This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2011 proceedings

Dynamic Resource Allocation in OFDMA Systems with Full-Duplex and Hybrid Relaying

Derrick Wing Kwan Ng and Robert Schober
The University of British Columbia

***Abstract*—In this paper, we formulate a joint optimization problem for resource allocation and scheduling in full-duplex orthogonal frequency division multiple access (OFDMA) relaying systems with amplify-and-forward (AF) and decode-and-forward (DF) relaying protocols. Our problem formulation takes into account heterogeneous data rate requirements for delay sensitive users. Besides, a theoretically optimal hybrid relaying, which allows a dynamic selection between AF relaying and DF relaying protocols with full-duplex relays or half-duplex relays, is also considered in the problem formulation and serves as a performance benchmark. A dual decomposition method is employed to solve the resulting optimization problem and a novel distributed iterative resource allocation and scheduling algorithm with closed-form power and subcarrier allocation is derived. Simulation results illustrate that the proposed distributed algorithm requires only a small number of iterations to achieves practically the same performance as the optimal centralized algorithm.**

I. Introduction

Orthogonal frequency division multiple access (OFDMA) is an important technique for high data rate wireless multiuser communication systems, such as 3GPP Long Term Evolution (LTE) and IEEE 802.16 Worldwide Interoperability for Microwave Access (WiMAX), not only because of its flexibility in resource allocation, but also because of its ability to exploit multiuser diversity. On the other hand, cooperative relaying for wireless networks has received considerable interest, as it provides coverage extension and reduced power consumption without incurring the high costs of additional base station (BS) deployment. There are two main types of relaying schemes, half-duplex (HD) relaying and full-duplex (FD) relaying. In the literature, a large amount of work has been devoted to HD relaying [1]-[3] as it facilities a low-complexity relay design. Nevertheless, HD relaying systems require additional resources to transmit data in a multi-hop manner which results in a loss in spectral efficiency. On the contrary, although FD relaying was considered impractical in the past because of its inherent loop interference problems, it has regained the attention of both industry [4], [5] and academia [6]-[8] as recent research shows that FD relaying is feasible by using interference cancellation techniques and transmit/receive antenna isolation [6]-[8]. Yet, efficient resource allocation and scheduling algorithms for multi-user FD relaying systems have not been reported in the literature so far.

Next generation wireless communication systems are required to support heterogeneous data rate services and to guarantee certain quality of service (QoS) requirements. The combination of relaying and OFDMA provides a viable solution for addressing these issues. In [1]-[3], best effort resource allocation and scheduling for homogeneous users in OFDMA HD relaying systems are studied for different system configurations. However, QoS requirements are driven by heterogeneous applications and users may demand different data rates, which best effort resource allocation cannot guarantee. Besides, FD relaying could provide a substantial performance gain and should not be overlooked in the system design. Furthermore, existing works focus on centralized resource allocation at the BS. As the numbers of users/relays and subcarriers in the system increase, brute force resource optimization may overload the BS which limits the system scalability. Therefore, a distributed resource allocation algorithm, which enables the exploitation of the possible advantages of different relaying protocols and duplexing schemes in different environments, fulfills the heterogeneous QoS requirement, and converges fast to the optimal solution is needed for practical implementation.

In this paper, we formulate the scheduling and resource allocation problem for OFDMA FD-relaying systems as an optimization. In addition, a theoretically optimal hybrid relaying scheme, which dynamically selects between AF relaying and DF relaying with full-duplex relays or half-duplex relays, is also considered in the problem formulation and serves as a performance benchmark. By using dual decomposition, the resulting optimization problem is decomposed into a master problem and several subproblems. The BS solves the master problem with a gradient method and updates the dual variables through the concept of pricing, while each relay solves its own subproblem by utilizing the dual variables and its local channel state information (CSI) without any help from other relays.

II. OFDMA relay network model

We consider an OFDMA downlink relay assisted packet transmission network which consists of one BS, M relays, and K mobile users which belong to one of two categories, namely, *delay sensitive* users and *non-delay sensitive* users. The *delay sensitive* users require a minimum constant data rate while *non-delay sensitive* users require only best-effort service. All transceivers are equipped with a single antenna. A cell is modeled by two concentric ring-shaped discs as shown in Figure 1. The cell coverage is divided into M areas corresponding to the M relays and each user is assigned to a relay. In this paper, we focus on resource allocation and scheduling for heterogeneous users who need the help of relays, i.e., cell edge users in the shaded region of Figure 1. We assume that there is no direct transmission between the BS and the mobile users due to heavy blockage and long distance transmission. A distributed algorithm is derived for resource allocation and scheduling purposes. Based on the CSI of users, the algorithm selects between four transmission strategies t on a per subcarrier basis for hybrid relaying, namely: $t = 1$; decode-and-forward full-duplex (DF-FD) relaying, $t = 2$, decode-and-forward half-duplex (DF-HD) relaying; $t = 3$, amplify-and-forward full-duplex (AF-FD) relaying; and $t = 4$, amplify-and-forward half-duplex (AF-HD) relaying.[1] A time-division channel allocation with two time slots is used to facilitate transmission. In the first time slot, the BS broadcasts a data packet to the relays. Then, in the second time slot, if subcarrier i is using FD relaying, the corresponding relay decodes/amplifies the previously received signal on subcarrier i and forwards it to the corresponding user while the BS transmits the next packet. If HD relaying is used on subcarrier i, the relays perform the same signal processing on subcarrier i as for FD transmission, however, the BS remains silent during the second time slot.

[1]The hybrid relaying serves as a theoretical upper bound of the system performance which provides useful system design insights such as the operating region of each transmission strategy. As we will see in the next section, we can always restrict the algorithm to select AF-FD relaying or DF-FD relaying for a practical implementation.

978-1-61284-233-2/11/$26.00 ©2011 IEEE

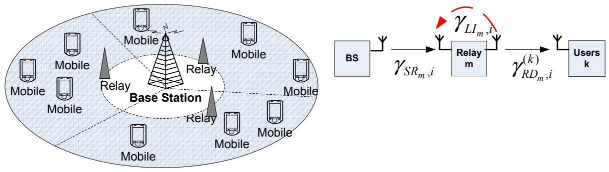

Fig. 1. The left hand side shows the model of a relay assisted packet transmission system with $K = 9$ users and $M = 3$ relays. The right hand side shows a block diagram of the transmission between the BS and user k via full-duplex relay m on subcarrier i with loop interference.

