Cross-Layer Design for Single-Cell OFDMA systems with Heterogeneous QoS and Partial CSIT

Charilaos C. Zarakovitis, Qiang Ni and Dionysios Skordoulis

School of Engineering & Design, Brunel University, Uxbridge, Middlesex, UB8 3PH, London, UK

Email: {charilaos.zarakovitis, qiang.ni, dionysios.skordoulis}@brunel.ac.uk

Abstract— **This paper proposes a novel cross-layer scheduling scheme for a single-cell orthogonal frequency division multiple access (OFDMA) wireless system with partial channel state information (CSI) at transmitter (CSIT) and heterogeneous user delay requirements. Previous research efforts on OFDMA resource allocation are typically based on the availability of perfect CSI or imperfect CSI but with small error variance. Either case consists to typify a non tangible system as the potential facts of channel feedback delay or large channel estimation errors have not been considered. Thus, to attain a more realistic resolution our cross-layer design determines optimal subcarrier and power allocation policies based on partial CSIT and individual user's quality of service (QoS) requirements. The simulation results show that the proposed cross-layer scheduler can maximize the system's throughput and at the same time satisfy heterogeneous delay requirements of various users with significant low power consumption.**

Keywords; multiple access, wireless networks, cross-layer analysis, OFDMA, resource allocation, QoS, CSI.

I. INTRODUCTION

OFDM is considered as one of the most efficient transmission techniques in a multi-path radio channel. The related multiple access scheme namely OFDMA is indicated as an innovative scheme for the next-generation wireless networks such as 3G LTE and WiMAX. This scheme provides multi-user diversity and high flexibility by assigning subsets of subcarriers to individual users according to the given QoS requirements. Emerging cross-layer strategies and algorithms on radio resource management (RRM) can control parameters such as transmit power and subcarrier allocation in order to utilize the network's resources as efficiently as possible. The optimal system characteristics and requirements include the efficient use of the available recourses ability to provide desired data rates in a harsh radio channel environment, support for various QoS requirements and mitigate the inter symbol interference (ISI) and co-channel interference (CCI).

In [6] the authors proposed a cross-layer scheduling algorithm targeting the minimization of average system delay in multi-access channel for homogeneous users. In [2] heterogeneous delay constraints imposed by various applications is considered for OFDMA systems under perfect CSIT knowledge. In the case of partial CSIT there will be systematic packet errors whenever the scheduled data rate exceeds the instantaneous mutual information despite the use of strong channel coding. The authors in [3] introduce a delay-sensitive cross-layer scheduler in order to overcome this effect by employing outdated CSIT. Alongside, bit- and power-loading algorithms were pursued in [7], [8], where partial CSIT was utilized to adapt the constellation size and power, adhering to a certain target bit-error-rate (BER) per subcarrier. A robust power-loading algorithm for an OFDM system with imperfect CSIT in order to maximize only throughput given a target outage probability under independent loading for different subcarriers is studied in [1]. Despite the related work, no study is carried on OFDMA optimization based on individual user's QoS requirements with partial CSIT by using a power-bit loading (PBL) algorithm.

This paper presents a novel cross-layer scheduler which utilizes a PBL algorithm that supports heterogeneous QoS requirements of both delay-sensitive and delay-insensitive users within a single-cell OFDMA system. Our technique maximizes the system throughput at the physical-layer under partial CSIT and aims for the satisfaction of a target outage probability and delay requirements specified by the applications given by the queuing model. For the packet throughput and the queues' delay analysis, the channel effect is described in the cross-layer design in terms of a time-varying server process using the M/G/1 model. A convex optimization problem is formulated from the adaptation of the delay constraints, which after relaxation results into an optimal subcarrier, power and rate allocation solution. The resulted optimal power allocation policy has a multi-user water-filling structure and has been produced by an adaptive power allocation (APA) algorithm. On the other hand the optimal subcarrier allocation strategy is based on a low complexity dynamic subcarrier allocation (DSA) greedy algorithm. In order to get the Lagrangian multipliers of DSA and APA algorithms a novel offline mechanism has been introduced that calculates fast and efficient values using the secant root-finding method. For this work both queuing and information theories are applied in order to merge the queue-model and the physical-layer (PHY) dynamics accounting the heterogeneity of the users' delay requirements.

The rest of this paper is organized as follows: Section II describes the system model including channel model, partial CSIT model and media access control (MAC) layer models. Section III presents the optimization problem formulation. The optimal subcarrier, power and rate allocation strategies are explained in Section IV with the joint APA and DSA algorithm structure. Simulation results are studied in Section V. Section VI concludes the paper.

