Abstract—In this paper we develop a time Scheduling, Subcarrier and Power Allocation scheme for Multi-Service Downlink OFDMA Systems. We propose a two step solution where time scheduling is separated from subcarrier and power allocation decisions. The system is modeled as a linear dynamic system with the aim of minimizing a quadratic cost function. The proposed scheduler is a Linear-Quadratic-Regulator (LQR) which achieves fairness among users by proposing an instantaneous data rate for each user in each time slot. The data rate vector proposed by the regulator is then fed as a constraint to the subcarrier and power allocation problem. This problem is formulated as a constrained convex optimization problem and we develop algorithms for subcarrier and power allocation to achieve the data rates proposed by LQR. Simulation results show good performance and better fairness among users.

I. INTRODUCTION

Quality-of-Service (QoS) guarantees and fairness among users is an important aspect of future resource allocation schemes. However time varying nature of the wireless channel, scarce bandwidth and limited power resources at the physical layer are the major challenges to achieving these objectives. Orthogonal Frequency Division Multiple Access (OFDMA) is a promising physical layer multiplexing technique which greatly aids in overcoming these challenges. This technique has several remarkable features such as the Inter Symbol Interference (ISI) mitigation in frequency selective fading channels, efficient exploitation of multiuser diversity, ability to multiplex several users on different subcarriers without interference, and enormous opportunities for dynamic resource allocation strategies [1]-[3]. Due to its superior performance, it has been has been adopted for both uplink and downlink air interfaces of WiMAX fixed and mobile standards, i.e IEEE802.16d and IEEE802.16e respectively [4], [5].

In this paper, we consider the problem of time scheduling, subcarrier and power allocation in multi-service downlink OFDMA system. The optimal solution to perform all these tasks simultaneously is not known in the literature. In [6]-[8] Proportional Fair (PF), Exponential Rule (Exp-rule) and Minimum Longest Weighted Delay First (MLWDF) schedulers are proposed. These schedulers do not perform power control and assume equal power allocation on all the subcarriers. Moreover PF, Exp-rule and MLWDF scheduling rules were developed for Time Division Multiple Access (TDMA) systems and their extension to OFDMA systems result in further sub-optimality in addition to the absence of power control.

In [9]-[11] this problem is formulated as a utility optimization problem. Utility functions of long term throughput, average queue length or average waiting time are used to optimize and allocate resources among different users. In this approach the choice of utility function plays a key role in the optimization process. However, appropriate utility functions which represent the level of user satisfaction according to their demanded service type are hard to find. Moreover since this approach was also developed for TDMA systems [12] without power control it is not yet known how to achieve a certain utility function in the presence of power control.

In this paper, we develop a two step solution to this joint scheduling and resource allocation problem. Since QoS demands of the users have to be achieved over multiple time slots while subcarrier and power allocation has to be performed during each time slot, therefore we can separate scheduling from resource allocation. Time scheduling problem then provides an instantaneous objective for the instantaneous subcarrier and power allocation problem. We model the dynamics of the multiuser system as a discrete time linear dynamic system. We develop the time scheduling problem as a fairness tracking control problem. This control problem is solved by a Linear-Quadratic-Regulator (LQR) [13]. The output of LQR is an instantaneous data rate for each user in each time slot. The data rate vector proposed by the LQR is then fed as a constraint to the resource allocation problem. This problem is formulated as a constrained sum-rate maximization problem and gives the optimal subcarrier and power allocation. The performance of our approach is then verified by simulations. (We use the terms “subcarrier and power allocation problem” and the “physical layer resource allocation problem” interchangeably.)

Throughout this paper we use uppercase boldface letters for matrices. The conjugate transpose of matrix $Q$ is denoted by $Q'$. $||\cdot||_Q$ denote the squared-Euclidean norm in an appropriate finite dimensional vector space weighted by matrix $Q$. $\text{diag}(\cdot)_K$ and $I_K$ represents the diagonal matrix and the Identity matrix of dimensions $K \times K$ respectively while $\Phi$ denotes an empty set.

