NETWORK MIMO FOR DOWNLINK IN-BAND RELAY TRANSMISSIONS WITH RELAYING PHASES OF FIXED DURATION

Adrián Agustin, Josep Vidal, Sandra Lagen, Eduard Valera

Dept. Signal Theory and Comunications, Universitat Politècnica de Catalunya (UPC)
C/ Jordi Girona 1-3, 08034, Barcelona, Spain
email: {adrian.agustin, josep.vidal}@upc.edu

ABSTRACT

A half-duplex relay station (RS)-based cellular system deployment is considered, where multiple base stations (BS) cooperate in the BS-RS in-band transmission for the downlink. The duration of the relay-receive and the relay-transmit phases are fixed beforehand, so that the interference induced by other cells is stationary during a transmission interval. With the optimization of the precoders and powers allocated to the wireless backhaul (relay-receive phase) and to the RS-MS access (relay-transmit phase), it is possible to exploit the benefits of network-MIMO (N-MIMO) along with combating the pathloss and shadowing effects thanks to the RSs. Consequently, an appealing significant enhanced spectral efficiency, power efficiency and coverage homogeneity are obtained. Results are benchmarked to the case where the relay phases are optimized.

Index terms- *Network-MIMO, Relay transmissions, QoS*

1. INTRODUCTION

With the advent of new sophisticated terminals and bandwidth-demanding services, system designers are pushed towards the challenge of enhancing system spectral efficiency and providing homogeneous coverage for wireless networks. Next generation standards are already considering that conventional paradigms need to be rethought. In this respect, mature enabling technologies (like MIMO) are considered an integral part of the system, while other (like RS-based deployments and coordinated BS transmissions, or Network-MIMO) are part of ambitious study items. Leveraging on the advantages offered by the joint use of all these techniques is a challenge faced by IEEE 802.16m [1] and LTE-A.

While implementation details of full-duplex RS are under investigation, relay-based enhancements in standards consider half-duplex relay operation, which incur a rate penalty as they require at least two timeslots to relay a message from source to destination [2][3]. It is therefore crucial to enhance the capacity of the in-band wireless backhaul between source and relay (in our case, the BS–RS link) to increase the information rate. One of the solutions usually assumed is that RSs are placed in specifically planned positions above roof-top or in lampposts, ensuring line-of-sight (LOS) conditions in the BS–RS link, and hence reducing the pathloss and shadowing effects. However, the price to pay is twofold: the likely LOS propagating conditions also to other-cell BSs (which will inject harmful interference) and the rank deficiency of the spatial channel when both BS and RS are equipped with multiple antennas. Both effects are detrimental to MIMO channel gains [4][5].

In this respect, N-MIMO seems especially suited to address in-band backhauling in relay transmissions for the downlink [6]. While coordination may be seen as an efficient way to combat the interference from neighbor cells, it also creates a virtual MIMO broadcast channel whose number of degrees of freedom is boosted (if compared to a conventional single-user MIMO under TDMA) and is hardly affected by the rank deficiency of single-user MIMO channels in LOS.

It has been observed that N-MIMO based on zero-forcing (BD–ZF) performs closely to dirty-paper coding [6] but, although its simplicity, it requires accurate channel knowledge from all involved links. However, N-MIMO is again appropriate for our problem thanks to the long channel coherence time of BS–RS links.

An additional way to improve the efficiency of relay transmissions is by optimizing the duration of the relay-receive and the relay-transmit phases [2][3]. In [13], we observed that the joint optimization of coordinated BS-RS links (through N-MIMO) and transmit duration phases brings large benefits. This approach is however not convenient when considering multiple coordinated cells: if each group of coordinated cells adapts the duration of the transmission independently, the interference power observed in each transmission slot may be time-varying, a harsh and undesirable situation for the cellular system.

The evaluation of the N-MIMO with in-band relaying for fixed duration of relay phases motivates this study. As compared to [15], the problem formulation is a particular case whose solution has a significantly reduced complexity: the number of variables to be optimized turns out to be independent on the number of transmission modes, thus making it easily amenable to multicarrier systems.

2. SYSTEM ASSUMPTIONS

Our system definition is based on the following practical assumptions:

1. The number of antennas at BS, RS and MS is n_B, n_R and n_M respectively, so MIMO performance gains are captured.
2. Perfect channel state information at the transmitter side (CSIT) is available at the BSs, so per-mode power loading is possible at the BS-RS link.
3. RSs are time half-duplexed terminals operating under Decode-and-Forward (DF).That is a suitable coding approach for BS-RS links, where high SNR is expected if LOS propagation is met [7].
4. Mobile stations (MSs) do not process the signals transmitted by the BS, only those transmitted by the RS. In other words simple forwarding relaying is assumed.
5. RS transmissions are not coordinated in the way BS transmissions are. Their transmissions are either interfered (if multiple RSs transmit simultaneously) or orthogonalized (if allocating one time slot per RS transmission). In the later case,the number of time slots is denoted by F.

