On Two Cross-layer Assisted Resource Allocation Schemes in Multiuser Multicarrier Systems

Qing Bai, Josef A. Nossek

Institute for Circuit Theory and Signal Processing, Technische Universität München
Theresienstr. 90, 80333 Munich, Germany
bai.qing@nws.ei.tum.de

***Abstract*—*Resource allocation* (RA) is a crucial task in the operation of wireless communication systems, which in many cases aims at meeting the users' *quality of service* (QoS) requirements with the minimum amount of resources. In this work, we consider the QoS-provisioning RA problem at the downlink of a multicarrier system, where the total transmit power consumption is to be minimized. Two RA schemes, one employing Lagrange dual methods and a greedy primal recovery scheme, and the other heuristic with a three-step approach, are proposed under the same cross-layer framework, where power allocation, adaptive modulation and coding as well as retransmission protocols are jointly modeled. Though both suboptimal, the two schemes are advantageous for their low complexity and small amount of online computations required. Their performances are illustrated and compared based on simulation results. The work also provides a quantitative comparison on the two commonly used retransmission protocols, namely *automatic repeat request* (ARQ) and *hybrid automatic repeat request* (HARQ).**

I. Introduction

As one of the key approaches to meet the increasing demand for better QoS in current and future wireless networks, *cross-layer optimization* has drawn much research attention from various aspects of communications. One of its main applications is to assist the radio resource allocations by jointly adapting variables from physical and link layers to optimize certain performance metrics, *e.g.*, the sum throughput or the sum transmit power. We term this kind of applications as *cross-layer assisted resource allocation* (CLARA). On the other hand, RA problems in multicarrier systems have long been studied, *e.g.*, [1][2]. With a more sophisticated cross-layer framework and target QoS parameters having a higher-level presentation, the relation and coupling between the optimization variables become more complicated. The mappings from resources to QoS parameters often lack differentiability, continuity, and even convexity, making the optimizations considerably challenging. The two algorithms presented in this work use look-up tables and stepwise variable fixing respectively to overcome these difficulties.

Unlike many other cross-layer models, retransmission protocols are included in our framework. Firstly, the time it takes for retransmissions is part of the latency a packet experiences until received correctly. Secondly, different ways of retransmitting a packet have an influence on the efficiency of radio resource utilization. Our simulation results prove the necessity to study various retransmission schemes and to set up appropriate models to evaluate their performances.

II. System Model

We consider the downlink scenario of an isolated single-cell multicarrier system with K users, each having one data stream to be served. The RA is done on a per *slot* basis, where a *slot* is a short time period of length T during which the wireless channel is assumed to stay constant. As information bits loaded onto consecutive slots are independently modulated and coded, a slot is formally referred to as a *Transmission Time Interval* (TTI), and the bit-loading procedure inherently includes *packetization* of the information bits. For every TTI, each data stream has a number of information bits to be transmitted, depending on its *throughput* requirement. The other relevant QoS parameter characterizing the data streams, the *latency*, is defined as:

Definition: The latency τ_k of a packet from user k is the delay it experiences until received correctly with an outage probability of no more than the predefined value $\pi^{(\text{out})}$. Let $f_k[m]$ be the probability that it takes exactly m TTI's to transmit a packet error-free, then $\tau_k = (M_k - 1)(\text{RTD} + T) + T$ where RTD represents *round trip delay*, and

$$M_k = \min_{M} M \quad \text{s.t.} \quad \sum_{m=1}^{M} f_k[m] \geq 1 - \pi^{(\text{out})}.$$

We derive in the following the mathematical descriptions of the regarded system components stemmed from [3], which lay the basis for cross-layer optimization.

A. Channel Model

The downlink broadcast channel is modeled as frequency-selective fading over the total system bandwidth and frequency-flat fading over each *subchannel*, which is consist of N_c adjacent subcarriers. FDMA is employed meaning the assignment of every subchannel is exclusive to one user, and *intercarrier interference* (ICI) is not taken into account. Moreover, we restrict ourselves here to the single-antenna case both at the base station (BS) and at the mobile stations (MS).

