Effective Capacity Maximization With Statistical Delay and Effective Energy Efficiency Requirements

Leila Musavian, *Member, IEEE*, and Qiang Ni, *Senior Member, IEEE*

***Abstract*—This paper presents the three-fold energy, rate and delay tradeoff in mobile multimedia fading channels. In particular, we propose a rate-efficient power allocation strategy for delay-outage limited applications with constraints on energy-per-bit consumption of the system. For this purpose, at a target delay-outage probability, the link-layer energy efficiency, referred to as effective-EE, is measured by the ratio of effective capacity (EC) and the total expenditure power, including the transmission power and the circuit power. At first, the maximum effective-EE of the channel at a target delay-outage probability is found. Then, the optimal power allocation strategy is obtained to maximize EC subject to an effective-EE constraint with the limit set at a certain ratio of the maximum achievable effective-EE of the channel. We then investigate the effect of the circuit power level on the maximum EC. Further, to set a guideline on how to choose the effective-EE limit, we obtain the transmit power level at which the rate of increasing EC (as a function of transmit power) matches a scaled rate of losing effective-EE. Analytical results show that a considerable EC-gain can be achieved with a small sacrifice in effective-EE from its maximum value. This gain increases considerably as the delay constraint becomes tight.**

***Index Terms*—Delay-outage probability constraint, energy-rate-delay tradeoff, effective capacity, effective energy efficiency, Nakagami fading.**

I. Introduction

Energy efficiency (EE), defined as the number of communicated bits per unit transmission power in the units of b/J/Hz, is considered as one of the design performance metrics for future wireless communications systems, e.g., 5G [1], [2]. In many cases, however, increasing EE comes at the price of decreasing the throughput [3]. The rate-energy tradeoff is studied in various research articles and projects [4]–[8]. In particular, [4] proposed an optimum design for joint transmission time and modulation parameters to reduce the total energy consumption for sending a given number of bits. Authors in [5] proposed an energy-efficient resource allocation in orthogonal frequency division multiple access (OFDMA) channels and showed that, similar to flat-fading channels, EE curve as a function of the transmission power is a bell-shape curve. Henceforth, increasing the transmission power beyond the corresponding point for the maximum EE, does not improve EE. Whereas, the rate is a monotonically increasing function of the transmission power. The results in [3], [5]–[8] are based on Shannon capacity which is mostly considered as the suitable capacity metric for systems with delay insensitive applications. The rate-energy tradeoff is shown to be more pronounced in systems with applications having delay requirements [9], [10], e.g., mobile multimedia communications.

Delay quality-of-service (QoS) requirement is indeed an essential factor for enabling mobile multimedia communications systems [11]. However, due to the variations in the wireless mobile fading channels, satisfying a deterministic delay constraint is either impossible or results in a very low transmission rate [11], [12]. Hence, in many systems, a certain delay-outage probability is tolerable, a characteristic that has paved the way for a cross-layer analysis of the wireless fading channel through a link-layer capacity model called effective capacity (EC) [11]. EC provides a measure for the maximum constant arrival rate under buffer-violation probability constraint by capturing the decaying rate of the buffer size probability for large queue lengths and incorporating it into the large deviations theory. By using the EC model, a rate-efficient power allocation technique under delay-outage probability constraint is proposed in [13]. The paper shows that to maintain the throughput while satisfying tighter delay constraints, higher transmission powers are required [11]. The maximum transmission power, however, is limited by different factors in wireless devices, e.g., by device size. Henceforth, energy-efficient transmission techniques are required.

Considering a deterministic delay constrained channel, [14] proposes an optimal power-rate allocation policy that minimizes the system energy consumption when the channel is considered to evolve as a Markov process. This work was extended to block-fading channels in [15], wherein an optimal power allocation strategy that determines the number of scheduled bits to be transmitted in each time-slot so that the total energy consumption is minimized while a deterministic delay constraint is satisfied was obtained. A game-theoretic approach is proposed in [16] wherein a de-centralized optimum rate and power allocation scheme is found when the utility function is defined as the difference between the throughput and consumed power when the experienced delay by the arrived packets is kept below a predefined threshold. The maximum achievable effective-EE of a flat-fading channel under a delay-outage probability constraint was obtained in [17] and [18], wherein the effective-EE is defined as the ratio of EC to the

Manuscript received May 27, 2014; revised September 24, 2014 and January 9, 2015; accepted February 22, 2015. Date of publication March 18, 2015; date of current version July 8, 2015. Part of this work was presented at the IEEE International Conference on Communications, Sydney, Australia, June 2014. This work was supported in part by the UK EPSRC under grant number EP/K011693/1 and the EU FP7 under grant number PIRSES-GA-2013-610524. The associate editor coordinating the review of this paper and approving it for publication was G. Mao.

The authors are with the School of Computing and Communications, Info-Lab21, Lancaster University, Lancaster, LA1 4WA, U.K. (e-mail: l.musavian@lancaster.ac.uk; q.ni@lancaster.ac.uk).

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/TWC.2015.2412542

transmission power without taking into account the effect of the circuit power consumption on effective-EE. In particular, [17] shows that the minimum received signal energy per bit for reliable communication is achieved when the signal-to-noise-ratio (SNR) tends to zero, i.e., either bandwidth tends to infinity or transmission power tends to zero. At near-zero transmission power, this minimum energy per bit is shown to be similar to systems with no delay requirements, studied in [19], whereas, for infinite-bandwidth systems, this minimum energy is much higher than that of a delay-unconstrained system [17]. The analysis in [17]–[19], however, does not account for the circuit power consumption, and hence, the rate-energy tradeoff could only be characterized under asymptotic conditions.

Taking into account non-zero circuit powers, the maximum achievable effective-EE under delay-outage probability constraints is found in [9], [20] for single carrier and multi-carrier channels, respectively. In particular, [9] considers a rate-independent circuit power and shows that, unlike systems with zero circuit power, the maximum effective-EE can be achieved at non-zero transmit powers. Further, [20] proposes an optimal power and subchannel allocation technique for achieving the maximum effective-EE by using a concave-convex fractional programming approach. On the other hand, assuming a delay-outage probability constrained channel, the maximum effective-EE was obtained in [21], wherein the optimal power allocation was obtained for high and low SNRs. Further, the maximum achievable effective-EE of a cellular system subject to minimum rate requirements is studied in [22]. Effective-EE as a function of the transmission power curve is shown to be a bell-shape curve in [9], [23] with sharper head when delay requirement becomes loose. To visualize the effective-EE and EC tradeoff more clearly, we plot the achievable effective-EE versus EC for two different target delay-outage probabilities in Fig. 1.[1] The figure reveals that while the slope of effective-EE curve is sharp at higher values of EC, it becomes slow around the maximum effective-EE value, i.e., $\eta_{\max}$. We conclude that a small reduction in effective-EE around its peak value results in a significant gain in EC. This gain, however, becomes less when the operating point departs from $\eta_{\max}$. Moreover, the figure shows that the peak in the curve is sharper in systems with looser delay requirement. Henceforth, in comparison between two systems with loose and tight delay constraints, the gain in the achievable EC, as a result of departing from the maximum effective-EE point, is larger when delay is tighter. For example, as shown in Fig. 1, at $\theta = 0.001$, with 10% loss in effective-EE, the gain in EC is 169%. Whereas, at $\theta = 0.1$ with the same loss rate in effective-EE (i.e., 10%), the EC-gain is 229% which is larger than 169%. These observations illustrate the important impact of delay constraint on the effective-EE and EC tradeoff in delay-limited systems which motivates the work to be presented in this paper.

[1]In this figure, θ represents the delay-outage probability constraint. Higher θ means more stringent delay requirement. Detailed explanation of θ will be given in Section II. This figure is generated for use in this paper. The general simulation parameters are the same as [9]. The channel is Rayleigh fading with unit-variance and the scaled circuit power value considered in this figure is 0 dB.

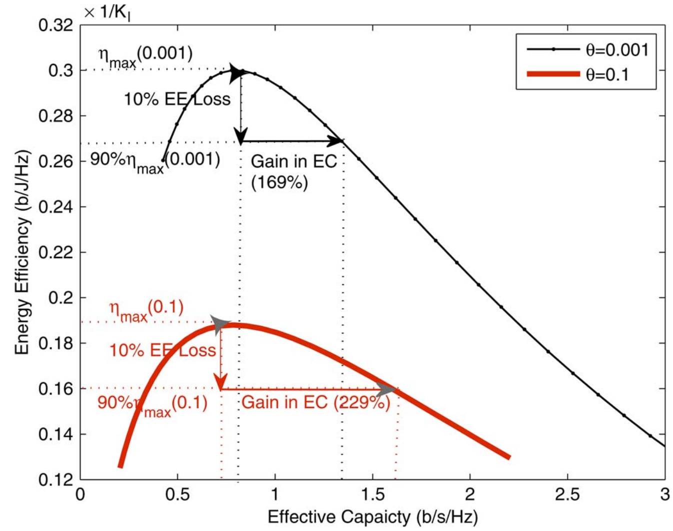

Fig. 1. Effective-EE versus EC for various delay requirements in Rician fading channels.

In this paper, we consider a delay-outage probability constrained system in Nakagami-m fading channels and investigate the maximum EC of such systems under an effective-EE constraint. The effective-EE constraint limit is set as a percentage of the maximum achievable effective-EE of the channel. For a target delay-outage probability, we represent the effective-EE, expressed in units of b/J/Hz, as the ratio of the achieved link EC to the sum of the transmission power and the circuit power. We start by proposing an optimal power allocation for achieving the maximum EC of the system with no input power constraint. The optimality of the power allocation is established by using the fact that the EC-maximization objective function is a strictly concave function in the transmission power and the effective-EE is a strictly quasi-concave function, and as such, its upper contour set is convex and a unique maximum exists. Using Lagrangian methods, we derive the optimal power allocation strategy which is shown to be similar to an EC-maximization problem subject to an input power limit set at P_{t}^{*}, which is the operating transmit power for satisfying the effective-EE constraint at equality. We further consider a system under joint constraints on effective-EE and input power, with a power limit $P_{\max}$, and show that when $P_{\max}$ is large enough for the problem to be feasible, the operating transmit power should be limited to $\min(P_{\max}, P_{\mathrm{t}}^{*})$.

