Joint Scheduling and Resource Allocation in Uplink OFDM Systems for Broadband Wireless Access Networks

Jianwei Huang, Vijay G. Subramanian, Rajeev Agrawal, and Randall Berry

Abstract—Orthogonal Frequency Division Multiplexing (OFDM) with dynamic scheduling and resource allocation is a key component of most emerging broadband wireless access networks such as WiMAX and LTE (Long Term Evolution) for 3GPP. However, scheduling and resource allocation in an OFDM system is complicated, especially in the uplink due to two reasons: (1) the discrete nature of subchannel assignments, and (2) the heterogeneity of the users' subchannel conditions, *individual* resource constraints and application requirements. We approach this problem using a gradient-based scheduling framework. Physical layer resources (bandwidth and power) are allocated to maximize the projection onto the gradient of a total system utility function which models application-layer Quality of Service (QoS). This is formulated as a convex optimization problem and solved using a dual decomposition approach. This solution has prohibitively high computational complexity but reveals guiding principles that we use to generate lower complexity sub-optimal algorithms. We analyze the complexity and compare the performance of these algorithms via simulation.

Index Terms—Orthogonal Frequency Division Multiplexing (OFDM), scheduling, resource allocation, optimization, dual decomposition, uplink communications

I. INTRODUCTION

Orthogonal Frequency Division Multiplexing (OFDM) is the core technology for most recent wireless data systems, including IEEE 802.16 (WiMAX), IEEE 802.11a/g (Wireless LANs), and LTE for 3GPP. In this paper, we analyze an uplink scheduling and resource allocation problem for OFDM wireless access networks. The specific problem is motivated by the WiMAX/802.16e standard, where there is a centralized scheduler that knows the QoS classes, and can estimate the queue-lengths on each mobile device. The WiMAX/802.16e standard specifies mechanisms for communicating this information to the scheduler and for conveying the scheduling decisions to the mobiles, both with low delays.[1]

Our approach is motivated by our previous work on down-link scheduling in CDMA systems [3] and OFDM systems [4]. As in [3], [4], we consider a *gradient-based* scheduling framework, which is described in detail in Section II along with our system model. In this framework, the time-varying gradient of a utility function is used to guide the resource allocation decisions and provide long-term Quality of Service (QoS) guarantees. In particular, we maximize a weighted sum-rate during each scheduling interval, where the weights are time-varying. The optimization variables are the assignment of OFDM subchannels to the users and the allocation of each user's power across the assigned subchannels. We highlight two challenging aspects of this problem in the OFDM uplink context. First, the discrete nature of subchannel assignments in OFDM systems usually leads to hard integer programming problems. Second, the per-user power constraints in the uplink make the problem even less tractable. We initially relax the integer constraints and allow multiple users to share one subchannel using orthogonalization (e.g. via time-sharing[2]). In Section III we derive an optimal solution to this re-laxed problem via a dual decomposition. Due to the per-user power constraints, the resulting algorithm has very high computational complexity. However, this provides insight into the structure of an optimal solution. In Section V we use the insights gained from the optimal solution to propose a family of sub-optimal algorithms that also take into account the integer constraints on subchannel allocations. Finally, in Section VI we these algorithms using simulation.

Most initial work on OFDM scheduling and resource allocation focused on the downlink case. The optimality conditions and algorithms given for the downlink, however, can not be directly applied to the uplink due to differences in the resource constraints (see Section IV). Recently, uplink OFDM resource allocation has received some attention, including [21]–[28]. In [21], an iterative OFDM resource allocation was proposed to find a Nash Bargaining solution. A heuristic algorithm that tries to minimize each user's transmission power while satisfying the individual rate constraints is given in [22]. In

---

[1] Our model is also appropriate for LTE [29] for 3GPP, UMB [30] for 3GPP2 and the FLASH OFDM system [19] from Qualcomm Flarion.

[2] While super-position coding would yield an even larger (and more tractable) capacity region, we do not consider it as it is still not practical.

[23], an algorithm for maximizing the sum-rate assuming Rayleigh fading is given; this is a special case of the problem considered here with equal weights. In our case no assumption on the fading distribution is made. In [24], an uplink problem with multiple antennas at the base station is considered; here, we focus on single antenna systems.

The work in [25]–[28] is closer to ours. In [25], weighted rate maximization problem is also considered for the uplink, but with static weights. The results in [25] are generalized in [26] to account for per-time slot fairness via a utility function of the instantaneous rate. Per time-slot fairness is also considered in [27]. Our work differs from these in that by using a gradient-based scheduler, we can consider *long-term* fairness, which depends on the average rate or queue sizes. For elastic data applications, long-term QoS evaluation is more reasonable than the short-term QoS evaluation during each time slot. It not only more faithfully reflects users' actual perceived performance, but also gives the system more flexibility in terms of exploiting multi-user diversity. Finally, [28] proposed a heuristic algorithm based on Lagrangian relaxation, which has high complexity due to a subgradient search of the dual variables. Here we use Lagrangian relaxation to give an optimal solution the uplink problem when time-sharing is allowed. Solving this problem provides both an upperbound on the actual system performance as well as the intuition we use to design good heuristic algorithms.

II. PROBLEM STATEMENT

We consider the problem of scheduling and resource allocation for the uplink of a OFDM cell in which a set $\mathcal{M} = \{1, \ldots, M\}$ of users are transmitting to the same base station. The total frequency band is divided into a set $\mathcal{N} = \{1, \ldots, N\}$ of subchannels (e.g., frequency bands). Let p_{ij} be the power user i allocates to subchannel j, which is subject to a per-user power constraint:

$$\sum_j p_{ij} \leq P_i, \forall i \in \mathcal{M}. \qquad (1)$$

Let x_{ij} be the fraction of subchannel j allocated to user i, where the total allocation across all users should be no larger than 1, i.e.,

$$\sum_i x_{ij} \leq 1, \forall j \in \mathcal{N}. \qquad (2)$$

We use bold symbols to denote vectors and matrices of these quantities, e.g., $\boldsymbol{P} = \{P_i, \forall i\}$, $\boldsymbol{e} = \{e_{ij}, \forall i, j\}$, $\boldsymbol{p} = \{p_{ij}, \forall i, j\}$, and $\boldsymbol{x} = \{x_{ij}, \forall i, j\}$.

Time is divided into equal length slots. At the beginning of every time slot, the scheduler seeks to maximize a (time-varying) weighted sum of the users' rates over a (time-varying) rate-region. We describe this rate-region next.

The scheduler is assumed to have knowledge of the normalized received signal-to-interference plus noise ratio (SINR) per unit transmit power, e_{ij}, for each user i and subchannel j.[3] The time-varying subchannel quality vector at time t is

denoted by $\boldsymbol{e}_t$. As in [4], this model incorporates various subchannelization schemes where the resource allocation is performed in terms of subchannels (i.e., sets of tones); e_{ij} represents a collective quality indicator for the subchannel, e.g., the (harmonic/geometric/arithmetic) average across the tones in the subchannel. This model also applies if resource allocation is done with a granularity of multiple symbols in the time domain.

