

Universidad Carlos III de Madrid

 e-Archivo

Institutional Repository

This document is published in:

*IEEE Trans. on Vehicular Technology* 63 (2014) 3, pp. 1178 - 1192

DOI: 10.1109/TVT.2013.2281367

© 2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

# Fairness-Adaptive Goodput-Based Resource Allocation in OFDMA Downlink with ARQ

José Joaquín Escudero-Garzás, *Member, IEEE*, Bertrand Devillers, and Ana García-Armada, *Senior Member, IEEE*

**Abstract**—We present a cross-layer resource-allocation (RA) scheme for the downlink in orthogonal frequency-division multiple-access (OFDMA) systems with fairness control among the users, where the resources to be allocated are power, bits per symbol, and subchannels. The use of *subchannels*, which are defined as group of subcarriers, leads to reducing the complexity of the bandwidth allocation compared with the commonly adopted subcarrier allocation. A goodput-based optimization function, which is derived by combining automatic repeat request (ARQ) and physical (PHY)-layer parameters, is used to perform RA for applications that demand error-free transmissions. Two transmission strategies are considered, with and without concatenation of subchannels, for which two different RA methods are developed, respectively. We also propose an algorithm that improves the complexity associated to both concatenation and nonconcatenation schemes, without appreciable performance loss.

**Index Terms**—Automatic repeat request (ARQ), cross-layer resource allocation (RA), fairness, goodput, multiuser orthogonal frequency-division multiplexing (OFDM), subchannel.

## I. INTRODUCTION

THE orthogonal frequency-division multiple-access (OFDMA) technique plays a key role in current wireless systems due to its resistance to multipath fading and high spectral efficiency. Many research efforts have been devoted to the crucial task of resource allocation (RA) in OFDMA to reach the maximum data rate, where the optimal solution may be achieved by exhaustive search over users, subcarriers, modulation and coding schemes (MCSs), and power levels. In particular, a widely adopted criterion in this type of RA schemes is the maximization of the sum throughput [1]–[7].

However, some high-bit-rate applications and services, such as broadcast and video streaming, demand error-free transmissions, and only error-free frames are kept by the receiver, whereas for the others, retransmission is required. In this case,

The work of A. García-Armada was supported by the GRE3N Project under Grant TEC2011-29006-C03-03.

J. J. Escudero-Garzás is with the Department of Systems and Information Engineering, University of Virginia, Charlottesville, 22904 VA USA (e-mail: jje6r@virginia.edu).

B. Devillers was with Centre Tecnològic de Telecomunicacions de Catalunya, Barcelona 08860, Spain. He is now with the European Patent Office, 2280 HV Rijswijk, The Netherlands (e-mail: bertrand.devillers@cttc.es).

A. García-Armada is with the Department of Signal Theory and Communications, Carlos III University of Madrid, Madrid 28911, Spain (e-mail: agarcia@tsc.uc3m.es).

a tradeoff between the bit rate achieved by the physical (PHY) layer and the error rate achieved by link-layer error-correcting schemes is desired, which is not realized by approaches based on throughput. Instead, *goodput*, which is defined as the number of error-free bits transmitted per unit of time, can be a very suitable metric, motivating a growing interest in the design of schemes for optimizing goodput-based functions, e.g., in [8]–[12]. On the other hand, a link-layer retransmission scheme is mandatory. Among the retransmission schemes, the automatic repeat request (ARQ) protocol has the advantage of being already present in existing OFDMA systems [13], [14] and is therefore extensively applied (see, e.g., [15]).

For the particular case of RA in goodput-oriented OFDMA systems, although the single-layer approach can be adopted [16], [17], cross-layer schemes are preferred as they enable global system performance optimization [18]–[25]. In their works, Aggarwal *et al.* [18]–[20] have proposed several methods for allocating subcarriers, power, and MCSs to maximize the expected-sum-goodput-based utility. In [18], they exploit the use of acknowledgment/negative acknowledgment (ACK/NAK) feedback. In [19], imperfect channel state information (CSI) at the transmitter (CSIT) is assumed, where Aggarwal *et al.* study the continuous (a single resource shared by several users) and discrete cases and show that the continuous case solves the discrete case in some situations, providing a bound for the others. The work in [19] is extended in [20] to the case when ACK/NAK feedback is utilized. Given that the optimal solution is a partially observable Markov decision process, which is impractical for implementation purposes, Aggarwal *et al.* propose a two-step approach that uses first a greedy resource allocation and afterward updates the subcarrier gain probability. Another related set of works has been published by Lau *et al.* [21]–[24], where they provide schemes and algorithms for OFDMA networks with imperfect CSIT. In [21], it is discussed the optimal RA in time-division duplex–OFDMA systems by exploiting the 1-bit limited feedback, showing that optimal rate, power, and user allocations that maximize the conditional average system goodput converge if the number of packets per slot is sufficiently large. In [22], Lau *et al.* present subband, power, rate, and user allocation for systems with delayed CSIT for maximizing the average total goodput. In [23] and [24], Hui and Lau propose, respectively, centralized and decentralized delay-sensitive cross-layer schemes to allocate power, rate, and subcarriers, where the CSIT is assumed to be outdated; both approaches use the average total goodput as an optimization function. A dual-decomposition-based solution to the problem of maximizing the average weighted sum goodput for OFDMA relay networks with

imperfect CSIT is presented in [25]. The problem is formulated in terms of rate, power, and subcarrier allocation policies, and it is shown how the system performance scales with the number of users and relays. However, the aforementioned schemes may lead to unfair RA when users with very good channel conditions are allocated the resources.

It is well known that OFDMA RA problems are computationally demanding; for example, OFDMA-based access networks are characterized by exploiting a large number of subcarriers (e.g., up to 2048 subcarriers in IEEE 802.16 standard-based systems [13], [14]) and render RA intractable from a computational viewpoint. The utilization of *subchannels* (also known in the literature as chunks [26] or bins [27]), which are defined as groups of subcarriers [13], [14], can alleviate the computational load associated with the RA [26]–[28].

Although some works claim that the transmitted frame is error free when a base station (BS) has perfect CSIT, powerful error-correcting coding (e.g., low-density parity-check or turbo codes) and sufficiently long frames are required in slow-fading channels [21], [23]. Moreover, in practical systems, even with perfect CSIT, the use of practical modulation schemes such as  $M$ -ary quadrature amplitude modulation (MQAM) inherently introduces penalization in the achievable bit rate, and the frame is not error free due to the bit error rate (BER) associated to a given modulation scheme. Throughout this paper, perfect CSIT is assumed, and MQAM modulation and convolutional coding are used.

In this paper, we address the optimization of the sum goodput transmitted by a BS in OFDMA networks. We propose a cross-layer RA scheme that combines PHY-layer parameters with the ARQ protocol to derive the goodput transmitted per frame expression and incorporates adaptive fairness and subchannel concatenation in the scheme. These aspects have been contemplated in a very preliminary form in an earlier conference paper [29]. Our proposal brings the following advantages.

- 1) Encoding frames are formed by using two possible strategies, i.e., with and without concatenation of subchannels, as in current OFDMA-based standards [14]. We show in Section VII the different results according to the selected concatenation strategy.
- 2) The use of subchannel allocation instead of subcarrier allocation reduces the complexity associated to the RA problem, which is here formulated as a subchannel, bit, and power allocation (SBPA) problem. As power and bits per symbol are allocated with subcarrier granularity, not at a subchannel level, the advantages of bit and power allocation are preserved.
- 3) The use of ARQ improves the performance of the system [30] and provides a connection among PHY-layer parameters (BER, power, and bits per symbol), link-layer parameters [frame success rate (FSR)], and goodput.
- 4) A fairness mechanism is included in the RA formulation. Fairness has been well addressed in the related literature, mainly in the form of proportional fairness (see, e.g., [31]). In this paper, fairness is incorporated in the theoretical framework such that the degree of fairness can be adapted by means of the  $\alpha$ -proportionally fair rule [32].

In general, goodput has been considered in its average form in the related literature [18]–[24]. Alternatively, we propose an approach based on the optimization of the sum of users' goodput achieved per frame [25], [33], or *goodput per transmitted frame*, since the problem constraints are per frame as far as the optimization problem is concerned. The goodput transmitted per frame expression is based on the FSR, which is a function of the BER per user and per subcarrier, providing a more accurate value of the FSR than if only the SNR distribution is known [18]. A related formulation is the outage probability per user per subcarrier and per packet used (see, e.g., [23]).

In this paper, we propose the use of two different encoding frame strategies based on whether the subchannels are concatenated (CS) or not concatenated (NCS). If CS is used, the problem (referred to as CS-RA) can be optimally solved by standard methods but at a high computational cost. We propose a two-step method, i.e., SBPA, based on decomposition theory that solves the CS-RA problem and notably reduces the computational complexity. For the NCS strategy, the NCS-RA problem is much less complex than the CS-RA problem and can be solved by adopting the SBPA method. To further reduce the computational cost associated to the NCS-RA problem, we propose a *rounding linear programming* (RLP) algorithm as an alternative to the usual integer constraint relaxation since this relaxation provides a noninteger solution that is not implementable in practice. The use of the RLP algorithm is extended to the CS case with the same advantages as for the NCS case. We also discuss whether a centralized or semidistributed approach is more adequate, depending on the encoding frame strategy.

The remainder of this paper is structured as follows. In Section II, we describe the system model. The optimization problem for CSs is formulated in Section III. The description of the two-step SBPA method for RA forms Section IV. Section V is devoted to the NCS scheme and the algorithm proposed for this case, including the concatenated case when the RLP algorithm is used. In Section VI, the suitability of the centralized and semidistributed schemes for implementation is discussed. The performance of the algorithms is presented in Section VII. Finally, in Section VIII, we extract some conclusions.

## II. SYSTEM MODEL

We consider a downlink single-hop OFDMA system with a BS and  $K$  active users that employ  $Q$  subcarriers. PHY- and data link (DL)-layer techniques are combined to optimize the goodput transmitted by the BS. At the PHY layer, adaptive modulation (comprising bit and power loading) and subchannel allocation determine the transmission scheme between the BS and the users. A forward error-correcting (FEC) scheme is also considered. At the DL layer, an ARQ protocol is applied to the frames exchanged between the BS and the users. A signaling scheme conveys information between the BS and the users to manage the adaptive transmission. The system is endowed with a feedback channel to generate CSIT, and efficient schemes to compress feedback information can be used to minimize the feedback rate [34]; this feedback channel is assumed to be error free. We assume that users have been previously scheduled

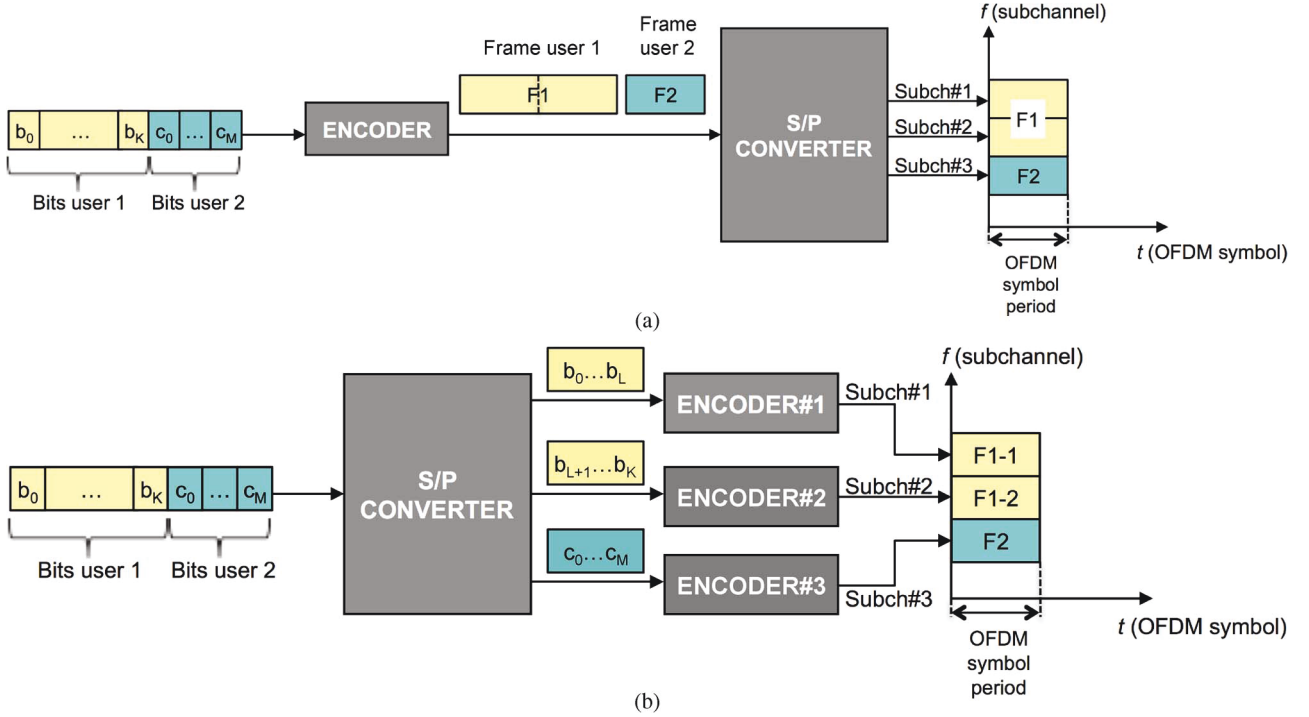


Fig. 1. Frame structure of the two encoding strategies. Example with two users and three subchannels. User 1 is assigned subchannels 1 and 2 and user 2 is assigned subchannel 3. (a) Concatenated frames. (b) Nonconcatenated frames.

within the actual frame. As multiuser diversity is exploited by optimizing the subchannel assignment to the scheduled users, a simple scheduling scheme, e.g., a round-robin (RR) method, can be used to avoid outage, although its design is outside the scope of this paper.

