RESEARCH   **Open Access**

Resource allocation for network MIMO systems with HARQ and partial channel knowledge

Stefano Tomasin

Abstract

In this article, we consider a cellular system, where multiple base stations (BSs) cooperate to decode signals coming from mobile terminals (MTs), using a hybrid automatic repeat request error protection mechanism. We focus on the problem of weighted network sum rate maximization, by addressing the following three resource allocation problems: (a) the scheduling of MT transmissions, (b) the scheduling of cooperating BSs for each MT, and (c) the allocation of resources on the backhaul network. Both a star and a ring architecture for the backhaul network are considered. We propose two approaches, where resource allocation on the backhaul is performed either *before* or *after* reception of the packet on the wireless link, with a either partial (statistical) or full knowledge of instantaneous channels. Since both assignments are NP-hard, we also propose greedy algorithms. Based on the numerical results we compare the performance of the various approaches in a typical cellular scenario. The main results, among others, are that (a) the perfect channel knowledge allows a reduction of the backhaul usage by 2.5 to 3 times with respect to the case of partial channel knowledge, and (b) the ring architecture turns out to be more flexible than the star architecture. In particular, the most flexible architecture is given by the RING topology with decoding at the BS, where the backhaul throughput can be allocated to the BSs in any proportion, while at the same time the transfer of information from the serving BS to the RNC is avoided.

1 Introduction

In cellular wireless networks, multicell processing (MCP) refers to a scenario where different base stations (BSs) cooperate in order to jointly decode packets coming in the uplink from mobile terminals (MTs). The resulting *macro diversity* due to the different location of BSs brings great benefits, especially for MTs at the edge of the cell. On the other hand, MCP requires the exchange of information among the cooperating BSs over a suitable *backhaul* communication infrastructure. However, constraints on the rates achieved on the backhaul constitute a bottleneck.

In order to take into account the backhaul constraint, some works have considered a fixed number of cooperating BSs. In this respect, the assignment of MTs to sets of cooperating BSs has been optimized in [1,2]. Close-to-optimal MT assignment is provided in [3], using a genetic algorithm and a sphere decoder. A heuristic approach, based on the distance of MTs from the BSs, is considered in [4]. Other works have focused on the minimization of

the amount of bits exchanged on the backhaul for a given rate on the wireless link. With this purpose distributed compression of the signal received at the BSs is proposed in [5], while a scheduling algorithm for the maximization of the network rate has been derived in [6]. In an effort to reduce the optimization complexity, Diehm et al. [7] propose first to group MTs according to both the wireless link capacity and the backhaul occupation, and then to schedule transmissions within each group. In [8], time-varying channels are considered in a scenario where a secondary BS is connected to a primary BS, and multiple MTs are served. The capacity of the macro diversity approach with backhaul constraints has been studied in [9] for the special case of a circular Wiener model. For an MCP scenario where MTs employ hybrid automatic repeat request (HARQ) and transmit at a fixed rate, the minimum number of cooperating BSs required to achieve a target outage probability after a fixed number of HARQ slots has been derived in [10].

In this article, we consider an MCP system, where MTs use HARQ. We consider two possible architectures for the backhaul network, namely star and ring architectures. In

Correspondence: tomasin@dei.unipd.it
Department of Information Engineering, University of Padova, via Gradenigo 6/B, Padova, Italy

the star architecture, each BS is connected directly to a radio network control (RNC) node, while in the ring architecture all BSs are connected among them and to the RNC by a ring, which operates as a single bus, i.e., any couple of BSs (or BS and RNC) may directly communicate without passing through the RNC. Decoding of packets coming from each MT may occur either at a BS or at the RNC.

Here, we focus on scheduling techniques for the maximization of the weighted network throughput. Note that other metrics can be used, such as the average packet delay or the packet loss rate [11]. However, by adjusting the weights we could take into account at least partially these issues (such as transmit buffer overflow) and also ensure fairness among users, as shown for example in [12]. Moreover, the weighted sum throughput is among the most important performance metrics and most literature focuses on it. Lastly, note that the maximization takes into account the HARQ outage, i.e., the possibility that the maximum number of packet retransmissions is achieved, since in this case the throughput is zero. In this article, the weighted sum rate maximization is achieved by (a) scheduling MT transmissions, (b) selecting cooperating BSs for MCP, and (c) allocating resources on the backhaul. For the first scheduling on the wireless link from MTs to BSs, we rely only on two types of knowledge: the path loss and the second-order statistics of fading. Sub-problems (b) and (c) can jointly be seen as a second scheduling on the backhaul, from BSs to the place where decoding occurs. This second scheduling can be performed either *before* or *after* wireless transmissions. When backhaul resources are scheduled *before* MT transmission, we can use only a partial channel state information (CSI), and we denote this case *fixed BS assignment—FBA*. When backhaul resources are scheduled *after* MT transmission, channels are estimated at the BSs through the received signals and therefore scheduling benefits from full CSI. This latter problem is indicated as *opportunistic BS assignment—OBA*, since cooperation among BSs is opportunistic with respect to actual channel conditions. Although OBA has the advantage of a more precise CSI, it yields (i) exchange of channel estimates, (ii) additional storage and processing requirements at the BS, and (iii) additional decoding delay.

For both OBA and FBAs, we formulate the maximization problem, considering the constraints on the rate supported by the backhaul. In particular, while OBA has as input the actual rates achieved on the wireless link, in FBA we consider the statistics of the rate on the wireless link, and we aim at maximizing the expected weighted throughput. It turns out that OBA is a mixed integer linear programming problem, while FBA is a multidimensional multiple-choice knapsack (MMCK) problem. In order to reduce the complexity of these approaches we also propose greedy algorithms.

The rest of the article is organized as follows. In Section 2, the MCP scenario with HARQ is described, providing the details of the models of both the wireless and the backhaul channels. The FBA problem is addressed in Section 3, while the OBA problem is considered in Section 4. The extension of the problems to a scenario including HARQ is provided in Section 5. Numerical results on a typical cellular scenario are presented in Section 6. Lastly, conclusions are outlined in Section 7.

2 System model

We consider a cellular system with K BSs that serve N MTs. The BSs are identified by indices in the set $\mathcal{K} = \{1, 2, \ldots, K\}$, while the MTs are identified by indices in the set $\mathcal{N} = \{1, 2, \ldots, N\}$. In this article, we focus on uplink communications from the MTs to the BSs. The MT with index n is denoted $\mathcal{T}(n)$. Each $\mathcal{T}(n)$ has a *serving BS* b_n, whose antenna main lobe illuminates the cell where the MT is located. Let $\mathcal{M}(k)$ be the set of MT indices served by BS k.

Data transmission from MTs is organized into packets. Each packet occupies a portion of time and a band. The time occupied by a packet is denoted *slot*, while the band occupied by a packet is denoted *resource block* (RB). In each RB and in each slot, more MTs may transmit simultaneously. The set of MTs transmitting on RB b is denoted $\mathcal{R}_b$. We will also identify the set of all the transmitting MTs as $\mathcal{U}$, regardless of the occupied RB. Clearly, we have $\mathcal{U} = \cup_b \mathcal{R}_b$. In the following, we focus on the transmission on a generic single slot.

A backhaul network is available to connect BSs. Moreover, the RNC ensures the exchange of data with the external communication network. Transmissions by the MTs are centrally scheduled by the RNC. We consider a cooperative scenario, where multiple BSs forward information on the received signal from the same MT over the backhaul network. Packet decoding is performed either at the RNC or at the serving BS (that collects the information from other BSs).

In the following, we consider capacity achieving codes and perfect coding over the backhaul network; therefore, we will focus on the mutual information of the message transmitted by the MT and that received by the BSs, as well as the mutual information between the message transmitted by the MT and the message reaching the RNC through the backhaul.

2.1 Channel model

As in nowadays cellular systems, we assume that MTs have a single antenna each, while BSs are equipped with M antennas each. Although we consider the case where all BSs have the same number of antennas, more general cases can easily be derived from our analysis. Within each RB, the channels from MTs to BSs are modeled as

flat block-fading, i.e., time-invariant over a *slot*, with independent realizations among consecutive slots, and with additive zero-mean Gaussian noise. The assignment of MTs to RBs is out of the scope of this article and is taken as input to our problems.

We assume that MTs transmit with unitary power. Let $H_{k,n}$ be the M-size column vector containing the channel from $\mathcal{T}(n)$ to BS k. The set of indices of BSs that cooperate for the decoding of the packet coming from $\mathcal{T}(n)$ is denoted as

$$\mathcal{C}_n = \{c_{1,n}, c_{2,n}, \ldots, c_{v_n,n}\}, \tag{1}$$

where $v_n = ||\mathcal{C}_n||$ is the number of cooperating BSs.