The simulation results in this work show that the value set for the effective-EE constraint limit affects the maximum achievable EC considerably. Furthermore, we prove that the effect of P_{c} on the maximum achievable EC depends on how the effective-EE constraint is set. These facts highlight the needs for setting guidelines on how to choose the required effective-EE limit. We, consequently, propose an EC-gain effective-EE-loss rate matching technique. Specifically, we set the effective-EE limit at a point at which the gain-rate of the EC is equal to a scaled loss-rate in the effective-EE as functions of the transmission power. Finally, we analytically prove that in systems with extremely stringent delay requirements, the achievable EC under effective-EE constraints increases monotonically as Nakagami-m fading parameter increases.

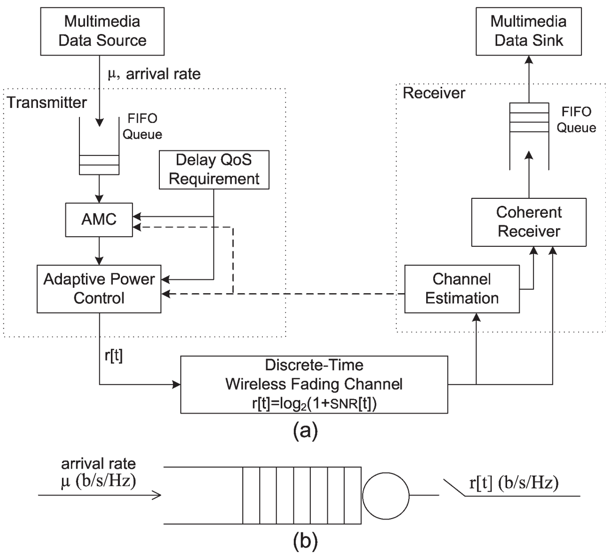

Fig. 2. System model. (a) System block diagram; (b) Equivalent queuing model.

The main contributions of this paper can be listed as:

- This paper obtains the maximum EC of a Nakagami-m fading channel when a certain effective-EE level is required by the system. The results are beneficial to system designers since the simulation results show that considerable EC gain can be achieved with a small reduction in achievable effective-EE.
- We analytically investigate the effect of the circuit power level on the maximum achievable EC under an effective-EE constraint.
- We prove that in systems with extremely stringent delay requirement, the maximum achievable capacity increases as the Nakagami-m fading parameter increases.
- Motivated by the fact that the achievable EC depends strongly on the required effective-EE level, we propose a method to set this value based on matching the gain-rate of EC with a scaled loss-rate of effective-EE.
- Simulation results indicate that with small decrease in effective-EE, delay-outage probability decreases considerably. Further it is shown that in Rayleigh and Rician fading channels, as delay becomes more stringent, the gain in EC increases. This trend, however, does not apply to Nakagami-m channels with $m = 0.5$.

The remainder of the paper is organized as follows. In Section II, we provide the system model. We obtain the optimal power allocation strategy to maximize EC subject to joint input power and effective-EE constraints in Section III. Further, we investigate the maximum achievable EC subject to an effective-EE-loss EC-gain rate matching technique. Finally, the numerical results are given in Section V, followed by conclusions in Section VI.

II. System Model

We consider a point-to-point communication system operating in a flat-fading channel similar to the one studied in [13], which is depicted in Fig. 2(a). The simplified equivalent queuing model for the system is provided in Fig. 2(b). User data initially enters a first-in-first-out (FIFO) buffer at a constant arrival rate μ (expressed in b/s/Hz). The transmitter sets its transmission rate and power based on the delay requirement of the system and the channel state information (CSI) obtained through a feedback channel. As a result, user data is read out of the FIFO buffer and transmitted over the wireless fading channel at a variable service rate. Ideal Nyquist transmission symbol rate is assumed, henceforth, the symbol duration T_{s} is equal to $T_{\mathrm{s}} = \frac{1}{B}$, where B is the system bandwidth. The wireless channel is considered to be block-fading, i.e., the channel response is fixed during a fading block, and independently changes from one fading block to another. The length of each fading block, denoted by T_{f}, is assumed to be an integer multiple of symbol duration T_{s}.

The service rate process $\{r[t], t = 1, 2, \ldots, T_{\mathrm{f}}\}$ over this block-fading channel using adaptive transmission is considered to be stationary and ergodic. Using Shannon capacity equation, the service rate (in b/s/Hz) is given as

$$r[t] = \log_2\left(1 + \frac{P_{\mathrm{t}}[t]\gamma[t]}{P_{\mathcal{L}}\sigma_{\mathrm{n}}^2}\right). \tag{1}$$

Here, $P_{\mathrm{t}}[t]$ is the transmission power in fading-block t, σ_{n}^2 indicates the noise power, $P_{\mathcal{L}}$ denotes the distance-based path-loss, and $\gamma[t]$ is the normalized channel power gain[2] of the considered unit-variance Nakagami-m block-fading channel with the probability density function (PDF) [24]

$$f_\gamma(\gamma) = \frac{m^m\gamma^{m-1}}{\Gamma(m)}\mathrm{e}^{-m\gamma},$$

where $\Gamma(m) = \int_0^\infty w^{m-1}\mathrm{e}^{-w}dw$ is the Gamma function [25].

A. Effective Capacity

Assuming that the delay-outage probability is limited, we use the link-layer effective capacity concept to find the achievable throughput of the fading channel. We note that the delay occurred in the signal transmission, is related on how long the data is kept in the FIFO buffer before it is read out to the wireless channel. In more detail, since in a FIFO buffer, the head of the queue will clear out first, the buffer queue length represents the delay in signal transmission. To analyze the buffer queue length overflow probability, and indirectly, the delay-outage probability, we incorporate the concept of EC [11], which provides a measure for the maximum constant arrival rate that a given service rate can support subject to a QoS exponent requirement defined by θ. Given that the assumptions for the Gartner-Ellis theorem [26, Pages 34–36] are satisfied,[3] the EC of an independent and identically distributed (i.i.d.)

[2]Hereafter, we omit the time index t wherever it is clear from the context.

[3]The Gartner-Ellis theorem assumptions are listed as the Gartner-Ellis function (i) exists for all real θ, (ii) is strictly convex, and (iii) is essentially smooth. A convex function $l(x)$ is essentially smooth if (i) its domain is non-empty, (ii) $l(x)$ is differentiable through its domain and (iii) $l(x)$ is a steep function [27, Page 44]. The conditions for a convex function being steep is given in [28, Page 30].

block-fading channel at a delay exponent θ can be expressed as [11][4]

$$\mathrm{EC}(\theta) = -\frac{1}{\theta T_\mathrm{f}} \ln\left(\mathbb{E}\left[\mathrm{e}^{-\theta T_\mathrm{f} r[t]}\right]\right), \tag{2}$$

where $\mathbb{E}[\cdot]$ indicates the expectation operator. We note that the EC formulation is given based on θ. The relation between θ and the buffer-length overflow probability is given through the large deviation principle theorem. In particular, assuming that the steady-state queue length, $q(\infty)$, exists, the probability that the queue length exceeds a certain threshold x is shown to decay exponentially fast according to [11], [12],

$$\Pr\left\{q(\infty) \geq x\right\} \approx \varepsilon \mathrm{e}^{-x\theta} \tag{3}$$

where θ is found when the maximum arrival rate is equal to the EC, i.e., $\mu = \mathrm{EC}(\theta)$ (given in (2)). Here, $\mathfrak{f}(x) \approx \mathfrak{g}(x)$ means that $\lim_{x\to\infty} \frac{\mathfrak{f}(x)}{\mathfrak{g}(x)} = 1$, and ε is the probability of a non-empty buffer, which can be approximated by the ratio of the constant arrival rate to the average service rate, i.e., $\varepsilon \sim \frac{\mu}{\mathbb{E}[r[t]]}$, [12], [13]. Large and small values of θ correspond to fast and slow decaying rates indicating stringent and loose QoS requirements, respectively. For example, when $\theta \to 0$, the system can tolerate an arbitrarily long delay, whereas the system cannot tolerate any delay when $\theta \to \infty$. Finally, the probability that the delay exceeds a maximum delay bound $D_{\max}$ is related to θ according to [11]

$$P^{\mathrm{out}}_{\mathrm{delay}} = \Pr\{\mathrm{Delay} \geq D_{\max}\} \approx \varepsilon \mathrm{e}^{-\theta\mu D_{\max}}, \tag{4}$$

where $D_{\max}$ is in units of a symbol period ($T_\mathrm{s} = 1/B$). Hence, a source that requires a delay-bound violation probability of at most $P^{\mathrm{out}}_{\mathrm{delay}}$, needs to limit its data rate to a maximum of μ, where μ is the solution to $\mu = \mathrm{EC}(\theta)$ given in (2).

B. Effective-EE

We formulate the effective-EE of a system under delay-outage probability constraint as the ratio of the EC to the total power consumption at the transmitter. The total transmission power can be modeled as $P_{\mathrm{tot}}[t] = \frac{P_\mathrm{t}[t]}{\epsilon} + P_\mathrm{c}$, where $0 \leq \epsilon \leq 1$ is the power amplifier (PA) efficiency, and $\frac{P_\mathrm{t}[t]}{\epsilon}$ is the total input power to the PA. P_c is the constant circuit power that corresponds to the power dissipation of the transmitter circuitry, which is considered to be independent of the transmission rate in this paper. The effective-EE can therefore be expressed as

$$\text{Effective-EE} = \eta(\theta) = \frac{\mathrm{EC}(\theta)}{P_\mathrm{c} + \frac{1}{\epsilon}\mathbb{E}[P_\mathrm{t}[t]]}, \tag{5}$$

where $\eta(\theta)$ indicates that effective-EE as a function of the delay exponent θ.

[4]It is shown that when the channel is block-fading, the Gartner-Ellis theorem assumptions are satisfied [11].

III. Optimal Power Allocation

Given that the transmit power of a system determines its rate and energy consumption, the performance in terms of effective-EE and EC tradeoff can be optimized by adaptively allocating the power over time based on the channel condition and the system delay requirement. To analyze the effective-EE and EC tradeoff characteristics, we formulate the optimization problem to maximize EC subject to constraints on the maximum average transmission power and required effective-EE. In other words, instead of maximizing effective-EE, we maximize EC constraining on the effective-EE, because the maximum achievable effective-EE of a system has a finite value. As shown in Fig. 1, a small sacrifice in the effective-EE from its maximum can result in a significant gain in EC (spectral efficiency) of the system. In this paper, we adapt the transmission power at the beginning of each fading-block based on the CSI and the delay constraint. Therefore, hereafter, we refer to the instantaneous transmit power $P_\mathrm{t}[t]$ by $P_\mathrm{t}(\gamma)$ to indicate that the transmission power is a function[5] of the channel power gain γ.