Let $H_{k,n}$ and $\sigma^2_{k,n}$ be the channel coefficient and Gaussian noise variance of user k on the nth subchannel, and p_n be the amount of power allocated on subchannel n. When assigned to user k, the *signal-to-noise-ratio* (SNR) on subchannel n is computed as $\gamma_{k,n} = \frac{|H_{k,n}|^2}{\sigma^2_{k,n}} p_n$. For the remaining part of this section we drop the subscripts k and n for simplicity.

We choose the TTI to be of length $T = 2$ ms and assume that one TTI contains $N_\mathrm{s} = 16$ symbols for data transmission. The *minimum allocation unit* (MAU) is an allocation region of one subchannel in the frequency dimension by one TTI in the time dimension, which contains $N_\mathrm{c} \times N_\mathrm{s}$ symbols.

B. FEC coding and modulation

We assume that modulation and coding across the subchannels are done independently, and with reference to the WiMAX standard 8 modulation and coding schemes (MCS) are chosen to form the candidate set $\mathcal{M}$, which are listed in Table I.

Table I
MODULATION AND CODING SCHEMES (MCS)

Index	Modulation Type	Alphabet Size A	Code Rate R	$R\log_2 A$
1	BPSK	2	$1/2$	0.5
2	QPSK	4	$1/2$	1
3	QPSK	4	$3/4$	1.5
4	16-QAM	16	$1/2$	2
5	16-QAM	16	$3/4$	3
6	64-QAM	64	$2/3$	4
7	64-QAM	64	$3/4$	4.5
8	64-QAM	64	$5/6$	5

With the absence of intersymbol interference in the system, each subchannel can be modeled as a *discrete memoryless channel* (DMC) over which the *noisy channel coding theorem* [6] can be applied. Let the modulation alphabet and coding rate on the subchannel under consideration be $\mathcal{A} = \{a_1, \ldots, a_A\}$ and R respectively. The *cutoff rate* of the subchannel with SNR γ can be expressed as

$$R_0(\gamma, A) = \log_2 A - \log_2 \left[1 + \frac{2}{A}\sum_{m=1}^{A-1}\sum_{l=m+1}^{A} e^{-\frac{1}{4}|a_l - a_m|^2\gamma}\right].$$

The noisy channel coding theorem states that there always exists a block code with block length l and binary code rate $R\log_2 A \le R_0(\gamma, A)$ in bits per subchannel use, such that with maximum likelihood decoding the error probability $\tilde{\pi}$ of a code word satisfies $\tilde{\pi} \le 2^{-l(R_0(\gamma,A) - R\log_2 A)}$.

In order to apply this upper bound to the extensively used turbo decoded convolutional code, quantitative investigations have been done in [3] and an expression for the *equivalent block length* is derived based on link level simulations as $n_\mathrm{eq} = \beta \ln L$, where parameter β is used to adapt this model to the specifics of the employed turbo code, and L is the coded packet length. Consequently, the transmission of L bits is equivalent to the sequential transmission of L/n_eq blocks of length n_eq and has an error probability of

$$\pi = 1 - (1 - \tilde{\pi})^{\frac{L}{n_\mathrm{eq}}} \le 1 - \left(1 - 2^{-n_\mathrm{eq}(R_0(\gamma,A) - R\log_2 A)}\right)^{\frac{L}{n_\mathrm{eq}}}.$$

C. Protocol

At the link layer retransmission protocols are studied. The data sequence transmitted in one MAU, *i.e.*, a *packet*, is used as the retransmission unit.

ARQ: The corrupted packets at the receiver are discarded, hence we assume that the *packet error probability* (PEP) of a retransmitted packet is the same as that of its original transmission, *i.e.*, $f[m] = \pi^{m-1}(1-\pi), m \in \mathbb{Z}^+$. Therefore the maximum allowable PEP is $\pi = \sqrt[M]{\pi^{(\mathrm{out})}}$ when the number of transmissions is M .