We define the $v_n M$-size column vector

$$\boldsymbol{H}_n = [\boldsymbol{H}_{c_{1,n},n}^T, \boldsymbol{H}_{c_{2,n},n}^T, \ldots, \boldsymbol{H}_{c_{v_n,n},n}^T]^T, \tag{2}$$

where ^{T} denotes the transpose operator, and we assume that the entries of $\boldsymbol{H}_n$ are zero-mean complex Gaussian variables with $v_n M \times v_n M$ correlation matrix

$$\boldsymbol{\Sigma}_n \triangleq \mathrm{E}[\boldsymbol{H}_n \boldsymbol{H}_n{}^H], \tag{3}$$

where ^{H} denotes the Hermitian operator. The correlation matrix includes the path loss factor

$$G(d_{k,n}) = G(1)d_{k,n}^{-3.4}, \tag{4}$$

due to distance $d_{k,n}$ of $\mathcal{T}(n)$ from BS k. In particular, when channels with respect to all antennas are uncorrelated, $\boldsymbol{\Sigma}_n$ is diagonal with entries

$$[\boldsymbol{\Sigma}_n]_{pM,pM} = G(d_{c_{p,n},n}), \tag{5}$$

where $p = 1, 2, \ldots, v_n$. Note that other scenarios can easily be obtained by simple changes. For example, in LTE the MT may have two antennas transmitting the same signal but with different phases. Assuming that the phases of the two signals perfectly compensate the channel phases (i.e., the two signals sum coherently at the serving BS), then the channel to the serving BS will be chi-distributed with four degrees of freedom. On the other hand, the channels with respect to the other BSs will be Gaussian with variance 2.

Let

$$\bar{\mathcal{R}}_n = \mathcal{R}_{\zeta_n} \setminus \{n\} \tag{6}$$

be the set of all MTs transmitting on RB ζ_n except $\mathcal{T}(n)$. Let $r_1^{(n)}, \ldots, r_{\bar{\mathcal{R}}_n}^{(n)}$, be the entries of $\bar{\mathcal{R}}_n$ and let us define the $M v_n \times ||\bar{\mathcal{R}}_n||$-size matrix of channels from all MTs transmitting on RB ζ_n, except the channel from $\mathcal{T}(n)$, as

$$\tilde{\boldsymbol{H}}_n = [\boldsymbol{H}_{r_1^{(n)}}, \boldsymbol{H}_{r_2^{(n)}}, \ldots, \boldsymbol{H}_{r_{||\bar{\mathcal{R}}_n||}^{(n)}}]. \tag{7}$$

Lastly, let us define the $M v_n \times ||\mathcal{R}_{\zeta_n}||$-size matrix containing the channels from all MTs transmitting on RB ζ_n as

$$\bar{\boldsymbol{H}}_n = [\boldsymbol{H}_n \tilde{\boldsymbol{H}}_n]. \tag{8}$$

If we had an infinite-capacity backhaul capable of supporting any transmission rate, we could consider the cooperating BSs as a single receiver with multiple antennas. Considering the simultaneous transmissions of MTs on the same RB, the capacity of the wireless link from $\mathcal{T}(n)$ to the cooperating BSs $\mathcal{C}_n$ is [13]

$$I_n^{(\mathrm{WL})}(\mathcal{C}_n, \mathcal{U}) = B \log_2 \det \left(\boldsymbol{I}_{M v_n} + \frac{1}{\sigma^2} \bar{\boldsymbol{H}}_n \bar{\boldsymbol{H}}_n^H \right)$$
$$- B \log_2 \det \left(\boldsymbol{I}_{M v_n} + \frac{1}{\sigma^2} \tilde{\boldsymbol{H}}_n \tilde{\boldsymbol{H}}_n^H \right), \tag{9}$$

where B is the RB bandwidth, σ^2 is the noise power, and $\boldsymbol{I}_A$ is the A-size identity matrix. Note that $I_n^{(\mathrm{WL})}(\mathcal{C}_n, \mathcal{U})$ depends also on the entire set of active MTs $\mathcal{U}$, due to the simultaneous transmission of other MTs active in the same RB. Note that Equation (9) is obtained in [13] by the definition of capacity as the difference of the entropy of the received signal and the conditional (to the transmitted signals) entropy of the received signal, therefore, is not related to a specific receiver. In order to reach this capacity, nonlinear receivers, e.g., using successive interference cancellation, may be needed.

As we will see in the following, the rate $I_n^{(\mathrm{WL})}(\mathcal{C}_n, \mathcal{U})$ is not achievable in the considered scenario, due to (a) limited backhaul capabilities, and (b) constraints introduced by both the data rate of the packet and the HARQ technique.

We consider the scenario where both BSs and RNC have only a partial CSI before transmission, i.e., they know only the channel statistics, including the correlation matrix $\boldsymbol{\Sigma}_n$. On the other hand, after the reception of packets in each slot, each BS properly estimates the channel with respect to the active MT, and we assume that the estimation is perfect.

2.2 Backhaul network model

Two topologies for the backhaul network are considered: star network (STAR) [14] and ring network (RING) [15,16]. In the STAR scenario, K links connect each BS directly to the RNC, supporting a maximum rate per BS of $\bar{I}^{(\mathrm{BH})}$. In the RING scenario, both BSs and RNC share a common backhaul medium, with a total maximum rate of $K\bar{I}^{(\mathrm{BH})}$, and direct communication is possible among BSs without passing through the RNC. With this choice, the total backhaul rate is the same for the two configurations, for the ease of comparison.

BS k reserves a bit rate $I_{k,n}^{(\mathrm{BH})}$ on the backhaul to forward information on the message received from $\mathcal{T}(n)$ to the BS k. Furthermore, a rate β is reserved for each MT whose information is forwarded, in order to send signaling information, e.g., channel estimation, that may be needed

for scheduling or decoding. Therefore, the reserved backhaul rate for the transmission of information relative to $\mathcal{T}(n)$ from BS k is $\beta + I_{k,n}^{(\text{BH})}$. We assume that transmission on the backhaul is performed using perfect source coding of the information to be transmitted. In the following, we will model the backhaul as a link of limited capacity and we will investigate the maximum rate that can be achieved in the transmission from the MT to the place where decoding is performed. A practical implementation of the transmission on the backhaul is left for future study.

Backhaul topologies yield different constraints on the rates of messages that can be exchanged among BSs. We now derive these constraints taking into account the place where decoding is performed.

2.2.1 STAR topology and decoding at the RNC (STAR-RNC)
In this case decoding is performed at the RNC, where signals from all cooperating BSs must be gathered. When BS k cooperates, it forwards the received signal using only the link from BS k to the RNC. For each link between RNC and BS $k \in \mathcal{K}$ we have the following constraint due to the limited backhaul availability:

$$\sum_{n \in \mathcal{U}} [I_{k,n}^{(\text{BH})} + \beta \delta_{k,n}] \leq \bar{I}^{(\text{BH})}, \tag{10a}$$

where $\delta_{k,n} = 0$ if $I_{k,n}^{(\text{BH})} = 0$, and $\delta_{k,n} = 1$ otherwise.

2.2.2 STAR topology and decoding at the serving BS (STAR-BS)
In this case, decoding is performed at the serving BS. When BS k cooperates with the serving BS b_n, it forwards the received signal using both the link from BS k to the RNC and the link from the RNC to BS b_n. Moreover, once the packet has successfully been decoded at the serving BS, data must be forwarded to the RNC. If the backhaul is full, the serving BS may forward the successfully decoded packet over more slots. Let ρ'_k be the rate of the decoded message that BS has to forward to the RNC. In particular, if decoding of a packet from $\mathcal{T}(n)$ occurs at BS k, ρ'_k is increased by the data rate of the message. The constraint due to the limited backhaul availability is therefore

$$\rho'_k + \sum_{n \in \mathcal{U}, b_n \neq k} [I_{k,n}^{(\text{BH})} + \beta \delta_{k,n}] + \sum_{k' \neq k} \sum_{n \in \mathcal{M}(k)} [I_{k',n}^{(\text{BH})} + \beta \delta_{k',n}] \leq \bar{I}^{(\text{BH})}, \tag{10b}$$

for all RNC-BS links $k \in \mathcal{K}$. Note that if $\rho'_k > \bar{I}^{(\text{BH})}$ the exceeding data are transmitted on the backhaul in the next slots.

