Downlink Mobile OFDMA Resource Allocation With Minimum User Rate Requests

Stelios Stefanatos[1], Christos Papathanasiou[2], Nikos Dimitriou[1]
[1]Institute of Accelerating Systems & Applications, University of Athens, Greece
[2] Computer and Communications Engineering, University of Thessaly, Greece

***Abstract*— The problem of resource allocation (RA) in a downlink OFDMA system with minimum user rate requests is examined under the realistic scenario of partial (imperfect) channel state information (CSI) at the base station. The challenge in this setting is the non-zero probability of outage events which may lead to significant performance degradation if algorithms assuming perfect CSI are utilized. In this paper the statistical description of the partial CSI is incorporated for deriving an (optimal) RA algorithm as the solution of an appropriate constrained optimization problem. Simplifications of the algorithm are utilized for significant complexity reduction with the additional property of allowing the system to adapt on–the–fly to time–varying minimum user rate requests.**

I. Introduction

OFDMA is currently adopted by wireless standards as a downlink modulation scheme that allows multiple parallel orthogonal flat–fading transmissions to system users [1]. The recent trend in the field of resource allocation (RA) is to exploit channel state information (CSI) available at the base station (BS) in order to adapt to the wireless channel fades and maximize the system's sum rate [2], [7], [8]. However, sum rate maximization results in a highly unfair RA since users with poor channel conditions are penalized, which may not be acceptable. As a remedy to this effect, RA has been posed as a (minimum rate) constrained optimization problem in, for example, [5], [10], and algorithms known from (constrained) optimization theory are utilized for the solution.

The aforementioned problem formulations and corresponding solutions assume *perfect* CSI which is unrealistic under (significant) user mobility, mainly because the available CSI at the BS becomes quickly outdated. Treating CSI as perfect in such a scenario can deteriorate system performance significantly, due to the non-zero probability of outage events (true channel state cannot support assigned rate). One simple remedy to this effect is to transmit at a reduced rate than the one dictated by the nominal CSI. However, this entails the risk of significant underutilization of the channel.

In this paper the problem of constrained RA under imperfect (partial) CSI at the BS is considered. Specifically, a statistical description (model) of the partial CSI is utilized along with the notion of goodput (instead of rate) for RA. An appropriate constrained optimization problem formulation is constructed and the optimal solution as well as simplified versions thereof are presented. The resulting simplified algorithm has the attractive property of allowing for system adaptation to time-varying minimum user requirements.

The paper is organized as follows: Section II describes the system and channel model. In Section III a constrained, goodput maximization problem is formulated and the optimal solution (RA algorithm) is derived. Section IV discusses computationally efficient approximations of the algorithm that also allow for simple system adaptation to varying user requests. Section V demonstrates simulated performance of the proposed scheme and Section VI concludes the paper.

This work has been supported by the WiMAGIC (FP7–ICT–215167) and NEWCOM++ (contract n. 216715) projects, partly funded by the European Union.

II. System and Channel Model

A downlink OFDMA system is considered, that employs N subcarriers for data transmission to the K system mobile users. All users are assumed to have full buffers and they request for a certain minimum transmission rate. Typically, this rate guarantee is required by delay-sensitive applications, whereas best effort services may not impose any rate constraints [5], [10]. The purpose of the RA algorithm is to share the system's resources among the various users in an "optimal" manner that takes into account the users' service requests as well as every information available on their respective channel state.

The (complex) channel gain corresponding to subcarrier $n \in \{1, \ldots, N\}$ of user $k \in \{1, \ldots, K\}$ is assumed to be quasi-static over the duration of one RA interval (frame) and is denoted by $h_{k,n}^{(m)}$, where m is the RA frame index. The channel gain may change between consecutive frames due to user mobility (block fading model). In addition to the fading channel effects, additive white Gaussian noise (AWGN) is also present at the receivers. In an effort to reduce the overhead associated with the CSI feedback process, CSI is fed back by the users' terminals in some periodic fashion, which, however, may not be fast enough to assume perfect CSI at the BS without performance degradation (feedback delay effect).

RA is performed on a per-subcarrier basis. Specifically, each user may be assigned one or more subcarriers for the duration of the RA frame that spans a number of OFDMA symbols. For simplicity reasons, each subcarrier is assigned to a single user (no superposition coding) and is encoded separately, i.e., joint coding across subcarriers is not considered. Each codeword length spans one RA frame duration which is assumed long enough to allow incorporation of capacity achieving coding schemes.