The rest of the paper is organized as follows. In section II system model is described and the problem is formulated. The LQR is developed in section III. Subcarrier and Power allocation problem is discussed in section IV. Simulation results are presented in section V while the paper is concluded.
in section VI.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider an OFDMA system with $K$ users and $F$ subcarriers. We assume that the total transmit power is constrained to $P_{\text{max}}$. Time is divided into discrete slots and during each time slot a data frame consisting of $D$ OFDM symbols is transmitted. User channels remain constant for the duration of a slot but may change from one time slot to another. We assume that perfect Channel State Information (CSI) is available at the BS. The channel gain to noise ratio $g_{k,f}$ of user $k$ on subcarrier $f$ during time slot $t$ is given by, $g_{k,f} = \frac{|h_{k,f}|^2}{\sigma_n^2 B}$, where, $|h_{k,f}|$ denotes the channel coefficient of user $k$ on subcarrier $f$ after Fast Fourier Transform (FFT), $\sigma_n^2$ is the power spectral density (PSD) of white noise and $B$ denotes the bandwidth of a subcarrier. Each user maintains a separate queue at the BS which receives data from the higher layer. The queue length of user $k$ evolves according to,

$$q_{k}^{t+1} = (q_{k}^{t} + f_{k}^{t} - p_{k}^{t})^+$$

where, $(a)^+ = \max(a, 0)$. $f_{k}^{t}$ is the input arrival rate during time slot $t$ and $p_{k}^{t}$ is the actual departure rate from the queue after subcarrier and power allocation. We differentiate between different services according to their QoS requirements on delay e.g., voice transmission and video/audio streaming services are delay sensitive whereas email and file transfer services are delay tolerant. Let $\bar{d}_k$ be the average delay requirement of the service demanded by user $k$ and $\bar{f}_k$ be the average input arrival rate to $k$th user queue, then by Little’s Law, $\bar{q}_k = \bar{d}_k \bar{f}_k$ is the target queue length of user $k$. We can define the fairness level $a_k$ of each user in terms of its target queue length,

$$a_k = \frac{1}{\sum_{k=1}^{K} 1/\bar{q}_k}$$

We now develop an optimization problem to determine the subcarrier and power allocation to different users. Since we consider an OFDMA system therefore we assume that $I_k^f$ are the subcarriers allocated to user $k$ during time slot $t$. Therefore, the data rate achieved by user $k$ on its allocated subcarrier set $I_k^t$ during time slot $t$ according to Shannon’s Capacity formula is given as,

$$R_k^t = \sum_{f \in I_k^t} \log \left( 1 + \frac{p_{k,f}^t g_{k,f}}{\gamma_{k,f}} \right) \text{ nats/s/Hz}$$

where $p_{k,f}^t$ is the power allocated to user $k$ on subcarrier $f$ during time slot $t$. We notice that due to exclusive assignment of a subcarrier to single user this problem has a combinatorial nature. This can be averted by using the familiar notion of time sharing of each subcarrier by different users [1],[3]. We introduce a time sharing factor $\gamma_{k,f} \in [0, 1]$ for user $k$ on subcarrier $f$ during time slot $t$. Thus during a time slot user $k$ is allowed to transmit on subcarrier $f$ for $\gamma_{k,f} D$ OFDM symbols. This is possible from resource allocation point of view because we have assumed that channel remains constant in each time slot. This assumption introduces the following constraint,

$$\sum_{k=1}^{K} \gamma_{k,f}^t \leq 1 \quad \forall t, f$$

Moreover by defining, $o_{k,f}^t = \gamma_{k,f}^t p_{k,f}^t$ as the average power allocated to user $k$ on subcarrier $f$ during time slot $t$ we get,

$$R_k^t(o_{k,f}^t, \gamma_{k,f}^t, g_{k,f}^t) = \gamma_{k,f}^t \log \left( 1 + \frac{o_{k,f}^t g_{k,f}^t}{\gamma_{k,f}^t} \right)$$

Eq (5) represents a concave function since its Hessian is negative semi-definite when $\gamma_{k,f}^t \geq 0$ and $o_{k,f}^t \geq 0$. Now with $R_k^t(. , .)$ as defined in (5), we can write the optimization problem as follows,

$$\max \ E \left[ \sum_{k=1}^{K} \sum_{f=1}^{F} R_k^t(o_{k,f}^t, \gamma_{k,f}^t, g_{k,f}^t) \right]$$

subject to the following constraints,

$$\sum_{k=1}^{K} \sum_{f=1}^{F} o_{k,f}^t \leq P_{\text{max}}, \quad \forall t$$

$$\sum_{k=1}^{K} \gamma_{k,f}^t \leq 1 \quad \forall t, f$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} E \left[ \sum_{t=1}^{T} q_k^t \right] \leq \bar{q}_k \quad \forall k$$

The objective of problem (6) is the long term throughput maximization or system capacity. Constraints (7) and (8) are the instantaneous or per time slot constraints. Constraint (7) demands that the total transmit power in any time slot $t$ should always be less than the maximum power available at the BS. Constraints (8) is the time sharing constraint on the subcarriers. Constraint (9) is the long term constraint and is related to the QoS and fairness requirements of the users. This constraint demands that the long term average queue length should be less than the target queue length.