#### A. Frame Structure: Concatenation of Subchannels

Based on the channel coding theorem [35], the concatenation of subchannels enables transmission of large encoding frames (or blocks) at rate  $R$  to achieve an error rate as low as possible. In current standards such as IEEE 802.16e [13], there are several possibilities for the number of subchannels per frame, i.e., the number of subchannels that are encoded with the same encoder. The possibilities range from one to many. In this paper, we have selected the two extreme possibilities, namely, only one encoder or as many encoders as the subchannels, and the effect of subchannel concatenation is analyzed using the following two extreme strategies.

- 1) *CS strategy*. All the subchannels allocated to a given user are split into two frames, F1-1 and F1-2, and these frames are independently transmitted in  $s\#1$  and  $s\#2$ , respectively. Note that the number of bits to be transmitted to each user,  $K$  and  $M$  in our example, is a decision made once the RA problem has been solved.
- 2) *NCS strategy*. Each subchannel is independently encoded so that one encoding frame comprises only one subchannel per  $N_s$  OFDM symbols.<sup>1</sup> Hence, each user receives as many frames as assigned subchannels.

Note that the actual implementation choice and, thus, the achieved performance will lie in between of the results here presented for the CS and NCS strategies. The selection of

the configuration to be used is understood to be made by the network operator. Moreover, the results of this paper, among other system-level considerations, may be used to evaluate the alternatives considering that there exists a tradeoff between complexity and performance, given that the NCS approach is much less computationally complex at the expense of a higher error probability as the frames are in general shorter than in the CS case, and resulting in a lower received goodput.

The following example with two users and three subchannels shown in Fig. 1(a) and (b) illustrates the differences between both schemes where the frames are transmitted in one OFDM symbol, i.e.,  $N_s = 1$ . For the concatenated case [see Fig. 1(a)], the bits to be transmitted to a given user are encoded in only one frame, and this frame is accommodated into the assigned subchannels. In our example, user 1 is assigned subchannels  $s\#1$  and  $s\#2$ , and user 2 is assigned subchannel  $s\#3$ . For the nonconcatenated case [see Fig. 1(b)], the bits are encoded in as many frames as subchannels that have been assigned to the user. As user 1 has been assigned  $s\#1$  and  $s\#2$ , the information bits are independently transmitted in  $s\#1$  and  $s\#2$ , respectively. Note that the number of bits to be transmitted to each user,  $K$  and  $M$  in our example, is a decision made once the RA problem has been solved.

In summary, for the concatenated case, we have as many frames as users and only one encoder, and for the nonconcatenated case, we have as many frames as encoders and subchannels.

#### B. Physical Layer and Channel Model

Subchannel, bit and power allocation (SBPA) is performed at the PHY level, with MQAM per subcarrier. The  $Q$  subcarriers

<sup>1</sup> $N_s$  is in general a function of the user and the allocated subchannels.



are evenly grouped into subchannels; rather than allocating subcarriers, the BS performs subchannel allocation to each user's downlink. The BS has  $N = Q/J$  subchannels to allocate, with  $J$  subcarriers per subchannel. Notice that, although the bandwidth is assigned to users at a subchannel level, bit and power allocation is done with subcarrier granularity, preserving the benefits of bit allocation and simultaneously simplifying the allocation process.

The distribution of subcarriers among the subchannels may also be part of the system design. However, we assume a pre-determined subcarrier arrangement, which reduces the system model complexity since it is kept fixed over time. Moreover, subchannel allocation to the best user for goodput optimization endows the system with multiuser diversity. One possibility is to arrange contiguous subcarriers into one subchannel [36]. Alternatively, optimal subcarrier allocation to subchannels may be done at the expense of the complexity (see, for instance, [37]), although this is out of the scope of this paper. For evaluation purposes, in this paper, a random distribution of the subcarriers over the subchannels is assumed as it can provide some frequency diversity.

We assume a multipath block fading channel, where  $H_{knj}$  denotes the channel gain associated with subcarrier  $j$  of subchannel  $n$  assigned to user  $k$ , in short notation  $(k, n, j)$ , and it is modeled as a complex zero-mean Gaussian random variable invariant over at least one frame period of  $N_s$  OFDM symbols. We denote by  $p_{knj}$  and  $m_{knj}$  the power and the number of bits per symbol allocated to  $(k, n, j)$ , which are constant over a frame. The noise samples are assumed to be complex Gaussian random variables of zero mean and variance  $\sigma^2$ . The maximum transmit power of the BS is  $P_T$ , and it is distributed among the  $Q$  subcarriers.

### III. PROBLEM STATEMENT FOR CONCATENATED SUBCHANNELS

Here, we analyze the case of CSs, where each user receives one frame consisting of the concatenated signal of the assigned subchannels.

The multicarrier channel is viewed by each user as a channel whose error probability is the average BER over the user's transmission subcarriers. We define the set  $\mathcal{N}_L$  as the  $L = 2^N - 1$  possible combinations of subchannels, i.e., the elements of  $\mathcal{N}_L$  are groups of subchannels to be assigned to users per frame, and  $\mathcal{N}_l$  represents the  $l$ th element of  $\mathcal{N}_L$  with cardinality  $N_l$ . For example, if  $N = 2$  subchannels, we have  $L = 3$ , being  $\mathcal{N}_L = \{\mathcal{N}_1, \mathcal{N}_2, \mathcal{N}_3\} = \{\{1\}, \{2\}, \{1, 2\}\}$ , and  $N_1 = N_2 = 1$ ,  $N_3 = 2$ . The error-correcting mechanism is implemented by a frame-based convolutional coding of rate  $r$  with hard-decision Viterbi decoding. For a given frame, if  $\varepsilon_{av}(k, l)$  denotes the average BER associated to user  $k$  when assigned the subchannels of  $\mathcal{N}_l \in \mathcal{N}_L$ , it can be expressed as

$$\varepsilon_{av}(k, l) = \frac{1}{\sum_{n \in \mathcal{N}_l} \sum_{j=1}^J m_{knj}} \sum_{n \in \mathcal{N}_l} \sum_{j=1}^J m_{knj} \varepsilon_{knj} \quad (1)$$

where  $\varepsilon_{knj}$ , which is the uncoded BER associated to the hard decision-making on the framed bits on  $(k, n, j)$  for MQAM

with Gray mapping, is approximated by [38]

$$\varepsilon_{knj} = 0.2 \exp\left(\frac{1.6|H_{knj}|^2 p_{knj}}{(2^{m_{knj}} - 1)\sigma^2}\right). \quad (2)$$

The frame received by user  $k$  spreads over  $N_s(k, l)$  OFDM symbols, where  $l$  denotes that user  $k$  is allocated the subchannels of  $\mathcal{N}_l$ , and  $m_{knj}$  is invariant over  $N_s(k, l)$  OFDM symbols.

#### A. Goodput Formulation

We recall that the goodput is defined as the number of error-free information bits transmitted per unit of time. In this paper, we deal with the *goodput obtained per transmitted frame*, which is formulated as a function of the FSR. The FSR is defined as the probability that a frame is correctly received after Viterbi decoding. For each user, the transmission is organized in frames of  $N_f$  information bits, where one or more frames are received according to one of the strategies defined in Section II-A. The FSR expression can be approximated as a function of  $\varepsilon_{av}$  by [39]

$$\text{FSR}(\varepsilon_{av}(k, l)) = d \exp\left(-\left(c_\tau \varepsilon_{av}^\tau + c_{\tau-1} \varepsilon_{av}^{\tau-1} + \dots + c_1 \varepsilon_{av}\right)\right) \quad (3)$$

where parameters  $\tau$ ,  $d$ , and  $c_1, \dots, c_\tau$  are adjusted so that the approximated FSR fits the true FSR curve. These parameters depend on the convolutional code used and the frame length.

At the link layer, an ARQ protocol performs the retransmission of the frames that are incorrectly received by the user. Expressions of the goodput for different ARQ protocols have been obtained for point-to-point OFDM links in [40]. Without loss of generality, we adopt a selective-repeat ARQ (SR-ARQ) protocol as it is straightforward to modify the formulas for other ARQ protocols taking the given expressions. The received goodput  $\chi_{kl}$  associated with the allocation of the subset of subchannels  $\mathcal{N}_l$  to user  $k$  is then expressed for the SR-ARQ as

$$\chi_{kl} = r \left( \sum_{n \in \mathcal{N}_l} \sum_{j=1}^J m_{knj} \right) \text{FSR}(\varepsilon_{av}(k, l)). \quad (4)$$

In view of (4), adaptive coding may be also jointly used with adaptive SBPA for each transmitted frame according to (3). However, adaptive coding is useful when the bit allocation remains fixed for the whole frame [41], whereas our scheme enables adaptive bit allocation within the frame. Then, although adaptive coding may be explored in this context,  $r$  remains unchanged for all the transmissions and subchannels.

In summary, we have formulated the individual goodput  $\chi_{kl}$  when an ARQ is used by relating the BER averaged over the subcarriers of a set of subchannels  $\varepsilon_{av}(k, l)$  with the FSR.

#### B. Utility Function and Subchannel Assignment Variables

The utility function  $U$  to be maximized may be defined as the sum of the goodput obtained by each user, but this may

lead to unfair situations in which some users are penalized. To avoid this situation, we introduce fairness by using, for each user, utility function  $u_k$  [42] as follows:

$$u_k = \frac{\chi_k[m]}{\bar{\chi}_k[m]} \quad (5)$$

where  $\chi_k[m]$  is the goodput achieved in frame  $m$  by user  $k$ , and  $\bar{\chi}_k[m]$  is the average goodput of user  $k$  over a past window of length  $W$ , which is calculated as [43]

$$\bar{\chi}_k[m] = \left(1 - \frac{1}{W}\right) \bar{\chi}_k[m-1] + \frac{1}{W} \chi_k[m-1]. \quad (6)$$

This class of utility function is shown to be equivalent to the proportional fair scheduling proposed in [43] for large values of  $W$ . We now introduce fairness control by means of the real-valued parameter  $\alpha$  [32], and the BS utility is formulated as

$$U[m] = \sum_{k=1}^K \frac{\chi_k[m]}{\bar{\chi}_k^{(1-\alpha)}[m]} \quad (7)$$

where parameter  $\alpha$  can be tuned to determine the degree of fairness. If  $\alpha = 0$ , the users are expected to receive the goodput approximating proportional fairness allocation. At the other extreme,  $\alpha = 1$  is expected to approximate maximum goodput allocation policy. In the sequel, the temporal dependence with  $m$  is removed from the given expressions to alleviate the notation once it is established that the problem is solved at each frame  $m$ .