It is assumed that the BS has an estimate $\{\hat{h}_{k,n}^{(m)}\}_{n=1}^{N}$ of the channel gains of each user which are, in general, not equal to $\{h_{k,n}^{(m)}\}_{n=1}^{N}$. The reliability of the kth user CSI is described in statistical terms by the corresponding conditional probability density function (pdf) $p(\{h_{k,n}^{(m)}\}_{n=1}^{N}|\{\hat{h}_{k,n}^{(m)}\}_{n=1}^{N},\{\hat{h}_{k,n}^{(m-1)}\}_{n=1}^{N},\ldots)$, which implies that the BS exploits information provided by the frequency and/or time correlations of the OFDM channel (independent channels are assumed among users). However, such an exact description is complicated and leads to impractical algorithms. Therefore, a simplified (suboptimal) version of this pdf is employed, namely $p(\{h_{k,n}^{(m)}\}_{n=1}^{N}|\{\hat{h}_{k,n}^{(m)}\}_{n=1}^{N},\{\hat{h}_{k,n}^{(m-1)}\}_{n=1}^{N},\ldots) \approx \prod_n p(h_{k,n}^{(m)}|\hat{h}_{k,n}^{(m)})$, i.e., channel correlations are ignored. Note that ignoring frequency correlation does not affect system performance since RA is restricted on a per-subcarrier level. In the following, superscript $(\cdot)^{(m)}$ will be dropped from notation for simplicity, since all quantities involved in the RA process correspond to the current (mth) frame.

Under perfect CSI, $p(h_{k,n}|\hat{h}_{k,n}) = \delta(h_{k,n} - \hat{h}_{k,n})$, where $\delta(\cdot)$ is the Dirac delta, whereas in the case of completely unknown CSI, $p(h_{k,n}|\hat{h}_{k,n}) = p(h_{k,n})$. When none of these two extremes are valid (realistic under periodic CSI feedback) the level of concentration of $p(h_{k,n}|\hat{h}_{k,n})$ on the nominal channel gain $\hat{h}_{k,n}$ reflects the accuracy of the partial CSI which will be exploited in next section for RA purposes. A typical model for $p(h_{k,n}|\hat{h}_{k,n})$ is the Gaussian [9], i.e.,

$$p(h_{k,n}|\hat{h}_{k,n}) = \mathcal{CN}\left(\hat{h}_{k,n}, \sigma^2_{\hat{h},k,n}\right), \tag{1}$$

where $\mathcal{CN}(m,s)$ denotes the circular symmetric complex Gaussian distribution of mean m and variance s, and $\sigma^2_{\hat{h},k,n}$ denotes the channel estimate variance. The value of $\sigma^2_{\hat{h},k,n}$ depends on channel estimation process accuracy and feedback/processing delay and can be either estimated or set to some fixed (e.g., worst case) value.

III. Constrained, Goodput-Maximizing Resource Allocation Problem

Assuming, without loss of generality, that the AWGN sample variance at every receiver is unity, the maximum bandwidth–normalized information rate (capacity) that can be reliably transmitted on a subcarrier experiencing channel gain $h_{k,n}$ is

$$C_{k,n} = \log_2\left(1 + |h_{k,n}|^2 p_{k,n}\right), \tag{2}$$

in bits/sec/Hz, where $p_{k,n}$ is the power allocated by the BS for transmission on this subcarrier [typically adjusted to the corresponding (perfect or partial) CSI]. Conditioned on the partial CSI, $C_{k,n}$ is a random variable, and assuming that the BS transmits at a rate $r_{k,n}(\hat{h}_{k,n}) \triangleq \hat{r}_{k,n}$ using a power $p_{k,n}(\hat{h}_{k,n}) \triangleq \hat{p}_{k,n}$,[1] there is a nontrivial outage probability $P_{\text{out}}(\hat{r}_{k,n},\hat{p}_{k,n}) \triangleq \Pr\{\hat{r}_{k,n} > C_{k,n}(h_{k,n},\hat{p}_{k,n})|\hat{h}_{k,n}\}$ that the true channel state cannot support the assigned rate. Under this observation, a natural measure for resource allocation purposes is the average *successfully* transmitted rate over a subcarrier (compactly referred to as goodput), defined as [6]

$$G_{k,n}(\hat{r}_{k,n},\hat{p}_{k,n}) \triangleq \hat{r}_{k,n}\bar{P}_{\text{out}}(\hat{r}_{k,n},\hat{p}_{k,n}). \tag{3}$$

in bits/sec/Hz, where $\bar{P}_{\text{out}}(\hat{r}_{k,n},\hat{p}_{k,n}) \triangleq 1 - P_{\text{out}}(\hat{r}_{k,n},\hat{p}_{k,n})$ is the probability of successful transmission given the available partial CSI. Clearly, one reasonable RA goal in a point–to–point communication setting is to transmit at a rate that maximizes goodput.