A. Channel Model

The channel impulse response is assumed to be time-invariant within a frame. We consider an OFDMA system with n_F subcarriers. In the first time slot, the (frequency domain) received symbol on subcarrier $i \in \{1, \dots, n_F\}$ at relay $m \in \{1, \dots, M\}$ for user $k \in \{1, \dots, K\}$ using transmission strategy t is given by

$$Y_{SR_m,i}^{(t,k)} = \sqrt{P_{SR_m,i}^{(t,k)} l_{SR_m}} H_{SR_m,i} X_i^{(k)} + Z_{R_m,i} + \sqrt{\varepsilon} H_{LI_m,i} U_{m,i}^{(k)} \quad (1)$$

$$\text{with} \quad U_{m,i}^{(t,k)} = \begin{cases} I_{m,i}^{(t,k)} & \text{for AF-FD relaying} \\ \sqrt{P_{RD_m,i}^{(t,k)}} W_i^{(k)} & \text{for DF-FD relaying} \end{cases}, \quad (2)$$

where $X_i^{(k)}$ is the unit-variance symbol transmitted from the BS to user k via relay m on subcarrier i, $I_{m,i}^{(k)}$ is the accumulated loop interference signal due to AF-FD relaying at relay m on subcarrier i, and $W_i^{(k)}$ is the unit-variance symbol transmitted to user k on subcarrier i from relay m for DF-FD relaying. $P_{SR_m,i}^{(t,k)}$ is the transmit power for the link between the BS and relay m in subcarrier i for user k. $P_{RD_m,i}^{(t,k)}$ is the transmit power for the link between relay m and user k in subcarrier i. l_{SR_m} represents the path loss between the BS and relay m. $Z_{R_m,i}$ is additive white Gaussian noise (AWGN) with distribution $\mathcal{CN}(0, N_0)$ on subcarrier i at relay m, where N_0 is the noise power spectral density and $\mathcal{CN}(\mu, \sigma^2)$ denotes a complex Gaussian random variable with mean μ and variance σ^2. In practice, both the BS and the relays are placed in relatively high positions and hence the number of blockages between them are limited and a strong line of sight is expected. Hence, the channel fading coefficients between the BS and relay m on subcarrier i, $H_{SR_m,i}$, are modeled as Rician fading with Rician factor κ, i.e., $H_{SR_m,i} \sim \mathcal{CN}(0, \sqrt{\kappa/(1+\kappa)}, 1/(1+\kappa))$. Assuming interference cancellation is employed at the relays, $\sqrt{\varepsilon} H_{LI_m,i}$ is the residual loop interference with loop interference channel coefficient $H_{LI_m,i} \sim \mathcal{CN}(0, \sigma_I^2)$ and $0 \le \varepsilon \le 1$ is the fraction of loop interference due to imperfect cancellation. If HD relaying is used, (1) is still valid if we set $\varepsilon = 0$, because of the orthogonal transmission between the BS and the relays.

If the AF protocol is used for transmission, the signal on subcarrier i is amplified by a gain factor $\sqrt{P_{RD_m,i}^{(t,k)} G_{RD_m,i}^{(t,k)}}$, and forwarded to the destination, where $P_{RD_m,i}^{(t,k)}$ is the transmit power for the link between relay m and user k on subcarrier i and $G_{RD_m,i}^{(t,k)}$ normalizes the input power of relay m on subcarrier i. The signal received at user k on subcarrier i from relay m using the AF protocol is given by

$$Y_{RD_m,i}^{(t,k)} = \sqrt{G_{RD_m,i}^{(t,k)} P_{RD_m,i}^{(t,k)} l_{RD_m}^{(k)}} H_{RD_m,i}^{(k)} \times \Big(\sqrt{P_{SR_m,i}^{(t,k)} l_{SR_m}} H_{SR_m,i} X_i^{(k)} + \sqrt{\varepsilon} H_{LI_m,i} I_{m,i}^{(k)} + Z_{R_m,i}\Big) + Z_i^{(k)}, \quad (3)$$

where variables $l_{RD_m}^{(k)}$, $H_{RD_m,i}^{(k)}$, and $Z_i^{(k)}$ are defined in a similar manner as the corresponding variables for the BS-to-relay links. To simplify the subsequent mathematical expressions and without loss of generality, we assume in the following a noise variance of $N_0 = 1$ in all receivers and the residual loop interference power is known at the relays. Assuming the AF-FD relays are in steady state, the variance of the accumulated loop interference at relay m in subcarrier i can be expressed as $(\sigma_{I_m,i}^{(t,k)})^2 = E[|I_{m,i}^{(t,k)}|^2] = P_{RD_m,i}^{(t,k)} G_{RD_m,i}^{(t,k)} \times$

$$\sum_{n=1}^{\infty} (P_{RD_m,i}^{(t,k)} G_{RD_m,i}^{(t,k)} \varepsilon |H_{LI_m,i}|^2)^{n-1} (P_{SR_m,i}^{(t,k)} l_{SR_m} |H_{SR_m,i}|^2 + 1)$$
$$= P_{RD_m,i}^{(t,k)} G_{RD_m,i}^{(t,k)} \frac{P_{SR_m,i}^{(t,k)} l_{SR_m} |H_{SR_m,i}|^2 + 1}{1 - \varepsilon |H_{LI_m,i}|^2 P_{RD_m,i}^{(t,k)} G_{RD_m,i}^{(t,k)}}, \quad (4)$$

where $E[\cdot]$ denotes statistical expectation and n represents the index of the infinite summation. The sum in (4) converges if $G_{RD_m,i}^{(t,k)} < 1/(P_{RD_m,i}^{(t,k)} \varepsilon |H_{LI_m,i}|^2)$ which holds for the popular choice $G_{RD_m,i}^{(t,k)} = (1 + P_{SR_m,i}^{(t,k)} l_{SR_m} |H_{SR_m,i}|^2 + \varepsilon P_{RD_m,i}^{(t,k)} |H_{LI_m,i}|^2)^{-1}$.