II. SYSTEM MODEL

Figure 1 illustrates the cross-layer OFDMA system for multi-user heterogeneous applications under partial CSIT, with N_F subcarriers and K users. The subcarrier allocation matrix is denoted by $\mathbf{S}_{N_F \times K} = [s_{ij}]$, where $s_{ij} = 1$ when the subcarrier i is allocated to the user j, and $s_{ij} = 0$ when the subcarrier i is not allocated to the user j. The power and rate allocation matrices are described by $\mathbf{P}_{N_F \times K} = [p_{ij}]$ and $\mathbf{R}_{N_F \times K} = [r_{ij}]$, where p_{ij} and r_{ij} indicate the power and the instantaneous data rate allocated for user j on subcarrier i respectively. Note that at one time one subcarrier i can only be occupied by one user j, hence $\sum_{j=1}^{K} s_{ij} = 1$. The average total transmit power of the base station (BS) should satisfy that $E\left[\sum_{j=1}^{K}\sum_{i=1}^{N_F} s_{ij} p_{ij}\right] \leq P_{TOTAL}$, where P_{TOTAL} is the total available power at the BS.

A. Channel Model

Let us consider a downlink OFDM transmission over frequency-selective multi-path fading channel. The transmission is time varying, due to multi-path propagation and Doppler shifts. After removing the cyclic prefix and performing fast Fourier transform (FFT) at the receiver, the relation between input and output at the frequency domain, is given as $\mathbf{y} = \mathbf{H}_{h_f}\mathbf{x} + \mathbf{z}$, where $\mathbf{y}$ and $\mathbf{x}$ denote the received and transmitted signal correspondingly. $\mathbf{H}_{h_f}$ is a diagonal matrix with diagonal elements $\mathbf{h}_f$ indicating the complex channel gain and $\mathbf{z}$ is the zero mean complex Gaussian noise $\mathbf{z} \sim \mathcal{CN}(\mathbf{0}, \sigma_Z^2\mathbf{I})$ with identity matrix $\mathbf{I}$. The transmitter acquires channel knowledge by channel estimation in time division duplex (TDD) operation.

B. CSIT Model

The channel state information received (CSIR) is assumed to be perfect in our system. The downlink CSIT at the BS is estimated from uplink dedicated training blocks sent by all K users. The partial CSIT $\hat{\mathbf{h}}_f \sim \mathcal{CN}(\mathbf{h}_f, \mathbf{M}_{\Delta h_f})$ is represented by $\hat{\mathbf{h}}_f = \mathbf{h}_f + \Delta\hat{\mathbf{h}}_f$, where $\mathbf{h}_f$ is the channel feedback and $\Delta\hat{\mathbf{h}}_f$ an error term with known probability density function (PDF). By adopting the channel modelling in [1], [6] and [9] it can be verified from the covariance matrix of the channel estimation error $\mathbf{M}_{\Delta h_f} = \left[\Delta\hat{\mathbf{h}}_f \Delta\hat{\mathbf{h}}_f^{\mathcal{H}}\right]$ that the channel estimation errors of different subcarriers are correlated. Also the CSIT error matrix $\Delta\mathbf{h}_f$ and the estimation matrix $\hat{\mathbf{h}}_f$ are uncorrelated having covariance matrix equal to zero $E\left[\Delta\mathbf{h}_f \hat{\mathbf{h}}_f\right] = 0$.
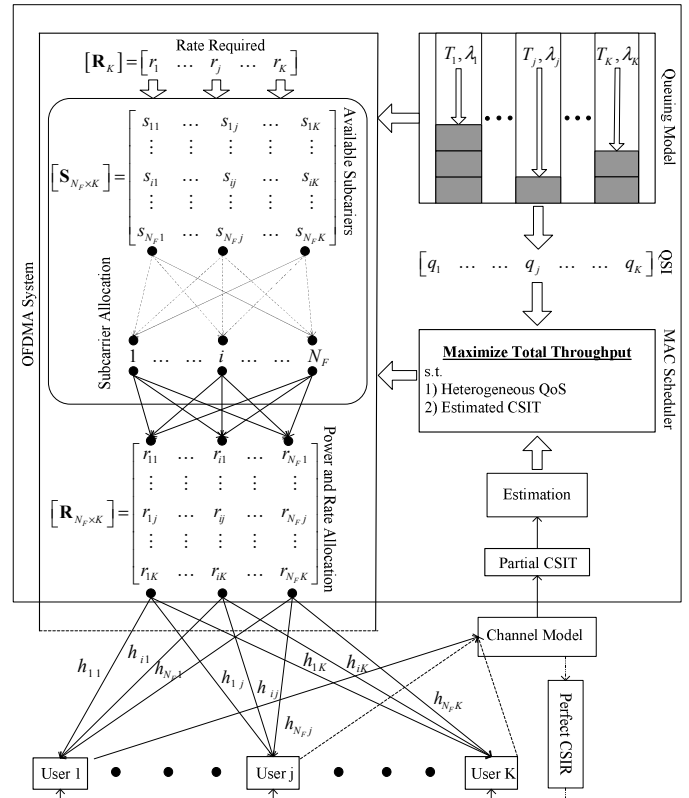

Figure 1. Cross layer scheduling and channel model for OFDMA system with heterogeneous QoS users and partial CSIT.