2.2.3 RING topology and decoding at the serving BS (RING-BS)
In this case, BSs communicate through a ring operating as a bus. In other words, BSs can communicate directly without passing through the RNC; however, all messages are

exchanged over a single line, the RING, thus we have now a single constraint on the sum of rates of all the messages exchanged over the backhaul. The RNC is also connected to the RING. When BS k cooperates with the serving BS b_n, it forwards the received signal directly to the serving BSs using the backhaul, in a single transmission. Therefore, with respect to the STAR topology in this case we do not need to go through the RNC. Once the packet has successfully been decoded at the serving BS, data are forwarded to the RNC still over the same backhaul. Also in this case, if the backhaul is full, the serving BS may forward the successfully decoded packet over more slots. Therefore, the constraint becomes

$$\sum_{k \in \mathcal{K}} \left\{ \rho'_k + \sum_{n \in \mathcal{U}, b_n \neq k} \left[I_{k,n}^{(\text{BH})} + \beta \delta_{k,n} \right] \right\} \leq K \bar{I}^{(\text{BH})}. \tag{10c}$$

Note that as decoding is performed at the serving BS, data received at that BS do not occupy the backhaul.

2.2.4 RING topology and decoding at the RNC (RING-RNC)
In this case, decoding is performed at the RNC. When BS k cooperates, it forwards the received signal using only the link from BS k to the RNC. Therefore, we have now the single constraint

$$\sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{U}} [I_{k,n}^{(\text{BH})} + \beta \delta_{k,n}] \leq K \bar{I}^{(\text{BH})}, \tag{10d}$$

where the only difference with respect to RING-BS is that now BS transmits the information on the backhaul also for the served MTs.

Note that for more general topologies we can write similar equations as (10), where we will have an equation for each link, including backhaul rate contributions of all BSs connected to that link. For the sake of a simpler explanation we focus here on the two mentioned topologies.

2.3 End-to-end achievable rate
Now we derive the capacity of the end-to-end link from the MT to the RNC that must forward the message to the external network. This capacity does not in general coincide with Equation (9), due to the constraint on the backhaul. Therefore, we indicate with $I_n^{(\text{max})}$ the maximum data rate for $\mathcal{T}(n)$ that can be achieved at the decoding site for given resource allocations on both the wireless link and the backhaul. This rate will be affected by the set of cooperating BSs, the channel conditions, and the constraints on the backhaul communications.

The considered transmission scenario from MTs to the BSs and from the BSs to the place where decoding is performed can be modeled as a primitive multiple-relay

channel, whose capacity is not known [17,18]. Therefore, we resort to the cut-set upper bound for its characterization. For decoding at the RNC, the maximum rate that can be achieved by $\mathcal{T}(n)$ is approximated as the minimum rate achieved for all possible cuts of set $\mathcal{C}_n$, i.e.,

$$I_n^{(\max)} \approx \min_{\mathcal{K}_1 \subseteq \mathcal{C}_n} I_n^{(\text{WL})}(\mathcal{K}_1, \mathcal{U}) + \sum_{k \in \mathcal{C}_n \setminus \mathcal{K}_1} I_{k,n}^{(\text{BH})}. \qquad (11)$$

Note that $I_n^{(\max)}$ depends on both channel conditions and the active MT set.

For decoding at the BS instead, the minimum inside Equation (11) must be taken over all subsets $\mathcal{K}_1 \subseteq \mathcal{C}_n$, such that $b_n \in \mathcal{K}_1$. We remark that the cut-set bound is rather loose and optimistic. However, closed form expressions of the capacity region may be available for special cooperation configurations. For example, Tandon and Poor [19] provide closed form expressions when the serving BS is assisted only by a second BS (i.e., $\mathcal{C}_n$ has only two elements) and the channel of the cooperating BS is degraded with respect to that of the serving BS. Therefore, the techniques proposed in this article are useful also when more precise bounds are available.

We first consider the scenario of a single transmission for the sake of a simpler notation, while in Section 5 we propose its extension to the case of multiple transmissions with HARQ. For a single transmission, $\mathcal{T}(n)$ encodes the message of $T_S \rho_n$ data bits (where T_S is the slot duration) with a capacity achieving code and modulation, and performs a transmission according to the scheduling.

Decoding is successful if

$$I_n^{(\max)} \geq \rho_n, \qquad (12)$$

and in this case a new message is queued for transmission at the next slot.

Due to the transmission rate ρ_n, the achievable rate $I_n^{(\max)}$ is reduced from Equation (11) to the *effective rate* of $\mathcal{T}(n)$ defined as

$$I_n = \min \left\{ \rho_n, I_n^{(\max)} \right\}. \qquad (13)$$

2.4 Cooperative BS assignment
We consider two scenarios for cooperative BS assignment, namely FBA and OBA. In both cases, the active MT set $\mathcal{U}$ is selected before MT transmission, based on the CSI statistics.

2.4.1 FBA
In FBA, the RNC allocates resources on the backhaul before MT transmission based on a partial CSI. A fixed amount of the backhaul rate is reserved to support the maximum rate that can be achieved on air according to the selected data rate and channel conditions. Note that

only a subset of BSs demodulate the signal coming from each MT and forward the information on the backhaul.

2.4.2 OBA
In OBA, the RNC allocates resources on the backhaul after MT transmission, based on the instantaneous channel conditions. In this case, all BSs receive and store the signals on all RBs. Then, backhaul rates are adapted to channel conditions. Although with this approach the use of backhaul is optimized to the effective channel conditions, all BSs must store received signals; therefore, additional memory is needed. Moreover, a delay in the decoding process is introduced, due to (a) performing channel estimates, (b) forwarding estimates to the RNC, and (c) waiting for BS scheduling. Lastly, the backhaul usage is further increased by the need of forwarding CSI of all MTs from the BSs to the RNC. Depending on the duration of the slot and on how many times the OBA is run, this overhead may be non-negligible. In any case it is a fixed backhaul occupation, thus it can be taken into account by properly decreasing $\bar{I}^{(\text{BH})}$ in Equation (10).

3 FBA optimization
With FBA, the choice of cooperating BSs for each MT is done before transmission, using only the CSI statistics. Together with the choice of the cooperating BSs, a fixed amount of the backhaul rate is reserved. This approach reduces information exchanges on CSI among BSs and RNC as only second-order statistics (which may be slowly time-varying) are needed. Lastly, since cooperating BSs immediately start forwarding information on the backhaul upon MT transmission, no further delay is introduced by MCP, except that introduced by the transmissions on the backhaul.

As discussed in Section 1, among the possible objectives of scheduling, we focus on the maximization of the weighted average network throughput, which accommodates various performance objectives, including maximum sum rate, proportionally fair allocation [20], and min–max criteria [12]. For FBA, the scheduling that maximizes the weighted average network throughput is obtained by solving

$$\max_{I_{k,n}^{(\text{BH})}, \mathcal{C}_n, \mathcal{U}} \sum_{n \in \mathcal{U}} \omega_n \mathrm{E}[I_n], \qquad (14)$$

subject to one of the constraints (10), and where ω_n are the weights. Note that problem (14) jointly determines the set of transmitting MTs, the set of cooperating BSs for each MT, and the rates reserved to transfer information coming from each active MT over the backhaul. The choice of the backhaul rate and the active MTs is performed knowing only the average wireless rate instead of

the instantaneous rate. The expectation in Equation (14) is taken with respect to the channel realizations.

In order to solve Equation (14) we propose to select a backhaul rate that does not constitute a bottleneck for the effective rate. By inspecting Equations (11) and (13) we note that—for given choices of both $\mathcal{C}_n$ and $\mathcal{U}$—by setting

$$I_{k,n}^{(\mathrm{BH})} = \rho_n - \sum_{t'=t_n^{(0)}}^{t-1} I_n(t') \quad \forall k \in \mathcal{C}_n, \tag{15}$$

I_n is maximized. Therefore, the backhaul, although with finite capacity, does not constitute a bottleneck and

$$E[I_n] = E\left[\min\left\{\rho_n - \sum_{t'=t_n^{(0)}}^{t-1} I_n(t'), I_n^{(\mathrm{WL})}(\mathcal{C}_n, \mathcal{U})\right\}\right]. \tag{16}$$

3.1 FBA problem revisited

Let $\mathcal{S}$ be the set of all possible cooperating BS sets, i.e., all subsets of $\mathcal{K}$. Let us define the following indicator function for all $\mathcal{C} \in \mathcal{S}$:

$$x_n(\mathcal{C}) = \begin{cases} 1 & \text{BSs with indices in } \mathcal{C} \text{ forward} \\ & \text{information on the backhaul for } \mathcal{T}(n) \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