On the other hand, if the DF protocol is used for transmission, the signal on subcarrier i is decoded and re-encoded at relay m, and then forwarded to the destination. The signal received at user k on subcarrier i from relay m is given by

$$Y_{RD_m,i}^{(t,k)} = \sqrt{P_{RD_m,i}^{(t,k)} l_{RD_m}^{(k)}} H_{RD_m,i}^{(k)} W_i^{(k)} + Z_i^{(k)}. \quad (5)$$

III. Resource Allocation and Scheduling Design

A. Instantaneous Channel Capacity and System Throughput

In this subsection, we define the adopted system performance measure. For the sake of notational simplicity, we define $\gamma_{RD_m,i}^{(k)} = l_{RD_m}^{(k)} |H_{RD_m,i}^{(k)}|^2$, $\gamma_{SR_m,i} = l_{SR_m} |H_{SR_m,i}|^2$, and $\gamma_{LI_m,i} = \varepsilon |H_{LI_m,i}|^2$. Given perfect CSI at the receiver (CSIR), the channel capacity between the BS and user k via relay m on subcarrier i of different transmission strategies, i.e., $C_{m,i}^{(t,k)}$, can be expressed as follows:

For $t = 1$, DF-FD relaying is selected and $C_{m,i}^{(1,k)} =$

$$\log_2 \left(1 + \min \left\{ \frac{P_{SR_m,i}^{(t,k)} \gamma_{SR_m,i}}{\gamma_{LI_m,i} P_{RD_m,i}^{(t,k)} + 1}, P_{RD_m,i}^{(t,k)} \gamma_{RD_m,i}^{(k)} \right\} \right); \quad (6)$$

for $t = 2$, DF-HD relaying is selected and $C_{m,i}^{(2,k)} =$

$$\frac{1}{2} \log_2 \left(1 + \min \left\{ \gamma_{SR_m,i} P_{SR_m,i}^{(t,k)}, P_{RD_m,i}^{(t,k)} \gamma_{RD_m,i}^{(k)} \right\} \right); \quad (7)$$

for $t = 3$, AF-FD relaying is is selected and $C_{m,i}^{(3,k)} =$

$$\log_2 \left(1 + \frac{\frac{P_{SR_m,i}^{(t,k)} \gamma_{SR_m,i}}{P_{RD_m,i}^{(t,k)} \gamma_{LI_m,i} + 1} P_{RD_m,i}^{(t,k)} \gamma_{RD_m,i}^{(k)}}{1 + P_{RD_m,i}^{(t,k)} \gamma_{RD_m,i}^{(k)} + \frac{P_{SR_m,i}^{(t,k)} \gamma_{SR_m,i}}{1 + P_{RD_m,i}^{(t,k)} \gamma_{LI_m,i}}} \right); \quad (8)$$

for $t = 4$, AF-HD relaying is selected and $C_{m,i}^{(4,k)} =$

$$\frac{1}{2} \log_2 \left(1 + \frac{\gamma_{SR_m,i} P_{SR_m,i}^{(t,k)} P_{RD_m,i}^{(t,k)} \gamma_{RD_m,i}^{(k)}}{1 + \gamma_{SR_m,i} P_{SR_m,i}^{(t,k)} + P_{RD_m,i}^{(t,k)} \gamma_{RD_m,i}^{(k)}} \right), \quad (9)$$

where the the pre-log factor $\frac{1}{2}$ in (7) and (9) is due to the two channel uses for transmitting one message with HD relaying.

Now, we define the instantaneous throughput (bit/s/Hz successfully delivered) for user k who is served by relay m as

$$\rho_m^{(k)} = \frac{1}{n_F} \sum_{i=1}^{n_F} \sum_{t=1}^{4} s_{m,i}^{(t,k)} C_{m,i}^{(t,k)}, \tag{10}$$

where $s_{m,i}^{(t,k)} \in \{0,1\}$ is the subcarrier allocation indicator. The *average weighted system throughput* is defined as the total average number of bit/s/Hz successfully decoded at the K users via the M relays and given by

$$U_{TP}(\mathcal{P}, \mathcal{S}) = \sum_{m=1}^{M} \sum_{k \in \mathcal{U}_m} w_k \rho^{(k)}, \tag{11}$$

where $\mathcal{P}$ and $\mathcal{S}$ are the power and subcarrier allocation policies, respectively, $\mathcal{U}_m$ is the set of users served by relay m, and w_k is a positive constant, which is specified in the media access control (MAC) layer and allows the schedulers to prioritize the users so as to achieve certain fairness objectives.

B. Problem Formulation

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

$$\begin{aligned}
&(\mathcal{P}^*, \mathcal{S}^*) = \arg\max_{\mathcal{P},\mathcal{S}} \quad U_{TP}(\mathcal{P}, \mathcal{S}) \\
\text{s.t.} \quad &\text{C1:} \quad \sum_{m=1}^{M} \sum_{k \in \mathcal{U}_m} \sum_{t=1}^{4} \sum_{i=1}^{n_F} \Big(P_{SR_m,i}^{(t,k)} + P_{RD_m,i}^{(t,k)}\Big) s_{m,i}^{(t,k)} \le P_T, \\
&\text{C2:} \quad \rho^{(k)} \ge R^{(k)}, \quad \forall k \in \mathcal{D} \\
&\text{C3:} \quad \sum_{m=1}^{M} \sum_{k \in \mathcal{U}_m} \sum_{t=1}^{4} s_{m,i}^{(t,k)} = 1, \quad \forall i \\
&\text{C4:} \quad s_{m,i}^{(t,k)} \in \{0,1\}, \ \forall m, i, t, k \\
&\text{C5:} \quad P_{SR_m,i}^{(t,k)}, P_{RD_m,i}^{(t,k)} \ge 0, \quad \forall m, i, k, t,
\end{aligned} \tag{12}$$

where $\mathcal{D}$ is the set of *delay sensitive* users requiring data rates of $R^{(k)}, \forall k \in \mathcal{D}$. Here, C1 is a joint power constraint for the BS and the relays with total maximum power P_T. Although in a practical system the BS and the relays have separate power supplies, a joint power optimization provides useful insight into the power usage of the whole system rather than the per hop required power. Moreover, for the BS and the relays with separate power constraints, obtaining a globally optimal solution in polynomial time does not seem possible due to the non-convexity of the problem. C2 guarantees the minimum required data rate $R^{(k)}$ for *delay sensitive* users which are chosen by the application layer. Constraints C3 and C4 are imposed to guarantee that each subcarrier is only used by one user. In other words, intra-cell interference does not exist in the system. Also, they ensure that each subcarrier can be transmitted with one strategy only. C5 is the positive power constraint.

C. Transformation of the Optimization Problem

In general, the considered problem is a mixed combinatorial and non-convex optimization problem. The combinatorial nature comes from the integer constraint for subcarrier allocation while the non-convexity is caused by the power allocation variables in the objective function. Thus, a brute force approach is needed to obtain the global optimal solution. However, such method is computationally infeasible for a large system and does not provide useful system design insight. In order to obtain an insightful closed-form solution for scheduling and resource allocation purposes, we introduce the following proposition.

