

Resource Allocation and Scheduling in Multi-Cell OFDMA Decode-and-Forward Relaying Networks

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 multi-cell orthogonal frequency division multiple access (OFDMA) half-duplex decode-and-forward (DF) relay assisted networks. Our problem formulation takes into account multi-cell interference and heterogeneous data rate requirements. We transform the resulting non-convex and combinatorial optimization problem into a standard convex optimization problem by introducing an additional interference constraint, which constitutes a performance lower bound for the original problem. Subsequently, the transformed optimization problem is solved by using dual decomposition and a centralized iterative resource allocation algorithm with closed-form power and subcarrier allocation policies is derived maximizing the average weighted system throughput (bit/s/Hz/base station). Simulation results illustrate that our proposed algorithm approaches the maximum achievable throughput and provides substantial performance gains compared to single-cell optimization.

I. INTRODUCTION

Orthogonal frequency division multiple access (OFDMA) is an important technique for high speed wireless multiuser communication networks, such as IEEE 802.22 Wireless Regional Area Networks (WRAN) and IEEE 802.16 Worldwide Interoperability for Microwave Access (WiMAX). In a single-cell OFDMA system, maximum system spectral efficiency can be achieved by selecting the best user for each subcarrier and adapting the corresponding power. On the other hand, a large amount of work has been devoted to cooperative relaying for wireless networks as it provides coverage extension and throughput gains [1]-[4]. Several efficient relaying protocols such as decode-and-forward (DF) and amplify-and-forward (AF) have been proposed in the literature to facilitate relaying.

Future broadband wireless communication systems are expected to provide heterogeneous data rate services with certain quality of service (QoS) requirements. The combination of OFDMA and DF relaying provides a possible solution for meeting these demanding requirements, particularly for users at the cell edge. In [1]-[4], best effort resource allocation and scheduling for homogeneous users in DF OFDMA systems are studied for different system configurations. However, in practice, users are heterogeneous with different QoS requirements such as the minimum required data rate, which best effort resource allocation cannot fulfill. Furthermore, [1]-[4] focus on single-cell systems by assuming that the co-channel interference caused by adjacent cells is negligible. However, aggressive frequency reuse with interference coordination techniques are a new trend for next generation communication systems since they achieve higher system capacity [5] and existing works considering a single-cell only may not be able to reveal the actual performance of a practical system. Base station coordination by only sharing channel state information (CSI) has been an active research area in the last few years as a major technique to mitigate co-channel interference. Resource allocation with base station coordination for single-hop multi-cell networks with single-carrier and multi-carrier modulation is considered in [6] and [7], [8], respectively. In [9], the sum rate performance of a time-division multiple access (TDMA) multi-cell system with half-duplex AF relays is studied. Yet, in all these works, fairness of users is not taken into account for the resource allocation which may result in starvation of weak cell edge users.

In this paper, we formulate the scheduling and resource allocation problem for multi-cell OFDMA DF relaying systems as an optimization problem with individual power constraints for the base stations and relays. The original non-convex, combinatorial problem is transformed into a convex one by imposing an additional maximum interference allowance constraint. By using dual decomposition, the transformed problem is separated into one master problem and several subproblems which leads to a low complexity iterative resource allocation algorithm.

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

II. MULTI-CELL OFDMA DF RELAY NETWORK MODEL

A. Multi-Cell System Model

We consider a multi-cell OFDMA DF half-duplex relay network which consists of a total of P base stations (BSs), M relays, and K mobile users belonging 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 have no data rate constraint. All transceivers are equipped with single-antennas. We assume universal frequency reuse and the P BSs share total bandwidth \mathcal{B} . Each cell is modeled by two concentric ring-shaped discs as shown in Figure 1. In this paper, we focus on resource allocation and scheduling with interference coordination 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 a separated resource for those users who do not need a relay. In the considered model, there is no direct link between the BSs and the users due to heavy blockage and long distance transmission. BSs are connected to a centralized unit via optical fiber backhaul links. Each relay only serves one BS and each user is only served by one relay and one BS. The channel is assumed to be time invariant within a transmission frame, but varying from one frame to another. The channel state information is assumed to be perfectly known at the centralized unit. The downlink transmission from the BSs to the users via the relays is accomplished in two phases. In the first phase, the BSs transmit their signals to the relay stations. Then, in the second phase, the relay stations decode the previously received signals and forward them to the corresponding users. Intra-cell interference does not exist since each subcarrier is only occupied by one user in each cell.

B. OFDMA Relay Channel Model

We consider an OFDMA system with n_F subcarriers. The channel impulse response is assumed to be time-invariant within each frame. Suppose user k is served by BS $p \in \{1, \dots, P\}$ and relay $m \in \{1, \dots, M\}$. The received symbol in the first

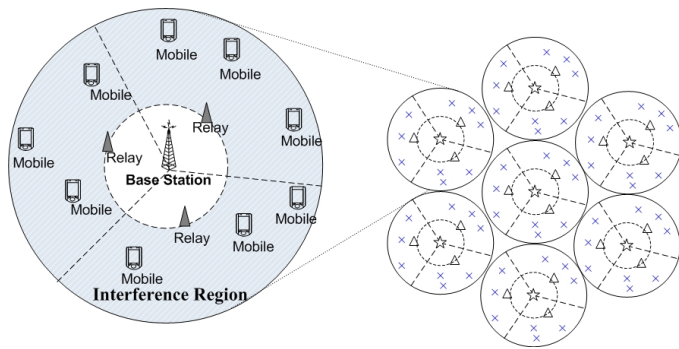


Fig. 1. A multi-cell system with $P = 7$ cells and $M = 21$ relays. There are in total $K = 63$ users in the system. Each cell is modeled by two concentric ring-shaped discs. The shaded part represents the multi-cell interference region.