We model the feasible rate region at time t by

$$\mathcal{R}(\boldsymbol{e}_t) = \left\{ \boldsymbol{r} : r_i = \sum_{j \in \mathcal{N}} f(x_{ij}, p_{ij} e_{ij}(t)) \right\}, \qquad (3)$$

where $(\boldsymbol{x}, \boldsymbol{p}) \in \mathcal{X}$ are chosen subject to (1) and (2) and the set

$$\mathcal{X} := \left\{ (\boldsymbol{x}, \boldsymbol{p}) \geq \boldsymbol{0} : 0 \leq x_{ij} \leq 1, \ p_{ij} \leq \frac{x_{ij} s_{ij}}{e_{ij}(t)} \ \forall i, j \right\}. \qquad (4)$$

Here, $f(a, b) = a \log(1 + \frac{b}{a})$ so that r_i is the achievable rate of user i in a Gaussian Multiple Access Channel using time-sharing (cf. [17, Section 15.3.6, pg. 547]). By continuity, we assume that $f(0, b) = 0$. The value of s_{ij} is a maximum SINR constraint on subchannel j for user i, which can model scenarios when users have limited choices of modulation and coding schemes.

In practical OFDM systems, x_{ij} is constrained to be an integer, i.e., we have the additional constraint $x_{ij} \in \{0, 1\}$ for all i, j. Initially, we ignore this constraint, and consider a system in which users can share each tone. If resource allocation is done on blocks of OFDM symbols, then fractional values of x_{ij} can be implemented by time-sharing the symbols in a block. Alternatively, this can also be implemented by frequency sharing (e.g., [31]), if there are a large number of subchannels with roughly equal gains. We will re-introduce these integer constraints in Sections IV and V.

Next we formulate the scheduling and resource allocation problem. Our approach is based on the gradient-based scheduling framework in [2], [10], [12]. Each user i is assigned a utility function $U_i(W_{i,t}, Q_{i,t})$ depending on their average throughput $W_{i,t}$ up to time t and their queue-length $Q_{i,t}$ at time t. This is used to quantify fairness and ensure stability of the queues. At the beginning of each time slot t, the scheduler chooses a $\boldsymbol{r}_t \in \mathcal{R}(\boldsymbol{e}_t)$ that maximizes a weighted sum of the users' rates, where the weights are determined by the gradient of the sum utility across all users, i.e., it solves

$$\max_{\boldsymbol{r}_t \in \mathcal{R}(\boldsymbol{e}_t)} \left(\nabla_{\boldsymbol{w}} U(\boldsymbol{W}_t, \boldsymbol{Q}_t) - \nabla_{\boldsymbol{q}} U(\boldsymbol{W}_t, \boldsymbol{Q}_t) \right)^T \boldsymbol{r}_t, \qquad (5)$$

where $U(\boldsymbol{W}_t, \boldsymbol{Q}_t) = \sum_{i=1}^{K} U_i(W_{i,t}, Q_{i,t})$. Further assuming that for each user i, $U_i(W_{i,t}, Q_{i,t}) = u_i(W_{i,t}) - \frac{d_i}{p}(Q_{i,t})^p$, then (5) is equivalent to

$$\max_{\boldsymbol{r}_t \in \mathcal{R}(\boldsymbol{e}_t)} \sum_i \left(\frac{\partial u_i(W_{i,t})}{\partial W_{i,t}} + d_i(Q_{i,t})^{p-1} \right) r_{i,t}, \qquad (6)$$

where, $u_i(W_{i,t})$ is a increasing concave function used to represent elastic data applications [1], [8], [13], [18], $d_i \geq 0$ is a QoS weight for user i's queue length, and $p > 1$ is a fairness parameter associated with the queue length,

---

[3]In both FDD and TDD systems this can be obtained using a combination of measurements made on the uplink pilots as well as past transmissions from the mobiles.

The broad class of policies in (6) can be tuned to yield good operating points by a proper choice of parameters. If $d_i = 0$ for all $i \in \mathcal{M}$, the resulting policy has been shown to yield utility maximizing solutions [2], [10], [12]. If $u_i(\cdot) \equiv 0$ with $d_i > 0$ for all $i \in \mathcal{M}$, then the policy has been shown to be stabilizing in a variety of settings [5]–[7]. The weights can also be adapted so as to maximize sum utility subject to stability [9] or (feasible) minimum throughput constraints. [11]

More generally, the optimization in (6) can be written as

$$\max_{\boldsymbol{r}_t \in \mathcal{R}(\boldsymbol{e}_t)} \sum_i w_{i,t} r_{i,t}, \qquad (7)$$

where $w_{i,t} \geq 0$ is a time-varying weight assigned to the ith user at time t. Our focus is on solving such a problem for an uplink OFDM system, i.e., when $\mathcal{R}(\boldsymbol{e}_t)$ is given by (3). Note that (7) must be re-solved at each scheduling instant because of changes in both the subchannel state, $\boldsymbol{e}_t$, and the weights. While in the above examples, the weights $w_{i,t}$ were given by the gradient of an utility function, our algorithms also apply to other methods for generating these weights.

III. OPTIMAL SOLUTION WITH FRACTIONAL ALLOCATIONS

We now consider the optimal solution to (7) when $\mathcal{R}(\boldsymbol{e}_t)$ is given by (3). Suppressing the time index, the problem is

$$\max_{(\boldsymbol{x},\boldsymbol{p}) \in \mathcal{X}} \sum_{i \in \mathcal{M}} w_i \sum_{j \in \mathcal{N}} x_{ij} \log\left(1 + \frac{p_{ij} e_{ij}}{x_{ij}}\right) \qquad \text{(UL)}$$

subject to the per subchannel assignment constraints in (2) and the per user power constraints in (1), where $\mathcal{X}$ is given in (4).

It can be shown that Problem UL has no duality gap and so we solve it via a dual formulation. We associate dual variables $\boldsymbol{\lambda} = (\lambda_i)_{i \in \mathcal{M}}$ with constraints (1) and $\boldsymbol{\mu} = (\mu_j)_{j \in \mathcal{N}}$ with constraints (2), resulting in the Lagrangian,

$$L(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{x}, \boldsymbol{p}) := \sum_{i,j} w_i x_{ij} \log\left(1 + \frac{p_{ij} e_{ij}}{x_{ij}}\right) \\ + \sum_i \lambda_i \left(P_i - \sum_j p_{ij}\right) + \sum_j \mu_j \left(1 - \sum_i x_{ij}\right). \quad (8)$$

Therefore, the optimal solution to Problem UL is given by

$$\min_{(\boldsymbol{\lambda}, \boldsymbol{\mu}) \geq \boldsymbol{0}} \max_{(\boldsymbol{x}, \boldsymbol{p}) \in \mathcal{X}} L(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{x}, \boldsymbol{p}). \qquad (9)$$

We solve this problem by the following steps. First, we analytically find the optimal $\boldsymbol{p}$ and $\boldsymbol{x}$ given fixed values of the dual variables. We then show that the optimal $\boldsymbol{\mu}$ is given by a search for the maximum value of a per-user metric on each subchannel. The final step is to numerically search for the optimal value of $\boldsymbol{\lambda}$.

