Power Allocation for Block-Fading Channels with Arbitrary Input Constellations

Khoa D. Nguyen, Albert Guillén i Fàbregas and Lars K. Rasmussen

***Abstract*—We consider power allocation strategies for arbitrary input channels with peak, average and peak-to-average power ratio (PAPR) constraints. We are focusing on systems with a fixed and finite input constellation, as encountered in most practical systems. Generalizing previous results, we derive the optimal power allocation scheme that minimizes the outage probability of block-fading channels with arbitrary input constellations, subject to PAPR constraints. We further show that the signal-to-noise ratio exponent for any finite peak-to-average power ratio is the same as that of the peak-power limited problem, resulting in an error floor. We also derive the optimal power allocation strategies that maximize the ergodic capacity for arbitrary input channels, subject to average and PAPR constraints. We show that capacities with peak-to-average power ratio constraints, even for small ratios, are close to capacities without peak-power restrictions. For both delay-limited and ergodic block-fading channels, the optimal power allocation strategies rely on the first derivative of the input-output mutual information, which may be computationally prohibitive for efficient practical implementation. To overcome this limitation, we develop suboptimal power allocation schemes that resemble the traditional water-filling technique. The suboptimal power allocation schemes significantly reduce computational and storage requirements, while enjoying minimal performance losses as compared to optimal schemes.**

***Index Terms*—Power allocation, block-fading channel, outage probability, outage diversity, channel capacity.**

I. Introduction

A major design challenge in wireless communication is to effectively deal with the varying nature of the channel, commonly referred to as fading [1]. When knowledge of the channel fading coefficients, also known as channel state information (CSI), is available at the transmitter, power allocation schemes can be employed to improve performance [2]–[14]. Transmitter CSI can be obtained by reusing the receiver CSI for transmission in time-division duplex (TDD) systems [2], or from a dedicated feedback channel [3], providing perfect or quantized CSI [4],[5].

Manuscript received February 21, 2008; revised June 25, 2008 and September 19, 2008; accepted October 11, 2008. The associate editor coordinating the review of this paper and approving it for publication was I.-M. Kim.

K. D. Nguyen is with Institute for Telecommunications Research, University of South Australia, Mawson Lakes Boulevard, Mawson Lakes 5095, South Australia, Australia, (e-mail: dangkhoa.nguyen@postgrads.unisa.edu.au).

A. Guillén i Fàbregas is with the Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK, (e-mail: albert.guillen@eng.cam.ac.uk).

L. K. Rasmussen was with Institute for Telecommunications Research, University of South Australia. He is now with Communication Theory Lab, Royal Institute of Technology, Stockholm, Sweden, (e-mail: lars.rasmussen@ee.kth.se).

This work was presented in part at the IEEE International Symposium on Information Theory, Toronto, Canada, July 2008.

This work has been supported by the Australian Research Council under ARC grants RN0459498, DP0558861 and DP0881160.

Digital Object Identifier 10.1109/TWC.2008.xxxxxx.

In a delay-limited system, codewords are transmitted over channels with B degrees of freedom, where B is finite. Example scenarios are transmission over slowly varying channels, or OFDM transmission over frequency-selective channels. The channel is conveniently modeled as a block-fading channel [6], [7], where each codeword is transmitted over B corresponding flat-fading blocks. In this case, the maximal achievable rate is a random variable, depending on the channel realization. For most fading statistics, the channel capacity is zero since there is a non-zero probability that any positive rate is not supported by the channel. A relevant performance measure in this case is the information outage probability [7], which is the probability that communication at a target rate R is not supported by the channel. The outage probability is also a lower bound on the word-error probability for communicating at rate R [8]. The optimal power allocation problem has been investigated in [8] for channels with Gaussian inputs, and in [9], [10], [11] for channels with arbitrary input constellations. References [8], [10] consider systems with peak (per-codeword) and average power constraints, and show that systems with average constraints perform significantly better than systems with peak constraints. However, systems with average constraints may employ large (possibly infinite) peak power. As a compromise, the optimal power allocation strategy for Gaussian input channels enforcing both peak and average power constraints is considered in [8], while the use of quantized power levels is investigated in [12].

For transmission over a fast-varying fading channel, the fading statistics are revealed within each codeword, and the channel is ergodic, i.e. it has infinite degrees of freedom ($B \to \infty$). In this case, adaptive techniques aim at maximizing the ergodic channel capacity, which is the maximum data rate that can be transmitted over the channel with vanishing error probability [15]. Optimal power allocation schemes, such as water-filling for channels with Gaussian inputs [8], [15] and mercury/water-filling for channels with an arbitrary input [9], have been developed for systems with average power constraints. The work in [13] derives the optimal power allocation strategy for Gaussian input channels with both peak and average power constraint, which results in a variation to the classical water-filling algorithm [15].

In this paper, we consider power allocation strategies for arbitrary input channels with peak, average and peak-to-average power ratio (PAPR) constraints. We are focusing on cases with a fixed and finite input constellation, as encountered in most practical systems. Generalizing the results in [8], [9] yields the optimal power allocation scheme that minimizes outage probability for transmission with arbitrary inputs over a block-fading channel, subject to PAPR constraints. The optimal power allocation schemes that maximize the ergodic capacity for arbitrary input channels, subject to average and PAPR constraints, are also derived. In both cases, the optimal power allocation strategies rely on the first derivative of the input-output mutual information, which may be computationally prohibitive for practical implementation. We therefore develop suboptimal power allocation schemes for systems with peak, average and PAPR constraints that significantly reduce the complexity, while enjoying minimal performance losses from optimality.

The paper is organized as follows. Sections II and III describe the system model and the information theoretic framework. Section IV discusses power allocation algorithms for minimizing the outage probability of delay-limited block-fading channels, while algorithms for maximizing the ergodic capacity is given in Section V. Concluding remarks are given in Section VI.

The following notation is used in the paper. Expectation with respect to a random variable ξ is denoted $\mathbb{E}_{\xi}\left[\cdot\right]$, while expectation with respect to $\xi \in \mathcal{R}$ is $\mathbb{E}_{\xi\in\mathcal{R}}\left[\cdot\right]$. Component-wise inequalities are denoted $\preceq, \succeq$. The exponential equalities $f(\xi) \doteq K\xi^{-d}$ indicates that $\lim_{\xi\to\infty} f(\xi)\xi^{d} = K$, with exponential inequalities $\dot{\leq}, \dot{\geq}$ similarly defined. Finally, $(f(\xi))_{+}$ denotes $\max\{f(\xi), 0\}$.

II. System Model

Consider transmission over a channel consisting of B blocks of L channel uses, in which, block $b, b = 1, \ldots, B$, undergoes an independent fading gain h_b, corresponding to a power fading gain $\gamma_b \triangleq |h_b|^2$. Assume that $\boldsymbol{h} = (h_1, \ldots, h_B)$ and $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_B)$ are available at the receiver and the transmitter, respectively. Suppose the transmit power is allocated following the rule $\boldsymbol{p}(\boldsymbol{\gamma}) = (p_1(\boldsymbol{\gamma}), \ldots, p_B(\boldsymbol{\gamma}))$. Then the corresponding complex base-band equivalent is

$$\boldsymbol{y}_b = \sqrt{p_b(\boldsymbol{\gamma})}h_b\boldsymbol{x}_b + \boldsymbol{z}_b, \quad b = 1, \ldots, B, \tag{1}$$

where $\boldsymbol{y}_b \in \mathbb{C}^L, \boldsymbol{x}_b \in \mathcal{X}^L$, with $\mathcal{X} \subset \mathbb{C}$ being the signal constellation set, are the received and transmitted signals in block b, respectively, and $\boldsymbol{z}_b \in \mathbb{C}^L$ is the additive white Gaussian noise (AWGN) vector with independently identically distributed circularly symmetric Gaussian entries, $Z \sim \mathcal{N}_{\mathbb{C}}(0, 1)$. Assume that the signal constellation $\mathcal{X}$ of size 2^M satisfies $\sum_{x\in\mathcal{X}} |x|^2 = 2^M$, then the instantaneous received signal-to-noise ratio (SNR) at block b is given by $p_b(\boldsymbol{\gamma})\gamma_b$. We consider systems with the following power constraints:

$$\text{Peak power}: \ \langle \boldsymbol{p}(\boldsymbol{\gamma})\rangle \triangleq \frac{1}{B}\sum_{b=1}^{B} p_b(\boldsymbol{\gamma}) \leq P_{\text{peak}},$$
$$\text{Average power}: \mathbb{E}_{\boldsymbol{\gamma}}\left[\langle \boldsymbol{p}(\boldsymbol{\gamma})\rangle\right] \leq P_{\text{av}}.$$

For the fully-interleaved ergodic case, the channel model can be obtained from (1) by letting $B \to \infty$ and $L = 1$. Due to ergodicity, power allocation for block b is only dependent on γ_b. For simplicity of notation, denote $p(\gamma)$ as the transmit power corresponding to the power fading gain γ. The following power constraints are considered:

$$\text{Peak power}: p(\gamma) \leq P_{\text{peak}},$$
$$\text{Average power}: \mathbb{E}_{\gamma}\left[p(\gamma)\right] \leq P_{\text{av}}.$$

Here we study the performance of systems with peak and average power constraints [8], [13], as well as systems with a *peak-to-average power ratio* constraint $\frac{P_{\text{peak}}}{P_{\text{av}}} \leq \text{PAPR}$. The fading gain h_b is assumed to have a Nakagami-m distributed magnitude and uniformly distributed phase, perfectly compensated for. The probability density function (pdf) of the fading gain $|h_b|$ is

$$f_{|h_b|}(\xi) = \frac{2m^m\xi^{2m-1}}{\Gamma(m)}e^{-m\xi^2}, \quad b = 1, \ldots, B,$$

where $\Gamma(a)$ is the Gamma function, $\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt$. The pdf of the power fading gain is

$$f_{\gamma}(\gamma) = \frac{m^m\gamma^{m-1}}{\Gamma(m)}e^{-m\gamma}, \quad \gamma \geq 0. \tag{2}$$

The Nakagami-m distribution represents a large class of practical fading statistics. In particular, we can recover the Rayleigh fading by setting $m = 1$ and approximate the Ricean fading with parameter K by setting $m = \frac{(K+1)^2}{2K+1}$ [1].