The EC-maximization problem of a system with delay-outage probability requirement under an average input power constraint and an effective-EE constraint can be mathematically expressed as

$$\mathrm{EC}_{\mathrm{opt}}(\theta) = \max_{P_\mathrm{t}(\gamma)\geq 0} -\frac{1}{\theta T_\mathrm{f}} \ln\left(\mathbb{E}_\gamma\left[\left(1 + \frac{P_\mathrm{t}(\gamma)\gamma}{K_\ell}\right)^{-\xi(\theta)}\right]\right) \tag{6a}$$

$$\text{subject to}: \quad \text{Effective-EE}$$

$$= \frac{-\frac{1}{\theta T_\mathrm{f}} \ln\left(\mathbb{E}_\gamma\left[\left(1 + \frac{P_\mathrm{t}(\gamma)\gamma}{K_\ell}\right)^{-\xi(\theta)}\right]\right)}{P_\mathrm{c} + \frac{1}{\epsilon}\mathbb{E}_\gamma\left[P_\mathrm{t}(\gamma)\right]} \geq \eta_{\min}(\theta) \tag{6b}$$

$$\mathbb{E}_\gamma\left[P_\mathrm{t}(\gamma)\right] \leq P_{\max}, \tag{6c}$$

where $K_\ell = P_\mathcal{L}\sigma_\mathrm{n}^2$ is the noise power that includes both additive white Gaussian noise (AWGN) and path loss, $\xi(\theta) = \frac{\theta T_\mathrm{f}}{\ln 2}$, and $\mathbb{E}_\gamma[\cdot]$ indicates the expectation over the PDF of γ. Further, $\eta_{\min}(\theta)$ denotes the required effective-EE limit and is chosen as a ratio of the maximum achievable effective-EE at the target delay-outage probability, $\eta_{\max}(\theta)$, which is formulated as

$$\eta_{\max}(\theta) = \max_{P_\mathrm{t}(\gamma)\geq 0} \frac{-\frac{1}{\theta T_\mathrm{f}} \ln\left(\mathbb{E}_\gamma\left[\left(1 + \frac{P_\mathrm{t}(\gamma)\gamma}{K_\ell}\right)^{-\xi(\theta)}\right]\right)}{P_\mathrm{c} + \frac{1}{\epsilon}\mathbb{E}_\gamma[P_\mathrm{t}(\gamma)]}. \tag{7}$$

Specifically, under a delay QoS exponent θ, the ratio of the required effective-EE limit ($\eta_{\min}(\theta)$) over the maximum achievable effective-EE ($\eta_{\max}(\theta)$) is referred to as the *EE-loss-rate* which is expressed as

$$\alpha_{\mathrm{EE}} = \frac{\eta_{\min}(\theta)}{\eta_{\max}(\theta)}, \tag{8}$$

where $0 \leq \alpha_{\mathrm{EE}} \leq 1$. We further define the EC that can be achieved corresponding to $\eta_{\max}(\theta)$ by $\mathrm{EC}_{\eta_{\max}}(\theta)$. The ratio of

[5]Note that the transmission power is also a function of the delay QoS exponent θ.

$\mathrm{EC_{opt}}(\theta)$ over $\mathrm{EC}_{\eta_{\max}}(\theta)$ is further defined as the *EC-gain-rate* and is formulated as

$$\alpha_{\mathrm{EC}} = \frac{\mathrm{EC_{opt}}(\theta)}{\mathrm{EC}_{\eta_{\max}}(\theta)}. \tag{9}$$

A. Optimal Power Allocation With No Input Power Constraint

First, the EC-maximization problem without considering the maximum input transmit power constraint is tackled, serving as a milestone towards finding an EC-optimal power allocation subject to the joint effective-EE and input power constraints. Further, to normalize the system performance with respect to the path loss and noise effect (K_ℓ), we scale the EC-maximization problem (6a) with K_ℓ. The EC-maximization problem, hence, can be expressed as

$$\mathrm{EC_{opt}}(\theta) = \max_{P_{\mathrm{r}}(\gamma)\geq 0} -\frac{1}{\theta T_{\mathrm{f}}} \ln\left(\mathbb{E}_\gamma\left[(1+P_{\mathrm{r}}(\gamma)\gamma)^{-\xi(\theta)}\right]\right) \tag{10a}$$

$$\text{subject to:} \quad \text{Effective-EE} = \frac{-\frac{1}{\theta T_{\mathrm{f}}}\ln\left(\mathbb{E}_\gamma\left[(1+P_{\mathrm{r}}(\gamma)\gamma)^{-\xi(\theta)}\right]\right)}{K_\ell\left(P_{\mathrm{c_r}}+\frac{1}{\epsilon}\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right]\right)} \geq \eta_{\min}(\theta), \tag{10b}$$

where $P_{\mathrm{r}}(\gamma) = \frac{P_{\mathrm{t}}(\gamma)}{K_\ell}$ is defined as the ratio of the transmit power to the path loss and noise power and $P_{\mathrm{c_r}} = \frac{P_{\mathrm{c}}}{K_\ell}$ represents the circuit to noise and path loss power ratio.[6]

The objective function in (10a), i.e., the EC function, is proved to be concave in the transmission power in [21]. On the other hand, since the denominator of the effective-EE constraint (10b) is affine, (10b) is a quasiconcave function in $P_{\mathrm{r}}(\gamma)$. From the definition of quasiconcavity, if the function $F:\mathbb{R}\rightarrow\mathbb{R}$, with $\mathbb{R}$ indicating the set of real numbers, is quasiconcave, then its upper contour sets are convex. In other words, F is quasiconcave iff S is convex where $S=\{x\in\mathbb{R}: F(x)\geq a\}, \forall\, a\in\mathbb{R}$. Since effective-EE is a strictly quasiconcave function in $P_{\mathrm{r}}(\gamma)$, the feasible set defined by (10b) is a convex set. Hence (10a) is a concave optimization problem and the Karush-Kuhn-Tucker (KKT) conditions are both sufficient and necessary for the optimal solution. Consider $\lambda\in\mathbb{R}$ is the Lagrange multiplier associated to (10b), then the Lagrangian is

$$\begin{aligned}\mathcal{L}\left(P_{\mathrm{r}}(\gamma),\lambda\right) = &-\frac{1}{\theta T_{\mathrm{f}}}\ln\left(\mathbb{E}_\gamma\left[(1+P_{\mathrm{r}}(\gamma)\gamma)^{-\xi(\theta)}\right]\right)\\ &+\lambda\left(-\frac{1}{\theta T_{\mathrm{f}}}\ln\left(\mathbb{E}_\gamma\left[(1+P_{\mathrm{r}}(\gamma)\gamma)^{-\xi(\theta)}\right]\right)\right.\\ &\left.-\eta_{\min}(\theta)\left(K_\ell\left(P_{\mathrm{c_r}}+\frac{1}{\epsilon}\mathbb{E}_\gamma[P_{\mathrm{r}}(\gamma)]\right)\right)\right).\end{aligned} \tag{11}$$

[6]Note that (10a) is equivalent to (6a), when re-formulated in terms of $P_{\mathrm{r}}(\gamma)$.

At the optimal power allocation $P^*_{\mathrm{r}}(\gamma)$, we have

$$\left.\frac{\partial\mathcal{L}(P_{\mathrm{r}}(\gamma),\lambda)}{\partial P_{\mathrm{r}}(\gamma)}\right|_{P_{\mathrm{r}}(\gamma)=P^*_{\mathrm{r}}(\gamma)} = 0, \Longrightarrow \tag{12a}$$

$$\frac{\xi(\theta)\gamma\,(1+P^*_{\mathrm{r}}(\gamma)\gamma)^{-\xi(\theta)-1}}{\theta T_{\mathrm{f}}\mathbb{E}_\gamma\left[(1+P^*_{\mathrm{r}}(\gamma)\gamma)^{-\xi(\theta)}\right]} + \lambda\frac{\xi(\theta)\gamma\,(1+P^*_{\mathrm{r}}(\gamma)\gamma)^{-\xi(\theta)-1}}{\theta T_{\mathrm{f}}\mathbb{E}_\gamma\left[(1+P^*_{\mathrm{r}}(\gamma)\gamma)^{-\xi(\theta)}\right]} - \lambda\frac{\eta_{\min}(\theta)K_\ell}{\epsilon} = 0 \Longrightarrow \tag{12b}$$

$$\xi(\theta)\gamma\,(1+P^*_{\mathrm{r}}(\gamma)\gamma)^{-\xi(\theta)-1} = \beta, \tag{12c}$$

where $\beta = \frac{\lambda\eta_{\min}(\theta)K_\ell\theta T_{\mathrm{f}}\mathbb{E}_\gamma[(1+P^*_{\mathrm{r}}(\gamma)\gamma)^{-\xi(\theta)}]}{\epsilon(\lambda+1)}$ is referred to as the scaled-Lagrangian-multiplier. From (12c), the optimal power distribution scheme can be found as

$$P^*_{\mathrm{r}}(\gamma) = \left[\frac{\xi(\theta)^{\frac{1}{1+\xi(\theta)}}}{\beta^{\frac{1}{1+\xi(\theta)}}\gamma^{\frac{\xi(\theta)}{1+\xi(\theta)}}} - \frac{1}{\gamma}\right]^+, \tag{13}$$

where $[x]^+ = \max\{0,x\}$. The solution to (10a) is hence given in (13) and the value for β can be found using the following Lemma:

Lemma 1: The scaled-Lagrangian-multiplier (β) is the solution of the effective-EE constraint to be satisfied at equality.

Proof: We prove the lemma by contradictory.

