Bandwidth Sharing for Relaying in Cellular Systems

Narayanan Krishnan, *Student Member, IEEE,* Roy D. Yates, *Fellow, IEEE,*
Narayan B. Mandayam, *Fellow, IEEE,* and Jignesh S. Panchal

***Abstract*—We investigate bandwidth allocation in next generation cellular systems employing relays similar to LTE advanced systems with type-I relays. We jointly optimize the bandwidth and power usage under constraints on required rate, bandwidth and transmit power. We study scenarios wherein, the relay acts as a forwarder for multiple User Equipments (UEs/users) in both uplink and/or downlink. This includes scenarios when the relay has its own data to send along with forwarding the data of other users. We examine the weighted power minimization problem for relaying with multiple users. We also show specific results with N user scenario and also single user case in order to understand how the bandwidth and power are allocated. Numerical evaluations with N users on a three sector LTE-A cell employing Fractional Frequency Reuse (FFR) indicate that power savings of at least 3 dB can be achieved by optimizing over both bandwidth and power.**

***Index Terms*—Bandwidth allocation, time sharing, OFDMA, multihop cellular system, resource management, convex optimization, power minimization, coverage extension.**

I. INTRODUCTION

NEXT generation cellular standards such as LTE-A and IEEE802.16j are poised to meet the requirements of International Mobile Telecommunications-Advanced (IMT-A) standards [1]. These systems are promising to provide peak data rates of up to 100 Mbps and 1 Gbps in high mobility environments and pedestrian environments, respectively and hence provide high speed mobile broadband access [2], [3]. One of the key technology components of next generation cellular systems is relaying, which has been shown to provide better throughput and increased coverage [4], [5]. Typically, relays help in forwarding data and based on their roles, they have been categorized into two types. The first type is used exclusively to extend the coverage to remote User Equipment (UE), beyond the service range of the base station (aka eNodeB in LTE-A jargon). These are called type-I relays by LTE-A specifications and non-transparent relays in IEEE802.16j specifications. One of the applications of “coverage extending” low powered relays is to provide coverage to indoor or office environments where the signal strength is weak. Apart from providing extended coverage, they also help in deployment of cells in areas where the cost of wired backhaul is prohibitive. On the other hand, the second category of relays, are used to help the UE within the service range of base station to improve its service quality and link capacity. These are called type-II relays by LTE-A or transparent relays in IEEE802.16j.

When relaying terminals aid the UE, they incur costs in the form of power expenditure and usage of relay bandwidth. To compensate the relay for these costs, some well known approaches include reputation based mechanisms, credit based incentives and mechanisms based on forwarding games [8]. A novel approach of bandwidth exchange was introduced in [9] in order to compensate the relay for its incurred costs. Here, the relay node is offered a portion of bandwidth of the destination node as a compensation for forwarding data. The relay can use the compensated bandwidth for purposes that it deems fit, while ensuring forwarding the data of the destination. It was shown that bandwidth exchange provides significant rate gains and improved coverage areas. In [12], a weighted power minimization problem was formulated for joint power, subcarrier allocation and subframe scheduling for downlink in-band [13] relaying systems with a focus on developing efficient algorithms. Our paper in contrast provides a theoretical insights into the nature of optimal solution in an out-band relaying system using which we intend to develop efficient algorithms.

In this work, we consider bandwidth sharing in the context of relaying in cellular systems and study optimization problems involving both bandwidth and transmit power under rate constraints. A key aspect of our work is that along with power, we consider the bandwidths allocated to links as optimization variables in the relaying system. We formulate a weighted power minimization problem under rate, bandwidth and power constraints assuming out-band relaying system [13] and develop theoretical insights into the nature of optimal solutions. We formulate the weighted power minimization problem for a system involving a base station and a relay with multiple UEs to be served. We also consider an in-band relaying system [13] in which the eNodeB-relay and relay-UE link use the same set of carrier frequencies but transmits at different time slots. We show that the problem of average power minimization in a time shared relaying system as in [4] can be reformulated to an equivalent bandwidth sharing problem and depending on the nature of power constraints the time sharing system performs worse or same as the bandwidth sharing system. We find that optimizing across power and bandwidth provides scope for better utilization of the available resources, such as minimizing the total power consumption by half while improving the coverage area of the relays. While the theoretical results of the relaying system apply to a wide range of systems which employ multicarrier schemes, the terminology and numerical results used to describe this work loosely conform to LTE-A standards.

Manuscript received November 27, 2010; revised July 9, 2011; accepted August 28, 2011. The associate editor coordinating the review of this paper and approving it for publication was K. B. Lee.

The authors are with the Dept. of ECE, Rutgers, The State University of New Jersey (e-mail: narayank@winlab.rutgers.edu).

Digital Object Identifier 10.1109/TWC.2011.110711.102126

II. SYSTEM MODEL

We start with system with an eNodeB, a relay and N UEs or users who are to be served. The relay can be a dedicated tower or another user that can potentially act as an intermediate node to help users forward its data. We assume that the relays can forward data simultaneously in uplink and downlink in different carrier frequencies. Our assumption is valid as the LTE-A and IEEE 802.16j standards [14] indeed has dual radio out-of-band relay system in which the relay nodes have two RF chains in order to transmit and receive simultaneously in different carriers. The system is allocated a fixed bandwidth B. The users $i \in \{1, 2, \ldots, N\}$ are subject to transmit at power below $P_{i,\max}$. Also, every UE needs to be served at a rate of $R^d_{i,\min}$ in the downlink and $R^u_{i,\min}$ in the uplink. The link gains are of the form $\rho_{ij} = \kappa d_{ij}^{-\beta}$, where κ is the proportionality constant, d_{ij} is the distance between the i^{th} transmitter and j^{th} receiver and the pathloss exponent is $\beta = 3$. The noise power spectral density is also absorbed into the constant κ that is used to calculate the link gain. In general, any variable with subscript $(\cdot)_{ij}$ corresponds to the variable across the directed link (i, j) from transmitter i to receiver j. We consider that the link gains $\rho_{ij} = \rho_{ji}$ for any links (i, j) and (j, i). Specifically, we denote the users with the subscript i, relay equipment by r and eNodeB by 0, as shown in Fig. 1. We look into cases when the relay is employed to help the users achieve its minimum data rate. We consider the transmit power at each node and the bandwidth allocated to each link as a variable to be optimized. Our objective is to minimize the weighted sum of power in the system. Optimizing over bandwidth is relevant as we are employing a multicarrier system. This is because in a multicarrier system like OFDMA the bandwidth allocated to each link is specified by a number of subcarriers. We assume that the subcarrier spacing is small enough so that the bandwidth variables can be approximated as to be continuous. We also assume in our work that the eNodeB does a centralized resource allocation of bandwidth and power.

The relay r can be employed in various possible modes such as, providing only downlink access to users, or providing only uplink access to users, or providing both uplink and downlink access to users and also other intermediate scenarios such as providing uplink support to some users while only downlink to other users. Along with the cases above the relay r may have uplink information to send to the eNodeB, or it may have downlink information to receive for its own. For example, the downlink and uplink information for relay r could be the control information needed for dedicated relaying. In this particular work, we formulate weighted sum power minimization problem when the relay is serving multiple users as shown in Fig. 2 and encompasses all possible cases discussed above. We assume that the eNodeB has the knowledge of channel gain between itself to the relays and between relays to UEs and solves the optimization problem in a centralized manner. Some special cases of the power minimization problem which which are of interest include:

1. The out-band relay r helps the users in forwarding both their downlink and uplink data.
2. The out-band relay r helps the users in forwarding both their downlink and uplink data along with transmitting its own data to the eNodeB.
3. The in-band relay r helps the users in forwarding both their uplink and downlink data.
4. The out-band relay r helps in forwarding both uplink and downlink data of a single user.

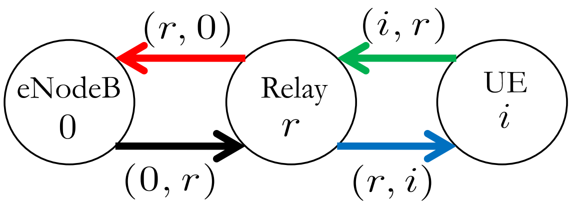


Fig. 1. Relay acts as forwarder in both directions: This corresponds to *Problem (4)* where a single relay serves a user in both uplink and downlink directions.

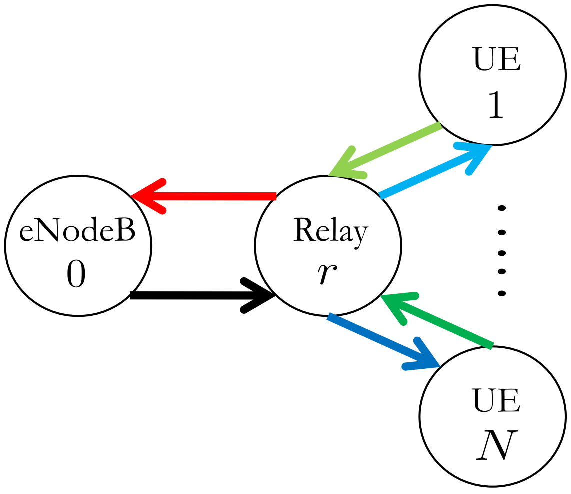


Fig. 2. Relaying for multiple users: This corresponds to *Problem (1)* where a single relay serves multiple users in both uplink and directions.

In the following section, these special cases are identified in *Problems (1), (2), (3)* and *(4)*. We consider the case of single user as separate problem as it provides some intuition on how the bandwidth is allocated in the uplink and downlink.

A more generic system model arises when the eNodeB is serving directly some UEs and is also be connected to UEs through more than a single relay. The system model in that case might be represented as a tree structure with eNodeB as the root, the UEs as leaf nodes and the relays as intermediate nodes. Although the mathematical formulation and evaluation of the results is cumbersome, this is a straightforward extension of a single hop scenario in the sense that the insights revealed by the weighted power minimization for a multiple relay scenario is qualitatively similar to the case of single relay system.