III. Outage Probability and Ergodic Capacity

Let $I_{\mathcal{X}}(\rho)$ be the input-output mutual information of an AWGN channel with input constellation $\mathcal{X}$ and received SNR ρ. Given a channel realization $\boldsymbol{\gamma}$ and a power allocation scheme $\boldsymbol{p}(\boldsymbol{\gamma})$ satisfying the power constraint P, the instantaneous input-output mutual information of the delay-limited block-fading channel given in (1) is

$$I_B(\boldsymbol{p}(\boldsymbol{\gamma}), \boldsymbol{\gamma}) = \frac{1}{B}\sum_{b=1}^{B} I_{\mathcal{X}}(p_b\gamma_b). \tag{3}$$

For a fixed transmission rate R, communication is in outage when $I_B(\boldsymbol{p}(\boldsymbol{\gamma}), \boldsymbol{\gamma}) < R$. The outage probability, which is a lower bound to the word error probability, is given by

$$P_{\text{out}}(\boldsymbol{p}(\boldsymbol{\gamma}), P, R) \triangleq \Pr(I_B(\boldsymbol{p}(\boldsymbol{\gamma}), \boldsymbol{\gamma}) < R). \tag{4}$$

Also, the capacity of an ergodic fading channel with input constellation $\mathcal{X}$ and power allocation rule $p(\gamma)$ is given by $C \triangleq \mathbb{E}_{\gamma}\left[I_{\mathcal{X}}(p(\gamma)\gamma)\right]$.

With Gaussian inputs, we have that $I_{\mathcal{X}_G}(\rho) = \log_2(1 + \rho)$, while for coded modulation over uniformly-distributed fixed discrete signal constellations, we have that

$$I_{\mathcal{X}}(\rho) = M - \frac{1}{2^M}\sum_{x\in\mathcal{X}} \mathbb{E}\left[\log_2\left(\sum_{x'\in\mathcal{X}} e^{-|\sqrt{\rho}(x-x')+Z|^2+|Z|^2}\right)\right]$$

where the expectation is taken over $Z \sim \mathcal{N}_{\mathbb{C}}(0, 1)$.

In deriving optimal power allocation schemes, a useful measure is the first derivative of the mutual information $I_{\mathcal{X}}(\rho)$ with respect to the SNR [9], [10]. From [16] we have that,

$$\frac{d}{d\rho}I_{\mathcal{X}}(\rho) = \frac{1}{\log 2}\text{MMSE}_{\mathcal{X}}(\rho),$$

where $\text{MMSE}_{\mathcal{X}}(\rho)$ is the minimum mean-square error (MMSE) in estimating an input symbol in $\mathcal{X}$ transmitted over an AWGN channel with SNR ρ. For Gaussian inputs, $\text{MMSE}_{\mathcal{X}_G}(\rho) = \frac{1}{1+\rho}$, while for a general constellation $\mathcal{X}$, we have that [9]

$$\text{MMSE}_{\mathcal{X}}(\rho) = \frac{1}{2^M}\sum_{x\in\mathcal{X}} |x|^2 - \frac{1}{\pi}\int_{\mathbb{C}} \frac{\left|\sum_{x\in\mathcal{X}} xe^{-|y-\sqrt{\rho}x|^2}\right|^2}{\sum_{x\in\mathcal{X}} e^{-|y-\sqrt{\rho}x|^2}}dy.$$

The mutual information and its first derivative for bit-interleaved coded modulation (BICM) using the classical non-iterative BICM decoder proposed by Zehavi in [17] can also be calculated as shown in [18], [19].

IV. Outage Probability Minimization

A. Peak Power Constraints

For power constraint $P_{\rm peak}$, the power allocation problem is [8]

$$\boldsymbol{p}^{\rm opt}(\boldsymbol{\gamma}) = \arg \min_{\substack{\langle \boldsymbol{p}(\boldsymbol{\gamma})\rangle \leq P_{\rm peak} \\ \boldsymbol{p}(\boldsymbol{\gamma}) \succeq 0}} P_{\rm out}\left(\boldsymbol{p}(\boldsymbol{\gamma}), P_{\rm peak}, R\right). \quad (5)$$

The solution is given by [9]–[11]

$$p_b^{\rm opt}(\boldsymbol{\gamma}) = \frac{1}{\gamma_b}{\rm MMSE}_{\mathcal{X}}^{-1}\left(\min\left\{{\rm MMSE}_{\mathcal{X}}(0), \frac{\eta}{\gamma_b}\right\}\right), \quad (6)$$

for $b = 1, \ldots, B$, where η is chosen such that the peak power constraint is met with equality.

The optimal power allocation schemes in (6) involves an inverse MMSE function, which may be excessively complex for practical implementation. Moreover the MMSE function provides little insight into the effects of system parameters. We propose suboptimal power allocation schemes similar to water-filling that tackle both drawbacks, leading to minor losses in outage performance as compared to the optimal solution.

The complexity of the solution in (6) is due to the complex expression of $I_{\mathcal{X}}(\rho)$. Low complexity approximations of the solution can be obtained by replacing $I_{\mathcal{X}}(\rho)$ in (5) with a simpler expression. For Gaussian input channels with $I_{\mathcal{X}_G}(\rho) = \log_2(1+\rho)$, solving (5) leads to the simple water-filling scheme [15]. This motivates approximating $I_{\mathcal{X}}(\rho)$ of systems with discrete input constellations by

$$I^{\beta}(\rho) \triangleq \min\{\log_2(1+\rho), \log_2(1+\beta)\}, \quad (7)$$

where β is a design parameter. Intuitively, β approximates a threshold SNR beyond which $I_{\mathcal{X}}(\rho)$ does not increase significantly with ρ. Details on choosing β will be given later in the section.

We obtain a suboptimal power allocation scheme given by

$$\boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma}) = \arg \max_{\substack{\langle \boldsymbol{p}(\boldsymbol{\gamma})\rangle \leq P_{\rm peak} \\ \boldsymbol{p}(\boldsymbol{\gamma}) \succeq 0}} \sum_{b=1}^{B} I^{\beta}(p_b\gamma_b). \quad (8)$$

Since $I^{\beta}(p_b\gamma_b) = \log_2(1+\beta)$ for $p_b \geq \frac{\beta}{\gamma_b}$, a solution of (8) can be obtained by solving

$$\boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma}) = \arg \max_{\substack{\langle \boldsymbol{p}(\boldsymbol{\gamma})\rangle \leq P_{\rm peak} \\ 0 \leq p_b \leq \frac{\beta}{\gamma_b}, b=1,\ldots,B}} \sum_{b=1}^{B} \log_2(1+p_b\gamma_b). \quad (9)$$

By applying the Karush-Kuhn-Tucker (KKT) conditions [20], we have that

$$p_b^{\rm tw} = \min\left\{\frac{\beta}{\gamma_b}, \left(\eta - \frac{1}{\gamma_b}\right)_+\right\}, \quad (10)$$

for $b = 1, \ldots, B$, where η takes the largest value such that $\langle \boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma})\rangle \leq P_{\rm peak}$. The resulting power allocation scheme is similar to water-filling, except for the truncation of the allocated power at $\frac{\beta}{\gamma_b}$. We refer to this scheme as truncated water-filling. The outage performance obtained by the truncated water-filling scheme depends on the choice of β.

Proposition 1: Consider transmission over the block-fading channel defined in (1) with input constellation $\mathcal{X}$ and the truncated water-filling power allocation scheme $\boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma})$ given in (10). Assume that the power fading gains follow the distribution given in (2). Then, for large $P_{\rm peak}$, the outage probability $P_{\rm out}\left(\boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma}), P_{\rm peak}, R\right)$ is given by

$$P_{\rm out}\left(\boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma}), P_{\rm peak}, R\right) \doteq \mathcal{K}_{\beta}^{\rm peak} P_{\rm peak}^{-md_{\beta}(R)}, \quad (11)$$

where $d_{\beta}(R) \triangleq 1 + \left\lfloor B\left(1 - \frac{R}{I_{\mathcal{X}}(\beta)}\right)\right\rfloor$.

Proof: See the Appendix. ■

Following [10], [21], the truncated water-filling obtains the optimal outage diversity when $d_{\beta}(R) = d(R) \triangleq 1 + \left\lfloor B\left(1 - \frac{R}{M}\right)\right\rfloor$, i.e.,

$$\beta \geq I_{\mathcal{X}}^{-1}\left(\frac{BR}{B - \left\lfloor B\left(1 - \frac{R}{M}\right)\right\rfloor}\right) \triangleq \beta_R. \quad (12)$$

Therefore, by letting $\beta \to \infty$, the truncated water-filling power allocation scheme given in (10), which now becomes the classical water-filling algorithm for Gaussian inputs, provides optimal outage diversity at any transmission rate. For any rate R such that $B\left(1 - \frac{R}{M}\right)$ is not an integer, we can design a truncated water-filling scheme that obtains optimal diversity by choosing $\beta \geq \beta_R$.