HARQ: The corrupted packets at the receiver are combined and jointly decoded using rate-compatible punctured convolutional codes. For the particular *incremental redundancy* (IR) scheme we employ where the retransmissions contain pure parity bits of the same length as the first transmission, the code rate for the mth transmission can be expressed as $R[m] = \frac{B}{m \cdot L} = \frac{1}{m}R$. Let $\tilde{m}$ denote the maximum number of transmissions determined by the mother code. The equivalent block length n_eq is then given by $n_\mathrm{eq} = \beta\ln(\tilde{m}L)$. The PEP expression for the mth transmission follows as

$$\pi[m] = 1 - \left(1 - 2^{-\beta\ln(\tilde{m}L)(R_0(\gamma) - \frac{1}{m}R\log_2 A)}\right)^{\frac{mL}{\beta\ln(\tilde{m}L)}},$$

and is approximated by

$$\pi[M] = \pi^{(\mathrm{out})}, \quad \pi[m] = 1, m = 1, \ldots, M-1,$$

when $R_0(\gamma)$ satisfies $\frac{1}{M}R\log_2 A < R_0(\gamma) \le \frac{1}{M-1}R\log_2 A$.

The system parameters are summarized in Table II.

Table II
SYSTEM PARAMETERS

Total bandwidth		10 MHz
Center frequency	f_c	2.5 GHz
FFT size		1024
Number of data subcarriers		720
Number of subchannels	N	30
Number of subcarriers per subchannel	N_c	$720/30 = 24$
Transmission Time Interval (TTI)	T	2 ms
Number of data symbols per TTI	N_s	16
Round Trip Delay (RTD)	RTD	10 ms
Maximum number of transmissions allowed	$\tilde{m}$	5
Turbo code dependent parameter	β	32
Outage probability	$\pi^{(\mathrm{out})}$	0.01

III. PROBLEM FORMULATION

Let the number of information bits intended for user k in the current TTI be b_k, and the maximum latency time for the transmission be $\tau_k^{(\mathrm{rq})}$. We formulate the transmit power minimization problem as

$$\begin{aligned} \min_{\boldsymbol{B},\boldsymbol{A},\boldsymbol{R},\boldsymbol{M}} & \sum_{n=1}^{N} p_n \\ \text{s.t.} \quad \boldsymbol{B} & \in \mathcal{B}, \\ (A_n, R_n) & \in \mathcal{M}, \quad n = 1, \ldots, N, \\ \textstyle\sum_{n=1}^{N} B_{k,n} & \ge b_k, \quad k = 1, \ldots, K, \\ \tau_k & \le \tau_k^{(\mathrm{rq})}, \quad k = 1, \ldots, K, \end{aligned} \tag{1}$$

where $\boldsymbol{B} \in \mathbb{Z}_{+,0}^{K\times N}$ represents the bit-loading matrix with its entry $B_{k,n}$ as the number of information bits for the kth user loaded onto the nth subchannel, and p_n, A_n, R_n, M_n are the transmit power, MCS and number of transmissions taken on the nth subchannel, respectively. The first constraint in (1) comes from FDMA in which $\mathcal{B} \subset \mathbb{Z}_{+,0}^{K\times N}$ represents the set of matrices that have only one nonzero entry in each of their columns, and the third and fourth constraints are the fulfillment of the QoS requirements of each user.

Dropping the subchannel indices, we let $s = \left\lceil \frac{B}{R \log_2 A} \right\rceil$ be the number of symbols occupied in one MAU. Transmit power p dependent on (B, A, R, M) can be written as

$$p = \left\lceil \frac{s}{N_s} \right\rceil \cdot \gamma(B, A, R, M) \cdot \frac{\sigma^2}{|H|^2}, \qquad (2)$$

where $|H|^2$ is the instantaneous channel gain and σ^2 is the noise power on one subcarrier. $\gamma(B, A, R, M)$ is the SNR required to convey B bits within M transmissions when MCS (A, R) is employed, which can be obtained from a binary search on the cutoff rate curve. For both ARQ and HARQ protocols, $\gamma(B, A, R, M^*) < \gamma(B, A, R, M)$ if $M^* > M$, which means the more transmissions, the less transmit power required. Therefore, to solve (1) the maximum number of transmissions M is fixed in the first place to $\left\lfloor \frac{\tau^{(\mathrm{rq})} - T}{\mathrm{RTD} + T} + 1 \right\rfloor$.