III. PROBLEM FORMULATION

A. Relaying in Uplink and Downlink to Multiple User Equipments

We consider the case when a relay helps in forwarding data to N users in both uplink and downlink directions. In terms of the power P_{ir}, bandwidth W_{ir}, the signal to noise ratio given by $\mathsf{SNR}_{ir} \triangleq \rho_{ir} P_{ir}/W_{ir}$ and the rate function defined as

$$\mathcal{R}(\mathsf{SNR}_{ir}, W_{ir}) \triangleq W_{ir} \log(1 + \mathsf{SNR}_{ir})$$

across the corresponding links (i, r), we formulate the problem of weighted power minimization as

minimize

$$\alpha_0 P_{0r} + \alpha_r \left(\sum_{i=1}^{N} P_{ri} + P_{r0} \right) + \sum_{i=1}^{N} \alpha_i P_{ir} \tag{1a}$$

subject to

$$\mathcal{R}(\mathsf{SNR}_{ir}, W_{ir}) \geq R^u_{i,\min} \quad i = 1, \ldots, N, \tag{1b}$$

$$\mathcal{R}(\mathsf{SNR}_{r0}, W_{r0}) \geq \sum_{i=1}^{N} \mathcal{R}(\mathsf{SNR}_{ir}, W_{ir}), \tag{1c}$$

$$\mathcal{R}(\mathsf{SNR}_{ri}, W_{ri}) \geq R^d_{i,\min} \quad i = 1, \ldots, N, \tag{1d}$$

$$\mathcal{R}(\mathsf{SNR}_{0r}, W_{0r}) \geq \sum_{i=1}^{N} \mathcal{R}(\mathsf{SNR}_{ri}, W_{ri}), \tag{1e}$$

$$\sum_{i=1}^{N} (W_{ir} + W_{ri}) + W_{r0} + W_{0r} \leq B, \tag{1f}$$

$$P_{r0} + \sum_{i=1}^{N} P_{ri} \leq P_{r,\max}, \tag{1g}$$

$$P_{0r} \leq P_{0,\max}, \tag{1h}$$

$$P_{ir} \leq P_{\mathrm{i},\max} \quad i = 1, \ldots, N, \tag{1i}$$

variables

$$P_{r0}, P_{0r}, P_{ir}, P_{ri}, W_{r0}, W_{0r}, W_{ir}, W_{ri} \quad i = 1, \ldots, N$$

In *Problem (1)*, B is the total available bandwidth and $\alpha_0, \alpha_i, \alpha_r$ are the weights associated with the power spent by the nodes with $\alpha_0 + \sum_{i=1}^{N} \alpha_i + \alpha_r = 1$. For example, when $\alpha_r = 1$ the problem is just relay power minimization and when $\alpha_0 = \alpha_i = \alpha_r = 1/(N+2)$ implies that the power of all the nodes are equally important and we are interested in system power minimization. Other values gives the engineer flexibility to design the system as required. Here, the constraints (1b), (1c) and (1d), (1e) represent the minimum rate requirement in the uplink and downlink simultaneously. By the perspective of a function argument [15, chapter 3], the rate constraints (1b), (1d) are convex in (P_{ir}, W_{ir}). The limited power constraints are represented by equations (1g), (1h) and (1i). Here, constraint (1g) implies that the relay has to use its limited power to provide the required rate in both the uplink and downlink directions for all users. $P_{r,\max}$ is the maximum power constraint on the relay r.

In our formulation we also consider bandwidth to be a variable that is to be optimized across the links. This is represented by the constraint (1f). Previous attempts in joint bandwidth and power optimization have been considered in the context of maximizing the achievable rates in relaying systems in [7]. In this problem we have a relay r which is helping forward data to N users in both uplink and downlink directions. The relaying problem is not convex because of the constraints (1c) and (1e). Although the problem is not convex we will later show that the optimal solution is same as that of an equivalent convex problem which enables us to exploit the advantages offered by convexity.

An important variant of the *Problem (1)* arises when the relay along with forwarding the data of the users has its own data to be sent to the eNodeB at some required rate. For example, this scenario arises when a relay has to allocate some resources (bandwidth and power) to send its control data at a rate of at least $R^u_{r,\min}$. The power minimization problem can then be formulated by replacing the constraint (1c) in *Problem (1)* by

$$\mathcal{R}(\mathsf{SNR}_{r0}, W_{r0}) \geq \sum_{i=1}^{N} \mathcal{R}(\mathsf{SNR}_{ir}, W_{ir}) + R^u_{r,\min} \tag{2}$$

We denote this problem as *Problem (2)*.

B. Relaying in Uplink and Downlink to Multiple User Equipments - Time Shared System

In this section we consider the case of an in-band relaying system with an eNodeB, a single relay and N users. We consider a frame structure similar to that employed in LTE-A and IEEE802.16m schemes where the time is also shared between eNodeB-relay and relay-UE links. Fig. 3 shows the typical frame structure of such a system with total frame time T. We denote T_{BS} as the time allocated to eNodeB to relay downlink frame. Similarly $T^{\mathrm{d}}_{\mathrm{RS}}$ stands for downlink relay to UE frame size, T_{UE} stands for UE to relay uplink frame size, $T^{\mathrm{u}}_{\mathrm{RS}}$ stands for relay to eNodeB uplink frame size. We also have $T_{\mathrm{BS}} + T^{\mathrm{d}}_{\mathrm{RS}} + T_{\mathrm{UE}} + T^{\mathrm{u}}_{\mathrm{RS}} \leq T$ and that all time slots be positive. During the time slot $T^{\mathrm{d}}_{\mathrm{RS}}$ relay transmits to all the UEs simultaneously in orthogonal bandwidth allocations and similarly during the time slot $T^{\mathrm{u}}_{\mathrm{RS}}$ all UEs transmit simultaneously to the relay in uplink in orthogonal bandwidth allocations. All the available bandwidth can be used in the each of the time slots. In terms of the instantaneous power S_{ir} and bandwidth B_{ir} across the link (i, r), we formulate the problem of weighted power minimization as follows,

minimize

$$\alpha_0 \frac{T_{\mathrm{BS}}}{T} S_{0r} + \alpha_r \left(\frac{T^{\mathrm{d}}_{\mathrm{RS}}}{T} \sum_{i=1}^{N} S_{ri} + \frac{T^{\mathrm{u}}_{\mathrm{RS}}}{T} S_{r0} \right) + \frac{T_{\mathrm{UE}}}{T} \sum_{i=1}^{N} \alpha_i S_{ir} \tag{3a}$$

subject to

$$T^{\mathrm{d}}_{\mathrm{RS}} \mathcal{R}(\mathsf{SNR}_{ri}, B_{ri}) \geq T R^d_{i,\min} \forall i = 1 \ldots N \tag{3b}$$

$$T_{\mathrm{BS}} \mathcal{R}(\mathsf{SNR}_{0r}, B_{0r}) \geq T^{\mathrm{d}}_{\mathrm{RS}} \left(\sum_{i=1}^{N} \mathcal{R}(\mathsf{SNR}_{ri}, B_{ri}) \right) \tag{3c}$$

$$T_{\mathrm{UE}} \mathcal{R}(\mathsf{SNR}_{ir}, B_{ir}) \geq T R^u_{i,\min} \forall i = 1 \ldots N \tag{3d}$$

$$T^{\mathrm{u}}_{\mathrm{RS}} \mathcal{R}(\mathsf{SNR}_{r0}, B_{r0}) \geq T_{\mathrm{UE}} \left(\sum_{i=1}^{N} \mathcal{R}(\mathsf{SNR}_{ir}, B_{ir}) \right) \tag{3e}$$

$$B_{0r} \leq B \tag{3f}$$

$$\sum_{i=1}^{N} B_{ri} \leq B \tag{3g}$$

$$\sum_{i=1}^{N} B_{ir} \leq B \tag{3h}$$

$$B_{r0} \le B \tag{3i}$$
$$T_{\mathrm{BS}} + T_{\mathrm{RS}}^{\mathrm{d}} + T_{\mathrm{UE}} + T_{\mathrm{RS}}^{\mathrm{u}} \le T \tag{3j}$$

variables

$$T_{\mathrm{BS}}, T_{\mathrm{RS}}^{\mathrm{d}}, T_{\mathrm{UE}}, T_{\mathrm{RS}}^{\mathrm{u}}, S_{0r}, S_{ri}, S_{r0}, S_{ir}, B_{0r}, B_{ri}, B_{r0}, B_{ir} \quad \forall i = 1 \dots N$$

In *Problem (3)*, the rate constraints are given by (3b), (3c), (3d), (3e) and the bandwidth constraints by (3g), (3f), (3h), (3i). The frame time constraint is given by (3j). The objective term corresponds to minimizing the weighted average power in the system in a single frame. The power constraint in the system can be either a peak power constraint or average power constraint. The peak power constraints take the form,

$$S_{0r} \le P_{0,\max}, \tag{3k-P}$$
$$S_{ir} \le P_{i,\max}, \tag{3l-P}$$
$$\sum_{i=1}^{N} S_{ri} \le P_{r,\max}, \tag{3m-P}$$
$$S_{r0} \le P_{r,\max}, \tag{3n-P}$$

while the average power constraints are,

$$\frac{T_{\mathrm{BS}}}{T} S_{0r} \le P_{0,\max} \tag{3k-A}$$
$$\frac{T_{\mathrm{UE}}}{T} S_{ir} \le P_{i,\max} \tag{3l-A}$$
$$\frac{T_{\mathrm{RS}}^{\mathrm{d}}}{T} \sum_{i=1}^{N} S_{ri} + \frac{T_{\mathrm{RS}}^{\mathrm{u}}}{T} S_{r0} \le P_{r,\max} \tag{3m-A}$$

While peak power constraints are important from the point of view of avoiding RF power amplifier non-linearities, the average power constraints becomes relevant when there are energy consumption limitations in individual nodes on a per frame basis. The power constraints for *Problem (1)* can be interpreted both in terms of peak and average power limits.