The assignment of subchannels to users is made using the binary variables  $\pi_{kl}$ , which indicate whether a group of subchannels  $\mathcal{N}_l$  is assigned to user  $k$  or not, i.e.,

$$\pi_{kl} = \begin{cases} 1, & \text{if set of subchannels } \mathcal{N}_l \text{ is allocated to user } k \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

It must be guaranteed that only one set of subchannels  $\mathcal{N}_l$  is assigned to each user within the current frame. Thus, only one of the values of  $\pi_{kl}$  is equal to 1 for user  $k$ , and  $\sum_{l=1}^L \pi_{kl} = 1$  must hold for each user  $k$ . We also define the binary variables  $\beta_{kn}$  to indicate if subchannel  $n$  is assigned to user  $k$ , i.e.,

$$\beta_{kn} = \begin{cases} 1, & \text{if subchannel } n \text{ is allocated to user } k \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

As subchannel  $n$  must be assigned to one and only one user,  $\sum_{k=1}^K \beta_{kn} = 1$  must hold  $\forall n$ .

These two groups of variables are related by  $\mathbf{B} = \mathbf{\Omega}\mathbf{\Pi}$ , where  $\mathbf{B}$  is a  $KN$  column vector with the  $\beta_{kn}$  variables,  $\mathbf{\Pi}$  is a  $KL$  column vector with the  $\pi_{kl}$  variables, and  $\mathbf{\Omega}$  is a  $KN \times KL$  matrix that makes the mapping of a group of subchannels onto the corresponding subchannels per user. This relationship makes it possible to solve the problem only in the  $\pi_{kn}$  variables since each constraint  $\sum_{k=1}^K \beta_{kn} = 1$  can be written as  $f_n(\mathbf{\Pi}) = \sum_{k=1}^K \beta_{kn} = 1$ , where if  $\beta_{kn}$  occupies the

$r$  position in vector  $\mathbf{B}$ , then  $\beta_{kn} = \omega_r \mathbf{\Pi}$ , with  $\omega_r$  being the  $r$ th row of  $\mathbf{\Omega}$ .<sup>2</sup>

### C. Problem Formulation

The optimization problem consists in maximizing the BS utility (7) finding the optimal bit, power level, and subchannel allocations (i.e.,  $m_{knj}$ ,  $p_{knj}$  and  $\pi_{kl}$ , respectively), and two necessary constraints are added: 1) The BS has the total power  $P_T$  to allocate among the subcarriers; and 2) each user  $k$  must receive a minimum required goodput denoted by  $\chi_{k \min}$ . The resulting problem is

$$\begin{aligned} \max_{\pi_{kl}, p_{klj}, m_{klj}} \quad & U = \sum_{k=1}^K \frac{\chi_k}{\bar{\chi}_k^{(1-\alpha)}} \\ \text{subject to} \quad & \sum_{k=1}^K \sum_{n=1}^N \sum_{j=1}^J p_{klj} \leq P_T \\ & \sum_{l=1}^L \pi_{kl} \chi_{kl} \geq \chi_{k \min}, \quad k = 1, \dots, K \\ & \sum_{l=1}^L \pi_{kl} = 1, \quad k = 1, \dots, K \\ & f_n(\mathbf{\Pi}) = 1, \quad n = 1, \dots, N \\ & \pi_{kl} \in \{0, 1\}, \quad k = 1, \dots, K, \quad l = 1, \dots, L. \end{aligned} \quad (10)$$

This and the forthcoming optimization problems are defined and solved on a per-frame basis.

Problem (10) is a nonlinear mixed integer programming (MIP) problem since  $p_{knj}$  and  $m_{knj}$  are real valued,  $\pi_{kl}$  is an integer, and the objective function is nonlinear. Nonlinear MIP problems with a large number of integer variables are characterized as  $\mathcal{NP}$ -complete, and exact methods, such as branch-and-bound, do not solve the problem efficiently [44]. Therefore, we propose a different approximation to the problem, as we describe in the following.

An analysis of our problem determines that the condition complicating the problem most is the total power constraint  $\sum_{k=1}^K \sum_{n=1}^N \sum_{j=1}^J p_{klj} \leq P_T$  as this condition couples the power variables of (10)  $\{p_{klj}\}$ . This type of constraint is referred to as ‘‘complicating constraints’’ [45] since, if this constraint were absent, the problem would be decoupled into  $K$  subproblems given that the objective function  $U$  is separable as it is the sum of the  $K$  per-user utility functions  $u_k$  and

<sup>2</sup>For instance, if  $K = 2$ ,  $N = 2$ , and for the definition of  $\mathcal{N}_L$  given in Section III, we have

$$\begin{pmatrix} \beta_{11} \\ \beta_{12} \\ \beta_{21} \\ \beta_{22} \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}}_{\mathbf{\Omega}} \begin{pmatrix} \pi_{11} \\ \pi_{12} \\ \pi_{13} \\ \pi_{21} \\ \pi_{22} \\ \pi_{23} \end{pmatrix}$$

with  $f_1 = \beta_{11} + \beta_{21} = \pi_{11} + \pi_{13} + \pi_{21} + \pi_{23}$ , and  $f_2 = \beta_{12} + \beta_{22} = \pi_{12} + \pi_{13} + \pi_{22} + \pi_{23}$ .

the other constraints are not coupling constraints among the users in  $\{p_{klj}\}$ . The conventional approach to this type of RA problems is the *decomposition theory* (see, for instance, [46]). The basis of decomposition is to decompose the original large problem into distributively solvable subproblems, which are then coordinated by a high-level master problem using some kind of signaling.

A first step toward achieving solvable subproblems is to use constant power allocation. Constant power allocation implies negligible throughput loss if the power is only allocated to channels with good channel gains [38], [47]. Given that, with multiuser adaptive RA, as it is our case, the channels are usually allocated to the users with favorable channel conditions, then power constant allocation has a negligible impact in the performance compared with the optimum [48], [49]. This is valid for throughput and is reasonably expected to fulfill also for the case of goodput optimization. Therefore, following [27], we relax the power constraint, assuming that  $P_T$  is equally divided among the  $N$  subchannels. The resulting total power allocated to group  $l$  formed by  $N_l$  subchannels is  $P_l = N_l(P_T/N)$ , and Constraint 1 becomes  $\sum_{k=1}^K \sum_{j=1}^J p_{klj} \leq P_l$ , giving the following RA problem:

$$\begin{aligned}
& \max_{\pi_{kl}, p_{klj}, m_{klj}} U = \sum_{k=1}^K \frac{\chi_k}{\bar{\chi}_k^{(1-\alpha)}} \\
& \text{subject to} \quad \sum_{k=1}^K \sum_{j=1}^{N_l J} p_{klj} \leq P_T \quad l = 1, \dots, L \\
& \quad \sum_{l=1}^L \pi_{kl} \chi_{kl} \geq \chi_{k \min} \\
& \quad \sum_{l=1}^L \pi_{kl} = 1, \quad k = 1, \dots, K \\
& \quad f_n(\mathbf{\Pi}) = 1, \quad n = 1, \dots, N \\
& \quad \pi_{kl} = \{0, 1\}, \quad k = 1, \dots, K, \quad l = 1, \dots, L
\end{aligned} \tag{11}$$

where  $\chi_k$  denotes the goodput received by user  $k$  and is calculated as  $\chi_k = \sum_{l=1}^L \pi_{kl} \chi_{kl}$ , with  $\chi_{kl}$  being the goodput given by (4). Even with the power constraint relaxation, (11) is still a nonlinear MIP problem with similar complexity compared with (10). In the following, we present a method that reduces this high complexity.

#### IV. SUBCHANNEL, BIT, AND POWER ALLOCATION FOR CONCATENATED SUBCHANNELS

We assume a centralized configuration given that the computational capacity associated to user devices is usually the limiting factor. Nevertheless, in Section VI, we discuss whether other configurations than the centralized scheme can be more suitable.

As pointed out earlier, the idea behind our approach is to decompose the original problem into several tractable subprob-

lems that can be then coordinated through a master problem. However, decomposition theory has been mainly used for continuous RA, and it cannot be in general straightforwardly employed for MIP problems. We propose the following two-step centralized scheme that first decouples the original problem (11) into  $KL$  subproblems ( $L$  possible combinations of subchannels times the number of users  $K$ ). The result of the first step is a matrix whose elements are the goodput per user and per group of subchannels. This matrix is used in the second step (master problem) to perform subchannel allocation. This scheme works as follows.

- 1) Each user performs channel estimation and provides CSI to the BS using the feedback channel.
- 2) The BS performs power and bit allocation for all users and every possible combination of subchannels, and it calculates the corresponding goodput vector  $\mathbf{x}_k$  for each user  $k$ . These goodput vectors form the goodput matrix  $\mathbf{X}$ .
- 3) The BS runs the subchannel allocation algorithm and sends the obtained SBPA to users in the forward channel.
- 4) The BS transmits the information to users according to the resulting allocation.

The two-step method, which is referred to as concatenated SBPA (C-SBPA), reduces the complexity of the original problem (11). The first step, which is detailed in Section IV-A and named the *goodput matrix calculation* (GMC), provides the power and bit allocation and goodput matrix  $\mathbf{X}$  by decomposing (11) into independent subproblems that can be computed in parallel. This first step corresponds to 2) in the previous enumerate. The second step is the *subchannel allocation problem* (SAP) and acts as the master problem, corresponding to 3).

*Remark 1:* Decomposition theory is based on finding alternative problem formulations that may reveal hidden decomposable structures, although representing the problem in a different way does not alter the optimal solution. In our proposed decomposition, the GMC subproblems provide the optimal solution, and the master problem (SAP) uses these solutions to optimally assign the subchannels. Hence, there is no loss of optimality due to this decomposition.

*Remark 2:* In general, decomposition-based algorithms iterate between the master problem (SAP) and the subproblems (GMC subproblems) until they converge. In our proposed method, such iterations are not necessary since there is no coupling variable between the SAP and GMC subproblems.

##### A. Goodput Matrix Calculation

The goal of this step is to decouple (11) in separable subproblems. This would be accomplished if the constraints  $\sum_{k=1}^K \sum_{j=1}^{N_l J} p_{klj} \leq N_l(P_T/N)$  are considered separately for user  $k = \{1, \dots, K\}$  and group of subchannels  $l = \{1, \dots, L\}$ ; therefore, we solve the  $KL$  subproblems as follows:

$$\begin{aligned}
& \max_{p_{klj}, m_{klj}} \chi_{kl} \\
& \text{s.t.} \quad \sum_{k=1}^K \sum_{j=1}^{N_l J} p_{klj} \leq N_l \frac{P_T}{N}
\end{aligned} \tag{12}$$

where the  $kl$ th GMC-C subproblem (12) has the goal of maximizing the goodput  $\chi_{kl}$  received by user  $k$  when it is allocated the  $l$ th group of subchannels given by (4), and the power constraint is for a given group of subchannels  $l$  formed by  $N_l$  subchannels. Therefore, each subproblem provides an  $L \times 1$  column goodput vector  $\mathbf{x}_k = [\chi_{k1}, \dots, \chi_{kL}]^T$  and the corresponding power and bit allocation, being the  $K \times L$  goodput matrix  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_K]^T$ .

For maximizing  $\chi_{kl}$ , we rely on the bit and power allocation algorithm EBPA described in [40]. This algorithm optimizes the goodput over a given set of subcarriers for the point-to-point OFDM downlink single-user case. We adapt this algorithm to perform the bit and power allocation within each group of subchannels for the OFDM downlink multiuser case. The modified algorithm, which is named the *modified bit and power allocation algorithm* (MBPA), is summarized in the following three steps.