phase at relay m for user k in subcarrier $i \in \{1, \dots, n_F\}$ is given by

$$Y_{R_{m,p}}^{(k)}(i) = \sqrt{P_{B_p,R_{m,p}}^{(k)}(i)l_{B_p,R_{m,p}}H_{B_p,R_{m,p}}(i)}X^{(k)}(i) + I_{R_{m,p}}(i) + Z_{R_{m,p}}(i), \quad (1)$$

where $P_{B_p,R_{m,p}}^{(k)}(i)$ and $X^{(k)}(i)$ are the transmit power and the transmit symbol for the link between BS p and relay m in subcarrier i in the first phase, respectively. $I_{R_{m,p}}(i)$ is the received multi-cell interference in subcarrier i at relay m . $l_{B_p,R_{m,p}}$ represents the path loss between BS p and relay m . $Z_{R_{m,p}}(i)$ is the additive white Gaussian noise (AWGN) in subcarrier i at relay m with zero mean and variance σ_z^2 . $H_{B_p,R_{m,p}}(i)$ is the small scale fading coefficient between the BS and relay m in subcarrier i . Relay m decodes the message $X^{(k)}(i)$ and forwards it to user k . Therefore, the signal received at user k in subcarrier i from relay m is given by

$$Y_{U_{m,p}}^{(k)}(i) = \sqrt{P_{R_{m,p}}^{(k)}(i)l_{R_{m,p}}^{(k)}H_{R_{m,p}}^{(k)}(i)}X^{(k)}(i) + I^{(k)}(i) + Z^{(k)}(i). \quad (2)$$

Variables $P_{R_{m,p}}^{(k)}(i)$, $l_{R_{m,p}}^{(k)}$, and $H_{R_{m,p}}^{(k)}(i)$ are defined in a similar manner as the corresponding variables for the BS-to-relay links except that the signalling direction is from relay m in cell p to user k . $Z^{(k)}(i)$ is the AWGN in subcarrier i at user k with zero mean and variance σ_z^2 . $I^{(k)}(i)$ is the received multi-cell interference of user k in subcarrier i and its variance is given by

$$\begin{aligned} \sigma_k^2(i) &= \mathcal{E}\{|I^{(k)}(i)|^2\} \\ &= \sum_{\substack{c=1 \\ c \neq p}}^P \sum_{a \in \mathcal{R}_c} \sum_{\substack{j=1 \\ j \neq k}}^{n_F} s_{a,c}^{(j)}(i) P_{R_{a,c}}^{(j)}(i) \left(l_{R_{a,c}}^{(k)} |H_{R_{a,c}}^{(k)}(i)|^2 \right) \end{aligned} \quad (3)$$

where $\mathcal{E}\{\cdot\}$ denotes statistical expectation. $s_{a,c}^{(j)}(i) \in \{0, 1\}$ is the subcarrier allocation indicator. $l_{R_{a,c}}^{(k)}$ and $H_{R_{a,c}}^{(k)}(i)$ are the path loss and small scale fading in subcarrier i between relay a in cell c and user k , respectively. \mathcal{R}_c is a set of relays which belong to BS c . In practice, users are located at random positions and hence a non-line-of-sight (NLoS) communication link is expected between the relays and the users. Thus, we model $H_{R_{m,p}}^{(k)}(i)$ as Rayleigh distributed, i.e., $H_{R_{m,p}}^{(k)}(i) \sim \mathcal{CN}(0, 1)$, where $\mathcal{CN}(\mu, \sigma^2)$ denotes a complex Gaussian random variable with mean μ and variance σ^2 . On the other hand, a strong line-of-sight (LoS) propagation channel is expected between the BS and the relays, since they are placed in relatively high positions in practice and the number of blockages between them are limited. Hence, $H_{B_p,R_{m,p}}(i)$ is modeled as Rician fading with Rician factor κ , i.e., $H_{B_p,R_{m,p}}(i) \sim \mathcal{CN}(\sqrt{\kappa/(1+\kappa)}, 1/(1+\kappa))$.

III. RESOURCE ALLOCATION AND SCHEDULING DESIGN

A. Instantaneous Channel Capacity and System Throughput

In this subsection, we introduce the adopted system performance measure. Given perfect CSI at the receiver (CSIR), the instantaneous channel capacity of the first hop between BS p and relay m for user k in using subcarrier i is given by

$$C_{B_p,R_{m,p}}^{(k)}(i) = \frac{1}{2} \log_2 \left(1 + \Gamma_{B_p,R_{m,p}}^{(k)}(i) \right) \quad (4)$$

with signal-to-interference-plus-noise ratio (SINR)

$$\Gamma_{B_p,R_{m,p}}^{(k)}(i) \approx \frac{P_{B_p,R_{m,p}}^{(k)}(i)l_{B_p,R_{m,p}}|H_{B_p,R_{m,p}}(i)|^2}{\sigma_z^2}, \quad (5)$$

where the pre-log factor $\frac{1}{2}$ in (4) is due the two channel uses required for transmitting one message and the approximation in (5) is because the channel capacity between the BS and the relay is assumed to be limited by channel noise. i.e., $\mathcal{E}\{|I_{R_{m,p}}(i)|^2\} \ll \sigma_z^2$. It can be verified by simulation that the multi-cell interference received at the relays is negligible compared with the noise variance for typical cell sizes, relay locations, and transmit powers. Similarly, the channel capacity for user k in using subcarrier i through relay m in cell p in the second hop is given by

$$C_{U_{m,p}}^{(k)}(i) = \frac{1}{2} \log_2 \left(1 + \Gamma_{U_{m,p}}^{(k)}(i) \right) \quad (6)$$

with received SINR

$$\Gamma_{U_{m,p}}^{(k)}(i) = \frac{P_{R_{m,p}}^{(k)}(i)l_{R_{m,p}}^{(k)}|H_{R_{m,p}}^{(k)}(i)|^2}{\sigma_k^2(i) + \sigma_z^2}. \quad (7)$$

Now, we define the instantaneous throughput (bit/s/Hz successfully delivered) for user k who is assigned to relay m and BS p as $\rho_{m,p}^{(k)} =$

$$\frac{1}{n_F} \min \left\{ \sum_{i=1}^{n_F} s_{m,p}^{(k)}(i) C_{B_p,R_{m,p}}^{(k)}(i), \sum_{i=1}^{n_F} s_{m,p}^{(k)}(i) C_{U_{m,p}}^{(k)}(i) \right\}. \quad (8)$$

The *average weighted system throughput* is defined as the total average number of bit/s/Hz/BS successfully decoded at the K users via the M relays and P BSs. It is given by

$$\mathcal{U}_{TP}(\mathcal{P}, \mathcal{S}) = \frac{1}{P} \sum_{p=1}^P \sum_{m \in \mathcal{R}_p} \sum_{k \in \mathcal{U}_{m,p}} w^{(k)} \rho_{m,p}^{(k)}, \quad (9)$$

where \mathcal{P} and \mathcal{S} are the power and subcarrier allocation policies, respectively. $\mathcal{U}_{m,p}$ is the set of users who are served by relay m and BS p . $w^{(k)}$ is a positive constant which can be used to enforce certain notions of fairness.

