Cross-Layer Scheduling for OFDMA Amplify-and-Forward Relay Networks

Derrick Wing Kwan Ng and Robert Schober
Department of Electrical and Computer Engineering, University of British Columbia, Canada
Email: wingn@ece.ubc.ca, rschober@ece.ubc.ca

***Abstract*— In this paper, we consider cross-layer scheduling for the downlink of amplify-and-forward (AF) relay assisted orthogonal frequency division multiple access (OFDMA) networks. The proposed cross-layer design takes into account the effects of imperfect channel state information at transmitter (CSIT) in slow fading. The rate adaptation, power adaptation, and subcarrier allocation policies are optimized to maximize the system goodput (bits/s/Hz successfully received by the mobiles). The optimization problem is solved by using dual decomposition resulting in a highly scalable distributed resource allocation algorithm. Simulation results illustrate that the proposed distributed cross-layer scheduler requires only a small number of iterations to achieves practically the same performance as the optimal centralized scheduler.**

I. Introduction

Orthogonal frequency division multiple access (OFDMA) is a promising candidate for high speed wireless communication networks, such as WiMAX and future Fourth Generation (4G) systems [1]. On the other hand, cooperative relaying is an attractive technique to increase the range of communication systems without incurring the high costs of additional base station deployment. Different relaying strategies such as amplify-and-forward (AF), compress-and-forward (CF), and decode-and-forward (DF) have been proposed in the literature. AF is particularly appealing as the relays only amplify and linearly process the received signal which leads to low-complexity transceiver designs. More importantly, AF relays are transparent to the adaptive modulation techniques that are typically employed at the base station. For these reasons, AF was selected as one of the possible relaying modes in IEEE 802.16J (Mobile Multihop Relay).

In recent years, there has been a growing interest in combing the concept of relaying with OFDMA/OFDM to enhance wireless system performance. In the existing literature, e.g. [2], [3], perfect global channel state information (CSI) of all links is assumed to be available at the base station and the relays such that resource allocation can be done optimally. However, in practice, perfect channel state information at the transmitter (CSIT) is difficult to obtain for the relay-to-users links due to the mobility of the users. Besides, existing works such as [4], [5] focus on centralized scheduling at the base station. In this case, the computational complexity at the base station increases exponentially with the number of users/relays and subcarriers, and the overhead for CSI feedback becomes significant which limits the scalability of the system in practice. Therefore, distributed scheduling algorithms which take into account imperfect CSIT and converge fast to the optimal solution are needed for practical implementation.

In this paper, we formulate the scheduling problem in AF relay assisted OFDMA systems as an optimization problem. By using dual decomposition, the problem is separated into a master problem and several subproblems. Each relay solves its own subproblem by utilizing its local CSI without any help from other relays while the base station updates the dual variables through the concept of pricing. Therefore, the computational complexity at the base station and the CSI feedback overhead are both significantly reduced compared to optimal centralized scheduling.

The rest of the paper is organized as follows. In Section II, we outline the model for the considered AF relay assisted OFDMA system. In Section III, we formulate the resource allocation as an optimization problem and solve it by dual decomposition. Section IV presents numerical performance results for the distributed algorithm. In Section V, we conclude with a brief summary of our results.

II. Relay assisted OFDMA network model

We consider a relay assisted OFDMA downlink packet transmission network which consists of one base station (BS), K mobile users, and M relays. All transceivers have single antennas. The cell coverage is divided into M areas corresponding to the M relays and each user is assigned to a relay according to the corresponding value of path loss. We assume that there is no direct transmission between the base station and mobile users due to heavy shadowing. A time-division channel allocation with two time slots is used to facilitate orthogonal relaying [6]. In the first time slot, the base station broadcasts its signal to the relays. Then, in the second time slot, the relays amplify the previously received signal and forward it to the associated users.