This problem has instantaneous and average constraints which make it difficult to solve. Since average target queue lengths $\bar{q}_k, \forall k$ can be achieved over multiple time slots we can split the problem into two sub-problems by proposing an instantaneous data rate for each queue in each time slot and then maximizing the instantaneous sum-rate. Thus, if these proposed data rates are achieved in each time slot then constraint (9) is ensured and if instantaneous sum-rate is maximized in each time slot then the long term objective in (6) is also guaranteed. A fairness tracking controller is developed which propose a data rate $\hat{R}_k^t$ for each user according to the changing wireless channel conditions and the QoS constraints. The instantaneous subcarrier and power allocation decisions are made by solving a constrained instantaneous sum-rate maximization problem. In this problem, the proposed data rate vector by the controller is an additional constraint along with
The instantaneous sum-rate problem is as follows,

$$\max_{\alpha_{k,f}, \gamma_{k,f}, \mu} \sum_{k=1}^{K} \sum_{f=1}^{F} R_{k,f}^{t} (\alpha_{k,f}, \gamma_{k,f}, \mu_{k,f})$$

subject to,

$$\sum_{k=1}^{K} \sum_{f=1}^{F} \alpha_{k,f} \leq P_{\text{max}}$$

$$\sum_{k=1}^{K} \gamma_{k,f} \leq 1, \quad \forall f$$

$$\sum_{f=1}^{F} R_{k,f}^{t} (\alpha_{k,f}, \gamma_{k,f}, \mu_{k,f}) \geq \tilde{R}_{k}^{t}, \quad \forall k$$

(13) is the instantaneous data rate constraint. If the data rates achieved by the resource allocation algorithm are equal to or greater then the proposed data rates \( \tilde{R}_{k} \), \( \forall k \), then fairness constraint is satisfied. However if in any time slot these data rates cannot be achieved due to bad channel conditions and power limitations then the error is fed to the controller which compensates for this loss according to the deviations from the target queue length. Hence time diversity in the wireless channel is utilized and constraint (9) is achieved in the longer run by tracking the controller decisions. Moreover, since we are maximizing the instantaneous sum-rate in (10) therefore the long term objective in (6) is also maximized which justifies the two step approach. The two step approach is detailed in Fig. 1. LQR (which will be explained later) derives an output data rate \( \tilde{R}_{k} \) for each user at each time slot \( t \). This data rate is fed to the resource allocation block for subcarrier and power assignment. A set of subcarriers \( I_{k} \) is allocated to user \( k \) and a data rate vector \( \pi = \{ \mu_{k} \}, \forall k \} \) is achieved according to the physical channel conditions. The allocation decisions are sent to the users via separate control channels which allow the users to recover their data. The achieved data rate vector \( \pi \) is sent back to the controller in order to track its decisions according to the changing wireless conditions. In the next section we develop the Fairness Tracking Control Problem and propose our LQR.

### III. Linear-Quadratic-Regulator (LQR)

In this section we develop the LQR problem to derive an instantaneous data rate constraint for the resource allocation problem. Since the actual achieved data rates \( \mu_{k} \) differ from the data rates \( \tilde{R}_{k} \) proposed by the controller, we can write the queue dynamics in (1) as,

$$q_{k}^{t+1} = (q_{k}^{t} + f_{k} - \tilde{R}_{k}^{t} + [\tilde{R}_{k}^{t} - \mu_{k}^{t}])^{+}$$

where, \( [\tilde{R}_{k}^{t} - \mu_{k}^{t}] \) can be seen as the tracking error which represents the difference between the proposed/controlled rate and the actual achieved rate after the subcarrier and power allocation. We define a control \( u_{k}^{t} \) which directly incorporates the fairness,

$$u_{k}^{t} = (\alpha_{k}^{t} - a_{k})C$$

where, \( \alpha_{k}^{t} \) is the instantaneous fairness level of user \( k \) at time \( t \), \( C \) is the average capacity of the underlying physical channel and \( a_{k} \) is the target fairness level of each user \( k \). Let, \( R_{k}^{t} = a_{k}^{t}C \) then (14) becomes,

$$q_{k}^{t+1} = (q_{k}^{t} + f_{k} - a_{k}C + [\tilde{R}_{k}^{t} - \mu_{k}^{t}])^{+}$$