IV. Two Resource Allocation Schemes

It is obvious that problem (1) is nonconvex and combinatorial with integer-valued variables. As a result, there are in general no standard optimization algorithms that can be directly applied to it. We propose two different ways to solve (1), both suboptimal but with tractable complexities.

A. Employing Lagrange Dual Methods

The employments of Lagrange dual decomposition and Lagrange dual methods to solving RA problems in multicarrier systems were studied, *e.g.*, in [4][5]. The key element in applying those methods to (1) within our cross-layer model is to establish a mapping from the optimization variables to the optimization objective which makes the evaluation of the dual function possible and fast.

1) The φ function: We define $\varphi(B, \tau^{(\mathrm{rq})})$ as the minimum power needed for the successful transmission of B bits within latency time $\tau^{(\mathrm{rq})}$, *i.e.*,

$$\begin{aligned} \varphi(B, \tau^{(\mathrm{rq})}) &= \min_{(A,R) \in \mathcal{M}} \left\lceil \frac{s}{N_s} \right\rceil \cdot \gamma(B, A, R, M) \cdot \frac{\sigma^2}{|H|^2} \\ &\stackrel{!}{=} \frac{\sigma^2}{|H|^2} \min_{(A,R) \in \mathcal{M}} \phi(B, A, R, M), \end{aligned}$$

where the function ϕ is independent of channel realizations.

Limited by the highest MCS, the number of information bits that can be loaded in one MAU is upper bounded by $B^{(\mathrm{u})} = 5 N_\mathrm{s} N_\mathrm{c}$. That means, for each $B \in [1, B^{(\mathrm{u})}]$ and $M \in [1, \tilde{m}]$, $\min_{(A,R) \in \mathcal{M}} \phi(B, A, R, M)$ can be computed by enumerating all the 8 MCS which at last results in a look-up table with $B^{(\mathrm{u})} \times \tilde{m}$ entries. This table is established offline and stored. At run time, multiplications by the noise-to-channel-gains-ratio to the retrieved table entries are sufficient to obtain $\varphi(B, \tau^{(\mathrm{rq})})$. Consequently, problem (1) can be equivalently written in a simpler form as

$$\begin{aligned} \min_{\boldsymbol{B} \in \mathcal{B}} \quad & \sum_{k=1}^{K} \sum_{n=1}^{N} \varphi_{k,n}(B_{k,n}, \tau_k^{(\mathrm{rq})}) \\ \text{s.t.} \quad & \textstyle\sum_{n=1}^{N} B_{k,n} \geq b_k, \quad k = 1, \ldots, K. \end{aligned} \qquad (3)$$

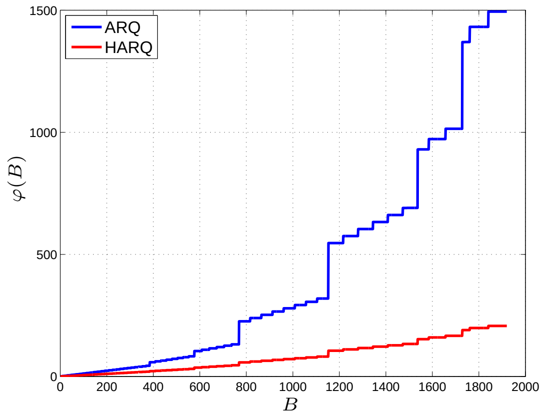

Figure 1. An exemplary φ function for ARQ and HARQ protocols

An exemplary φ function with varying B is shown in Fig. 1, where $\tau^{(\mathrm{rq})}$ is fixed to 20 ms. The visualization shows the monotonicity of required transmit power as a function of B, and as B increases, the power increments for the same increment in B become larger. Both properties are in accordance with basic knowledge from information theory. However, the φ function is not convex due to its discrete inputs and the changes of the optimum MCS at some values of B. As a result, optimization (3) is not convex in both objective and constraints and has a nonzero duality gap when dual methods are applied. Also note that the power consumption in the HARQ case is much less than that of the ARQ case.