A. Channel Model

The channel impulse response has W taps and is time-invariant (slow fading) within the scheduling slot. Assuming an OFDMA system with n_F subcarriers, in the first time slot, the (frequency domain) received symbol in subcarrier $i \in \{1, \dots, n_F\}$ at relay $m \in \{1, \dots, M\}$ for user $k \in \{1, \dots, K\}$ is given by

$$Y_{SR_m,i}^{(k)} = \sqrt{P_{SR_m,i}^{(k)} l_{SR_m}} H_{SR_m,i} X_i^{(k)} + Z_{SR_m,i} \quad (1)$$

where $P_{SR_m,i}^{(k)}$ is the transmit signal-to-noise ratio (SNR) for user k in subcarrier i for the link between the base station and relay m. $X_i^{(k)}$ is the symbol transmitted to user k on subcarrier i in the first time slot. In practice, both the BS and the relays are placed in relatively high positions and hence the number of

978-1-4244-2515-0/09/$25.00 ©2009 IEEE

blockages or scatterers between them are limited and a strong line of sight is expected. Hence, the channel fading coefficients between the BS and relay m in subcarrier i, $H_{SR_m,i}$, are modeled as Rician fading with Rician factor κ, i.e., $H_{SR_m,i} \sim \mathcal{CN}(\sqrt{\kappa/(1+\kappa)}, 1/(1+\kappa))$. l_{SR_m} represents the path loss between the base station and relay m, and $Z_{SR_m,i}$ is additive white Gaussian noise (AWGN) with distribution $\mathcal{CN}(0, N_0)$ in subcarrier i at relay m. Here, $\mathcal{CN}(\mu, \sigma^2)$ denotes a complex Gaussian random variable with mean μ and variance σ^2. In order to simplify the subsequent mathematical expressions and without loss of generality, we assume in the following a normalized noise variance of $N_0 = 1$ at all relay and user stations. To improve performance, the signal received at relay m in subcarrier i in the first time slot is mapped to subcarrier $j \in \{1, \ldots, n_F\}$ in the second time slot [7]. Furthermore, the received signal is multiplied by a gain $G^{(k)}_{RD_m,j}$ and forwarded to the destination. The signal received at user k in subcarrier j from relay m is given by

$$Y^{(k)}_{RD_m,j} = G^{(k)}_{RD_m,j}\sqrt{P^{(k)}_{RD_m,j} l^{(k)}_{RD_m}}\, H^{(k)}_{RD_m,j} Y^{(k)}_{SR_m,i} + Z^{(k)}_j \quad (2)$$

where variables $P^{(k)}_{RD_m,j}$, $l^{(k)}_{RD_m}$, $H^{(k)}_{RD_m,j}$, and $Z^{(k)}_j$ are defined in a similar manner as the corresponding variables for the base station-to-relay links. Since the users are generally surrounded by a large number of scatterers, we model the small scale fading coefficients between relay m and user k as Rayleigh distributed, i.e., $H^{(k)}_{RD_m,j} \sim \mathcal{CN}(0, 1)$. Following [6], the gain is chosen as

$$|G^{(k)}_{RD_m,j}|^2 = 1/(1 + P^{(k)}_{SR_m,i} l_{SR_m} |H_{SR_m,i}|^2). \quad (3)$$

Assuming that all transceivers have the same noise variance and operate at high SNR, based on (1)–(3) the equivalent SNR after two hops at user k can be approximated as[1] $\Gamma^{(k)}_{eq_m,i,j} \approx$

$$\frac{P^{(k)}_{SR_m,i} l_{SR_m} |H_{SR_m,i}|^2 P^{(k)}_{RD_m,j} l^{(k)}_{RD_m} |H^{(k)}_{RD_m,j}|^2}{P^{(k)}_{SR_m,i} l_{SR_m} |H_{SR_m,i}|^2 + P^{(k)}_{RD_m,j} l^{(k)}_{RD_m} |H^{(k)}_{RD_m,j}|^2} \quad (4)$$

B. Channel State Information

We assume that the local path loss information is perfectly known in all devices due to accurate long term measurements. Furthermore, since we assume that both the base station and the relays are static, the associated channel is time-invariant and can be accurately estimated. Therefore, we can assume perfect CSIT for the BS–to–relay links. For the relay–to–user link, we also assume that the user has perfect CSI. However, due to the mobility of the users and the duplexing delay, this CSI may be outdated when it reaches the relays. To capture this phenomenon, we model the estimated CSIT (at the relay) for the link between relay m and user k in subcarrier j as

$$\hat{H}^{(k)}_{RD_m,j} = H^{(k)}_{RD_m,j} + \Delta H^{(k)}_{RD_m,j} \quad (5)$$

where $H^{(k)}_{RD_m,j} \sim \mathcal{CN}(0, 1)$ is the actual CSI and $\Delta H^{(k)}_{RD_m,j} \sim \mathcal{CN}(0, \sigma_e^2)$ represents the CSIT error with variance σ_e^2.