With the results above, we choose β as follows. For a transmission rate R such that $B\left(1 - \frac{R}{M}\right)$ is not an integer, we perform a simulation to compute the outage probability obtained by truncated water-filling with various $\beta \geq \beta_R$ and pick the β that gives the best outage performance. The dashed lines in Figure 1 illustrate the performance of the obtained schemes for block-fading channels with $B = 4$ and QPSK input under Rayleigh fading. At all rates of interest, the truncated water-filling schemes suffer only minor losses in outage performance as compared to the optimal schemes (solid lines), especially at high SNR. We also observe a remarkable difference with respect to pure water-filling for Gaussian inputs (dotted lines). As a matter of fact, pure water-filling performs worse than uniform power allocation.

For rate R such that $B\left(1 - \frac{R}{M}\right)$ is an integer, (12) requires $\beta \to \infty$ for optimal outage diversity. Therefore, especially at high operating SNR, β needs to be relatively large to maintain diversity. While $I_{\mathcal{X}}(\rho)$ saturates at M, $I^{\beta}(\rho)$ saturates at $\log_2(1+\beta)$. Hence, a large β results in a large discrepancy between $I^{\beta}(\rho)$ and $I_{\mathcal{X}}(\rho)$, which in turns causes a large outage performance loss of the truncated water-filling scheme. For $\beta = 15$, the sub-optimality of the truncated water-filling scheme is illustrated by gap between the dashed and the solid lines in Figure 2. In the extreme case where $\beta \to \infty$ (pure water-filling), we observe a significant loss in outage performance as illustrated in Figure 1. To reduce this loss, we propose a refined truncated water-filling scheme $\boldsymbol{p}^{\rm ref}(\boldsymbol{\gamma})$, which is based on a more accurate approximation,

$$I^{\rm ref}(\rho) = \begin{cases} \log_2(1+\rho), & \rho \leq \alpha \\ \min\{\kappa\log_2(\rho) + a, \kappa\log_2(\beta) + a\}, & \text{otherwise}, \end{cases}$$

where κ and a are chosen such that (in a dB scale) $\kappa\log_2(\rho) + a$ is a tangent to $I_{\mathcal{X}}(\rho)$ at a given point ρ_0; α is chosen such

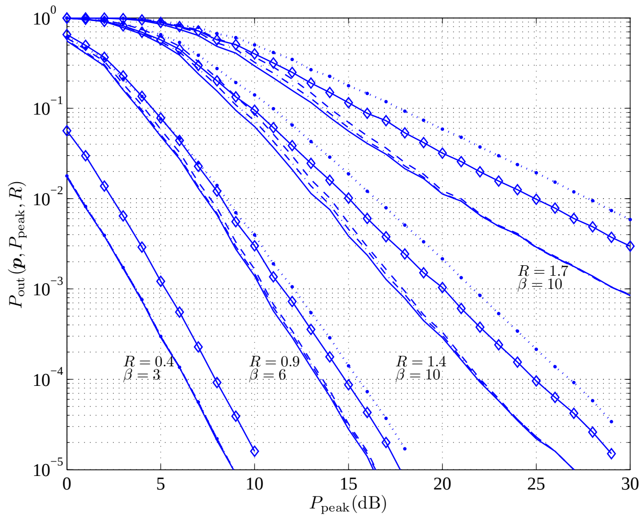

Fig. 1. Outage performance of various short-term power allocation schemes for QPSK-input block-fading channels with $B = 4$ and Rayleigh fading. The solid-lines represent the optimal scheme; the solid lines with $\diamond$ represent uniform power allocation; the dashed lines and dashed-dotted lines represent truncated water-filling and its corresponding refinement, respectively; the dotted lines represent the classical water-filling scheme.

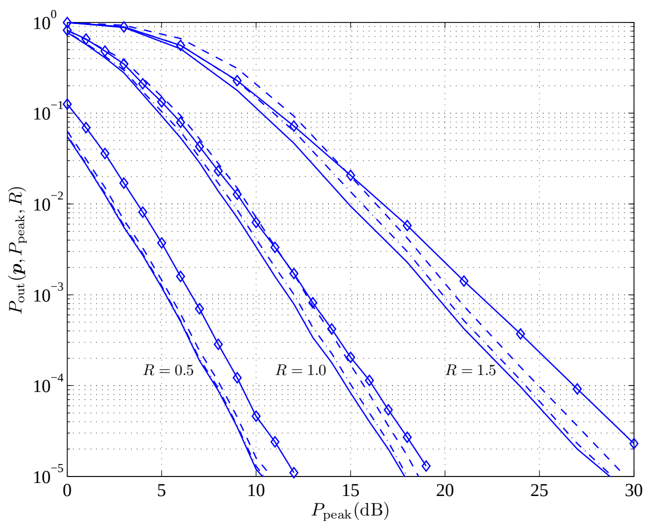

Fig. 2. Outage performance of various short-term power allocation schemes for QPSK-input block-fading channels with $B = 4$ and Rayleigh fading. The solid-lines represent the optimal scheme; the solid lines with $\diamond$ represent the uniform power allocation; the dashed lines and dashed-dotted lines correspondingly represent the truncated water-filling and its refinement with $\beta = 15$.

that $\kappa \log_2(\alpha) + a = \log_2(1+\alpha)$, and β is a design parameter. Parameters of $I^{\text{ref}}(\gamma)$ for some schemes are collected in Table I. Similar to the truncated water-filling case,

$$p_b^{\text{ref}} = \min\left\{\frac{\beta}{\gamma_b}, \kappa\eta, \frac{\alpha}{\gamma_b}, \left(\eta - \frac{1}{\gamma_b}\right)_+\right\}, \tag{13}$$

where η takes the largest value such that $\left\langle \boldsymbol{p}^{\text{ref}}(\boldsymbol{\gamma})\right\rangle \le P_{\text{peak}}$.

The refined scheme provides additional gain over the truncated water-filling scheme, especially when the transmission rate requires a relatively large β to maintain the outage diversity. The dashed-dotted lines in Figure 2 show the outage performance of the refined truncated water-filling scheme for block-fading channels with $B = 4$, and QPSK input under Rayleigh fading. The outage performance of the refined truncated water-filling scheme is close to the outage performance of the optimal case even at rates where the Singleton bound is discontinuous.

B. Average Power Constraint

Under an average power constraint, the power allocation problem is

$$\boldsymbol{p}_{\text{av}}^{\text{opt}} = \arg \min_{\substack{\mathbb{E}[\langle \boldsymbol{p}(\boldsymbol{\gamma})\rangle] \le P_{\text{av}} \\ \boldsymbol{p}(\boldsymbol{\gamma}) \succeq 0}} \Pr(I_B(\boldsymbol{p}(\boldsymbol{\gamma}), \boldsymbol{\gamma}) \le R). \tag{14}$$

From [8], [10], the solution $\boldsymbol{p}_{\text{av}}^{\text{opt}}(\boldsymbol{\gamma})$ of (14) is given by

$$\boldsymbol{p}_{\text{av}}^{\text{opt}}(\boldsymbol{\gamma}) = \begin{cases} \boldsymbol{\wp}^{\text{opt}}(\boldsymbol{\gamma}), & \langle \boldsymbol{\wp}^{\text{opt}}(\boldsymbol{\gamma})\rangle \le s^\star \\ \boldsymbol{0}, & \text{otherwise}, \end{cases} \tag{15}$$

where $\boldsymbol{\wp}(\boldsymbol{\gamma})$ is the transmission strategy that satisfies the rate constraint with minimum power,

$$\boldsymbol{\wp}^{\text{opt}}(\boldsymbol{\gamma}) = \arg \min_{\substack{I_B(\boldsymbol{\wp}(\boldsymbol{\gamma}), \boldsymbol{\gamma}) \ge R \\ \boldsymbol{\wp}(\boldsymbol{\gamma}) \succeq 0}} \langle \boldsymbol{\wp}(\boldsymbol{\gamma})\rangle, \tag{16}$$

and $s^\star$ is given by

$$s^\star = \max\{s : \mathcal{P}(\boldsymbol{\wp}^{\text{opt}}(\boldsymbol{\gamma}), s) \le P_{\text{av}}\} \tag{17}$$

with $\mathcal{P}(\boldsymbol{\wp}(\boldsymbol{\gamma}), s) \triangleq \mathbb{E}_{\langle \boldsymbol{\wp}(\boldsymbol{\gamma})\rangle \le s}[\langle \boldsymbol{\wp}(\boldsymbol{\gamma})\rangle]$ being the average power consumed by using $\boldsymbol{\wp}$ given a peak power constraint s. The outage probability is given by [10] $P_{\text{out}}(\boldsymbol{p}_{\text{av}}^{\text{opt}}(\boldsymbol{\gamma}), P_{\text{av}}, R) = P_{\text{out}}(\boldsymbol{p}^{\text{opt}}(\boldsymbol{\gamma}), s^\star, R)$. For completeness, $\boldsymbol{\wp}^{\text{opt}}(\boldsymbol{\gamma})$ is given by [10]

$$\wp_b^{\text{opt}} = \frac{1}{\gamma_b}\text{MMSE}_{\mathcal{X}}^{-1}\left(\min\left\{\text{MMSE}_{\mathcal{X}}(0), \frac{\eta}{\gamma_b}\right\}\right), \tag{18}$$

with η chosen such that $I_B(\boldsymbol{\wp}^{\text{opt}}(\boldsymbol{\gamma}), \boldsymbol{\gamma}) = R$.

