

Research Article

Practical Radio Link Resource Allocation for Fair QoS-Provision on OFDMA Downlink with Partial Channel-State Information

Ayman Alsawah and Inbar Fijalkow

ETIS, CNRS, ENSEA, University Cergy-Pontoise, 6 Avenue du Ponceau, 95000 Cergy-Pontoise, France

Correspondence should be addressed to Ayman Alsawah, ayman.alsawah@ensea.fr

Received 13 February 2008; Revised 6 June 2008; Accepted 22 July 2008

Recommended by Shinsuke Hara

We address the problem of resource allocation on the downlink of an OFDMA single-cell system under fairness constraints with limited channel state information (CSI). Target QoS corresponds to a minimum user data rate, a target bit-error rate and a maximum BER-outage probability. The channel model includes path-loss, shadowing, and fading. The only available CSI is the channel average gain of each user. This partial CSI defines a *shadowed path-loss* that yields a modified user distribution. Resource allocation is based on the *shadowed user distribution* that we characterize analytically. Thus, under the target QoS, we provide the optimal resource allocation that maximizes the user rate. Compared to full-CSI-based allocation schemes, our solution offers a significant complexity and feedback reduction. Finally, the performance of our method is compared to other existing methods and the robustness of its outage performance to CSI errors is shown.

Copyright © 2008 A. Alsawah and I. Fijalkow. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

In cellular systems, service providers are interested in offering a given quality of service (QoS) to all users regardless of their locations in the cell. This leads to the challenging problem of resource allocation under fairness constraints. In this paper, we focus on the case of an OFDMA downlink. OFDMA is a promising modulation and multiple-access technique that has been adopted in key standards like 802.16d/e (WiMax) [1, 2].

The problem of resource allocation in OFDMA under fairness constraints continues to be an active research area [3–7]. Depending on the target application, the proposed solutions differ in the fairness requirement and in the adopted objective function. However, the common point is that some available *channel state information* (CSI) is used by the base station in order to maximize the objective function. The optimization operates on some degrees of freedom like subcarrier, rate, and power allocation schemes. Obviously, the ultimate optimal performance is obtained if the base station jointly optimizes the available degrees of freedom while having an instantaneous full CSI knowledge. Full CSI allows the system to exploit the different forms of diversity like time,

frequency, or multiuser diversity. In practice, this knowledge is very expensive in terms of the bandwidth required on feedback channels. Moreover, even with full and perfect CSI, finding the optimal solution usually involves prohibitive computational complexity. These drawbacks usually make the proposed solutions unsuitable for practical use in real systems. Therefore, the goal of the present work is to propose a low-complexity resource allocation algorithm assuming partial and imperfect CSI.

This paper focuses on the downlink of a single-cell OFDMA system with adaptive modulation and subcarrier allocation. Fairness constraints correspond to a target QoS common to all the users in the cell. This target QoS is defined by a minimum data rate and a target bit-error rate (BER). In addition to the path-loss, the channel model includes log-normal shadowing and Rayleigh fading. We suppose that the only CSI available to the base station is the channel average power gain of each user. This partial CSI can be estimated by averaging the received power over the different subcarriers assuming equal subcarriers' powers. It can be considered as a shadowed path-loss that corresponds to modified users' distances that we call *shadowed distances*. Resource allocation is carried out by the base station based on the notion of

user shadowed-distance that we introduce and characterize analytically. The reason why the shadowed distance is considered rather than the channel average power-gain is that the resource allocation algorithm that we propose is closely related to user distribution. This simple CSI results in a feedback overhead reduction factor equal to the number of subcarriers, that is, between 64 and 1024, compared to full-CSI required overhead. Different methods for CSI feedback reduction were proposed in literature [8–17]. Some [10–12] are based on optimizing CSI quantization. Others [15] exploit available channel correlation properties coupled with channel prediction. Our shadowed-distance-based approach belongs to the family of methods using channel statistics and system load knowledge [17]. By exploiting the assumption of uniform user distribution a set of interesting analytic results as well as a practical resource allocation algorithm are provided in this work.

Thus, under the target QoS, a total peak power constraint, and a given number of users uniformly distributed over the cell of a given radius, we provide the optimal subcarrier and rate allocation that offers the maximum data rate per user. The presence of unknown fading leads to an outage each time the BER achieved by a given user exceeds the target BER. Therefore, a maximum BER-outage probability constraint is added to the target QoS specification. This BER-outage constraint requires considering a fading power margin. This means that a portion of the total power is dedicated to fading mitigation and that resource allocation is based on the remaining power. Moreover, due to shadowing, some users may fall, in terms of shadowed distance, out of the coverage range. This corresponds to a rate-outage probability that we characterize as well.

Simulation results allow us to evaluate the achieved performance for a typical parameter setting. They show that our algorithm yields a significant spectral efficiency enhancement compared to a traditional static resource allocation. Meanwhile, the loss in terms of average spectral efficiency with respect to a full-CSI-based opportunistic allocation remains acceptable. This loss is counterweighted by the complexity and feedback overhead reduction offered by our approach. Finally, simulations reveal that the overall outage performance is very robust to CSI estimation errors.

The remaining of this paper is organized as follows. The next section describes the system model and introduces the notion of shadowed distance. Section 3 is devoted to shadowed-distance estimation issue. Then, the considered optimization problem is stated in Section 4 where a brief state-of-the-art is provided as well. The main analytic results on optimal resource allocation are derived in Section 5. Some additional equations characterizing the average achieved performance are given in Section 6. Section 7 focuses on practical aspects of the resource allocation algorithm. Simulation results are presented in Section 8 before the concluding remarks of Section 9.

2. SYSTEM MODEL

We consider an uncoded OFDMA downlink from a base station to U uniformly-distributed users in a single circular

cell of radius R . A total peak power P_{tot} is available for transmission over S subcarriers. The transmitted signal received by user u experiences a frequency-selective slow-fading channel characterized by S identically distributed random variables $g_{u,s}$ ($s = 1, \dots, S$). These $g_{u,s}$ represent the channel power gains over the different subcarriers. Channel realizations are independent from a user to another. Each coefficient $g_{u,s}$ accounts for a deterministic propagation path-loss $G(x_u)$ that depends on the distance x_u of user u to the base station, in addition to a *log-normal shadowing* $10^{0.1\xi_u}$ and to a multipath *squared-Rayleigh power fading* $\phi_{u,s}^2$ with $E[\phi_{u,s}^2] = 1$. So, if $p_{u,s}$ denotes the transmitted power on subcarrier s , the received *signal-to-noise ratio* (SNR) at the u th user on this subcarrier is given by

$$\gamma_{u,s} = \frac{p_{u,s}g_{u,s}}{BN_0} = \frac{p_{u,s}}{BN_0}G(x_u)10^{0.1\xi_u}\phi_{u,s}^2, \quad (1)$$

where B is the subcarrier spacing and N_0 is the AWGN power spectral density. The path-loss $G(x_u)$ represents the long-term average, called also the *area mean*, of the channel power gain at distance x_u . Thus, the quantity

$$\bar{\bar{\gamma}}_{u,s} = \frac{p_{u,s}}{BN_0}G(x_u) \quad (2)$$

is the *area-mean received SNR* at user u on subcarrier s . The double bar in $\bar{\bar{\gamma}}_{u,s}$ corresponds to the expectation of the instantaneous SNR with respect to both shadowing and fading ($\bar{\bar{\gamma}}_{u,s} = E_{\xi,\phi}[\gamma_{u,s}]$).

We assume that the logarithmic path-loss $G_{\text{dB}}(x_u) = -10 \log_{10}G(x_u)$ follows the *exponent model* [18] defined by

$$G_{\text{dB}}(x_u) = G_{\text{dB}}(x_0) + 10\alpha \log_{10} \frac{x_u}{x_0}. \quad (3)$$

The term $G_{\text{dB}}(x_0)$ is the path-loss at a *reference distance* x_0 while $\alpha \geq 2$ is the *path-loss exponent*. The value of α depends on the terrain nature and on the base station antenna height [18]. We have

$$G_{\text{dB}}(x_0) = 20 \log_{10} \frac{4\pi f x_0}{c}, \quad (4)$$

where f is the frequency of the considered subcarrier and c is the light speed. By taking a unit reference distance ($x_0 = 1$), the path-loss can be written simply as follows:

$$G(x_u) = \frac{G_0}{x_u^\alpha} \quad (5)$$

with $G_0 = (c/4\pi f)^2$. Since the OFDM total bandwidth SB is typically small compared to the center frequency f_c , we consider that G_0 is independent from the subcarrier index. This parameter is usually calculated using the high end of the bandwidth, that is, at frequency $f_c + SB/2$ (worst case value).

In (1), the log-normal shadowing $10^{0.1\xi_u}$ is the mid-term average gain called the *local-mean* gain. It is characterized by the *shadowing standard deviation* σ which is the standard deviation of the zero-mean Gaussian random variable ξ_u . We

define the *local-mean received SNR* (or the shadowed SNR) by

$$\bar{\gamma}_{u,s} = \frac{p_{u,s}}{BN_0} G(x_u) 10^{0.1\xi_u}. \quad (6)$$

The single bar in $\bar{\gamma}_{u,s}$ is related to the expectation of the instantaneous SNR $\gamma_{u,s}$ with respect to the fading process only ($\bar{\gamma}_{u,s} = E_\phi[\gamma_{u,s}]$). This shadowed SNR follows a log-normal distribution $\mathcal{LN}(\mu_{u,s}, \sigma^2)$ where the logarithmic average is simply the area-mean received SNR (2) in decibels $\mu_{u,s} = 10 \log_{10} \bar{\gamma}_{u,s}$.

Now we define the *shadowed path-loss* for user u by

$$\mathcal{G}_u = G(x_u) 10^{0.1\xi_u} = \frac{G_0}{x_u^\alpha} 10^{0.1\xi_u}. \quad (7)$$

The reason for calling \mathcal{G}_u the *shadowed path-loss* is that a *user equivalent-distance* d_u can be introduced so that \mathcal{G}_u is obtained by the same path-loss model (5) as follows:

$$\mathcal{G}_u = \frac{G_0}{d_u^\alpha}. \quad (8)$$

Thus, from (7) and (8) we get

$$d_u = x_u 10^{-0.1\xi_u/\alpha}. \quad (9)$$

In the sequel, we call d_u the *shadowed distance* of user u . Shadowed-distance estimation issue is considered in Section 3.