C. Relaying for a Single User on Uplink and Downlink

For simplicity and understanding of the bandwidth sharing process we also consider the case of the weighted power minimization problem in relaying when there is only one user i.e., $N = 1$. The relay r helps in forwarding data for the UE i in both uplink and downlink direction as shown in the Fig. 1. We formulate the problem of minimizing the weighted sum of powers spent by the system in ensuring the minimum required data rates in both uplink and downlink for the UE i as

minimize $\quad \alpha_0 P_{0r} + \alpha_i P_{ir} + \alpha_r (P_{ri} + P_{r0}) \qquad$ (4a)

subject to

$$\mathcal{R}(\mathsf{SNR}_{ir}, W_{ir}) \ge R_{i,\min}^{u}, \tag{4b}$$
$$\mathcal{R}(\mathsf{SNR}_{r0}, W_{r0}) \ge R_{i,\min}^{u}, \tag{4c}$$
$$\mathcal{R}(\mathsf{SNR}_{ri}, W_{ri}) \ge R_{i,\min}^{d}, \tag{4d}$$
$$\mathcal{R}(\mathsf{SNR}_{0r}, W_{0r}) \ge R_{i,\min}^{d}, \tag{4e}$$
$$W_{ir} + W_{0r} + W_{ri} + W_{r0} \le B, \tag{4f}$$
$$P_{ri} + P_{r0} \le P_{r,\max}, \tag{4g}$$
$$P_{0r} \le P_{0,\max}, \tag{4h}$$
$$P_{ir} \le P_{i,\max}, \tag{4i}$$

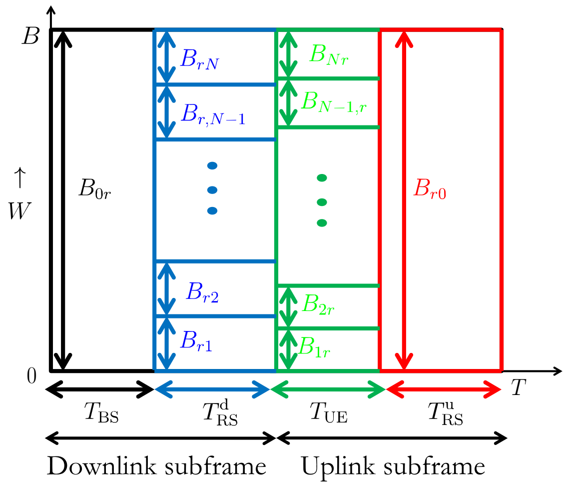


Fig. 3. The diagram shows the resource block division for a time shared system represented by *Problem (3)*. While the total frame time T is divided into uplink and downlink subframe time zones each of them is further divided for enabling single hop relaying. T_{BS} is the time allocated to eNodeB for the downlink frame, $T_{\mathrm{RS}}^{\mathrm{d}}$ stands for downlink relay to UE frame size, T_{UE} stands for UE to relay uplink frame size, $T_{\mathrm{RS}}^{\mathrm{u}}$ stands for relay to eNodeB uplink frame size. In the time slot when the relay transmit to all UEs and also when all UEs transmit to relays the bandwidth is shared orthogonally as shown in the figure.

variables $\quad P_{0r}, P_{ri}, P_{ir}, P_{r0}, W_{0r}, W_{ri}, W_{ir}, W_{0r}.$

So far we have defined four related optimization problems in the relaying system which we denote as *Problems (1), (2), (3)* and *(4)* throughout this paper. Table I provides brief descriptions of these problems for quick reference. This table also includes additional optimization problems that we will define later in the paper.

IV. RESULTS AND DISCUSSION

Lemma 1. *When the constraint set is feasible, an optimal solution to* Problem (1) *has the following characteristics: The rate constraints (1b), (1c), (1d), (1e) and the bandwidth constraint (1f) are tight.*

Proof: The proof is given in the Appendix. ■

As a consequence of Lemma 1 we can formulate an equivalent convex formulation for the *Problem (1)*. This is explained in the following theorem.

Theorem 2. *The optimal solution to* Problem (1) *is the same as that of the optimal solution of* Problem (1R) *in which the constraints (1c) and (1e) are replaced by,*

$$\mathcal{R}(\mathsf{SNR}_{r0}, W_{r0}) \ge \sum_{i=1}^{N} R_{i,\min}^{u}, \tag{1R-c}$$
$$\mathcal{R}(\mathsf{SNR}_{0r}, W_{0r}) \ge \sum_{i=1}^{N} R_{i,\min}^{d}, \tag{1R-e}$$

respectively.

Proof: Let $p^\star$ denote the optimal objective to *Problem (1)* and $q^\star$ the optimal objective to the reformulated *Problem (1R)* with the constraint (1c) replaced by (1R-c) and (1e) replaced by (1R-e). It follows that the feasible set of *Problem (1)* is a

TABLE I

FOR EACH NUMBERED OPTIMIZATION PROBLEMS THIS TABLE GIVES A BRIEF DESCRIPTION OF THE PROBLEM FOR CONVENIENCE. HERE BIDIRECTIONAL MEANS THAT THE RELAY IS SUPPORTING DATA IN BOTH DOWNLINK AND UPLINK FOR THE UEs.

Problem	**Description**	**# of relays**	**# of UEs**	**Max. Power Constraint**
(1)	Bidirectional Relaying	1	N	Yes
(2)	Bidirectional Relaying - with relay uplink data	1	N	Yes
(3)	Bidirectional Relaying - Time Sharing	1	N	Yes
(4)	Bidirectional Relaying	1	1	Yes
(1R)	Bidirectional Relaying - convex formulation	1	N	Yes
(5)	Bidirectional Relaying - dual formulation	1	N	No
(6)	Link specific optimization problem	1	N	No
(14)	Multiple Relays & multiple UEs	M	N	No

subset of the feasible set of *Problem (1R)*. Therefore, $p^\star \geq q^\star$. However, from Lemma 1 since the rate constraints and the bandwidth constraints are tight at optimum, the optimal solution to *Problem (1R)* is a feasible solution to *Problem (1)*. This implies that there is feasible solution for *Problem (1)* with objective value $q^\star$. As $p^\star$ is the optimal solution to *Problem (1)*, by definition of an optimal solution $q^\star \geq p^\star$. Therefore, $q^\star = p^\star$ and optimal points for both problems are identical. ■

Theorem 2 shows that the relaying problem for multiple UEs, which is non-convex by nature can be solved optimally by an equivalent convex formulation. The convex formulation allows the optimal solution to be obtained numerically. Although convex, we have only limited analytical insights into the nature of the optimal solution because of the power constraints (1g), (1h), (1i) on the individual nodes. However, under most problem instances specified by relay and UE positions in the cellular system, the nodes will be transmitting strictly below maximum power. In the next section we look into the nature of the optimal solution assuming that the power constraints are slack at optimum. The solution obtained to *Problem (1)* when the power constraints are slack are the same as when we assume that each of the nodes have unlimited maximum power.

A. Power Minimization - Unlimited Maximum Power

In this section we consider a situation when the maximum power constraints in *Problem (1)* are eliminated. This is the same as assuming that each of the nodes have unlimited maximum power at their disposal. Henceforth we denote *Problem (1R)* as *Problem (1)* as both have the same optimal solution. We look into this problem to get some insight into the optimal solution in situations when none of the power constraints are tight at optimum. Writing the partial Lagrangian with respect to the bandwidth constraint (1f) in *Problem (1R)* results in **minimize**

$$\alpha_0 P_{0r} + \alpha_r \left(\sum_{i=1}^{N} P_{ri} + P_{r0} \right) + \sum_{i=1}^{N} \alpha_i P_{ir} + \lambda \left(\sum_{i=1}^{N} (W_{ir} + W_{ri}) + W_{r0} + W_{0r} - B \right) \tag{5a}$$

subject to

$$\mathcal{R}(\mathsf{SNR}_{ir}, W_{ir}) \geq R^u_{i,\min} \quad \forall i = 1, \ldots, N, \tag{5b}$$

$$\mathcal{R}(\mathsf{SNR}_{0r}, W_{0r}) \geq \sum_{i=1}^{N} R^d_{i,\min}, \tag{5c}$$

$$\mathcal{R}(\mathsf{SNR}_{ri}, W_{ri}) \geq R^d_{i,\min} \quad \forall i = 1, \ldots, N, \tag{5d}$$

$$\mathcal{R}(\mathsf{SNR}_{r0}, W_{r0}) \geq \sum_{i=1}^{N} R^u_{i,\min}, \tag{5e}$$

where, we have assumed that there are no maximum power constraints. The partial Lagrangian helps us to decouple the problem on a link by link basis where previously it was coupled across the links because of the bandwidth constraint (1f). For a given $\lambda \geq 0$ we can split the problem into $2N+2$ different *Link Specific Optimization* problems of the form

$$\begin{array}{rll} \textbf{minimize} & \alpha_l P_l + \lambda W_l, & (6a) \\ \textbf{subject to} & \mathcal{R}(\mathsf{SNR}_l, W_l) \geq R_l, & (6b) \\ \textbf{variables} & P_l, W_l, & (6c) \end{array}$$

corresponding to each directed link $l \in \{(i,r),(r,i)\}$ in the system. The parameters α_l, ρ_l, R_l are link specific and α_l corresponds to the weights in the weighted power minimization and is associated with the transmitter of link l, ρ_l corresponds to the link gain, and R_l is the minimum rate that needs to be supported on link l. For example, in the link optimization problem corresponding to link $(0,r)$, i.e. downlink from the base station node 0 and relay node r, the parameters correspond to $\alpha_l = \alpha_0$, $\rho_l = \rho_{0r}$ and $R_l = \sum_{i=1}^{N} R^d_{i,\min}$.

Since strong duality holds for *Problem (1)* we have at $\lambda = \lambda^\star$, that each *Link Optimization Problem (6)* gives the primal optimal points. Lemma 1 implies that the rate constraints are tight. Thus *Problem (6)* is the same as

$$\min_{W_l} \alpha_l \frac{W_l}{\rho_l} \left(\exp\left(\frac{R_l}{W_l}\right) - 1 \right) + \lambda^\star W_l$$

Therefore, at optimality for every link l in the system

$$\exp\left(\frac{R_l}{W_l}\right)\left(\frac{R_l}{W_l} - 1\right) + 1 - \rho_l \lambda^\star / \alpha_l \;=\; 0. \tag{7}$$

From equation (7) we can gather that the spectral efficiency R_l/W_l depends only the term $\rho_l \lambda^\star/\alpha_l$. This suggests that all the three factors - ρ_l, the link gain, λ, the dual bandwidth cost, and α_l, the node weight decides the spectral efficiency of the link l. This specific property of the *Problem (6)* enable us to prove the following claim specified as Lemma 3. Also, equation (7) along with bandwidth constraint provide

$N+3$ non-linear equations in $N+3$ variables to solve the optimization problem.