Let us define $\nu=\frac{1}{\beta}$. Now, we assume that at optimal ν^* maximum EC can be achieved while the effective-EE constraint at ν^* is satisfied with strict inequality. We start by referring to the fact that the effective-EE is a continuous function of ν. Therefore, for a small value of $\Delta\nu>0\in\mathbb{R}$, the effective-EE constraint still holds at $\nu^*+\Delta\nu$. On the other hand, we note that $\mathrm{EC}(\theta)$ is a monotonically increasing function of ν since at each γ the power $P_{\mathrm{r}}(\gamma)$ is a monotonically increasing function of ν. In other words, $\mathrm{EC}(\theta)|_{\nu=\nu^*+\Delta\nu} > \mathrm{EC}(\theta)|_{\nu=\nu^*}$. Henceforth, ν^* does not correspond to the maximum EC, which contradicts the assumption. Hence, we conclude the proof for the lemma. ■

Therefore, optimal value for β (referred to as β^*) is found such that the constraint (10b) is satisfied with equality. Now, by replacing $P^*_{\mathrm{r}}(\gamma)$ from (13) into (10b) and changing the inequality operation with equality, we get

$$\begin{aligned}&-\frac{1}{\theta T_{\mathrm{f}}}\ln\left(\mathbb{E}_\gamma\left[\left(1+\left[\frac{(\gamma\xi(\theta))^{\frac{1}{1+\xi(\theta)}}}{\beta^{*\frac{1}{1+\xi(\theta)}}}-1\right]^+\right)^{-\xi(\theta)}\right]\right)\\ &-\eta_{\min}(\theta)\left(K_\ell\left(P_{\mathrm{c_r}}+\frac{1}{\epsilon}\mathbb{E}_\gamma\left[\frac{\xi(\theta)^{\frac{1}{1+\xi(\theta)}}}{\beta^{*\frac{1}{1+\xi(\theta)}}\gamma^{\frac{\xi(\theta)}{1+\xi(\theta)}}}-\frac{1}{\gamma}\right]^+\right)\right)\\ &=0.\end{aligned} \tag{14}$$

The closed-form expressions for the expectation operations in (14) can be obtained by using the results of [30]. When the expectations are replaced with the closed-form expressions, the optimal value for β, i.e., β^*, can be solved from (14) using root-finding functions, e.g., fzero in Matlab. Once β^* is found, the optimum power allocation to solve (10a) is $P^*_{\mathrm{t}}(\gamma) = K_\ell\,\times$

$P_\mathrm{r}^*(\gamma)|_{\beta=\beta^*}$, where $P_\mathrm{r}^*(\gamma)$ is given in (13). Hence, concluding the solution for (10a).

We now define the optimum average input power level, P_t^*, which is related to $P_\mathrm{r}^*(\gamma)$ through

$$P_\mathrm{t}^* = K_\ell \times \mathbb{E}_\gamma \left[P_\mathrm{r}^*(\gamma)\right]\big|_{\beta=\beta^*}. \tag{15}$$

The *EC-gain-rate*(α_EE) can hence be found as

$$\alpha_\mathrm{EC} = \frac{\mathrm{EC}_\mathrm{opt}(\theta)}{\mathrm{EC}_{\eta_\mathrm{max}}(\theta)} \tag{16}$$

$$= \frac{\eta_\mathrm{min}(\theta)\left(P_\mathrm{c} + P_\mathrm{t}^*/\epsilon\right)}{\eta_\mathrm{max}(\theta)\left(P_\mathrm{c} + P_{\eta_\mathrm{max}}^*/\epsilon\right)}, \tag{17}$$

where $P_{\eta_\mathrm{max}}^*$ is the average input power at which the maximum effective-EE can be achieved. Using (8), the *EC-gain-rate* can further be simplified to

$$\alpha_\mathrm{EC} = \alpha_\mathrm{EE} \times \frac{P_\mathrm{c} + P_\mathrm{t}^*/\epsilon}{P_\mathrm{c} + P_{\eta_\mathrm{max}}^*/\epsilon}. \tag{18}$$

B. Optimal Power Allocation With Input Power Constraint

In this subsection, we aim to solve the optimization problem (6a)–(6c) using the results of Section III-A. We start by reformulating the optimization problem (6a)–(6c) in terms of $P_\mathrm{r}(\gamma)$, yielding

$$\mathrm{EC}_\mathrm{opt}(\theta) = \max_{P_\mathrm{r}(\gamma)\geq 0} -\frac{1}{\theta T_\mathrm{f}} \ln\left(\mathbb{E}_\gamma\left[(1+P_\mathrm{r}(\gamma)\gamma)^{-\xi(\theta)}\right]\right) \tag{19a}$$

$$\text{subject to:}\quad \frac{-\frac{1}{\theta T_\mathrm{f}} \ln\left(\mathbb{E}_\gamma\left[(1+P_\mathrm{r}(\gamma)\gamma)^{-\xi(\theta)}\right]\right)}{K_\ell\left(P_{\mathrm{c}_\mathrm{r}} + \frac{1}{\epsilon}\mathbb{E}_\gamma\left[P_\mathrm{r}(\gamma)\right]\right)} \geq \eta_\mathrm{min}(\theta) \tag{19b}$$

$$\mathbb{E}_\gamma\left[P_\mathrm{r}(\gamma)\right] \leq \frac{P_\mathrm{max}}{K_\ell}, \tag{19c}$$

Using results in (13) to (15), the power-constrained EC-maximization problem with joint effective-EE and input power constraints reduces to an EC-maximization problem with two input power constraints with the power limits set at P_t^* and P_max. Let us define the optimal operating power to solve (19a) as P_con^*. Then, P_con^* can be found as

$$P_\mathrm{con}^* = \begin{cases} P_\mathrm{t}^* & \text{when } P_\mathrm{max} \geq P_\mathrm{t}^* \\ P_\mathrm{max} & \text{when } P_\mathrm{max} < P_\mathrm{t}^* \\ & \text{and } \eta(\theta)|_{\mathbb{E}_\gamma[P_\mathrm{t}(\gamma)]=P_\mathrm{max}} \geq \eta_\mathrm{min} \\ \text{Not feasible} & \text{when } P_\mathrm{max} < P_\mathrm{t}^* \\ & \text{and } \eta(\theta)|_{\mathbb{E}_\gamma[P_\mathrm{t}(\gamma)]=P_\mathrm{max}} < \eta_\mathrm{min}. \end{cases} \tag{20}$$

In other words, when P_max is very small such that for any input power smaller than P_max, η_min cannot be achieved, the problem is not feasible. Otherwise, the optimum average transmission power is equal to $\min(P_\mathrm{t}^*, P_\mathrm{max})$. Therefore, the optimal power allocation to solve (6a)–(6c) is according to (13), wherein, optimal $\beta(\beta^*)$ is found such that $K_\ell \mathbb{E}_\gamma[P_\mathrm{r}(\gamma)]|_{\beta=\beta^*} = P_\mathrm{con}^*$.

To summarize, the Pseudocode of the optimal power allocation process to solve (6a)–(6c) is illustrated as

---

Initializations:

P_max: The input power limit.

α_EE: The ratio between the required effective-EE level and the maximum achievable effective-EE of the channel.

P_c: The circuit power.

ϵ: The power amplifier efficiency.

$K_\ell = P_\mathcal{L}\sigma_\mathrm{n}^2$: The pathloss and noise factor.

T_f: The fading block duration.

m: The Nakagami fading parameter.

Step 1:

Find the maximum achievable effective-EE of the channel: η_max.

Calculate $\eta_\mathrm{min} = \alpha_\mathrm{EE}\eta_\mathrm{max}$.

Create (14), using closed-form expressions given in [27].

Using root-finding functions, e.g., fzero in Matlab, find β^* that solves (14).

Calculate $P_\mathrm{t}^* = K_\ell \mathbb{E}_\gamma[P_\mathrm{r}(\gamma)]|_{\beta=\beta^*}$, where $P_\mathrm{r}(\gamma)$ is given in (13).

Step 2:

If $P_\mathrm{max} \leq P_\mathrm{t}^*$ *and the achievable effective-EE at* P_max *is less than* η_min:

Problem is not feasible. Terminate calculations.

If $P_\mathrm{max} \leq P_\mathrm{t}^*$, *and the achievable effective-EE at* P_max *is larger than or equal to* η_min:

Create $\mathbb{E}_\gamma[P_\mathrm{r}(\gamma)] = \frac{P_\mathrm{max}}{K_\ell}$, where $P_\mathrm{r}(\gamma)$ is given in (13).

Update β^* to solve the above equation.

Step 3: Calculate outputs

Calculate $\mathrm{EC}(\theta)$ given in (2) by using power allocation provided in (13) and calculated β^*.

Calculate effective-EE, using (5).

C. Effect of P_c on the Effective-EE and EC Tradeoff

In this subsection, we aim to investigate the effect of the circuit power on the tradeoff between the effective-EE and EC. Specifically, we investigate the maximum achievable EC of two systems with different circuit power values, when the system requirements in terms of the target delay-outage probability and effective-EE limit are un-changed. In doing so, we study two different cases.

In **Case 1**, we assume that the required effective-EE level is fixed and does not depend on the maximum achievable effective-EE of the channel, i.e., η_max. In this case, using the fact that effective-EE is a strictly quasi-concave function of the transmission power, and a monotonically decreasing function of the circuit power, a certain effective-EE level can be achieved at a higher transmission power in a system with a lower circuit power value. Henceforth, a higher EC can be achieved.

In **Case 2**, the required effective-EE is assumed to be a certain percentage of the maximum achievable effective-EE of the channel. Assume $P_{\mathrm{c},1}$ and $P_{\mathrm{c},2}$ represent the circuit power of two systems under consideration with $P_{\mathrm{c},1} \leq P_{\mathrm{c},2}$. We note that $\eta_{\mathrm{max},1} \geq \eta_{\mathrm{max},2}$, with $\eta_{\mathrm{max},1}$ and $\eta_{\mathrm{max},2}$ representing the maximum achievable effective-EE of the systems with $P_{\mathrm{c},1}$ and $P_{\mathrm{c},2}$, respectively. The effective-EE of the two systems at an input power $\mathbb{E}_{\gamma}[P_{\mathrm{t}}(\gamma)] = P_1$ can be now formulated as

$$\eta_1 = \frac{\mathrm{EC}_1}{P_{\mathrm{c},1} + \frac{P_1}{\epsilon}}, \tag{21}$$

$$\eta_2 = \frac{\mathrm{EC}_1}{P_{\mathrm{c},2} + \frac{P_1}{\epsilon}}, \tag{22}$$

where $\mathrm{EC}_1 = \mathrm{EC}|_{\mathbb{E}_{\gamma}[P_{\mathrm{t}}(\gamma)]=P_1}$, and η_1 and η_2 are the achievable effective-EE of systems with $P_{\mathrm{c},1}$ and $P_{\mathrm{c},2}$ at $\mathbb{E}_{\gamma}[P_{\mathrm{t}}(\gamma)] = P_1$, respectively. We now define $\alpha_{\mathrm{EE},1} = \frac{\eta_1}{\eta_{\max,1}}$ and $\alpha_{\mathrm{EE},2} = \frac{\eta_2}{\eta_{\max,2}}$. We further define

$$\mathcal{R} = \frac{\alpha_{\mathrm{EE},1}}{\alpha_{\mathrm{EE},2}} \tag{23}$$

$$= \frac{\eta_{\max,2}\left(P_{\mathrm{c},2} + \frac{P_1}{\epsilon}\right)}{\eta_{\max,1}\left(P_{\mathrm{c},1} + \frac{P_1}{\epsilon}\right)} \tag{24}$$