The FBA problem can be written as

$$\max_{\{x_n(\mathcal{C})\}} \sum_{n \in \mathcal{N}} \sum_{\mathcal{C} \in \mathcal{S}} \omega_n x_n(\mathcal{C}) \mathrm{E}[I_n] \tag{18a}$$

subject to Equations (15), (16),

$$(1) \quad \sum_{\mathcal{C} \in \mathcal{S}} x_n(\mathcal{C}) = 1, \quad n \in \mathcal{N}, \tag{18b}$$

$$(2) \quad x_n(\mathcal{C}) = \{0, 1\}, \quad n \in \mathcal{N}, \mathcal{C} \in \mathcal{S}, \tag{18c}$$

and, for STAR-RNC topology subject to Equations (15), (16),

$$(3) \quad \sum_{n \in \mathcal{N}} \sum_{\mathcal{C} \in \mathcal{S}: k \in \mathcal{C}} [I_{k,n}^{(\mathrm{BH})} + \beta] x_n(\mathcal{C}) \leq \bar{I}^{(\mathrm{BH})}, \quad k \in \mathcal{K}. \tag{18d}$$

For the STAR-BS case the last constraint is replaced by

$$(3) \quad \sum_{n \in \mathcal{N}: b_n \neq k} \sum_{\mathcal{C} \in \mathcal{S}: k \in \mathcal{C}} [I_{k,n}^{(\mathrm{BH})} + \beta] x_n(\mathcal{C}) +$$
$$\sum_{n \in \mathcal{M}(k)} \sum_{\mathcal{C} \in \mathcal{S}} \sum_{k' \in \mathcal{C}, k' \neq k} [I_{k',n}^{(\mathrm{BH})} + \beta] x_n(\mathcal{C}) \leq \bar{I}^{(\mathrm{BH})} - \rho'_k, \quad k \in \mathcal{K}. \tag{19}$$

For the RING-RNC case we have

$$(3) \quad \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}} \sum_{\mathcal{C} \in \mathcal{S}: k \in \mathcal{C}} x_n(\mathcal{C})[I_{k,n}^{(\mathrm{BH})} + \beta] \leq K \bar{I}^{(\mathrm{BH})}. \tag{20}$$

Lastly, for the RING-BS case we have

$$(3) \quad \sum_{k \in \mathcal{K}} \sum_{n \in \mathcal{N}, b_n \neq k} \sum_{\mathcal{C} \in \mathcal{S}: k \in \mathcal{C}} x_n(\mathcal{C})[I_{k,n}^{(\mathrm{BH})} + \beta] \leq K \bar{I}^{(\mathrm{BH})} - \sum_{k \in \mathcal{K}} \rho'_k. \tag{21}$$

The problem can be seen as a MMCK problem [21], where for each MT we extract one and only one BS configuration $\mathcal{C}$ from each class $\mathcal{S}$, in order to maximize the total utility provided by (18a), while satisfying the constraints.

Let $x_n^*(\mathcal{C})$ be the solution of Equation (18). Then the set of active MTs is

$$\mathcal{U} = \{n : \exists \mathcal{C} \neq \emptyset : x_n^*(\mathcal{C}) = 1\}, \tag{22}$$

and the set of cooperating BSs for $\mathcal{T}(n)$ is

$$\mathcal{C}_n = \{\mathcal{C} : x_n^*(\mathcal{C}) = 1\}, \quad n \in \mathcal{U}. \tag{23}$$

3.2 Greedy solution

Since the MMCK problem is a NP-hard problem [21], we consider here an approximated solution obtained using a greedy approach. In particular, we order the objects to be inserted into the knapsack (the cooperative BSs for each MT) according to the *efficiency ratio* between the provided utility and the related cost. For decoding at the RNC we define the efficiency ratio as

$$\lambda_n(\mathcal{C}) \triangleq \frac{\mathrm{E}[I_n]}{\sum_{k \in \mathcal{C}} I_{k,n}^{(\mathrm{BH})}}, \tag{24}$$

while for decoding at the BS the definition is

$$\lambda_n(\mathcal{C}) \frac{\mathrm{E}[I_n]}{\sum_{k \in \mathcal{C}, k \neq b_n} I_{k,n}^{(\mathrm{BH})}}, \tag{25}$$

where $\mathrm{E}[I_n]$ is provided by (16). Note that Equations (24) and (25) are one of the various possible ways of combining utility ($\mathrm{E}[I_n]$) with the cost (backhaul occupation). Other approaches may be possible and would not substantially alter the operation of the greedy algorithm, while results will be depending on the adopted definition of efficiency ratio.

MTs and their associated cooperative BSs are chosen iteratively as described in Algorithm 1. In particular, at each iteration of the algorithm we select the MT with index n^* and the set of cooperating BSs $\mathcal{C}^*$ compatible with both the already selected MT-cooperating BS set pairs and the backhaul constraints. The selection is performed starting from the configurations having the highest efficiency ratio, which provide the most cost-effective solution, i.e., yield the highest gain of the objective function for the same impact on the backhaul occupancy. The

algorithm is greedy, in the sense that all solutions that are not compatible with the configurations selected at the previous iterations are discarded. However, if a configuration with a lower efficiency ratio for $\mathcal{T}(n)$ is compatible with the backhaul constraints induced by the selected configuration for MTs $n' \neq n$, the set of cooperating BSs previously selected for $\mathcal{T}(n)$ can be replaced by the new set as it yields a higher value of the objective function.

In details, at any iteration, let $\{x_n^*(\mathcal{C})\}$ be the current feasible solution of the maximization problem. The set of cooperative BSs $\mathcal{C}$ of $\mathcal{T}(n)$ to be selected by the greedy algorithm must satisfy Equation (18d) or (19) or (21) for

$$x_n(\mathcal{C}) = \begin{cases} x_{n'}^*(\mathcal{C}) & n \neq n', \forall \mathcal{C} \\ x_{n'}^*(\mathcal{C}') & n = n', \forall \mathcal{C}' \neq \mathcal{C}. \end{cases} \quad (26)$$

In fact, if $\mathcal{T}(n)$ has already been associated with a set of BSs, we can replace this association if compatible with the other constraints. Moreover, in selecting the set of cooperating BSs at each iteration we attempt to add (or substitute) solutions having a lower efficiency. In fact, a lower efficiency yields not only a higher occupation of the backhaul, but also a higher rate on the wireless link.

Algorithm 1 Greedy algorithm for FBA optimization
1   $\bar{\lambda} = \infty, x_n(\mathcal{C}) = 0, n \in \mathcal{N}$ and $\mathcal{C} \in \mathcal{S}$;
2   **repeat**
3   **for** $n \leftarrow 1$ **to** N **do**
4         $\mathcal{C}_n^{(\max)} = \operatorname{argmax}_{\mathcal{C} \in \mathcal{S}}\{\lambda_{n(\mathcal{C})} :$
            (18d) or (19) or(21) is satisfied and
            $\lambda_{n(\mathcal{C})} < \bar{\lambda}\}$;
5   **end**
6      $n^* = \operatorname{argmax}_n \lambda_n(\mathcal{C}_n^{(\max)})$;
7   **if** $||\mathcal{C}_{n^*}^{(\max)}|| > 0$ **then**
8         $x_{n^*}(\mathcal{C}_{n^*}^{(\max)}) = 1$;
9         $\bar{\lambda} = \lambda_{n^*}(\mathcal{C}_{n^*}^{(\max)})$;
10   **end**
11 **until** $||\mathcal{C}_{n^*}^{(\max)}|| > 0$;

Note that the computation of the expected effective rate in Equation (16) requires the complete statistical description of the random variable $I_n^{(\mathrm{WL})}(\mathcal{C}_n, \mathcal{U})$, and not just of its average, due to the presence of the minimum in the definition (13) of I_n. Unfortunately, the derivation of the complete statistical description of $I_n^{(\mathrm{WL})}(\mathcal{C}_n, \mathcal{U})$ is a hard problem, which at the moment has not yet been solved in the literature. We recall that in [22] the asymptotic statistics of the achievable rate for spatially correlated Rician fading multiple-input multiple-output (MIMO) channels with interference have been computed. However, all interfering MTs have the same received correlation matrix, i.e., the path loss is the same for all MTs, which is not the scenario considered in this article.

Hence, we resort to the Jensen inequality, and observing that $\min\{x, c\}$ (with c constant) is a concave function of x, we upper bound the expected effective rate as

$$\mathrm{E}[I_n] \leq \min\left\{\rho_n - \sum_{t'=t_n^{(0)}}^{t-1} I_n(t'), \mathrm{E}\left[I_n^{(\mathrm{WL})}(\mathcal{C}_n, \mathcal{U})\right]\right\}. \quad (27)$$

For the computation of the expectation in Equation (27) we first observe that from Equation (9) the expectation can be written as

$$\mathrm{E}[I_n^{(\mathrm{WL})}(\mathcal{C}_n, \mathcal{U})] = B\mathrm{E}\left[\log_2 \det\left(\boldsymbol{I}_{M\nu_n} + \frac{1}{\sigma^2}\bar{\boldsymbol{H}}_n\bar{\boldsymbol{H}}_n^H\right)\right]$$
$$- B\mathrm{E}\left[\log_2 \det\left(\boldsymbol{I}_{M\nu_n} + \frac{1}{\sigma^2}\tilde{\boldsymbol{H}}_n\tilde{\boldsymbol{H}}_n^H\right)\right]. \quad (28)$$

Then for the computation of the two expectations in Equation (28) we resort to the asymptotic results of [23], which provide a good approximation with a finite (even small) number of antennas. Note that in Equation (28) for each expectation the number of transmit antennas is the number of active MTs (or this number minus one, for the second term).