Proposition 1 (Equivalent Capacity): Without loss of generality, we define a new variable $P_{m,i}^{(t,k)}$ such that $P_{m,i}^{(t,k)} = P_{SR_m,i}^{(t,k)} + P_{RD_m,i}^{(t,k)}$, which represents the power consumption in subcarrier i for user k via relay m. Let $\Gamma_{m,i}^{(t,k)}$ be the corresponding equivalent received signal-to-noise ratio (SNR) or signal-to-interference-plus-noise ratio (SINR) at user k on subcarrier i via relay m using transmission strategy t. Then, the maximized channel capacity on subcarrier i for user k via relay m of each transmission strategy is given as follows:

For $t = 1$, $C_{m,i}^{(1,k)} = \log_2(1 + \Gamma_{m,i}^{(1,k)})$ with

$$\Gamma_{m,i}^{(1,k)} = \frac{\sqrt{(\Upsilon_{m,i}^{(k)})^2 + 4\Psi_{m,i}^{(k)} P_{m,i}^{(1,k)}} - (\Upsilon_{m,i}^{(k)})}{2\gamma_{LI_m,i}}; \tag{13}$$

for $t = 2$, $C_{m,i}^{(2,k)} = \frac{1}{2}\log_2(1 + \Gamma_{m,i}^{(2,k)})$ with

$$\Gamma_{m,i}^{(2,k)} = P_{m,i}^{(2,k)} \frac{\gamma_{SR_m,i}\gamma_{RD_m,i}^{(k)}}{\Upsilon_{m,i}^{(k)}}; \tag{14}$$

for $t = 3$, $C_{m,i}^{(3,k)} = \log_2(1 + \Gamma_{m,i}^{(3,k)})$ with

$$\Gamma_{m,i}^{(3,k)} \approx \frac{\sqrt{\gamma_{RD_m,i}^{(k)}\gamma_{SR_m,i} P_{m,i}^{(3,k)}}}{2\sqrt{\gamma_{LI_m,i}}}; \tag{15}$$

for $t = 4$, $C_{m,i}^{(4,k)} = \frac{1}{2}\log_2(1 + \Gamma_{m,i}^{(4,k)})$ with

$$\Gamma_{m,i}^{(4,k)} \approx \frac{P_{m,i}^{(4,k)}\gamma_{SR_m,i}\gamma_{RD_m,i}^{(k)}}{(\sqrt{\gamma_{SR_m,i}} + \sqrt{\gamma_{RD_m,i}^{(k)}})^2}, \tag{16}$$

where $\Psi_{m,i}^{(k)} = \gamma_{LI_m,i}\gamma_{RD_m,i}^{(k)}\gamma_{SR_m,i}$ and $\Upsilon_{m,i}^{(k)} = \gamma_{SR_m,i} + \gamma_{RD_m,i}^{(k)}$. Hence, we substitute (13)-(16) into the original objective function in (12) to obtain a new objective function.

Proof: Please refer to the Appendix.

Note that the above proposition is a key step to simplify the optimization problem. The objective function is now concave with respect to (w.r.t.) the new optimization variables $P_{m,i}^{(t,k)}$ for all transmission strategies. The next step is to handle the combinatorial subcarrier assignment constraint. Brute force optimization is a possible way to obtain a global optimal solution but it is computational infeasible for even moderate system sizes. In order to strike a balance between optimality and complexity, we follow the approach in [9] and relax $s_{m,i}^{(t,k)}$ in constraint C4 to be a real value between zero and one instead of a Boolean. Then, $s_{m,i}^{(t,k)}$ can be interpreted as a time sharing factor for the K users to utilize subcarrier i through relay m with different transmission strategies. It can be shown that the relaxation is optimal [10] for large numbers of subcarriers. Thus, the optimization problem can be rewritten as

$$\begin{aligned}
&(\mathcal{P}^*, \mathcal{S}^*) = \arg\max_{\mathcal{P},\mathcal{S}} \quad \sum_{m=1}^{M} \sum_{k \in \mathcal{U}_m} \sum_{t=1}^{4} \sum_{i=1}^{n_F} w^{(k)} s_{m,i}^{(t,k)} \widetilde{C}_{m,i}^{(t,k)} \\
\text{s.t.} \quad &\text{C1:} \quad \sum_{m=1}^{M} \sum_{k \in \mathcal{U}_m} \sum_{t=1}^{4} \sum_{i=1}^{n_F} \widetilde{P}_{m,i}^{(t,k)} \le P_T, \\
&\qquad \text{C2, C3, C5} \\
&\text{C4:} \quad 0 \le s_{m,i}^{(t,k)} \le 1 \quad \forall m, i, t, k,
\end{aligned} \tag{17}$$

where $\widetilde{P}_{m,i}^{(t,k)} = P_{m,i}^{(t,k)} s_{m,i}^{(t,k)} = \widetilde{P}_{SR_m,i}^{(t,k)} + \widetilde{P}_{RD_m,i}^{(t,k)}$, $\widetilde{P}_{SR_m,i}^{(t,k)} = P_{SR_m,i}^{(t,k)} s_{m,i}^{(t,k)}$, and $\widetilde{P}_{RD_m,i}^{(t,k)} = P_{RD_m,i}^{(t,k)} s_{m,i}^{(t,k)}$ are auxiliary power allocation variables. $\widetilde{C}_{m,i}^{(t,k)} = C_{m,i}^{(t,k)}|_{P_{m,i}^{(t,k)}=\widetilde{P}_{m,i}^{(t,k)}/s_{m,i}^{(t,k)}}$ is the time-shared channel capacity for each transmission strategy. It can be shown that (17) is now jointly concave w.r.t. all optimization variables and the duality gap is equal to zero under some mild conditions [11, Chapter 5]. Hence, solving the dual problem is equivalent to solving the original primal problem.

D. Dual Problem Formulation

In this subsection, we formulate the dual for the transformed optimization problem. For this purpose, we first need the Lagrangian function of the primal problem. Upon rearranging terms, the Lagrangian can be written as

$$\begin{aligned}\mathcal{L}(\lambda,\boldsymbol{\delta},\boldsymbol{\beta},\mathcal{P},\mathcal{S}) &= \sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{t=1}^{4}(w_k+\delta_k)s_{m,i}^{(t,k)}\widetilde{C}_{m,i}^{(t,k)}\\ &- \lambda\sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{t=1}^{4}\widetilde{P}_{m,i}^{(t,k)} - \sum_{k=1}^{K}R^{(k)}\delta_k\\ &- \sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{t=1}^{4}\beta_i s_{m,i}^{(t,k)} + \lambda P_T + \sum_{i=1}^{n_F}\beta_i, \end{aligned} \quad (18)$$

where λ is the Lagrange multiplier corresponding to the joint power constraint and $\boldsymbol{\beta}$ is the Lagrange multiplier vector associated with the subcarrier usage constraints with elements $\beta_i,\ i\in\{1,\dots,n_F\}$. $\boldsymbol{\delta}$ is the Lagrange multiplier vector corresponding to the data rate constraint and has elements δ_k, $k\in\{1,\dots,K\}$. Note that $\delta_k=0$ for *non-delay sensitive* users, i.e., $k\notin\mathcal{D}$. The boundary constraints C4 and C5 will be absorbed into the Karush-Kuhn-Tucker (KKT) conditions when deriving the optimal solution in Section III-E.