C. Physical Layer Model

In this section, the data rate required for user j on subcarrier i will be achieved by following the Shannon's capacity theory.

The mutual information between the input and the output $I\left(X_{ij}, Y_{ij} \mid \hat{h}_{ij}\right)$ is a random variable denoted by [5]

$$I_{ij} = \log\left(1 + p_{ij}\left|\hat{h}_{ij}\right|^2 / \sigma_Z^2\right) \tag{1}$$

, where $\hat{h}_{ij} \sim \mathcal{CN}\left(h_{ij}, \sigma_h^2\right)$ the imperfect CSIT realization σ_h^2 the CSIT error variance and σ_Z^2 the noise variance. Given a target outage probability P_{out}, the data rate r_{ij} of user j on subcarrier i must satisfy that $\Pr\left(I_{ij} < r_{ij} \mid \hat{h}_{ij} = h_{ij}\right) \leq P_{out}$. Since $\hat{h}_{ij} \sim \mathcal{CN}\left(h_{ij}, \sigma_h^2\right)$, $\left|\hat{h}_{ij}\right|^2$ follows a non-central chi squared distribution with 2 degrees of freedom and a non-central parameter $\left|h_{ij}\right|^2 / \sigma_h^2$. The PDF of I_{ij} in (1) is expressed as

$$p\left(I_{ij} \mid \hat{h}_{ij} = h_{ij}\right) = \frac{\sigma_Z^2}{p_{ij}\sigma_h^2} \cdot e^{I_{ij} - \frac{\left|\hat{h}_{ij}\right|}{\sigma_h^2} - \frac{\sigma_Z^2}{p_{ij}\sigma_h^2}(e^{I_{ij}} - 1)} I_0\left(2\sqrt{\frac{\left|\hat{h}_{ij}\right|^2}{p_{ij}\sigma_h^2}(e^{I_{ij}} - 1)}\right)$$

, where $I_0(\cdot)$ is the zero-order modified Bessel function of the first kind. Continuously the maximum data rate r_{ij} must satisfy

$$\int_0^{r_{ij}} p\left(I_{ij}\big|\hat{h}_{ij}=h_{ij}\right)dI = P_{out} \tag{2}$$

According to [1], for large $\left|\hat{h}_{ij}\right|^2/\sigma_h^2$ the probability distribution of the mutual information I_{ij} in (1) can be approximated by the Gaussian distribution as follows

$$I_{ij} \sim \mathcal{N}\left(\log\left(1+\frac{\left|\hat{h}_{ij}\right|^2 p_{ij}}{\sigma_Z^2}\right), \frac{2\left|\hat{h}_{ij}\right|^2 p_{ij}^2\sigma_h^2}{\left(\left|\hat{h}_{ij}\right|^2 p_{ij}+\sigma_Z^2\right)^2}\right) \tag{3}$$

From the definition of r_{ij} in (2) and from (3) the instant outage rate is obtained

$$r_{ij} = \log\left(1+\frac{\left|\hat{h}_{ij}\right|^2 p_{ij}}{\sigma_Z^2}\right) - \frac{\sqrt{2}\left|\hat{h}_{ij}\right| p_{ij}\sigma_h}{\left|\hat{h}_{ij}\right|^2 p_{ij}+\sigma_Z^2}Q^{-1}\left(P_{out}\right) \tag{4}$$

, where $Q(\cdot)$ is the complementary Gaussian cumulative distribution function (CCDF).

D. Queuing System Model

Let us assume that the considered system is modelled as a M/G/1 queue, and packets arrive at the jth user's buffer according to a Poisson arrival process with independent rate λ_j (in packets per time slots) with packets of fixed size F (in bits). Note that the packet size can be varied but due to brevity of the queuing model analysis in this paper is considered constant. We introduce K tuples $\left[F,\lambda_j,T_j^{\max}\right]$ to describe the QoS parameters that K users may require, where $T_j^{\max}$ (in time slots) denotes the maximum delay of the jth user. The queue state information (QSI) of each user is denoted by B_K. We define $B_K = q_j$ the $K\times1$ vector with the jth component denoting the number of packets remaining in user's buffer.