So, replacing user actual distance x_u by the shadowed distance d_u allows us to make abstraction of the shadowing as if the random component of the channel is reduced to fading. Obviously, the uniform user distribution is transformed into a different one (see Figure 2) that we characterize farther in our analysis. In the following, we consider that the available CSI about user u is the corresponding shadowed distance d_u defined in (9). Based on this CSI, the base station suitably allocates system resources (subcarriers, powers and rates) in order to offer the maximum common rate while satisfying the minimum required QoS. The minimum QoS is defined by the minimum data rate per user D_0 , the target BER b and the maximum BER-outage probability ε . The resource allocation is done on a frame-basis with a given frame length L_f expressed in OFDM symbols. We consider that the channel variation rate is small compared to the frame rate so that the CSI is invariant during the same frame (slow fading assumption).

Finally, we assume that the users terminals demodulate each subcarrier using a coherent receiver with a Nyquist-matched filter. This means that the subcarrier spacing B is equal to the symbol rate. Consequently, the quantity $\log_2 M$ represents the spectral efficiency in bps/Hz of the M -QAM constellation. Throughout this paper, the term M -QAM is also used for the BPSK case where $M = 2$. We suppose that the set $\mathcal{M} = \{M_1, M_2, \dots, M_Q\}$ of available constellations' orders is an ordered set, that is, $M_1 > M_2 > \dots > M_Q$ where $Q = \text{card}(\mathcal{M})$. Table 1 summarizes the main notations.

Before formulating the optimization problem in Section 4, we provide in the following section some details about the shadowed-distance estimation.

3. SHADOWED-DISTANCE ESTIMATION

The shadowed distance defined in (9) can be estimated as follows. Assume that a dedicated OFDM symbol with known and equal subcarriers' powers $p_{u,s} = p_0$ is received by user u so that the later is able to measure the instantaneous received SNR $\gamma_{u,s}$ on each subcarrier. Remember that the shadowed SNR $\bar{\gamma}_{u,s}$ defined in (6) is the expectation of $\gamma_{u,s}$ with respect to the power fading $\phi_{u,s}^2$, that is,

$$\gamma_{u,s} = \bar{\gamma}_{u,s} \phi_{u,s}^2. \quad (10)$$

With equal subcarriers' powers p_0 , using (8) and (9), the shadowed SNR becomes

$$\bar{\gamma}_{u,s} = \frac{p_0 G_0}{BN_0 d_u^\alpha}. \quad (11)$$

From (10) and (11) we get

$$\gamma_{u,s} = \frac{p_0 G_0}{BN_0 d_u^\alpha} \phi_{u,s}^2. \quad (12)$$

By averaging $\gamma_{u,s}$ in (12) over the S subcarriers we obtain

$$\frac{1}{S} \sum_{s=1}^S \gamma_{u,s} = \frac{p_0 G_0}{BN_0 d_u^\alpha} \frac{1}{S} \sum_{s=1}^S \phi_{u,s}^2. \quad (13)$$

By considering that the sequence $\phi_{u,1}^2, \dots, \phi_{u,S}^2$ forms a realization of the ergodic random process $\phi_{u,s}^2$, we can approximate $(1/S) \sum_{s=1}^S \phi_{u,s}^2$ in (13) by the expectation $E[\phi_{u,s}^2] = 1$ (this is a good approximation since $S \gg 1$). Thus, given an estimator \hat{N}_0 of the noise spectral density and knowing the values of constants α and G_0 , we obtain from (13) the following estimator of the shadowed distance:

$$\hat{d}_u = \left(\frac{p_0 G_0 S}{B \hat{N}_0 \sum_{s=1}^S \gamma_{u,s}} \right)^{1/\alpha}. \quad (14)$$

From (13) and (14) we can derive the following relationship between the actual shadowed distance and its estimator:

$$\hat{d}_u = \left(\frac{N_0 S}{\hat{N}_0 \sum_{s=1}^S \phi_{u,s}^2} \right)^{1/\alpha} d_u. \quad (15)$$

With perfect noise power estimation ($\hat{N}_0 = N_0$) and sufficiently large number of subcarriers so that $(1/S) \sum_{s=1}^S \phi_{u,s}^2 = 1$, we get $\hat{d}_u = d_u$. In the following we suppose that the shadowed distances are perfectly estimated by users and relayed to the BS over an error-free feedback channel.

In Section 4 we formulate the considered optimization problem with arbitrary CSI in order to compare our approach based on partial CSI to other existing full-CSI-based approaches.

4. PROBLEM FORMULATION AND RELATED WORK

Assume that the base station allocates subcarrier s to user u with power $p_{u,s}$, so that this user achieves a data rate $r_{u,s}$ on

TABLE 1: Main notations

f_c	Carrier frequency	Hz
S	Number of subcarriers	—
B	Subcarrier spacing (\Rightarrow total OFDM bandwidth = SB)	Hz
\mathcal{M}	Set of available modulation orders (ex. $\mathcal{M} = \{4, 2\} \Leftrightarrow \{\text{QPSK}, \text{BPSK}\}$)	—
P_{tot}	Maximum total transmit power	W
$p_{u,s}$	Power allocated to user u on subcarrier s	W
x_u	Distance between the u th user and the base station	m
N_0	White noise power spectral density	W/Hz
α	Path-loss exponent	—
σ	Log-normal shadowing logarithmic standard deviation	dB
T_s	OFDM symbol duration	sec
L_f	Frame length	symbols
D_0	Target minimum data rate per user	bps
b	Target bit-error rate	—
ε	Target maximum BER-outage probability	—
U	Total number of users	users
R	Cell radius	m

that subcarrier. This rate depends on the target BER b , on the maximum outage probability ε , on the allocated power $p_{u,s}$, and on the available CSI denoted as $h_{u,s}$. This $h_{u,s}$ may represent the actual channel gain $g_{u,s}$, in the case of full CSI, or any other information derived from $g_{u,s}$ in the case of partial or imperfect CSI. Let us describe the rate adaptation by a function Φ as follows:

$$r_{u,s} = \Phi(b, \varepsilon, p_{u,s}, h_{u,s}). \quad (16)$$

The subcarrier allocation can be described by the matrix $A = [a_{u,s}]$ where $a_{u,s} = 1$ if subcarrier s is allocated to user u and $a_{u,s} = 0$ otherwise. Hence, user u achieves a total data rate

$$r_u = \sum_{s=1}^S a_{u,s} \Phi(b, \varepsilon, p_{u,s}, h_{u,s}). \quad (17)$$

Let $P = [p_{u,s}]$ be the power allocation matrix. Maximizing the common data rate for a given number of users U can be written as a problem of sum-rate maximization under fairness constraints as follows:

$$\begin{aligned} & \max_{A,P} \sum_{u=1}^U \sum_{s=1}^S a_{u,s} \Phi(b, \varepsilon, p_{u,s}, h_{u,s}) \\ \text{subject to} & \quad [\text{c1}] : a_{u,s} \in \{0, 1\} \quad \forall u, s, \\ & \quad [\text{c2}] : a_{u,s} a_{u',s} = 0 \quad \forall s, \forall u \neq u', \\ & \quad [\text{c3}] : \sum_{u=1}^U \sum_{s=1}^S a_{u,s} p_{u,s} \leq P_{\text{tot}}, \\ & \quad [\text{c4}] : r_1 = r_2 = \dots = r_U, \\ & \quad [\text{c5}] : r_u \geq D_0 \quad \forall u. \end{aligned} \quad (18)$$

In this formulation, [c3] represents the total power constraint. Fairness is defined by [c4] while [c5] corresponds to a minimum rate requirement. Finally, constraints [c1] and [c2] mean that during each OFDM symbol, a given

subcarrier can be allocated to at most one user. This condition is very stringent and may render the problem unfeasible because of the discrete nature of variables $a_{u,s}$. If this condition is relaxed, we obtain the following problem where all the variables are continuous:

$$\begin{aligned} & \max_{A,P} \sum_{u=1}^U \sum_{s=1}^S a_{u,s} \Phi(b, \varepsilon, p_{u,s}, h_{u,s}) \\ \text{subject to} & \quad [\text{c1}] : a_{u,s} \in [0, 1] \quad \forall u, s, \\ & \quad [\text{c2}] : \sum_{u=1}^U a_{u,s} \leq 1 \quad \forall s, \\ & \quad [\text{c3}] : \sum_{u=1}^U \sum_{s=1}^S a_{u,s} p_{u,s} \leq P_{\text{tot}}, \\ & \quad [\text{c4}] : r_1 = r_2 = \dots = r_U, \\ & \quad [\text{c5}] : r_u \geq D_0 \quad \forall u. \end{aligned} \quad (19)$$

Allowing variables $a_{u,s}$ to take real values in the range $[0, 1]$ does not necessarily mean that the exclusive subcarrier assignment is violated. In fact, sharing the same subcarrier can be carried out by time-division over several OFDM symbols assuming that the channel state remains invariant meanwhile. For example, if the optimal solution of (19) gives $a_{1,s} = a_{2,s} = 0.5$, then subcarrier s has to be shared by both user 1 and user 2 during 50% of the time each.

Let us go back to (16) that describes the rate adaptation. Without any CSI, one must allocate subcarriers and powers in a static way based on a worst-case design (adaptation to edge-user condition). With full CSI, where $h_{u,s}$ represents the actual channel power-gain $g_{u,s}$ in (1), a zero-outage ($\varepsilon = 0$) can be achieved since the allocated powers can compensate for the actual channel attenuation. This allows the system to get benefit of the frequency diversity and to achieve higher data rates. In literature, authors often consider the case of full CSI with continuous modulation and zero-outage [3–7]. Some authors [3, 6, 7] consider the error-

free Shannon capacity so that $r_{u,s} = B \log_2(1 + \gamma_{u,s})$. Such information-theoretical approaches provide upper bounds on the achievable performance but are not directly connected to real implementations. Other authors [4, 5] consider the case of a positive bit error rate ($b > 0$) with M -QAM constellations. In this case, a BER approximation for M -QAM performance is used. The SNR gap [19], used in [4], is an example of a BER approximation in the uncoded M -QAM case. Another example for encoded M -QAM can be found in [20].

In all cases, finding the optimal solution to (19) involves excessive computational complexity [7]. Hence, several suboptimal solutions have been proposed. Some are based on separating the subcarrier allocation step from the power allocation step as in [5]. Others assume equal powers to simplify the subcarrier assignment step as in [7]. Nevertheless, the assumption of full CSI remains unrealistic in practical systems especially for a large number of users. This is due to the prohibitive overhead required for CSI feedback. Moreover, in real systems, whatever the CSI estimation method is, the obtained CSI is imperfect due to estimation errors, feedback errors, feedback delay, and quantization noise. Thus, the aim of this work is to propose a low-complexity suboptimal solution based on partial and imperfect CSI so that the resulting algorithm can be implemented in real systems with acceptable performance.

