This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2011 proceedings

Optimal QoS-constrained Resource Allocation in Downlink and Uplink Multicarrier Systems

Qing Bai and Josef A. Nossek
Institute for Circuit Theory and Signal Processing
Technische Universität München, Munich, Germany
Email: {bai.qing, nossek}@nws.ei.tum.de

***Abstract*—When formulated as constrained optimizations of various kinds, resource allocation (RA) in wireless communication systems involves practically not only continuous variables such as transmit power, but also discrete variables such as modulation formats and coding rates. Multicarrier systems with frequency division multiple access (FDMA), *e.g.*, OFDMA, require exclusive usage of their subcarriers (or subchannels) by a single user during a certain time unit, and therefore render the assignment of subcarriers a combinatorial problem which is a crucial part of RA. In this paper, we formulate the *quality of service* (QoS) constrained RA problems in multicarrier systems as optimizations over a finite set of available *modes of operations* (MOP) and employ integer programming techniques to solve them. Specifically, optimal algorithms are developed based on the *branch and bound* (BAB) and the *branch and price* (BAP) methods for three types of RA problems: the downlink transmit power minimization, the uplink sum transmit power minimization, and the downlink transmit power constrained energy minimization. Besides justifying the optimality of these algorithms, simulation results also provide performance comparisons on ARQ and HARQ protocols which serve as an application example of the algorithms proposed, *i.e.*, providing comparison benchmark for different system, protocol, or suboptimal algorithm design.**

I. Introduction

In wireless communication systems, effective RA in the presence of perfect channel state information (CSI) enables better utilization of the wireless channel and the available radio resources as well as QoS provisioning with lower cost. Therefore, to find the optimal RA strategy has been one of the main research concerns in the area. As far as multicarrier systems with single antenna are concerned, optimal subcarrier and power allocation strategies were obtained in *e.g.*, [1][2][3], where the weighted sum rate maximization and weighted sum power minimization problems were solved in the dual domain, and optimality conditions of the algorithm have been proven. These works, as well as many of the others proposing suboptimal RA algorithms under certain scenarios and assumptions (*e.g.*, [4]), or employing cross-layer techniques to optimize certain QoS measures or utility functions (*e.g.*, [5]), usually adopt a capacity based approach where the effect of discrete modulation and coding levels on the achievable rates and coded bit error probability is not taken into account. In the cross-layer framework we adopt from [6], the system is restricted to using only a few *modulation and coding schemes* (MCS) which is a rather practical limitation, and the dependence of the *packet error probability* (PEP) on MCS and signal-to-noise-ratio is characterized by the cutoff rate and the channel coding theorem. Different from RA problems based on a capacity model, the RA problems we have formulated in previous contributions [7]-[10] are not only nonconvex but also discrete, causing a nonzero duality gap and thus prohibiting optimal solutions to be obtained from solving the dual problems. The lack of optimal solutions also makes evaluation of suboptimal algorithms difficult and implausible. In this work, we reformulate the RA problems into integer programming problems and develop optimal algorithms based on the BAB and BAP methods to tackle them.

The remaining of the paper is organized as follows: the scenario considered and the system model are described in Sec. II, while the original formulations of the downlink and uplink RA problems are introduced in Sec. III. In Sec. IV, the development and implementation of the optimal RA algorithms are explained in detail. Simulation results and the corresponding analysis are provided in Sec. V before we conclude the work in Sec. VI.

II. System Model

Consider K users in an isolated single-cell each having one service flow to receive from or to transmit to the BS. RA is done for each *transmission time interval* (TTI), and consecutive transmissions of data are assumed to be independent from TTI to TTI. Depending on its *throughput* requirement, each service flow may have a number of information bits, in form of *packet*, to transmit at the beginning of a TTI. The other relevant QoS parameter characterizing the service flows, *latency*, is defined as the delay a packet experiences until received correctly with an outage probability of no more than the predefined value $\pi^{(\mathrm{out})}$. Mathematically, let $f_k[m]$ be the probability that it takes exactly m TTI's to transmit a packet of user k error-free, then the latency of the packet is computed as $\tau_k = (M_k - 1)(T_\mathrm{R} + T_\mathrm{I}) + T_\mathrm{I}$ where T_I and T_R represent the length of a TTI and the *round trip delay*, and $M_k = \min M$ such that $\sum_{m=1}^{M} f_k[m] \geq 1 - \pi^{(\mathrm{out})}$.