The value of $\boldsymbol{p}$ which maximizes $L(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{x}, \boldsymbol{p})$ given $\boldsymbol{x}$, $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ is given bu

$$p_{ij}^* = \frac{x_{ij}}{e_{ij}} \min\left\{ \left(\frac{w_i e_{ij}}{\lambda_i} - 1\right)^+, s_{ij} \right\}, \qquad (10)$$

where $(\cdot)^+ = \max(\cdot, 0)$. Substituting $\boldsymbol{p}^*$ into $L(\cdot, \cdot, \cdot, \cdot)$ yields

$$L(\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{x}) = \sum_{ij} x_{ij} \left(w_i h\left(\lambda_i, w_i e_{ij}, s_{ij}\right) - \mu_j\right) \\ + \sum_j \mu_j + \sum_i \lambda_i P_i, \quad (11)$$

where we have used the function $h(\cdot, \cdot, \cdot)$ from [3]; namely,

$$h(a, b, c) = \begin{cases} 0 & \text{if } a \geq b; \\ \frac{a}{b} - 1 - \log\frac{a}{b} & \text{if } \frac{b}{1+c} \leq a < b; \\ \log(1 + c) - \frac{a}{b} c & \text{if } a < \frac{b}{1+c}, \end{cases} \quad (12)$$

where $a \geq 0$, $b > 0$ and $c \geq 0$. Optimizing (11) over $\boldsymbol{x}$ such that $x_{ij} \in [0, 1]$ yields

$$L(\boldsymbol{\lambda}, \boldsymbol{\mu}) = \sum_j \mu_j + \sum_i \lambda_i P_i \\ + \sum_{ij} \left(w_i h\left(\lambda_i, w_i e_{ij}, s_{ij}\right) - \mu_j\right)^+, \quad (13)$$

where the optimal subchannel allocation has the following structure

$$x_{ij}^*(\mu_j) = \begin{cases} 1, & \text{if } w_i h\left(\lambda_i, w_i e_{ij}, s_{ij}\right) > \mu_j; \\ [0,1], & \text{if } w_i h\left(\lambda_i, w_i e_{ij}, s_{ij}\right) = \mu_j; \\ 0, & \text{if } w_i h\left(\lambda_i, w_i e_{ij}, s_{ij}\right) < \mu_j. \end{cases} \quad (14)$$

Since the cost function in (13) is separable, by defining $\mu_{ij}(\cdot) := w_i h(\cdot, w_i e_{ij}, s_{ij})$ as in [3], we can minimize $L(\boldsymbol{\lambda}, \boldsymbol{\mu})$ over $\boldsymbol{\mu}$ for a given $\boldsymbol{\lambda}$ by setting $\mu_j = \mu_j^*(\boldsymbol{\lambda})$ given by

$$\mu_j^*(\boldsymbol{\lambda}) = \max_i \mu_{ij}(\lambda_i). \qquad (15)$$

From (14) and (15), it is clear that $x_{ij}^*(\mu_j^*(\boldsymbol{\lambda})) \equiv 0$ if $i \notin \arg\max_{i \in \mathcal{M}} \mu_{ij}(\lambda_i)$, i.e., users not maximizing a specific subchannel metric are not allocated the subchannel. There will be ties when multiple users achieve the value μ_j^* on subchannel j. These can be broken arbitrarily for optimizing the dual function. Substituting $\boldsymbol{\mu}^*$ into $L(\boldsymbol{\lambda}, \boldsymbol{\mu})$, and noticing that $\boldsymbol{\mu}^*$, $\boldsymbol{x}^*$, $\boldsymbol{p}^*$ are all functions of $\boldsymbol{\lambda}$, we have

$$L(\boldsymbol{\lambda}) := \sum_j \max_i \mu_{ij}(\lambda_i) + \sum_i \lambda_i P_i.$$

The solution to (9) is given by numerically minimizing $L(\boldsymbol{\lambda})$ over $\boldsymbol{\lambda} \geq \boldsymbol{0}$. For this we use a subgradient-based search and update $\boldsymbol{\lambda}$ by

$$\lambda_i(t+1) = \left[\lambda_i(t) - \kappa(t)\left(P_i - \sum_j p_{ij}^*(t)\right)\right]^+, \forall i \in \mathcal{M}.$$

where p_{ij}^* is given by (10) and x_{ij} are given by (14) and additionally satisfy the feasibility constraint (2) in case of ties. The algorithm will converge when $\kappa(t)$ is chosen appropriately, e.g., [20, Exercise 6.3.2]. Given an optimal $\boldsymbol{\lambda}^*$, by duality, $L(\boldsymbol{\lambda}^*)$ is the optimal objective value to Problem UL.

However, to implement this algorithm, the scheduler must specify the corresponding optimal primal values of $(\boldsymbol{x}^*, \boldsymbol{p}^*)$. Here, as in [4], more care is required. Specifically, when ties occur in (15), it is often needed to split the subchannel among several users (i.e., allowing fractional values of $\boldsymbol{x}^*$). Following

a similar approach as in [3] (for the downlink), the optimal fractional values can be found by solving a linear program whose size increases with the number of users and tones involved in each tie. As discussed below, this number can be quite large in the uplink setting. Moreover, as noted earlier, one is typically interested in an integer allocation in practice. We consider this problem next.

IV. INTEGER SUBCHANNEL ALLOCATION BASED ON OPTIMAL ALGORITHM

We now address the problem:

$$\max_{\substack{(\boldsymbol{x},\boldsymbol{p})\in\mathcal{X}, \\ x_{ij}\in\{0,1\},\forall i,j}} \sum_{i\in\mathcal{M}} w_i \sum_{j\in\mathcal{N}} x_{ij} \log\left(1 + \frac{p_{ij}e_{ij}}{x_{ij}}\right), \quad \text{(UL-Int)}$$

subject to per user power constraints (1). Initially, consider the following heuristic for Problem UL-Int: (*i*) Solve Problem UL as in the previous section; and (*ii*) "break" any ties on all subchannels, i.e., whenever there is a fractional x_{ij}^* value, choose one user in the tie and allocate subchannel j to that user only. Clearly, if there are no ties, this algorithm gives the optimal solution to Problem UL-Int. After all ties are broken, we can then re-optimize the power allocation for each user using a finite-time *water-filling* algorithm as in [4].

In [4] a similar procedure is used for a downlink OFDM scheduling problem. However, there are several major differences between the uplink and downlink settings that make this approach less appealing for the uplink. First, in the downlink case there is a single power constraint $\sum_{i,j} p_{ij} \leq P$ for the base station instead of the per-user constraints in (1). Hence, in the downlink $L(\lambda)$ is a function of only a single dual variable λ, which simplifies the numerical search for the optimizer. In the uplink setting, the convergence of the subgradient search is too slow to be useful for scheduling on a fast time-scale.

Second, even if the optimal $\boldsymbol{\lambda}$ can be found, breaking ties is more difficult than in the downlink case. Scalar subgradients of $L(\lambda)$ in the downlink case can be used to devise simple tie-breaking rules [4], while in the uplink case, the subgradients are vectors, complicating such an approach. Also, the uplink case can be more sensitive to how ties are resolved. For example, if two users, i and l, have the same weights $(w_i = w_l)$ and the same gains on subchannel j $(e_{ij} = e_{lj})$, then allocating subchannel j to either yields the same total weighted rate and the same total power usage in the downlink case. On the other hand, different allocations lead to different individual power consumptions in the uplink case, and thus may lead to totally different solutions.