Lemma 3. ***Symmetry of Spectral Efficiency:*** *When the weights* $\alpha_0 = \alpha_r = \alpha_i = \alpha$ *are the same, the optimal solution to* Problem (1) *satisfies*

$$\frac{\sum_{i=1}^{N} R^d_{i,\min}}{W^\star_{0r}} = \frac{\sum_{i=1}^{N} R^u_{i,\min}}{W^\star_{r0}}, \tag{8}$$

$$\frac{R^d_{i,\min}}{W^\star_{ri}} = \frac{R^u_{i,\min}}{W^\star_{ir}} \quad \forall i. \tag{9}$$

Proof: The optimality conditions (7) imply

$$\exp\left(\frac{\sum_{i=1}^{N} R^d_{i,\min}}{W^\star_{0r}}\right)\left(\frac{\sum_{i=1}^{N} R^d_{i,\min}}{W^\star_{0r}} - 1\right) = \exp\left(\frac{\sum_{i=1}^{N} R^u_{i,\min}}{W^\star_{r0}}\right)\left(\frac{\sum_{i=1}^{N} R^u_{i,\min}}{W^\star_{r0}} - 1\right) \tag{10}$$

because, the links $l \in \{(0,r),(r,0)\}$ have equal weights $\alpha_l = \alpha$ and equal link gain $\rho_l = \rho_{0r} = \rho_{r0}$. Since $e^x(x-1)$ is monotonic and positive derivative for all $x \geq 0$, (10) implies,

$$\frac{\sum_{i=1}^{N} R^d_{i,\min}}{W^\star_{0r}} = \frac{\sum_{i=1}^{N} R^u_{i,\min}}{W^\star_{r0}}.$$

Similarly, using equation (7) corresponding to the links (r,i) and (i,r), we obtain

$$\frac{R^d_{i,\min}}{W^\star_{ri}} = \frac{R^u_{i,\min}}{W^\star_{ir}} \tag{11}$$

for all $i = 1, \ldots, N$. ■

Lemma 1 also implies that,

$$\frac{\sum_{i=1}^{N} R^d_{i,\min}}{W^\star_{0r}} = \frac{\sum_{i=1}^{N} R^u_{i,\min}}{W^\star_{r0}} = \log\left(1 + \frac{\rho_{0r} P^\star_{0r}}{W^\star_{0r}}\right) = \log\left(1 + \frac{\rho_{r0} P^\star_{r0}}{W^\star_{r0}}\right), \tag{12}$$

where $\log(1 + \rho_{0r} P^\star_{0r}/W^\star_{0r})$ is the spectral efficiency of the link $(0,r)$. Therefore, $\sum_{i=1}^{N} R^d_{i,\min}/W^\star_{0r}$ represents the spectral efficiency of the link $(0,r)$, where $\sum_{i=1}^{N} R^d_{i,\min}$ is the rate that needs to be supported in that link. Similarly, $\sum_{i=1}^{N} R^u_{i,\min}/W^\star_{r0}$ represent the spectral efficiency of the link $(r,0)$ where $\sum_{i=1}^{N} R^u_{i,\min}$ is the rate that needs to supported in that link.

Corollary 4. *When the weights* $\alpha_0 = \alpha_r = \alpha_i = \alpha$ *are the same, the optimal solution to* Problem (2) *has the property that*

$$\frac{\sum_{i=1}^{N} R^d_{i,\min}}{W^\star_{0r}} = \frac{\sum_{i=1}^{N} R^u_{i,\min} + R^u_{r,\min}}{W^\star_{r0}}. \tag{13}$$

Proof: The result follows from Lemma 3. ■

The *Symmetry of Spectral Efficiency* is preserved even when the relay has its own data to deliver to the eNodeB. As a consequence of Lemma 3 we can reduce the number of equations in the optimality conditions (7) by half from $2N+2$ to $N+1$.

B. Extension to Multihop Relays - An Example

Earlier in section II we had mentioned that the results that will be obtained for a single relay can be extended to a generic system model represented by a tree structure with eNodeB as the root, the UEs being the leaf nodes and the relays as the intermediate nodes. This follows from the above fact that we can formulate a link specific optimization problem by taking the Lagrangian dual with respect to the bandwidth constraint when we know the minimum rate that each link needs to transmit. This rate depends upon the specific tree structure i.e. the route from eNodeB to different UEs through the relays and the rate demanded in uplink and downlink by each UEs. Fig. 4 shows the example of a three sector single cell system with one relay per sector serving the edge UEs as in LTE-A system. The interior users are those who lie within a radius of r_{int} from the eNodeB and others which lie outside are the exterior users. We assume that the interior UEs are served directly by eNodeB and the exterior UEs in a sector are served by their corresponding relays. By defining the dual bandwidth cost $\lambda \geq 0$ for the bandwidth constraint, this system can also be decomposed into link specific optimization problems. In fact the same principle can be extended to a mesh network where given that the routes from different sources to its destination are defined, we can decompose the network into link specific optimization problems which are as many in number as the number of links in the system, for bandwidth and power optimization. However, we focus on the single relay system because as the main problem in larger networks is the identification of the optimal relay node or the optimal route to reach the destination. Also, the insights for the single relay system which, hinge on Lagrangian dual of the bandwidth constraint, can be easily extended into multihop networks. As a specific and easy example, consider a set of N users that needs to served by a M relays as shown in Fig. 5 (with all the users being served by the M^{th} relay). If r_j is the index of relays with $j = 1, 2, \ldots, M$, $P_{r_j r_{j+1}}$ the power transmitted by the j^{th} relay to $j+1^{th}$ relay and $W_{r_j r_{j+1}}$ the bandwidth allocated to the link (r_j, r_{j+1}), the optimization problem can be formulated as, *Problem (14)*

minimize

$$\sum_{j=0}^{M-1} \left(\alpha_{r_j} P_{r_j, r_{j+1}} + \alpha_{r_{j+1}} P_{r_{j+1}, r_j}\right) + \alpha_{r_M} \sum_{i=1}^{N} P_{r_M i} + \sum_{i=1}^{N} \alpha_i P_{i r_M} \tag{14a}$$

subject to

$$\mathcal{R}(\mathsf{SNR}_{i r_M}, W_{i r_M}) \geq R^u_{i,\min} \quad i = 1, \ldots, N, \tag{14b}$$

$$\mathcal{R}(\mathsf{SNR}_{r_M i}, W_{r_M i}) \geq R^d_{i,\min} \quad i = 1, \ldots, N, \tag{14c}$$

$$\mathcal{R}(\mathsf{SNR}_{r_{j+1}, r_j}, W_{r_{j+1}, r_j}) \geq \sum_{i=1}^{N} \mathcal{R}(\mathsf{SNR}_{i r_M}, W_{i r_M}) \quad j = 0, \ldots, M-1, \tag{14d}$$

$$\mathcal{R}(\mathsf{SNR}_{r_j, r_{j+1}}, W_{r_j, r_{j+1}}) \geq \sum_{i=1}^{N} \mathcal{R}(\mathsf{SNR}_{r_M i}, W_{r_M i}) \quad j = 0, \ldots, M-1, \tag{14e}$$

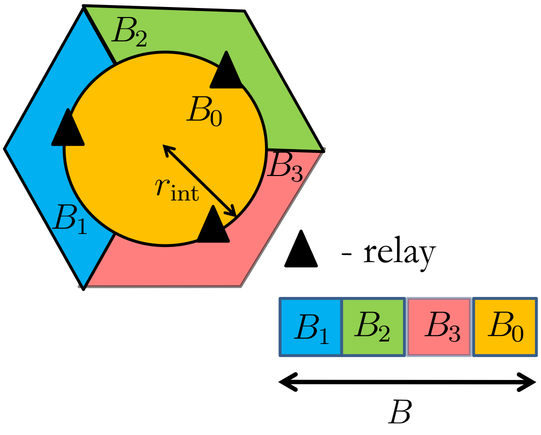


Fig. 4. Three sector single cell LTE-A system with one relay per sector. The UEs are classified as interior users and exterior users. The interior users are those who lie within a radius of r_{int} from the eNodeB and others which lie outside are the exterior users. The interior UEs are served directly by eNodeB and the exterior users in a sector are served by their corresponding relays. The relays in each sector are located at a point of intersection of edge of the core radius circle and line which bisects the sector into two equal halves.

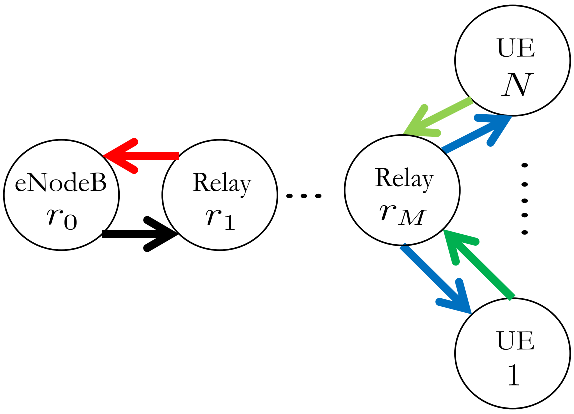


Fig. 5. Example of a multihop relaying system.

$$\sum_{i=1}^{N}(W_{ir_M}+W_{r_Mi})+\sum_{j=0}^{M-1}(W_{r_j,r_{j+1}}+W_{r_{j+1},r_j})\le B, \tag{14f}$$

variables

$$P_{ir_M}, W_{ir_M}, P_{r_Mi}, W_{r_Mi}, P_{r_j,r_{j+1}}, W_{r_j,r_{j+1}}$$

The equations (14b), (14c) represent the rate constraints on the uplink and downlink between the relay r_M and UEs. The rate constraints (14d), (14e) represent the relay to relay rate constraint requirements in which each relay has to support the end UEs data rate requirement in uplink and downlink. The equation (14f) represent the system bandwidth constraint. For the case when $M = 1$ this reduces to the *Problem (1)*. The conclusions made in lemma 1 and theorem 2 for a single relay system can be extended to *Problem (14)* and based on the Lagrangian dual of the bandwidth constraint lemma 3 can also be derived.

C. Bandwidth Shared System Vs Time Shared System

In this section we show the relation between weighted average power minimization in time shared relaying (*Problem 3*) as in 4G systems and weighted power minimization in bandwidth shared system as in *Problem (1)*. As mentioned in section III-B the time shared system with peak power constraints is represented by *Problem (3)* with the constraints (3k-P), (3l-P), (3m-P) and (3n-P). The time shared system with average power constraints is represented by *Problem (3)* with the constraints (3k-A), (3l-A) and (3m-A). The behaviour of the two systems as compared to the bandwidth sharing system is explained in the following theorem.