For the multiuser setting associated with the downlink OFDMA system considered in this paper, it is natural to employ the *weighted sum goodput* computed by summing every user's weighted individual goodput over all subcarriers, as a measure of the system's welfare. The RA performed by the BS should try to (optimally) schedule resources (namely, subcarriers, power and rate) to the various users of the system so as the weighted sum goodput is maximized while at the same time satisfy constraints related to the consumed power and user rate requests. Note that, based on the previous discussion, it is natural to consider individual user requests in terms of goodput rather than rates. The RA problem can be stated mathematically as follows:

$$\begin{cases} \max_{\{\hat{r}_{k,n}\},\{\hat{p}_{k,n}\},\{\hat{a}_{k,n}\}} G_{\text{tot}}, \\ \quad \text{where } G_{\text{tot}} \triangleq \sum_{k=1}^{K} w_k \sum_{n=1}^{N} \hat{a}_{k,n} G_{k,n}(\hat{r}_{k,n},\hat{p}_{k,n}) \\ \text{subject to:} \\ \text{C1. } \sum_{k=1}^{K}\sum_{n=1}^{N} \hat{a}_{k,n}\hat{p}_{k,n} \le \mathcal{P}; \\ \text{C2. } \sum_{n=1}^{N} \hat{a}_{k,n} G_{k,n}(\hat{r}_{k,n},\hat{p}_{k,n}) \ge \mathcal{G}_k, \ \forall k; \\ \text{C3. } \sum_{k=1}^{K} \hat{a}_{k,n} \le 1, \ \forall n; \\ \text{C4. } \hat{p}_{k,n} \ge 0, \ \forall k,n; \\ \text{C5. } \hat{r}_{k,n} \ge 0, \ \forall k,n; \\ \text{C6. } \hat{a}_{k,n} \ge 0, \ \forall k,n; \end{cases} \tag{4}$$

In (4), $\{w_k\}$ are the user weights selected in order to provide (soft) priorities among them, $\mathcal{G}_k$ denotes the minimum goodput requested by the kth user (which is obtained by some appropriate transformation of the minimum rate request), whereas $\mathcal{P}$ is the maximum available power at the BS for downlink transmission. Also note that in the formulation of the optimization problem, the *occupation* index of the kth user on the nth subcarrier $\hat{a}_{k,n} \in [0,1]$ has been introduced, i.e., users are allowed to *time–share* a specific subcarrier during a RA frame. Although such a sharing may be difficult to implement in practice, this formulation allows to show explicitly that the optimal subcarrier allocation is indeed a greedy one (i.e., each subcarrier is occupied by a single/winner user).

Assuming that problem (4) is feasible and introducing the *non-negative* Lagrange multipliers λ_1, $\{\lambda_{2,k}\}$, $\{\lambda_{3,n}\}$, $\{\lambda_{4,k,n}\}$, $\{\lambda_{5,k,n}\}$, $\{\lambda_{6,k,n}\}$ associated with constraints C1–C6, respectively, the following Karush-Kuhn-Tucker (KKT) conditions (in addition to the problem constraints) should be

[1] In the following, notation $\hat{x}$ indicates that the value of x depends on the current partial CSI.

satisfied at the optimal solution [3] (superscript $(\cdot)^*$ denotes optimal values)

$$\hat{a}^*_{k,n}\frac{\partial\phi_{k,n}(\hat{r}^*_{k,n},\hat{p}_{k,n},\lambda^*_1,\lambda^*_{2,k})}{\partial\hat{p}_{k,n}}\Big|_{\hat{p}_{k,n}=\hat{p}^*_{k,n}}+\lambda^*_{4,k,n}=0,\forall k,n,\quad(5)$$

$$\hat{a}^*_{k,n}\frac{\partial\phi_{k,n}(\hat{r}_{k,n},\hat{p}^*_{k,n},\lambda^*_1,\lambda^*_{2,k})}{\partial\hat{r}_{k,n}}\Big|_{\hat{r}_{k,n}=\hat{r}^*_{k,n}}+\lambda^*_{5,k,n}=0,\forall k,n,\quad(6)$$