[1]The assumption of high SNR is necessary since otherwise the optimization problem in (11) is not convex.

III. Cross-Layer Design

A. Instantaneous Mutual Information and System Goodput

Given perfect CSI at the receiver, the instantaneous mutual information between the base station and user k using subcarrier pair (i, j) in the first and second time slots via relay m can be approximated for high SNR as

$$C^{(k)}_{m,i,j} \approx \frac{1}{2}\log_2(1 + \Gamma^{(k)}_{eq_m,i,j}). \quad (6)$$

In slow fading with imperfect CSIT, a *packet outage* occurs whenever the data rate $r^{(k)}_{m,i,j}$ exceeds the instantaneous mutual information even if channel capacity achieving coding is applied for error protection. This is because, due to the imperfect CSIT, the instantaneous mutual information is not known at the relays and the base station. In order to model the errors due to *packet outage*, we consider the performance in terms of the system goodput rather than the ergodic capacity. Define $\mathcal{U}_m$ as the set of users associated with relay m. The *weighted average system goodput* is defined as the total average bits/s/Hz successfully delivered to the K mobile stations through the M relays (averaged over multiple scheduling slots) and is given as $U_{goodput}(\mathcal{P}, \mathcal{R}, \mathcal{S}) =$

$$\frac{1}{n_F} E_{\mathbf{H}_{SR_m}, \hat{\mathbf{H}}_{RD_m}} \Bigg\{ \sum_{m=1}^{M} \sum_{k\in\mathcal{U}_m} w_k \sum_{i=1}^{n_F}\sum_{j=1}^{n_F} s^{(k)}_{m,i,j} r^{(k)}_{m,i,j} \times \Pr[r^{(k)}_{m,i,j} \le C^{(k)}_{m,i,j} | \mathbf{H}_{SR_m}, \hat{\mathbf{H}}_{RD_m}, \mathbf{L}_m] \Bigg\} \quad (7)$$

where $E_X\{\cdot\}$ denotes statistical expectation with respect to random variable X, and $\mathcal{P}, \mathcal{R}$, and $\mathcal{S}$ are the power, rate, and subcarrier allocation policies, respectively. $s^{(k)}_{m,i,j} \in \{0, 1\}$ is the subcarrier allocation indicator. w_k is a positive constant to prioritize different users and to enforce certain notions of fairness. Vectors $\hat{\mathbf{H}}_{RD_m}$, $\mathbf{H}_{SR_m}$, and $\mathbf{L}_m$ contain the estimated CSIT, $\hat{H}^{(k)}_{RD_m,j}$, for all links from relay m to users $k \in \mathcal{U}_m$, the actual CSIT, $H_{SR_m,i}$, for the link between the base station and relay m, and the path loss for all links involving relay m, respectively.

B. Optimization Problem Formulation

The optimal power allocation policy, $\mathcal{P}^*$, rate allocation policy, $\mathcal{R}^*$, and subcarrier allocation policy, $\mathcal{S}^*$, are given by

$$\begin{aligned}
&(\mathcal{P}^*, \mathcal{R}^*, \mathcal{S}^*) = \arg\max_{\mathcal{P},\mathcal{R},\mathcal{S}} U_{goodput}(\mathcal{P}, \mathcal{R}, \mathcal{S})\\
\text{s.t.}\quad &\text{C1: } \Pr[r^{(k)}_{m,i,j} \le C^{(k)}_{m,i,j} | \mathbf{H}_{SR_m}, \hat{\mathbf{H}}_{RD_m}, \mathbf{L}_m] = \varepsilon, \ \forall k\\
&\text{C2: } \sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{j=1}^{n_F} s^{(k)}_{m,i,j}(P^{(k)}_{RD_m,j} + P^{(k)}_{SR_m,i}) \le P_{\mathrm{t}}\\
&\text{C3: } \sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F} s^{(k)}_{m,i,j} = 1, \qquad \forall j\\
&\text{C4: } \sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{j=1}^{n_F} s^{(k)}_{m,i,j} = 1, \qquad \forall i\\
&\text{C5: } P^{(k)}_{SR_m,i}, P^{(k)}_{RD_m,j} \ge 0, \qquad \forall m, i, j, k\\
&\text{C6: } s^{(k)}_{m,i,j} \in \{0, 1\}, \quad \forall m, i, j, k
\end{aligned} \quad (8)$$

In C1, ε denotes the required outage probability for the transmission to user k, i.e., C1 represents a quality of service (QoS) metric. C2 is a joint power constraint for the base station and the relays with total power P_{t}. Constraints C3, C4, and C6 are imposed to guarantee that each subcarrier will be used at most once in each time slot.