We derive in the following the mathematical descriptions of the regarded system components stemmed from [6], which lay the basis for cross-layer optimization.

A. Channel Model

The wireless channel is modeled as frequency-selective fading over its whole bandwidth and frequency-flat fading over each *subchannel*, which consists of N_c adjacent subcarriers.

The assignment of any subchannel is exclusive, and *inter-carrier interference* is not taken into account. Moreover, we restrict ourselves here to the single-antenna case both at BS and MS. On a particular subchannel n, let $H_{k,n}$ and $\sigma^2_{k,n}$ be the channel coefficient and Gaussian noise variance of user k, and p_n be the amount of power being allocated. When assigned to user k, the *signal-to-noise-ratio* (SNR) 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. Assuming that one TTI contains N_s symbols for data transmission, the *minimum allocation unit* (MAU) is defined as an allocation region of one subchannel in the frequency dimension by one TTI in the time dimension, which contains $N_c N_s$ symbols.

B. FEC coding and modulation

With reference to the WiMAX standard, 8 MCS are chosen as candidates to be employed by the MAU's, 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 is discrete and memoryless over which the *noisy channel coding theorem* [11] 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 [6] and an expression for the *equivalent block length* is derived based on link level simulations as $n_{eq} = \beta \ln L$, 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

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

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 simply discarded, hence we assume that the PEP of a retransmitted packet is the same as that of its original transmission, *i.e.*, $f[m] = \pi^{m-1}(1-\pi)$, where $m \in \mathbb{Z}^+$, $\pi = \pi[1]$.

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 $n_{eq} = \beta \ln(\tilde{m} L)$. The PEP for the mth transmission can be approximated by

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

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 subchannels	N	
Number of subcarriers per subchannel	N_c	24
Transmission Time Interval (TTI)	T_I	2 ms
Symbol duration	T_s	
Number of data symbols per TTI	N_s	16
Round Trip Delay (RTD)	T_R	10 ms
Maximum number of transmissions allowed	$\tilde{m}$	5
Turbo code dependent parameter	β	32
Outage probability	$\pi^{(out)}$	0.01

III. DOWNLINK AND UPLINK RESOURCE ALLOCATION PROBLEMS

Consider data transmission over one MAU where we assume that every symbol and subcarrier in the MAU is loaded with data. That means, the number of loaded information bits B is determined by the MCS employed, *i.e.*, $B = N_c N_s R\log_2 A$. The transmit power required for the current transmission is given by

$$P = N_c \cdot \gamma(A, R, M) \cdot \frac{\sigma^2}{|H|^2},$$

and the expected energy consumption for the successful data transmission on the MAU is expressed as

$$E = T_s N_c N_s \cdot \gamma(A, R, M) \cdot \sum_{m=1}^{M} f[m] \left(\frac{\sigma^2}{|H|^2} + \frac{(m-1)\sigma^2}{|H|^2_{avg}} \right),$$

where $\gamma(A, R, M)$ is the SNR required to transmit the data successfully within M transmissions using MCS (A, R), and $|H|^2$ and $|H|^2_{avg}$ are the instantaneous and average channel gains. The triple (A, R, M) represents one MOP of the system and is denoted by q. Let the set of all available MOP's be $\mathcal{M}$.