This work was supported in part by the EC through FREEDOM FP7 ICT-2007-4-248891, by FEDER funds and the Spanish Science and Technology Commission through projects: TEC2010-19171/TCM, TEC2006-06481/TCM, and CONSOLIDER CSD2008-00010 COMONSENS; also by project 2009SGR1236 (AGAUR) of the Catalan Administration.

6. Each RS transmits to a single associated MS and therefore it is considered as a single user (possibly interfered) MIMO link. Full CSIT may be exploited at the RS if sufficient feedback rate from the MS is allowed. Otherwise, only average CSIT is assumed. Even if coding is done across multiple states channel (as in a multicarrier case) and interference is white, the maximum rate is given by the MIMO ergodic capacity, for which exact expressions are known [8].
7. It is assumed that the duration of slots for the BS-RS link and for the RS-MS link are not optimized.

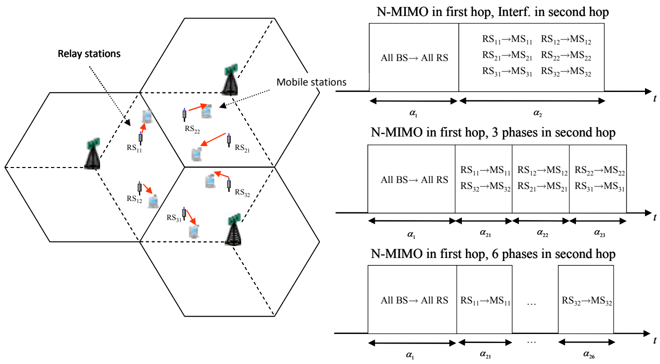

Figure 1 - Access modes considered for N-MIMO over half-duplex relay transmissions in the downlink (B=3, R=6, F=1, 3 and 6).

3. NETWORK-MIMO IN RELAY TRANSMISSION

3.1. Signal model on the first hop

We shall adopt a downlink transmission setup where B BSs coordinate their transmissions and are assisted by R RSs to transmit messages to R MSs. Each MS is associated to a single RS. All BSs transmit on a fixed fraction of time α_1 on the first hop to the RSs following a N-MIMO strategy based on BD-ZF [6][9] (Figure 1), which is appropriate for BS-RS links in LOS conditions (MMSE precoding provides improved performance only at low SNR [10]).

The signal transmitted by all $n_B \cdot B$ antennas is given by

$$\mathbf{x} = \sum_{i=1}^{R} \mathbf{Q}_i \mathbf{b}_i \in \mathbb{C}^{n_B \cdot B \times 1}$$

where $\mathbf{b}_i$ is the symbol stream with m_i components associated to the i-th RS and $\mathbf{Q}_i$ is its associated precoding matrix. We adopt a conventional BD-ZF precoding [6] defined by three matrices,

$$\mathbf{Q}_i = \mathbf{V}_i \mathbf{W}_i \mathbf{P}_i^{\frac{1}{2}} \tag{1}$$

where $\mathbf{P}_i$ is a diagonal matrix describing the power allocated per symbol stream $\mathbf{b}_i$, while $\mathbf{V}_i$ is the $B \cdot n_B \times (B \cdot n_B - (R-1) \cdot n_R)$ BD-ZF precoding matrix. By virtue of the ZF precoding, the signal received by the i-th RS is affected by the $n_R \times B \cdot n_B$ channel matrix $\mathbf{H}_i$ (containing the channel gains between the transmitting antennas at the B BSs and its receiving antennas):

$$\mathbf{y}_i = \mathbf{H}_i \left(\mathbf{V}_i \mathbf{W}_i \mathbf{P}_i^{\frac{1}{2}} \mathbf{b}_i + \sum_{j=1, i \neq j}^{R} \mathbf{V}_j \mathbf{W}_j \mathbf{P}_j^{\frac{1}{2}} \mathbf{b}_j \right) + \mathbf{n}_i \in \mathbb{C}^{n_R \times 1} \tag{2}$$

where m_i denotes the number of symbol streams associated to the i-th RS. Regarding matrix $\mathbf{W}_i$, if we decide to maximize the transmission rate, its optimal design has been recently derived in [15] when individual power constraints per BS are considered. However, the improvement over SVD-based precoding is modest at the expenses of increasing the computational complexity. Consequently, we define $\mathbf{W}_i$ as the matrix containing the m_i right singular vectors of $\mathbf{H}_i \mathbf{V}_i$ associated to the largest singular vectors.