In OFDMA-based systems like WiMax, only a reduced number of bits per frame is dedicated to CSI feedback [9] on the uplink. Thus, the CSI must be efficiently quantized and the feedback amount of the resulting quantized CSI must be as small as possible in order to limit the feedback rate and capture the channel variations. In literature, several feedback reduction techniques are proposed (e.g., [8–17] and references therein). CSI quantization on a reduced number of bits is shown [10] to have a marginal effect on the achieved capacity compared to full-CSI case even under some fairness constraints. It was reported [11] that a single-bit feedback per user per subcarrier is able to capture the double-logarithmic capacity growth with the number of users. This assumes an optimized SNR-thresholding [12] on each subcarrier. Further feedback reduction can be achieved by grouping adjacent subcarriers into clusters [13] and reporting one value per cluster (the average or the worst-case SNR). Another partial CSI approach [14] consists in reporting only the k -best subcarriers CSI ($k = 1, 2, \dots$). This technique becomes particularly interesting when it is coupled with opportunistic user scheduling [15]. In [16] the authors introduce the notion of *selective multuser diversity* based on eliminating from the feedback pool those users who have no chance to be scheduled. In time and/or frequency correlated channels the available CSI at the receiver (user-side) can be undersampled to reduce the feedback rate and then the transmitter (base-station-side) uses interpolation to retrieve the missing values [15]. Finally, the CSI quantization size and thresholds can also be optimized based on known channel statistics and system load (active number of users) instead of instantaneous SNRs [17].

In this paper, we suppose that the only available CSI is the users shadowed distances defined in (9). Furthermore, we

show later that even a coarse CSI estimation can be sufficient as the obtained performance is shown to be very robust against CSI estimation errors. The reason why the shadowed distance d_u is considered rather than the shadowed path-loss (7) is that the resource allocation algorithm that we propose is based on user distribution.

To reduce the complexity of the proposed solution, we adopt the *equal-power scheme*, that is

$$p_{u,s} = \frac{P_{\text{tot}}}{S} \quad \forall u, s. \quad (20)$$

With equal powers, (19) is reduced to a subcarrier and rate allocation problem. Moreover, since the shadowed distance d_u is not a frequency-selective information, the subcarrier allocation is transformed into a bandwidth allocation. Thus, resolving (19) consists in deciding how many subcarriers each user does need and which M -QAM constellation has to be used on these subcarriers. Let us replace the subcarrier assignment matrix A by the vector $W = [W(1), \dots, W(U)]$ where $W(u)$ represents the number of subcarriers allocated to user u , that is,

$$W(u) = \sum_{s=1}^S a_{u,s}. \quad (21)$$

With the equal power scheme (20), we obtain from ((19).c3) and (21) the following condition:

$$\sum_{u=1}^U W(u) \leq S. \quad (22)$$

We assume also that the same constellation of order $M(u)$ is used on the $W(u)$ subcarriers of user u . With uncoded M -QAM, the rate adaptation function (16) becomes

$$\Phi\left(b, \varepsilon, \frac{P_{\text{tot}}}{S}, d_u\right) = B \log_2 M(u). \quad (23)$$

Notice that $\Phi(b, \varepsilon, P_{\text{tot}}/S, d_u)$ is now independent from the subcarrier index s so that, from (17) and (21), the data rate achieved by user u is

$$\begin{aligned} r_u &= W(u) \Phi\left(b, \varepsilon, \frac{P_{\text{tot}}}{S}, d_u\right) \\ &= W(u) B \log_2 M(u). \end{aligned} \quad (24)$$

So, under these assumptions, the optimization problem (19) becomes

$$\begin{aligned} &\max_W \sum_{u=1}^U W(u) \Phi\left(b, \varepsilon, \frac{P_{\text{tot}}}{S}, d_u\right) \\ &\text{subject to} \quad [\text{c1}] : W(u) \in]0, S[\quad \forall u, \\ &\quad \quad \quad [\text{c2}] : \sum_{u=1}^U W(u) \leq S, \\ &\quad \quad \quad [\text{c3}] : r_1 = r_2 = \dots = r_U, \\ &\quad \quad \quad [\text{c4}] : r_u \geq D_0 \quad \forall u. \end{aligned} \quad (25)$$

Note that $W(u)$ in (25) is allowed to take real values. This can be achieved in practice by time-multiplexing as explained later in Section 7. If we denote by D the common data rate, then (25) can be transformed into a common rate maximization problem as follows:

$$\begin{aligned} & \max_W D \\ & \text{subject to} \quad [c1] : W(u) \in]0, S[\quad \forall u, \\ & \quad [c2] : \sum_{u=1}^U W(u) \leq S, \\ & \quad [c3] : W(u) \Phi\left(b, \varepsilon, \frac{P_{\text{tot}}}{S}, d_u\right) = D \quad \forall u, \\ & \quad [c4] : D \geq D_0. \end{aligned} \quad (26)$$

The rate adaptation, defined by the function $\Phi(b, \varepsilon, P_{\text{tot}}/S, d_u)$ or equivalently by $M(u)$, must ensure the target BER b for all users with probability $1 - \varepsilon$ at least. The function $M(u)$ is derived in Section 5 as well as the optimal subcarrier allocation that corresponds to the maximum common rate D_{max} .

5. OPTIMAL RESOURCE ALLOCATION

In this section, we resolve the optimization problem (26) formulated in Section 4.

5.1. Rate allocation

Since the partial CSI d_u does not depend on the subcarrier index, subcarrier allocation consists of deciding how many subcarriers each user does need in order to achieve the common data rate D that we want to maximize. Obviously, this depends on which M -QAM constellation is used on the subcarriers allocated to the user of interest. Maximizing the common data rate corresponds to the case where the constellation of the highest possible order is chosen on each subcarrier. Unfortunately, the choice of constellation is subject to the BER-outage constraint. Let us find the expression of the outage probability $p_{\text{out}}(u)$ for user u as a function of the target BER b , the distance d_u , and the chosen constellation order $M(u)$. We have

$$p_{\text{out}}(u) = \text{Proba}[\beta_{M(u)}(\gamma_{u,s}) > b], \quad (27)$$

where $\beta_{M(u)}(\gamma)$ is the function describing the achieved BER versus the SNR γ and the modulation order $M(u)$. Since this BER function is decreasing with respect to the SNR argument γ , we have

$$p_{\text{out}}(u) = \text{Proba}[\gamma_{u,s} < \beta_{M(u)}^{-1}(b)] \quad (28)$$

with $\beta_{M(u)}^{-1}(\cdot)$ being the inverse function of $\beta_{M(u)}(\cdot)$ that provides the minimum SNR required by an M -QAM constellation to achieve a given BER. This outage probability can be expressed using the *cumulative distribution function* $F_{\gamma_{u,s}}(\cdot)$ of the instantaneous SNR $\gamma_{u,s}$ which, conditionally

to the shadowed SNR (6), follows a chi-square law so that $F_{\gamma_{u,s}}(\gamma) = 1 - \exp(\gamma/\bar{\gamma}_{u,s})$. Thus we get

$$\begin{aligned} p_{\text{out}}(u) &= F_{\gamma_{u,s}}(\beta_{M(u)}^{-1}(b)) \\ &= 1 - \exp\left(\frac{-\beta_{M(u)}^{-1}(b)}{\bar{\gamma}_{u,s}}\right). \end{aligned} \quad (29)$$

Using the equal-power scheme (20) along with (8) and (7), the shadowed SNR (6) becomes

$$\bar{\gamma}_{u,s} = \frac{P_{\text{tot}}G_0}{SBN_0d_u^\alpha}. \quad (30)$$

Replacing (30) into (29) provides

$$p_{\text{out}}(u) = 1 - \exp\left(-\frac{\beta_{M(u)}^{-1}(b)SBN_0d_u^\alpha}{P_{\text{tot}}G_0}\right). \quad (31)$$

For a given value of $M(u)$, this outage probability increases with distance d_u . Therefore, the maximum outage probability constraint

$$p_{\text{out}}(u) \leq \varepsilon \quad (32)$$

means that each constellation M_q -QAM ($q = 1, \dots, Q$) can be used up to a maximum distance R_q which is the solution of

$$1 - \exp\left(-\frac{\beta_{M_q}^{-1}(b)SBN_0R_q^\alpha}{P_{\text{tot}}G_0}\right) = \varepsilon. \quad (33)$$

We obtain

$$R_q = \left[\frac{(P_{\text{tot}}/F)G_0}{SBN_0\beta_{M_q}^{-1}(b)}\right]^{1/\alpha} \quad (34)$$

with the parameter F given by

$$F = \frac{-1}{\log(1 - \varepsilon)}. \quad (35)$$

This parameter F represents the *fading power margin* that guarantees a BER-outage probability bounded by ε .

In (34), the SNR-threshold function $\beta_{M_q}^{-1}(b)$ is increasing with M_q (a higher-order modulation requires higher SNR to achieve the same BER). So, for the complete set of available constellations we have $R_1 < R_2 < \dots < R_Q$. Remember that maximizing the common rate requires using for each user the constellation of the highest possible order. Consequently, the M_q -QAM constellation must be allocated to users whose shadowed distances d_u are in $]R_{q-1}, R_q]$ with $R_0 = 0$. This defines the optimal rate allocation as follows:

$$M(u) = \max_q \left\{ M_q \in \mathcal{M} : d_u \leq R_q = \left[\frac{(P_{\text{tot}}/F)G_0}{SBN_0\beta_{M_q}^{-1}(b)}\right]^{1/\alpha} \right\}. \quad (36)$$

Thus, each constellation M_q -QAM covers an *annular zone* of internal (resp., external) radius R_{q-1} (resp., R_q). This is depicted in Figure 1 in the case of three modulations ($Q = 3$). In the following, the q th annular zone is referred to as *zone* q .