C. Transformation of the Optimization Problem

The first step to simplify optimization problem (8) is to incorporate the outage probability C1 in the objective function. We introduce the following Lemma.

Lemma 1 (Equivalent rate constraint): For a given outage probability ε in C1, the equivalent data rate at high SNR is given by

$$\text{C1} \Rightarrow r_{m,i,j}^{(k)} = \frac{1}{2}\log_2(1+\Lambda_{eq_m,i,j}^{(k)}) \tag{9}$$

with $\Lambda_{eq_m,i,j}^{(k)} =$

$$\frac{P_{SR_m,i}^{(k)} l_{SR_m} |H_{SR_m,i}|^2 P_{RD_m,j}^{(k)} l_{RD_m}^{(k)} F_{RD_m,j}^{-1(k)}(\varepsilon)}{P_{SR_m,i}^{(k)} l_{SR_m} |H_{SR_m,i}|^2 + P_{RD_m,j}^{(k)} l_{RD_m}^{(k)} F_{RD_m,j}^{-1(k)}(\varepsilon)} \tag{10}$$

where $F_{RD_m,j}^{-1(k)}(\varepsilon)$ denotes the inverse cumulative distribution function (cdf) of a non-central chi-square random variable with 2 degrees of freedom and non-centrality parameter $|\hat{H}_{RD_m,j}^{(k)}|^2/\sigma_e^2$.

Proof 1: Due to page limitation, only a sketch of the proof is provided. We start by calculating the outage probability in C1 in (8). Since CSI is available at the scheduler, the only randomness is due to the imperfect CSIT of the relay-to-user link. Thus, the outage probability can be written as a function of the cdf of a non-central chi-square random variable, with 2 degrees of freedom and non-centrality parameter $|\hat{H}_{RD_m,j}^{(k)}|^2/\sigma_e^2$. Then by equating the cdf with the required outage probability ε and solving for $r_{m,i,j}^{(k)}$, (9) follows immediately. □

By substituting (9) into (7) a modified objective function is obtained and the cross-layer scheduling problem becomes an $\mathcal{NP}$-hard mixed combinatorial and convex optimization problem. Therefore, we follow the approach in [8] and relax constraint C6 in (7). In particular, we allow $s_{m,i,j}^{(k)}$ to assume any real value between zero and one. Then, $s_{m,i,j}^{(k)}$ can be interpreted as a time sharing factor for the K users for utilizing subcarrier pair (i,j) through relay m. Therefore, using the modified objective function in Lemma 1, the auxiliary powers $\tilde{P}_{SR_m,i}^{(k)} = P_{SR_m,i}^{(k)} s_{m,i,j}^{(k)}$ and $\tilde{P}_{RD_m,j}^{(k)} = P_{RD_m,j}^{(k)} s_{m,i,j}^{(k)}$, and the continuous relaxation of C6, we can rewrite problem (8) as

$$\begin{aligned}
&\arg\max_{\mathcal{P},\mathcal{R},\mathcal{S}} \sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m} w_k \sum_{i=1}^{n_F}\sum_{j=1}^{n_F} \frac{s_{m,i,j}^{(k)}}{2}\log_2\left(1+\frac{\Lambda_{eq_m,i,j}^{(k)}}{s_{m,i,j}^{(k)}}\right)\\
&\text{s.t. C2: } \sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{j=1}^{n_F}(\tilde{P}_{RD_m,j}^{(k)}+\tilde{P}_{SR_m,i}^{(k)}) \le P_{\mathrm{t}},\\
&\quad \text{C3},\quad \text{C4},\quad \text{C5},\\
&\quad \text{C6: } 0 \le s_{m,i,j}^{(k)} \le 1, \qquad \forall m,i,j,k
\end{aligned} \tag{11}$$

Since the power constraint is instantaneous, average weighted system goodput maximization is identical to the maximization of the instantaneous weighted goodput for each set of channel gains. The new problem is jointly concave with respect to all optimization variables, under some mild conditions [9], it can be shown that solving the dual problem is equivalent to solving the primal problem.