B. Problem Formulation

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

$$\begin{aligned} (\mathcal{P}^*, \mathcal{S}^*) &= \arg \max_{\mathcal{P}, \mathcal{S}} \mathcal{U}_{TP}(\mathcal{P}, \mathcal{S}) \\ \text{s.t. C1: } &\sum_{m \in \mathcal{R}_p} \sum_{k \in \mathcal{U}_{m,p}} \sum_{i=1}^{n_F} P_{B_p,R_{m,p}}^{(k)}(i) s_{m,p}^{(k)}(i) \leq P_{B_T}, \quad \forall p \\ \text{C2: } &\sum_{k \in \mathcal{U}_{m,p}} \sum_{i=1}^{n_F} P_{R_{m,p}}^{(k)}(i) s_{m,p}^{(k)}(i) \leq P_{R_T}, \quad \forall m, p \\ \text{C3: } &\rho_{m,p}^{(k)} \geq r^{(k)}, \quad \forall k \in \mathcal{D}_{m,p} \end{aligned}$$

$$\begin{aligned}
 \text{C4: } & \sum_{m \in \mathcal{R}_p} \sum_{k \in \mathcal{U}_{m,p}} s_{m,p}^{(k)}(i) = 1, \quad \forall p, i \\
 \text{C5: } & s_{m,p}^{(k)}(i) \in \{0, 1\}, \quad \forall m, p, k, i \\
 \text{C6: } & P_{B_p, R_{m,p}}^{(k)}(i), P_{R_{m,p}}^{(k)}(i) \geq 0, \quad \forall m, p, i, k
 \end{aligned} \quad (10)$$

where $\mathcal{D}_{m,p}$ is the set of *delay sensitive* users who are served by relay m and BS p . Here, C1 (C2) represents the individual power constraint for each BS (relay). C3 specifies the minimum required data rate for *delay sensitive* users which are chosen by the application layer. Constraints C4 and C5 are imposed to guarantee that each subcarrier is only used by one user in each cell. In other words, intra-cell interference is avoided in the system.

C. Transformation of the Optimization Problem

The considered problem is a mixed combinatorial and non-convex optimization problem which is NP hard. In order to make the problem tractable, we perform a three-step transformation to simplify the problem. The first step is to relax the subcarrier selection variable $s_{m,p}^{(k)}(i)$ in C5 to be a real value between zero and one instead of a Boolean, i.e., $0 \leq s_{m,p}^{(k)}(i) \leq 1$, which is known as time-sharing. The time-sharing relaxation is asymptotically optimal for a large number of subcarriers [10]. In addition, we introduce two new variables and define them as $\tilde{P}_{B_p, R_{m,p}}^{(k)}(i) = P_{B_p, R_{m,p}}^{(k)}(i) s_{m,p}^{(k)}(i)$ and $\tilde{P}_{R_{m,p}}^{(k)}(i) = P_{R_{m,p}}^{(k)}(i) s_{m,p}^{(k)}(i)$. Secondly, we add an additional constraint C7 to the original problem:

$$\text{C7: } s_{m,p}^{(k)}(i) \tilde{\sigma}_k^2(i) \leq I \quad \forall k, i, m, p, \quad (11)$$

$\tilde{\sigma}_k^2(i) = \sigma_k^2(i) |_{\tilde{P}_{R_{m,p}}^{(k)}(i) = P_{R_{m,p}}^{(k)}(i) s_{m,p}^{(k)}(i)}$. C7 is the maximum multi-cell interference allowance in each subcarrier. By varying¹ the value of I , the resource allocator is able to control the maximum amount of interference in each subcarrier of each user. We incorporate constraint C7 into equation (7). As a result, the channel capacities of the first and second hop for user k through relay m in subcarrier i in cell p are given by $\tilde{C}_{B_p, R_{m,p}}^{(k)}(i) = \frac{1}{2} \log_2 \left(1 + \frac{\tilde{\Gamma}_{B_p, R_{m,p}}^{(k)}(i)}{s_{m,p}^{(k)}(i)} \right)$ and $\tilde{C}_{U_{m,p}}^{(k)}(i) = \frac{1}{2} \log_2 \left(1 + \frac{\tilde{P}_{R_{m,p}}^{(k)}(i) |H_{R_{m,p}}^{(k)}(i)|^2}{(I + \sigma_z^2) s_{m,p}^{(k)}(i)} \right)$, respectively, where $\tilde{\Gamma}_{B_p, R_{m,p}}^{(k)}(i) = \Gamma_{B_p, R_{m,p}}^{(k)}(i) |_{P_{B_p, R_{m,p}}^{(k)}(i) = \tilde{P}_{B_p, R_{m,p}}^{(k)}(i) / s_{m,p}^{(k)}(i)}$. Note that the actual channel capacity is larger than the scheduled capacity in the second hop, i.e., $C_{U_{m,p}}^{(k)}(i) \geq \tilde{C}_{U_{m,p}}^{(k)}(i)$. In fact, $\tilde{C}_{U_{m,p}}^{(k)}(i)$ can be viewed as the worst case capacity for resource allocation in the second hop. Finally, the max-min formulation in the objective function can be handled by introducing auxiliary variables $z_{m,p}^{(k)}$, $m \in \{1, \dots, M\}$, $p \in \{1, \dots, P\}$, $k \in \{1, \dots, K\}$. Then, the original optimization problem is transformed into

$$\begin{aligned}
 & \arg \max_{\mathcal{P}, \mathcal{S}, z_{m,p}^{(k)}} \sum_{p=1}^P \sum_{m \in \mathcal{R}_p} \sum_{k \in \mathcal{U}_{m,p}} w^{(k)} z_{m,p}^{(k)} \\
 & \text{s.t. } \text{C4, C6, C7} \\
 \text{C1: } & \sum_{m \in \mathcal{R}_p} \sum_{k \in \mathcal{U}_{m,p}} \sum_{i=1}^{n_F} \tilde{P}_{B_p, R_{m,p}}^{(k)}(i) \leq P_{B_T}, \quad \forall p
 \end{aligned}$$

¹The maximum multi-cell interference allowance variable I is not an optimization variable in the proposed framework.