As $\gamma(q)$ decreases monotonically with increasing M when A, R are fixed, transmit power P also decreases with more transmission trials. Therefore, M should be set to its maximal possible value when minimizing P, *i.e.*, $M_k^{(\mathrm{rq})} \triangleq \min\left\{\tilde{m}, \left\lfloor \frac{\tau_k^{(\mathrm{rq})} - T_\mathrm{I}}{T_\mathrm{R} + T_\mathrm{I}} \right\rfloor + 1\right\}$, where $\tau_k^{(\mathrm{rq})}$ is the maximum latency for data transmission of user k. Let the number of information bits to be transmitted to user k be $b_k^{(\mathrm{rq})}$. We formulate the DL transmit power minimization problem as

$$\begin{aligned} \min_{\boldsymbol{Q}\in\mathcal{Q}} \quad & \sum_{k=1}^{K}\sum_{n=1}^{N} P_{k,n}(Q_{k,n}) \\ \text{s.t.} \quad & \sum_{n=1}^{N} B(Q_{k,n}) \ge b_k, \quad k = 1, \dots, K, \\ & M(Q_{k,n}) \le M_k^{(\mathrm{rq})}, \quad k = 1, \dots, K, \end{aligned} \tag{1}$$

where $\boldsymbol{Q} \in \mathcal{M}^{K\times N}$ is the MOP matrix with its entry $Q_{k,n}$ as the MOP chosen for the kth user on the nth subchannel, and $\mathcal{Q} \subset \mathcal{M}^{K\times N}$ represents the set of matrices that have only one nonzero entry in each of their columns. $B(q)$, $M(q)$ and $P_{k,n}(q)$ denote the number of information bits, number of transmissions and transmit power required depending on MOP q, respectively. Note that at optimality, $M(Q_{k,n}) = M_k^{(\mathrm{rq})}, k = 1, \dots, K$ for nonzero $Q_{k,n}$, which means M is actually not considered as an optimization variable.

The UL sum transmit power minimization problem can be formulated similarly, except that each mobile user is limited by its own transmit power constraint, denoted by $P_k^{(\mathrm{av})}$:

$$\begin{aligned} \min_{\boldsymbol{Q}\in\mathcal{Q}} \quad & \sum_{k=1}^{K}\sum_{n=1}^{N} P_{k,n}(Q_{k,n}) \\ \text{s.t.} \quad & \sum_{n=1}^{N} B(Q_{k,n}) \ge b_k^{(\mathrm{rq})}, \quad k = 1, \dots, K, \\ & M(Q_{k,n}) \le M_k^{(\mathrm{rq})}, \quad k = 1, \dots, K, \\ & \sum_{n=1}^{N} P_{k,n}(Q_{k,n}) \le P_k^{(\mathrm{av})}, \quad k = 1, \dots, K. \end{aligned} \tag{2}$$

It has been pointed out in [10] based on analysis and simulation results from [8], that to allow for more transmissions is energy inefficient in general. In fact, transmit power and energy consumption as defined are two competing objectives in the QoS-constrained RA problem. This conclusion has given rise to the DL energy consumption minimization problem under transmit power constraint formulated as

$$\begin{aligned} \min_{\boldsymbol{Q}\in\mathcal{Q}} \quad & \sum_{k=1}^{K}\sum_{n=1}^{N} E_{k,n}(Q_{k,n}) \\ \text{s.t.} \quad & \sum_{n=1}^{N} B(Q_{k,n}) \ge b_k^{(\mathrm{rq})}, \quad k = 1, \dots, K, \\ & M(Q_{k,n}) \le M_k^{(\mathrm{rq})}, \quad k = 1, \dots, K, \\ & \sum_{k=1}^{K}\sum_{n=1}^{N} P_{k,n}(Q_{k,n}) \le P^{(\mathrm{tx})}, \end{aligned} \tag{3}$$

where $P^{(\mathrm{tx})}$ denotes the total available transmit power at the BS. In (3), not only the MCS but also the number of transmissions are active optimization variables.

Problems (1)(2)(3) clearly have very similar structures, differing mainly in a few constraints. However, such a difference has already much influence on heuristic algorithm design. In the next section we will see that the BAB and BAP methods provide a unified framework for solving these RA problems optimally.