Theorem 5. *In terms of minimizing the weighted transmit power in the system:*

- *Time sharing with average power constraints is equivalent to bandwidth sharing.*
- *Time sharing with peak power constraints will yield a weighted power objective at least as large as bandwidth sharing.*

Proof: We replace the instantaneous power variables S_{0r}, S_{ri}, S_{r0}, S_{ir} and the bandwidth variables B_{0r}, B_{ri}, B_{r0}, B_{ir} in *Problem (3)* by their time averaged power variables P_{0r}, P_{ri}, P_{r0}, P_{ir} and average bandwidth variables W_{0r}, W_{ri}, W_{r0}, W_{ir} respectively as shown in the following the two steps,

1. $TP_{0r} = T_{\text{BS}}S_{0r}$, $TP_{ri} = T_{\text{RS}}^{\text{d}}S_{ri}$, $TP_{r0} = T_{\text{RS}}^{\text{u}}S_{r0}$, $TP_{ir} = T_{\text{UE}}S_{ir}$
2. $TW_{0r} = T_{\text{BS}}B_{0r}$, $TW_{ri} = T_{\text{RS}}^{\text{d}}B_{ri}$, $TW_{r0} = T_{\text{RS}}^{\text{u}}B_{r0}$, $TW_{ir} = T_{\text{UE}}B_{ir}$

The reformulation will result in a convex problem with the resulting objective and the rate constraints the same as that of *Problem (1)*. The bandwidth constraints in *Problem (3)* will be reformulated into,

$$\begin{aligned} W_{0r} &\le \frac{T_{\text{BS}}}{T}B, \\ \sum_{i=1}^{N} W_{ri} &\le \frac{T_{\text{RS}}^{\text{d}}}{T}B, \\ \sum_{i=1}^{N} W_{ir} &\le \frac{T_{\text{UE}}}{T}B, \\ W_{r0} &\le \frac{T_{\text{RS}}^{\text{u}}}{T}B. \end{aligned} \tag{15}$$

However, combining those with the frame time constraint (3j), will result in a single constraint given by $\sum_{i=1}^{N}(W_{ir}+W_{ri})+W_{r0}+W_{0r}\le B$ which is the same as equation (1f).

When the average power constraints given by equations (3k-A), (3l-A) and (3m-A) are employed in the time sharing system, the above reformulation of variables will result in *Problem (1)* implying that minimizing the weighted power in the time sharing system with average power constraints will result in the same objective value as in bandwidth sharing system.

When the peak power constraints given by equations (3k-P), (3l-P), (3m-P) and (3n-P) are employed the power constraints will be reformulated into

$$\begin{aligned} P_{0r} &\le \frac{T_{\text{BS}}}{T} P_{0,\max}, \\ P_{ir} &\le \frac{T_{\text{UE}}}{T} P_{i,\max}, \\ \sum_{i=1}^{N} P_{ri} &\le \frac{T_{\text{RS}}^{\text{d}}}{T} P_{r,\max}, \\ P_{r0} &\le \frac{T_{\text{RS}}^{\text{u}}}{T} P_{r,\max} \end{aligned} \tag{16}$$

Clearly these power constraints result in a lesser search space for the reformulated time sharing optimization problem than the power constraints in *Problem (1)*. Therefore, the optimal objective of the time sharing system of *Problem (3)* with peak power constraints is at least as much as that obtained from the bandwidth sharing system of *Problem (1)*. ■

When the nodes are transmitting for only a fraction of the frame time in a time sharing system, the effective rate with which they need to communicate is higher than the minimum prescribed rate to transmit the same amount of data in a frame duration. For example, for the rate constraints (3b) the effective rate at which the system has to download data in time slot T_{RS}^{d} to the users is $TR_{i,\min}^{d}/T_{\text{RS}}^{\text{d}}$ and in order to accomplish that the power has to be increased. In this situation the average power constraints will not affect the system when compared with the bandwidth shared system. However, with peak power constraints, increasing the power levels may result in hitting the power constraints and may perform worse than a bandwidth shared system. Also when we are not considering power constraints time sharing is equivalent to bandwidth sharing and hence the lemma 1, theorem 2, and lemma 3 obtained for bandwidth sharing is also relevant for time sharing system.

D. Power Minimization for a Single User Case

We consider the special case when there is only one user ($N = 1$) in the system and develop an intuitive understanding of the bandwidth sharing mechanism. The weighted power minimization problem for a single user is given by *Problem (4)* but without the power constraints. We define the bandwidth allocated in the downlink for *Problem (4)* to be $B^d = W_{0r} + W_{ri}$ and similarly that allocated in the uplink to be $B^u = W_{ir} + W_{r0}$.

Theorem 6. ***Rate Proportional Bandwidth Allocation:*** *When the weights $\alpha_0 = \alpha_r = \alpha_i = \alpha$ are equal, the bandwidths allocated in the downlink and uplink at the optimum for* Problem (4) *are independent of the link gains and are given by*

$$\begin{aligned} B^{d\star} &= \frac{R_{i,\min}^{d}}{R_{i,\min}^{u} + R_{i,\min}^{d}} B, \\ B^{u\star} &= \frac{R_{i,\min}^{u}}{R_{i,\min}^{u} + R_{i,\min}^{d}} B. \end{aligned}$$

Proof: By Lemma 1 for $N = 1$, we have at optimality

$$(W_{0r}^{\star} + W_{r0}^{\star}) + (W_{ri}^{\star} + W_{ir}^{\star}) = B.$$

Similarly for $N = 1$, using Lemma 3 implies

$$(W_{0r}^{\star} + W_{ri}^{\star}) \left(1 + \frac{R_{i,\min}^{u}}{R_{i,\min}^{d}}\right) = B.$$

Since $B^{d\star} = W_{0r}^{\star} + W_{ri}^{\star}$ the result follows. Similarly, we get $B^{u\star}$ from $B^{u\star} = B - B^{d\star}$. ■

Thus, since the optimal bandwidth split between uplink and downlink is known, the *Problem (4)* can be split into two problems one for the downlink and other for the uplink direction. Also, the bandwidth allocated to the downlink problem out of the total available bandwidth, is proportional to that fraction of the downlink rate requirement out of the total rate.

Theorem 7. *When the weights $\alpha_0 = \alpha_r = \alpha_i = \alpha$ are equal, the optimal bandwidth allocated to each of the links for* Problem (4) *depends only on the ratio of link gains ρ_{0r}/ρ_{ri}.*

Proof: Using optimality conditions (7) for links $(0, r)$, (r, i) and Theorem 6 we get,

$$\begin{aligned} &e^{\frac{R_{i,\min}^{d}}{W_{0r}}} \left(\frac{R_{i,\min}^{d}}{W_{0r}} - 1\right) + 1 \\ &\quad = \beta \left\{ e^{\frac{R_{i,\min}^{d}}{B^{d\star} - W_{0r}}} \left(\frac{R_{i,\min}^{d}}{B^{d\star} - W_{0r}} - 1\right) + 1 \right\}, \end{aligned} \tag{17}$$

where $\beta = \rho_{0r}/\rho_{ri}$. Since, the LHS is a decreasing function and the RHS is an increasing function in W_{0r}, the above equation has only one point of intersection which is dependent only on β. Therefore, the optimal W_{0r} only depends on the ratio of the link gains. Similarly, we get

$$\begin{aligned} &e^{\frac{R_{i,\min}^{u}}{W_{r0}}} \left(\frac{R_{i,\min}^{u}}{W_{r0}} - 1\right) + 1 \\ &\quad = \beta \left\{ e^{\frac{R_{i,\min}^{u}}{B^{u\star} - W_{r0}}} \left(\frac{R_{i,\min}^{u}}{B^{u\star} - W_{r0}} - 1\right) + 1 \right\}. \end{aligned} \tag{18}$$

■

The optimal bandwidth allocations $W_{0r}^{\star}, W_{r0}^{\star}, W_{ri}^{\star}, W_{ir}^{\star}$ can be found numerically from equations (17) and (18).

E. Minimizing the Relay Power - Single User Case

There may also arise situations when the relay power is precious and hence we only intend to minimize the relay power. Again we show the results for a single user case but the results for this section are easily extendible for the case for multiple users.

Minimizing Relay Power for Problem (4)

We set the value of $\alpha_r = 1$, $\alpha_0 = 0$ and $\alpha_i = 0$, $i \neq r$ when we want to minimize the relay power. When $\alpha_r = 1$, $P_{0r}^{\star} = P_{0,\max}, P_{ir}^{\star} = P_{i,\max}$. Consequently, $W_{0r}^{\star}$ and $W_{ir}^{\star}$ will be the unique solution to,

$$\begin{aligned} W_{0r} \log\left(1 + \frac{\rho_{0r} P_{0,\max}}{W_{0r}}\right) &= R_{i,\min}^{d} \qquad \text{and,} \\ W_{ir} \log\left(1 + \frac{\rho_{ir} P_{i,\max}}{W_{ir}}\right) &= R_{i,\min}^{u}, \end{aligned} \tag{19}$$

respectively. Intuitively, the eNodeB and the UEs transmit at the maximum power so that the least amount out of total available bandwidth is necessary to maintain the minimum rate from eNodeB to relay and UE to relay links. The rest of the bandwidth is used up by the relay so that it expends minimum power to maintain the minimum rate constraint in uplink and downlink. The relay power minimization problem turns out to be

minimize

$$P_{ri} + P_{r0} \tag{20a}$$

subject to

$$\mathcal{R}(\mathsf{SNR}_{r0}, W_{r0}) \geq R^u_{i,\min}, \tag{20b}$$

$$\mathcal{R}(\mathsf{SNR}_{ri}, W_{ri}) \geq R^d_{i,\min}, \tag{20c}$$

$$W_{ri} + W_{r0} \leq B - (W^\star_{0j} + W^\star_{ir}) = B', \tag{20d}$$

$$P_{ri} + P_{r0} \leq P_{r,\max}, \tag{20e}$$

variables

$$P_{ri}, P_{r0}, W_{r0}, W_{ri}.$$

This can be solved in a centralized manner at the relay.