The threshold $s^\star$ in (17) is fixed for a given P_{av} and fading statistic $f_{\boldsymbol{\gamma}}(\boldsymbol{\gamma})$. Consequently, the complexity of the scheme $\boldsymbol{p}_{\text{av}}^{\text{opt}}(\boldsymbol{\gamma})$ is governed by the complexity of $\boldsymbol{\wp}^{\text{opt}}(\boldsymbol{\gamma})$. Therefore, suboptimal alternatives can be employed in (15) to reduce the complexity of $\boldsymbol{p}_{\text{av}}^{\text{opt}}(\boldsymbol{\gamma})$. Following the approach in Section IV-A, we consider

$$\boldsymbol{\wp}^\star(\boldsymbol{\gamma}) = \arg \min_{\substack{\sum_{b=1}^B \log_2(1+\wp_b\gamma_b) \ge BR \\ 0 \le \wp_b \le \frac{\beta}{\gamma_b},\quad b=1,\dots,B}} \langle \boldsymbol{\wp}(\boldsymbol{\gamma})\rangle. \tag{19}$$

Applying the KKT conditions, we have

$$\wp_b^\star = \min\left\{\frac{\beta}{\gamma_b}, \left(1 - \frac{\eta}{\gamma_b}\right)_+\right\}, \tag{20}$$

where η is chosen such that $\sum_{b=1}^B \log_2(1+\wp_b^\star\gamma_b) = BR$. The power allocation rule $\boldsymbol{\wp}^\star(\boldsymbol{\gamma})$ does not meet the rate constraints since $\log_2(1+\wp_b^\star\gamma_b) \ge I_{\mathcal{X}}(\wp_b^\star\gamma_b)$. By adjusting η, we obtain

$$\wp_b^{\text{tw}} = \min\left\{\frac{\beta}{\gamma_b}, \left(1 - \frac{\eta}{\gamma_b}\right)_+\right\}, \tag{21}$$

where η is chosen such that $I_B(\boldsymbol{\wp}^{\text{tw}}(\boldsymbol{\gamma}), \boldsymbol{\gamma}) = R$. The truncated water-filling scheme for systems with average power

TABLE I
PARAMETERS ρ_0, κ, a AND α FOR THE REFINED POWER ALLOCATION SCHEME.

	Modulation Scheme							
	QPSK		8-PSK		16-QAM		64-QAM	
	CM	BICM	CM	BICM	CM	BICM	CM	BICM
ρ_0	3	3	7	7	15	15	63	63
κ	0.3528	0.3528	0.4693	0.4744	0.56	0.5608	0.6581	0.6460
a	1.1327	1.1327	1.1397	1.1234	1.347	1.3452	1.5255	1.5978
α	1.585	1.585	2.1677	2.0922	5.8884	5.8264	18.954	19.8884

constraints can be obtained by employing $\wp^{\rm tw}(\boldsymbol{\gamma})$ instead of $\wp^{\rm opt}(\boldsymbol{\gamma})$ in (15),

$$\boldsymbol{p}_{\rm av}^{\rm tw}(\boldsymbol{\gamma}) = \begin{cases} \wp^{\rm tw}(\boldsymbol{\gamma}), & \langle \wp^{\rm tw}(\boldsymbol{\gamma})\rangle \le s^{\rm tw} \\ \boldsymbol{0}, & \text{otherwise}, \end{cases} \tag{22}$$

where $\wp^{\rm tw}(\boldsymbol{\gamma})$ is given in (21) and $s^{\rm tw}$ satisfies

$$s^{\rm tw} = \max\{s : \mathcal{P}(\wp^{\rm tw}(\boldsymbol{\gamma}), s) \le P_{\rm av}\}. \tag{23}$$

The achieved outage probability is given by $P_{\rm out}\left(\boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma}), P_{\rm av}, R\right) = \Pr(\langle \wp^{\rm tw}(\boldsymbol{\gamma})\rangle > s^{\rm tw})$. The following Proposition helps analyzing the performance the power allocation rule $\boldsymbol{p}_{\rm av}^{\rm tw}(\boldsymbol{\gamma})$.

Proposition 2: Consider transmission at rate R over the block-fading channel given in (1) with power allocation scheme $\boldsymbol{p}_{\rm av}^{\rm tw}(\boldsymbol{\gamma})$ and long-term power constraint $P_{\rm av} = \mathcal{P}(\wp^{\rm tw}(\boldsymbol{\gamma}), s)$. Then, independent of the fading statistics, the outage probability satisfies

$$\Pr(\left\langle \wp^{\rm tw}(\boldsymbol{\gamma})\right\rangle > s) = P_{\rm out}\left(\boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma}), s, R\right). \tag{24}$$

Proof: Given a $\boldsymbol{\gamma}$, $\wp^{\rm tw}(\boldsymbol{\gamma}), \boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma})$ are determined by an increasing function of η. Therefore, if $\langle \wp^{\rm tw}(\boldsymbol{\gamma})\rangle > s \ge \langle \boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma})\rangle$, we have $\boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma}) \prec \wp^{\rm tw}(\boldsymbol{\gamma})$, which induces $I_B(\wp^{\rm tw}(\boldsymbol{\gamma}), \boldsymbol{\gamma}) < R$. Similarly, $I_B(\boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma}), \boldsymbol{\gamma}) < R$ induces $\langle \wp^{\rm tw}(\boldsymbol{\gamma})\rangle > s$. ■

Therefore, from Proposition 1, under Nakagami-m fading statistic,

$$P_{\rm out}\left(\boldsymbol{p}_{\rm av}^{\rm tw}(\boldsymbol{\gamma}), \mathcal{P}(\wp^{\rm tw}(\boldsymbol{\gamma}), s), R\right) \doteq \mathcal{K} s^{-md_\beta(R)}. \tag{25}$$

As shown in [10], the large-SNR performance of power allocation with average power constraint is determined from the large-SNR behavior of the corresponding power allocation scheme with peak-power constraints. Using (25), and applying the result from [10], the delay-limited capacity of systems employing $\boldsymbol{p}_{\rm av}^{\rm tw}(\boldsymbol{\gamma})$ is positive if $md_\beta(R) > 1$, while if $md_\beta(R) < 1$, the outage diversity with respect to $P_{\rm av}$ is given by $\frac{md_\beta(R)}{1-md_\beta(R)}$. Optimal outage diversity is then guaranteed if $d_\beta(R) = d(R)$.

Similar to Section IV-A, we propose the refinement of the truncated water-filling scheme, which is obtained by employing the structure in (15) with $\wp^{\rm ref}(\boldsymbol{\gamma})$, where

$$\wp_b^{\rm ref} = \min\left\{\frac{\beta}{\gamma_b}, \kappa\eta, \frac{\alpha}{\gamma_b}, \left(\eta - \frac{1}{\gamma_b}\right)_+\right\}, \tag{26}$$

and η is chosen such that $I_B(\wp^{\rm ref}(\boldsymbol{\gamma}), \boldsymbol{\gamma}) = R$.

Determining η for the power allocation rules in (21) and (26) requires evaluating or tabulating $I_{\mathcal{X}}(\rho)$, which may be computationally complex. This can be avoided by using an approximation $\tilde{I}_{\mathcal{X}}(\rho)$ of $I_{\mathcal{X}}(\rho)$. Let $\Delta R \triangleq \max_\rho(\tilde{I}_{\mathcal{X}}(\rho) - I_{\mathcal{X}}(\rho))$, then, a suboptimal $\wp(\boldsymbol{\gamma})$ satisfies the rate constraint when η is chosen such that $\sum_{b=1}^{B} \tilde{I}_{\mathcal{X}}(\wp_b\gamma_b) = B(R + \Delta R)$. Following [22], we propose $\tilde{I}_{\mathcal{X}}(\rho) = M(1-e^{-c_1\rho^{c_2}})^{c_3}$. Using numerical optimization to minimize the mean-squared-error between $I_{\mathcal{X}}(\rho)$ and $\tilde{I}_{\mathcal{X}}(\rho)$, the parameters c_1, c_2, c_3 for various modulation schemes are showed in Table II. The performance of the suboptimal schemes is illustrated by the dashed lines in Figure 3. As we observe, the performance of the truncated water-filling scheme is very close to that of the optimal scheme; the refined truncated water-filling scheme provides additional gains, especially for systems operate at rates close to M.

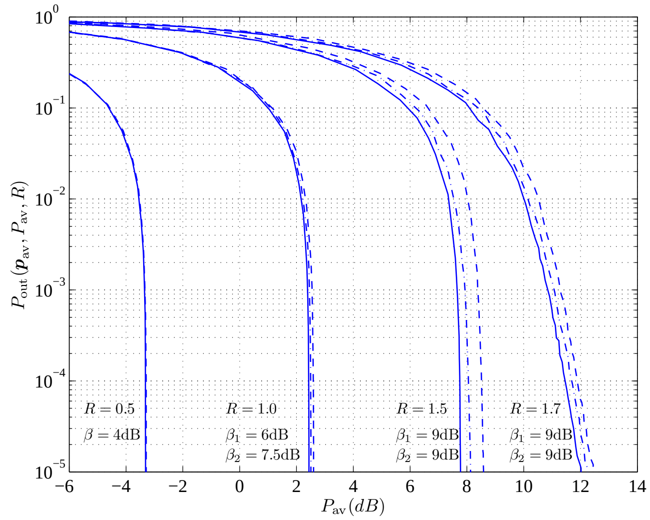

Fig. 3. Outage performance of various long-term power allocation schemes for QPSK-input block-fading channels with $B = 4$ and Rayleigh fading. The solid-lines represent the optimal scheme; the dashed lines and dashed-dotted lines correspondingly represent the long-term truncated water-filling $\left(\boldsymbol{p}_{\rm av}^{\rm tw}(\boldsymbol{\gamma}) \text{ with } \beta_1\right)$ and its refinement $\left(\boldsymbol{p}_{\rm av}^{\rm ref}(\boldsymbol{\gamma}) \text{ with } \beta_2\right)$.