Thus, the dual problem is given by

$$\min_{\lambda,\boldsymbol{\delta},\boldsymbol{\beta}\geq 0}\ \max_{\mathcal{P},\mathcal{S}}\quad \mathcal{L}(\lambda,\boldsymbol{\delta},\boldsymbol{\beta},\mathcal{P},\mathcal{S}). \quad (19)$$

E. Distributed Solution - Subproblem for Each Relay

By dual decomposition, the dual problem in (19) can be decomposed into a master problem and serval subproblems. The dual problem can be solved iteratively where in each iteration each relay solves its own subproblem without using the CSI from the other relays and passes its local solution to the BS for solving the master problem. The subproblem to be solved at relay m is given by

$$\max_{\mathcal{P},\mathcal{S}}\quad \mathcal{L}_m(\lambda,\boldsymbol{\delta},\boldsymbol{\beta},\mathcal{P},\mathcal{S})\quad \text{with} \quad (20)$$

$$\begin{aligned}\mathcal{L}_m(\lambda,\boldsymbol{\delta},\boldsymbol{\beta},\mathcal{P},\mathcal{S}) &= \sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{t=1}^{4}(w_k+\delta_k)s_{m,i}^{(t,k)}\widetilde{C}_{m,i}^{(t,k)}\\ -\lambda\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{t=1}^{4}\widetilde{P}_{m,i}^{(t,k)} &- \sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{t=1}^{4}\beta_i s_{m,i}^{(t,k)}, \end{aligned} \quad (21)$$

where the Lagrange multipliers λ, $\boldsymbol{\delta}$, and $\boldsymbol{\beta}$ are provided by the BS. Using standard optimization techniques and the KKT conditions, the optimal power allocation for subcarrier i via relay m for different transmission strategies is obtained as:

DF Full-Duplex Relaying, i.e., $t=1$,

$$P_{m,i}^{(1,k)*} = \Big[\frac{4\Psi_{m,i}^{(k)}}{(\Theta_{m,i}^{(k)}+\Omega_{m,i}^{(k)})^2} - \frac{(\Upsilon_{m,i}^{(k)})^2}{4\Psi_{m,i}^{(k)}}\Big]^+, \quad (22)$$

$$P_{RD_m,i}^{(1,k)*} = \Big[\frac{\Upsilon_{m,i}^{(k)} - \sqrt{4\Psi_{m,i}^{(k)}P_{m,i}^{(1,k)*} + (\Upsilon_{m,i}^{(k)})^2}}{2\gamma_{LI_m,i}\gamma_{RD_m,i}^{(k)}}\Big]^+, \quad (23)$$

where $\Theta_{m,i}^{(k)} = \sqrt{\frac{\lambda\ln(2)}{(w_k+\delta_k)}(8\Psi_{m,i}^{(k)} + \frac{\lambda\ln(2)}{(w_k+\delta_k)}A_{m,i}^{(k)})}$, $A_{m,i}^{(k)} = (\Upsilon_{m,i}^{(k)} - 2\gamma_{LI_m,i})^2$, $\Omega_{m,i}^{(k)} = \frac{\lambda}{(w_k+\delta_k)}\ln(2)(2\gamma_{LI_m,i} - \Upsilon_{m,i}^{(k)})$, and $[x]^+ = \max\{0,x\}$.

DF Half-Duplex Relaying, i.e., $t=2$,

$$P_{m,i}^{(2,k)*} = \Big[\frac{w_k+\delta_k}{2\lambda\ln(2)} - \frac{\gamma_{SR_m,i}+\gamma_{RD_m,i}^{(k)}}{\gamma_{SR_m,i}\gamma_{RD_m,i}^{(k)}}\Big]^+, \quad (24)$$

$$P_{RD_m,i}^{(2,k)*} = \Big[\frac{(w_k+\delta_k)\gamma_{SR_m,i}}{2\lambda\ln(2)\big(\gamma_{RD_m,i}^{(k)}+\gamma_{SR_m,i}\big)} - \frac{1}{\gamma_{RD_m,i}^{(k)}}\Big]^+. \quad (25)$$

AF Full-Duplex Relaying, i.e., $t=3$,

$$P_{m,i}^{(3,k)*} = \frac{\Psi_{m,i}^{(k)}}{\Big(2\lambda/(w_k+\delta_k)\ln(2)\gamma_{LI_m,i} + \Phi_{m,i}^{(k)}\Big)^2}, \quad (26)$$

$$P_{RD_m,i}^{(3,k)*} = \sqrt{P_{m,i}^{(3,k)*}}\sqrt{\frac{\gamma_{SR_m,i}}{\gamma_{LI_m,i}\gamma_{RD_m,i}^{(k)}}}, \quad (27)$$

where $\Phi_{m,i}^{(k)} = \sqrt{4(\frac{\lambda}{(w_k+\delta_k)})^2\ln^2(2)\gamma_{LI_m,i}^2 + \frac{2\ln(2)\lambda}{(w_k+\delta_k)}\Psi_{m,i}^{(k)}}$.