Let, \( \lambda_{k}^{t} = f_{k} - a_{k}C + [\tilde{R}_{k}^{t} - \mu_{k}^{t}] \), then,

$$q_{k}^{t+1} = (q_{k}^{t} - u_{k}^{t} + \lambda_{k}^{t})^{+}$$

\( \lambda_{k}^{t} \) can be seen as the disturbance in the system dynamics. It contains the combined effect of tracking error as well as the variations in the input arrival rate. In order to incorporate QoS requirements of the users according to their target queue lengths we define a new state variable,

$$x_{k}^{t} = q_{k}^{t} - \bar{q}_{k}$$

\( x_{k}^{t} \) measures the instantaneous deviation of user queue length from its target queue length. In terms of state variable \( x_{k}^{t} \), (17) can now be written as,

$$x_{k}^{t+1} = x_{k}^{t} - u_{k}^{t} + \lambda_{k}^{t}$$

Since there are \( K \) users in the system, we define, \( X^{t} = [x_{1}^{t}, \ldots, x_{K}^{t}] \), \( U^{t} = [u_{1}^{t}, \ldots, u_{K}^{t}] \) and \( \Gamma^{t} = [\lambda_{1}^{t}, \ldots, \lambda_{K}^{t}] \) and write the system state as,

$$X_{k}^{t+1} = AX^{t} + BU^{t} + DI^{t}$$

where, we define the following matrices, \( A = I_{K} \), \( B = -I_{K} \) and \( D = I_{K} \). Equation (20) represents a discrete time linear dynamic system. In order to ensure the fairness constraint (9) and to reduce the tracking error we define the following quadratic cost function,

$$J = \lim_{T \to \infty} \frac{1}{T} E \left[ \sum_{t=1}^{T} \left( ||X^{t}||_{W}^{2} + ||U^{t}||_{V}^{2} \right) \right]$$
where, $V = \text{diag}(\alpha, \ldots, \alpha)_K$ and $W = I_K$ are the weighting matrices. Moreover, we define $\alpha = \exp\{\rho \sum_{k=1}^K (R_k^t - \mu_k^t)\}$, where $\rho$ is a scaling constant. There are two terms in the cost function, i) the first term represents the penalty for deviating from the target queue length. Since target queue length depends on the service type therefore, this term directly affects the QoS requirements of the users and ii) the second term contains the variable $\alpha$ in the weighting matrix which depends on the scaled difference between the demanded and the achieved data rates by the users. This term tracks the error in the controller decisions according to the changing wireless conditions at the physical layer. If the physical layer capacity is less than the sum of demanded data rates by the controller, $\alpha$ increases in the cost function. Therefore, the control action $U^t$ decrease the demanded data rates. Similarly, when the actual physical layer capacity is higher than the demanded data rates, then $\alpha$ decreases. In this case, $U^t$ can either increase or decrease depending on the queue lengths. In other words, since enough physical layer capacity is available, the control $U^t$ can increase or decrease the data rates of the users depending on the QoS constraints and the queue lengths of the users.

The control problem with discrete time linear dynamic system state (20) and quadratic cost (21) can be converted into a standard LQR problem by absorbing the disturbance term $\Gamma t$ in the state equation. We define a new state variable $Z^t$,

$$Z^t = \begin{pmatrix} X^t \\ 1 \end{pmatrix}$$

Since the dimensions of the system have increased, therefore, the state equation (20) and cost function (21) can now be written as,

$$Z^{t+1} = A_1 Z^t + B_1 U^t$$

$$J = \lim_{T \to \infty} \frac{1}{T} E \left[ \sum_{t=1}^T \left( \|Z^t\|_W^2 + \|U^t\|_V^2 \right) \right]$$

where,

$$A_1 = \begin{pmatrix} A & \Gamma t \\ 0 & 1 \end{pmatrix}, \quad B_1 = \begin{pmatrix} B \\ 0 \end{pmatrix}$$

$$W_1 = \begin{pmatrix} W & 0 \\ 0 & 0 \end{pmatrix}, \quad V_1 = \begin{pmatrix} V \\ 0 \end{pmatrix}$$