2) Dual Methods: We follow a similar procedure as proposed in [4] to find the dual optimum solution of (3). Introducing Lagrange multipliers $\boldsymbol{\lambda} \in \mathbb{R}^{K \times 1}$ to the K bit-loading constraints gives the Lagrangian

$$L(\boldsymbol{B}, \boldsymbol{\lambda}) = \sum_{k=1}^{K} \sum_{n=1}^{N} \varphi_{k,n}(B_{k,n}, \tau_k^{(\mathrm{rq})}) + \sum_{k=1}^{K} \lambda_k \Big(\sum_{n=1}^{N} B_{k,n} - b_k \Big),$$

and the dual function $g(\boldsymbol{\lambda}) = \inf_{\boldsymbol{B} \in \mathcal{B}} L(\boldsymbol{B}, \boldsymbol{\lambda})$ can be decomposed into N independent optimization problems

$$g_n(\boldsymbol{\lambda}) = \inf_{\boldsymbol{B} \in \mathcal{B}} \sum_{k=1}^{K} \left(\varphi_{k,n}(B_{k,n}, \tau_k^{(\mathrm{rq})}) + \lambda_k B_{k,n} \right)$$

plus a $\boldsymbol{B}$-independent term. The dual problem to (3) reads

$$\max g(\boldsymbol{\lambda}) \quad \text{s.t.} \quad \boldsymbol{\lambda} \succeq \boldsymbol{0}. \qquad (4)$$

The ellipsoid method is employed to efficiently update the dual variable $\boldsymbol{\lambda}$. Denote the optimal value and solution to (4) as d^* and $\boldsymbol{\lambda}^*$ respectively, and the bit-loading matrix obtained with $\boldsymbol{\lambda}^*$ as $\tilde{\boldsymbol{B}}$. By weak duality, d^* gives a lower bound on the primal optimal value. Yet $\tilde{\boldsymbol{B}}$ is not optimum and often primal-infeasible which makes primal recovery necessary.

3) Primal Recovery Scheme: Although the bit-loading matrix $\tilde{\boldsymbol{B}}$ mostly fail to meet all the bit-loading constraints, the *subchannel assignment* (SA) it implies ($\tilde{B}_{k,n} > 0$ indicates that subchannel n is assigned to user k) still suggests an efficient way of allocating the whole set of subchannels.

However, as $B_{k,n}$ is limited by $B^{(\mathrm{u})}$ from above, the dual optimum SA can be infeasible, especially when the total number of information bits to be loaded is large. Therefore, in order to perform primal recovery based on the dual optimum SA, we have to assure its feasibility first.

The minimum number of subchannels needed by user k can be computed as $N_k^{(\mathrm{l})} = \lceil \frac{b_k}{B^{(\mathrm{u})}} \rceil$. Due to FDMA, the condition $\sum_{k=1}^{K} N_k^{(\mathrm{l})} \le N$ should be examined before we start solving (1)[1]. Let the set of subchannels assigned to user k by the dual optimum SA be $\mathcal{S}_k$, *i.e.*, $\mathcal{S}_k = \{n : \tilde{B}_{k,n} > 0\}$. If $\exists k$ with $|\mathcal{S}_k| < N_k^{(\mathrm{l})}$, then the dual optimum SA is infeasible.