3.3 Local decoding
Zennaro et al. [10] proposed a *local decoding* strategy. In their study each cooperating BS attempts to decode the packet transmitted by a subset of MTs independently from the other BSs, i.e., without exchange of information with cooperating BSs. The first BS that successfully decodes the packet forwards the decoded bits to the RNC. In this case, the constraint is not on the backhaul rate but on the maximum number of MTs, whose messages can be decoded by each BS.

The local decoding strategy fits problem (18) by reinterpreting constraint (18d) as the decoding capability constraint. In particular, let Y be the maximum number of packets that can simultaneously be decoded by each BS. By setting

$$\bar{I}_k^{(\mathrm{BH})} = Y, \quad I_{n,n}^{(\mathrm{BH})} = 1, \quad I_{k,n}^{(\mathrm{BH})} = 0, \quad k \neq n, \quad (29)$$

problem (18) provides the maximization of the weighted sum rate for local decoding.

4 OBA optimization
With OBA, backhaul resources are allocated after MT transmission, based on the instantaneous channel conditions. Hence, the set of active MTs $\mathcal{U}$ must be chosen beforehand, e.g., by performing FBA (and then discarding the allocation of the backhaul resources).

All BSs must estimate the channel and forward the CSI to the RNC or the serving BS, while storing the received signals. However, only a subset $\mathcal{C}_n$ of BSs forwards information on the backhaul for decoding purposes. The effective rate I_n is not anymore a random variable but a deterministic function of both $I_{k,n}^{(BH)}$ and $I_n^{(WL)}(\mathcal{C}_n, \mathcal{U})$.

For OBA we aim at selecting (a) the set of cooperating BSs $\mathcal{C}_n$ for each MT with index $n \in \mathcal{U}$, and (b) the backhaul occupation $I_{k,n}^{(BH)}$ that maximizes the weighted sum rate, i.e.,

$$\max_{\mathcal{C}_n, I_{k,n}^{(BH)}} \sum_{n \in \mathcal{U}} \omega_n I_n. \tag{30}$$

Since in the OBA scenario the channel gains are perfectly known, the optimization problem (30) for decoding at the RNC can be written as

$$\max_{I_n, I_{k,n}^{(BH)}} \sum_{n \in \mathcal{U}} \omega_n I_n \tag{31a}$$

subject to

$$I_n \leq \rho_n, \quad n \in \mathcal{U} \tag{31b}$$

$$I_n \leq I_n^{(WL)}(\mathcal{K}_1, \mathcal{U}) + \sum_{k \in \mathcal{K} \setminus \mathcal{K}_1} I_{k,n}^{(BH)}, \quad n \in \mathcal{U}, \quad \forall \mathcal{K}_1 \subseteq \mathcal{K} \tag{31c}$$

and one of the constraints (10). Constraint (31b) ensures that the achieved rate is not larger than the packet rate. Constraint (31c) is the cut-set upper bound constraint.

The set of BSs that forward the information on the backhaul for $\mathcal{T}(n)$ is $\mathcal{C}_n = \{k : I_{k,n}^{(BH)} > 0\}$. When decoding is performed at the BS, constraint (31c) must hold only for $n \in \mathcal{U}, \forall \mathcal{K}_1 \subseteq \mathcal{K}$ such that $b_n \in \mathcal{K}_1$.

We observe that, for a given set of values $\delta_{k,n}$, (31) is a linear programming problem that can be solved by standard optimization tools. However, since $\delta_{k,n}$ are unknown, we obtain a mixed integer programming problem, that we solve again with a greedy approach. First, we solve the linear programming problem assuming $\delta_{k,n} = 0$. The resulting values of backhaul occupation are denoted $\hat{I}_{k,n}^{(BH)}$. An iterative process is then started, and at each iteration we consider a feasible solution, denoted $\delta_{k,n}^*$.

We start with $\delta_{k,n}^*$ as provided by the FBA solution. Then we consider the couple (k, n) that maximizes the backhaul occupation among the links not yet inserted in the set identified by $\delta_{k,n}^*$, i.e.,

$$(\hat{k}, \hat{n}) = \text{argmax}_{(k,n)} \hat{I}_{k,n}^{(BH)} (1 - \delta_{k,n}^*). \tag{32}$$

We then solve the linear programming problem, where $\delta_{\hat{k},\hat{n}}^* = 1$. If the solution yields a higher objective function than that of the previous iteration, we proceed with the next iteration. Otherwise, $\delta_{\hat{k},\hat{n}}^* = 0$, the solution of the

previous iteration is taken as final and the iterative process is terminated. Note that in this process at most $K||\mathcal{U}||$ iterations are performed.

5 Use of HARQ

In order to cope with fading, here we consider the extension of the model to HARQ. Now all variables depend on the time slot t at which transmission occurs.

For the HARQ retransmission process, let $t_n^{(0)}$ be the slot at which the first transmission from $\mathcal{T}(n)$ occurs. Decoding is successful at slot $t_n^{(0)}$ if (see Equation 12)

$$I_n^{(max)}(t_n^{(0)}) \geq \rho_n, \tag{33}$$

and in this case a new message is queued for transmission at the next slot. If decoding fails, the node where decoding occurs (i.e., either the serving BS or the RNC) stores the information received by cooperating BSs and at the next slot a block of redundancy bits is transmitted by the MT. The process is stopped at slot τ_n, either when the maximum number of retransmissions τ_{max} is exceeded, i.e., when

$$\tau_n = t_n^{(0)} + \tau_{max} - 1, \tag{34}$$

or when the rate of the retransmissions exceeds ρ_n for the first time, i.e.,

$$\sum_{t=t_n^{(0)}}^{\tau_n - 1} I_n^{(max)} < \rho_n, \text{ and } \sum_{t=t_n^{(0)}}^{\tau_n} I_n^{(max)} \geq \rho_n, \tag{35}$$

in which case decoding is successful.

Due to both HARQ and the rate ρ_n of the packets transmitted by $\mathcal{T}(n)$, the achievable rate $I_n^{(max)}$ is further reduced from Equation (13) to the effective rate of $\mathcal{T}(n)$, which now becomes

$$I_n = \min \left\{ \rho_n - \sum_{t'=t_n^{(0)}}^{t-1} I_n(t'), I_n^{(max)} \right\}. \tag{36}$$

Moreover, Equation (31b) becomes

$$I_n \leq \rho_n - \sum_{t'=t_n^{(0)}}^{t-1} I_n(t'), \quad n \in \mathcal{U}. \tag{37}$$

On the choice of ρ_n:

The value of ρ_n must be chosen according to the desired HARQ performance and can be adapted to both channel and traffic conditions. In particular, ρ_n can be chosen in order to ensure an average packet decoding delay of $\tau_0 < \tau_{max}$ slots. The adaptation process works as follows:

- if early decoding occurs, i.e., the packet is decoded before τ_0 slots, then ρ_n for the next data packet is increased;

- if decoding either occurs after τ_0 or fails after τ_{max} slots, ρ_n must be reduced for the next transmission.
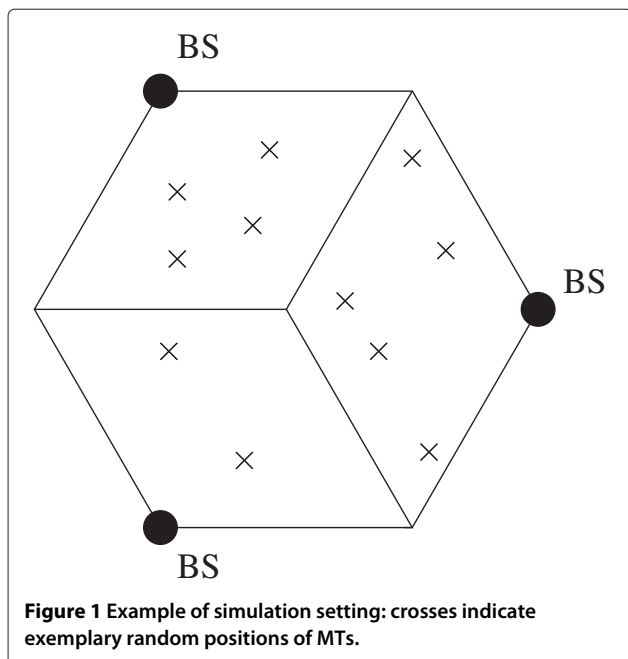
ρ_n is in general taken from a discrete set of rates associated to combinations of constellation sizes and code rates.