This is a standard discrete time linear quadratic regulator problem, which admits the unique solution [13],

$$U^t = -F Z^t$$

where, the feedback control gain is, $F = (V_1 + B_1 S B_1) A_1$ and $S$ is the solution to the discrete time algebraic Riccati equation,

$$S = W_1 + A_1 S - S B_1 (V_1 + B_1 S B_1)$$

(28)

$S$ is a symmetric matrix and let us define,

$$S = \begin{pmatrix} s_1 & s_2 & s_3 \\ s_2 & s_4 & s_5 \\ s_3 & s_5 & s_6 \end{pmatrix}$$

(29)

where, $s_1, s_2$ and $s_3$ are matrices of appropriate dimensions. By using equations (28) and (29), it is easy to see that we have following equations for $s_1$ and $s_2$ respectively,

$$s_1 = W + A^t \left[ s_1 - s_1 B (V + B s_1 B)^{-1} B s_1 A \right]$$

(30)

$$s_2 = A^t \left[ (s_1 - s_1 B (V + B s_1 B)^{-1} B s_1) \right]$$

(31)

and feedback control gain $F$ and the control $U^t$ becomes,

$$F = (V + B s_1 B)^{-1} B s_1 A B s_1 \Gamma^t + B' s_2$$

(32)

$$U^t = -(V + B s_1 B)^{-1} B' s_1 X^t + F s_2$$

(33)

(34)

Equation (35) represents a state feedback controller. Since at any given time $t$, the state $X^t$ and $\Gamma^t$ are not available to the controller therefore we use Kalman filter for the estimation. Finally, the data rate $R_k^t$ proposed by the controller for user $k$ at time $t$ is,

$$\tilde{R}_k^t = u_k^t + a_k C$$

(36)

In the next step these data rates are fed to the instantaneous resource allocation problem (10).

IV. SUBCARRIER AND POWER ALLOCATION

In this section we discuss the instantaneous subcarrier and power allocation algorithm by solving the constrained optimization problem (10). This is a convex optimization problem with linear and convex differentiable constraints we can solve this problem by using Convex Optimization theory [14]. Let, $\eta$, $\psi^t_f$, $P_{max}$ be the Lagrange multipliers associated with constraints (11), (12) and (13) respectively in problem (10). The Lagrangian of the problem is,

$$L \{ \eta, \psi^t_f, \gamma^t_k, \delta_k, P_{max} \} = \sum_{k=1}^K \sum_{f=1}^F \gamma^t_{k,f} \log \left( 1 + \frac{\delta^t_{k,f} g^t_{k,f}}{\gamma^t_{k,f}} \right)$$

(37)

Since the objective and the constraint functions are convex differentiable, therefore, the duality gap is zero. The solution can be obtained by solving the Karush-Kuhn-Tucker (KKT)
conditions which can be obtained by setting \( \frac{\partial C(\bar{x})}{\partial x_{k}} = 0 \) and \( \frac{\partial C(\bar{x})}{\partial y_{k}} = 0 \). From these conditions we get,

\[
p^t_{k,f} = \left( \frac{1 + \delta_k}{\eta} - \frac{1}{y_{k,f}} \right)^+ 
\]

(38)

\[
(1+\delta_k) \left( \log \left( \frac{(1+\delta_k)y_{k,f}}{\eta} \right)^+ - \left( 1 - \frac{\eta}{(1+\delta_k)y_{k,f}} \right)^+ \right) = \psi_f
\]

(39)

Based on the equations (38) and (39), efficient resource allocation algorithms can be developed. Since the physical layer resource allocation problem in this paper is same as that considered in [15]. Therefore, we can use the algorithms developed in [15] for subcarrier and power allocation. We use the optimal algorithm developed in [15]. This algorithm achieves the rate and power constraints simultaneously. This algorithm consists of an inner loop and an outer loop. The outer loop starts with a small value of \( \eta > 0 \) and increments it in small steps. For each value of \( \eta \), the inner loop allocates the power and subcarriers to the users according to their data rate constraints. This process is repeated till all the constraints are satisfied. We produce this algorithm from [15] in Table I.