AF Half-Duplex Relaying, i.e., $t=4$,

$$P_{m,i}^{(4,k)*} = \Big[\frac{w_k+\delta_k}{2\lambda\ln(2)} - \frac{\Big(\sqrt{\gamma_{SR_m,i}}+\sqrt{\gamma_{RD_m,i}^{(k)}}\Big)^2}{\gamma_{SR_m,i}\gamma_{RD_m,i}^{(k)}}\Big]^+, \quad (28)$$

$$P_{RD_m,i}^{(4,k)*} = \Big[P_{m,i}^{(4,k)*}\frac{\gamma_{SR_m,i} - \sqrt{\gamma_{SR_m,i}\gamma_{RD_m,i}^{(k)}}}{\gamma_{SR_m,i} - \gamma_{RD_m,i}^{(k)}}\Big]^+. \quad (29)$$

On the other hand, $P_{SR_m,i}^{(t,k)*}$ can be calculated from $P_{SR_m,i}^{(t,k)*} = P_{m,i}^{(t,k)*} - P_{RD_m,i}^{(t,k)*}, \forall t$. To obtain the optimal subcarrier allocation, we calculate the marginal benefit $Q_{m,i}^{(t,k)}$ [12] of each transmission strategy by solving $\frac{\partial\mathcal{L}_m}{\partial s_{m,i}^{(t,k)}}|_{P_{m,i}^{(t,k)}=P_{m,i}^{*(t,k)}} = Q_{m,i}^{(t,k)} - \frac{\beta_i}{w_k+\delta_k}$, which yields

$$Q_{m,i}^{(1,k)} = C_{m,i}^{(1,k)*} - \frac{P_{m,i}^{(1,k)*}\gamma_{RD_m,i}^{(k)}\gamma_{SR_m,i}/(1+\Gamma_{m,i}^{(1,k)*})}{\ln(2)(2\Gamma_{m,i}^{(1,k)*}\gamma_{LI_m,i} + \Upsilon_{m,i}^{(k)})}, \quad (30)$$

$$Q_{m,i}^{(t,k)} = C_{m,i}^{(t,k)*} - \frac{\Gamma_{m,i}^{(t,k)*}}{2\ln(2)(1+\Gamma_{m,i}^{(t,k)*})},\ t=2,3,4, \quad (31)$$

where $C_{m,i}^{(t,k)*}$ and $\Gamma_{m,i}^{(t,k)*}$ are obtained by substituting the optimal powers into (13)-(16). Thus, the subcarrier selection determined by relay m is given by

$$s_{m,i}^{(t,k)*} = \begin{cases} 1 & \text{if } Q_{m,i}^{(t,k)} = \max\limits_{a,b} Q_{m,i}^{(a,b)} \text{ and } Q_{m,i}^{(a,b)} \geq \frac{\beta_i}{w_b+\delta_b}\\ 0 & \text{otherwise.}\end{cases} \quad (32)$$

The dual variables β_i act as the global price in using subcarrier i in the system. On the other hand, dual variables δ_k and w_k not only affect the power allocation by changing the water-level of user k in (24)–(29), but also force the scheduler to assign more subcarriers to *delay sensitive* users and higher priority users by lowering the price in the selection process. It can be seen from (24)–(29) and (32) that relay m, $m\in\{1,\dots,M\}$, only requires the CSI of its own BS-to-relay link, the CSI of the relay-to-user links of the corresponding users, and the dual variables λ, β_i, $i\in\{1,\dots,n_F\}$, and δ_k, $k\in\{1,\dots,K\}$, supplied by the BS.

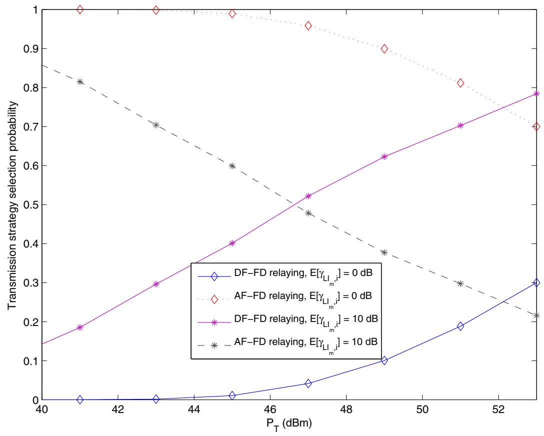

Fig. 2. Transmission strategy selection probability versus total transmit power for $K = 15$ users and low to moderate loop interference, i.e., $E[\gamma_{LI_m,i}] \leq 10$ dB.

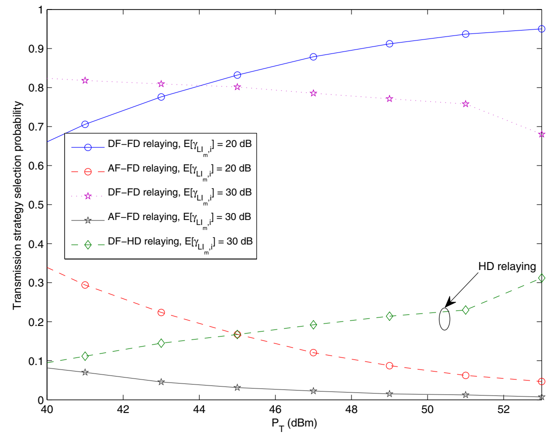

Fig. 3. Transmission strategy selection probability versus total transmit power for $K = 15$ users and different values of strong loop interference, i.e., $E[\gamma_{LI_m,i}] \geq 20$ dB.

F. Solution of the Master Dual Problem at the BS

For solving the master problem at the BS, each relay calculates the local resource allocations and passes this information, i.e., $C_{m,i}^{(t,k)*}, P_{SR_m,i}^{(t,k)*}$, and $P_{RD_m,i}^{(t,k)*}$, to the BS. The gradient method can be used to solve the minimization in the master problem in (19). The solution is given by

$$\beta_i(n+1) = \left[\beta_i(n) - \xi_1(n)\Big(1 - \sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{t=1}^{4} s_{m,i}^{(t,k)}\Big)\right]^+, \ \forall i$$

$$\lambda(n+1) = \left[\lambda(n) - \xi_2(n)\Big(P_T - \sum_{m=1}^{M}\sum_{k\in\mathcal{U}_m}\sum_{i=1}^{n_F}\sum_{t=1}^{4} \widetilde{P}_{m,i}^{(t,k)}\Big)\right]^+,$$

$$\delta_k(n+1) = \left[\delta_k(n) - \xi_3(n)\Big(\rho^{(k)} - R^{(k)}\Big)\right]^+, \ \forall k \in \mathcal{D} \qquad (33)$$

where n is the iteration index and $\xi_1(n)$, $\xi_2(n)$, and $\xi_3(n)$ are positive step sizes. Convergence to the optimal solution is guaranteed if the chosen step sizes satisfy the general conditions stated in [11, Chapter 1.2]. By combining the gradient update equations at the BS and the subcarrier selection criterion in (32) at the relays, a selected subcarrier will be occupied by one user with one transmission strategy only eventually.