V. NUMERICAL RESULTS

In order to further understand the engineering implications of the resource allocation, we numerically solve the weighted power minimization problem for a single user scenario. We numerically solve the *Problem (4)* which represents the single user case and also do simulations for *Problem (1)* which represents the N user case. Specifically we use parameters representative of an LTE-A system in the numerical results that follow.

A. Relaying in both Directions - Uplink and Downlink

We solve the *Problem (4)* and plot the resulting bandwidth shared, and power consumed by the links for maintaining a downlink rate of 7 Mb/s with 10 MHz bandwidth. We fix the maximum eNodeB power at 43 dBm, and the maximum power limit for relay and UE at 23 dBm which is the standard for LTE-A systems as given in [17], [18]. Also, the type-I relay is fixed at a distance of 1000 m from the eNodeB. The bandwidth allocation and power consumption is plotted for increasing ρ_{0i}/ρ_{ri} (dB) values until the optimization problem becomes infeasible. Also, increasing sequence of ρ_{0i}/ρ_{ri} (dB) values indicate increasing distance between relay r and UE i when the eNodeB to relay distance is fixed. The optimization problem finds a feasible point at maximum relay power untill the eNodeB reaches its maximum power and thereafter the system is infeasible. Figs. 6 and 7 plot the bandwidth partition and power expenditure at each of the links. The plots are for increasing ρ_{0i}/ρ_{ri} (dB) values until the optimization problem becomes infeasible. From Fig. 7 we can deduce that the reason for infeasibility is the relay power constraint. While the eNodeB power for downlink P_{0r} and UE i power P_{ir} in the uplink have not reached their maximum but the relay power $(P_{ri} + P_{r0})$ is at its maximum value of 23 dBm. Just before the system becomes infeasible the optimization problem tries to allocate as much bandwidth as possible for the (r, i) and $(r, 0)$ links, so that even with the limited 23 dBm relay power the system can still maintain the required data rate. This can be seen from the increasing W_{ri},W_{r0} and the decreasing W_{0r}, W_{ir} near the infeasibility region in the Fig. 6. This can also be verified from Fig. 7 where there is a corresponding increase in the powers, P_{0r} and P_{ir} to maintain the required data rate as the system approached infeasibility. To summarize, the bandwidth sharing utilizes that power that would otherwise be unused in the eNodeB and the UE i while transferring some extra bandwidth to the relatively power constrained relay to maintain required rate. When there is no bandwidth sharing, each link is given a fixed 2.5 MHz bandwidth, and the relay power optimized between uplink and downlink. However, when there is no bandwidth sharing, it is found to be infeasible after when $\rho_{0i}/\rho_{ri} > -16$ dB, while with sharing infeasible region starts when $\rho_{0i}/\rho_{ri} > -14$ dB (Fig. 8). This is because there is not enough relay power to maintain the required rate in uplink and downlink. Converting to distances, this corresponds to UE being anywhere in the region of 293 m from relay when there is no bandwidth sharing and 342 m when there is bandwidth sharing. Thus bandwidth sharing results in a 49 m increase in coverage area. This is illustrated in Fig. 9. Also, when both the schemes are feasible the total power consumption with bandwidth sharing is 3 dB less than when there is no bandwidth sharing.

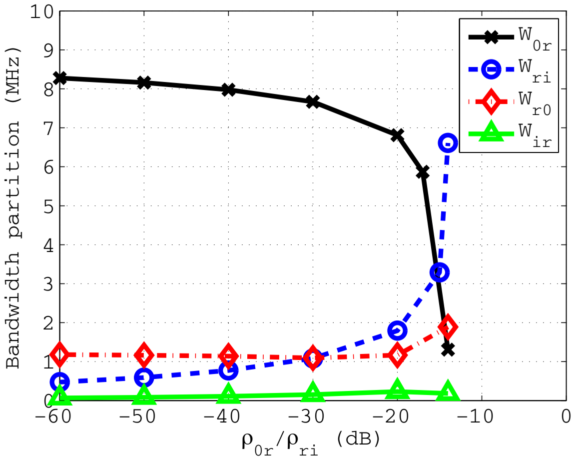


Fig. 6. Numerical Solution to *Problem (4)* showing bandwidth sharing among links. Notice the increased bandwidth allocation W_{r0} and W_{ri} at $\rho_{0r}/\rho_{ri} \geq -18$ dB. This is because the relay power has reached its maximum prescribed limit and minimum required rate is maintained by allocating more bandwidth to the links.

B. Numerical Results for Downlink - Multiple Users

In this section we extend the numerical simulation results to a N-user downlink scenario. We consider a LTE-A system with strict Fractional Frequency Reuse (FFR) scheme. Fig. 4 shows the diagram of a three sector single cell LTE-A system with one relay per sector. The N users are assumed to be uniformly distributed within the cell and they are classified as interior users and exterior users. The interior users are those who lie within a radius of r_{int} from the eNodeB and others which lie outside are the exterior users. We assume that exterior users in a sector are served by their corresponding

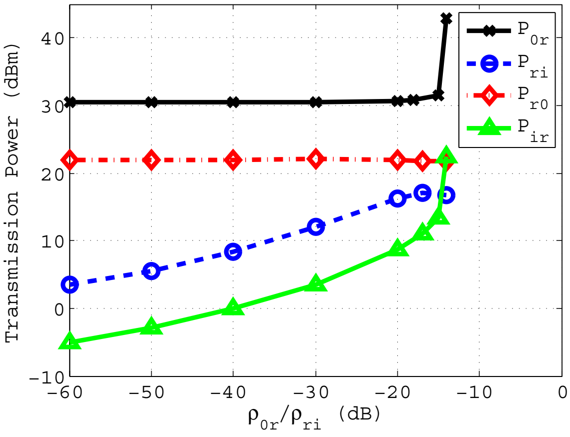


Fig. 7. Numerical Solution to *Problem (4)* showing power consumption on individual nodes. The relay power consumption is the sum P_{r0} + P_{ri}. Notice the increased power consumption in eNodeB and UE after $\rho_{0r}/\rho_{ri} \geq -18$ dB i.e. once the relay power reached the maximum prescribed limit. This is because a large portion of the bandwidth is allocated to links $(r, 0)$ and (r, i) and consequently the links $(0, r)$ and $(r, 0)$ are allocated less bandwidth out the total.

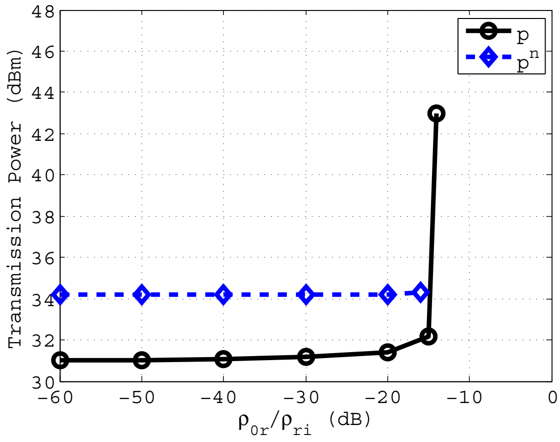


Fig. 8. Numerical Solution to *Problem (4)* showing total Power consumption in the system. P represents the total power consumed when bandwidth is shared and P^n for fixed allocation of bandwidth.

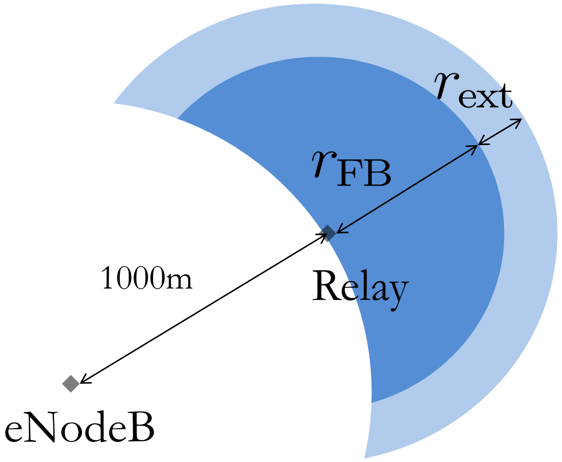


Fig. 9. The diagram shows coverage extension when bandwidth is shared as compared to fixed allocation of bandwidth. The relay is placed at position of 1000 m from the eNodeB. We assume that the UE does not have direct access to or from the eNodeB and communicates only through the relay. Here r_{FB} denotes the relay coverage till infeasibility for fixed bandwidth allocation scheme (dark blue region) and r_{ext} the coverage extension due to bandwidth sharing (light blue region). For *Problem (4)*, $r_{\text{FB}} = 293$ m and $r_{\text{ext}} = 49$ m. Also within the region where both schemes are feasible bandwidth sharing for *Problem (4)* consumes up to 3 dB less power.

relays. The relays in each sector are located at a point of intersection of edge of the core radius circle and line which bisects the sector into two equal halves. The total system bandwidth B is split into four chunks B_0, B_1, B_2 and B_3 as per the LTE-A strict FFR schemes [19], [20]. B_0 is the total bandwidth allocated to the cell interior users and B_1, B_2, B_3 corresponds to that allocated to the cell exterior users in sector 1, 2, 3 respectively. In [19], the fraction of the total bandwidth B, allocated to the interior users (B_0) is based on the area covered by the interior radius and is given by $B_0 = BA_{\text{int}}/A_{\text{tot}}$ as a result of the uniform distribution of the users, where $A_{\text{int}} = \pi r_{\text{int}}^2$ is the interior area, A_{tot} is the area of the cell and $B_r = (B - B_0)/3$. The eNodeB allocates bandwidth to the interior users flexibly from B_0 and the relay r allocates bandwidth flexibly to the exterior users in the corresponding sector r from the band B_r. While it is clear that the interior users are allocated bandwidth from B_0 and the exterior users in each sector are served from B_r, $r = 1, 2, 3$, it is not apparent from where the bandwidth for the eNodeB-relay, $(0, r)$ links should be allocated. In one case, the three relays can be allocated bandwidth from B_0 reserved for the interior users as each relay can be viewed as another core user. While in another case the $(0, r)$ link can be allocated bandwidth from B_r as each relays serves only the exterior users. As it is not clear which is better in terms of minimizing the total power in the system, we classify those into two allocation methods,

I The $(0, r)$ links for $r = 1, 2, 3$ are allocated bandwidth from B_0.
II The $(0, r)$ link is allocated bandwidth from the exterior user bandwidth B_r.