Finally, the number of ties is typically much larger in the uplink case than in the downlink case. Consider a simple scenario with two users and two subchannels. Each user has the same gain over both subchannels, i.e., $e_{i1} = e_{i2} = e_i$ for $i = 1, 2$, and $P = P_1 = P_2$, where P is the total power constraint in the downlink case. Assume user 2 has a much better subchannel than user 1 so that in the downlink case, the unique optimal solution is to allocate both subchannels to user 2, and there is no tie. However, in the uplink case, it can be shown that at the optimal dual solution, λ_1 and λ_2 will satisfy $\mu_{1j}(\lambda_1) = \mu_{2j}(\lambda_2)$ for $j = 1$ and 2, i.e., there is a tie

in each subchannel and four possible subchannel allocations must be considered to determine how to break the tie. This can be easily extended to M users and N subchannels, with each user having the same gain over all its subchannels, resulting in M^N ties even in this simplistic setting.

V. LOW COMPLEXITY SUBOPTIMAL ALGORITHMS

Due to the issues discussed in the previous section, finding the optimal dual solution to Problem UL and breaking any ties to get an integer allocation is computationally difficult, even for a moderately sized system. Thus, we now present a family of sub-optimal algorithms (SOAs), which try to reduce this complexity using the structure revealed by the optimal algorithm while enforcing an integer-tone allocation and exhibiting good performance.

In the optimal algorithm, given the optimal $\boldsymbol{\lambda}^*$, the optimal subchannel allocation up to any ties is determined by sorting the users on each tone according to the metric $\mu_{ij}(\boldsymbol{\lambda}^*)$ as in (14). Given an optimal subchannel allocation, the optimal power allocation is given by a per-user water-filling allocation as in (10). In each SOA, we use the same two phases but modify them to reduce the complexity. Specifically, we begin with a *subChannel Allocation (CA)* phase which assigns each subchannel to at most one user. Two different implementations of the CA phase are given. In SOA1, instead of using $\mu_{ij}(\boldsymbol{\lambda}^*)$, we consider metrics based on a constant power allocation over all subchannels assigned to a user. In SOA2, we again use a dual based approach, but here we first determine the number of subchannels assigned to each user and then match specific subchannels and users. After the CA phase in both SOAs, we execute the *Power Allocation (PA)* phase in which each user's power is allocated across the assigned subchannels as in the optimal algorithm.

A. CA in SOA1: Progressive Subchannel Allocation based on Metric Sorting

In this family of SOAs, subchannels are assigned sequentially in one pass based on a per user metric for each subchannel. Let $\mathcal{K}_i(n)$ denote the set of subchannels assigned to user i after the nth iteration. Let $g_i(n)$ denote user i's metric during the nth iteration and let $l_i(n)$ be the subchannel index that user i would like to be assigned if he/she is assigned the nth subchannel. The resulting CA algorithm is given in Algorithm 1. Note that all the user metrics are updated after each subchannel is assigned.

We consider several variations of Algorithm 1 which correspond to different choices for steps 4 and 5. The choices for step 4 are:

(4A): Sort the subchannels based on the best channel condition among *all* users. This involves two steps. First, for each subchannel j, find the best channel condition among all users and denote it by $\tilde{\mu}_j := \max_i e_{ij}$. Second, find a subchannel permutation $\{\alpha_j\}_{j\in\mathcal{N}}$ such that $\tilde{\mu}_{\alpha_1} \geq \tilde{\mu}_{\alpha_2} \geq \cdots \geq \tilde{\mu}_{\alpha_N}$, and set $l_i(n) = \alpha_n$ for each user i at the nth iteration. Each $\max$ operation has complexity of $O(M)$, and the sorting operation has a complexity of $O(N \log(N))$. The total complexity is

Algorithm 1 CA Phase for SOA1

1: Initialization: set $n = 0$ and $\mathcal{K}_i(n) = \emptyset$ for each user i.
2: **while** $n < N$ **do**
3:     $n = n + 1$.
4:     Update subchannel index $l_i(n)$ for each user i.
5:     Update metric $g_i(n)$ for each user i.
6:     Find $i^*(n) = \arg\max_i g_i(n)$ (break ties arbitrarily).
7:     Assign the nth subchannel to user $i^*(n)$:

$$\mathcal{K}_i(n) = \begin{cases} \mathcal{K}_i(n-1) \cup \{l_i(n)\}, & \text{if } i = i^*(n); \\ \mathcal{K}_i(n-1), & \text{otherwise.} \end{cases}$$

8: **end while**

$O(NM + N\log N)$. We note that this is a one-time "pre-processing" that needs to done before the CA phase starts. During the subchannel allocation iterations, the users just choose the subchannel index from the sorted list.

(4B): Sort the subchannels based on the channel conditions for each *individual* user. For each user i at the nth iteration, set $l_i(n)$ to be the subchannel index with the largest gain among all unassigned subchannels, i.e., $l_i(n) = \arg\max_{j \in \mathcal{N} \setminus \cup_i \mathcal{K}_i(n-1)} e_{ij}$. This requires M sorts (one per user) that only need to be performed once, since each subchannel assignment does not change a user's ordering of the remaining subchannels. The total complexity of the M sorting operations is $O(MN\log N)$, which is higher than that in *(4A)*.

Let $k_i(n) = |\mathcal{K}_i(n)|$. The choices for Line 5 are:

(5A): Set $g_i(n)$ to be the total increase in user i's utility if assigned subchannel $l_i(n)$, assuming constant power allocation over all assigned subchannels, i.e.,

$$g_i(n) = w_i \Bigg[\sum_{j \in \mathcal{K}_i(n-1) \cup \{l_i(n)\}} \log\left(1 + \frac{P_i e_{ij}}{k_i(n-1)+1}\right) \\ - \sum_{j \in \mathcal{K}_i(n-1)} \log\left(1 + \frac{P_i e_{ij}}{k_i(n-1)}\right) \Bigg]. \quad (16)$$

(5B): Set $g_i(n)$ to be user i's gain from only subchannel $l_i(n)$, assuming constant power allocation, i.e.,

$$g_i(n) = w_i \log\left(1 + \frac{P_i}{k_i(n-1)+1} e_{i,l_i(n)}\right).$$

Compared with $(5A)$, this metric ignores the change in user i's utility due to the decrease in power allocated to any subchannels in $\mathcal{K}_i(n-1)$.

The complexity of either of these choices over N iterations is $O(NM)$, and so the total complexity for the CA phase [4] is $O(NM + N\log N)$ (if $(4A)$ is chosen) or $O(MN\log N)$ (if $(4B)$ is chosen).

B. CA in SOA2: subchannel Number Assignment & subchannel User Matching

As summarized in Algorithm 2, SOA2 implements the CA phase through two steps: subchannel number assignment

[4]We note that SOA1 with $(4B)$ and $(5B)$ is similar to the algorithms proposed in [25]. In Section VI we show that other variations of SOA1 ($(4B)$ and $(5A)$) and SOA2 can achieve better performance with similar or slightly higher complexity.

(CNA) and subchannel user matching (CUM).