6 Numerical results

In order to evaluate the performance of the proposed resource allocation techniques we consider a cellular scenario with $K = 3$ BSs covering an hexagonal area as shown in Figure 1. No further out of cell interference is considered. We assume that MTs are randomly dropped into the area covered by the BSs.

The noise level, the gain at unitary distance $G(1)$, and the size of the cell are chosen in order to ensure an average signal-to-noise ratio (SNR) of 10 dB at the edge of the cell. Each BS is equipped with $M = 2$ antennas and channels with respect to all the antennas are assumed uncorrelated, Rayleigh fading as from Equations (4) and (5). The total bandwidth of 20 MHz is split among 12 RBs thus yielding a band of $B = 1.7$ MHz per RB. We assume that MTs served by the same BS transmit on different RBs, while many MTs served by different BSs may be transmitting on the same RB. The maximum decoding delay is set at $\tau_{max} = 6$ slots. For both FBA and OBA, we consider the same backhaul constraint, thus ignoring the overhead due to the extra signaling required by OBA to transmit CSI to the RNC at each slot. MTs are assigned at random to the BSs. Selection of active MTs for OBA is performed by first applying FBA.

Maximization of the network throughput is considered (i.e., $\omega_n = 1$). We set $\rho_n/B = \bar{\rho} = 15$ bit/s/Hz, equal for all MTs. Note that this is the maximum rate achievable if the first transmission is successful. The average rate will be reduced by retransmissions due to channel fading.
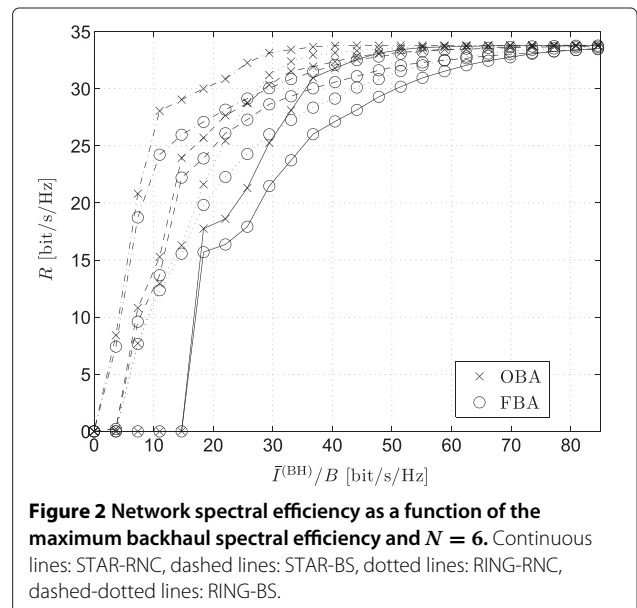
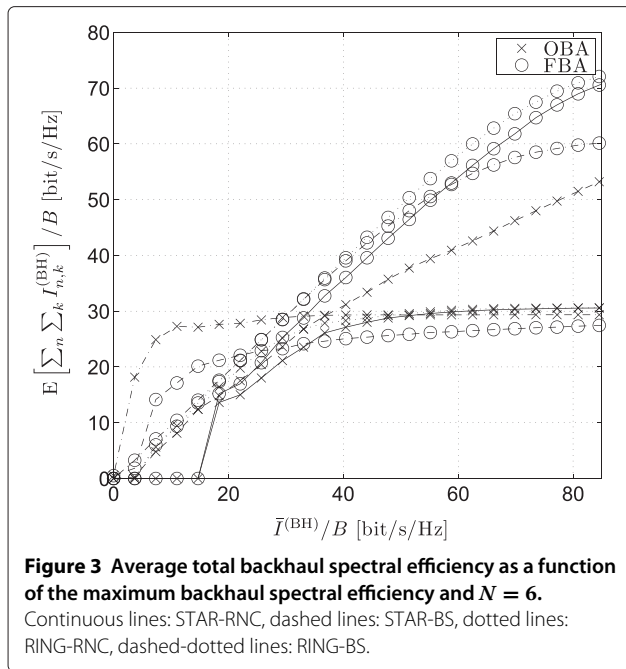We first consider the network spectral efficiency, i.e.,

$$R = \mathrm{E}\left[\frac{Q}{T \cdot B}\right], \tag{38}$$

where Q is the total number of data bits successfully decoded and forwarded to the RNC, over the observation time T. Figure 2 shows R as a function of maximum backhaul spectral efficiency $\bar{I}^{(BH)}/B$ for $N = 6$ MTs in the system. Correspondingly, Figure 3 shows the average total backhaul spectral efficiency, i.e.,

$$\frac{1}{B}\mathrm{E}\left[\sum_{n \in \mathcal{U}}\sum_{k \in \mathcal{K}} I_{k,n}^{(BH)}\right], \tag{39}$$

as a function of the maximum backhaul spectral efficiency and $N = 6$. From Figure 2 we observe that for high values of $\bar{I}^{(BH)}$, i.e., when the backhaul does not represent a bottleneck, all techniques achieve the same performance, except for the fact that when decoding is performed at the BS, an additional slot is needed to forward data to the RNC, thus reducing the spectral efficiency. Two general observations are in place: (a) OBA is more flexible in scheduling transmissions on the backhaul than FBA, and (b) the RING topology is more flexible than the STAR topology, where the share of the total backhaul rate is pre-assigned to each BS. Note also that for high $\bar{I}^{(BH)}$ decoding at RNC outperforms decoding at the BS, since the latter approach requires additional overhead to forward the decoded packet to the RNC. However, for limited backhaul rates, decoding at the BS has the advantage of not requiring continuous transmissions on the backhaul.

Figure 1 Example of simulation setting: crosses indicate exemplary random positions of MTs.

Figure 2 Network spectral efficiency as a function of the maximum backhaul spectral efficiency and $N = 6$. Continuous lines: STAR-RNC, dashed lines: STAR-BS, dotted lines: RING-RNC, dashed-dotted lines: RING-BS.

Figure 3 Average total backhaul spectral efficiency as a function of the maximum backhaul spectral efficiency and $N = 6$. Continuous lines: STAR-RNC, dashed lines: STAR-BS, dotted lines: RING-RNC, dashed-dotted lines: RING-BS.

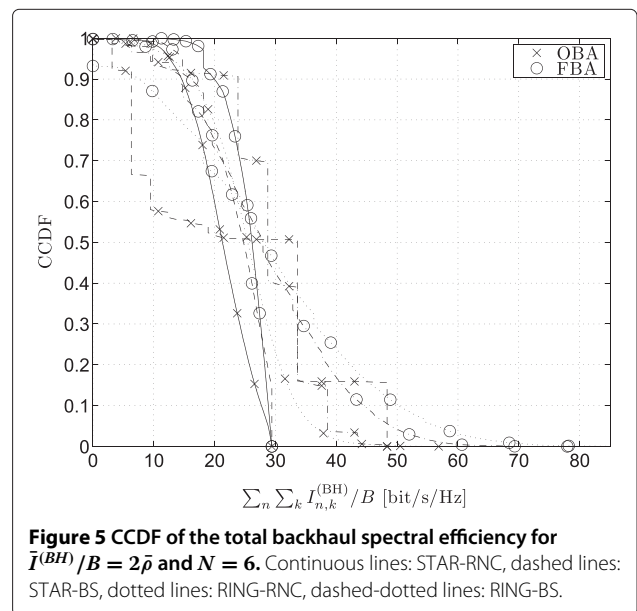**Figure 4 CCDF of the network spectral efficiency.**