In method-I allocation scheme the eNodeB to the interior users link have to operate at higher powers so as to compensate for the bandwidth lost to the $(0, r)$ links. In contrast, for Method-II B_0 is used exclusively reserved to the interior users and B_r is used to support the sum of the data rate of the exterior users served by the relay r in the $(0, r)$ link in addition to the exterior users. This leads to a increase in total power consumption for the exterior users. We compare the above two possible schemes using the bandwidth partition based on [19] for varying interior radius. The method-I consumes more power than method-II when the interior area is small as compared to the exterior area. This is because when r_{int} is small B_0 will be consequently less and with the less availability of bandwidth the eNodeB has to support both the interior users and the three relays. Coupled with the less availability of B_0, since the number of users classified as exterior is larger in number and that the $(0, r)$ links also has to support the sum of the exterior user rates makes method-I consume more power than method-II. Method-II on the other hand clearly consumes more power when the interior area is large as compared to the exterior area. As the interior area is large, B_r is very low and more power is required to satisfy the both $(0, r)$ link (whose link gain is weaker as the relay is in the edge of interior radius) and the corresponding exterior users in sector-r. This obvious disadvantage of the simple area based bandwidth partitioning can be mitigated by considering some additional bandwidth for $(0, r)$ links given to B_0 or B_r depending on whether we employ method-I or method-II. This is given in the next section.

1) Area Based Bandwidth Partition: Let δ be the expected number of users per unit area. Therefore the expected number of users in the interior region can be given as $n_0 = \delta A_{\text{int}}$, and the expected number of users served by each relay can be given as $n_1 = n_2 = n_3 = \delta(A_{\text{tot}} - A_{\text{int}})/3$. Therefore, the expected number of links in the system is given by $n_0 + 2\sum_{r=1}^{3} n_r$ with one link per directly-served users and two links per relay-served users. The bandwidth allocation B_0 is proportional to the number of links that is served by the eNodeB. For method-I we have,

$$B_0 = B\frac{n_0 + \sum_{r=1}^{3} n_r}{n_0 + 2\sum_{r=1}^{3} n_r} = B\frac{1}{2 - A_{\text{int}}/A_{\text{tot}}} \tag{21}$$

and,

$$B_r = \frac{B}{3}\frac{1 - A_{\text{int}}/A_{\text{tot}}}{2 - A_{\text{int}}/A_{\text{tot}}} \tag{22}$$

We denote the above partition “A-I” which stands for area based bandwidth partition using method-I. Similarly, for method-II we have,

$$B_0 = B\frac{n_0}{n_0 + 2\sum_{r=1}^{3} n_i} = B\frac{A_{\text{int}}/A_{\text{tot}}}{2 - A_{\text{int}}/A_{\text{tot}}} \tag{23}$$

$$B_r = \frac{2B}{3}\frac{1 - A_{\text{int}}/A_{\text{tot}}}{2 - A_{\text{int}}/A_{\text{tot}}} \tag{24}$$

We denote the above the partition as “A-II”. Fig. 10 shows the plots of total bandwidth allocated to the interior region for different values of core radius. The bandwidth partition “A-I” allocates sufficient bandwidth to the interior region even for low interior radius by always considering that it is also used to support the $(0, r)$ links. Also, bandwidth for partition “A-II” reduces B_0 when the interior area is large and hence ensures that there is sufficient bandwidth $B_r = (B - B_0)/3$ to support the $(0, r)$ links. We then compare this with a completely flexible bandwidth allocation where B_0, B_1, B_2 and B_3 are also variables in the optimization process with $B_0 + \sum_{r=1}^{3} B_r \leq B$. In contrast to the area based bandwidth partition, the flexible bandwidth allocation scheme can be considered as a link gain based optimal allocation of bandwidth per links for minimizing the system power. Fig. 10 also shows the sum of bandwidth allocated to the direct links and eNodeB-relay links in a flexible scheme. It is observed that in A-I the link based partition allocates more than optimal bandwidth to the interior region when the interior radius is small but less than optimal bandwidth when the interior radius is large. We also implement a fixed bandwidth allocation where each active links is allotted a fixed bandwidth. In the fixed bandwidth allocation scheme B_0 is given by equation (23), $B_r, r = 1, 2, 3$ is partitioned as in equation (22) and each $(0, r)$ link is assigned a third of the rest which in turn is equal to B_r itself. Each of the links within the partitioned region is assigned a fixed bandwidth equal to the ratio of bandwidth in the region to the number of users in the region.

The simulations were done for 10 UEs in the system uniformly distributed in the cell. The total bandwidth availability is 50 MHz and the minimum downlink rate requirement is uniformly distributed between $0 - 10$ Mbps. Fig. 11 shows the average power consumption for A-I, A-II, flexible and fixed bandwidth allocation schemes. While is it clear that a completely flexible bandwidth allocation should consume the least power, it also achieves at least 3 dB reduction in average power than the fixed bandwidth allocation for all values of interior radius. This is a direct extension and stronger result as compared to the single user case where the gain was up to 3 dB (Fig. 8). At interior radius of 200 m the fixed bandwidth allocation scheme seems to consume much more power than for other values of interior radius. The behaviour is due to the fact that the bandwidth allocated to interior users B_0 according to equation (23) (same as “A-II’ curve in Fig. 10) is very low resulting in high power consumption. Although the bandwidth allocation is exactly same for “A-II” scheme the average power consumed is not as much as the fixed scheme because within the interior region bandwidth is flexibly allocated.

The average power consumption for area based method-I and II schemes lie between that of fixed and flexible schemes. They represent an intermediate and implementable method for LTE-A systems as compared to flexible allocation and fixed allocation schemes. The flexible allocation scheme may not be realistic to implement and fixed allocation results in at least 3 dB additional power consumption which could be avoided. The flexible scheme is unrealistic from the point of view that B_0 and $B_r, r = 1, 2, 3$ are varying for every snapshot instance of the optimization problem while the FFR in LTE-A requires it to fixed. From Fig. 11, it is found there is less than 1 dB difference between the area based method-I and flexible allocation scheme for values of interior radius between 500 to 800 m. This can be attributed to the fact that in Fig. 10 there is not much difference in bandwidth allotted by the flexible scheme as compared with the method-I. The area based method-II scheme on the other hand has the average power consumption more than that of method-I primarily because of its inherent bandwidth partitioning technique. In method-II the B_0 allotted seems to be more than enough for its direct links while B_r is comparatively less to support the $(0, r)$ links as well as relay to UE links. A supporting viewpoint to this argument is that as the interior radius increases B_0 increases and difference in average power consumption between method-I and method-II increases as seen in Fig. 11.

VI. CONCLUSION

In this paper we introduced bandwidth and power optimization in a cellular system employing type-I relaying system. We formulated a weighted power minimization problem, optimizing over both power and bandwidth under rate, bandwidth and power constraints for serving multiple users We developed theoretical insights into the nature of optimal solution when

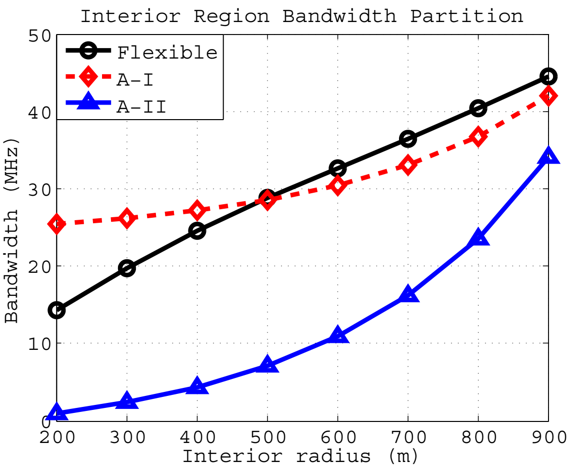


Fig. 10. The plot shows the bandwidth partitioned for the interior region with link based, area based as well as flexible allocation. For the flexible partition the sum of bandwidth allocated to the interior users and the eNodeb to relay links is plotted.

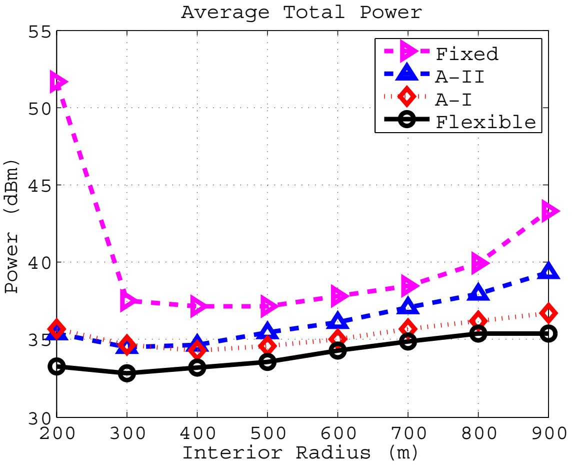


Fig. 11. The plot the shows the average power consumption for four different schemes - Flexible, Fixed, A-I and A-II. It is seen that the flexible bandwidth allocation consumes at least 3 dB less power as compared to fixed bandwidth allocation.

the system has unconstrained maximum power. From the implementation perspective, it is also seen that bandwidth sharing provides total power gain of about 3.5 dB as compared to baseline scheme of fixed allocation of bandwidth per link. This is a significant gain in view of the actual amount of eNodeB power in watts saved and fact that eNodeB power expenditure forms 80% of the total power spent in the cellular system [10]. Future directions of interest include cases with multiple cells and when Inter Cell Interference (ICI) is significant.