Now, by taking the derivative of the right-hand-side (RHS) of (24) with respect to the transmission power P_1, we note that the RHS of (24) is a monotonically decreasing function of P_1. Henceforth, $\mathcal{R}$ is a monotonically decreasing function of P_1. Now, we refer to the fact that the maximum effective-EE is achieved at a lower input transmit power in a system with a lower circuit power [27]. In other words, if P_2^* is the input power at which $\eta_{\max,2}$ can be achieved, then $\eta_{\max,1}$ is achieved at a power lower than P_2^*. Therefore, at $\mathbb{E}_{\gamma}[P_{\mathrm{t}}(\gamma)] = P_2^*$, $\alpha_{\mathrm{EE},1} < 1$ and $\alpha_{\mathrm{EE},2} = 1$, hence at this point, $\mathcal{R} \leq 1$. We can now conclude that

$$\alpha_{\mathrm{EE},1} \leq \alpha_{\mathrm{EE},2}, \qquad \forall \quad \mathbb{E}_{\gamma}\left[P_{\mathrm{t}}(\gamma)\right] \geq P_2^*. \tag{25}$$

Therefore, when the required effective-EE level is set as a certain ratio of the maximum achievable effective-EE of the channel, a system with lower circuit power achieves the effective-EE limit at a lower transmission power. Henceforth, a lower EC can be achieved at this case.

D. Effective-EE and EC Tradeoff When $\theta \to \infty$ (Representing Extremely Stringent Delay Requirement)

Here, we consider a system with extremely stringent delay requirement. In this case, EC equals the zero-outage capacity, and the optimum transmission strategy is to maintain a constant received-SNR, δ, or a constant rate [30]. Therefore, the optimum power allocation is the channel inversion with fixed rate (*cifr*) transmission technique [30]

$$P_{\mathrm{r}}(\gamma) = \frac{\delta}{\gamma}.$$

Therefore, the EC-maximization problem considered in this paper simplifies to

$$\mathrm{EC}_{\mathrm{opt}}(\theta)|_{\theta\to\infty} = \max_{\delta\geq 0} \ln(1+\delta) \tag{26}$$

$$\text{subject to:} \quad \eta_{\mathrm{cifr}}(\delta)$$

$$= \left.\frac{\ln(1+\delta)}{K_{\ell}\left(P_{\mathrm{c_r}} + \frac{1}{\epsilon}\delta\mathbb{E}_{\gamma}\left[\frac{1}{\gamma}\right]\right)} \geq \eta_{\min}(\theta)\right|_{\theta\to\infty}. \tag{27}$$

Hereafter, for the ease of notation, we use $\eta_{\min}(\infty) = \eta_{\min}(\theta)|_{\theta\to\infty}$.

For Nakagami-m channels,

$$\mathbb{E}_{\gamma}\left[\frac{1}{\gamma}\right] = \begin{cases} \frac{m}{m-1}, & m > 1, \\ \infty, & m \leq 1, \end{cases}$$

which means that the achievable effective-EE with *cifr* transmission policy in Nakagami fading channels with $m \leq 1$ is zero. Note that when $m > 1$, $\mathbb{E}_{\gamma}[\frac{1}{\gamma}]$ decreases with m.

Since at the optimal point for (26), the constraint (27) is satisfied with equality, we simplify the constraint according to

$$\frac{m}{m-1} = \frac{\ln(1+\delta) - K_{\ell}\eta_{\min}(\infty)P_{\mathrm{c_r}}}{\frac{K_{\ell}}{\epsilon}\eta_{\min}(\infty)\delta} = F(\delta). \tag{28}$$

Now, the first derivative of $F(\delta)$ with respect to δ can be obtained as

$$F'(\delta) = \frac{\partial F(\delta)}{\partial \delta} = \frac{\frac{\delta}{1+\delta} - \ln(1+\delta)}{\frac{K_{\ell}}{\epsilon}\eta_{\min}(\infty)\delta^2} + \frac{\epsilon P_{\mathrm{c_r}}}{\delta^2}. \tag{29}$$

We note that nominator of $F'(\delta)$ monotonically deccreases with δ, whereas its denominator is a non-negative monotonically increasing function of δ. On the other hand, we note that $F'(\delta)|_{\delta=0} \geq 0$ and $F'(\delta)|_{\delta\to\infty} \leq 0$. Therefore, $F'(\delta)$ changes sign only once and hence, $F(\delta)$ is bell-shape in δ, meaning that it increases monotonically until it reaches its maximum, and then it is a monotonically decreasing function of δ.

Now, let us assume that at a certain Nakagami-m parameter, the condition in (28) is feasible. Therefore, two possible solutions exist. Since the objective function in (26) is to maximize the achievable rate, and the *cifr* capacity is monotonically increasing with δ, hence, the acceptable solution for δ is the bigger one, which is after $F(\delta)$ peak. Therefore, $F(\delta)$ is a decreasing function at this point to infinity. Recall that the left-hand-side (LHS) of (28) is a decreasing function of m. Therefore, as m increases, $F(\delta)$ decreases, and therefore, δ increases and as such, the achievable capacity increases. Or in brief, as Nakagami-m fading parameter increases, the achievable EC increases.

IV. Effective-EE-EC Loss-Gain Matching Power Allocation

The aim of this subsection is to provide a guideline on how to choose the parameter α_{EE}, bearing in mind its significant effect on the achievable effective-EE and EC. Recall that, the effective-EE curve is a monotonically increasing function of EC (presented in Fig. 1) until it reaches its maximum point. After this point, the effective-EE monotonically decreases with the increase of the power and with the increase of the EC. The slope of effective-EE curve close to the peak effective-EE value is small. This translates into achieving a big gain in EC, while losing effective-EE slightly. As the operational point departs from the maximum effective-EE point, the slope of the effective-EE curve becomes sharp and henceforth, the gain that can be achieved in EC, at the cost of losing effective-EE, will reduce. These observations cannot be directly studied from the exact values of the EC and effective-EE themselves, hence, we study the EC-gain slope versus the effective-EE-loss slope.

Here, we propose to choose the operational power where the EC-gain-slope matches a scaled effective-EE-loss-slope. In a system design level, the loss in the effective-EE can be considered as the extra cost of energy that the system is willing to pay for increasing its achievable EC. This scaling factor shows the ratio of the increasing-rate of the cost that the system has to pay versus the increasing-rate of its achievable EC.

Now, we formulate the optimization problem according to the effective-EE-EC loss-gain matching method explained above. Considering a target delay-outage probability, we aim to propose a power allocation strategy under which the gain-rate in the achievable EC and the loss-rate in the achievable effective-EE as a function of the input transmit power matches by a scaling factor defined by $\mathcal{K}$. Mathematically, the optimization problem can be expressed as

$$\mathrm{EC_{opt}}(\theta) = \max_{P_{\mathrm{r}}(\gamma)\geq 0} -\frac{1}{\theta T_{\mathrm{f}}} \ln\left(\mathbb{E}_\gamma\left[(1+P_{\mathrm{r}}(\gamma)\gamma)^{-\xi(\theta)}\right]\right) \tag{30a}$$

$$\text{subject to:}\quad \mathcal{K}\frac{\partial\eta(\theta)}{\partial\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right]} + \frac{\partial\mathrm{EC}(\theta)}{\partial\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right]} = 0. \tag{30b}$$

We start by studying the constraint (30b) which can be expanded as

$$\frac{\frac{\partial\mathrm{EC}(\theta)}{\partial\mathbb{E}_\gamma[P_{\mathrm{r}}(\gamma)]}}{\mathrm{EC}(\theta)} = \frac{1/\epsilon}{P_{\mathrm{c_r}} + \frac{1}{\epsilon}\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right]} - \frac{1/\epsilon}{P_{\mathrm{c_r}} + \frac{1}{\epsilon}\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right] + \frac{\mathcal{K}}{K_\ell}}, \tag{31}$$

where it can further be simplified to

$$\frac{\partial\ln\left(\mathrm{EC}(\theta)\right)}{\partial\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right]} = \frac{\partial\ln\left(\frac{P_{\mathrm{c_r}}+\frac{1}{\epsilon}\mathbb{E}_\gamma[P_{\mathrm{r}}(\gamma)]}{P_{\mathrm{c_r}}+\frac{1}{\epsilon}\mathbb{E}_\gamma[P_{\mathrm{r}}(\gamma)]+\frac{\mathcal{K}}{K_\ell}}\right)}{\partial\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right]}. \tag{32}$$

Assuming $\frac{P^*}{K_\ell}$ is the solution to (32), the optimization problem in (30a) simplifies to an EC-maximization problem with an input power constraint at level P^*. Henceforth, the optimal power allocation strategy follows (13). Since the $\mathbb{E}_\gamma[P_{\mathrm{r}}(\gamma)]$ is a monotonically increasing function of β, the constraint (32) can be re-written as

$$\frac{\partial\ln\left(\mathrm{EC}(\theta)\right)}{\partial\beta} = \frac{\partial\ln\left(\frac{P_{\mathrm{c_r}}+\frac{1}{\epsilon}\mathbb{E}_\gamma[P_{\mathrm{r}}(\gamma)]}{P_{\mathrm{c_r}}+\frac{1}{\epsilon}\mathbb{E}_\gamma[P_{\mathrm{r}}(\gamma)]+\frac{\mathcal{K}}{K_\ell}}\right)}{\partial\beta}. \tag{33}$$

The optimal solution for β^* can now be found numerically from (33).