IV. Results

In this section, we evaluate the system performance using simulations. Each cell is modeled as two concentric ring-shaped discs. The outer boundary and the inner boundary have radii of 1 km and 500 m, respectively. There are $M = 3$ relays 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. Unless specified otherwise, there are always 3 *delay sensitive* users with data rate requirement $R^{(k)} = 1$ bit/s/Hz in the system, while the remaining users are *non-delay sensitive*. The number of subcarriers is $n_F = 64$ with carrier center frequency 2.5 GHz, system bandwidth $\mathcal{B} = 5$ MHz, and $w_k = 1, \forall k$. Each subcarrier has a bandwidth of 78 kHz. The 3GPP path loss model is used. The small scale fading coefficients of the BS-to-relay links are generated as independent and identically distributed (i.i.d.) Rician random variables with $\kappa = 6$ dB, while the small scale fading coefficients of the relay-to-user links are i.i.d. Rayleigh fading. The weighted average system throughput is obtained by counting the number of packets which are successfully decoded by the users averaged over both macroscopic and microscopic fading.

A. Transmission Strategy Selection Probability

Figures 2 and 3 illustrate the transmission strategy selection probability versus P_T for different loop interference power levels. The number of iterations for the proposed distributed resource allocation algorithm is 5. For low to moderate loop interference power, i.e., $E[\gamma_{LI_m,i}] \leq 10$ dB, as demonstrated in Figure 2, AF-FD relaying and DF-FD relaying, i.e., $t = 1, 3$, dominate the system performance. In the low transmit power regime, AF-FD relaying is preferred over DF-FD relaying as the latter is always limited by the bottleneck link as shown in (6). However, when the total power in the system increases, DF-FD relaying becomes more attractive since AF-FD relaying suffers from strong interference amplification. On the other hand, when the loop interference is strong as depicted in Figure 3, DF-HD relaying comes into the picture and the selection probability of AF-FD relaying drops dramatically, since AF-FD relaying is sensitive to loop interference. Besides, it can be observed that DF-HD relaying is more attractive than AF-HD in the considered practical operating region as the DF protocol avoids noise amplification by signal regeneration.

B. Average System Throughput versus Transmit Power

Figure 4 depicts the average weighted system throughput versus the total transmit power for a total of 15 users with loop interference power $E[\gamma_{LI_m,i}] = 10$ dB. The proposed distributed algorithm with 5 iterations is compared with the optimal centralized resource allocation and scheduling algorithm. The resource allocation algorithms for AF-FD relaying and DF-FD relaying can be obtained by restricting the hybrid relaying resource allocation algorithm to select the corresponding objective function only. These algorithms are compared with two baseline schemes used in the literature, namely, DF-HD relaying and AF-HD relaying. It can be observed that the performance of the proposed distributed algorithms closely approaches that of the optimal centralized algorithm in all considered cases even after only 5 iterations, which confirms the practicality of the distributed algorithm. On the other hand, it is not surprising that both AF-FD relaying and DF-FD relaying have a close-to-optimal performance since FD relaying dominates the system performance. On the other hand, the two baseline schemes are always worse than the proposed AF-FD, DF-FD, and hybrid relaying schemes. This is because despite the loop interference, FD relays have a better spectral efficiency by allowing the BS and the relays to transmit simultaneously in two phases, while the HD baseline scheme relays use two phases to transmit one message.

C. Average System Throughput versus Number of Users

Figure 5 demonstrates the average weighted system throughput versus the number of users with different user data rate requirements for transmit power $P_T = 53$ dBm. The number of iterations for the proposed algorithm for both FD relaying and hybrid relaying is 5. It can be observed that the average system

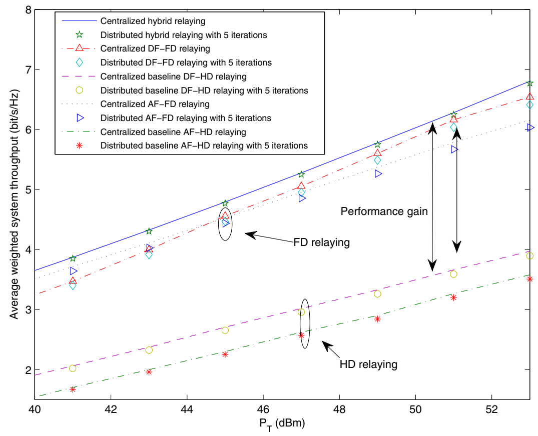

Fig. 4. Average weighted system throughput versus total transmit power for loop interference power $E[\gamma_{LI_{m,i}}] = 10$ dB and different resource allocation and scheduling algorithms with $K = 15$.

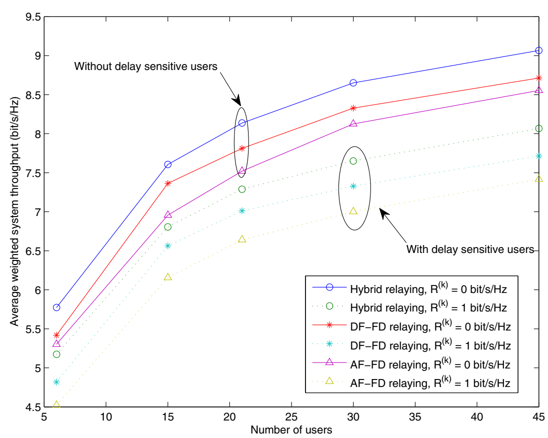

Fig. 5. Average weighted system throughput versus number of users with different data rate requirements for transmit power $P_T = 53$ dBm and loop interference power $E[\gamma_{LI_{m,i}}] = 10$ dB.

throughput for both relaying schemes increase with the number of users. This is because as the number of users increase, the proposed distributed algorithm has a higher chance to select users who have strong channels in both hops. This effect is known as multi-user diversity (MUD). However, the scheduler looses degrees of freedom for scheduling and resource allocation as the data rate requirements of *delay sensitive* users become more stringent, since most of the resources are taken by them regardless of their possibly poor channel qualities.

V. Conclusions

In this paper, we formulated the dynamic resource allocation and scheduling design for OFDMA systems with full duplex relays as a non-convex and combinatorial optimization problem, in which heterogeneous users and hybrid relaying were taken into consideration. An efficient iterative distributed resource allocation algorithm with closed-form power and subcarrier allocation policies was derived by dual decomposition. Simulation results not only showed that the performance of the proposed algorithm approaches the optimal performance within a small number of iterations, but also demonstrated the possible performance gains obtained by FD relaying compared to HD relaying.