D. Dual Problem Formulation

In this subsection, we solve the cross-layer scheduling optimization problem by solving its dual. For this purpose, we first need the Lagrangian function of the primal problem. Upon rearranging terms, the Lagrangian can be written as

$$\begin{aligned}
&\mathcal{L}(\lambda,\boldsymbol{\gamma},\boldsymbol{\beta},\mathcal{P},\mathcal{R},\mathcal{S})\\
=&\sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m} w_k \sum_{i=1}^{n_F}\sum_{j=1}^{n_F}\frac{s_{m,i,j}^{(k)}}{2}\log_2\left(1+\frac{\Lambda_{eq_m,i,j}^{(k)}}{s_{m,i,j}^{(k)}}\right)\\
&-\lambda\sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{j=1}^{n_F}(\tilde{P}_{RD_m,j}^{(k)}+\tilde{P}_{SR_m,i}^{(k)})\\
&-\sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{j=1}^{n_F}\beta_i s_{m,i,j}^{(k)} - \sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{j=1}^{n_F}\gamma_j s_{m,i,j}^{(k)}\\
&+\lambda P_t + \sum_{j=1}^{n_F}\gamma_j + \sum_{i=1}^{n_F}\beta_i
\end{aligned} \tag{12}$$

where λ is the Lagrange multiplier corresponding to the joint power constraint, and $\boldsymbol{\gamma}$ and β are Lagrange multiplier vectors associated with the subcarrier usage constraints with elements $\gamma_i,\ i\in\{1,\ldots,n_F\}$, and $\beta_j,\ j\in\{1,\ldots,n_F\}$, respectively. Thus, the dual problem is given by

$$\min_{\lambda,\boldsymbol{\gamma},\boldsymbol{\beta}\ge 0}\ \max_{\mathcal{P},\mathcal{R},\mathcal{S}}\quad \mathcal{L}(\lambda,\boldsymbol{\gamma},\boldsymbol{\beta},\mathcal{P},\mathcal{R},\mathcal{S}). \tag{13}$$

In the following sections, we solve the above dual problem by decomposing it into two parts: the first part is a subproblem to be solved by each relay station; the second part is the master dual problem to be solved by the base station.

E. Distributed Solution - Subproblem for Each Relay Station

By dual decomposition, relay station m can solve the following subproblem without assistance from other relays

$$\max_{\mathcal{P},\mathcal{R},\mathcal{S}}\quad \mathcal{L}_m(\lambda,\boldsymbol{\gamma},\boldsymbol{\beta},\mathcal{P},\mathcal{R},\mathcal{S}) \tag{14}$$

with

$$\begin{aligned}
&\mathcal{L}_m(\lambda,\boldsymbol{\gamma},\boldsymbol{\beta},\mathcal{P},\mathcal{R},\mathcal{S})\\
=&\sum_{k\in\mathcal{U}_m} w_k \sum_{i=1}^{n_F}\sum_{j=1}^{n_F}\frac{s_{m,i,j}^{(k)}}{2}\log_2\left(1+\frac{\Lambda_{eq_m,i,j}^{(k)}}{s_{m,i,j}^{(k)}}\right)\\
&-\lambda\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{j=1}^{n_F}\left(\tilde{P}_{RD_m,j}^{(k)}+\tilde{P}_{SR_m,i}^{(k)}\right)\\
&-\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{j=1}^{n_F}\beta_i s_{m,i,j}^{(k)} - \sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{j=1}^{n_F}\gamma_j s_{m,i,j}^{(k)}
\end{aligned} \tag{15}$$

where the Lagrange multipliers λ, $\boldsymbol{\gamma}$, and $\boldsymbol{\beta}$ are provided by the base station. Using standard optimization techniques and the Karush-Kuhn-Tucker (KKT) condition, the optimal power allocation for subcarrier pair (i,j) is obtained as