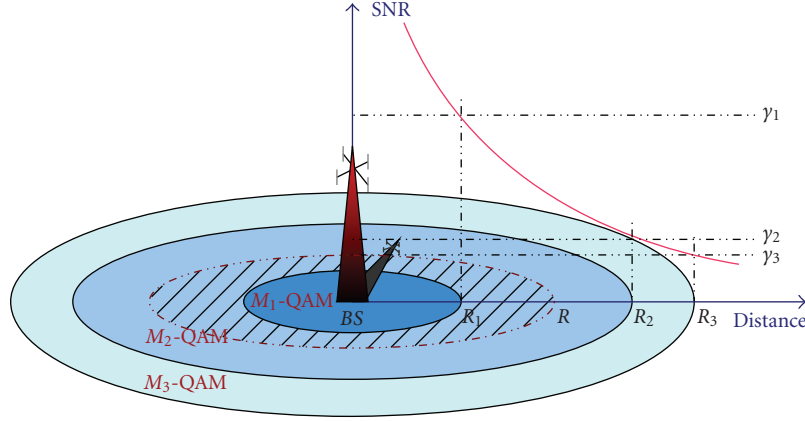


FIGURE 1: Modulation zones and SNR thresholds ($Q = 3$ modulations).

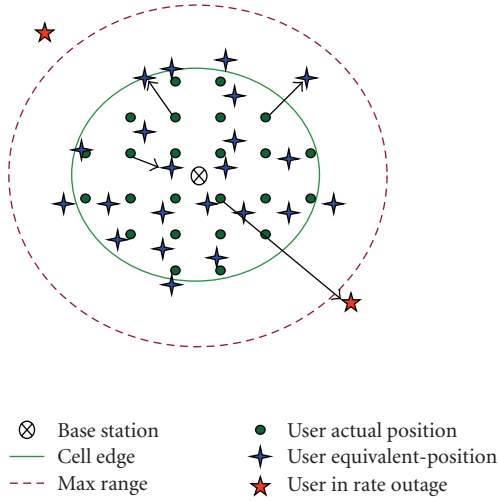


FIGURE 2: Initial uniform and shadowed user distribution.

5.2. Rate-outage probability

The coverage range corresponds to R_Q , the range of the lowest-order modulation. Although users are actually located inside the circle of radius R , their shadowed distances may be greater than R or even than the maximum range R_Q (see Figure 2). As long as the user shadowed distance is smaller than R_Q , this user can be served with the required QoS. In the opposite case, this user is said to be in *rate-outage*. This corresponds to the situation where the base station decides to use all the available constellations and to serve all the users whose shadowed distances are smaller than R_Q . But, in order to improve the overall QoS of the provisioned users, the base station may decide not to serve those whose shadowed distances are greater than a given *cutoff distance* $R_{\text{cut}} \in [R, R_Q]$. Introducing the rate-outage corresponds to a kind of relaxation of the rate-fairness constraint ((26).c3) and the minimum-rate constraint ((26).c4). This relaxation prevents users in bad channel conditions (large shadowed distances) to penalize the achieved rate of the remaining users by requiring excessive bandwidth. Thus, tolerating some rate-

outage, with bounded outage probability, is a pragmatic approach that allows us to improve the system-wide spectral efficiency.

We define the *rate-outage probability* for a user u by $\rho_u(R_{\text{cut}}) = \text{Proba}[d_u > R_{\text{cut}}]$. This probability can be derived given the statistics of the random variable d_u defined in (9). It is easier to consider the random variable $10 \log_{10} d_u = 10 \log_{10} x_u - \xi_u/\alpha$ which follows a Gaussian law $\mathcal{N}(10 \log_{10} x_u, \sigma^2/\alpha^2)$. The corresponding cumulative distribution function is

$$F_{10 \log_{10} d_u}(y) = 0.5 + 0.5 \operatorname{erf}\left(\frac{y - 10 \log_{10} x_u}{\sigma \sqrt{2}/\alpha}\right), \quad (37)$$

where $\operatorname{erf}(\cdot)$ is the *error function* defined by

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (38)$$

So, the rate-outage probability of user u becomes

$$\begin{aligned} \rho_u(R_{\text{cut}}) &= 1 - F_{10 \log_{10} d_u}(10 \log_{10} R_{\text{cut}}) \\ &= 0.5 - 0.5 \operatorname{erf}\left(\frac{10 \log_{10}(R_{\text{cut}}/x_u)}{\sigma \sqrt{2}/\alpha}\right). \end{aligned} \quad (39)$$

For a given R_{cut} , the rate-outage probability reaches its maximum for edge users $x_u = R$,

$$\rho_{\max}(R_{\text{cut}}) = 0.5 - 0.5 \operatorname{erf}\left(\frac{10 \log_{10}(R_{\text{cut}}/R)}{\sigma \sqrt{2}/\alpha}\right). \quad (40)$$

This worst-case rate-outage probability is minimized if the base station decides to serve the maximum number of users by setting R_{cut} to the lowest-order modulation range R_Q . From (34), we have

$$R_Q = \left[\frac{(P_{\text{tot}}/F) G_0}{SBN_0 \beta_{M_q}^{-1}(b)} \right]^{1/\alpha}. \quad (41)$$

We suppose that this maximum range is greater than the cell radius R . So, from (40) and (41), the minimum value of the worst-case (edge-user) rate-outage probability is given by

$$\min \rho_{\max} = 0.5 - 0.5 \operatorname{erf} \left(\frac{10}{\sigma\sqrt{2}} \log_{10} \left(\frac{SBN_0\beta_{M_Q}^{-1}(b)R_{\text{cut}}^\alpha}{(P_{\text{tot}}/F)G_0} \right) \right). \quad (42)$$

Adjusting the maximum rate-outage probability (40) through the parameter R_{cut} allows the base station to find a satisfying tradeoff between the average number of served users and the QoS that can be offered to them. In fact, extending the service range by increasing R_{cut} means that the spectral resource (subcarriers) has to be shared among a greater number of users.

5.3. Subcarrier allocation

Let U_q be the number of users who, due to the shadowing, seem to fall inside the q th modulation zone of internal (resp., external) radius R_{q-1} (resp. R_q), that is,

$$U_q = \operatorname{card}\{u = 1, \dots, U : R_{q-1} < x_u 10^{-0.1\xi_u/\alpha} \leq R_q\}. \quad (43)$$

Zones' radii are given in (34). The quantity U_q is a random variable. This means that the user partition over the modulation zones varies each time a new shadowing realization occurs. We saw earlier that the subcarrier allocation under our assumptions is reduced to a user-wise bandwidth allocation defined by the vector $W = [W(1), \dots, W(U)]$. Suppose that users achieve an *unknown* data rate D_{\max} each ($r_u = D_{\max}, \forall u$). From (24) we get

$$W(u) = \frac{D_{\max}}{(B \log_2 M(u))}. \quad (44)$$

According to (36), the optimal constellation order $M(u)$ is equal to M_q for all users in zone q . Consequently, each user in zone q needs $D_{\max}/(B \log_2 M_q)$ subcarriers. The number of users contained in zone q is U_q defined in (43). Thus, the number of subcarriers required by zone q in order to satisfy its U_q users is

$$S_q = \frac{D_{\max} U_q}{B \log_2 M_q}. \quad (45)$$

The issue now is to determine the required number of modulation zones, denoted by $Z \leq Q$, as a function of the cutoff distance $R_{\text{cut}} \in [R, R_Q]$. We have

$$Z = \min\{q \in \{1, \dots, Q\} : R_{\text{cut}} \leq R_q\}. \quad (46)$$

So, from the constraint of the total number of subcarriers

$$\sum_{q=1}^Z S_q = S \quad (47)$$

and from (45), it follows that

$$D_{\max} = \frac{BS}{\sum_{q=1}^Z U_q / \log_2 M_q} \quad (48)$$

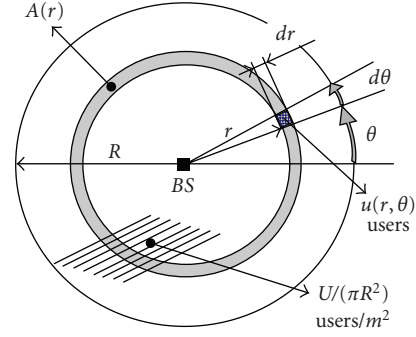


FIGURE 3: Differential annular zone $\mathcal{A}(r)$.

which is the maximum common rate. By substituting D_{\max} from (48) back into (45) we get the following expression of the *optimal zone-wise subcarrier allocation*:

$$S_q = \frac{U_q / \log_2 M_q}{\sum_{k=1}^Z U_k / \log_2 M_k} S. \quad (49)$$

Note that S_q is a random variable because U_q is. This means that the zone-wise bandwidth reservation depends on the shadowing realization. In other words, the base station must update the variables S_q 's each time it gets a new feedback about the users shadowed distances. The maximum user rate D_{\max} does change accordingly as well.

Now, the aim of what follows is to evaluate the achieved average performance in terms of rate-outage, user rate, spectral efficiency, and bearable number of users.

6. AVERAGE ACHIEVED PERFORMANCE

In order to evaluate the average performance of our method, we need to evaluate the expectation of the shadowed zone-wise number of users U_q defined in (43).

6.1. Average shadowed user distribution

Starting from the assumption of a deterministic uniform user distribution (actual users' positions), we want to derive the average user distribution when user distance x_u is replaced by the shadowed distance d_u defined in (9). Averaging is relative to the log-normal shadowing process.

Consider a differential annular zone centered at the base station (Figure 3) with an internal (resp., external) radius of r (resp., $r + dr$). We denote this annular zone by $\mathcal{A}(r)$. The area of $\mathcal{A}(r)$ is $\pi(r + dr)^2 - \pi r^2 \simeq 2\pi r dr$. When U users are uniformly distributed over the cell of radius R , we have a constant user density of $U/(\pi R^2)$. Thus, the annular zone $\mathcal{A}(r)$ contains $2Ur dr/R^2$ users. These users are uniformly distributed over 2π radians around the base station. This means that the number of users contained within the differential angular sector $[\theta, \theta + d\theta]$ radians at distances in $[r, r + dr]$ (see Figure 3) is equal to

$$u(r, \theta) = \frac{(2Ur/R^2) dr d\theta}{2\pi} = \frac{U}{\pi R^2} r dr d\theta. \quad (50)$$

Since the shadowing is closely related to the surrounding environment topology, we assume that these $u(r, \theta)$ users are subject to the same shadowing realization. In other words, they have the same shadowed distance $d(r, \theta) = r 10^{-0.1 \xi(r, \theta)/\alpha}$ where $\xi(r, \theta)$ follows a normal distribution $\mathcal{N}(0, \sigma^2)$. The cumulative distribution function of $d(r, \theta)$ is given by

$$\begin{aligned} F_{d(r, \theta)}(y) &= \text{Proba}[d(r, \theta) \leq y] \\ &= 0.5 + 0.5 \operatorname{erf}\left(\frac{10 \log_{10}(y/r)}{\sigma \sqrt{2}/\alpha}\right). \end{aligned} \quad (51)$$

Thus, the average ratio of users whose shadowed distances are in $]0, R_q]$ is

$$\bar{u}(R_q) = \frac{1}{U} \int_0^R \int_0^{2\pi} \text{Proba}[d(r, \theta) \leq R_q] u(r, \theta). \quad (52)$$

Using (50) and (51) we prove in the appendix that

$$\begin{aligned} \bar{u}(R_q) &= \frac{1}{2} \left[1 + \operatorname{erf}\left(C \log \frac{R_q}{R}\right) \right. \\ &\quad \left. + \frac{R_q^2}{R^2} e^{1/C^2} \left(1 - \operatorname{erf}\left(C \log \frac{R_q}{R} + \frac{1}{C}\right) \right) \right], \end{aligned} \quad (53)$$

where $C = 10\alpha/(\sigma\sqrt{2}\log 10)$. The product $\bar{u}(\cdot)U$ represents the average cumulative distribution function of the number of users when actual distances are replaced by the shadowed distances. It is the expectation, with respect to the shadowing random process, of U_q defined in (43). This important theoretical result is confirmed by simulation in Section 8.