$$\phi_{k,n}(\hat{r}^*_{k,n},\hat{p}^*_{k,n},\lambda^*_1,\lambda^*_{2,k})-\lambda^*_{3,n}+\lambda^*_{6,k,n}=0,\ \ \forall k,n,\quad(7)$$

$$\lambda^*_1\left(\mathcal{P}-\sum_{k=1}^K\sum_{n=1}^N\hat{a}_{k,n}\hat{p}_{k,n}\right)=0,\quad(8)$$

$$\lambda^*_{2,k}\left(\sum_{n=1}^N\hat{a}_{k,n}G_{k,n}(\hat{r}_{k,n},\hat{p}_{k,n})-\mathcal{G}_k\right)=0,\ \ \forall k,\quad(9)$$

$$\lambda^*_{3,n}\left(1-\sum_{k=1}^K\hat{a}_{k,n}\right)=0,\ \ \forall n,\quad(10)$$

$$\lambda^*_{4,k,n}\hat{p}^*_{k,n}=0,\ \lambda^*_{5,k,n}\hat{r}^*_{k,n}=0,\ \lambda^*_{6,k,n}\hat{a}^*_{k,n}=0,\forall k,n\quad(11)$$

where $\phi_{k,n}(\hat{r}^*_{k,n},\hat{p}^*_{k,n},\lambda^*_1,\lambda^*_{2,k})\triangleq(w_k+\lambda^*_{2,k})G_{k,n}(\hat{r}^*_{k,n},\hat{p}^*_{k,n})-\lambda^*_1\hat{p}^*_{k,n}$ is the link quality indicator [7] associated with the nth subcarrier of the kth user. From the nonnegativity of $\lambda^*_{6,k,n}$ and (8) it follows that

$$\phi_{k,n}(\hat{r}^*_{k,n},\hat{p}^*_{k,n},\lambda^*_1,\lambda^*_{2,k})\le\lambda^*_{3,n},\ \forall k,n,\quad(12)$$

with equality when $\lambda^*_{6,k,n}=0$. If $\hat{a}^*_{k_0,n}>0$ (i.e., subcarrier n is employed for user k_0), it follows from the third complementarity condition of (11) that $\phi_{k_0,n}(\hat{r}^*_{k_0,n},\hat{p}^*_{k_0,n},\lambda^*_1,\lambda^*_{2,k_0})$ achieves the upper bound of (12). Under mild conditions (continuous $p(h_{k,n}|\hat{h}_{k,n})$; typical for wireless channel models) it is almost sure (a.s.) that different users will have different link quality indicators for the same subcarrier [7], which implies that $\phi_{k,n}(\hat{r}^*_{k,n},\hat{p}^*_{k,n},\lambda^*_1,\lambda^*_{2,k})<\lambda^*_{3,n},\forall k\neq k_0$, and, in turn, $\lambda^*_{6,k,n}>0,\forall k\neq k_0$. Employing the third complementarity condition of (11) once more, it follows that $\hat{a}^*_{k,n}=0,k\neq k_0$. In summary, the following proposition has been proved:

Proposition: The optimal subcarrier allocation for problem (4) is a greedy one, i.e., subcarrier n is exclusively reserved for transmission to the user with the largest corresponding link quality indicator $\phi_{k,n}(\hat{r}^*_{k,n},\hat{p}^*_{k,n},\lambda^*_1,\lambda^*_{2,k})$.

Based on the above proposition the optimal RA consists of the following steps:

1) Given the optimal λ^*_1 and $\{\lambda^*_{2,k}\}$ perform a two-dimensional maximization of $\phi_{k,n}(\hat{r}_{k,n},\hat{p}_{k,n},\lambda^*_1,\lambda^*_{2,k})$ over $\hat{r}_{k,n}\in\mathbb{R}^\dagger$ and $\hat{p}_{k,n}\in\mathbb{R}^\dagger$, $\forall k,n$, where $\mathbb{R}^\dagger$ denotes the non-negative orthant.[2]
2) Associate each subcarrier n with its corresponding winner user $k^*_n=\arg\max_k\phi_{k,n}(\hat{r}^*_{k,n},\hat{p}^*_{k,n},\lambda^*_1,\lambda^*_{2,k})$ transmitting at rate $\hat{r}^*_{k^*_n,n}$ with power $\hat{p}^*_{k^*_n,n}$.