- 1) For a group of subchannels  $l$  and user  $k$ , sort the subcarriers such that  $|H_{kl1}|^2 \geq \dots \geq |H_{klJ'}|^2$ , with  $J' = N_l J$ . Set  $j^* = J'$ .
- 2) Solve

$$\begin{aligned} & \frac{2^{m_{klj^*}}}{|H_{klj^*}|^2} \left( \sum_{i=1}^{J'} \log_2 \left( \frac{|H_{kli}|^2}{|H_{klj^*}|^2} \right) + m_{klj^*} J' \right) \\ & \times (\tau c_\tau \varepsilon_{\text{av}}^\tau + \dots + c_1 \varepsilon_{\text{av}}) \left( \frac{c_2 P_l \ln 2}{\sigma_n^2} \right) \\ & = \left( J' \frac{2^{m_{klj^*}}}{|H_{klj^*}|^2} - \sum_{i=1}^{J'} \frac{1}{|H_{kli}|^2} \right)^2 \end{aligned} \quad (13)$$

where  $P_l = N_l(P_T/N)$ . If there is no positive solution for  $m_{klj^*}$ , then  $m_{klj^*} = 0$ ,  $j^* \leftarrow j^* - 1$ , and go to Step 2. Else, go to Step 3.

- 3)  $\forall j \leq j^* : m_{klj} = \log_2((2^{m_{klj^*}} |H_{klj}|^2) / (|H_{klj^*}|^2))$ ,  $p_{klj} = ((2^{m_{klj}} - 1) \sigma_n^2) / (c_2 |H_{klj}|^2) P_l / \sum_{i=1}^{J'} (((2^{m_{kli}} - 1) \sigma_n^2) / (c_2 |H_{kli}|^2))$ .

The obtained  $\{m_{klj}\}$  and  $\{p_{klj}\}$  guarantee that the goodput calculated for the set of subchannels  $\mathcal{N}_l$  is near optimal. Nevertheless,  $\{m_{klj}\}$  are real valued, and for feasible implementations, the number of bits per symbol must take discrete values for MQAM modulations. Hence,  $m_{klj}$  is obtained by rounding down the real-valued  $m_{klj}$  to the nearest element of the set of allowed integer values  $\mathcal{M}$  as the loss in terms of goodput is almost negligible [40]. By inserting the obtained  $m_{klj}$  and  $p_{klj}$  into (1)–(4), we get each  $\chi_{kl}$  of  $\mathbf{X}$ .

### B. Subchannel Allocation Problem

This step consists in finding the optimal vector  $\mathbf{\Pi} = [\pi_{11}, \dots, \pi_{KL}]$ , and a new optimization problem (14) is derived from (11) once  $\mathbf{X}$  is known. To do this, we impose that  $\chi_{kl}$  is one of the elements of  $\mathbf{X}$ , instead of considering  $m_{klj}$  and  $p_{klj}$  as variables, and the new problem is

$$\max_{\pi_{kl}} U = \sum_{k=1}^K \frac{1}{(1-\alpha)^{\chi_k}}$$

$$\begin{aligned} \text{subject to } & \sum_{l=1}^L \pi_{kl} \chi_{kl} \geq \chi_{k \min}, \quad k = 1, \dots, K \\ & \chi_{kl} \in \mathbf{x}_k^T, \quad k = 1, \dots, K \\ & \sum_{l=1}^L \pi_{kl} = 1, \quad k = 1, \dots, K. \\ & f_n(\mathbf{\Pi}) = 1, \quad n = 1, \dots, N \\ & \pi_{kl} = \{0, 1\}, \quad k = 1, \dots, K, \quad l = 1, \dots, L. \end{aligned} \quad (14)$$

In the literature, we can find different types of algorithms to optimally solve the integer programming (IP) problem (14), yet two of them are mainly used [50]: branch-and-bound algorithms and cutting plane methods. Problem (14) is  $\mathcal{NP}$ -hard, none of these approaches solves it polynomially, and standard branch-and-bound algorithms are commonly preferred for the sake of simplicity. Still, finding the optimal solution of (14) implies a high computational cost. In the following, we introduce the NCS strategy that significantly reduces the subchannel allocation complexity.

### V. SUBCHANNEL, BIT, AND POWER ALLOCATION FOR NONCONCATENATED SUBCHANNELS

We now consider that the bits corresponding to different subchannels are encoded in different frames, i.e., one subchannel per frame. In other words, the subchannels are encoded independently, and  $\mathcal{N}_L = \{1, \dots, N\}$ . We use the nonconcatenated SBPA (NC-SBPA) to refer to this approach. Both transmitter (BS) and receiver (user) process each subchannel separately, requiring as many encoders and decoders as subchannels, and the system model is adapted as follows.

The formulation is based on the goodput received by user  $k$  when allocated subchannel  $n$ , which is denoted by  $\chi_{kn}$ , resulting in  $N_s(k, n)$  OFDM symbols per frame. This goodput  $\chi_{kn}$  is calculated as

$$\chi_{kn} = r \left( \sum_{j=1}^J m_{knj} \right) \text{FSR}(\varepsilon_{\text{av}, kn}) \quad (15)$$

assuming  $m_{knj}$  is constant over the  $N_s(k, n)$  OFDM symbols. Contrary to the CS case, only the subchannel assignment  $\beta_{kn}$  variables (9) are necessary, and we can ignore the  $\pi_{kn}$  variables, by calculating the total goodput received by user  $k$  as the sum of goodput per subchannel  $\chi_k = \sum_{n=1}^N \beta_{kn} \chi_{kn}$ . By properly modifying (1)–(3), the expressions for the FSR and the average BER are

$$\varepsilon_{\text{av}}(k, n) = \frac{1}{\sum_{j=1}^J m_{knj}} \sum_{j=1}^J m_{knj} 0.2 \exp \left( -\frac{1.6 |H_{knj}|^2 p_{knj}}{(2^{m_{knj}} - 1) \sigma^2} \right) \quad (16)$$

$$\text{FSR}(\varepsilon_{\text{av}}) = d \exp \left( - (c_\tau \varepsilon_{\text{av}}^\tau + c_{\tau-1} \varepsilon_{\text{av}}^{\tau-1} + \dots + c_1 \varepsilon_{\text{av}}) \right) \quad (17)$$



which results in the utility function  $U = \sum_{j=1}^K (1/(\bar{\chi}_k^{(1-\alpha)})) \sum_{n=1}^N \beta_{kn} \chi_{kn}$ . Assuming again that  $P_T$  is equally distributed among the  $N$  subchannels, the new problem is

$$\begin{aligned}
& \max_{\beta_{kn}, p_{knj}, m_{knj}} U = \sum_{k=1}^K \frac{1}{\bar{\chi}_k^{(1-\alpha)}} \sum_{n=1}^N \chi_{kn} \beta_{kn} \\
& \text{subject to} \quad \sum_{k=1}^K \sum_{j=1}^J p_{knj} \leq \frac{P_T}{N}, \quad n = 1, \dots, N \\
& \quad \sum_{k=1}^K \beta_{kn} = 1 \quad \forall n = 1, \dots, N \\
& \quad \sum_{n=1}^N \beta_{kn} \chi_{kn} \geq \chi_{k \min} \quad \forall k = 1, \dots, K \\
& \quad \beta_{kn} = \{0, 1\}, \quad k = 1, \dots, K, \quad n = 1, \dots, N
\end{aligned} \tag{18}$$

which is solved by adapting the two-step approach of Section IV, as we detail in the following.

#### A. GMC and SAP for NC-SBPA

For the GMC step, (18) is decoupled for user  $k = \{1, \dots, K\}$  and subchannel  $n = \{1, \dots, N\}$  into  $KN$  separable subproblems of the form

$$\begin{aligned}
& \max_{p_{knj}, m_{knj}} \chi_{kn} \\
& \text{s.t.} \quad \sum_{k=1}^K \sum_{j=1}^J p_{knj} \leq \frac{P_T}{N}
\end{aligned} \tag{19}$$

where the  $(k, n)$ th GMC-NC subproblem (19) has the goal of maximizing the goodput  $\chi_{kn}$  received by user  $k$  when he is allocated the  $n$ th subchannel. Now, the  $k$ th row of goodput matrix  $\mathbf{X}$  is  $\mathbf{x}_k^T = [\chi_{k1}, \dots, \chi_{kN}]$ ; hence,  $\mathbf{X}$  has  $K \times N$  dimensionality. Proceeding similarly to Section IV, the bit and power allocations  $m_{knj}$  and  $p_{knj}$ , respectively, are found, and each element of  $\mathbf{X}$  is obtained by inserting (16)–(17) into (15).

At the second step (SAP), we assign the subchannels to users, finding the optimal  $\beta_{kn}$  elements of  $\mathbf{B}$ , with  $\mathbf{B}$  being the  $K \times N$  subchannels assignment matrix. Given that  $L \gg N$  and that we avoid the calculation of matrix  $\mathbf{\Pi}$ , this new problem is much less complex than (14), i.e.,

$$\begin{aligned}
& \max_{\beta_{kn}} U = \sum_{k=1}^K \frac{1}{\bar{\chi}_k^{(1-\alpha)}} \sum_{n=1}^N \chi_{kn} \beta_{kn} \\
& \text{subject to} \quad \sum_{k=1}^K \beta_{kn} = 1 \quad \forall n = 1, \dots, N \\
& \quad \sum_{n=1}^N \beta_{kn} \chi_{kn} \geq \chi_{k \min} \quad \forall k = 1, \dots, K \\
& \quad \beta_{kn} = \{0, 1\}, \quad k = 1, \dots, K, \quad n = 1, \dots, N
\end{aligned} \tag{20}$$

where  $\chi_{kn}$  are given from  $\mathbf{X}$ . Two standard approaches can be applied to solve (20). With the *IP solution*, for which the same discussion as in Section IV applies, the subchannels assignment matrix  $\mathbf{B}_{\text{IP}}$  represents the solution to (20) directly obtained using branch-and-bound algorithms, and  $U(\mathbf{B}_{\text{IP}})$  refers to the corresponding value of the utility function. Nevertheless, this optimal approach still has a high computational cost. *Integer constraint relaxation*, which is widely used for solving binary integer programming problems, can be applied to obtain a linear programming (LP) problem, where  $\beta_{kn} \in [0, 1]$  substitutes constraint  $\beta_{kn} \in \{0, 1\}$ . The LP problem solution, which is denoted by  $\mathbf{B}_{\text{LP}}$ , provides the utility value  $U(\mathbf{B}_{\text{LP}})$ , which is an upper bound for  $U(\mathbf{B}_{\text{IP}})$ . In the LP literature, primal–dual interior point algorithms are preferred [51], and among them, the algorithm of [52] has proved very successful in practice. However, the solution to the relaxed problem is not valid for practical purposes:  $\beta_{kn}$  must take real values, and although it can be interpreted as the time-sharing factor for the  $k$ th user of the  $n$ th subchannel, this is not coherent with the original problem formulation. An integer solution is obtained if these values are rounded, as is similarly done in [1], but there is no guarantee about satisfying the problem constraints, particularly the minimum goodput constraints. Still, this noninteger solution  $\mathbf{B}_{\text{LP}}$  can be used to generate an integer solution, as we describe in the following.

#### B. Rounded Linear Programming Algorithm

Given the complexity of the branch-and-bound approach and that the relaxed solution is not suitable for practical implementations, we propose an algorithm, which is referred to as *rounded linear programming* (RLP), having the following characteristics:

- 1) provides a valid integer solution, which is denoted by  $\mathbf{B}_{\text{RLP}}$ , starting from the  $\mathbf{B}_{\text{LP}}$  solution;
- 2) fulfills the constraint on the minimum required goodput;
- 3) achieves a value of utility function  $U(\mathbf{B}_{\text{RLP}})$  as close as possible to  $U(\mathbf{B}_{\text{LP}})$ ;
- 4) has the inherent low complexity of the LP approach.

We denote this approach as NC-RLP. Note that simply rounding  $\mathbf{B}_{\text{LP}}$  does not ensure to always satisfy characteristics 2 and 3. The RLP algorithm is based on the observation that the allocation variables  $\beta_{kn}$  corresponding to  $\mathbf{B}_{\text{IP}}$  and  $\mathbf{B}_{\text{LP}}$  largely coincide. We refer here to the element  $(k, n)$  of  $\mathbf{B}_{\text{LP}}$  and  $\mathbf{B}_{\text{RLP}}$  as  $\beta_{\text{LP}}(k, n)$  and  $\beta_{\text{RLP}}(k, n)$ , respectively.