6.2. Rate-outage average performance

Thanks to (53), we can now evaluate \bar{U}_{out} , the average number of users that can never be served (users in rate outage) because their shadowed distances exceed the range R_Q , given in (41), of the lowest-order modulation. We have

$$\bar{U}_{\text{out}}(R_Q) = (1 - \bar{u}(R_Q))U. \quad (54)$$

Moreover, the average number of users in rate outage versus an arbitrary cutoff distance $R_{\text{cut}} \in [R, R_Q]$ is given by

$$\bar{U}_{\text{out}}(R_{\text{cut}}) = (1 - \bar{u}(R_{\text{cut}}))U. \quad (55)$$

As mentioned earlier, the cutoff distance R_{cut} provides a way to tradeoff the maximum rate-outage probability (40) and the average number of users in rate-outage (55) against the data rate (48) offered to each served user. The effect of R_{cut} is investigated by simulation in Section 8.

6.3. Average user-rate and spectral efficiency

In order to derive an approximate bound for the average maximum rate \bar{D}_{max} , we replace in (48) each variable U_q by its expectation as follows:

$$\bar{D}_{\text{max}} = \frac{BS}{\sum_{q=1}^Z \bar{U}_q / \log_2 M_q}. \quad (56)$$

From (53) we get

$$\bar{U}_q = (\bar{u}(R_q) - \bar{u}(R_{q-1}))U. \quad (57)$$

What about the average spectral efficiency achieved by our allocation method? Note that the average number of served users is

$$\bar{U}_{\text{srv}} = \sum_{q=1}^Z \bar{U}_q. \quad (58)$$

These \bar{U}_{srv} users achieve an aggregate data rate of $\bar{U}_{\text{srv}} \bar{D}_{\text{max}}$ on average using a total bandwidth of BS. So, from (56) we find that the average spectral efficiency is given by

$$\bar{\eta} = \frac{\sum_{q=1}^Z \bar{U}_q}{\sum_{q=1}^Z \bar{U}_q / \log_2 M_q}. \quad (59)$$

Notice that the constraint of the minimum data rate per user D_0 has not been considered yet. Considering this additional constraint provides a criterion for *admission control* as discussed in Section 6.2.

6.4. System capacity and admission control

Having $D_{\text{max}} \geq D_0$ along with (48) and (57) gives the following *upper-bound on the bearable number of users*

$$U_{\text{max}} = \frac{BS}{D_0 \sum_{q=1}^Z (\bar{u}(R_q) - \bar{u}(R_{q-1})) / \log_2 M_q}. \quad (60)$$

This U_{max} corresponds to a kind of *maximum system load* in terms of number of users given the minimum required QoS. When the system is fully loaded, that is, $U = U_{\text{max}}$, it cannot offer to each user better than the minimum required data rate D_0 . In this case, any additional user that requests an access to the service is rejected by the base station. So, the ratio U/U_{max} can be considered as a metric for the system load. In brief, for $U \leq U_{\text{max}}$ we have

$$D_{\text{max}} = \frac{U_{\text{max}}}{U} D_0. \quad (61)$$

Section 7 describes the proposed resource allocation algorithm from a practical point of view.

7. RESOURCE ALLOCATION ALGORITHM IN PRACTICE

In the following we assume that $U \leq U_{\text{max}}$. We saw that the optimal user-wise subcarrier allocation is defined by (49). However, one must take into account the fact that, in practice, only an integer number of subcarriers can be assigned to a given modulation zone. Thus, the S_q 's have to be rounded to integer numbers. Let

$$\tilde{S}_q = \mathcal{I}(S_q) = \mathcal{I}\left(\frac{U_q / \log_2 M_q}{\sum_{k=1}^Z U_k / \log_2 M_k} S\right), \quad (62)$$

where $\mathcal{I}(x)$ is the nearest integer to x . These \tilde{S}_q 's define Z zones inside the frame where the first zone, for example, is

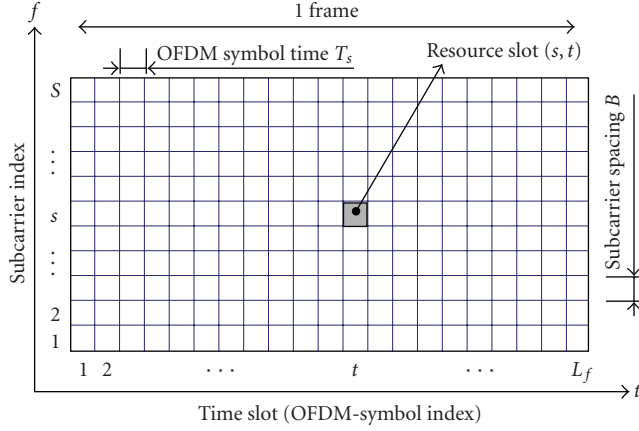


FIGURE 4: Frame structure and elementary resource slot.

constituted of slots modulated by the M_1 -QAM constellation.

From (44), the number of subcarriers that a user u in zone q needs to achieve the common rate is equal to $W(u) = D_{\max}/(B \log_2 M_q)$. So, another concern is that $W(u)$ is not necessarily an integer. In practice, users are mapped to a frame of length L_f OFDM symbols (see Figure 4). With S subcarriers, a frame is composed of SL_f slots. Let N_q be the integer number of slots allocated to each user in zone q . Hence, each user obtains on average N_q/L_f subcarriers per frame. So, to cope with a noninteger $W(u)$, the value of N_q must be chosen so that the difference $|N_q/L_f - W(u)|$ is minimized. This gives

$$N_q = \mathcal{I}(L_f W(u)) = \mathcal{I}\left(\frac{L_f D_{\max}}{B \log_2 M_q}\right). \quad (63)$$

The difference between N_q/L_f and $W(u)$ makes the obtained average data rate per user $N_q B \log_2 M_q / L_f$ slightly different from the theoretical value D_{\max} . This difference decreases with the frame length L_f . Moreover, the actual number of users that can be mapped to the $\tilde{S}_q L_f$ slots of the M_q -QAM zone in the frame is equal to

$$\tilde{U}_q = \tilde{S}_1 \div N_q, \quad (64)$$

where \div is the integer division operator.

Now we describe the proposed resource allocation algorithm. We assume that the shadowed distances of users are known to the base station. So, the base station can sort the users in a vector \mathcal{U} according to their increasing shadowed distances. Under these assumptions, the proposed resource allocation algorithm consists of two steps. The first step can be done once of-line while the second needs to be carried out dynamically according to the system load and to the CSI update rate.

Step 1 (offline resource reservation).

- (1) Find the required power margin F using (35).
- (2) Use (34) to find the maximum range R_q for each M -QAM constellation.

- (3) Calculate from (40) the cutoff distance R_{cut} that yields an acceptable maximum rate-outage probability.
- (4) Deduce the required number of zones Z from (46) and the corresponding set of constellations (i.e., the Z first high-order constellations in \mathcal{M}).

Step 2 (online resource allocation).

- (1) Using the vector \mathcal{U} of users sorted according to their shadowed distances, find U_q , the number of users belonging to each modulation zone.
- (2) Deduce the zone-wise number of subcarriers \tilde{S}_q for $q = 1, \dots, Z$ using (62).
- (3) Compute the maximum common rate D_{\max} from (48).
- (4) Use (63) and (64) to calculate N_q , the required number of slots per user as well as the number of bearable users \tilde{U}_q , respectively.
- (5) Map users to the frame slots as follows: the first \tilde{U}_1 users in \mathcal{U} are mapped to the first \tilde{S}_1 subcarriers corresponding to the highest-order modulation zone. The first user is granted the first N_1 time slots on the first subcarrier. If $N_1 > L_f$, additional time slots on the second subcarrier is granted to this user until N_1 is reached. Then, the next user is allocated the next time slots and so on. These operations are repeated for the next \tilde{U}_2 users in \mathcal{U} , that belong to the second modulation zone, and so on until all slots are occupied.

Figure 5 illustrates the online step of the allocation process.

The available CSI, limited to the users shadowed distances, is implicitly used by the base station during the slot allocation stage in Step 2. One should wonder how precise does this CSI need to be? In other words, how the achieved performance can be affected by errors on the CSI? Errors may come from imperfect CSI estimation as well as from quantization noise on the feedback channel. Ideally, users are sorted in vector \mathcal{U} according to their shadowed distances. Suppose that we modify the order in \mathcal{U} of a subset of users belonging to the same modulation zone. This kind of perturbation has no effect on the expected performance since these users continue to get the same resources. On the contrary, some performance degradation may appear when the CSI errors shift some users from a modulation zone to another. The effect of imperfect CSI is evaluated in the numerical results presented in Section 8.