The BD-ZF precoder design requires $\mathbf{V}_i \in \text{kernel}(\tilde{\mathbf{H}}_i)$, where:

$$\tilde{\mathbf{H}}_i = \begin{bmatrix} \mathbf{H}_1^T & \cdots & \mathbf{H}_{i-1}^T & \mathbf{H}_{i+1}^T & \cdots & \mathbf{H}_R^T \end{bmatrix}^T \tag{3}$$

The existence of the kernel requires $B \cdot n_B > (R-1) \cdot n_R$, and hence:

$$\text{rank}\left(\mathbf{H}_i \mathbf{V}_i\right) \leq \min\left(n_R, B \cdot n_B - (R-1) n_R\right)$$

Additionally, symbol decidability at the receivers requires

$$\begin{aligned} &\sum_{i=1}^{R} m_i \leq B \cdot n_B \\ &m_i \leq \text{rank}\left(\mathbf{H}_i \mathbf{V}_i\right) \qquad i = 1, \ldots, R \end{aligned} \tag{4}$$

It must be remarked that in the eventual case the i-th RS observes all coordinated BSs in LOS (hence BS-RS link channels are rank deficient) the rank of $\mathbf{H}_i$ grows up to full-rows rank if channels to the B BSs are linearly independent.

Once $\mathbf{W}_i$ have been selected, the achievable rate for messages intended to the i-th RS becomes,

$$r_{1i} = \log_2 \left| \mathbf{I} + \mathbf{S}_i \mathbf{P} \mathbf{S}_i^H \right| = \sum_{j=1}^{m_i} \log_2 \left(1 + s_{ij}^2 p_{ij} \right) \tag{5}$$

where $\mathbf{S}_i$=diag($s_{i1}, \ldots, s_{im_i}$) contains the singular values of $\mathbf{N}_i^{-1/2} \mathbf{H}_i \mathbf{V}_i$ (being $\mathbf{N}_i$ the correlation matrix of the noise plus external interference) and $\mathbf{P}_i$= diag($p_{i1}, \ldots, p_{im_i}$).

The total power transmitted by the k-th BS is given by:

$$\begin{aligned} P^k &= \text{tr}\, E\left\{ \mathbf{x}_k \mathbf{x}_k^H \right\} = \text{tr} \sum_{i=1}^{R} \tilde{\mathbf{W}}_i^k \mathbf{P}_i \tilde{\mathbf{W}}_i^{kH} \\ &= \sum_{i=1}^{R} \sum_{j=1}^{m_i} p_{ij} \tilde{\mathbf{w}}_{ij}^{kH} \tilde{\mathbf{w}}_{ij}^k = \sum_{i=1}^{R} \sum_{j=1}^{m_i} p_{ij} \delta_{ij}^k \end{aligned} \tag{6}$$

where $\mathbf{x}_k$ is the signal transmitted by the k-th BS, $\tilde{\mathbf{W}}_i^k$ contains the n_B rows of $\mathbf{V}_i \mathbf{W}_i$ used by the k-th BS in the transmission of message to the i-th RS and $\tilde{\mathbf{w}}_{ij}^k$ is the j-th column of $\tilde{\mathbf{W}}_i^k$. Moreover, the power transmitted by l-th antenna at the k-th BS is:

$$P_l^k = E\left\{ \left| x_{kl} \right|^2 \right\} = \sum_{i=1}^{R} \tilde{\mathbf{w}}_i^{kl\,T} \mathbf{P}_i \tilde{\mathbf{w}}_i^{kl} \tag{7}$$

where x_{kl} is the l-th element of $\mathbf{x}_k$ and $\tilde{\mathbf{w}}_i^{kl}$ is the l-th column of $\tilde{\mathbf{W}}_i^k$.

3.2. Signal model on the second hop

On the second hop, each RS transmits to its associated MS on a fixed fraction of time $\alpha_2 = 1 - \alpha_1$, subframe which may be split over F time slots (being F an integer submultiple of R) of durations $\alpha_{21}, \ldots, \alpha_{2F}$ (see Figure 1). On each time slot, R/F relays can transmit. In this way we reduce interference at the expenses of some loss in spectral efficiency.

As we are assuming no coordination among RSs and simple receivers at the MS, only single user MIMO transmissions can be appointed. The achievable rate for each RS-MS link, r_2, follows the conventional MIMO capacity expression, affected by the presence of interference from other RS transmissions:

$$r_{2iq} = \log_2 \left(1 + \frac{P_{iq}^R}{n_R} \mathbf{h}_{ii}^H \mathbf{h}_{ii} \left(\sigma^2 + \sum_{j=1, j \neq i}^{R} \frac{P_{jq}^R}{n_R} \mathbf{h}_{ji}^H \mathbf{h}_{ji} \right)^{-1} \right) \tag{8}$$

where $i = 1, \ldots, R$; $q = 1, \ldots, F$; r_{2iq} denotes the rate in the i-th RS-MS link, which has been scheduled in timeslot q; P^R_{iq} is the power transmitted by the i-th RS in the q-th time slot to its MS and P^R_{jq} defines the power transmitted by the j-th RS on the same time slot.