Maximization of $\phi_{k,n}$ requires evaluation of $P_{\text{out}}(\hat{r}_{k,n},\hat{p}_{k,n})$. In [9], the partial–CSI–conditioned pdf of the capacity $p\left(C_{k,n}|\hat{h}_{k,n};\hat{p}_{k,n}\right)$, for a given power assignment $\hat{p}_{k,n}$, is derived based on the channel uncertainty model of (1), which can, in principle, be employed to evaluate P_{out} by numerical integration (c.f. Sec. IV).

Since the optimal λ^*_1 and $\{\lambda^*_{2,k}\}$ are not known in advance, one has to resort on numerical (iterative) procedures. For simplicity, the (non-negative) values $\lambda^*_1,\{\lambda^*_{2,k}\}$ that minimize the dual function of problem (4) may be employed since they can be found by a line search over the unconstrained dual-objective function [3]. The advantage of this approach is that the dual–minimization problem is convex and a unique solution can always be found by a numerical line search algorithm. One common approach is the method of subgradients [5], [7], [10] in which optimal values can be easily shown in this case to be iteratively updated (starting from some initial value) as

$$\lambda_1^{*(i)}=\lambda_1^{*(i-1)}+\mu_1\left(\hat{\mathcal{P}}^{(i-1)}-\mathcal{P}\right)^\dagger,\quad(13)$$

$$\lambda_{2,k}^{*(i)}=\lambda_{2,k}^{*(i-1)}+\mu_{2,k}\left(\mathcal{G}_k-\hat{\mathcal{G}}_k^{(i-1)}\right)^\dagger,\forall k,\quad(14)$$

where i is the iteration index, μ_1, $\{\mu_{2,k}\}$ are small positive numbers, $\hat{\mathcal{P}}^{(i-1)}\triangleq\sum_{k=1}^K\sum_{n=1}^N\hat{a}^{*(i-1)}_{k,n}\hat{p}^{*(i-1)}_{k,n}$, $\hat{\mathcal{G}}_k^{(i-1)}\triangleq\sum_{n=1}^N\hat{a}^{*(i-1)}_{k,n}G_{k,n}(\hat{r}^{*(i-1)}_{k,n},\hat{p}^{*(i-1)}_{k,n})$, $\hat{r}^{*(i)}_{k,n},\hat{p}^{*(i)}_{k,n}$ are the $\phi_{k,n}$ maximizers corresponding to the $\lambda_1^{*(i)}$, $\lambda_{2,k}^{*(i)}$ values, and $(x)^\dagger\triangleq\max(x,0)$.

Unfortunately, problem (4) is not concave which means that the dual minimizing $\lambda_1,\{\lambda_{2,k}\}$ are not in general equal to the optimal multiplier values that solve the original RA problem [3]. Therefore, it may well turn out that the power and rate constraints (C1 and C2, respectively) are not satisfied by employing them. In the case of exceeded power, a simple method for obtaining a feasible solution can be employed, namely, the assigned powers are normalized so that C1 is satisfied with equality [8]. More innovative (and complicated) procedures are required when C2 is not satisfied. Note, however, that due to the randomness of the wireless channel it is possible that a non-feasible solution may not exist due to excessive individual rate requirements, therefore C2 cannot be satisfied for any value of $\{\lambda_{2,k}\}$. Unfortunately detecting non-feasibility is a difficult task, which is not considered explicitly in this paper.

IV. Low Complexity, Adaptive RA algorithm

The RA algorithm described above is computationally expensive due to the following reasons:

- Maximization of $\phi_{k,n}(\hat{r}_{k,n},\hat{p}_{k,n},\lambda^*_1,\lambda^*_{2,k})$ over $\hat{r}_{k,n}$ and $\hat{p}_{k,n}$ has to be performed numerically.
- Numerical search is also required for obtaining the optimal values (estimates) of $\lambda_1,\{\lambda_{2,k}\}$.

These issues make the practical applicability of the algorithm questionable. In the following subsections various approximations and algorithms are invoked in order to come up with a simplified version of the optimal RA algorithm.

A. Closed form approximation of optimal power-rate allocation

In [4], a number of algebraic approximations/simplifications were employed that result in a closed form solution of the optimal rate-power pair ($\phi_{k,n}$ maximizers) for a given set of Lagrange multipliers, that are summarized in the following.

Towards obtaining a closed form solution for $\hat{r}^*_{k,n}$ the Gaussian approximation of $p(C_{k,n}|\hat{h}_{k,n};\hat{p}_{k,n})$ proposed in [9] was employed that results in a closed form expression of

[2]This follows from (5), (6).