$$\begin{aligned}
 \text{C2: } & \sum_{k \in \mathcal{U}_{m,p}} \sum_{i=1}^{n_F} \tilde{P}_{R_{m,p}}^{(k)}(i) \leq P_{R_T}, \quad \forall m, p \\
 \text{C3: } & \tilde{\rho}_{m,p}^{(k)} \geq r^{(k)}, \quad \forall k \in \mathcal{D}_{m,p} \\
 \text{C5: } & 0 \leq s_{m,p}^{(k)}(i) \leq 1, \quad \forall m, p, k, i \\
 \text{C8: } & \sum_{i=1}^{n_F} s_{m,p}^{(k)}(i) \tilde{C}_{B_p, R_{m,p}}^{(k)}(i) \geq z_{m,p}^{(k)}, \quad \forall p, m, k \\
 \text{C9: } & \sum_{i=1}^{n_F} s_{m,p}^{(k)}(i) \tilde{C}_{U_{m,p}}^{(k)}(i) \geq z_{m,p}^{(k)}, \quad \forall p, m, k,
 \end{aligned} \quad (12)$$

where $\tilde{\rho}_{m,p}^{(k)} = \rho_{m,p}^{(k)} |_{C_{U_{m,p}}^{(k)}(i) = \tilde{C}_{U_{m,p}}^{(k)}(i)}$. Note that the constant term $\frac{1}{n_F \times P}$ is removed from the transformed objective function for notation simplicity as it does not affect the values of the arguments which maximize the objective function. The transformed problem is now jointly concave with respect to the optimization variables. Thus, the considered optimization problem can be solved in its dual domain [11].

D. Dual Problem Formulation

In this subsection, the transformed optimization problem is solved by *Lagrange dual decomposition*. For this purpose, we first need the Lagrangian function of the primal problem. Upon rearranging terms, we obtain the Lagrangian $\mathcal{L}(\lambda, \beta, \gamma, \mu, \nu, \theta, \delta, \eta, \mathcal{P}, \mathcal{S}, z_{m,p}^{(k)}) =$

$$\begin{aligned}
 & \sum_{p=1}^P \sum_{m \in \mathcal{R}_p} \sum_{k \in \mathcal{U}_{m,p}} (w^{(k)} - (\mu_{m,p}^{(k)} + \nu_{m,p}^{(k)})) z_{m,p}^{(k)} \\
 & + \sum_{p=1}^P \sum_{m \in \mathcal{R}_p} \sum_{k \in \mathcal{U}_{m,p}} \sum_{i=1}^{n_F} s_{m,p}^{(k)}(i) \times \left((\eta^{(k)} + \mu_{m,p}^{(k)}) \tilde{C}_{B_p, R_{m,p}}^{(k)}(i) \right. \\
 & \left. + (\delta^{(k)} + \nu_{m,p}^{(k)}) \tilde{C}_{U_{m,p}}^{(k)}(i) \right) + \sum_{p=1}^P \sum_{m \in \mathcal{R}_p} \gamma_{m,p} P_{R_T} \\
 & - \sum_{p=1}^P \lambda_p \left[\sum_{m \in \mathcal{R}_p} \sum_{k \in \mathcal{U}_{m,p}} \sum_{i=1}^{n_F} \tilde{P}_{B_p, R_{m,p}}^{(k)}(i) - P_{B_T} \right] \\
 & - \sum_{p=1}^P \sum_{m \in \mathcal{R}_p} \gamma_{m,p} \left[\sum_{k \in \mathcal{U}_{m,p}} \sum_{i=1}^{n_F} \tilde{P}_{R_{m,p}}^{(k)}(i) \right] - \sum_{k \in \mathcal{D}_{m,p}} r^{(k)} (\eta^{(k)} + \delta^{(k)}) \\
 & - \sum_{p=1}^P \sum_{i=1}^{n_F} \beta_p(i) \left[\sum_{m \in \mathcal{R}_p} \sum_{k \in \mathcal{U}_{m,p}} s_{m,p}^{(k)}(i) - 1 \right] \\
 & - \sum_{p=1}^P \sum_{m \in \mathcal{R}_p} \sum_{i=1}^{n_F} \sum_{k \in \mathcal{U}_{m,p}} \theta_{m,p}^{(k)}(i) (s_{m,p}^{(k)}(i) \tilde{\sigma}_k^2(i) - I),
 \end{aligned} \quad (13)$$

where λ is the Lagrange multiplier vector associated with the individual BS power constraints with elements λ_p , $p \in \{1, \dots, P\}$. γ is the Lagrange multiplier vector corresponding to the individual relay power constraints with elements $\gamma_{m,p}$, $m \in \{1, \dots, M\}$. δ and η are the Lagrange multiplier vectors corresponding to the data rate constraint in the two hops and have elements $\delta^{(k)}$ and $\eta^{(k)}$, $k \in \{1, \dots, K\}$, respectively. $\delta^{(k)} = 0$ and $\eta^{(k)} = 0$ for *non-delay sensitive* users. Lagrange multiplier vector β with elements $\beta_p(i)$, $i \in \{1, \dots, n_F\}$, is connected to the subcarrier usage constraints. μ and ν are the Lagrange multiplier vectors for constraints C8 and C9 in (12) with elements $\mu_{m,p}^{(k)}$ and $\nu_{m,p}^{(k)}$, respectively. θ is the Lagrange multiplier vector for the maximum received interference allowance constraint in each subcarrier with elements $\theta_{m,p}^{(k)}(i)$. The boundary constraints C5 and C6 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

$$\min_{\lambda, \beta, \gamma, \mu, \nu, \theta, \delta, \eta \geq 0} \max_{\mathcal{P}, \mathcal{S}, z_{m,p}^{(k)}} \mathcal{L}(\lambda, \beta, \gamma, \mu, \nu, \theta, \delta, \eta, \mathcal{P}, \mathcal{S}, z_{m,p}^{(k)}). \quad (14)$$

In general, the above dual problem is unbounded if $z_{m,p}^{(k)} \rightarrow \infty$. Consider the parts of the dual function in the inner maximization which are related to $z_{m,p}^{(k)}$:

$$\begin{aligned} & \max_{z_{m,p}^{(k)}} \sum_{p=1}^P \sum_{m \in \mathcal{R}_p} \sum_{k \in \mathcal{U}_{m,p}} (w^{(k)} - (\mu_{m,p}^{(k)} + \nu_{m,p}^{(k)})) z_{m,p}^{(k)} \\ & = \begin{cases} 0 & \text{if } \mu_{m,p}^{(k)} + \nu_{m,p}^{(k)} = w^{(k)} \\ \infty & \text{otherwise} \end{cases}. \end{aligned} \quad (15)$$