8. NUMERICAL RESULTS

In Table 2, an example of a typical parameter setting is provided. We assume that the available constellations are 64-QAM, 16-QAM, QPSK, and BPSK. Concerning the SNR thresholds $\gamma_q = \beta_{M_q}^{-1}(b)$ that correspond to the target BER b for these constellations, we know that for the BPSK, the error

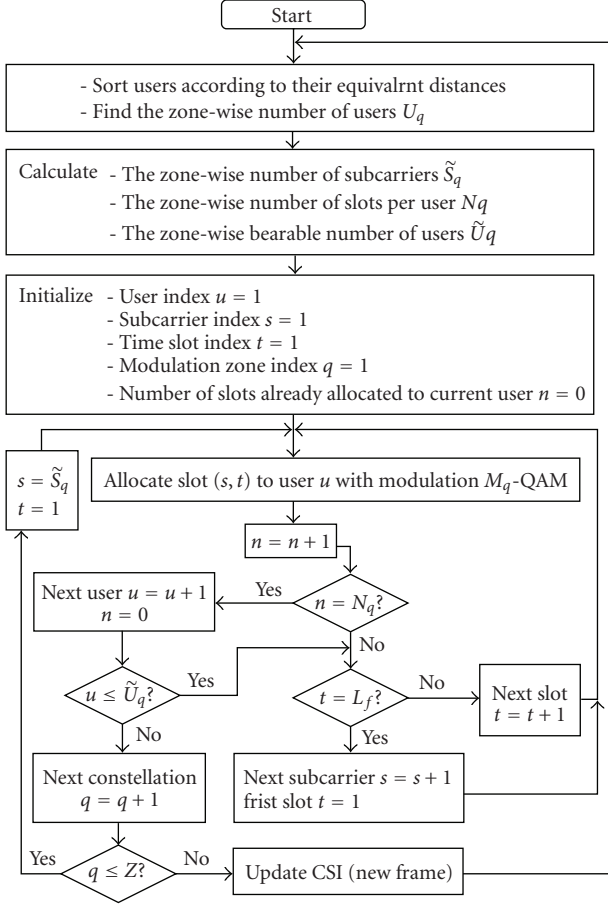


FIGURE 5: Flow chart of the resource allocation online step.

probability can be expressed using the error function (38) by $b = 0.5 \operatorname{erf}(\sqrt{\gamma_{\text{BPSK}}})$. Thus, we have

$$\gamma_{\text{BPSK}} = [\operatorname{erf}^{-1}(2b)]^2. \quad (65)$$

For higher-order modulations ($M_q > 2$), a good approximation of the SNR threshold for an M_q -QAM modulation with $b \leq 10^{-3}$ is given in [19] by

$$\gamma_q = (M_q - 1)\Gamma(b), \quad (66)$$

where $\Gamma(b) = -\log(5b)/1.6$. Table 3 gives for each constellation the SNR-threshold γ_q corresponding to a BER of 10^{-3} .

In the following, the performance of our allocation method is characterized versus the *worst-case average SNR* (WASNR) defined by

$$\gamma_{\text{wa}} = \frac{P_{\text{tot}} G_0}{SBN_0 R^\alpha}. \quad (67)$$

This WASNR corresponds to the area-mean SNR on the cell edge. Given the parameter setting in Table 2, we find using (35) that the required power margin in logarithmic units is $F_{\text{dB}} = 10 \log_{10} F \approx 12.9$ dB. In Figure 6 the maximum attainable range (BPSK coverage) for $\gamma_{\text{wa}} \in [5, 25]$ dB is plotted using (34) and (67). Taking into account the target

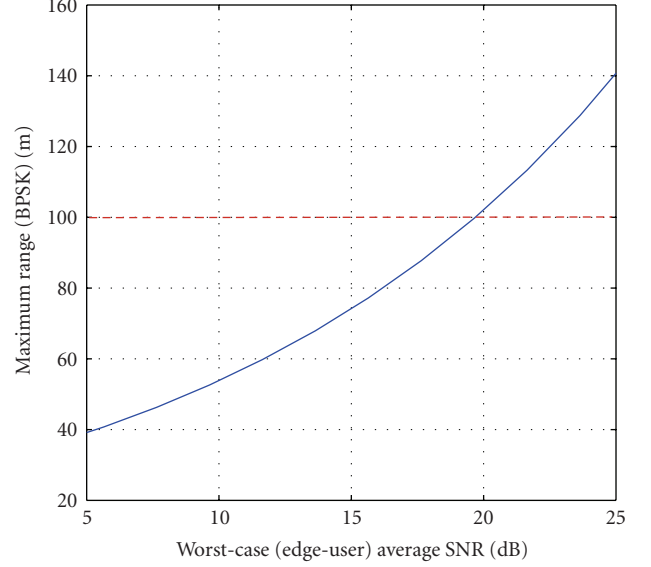


FIGURE 6: Maximum attainable range (BPSK coverage) versus worst-case (edge-user) average SNR (WASNR).

cell radius $R = 100$ m, we see that the minimum acceptable WASNR value is $\gamma_{\text{wa}} \approx 19.64$ dB ($P_{\text{tot}} \approx 2.73$ W). Moreover, if we assume that the maximum possible value for the total power is $P_{\text{tot}} = 10$ W, we find from (67) that γ_{wa} must be limited to about 25.6 dB. Hence, in the following, we let the WASNR varies in the range [20, 25] dB.

Let us start by checking the validity of the average shadowed user distribution defined in (53). Remember that the product $\bar{u}(x)U$ represents the average cumulative distribution function of the number of users whose shadowed distances are within $[0, x]$. In other words, the average histogram of the shadowed distances must coincide with $(\bar{u}(x + \Delta x) - \bar{u}(x))U$ for $\Delta x \ll 1$. This is validated by simulation results depicted in Figure 7 where the curves are normalized to the total number of users and averaged on 1000 shadowing realizations.

Now we compare the average maximum achieved user rate for 100 users to the analytic expression (56). This is shown in Figure 8 for two different cutoff distances. When R_{cut} is set to the BPSK range R_{BPSK} , the maximum number of users is served except those who fall beyond R_{BPSK} in terms of shadowed distance. The achieved data rate is improved in the opposite case when the base station decides not to serve users beyond the cell edge by setting $R_{\text{cut}} = R$. The improvement is significant for high SNRs while it vanishes near SNR = 20 dB where $R_{\text{BPSK}} \approx R$. The gain in user rate is paid for in terms of the average number of users in rate outage as expected by (55). This is shown in Figure 9. So, a tradeoff has to be found, via R_{cut} , between user rate and rate outage. By varying R_{cut} for a fixed total power $P_{\text{tot}} = 10$ W, we show in Figure 10 the average percentage of users in rate outage versus the average maximum user rate.

Now we compare the average spectral efficiency (59) of our allocation method to the average spectral efficiency of two other methods. The first one, used if no CSI is available,

TABLE 2: Simulation parameters' values.

Center frequency f_c	3.5	GHz
Total transmit power P_{tot}	10	W
Number of subcarriers S	256	subcarriers
Total bandwidth B_{tot}	20	MHz
Subcarrier spacing $B = B_{\text{tot}}/S$	78.125	KHz
OFDM symbol duration $T_s = 1/B$	12.8	μs
Frame length L_f	100	symbols
Frame duration $L_f T_s$	1.28	ms
Noise power spectral density N_0	-174	dBm/Hz
Path-loss exponent α	3.6	—
Shadowing standard deviation σ	5	dB
Target minimum data rate D_0	100	Kbps
Target BER b	10^{-3}	—
Target maximum outage probability ε	0.05	—
Target cell radius R	100	m

TABLE 3: Available modulations' SNR thresholds for $b = 10^{-3}$.

Modulation	64-QAM	16-QAM	QPSK	BPSK
Zone index q	1	2	3	4
γ_q (dB)	23.2	17	10	6.8

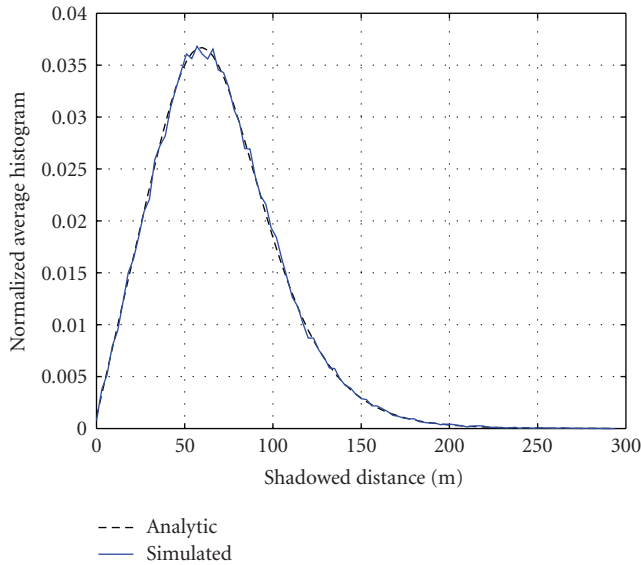
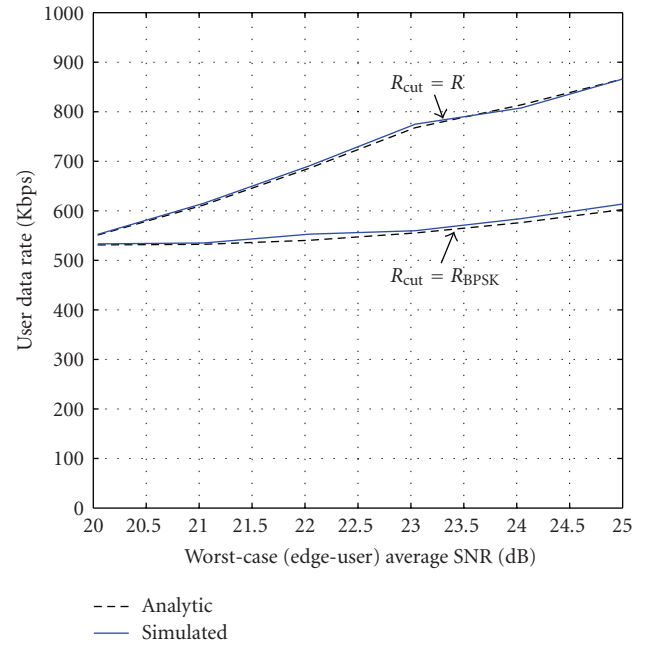


FIGURE 7: Average histogram of user shadowed distance.

corresponds to a traditional “static” allocation of subcarriers and rates based on a worst-case design. The second allocation method, called “Max-Min” method, is the one introduced in [7] and reused in [6]. It is based on the assumption of full CSI knowledge that allows an improved spectral efficiency but requires excessive feedback overhead.