$P_{\text{out}}(\hat{r}_{k,n}, \hat{p}_{k,n})$ and, in turn, of $\phi_{k,n}(\hat{r}_{k,n}, \hat{p}_{k,n}, \lambda_1^*, \lambda_{2,k}^*)$. Substituting this expression in (6) and employing a second order Taylor approximation the optimal rate can be approximated as

$$\hat{r}_{k,n}^* = \begin{cases} \hat{C}_{k,n} - \Delta\hat{C}_{k,n}, & \text{for } \hat{C}_{k,n} \geq \sigma_{\hat{C},k,n}\sqrt{\pi/2}, \\ 0, & \text{for } \hat{C}_{k,n} < \sigma_{\hat{C},k,n}\sqrt{\pi/2}, \end{cases} \quad (15)$$

where $\Delta\hat{C}_{k,n}$ is a (small) non-negative value shown at the bottom of the page, representing the difference of the optimal rate from the nominal capacity $\hat{C}_{k,n}(\hat{p}_{k,n}) \triangleq \log_2(1+|\hat{h}_{k,n}|^2\hat{p}_{k,n})$ (capacity assuming the channel estimate is perfect), where $\sigma^2_{\hat{C},k,n}(\hat{p}_{k,n}) \triangleq a_{k,n}2\ln 2/(a_{k,n}+b_{k,n})^2$ is the uncertainty of the capacity conditioned on the partial CSI with $a_{k,n} \triangleq |\hat{h}_{k,n}|^2/\sigma^2_{\hat{h},k,n}$, and $b_{k,n} \triangleq 1/(\hat{p}_{k,n}\sigma^2_{\hat{h},k,n})$. Note that a non-zero optimal rate is always smaller that the nominal capacity in order to reduce outage probability, with $\Delta\hat{C}_{k,n} = 0$ when $\sigma^2_{\hat{C},k,n} = 0$ (perfect CSI) as expected.

In order to avoid complications due to dependence of the rate formula of (15) on the allocated power, uniform power allocation (UPA) was proposed irrespective of user and corresponding channel, i.e., allocate the same power, $\hat{p}_{k,n}^* = \mathcal{P}/N\ \forall k, n$. This approach discards the need of optimizing over $\hat{p}_{k,n}$ since constraint C1 is automatically satisfied with equality ($\lambda_1 = 0$ in this case). Clearly, imposing a UPA will not reach the performance of the system under optimal power allocation, however, it is well known that the latter is advantageous only under operation in very low signal-to-noise ratios (SNR) [1], [8], as also demonstrated by simulation results in [4].

B. On-line Update of $\{\lambda_{2,k}^*\}$ Multipliers

The closed-form approximation of the optimal power and rate allocation significantly reduces complexity of the RA algorithm. However, the need for obtaining optimal values for the multipliers associated with the rate constraints C2 via a numerical search may still be prohibitive for practical implementation. Especially in the case where the channel, user weights and rate requests vary significantly between successive RA frames, a new set of $\{\lambda_{2,k}^*\}$ values must be found in each frame. In order to obtain an efficient algorithm for obtaining $\{\lambda_{2,k}^*\}$ values, the following observation on the importance of $\{\lambda_{2,k}^*\}$ for the *simplified* RA process is crucial, namely, the closed form solution (15) for rate allocation *does not* depend on $\{\lambda_{2,k}^*\}$.[3] In other words, $\{\lambda_{2,k}^*\}$ impose the user goodput constraints by biasing the *subcarrier allocation* process in favoring users with larger goodput constraints, effectively providing them with a larger "total" user weight $w_k + \lambda_{2,k}^*$ and corresponding link quality indicator, compared to users with smaller goodput requirements.

This observation suggests the idea of employing some simple ad-hoc algorithm [instead of the iterations of Eq. (14)] that provides $\{\lambda_{2,k}^*\}$ values that are larger for users with larger goodput requirements. In this paper an adaptive method is proposed that achieves this, where adaptation is performed on a per-RA-frame interval,[4] i.e., based on the $\{\lambda_{2,k}^*\}$ values employed in the previous RA frame the current $\{\lambda_{2,k}^*\}$ values are updated similar to (14) where now i stands for RA frame index and $\hat{\mathcal{G}}_k^{(i)}$ is an estimate of the goodput achieved by the user up to the ith RA frame (where this estimate may simply be the goodput provided in the previous RA frame or even a running average of the goodput provided over the past RA frames). The advantage of this scheme is that requires a single update (iteration) for $\{\lambda_{2,k}^*\}$ evaluation while allowing for adaptation of the system to time–varying goodput requests (c.f. Sec. V).