APPENDIX A
PROOF OF LEMMA 1

We will prove that the rate constraints are always tight for *Problem (1)* at optimum by a simple contradiction statement. Assume that for *Problem (1)*, the constraint (1b) is slack at optimum. Therefore,

$$W_{ir}^\star \log\left(1+\frac{\rho_{ir}P_{ir}^\star}{W_{ir}^\star}\right) > R_{i,\min}^u$$

However, if $P_{ir}^\star$ exists then there always exists $P_{ir} < P_{ir}^\star \leq P_{i,\max}$ such that

$$W_{ir}^\star \log\left(1+\frac{\rho_{ir}P_{ir}}{W_{ir}^\star}\right) = R_{i,\min}^u,$$

and also has a lesser objective value. Therefore, P_{ir} is the optimum and not $P_{ir}^\star$ and at the optimum value the rate constraint is tight. This is true for all rate constraints (1b) and (1d) in *Problem (1)*. For constraint (1c), since

$$W_{ir}^\star \log\left(1+\frac{\rho_{ir}P_{ir}^\star}{W_{ir}^\star}\right) = R_{i,\min}^u, \quad \forall i$$

and using the same arguments as before we can say that at optimum,

$$W_{r0}^\star \log\left(1+\frac{\rho_{r0}P_{r0}^\star}{W_{r0}^\star}\right) = \sum_{i=1}^{N} R_{i,\min}^u.$$

Similarly, for *Problem (1)* we have at optimum,

$$W_{0r}^\star \log\left(1+\frac{\rho_{0r}P_{0r}^\star}{W_{0r}^\star}\right) = \sum_{i=1}^{N} R_{i,\min}^d.$$

Again, we prove that the bandwidth constraint is tight at optimum by contradiction. Assume that the bandwidth constraint is not tight at optimum. Therefore,

$$\sum_{i=1}^{N}(W_{ir}^\star + W_{ri}^\star) + W_{r0}^\star + W_{0r}^\star < B.$$

Consider the constraint (1b) and from the earlier conclusion we know that at optimum the rate constraints are tight. Therefore,

$$P_{1r}^\star = W_{1r}^\star\left(e^{\frac{R_{1,\min}^u}{W_{1r}^\star}} - 1\right) \tag{25}$$

for user 1. Assume the optimal objective of the *Problem (1)* is $p^\star$. However, clearly we can find another bandwidth allocation $W_{1r}^\star + \delta W_{1r}$ with $\delta W_{1r} > 0$ such that the corresponding power P_{1r} consumed to meet the rate $R_{1,\min}^u$ is less than $P_{1r}^\star$ and $\sum_{i=1}^{N}(W_{ir}^\star + W_{ri}^\star) + W_{r0}^\star + W_{0r}^\star + \delta W_{1r} = B$. Consequently the new objective value $p < p^\star$. This implies that our initial assumption that $p^\star$ is the optimum is not true and any situation where the bandwidth constraint is slack does not give the optimum solution. Hence, if an optimum solution exist for the *Problem (1)* then at that solution the bandwidth constraint has to be tight.

ACKNOWLEDGEMENTS

This work is supported by the NSF grants CCF-0634973 and CNS-0721826. The authors also extend their acknowledgements to Mikael Prytz from Ericsson for his valuable comments that helped to improve this work.

REFERENCES

[1] S. Kumar and N. Marchetti, "IMT-Advanced: technological requirements and solution components," in *Proc. 2nd Int. Workshop Cognitive Radio Advanced Spectrum Manage.*, pp. 1–5, May 2009.

[2] S. Parkvall, E. Dahlman, A. Furuskar, Y. Jading, M. Olsson, S. Wanstedt, and K. Zangi, "LTE-Advanced–evolving LTE towards IMT-Advanced," in *Proc. IEEE* 68^{th} *Vehicular Tech. Conf.* pp. 1–5, Sep. 2008.

[3] D. Astely, E. Dahlman, A. Furuskar, Y. Jading, M. Lindstrom, and S. Parkvall, "LTE: the evolution of mobile broadband," *IEEE Commun. Mag.*, vol. 47, pp. 44–51, Apr. 2009.

[4] S. W. Peters and R. W. Heath, "The future of WiMAX: multihop relaying with IEEE 802.16j," *IEEE Commun. Mag.*, vol. 47, pp. 104–111, Jan. 2009.

[5] Y. Yang, H. Hu, and Jing Xu, "Relay technologies for WiMAX and LTE-Advanced mobile systems," *IEEE Commun. Mag.*, Oct. 2009.

[6] L. Long and E. Hossain, "Multihop cellular networks: potential gains, research challenges, and a resource allocation framework," *IEEE Commun. Mag.*, vol. 45, pp. 66–73, Sep. 2007.

[7] C. T. K. Ng and G. J. Foschini, "Transmit signal and bandwidth optimization in multiple-antenna relay channels." Available: http://arxiv.org/PS_cache/arxiv/pdf/1001/1001.2938v1.pdf

[8] D. Zhang, O. Ileri, and N. B. Mandayam, "Bandwidth exchange as an incentive for relaying," in *Proc. 42nd Annual Conference on Information Sciences and Systems*, pp. 749–754, Mar. 2008.

[9] D. Zhang, R. Shinkuma, and N. B. Mandayam, "Bandwidth exchange: an energy conserving incentive mechanism for cooperation," *IEEE Trans. Wireless Commun.*, vol. 9, no. 6, June 2010.

[10] F. Richter, A. J. Fehske, and G. P. Fettweis, "Energy efficiency aspects of base station deployment strategies for cellular networks," in *Proc. IEEE Veh. Tech. Conf.*, pp. 1–5, Sep. 2009.

[11] J. Acharya, "Utility maximization in dynamic spectrum allocation," Ph.D. thesis, Rutgers University, Dept. of Electrical and Computer Engineering, May 2009.

[12] T. Girici, "Joint power, subcarrier and subframe allocation in Multihop relay networks," *Int. J. Commun. Syst.*, vol. 22, pp. 835–855, Jan. 2009.

[13] 3GPP TR 36.814, "Further advancements for E-UTRA Physical Layer Aspects (Release 9)." Available: http://www.3gpp.org/ftp/Specs/html-info/36-series.htm

[14] M. Hart *et al.*, "Out-of-band relay clarification," IEEE C802.16j-08/079r1, 2008.

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

[16] M. Grant and S. Boyd, "CVX: Matlab Software for Disciplined Convex Programming, version 1.21," http://cvxr.com/cvx, May 2010.

[17] 3GPP TS 36.101, "User equipment (UE) radio transmission and reception." Available: http://www.3gpp.org/ftp/Specs/html-info/36-series.htm.

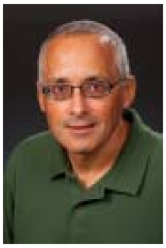
[18] 3GPP TS 36.104, "Base station (BS) radio transmission and reception." Available: http://www.3gpp.org/ftp/Specs/html-info/36-series.htm.

[19] T. Novlan, J. G. Andrews, I. Sohn, R. K. Ganti, and A. Ghosh, "Comparison of fractional frequency reuse approaches in the OFDMA cellular downlink," in *Proc. IEEE Globecom Conf.*, pp. 1–5, Dec. 2010.

[20] N. Chun, L. Pei, and S. Panwar, "Interference management using frequency planning in an OFDMA based wireless network," in *Proc. IEEE Wireless Commn. Netw. Conf.*, pp. 998–1003, Mar. 2011.

[21] N. Krishnan, J. S. Panchal, N. B. Mandayam, and R. D. Yates, "Bandwidth sharing in LTE-A systems," 48^{th} *Annual Allerton Conf. on Commun. Control and Computing*, Oct. 2010.

Narayanan Krishnan received his B.Tech degree in Electronics and Communications Engineering from College of Engineering, Trivandrum in India in 2004. From 2004 till 2007, he was with Infosys Technologies Inc, Bangalore as a network test engineer. He then completed his Masters degree at Kansas State University and is currently a Ph.D. candidate at WINLAB, Rutgers University. His research interests are in resource allocation for 4G cellular systems.



Roy D. Yates received the B.S.E. degree in 1983 from Princeton and the S.M. and Ph.D. degrees in 1986 and 1990 from MIT, all in Electrical Engineering. Since 1990, he has been with the Wireless Information Networks Laboratory (WINLAB) and the ECE department at Rutgers University. He presently serves as an Associate Director of WINLAB and a Professor in the ECE Dept. He also serves as an associate editor of the IEEE TRANSACTIONS ON INFORMATION THEORY. He is a co-author (with David Goodman) of the text *Probability and Stochastic Processes: A Friendly Introduction for Electrical and Computer Engineers* published by John Wiley and Sons. He is a 2011 IEEE Fellow and a recipient of the 2003 IEEE Marconi Prize Paper Award in Wireless Communications, the best paper award for the ICC 2006 Wireless Communications Symposium, and the 2011 Rutgers University Teacher-Scholar award. His research in wireless networks includes data dissemination, interference mitigation, secret communication, and spectrum regulation.

Narayan B Mandayam is currently the Peter D. Cherasia Faculty Scholar at Rutgers University. He received the B.Tech (Hons.) degree in 1989 from the Indian Institute of Technology, Kharagpur, and the M.S. and Ph.D. degrees in 1991 and 1994 from Rice University, all in electrical engineering. From 1994 to 1996, he was a Research Associate at the Wireless Information Network Laboratory (WINLAB), Rutgers University before joining the faculty of the Electrical and Computer Engineering department at Rutgers where he became Associate Professor in 2001 and Professor in 2003. Currently, he also serves as Associate Director at WINLAB. He was a visiting faculty fellow in the Department of Electrical Engineering, Princeton University in 2002 and a visiting faculty at the Indian Institute of Science in 2003. His research interests are in various aspects of wireless data transmission including system modeling and performance, signal processing and radio resource management with emphasis on techniques for cognitive radio networks.

Dr. Mandayam is a recipient of the Fred W. Ellersick Prize from the IEEE Communications Society in 2009 along with O. Ileri for their work on dynamic spectrum access models and spectrum policy. He is also a recipient of the Institute Silver Medal from the Indian Institute of Technology in 1989 and the National Science Foundation CAREER Award in 1998. He is a coauthor with C. Comaniciu and H. V. Poor of the book *Wireless Networks: Multiuser Detection in Cross-Layer Design* (Springer, NY). He has served as an Editor for the journals IEEE COMMUNICATION LETTERS and IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He has also served as a guest editor of the IEEE JSAC Special Issues on Adaptive, Spectrum Agile and Cognitive Radio Networks (2007) and Game Theory in Communication Systems (2008). He is a Fellow of the IEEE.

Jignesh S. Panchal is a RF performance engineer at Alcatel-Lucent, Murray Hill, NJ. He received his B.E. in Electronics and Communication Engineering from LD College of Engineering, Gujarat University, India in 1996 and M.S. in Electrical and Computer Engineering from Rutgers University in 1998. He is currently a Ph.D. candidate in WINLAB, Rutgers University advised by Prof. Roy D. Yates. His research interests include cellular resource scheduling and sharing in 4G LTE and future cellular network design and architecture.