$$\begin{aligned} P^{(k)*}_{SR_m,i} &= s^{(k)}_{m,i,j}\Big(w_k/\lambda - \Phi^{(k)}_{m,i,j}\Big)^+\Big/\Big(1+\Omega^{(k)}_{m,i,j}\Big) \\ P^{(k)*}_{RD_m,j} &= s^{(k)}_{m,i,j}\Big(w_k/\lambda - \Phi^{(k)}_{m,i,j}\Big)^+\Big/\Big(1+1/\Omega^{(k)}_{m,i,j}\Big) \end{aligned} \quad (16)$$

where $\Phi^{(k)}_{m,i,j} = \frac{\left(\sqrt{l_{SR_m}|H_{SR_m,i}|^2}+\sqrt{l^{(k)}_{RD_m}F^{-1(k)}_{RD_m,j}(\varepsilon)}\right)^2}{l_{SR_m}|H_{SR_m,i}|^2 l^{(k)}_{RD_m}F^{-1(k)}_{RD_m,j}(\varepsilon)}$, $\Omega^{(k)}_{m,i,j} = \sqrt{\frac{l_{SR_m}|H_{SR_m,i}|^2}{l^{(k)}_{RD_m}F^{-1(k)}_{RD_m,j}(\varepsilon)}}$, and $(x)^+ = \max\{0,x\}$. In order to obtain the optimal subcarrier allocation, we take the derivative of the subproblem with respect to $s^{(k)}_{m,i,j}$ and set it to zero, i.e, $\frac{\partial \mathcal{L}_m}{\partial s^{(k)}_{m,i,j}} = 0$, which yields

$$\underbrace{\log_2\left(1+\frac{\Lambda^{(k)}_{eq_m,i,j}}{s^{(k)}_{m,i,j}}\right) - \frac{\Lambda^{(k)}_{eq_m,i,j}}{s^{(k)}_{m,i,j}+\Lambda^{(k)}_{eq_m,i,j}}}_{A^{(k)}_{m,i,j}} = \frac{2(\gamma_i+\beta_j)}{w_k}$$

Thus, the subcarrier pair selection determined by relay m is given by

$$s^{(k)}_{m,i,j} = \begin{cases} 1 & \text{if } A^{(k)}_{m,i,j} \geq \frac{2(\gamma_i+\beta_j)}{w_k} \\ 0 & \text{otherwise} \end{cases}. \quad (17)$$

The dual variables γ_i and β_j act as prices for using the subcarrier pair (i,j). Only that user who has large weights w_k and subcarrier pairs with large channel gains is able to pay the price and is selected by the scheduler. Finally, the optimal rate allocation $r^{(k)*}_{m,i,j}$ is obtained by substituting (16) into the equivalent rate constraint in (9) for the subcarriers with $s^{(k)}_{m,i,j} = 1$.

F. Solution of the Master Dual Problem at the Base Station

The dual function is differentiable and, hence, the gradient method can be used to solve the minimization of the master problem in (13)

$$\gamma_i(t+1) = \left[\gamma_i(t) - \xi_1(t)(1-\sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{j=1}^{n_F} s^{(k)}_{m,i,j})\right]^+, \forall i$$

$$\beta_j(t+1) = \left[\beta_j(t) - \xi_2(t)(1-\sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F} s^{(k)}_{m,i,j})\right]^+, \forall j$$

$$\lambda(t+1) = \Bigg[\lambda(t) - \xi_3(t)\times (P_{\mathrm{t}} - \sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{j=1}^{n_F}\tilde{P}^{(k)}_{RD_m,j} + \tilde{P}^{(k)}_{SR_m,i})\Bigg]^+ \quad (18)$$

where t is the iteration index. By choosing appropriate step sizes $\xi_1(t), \xi_2(t)$, and $\xi_3(t)$, convergence to the optimal solution is guaranteed [9]. Therefore, in each scheduling time slot, the base station initializes all Lagrange multipliers and broadcasts the values to the relays. Then, each relay solves subproblem (14) and feeds back the solution to the base station. Subsequently, the base station updates the Lagrange multipliers using (18) and broadcasts the new values to the relays. This process is repeated until convergence is achieved.