C. Summary of Simplified Algorithm

Based on the discussion of the previous subsections the simplified algorithm steps are as follows (executed on the mth RA frame):

- Update multipliers $\{\lambda_{2,k}^{*(m)}\}$ as per (14)[5].
- Compute rate allocation $\hat{r}_{k,n}^*$ as per (15) for uniform power allocation $\hat{p}_{k,n}^* = \mathcal{P}/N\ \forall\ k,\ n$.
- Assign subcarrier n to user with the largest value of $\phi_{k,n}(\hat{r}_{k,n}^*, \hat{p}_{k,n}^*, 0, \lambda_{2,k}^{*(m)})$.
- Update goodput estimates $\hat{\mathcal{G}}_k^{(m)}\ \forall k$.

V. Simulation Results

A downlink OFDMA system was simulated in order to assess the effectiveness of the proposed simplified RA algorithm described in Sec. IV. A in all cases.[6] Each user is assumed to experience independent frequency selective fading, with the subcarrier channel gains distributed as $p(h_{k,n}) = \mathcal{CN}(0, \text{SNR}_k), \forall k, n$, where SNR_k is the (average) subcarrier SNR of the kth user (same for all subcarriers). The RA algorithm employs channel estimates $\hat{h}_{k,n}$ that result in the conditional pdf of (1) with $\sigma^2_{\hat{h},k,n} = 0.1$.

Achievable Rate Region: Figure 1 depicts the achievable rate (goodput) region for $N = 8$ with $K = 2$ users experiencing SNRs of $\text{SNR}_1 = 10$dB and $\text{SNR}_2 = 5$dB, obtained by varying the user weights in the space

[3] This is not the case when numerical maximization of $\phi_{k,n}$ is performed.

[4] Similar to the on-line algorithm first proposed in [7]. See also [10].

[5] With the corresponding quantities as interpreted in Scc. IV. B

[6] The reader is referred to [4] for a comparison of the various versions of the RA algorithm in a non-goodput-constrained setting.

$$\Delta\hat{C}_{k,n} = \frac{\sqrt{\sigma^2_{\hat{C},k,n}\left(2\hat{C}^2_{k,n} - \hat{C}_{k,n}\sqrt{2\pi\sigma^2_{\hat{C},k,n}} + 4\sigma^2_{\hat{C},k,n}\right)} - 2\sigma^2_{\hat{C},k,n}}{\hat{C}_{n,k}} \quad (17)$$

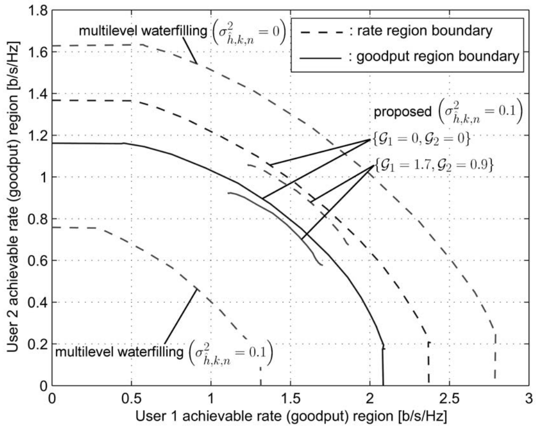

Fig. 1. Two–user achievable ergodic rate (goodput) region.

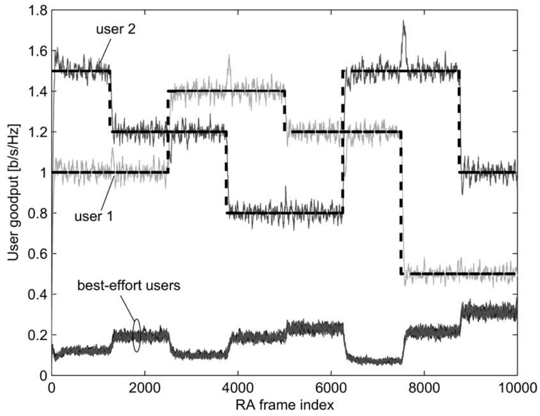

Fig. 2. Adaptation of proposed RA algorithm over users' minimum goodput requests.

