lambda_0 E\left[Sr \log\left(\frac{1+\varepsilon_\mu(Sr)Sr}{1+p_1 Sr}\right)\middle|\mathcal{R}_1\right] = 0$. For any given value of $\mathcal{L}$, the region $\mathcal{R}_1$ can be divided into two parts $\mathcal{R}_{11}$ and $\mathcal{R}_{12}$, where $\mathcal{R}_{11} = [0, s_t)$ and $\mathcal{R}_{12} = [s_t, s_2)$. As $\mathcal{L}$ becomes arbitrarily large, $\mathcal{R}_{12}$ becomes vanishingly small, and then

$$\begin{aligned}&\lambda_0 E\left[Sr \log\left(\frac{1+\varepsilon_\mu(Sr)Sr}{1+p_1 Sr}\right)\middle|\mathcal{R}_1\right]\\ &\to \lambda_0 E\left[Sr \log\left(\frac{1+\varepsilon_\mu(Sr)Sr}{1+\varepsilon_\mu(s_t r)Sr}\right)\middle|\mathcal{R}_{11}\right] < 0,\end{aligned} \tag{63}$$

which is a contradiction to the optimality condition for $p_1 > 0$. Thus, p_1 must be zero, which implies $p_1 = \varepsilon_\mu(s_{\mathrm{lb}} r)$ for a sufficiently large $\mathcal{L}$. Since $\lim_{\mathcal{L}\to\infty} p_1 = \varepsilon_\mu(s_t r) = \varepsilon_\mu(s_{\mathrm{lb}} r)$, we must have $\lim_{\mathcal{L}\to\infty} s_2 = s_t = s_{\mathrm{lb}}$.

REFERENCES

[1] H. A. Willebrand and B. S. Ghuman, "Fiber optics without fiber," *IEEE Spectrum*, vol. 38, no. 8, pp. 41–45, Aug. 2001.

[2] K. Chakraborty and P. Narayan, "The Poisson fading channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2349–2364, July 2007.

[3] K. Chakraborty and S. Dey, "Service-outage-based power and rate control for Poisson fading channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2304–2318, May 2009.

[4] 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.

[5] K. Chakraborty, "Capacity of the MIMO optical fading channel," in *Proc. 2005 Int. Symp. Inf. Theory*, pp. 530–534.

[6] K. Chakraborty, S. Dey, and M. Franceschetti, "Outage capacity of MIMO Poisson fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 4887–4907, Nov. 2008.

[7] N. Letzepis and A. Guillén i Fábregas, "Outage probability of the Gaussian MIMO free-space optical channel with PPM," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3682–3690, Dec. 2009.

[8] Y. Y. He and S. Dey, "Power allocation in spectrum sharing cognitive radio networks with quantized channel information," *IEEE Trans. Commun.*, vol. 59, no. 6, pp. 1644–1656, June 2011.

[9] P. A. Chou, T. Lookabaugh, and R. M. Gray "Entropy-constrained vector quantization," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 31, no. 1, pp. 31–42, Jan. 1989.

[10] T. T. Kim and M. Skoglund, "On the expected rate of slowly fading channels with quantized side information," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 820–829, Apr. 2007.

[11] L. Lin, R. D. Yates, and P. Spasojević, "Adaptive transmission with discrete code rates and power levels," *IEEE Trans. Commun.*, vol. 51, no. 12, pp. 2115–2125, Dec. 2003.

[12] K. Huang and R. Zhang "Cooperative feedback for multiantenna cognitive radio networks," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 747–758, Feb. 2011.

[13] R. Cendrillon, W. Yu, M. Noonen, J. Verlinden, and T. Bostoen, "Optimal multiuser spectrum balancing for digital subscriber lines," *IEEE Trans. Commun.*, vol. 54, no. 5, pp. 922–933, May 2006.

Yuanyuan He (S'08-M'12) received her B.Eng. degree in communication engineering from Yanshan University, China, in 2005, the M.Eng. degree in electronic engineering from Harbin Institute of Technology, China, in 2007, and the Ph.D. degree in the Department of Electrical and Electronic Engineering, University of Melbourne, Melbourne, Australia, in 2011, where she is has been a Postdoctoral Research Fellow since 2011. Her research interests lie in the area of wireless communications including resource allocation, cognitive radio, and limited feedback.

Subhrakanti Dey (SM'06) was born in India in 1968. He received the B.Tech. and M.Tech. degrees from the Department of Electronics and Electrical Communication Engineering, Indian Institute of Technology, Kharagpur, India, in 1991 and 1993, respectively, and the Ph.D. degree from the Department of Systems Engineering, Research School of Information Sciences and Engineering, Australian National University, Canberra, Australia, in 1996.

He has been with the Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Australia, since February 2000, where he is currently a full Professor. From September 1995 to September 1997 and September 1998 to February 2000, he was a postdoctoral Research Fellow with the Department of Systems Engineering, Australian National University. From September 1997 to September 1998, he was a post-doctoral Research Associate with the Institute for Systems Research, University of Maryland, College Park. His current research interests include networked control systems, wireless communications and networks, signal processing for sensor networks, and stochastic and adaptive estimation and control.

Prof. Dey currently serves on the Editorial Board of *Elsevier Systems and Control Letters*. He was also an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He is a Senior Member of IEEE.