According to the above M/G/1 queue assumption, the average delay of each user $E[D_j]$ can be obtained by

$$E\left[D_j\right] = \frac{\lambda_j E\left[X_j^2\right]}{2\left(1-\lambda_j E\left[X_j\right]\right)} \tag{5}$$

$E\left[X_j\right]$ represents the average service time of user j, having variance $E\left[X_j^2\right]$. The service time X_j of user j in an OFDMA system can be characterized as the number of the subcarriers M, allocated to user j as described below

$$\text{M} = \left\{N_F : U_j = F\right\} \tag{6}$$

, where $U_j \overset{\Delta}{=} \sum_{i=1}^{N_F} q_{ij}$ shows the number of bits loaded to N_F subcarriers and q_{ij} the number of identically distributed bits of user j loaded to subcarrier i. The data streams can be grouped in a serial or parallel manner. According to (6) and conditioning on $U_j = F$ we get $X_j = N_F$. The average value and the variance of X_j correspondingly are

$$E\left[X_j\right] = F/E\left[q_{ij}\big|U_j=F\right] \tag{7}$$

$$E\left[X_j^2\right] \geq F/E\left[q_{ij}^2\big|U_j=F\right] \tag{8}$$

By substituting (7), (8) at (5) and applying the quadratic formula is resulted that the arrival rate at the queue of user j is

$$q_j \geq F\left(\lambda_j T_j^{\max}+\sqrt{T_j^{\max}\lambda_j\left(T_j^{\max}\lambda_j+2\right)}\right)/2T_j^{\max} \tag{9}$$

In order to guarantee the QoS requirements, the average data rate of each user must satisfy the condition $\sum_{i=1}^{N_F} r_{ij} \geq q_j\left(F,T_j^{\max},\lambda_j\right)$, which means that the instantaneous data rate of user j must not be smaller than the incoming traffic arrival rate to user j's queue in order to achieve the stability of the queues. Thus even if a user doesn't have delay requirements the system provides an average scheduled data rate of at least the same as the bits arrival rate to user's buffer.

Since BW/N_F is the bandwidth allocation for each subcarrier, where BW denotes the total system's bandwidth, the maximum lower bound of the average scheduled data rate, required by a delay- sensitive user j, is expressed by

$$\sum_{i=1}^{N_F} r_{ij} \geq \frac{N_F}{BW}q_j\left(F,T_j^{\max},\lambda_j\right) \tag{10}$$

It should be noted that a necessary condition, $\sum_{i=1}^{N_F} r_{ij} \geq F\lambda_j$ must hold for the delay-insensitive users in (9) when, according to the general stability condition $T_j^{\max}\to\infty$. Even without any delay requirement the average data rate of user j, $\sum_{i=1}^{N_F} r_{ij}$ should be at least the same to user's j queue $F\lambda_j$. This results stabilization of all users' queues since the data rate of delay-insensitive and delay-sensitive users must meet the corresponding queue rate, defined by the queue state q_j of each user.

E. MAC Layer Model

The MAC layer is responsible for the cross-layer resource allocation scheduling of the system. At the beginning of every OFDM frame consisted by nL_F OFDM blocks, the receiver estimates the CSIT as described in Section II.A. Based on the CSIT estimation and the QoS requirements of the users described in Section II.D, the scheduler determines the

subcarrier allocation policy $\mathcal{S}_{N_F \times K}\left[\hat{H}_{N_F \times K}, q_K\right]$, the power allocation policy $\mathcal{P}_{N_F \times K}\left[\hat{H}_{N_F \times K}, q_K\right]$ and the corresponding data rating allocation policy $\mathcal{R}_{N_F \times K}\left[\hat{H}_{N_F \times K}, q_K\right]$, where $\hat{H}_{N_F \times K}$ is the estimated CSIT model and q_K a $K \times 1$ vector, which describes the queue state of each user of the system.

III. PROBLEM FORMULATION

Based on the system model presented so far, the OFDMA cross-layer design with heterogeneous QoS and partial CSIT can be formulated as follows;

Find optimal resource allocation policies
$\mathcal{S}_{N_F \times K}\left[\hat{H}_{N_F \times K}, q_K\right], \mathcal{P}_{N_F \times K}\left[\hat{H}_{N_F \times K}, q_K\right]$ and $\mathcal{R}_{N_F \times K}\left[\hat{H}_{N_F \times K}, q_K\right]$
s.t.

$$\max_{\mathcal{S}, \mathcal{P}, \mathcal{R}} E\left[\sum_{j=1}^{K}\sum_{i=1}^{N_F} s_{ij}\left(\log\left(1 + \frac{|\hat{h}_{ij}|^2 p_{ij}}{\sigma_Z^2}\right) - \frac{\sqrt{2}|\hat{h}_{ij}| p_{ij}\sigma_h}{|\hat{h}_{ij}|^2 p_{ij} + \sigma_Z^2}Q^{-1}\left(P_{out}\right)\right)\right]$$

s.t. (A1) $s_{ij} = \{0,1\}, p_{ij} \in (0,1)$

(A2) $\sum_{j=1}^{K} s_{ij} = 1$ (11)