Now, to obtain the *EE-loss-rate* α_{EE} that is associated to a particular $\mathcal{K}$, we insert the equality $\mathrm{EC}(\theta) = \eta(\theta)K_\ell(P_{\mathrm{c_r}} + \frac{1}{\epsilon}\mathbb{E}_\gamma[P_{\mathrm{r}}(\gamma)])$ into (30b) yielding

$$\frac{\partial\eta(\theta)}{\partial\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right]}K_\ell\left(P_{\mathrm{c_r}} + \frac{1}{\epsilon}\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right]\right) + K_\ell\frac{\eta(\theta)}{\epsilon} = -\mathcal{K}\frac{\partial\eta(\theta)}{\partial\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right]}, \tag{34}$$

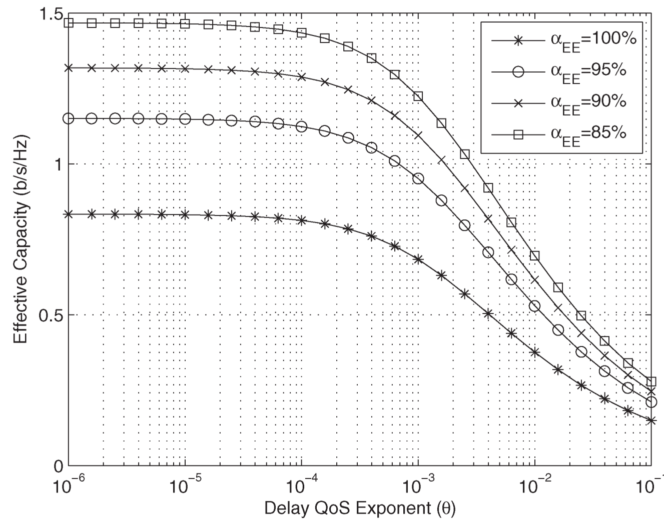

Fig. 3. EC versus delay QoS exponent θ for various α_{EE} in Rayleigh fading channels.

which further simplifies to

$$\frac{\partial\ln\left(\eta(\theta)\right)}{\partial\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right]} = \frac{-1}{\epsilon\left(P_{\mathrm{c_r}} + \frac{1}{\epsilon}\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right] + \frac{\mathcal{K}}{K_\ell}\right)}. \tag{35}$$

Using (8), the *EE-loss-rate* α_{EE} with respect to $\mathcal{K}$ can be found as

$$\frac{\partial\ln\left(\alpha_{\mathrm{EE}}\right)}{\partial\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right]} = \frac{-1}{\epsilon\left(P_{\mathrm{c_r}} + \frac{1}{\epsilon}\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right] + \frac{\mathcal{K}}{K_\ell}\right)}. \tag{36}$$

In addition, direct relation in between α_{EE} and α_{EC} with consideration of effective-EE-EC loss-gain matching technique can be found from

$$\mathcal{K}\left(P_{\mathrm{c}} + \frac{K_\ell}{\epsilon}P^*_{\max}\right)\frac{\partial\alpha_{\mathrm{EE}}}{\partial\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right]} + \frac{\partial\alpha_{\mathrm{EC}}}{\partial\mathbb{E}_\gamma\left[P_{\mathrm{r}}(\gamma)\right]} = 0. \tag{37}$$

V. Numerical Results

In this section, we numerically evaluate the maximum achievable EC (in b/s/Hz) of fading channels under delay-outage probability and effective-EE constraints. We further investigate the effects of the circuit power, the fading-block duration, and the fading severeness on the maximum achievable EC and the effective-EE and EC trade-off of the communications link. In the following figures, the fading-block duration $T_{\mathrm{f}} = 500$, the circuit-to-noise power ratio $P_{\mathrm{c_r}} = 0$ dB, and the Nakagami fading parameter $m = 1$ (i.e., Rayleigh fading), unless otherwise indicated.

We start by plotting the normalized EC in b/s/Hz versus the delay QoS exponent θ for various values of α_{EE} in a Rayleigh fading channel in Fig. 3. For example, the figure shows that by reducing the α_{EE} from 100% to 95% (with only 5% loss in effective-EE), considerable gain in the EC can be achieved at any value of θ. Specifically, with $\theta \leq 10^{-4}$, the EC increases from around 0.8 b/s/Hz to around 1.2 b/s/Hz, hence 50% increase in the achievable EC is observed compared to 5%

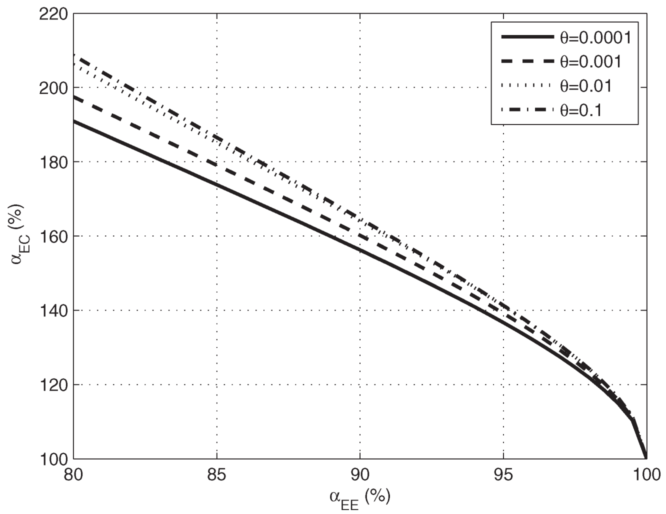

Fig. 4. α_{EC} in percentage versus α_{EE} in percentage for various delay exponent values θ in Rayleigh fading channels.

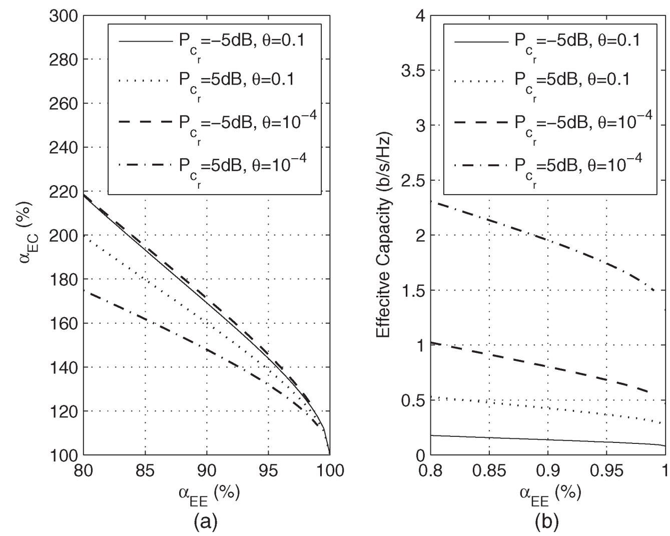

Fig. 5. α_{EC} (in percentage) and maximum achievable EC versus α_{EE} (in percentage) for various circuit power levels and delay exponent values in Rayleigh fading channels.

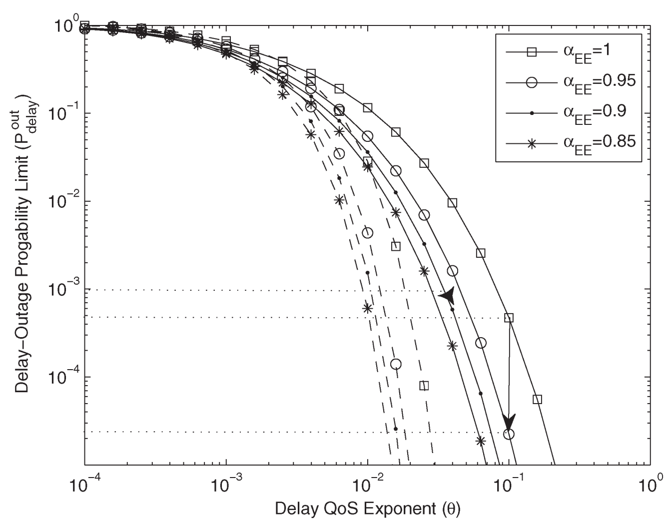

Fig. 6. Achievable delay-outage probability versus θ for various α_{EE} in Rayleigh (solid lines) and Rician (dashed lines) fading channels.

loss in effective-EE. Moreover, it is shown that when α_{EE} is further reduced, the curves get closer to each other and the gain in EC in between the two consecutive cases of different α_{EE} reduces.

To observe in a clear picture the variations in α_{EC} as a function of α_{EE}, Fig. 4 includes the plots for the α_{EC} in percentage versus the α_{EE} in percentage for various delay QoS exponent values. The figure shows that α_{EC} monotonically increases with the decrease of α_{EE}. We can also see that a tiny reduction in effective-EE around its maximum (i.e., when α_{EE} is very close to 100%) generates a significant gain in EC (i.e., α_{EC} increasing rapidly). When α_{EE} is further reduced (say less than 98%), the increasing speed of EC gain over effective-EE loss (α_{EC} vs. reduction of α_{EE}) becomes slower. Furthermore, higher EC gain is observed in systems with tighter delay constraints (i.e., with larger value of θ) than those with loose delay constraints. These observations match the findings obtained from Fig. 3.

In Fig. 5, two systems with different circuit power values are considered and the plots for α_{EC} (Fig. 5(a)) and the maximum achievable EC (Fig. 5(b)) with various delay requirements are provided. The figures show that in comparison between two systems with different $P_{\mathrm{c_r}}$s, a system with a higher circuit power achieves less gain in α_{EC}. On the other hand, higher EC can be achieved in a system with a larger P_{c}. The latter finding confirms the derivations in Section III-C, Case 2. The gap in between the two curves for the α_{EC} with different circuit power is more pronounced when the delay requirement becomes tighter, i.e., higher θ.

Fig. 6 includes the plots for the delay-outage probability limit, $P_{\mathrm{delay}}^{\mathrm{out}}$, versus the delay exponent, θ, for various α_{EE} with the maximum tolerable delay threshold $D_{\max} = 500$. Two different channel fading types, i.e., Rayleigh (solid lines) and Rician (dashed lines), are considered in this figure which shows that the delay-outage probability is smaller in Rician fading channels compared to Rayleigh fading channels. For a target $P_{\mathrm{delay}}^{\mathrm{out}}$, the corresponding value for θ found from Fig. 6 for Rayleigh fading channel curves can be used in Fig. 3 to obtain the EC of the system. For example, for the target delay-outage probability of 0.1% and $\alpha_{\mathrm{EE}} = 90\%$, the delay exponent is $\theta = 0.04$ (see Fig. 6), and the corresponding achieved EC is 0.35 b/s/Hz (see Fig. 3). The figure further reveals that reducing α_{EE} reduces the achieved delay-outage probability significantly. For example, in Rayleigh fading channel with $\theta = 0.1$, decreasing α_{EE} from 100% to 95%, reduces $P_{\mathrm{delay}}^{\mathrm{out}}$ by more than 10 times.