Algorithm 2 CA Phase of SOA2

1: CNA step: determine the number of subchannels n_i allocated to each user i such that $\sum_{i \in \mathcal{M}} n_i \leq N$.
2: CUM step: determine the subchannel assignment $x_{ij} \in \{0,1\}$ for all i and j, such that $\sum_{j \in \mathcal{N}} x_{ij} = n_i$.

CNA Step: The CNA step determines the number of subchannels n_i assigned to each user $i \in \mathcal{M}$ based on the approximation that each user sees a flat wide-band fading subchannel. This step does not specify which subchannel is allocated to which user; such a mapping is left to the CUM step. The CNA step is further divided into two stages: a basic assignment stage and an assignment improvement stage.

Stage 1, Basic Assignment: Here, we model each user i as having a normalized SINR $\bar{e}_i = \frac{1}{N}\sum_{j \in \mathcal{N}} e_{ij}$, and then determine the number of subchannels assigned n_i for all i by solving:

$$\max_{\{n_i \geq 0, i \in \mathcal{M}\}} \sum_{i \in \mathcal{M}} w_i n_i \log\left(1 + \frac{P_i}{n_i}\bar{e}_i\right)$$
$$\text{subject to: } \sum_{i \in \mathcal{M}} n_i \leq N. \quad \text{(SOA2-CNA)}$$

It can be shown that Problem SOA2-CNA is a standard concave maximization problem over a convex set with a unique and possibly non-integer solution; we use a dual relaxation method to find this solution. The optimal Lagrange multiplier for the subchannel constraint and any intermediate optimal n_i allocation can be found by a line-search, over ranges $[0, \max_i (w_i \log(1 + P_i\bar{e}_i))]$ and $[0, N]$, respectively. Hence, the worst case complexity of the solving each subproblem is independent of M or N. Since we need to determine the value of n_i for every user, the complexity of the basic assignment step is $O(M)$. If the resultant channel allocations contain non-integer values, we will approximate with an integer solution that satisfies $\sum_{i \in \mathcal{M}} n_i = N$.

Since each user is allocated only a subset of the subchannels, the normalized SINR $\bar{e}_i = \frac{1}{N}\sum_{j \in \mathcal{N}} e_{ij}$ is typically a pessimistic estimate of the averaged subchannel conditions over the allocated subset. This motivates the following assignment improvement stage of CNA.

Stage 2, Assignment Improvement: Here, we iteratively solve the following variation of Problem SOA2-CNA:

$$\max_{\{n_i(t) \geq 0, i \in \mathcal{M}\}} \sum_{i \in \mathcal{M}} w_i n_i(t) \log\left(1 + \frac{P_i}{n_i(t)}\bar{e}_i(t)\right)$$
$$\text{subject to: } \sum_{i \in \mathcal{M}} n_i(t) \leq N, \quad \text{(SOA2-CNA-t)}$$

for $t = 1, 2, \ldots$. During the t-th iteration, $\bar{e}_i(t)$ is a refined estimate of the normalized SINR based on the best $n_i(t-1)$ subchannels of user i ($n_i(0) = N$). The iterative procedure stops when the subchannel allocation converges or the maximum number of iterations allowed is reached. At the end an integer approximation will be performed, if needed.

The complete algorithm for the CNA phase of SOA2 is given in Algorithm 3. In order to perform the assignment

improvement, we need to perform M sorting operations once, with a total complexity $O(MN \log(N))$. Step 4 of each iteration has complexity of $O(M)$ due to solving M subproblems for a fixed dual variable. The maximum number of iterations is fixed and thus is independent of N or M. The integer approximation stage (typically) requires a sorting with the complexity of $O(M \log(M))$. So the total complexity for the CNA phase of SOA2 is $O(MN \log(N) + M \log(M))$.

Algorithm 3 CNA Phase of SOA2

1: Initialization: integer MaxIte> 0, $t = 0$, $n_i(0) = N$ and $n_i(1) = N/2$ for each user i.
2: **while** ($n_i(t+1) \neq n_i(t)$ for some i) & ($t <$MaxIte) **do**
3:    $t = t + 1$.
4:    For each user i, $\overline{e}_i(t) =$ average gain of user i's best $n_i(t-1)$ subchannels.
5:    Solve Problem (SOA2-CNA-t) to determine the optimal $n_i(t)$ for each user i.
6: **end while**
7: let $n_i^* = n_i(t)$ for each user i.

CUM Step: The CNA step determines how many subchannels are to be allocated to each user but not the exact subchannel assignment. This is accomplished in the CUM step by finding a subchannel assignment that maximizes the weighted-sum rate assuming each user employs a flat power allocation, i.e. we solve the problem:

$$\max_{x_{ij} \in \{0,1\}} \sum_{i \in \mathcal{M}} \sum_{j \in \mathcal{N}} x_{ij} w_i \log\left(1 + \frac{P_i}{n_i^*} e_{ij}\right)$$
$$\text{subject to: } \sum_{j \in \mathcal{N}} x_{ij} = n_i^*, \forall i \in \mathcal{M}, \quad \text{(SOA2-CUM)}$$
$$\sum_{i \in \mathcal{M}} x_{ij} = 1, \forall j \in \mathcal{N},$$

where $\boldsymbol{n}^* = (n_i^*, i \in \mathcal{M})$ is the integer subchannel allocation obtained in the CNA step.

Problem SOA2-CUM is an integer *Assignment Problem* whose *optimal* solution can be found by using the *Hungarian Algorithm* [15].[5] To use the Hungarian algorithm here, we need to perform the following "virtual user splitting": For user i, let $r_{ij} = w_i \log\left(1 + \frac{P_i}{n_i^*} e_{ij}\right)$, and let

$$\boldsymbol{r}_i = [r_{i1}, r_{i2}, \cdots, r_{iN}]$$

be user i's achievable rates over all possible subchannels. We can then form an $M \times N$ matrix $\boldsymbol{R} = \left[\boldsymbol{r}_1^T, \boldsymbol{r}_2^T, \cdots, \boldsymbol{r}_M^T\right]^T$.[6] Next, we split each user i into n_i^* virtual users by adding $n_i^* - 1$ copies of the row vector $\boldsymbol{r}_i$ to the matrix $\boldsymbol{R}$ giving a $N \times N$ square matrix. Solving Problem SOA2-CUM is then equivalent to finding a *permutation matrix* $\boldsymbol{C}^* = [c_{ij}]_{N \times N}$

---
[5] A similar idea has been used to solve other OFDM resource allocation problems, e.g., [16], [21].

[6] Here we assume that each user is allocated at least one subchannel. If only $\tilde{M} < M$ users are allocated positive amount of channels, we can replace M by $\tilde{M}$ in the discussions.