### TABLE I

**PHYSICAL LAYER RESOURCE ALLOCATION ALGORITHM**

<table>
<thead>
<tr>
<th>1) Initialization:</th>
<th>( \eta = 0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outer loop:</td>
<td>( \eta = \eta + \Delta \eta )</td>
</tr>
<tr>
<td>Inner loop:</td>
<td>( \delta_k = \min_{\bar{x}<em>{k,f}} \frac{1}{\bar{x}</em>{k,f}} ), ( \forall k, \phi_k, f = 0 ), ( \forall k, f )</td>
</tr>
<tr>
<td>1) Initialization:</td>
<td>( \gamma_{k,f} = 0 ), ( \forall k, f ), ( \Gamma = { \Gamma_1, \ldots, \Gamma_K } = 0 )</td>
</tr>
<tr>
<td>2) Repeat till all the rate constraints are achieved.</td>
<td></td>
</tr>
<tr>
<td>3a) Repeat till ( k^{th} ) user rate constraint is achieved.</td>
<td></td>
</tr>
<tr>
<td>3b) Increase waterlevel of user ( k ) in small steps, ( \delta_k = \delta_k + \Delta \eta ).</td>
<td></td>
</tr>
<tr>
<td>3c) On all the subcarriers compute, ( \phi_{k,f} = \frac{1}{\eta} )</td>
<td></td>
</tr>
<tr>
<td>3d) Allocate subcarrier to this user if ( \phi_{k,f} ) is maximum and set ( \gamma_{f,f} = 1 ) otherwise set ( \gamma_{f,f} = 0 ).</td>
<td></td>
</tr>
<tr>
<td>4) Compute the achieved data rate according to,</td>
<td></td>
</tr>
<tr>
<td>( \Gamma_k = \sum_{f=1}^{F} \gamma_{k,f} \left( \log \left( \frac{(1+\delta_k)y_{f,k}}{\eta} \right)^+ \right) )</td>
<td></td>
</tr>
<tr>
<td>5) Compute the total power,</td>
<td></td>
</tr>
<tr>
<td>( P_{total} = \sum_{k=1}^{K} \sum_{f=1}^{F} \gamma_{k,f} \left( \frac{1+\delta_k}{\eta} - \frac{1}{y_{k,f}} \right)^+ )</td>
<td></td>
</tr>
</tbody>
</table>

A. Non-Feasibility

The problem in (10) is considered feasible if all the minimum rate constraints (13) can be satisfied within the given amount of power. The non-feasibility in the problem can be detected from the sum-power minimization algorithm developed in [15]. This algorithm gives us a method to detect the non-feasibility in the problem but does not tell us how to resolve it. In case of non-feasibility there is not enough instantaneous capacity at the physical layer hence the minimum rate constraints proposed by the controller cannot be achieved for all the users. Therefore, we have to decrease the data rates of some of the users. Since state variable \( x' \) measures the deviation of current queue length of the user from its target queue length, therefore, we decrease the demanded data rate of the user with minimum value of \( x' \). We adopt this criterion to further ensure fairness among users in terms of achieved data rates when there is non-feasibility in the problem. In case, there are more than one users with the same minimum value of \( x' \), we select the user \( k' \) which consumes the maximum amount of power to achieve its data rate. Thus decreasing the data rate of user \( k' \) will result in maximum performance improvement. This algorithm is developed below.

1. Find the user \( f \) with minimum value of \( x' \) and denote this set by \( \Pi \).
2. Select the user \( k' \) such that \( k' = \max_{k \in \Pi} \frac{P_{k}^t}{P_{f}^t} \).
3. Decrease the proposed data rate \( \tilde{R}_k^t \) of user \( k' \) by an amount \( \nu \).
4. Use the algorithm in Table I to achieve the new rate constraints.
5. Repeat this process till the problem becomes feasible and convergence is achieved.

The number of iterations of this algorithm depends on the optimal step-size \( \nu \). If this value is optimally determined then the algorithm converges relatively quickly in few iterations.

### V. SIMULATION RESULTS

We consider a downlink OFDMA system with 10 users and 24 subcarriers. We assume that the bandwidth of each subcarrier is 375 KHz and BS has a peak power constraint of 43 dBm. We consider a frequency selective Rayleigh fading channel with exponential delay profile. The power spectral density of noise is -174 dBm/Hz. Time is divided into slots and duration of each Transmission Time Interval (TTI) is 1ms. We assume that packets are generated according to poisson distribution with packet size of 1kbits. The users are uniformly distributed in a cell of radius 500m. Path losses are calculated according to Cost-Hata Model [16]. The users experience different path loss due to varying distances from the BS which greatly impact their ability to achieve different rates. In all the simulations we assume that user 1 is closest to the BS and is located at a distance of 50m. Similarly user 10 is the worst user and is located at the cell edge in all the simulations.