(A3) $p_{ij} \geq 0$

(A4) $E\left[\sum_{i=1}^{N_F} s_{ij}\left(\log\left(1 + \frac{|\hat{h}_{ij}|^2 p_{ij}}{\sigma_Z^2}\right) - \frac{\sqrt{2}|\hat{h}_{ij}| p_{ij}\sigma_h}{|\hat{h}_{ij}|^2 p_{ij} + \sigma_Z^2}Q^{-1}\left(P_{out}\right)\right)\right]$
$\geq \frac{N_F}{t_s \cdot BW} q_j\left(F, T_j^{\max}, \lambda_j\right)$

(A5) $\sum_{j=1}^{K}\sum_{i=1}^{N_F} s_{ij} p_{ij} \leq P_{TOTAL}$

(A6) $P_{out} = P_{application}$

, where $i = 1,...,N_F$ and $j = 1,...,K$. Constraints (A1) and (A2) are used to ensure that only one user can occupy a subcarrier i at one time. Constraint (A3) is used to ensure that the power takes only positive values, (A4) is the average delay constraint, (A5) is the average total power constraint and (A6) is used to ensure that the outage probability P_{out} satisfies a target outage probability $P_{application}$ specified by the QoS requirements of the user.

IV. SUBCARRIER, POWER AND RATE ALLOCATION STRATEGIES

A. Optimal Subcarrier, Power and Rate Allocation

Each frame is allocated to a user with the best channel gain. The optimization problem in (11) is a mixed integer and convex optimization problem. The convexity of the objective function has been proved with a condition on the target outage probability should satisfy $Q^{-1}\left(P_{out}\right) > \sqrt{\theta}\left[1+\varphi\right]/8$, where

$\theta = |\hat{h}_{ij}|^2 / \sigma_h^2$ the non central parameter of the chi-squared distribution of the mutual information $F_{\chi^2}\left(I_i\right)$ on subcarrier i and $\varphi = p_{ij}|h_{ij}|^2 / \sigma_Z^2$. We can not present the proof here due to space limitation. By using the Lagrangian function method and defining the Karush-Kuhn-Tucker (KKT) [4] conditions for (11), the optimal power allocation policy can be achieved as $\mathcal{P}_{N_F \times K}^* = \left[p_{ij}^*\right]$, which results to the minimum power needed for each user according to its QoS requirements is

$$\tilde{p}_{ij}^* = \begin{cases} \left(\beta\left(\xi_j, \mu\right) - \frac{\sigma_Z^2}{|\hat{h}_{ij}|^2}\right)^+, & \forall s_{ij} = 1 \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

, with

$$\beta\left(\xi_j, \mu\right) = \frac{1}{2}\left(\omega_j + \frac{\sqrt{|\hat{h}_{ij}|\omega_j\left(|\hat{h}_{ij}|^3\omega_j - 4\sqrt{2}\sigma_h\sigma_Z^2 Q^{-1}\left(P_{out}\right)\right)}}{|\hat{h}_{ij}|^2}\right)^+$$

, where $\omega_j = \xi_j + 1/\mu$ and $(x)^+$ means $\max(0, x)$. The corresponding optimal subcarrier allocation policy $\mathcal{S}_{N_F \times K}^* = \left[s_{ij}^*\right]$ among the N_F subcarriers of the system is given by

$$j^* = \arg\max_{j \in [1,K]} \hat{H}_{ij}\left(\xi_j, \mu\right)$$
$$\text{for } s_{ij}^* = \begin{cases} 1, & j = j^* \\ 0, & j^* \text{does not exist} \end{cases} \quad (13)$$

, where

$$\hat{H}_{ij}\left(\xi_j, \mu\right) = \left(\xi_j + 1\right)\left(\log\left(\frac{|\hat{h}_{ij}|^2\left(\beta\left(\xi_j, \mu\right)\right)}{\sigma_Z^2}\right)^+ + \left(\frac{\sigma_Z^2}{|\hat{h}_{ij}|^2\left(\beta\left(\xi_j, \mu\right)\right)} - 1\right)^+\right)$$
$$+ \frac{\sqrt{2}\sigma_h Q^{-1}\left(P_{out}\right)}{|\hat{h}_{ij}|}\left(\frac{2\sigma_Z^2}{|\hat{h}_{ij}|^2\left(\beta\left(\xi_j, \mu\right)\right)} - \frac{\sigma_Z^4}{|\hat{h}_{ij}|^4\left(\beta\left(\xi_j, \mu\right)\right)^2} - 1\right)^+$$

Note, that the search of $\{s_{ij}^*\}$ is computationally efficient with linear complexity $N_F \times K$.