In order to have a bounded dual function, the dual variables $\mu_{m,p}^{(k)}$ and $\nu_{m,p}^{(k)}$ must satisfy $\mu_{m,p}^{(k)} + \nu_{m,p}^{(k)} = w^{(k)}$. Thus, the dual problem is simplified to

$$\min_{\lambda, \beta, \gamma, \mu, \theta, \delta, \eta \geq 0} \max_{\mathcal{P}, \mathcal{S}} \tilde{\mathcal{L}}(\lambda, \beta, \gamma, \mu, \theta, \delta, \eta, \mathcal{P}, \mathcal{S}), \quad (16)$$

where $\tilde{\mathcal{L}}(\lambda, \beta, \gamma, \mu, \theta, \delta, \eta, \mathcal{P}, \mathcal{S}) = \mathcal{L}(\lambda, \beta, \gamma, \mu, \nu, \theta, \delta, \eta, \mathcal{P}, \mathcal{S}, z_{m,p}^{(k)})|_{\nu_{m,p}^{(k)} = w^{(k)} - \mu_{m,p}^{(k)}}$. Note that the auxiliary variables $z_{m,p}^{(k)}$ vanish when we set $\nu_{m,p}^{(k)} = w^{(k)} - \mu_{m,p}^{(k)}$ in (13).

E. Dual Decomposition and Solution

By dual decomposition, the dual problem is decomposed into a master problem and $M \times P \times n_F$ subproblems with identical structures. The subproblems can be expressed as

$$\max_{\mathcal{P}, \mathcal{S}} \tilde{\mathcal{L}}_{m,p,i}(\lambda, \beta, \gamma, \mu, \theta, \delta, \eta, \mathcal{P}, \mathcal{S}) \quad (17)$$

with $\tilde{\mathcal{L}}_{m,p,i}(\lambda, \beta, \gamma, \mu, \theta, \delta, \eta, \mathcal{P}, \mathcal{S}) =$

$$\begin{aligned} & \sum_{k \in \mathcal{U}_{m,p}} s_{m,p}^{(k)}(i) \left((\eta^{(k)} + \mu_{m,p}^{(k)}) \tilde{C}_{B_p, R_{m,p}}^{(k)}(i) \right. \\ & \left. + (\delta^{(k)} + \nu_{m,p}^{(k)}) \tilde{C}_{U_{m,p}}^{(k)}(i) \right) + \gamma_{m,p} P_{RT} \\ & - \beta_p(i) \left[\sum_{k \in \mathcal{U}_{m,p}} s_{m,p}^{(k)}(i) - 1 \right] - \sum_{k \in \mathcal{D}} r^{(k)} (\eta^{(k)} + \delta^{(k)}) \\ & - \lambda_p \left[\sum_{k \in \mathcal{U}_{m,p}} \tilde{P}_{B_p, R_{m,p}}^{(k)}(i) - P_{BT} \right] - \gamma_{m,p} \left[\sum_{k \in \mathcal{U}_{m,p}} \tilde{P}_{R_{m,p}}^{(k)}(i) \right] \\ & - \sum_{k \in \mathcal{U}_{m,p}} \theta_{m,p}^{(k)}(i) \left(s_{m,p}^{(k)}(i) \tilde{\sigma}_k^{2*}(i) - I \right) \\ & - \sum_{g \neq p} \sum_{q \in \mathcal{R}_g} \sum_{b \in \mathcal{U}_{q,g}} \theta_{q,g}^{(b)}(i) (s_{q,g}^{(b)}(i) \tilde{P}_{R_{m,p}}^{(b)}(i) l_{R_{m,p}}^{(b)} |H_{R_{m,p}}^{(b)}(i)|^2 - I) \end{aligned} \quad (18)$$

for a given set of Lagrange multipliers. Let $\tilde{P}_{B_p, R_{m,p}}^{(k)*}(i)$, $\tilde{P}_{R_{m,p}}^{(k)*}(i)$, and $s_{m,p}^{(k)*}(i)$ denote the optimal solution of the subproblem. Using standard optimization techniques and the Karush-Kuhn-Tucker (KKT) conditions, the optimal power allocation at BS p for transmission to user k in subcarrier i via relay m can be obtained as

$$\begin{aligned} & \tilde{P}_{B_p, R_{m,p}}^{(k)*}(i) = s_{m,p}^{(k)}(i) P_{B_p, R_{m,p}}^{(k)*}(i) \\ & = s_{m,p}^{(k)}(i) \left(\frac{(\mu_{m,p}^{(k)} + \eta^{(k)})}{\lambda_p 2 \ln 2} - \frac{\sigma_z^2}{|H_{B_p, R_{m,p}}(i)|^2 l_{B_p, R_{m,p}}(i)} \right)^+, \end{aligned} \quad (19)$$

$$\begin{aligned} & \tilde{P}_{R_{m,p}}^{(k)*}(i) = s_{m,p}^{(k)}(i) P_{R_{m,p}}^{(k)*}(i) \\ & = s_{m,p}^{(k)}(i) \left(\frac{(\nu_{m,p}^{(k)} + \delta^{(k)})}{2 \ln 2 (\gamma_{m,p} + \Theta^{(k)}(i))} - \frac{\sigma_z^2 + I}{l_{R_{m,p}}^{(k)} |H_{R_{m,p}}^{(k)}(i)|^2} \right)^+ \end{aligned} \quad (20)$$

where $(x)^+ = \max\{0, x\}$. $\Theta^{(k)}(i) = \sum_{g \neq p} \sum_{q \in \mathcal{R}_g} \sum_{b \in \mathcal{U}_{q,g}} \theta_{q,g}^{(b)}(i) (s_{q,g}^{(b)}(i) l_{R_{m,p}}^{(b)} |H_{R_{m,p}}^{(b)}(i)|^2)$. The power allocation in (19) and (20) can be interpreted as a *multi-level* water-filling scheme as the water levels of different users can be different. Specifically, the water-levels of *delay-sensitive* users, i.e., $\frac{(\mu_{m,p}^{(k)} + \eta^{(k)})}{\lambda_p 2 \ln 2}$ and $\frac{(\nu_{m,p}^{(k)} + \delta^{(k)})}{2 \ln 2 (\gamma_{m,p} + \Theta^{(k)}(i))}$, are generally higher than those of *non-delay sensitive* users, in order to satisfy constraint C3 in (10), since $\eta^{(k)}, \delta^{(k)} > 0$ for *delay sensitive* users. On the other hand, the allocation of subcarrier i at BS p to user k through relay m is specified by