The worst performing scheme is FBA with STAR topology and decoding at the RNC. A comparison between FBA and OBA clearly shows the better performance of OBA, which however comes at the cost of both an increased activity of the BS (number of decoded packets) and an additional exchange of control information for real-time scheduling of decoding at the serving BS. We observe that at $\bar{I}^{(\mathrm{BH})}/B = 4\bar{\rho}$ bit/s/Hz FBA with decoding at the BS is very close to the maximum network spectral efficiency, while the same performance is obtained by OBA with decoding at the BS for just $\bar{I}^{(\mathrm{BH})}/B = 2\bar{\rho}$ bit/s/Hz. For STAR topology, the increase of the network spectral efficiency with respect to the backhaul spectral efficiency is much reduced when compared to other schemes, due to the additional overhead in forwarding decoded packets to the RNC. Note that when decoding is performed at the RNC and STAR topology is used, for low $\bar{I}^{(\mathrm{BH})}$ no communication is possible. In fact, with FBA we reserve the backhaul according to Equation (15), and therefore no MT is allowed to transmit as long as $\bar{I}^{(\mathrm{BH})}/B < \bar{\rho}$. Then, since OBA is performed after FBA, also for OBA no communication is possible for low values of the backhaul constraint.
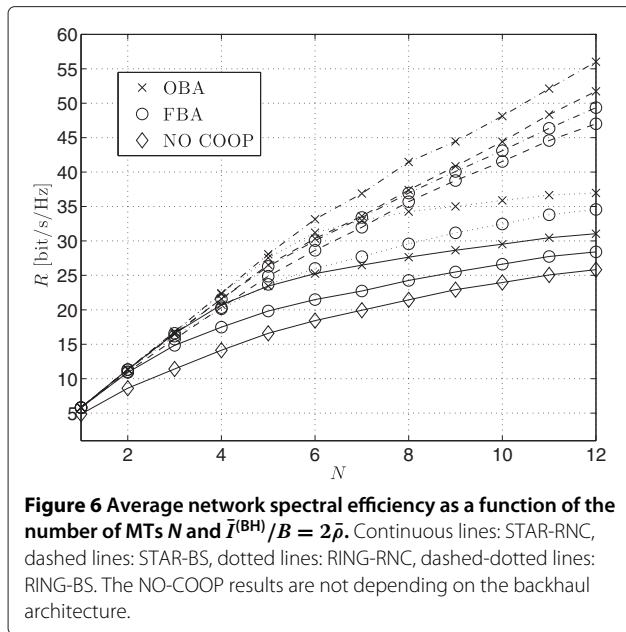
In order to better understand the behavior of the various approaches, Figure 4 shows the complementary cumulative distribution function (CCDF) of I_n/B when $\bar{I}^{(\mathrm{BH})}/B = 2\bar{\rho}$ or $\bar{I}^{(\mathrm{BH})}/B = 3\bar{\rho}$, decoding is performed at the BS, and STAR topology is used. We observe that OBA outperforms FBA in both cases since the cooperating BSs are chosen after the MT transmission, thus providing the best channel among all possible BSs, including the
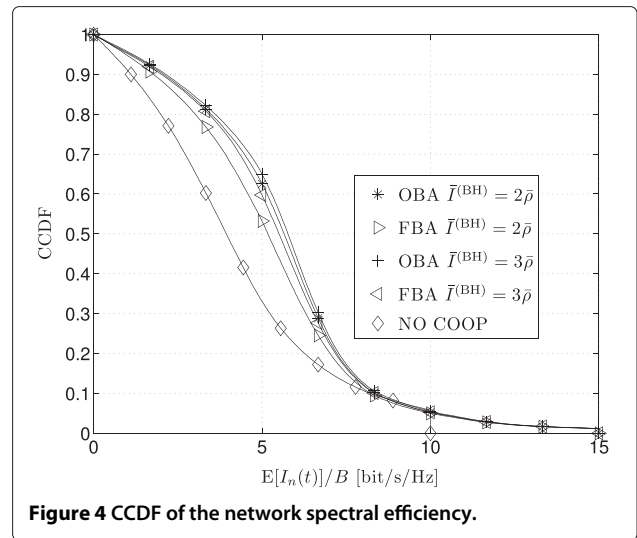
effects of fading. We also report the CCDF of I_n/B for a non-cooperative system, i.e., where the signal coming from each MT is decoded at the serving BS without the assistance of the other BSs. We observe that the cooperative systems always outperform the non-cooperative system. Note also that for the non-cooperative system we set $\rho_n/B = \bar{\rho} = 10$ bit/s/Hz, in order to maximize its achievable rate considering the maximum number of retransmissions available for HARQ.

Figure 5 shows the CCDF of the total backhaul spectral efficiency for $\bar{I}^{(\mathrm{BH})}/B = 2\bar{\rho}$ and $N = 6$. Note that while the maximum value of total backhaul occupancy is limited to $2\bar{\rho}$ for the STAR topology, for RING topology the limit is $6\bar{\rho}$ as from Equations (10c) and (10d). For
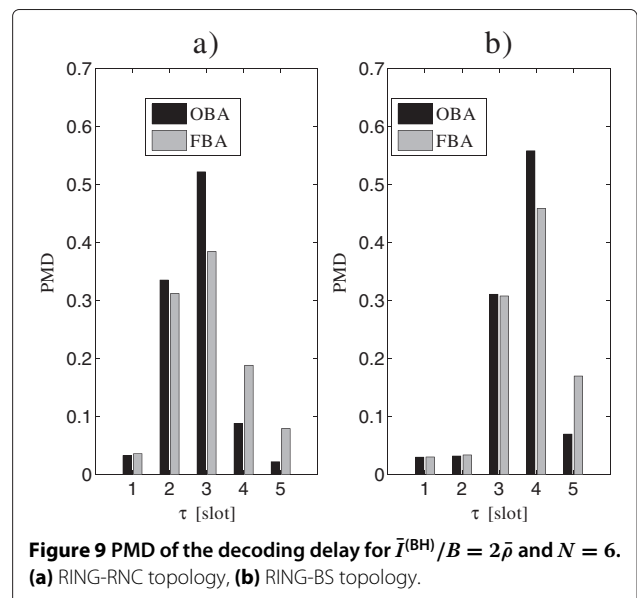

Figure 5 CCDF of the total backhaul spectral efficiency for $\bar{I}^{(BH)}/B = 2\bar{\rho}$ and $N = 6$. Continuous lines: STAR-RNC, dashed lines: STAR-BS, dotted lines: RING-RNC, dashed-dotted lines: RING-BS.

Figure 6 Average network spectral efficiency as a function of the number of MTs *N* and $\bar{I}^{(BH)}/B = 2\bar{\rho}$. Continuous lines: STAR-RNC, dashed lines: STAR-BS, dotted lines: RING-RNC, dashed-dotted lines: RING-BS. The NO-COOP results are not depending on the backhaul architecture.

Figure 8 PMD of the decoding delay for $\bar{I}^{(BH)}/B = 2\bar{\rho}$ and $N = 6$. **(a)** STAR-RNC topology, **(b)** STAR-BS topology.

FBA with RING topology, we observe that when decoding is performed at the RNC the sum backhaul spectral efficiency shows a higher variance than that of the system with decoding at the BS, having a more efficient exploitation of the statistics of the channel, which are reflected into a higher network spectral efficiency (see Figure 2). We observe that OBA exploits at best the backhaul, showing a very limited saturation probability and a reduced variance.

Figure 6 shows the average network spectral efficiency as a function of the number of MTs *N* and $\bar{I}^{(BH)}/B =$
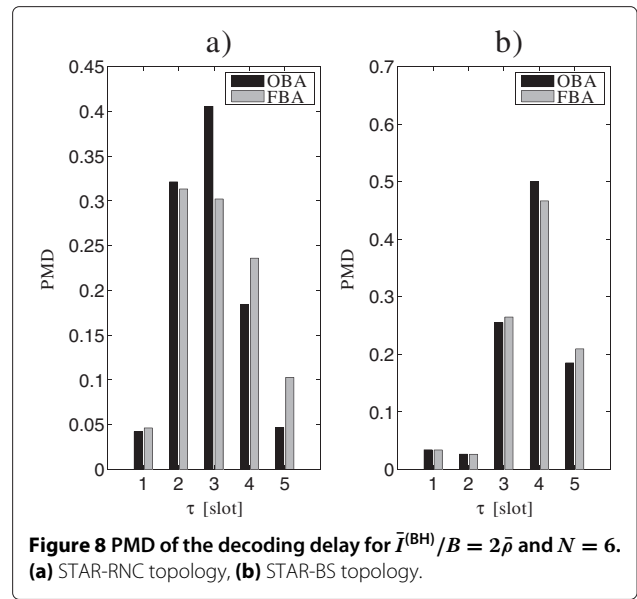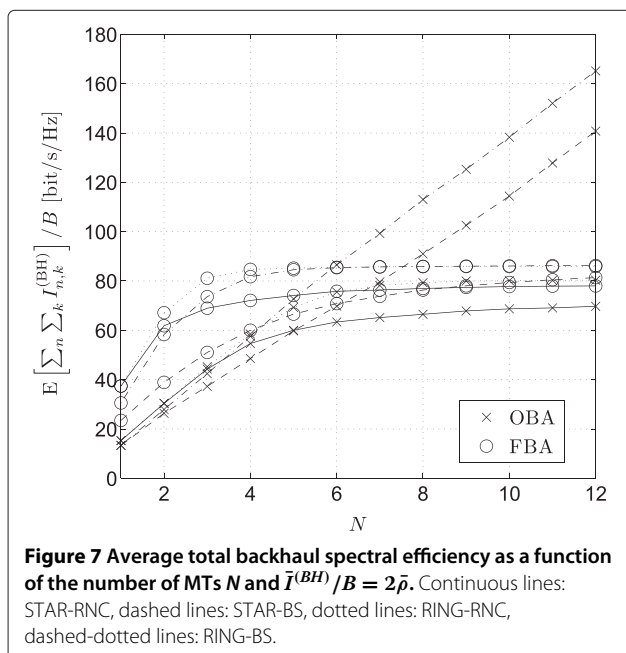
$2\bar{\rho}$ bit/s/Hz. We observe that OBA significantly outperforms FBA for all topologies except the STAR topology with decoding at the BS, for the reasons already explained. Indeed, OBA has an almost linear increase of the network spectral efficiency, thus being able of fully exploiting the multiuser diversity. On the other hand, the network spectral efficiency of FBA saturates as the number of MTs increases due to backhaul constraint. We also report the average network spectral efficiency of the non-cooperative system. We observe that the cooperative systems provide a significant gain with respect to the case of no cooperation.