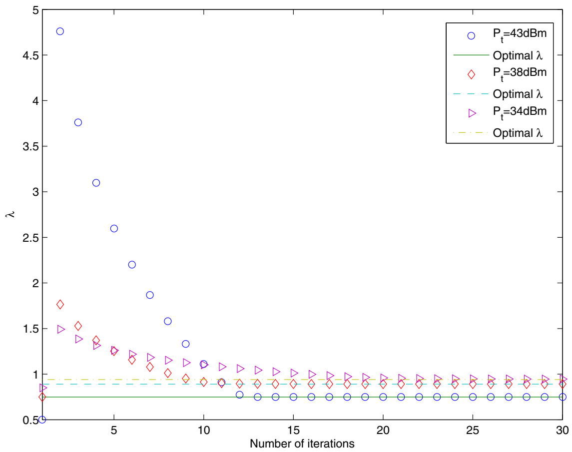

Fig. 1. Dual variable λ versus number of iterations for $P_{\mathrm{t}} = 43$ dBm, 38 dBm, and 34 dBm for a system with $M = 3$ relays, $K = 10$ users, and $n_F = 64$ subcarriers.

IV. Results and Discussions

In this section, we evaluate the system performance using simulations. A single cell with two ring-shaped boundary regions is considered. The outer boundary and the inner boundary have radii of 1 km and 500 m, respectively. There are $M = 3$ relay stations equally distributed on the inner cell boundary for assisting the transmission and K active users are uniformly distributed between the inner and the outer boundaries. The number of subcarriers is $n_F = 64$ and $w_k = 1, \forall k$, for illustration. The number of channel taps is $W = 5$ and the 3GPP path loss model is used [10]. The small scale fading coefficients of the base station-to-relay links are modeled as i.i.d. Rician random variables with Rician factor equal to 6 dB, while the small scale fading coefficients of the relay-to-user links are i.i.d. $\mathcal{CN}(0,1)$. The target packet outage probability is set to $\varepsilon = 0.01$. The weighted average system goodput is obtained by counting the number of packets successfully decoded by all users averaged over both the macroscopic and microscopic fading.

A. Convergence of Distributed Algorithm

Figure 1 illustrates the convergence of λ in time for different P_{t}. Most of the time, the distributed algorithm can approach 90-95% of the optimal value after 20 iterations. It should be noted that the proposed distributed solution provides significant reduction in computational complexity at the base station compared to optimal centralized scheduling and the signaling overhead for CSIT feedback is comparatively low since only local CSI is needed at each relay.

B. System Goodput versus Transmit Power

Figure 2 illustrates the average system goodput versus the total transmit power for $K = 10$ users with channel estimation error variance $\sigma_e^2 = 0.01$. The performance of the proposed

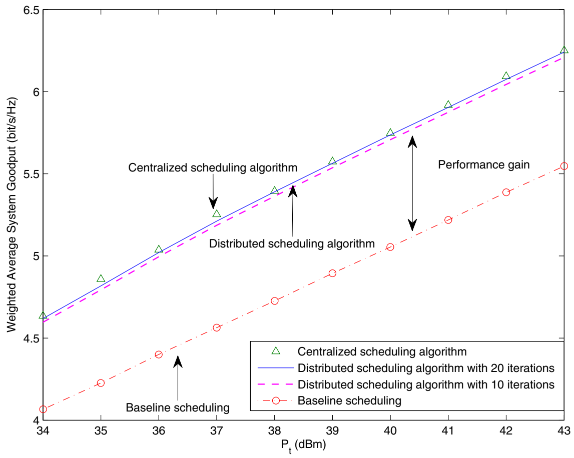

Fig. 2. Average system goodput versus total transmit power for different scheduling algorithms and CSIT errors $\sigma_e^2 = 0.01$.

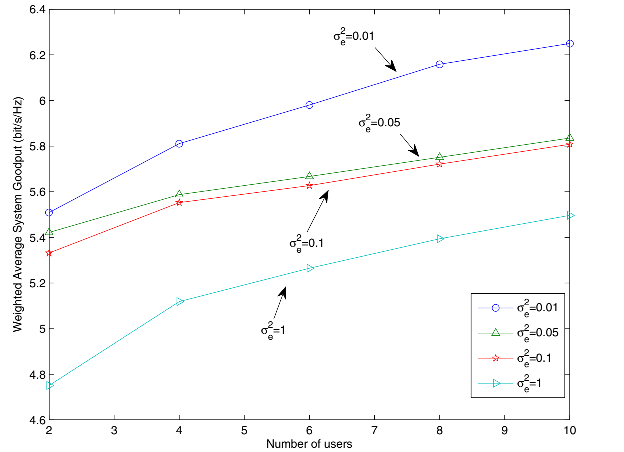

Fig. 3. Average system goodput versus number of users for different CSIT errors ($\sigma_e^2 = 0.01, 0.05, 0.1, 1$) and $P_{\mathrm{t}} = 43$ dBm.