Denote the set of users with $|\mathcal{S}_k| < N_k^{(\mathrm{l})}$ and $|\mathcal{S}_k| > N_k^{(\mathrm{l})}$ as $\mathcal{K}_\mathrm{u}$ and $\mathcal{K}_\mathrm{o}$, respectively, and the set of assignable subchannels as $\mathcal{N}_\mathrm{a} = \{n : n \in \mathcal{S}_k, k \in \mathcal{K}_\mathrm{o}\} \cup \{n : n \notin \mathcal{S}_k, \forall k\}$, *i.e.*, the union of unoccupied subchannels and those currently occupied by users from set $\mathcal{K}_\mathrm{o}$. Intuitively, we solve

$$(k^*, n^*) = \underset{k \in \mathcal{K}_\mathrm{u}, n \in \mathcal{N}_\mathrm{a}}{\operatorname{argmin}} \varphi_{k,n}(B^{(\mathrm{u})}, \tau_k^{(\mathrm{rq})}),$$

assign subchannel n^* to user k^*, update $\{\mathcal{S}_k\}$ and check whether the new subchannel assignment is feasible or not. The procedure terminates when $\mathcal{K}_\mathrm{u}$ becomes empty.

Fixing the obtained feasible SA, we have K decoupled minimization problems, one for each user, as

$$\min_{\{B_{k,n} : n \in \mathcal{S}_k\}} \sum_{n \in \mathcal{S}_k} \varphi_{k,n}(B_{k,n}, \tau_k^{(\mathrm{rq})}) \quad \text{s.t.} \quad \sum_{n \in \mathcal{S}_k} B_{k,n} \ge b_k,$$

which can again be solved in the dual domain. Let the dual optimal bit-loading be $\{B^*_{k,n} : n \in \mathcal{S}_k\}$. If $\sum_{n \in \mathcal{S}_k} B^*_{k,n} \ne b_k$, we can load or unload the extra bits one by one on the subchannel that leads to the minimum power increment or the maximum power decrement, until $\sum_{n \in \mathcal{S}_k} B^*_{k,n} = b_k$ is satisfied. Such a recovery scheme is simple, but greedy and performance-degrading.

B. A Heuristic Method

In [7] we proposed a heuristic three-step approach to solve the transmit power minimization problem in multicarrier systems, where each user requires a minimum data throughput and a maximum latency time. The scenario is different from what we consider here in that the number of information bits intended for each user was assumed to be infinity. Therefore bit-loading is simply determined by SA and the choices of MCS as each of the $N_\mathrm{s} N_\mathrm{c}$ symbols in one MAU are taken. To accommodate the finite number of information bits b, the *branch and bound* (BAB) method is applied to find the exact number of symbols occupied in one MAU.

1) Subchannel Assignment (SA): In this step we assume the same MCS is used on every subchannel. A power matrix $\boldsymbol{P} \in \mathbb{R}_+^{K \times N}$ can be computed, with its entry $p_{k,n}$ being the minimum power needed to achieve the required PEP of user k on subchannel n. Let $N_k^{(\mathrm{u})} = \left\lceil \frac{b_k}{0.5 N_\mathrm{s} N_\mathrm{c}} \right\rceil$ be the maximum number of subchannels user k could possibly use. The SA problem is formulated as picking from each column of $\boldsymbol{P}$ one entry such that the kth row has between $N_k^{(\mathrm{l})}$ and $N_k^{(\mathrm{u})}$ picked entries, and the sum of all picked entries is minimized.

2) Bit and Power Allocation (BPA): With the SA result as input, bit and power allocation is no longer coupled among the users and boils down for each user to

$$\min_{\boldsymbol{s}, \boldsymbol{R}, \boldsymbol{A}} \sum_{n \in \mathcal{S}_k} p_n \quad \text{s.t.} \quad \sum_{n \in \mathcal{S}_k} s_n R_n \log_2 A_n \ge b_k, \tag{5}$$

where $s_n \in [0, N_\mathrm{s} N_\mathrm{c}]$ is the number of symbols occupied on subchannel n, and the dependence of p_n on s_n is indicated by (2). Firstly we look for all feasible and efficient MCS combinations on $n \in \mathcal{S}_k$ fixing s_n to $N_\mathrm{s} N_\mathrm{c}$. Then for each MCS combination, s_n are the only optimization variables in (5). Relaxing integer valued s_n to real numbers, the problem can be solved using standard linear programming techniques. Yet directly rounding the solution does not give us the optimum solution to the original problem in general.