The Lagrange multipliers $\left(\xi_j, \mu, \nu_i\right)$ satisfy the power constraint in (12) and the delay constraint in (A4) for all the users. The corresponding optimal rate allocation is given, in the case where subcarrier allocation has been performed, by substitution of (12) into (4).

B. Joint APA and DSA Algorithm

The subcarrier allocation in (13) has been implemented by a low $N_F \times K$ complexity greedy algorithm. The optimal power allocation expressed in (12) can be applied by a multi-level water-filling strategy, where $\xi_j + 1/\mu$ is the water level

of user j. Users with urgent packets/delay-sensitive are allowed to transmit at higher power level $\xi_j + 1/\mu$ according to their Class/urgency, while non-urgent users/delay-insensitive are allocated with power level $1/\mu$.

The minimum power $P_{\min}$ required to support all delay constraints of users in (9) can be found by solving the following system of equations

$$\begin{cases} P_{\min} = E\left[\sum_{i=1}^{N_F}\sum_{j=1}^{K} s_{ij}\left(\beta(\xi_j,\mu) - \frac{\sigma_Z^2}{|\hat{h}_{ij}|^2}\right)^+\right] \\ r_{ij}^* = \frac{N_F}{t_s BW} q_j\left(F, T_j^{\max}, \lambda_j\right), \forall j \end{cases}$$

Suppose that $P_{TOT} \geq P_{\min}$, the Lagrange multipliers ξ_j, μ can be calculated by an offline part of the scheduling algorithm by solving the following system of equations

$$\begin{cases} E\left[\sum_{i=1}^{N_F}\sum_{j=1}^{K} s_{ij}\left(\beta(\xi_j,\mu) - \frac{\sigma_Z^2}{|\hat{h}_{ij}|^2}\right)^+\right] - P_{TOTAL} = 0 \\ \xi_j\left(r_{ij}^* - \frac{N_F}{t_s BW} q_j\left(F, T_j^{\max}, \lambda_j\right)\right) = 0 \end{cases} \quad (14)$$

The Lagrangian multipliers update in (14) has been implemented by an offline algorithm based on the secant and bisection methods. The iteration mechanism can be described by the following pseudo code:

Offline procedure for Dynamic Subcarrier Allocation

I) Initialization; choose an arbitrary μ.

Initialize a set of $\boldsymbol{\xi} = \{\xi_1,...,\xi_K\}$

Obtain $\boldsymbol{\xi_0}$ and $\boldsymbol{\xi_1}$ such as $\begin{cases} f_j(\mu,\xi_0) < 0 \\ f_j(\mu,\xi_1) > 0 \end{cases}, \forall j \in K$

II) Secant method: for each user, update $\boldsymbol{\xi_n}$ until:

$$\left|\frac{\xi_{j,n} - \xi_{j,n-1}}{f_j(\mu,\xi_n) - f_j(\mu,\xi_{n-1})} f_j(\mu,\xi_n)\right| < \frac{\delta}{K}$$

, where δ is a sufficiently small number.

$$\xi_{j,n+1} = \xi_{j,n} - \frac{\xi_{j,n} - \xi_{j,n-1}}{f_j(\mu,\xi_n) - f_j(\mu,\xi_{n-1})} f_j(\mu,\xi_n)$$

III) Find optimal region of ξ^*: Repeat step II) until find ξ^*

such that $\|f(\mu,\xi^*)\|^2 < \delta$, where

$$f(\mu,\xi^*) = \{f_1(\mu,\xi^*),...,f_K(\mu,\xi^*)\}$$

Offline procedure for Adaptive Power Allocation

IV) Check if $P(\mu,\xi^*) < 0$ as it is infeasible.

If $F(\mu,\xi^*) = 0$, then $\mu = \mu^*$ and $F(\mu,\xi^*)$ is the solution.

If $F(\mu,\xi^*) > 0$ then go to V)

V) Run a bisection algorithm with a feasible search region

of μ^*, denoted as $\left[\underline{\mu_0}, \overline{\mu_0}\right]$ such that $\begin{cases} P\left(\overline{\mu_0}, \xi^*\right) > 0 \\ P\left(\underline{\mu_0}, \xi^*\right) < 0 \end{cases}$

The search of μ^* is based on the following;

$$\mu_n = \frac{\underline{\mu_n} + \overline{\mu_n}}{2} \text{ and } \underline{\mu_{n+1}} = \begin{cases} \mu_n \text{ if } P(\mu_n,\xi^*) > 0 \\ \underline{\mu_n} \text{ if } P(\mu_n,\xi^*) < 0 \end{cases},$$

$$\overline{\mu_{n+1}} = \begin{cases} \mu_n \text{ if } P(\mu_n,\xi^*) > 0 \\ \overline{\mu_n} \text{ if } P(\mu_n,\xi^*) < 0 \end{cases}.$$

The iteration on μ_n continues until $\left|P\left(\mu_n, \xi^*(\mu_n)\right)\right|^2 < \delta$.