C. Peak-to-Average Power Ratio Constraints

For systems with average power $P_{\rm av}$ and peak-to-average power ratio constraint $\rm PAPR$, the power allocation rule is subject to both average power constraint $P_{\rm av}$ and peak power constraint $P_{\rm peak} = {\rm PAPR} \cdot P_{\rm av}$, the optimal power allocation scheme solves [8],

$$\hat{\boldsymbol{p}}(\boldsymbol{\gamma}) = \arg \min_{\substack{\langle \boldsymbol{p}(\boldsymbol{\gamma})\rangle \le P_{\rm peak} \\ \mathbb{E}[\langle \boldsymbol{p}(\boldsymbol{\gamma})\rangle] \le P_{\rm av} \\ \boldsymbol{p}(\boldsymbol{\gamma}) \succeq 0}} \Pr(I_B(\boldsymbol{p}(\boldsymbol{\gamma}), \boldsymbol{\gamma}) < R) \tag{27}$$

TABLE II
OPTIMIZED c_1, c_2 AND c_3 PARAMETERS FOR THE APPROXIMATION OF $\tilde{I}_{\mathcal{X}}(\rho)$.

	Modulation Scheme							
	QPSK		8-PSK		16-QAM		64-QAM	
	CM	BICM	CM	BICM	CM	BICM	CM	BICM
c_1	0.77	0.77	0.61	0.81	0.48	0.59	0.47	0.4
c_2	0.87	0.87	0.68	0.06	0.61	0.06	0.44	0.05
c_3	1.16	1.16	1.45	1.75	1.48	1.65	1.87	1.63
ΔR	0.0033	0.0033	0.0241	0.0223	0.0414	0.0259	0.0977	0.0656

Following the arguments in [8], the optimal power allocation rule $\hat{\boldsymbol{p}}^{\rm opt}(\boldsymbol{\gamma})$ is given by

$$\hat{\boldsymbol{p}}^{\rm opt}(\boldsymbol{\gamma}) = \begin{cases} \wp^{\rm opt}(\boldsymbol{\gamma}), & \langle \wp^{\rm opt}(\boldsymbol{\gamma})\rangle \le \hat{s} \\ \mathbf{0}, & \text{otherwise}, \end{cases} \tag{28}$$

where $\wp^{\rm opt}(\boldsymbol{\gamma})$ is given in (18) and $\hat{s} = \min\{s^\star, P_{\rm peak}\}$ with $s^\star$ defined as in (17).

We observe that, depending on $P_{\rm av}$ and the PAPR (which is fixed), one of the power constraints is redundant and the outage performance is dependent on the remaining constraint. In particular we have that

$$P_{\rm out}\left(\hat{\boldsymbol{p}}^{\rm opt}(\boldsymbol{\gamma}), P_{\rm av}, R\right) = \begin{cases} P_{\rm out}\left(\boldsymbol{p}^{\rm opt}(\boldsymbol{\gamma}), P_{\rm peak}, R\right), & s^\star > P_{\rm peak} \\ P_{\rm out}\left(\boldsymbol{p}_{\rm av}^{\rm opt}(\boldsymbol{\gamma}), P_{\rm av}, R\right), & s^\star \le P_{\rm peak}. \end{cases} \tag{29}$$

Consequently, the outage probability can also be evaluated as

$$P_{\rm out}\left(\hat{\boldsymbol{p}}^{\rm opt}(\boldsymbol{\gamma}), P_{\rm av}, R\right) = \max\left\{P_{\rm out}\left(\boldsymbol{p}^{\rm opt}(\boldsymbol{\gamma}), P_{\rm peak}, R\right), P_{\rm out}\left(\boldsymbol{p}_{\rm av}^{\rm opt}(\boldsymbol{\gamma}), P_{\rm av}, R\right)\right\}. \tag{30}$$

Suboptimal schemes based on truncated and refined truncated water-filling follow directly from the approach described above. For example, the truncated water-filling solution $\hat{\boldsymbol{p}}^{\rm tw}(\boldsymbol{\gamma})$ is obtained from (22) by replacing $s^{\rm tw}$ with $\hat{s}^{\rm tw} = \min\{P_{\rm peak}, s^{\rm tw}\}$. For large $P_{\rm av}$, we have the following.

Proposition 3: Consider transmission at rate R over the block-fading channel given in (1) with power allocation scheme $\hat{\boldsymbol{p}}^{\rm opt}(\boldsymbol{\gamma})$ (or $\hat{\boldsymbol{p}}^{\rm tw}(\boldsymbol{\gamma})$). Assume input constellation $\mathcal{X}$ of size 2^M. Further assume that the power fading gains $\boldsymbol{\gamma}$ follow the Nakagami-m distribution given in (2). Then, for large $P_{\rm av}$ and any $\mathrm{PAPR} < \infty$, the outage probability behaves like

$$P_{\rm out}\left(\hat{\boldsymbol{p}}^{\rm opt}(\boldsymbol{\gamma}), P_{\rm av}, R\right) \doteq \mathcal{K} P_{\rm av}^{-md(R)} \tag{31}$$

$$P_{\rm out}\left(\hat{\boldsymbol{p}}^{\rm tw}(\boldsymbol{\gamma}), P_{\rm av}, R\right) \doteq \mathcal{K}_\beta P_{\rm av}^{-md_\beta(R)}. \tag{32}$$

Proof: For sufficiently large $P_{\rm peak}$ we have that [10]

$$P_{\rm out}\left(\boldsymbol{p}^{\rm opt}(\boldsymbol{\gamma}), P_{\rm peak}, R\right) \doteq \mathcal{K}^{\rm peak} P_{\rm peak}^{-md(R)}. \tag{33}$$

Let $\mathcal{P}(\wp^{\rm opt}(\boldsymbol{\gamma}), s)$ be the average power constraint as a function of the threshold s in the allocation scheme $\boldsymbol{p}_{\rm av}^{\rm opt}(\boldsymbol{\gamma})$ in (15). Asymptotically with s [10],

$$\frac{d}{ds}\mathcal{P}(\wp^{\rm opt}(\boldsymbol{\gamma}), s) \doteq \mathcal{K}^{\rm peak} d(R) s^{-d(R)}.$$

Therefore, from L'Hôpital's rule, we have for any PAPR

$$\lim_{s\to\infty} \frac{\mathrm{PAPR}\cdot\mathcal{P}(\wp^{\rm opt}(\boldsymbol{\gamma}), s)}{s} = \lim_{s\to\infty} \frac{d}{ds}\mathrm{PAPR}\cdot\mathcal{P}(\wp^{\rm opt}(\boldsymbol{\gamma}), s) = 0.$$

It follows that for any PAPR, there exists an s_0 and the corresponding average power constraint $P_0 = \mathcal{P}(\wp^{\rm opt}(\boldsymbol{\gamma}), s_0)$ such that $s_0 = \mathrm{PAPR}\cdot P_0$ and $s > \mathcal{P}(\wp^{\rm opt}(\boldsymbol{\gamma}), s)\cdot\mathrm{PAPR}$ if $\mathcal{P}(\wp^{\rm opt}(\boldsymbol{\gamma}), s) > P_0$. Consequently, $P_{\rm out}\left(\hat{\boldsymbol{p}}^{\rm opt}(\boldsymbol{\gamma}), P_{\rm av}, R\right) = P_{\rm out}\left(\boldsymbol{p}^{\rm opt}(\boldsymbol{\gamma}), \mathrm{PAPR}\cdot P_{\rm av}, R\right)$ for $P_{\rm av} > P_0$. Thus, together with (33), at large $P_{\rm av}$, we have

$$\begin{aligned} P_{\rm out}\left(\hat{\boldsymbol{p}}^{\rm opt}(\boldsymbol{\gamma}), P_{\rm av}, R\right) &\doteq P_{\rm out}\left(\boldsymbol{p}^{\rm opt}(\boldsymbol{\gamma}), \mathrm{PAPR}\cdot P_{\rm av}, R\right) \\ &\doteq \mathcal{K}^{\rm peak}\mathrm{PAPR}^{-md(R)} P_{\rm av}^{-md(R)}. \end{aligned}$$

By noting that $P_{\rm out}\left(\boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma}), P_{\rm peak}, R\right) \doteq \mathcal{K}_\beta^{\rm peak} P_{\rm peak}^{-md_\beta(R)}$, the proof for the suboptimal scheme $\hat{\boldsymbol{p}}^{\rm tw}(\boldsymbol{\gamma})$ follows using the same arguments as above. ■

The threshold P_0 in the proof is the average power constraint such that the threshold s in (17) satisfies $s = \mathrm{PAPR}\cdot P_0$. Equivalently, P_0 satisfies

$$\int_{\boldsymbol{\gamma}:\langle\wp^{\rm opt}(\boldsymbol{\gamma})\rangle \le \mathrm{PAPR}\cdot P_0} \left\langle \wp^{\rm opt}(\boldsymbol{\gamma})\right\rangle dF_{\boldsymbol{\gamma}}(\boldsymbol{\gamma}) = P_0, \tag{34}$$

where $F_{\boldsymbol{\gamma}}(\boldsymbol{\gamma})$ is the joint pdf of $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_B)$. We therefore have that

$$P_{\rm out}\left(\hat{\boldsymbol{p}}^{\rm opt}(\boldsymbol{\gamma}), P_{\rm av}, R\right) = P_{\rm out}\left(\boldsymbol{p}^{\rm opt}(\boldsymbol{\gamma}), \mathrm{PAPR}\cdot P_{\rm av}, R\right)$$

for $P_{\rm av} > P_0$. Thus, for asymptotically large $P_{\rm av}$, the outage probability for systems with a PAPR constraint is determined by the outage probability of systems with peak power constraint $P_{\rm peak} = \mathrm{PAPR}\cdot P_{\rm av}$. As a consequence, we have that the delay-limited capacity [23] is zero for any finite PAPR. For simplicity, we first consider the outage performance of systems with $B = 1$ under Nakagami-m fading statistic. Let $F_\gamma(\gamma)$ be the cumulative distribution function (cdf) of γ. For $P_{\rm av} > P_0$ $(s > \mathrm{PAPR}\cdot P_{\rm av})$ the outage probability is

$$\begin{aligned} P_{\rm out}\left(\boldsymbol{p}^{\rm opt}(\boldsymbol{\gamma}), \mathrm{PAPR}\cdot P_{\rm av}, R\right) &= \Pr\left(\gamma < \frac{I_{\mathcal{X}}^{-1}(R)}{\mathrm{PAPR}\cdot P_{\rm av}}\right) \\ &= F_\gamma\left(\frac{I_{\mathcal{X}}^{-1}(R)}{\mathrm{PAPR}\cdot P_{\rm av}}\right). \end{aligned}$$