$\left\{(w_1, w_2) \in [0,1]^2; w_1 + w_2 = 1\right\}$ and averaging over independent channel realizations (ergodic performance). Specifically, performance of the proposed algorithm with $\{\lambda^*_{2,k}\}$ found as per (15) for each channel realization is depicted, both in terms of achievable goodput (solid curves) as well as successfully transmitted rate (dotted curves), for various values of goodput constraints. For comparison purposes, performance of the optimal RA algorithm with no rate (goodput) constraints under perfect CSI assumption (multilevel waterfilling) [2] is also shown, both for $\sigma^2_{\hat{h},k,n} = 0.1$ and $\sigma^2_{\hat{h},k,n} = 0$ (perfect CSI) employing UPA in the case also. It can be seen that imperfect CSI severely decreases the achievable rate region if not properly accounted for (as is the case with the algorithm of [2]). On the other hand, the proposed algorithm compensates for a large percentage of the loss incurred by the partial CSI, when no goodput constraints are considered. In addition, it can be seen that the algorithm successfully provides the minimum user goodput requests (when imposed) with a corresponding reduction of the achievable region as expected. It is interesting to note that, although the algorithm may not provide the minimum user guarantees in every RA frame (either due to infeasibility of the solution or nonconvexity of the problem) constraints are satisfied when channel-averaged (ergodic) metrics are considered.

User Goodput Adaptation: Performance of an $N = 32$ subcarriers system with $K = 8$ users, all of them experiencing an SNR $= 10$ dB, is examined in this simulation. Users 1 and 2 request for minimum goodputs which vary with time, whereas the remaining users have no minimum goodput requirements (best–efort users). All user weights are set to $w_k = 1$. Figure 2 depicts the instantaneously achieved goodput for each user over a period of 10000 RA frames, provided by the simplified–adaptive algorithm of Sec. IV. C. For each RA frame, an independent channel realization is generated for each user, and the performance depicted is by averaging over 10 independent simulation runs of the whole RA period in order to smoothen the curves. As can be seen the RA algorithm manages to "track" the variable user goodput requests (indicated by the slash curves) with a certain fluctuation. As expected, performance of best effort users is directly affected by how stringent goodput requirements are.

VI. Conclusion

The problem of RA in downlink OFDMA under partial CSI at the BS and minimum user requests was investigated. By adopting the notion of goodput as an appropriate measure of system performance and exploiting the statistical description of the partial CSI optimal, a greedy subcarrier allocation was shown to be optimum. Various simplifications were invoked in order to derive a low–complexity RA algorithm. The resulting scheme also allows for simple adaptation to time–varying user requirements.

References

[1] L. Hoo, B. Halder, J. Tellado, and J. Cioffi, "Multiuser transmit optimization for multicarrier broadcast channels: asymmptotic FDMA capacity region and algorithms," *IEEE Trans. Commun.*, vol. 52, pp. 922–930, June 2004.

[2] G. Li and H. Liu, "On the optimality of the OFDMA network," *IEEE Commun. Lett.*, vol. 9, pp. 438–440, May 2005.

[3] J. Nocedal and S. Wright, *Numerical Optimization*, 2nd ed., Springer, 2006.

[4] S. Stefanatos and N. Dimitriou, "Downlink OFDMA resource allocation under partial channel state information," accepted for publication in IEEE ICC 09, Dresden, Germany, June 2009.

[5] M. Tao, Y.-C. Liang, and F. Zhang, "Resource allocation for delay differentiated traffic in multiuser OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 2190–2201, June 2008.

[6] R. Wang and V. K. N. Lau, "Cross layer design of downlink multi-antenna OFDMA systems with imperfect CSIT for slow fading channels," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2417–2421, July 2007.

[7] X. Wang, G. B. Giannakis, and Y. Yu, "Channel–adaptive optimal OFDMA scheduling," *Proc. of 41st Conf. on Info. Sciences and Systems*, pp. 536–541, Johns Hopkins Univ. NJ, March 14–16, 2007.

[8] I. C. Wong and B. L. Evans, "Optimal downlink OFDMA resource allocation with linear complexity to maximize ergodic rates," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 962–971, March 2008.

[9] Y. Yao and G. B. Giannakis, "Rate–maximizing power allocation in OFDM based on partial channel knowledge," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 1073–1083, May 2005.

[10] R. Zhang, "Optimal dynamic resource allocation for multi-antenna broadcasting with heterogeneous delay constrained traffic," *IEEE J. Select. Areas Signal Processing*, vol. 2, pp. 243–255, April 2008.