$$s_{m,p}^{(k)*}(i) = \begin{cases} 1 & \text{if } X_{m,p}^{(k)}(i) = \max_{j,d} X_{d,p}^{(j)}(i) \\ 0 & \text{otherwise} \end{cases} \quad (21)$$

$$\begin{aligned} & \text{where } X_{d,p}^{(j)}(i) = \\ & (\mu_{m,p}^{(k)} + \eta^{(k)}) \left[\log_2 \left(1 + \Gamma_{B_p, R_{m,p}}^{(k)*}(i) \right) - \frac{\Gamma_{B_p, R_{m,p}}^{(k)*}(i)}{\sigma_z^2 + \Gamma_{B_p, R_{m,p}}^{(k)*}(i)} \right] \\ & + (\nu_{m,p}^{(k)} + \delta^{(k)}) \left[\log_2 \left(1 + \Gamma_{U_{m,p}}^{(k)*}(i) \right) - \frac{\Gamma_{U_{m,p}}^{(k)*}(i)}{\sigma_z^2 + I + \Gamma_{U_{m,p}}^{(k)*}(i)} \right] \\ & - \theta_{m,p}^{(k)}(i) \tilde{\sigma}_k^{2*}(i) - \beta_p(i), \end{aligned} \quad (22)$$

$$\begin{aligned} & \text{where } \Gamma_{B_p, R_{m,p}}^{(k)*}(i) = \Gamma_{B_p, R_{m,p}}^{(k)}(i) |_{P_{B_p, R_{m,p}}^{(k)}(i) = P_{B_p, R_{m,p}}^{(k)*}(i)}, \\ & \Gamma_{U_{m,p}}^{(k)*}(i) = \frac{P_{R_{m,p}}^{(k)*}(i) l_{R_{m,p}}^{(k)} |H_{R_{m,p}}^{(k)}(i)|^2}{\tilde{\sigma}_k^{2*}(i) + \sigma_z^2}, \text{ and } \tilde{\sigma}_k^{2*}(i) = \\ & \tilde{\sigma}_k^2(i) |_{P_{B_p, R_{m,p}}^{(k)}(i) = P_{B_p, R_{m,p}}^{(k)*}(i)}. \end{aligned}$$

F. Solution of the Master Problem

The gradient method is used to solve the minimization of the master problem in (14). The update algorithm is given by

$$\begin{aligned} \mu_{m,p}^{(k)}[n+1] & = \left[\mu_{m,p}^{(k)}[n] - \xi_1[n] \times \left(\sum_{i=1}^{n_F} s_{m,p}^{(k)}(i) \right. \right. \\ & \quad \left. \left. \times (\tilde{C}_{B_p, R_{m,p}}^{(k)}(i) - \tilde{C}_{U_{m,p}}^{(k)}(i)) \right) \right]_{\mathcal{U}_{m,p}^{(k)}}^+, \forall k, m, p \\ \gamma_{m,p}[n+1] & = \left[\gamma_{m,p}[n] - \xi_2[n] \right. \\ & \quad \left. \times [P_{RT} - \sum_{k \in \mathcal{U}_{m,p}} \sum_{i=1}^{n_F} \tilde{P}_{R_{m,p}}^{(k)}(i)] \right]^+, \forall m, p \\ \lambda_p[n+1] & = \left[\lambda_p[n] - \xi_3[n] \right. \\ & \quad \left. \times [P_{BT} - \sum_{m \in \mathcal{R}_p} \sum_{k \in \mathcal{U}_{m,p}} \sum_{i=1}^{n_F} \tilde{P}_{B_p, R_{m,p}}^{(k)}(i)] \right]^+, \forall p \\ \eta_k[n+1] & = \left[\eta_k[n] - \xi_4[n] \right. \\ & \quad \left. \times \left[\sum_{i=1}^{n_F} s_{m,p}^{(k)}(i) \tilde{C}_{B_p, R_{m,p}}^{(k)}(i) - r^{(k)} \right] \right]^+, \forall k \in \mathcal{D}_{m,p} \end{aligned}$$

TABLE I

Algorithm 1 Centralized resource allocation algorithm

```

1: Initialize  $L_{max}$ ,  $\lambda$ ,  $\gamma$ ,  $\mu$ ,  $\nu$ ,  $\theta$ ,  $\delta$ ,  $\eta$ , and set iteration index  $n = 0$ 
2: Initialize  $\mathcal{P}_n$  and compute  $\mathcal{S}_n$  according to (21) for  $n = 0$ 
3: repeat {Outer Loop}
4:   for  $i = 1$  to  $n_F$  do
5:     repeat {Inner Loop}
6:       for  $p = 1$  to  $P$  do
7:         Update  $P_{B_p, R_{m,p}}^{(k)}$  and  $P_{R_{m,p}}^{(k)}$  according to (19) and
           (20), respectively, for  $k \in \mathcal{U}_{m,p}$ 
8:       end for
9:       Update  $\mathcal{S}_n(i)$  for each BS according to (21) while assuming the
           subcarrier allocation in other cells remains unchanged
10:      until  $\mathcal{P}_n(i)$  and  $\mathcal{S}_n(i)$  converge
11:    end for
12:    Update  $\lambda$ ,  $\gamma$ ,  $\mu$ ,  $\nu$ ,  $\theta$ ,  $\delta$ ,  $\eta$  according to (23) and set  $n = n + 1$ 
13: until convergence or  $n = L_{max}$ 
    
```

$$\begin{aligned}
 \delta_k[n+1] &= \left[\delta_k[n] - \xi_5[n] \right. \\
 &\quad \times \left. \left[\sum_{i=1}^{n_F} s_{m,p}^{(k)}(i) \tilde{C}_{U_{m,p}}^{(k)}(i) - r^{(k)} \right] \right]^+, \forall k \in \mathcal{D}_{m,p} \\
 \theta_{m,p}^{(k)}(i)[n+1] &= \left[\theta_{m,p}^{(k)}(i)[n] - \xi_6[n] \right. \\
 &\quad \times \left. \left(s_{m,p}^{(k)}(i) \tilde{\sigma}_k^2(i) - I \right) \right]^+, \forall m, p, k, i
 \end{aligned} \quad (23)$$

where $\mathbb{U}_{m,p}^{(k)}$ denotes the projection operator on the feasible set $\mathbb{U}_{m,p}^{(k)} = \{\mu_{m,p}^{(k)} | 0 \leq \mu_{m,p}^{(k)} \leq w^{(k)}\}$. $\nu_{m,p}^{(k)}$ can be obtained from $\nu_{m,p}^{(k)} = w^{(k)} - \mu_{m,p}^{(k)}$. n is the iteration index and $\xi_j[n]$, $j \in \{1, \dots, 6\}$, are positive step sizes. Updating $\beta(i)$ is not necessary as it has the same value for each user and relay served by the same BS and it does not affect the subcarrier allocation in (21). It is guaranteed that the algorithm converges to the optimal solution for the transformed convex problem, if constant step sizes are chosen for a continuous dual function [11].