The pseudocode of the algorithm and its main steps are provided in the following. The number of elements of set  $\mathcal{S}$  is denoted by  $\|\mathcal{S}\|$ . We define  $\mathcal{S}_K$  as the set of users and  $\mathcal{S}_N$  as the set of available subchannels (Step 1). Subchannels are previously assigned to the user with highest  $\beta_{\text{LP}}$  for each subchannel (Step 2). This may cause that the minimum goodput constraint is not achieved for some users. Next, subchannels are only assigned to users who obtain the minimum required goodput with a single allocated subchannel (Step 3), and these users form the set  $\mathcal{S}'_K$ , with  $\mathcal{S}'_N$  being the set of subchannels allocated to users of  $\mathcal{S}'_K$ . We form two more sets (Step 3), i.e.,  $\mathcal{S}''_K = \mathcal{S}_K - \mathcal{S}'_K$  and  $\mathcal{S}''_N = \mathcal{S}_N - \mathcal{S}'_N$ , that represent,

respectively, the users whose resulting goodput, with single subchannel allocation, is less than  $\chi_{k \min}$ , and the subchannels that are still to be assigned. Now, the subchannels of  $\mathcal{S}_N''$  have to be allocated to users of  $\mathcal{S}_K''$ . The key for this allocation is to assign each subchannel  $n''$  of  $\mathcal{S}_N''$  to the user  $k'' \in \mathcal{S}_K''$  obtaining the maximum goodput at  $n''$ , but only considering the  $\mathbf{B}_{LP}$  solution (Steps 4–6). Among the users of  $\mathcal{S}_K''$  who share  $n''$  in the  $\mathbf{B}_{LP}$  solution (users  $k''$  such that  $\beta_{LP}(k'', n'') \neq 0$ ), we select the user  $k^*$  that obtains the largest goodput from  $n''$ . However, some users may still not achieve  $\chi_{k \min}$ . We define the set  $\mathcal{S}_K'''$  containing those users (Step 7),  $\mathcal{S}_K''' = \{k \in \mathcal{S}_K'' \mid \chi_k < \chi_{k \min}\}$ , and the solution  $\mathbf{B}_{RLP}$  is refined in two phases.

The coarse refinement of the allocation (Steps 8–18) is fulfilled if the number of subchannels to be reallocated is greater than one ( $\|\mathcal{S}_N''\| > 1$ ). Swapping the subchannels between  $k \in \mathcal{S}_K'''$  and the remaining users of  $\mathcal{S}_K'''$  intends to achieve  $\chi_k > \chi_{k \min}$  for all users, with a similar criterion of that used in Steps 4–6. A later fine-tuning (Steps 19–33) may be necessary if  $\mathcal{S}_K'''$  is not empty yet. Swapping the subchannel between pairs of users is performed, based on ordering the subchannel/user pairs according to the heuristic parameter  $v$ . This parameter guarantees a deviation as small as possible from the  $\mathbf{B}_{LP}$  solution. Thus, for calculating  $v$ , we retake  $\mathbf{B}_{LP}$ ; all pairs  $(k, n)$ , such that  $\text{round}(\beta_{LP}(k, n)) = 1$  in step 1, are now considered for subchannel reallocation. We define  $v$  for these pairs  $(k, n)$  as  $v(k, n) = ((\chi_{kn} - \chi_{k \min})/\hat{\chi}_k)$ , with  $\hat{\chi}_k$  being the total goodput obtained by user  $k$  after Step 6 (if  $\hat{\chi}_k = 0$ ,  $v(k, n) = \chi_{kn} - \chi_{k \min}$  is used), and the first option for user  $k_1$  not achieving  $\chi_{k \min}$  is to swap his subchannel  $n_1$  with user  $k$  with maximum  $v$ .

---

### Algorithm 1 RLP

---

1: Initialize  $\mathcal{S}_K = \{1, \dots, K\}$ ,  $\mathcal{S}_N = \{1, \dots, N\}$ , and  $\mathbf{B}_{RLP}$  as rounded  $\mathbf{B}_{LP}$ , and do,  $\forall n$ ,

$$\beta_{RLP}(k, n) = \begin{cases} 1, & \text{if } k = \arg \max_{k \in \mathcal{S}_K} \beta_{LP}(k, n) \\ 0, & \text{otherwise.} \end{cases}$$

2: Set  $\mathcal{A} = \{(k, n) \mid \beta_{RLP}(k, n) = 1, \chi_{kn} \geq \chi_{k \min}\}$ . Update the initial assignment as

$$\beta_{RLP}(k, n) = \begin{cases} 1, & \text{if } (k, n) \in \mathcal{A} \\ 0, & \text{otherwise.} \end{cases}$$

3: Set  $\mathcal{S}_K' = \{k \in \mathcal{S}_K : \exists n \in \mathcal{S}_N \mid (k, n) \in \mathcal{A}\}$  and  $\mathcal{S}_K'' = \mathcal{S}_K - \mathcal{S}_K'$ . Set also  $\mathcal{S}_N' = \{n \in \mathcal{S}_N : \exists k \in \mathcal{S}_K \mid (k, n) \in \mathcal{A}\}$  and  $\mathcal{S}_N'' = \mathcal{S}_N - \mathcal{S}_N'$ .

4: **for** each  $n'' \in \mathcal{S}_N''$  **do**

5:  $\beta_{RLP}(k^*, n'') = 1 \Leftrightarrow k^* = \arg \max_{k'' \in \mathcal{S}_K''} \chi_{k''n''}$  and  $\beta_{LP}(k^*, n'') \neq 0$

6: **end for**

7: Identify users not receiving  $\chi_{k \min}$ :  $\mathcal{S}_K''' = \{k \in \mathcal{S}_K'' \mid \chi_k < \chi_{k \min}\}$

8: **if**  $\|\mathcal{S}_N''\| > 1$  **then**

9: **for all**  $k \in \mathcal{S}_K'''$  **do**

10:  $\mathcal{S}_k^n = \mathcal{S}_N'' \cup \{n \in \mathcal{S}_N \mid \beta_{RLP}(k, n) = 1\}$

11: **while**  $\chi_k < \chi_{k \min}$  and  $\mathcal{S}_k^n \neq \emptyset$  **do**

12: Take  $n \in \mathcal{S}_k^n$  randomly and remove  $n$  from  $\mathcal{S}_k^n$

13: Find  $r \in \{n \in \mathcal{S}_N \mid \beta_{RLP}(q, n) = 1, q \in \mathcal{S}_K''\}$ ,  $n^* = \arg \max_r \chi_{kr}$ ,  $k^* = \arg \max_q \{q \in \mathcal{S}_K'' \mid \beta_{RLP}(q, n^*) = 1\}$

14: Swap subchannels  $n$  and  $n^*$  between  $k$  and  $k^*$ :  $\beta_{RLP}(k, n) = 0$ ,  $\beta_{RLP}(k, n^*) = 1$ ;  $\beta_{RLP}(k^*, n^*) = 0$ ,  $\beta_{RLP}(k^*, n) = 1$

15: **end while**

16: Update  $\mathcal{S}_K'''$

17: **end for**

18: **end if**

19: **if**  $\mathcal{S}_K''' \neq \emptyset$  **then**

20: **for all**  $k_1 \in \mathcal{S}_K'''$  **do**

21: **find** all pairs  $(k, n)$ ,  $k \neq k_1$ , such that rounded  $\beta_{LP}(k, n) = 1$ , and form  $\mathcal{S}_{k_1}^{n_1} = \{n \in \mathcal{S}_N \mid \beta_{RLP}(k_1, n) = 1\}$

22: Calculate  $v$  for all  $(k, n)$  pairs obtained in step 21:  $v(k, n) = ((\chi_{kn} - \chi_{k \min})/\hat{\chi}_k)$ , with  $\hat{\chi}_k$  being the total goodput obtained by user  $k$  after step 6.

23: Form  $\mathbf{v}^T = (v_1, \dots, v_j, \dots, v_{N - \|\mathcal{S}_{k_1}^{n_1}\|})$ , such that  $v_1 \geq \dots \geq v_{N - \|\mathcal{S}_{k_1}^{n_1}\|}$

24: Initialize  $j = 1$ . Take  $n_1$  as the first element of  $\mathcal{S}_{k_1}^{n_1}$

25: **while**  $\chi_{k_1} < \chi_{k \min}$  and  $\mathcal{S}_{k_1}^{n_1} \neq \emptyset$  **do**

26: Swap  $n_1$  with the corresponding  $n_j$  of the  $j$ th element of  $\mathbf{v}^T$ ,  $v_j$ , between  $k_1$  and  $k_j$

27: **if**  $\chi_k < \chi_{k \min}$  **then**

28: Undo swap;  $j = j + 1$

29: **end if**

30: Take next  $n_1 \in \mathcal{S}_{k_1}^{n_1}$  if  $j = N - \|\mathcal{S}_{k_1}^{n_1}\| + 1$

31: **end while**

32: **end for**

33: **end if**

---

### C. RLP for Concatenated Subchannels

At this point, an interesting question emerges. Can the subchannel allocation obtained by the RLP algorithm be used in CS transmission without significant performance loss with respect to the optimum allocation of Section IV? This implies that the number of subproblems to be solved in the GMC step decreases from  $KL$  to  $K$ , therefore notably reducing the processing time and achieving a very good tradeoff between complexity and performance, as we show in Section VII. We refer to this suboptimal approach as concatenated RLP (C-RLP) algorithm.

The algorithm works as follows. First, the RLP algorithm is used to obtain the subchannel allocation  $\mathbf{B}_{RLP}$ . Second, the bit and power allocation  $\{m_{klj}, p_{klj}\}$  is calculated only for the concatenation of those allocated subchannels resulting from the RLP algorithm, i.e., considering  $\mathbf{B}_{RLP}$ , and not for the total  $KL$  possible combinations of subchannels. In this step, the MBPA in Section IV-A is used to obtain  $\{m_{klj}, p_{klj}\}$ . Note that  $l$  represents the set of subchannels allocated to user  $k$  that are transmitted concatenated; therefore, only  $K$  allocation pairs  $\{m_{klj}, p_{klj}\}$  are calculated. Finally, the concatenation strategy is used to transmit.

---

**Algorithm 2** C-RLP
 

---

- 1: Run the RLP algorithm to obtain the subchannel allocation  $\mathbf{B}_{\text{RLP}}$ .
  - 2: **for all**  $k = 1, \dots, K$  **do**
  - 3:    $\mathcal{N}_l = \{n : \beta_{\text{RLP}}(k, n) = 1\}$ .
  - 4:   Obtain  $\{m_{klj}, p_{klj}\}$  by using the MBPA algorithm corresponding to the subchannel concatenation of  $\mathcal{N}_l$ .
  - 5: **end for**
  - 6: Transmit the bits using concatenation.
- 

## VI. IMPLEMENTATION: CENTRALIZED VERSUS SEMIDISTRIBUTED CONFIGURATION

The methods developed earlier can be implemented in centralized or distributed mode, both endowed with a feedback channel. Compared with the centralized approach assumed in Section IV, in some situations, the distribution of the tasks among users and the BS can be advantageous. In this case, a two-step semidistributed approach may be suitable. This scheme differs from the centralized scheme of Section IV in the calculation of the goodput vectors  $\{\mathbf{x}_k\}$  (Step 2). Now, each user  $k$  executes the corresponding GMC stage to obtain  $\mathbf{x}_k$ . As a result, the  $K$  goodput vectors are obtained in parallel. Each user's device then sends the necessary information ( $\mathbf{x}_k$  and bit and power allocation) to the BS through the feedback channel. Next, the BS runs the subchannel allocation algorithm and sends the resulting subchannel allocation to users through the forward channel. Finally, the BS transmits the information bits in the frames organized accordingly.