IV. Optimal Resource Allocation Algorithms

A. Reformulation of the RA problems

Since no time- or frequency-sharing of MOP is allowed within one MAU, 0-1 variables can be used to indicate whether a MOP is used. To this end, a vector of $N|\mathcal{M}|$ binary variables completely characterizes the transmission mode of one user, *i.e.*, on which subchannel it is transmitting or receiving with which MOP. Corresponding to such a vector for user k, which is denoted by $\boldsymbol{x}_k$, there is a power vector $\boldsymbol{p}_k$ whose elements are the transmit power values that are required to support the respective MOP and the QoS constraint of the user. Putting the K vectors together, we obtain the *selection matrix* $\boldsymbol{X}$ and the *power matrix* $\boldsymbol{P}$, both of dimension $N|\mathcal{M}| \times K$, as

$$\boldsymbol{X} = [\boldsymbol{x}_1, \dots, \boldsymbol{x}_K], \quad \boldsymbol{P} = [\boldsymbol{p}_1, \dots, \boldsymbol{p}_K].$$

With these notations the RA problems can be reformulated as 0-1 integer programming problems. We show here only the reformulation of (1) as an example: Problem (1) is equivalent to the integer programming

$$\begin{aligned} P_{\min} \triangleq \min_{\boldsymbol{X}} \quad & \mathrm{tr}\left(\boldsymbol{P}^\mathrm{T}\boldsymbol{X}\right) \\ \text{s.t.} \quad & \boldsymbol{S}\boldsymbol{X}\mathbf{1}_K \preceq \mathbf{1}_N, \\ & \boldsymbol{b}^\mathrm{T}\boldsymbol{x}_k \ge b_k^{(\mathrm{rq})}, \quad k = 1, \dots, K, \\ & \boldsymbol{X} \in \{0,1\}^{N|\mathcal{M}|\times K}, \end{aligned} \tag{4}$$

where $\boldsymbol{S} \in \{0,1\}^{N\times N|\mathcal{M}|}$ is a summation matrix for calculating the usage of each subchannel, *i.e.*, $S_{n,m} = 1$ for $n = 1, \dots, N, m = (n-1)|\mathcal{M}| + 1, \dots, n|\mathcal{M}|$, and $\boldsymbol{b}$ is the column vector containing the numbers of information bits when each MOP is employed.

B. BAB and BAP methods

The BAB algorithm is one of the most commonly used integer programming methods [12]. Being an enumeration method in essence, it promises to give the global optimal solution irrespective to the convexity of the problem. The basic idea is to partition the feasible set of the problem into convex sets and find upper and lower bounds on the objective function for each set [13]. Global upper and lower bounds are formed and if they are close enough, the algorithm terminates. If not, one of the sets is chosen to be further partitioned. The BAB method converges often slowly and has exponential worst case performance.

Over the past few decades, column generation has proven to be one of the most successful approaches for solving large-scale integer programming problems [14][15]. Among its numerous applicatios, the BAP algorithm for the *generalized assignment problem* (GAP) [16][17] appears to be the most promising one to be adapted to solve our RA problem due to the similarity in problem structure. The BAP integrates the

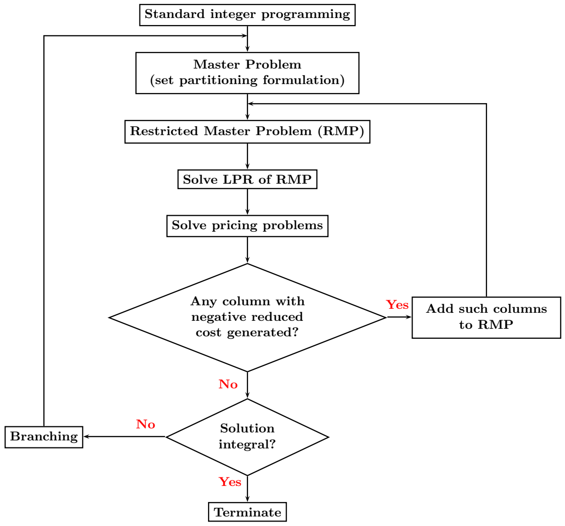

Figure 1. Flow chart of the BAP algorithm

BAB and the column generation methods and is more efficient than BAB, for it works with a Dantzig-Wolfe decomposed formulation of the original problem which has a tighter *linear programming relaxation* (LPR). A flow chart of the BAP algorithm is provided in Fig. 1, and the explanation of the whole procedure is left to the next subsection when it is applied to solve Problem (4).