For $P_{\rm av} < P_0$ $(s < \mathrm{PAPR}\cdot P_{\rm av})$, s in (15) is obtained by solving

$$\frac{mI_{\mathcal{X}}^{-1}(R)}{\Gamma(m)}\Gamma\left(m-1, \frac{mI_{\mathcal{X}}^{-1}(R)}{s}\right) = P_{\rm av},$$

and the outage probability is given by

$$\begin{aligned} P_{\rm out}\left(\boldsymbol{p}_{\rm av}^{\rm opt}(\boldsymbol{\gamma}), P_{\rm av}, R\right) &= \Pr\left(\gamma < \frac{I_{\mathcal{X}}^{-1}(R)}{s}\right) \\ &= F_\gamma\left(\frac{I_{\mathcal{X}}^{-1}(R)}{s}\right). \end{aligned}$$

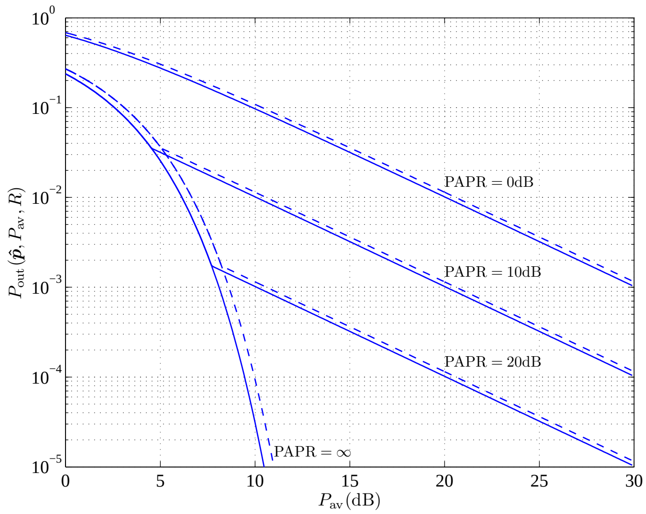

Fig. 4. Outage probability for systems with PAPR constraints over Nakagami-m block-fading channels $B = 1, m = 1, R = 1$, 16-QAM inputs. The solid and dashed lines correspondingly represent outage probability of systems with coded modulation and BICM.

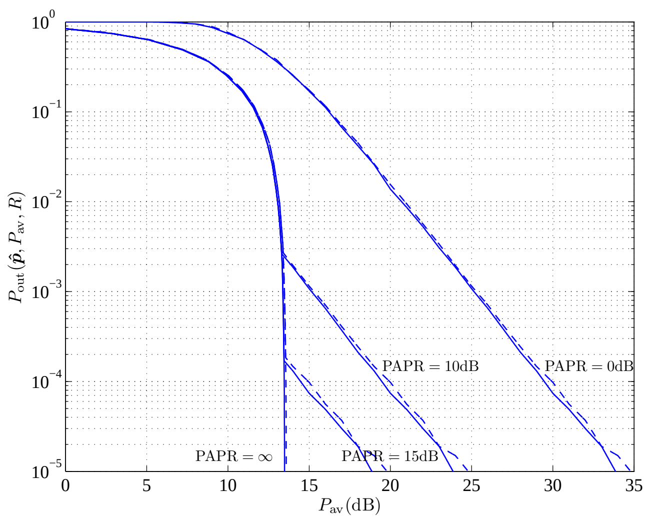

Fig. 5. Outage probability for systems with PAPR constraints over Nakagami-m block-fading channels $B = 4, m = 1, R = 3$, 16-QAM inputs. The solid and dashed lines correspondingly represent outage probability of systems with coded modulation and BICM.

The analytical result for $B = 1$ is illustrated in Figure 4 for a 16-QAM input, Rayleigh fading channel at rate $R = 1$. We observe that as we increase the PAPR constraint, the error floor occurs at lower error probability values, and eventually, at values below a target quality-of-service error rate. We also observe that the loss incurred by BICM is minimal.

For systems with $B > 1$, analytical results are not available in closed form. However, from (30), the outage probability of systems with PAPR constraints can be obtained from systems with peak power constraints and systems with average power constraints separately. Moreover, at high P_{av}, the outage probability can be obtained by the outage probability of systems with only a peak power constraint $P_{\mathrm{av}} \cdot \mathrm{PAPR}$. Simulation results for a 16-QAM input, Rayleigh fading channel with $B = 4$ blocks at rate $R = 3$ are given in Figure 5. In both

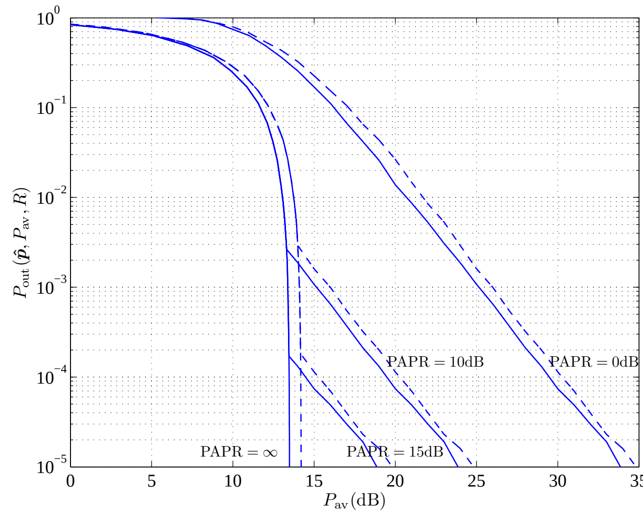

Fig. 6. Outage probability for systems with peak and average power constraints using 16-QAM input constellation over Nakagami-m block-fading channels with $B = 4, m = 1, R = 3$ and peak-to-average power ratio PAPR. The solid and dashed lines correspondingly represent outage probability of systems with optimal and truncated water-filling schemes with $\beta = 19$ dB.

cases the outage probability at high P_{av} resulting from the optimal power allocation scheme is governed by the peak power constraints, and therefore, the optimal outage diversity is given by the Singleton bound.

The outage performance of systems with 16-QAM inputs, Rayleigh fading channel with $B = 4, R = 3$, employing the truncated water-filling scheme is illustrated in Figure 6. As pointed out in Section IV-A, relatively high β (19 dB) is required to maintain the optimal outage diversity up to the P_{av} of interest. In this case, the refined truncated water-filling power allocation scheme, as discussed in the previous section, can be used to reduce the gap from optimality. For rates where smaller β can be used, small gaps from the optimal outage probability are observed, as for systems with peak and with average power constraints.

V. Ergodic Capacity Maximization

We now consider the capacity of the ergodic channel, where the number of blocks B is sufficiently large to reveal the statistics of the channel within one codeword. For a given power allocation rule $p(\gamma)$, the ergodic capacity of the channel is

$$C = \mathbb{E}_{\gamma}\left[I_{\mathcal{X}}(p(\gamma)\gamma)\right] = \int_{\gamma>0} I_{\mathcal{X}}(p(\gamma)\gamma) f_{\gamma}(\gamma) d\gamma. \qquad (35)$$

A. Average Power Constraint

For a system with an average power constraint P_{av}, the power allocation problem is

$$p^{\mathrm{opt}}(\gamma) = \arg \max_{\substack{\mathbb{E}_{\gamma}[p(\gamma)] \le P_{\mathrm{av}} \\ p(\gamma) \ge 0}} \mathbb{E}_{\gamma}\left[I_{\mathcal{X}}(p(\gamma)\gamma)\right]. \qquad (36)$$

The solution is given by [9]

$$p^{\mathrm{opt}}(\gamma) = \frac{1}{\gamma} \mathrm{MMSE}_{\mathcal{X}}^{-1}\left(\min\left\{\mathrm{MMSE}_{\mathcal{X}}(0), \frac{\eta}{\gamma}\right\}\right), \qquad (37)$$

where η is chosen such that $\mathbb{E}_\gamma\left[p^{\text{opt}}(\gamma)\right] = P_{\text{av}}$. The resulting capacity is

$$C^{\text{opt}} = \int_{\frac{\eta}{\text{MMSE}_{\mathcal{X}}(0)}}^{\infty} I_{\mathcal{X}}\left(\text{MMSE}_{\mathcal{X}}^{-1}\left(\frac{\eta}{\gamma}\right)\right) f_\gamma(\gamma) d\gamma. \quad (38)$$

As before, the suboptimal truncated water-filling scheme can be obtained by solving

$$p^{\text{tw}}(\gamma) = \arg \max_{\substack{\mathbb{E}_\gamma[p(\gamma)] \le P_{\text{av}} \\ 0 \le p(\gamma) \le \frac{\beta}{\gamma}}} \mathbb{E}_\gamma\left[\log_2(1 + p(\gamma)\gamma)\right]. \quad (39)$$