For each μ_n update $\xi^*(\mu_n)$ by repeating IV) and V).

V. SIMULATION RESULTS

The simulations are performed by using Monte Carlo simulation methods.

A. Simulation Modelling

We consider a single-cell OFDMA system with $BW = 80$ KHz and $N_F = 6$. The duration of a scheduling time slot is 0.002 sec. The channel fading between different users and different subcarriers has been modelled as independent identically distributed (i.i.d.) complex Gaussian process with unit variance. Four classes of users are considered specified by the tuples $(\mathbf{F}, \boldsymbol{\lambda}, \mathbf{T}) = \{(80,0.25,2),\ (80,0.4,4),\ (80,0.45,8),$ $(80,0.5,1000)\}$. Class 1, Class 2 and Class 3 users represent the delay sensitive traffic with heterogeneous delay requirements while Class 4 users represent the delay insensitive applications with heterogeneous traffic loading. The packet size of the queuing model is chosen to be constant at 80 bits.

B. Simulation Results

Figure 2 illustrates the total system's data rate in relation to the average total transmit power under a given target outage probability $P_{out} = 0.01$. For large CSIT error $\left(\sigma_h^2 = 0.1, 0.2, 0.3\right)$ and higher data rate requirements the minimum total power increases up to 3.6dB for the case of $\sigma_h^2 = 0.3$, which is something expected. The difference of the minimum total power required when the error variance is $\sigma_h^2 = 0.3$ compared with almost perfect CSIT conditions $\sigma_h^2 = 0.01$ is 1.91dB, which is significantly small concerning the difference of the CSIT imperfectness between these two cases. This makes the system to be low power consumption even under large CSIT error, which offers great benefits to the modern wireless systems.

Figure 3 presents the minimum required total power in relation to the error variance σ_h^2 given different P_{out}. The

minimum required total power increases as the error uncertainty increases especially when the outage probability is small $P_{out} = 0.001$. When the CSIT uncertainty is large $\sigma_h^2 = 0.5$ the difference between the minimum power requirement for $P_{out} = 0.001$ and $P_{out} = 0.1$ is 1.2dB. The system's behaviour is stable with small syncretic differences of power requirements under different target outage probabilities.

Figure 4 presents the average delay versus the CSIT error variance σ_h^2 under target outage probability $P_{out} = 0.01$ given total power $P_{TOT} = 1.5$dB. It is shown that the delay requirements of the delay-sensitive users are satisfied for maximum channel error variance $\sigma_h^2 = 0.1$ and the delay of the delay-insensitive user increases when the error variance increases. For high CSIT error the delay of the sensitive users is steady without overcoming their maximum delay tolerance and the delay of the insensitive user increases dramatically since the scheduler gives priority to the sensitive users.

VI. CONCLUSION

This paper presents an innovative cross-layer design for single-cell OFDMA systems with heterogeneous QoS requirements for both delay-sensitive and delay-insensitive users, under imperfect CSIT condition. The successful optimization of the power loading function subject to the system dynamics results to the optimal subcarrier, power and rate allocation policies. The design and performance of the joint APA and DSA mechanism has been thoroughly described throughout this paper. The simulation results show that the proposed cross-layer scheduler provides maximization of the overall system's throughput and efficiently serves the QoS differentiation of the mobile users, with real-time and robust performance having low power consumption even at large CSIT error. The proposed scheduler has been designed concerning realistic systems' characteristics and can be adopted by the next generation wireless standards.

REFERENCES

[1] Yingwei Yao, and Georgios B. Giannakis "Rate Maximizing Power Allocation in OFDM based on Partial Channel Knowledge" *IEEE Trans. on Wireless Comm.,* vol. 4, no. 3, pp. 1073-1083, May 2005

[2] David Hui, Vincent K. N. Lau, and Wong H. Lang "Cross-Layer Design for OFDMA Wireless Systems With Heterogeneous Delay Requirements" *IEEE Trans. on Wireless Comm.,* vol. 6, no. 8, pp. 2872-2880, Aug. 2007

[3] David Hui, and Vincent K. N. Lau "Delay-Sensitive Cross-Layer Designs for OFDMA Systems with Outdated CSIT" *WCNC 2007* proceedings.