When the solution we get from linear programming is fractional, the BAB method is applied which branches on a fractional value and generates two subproblems. For example, $s_1 = 8.5$ adds constraint $s_1 \ge 9$ or $s_1 \le 8$ to the original problem. As soon as an integer valued solution is obtained in branching and solving the subproblems, the corresponding objective value is used as the bound to cut off inactive subproblem branches, *e.g.*, those that are worse than the current best solution. The procedure terminates when there are no more active subproblems.

3) Adjustments: The outcome of PA might indicate zero MCS on some subchannels, which means these subchannels are released from occupation and can be assigned to other users. As higher MCS are much more power consuming than lower MCS, we find the subchannels using the relatively highest MCS as well as their possessors, and compare each alternative of assigning the empty subchannels to these users.

V. Simulation Results

For simulations, $K = 10$ users uniformly located in a cell of radius 2 km are assumed. The wireless channel is modeled as a frequency-selective fading channel consisting of six independent Rayleigh multipaths with an exponentially decaying power profile. The delay spreads are uniformly distributed within 1 μs, resulting in a rms delay spread of about 0.3 μs which is consistent with the assumed channel coherence bandwidth. The path loss in dB is computed as $PL(d) = 140.6 + 35.0 \log_{10} d$ following the COST-Hata model, where d is the distance between MS and BS in km, and the receiver noise level is assumed to be -174 dBm/Hz.

Each user's QoS requirements are listed in Table III, where the unit for b_k is bit and the unit for τ_k is ms, and α is a scalar that takes values from $\{0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4\}$. Besides the algorithms discussed previously, a static RA scheme is simulated for comparison purpose. The static scheme first assigns each user with a fixed set of adjacent subchannels and then performs the greedy bit-loading, in the same way as

[1] In order to provide the resource allocation entity with appropriate traffic loads, a scheduling component on its top is necessary.

Table III
QoS REQUIREMENTS OF 10 USERS FOR SIMULATION

User	b_k	τ_k	User	b_k	τ_k	User	b_k	τ_k
1-4	$512 \cdot \alpha$	20	5-7	$800 \cdot \alpha$	40	8-10	$1600 \cdot \alpha$	80

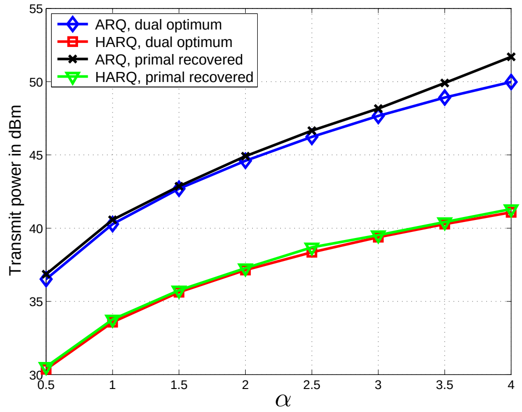

Figure 2. Dual optimum and primal recovered average power at each α

used for primal recovery. Each test scenario has been simulated under 1000 independent channel realizations.

Fig. 2 shows the difference between the dual optimum and the primal recovered average power consumption over 1000 simulations for the first RA scheme, which is satisfactorily small. For ARQ protocol, the overall difference is 0.6 dBm whereas for HARQ the value is slightly less than 0.2 dBm. Note that the actual optimal transmit power curves lie between the dual optimum and the primal recovery curves. The next two figures illustrate the statistics of power consumptions for the three RA schemes and the two retransmission protocols, where Fig. 3 shows the cumulative distributions of the transmit power with $\alpha = 2$, and Fig. 4 presents the average power consumption for different α values.