We further plot the achieved delay-outage probability versus α_{EE} for various values of θ in Rayleigh fading channels with maximum tolerable delay threshold $D_{\max} = 500$ in Fig. 7. The figure shows that for loose delay-constrained systems (e.g., $\theta = 10^{-4}$) reducing α_{EE} does not affect the achieved delay-outage probability. On the other hand, when delay constraint becomes tighter (e.g., $\theta = 0.1$), delay-outage probability increases sharply with α_{EE}. This indicates the pronounced tradeoff between the achieved effective-EE and delay-constrained system performance.

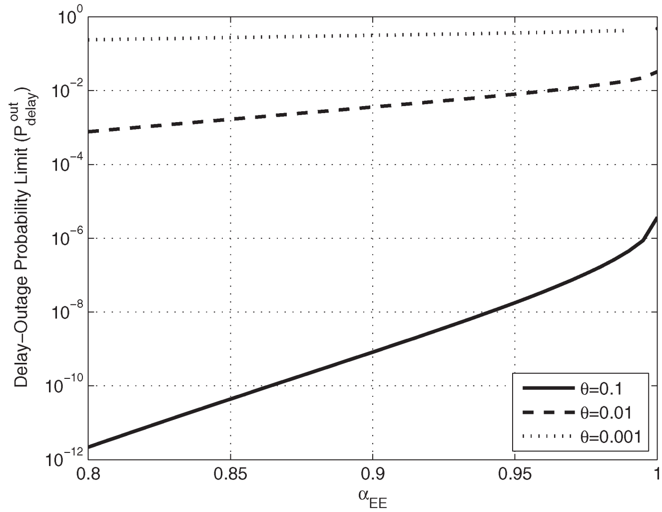

Fig. 7. Achievable delay-outage probability versus α_{EE} for various θ values with fading-block duration, $T_{\mathrm{f}} = 500$, $D_{\max} = 500$.

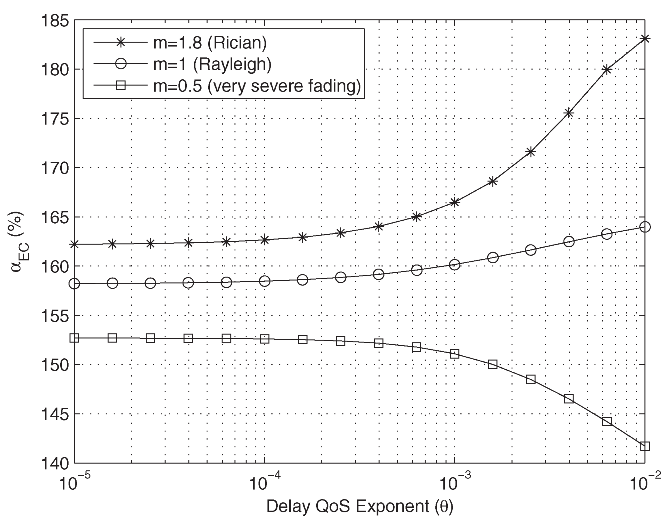

Fig. 8. α_{EC} in percentage versus delay QoS exponent θ for various Nakagami-m fading parameters m.

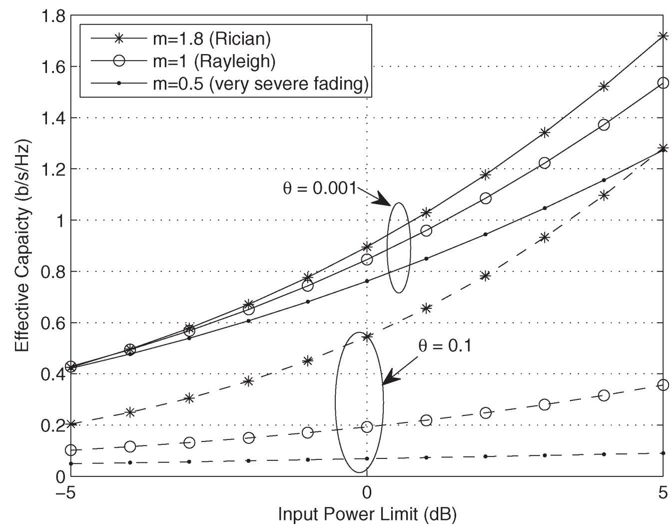

Fig. 9. EC versus input power limit in dB for various delay exponent θ and Nakagami-m fading parameter m.

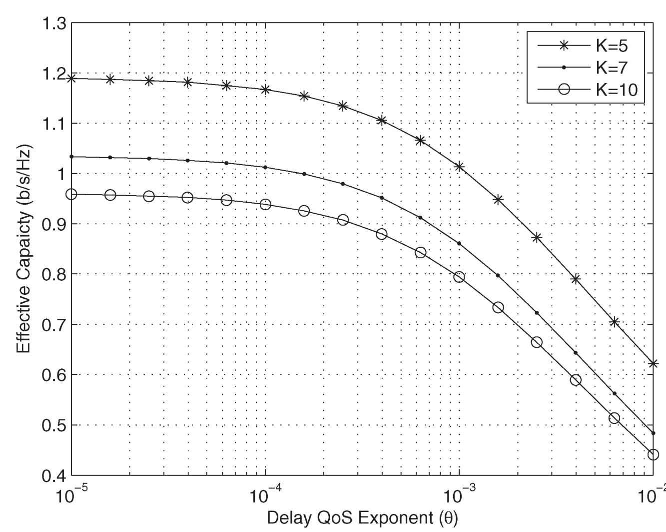

Fig. 10. EC versus delay QoS exponent θ for various $\mathcal{K}$ in Rayleigh fading channels.

We study the effects of different types of mobile fading channels on the *EC-gain* of delay-outage constrained systems by presenting plots for α_{EC} versus the delay QoS exponent for various values of Nakagami-m fading parameter in Fig. 8. The value for α_{EE} is set to 90% for the plots included in this figure. The figure shows that for Rayleigh $(m = 1)$ and Rician $(m = 1.8)$ fading channels the *EC-gain* increases when θ increases, or equivalently, when delay constraint becomes tighter. The behavior of the curve with $m = 0.5$ (i.e. very severe fading channel) is, however, the opposite. To understand the results behind this behavior, we plot the EC of these three different fading channels versus the input power for two different values of θ in Fig. 9. The figure shows that in Rician or Rayleigh fading channels, the EC increases with almost the same, or similar slope for $\theta = 0.001$ and $\theta = 0.1$ cases. In the Nakagami-m fading channel with $m = 0.5$, however, the EC increases with a much slower slope when $\theta = 0.1$ compared to the case when $\theta = 0.001$. In other words, in this case $(m = 0.5)$, when delay constraint is tight, the EC increases slowly with input power, hence increasing α_{EE}, or equivalently, increasing the operating transmit power, results in a small gain in EC and hence samll α_{EC}. Hence, the choice for α_{EE}, not only depends on the delay requirement, but also on how quickly the EC increases with the transmit power.

In Fig. 10–Fig. 12, we study the EC, the effective-EE, and the *EE-loss-rate* of a Rayleigh fading channel when the choice of α_{EE} is chosen based on the slopes of the EC and effective-EE curves versus power. In particular, the plots for the achievable EC and effective-EE versus θ when the slopes match at different factors, i.e., $\mathcal{K}$ in (30b), are considered in Fig. 10 and Fig. 11, respectively. The figures show that increasing $\mathcal{K}$ reduces the achievable EC, while increasing the achievable effective-EE. Note that in this figure, the operational point is found such that (according to (30b)) the gain-rate in the EC (i.e., the slope of the EC versus power) is $\mathcal{K}$ times larger than the loss-rate in the effective-EE. Further recall that the slope of the effective-EE curve is small around its maximum value $(\eta_{\max})$ and it becomes

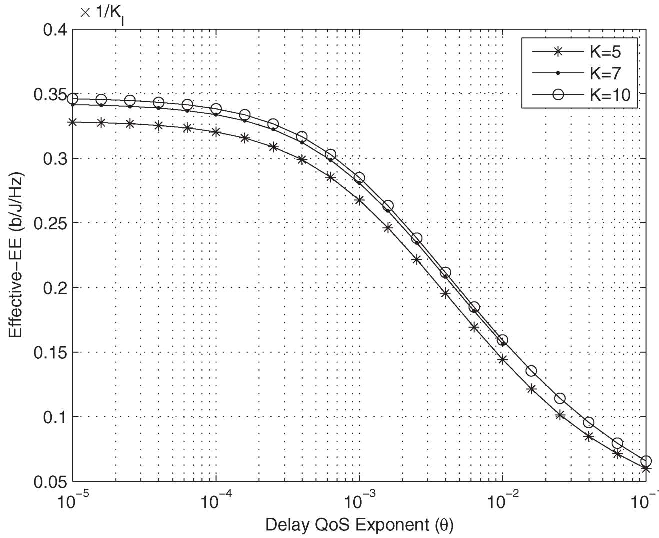

Fig. 11. Achievable effective-EE versus delay QoS exponent θ for various $\mathcal{K}$ in Rayleigh fading channels.

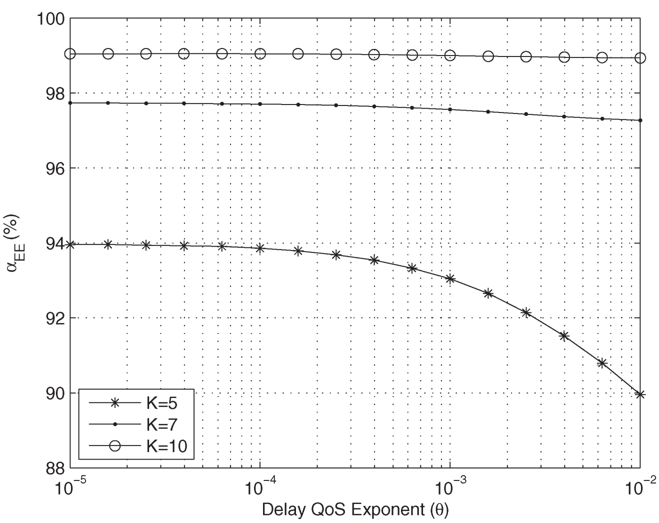

Fig. 12. *Effective-EE-loss-rate*α_{EE} versus delay QoS exponent θ for various $\mathcal{K}$ in Rayleigh fading channels.

larger as power increases. Henceforth, decreasing $\mathcal{K}$ will push the operational point away from the peak effective-EE point, and as such achieving higher EC. On the other hand, the values for the corresponding α_{EE} to different $\mathcal{K}$ values matching the effective-EE-EC loss-gain can be read from Fig. 12. The figure shows that with $\mathcal{K} = 10$, α_{EE} is close to 100%, meaning that the operational effective-EE constraint limit is close to the maximum achievable effective-EE. On the other hand, when $\mathcal{K}$ reduces to 5, α_{EE} decreases significantly.