Since we have dynamic power control in our resource allocation algorithm hence to compare the performance of our controller which is acting here as a scheduler we need algorithms with power control. We compare the performance of our approach with PF and M-LWDF schedulers. Dynamic power control can be achieved by calculating the scheduling weights of PF and M-LWDF schedulers and then using a weighted sum rate maximization algorithm. This gives us PF and M-LWDF schedulers with power control which then gives us enough base for comparison with our controller. In Figure 2 we assume that the average demanded delay constraint of each user is 10 TTI and we plot the mean delay achieved by all the users in the systems versus the input arrival rate. As the input arrival rates increase the average delay achieved by the system also increases for all three schemes. The performance
of PF scheduler is the worst because it is not designed to handle delay constrained traffic. The performance of M-LWDF is better than the PF scheduler but it can only guarantee the mean demanded delay till 6 Packets/TTI/User and for input arrival rates higher than this value the achieved delay is greater than the demanded delay. It is evident from the figure that our algorithm outperforms both these algorithms and can guarantee the mean delay for input arrival rates greater than 8 Packets/TTI/User. Since we have incorporated power control in the PF and M-LWDF scheduling rules we can safely say that this superior performance results from the use of our LQR along with the dynamic resource allocation algorithm.

In order to see the impact on the achieved delays of different users which are located at varying distances from the BS we plot figure 3. In this figure we compare the performance of the best and the worst user in the system by plotting the difference between the mean delays achieved by these two users. The average achieved delay of the worst user is always greater than that for the best user due to the difference between the attenuations caused by the corresponding path losses. However, this difference is not huge and is almost 5.5 ms for the plotted input arrival rates. The superior performance of our approach compared to the M-LWDF algorithm is due to the fact that the rates proposed by the LQR are according to the delays experienced by the individual users in the system. Thus the worst user is compensated by the controller and a higher data rate is proposed for this user compared to the best user in most of the time slots. Moreover, if there is not enough capacity at the physical layer to satisfy the demanded rates then again in the non-feasible algorithm the worst user is given preference over the best user. The demand of higher data rate for the worst user is taken care of by the algorithm is given preference over the best user. The demand of higher data rate then again in the non-feasible algorithm the worst user higher data rate is proposed for this user compared to the best user. Thus the worst user is compensated by the controller and a higher data rate is proposed for this user compared to the best user in most of the time slots. Moreover, if there is not enough capacity at the physical layer to satisfy the demanded rates then again in the non-feasible algorithm the worst user is given preference over the best user. The demand of higher data rate for the worst user is taken care of by the algorithm is given preference over the best user. The demand of higher data rate then again in the non-feasible algorithm the worst user higher data rate is proposed for this user compared to the best user. Thus the worst user is compensated by the controller and a higher data rate is proposed for this user compared to the best user in most of the time slots. Moreover, if there is not enough capacity at the physical layer to satisfy the demanded rates then again in the non-feasible algorithm the worst user is given preference over the best user. The demand of higher data rate for the worst user is taken care of by the algorithm is given preference over the best user. The demand of higher data rate then again in the non-feasible algorithm the worst user higher data rate is proposed for this user compared to the best user. Thus the worst user is compensated by the controller and a higher data rate is proposed for this user compared to the best user in most of the time slots. Moreover, if there is not enough capacity at the physical layer to satisfy the demanded rates then again in the non-feasible algorithm the worst user is given preference over the best user. The demand of higher data rate for the worst user is taken care of by the algorithm.

Fig. 2. Average Achieved Delay for all the users in the system vs Average Input Arrival Rate. Target delay deadline = 10 TTI

VI. CONCLUSION

We developed a two step solution to joint scheduling and resource allocation problem in multi-service OFDMA systems. An LQR is proposed which derives an instantaneous data rate for each user depending on its demanded service type. Subcarrier and power allocation decisions are made based on these proposed data rates. We discussed the feasibility of the problem and developed efficient algorithms. Simulation results show good performance and better fairness among users.

REFERENCES