TABLE I
WORST CASE COMPUTATIONAL COMPLEXITY OF SUBOPTIMAL
ALGORITHMS

Suboptimal Algorithms			Worst Case Complexity
SOA1	subChannel Allocation	4A 5A	$O(NM + N \log N)$
		4A 5B	$O(NM + N \log N)$
		4B 5A	$O(MN \log N)$
		4B 5B	$O(MN \log N)$
	Power Allocation		$O(MN)$
	Total		$\boldsymbol{O(MN \log N)}$
SOA2	subChannel Allocation	CNA	$O(MN \log N + M \log M)$
		CUM	$O(N^3)$
	Power Allocation		$O(MN)$
	Total		$\boldsymbol{O(N^3 + MN \log N + M \log M)}$

such that

$$\boldsymbol{C}^* = \arg\min_{\boldsymbol{C} \in \mathcal{C}} -\boldsymbol{C} \cdot \boldsymbol{R} := \arg\min_{\boldsymbol{C} \in \mathcal{C}} -\sum_{i=1}^{N} \sum_{j=1}^{N} c_{ij} r_{ij}. \quad (17)$$

This problem can be solved by the standard Hungarian algorithm which has a computational complexity of $O(N^3)$. After obtaining $\boldsymbol{C}^*$, we can calculate the corresponding subchannel allocation $\boldsymbol{x}^*$, e.g., if $c_{kj}^* = 1$ and virtual user k corresponds to the actual user i, then $x_{ij}^* = 1$.

C. Power Allocation (PA) phase with Single-user Water-filling

In this phase each user optimally allocates its power across the subchannel assigned to it in the CA phase. For user i, this can be formulated as the following problem

$$\max_{\boldsymbol{p}_i \in \mathcal{P}_i} \sum_j x_{ij}^* \log(1 + p_{ij} e_{ij}), \quad \text{(PA}_i\text{)}$$

where $\mathcal{P}_i = \left\{\boldsymbol{p}_i \geq 0 : p_{ij} \leq \frac{s_{ij}}{e_{ij}}, \sum_{j \in \mathcal{N}} p_{ij} \leq P_i\right\}$.

The solution to Problem PA$_i$ will be a water-filling type of power allocation, taking into account the per subchannel SINR constraint. Specifically, if $\sum_{j \in \mathcal{N}} x_{ij}^* \frac{s_{ij}}{e_{ij}} \leq P_i$, we let $p_{ij}^* = \frac{s_{ij}}{e_{ij}}$. Otherwise, the optimal power allocation is determined by

$$p_{ij}^* = \min\left(\left(\frac{x_{ij}^*}{\nu_i} - \frac{1}{e_{ij}}\right)^+, \frac{s_{ij}}{e_{ij}}\right),$$

where the constant ν_i is chosen so that $\sum_{j \in \mathcal{N}} p_{ij}^* = P_i$. Note that it is possible that some subchannel is allocated to user i but gets no power due to its (relatively) poor subchannel gain. The optimal value of ν_i can be found through a simple line search. However, it is possible to devise a finite-time algorithm (with a maximum of $2n_i^*$ steps) to calculate the exact value of v_i as in [3], [4] with the difference here being that this procedure needs to be executed for every user who is allocated a subchannel. Since $\sum n_i^* = N$, it follows that the total worst case computational complexity for this approach is $O(MN)$.

D. Complexity Summary of Suboptimal Algorithms

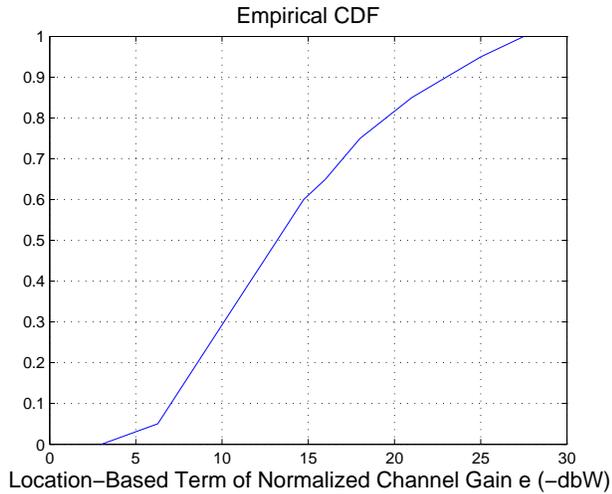The worst case computational complexities of SOA1 and SOA2 are summarized in Table I.

Empirical CDF

Fig. 1. Empirical CDF of the location-based term in users' normalized channel condition e_{ij}'s for all user i and subchannel j
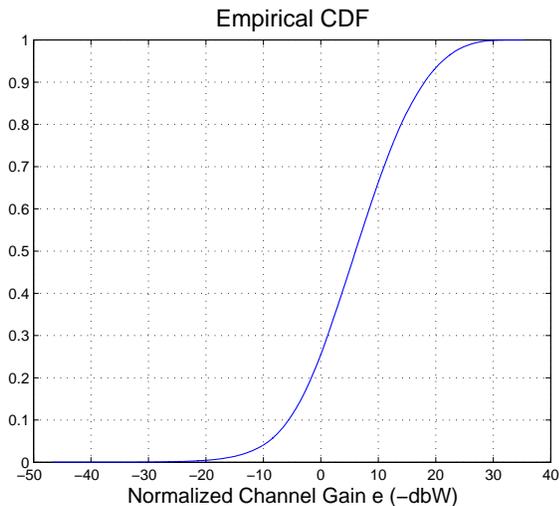
Empirical CDF

Fig. 2. Empirical CDF of users' normalized channel condition e_{ij}'s for all user i and subchannel j

VI. SIMULATION RESULTS

We report simulation results for the following four algorithms: 1) (Integer-Dual) integer subchannel allocation (with tie breaking) based on optimal algorithm as in Section IV and power control as in Section V-C, inspecting up to 128 ways of breaking the ties with an integer allocation and selecting the allocation among these with the largest weighted sum rate (before reallocating the power); 2) (SOA1) all four versions of subchannel allocation procedure as in Section V-A, and power control as in Section V-C; 3) (SOA2) subchannel allocation as in Section V-B (with up to 10 iterations) and power control as in Section V-C; and 4) (Base-line) each subchannel j is allocated to the user i with the highest e_{ij}, without considering the weights w_i's and the power constraints, while each user's power is still allocated as in Section V-C.

We consider a system bandwidth of 5MHz consisting of 512 OFDM tones, grouped into 64 subchannels (8 adjacent tones

per subchannel, i.e., corresponding to the "Band AMC mode" of 802.16 d/e.). Each user's subchannel gains are the product of a constant location-based term, picked using an empirically obtained distribution, and a fast fading term, generated using a block-fading model and a standard mobile delay-spread model with a delay spread of 10μsec. The fast-fading component for each multi-path component is held fixed for 2msec and an independent value is generated for the next block, which corresponds to a 250Hz Doppler. The empirical cumulative distribution function (CDF) of the location-based term of users' location-based term of the e_{ij}'s is given in Fig. 1, and the empirical CDF of the net e_{ij}'s (i.e., including the fast-fading component) is given in Fig. 2. Since the units of e_{ij} are in 1/watt, the x-axis in both figures is measured in -dBW. The symbol duration is 100μsec with a cyclic prefix of 10μsec, which roughly corresponds to 20 OFDM symbols per fading block (i.e., 2msec). This is one of the allowed configurations in the IEEE 802.16 standards [14]. Resource allocation (solving Problem UL) is done once per fading block. All the results are averaged over the last 1000 fading blocks. For the sake of illustration and to keep things simple, we assume that all users are infinitely back-logged, $d_i = 0$, and have the same isoelastic utility function $u_i(W_{i,t}) = (W_{i,t})^\alpha/\alpha$, where $W_{i,t}$ is the long term average throughput of user i up to time t. All users also have the same maximum power constraint of $P_i = 2$W. We calculate the achievable rate of user i on subchannel j as

$$r_{ij} = Bx_{ij}\log\left(1 + \frac{p_{ij}e_{ij}}{x_{ij}}\right),$$

where B is the subchannel bandwidth.