It can be verified whether the exchanged information (number of variables) is lower in the semidistributed or in the centralized scheme. The semidistributed configuration [see Fig. 2(a)] requires for the concatenated scheme  $Q$  bit allocation and  $Q$  power allocation variables per group of subchannel and per user, and the goodput vector  $\mathbf{x}_k$  with  $L$  components, for each user. Having  $K$  users and  $L$  possible groups of subchannels,  $QKL$  power variables,  $QKL$  bit allocation variables, and  $KL$  goodput vector components are received by the BS from the  $K$  users. The BS informs the users their subchannel allocation  $\beta_{kn}$ , which are used for this  $KN$   $\beta$  variables. This results in a total of  $(2Q + 1)KL + KN$  variables. Similarly, for the semidistributed configuration with a nonconcatenated scheme,  $2QKN + 2KN$  variables are required, given that the goodput vector  $\mathbf{x}_k$  has now  $N$  components, and the bit and power allocation variables are now determined for each of the  $N$  subchannels individually.

For the centralized scheme [see Fig. 2(b)], the same number of variables is exchanged independently of the selected transmission scheme. The CSI requires  $QKN$  parameter values for the channel gains  $\{H_{knj}\}$ , and  $2QKN$  bit and power allocation variables, and  $KN$  subchannel variables are required, giving a total of  $3QKN + KN$ .

The given results are summarized in Table I and then analyzed considering the usual case that  $Q \gg 1$ . For the concatenated case, the difference between the semidistributed and

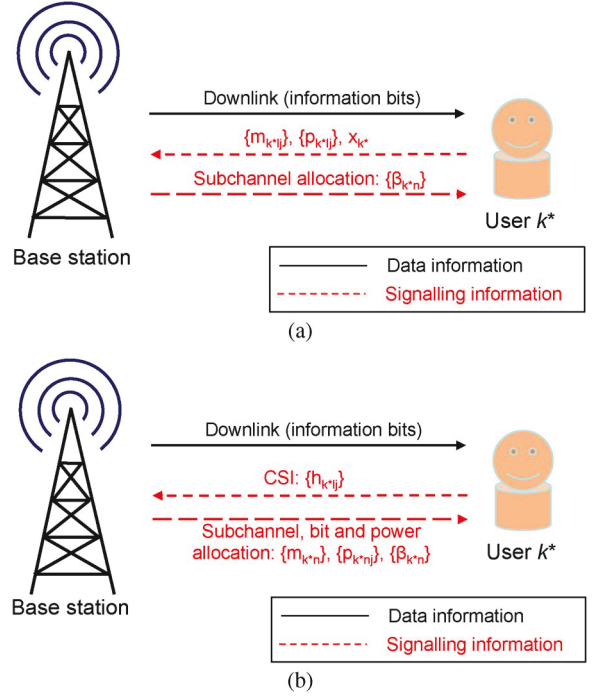


Fig. 2. Signaling exchange (variables and CSI) between the BS and the users depending on the configuration. (a) Semidistributed configuration with a concatenated transmission scheme. For the nonconcatenated scheme, the exchanged information is the same, replacing  $m_{k^*lj}$  and  $p_{k^*lj}$  by  $m_{k^*nj}$  and  $p_{k^*nj}$ , respectively. (b) Centralized configuration with a concatenated or nonconcatenated transmission scheme.

TABLE I  
COMPARATIVE OF THE EXCHANGED NUMBER OF VARIABLES FOR THE SEMIDISTRIBUTED AND CENTRALIZED CONFIGURATIONS

	Concatenated	Non-concatenated
Centralized	$3QKN + KN$	$3QKN + KN$
Semi-distributed	$(2Q + 1)KL + KN$	$2QKN + 2KN$

centralized schemes is  $(2L - 3N)QK$ . Recalling that  $L = 2^N - 1$ , we observe that, in this case, the parallel computation of power and bit allocation done by each user in the semidistributed configuration likely leads to faster and more robust systems at the expense of a larger information exchange. On the other hand, for the nonconcatenated case, the difference is  $(1 - Q)KN$ ; therefore, the semidistributed scheme involves less information exchange and can be preferable. If the feedback channel is nonerror free, we can hypothesize that the centralized configuration would be preferred when the concatenation scheme is adopted since the lower number of exchanged variables the lower number of expected errors are due to feedback channel unreliability. Analogously, the semidistributed configuration would be preferred when the nonconcatenated scheme is adopted.

Computational capabilities must be also taken into consideration. While BSs are in general computationally powerful, user's units may have processing limitations. Therefore, the semidistributed configuration could be only adopted if certain computational capacity is guaranteed in the user end.

Note that, with respect to the presented resolution methods, performance is not altered whether the centralized or

the semidistributed approach is used, provided that feedback channel is assumed to be error free.

## VII. NUMERICAL RESULTS

Simulation results are presented to evaluate and compare the performance of the algorithms for a fast varying channel, where the channel gain coefficients  $H_{knj}$  vary randomly and independently from one frame to the following frame. Each user is assumed to experience an independent complex Gaussian channel gain per subcarrier, and path loss is not considered. The simulation scenario is based on the 1024 mode defined in [13], which provides 768 data subcarriers and 48 subcarriers per subchannel ( $N = 16$  subchannels), with a bandwidth of 10937.5 Hz per subcarrier and OFDM symbol period of 102.9  $\mu$ s. Convolutional coding with generator polynomials [133, 171] is used with a code rate of 1/2, adaptive modulation is performed by 4/16/64QAM (i.e.,  $\mathcal{M} = \{0, 2, 4, 6\}$ ) as per [13], [14], and the FSR expression, given by (3), is [40]  $\text{FSR} = 0.879 \exp(1166326\epsilon_{\text{av}}^4 - 77313\epsilon_{\text{av}}^3 + 2216\epsilon_{\text{av}}^2 - 28.14\epsilon_{\text{av}})$ , for a frame length of 4096 bits. We consider a noise spectral density of  $N_0 = -80$  dBm/Hz, with  $N_0 = \sigma^2/B$  and  $B = 5.6$  MHz, and the ratio of maximum received power per subchannel to noise spectral density is 40 dB  $\cdot$  Hz. The minimum required goodput per user is 875 kb/s (90 information bits per OFDM symbol). The average goodput  $\bar{\chi}_k$  given by (5) is updated at each frame with  $W = 50$ . For all simulations, the results are obtained by averaging over 2000 random channel realizations.

We first assess the validity of the power constraint relaxation by simulating the optimal solution, which is compared with the C-SBPA solution. This optimal solution is obtained by solving the goodput optimization problem when the total power  $P_T$  is distributed within all the available subcarriers according to the MBPA algorithm, and the subcarriers are individually assigned to the users. Instead of the scenario described earlier, we have considered a simpler simulation scenario with four users and 48 subcarriers, with eight subcarriers per subchannel. We can observe in Fig. 3(a) that the loss of C-SBPA with respect to the optimal (labeled as OPT) is less than 0.5% in terms of the maximum achievable goodput. In terms of fairness, it is shown in Fig. 3(b) that, again, the loss of C-SBPA has very little significance with respect to the optimum. To compare the computational time, we have considered that the parallel computation of the  $K \times L$  problems may become unrealistic; however, it is very reasonable to assume that  $K$  processors (at either the BS or considering each user) compute sequentially the corresponding  $L$  subproblems. This is labeled as parallel in the figure, whereas the case that the  $K \times L$  subproblems are solved by the same entity is the “sequential” case. The results presented in Fig. 3(c) reveals that, even in this simple example, it takes above five times more to compute the optimal solution, and the C-SBPA sequential case outperforms the optimal case. Therefore, hereafter, we consider C-SBPA as the upper bound for our simulation scenario.

Simulations show the difference between the concatenated and nonconcatenated proposed methods in Figs. 4–6 for the case of four users and the 1024 mode parameters. We see

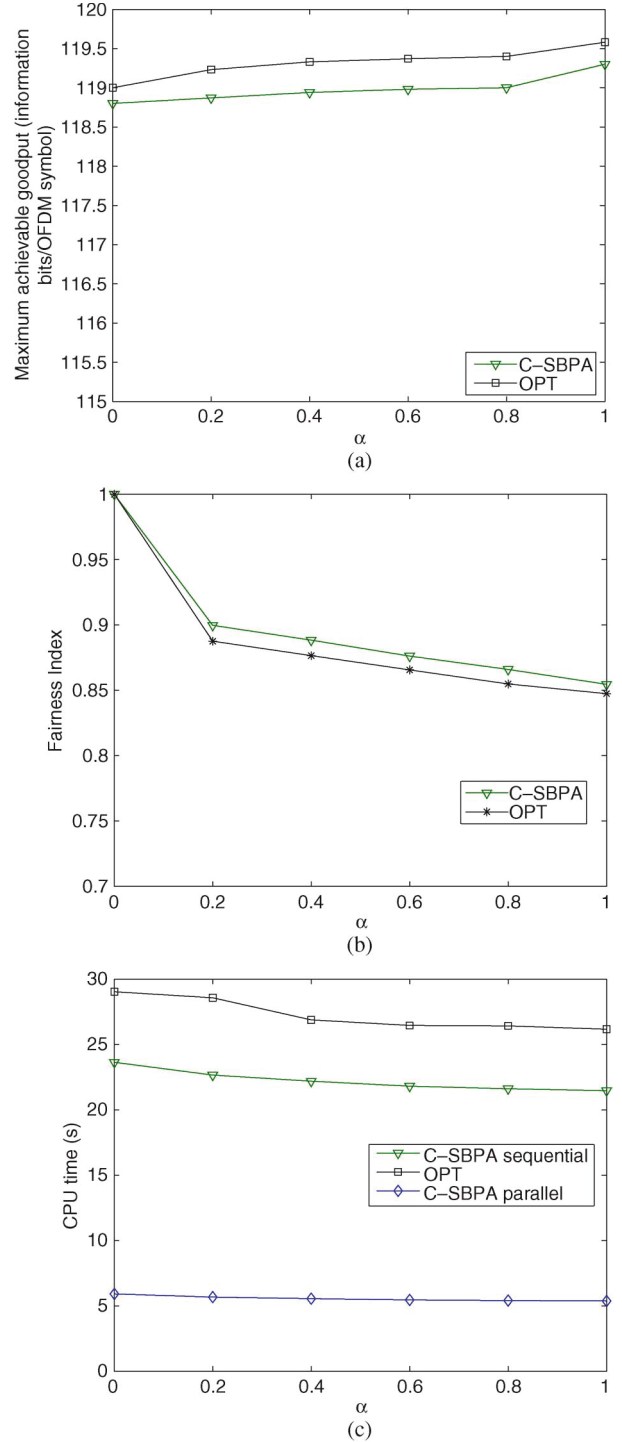


Fig. 3. Comparison between optimal power allocation (no power constraint relaxation) OPT and C-SBPA. (a) Maximum achievable goodput. (b) Fairness. (c) Complexity as CPU time.

that, although the C-SBPA algorithm provides slightly better results, as Figs. 4 and 5 show, this approach is intractable in practice since the computational time required is much larger even if it is compared with NC-SBPA, the most complex of the other algorithms, as Fig. 6 illustrates. Given that C-SBPA has been shown to be very close to the optimal in Fig. 3(a) and (b), it is a valid benchmark for the other proposed algorithms. Therefore, we focus on the nonconcatenated approaches



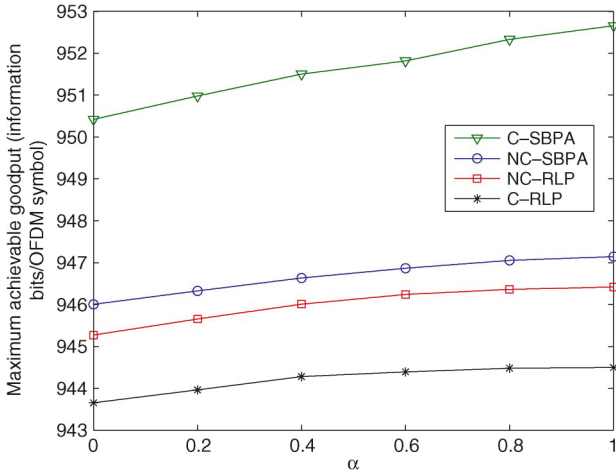


Fig. 4. Maximum averaged sum goodput received by users with C-SBPA, NC-SBPA, NC-RLP, and C-RLP algorithms, with  $\alpha \in [0, 1]$ , for the case of four users.