C. Optimal RA algorithms and their implementations

We first apply the BAB algorithm to (4) which is summarized in Algorithm 1.

Algorithm 1 Branch and Bound (BAB) algorithm

1: Solve the LPR of (4) by using any standard algorithm
2: Terminate if LPR of (4) is infeasible
3: Compute global upper and lower bound U and L on $P_{\min}$
4: Terminate if $\frac{U-L}{L} < \epsilon$
5: Select a fractional variable x and create two branches by setting $x = 0$ and $x = 1$
6: Solve the LPR's of the two branches and compute the respective upper and lower bounds
7: Update global upper and lower bounds if necessary
8: Terminate if $\frac{U-L}{L} < \epsilon$
9: Choose the branch with lower LPR objective and repeat steps 5–8

If the elements of the selection matrix $\boldsymbol{X}$ can take real values between 0 and 1, (4) becomes a linear optimization that can be solved by any standard linear optimizer, and the optimal value of this relaxed problem provides a lower bound on $P_{\min}$, as its domain is larger than (4). Note that if the relaxed problem is infeasible, (4) can be determined to be infeasible. As any feasible solution to (4) provides an upper bound on the problem, at each branch, we try to recover a feasible solution based on the fractional solution to LPR. A heuristic algorithm is designed to accomplish this.

If the number of fractional variables in the LPR solution is very small, all possible combinations of them can be enumerated and the one that leads to the minimum objective is taken. Otherwise, we can find the set of subchannels that are shared by different users and determine the owners of these subchannels according to the fractions each user gets on them as well as how important they are to each user. For example, if a user only holds a fraction of one particular subchannel and no other subchannels, this subchannel has to be assigned to the user; without such special conditions a shared subchannel is assigned to the user with the largest fraction of it. After one shared subchannel is assigned, the reduced LPR is resolved and the procedure repeats until there are no more shared subchannels. Although this recovery scheme is not optimal, it generates in many cases good feasible solution in the early phase of the BAB algorithm. Note that the optimality of BAB is not destroyed by a suboptimal recovery scheme. The algorithm is implemented recursively and requires a lot of memory when executed. At each node of the BAB tree, at least one linear optimization of size $K \cdot N \cdot |\mathcal{M}|$ is solved.

As the first step of applying the BAP algorithm, the transmit power minimization problem as in (4) has to be equivalently written in the so-called *set partioning formulation*. Let $\mathcal{L}_k = \{\boldsymbol{a}_1^k, \ldots, \boldsymbol{a}_{L_k}^k\}$ be the set of all feasible bit-loadings of user k, where any element $\boldsymbol{a}_j^k \in \{0,1\}^{N|\mathcal{M}|\times 1}$ satisfies

$$\boldsymbol{b}^{\mathrm{T}}\boldsymbol{a}_j^k \geq b_k^{(\mathrm{rq})}, \quad \boldsymbol{S}\boldsymbol{a}_j^k \preceq \boldsymbol{1}_N,$$

and corresponds to the transmit power $\boldsymbol{p}_k^{\mathrm{T}}\boldsymbol{a}_j^k$. Such a vector is referred to as a *pattern*. If patterns of all users form the columns of a large matrix $\tilde{\boldsymbol{A}}$ with a certain order and the transmit power values associated with these patterns are placed with the same order in a vector $\tilde{\boldsymbol{c}}$, the set partioning formulation of (4) is obtained as

$$\begin{array}{lrcl} \min\limits_{\tilde{\boldsymbol{y}}} & \tilde{\boldsymbol{c}}^{\mathrm{T}}\tilde{\boldsymbol{y}} & & \\ \text{s.t.} & \boldsymbol{S}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{y}} & \preceq & \boldsymbol{1}_N, \\ & \tilde{\boldsymbol{G}}\tilde{\boldsymbol{y}} & = & \boldsymbol{1}_K, \\ & \tilde{y}_j & \in & \{0,1\}, \quad j = 1, \ldots, \sum_{k=1}^{K} L_k, \end{array} \tag{5}$$