In all tables, the "Utility" column denotes $\sum_i U_i(W_{i,t})$ (in scientific notations when appropriate), where t is the end of simulation time. The "Log U" column denotes the logarithmic utility function, $\sum_i \log(W_{i,t})$, which provides an alternate characterization of fairness among users. The "Rate" column denotes the total rate achieved by users in Mbps. The "User #" column denotes the average number of users who receive positive rates within one scheduling interval. The "Opt. Ratio" for the Integer-Dual algorithm denotes the ratio between the objective value of Problem UL-Int achieved with the algorithm and the maximum objective value of Problem UL averaged over all scheduling instances. This gives us an idea of how good the performance of the Integer-Dual algorithm is. Furthermore, we emphasize that this value is based on the average performance of solving problem UL in each scheduling instance; this does not translate directly into a bound on the *average utility* under a optimal scheduling rule and under the given policy as the trajectory of the scheduling weights will be different under the two policies. Since other suboptimal algorithms do not find the optimal dual value of Problem UL at each scheduling instance, the "Optimality Ratio" does not apply to them.

Table II shows results for all the algorithms (summed over all users) when scheduling decisions are made every 20 OFDM symbols (i.e., a fading block of 2msec). The utility parameter is $\alpha = 0.5$. In Table II, SOA1 (with 4B & 5A) and SOA2 achieve the best performance in terms of total utility. Their performance is even better than the Integer-Dual approach,

TABLE II
ALGORITHM PERFORMANCE FOR SCHEDULING EVERY 20 OFDM
SYMBOLS, $\alpha = 0.5$

Algorithms		Utility	Log U	Rate	User #	Opt. Ratio
Integer-Dual		53922	514.0	21.56	37.5	0.9412
SOA 1	4A 5A	52494	510.7	22.86	34.6	N/A
	4A 5B	51697	509.2	20.22	28.1	N/A
	4B 5A	54165	513.3	22.25	35.0	N/A
	4B 5B	53156	511.4	21.43	28.6	N/A
SOA 2		54316	513.6	22.33	35.1	N/A
Base Line		21406	-1960.5	16.13	2.66	N/A

TABLE III
ALGORITHM PERFORMANCE FOR SCHEDULING EVERY 20 OFDM
SYMBOLS, $\alpha = 0$

Algorithm		Utility	Log U	Rate	User #	Opt. Ratio
Integer-Dual		515	515	19.80	39.0	0.9715
SOA 1	4A 5A	511	511	18.47	36.9	N/A
	4A 5B	509	509	21.56	28.6	N/A
	4B 5A	515	515	20.04	37.4	N/A
	4B 5B	512	512	17.80	29.3	N/A
SOA 2		515	515	19.91	37.4	N/A
Base Line		-1961	-1961	16.13	2.66	N/A

which was obtained based on the optimal value of the relaxed problem. This is likely because only 128 ways to break ties are considered, which is typically not sufficient. Since the Integer-Dual algorithm achieves an optimality ratio of 0.9412, this suggests that SOA1 and SOA2 achieve very close to optimal performance as well. The base-line algorithm always has poor performance.

Tables III and IV give the results of the algorithms when the utility parameter α is equal to 0 (proportional fair allocation) and 1 (maximum rate allocation), respectively. It is clear that there is a trade-off between fairness (measured by Log U) and efficiency (measured by total rate). The value of $\alpha = 0$ gives a throughput allocation that is the fairest and the least efficient, while $\alpha = 1$ is the most efficient and least fair. Once again we note that all of the heuristics have good performance with SOA1 (with 4B & 5A) and SOA2 achieving the best performances in terms of total utility.

In each case simulated, all of the SOA's have good performance with SOA1 (with 4B & 5A) and SOA2 consistently achieving the best performance in terms of total utility. From

TABLE IV
ALGORITHM PERFORMANCE FOR SCHEDULING EVERY 20 OFDM
SYMBOLS, $\alpha = 1$

Algorithm		Utility	Log U	Rate	User #	Opt. Ratio
Integer-Dual		23.24e6	472.56	23.37	20.43	0.82541
SOA 1	4A 5A	23.19e6	448.99	22.28	22.6	N/A
	4A 5B	23.11e6	-136.03	23.20	15.6	N/A
	4B 5A	24.31e6	444.02	24.42	32.7	N/A
	4B 5B	23.95e6	-195.10	24.05	15.6	N/A
SOA 2		24.46e6	372.95	24.57	21.8	N/A
Base Line		16.08e6	-1961	16.13	2.66	N/A

the analysis in Section V-D, we note that these have slightly higher complexity than some of the other SOA's. Hence if lower complexity is desired, this can be provided with only a slight loss in performance. We also note that in each of the algorithms (except the base line one) a large number of users are scheduled in each time-slot, which may lead to a high signaling overhead. This can be addressed by adding a penalty term to our objective which increases with the number of users scheduled.

VII. CONCLUSIONS

We presented an optimization-based formulation for scheduling and resource allocation in the uplink OFDM access network. Compared to the downlink, we showed that the uplink was computationally more challenging due in part to the per-user power constraints. A (high complexity) optimal algorithm was given as well as a family of low complexity heuristics. The heuristics were shown to have good performance via simulations for a range of different user utilities and scheduling time-scales. Two algorithms from this family consistently achieved the best performance, but had a slightly higher complexity than some of the other algorithms, enabling complexity to be traded off with performance within this family of algorithms.

REFERENCES

[1] R. Agrawal, A. Bedekar, R. La, V. Subramanian, "A Class and Channel-Condition based Weighted Proportionally Fair Scheduler," in *Proc. ITC*, Salvador, Brazil, Dec. 2001.

[2] R. Agrawal and V. Subramanian, "Optimality of Certain Channel Aware Scheduling Policies," in *Proc. of Allerton Conference*, Oct. 2002.

[3] R. Agrawal, V. Subramanian and R. Berry, "Joint Scheduling and Resource Allocation in CDMA Systems," in *Proc. of WiOpt '04*, Cambridge, UK, March 2004.

[4] J. Huang, V. G. Subramanian, R. Agrawal and R. Berry, "Downlink scheduling and resource allocation for OFDM systems," *IEEE Trans. on Wireless*, to appear.

[5] L. Tassiulas and A. Ephremides, "Dynamic server allocation to parallel queue with randomly varying connectivity", *IEEE Trans. IT*, vol. 39, pp. 466-478, March 1993.

[6] R. Leelahakriengkrai and R. Agrawal, "Scheduling in Multimedia Wireless Networks," in *Proc. ITC*, Salvador, Brazil, pp. 556–564, Dec. 2001.

[7] M. Andrews, K. K. Kumaran, K. Ramanan, A. L. Stolyar, R. Vijayakumar, and P. Whiting, "Scheduling in a queueing system with asynchronously varying service rates," *Probability in Engineering and Informational Sciences*, vol. 18, No. 2, pp. 191–217, 2004.