In the static allocation case, the power margin must account for both effects of shadowing and fading. We found by simulation that under a composite log-normal-Rayleigh channel, the required power margin for the specified BER

FIGURE 8: Maximum user rate versus worst-case average SNR for $U = 100$ users.

and outage probability is $F = 14.8$ dB. This value can be retrieved analytically using results in [21] where it was shown that a composite log-normal-Rayleigh distribution is equivalent to a modified log-normal distribution. Without any CSI, the same modulation must be used over all subcarriers. The modulation order M must guarantee to edge users (worst-case) the required BER. So, M can be obtained by setting d_u in (36) to the cell radius R . Then, the S subcarriers must be equally partitioned among the users so that each user achieves a data rate $(S/U)B \log_2 M$. Concerning the “max-min” method [6], its goal is to maximize the minimum user rate under a total power constraint. Maximizing the minimum rate is equivalent to maximizing the sum rate with

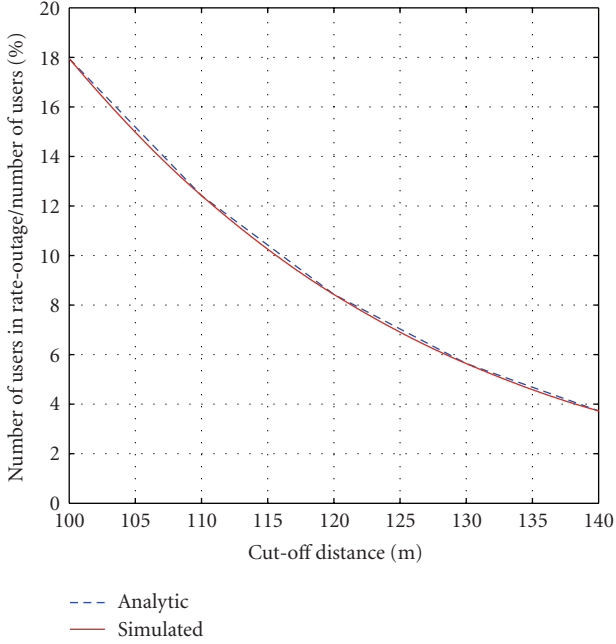


FIGURE 9: Average percentage of users in rate outage versus the cutoff distance for $U = 100$ users and $P_{\text{tot}} = 10$ W (WASNR ≈ 25 dB).

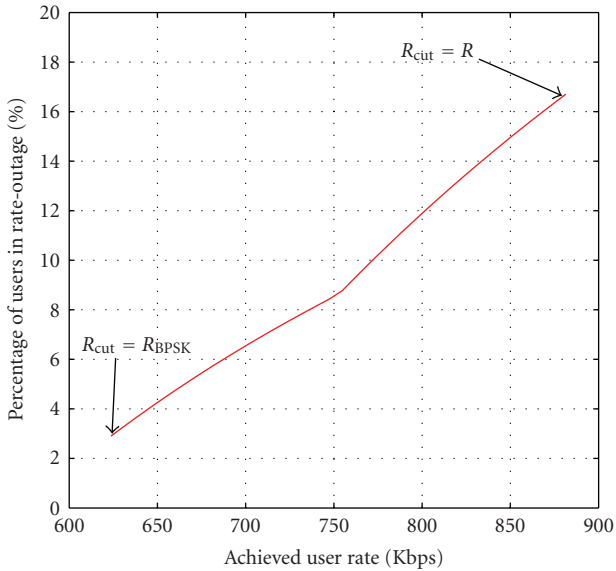


FIGURE 10: Tradeoff between the achieved user rate and the percentage of users in rate outage for $U = 100$ users and $P_{\text{tot}} = 10$ W (WASNR ≈ 25 dB).

users having equal rates. In Figures 11 and 12, the spectral efficiency (59) of our allocation method is compared to the spectral efficiency of the above-mentioned methods for 10 and 100 users respectively. We choose a cutoff distance of 120 m (the mid-point between the cell radius and the BPSK range). According to Figure 9, this cutoff distance yields about 8.5% of users in rate-outage. This, from Figure 10, corresponds to an average user rate of 760 kbps. To make

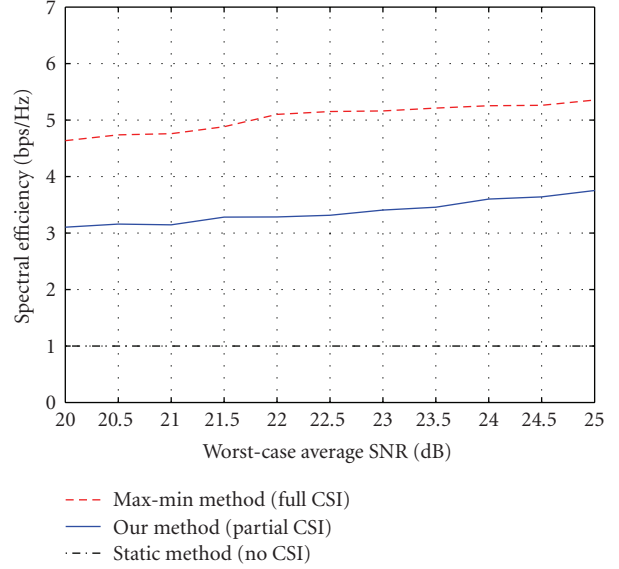


FIGURE 11: Average spectral efficiency of our allocation method compared to the static and max-min methods for $U = 10$ users and $R_{\text{cut}} = 120$ m.

the comparison fair, the CSI is quantized on 3 bits. This reduces the feedback overhead required by the “max-min” method to 3S bits per user. However, our allocation method requirement in terms of CSI precision is significantly less as shown in the sequel. Due to multiuser diversity gain, the “Max-Min” method exhibits a spectral efficiency gain of about 1 bps/Hz when the number of users passes from 10 to 100. On the otherhand, the average spectral efficiency of our method is unchanged as expected by (59). In both cases, for the static method, the whole cell is covered using the BPSK of spectral efficiency 1 bps/Hz. We see that our allocation method offers a significant spectral efficiency gain compared to the static method.

As mentioned earlier, the main advantages of our method are its simplicity and the limited CSI feedback it requires. The price to be paid is some degradation in spectral efficiency compared to the full-CSI-based “max-min” method as shown in Figures 11 and 12. This loss is compensated by the complexity reduction and the limited feedback overhead that our algorithm requires.

Concerning the rate outage, note that the static method yields zero rate outage at the expense of the achieved spectral efficiency. If the BPSK is replaced by the QPSK, the rate outage remains null (since the BS has not any criterion for rejecting some users) while the BER-outage probability constraint will be violated for near-edge users which is not compatible with the target of this work. The simulated average percentage of users in rate outage is plotted in Figure 13 versus the WASNR and compared to the expected analytic results from (55).

Finally, we want to characterize the sensitivity of our allocation method to CSI accuracy. Imperfect CSI is modeled by adding a zero-mean Gaussian error to the user shadowed distance. We assume that errors for different users are

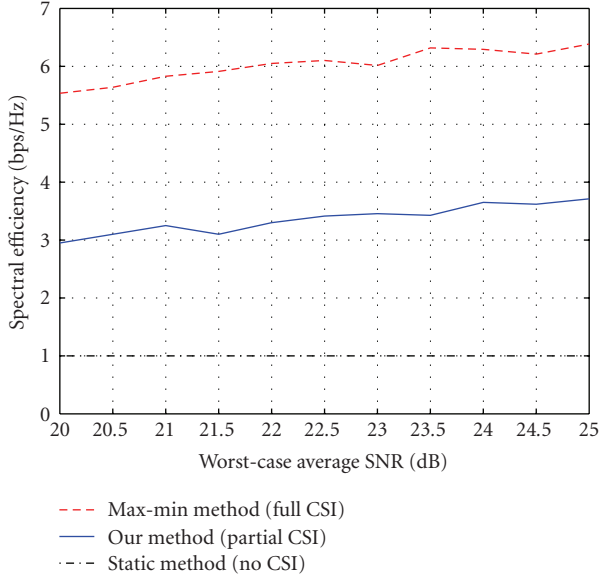


FIGURE 12: Average spectral efficiency of our allocation method compared to the static and max-min methods for $U = 100$ users and $R_{\text{cut}} = 120$ m.

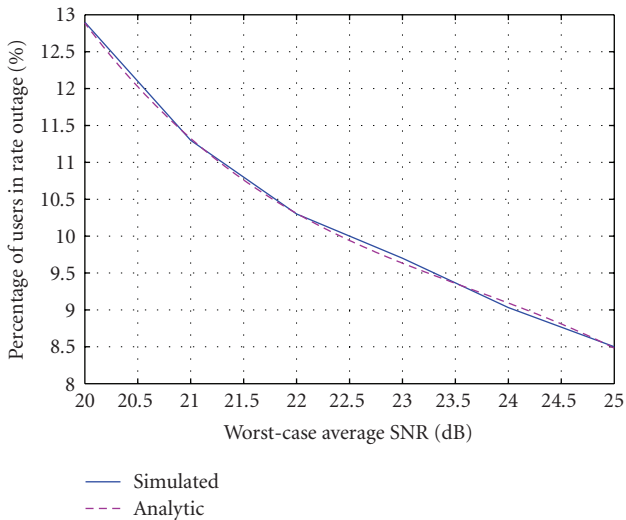


FIGURE 13: Average percentage of users in rate outage versus the worst-case average SNR for $U = 100$ users and $R_{\text{cut}} = 120$ m.

independent. Thus, if d_u is the shadowed distance of user u , the estimated distance is $\hat{d}_u = d_u + e_u$. The error e_u follows a Gaussian law of standard deviation $\sigma_u = aR$ so that a represents the error standard deviation normalized to the cell radius R . This parameter measures the CSI accuracy (a perfect CSI corresponds to $a = 0$). Errors on shadowed distances disturb the rate allocation decision leading to unexpected BER outage events. We use the *average percentage of users in BER-outage per frame* as an overall performance metric. In Figure 14, this metric is plotted versus the accuracy parameter a for a fully loaded system ($U = U_{\text{max}}$) and a total power $P_{\text{tot}} = 10$ W. We notice that the percentage of

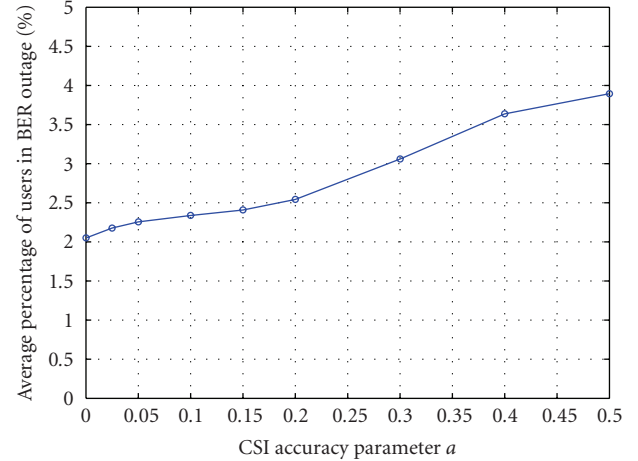


FIGURE 14: Effect of CSI accuracy measured by the parameter a on the average percentage of users in BER outage ($U = U_{\text{max}}$, $P_{\text{tot}} = 10$ W).

users in BER outage for perfect CSI ($a = 0$) is about 2%. The degradation does not exceed 2% even at $a = 0.5$ which corresponds to a significantly-degraded CSI. This shows the robustness of the proposed resource allocation method to CSI estimation errors. So, even with coarse estimates of users' shadowed distances, the achieved overall performance remains acceptable. In fact, all the base station needs in order to properly allocate the resources is the index of the modulation zone each user belongs to. Assume that the base station broadcasts the modulation zones' radii to all the users in a dedicated frame header and that the shadowed distances are estimated by users themselves. In this case, each user can find the index of his modulation zone and then feedback this value to the base station on the uplink. This feedback information is simply a discrete value between 1 and Z , the number of modulation zones. So, the feedback requires about $\log_2 Z$ information bits per user. In our example above where $Z = 4$ zones, two bits per user are needed. This approach is equivalent to quantizing the CSI, or users' shadowed distances, using irregular thresholds which are the zones radii ($\mathcal{R}_q = 51, 76, 119, 146$ (m) in our example). In full CSI approaches, if the CSI is quantized on N bits, the feedback overhead is NS bits per user.