where the binary optimization variables $\tilde{\boldsymbol{y}}$ indicate the usage of each pattern, the first set of constraints states that at most one MOP should be present on each MAU, and the second set of constraints enforces exactly one pattern to be chosen for one user (matrix $\tilde{\boldsymbol{G}}$ records the owner of each pattern). Problem (5) is referred to as the *master problem* (MP) and it has prohibitively many variables. As most of the patterns are not chosen at optimality, we work with only a small subset of explicitly given patterns called the *basis*, which form matrix $\boldsymbol{A}$, and deal with the *restricted master problem* (RMP)

$$\begin{array}{lrcl} \min\limits_{\boldsymbol{y}} & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{y} & & \\ \text{s.t.} & \boldsymbol{S}\boldsymbol{A}\boldsymbol{y} & \preceq & \boldsymbol{1}_N, \\ & \boldsymbol{G}\boldsymbol{y} & = & \boldsymbol{1}_K, \\ & y_j & \in & \{0,1\}, \quad j = 1, \ldots, L, \end{array} \tag{6}$$

where L is the size of the basis and all quantities irrelevant to the basis have been taken out as compared to (5).

With a proper initial basis the LPR of the RMP can be solved and the optimal dual variables associated with each of the constraints can be obtained. Additional columns for the RMP can be generated by solving the *pricing problem*

$$\begin{aligned} &\min_k z_k - v_k \quad \text{where} \\ &z_k = \min_{\boldsymbol{a}\in\{0,1\}^{N|\mathcal{M}|\times 1}} \left(\boldsymbol{p}_k - \boldsymbol{S}^{\mathrm{T}}\boldsymbol{u}\right)^{\mathrm{T}} \boldsymbol{a} \\ &\qquad \text{s.t.} \quad \boldsymbol{b}^{\mathrm{T}}\boldsymbol{a} \geq b_k^{(\mathrm{rq})}, \\ &\qquad\qquad \boldsymbol{S}\boldsymbol{a} \preceq \mathbf{1}_N, \end{aligned} \tag{7}$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ are the optimal dual variables associated with the first and second set of constraints in (6) respectively. Objective of (7) is referred to as the *reduced cost* associated with the generated pattern. If the optimal value of (7) is non-negative, then any generated pattern can not help in reducing the objective of the RMP, which means that the LPR of the MP has been solved. Otherwise, the pattern that optimally solves (7), or any set of patterns with negative reduced costs, can be added to the RMP. The remaning part of the BAP algorithm as indicated in Fig. 1, is the BAB procedure. Note that here the branching rule has to be compatible with the pricing problem, which is accomplished by branching using the standard formulation of the problem, while solving the LPR of the set partitioning formulation of the problem.

To ensure that the LPR of RMP is always feasible, we keep one dummy column for every user at each node of the BAB tree in implementation. That is, a column with all zero entries but a very large transmit power. To this end, we have the option to start the algorithm with only dummy columns, *i.e.*, an empty basis. As an alternative, we can also initialize the RMP with patterns generated using a heuristic algorithm, *e.g.*, from [9]. Analysis on both schemes is given in the next section. The pricing problem is solved optimally using the BAB method which is quite expensive as in each iteration of the algorithm, K pricing problems of dimension $N|\mathcal{M}|$ have to be solved. As the cost of solving the LPR of RMP is relatively lower, we have chosen to let all generated patterns with a negative reduced cost to enter the basis.

V. Simulation Results

We choose two test cases as listed in Table III for simulations. All the other parameters remain the same as those taken in [8], otherwise stated when encountered.

Table III
Key test parameters

Test	K	N	$(b_k^{(\mathrm{rq})}, \tau_k^{(\mathrm{rq})})$ / (bits, ms)	Scaling factor α
I	3	8	$(128, 20)$, $(400, 40)$, $(800, 100)$	$1, 2, \ldots, 8$
II	2	4	$(128, 20)$, $(400, 40)$	8

First, test case II is used to validate the algorithms implemented, for the optimal solution to (4) can be obtained with enumeration. Then, a heuristic algorithm from [9], the BAB and the BAP algorithms are applied to solve (4) with test case I ($\alpha = 8$), where the two initialization schemes of the BAP algorithm as discussed in Sec. IV are both simulated. Besides,