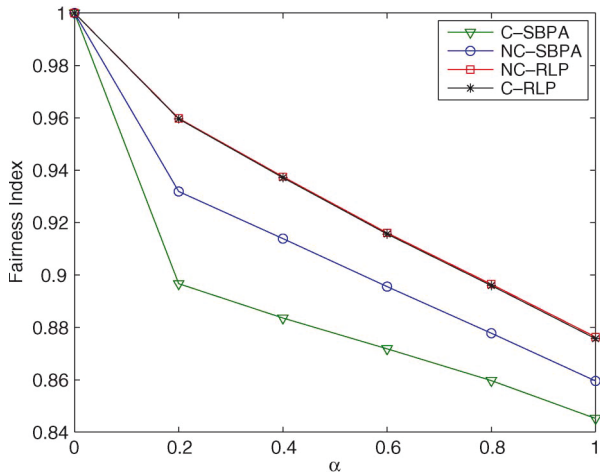


Fig. 5. FI with C-SBPA, NC-SBPA, NC-RLP, and C-RLP algorithms, with  $\alpha \in [0, 1]$ , for the case of four users.

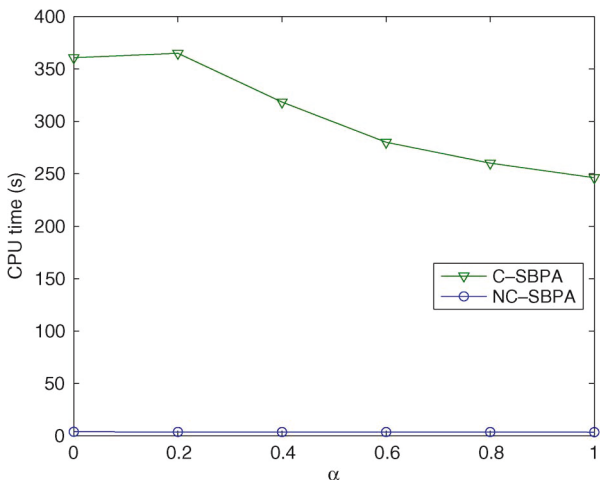


Fig. 6. Comparison in CPU time for the C-SBPA and NC-SBPA algorithms, with  $\alpha \in [0, 1]$ , for the case of four users.

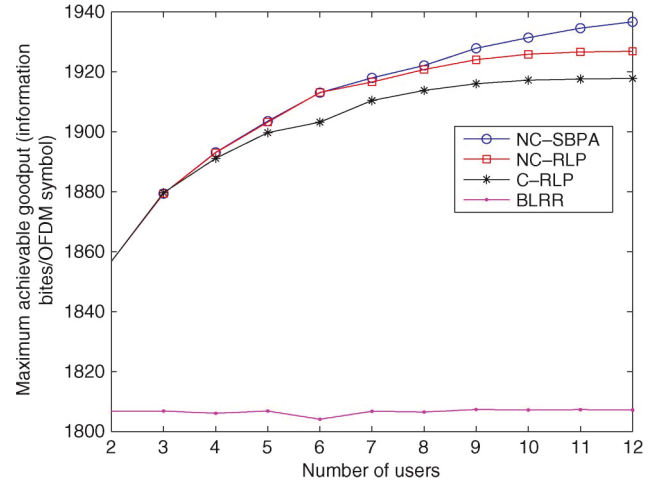


Fig. 7. Maximum averaged sum goodput received by users with NC-SBPA, NC-RLP, and C-RLP algorithms, with  $\alpha = 0$ , and BLRR scheduling.

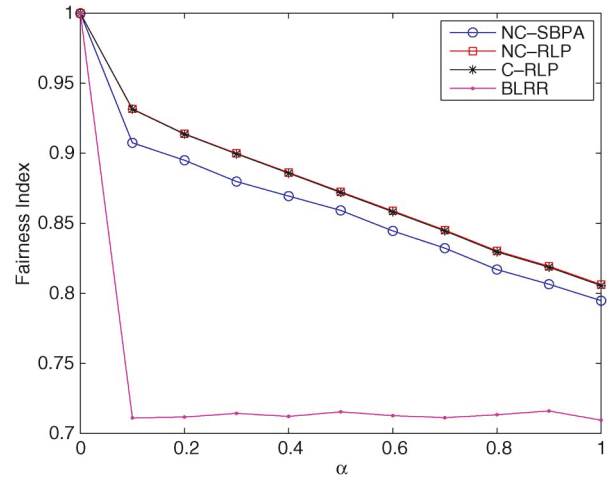


Fig. 8. FI variation with the fairness parameter  $\alpha$  for the NC-SBPA, NC-RLP, C-RLP, and BLRR algorithms.

(NC-SBPA and NC-RLP) and the C-RLP scheme in the remainder of this section.

We now compare the algorithms that can be considered computationally tractable, namely, NC-SBPA, NC-RLP, and C-RLP through Figs. 7–9. Fig. 7 shows the performance in terms of maximum average sum goodput transmitted by the BS per OFDM symbol, which is calculated as the sum of the maximum goodput that can be received by users per OFDM symbol. We display the results for  $\alpha = 0$  since we have observed a very similar behavior for the remaining values of  $\alpha$ . For comparison purposes, results for a RR NCSs scheduling are also depicted. In this RR scheduling, users in turn are randomly assigned a subchannel until all subchannels are allocated, and for fair comparison, the same bit loading algorithm of Section IV-A is used. We denote it as bit loading RR (BLRR) scheduling. The performance of the NC-SBPA, NC-RLP, and C-RLP algorithms are very close, whereas the loss is as much as 8% for BLRR (12 users). It can be also observed that, due to the exploitation of multiuser diversity, the average sum goodput rises with the number of users for the three proposed algorithms, whereas it keeps approximately constant for BLRR.

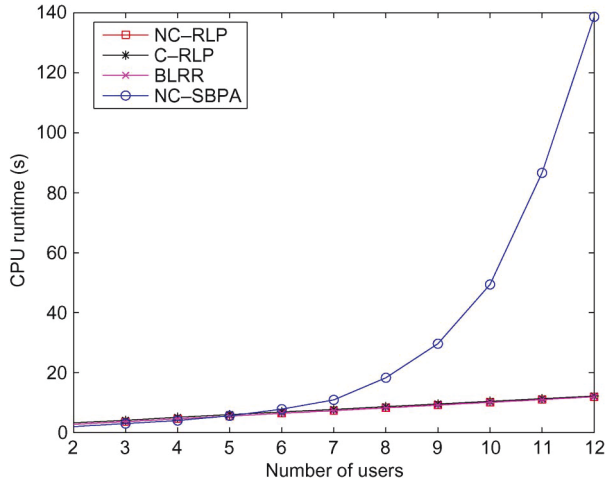


Fig. 9. CPU time for the NC-SBPA, NC-RLP, and C-RLP algorithms and BLRR scheduling, with  $\alpha = 1$ .

Fairness control over the goodput achieved per user is assessed in Fig. 8 through the well-known fairness index (FI), which is defined as [53]

$$\text{FI} = \frac{\left(\sum_{k=1}^K \chi_k\right)^2}{K \sum_{k=1}^K \chi_k^2} \quad (21)$$

where  $\chi_k$  represents the goodput for user  $k$ . The random sequential variation of the channel coefficients from one frame to the following frame causes the channels experienced by the users to behave in a similar way in the long term in average, and consequently, similar averaged goodput would be expected without fairness control. For the proposed algorithms, the fairness in the achieved goodput per user is decreasing with  $\alpha$ , where the fairest result corresponds to  $\alpha = 0$  and the lowest is achieved for  $\alpha = 1$ , whereas no fairness degree is accomplished with RR scheduling, as Fig. 8 reveals for the case of seven users. We see also that the difference among the proposed algorithms is small in terms of fairness; in fact, C-RLP and NC-RLP manifest almost exactly the same values. Similar comparisons are obtained for different numbers of users.

Finally, the algorithms are compared in terms of computational complexity, as it is shown in Fig. 9, where the complexity associated to the algorithms is represented by the average runtime (CPU time) in a PentiumM 1.6-GHz computer. We have computed the processing time considering, as previously, that  $K$  processors (at either the BS or each user) compute sequentially the corresponding  $N$  subproblems. The RLP-based algorithms exhibit practically the same computational complexity, and the difference is negligible with respect to BLRR. Therefore, C-RLP and NC-RLP achieve higher goodput and fairness control at the expense of a very small increase in CPU time with respect to BLRR. The NC-SBPA algorithm shows by far the worst computational time, which is the price to pay for achieving the highest sum goodput. Although CPU time is slightly larger, the C-RLP algorithm has the advantage that the transmitter–receiver architecture is considerably simpler as less encoder/decoder pairs are needed [see Fig. 1(a) and (b)].

In light of these results, we can conclude the following. The SBPA algorithm provides the best results for both the concatenated and nonconcatenated strategies. However, the complexity is significantly high, particularly for the C-SBPA algorithm. The RLP algorithms provide very good results when compared with the NC-SBPA algorithm, with much lower complexity when the number of users is large ( $K > 7$ ). We see that the NC-RLP algorithm accomplishes a slightly better performance than C-RLP in terms of goodput but at the expense of a more complicated transmitter architecture.

## VIII. CONCLUSION

In this paper, we have studied the optimization of the goodput transmitted by a BS for the downlink in OFDMA systems considering fairness control. This goodput optimization is formulated as a cross-layer RA problem based on PHY- and link-layer parameters. Two different transmission strategies are considered for the encoding frame: CSs and assuming independently encoded subchannels. We present a two-step RA method based on decomposition theory that reduces the complexity of the allocation problem and can be applied to both transmission schemes (C-SBPA and NC-SBPA).

Together with the two-step method, our main contribution is the development of the RLP algorithm. This algorithm provides an implementable solution for the NC case very close to the optimal. The RLP algorithm can be also applied to concatenated transmission (C-RLP), providing a two-fold benefit. First, the complexity of the algorithm is similar to the complexity of the corresponding relaxed LP problem, thus much lower than the complexity of the optimal solution. Second, the subchannel allocation solution is an integer, which makes its implementation in real systems possible.

Simulations reveal the difference between the concatenated and nonconcatenated approaches. We show that the proposed algorithms largely outperform an improved RR scheduling. Although not implementable in practice, the SBPA-based algorithms provide close benchmarking for the RLP-based algorithms since they constitute an upper bound of the performance that can be achieved. Moreover, while the concatenated approach C-RLP accomplishes a little inferior goodput with the same computational time than NC-RLP, the transmitter architecture is simpler as it requires only one encoder. Therefore, the C-RLP algorithm appears to be the only feasible choice if we consider the complexity of having as many encoders as subchannels required for NC-RLP. However, as in Section II, these two extremes are idealizations of the actual situation where most likely the number of subchannels per encoder would be a reduced number, therefore allowing an NC-RLP strategy to be implemented. Its performance will lie between the two extremes analyzed in this paper.

## REFERENCES

- [1] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.
- [2] J. Jang and K. B. Lee, "Transmit power adaptation for multiuser OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 2, pp. 171–178, Feb. 2003.