[4] D. Bertsekas and R. Gallager, *Data Networks*, Pretice-Hall, 2nd ed., 1992

[5] I. E. Telatar, "Capacity of multi-antenna Gaussian channel," AT&T Bell Labs. Tech. Memo, 1995.

[6] E. M. Yeh, *Multi-access and Fading in Communication Networks*, Ph.D. Thesis, Department of Electrical Engineering and Computer Science, MIT, 2001.

[7] P. Xia, S. Zhou, and G. B. Giannakis, "MIMO-OFDM with ST coding and beam-forming adapted to partial CSI," in *Proc. 37th Conf. Information Sciences Systems*, Baltimore, MD, Mar. 2003.

[8] S. Ye, R. Blum, and L. Cimini, "Adaptive modulation for variable rate OFDM systems with imperfect channel information," in *Proc. IEEE Vehicular Technology Conf.*, vol. 2, Birmingham, AL, May 2002, pp. 767–771.

[9] W. C. Jakes, *Microwave Mobile Communications*. New York: Wiley, 1974.

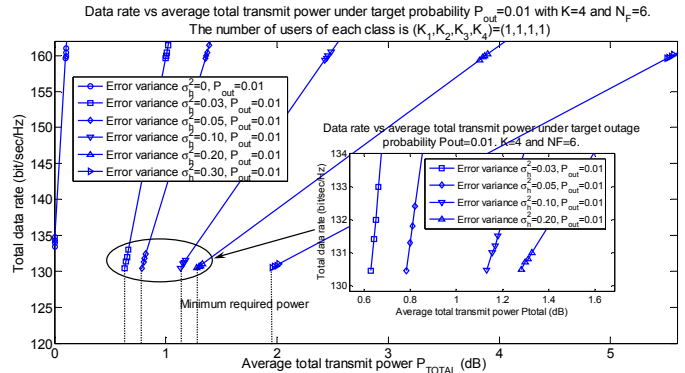[10] Slawomir Pietrzyk, *OFDMA for Broadband Wireless Access,* Artech House, London, 2006.

Figure 2: Total data rate vs. average total transmit power under probability $P_{out} = 0.01$ of a system consisted by $K = 4$, $N_F = 6$ and the classes of the system's users are $(K_1, K_2, K_3, K_4) = (1,1,1,1)$.
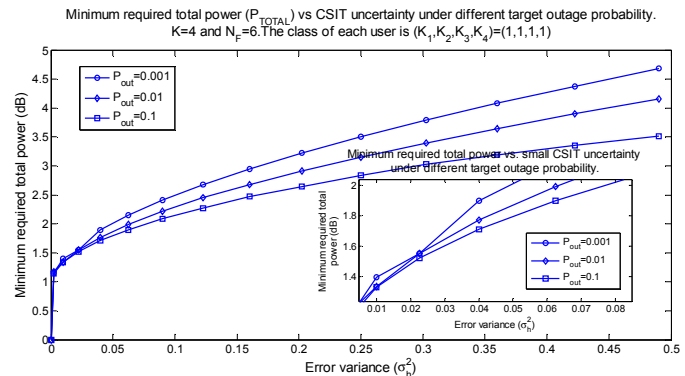

Figure 3: Minimum required total power vs. CSIT error variance σ_h^2 under different target outage probabilities $(P_{out} = 0.1, 0.01, 0.001)$ of a system consisted by $K = 4$, $N_F = 6$ and the classes of the system's users are $(K_1, K_2, K_3, K_4) = (1,1,1,1)$.

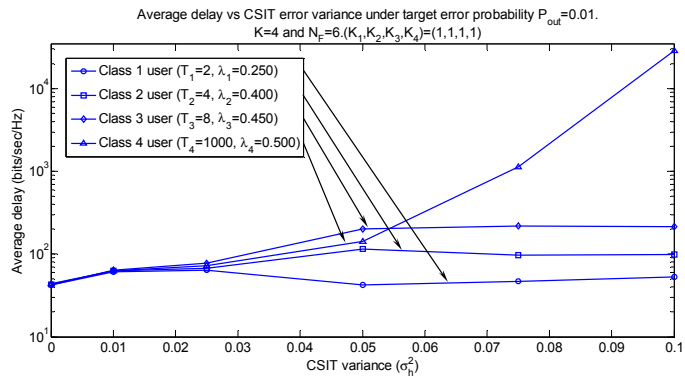

Figure 4: Average delay vs. CSIT error variance σ_h^2 under target outage probability $P_{out} = 0.01$ of a system consisted by $K = 4$, $N_F = 6$ and the classes of the system's users are $(K_1, K_2, K_3, K_4) = (1,1,1,1)$.