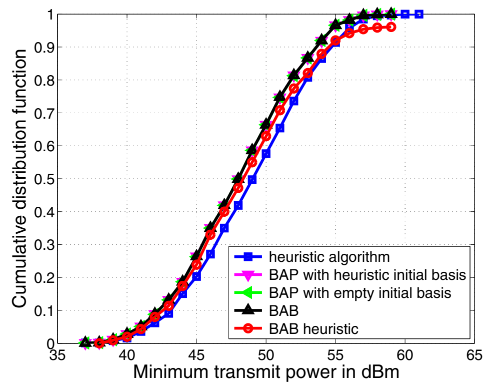

Figure 2. CDF of objective value of (1), Test I with ARQ protocol

when the BAB algorithm is applied, the global upper bound U at the root node of the BAB tree is also recorded. The cumulative distribution of the minimized transmit power under 1000 independent channel realizations are shown in Fig. 2. The three curves representing the results of BAB, BAP with heuristic initial basis, and BAP with empty initial basis are almost identical, except for the negligible difference caused by the termination accuracy $\epsilon = 0.05$ which might not be small enough for some cases. The BAB algorithm takes a significantly larger amount of time to find the optimal solution as compared to BAP. A little surprisingly, the BAP algorithm with empty initial basis is not much slower than that with heuristic initial basis on average. The explanation is that when a pattern identified by the heuristic algorithm is not optimal, it does not necessarily give benifit to the generation of optimal patterns but increases the dimension of the RMP. The heuristic solution given as a byproduct of BAB performs in many cases better than the heuristic algorithm, but it requires more time to be computed and there is no guarantee that a feasible solution can be found in the first iteration of BAB.

Traffic density is varied by multiplying $b_k^{(\mathrm{rq})}, k = 1, \ldots, K$ with the scaling factor α. In Fig. 3, the minimum transmit power required under different traffic densities obtained with the heuristic algorithm from [9] and the BAP algorithm are shown, where both ARQ and HARQ retransmission protocols are simulated. The transmit power at each marked point in the figure is the averaged value over 100 independent channel realizations. Obviously, the HARQ protocol outperforms the ARQ protocol and its advantage is even more significant when the traffic density is high. Although the heuristic algorithm exhibits near-optimal performance especially in the low traffic density region, this conclusion is drawn more convincingly with the optimal results provided by BAP.

The successful implementation of the BAB and BAP algorithms to the UL sum transmit power minimization problem (2) and the DL transmit power constrained energy minimization problem (3) are demonstrated by Fig. 4 and Fig. 5, respectively. Fig. 4 shows the cumulative distribution of the obtained optimal value of (2) for test case I ($\alpha = 8$), where $P_1^{(\mathrm{av})} = P_2^{(\mathrm{av})} = 200$ mW and $P_3^{(\mathrm{av})} = 400$ mW. For