Finally, in Fig. 13, we consider a system with extremely stringent delay requirement $(\theta \to \infty)$ and plot the maximum achievable EC versus the Nakagami fading parameter m for various values of circuit power and effective-EE constraint limit, $\eta_{\min}$. The figure shows that EC increases with m which confirms the derivations of Section III-D. Further, the figure shows that when $\eta_{\min}$ is chosen as a fixed value, EC decreases with the circuit power, which confirms the results provided in Section III-C, Case 1.

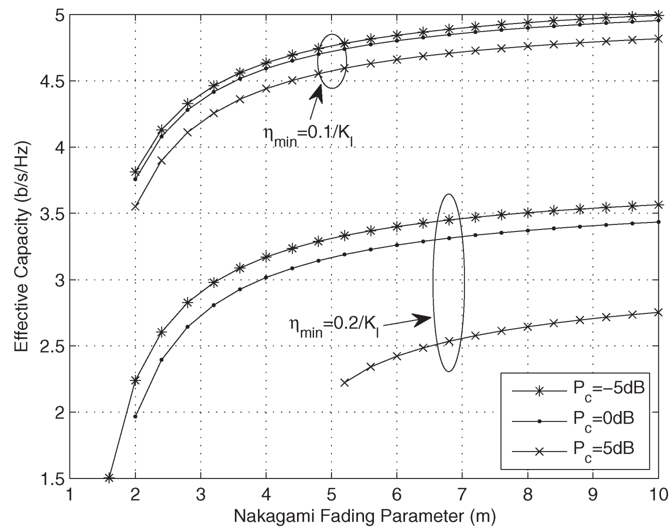

Fig. 13. Achievable EC versus Nakagami fading parameter m for various circuit power values (P_{c}) and effective-EE constraint levels $(\eta_{\min})$ when θ approaches infinity.

VI. Conclusion

We proposed an effective-EE-constrained rate-optimal power allocation technique for Nakagami-m flat-fading channels under delay-outage probability constraints when taking into account non-zero circuit power consumption during signal transmission. At any target delay-outage probability, we first obtained the maximum achievable effective-EE of the channel. The required effective-EE limit of the channel is set at a certain percentage of the maximum achievable effective-EE. We showed that the maximum EC can be achieved at the average input power level (P_{t}^*) at which the effective-EE constraint is satisfied at equality. When an input power constraint is present, the power should be distributed optimally over time based on the minimum of P_{t}^* and the maximum transmit power limit. We then prove that when the required effective-EE level is set as a certain ratio of $\eta_{\max}$, a system with lower P_{c} achieves less EC. We further investigate the effect of Nakagmai-m fading parameter on the achievable EC of a system with extremely stringent delay requirement. To set a guideline on how to choose the value of the required effective-EE, we proposed an effective-EE-loss EC-gain rate matching approach. In detail, we obtained the operating power at which the gain-rate in the achievable EC is equal to a scaled loss-rate of the effective-EE. The numerical results showed that in stringent delay-limited systems, the effective-EE and EC tradeoff is more pronounced compared to loose delay-constrained systems in the sense that reducing the required effective-EE threshold, increases EC significantly.

References

[1] "5G radio access, research and vision," Kista, Sweden, white paper, Jun. 2013. [Online]. Available: http://www.climateactionprogramme.orgimages/uploads/documents/Sustainability_and_energy_efficiency_-_5G_radio_access.pdf

[2] L. Rodríguez-Roselló, "R&D to shape the networks and services of the future," Eur. Commission, Brussels, Belgium, Jun. 2013, [Online]. Available: http://icc2013.ieee-icc.org/P8_2_H2020_ICC2013_Rosello_EC.pdf

[3] G. Li *et al.*, "Energy-efficient wireless communications: Tutorial, survey, open issues," *IEEE Wireless Commun. Mag.*, vol. 18, no. 6, pp. 28–35, Dec. 2011.
[4] S. Cui, A. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sep. 2005.
[5] G. Miao, N. Himayat, Y. Li, and D. Bormann, "Energy efficient design in wireless OFDMA," in *Proc. IEEE ICC*, May 2008, pp. 3307–3312.
[6] G. Miao, N. Himayat, G. Li, and S. Talwar, "Distributed interference-aware energy-efficient power optimization," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1323–1333, Apr. 2011.
[7] R. Prabhu and B. Daneshrad, "An energy-efficient water-filling algorithm for OFDM systems," in *Proc. IEEE ICC*, Cape Town, South Africa, May 2010, pp. 1–5.
[8] C. Isheden and G. P. Fettweis, "Energy-efficient link adaptation on a Rayleigh fading channel with receiver CSI," in *Proc. IEEE ICC*, Kyoto, Japan, Jun. 2011, pp. 1–5.
[9] L. Musavian and T. Le-Ngoc, "Energy-efficient power allocation for delay-constrained systems," in *Proc. GLOBECOM*, 2012, pp. 3554–3559.
[10] W. Cheng, X. Zhang, and H. Zhang, "Joint spectrum and power efficiencies optimization for statistical QoS provisionings over SISO/MIMO wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 903–915, May 2013.
[11] D. Wu and R. Negi, "Effective capacity: A wireless link model for support of quality of service," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 630–643, Jul. 2003.
[12] C.-S. Chang, "Stability, queue length, delay of deterministic and stochastic queueing networks," *IEEE Trans. Autom. Control*, vol. 39, no. 5, pp. 913–931, May 1994.
[13] J. Tang and X. Zhang, "Quality-of-service driven power and rate adaptation over wireless links," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 3058–3068, Aug. 2007.
[14] M. Zafer and E. Modiano, "Delay-constrained energy efficient data transmission over a wireless fading channel," in *Inf. Theory Appl. Workshop*, 2007, pp. 289–298.
[15] J. Lee and N. Jindal, "Energy-efficient scheduling of delay constrained traffic over fading channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1866–1875, Apr. 2009.
[16] F. Meshkati, H. Poor, and S. Schwartz, "Energy-efficient resource allocation in wireless networks," *IEEE Signal Process. Mag.*, vol. 24, no. 3, pp. 58–68, May 2007.
[17] M. Gursoy, D. Qiao, and S. Velipasalar, "Analysis of energy efficiency in fading channels under QoS constraints," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4252–4263, Aug. 2009.
[18] D. Qiao, M. Gursoy, and S. Velipasalar, "Energy efficiency in the low-SNR regime under queueing constraints and channel uncertainty," *IEEE Trans. Commun.*, vol. 59, no. 7, pp. 2006–2017, Jul. 2011.
[19] S. Verdu, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.
[20] A. Helmy, L. Musavian, and T. Le-Ngoc, "Energy-efficient power allocation for multicarrier systems with delay-outage probability constraints," in *Proc. IEEE ICC*, Budapest, Hungary, Jun. 2013, pp. 5872–5877.
[21] L. Liu, "Energy-efficient power allocation for delay-sensitive traffic over wireless systems," in *Proc. IEEE ICC*, Ottawa, Canada, Jun. 2012, pp. 5901–5905.
[22] B. Liu, L. Qiu, and J. Xu, "QoS aware energy efficiency analysis in the cellular networks," *IEICE Trans.*, vol. 94-B, no. 10, pp. 2925–2928, Oct. 2011.
[23] L. Musavian and Q. Ni, "Delay-QoS-driven spectrum and energy efficiency tradeoff," in *Proc. IEEE ICC*, 2014, pp. 4981–4986.
[24] M. K. Simon and M.-S. Alouini, *Digital Communication Over Fading Channels: A Unified Approach to Performance Analysis*. New York, NY, USA: Wiley, 2000.
[25] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*. New York, NY, USA: Dover, 1965.
[26] J. A. Bucklew, *Introduction to Rare Event Simulation*. New York, NY, USA: Springer-Verlag, 2004.
[27] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*., 2nd ed. New York, NY, USA: Springer-Verlag, 1998.
[28] B. Simon, *Convexity: An Analytic Viewpoint*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[29] L. Musavian and T. Le-Ngoc, "Energy-efficient power allocation over Nakagami-m fading channels under delay-outage constraints," *IEEE Trans. Wireless Commun.*, vol. 13, no. 8, pp. 4081–4091, Aug. 2014.
[30] A. J. Goldsmith and P. P. Varaiya, "Capacity of fading channels with channel side information," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1986–1992, Nov. 1997.

Leila Musavian (S'05–M'07) received the Ph.D. degree in telecommunications from King's College London, U.K., in 2006. She was a Post-doctoral Fellow at the National Institute of Scientific Research-Energy, Materials, and Telecommunications (INRS-EMT), University of Quebec, Canada, from 2006 to 2008. Between 2009 and 2010, she was with Loughborough University and investigated adaptive transmission techniques for delay quality-of-service (QoS) provisioning in cognitive radio networks (CRNs). In 2010–2012, she was a Research Associate in the Department of Electrical and Computer Engineering, McGill University, and was working on developing energy-efficient resource allocation techniques for multi-user wireless communications systems. She is currently working as a Lecturer in communications in the School of Computing and Communications, InfoLab21, Lancaster University. Her research interests lie in the area of wireless communications and include radio resource management for next generation wireless networks, CRNs, green communication and energy-efficient transmission techniques, cross-layer design for delay QoS provisioning in spectrum-sharing channels.

Qiang Ni (SM'08) received the B.Sc., M.Sc., and Ph.D. degrees from Huazhong University of Science and Technology, Wuhan, China, all in engineering. He led the Intelligent Wireless Communication Networking Group, Brunel University London, U.K. Currently, he is a Professor in communication systems with the School of Computing and Communications, Lancaster University, InfoLab21, Lancaster, U.K. His main research interests lie in the area of wireless communications and networking, including Green Communications, Cognitive Radio Systems, 5G, IoTs and Vehicular Networks. He was an IEEE 802.11 Wireless Standard Working Group Voting member and a contributor to the IEEE Wireless Standards.