Using the KKT conditions, we have that

$$p^{\text{tw}}(\gamma) = \min\left\{\frac{\beta}{\gamma}, \left(\eta - \frac{1}{\gamma}\right)_+\right\}, \quad (40)$$

where η is chosen such that $\mathbb{E}_\gamma\left[p^{\text{tw}}(\gamma)\right] = P_{\text{av}}$. The resulting capacity is

$$C^{\text{tw}} = \int_{\frac{1}{\eta}}^{\frac{\beta+1}{\eta}} I_{\mathcal{X}}(\eta\gamma - 1) f_\gamma(\gamma) d\gamma + I_{\mathcal{X}}(\beta)\left(1 - F_\gamma\left(\frac{\beta+1}{\eta}\right)\right). \quad (41)$$

B. Peak-to-Average Power Constraint

For systems with a PAPR constraints, the optimal power allocation rule is given by

$$p_{\text{papr}}^{\text{opt}}(\gamma) = \arg \max_{\substack{\mathbb{E}_\gamma[0 \le p(\gamma)] \le P_{\text{av}} \\ p(\gamma) \le P_{\text{peak}}}} \mathbb{E}_\gamma\left[I_{\mathcal{X}}(p(\gamma)\gamma)\right], \quad (42)$$

where $P_{\text{peak}} = \text{PAPR} \cdot P_{\text{av}}$. Applying the KKT conditions, the optimal power allocation is

$$p_{\text{papr}}^{\text{opt}}(\gamma) = \min\left\{P_{\text{peak}}, \frac{1}{\gamma}\text{MMSE}_{\mathcal{X}}^{-1}\left(\min\left\{\text{MMSE}_{\mathcal{X}}(0), \frac{\eta}{\gamma}\right\}\right)\right\},$$

where η is chosen such that $\mathbb{E}_\gamma\left[p_{\text{papr}}^{\text{opt}}(\gamma)\right] = P_{\text{av}}$.

The suboptimal power allocation rule based on the truncated water-filling algorithm is

$$p_{\text{papr}}^{\text{tw}}(\gamma) = \arg \max_{\substack{\mathbb{E}_\gamma[p(\gamma)] \le P_{\text{av}} \\ 0 \le p(\gamma) \le \min\left\{P_{\text{peak}}, \frac{\beta}{\gamma}\right\}}} \mathbb{E}_\gamma\left[\log_2(1 + p(\gamma)\gamma)\right].$$

Letting $\alpha(\gamma) = \min\left\{P_{\text{peak}}, \frac{\beta}{\gamma}\right\}$, then a truncated water-filling scheme is given by

$$p_{\text{papr}}^{\text{tw}}(\gamma) = \min\left\{\alpha(\gamma), \left(\eta - \frac{1}{\gamma}\right)_+\right\}, \quad (43)$$

where η is chosen such that $\mathbb{E}_\gamma\left[p_{\text{papr}}^{\text{tw}}(\gamma)\right] = P_{\text{av}}$. It can be seen that if $\eta \le P_{\text{peak}}$ or $\frac{\beta+1}{\eta} \le \frac{1}{\eta - P_{\text{peak}}}$, (43) is equivalent to (40). Therefore, the resulting ergodic capacity is given in (41). Otherwise, let $a = \frac{1}{\eta - P_{\text{peak}}}$ and $b = \frac{\beta}{P_{\text{peak}}}$, then the resulting ergodic capacity can be written as

$$C_{\text{papr}}^{\text{tw}} = \int_{1/\eta}^{a} I_{\mathcal{X}}(\eta\gamma - 1) f_\gamma(\gamma) d\gamma + \int_a^b I_{\mathcal{X}}(P_{\text{peak}}\gamma) f_\gamma(\gamma) d\gamma + \left(1 - F_\gamma(b)\right) I_{\mathcal{X}}(\beta).$$

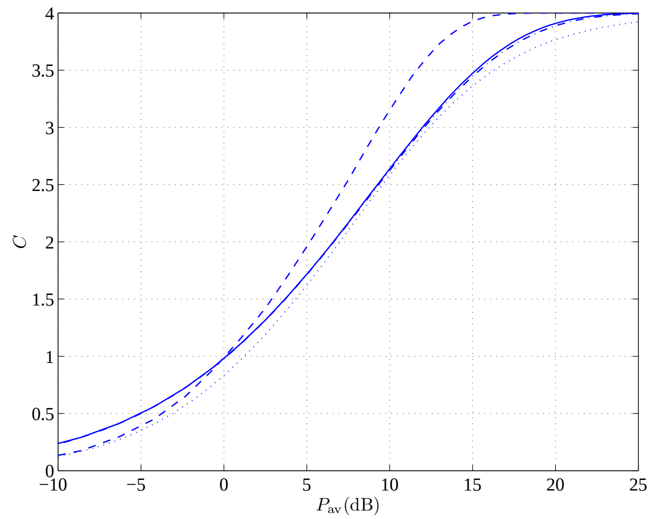

Fig. 7. Capacity of ergodic fading channel with $m = 1$, 16-QAM coded modulation inputs and average power constraint. The dashed line represents capacity of the unfaded AWGN channel, and the solid, dashed-dotted and dotted lines correspondingly represent capacities with optimal, truncated water-filling and uniform power allocation.

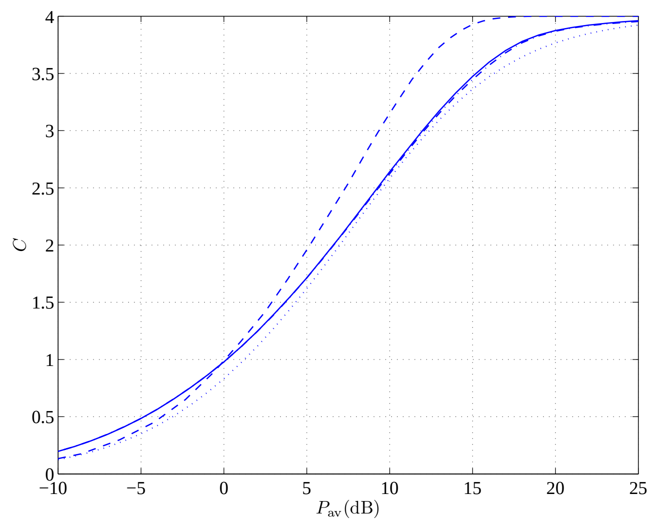

Fig. 8. Capacity of ergodic fading channel with $m = 1$, 16-QAM coded modulation inputs and $\text{PAPR} = 3$dB. The dashed line represents capacity of the unfaded AWGN channel and the solid, dashed-dotted and dotted lines correspondingly represent 16-QAM capacities with optimal, truncated water-filling and uniform power allocation.

C. Numerical Results

Numerical results for the ergodic capacity of Rayleigh fading channels with 16-QAM inputs are presented in Figures 7, 8, 9. Figures 7 and 8 show the performance of the truncated water-filling scheme with average power constraints and PAPR constraints, respectively, where β has been chosen to maximize capacity at each P_{av}. The results show that the performance of the truncated water-filling scheme is close to optimal for both systems with average power constraints and systems with PAPR constraints. Figure 9 shows the ergodic capacity for various PAPR constraints. Loss in capacity due to PAPR constraints occurs at low and high P_{av}. Still, the loss is minimal, even with relatively small PAPR (4dB).

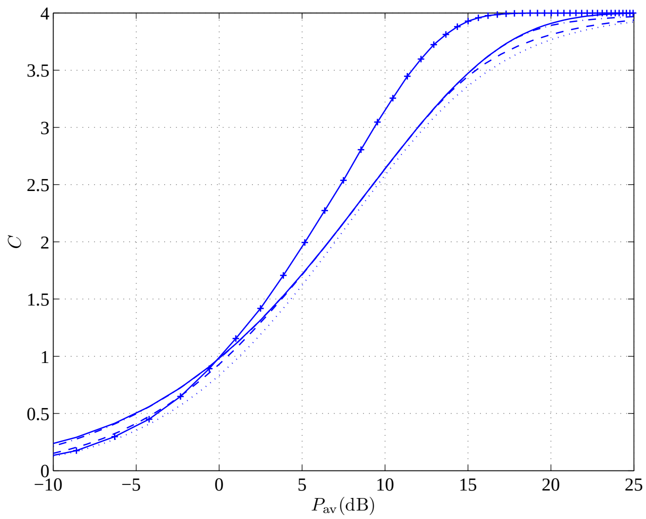

Fig. 9. Capacity of ergodic fading channel with $m = 1$, 16-QAM coded modulation inputs and PAPR constraints. The solid line with crosses represents capacity of the unfaded AWGN channel; the dotted line represents the capacity with uniform power allocation and the solid, dashed-dotted and dashed lines correspondingly represent 16-QAM capacities with $\mathrm{PAPR} = \infty, 4, 1$dB.

VI. Conclusions

We have studied power allocation schemes under peak, average and peak-to-average power constraints for ergodic and delay-limited block-fading channels with arbitrary input distributions. We propose suboptimal schemes with low computational and storage capabilities, while still performing close to optimal. In the delay-limited block-fading case, we have shown that the suboptimal scheme maintains the optimal outage diversity in most cases. A refined truncated water-filling scheme is proposed to reduce the performance loss when optimal outage diversity cannot be obtained by truncated water-filling. For systems with PAPR constraints, the optimal and suboptimal solutions can be easily computed from the corresponding solutions with independent peak and average power constraints. The asymptotic performance for finite PAPR is always determined by the peak power, and the exponent is therefore given by the exponent of systems with peak power constraints. In the ergodic case, the truncated water-filling scheme yields insignificant loss compare to the optimal scheme for the entire SNR range. Also, even small PAPR values entail minimal loss to the ergodic capacity.