When the total backhaul spectral efficiency is considered, as a function of the number of MTs, we observe

Figure 7 Average total backhaul spectral efficiency as a function of the number of MTs *N* and $\bar{I}^{(BH)}/B = 2\bar{\rho}$. Continuous lines: STAR-RNC, dashed lines: STAR-BS, dotted lines: RING-RNC, dashed-dotted lines: RING-BS.

Figure 9 PMD of the decoding delay for $\bar{I}^{(BH)}/B = 2\bar{\rho}$ and $N = 6$. **(a)** RING-RNC topology, **(b)** RING-BS topology.

from Figure 7 that for OBA the occupation of the backhaul grows almost linearly with the number of MTs, as it schedules a priori with respect to MT transmissions, while for FBA the constraints have impact on the performance, determining a saturation of the curve. The STAR topology with decoding at the BS shows a saturation of the backhaul occupancy already for a small number of MTs in the system.

In order to assess the performance of the proposed resource allocation techniques for HARQ transmissions, we have considered the decoding delay. Figures 8 and 9 show the probability mass distribution (PMD) of the decoding delay for various configurations and $N = 6$ MTs. We observe that OBA with decoding at RNC and RING topology decodes within four slots with probability close to one, while all other approaches require five slots with significant probability.

7 Conclusions

We have considered a network MIMO system where BSs exchange information for jointly decoding signals coming from MTs, using also HARQ for error protection. The scheduling of transmissions on both the wireless links and the backhaul network is based on either a partial or a full channel knowledge. From numerical results various conclusions can be driven.

In the comparison between FBA and OBA, we conclude that the perfect channel knowledge exploited in FBA allows for a better usage of the backhaul, with a reduction to a half or one-third of the occupancy required by OBA. Moreover, the sum rate achieved by the system using FBA is larger by at least 10% with respect to OBA, for the same constraint on the backhaul. These benefits come at the expense of backhaul scheduling after each packet reception, in order to keep track of channel variations. As a consequence, delay in the decoding may be experienced, and potentially larger memories are needed at the BSs.

A second conclusion pertains both the backhaul topologies and the place where decoding occurs. In this respect, the most flexible architecture is given by the RING topology with decoding at the BS, where the backhaul throughput can be allocated to the BSs in any proportion, while at the same time the transfer of information from the serving BS to the RNC is avoided. On the other hand, the RING topology requires that all BSs are connected on a single bus. The STAR topology (where there is only a direct connection between the RNC and each BS, but not among BSs directly) suffers from a rigidity in allocating the total available backhaul rate to the BSs' needs.

For both FBA and OBA problems, we have proposed greedy algorithms, that allow to reduce the complexity with respect to exhaustive search algorithms. Overall, we can conclude that network MIMO exploiting a RING topology provides a significant performance advantage

over existing non-cooperative decoding in cellular system, both when partial channel knowledge is available (and OBA is used) and when perfect channel knowledge is exploited (and FBA is used).

References
1. S Venkatesan, in *Proc. IEEE International Symposium on Personal Indoor and Mobile Radio Communication (PIMRC 2007)*. Coordinating base stations for greater uplink spectral efficiency: proportional fair user rates. September 2007
2. A Papadogiannis, D Gesbert, E Hardouin, in *Proceedings of the International Conference on Communication (ICC)*. A dynamic clustering approach in wireless networks with multi-cell cooperative processing. June 2008, pp. 4033–4037
3. M Kamoun, L Mazet, in *Proceedings of the European Wireless Conference (EWC)*. Base-station selection in cooperative single frequency cellular network. April 2007
4. P Cheng, F Zhou, C Xu, J Wang, G Li, X Zhang, in *Proceedings of the International Conference on Wireless Communication Networking and Mobile Computing (WiCOM)*. Performance evaluation of uplink base station cooperation in multi-cell MIMO systems. September 2010, pp. 1–5
5. A del Coso, S Simoens, in *Proceedings of the IEEE Workshop on Signal Processing Advances in Wireless Communication (SPAWC)*. Distributed compression for the uplink channel of a coordinated cellular network with a backhaul constraint. January 2008, pp. 301–305
6. P Marsch, G Fettweis, in *Proceedings of the IEEE International Conference on Communication (ICC)*. A framework for optimizing the uplink performance of distributed antenna systems under a constrained backhaul. June 2007, pp. 975–979
7. F Diehm, P Marsch, G Fettweis, B Ramamurthi, in *Proceedings of the IEEE Global Communication Conference (GLOBECOM)*. A low-complexity algorithm for uplink scheduling in cooperative cellular networks with a capacity-constrained backhaul infrastructure. November 2009
8. M Grieger, P Marsh, G Fettweis, in *Proceedings of the IEEE International Conference on Communication (ICC)*. Ad hoc cooperation for the cellular uplink with capacity constrained backhaul. June 2010
9. A Sanderovich, O Somekh, HV Poor, S Shamai (Shitz), Uplink macro diversity of limited backhaul cellular network. IEEE Trans. Inf. Theory. **55**(8), 3457–3478 (2009)
10. S D Zennaro, L Tomasin, Vangelista, Base station selection in uplink macro diversity cellular systems with hybrid ARQ. IEEE J. Sel. Areas Commun. **29**(6), 1249–1259 (2011)
11. J Ramis, G Femenias, Cross-layer design of adaptive multirate wireless networks using truncated HARQ. IEEE Trans. Veh. Technol. **60**(3), 944–954 (2011)
12. P Svedman, SK Wilson, LJ Cimini Jr., B Ottersten, Opportunistic beamforming and scheduling for OFDMA systems. IEEE Trans. Commun. **55**(5), 941–952 (2007)
13. M Chiani, MZ Win, S Hyundong, in *Proceedings of the IEEE Global Communication Conference (GLOBECOM)*. Capacity of MIMO systems in the presence of interference. December 2006, pp. 1–6
14. F Boccardi, H Huang, A Alexiou, in *Proceedings of the Asilomar Conference on Signals, Systems and Computers*. Network MIMO with reduced backhaul requirements by MAC coordination. October 2008, pp. 1125–1129
15. C Charnsripinyo, N Wattanapongsakorn, in *Proceedings of the EEE Region 10 TENCON Conference*, vol. B. A model for reliable wireless access network topology design. 21–24 November, 2004, pp. 561–564
16. Next Generation Mobile Network, white paper, LTE backhauling deployment scenarios. http://www.ngmn.org/uploads/media/NGMN_Whitepaper_LTE_Backhauling_Deployment_Scenarios.pdf
17. Y-H Kim, in *Proceedings of the Allerton Conference*. Coding techniques for primitive relay channels. September 2010–17

18. R Tandon, S Ulukus, in *Proceedings of the Allerton Conference*. Capacity of a class of semi-deterministic primitive relay channels. September 2010
19. R Tandon, HV Poor, in *Proceedings of the Annual Conference on Information Sciences and Systems (CISS)*. On the capacity region of multiple-access relay channels (MD, Baltimore, 2011), pp. 1–5
20. A Jalai, R Padovani, R Pankaj, in *Proceedings of the Vehicular Technology Conference (VTC)*. Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system. May 2000
21. H Kellerer, U Pferschy, D Pisinger, *Knapsack Problems*. (Springer, New York, 2004)
22. E Riegler, G Taricco, Asymptotic statistics of the mutual information for spatially correlated Rician fading MIMO channels with interference. IEEE Trans. Inf. Theory. **56**(4), 1542–1559 (2010)
23. D Aktas, M Naeem Bacha, JS Evans, SV Hanly, Scaling results on the sum capacity of cellular networks with MIMO links. IEEE Trans. Inf. Theory. **52**(7), 3264–3274 (2006)