distributed scheduling algorithm with 10 and 20 iterations is compared with that of the optimal centralized scheduling algorithm and a baseline round robin scheduler for the same target outage probability for all users. For the centralized scheduler, the base station uses the global CSI of each link to perform an exhaustive search to find the optimal subcarrier allocation and uses a standard water-filling procedure to obtain the optimal power allocation and rate adaption. As can be observed, the performance difference between the distributed and the optimal centralized schedulers is negligible even for the case of 10 iterations. For the baseline round robin scheduler subcarrier mapping is not performed and the power is optimally allocated in centralized manner. The performance of the baseline scheme is always worse than that of the proposed distributed scheduler because the proposed scheduler fully exploits the available CSI for resource allocation while the baseline scheduler does not.

C. System Goodput versus Number of Users

Figure 3 depicts the average system goodput versus the number of users for different CSIT error variances σ_e^2. The total transmit power is 43 dBm and the number of iterations for the distributed scheduler is 20. It can be observed that the system goodput grows with the number of users since the proposed scheduling algorithm is able to exploit multi-user diversity (MUD). Typically, schedulers cannot exploit MUD if $\sigma_e^2 = 1$ since the CSIT for the small scale fading is independent of the actual channel. However, the average system goodput of the proposed scheduler still grows with the number of users. This is because the proposed scheduler takes into account both small scale and large scale fading. Although the scheduler does not have any information concerning the small scale fading in the relay-to-user links if $\sigma_e^2 = 1$, the path loss information is still beneficial for power and rate adaptation. Therefore, the scheduler can still exploit MUD by selecting users who are closer to relay station.

V. Conclusion

In this paper, taking into account imperfect CSIT, the cross-layer design of scheduling for AF relay assisted OFDMA downlink transmission is formulated as a mixed combinatorial and convex optimization problem. Based on dual decomposition of the primal problem, a distributed resource allocation algorithm is derived. Our simulation results show that the performance of the distributed algorithm approaches that of the optimal centralized scheduler in a small number of iterations which confirms the practicality of the proposed scheduler.

References

[1] A. Doufexi and S. Armour, "Design considerations and physical layer performance results for a 4G OFDMA System Employing Dynamic Subcarrier Allocation," in *Proc. IEEE Intern. Symp. Personal, Indoor and Mobile Radio Commun.*, vol. 1, Sep 2005, pp. 357–361.

[2] I. Hammerstrom and A. Wittneben, "Power allocation schemes for amplify-and-forward MIMO-OFDM relay links," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2798–2802, Aug 2007.

[3] T. C. Y. Ng and W. Yu, "Joint optimization of relay strategies and resource allocations in cooperative cellular networks," *IEEE J. Select. Areas Commun.*, vol. 25, pp. 328 – 339, Feb. 2007.

[4] L. You, M. Song, J. Song, Q. Miao, and Y. Zhang, "Resource allocation for OFDMA based relay enhanced cellular networks," in *Proc. IEEE Vehicular Technol. Conf.*, May 2008, pp. 1925–1929.

[5] S. Xuemin and M. Awad, "OFDMA based two-hop cooperative relay network resources allocation," in *Proc. IEEE Intern. Conf. on Commun.*, May 2008, pp. 4414–4418.

[6] J. Laneman, D. Tse, and G.W.Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inform. Theory*, vol. 50, pp. 3062– 3080, Dec. 2004.

[7] M. Herdin, "A chunk based OFDM amplify-and-forward relaying scheme for 4G mobile radio systems," in *Proc. IEEE Wireless Commun. and Networking Conf.*, vol. 10, June 2006, pp. 4507–4512.

[8] C. Y. Wong, R. Cheng, K. Lataief, and R. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Select. Areas Commun.*, vol. 17, 1999.

[9] S. Boyd and L.Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[10] "Spatial channel model for multiple input multiple output (MIMO) simulations," 3GPP TR 25.996 V7.0.0 (2007-06), Tech. Rep.