Appendix

Let

$$I_1^\beta(\rho) \triangleq \begin{cases} I_{\mathcal{X}}(\beta) & \rho > \beta \\ 0, & \text{otherwise.} \end{cases} \tag{44}$$

Since $p_b^{\rm tw}\gamma_b \le \beta$, we have that,

$$\sum_{b=1}^B I_1^\beta(p_b^{\rm tw}\gamma_b) \le \sum_{b=1}^B I_B(\boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma}),\boldsymbol{\gamma}) \le \sum_{b=1}^B I^\beta(BP_{\rm peak}\gamma_b) \tag{45}$$

We further lower bound $I_B(\boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma}),\boldsymbol{\gamma})$ by the following Proposition.

Proposition 4: Consider the truncated water-filling scheme given in (10). We have that

$$\sum_{b=1}^B I_1^\beta(p_b^{\rm tw}\gamma_b) \ge \sum_{b=1}^B I_1^\beta(P_{\rm peak}\gamma_b) \tag{46}$$

for any channel realization $\boldsymbol{\gamma}$, where $I_1^\beta(\rho)$ is given in (44).

Proof: According to (44), $I_{\mathcal{X}}^\beta(P_{\rm peak}\gamma_b)$ is non-zero only if $\gamma_b \ge \frac{\beta}{P_{\rm peak}}$. Therefore, we need to prove that if $\gamma_b \ge \frac{\beta}{P_{\rm peak}}$ then $p_b^{\rm tw} \ge \frac{\beta}{\gamma_b}$ for all realization of $\boldsymbol{\gamma}$. Without loss of generality, assume that $\gamma_1 \le \ldots \le \gamma_B$. If $\gamma_B < \frac{\beta}{P_{\rm peak}}$, (46) is certainly true. Otherwise, there exists a $k, 1 \le k \le B$, such that $\gamma_{k-1} < \frac{\beta}{P_{\rm peak}} \le \gamma_k \le \ldots \le \gamma_B$. Consider the following two cases.

- If $\sum_{b=1}^B \frac{1}{\gamma_b} < \frac{BP}{\beta}$ then from (10), $p_b^{\rm tw} = \frac{\beta}{\gamma_b}$, $b = 1,\ldots,B$.
- Otherwise, from (10), the power allocation solution is given by

$$p_b^{\rm tw} = \min\left\{\frac{\beta}{\gamma_b}, \left(\eta - \frac{1}{\gamma_b}\right)_+\right\},\ b = 1,\ldots,B, \tag{47}$$

where η is chosen such that $\sum_{b=1}^B p_b^{\rm tw} = BP$. Since $\gamma_k \ge \frac{\beta}{P}$, we have from (10) that $p_b^{\rm tw} \le \frac{\beta}{\gamma_b} \le \frac{\beta}{\gamma_k} \le P$, $b = k,\ldots,B$. Therefore,

$$\sum_{b=1}^k p_b^{\rm tw} = \sum_{b=1}^B p_b^{\rm tw} - \sum_{b=k+1}^B p_b^{\rm tw} \ge kP. \tag{48}$$

Now, suppose $\eta < \frac{\beta+1}{\gamma_k}$, then, for $b = 1,\ldots,k$, $p_b^{\rm tw} \le \eta - \frac{1}{\gamma_b} < \frac{\beta+1}{\gamma_k} - \frac{1}{\gamma_k} = \frac{\beta}{\gamma_k} \le P$. Thus, $\sum_{b=1}^k p_b^{\rm tw} < kP$, which contradicts to (48). Therefore, assumption $\eta < \frac{\beta+1}{\gamma_k}$ is invalid. We then conclude that $\eta \ge \frac{\beta+1}{\gamma_k} \ge \frac{\beta+1}{\gamma_b}, b = k,\ldots,B$. Therefore from (47), $p_b^{\rm tw} = \frac{\beta}{\gamma_b}, b = k,\ldots,B$.

Thus, in all cases, we have $p_b^{\rm tw} = \frac{\beta}{\gamma_b}$ if $\gamma_b \ge \frac{\beta}{P}$. This concludes the proof of the proposition. ■

Therefore, from (45), we have that

$$\Pr\left(\sum_{b=1}^B I^\beta(BP_{\rm peak}) < BR\right) \le P_{\rm out}\left(\boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma}), P_{\rm peak}, R\right) \le \Pr\left(\sum_{b=1}^B I_1^\beta(P_{\rm peak}\gamma_b) < BR\right)$$

With similar arguments to the analysis in [21], we have that with Nakagami-m fading,

$$\Pr\left(\sum_{b=1}^B I^\beta(BP_{\rm peak}\gamma_b) < BR\right) \doteq \mathcal{K}_\ell P_{\rm peak}^{-md_\beta(R)}$$
$$\Pr\left(\sum_{b=1}^B I_1^\beta(P_{\rm peak}\gamma_b) < BR\right) \doteq \mathcal{K}_u P_{\rm peak}^{-md_\beta(R)},$$

which leads to $P_{\rm out}\left(\boldsymbol{p}^{\rm tw}(\boldsymbol{\gamma}), P_{\rm peak}, R\right) \doteq \mathcal{K}_\beta^{\rm peak} P_{\rm peak}^{-md_\beta(R)}$ as required.

REFERENCES

[1] J. G. Proakis, *Digital Communications*, 4th ed. McGraw Hill, 2001.

[2] R. Knopp and G. Caire, "Power control and beamforming for systems with multiple transmit and receive antennas," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 638-648, Oct. 2002.

[3] V. K. N. Lau, Y. Liu, and T.-A. Chen, "Capacity of memoryless channels and block-fading channels with designable cardinality-constrained channel state feedback," *IEEE Trans. Inform. Theory*, vol. 50, no. 9, pp. 2038-2049, Sept. 2004.

[4] S. Bhashyam, A. Sabharwal, and B. Aazhang, "Feedback gain in multiple antenna systems," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 85-798, May 2002.

[5] 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.

[6] L. H. Ozarow, S. Shamai, and A. D. Wyner, "Information theoretic considerations for cellular mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 359-378, May 1994.

[7] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: Informatic-theoretic and communications aspects," *IEEE Trans. Inform. Theory*, vol. 44, no. 6, pp. 2619-2692, Oct. 1998.

[8] G. Caire, G. Taricco, and E. Biglieri, "Optimal power control over fading channels," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1468-1489, July 1999.

[9] A. Lozano, A. M. Tulino, and S. Verdú, "Optimum power allocation for parallel Gaussian channels with arbitrary input distributions," *IEEE Trans. Inform. Theory*, vol. 52, no. 7, pp. 3033-3051, July 2006.

[10] K. D. Nguyen, A. Guillén i Fàbregas, and L. K. Rasmussen, "Large-SNR outage analysis of power allocation in block-fading channels," *submitted to IEEE Trans. Inform. Theory*, May 2008.

[11] G. Caire and K. R. Kumar, "Information-theoretic foundations of adaptive coded modulation," in *Proc. IEEE*, vol. 95, no. 12, pp. 2274-2298, Dec. 2007.

[12] A. Khoshnevis and A. Sabharwal, "Performance of quantized power control in multiple antenna systems," in *IEEE Int. Conf. Commun.*, Paris, France, June 2004, pp. 803-807.

[13] M. A. Khojastepour and B. Aazhang, "The capacity of average and peak power constrained fading channels with channel side information," in *2004 IEEE Wireless Commun. Networking Conf.*, Atlanta, Georgia, USA, 24-29, Oct. 2004.

[14] C. Steger, A. Khoshnevis, A. Sabharwal, and B. Aazhang, "The case for transmitter training," in *IEEE Int. Symp. Inform. Theory*, Seatle, Washington, USA, July 2006, pp. 35-39.

[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley and Sons, 2006.

[16] D. Guo, S. Shamai, and S. Verdú, "Mutual information and minimum mean-square error in Gaussian channels," *IEEE Trans. Inform. Theory*, vol. 51, no. 4, pp. 1261-1282, Apr. 2005.

[17] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Trans. Commun.*, vol. 40, no. 5, pp. 873-884, May 1992.

[18] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 927-946, May 1998.

[19] A. Guillén i Fàbregas and A. Martinez, "Derivative of BICM mutual information," *IET Electronics Letters*, vol. 43, no. 22, pp. 1219-1220, Oct. 2007.

[20] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[21] K. D. Nguyen, A. Guillén i Fàbregas, and L. K. Rasmussen, "A tight lower bound to the outage probability of block-fading channels," *IEEE Trans. Inform. Theory*, vol. 53, no. 11, pp. 4314-4322, Nov. 2007.

[22] F. Brännström, L. K. Rasmussen, and A. J. Grant, "Convergence analysis and optimal scheduling for multiple concatenated codes," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3354-3364, Sept. 2005.

[23] S. V. Hanly and D. N. C. Tse, "Multiaccess fading channels-Part II: Delay-limited capacities," *IEEE Trans. Inform. Theory*, vol. 44, no. 7, pp. 2816-2831, Nov. 1998.

PLACE
PHOTO
HERE

Khoa Nguyen (S' 06) received the B.E. in Electrical Engineering at the University of Melbourne, Victoria, Australia in 2005. Since 2006, he has been pursuing the Ph.D. degree in electrical engineering at the Institute for Telecommunications Research, University of South Australia, Australia. His research interests are in the areas of information theory and error control codes. He has mainly worked on adaptive techniques and theoretical limits of communications with practical constraints.