G. Iterative Algorithm for Practical Implementation

Theoretically, equations (19)-(23) provide a complete solution for the considered multi-cell resource allocation problem. However, (20) and (21) require non-causal knowledge of the resource allocation policies in other cells which is a hurdle for practical implementation. In this section, we present an iterative algorithm to bridge the gap between theory and practice. The algorithm is outlined in Table I. In Table I, L_{max} is the maximum number of iterations. \mathcal{P}_n and \mathcal{S}_n are the power allocation and subcarrier allocation policies in the n th iteration, respectively. $\mathcal{P}_n(i)$ and $\mathcal{S}_n(i)$ are the power allocation and subcarrier allocation policy vectors in subcarrier i for all BSs and all relays in the n th iteration, respectively. The overall iterative algorithm is implemented by two nested loops. The inner loop, i.e., line 5 to line 10, is solving the maximization in (16) for a given set of Lagrange multipliers for subcarrier i . In line 7, we first keep the subcarrier allocation in subcarrier i in other cells fixed and then optimize $P_{B_p, R_{m,p}}^{(k)}$ and $P_{R_{m,p}}^{(k)}$ (i). Then, we use the optimized variables in cell p to optimize the power allocation variables in subcarrier i for cell $p + 1$, and so on. The same logic is also applied in line 9. In other words, a multi-variable function is optimized over each variable separately in the inner loop which is known as coordinate ascent method [11]. Convergence to the optimal solution for a given set of Lagrange multipliers is ensured [11] for convex optimization problems. On the other hand, the outer loop, i.e., line 3 to line 13, is trying to solve the minimization of the master problem by updating the dual variables.

IV. RESULTS AND DISCUSSIONS

In this section, we evaluate the system performance with the proposed resource allocation and scheduling algorithm using

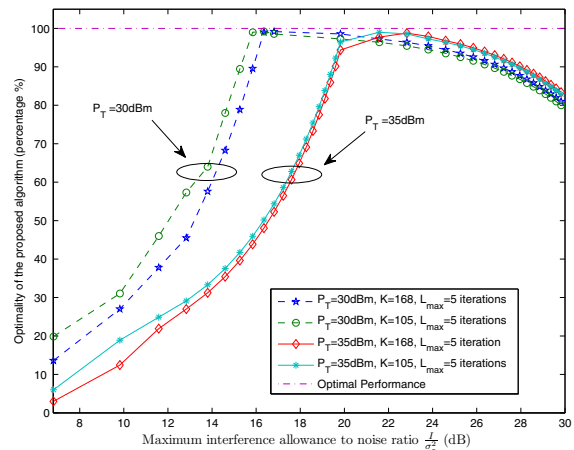


Fig. 2. The performance of the proposed algorithm versus the maximum interference to noise ratio $\frac{I}{\sigma_z^2}$ for different value of P_T and number of users.

simulations. A multi-cell system with 7 cells is considered. Each cell is modeled as two concentric ring-shaped discs. The outer boundary has a radius of 2 km and the inner boundary a radius of 1 km, cf. Figure 1. There are $M = 21$ relay stations in the system and each cell has 3 relays which are equally distributed at the inner cell boundary of each cell for assisting the transmission. There are K/P active cell edge users uniformly distributed in the outer disc of each cell. 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 and the noise variance is $\sigma_z^2 = -125$ dBm. The 3GPP path loss model is used [12]. 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. We assume that the maximum transmit power per cell is P_T and each transmission device, i.e., BS and relay, has a maximum transmit power of $P_{R_T} = P_{B_T} = \frac{P_T}{N+1}$, where N is the number of relays per cell. The average weighted 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. System Throughput versus I

Figure 2 illustrates the performance of the proposed algorithm versus the value of I for different P_T and different numbers of users K . The y-axis is normalized by the performance of the optimal solution to the original non-convex problem in (10)². The x-axis represents the maximum interference allowance to noise ratio, i.e., $\frac{I}{\sigma_z^2}$. We assume that there are always 7 delay sensitive users with data rate requirement $r^{(k)} = 0.1$ bit/s/Hz in the system³, while the remaining users are non-delay sensitive. The number of iterations for the proposed algorithm is 5. It can be seen that for a wide range of $\frac{I}{\sigma_z^2}$, we can achieve more than 90% of the optimal performance but benefit from the convexity of the transformed problem. Furthermore, the choice of I is not sensitive to the number of users for a given value of P_T . On the other hand, as expected, the optimal value of I strongly depends on P_T since the amount of multi-cell interference increases with P_T , a higher value of I is needed to reflect the actual interference level.

²The optimal solution of (10) can be obtained by following a similar approach as in [10]. Note, however, that the transformed problem is guaranteed to be solved in polynomial time with linear complexity in each iteration while the original problem in (10) has an exponential complexity in the number of M , P , and n_F .

³The target cell-edge performance for 4G communication systems is 0.09 bit/s/Hz [13].

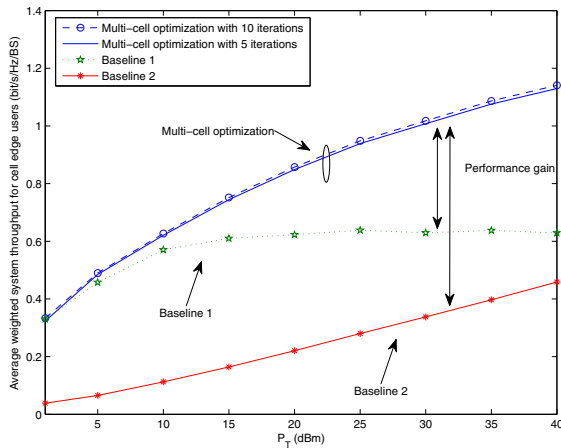


Fig. 3. Average weighted system throughput versus total transmit power per cell for different resource allocation and scheduling algorithms with $K = 105$.