Appendix-Proof of Proposition 1

Due to page limitation, we provide only a sketch of the proof. For $t = 2$, i.e., DF-HD relaying, the channel capacity is maximized when the amounts of received information at the relay and the destination are the same, i.e., $\gamma_{SR_{m,i}} P_{SR_{m,i}}^{(2,k)} = P_{RD_{m,i}}^{(2,k)} \gamma_{RD_{m,i}}^{(k)}$. By combing this fact with $P_{m,i}^{(t,k)} = P_{SR_{m,i}}^{(t,k)} + P_{RD_{m,i}}^{(t,k)}$, we can express $P_{RD_{m,i}}^{(t,k)}$ in terms of $P_{m,i}^{(t,k)}$ and the channel coefficients, i.e., $P_{RD_{m,i}}^{(2,k)} = \frac{P_{m,i}^{(2,k)}\gamma_{SR_{m,i}}}{\gamma_{SR_{m,i}}+\gamma_{RD_{m,i}}^{(k)}}$. Then, we substitute the new expression for $P_{RD_{m,i}}^{(2,k)}$ back into the original capacity equation (7), and the result follows immediately. Equation (13) for $t = 1$ can be obtained in a similar manner. On the other hand, for the AF protocol in $t = 3, 4$, we assume a high SNR in order to obtain a tractable result. Then, the SINR/SNR for $t = 3, 4$ can be approximated as

$$\Gamma_{m,i}^{(3,k)} \approx \frac{\frac{P_{SR_{m,i}}^{(3,k)}\gamma_{SR_{m,i}}}{P_{RD_{m,i}}^{(3,k)}\gamma_{LI_{m,i}}} P_{RD_{m,i}}^{(3,k)}\gamma_{RD_{m,i}}^{(k)}}{P_{RD_{m,i}}^{(3,k)}\gamma_{RD_{m,i}}^{(k)} + \frac{P_{SR_{m,i}}^{(3,k)}\gamma_{SR_{m,i}}}{P_{RD_{m,i}}^{(3,k)}\gamma_{LI_{m,i}}}}, \quad (34)$$

$$\Gamma_{m,i}^{(4,k)} \approx \frac{\gamma_{SR_{m,i}} P_{SR_{m,i}}^{(4,k)} P_{RD_{m,i}}^{(4,k)}\gamma_{RD_{m,i}}^{(k)}}{\gamma_{SR_{m,i}} P_{SR_{m,i}}^{(4,k)} + P_{RD_{m,i}}^{(4,k)}\gamma_{RD_{m,i}}^{(k)}}, \quad (35)$$

respectively. Furthermore, we use $P_{SR_{m,i}}^{(t,k)} + P_{RD_{m,i}}^{(t,k)} = P_{m,i}^{(t,k)}$ in (34) and (35) and take the derivative of $\Gamma_{m,i}^{(t,k)}$ w.r.t. $P_{SR_{m,i}}^{(t,k)}$ for $t = 3, 4$. Again, we can express $P_{RD_{m,i}}^{(t,k)}$ in terms of $P_{m,i}^{(t,k)}$ and the channel coefficients, e.g., $P_{RD_{m,i}}^{(4,k)} = \frac{P_{m,i}^{(4,k)}(\gamma_{SR_{m,i}} - \sqrt{\gamma_{SR_{m,i}}\gamma_{RD_{m,i}}^{(k)}})}{\gamma_{SR_{m,i}} - \gamma_{RD_{m,i}}^{(k)}}$ for $t = 4$. Then, (16) can be obtained by substituting the new expression for $P_{RD_{m,i}}^{(t,k)}$ into (34) for $t = 4$. For deriving (15) in case of $t = 3$, we follow the same steps as for $t = 4$, and further approximate the resulting SINR as

$$\Gamma_{m,i}^{(3,k)} \approx \frac{(P_{m,i}^{(3,k)})^2 \gamma_{SR_{m,i}}\gamma_{RD_{m,i}}^{(k)}}{P_{m,i}^{(3,k)}\gamma_{SR_{m,i}} + 2\sqrt{(P_{m,i}^{(3,k)})^3 \Psi_{m,i}^{(k)}}} \quad (36)$$

$$\approx \frac{\sqrt{\gamma_{RD_{m,i}}^{(k)}\gamma_{SR_{m,i}} P_{m,i}^{(3,k)}}}{2\sqrt{\gamma_{LI_{m,i}}}}, \text{for} \quad P_{m,i}^{(3,k)} \to \infty \quad (37)$$

to make the power allocation tractable.

References

[1] M. K. Awad and S. Xuemin, "OFDMA Based Two-Hop Cooperative Relay Network Resources Allocation," in *Proc. IEEE Intern. Conf. on Commun.*, May 2008, pp. 4414–4418.

[2] D. W. K. Ng and R. Schober, "Cross-Layer Scheduling for OFDMA Amplify-and-Forward Relay Networks," *IEEE Trans. Veh. Technol.*, vol. 59, pp. 1443–1458, Mar 2010.

[3] W. Nam, W. Chang, S.-Y. Chung, and Y. H. Lee, "Transmit Optimization for Relay-Based Cellular OFDMA Systems," in *Proc. IEEE Intern. Commun. Conf.*, Jun 2007, pp. 5714 – 5719.

[4] "Full Duplex Configuration of Un and Uu Subframes for Type I Relay," 3GPP TSG RAN WG1 R1-100139, Tech. Rep., Jan 2010.

[5] "Text Proposal on Inband Full Duplex Relay for TR 36.814," 3GPP TSG RAN WG1 R1-101659, Tech. Rep., Feb 2010.

[6] H. Ju, E. Oh, and D. Hong, "Catching Resource-Devouring Worms in Next-Generation Wireless Relay Systems: Two-Way Relay and Full-Duplex Relay," *IEEE Commun. Magazine*, vol. 47, pp. 58 – 65, Sep. 2009.

[7] ——, "Improving Efficiency of Resource Usage in Two-Hop Full Duplex Relay Systems Based on Resource Sharing and Interference Cancellation," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 3933 – 3938, Aug. 2009.

[8] T. Riihonen, S. Werner, and R. Wichman, "Optimized Gain Control for Single-Frequency Relaying with Loop Interference," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 2801 – 2806, Jun. 2009.

[9] C. Y. Wong, R. S. Cheng, K. B. Lataief, and R. D. Murch, "Multiuser OFDM with Adaptive Subcarrier, Bit, and Power Allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, pp. 1747–1758, Oct 1999.

[10] W. Yu and R. Lui, "Dual Methods for Nonconvex Spectrum Optimization of Multicarrier Systems," *IEEE Trans. Commun.*, vol. 54, pp. 1310 – 1321, July 2006.

[11] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Athena Scientific, 1999.

[12] W. Yu and J. M. Cioffi, "FDMA Capacity of Gaussian Multiple-Access Channels with ISI," *IEEE Trans. Commun.*, vol. 50, pp. 102 – 111, Jan 2002.