It is clear from the figures that the algorithm employing Lagrange dual methods outperforms the heuristic method. The exact average difference is 1.4 dBm for ARQ and 0.9 dBm for

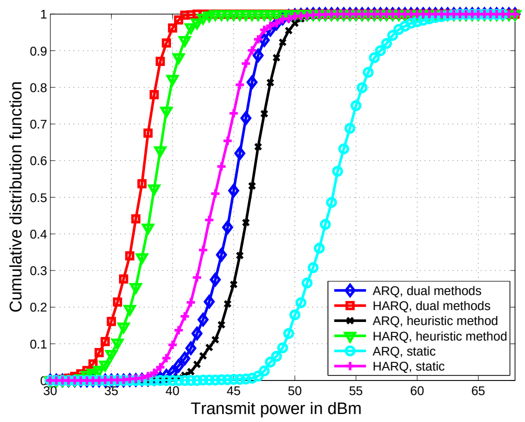

Figure 3. CDF of transmit power for 3 RA schemes at $\alpha = 2$

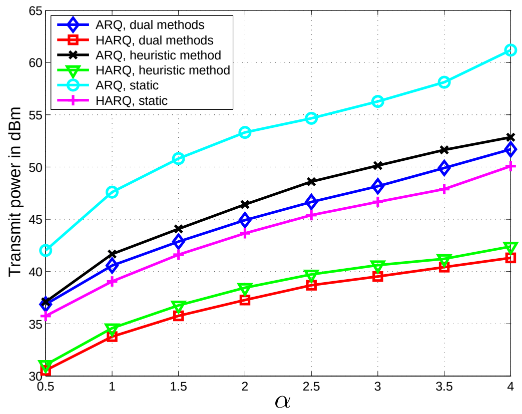

Figure 4. Average transmit power for 3 RA schemes at different α's

HARQ. The static RA scheme is still much worse, spending 6.4 dBm and 5.7 dBm more power for the two protocols than the heuristic method. Yet the biggest performance gap comes from the ARQ and HARQ protocols, being $8, 8.5$ and 9.2 dBm for the three RA schemes, which is in accordance with the situation present in Fig. 1.

VI. Conclusions

In this work we have presented two CLARA algorithms to solve the transmit power minimization problem under QoS constraints in multicarrier systems. Both algorithms provide suboptimal solutions but are highlighted for their low complexity. On the other hand, the fairly big advantage of using retransmission protocol HARQ over ARQ has been demonstrated. Although from another perspective of the resource minimization problem, *i.e.*, minimizing energy consumption [8], the advantage of HARQ is not as significant as here where we allow for the maximum number of retransmissions, it is still worthwhile to consider employing HARQ for more efficient resource usage at the expense of a higher coding complexity.

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 Journal on Select. Areas of Commu., pp. 1747-1758, vol. 17, 1999.

[2] I. C. Wong, Z. Shen, B. L. Evans and J. G. Andrews, *A low complexity algorithm for proportional resource allocation in OFDMA systems*, in IEEE Workshop on Signal Processing Systems, Oct. 2004, pp. 1-6.

[3] B. Zerlin, M. T. Ivrlač, W. Utschick, J. A. Nossek, I. Viering and A. Klein, *Joint optimization of radio parameters in HSDPA*, in IEEE 61st Vehicular Technology Conference VTC 2005-Spring, vol. 1, pp. 295-299.

[4] K. Seong, M. Mohseni and J. M. Cioffi, *Optimal Resource Allocation for OFDMA Downlink Systems*, in IEEE Int. Symposium on Information Theory, July 2006, pp. 1394-1398.

[5] W. Yu and R. Lui, *Dual Methods for Nonconvex Spectrum Optimization of Multicarrier Systems*, in IEEE Transactions on Communications, July 2006, vol. 54, no. 7, pp. 1310-1322.

[6] R. G. Gallager, *Information Theory and Reliable Communication*, John Wiley and Son, 1968.

[7] Q. Bai, J. A. Nossek, *A QoS-providing Resource Allocation Scheme in Multiuser Multicarrier Systems*, 2nd International Workshop on Cross-Layer Design, June 2009.

[8] Q. Bai, M. T. Ivrlač, J. A. Nossek, *QoS-constrained Energy Minimization in Multiuser Multicarrier systems*, 1st International Conference on Mobile Lightweight Wireless Systems, May 2009.