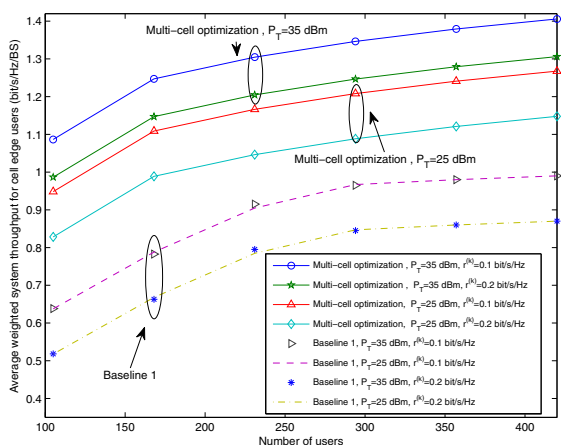


Fig. 4. Average weighted system throughput versus number of users with different data rate requirements and transmit powers for different resource allocation algorithms.

B. System Throughput versus Transmit Power

Figure 3 illustrates the average weighted system throughput versus the total transmit power in each cell, P_T , for a total of 105 users of which 7 are *delay sensitive* users with data rate requirement $r^{(k)} = 0.1$ bit/s/Hz and the other users are *non-delay sensitive*. The value of I in the proposed algorithm is chosen to guarantee that always more than 95% of the optimal performance can be achieved. The performance of the proposed resource allocation algorithm with 5 and 10 iterations is compared with two baselines schemes. In baseline scheme 1, each BS performs its own resource allocation for the DF relaying system and ignores the multi-cell interference. For baseline scheme 2, interference is completely avoided by setting the frequency reuse factor to $\frac{1}{P}$. Interestingly, in the low transmit power regime, i.e., $P_T < 10$ dBm, baseline scheme 1 achieves almost the same performance as the proposed resource allocation algorithm. This is because for low transmit powers, noise is the dominant factor affecting the system performance and the multi-cell interference is negligible. Yet, the operating point of the system is shifting from noise limited to interference limited as the total transmit power in the system increases and the proposed algorithm achieves a substantial performance gain compared with baseline scheme 1, since the latter one does not scale with transmit power due to the strong interference. On the other hand, the proposed algorithm achieves a higher spectral efficiency than baseline scheme 2 since the latter one does not fully utilize the spectrum to avoid multi-cell interference.

C. System Throughput versus Number of Users

Figure 4 depicts the average weighted system throughput versus the number of users with different transmit power and

user data rate requirements. We compare the proposed algorithm with baseline scheme 1. The number of iterations for the proposed algorithm is 5. It can be observed that the average system throughput increases with the number of users for the proposed resource allocation and scheduling algorithm. This is because as the number of active users increase, the scheduler has a higher chance to select users who have both a strong channel with respect to their home cell and weak channels with respect to other cells. This effect can be interpreted as the multi-user diversity (MUD) in multi-cell systems. However, when the data rate requirements become more stringent, the scheduler loses degrees of freedom for resource allocation and scheduling since it needs to serve the *delay sensitive* users despite their possibly poor channel qualities, which diminishes the MUD gain. On the other hand, the performances of baseline scheme 1 for $P_T = 25$ dBm and $P_T = 35$ dBm are virtually identical for increasing number of users, since the performance of baseline scheme 1 is limited by the strong interference in this case.

V. CONCLUSIONS

In this paper, we formulate the resource allocation and scheduling design for multi-cell OFDMA DF relay assisted networks as a non-convex and combinatorial optimization problem, in which interference from other cells and data rate requirements from heterogeneous users are taken into consideration. By imposing an additional interference allowance constraint and relaxing the subcarrier allocation constraints, the considered problem can be transformed into a convex problem. An iterative resource allocation algorithm with closed-form power and subcarrier allocation policies is derived by dual decomposition. Simulation results show that the performance of the proposed algorithm approaches the optimal performance in a small number of iterations.

REFERENCES

- [1] Y. Cui, V. K. N. Lau, and R. Wang, "Distributive Subband Allocation, Power and Rate Control for Relay-Assisted OFDMA Cellular System with Imperfect System State Knowledge," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 5096 – 5102, Oct 2009.
- [2] H.-X. Li, H. Yu, H.-W. Luo, J. Guo, and C. Li, "Dynamic Subchannel and Power Allocation in OFDMA-Based DF Cooperative Relay Networks," in *Proc. IEEE Global Telecommun. Conf.*, Dec 2008, pp. 1 – 5.
- [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] L. Weng and R. D. Murch, "Cooperation Strategies and Resource Allocations in Multiuser OFDMA Systems," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 2331 – 2342, Jun 2009.
- [5] "Downlink Inter-Cell Interference Co-ordination/Avoidance - Evaluation of Frequency Reuse," 3GPP TSG-RAN WG1 Contribution R1-061374, Tech. Rep., 2006.
- [6] A. Gjendemsj, D. Gesbert, G. E. Oien, and S. G. Kiani, "Binary Power Control for Sum Rate Maximization over Multiple Interfering Links," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 3164 – 3173, Aug 2008.
- [7] Z. Liang, Y. H. Chew, and C. C. Ko, "A Linear Programming Solution to Subcarrier, Bit and Power Allocation for Multicell OFDMA Systems," in *Proc. IEEE Wireless Commun. and Networking Conf.*, Apr 2008, pp. 1273 – 1278.
- [8] L. Venturino, N. Prasad, and X. Wang, "Coordinated Scheduling and Power Allocation in Downlink Multicell OFDMA Networks," *IEEE Trans. Veh. Technol.*, vol. 58, pp. 2835 – 2848, Jul 2009.
- [9] O. Simeone, O. Somekh, Y. Bar-Ness, and U. Spagnolini, "Uplink Throughput of TDMA Cellular Systems with Multicell Processing and Amplify-and-Forward Cooperation Between Mobiles," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2942 – 2951, Aug 2007.
- [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] "Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations," 3GPP TR 25.996 V7.0.0 (2007-06), Tech. Rep.
- [13] A. Technologies and M. Rumney, *LTE and the Evolution to 4G Wireless: Design and Measurement Challenges*, 1st ed. Agilent Technologies, 2009.