- [3] Z. Shen, J. G. Andrews, and B. L. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2726–2737, Nov. 2005.
- [4] K. E. Baamrani, A. A. Ouahman, V. P. G. Jiménez, A. G. Armada, and S. Allaki, "Subcarrier and power allocation for the downlink of multiuser OFDM transmission," *Wireless Pers. Commun.*, vol. 39, no. 4, pp. 457–465, Dec. 2006.
- [5] M. Pischella and J.-C. Belfiore, "Weighted sum throughput maximization in multicell OFDMA networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 896–905, Feb. 2010.
- [6] M. K. Awad, V. Mahinthan, M. Mehrjoo, X. Shen, and J. W. Mark, "A dual-decomposition-based resource allocation for OFDMA networks with imperfect CSI," *IEEE Trans. Veh. Technol.*, vol. 59, no. 5, pp. 2394–2403, Jun. 2010.
- [7] J. F. Schmidt, J. E. Cousseau, R. Wichman, and S. Werner, "Bit loading using imperfect CSIT for prediction-based resource allocation in mobile OFDMA," *IEEE Trans. Veh. Technol.*, vol. 60, no. 8, pp. 4082–4088, Oct. 2011.
- [8] N. Mastronarde, D. S. Turaga, and M. van der Schaar, "Collaborative resource exchanges for peer-to-peer video streaming over wireless mesh networks," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 1, pp. 108–118, Jan. 2007.
- [9] K.-H. Liu, H. Rutagemwa, X. Shen, and J. W. Mark, "Efficiency and goodput analysis of Dly-ACK in IEEE 802.15.3," *IEEE Trans. Veh. Technol.*, vol. 56, no. 6, pp. 3888–3898, Nov. 2007.
- [10] N. Baldo, F. Maguolo, S. Merlin, A. Zanella, M. Zorzi, D. Melpignano, and D. Siorpaes, "GORA: Goodput optimal rate adaptation for 802.11 using medium status estimation," in *Proc. IEEE Int. Conf. Commun.*, May 2008, pp. 4916–4921.
- [11] L.-C. Wang, W.-C. Liu, A. Chen, and K.-N. Yen, "Joint rate and power adaptation for wireless local area networks in generalized Nakagami fading channels," *IEEE Trans. Veh. Technol.*, vol. 58, no. 3, pp. 1375–1386, Mar. 2009.
- [12] K.-T. Feng, P.-T. Lin, and W.-J. Liu, "Frame-aggregated link adaptation protocol for next generation wireless local area networks," *EURASIP J. Wireless Commun. Netw.*, vol. 2010, no. 1, pp. 164651-1–164651-13, Jun. 2010.
- [13] *IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems Amendment 2: Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands and Corrigendum 1*, IEEE Std. 802.16e-2005 and IEEE Std 802.16-2004/Cor 1-2005 (Amendment and Corrigendum to IEEE Std 802.16-2004), 2006, pages 1–822.
- [14] *IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems*, IEEE Std. 802.16-2009, 2009, pages 1–2080.
- [15] I. Stupia, L. Vandendorpe, J. Louveaux, F. Giannetti, V. Lottici, and N. A. D'Andrea, "Power allocation for goodput optimization in BICM-OFDM systems," in *Proc. IEEE Int. Conf. Commun.*, 2008, pp. 3604–3608.
- [16] S. Stefanatos and N. Dimitriou, "Downlink OFDMA resource allocation under partial channel state information," in *Proc. IEEE Int. Conf. Commun.*, 2009, pp. 1–5.
- [17] S. Stefanatos, C. Papanthasiou, and N. Dimitriou, "Downlink mobile OFDMA resource allocation with minimum user rate requests," in *Proc. IEEE GLOBECOM Conf.*, 2010, pp. 1–5.
- [18] R. Aggarwal, M. Assaad, C. E. Koksal, and P. Schniter, "OFDMA downlink resource allocation via ARQ feedback," in *Proc. 43rd Asilomar Conf. Signals, Syst. Comput.*, 2010, pp. 1493–1497.
- [19] R. Aggarwal, M. Assaad, C. E. Koksal, and P. Schniter, "Joint scheduling and resource allocation in the OFDMA downlink: Utility maximization under imperfect channel-state information," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5589–5604, Nov. 2011.
- [20] R. Aggarwal, C. E. Koksal, and P. Schniter, "Joint scheduling and resource allocation in OFDMA downlink systems via ACK/NAK feedback," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 3217–3227, Jun. 2012.
- [21] R. Wang and V. K. N. Lau, "Robust optimal cross-layer designs for TDDOFDMA systems with imperfect CSIT and unknown interference: State space approach based on 1-bit ACK/NAK feedbacks," *IEEE Trans. Commun.*, vol. 56, no. 5, pp. 754–761, May 2008.
- [22] V. Lau, W. K. Ng, and D. S. W. Hui, "Asymptotic tradeoff between cross-layer goodput gain and outage diversity in OFDMA systems with slow fading and delayed CSIT," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2732–2739, Jul. 2008.
- [23] D. Hui and V. Lau, "Design and analysis of delay-sensitive cross-layer OFDMA systems with outdated CSIT," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3484–3491, Jul. 2009.
- [24] D. S. W. Hui and V. K. N. Lau, "Design and analysis of delay-sensitive decentralized cross-layer OFDMA systems with efficient feedback algorithm," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5844–5851, Dec. 2009.
- [25] D. W. K. Ng and R. Schober, "Cross-layer scheduling for OFDMA amplify-and-forward relay networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 3, pp. 1443–1458, Mar. 2010.
- [26] S. Pfletschinger, A. Piatyszek, and S. Stiglmayr, "Frequency-selective link adaptation using duo-binary turbo codes in OFDM systems," in *Proc. IST 16th Mobile Wireless Commun. Summit*, Jul. 2007, pp. 1–5.
- [27] M. Sternad, T. Svensson, T. Ottosson, A. Ahlén, A. Svensson, and A. Brunstrom, "Towards systems beyond 3G based on adaptive OFDMA transmission," *Proc. IEEE*, vol. 95, no. 12, pp. 2432–2455, Dec. 2007.
- [28] M. M. Rashid and V. K. Bhargava, "A model-based downlink resource allocation framework for IEEE 802.16e mobile WiMAX systems," *IEEE Trans. Veh. Technol.*, vol. 59, no. 8, pp. 4026–4042, Oct. 2010.
- [29] J. Escudero-Garzas, B. Devillers, L. Vandendorpe, and A. Garcia-Armada, "Subchannel, bit and power allocation in multiuser OFDM systems for goodput optimization with fairness," presented at the 12th International Symposium WPMC, Sendai, Japan, 2009, Paper S04-3.
- [30] D. Qiao, S. Choi, and K. G. Shin, "Goodput analysis and link adaptation for IEEE 802.11a wireless LANs," *IEEE Trans. Mobile Comput.*, vol. 1, no. 4, pp. 278–292, Oct.–Dec. 2002.
- [31] G. Caire, R. R. Muller, and R. Knopp, "Hard fairness versus proportional fairness in wireless communications: The single-cell case," *IEEE Trans. Inf. Theory*, vol. 53, no. 4, pp. 1366–1385, Apr. 2007.
- [32] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Netw.*, vol. 8, no. 5, pp. 556–567, Oct. 2000.
- [33] F. Brah, J. Louveaux, and L. Vandendorpe, "CDIT-based constrained resource allocation for mobile WiMAX systems," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, no. 1, pp. 425367-1–425367-8, Jan. 2009.
- [34] T. Eriksson and T. Ottosson, "Compression of feedback for adaptive transmission and scheduling," *Proc. IEEE*, vol. 95, no. 12, pp. 2314–2321, Dec. 2007.
- [35] T. M. Cover and Thomas, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 1991.
- [36] H. Zhu and J. Wang, "Chunk-based resource allocation in OFDMA systems-Part I: Chunk allocation," *IEEE Trans. Commun.*, vol. 57, no. 9, pp. 2734–2744, Sep. 2009.
- [37] M. Kaneko, P. Popovski, and J. Dahl, "Proportional fairness in multicarrier system with multi-slot frames: Upper bound and user multiplexing algorithms," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 22–26, Jan. 2008.
- [38] S. T. Chung and A. J. Goldsmith, "Degrees of freedom in adaptive modulation: A unified view," *IEEE Trans. Commun.*, vol. 49, no. 9, pp. 1561–1571, Sep. 2001.
- [39] Q. Liu, S. Zhou, and G. B. Giannakis, "Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1746–1755, Sep. 2004.
- [40] B. Devillers, J. Louveaux, and L. Vandendorpe, "Bit and power allocation for goodput optimization in coded parallel subchannels with ARQ," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3652–3661, Aug. 2008.
- [41] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [42] G. Song and Y. Li, "Cross-layer optimization for ofdm wireless networks-Part II: algorithm development," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 625–634, Mar. 2005.
- [43] P. Viswanath, D. N. C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1277–1294, Jun. 2002.
- [44] C. A. Floudas, *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications: Fundamentals and Applications*. New York, NY, USA: Oxford Univ. Press, 1995.
- [45] S. Boyd, L. Xiao, A. Mutapic, and J. Mattingley, "Notes on decomposition methods," Stanford Univ., Stanford, CA, USA, Notes for EE364B, 2007.
- [46] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.
- [47] W. Rhee and J. M. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," in *Proc. IEEE 51st VTC-Spring*, Tokyo, Japan, 2000, vol. 2, pp. 1085–1089.
- [48] P. S. Chow, "Bandwidth optimize digital transmission techniques for spectrally shaped channels," Ph.D. dissertation, Stanford Univ., Stanford, CA, USA, 1993.

- [49] W. Yu and J. M. Cioffi, "Constant-power waterfilling: Performance bound and low-complexity implementation," *IEEE Trans. Commun.*, vol. 54, no. 1, pp. 23–28, Jan. 2006.
- [50] D. Bertsimas and R. Weismantel, *Optimization Over Integers*. Belmont, MA, USA: Dynamic Ideas, 2005.
- [51] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA, USA: Athena Scientific, 1999.
- [52] S. Mehrotra, "On the implementation of a primal-dual interior point method," *SIAM J. Optim.*, vol. 2, no. 4, pp. 575–601, 1992.
- [53] R. Jain, D. M. Chiu, and W. R. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer system," Eastern Res. Lab., Digital Equipment Corp., Maynard, MA, USA, 1984.



**José Joaquín Escudero-Garzás** (S'06–M'10) received the Ph.D. degree in electrical engineering from Carlos III University of Madrid, Madrid, Spain, in 2010.

From 1997 to 2002, he was a Provisioning Engineer and the Head of the Telecommunications Network Maintenance Department with Unión Fenosa, S.A., and Jazztel. In 2002, he was an Assistant Lecturer with Carlos III University of Madrid, where he taught more than 12 courses related to telecommunication systems and wired/wireless communications

and participated in several competitive national and European projects. From March to August 2008, he was a Guest Researcher with the Communications and Remote Sensing Laboratory, Université Catholique de Louvain, Louvain-La-Neuve, Belgium. From October 2010 to September 2012, he was a Post-doctoral Fellow with the Department of Telecommunications and Systems Engineering, Autonomous University of Barcelona, Barcelona, Spain. He is currently with the Department of Systems and Information Engineering, University of Virginia, Charlottesville, VA, USA. His main research interests include optimization and game-theory-based techniques applied to smart grid and wireless communication systems.



**Bertrand Devillers** received the B.S. degree in electrical engineering and the Ph.D. degree from the Université Catholique de Louvain (UCL), Louvain-La-Neuve, Belgium, in June 2004 and March 2009, respectively.

From October 2004 to September 2008, he was a Research Fellow of the Belgian National Fund for Scientific Research (FRS-FNRS) with the Digital Communication Group, Communications and Remote Sensing Laboratory, UCL. From October 2008 to March 2009, he was a Research Assistant with

UCL. From April 2009 to February 2013, he was a Research Associate with the Centre Tecnològic de Telecomunicacions de Catalunya, Barcelona, Spain. He is currently a patent examiner with the European Patent Office, The Hague, The Netherlands. His research interests include communication theory and signal processing for digital terrestrial and satellite communications, particularly multiple-input–multiple output systems, multicarrier and cyclically prefixed single-carrier modulations, and energy efficiency.



**Ana García-Armada** (S'96–A'98–M'00–SM'08) received the Ph.D. degree in electrical engineering from the Technical University of Madrid, Madrid, Spain, in February 1998.

She is currently a Professor with Carlos III University of Madrid, where she has occupied a variety of management positions, particularly in the Communications Research Group and the Laboratory of Communication Systems for Security and Space in the Scientific Park. She has participated in several national and international research projects

related to wireless communications. She is the author of more than 35 papers in international journals and more than 50 contributions to international conference proceedings, and the co-author of eight book chapters on wireless communications and signal processing. She has contributed to international organizations such as the International Telecommunication Union and the European Telecommunications Standards Institute. Her main interests include orthogonal frequency-division multiplexing and multiple-input–multiple-output techniques, as well as signal processing applied to wireless communications.