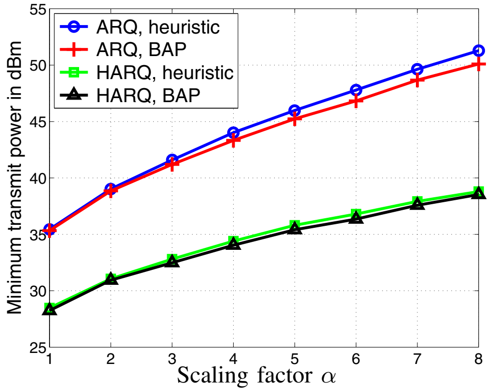

Figure 3. Averaged objective value of (1) under different traffic densities

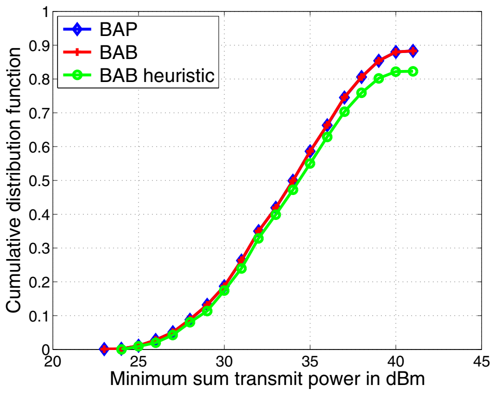

Figure 4. CDF of objective value of (2), Test I with ARQ protocol

problem (3) where $|\mathcal{M}|$ is $\tilde{m}$ times larger than that of the other two problems, test case II is simulated where the transmit power constraint is set to 48 and 36 dBm. The ratio between the minimum energy consumption obtained with the heuristic algorithm [10] and with the BAP algorithm is calculated for 1000 simulations and its probability density is shown in Fig. 5 in logarithmic scale. In addition to providing the suboptimality measure, the optimal results also indicate whether the feasibility decision made by the heuristic algorithm is true or not. For example, in the case of ARQ and $P^{(\mathrm{tx})} = 48$ dBm, the heuristic algorithm fails to find any feasible solutions in 270 simulations, but the optimal algorithm determines that only 250 of these problems are truly infeasible.

VI. Conclusions

In a multiuser multicarrier system with discrete MOP's, the QoS-constrained RA problems are formulated as 0-1 integer programming problems where each optimization variable corresponds to the employment of one MOP. Optimal RA algorithms are developed and implemented based on the BAB and BAP methods. Although suffering from high complexity, the proposed optimal algorithms provide benchmarks for comparative studies on different system designs or suboptimal RA methods. What is more, the algorithms can be further

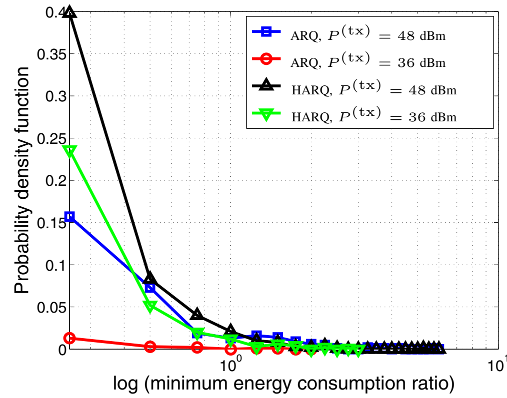

Figure 5. PDF of ratio between objective values of (3) from heuristic and BAP algorithm in logarithmic scale

accelerated by an efficient integral solution recovery scheme, which, perhaps even more favorably, enables early termination of the search procedure with a good intermediate solution.

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] 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.

[3] 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.

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

[5] G. Song and Y. Li, *Cross-layer optimization for OFDM wireless networks-part I: theoretical framework*, in IEEE Trans. on Wireless Communications, March 2005.

[6] 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.

[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.

[9] Q. Bai and J. A. Nossek, *On Two Cross-layer Assisted Resource Allocation Schemes in Multiuser Multicarrier Systems*, 6th Int. Symposium on Wireless Communication Systems, Sept. 2009.

[10] Q. Bai and J. A. Nossek, *On Energy Efficient Cross-layer Assisted Resource Allocation in Multiuser Multicarrier Systems*, 20th Int. Symposium on Personal, Indoor and Mobile Radio Communications, Sept. 2009.

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

[12] H. M. Salkin and K. Mathur, *Foundations of integer programming*, Chapter 8, North-Holland.

[13] S. Boyd, *Branch and Bound Methods, EE364b, Stanford University*, http://www.stanford.edu/class/ee364b/lectures/bb_slides.pdf.

[14] W. E. Wilhelm, *A Technical Review of Column Generation in Integer Programming*, Optimization and Engineering, Vol. 2, pp. 159-200, 2001.

[15] M. E. Lübbecke and J. Desrosiers, *Selected Topics in Column Generation*, Operations Research, Vol. 53, No. 6, pp. 1007-1023, 2005.

[16] M. Savelsbergh, *A branch-and-price algorithm for the generalized assignment problem*, Operations Research, Vol. 45, No. 6, 1997.

[17] C. Barnhart, E. L. Johnson, G. L. Nemhauser, M. W. P. Savelsbergh and P. H. Vance, *Branch-and-price: Column generation for solving huge integer programs*, Operations Research, Vol. 46, No. 3, 1998.