[8] A. Jalali, R. Padovani, R. Pankaj, "Data throughput of CDMA-HDR a high efficiency - high data rate personal communication wireless system.," in *Proc. IEEE VTC*, Spring 2000, vol. 3, pp. 1854–1858.

[9] A. L. Stolyar, "Maximizing Queueing Network Utility subject to Stability: Greedy Primal-Dual Algorithm," *Queueing Systems*, vol. 50, pp. 401–457, 2005.

[10] A. L. Stolyar, "On the asymptotic optimality of the gradient scheduling algorithm for multiuser throughput allocation," *Operations Research*, vol. 53, No. 1, pp. 12–25, 2005.

[11] M. Andrews, L. Qian, and A. Stolyar, "Optimal utility based multi-user throughput allocation subject to throughput constraints," in *Proc. IEEE INFOCOM*, March 2005, vol. 4, pp. 2415–2424.

[12] H. Kushner and P. Whiting, "Asymptotic properties of proportional-fair sharing algorithms," in *Proc. Allerton Conference*, Oct. 2002.

[13] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE/ACM Trans. Networking*, vol. 8, Issue 5, pp. 556–567, Oct. 2000.

[14] "IEEE 802.16e-2005 and IEEE Std 802.16-2004/Cor1-2005," 2005, Available: http://www.ieee802.org/16/.

[15] E. D. Nering and A. W. Tucker, *Linear Programs and Related Problems*. Academic Press Inc., 1993.

[16] H. Yin and H. Liu, "An efficient multiuser loading algorithm for OFDM-based broadband wireless systems," in *Proc. IEEE Globecom*, vol. 1, pp. 103–107, Dec. 2000.

[17] T. M. Cover and J. A. Thomas, "Elements of information theory," Second edition, Wiley-Interscience, Hoboken, NJ, 2006.

[18] S. Shenker, "Fundamental design issues for the future Internet," *IEEE JSAC*, vol. 13, no. 7, pp. 1176–1188, 1995.

[19] FLASH-OFDM Technology, http://www.qualcomm.com/technology/flash-ofdm

[20] D. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, Massachusetts: Athena Scientific, 1999.

[21] Z. Han, Z. Ji, and K. Liu, "Fair Multiuser Channel Allocation for OFDMA Networks Using Nash Bargaining Solutions and Coalitions," *IEEE Trans. Comm.*, vol. 53, no. 8, pp. 1366–1376, 2005.

[22] S. Pfletschinger, G. Muenz, and J. Speidel, "Efficient subcarrier allocation for multiple access in OFDM systems," in *Proc. International OFDM-Workshop*, 2002.

[23] Y. Ma, "Constrained Rate-Maximization Scheduling for Uplink OFDMA," in *Proc. IEEE MILCOM*, pp. 1–7, Oct. 2007.

[24] B. Da and C. Ko, "Dynamic subcarrier sharing algorithms for uplink OFDMA resource allocation," in *Proc. 6th ICICS*, pp. 1–5, 2007.

[25] K. Kim, Y. Han, and S. Kim, "Joint subcarrier and power allocation in uplink OFDMA systems," *IEEE Comms. Letters*, vol. 9, no. 6, pp. 526–528, 2005.

[26] C. Ng and C. Sung, "Low complexity subcarrier and power allocation for utility maximization in uplink OFDMA systems," *IEEE Trans. Wireless Comm.*, vol. 7, no. 5 Part 1, pp. 1667–1675, 2008.

[27] L. Gao and S. Cui, "Efficient subcarrier, power, and rate allocation with fairness consideration for OFDMA uplink," *IEEE Trans. Wireless Comm.*, vol. 7, no. 5 Part 1, pp. 1507–1511, 2008.

[28] K. Kwon, Y. Han, and S. Kim, "Efficient Subcarrier and Power Allocation Algorithm in OFDMA Uplink System," *IEICE Trans. Comm.*, vol. 90, no. 2, pp. 368–371, 2007.

[29] "3GPP Long Term Evolution," 2008, Available: http://www.3gpp.org/Highlights/LTE/LTE.htm .

[30] "3GPP2 Ultra Mobile Broadband," 2007, Available: http://www.3gpp2.org/ .

[31] W. Yu, R. Lui, and R. Cendrillon, "Dual optimization methods for multiuser orthogonal frequency division multiplex systems," in *Proc. IEEE Globecom*, vol. 1, 2004, pp. 225–229.

Vijay G. Subramanian (M'01) received his Ph.D. degree in Electrical Engg. from the University of Illinois at Urbana-Champaign, in 1999. From 1999 to 2006, he was with the Networks Business, Motorola, Arlington Heights, IL, USA. Since May 2006 he is a Research Fellow at the Hamilton Institute, NUIM, Ireland. His research interests include information theory, communication networks, queueing theory, and applied probability and stochastic processes.

Rajeev Agrawal is a Fellow of the Technical Staff at Motorola where his responsibilities include the architecture, design and optimization of Motorola's next generation wireless systems. Prior to joining Motorola in 1999, Rajeev was Professor of Electrical and Computer Engg. and Computer Science departments at the University of Wisconsin - Madison. He also spent a sabbatical year at IBM TJ Watson Research, British Telecom Labs, and INRIA-Sophia Antipolis. Rajeev received his M.S. (1987) and Ph.D. (1988) degrees in Electrical Engg.-systems from the University of Michigan, Ann Arbor and his B.Tech. (1985) degree in Electrical Engg. from the Indian Institute of Technology, Kanpur.

Jianwei Huang (S'01-M'06) is an Assistant Professor in Information Engg. Department at the Chinese University of Hong Kong. He received the M.S. and Ph.D. degrees in Electrical and Computer Engg. from Northwestern University in 2003 and 2005, respectively. From 2005 to 2007, He worked as a Postdoctoral Research Associate in the Department of Electrical Engg. at Princeton University. His main research interests lie in the area of modeling and performance analysis of communication networks, including cognitive radio networks, OFDM and CDMA systems, wireless medium access control, multimedia communications, network economics, and applications of optimization theory and game theory. Dr. Huang is an Associate Editor of Journal of Computer & Electrical Engg., the Lead Guest Editor of the IEEE Journal of Selected Areas in Communications special issue on "Game Theory in Communication Systems", the Lead Guest Editor of the Journal of Advances in Multimedia special issue on "Collaboration and Optimization in Multimedia Communications", and a Guest Editor of the Journal of Advances in Multimedia special issue on "Cross-layer Optimized Wireless Multimedia Communications", the Technical Program Committee Co-Chair of the International Conference on Game Theory for Networks (GameNets'09).

Randall A. Berry (S'93-M'00) received the M.S. and PhD degrees in Electrical Engg. and Computer Science from the Massachusetts Institute of Technology in 1996 and 2000, respectively. In September 2000, he joined the faculty of Northwestern University, where he is currently an Associate Professor of Electrical Engg. and Computer Science. Dr. Berry is the recipient of a 2003 NSF CAREER award. He is currently serving on the editorial boards of the IEEE Transactions on Information Theory and the IEEE Transactions on Wireless Communications.