9. CONCLUSION

In this paper, we considered the problem of resource allocation on the downlink of a single-cell OFDMA system under QoS fairness constraints with limited channel state information (CSI). Fairness was defined by a minimum user data rate, a target BER, and a maximum BER-outage probability. We supposed that the only CSI available to the base station is a coarse estimation of the users shadowed distances that we defined. Thus, under the fairness constraint, a total peak power constraint and a given number of users uniformly distributed over the cell of a given radius, we derived the optimal subcarrier and rate allocation that offers the maximum data rate per user.

Simulation results showed that our resource allocation algorithm yields a significant spectral efficiency enhancement compared to the traditional static scheme. Meanwhile, the loss in terms of average spectral efficiency with respect to a full-CSI-based opportunistic allocation was shown to be acceptable. In fact, this performance degradation is the price to be paid for the reduced complexity and the low feedback overhead that our solution requires. Finally, the robustness of our algorithm to CSI estimation errors was shown which proving that even a coarse CSI is sufficient for our algorithm to operate efficiently.

APPENDIX

Here we provide the details of derivation of (53) from (52). After replacing (50) and (51) into (52) and integrating with respect to θ , the equation to be proved becomes

$$\begin{aligned} & \int_0^R \operatorname{erf}\left(\frac{10 \log_{10}(R_q/r)}{\sigma\sqrt{2}/\alpha}\right) r dr \\ &= \frac{1}{2} R^2 \operatorname{erf}\left(C \log \frac{R_q}{R}\right) \\ &+ \frac{1}{2} R_q^2 e^{1/C^2} \left(1 - \operatorname{erf}\left(C \log \frac{R_q}{R} + \frac{1}{C}\right)\right). \end{aligned} \quad (\text{A.1})$$

The positive constant C was defined earlier by $C = 10\alpha/(\sigma\sqrt{2} \log 10)$. Let $A_q = C \log R_q$. Then, the left-hand side (LHS) of the equation above becomes

$$\text{LHS} = \int_0^R \operatorname{erf}(A_q - C \log r) r dr. \quad (\text{A.2})$$

Considering the variable change $y = A_q - C \log r$ we obtain

$$\text{LHS} = \frac{1}{C} e^{2A_q/C} \int_{A_q - C \log R}^{\infty} e^{-2y/C} \operatorname{erf}(y) dy. \quad (\text{A.3})$$

Now we integrate by parts assuming that $g(y) = \operatorname{erf}(y)$ and $df(y)/dy = -(2/C)e^{-(2/C)y}$. It follows that $dg(y)/dy = (2/\sqrt{\pi})e^{-y^2}$ and $f(y) = e^{-(2/C)y}$. So, we obtain

$$\begin{aligned} \text{LHS} &= -\frac{1}{2} e^{2A_q/C} \left[[f(y)g(y)]_{A_q - C \log R}^{\infty} \right. \\ &\quad \left. - \int_{A_q - C \log R}^{\infty} f(y) \frac{dg(y)}{dy} dy \right] \\ &= \frac{1}{2} e^{2A_q/C} \left[e^{-2(A_q - C \log R)/C} \operatorname{erf}(A_q - C \log R) \right. \\ &\quad \left. + \frac{2}{\sqrt{\pi}} \int_{A_q - C \log R}^{\infty} e^{-(2/C)y - y^2} dy \right]. \end{aligned} \quad (\text{A.4})$$

A new variable change $x = y + 1/C$ in the second term integral leads to

$$\begin{aligned} \text{LHS} &= \frac{1}{2} e^{2A_q/C} \left[e^{-2(A_q - C \log R)/C} \operatorname{erf}(A_q - C \log R) \right. \\ &\quad \left. + e^{1/C^2} \frac{2}{\sqrt{\pi}} \int_{A_q - C \log R + 1/C}^{\infty} e^{-x^2} dx \right] \\ &= \frac{1}{2} R^2 \operatorname{erf}(A_q - C \log R) \\ &\quad + \frac{1}{2} e^{1/C^2} e^{-2A_q/C} \left(1 - \operatorname{erf}\left(A_q - \frac{C \log R + 1}{C}\right)\right). \end{aligned} \quad (\text{A.5})$$

Remembering that $A_q = C \log R_q$, the wanted formula is obtained.

ACKNOWLEDGMENT

This work was supported by the French ANR Telecom project ORMAC.

REFERENCES

- [1] IEEE std 802.16d, "Air Interface for Fixed Broadband Access Systems," 2004.
- [2] WiMax Forum, "Fixed, Nomadic, Portable and Mobile Applications for 802.16-2004 and 802.16e WiMAX Networks," White paper, November 2005.
- [3] A. Pandharipande, M. Kountouris, H. Yang, and H. Park, "Subcarrier allocation schemes for multiuser OFDM systems," in *Proceedings of the 7th IEEE International Conference on Signal Processing and Communications (SPCOM '04)*, pp. 540–544, Bangalore, India, December 2004.
- [4] I. C. Wong, Z. Shen, B. L. Evans, and J. G. Andrews, "A low complexity algorithm for proportional resource allocation in OFDMA systems," in *Proceedings of IEEE Workshop on Signal Processing Systems (SIPS '04)*, pp. 1–6, Austin, Tex, USA, October 2004.
- [5] H. Yin and H. Liu, "An efficient multiuser loading algorithm for OFDM-based broadband wireless systems," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '00)*, vol. 1, pp. 103–107, San Francisco, Calif, USA, November-December 2000.
- [6] Z. Shen, J. G. Andrews, and B. L. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Transactions on Wireless Communications*, vol. 4, no. 6, pp. 2726–2737, 2005.
- [7] W. Rhee and J. M. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," in *Proceedings of the 51st IEEE Vehicular Technology Conference (VTC '00)*, vol. 2, pp. 1085–1089, Tokyo, Japan, May 2000.
- [8] K. I. Pedersen, G. Monghal, I. Z. Kovacs, et al., "Frequency domain scheduling for OFDMA with limited and noisy channel feedback," in *Proceedings of the 66th IEEE Vehicular Technology Conference (VTC '07)*, pp. 1792–1796, Baltimore, Md, USA, September-October 2007.
- [9] C. Mehlführer, S. Caban, and M. Rupp, "Experimental evaluation of adaptive modulation and coding in MIMO WiMAX with limited feedback," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, Article ID 837102, 12 pages, 2008.

- [10] J. Chen, R. Berry, and M. Honig, "Large system performance of downlink OFDMA with limited feedback," in *Proceedings of IEEE International Symposium on Information Theory (ISIT '06)*, pp. 1399–1403, Seattle, Wash, USA, July 2006.
- [11] J. Chen, R. Berry, and M. Honig, "Asymptotic analysis of downlink OFDMA capacity," in *Proceedings of 44th Annual Allerton Conference on Communication, Control and Computing*, Monticello, Ill, USA, September 2006.
- [12] S. Sanayei and A. Nosratinia, "Opportunistic downlink transmission with limited feedback," *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4363–4372, 2007.
- [13] J. Chen, R. Berry, and M. Honig, "Performance of limited feedback schemes for downlink OFDMA with finite coherence time," in *Proceedings of IEEE International Symposium on Information Theory (ISIT '07)*, pp. 2751–2755, Nice, France, June 2007.
- [14] Z.-H. Han and Y.-H. Lee, "Opportunistic scheduling with partial channel information in OFDMA/FDD systems," in *Proceedings of the 60th IEEE Vehicular Technology Conference (VTC '04)*, vol. 1, pp. 511–514, Los Angeles, Calif, USA, September 2004.
- [15] P. Svedman, D. Hammarwall, and B. Ottersten, "Sub-carrier SNR estimation at the transmitter for reduced feedback OFDMA," in *Proceedings of the 14th European Signal Processing Conference (EUSIPCO '06)*, Florence, Italy, September 2006.
- [16] D. Gesbert and M. S. Alouini, "Selective multi-user diversity," in *Proceedings of the 3rd IEEE International Symposium on Signal Processing and Information Technology (ISSPIT '03)*, pp. 162–165, Darmstadt, Germany, December 2003.
- [17] M. Johansson, "On scheduling and adaptive modulation with limited channel feedback," Tech. Rep., Signals and Systems Group, Uppsala University, Uppsala, Sweden, 2004.
- [18] V. S. Abhayawardhana, I. J. Wassell, D. Crosby, M. P. Sellars, and M. G. Brown, "Comparison of empirical propagation path loss models for fixed wireless access systems," in *Proceedings of the 61st IEEE Vehicular Technology Conference (VTC '05)*, vol. 1, pp. 73–77, Stockholm, Sweden, May-June 2005.
- [19] A. J. Goldsmith and S.-G. Chua, "Variable rate variable power MQAM for fading channels," *IEEE Transactions on Communications*, vol. 45, no. 10, pp. 1218–1230, 1997.
- [20] D. Kivanc, G. Li, and H. Liu, "Computationally efficient bandwidth allocation and power control for OFDMA," *IEEE Transactions on Wireless Communications*, vol. 2, no. 6, pp. 1150–1158, 2003.
- [21] M.-S. Alouini and A. J. Goldsmith, "Area spectral efficiency of cellular mobile radio systems," *IEEE Transactions on Vehicular Technology*, vol. 48, no. 4, pp. 1047–1066, 1999.