When F=R (interference is avoided at RS transmissions), the best solution is setting P^R_{iq} to the maximum power on each RS (P^R_{max}). Otherwise, when F<R we can adapt the power transmitted by each RS in such a way that the interference generated to other MS on q-th time slot is reduced, and hence r_{2iq} increases. To that end, we propose the following optimization for each q-th time slot:

$$\min_{P_{iq}^R} \; -f\left(r_{21q}, \ldots, r_{2Rq}\right) \qquad \text{s.t.} \quad P_{iq}^R \geq 0 \quad i = 1, \ldots, R \tag{9}$$

a problem that is not convex in P^R_{iq} even for concave target functions $f(,)$. However, it is sure that there is a better option than all RSs transmitting at P^R_{max} which can be obtained by applying interior point methods [11] initializing P^R_{iq} with P^R_{max}.

To preserve information flow through RSs, the rate at the i-th MS served by the i-th RS is constrained by the minimum of rates in both hops:

$$r_i \le \min\left(\alpha_1 r_{1i}, \alpha_{2q} r_{2iq}\right) \tag{10}$$

where $i = 1,\ldots,R$; $q = 1,\ldots,F$; and r_{2iq} is the rate in the second hop obtained from the optimization in (9). Equation (10) can also be written as two simultaneous inequalities:

$$\alpha_1 \sum_{j=1}^{m_i} \log_2\left(1+s_{ij}^2 p_{ij}\right) \ge r_i \qquad \alpha_{2q} r_{2iq} \ge r_i \tag{11}$$

3.3. WSR-based resource allocation

We allocate the resources based on the maximization of the weighted sum-rate (WSR) criterion that allows adding certain QoS over the served users depending on priorities μ_i:

$$\left(\mathrm{P_{WSR}}\right): \quad \min_{\{r_i\},\{p_{ij}\}} \; -\sum_{i=1}^{R} \mu_i r_i$$
$$s.t. \begin{cases} r_i - \alpha_1 \sum_{j=1}^{m_i} \log_2\left(1+s_{ij}^2 p_{ij}\right) \le 0 & i=1,\ldots,R \\ \sum_{i=1}^{R}\sum_{j=1}^{m_i} p_{ij}\delta_{ij}^k - P_{\max}^k \le 0 & k=1,\ldots,B \\ r_i - \alpha_{2q} r_{2iq} \le 0 & i=1,\ldots,R \\ -r_i \le 0 & i=1,\ldots,R \\ -p_{ij} \le 0 & i=1,\ldots,R \quad j=1,\ldots,m_i \end{cases} \tag{12}$$

where δ_{ij}^k is defined in equation (6). Note that problem ($\mathrm{P_{WSR}}$) is convex and can be solved using standard convex optimization techniques, like interior point methods [11]. Nevertheless, we can further elaborate towards an efficient numerical algorithm based on the dual update methods [11][14] that will define a polynomial complexity algorithm along with a reduction of the number of variables to be optimized.

The main difficulty in solving ($\mathrm{P_{WSR}}$) is caused by the max-rate constraints imposed by the transmissions in the second hop (third constraint in equation (12)). When they are active, the unique maximum WSR is attained by many power allocation strategies. For example, all the available power at the BSs could be used while adopting a bitrate lower than the Shannon rate bound. Power would therefore be wasted.

We deal with that drawback by transforming the max-rate constraints into power constraints per stream and reformulating the optimization problem ($\mathrm{P_{WSR}}$), see the details in the Appendix. The rate and power allocation thus obtained become,

$$\begin{cases} p_{ij}\left(\lambda_k^*, \tilde{\lambda}_i^*\right) = \left[\dfrac{\alpha_1}{\ln 2}\dfrac{\mu_i}{\sum_{k=1}^{B}\lambda_k^*\delta_{ij}^k + \tilde{\lambda}_i^*} - \dfrac{1}{s_{ij}^2}\right]^+, & i=1,\ldots,R \\ r_i = \alpha_1 \sum_{j=1}^{m_i}\log_2\left(1+s_{ij}^2 p_{ij}\left(\lambda_k^*, \tilde{\lambda}_i^*\right)\right) & k=1,\ldots,B \end{cases} \tag{13}$$

The values of $\lambda_k^*, \tilde{\lambda}_i^*$ are calculated using algorithm in Table I, which is based on the bisection method for λ_k^*, and the ellipsoid method [14] for $\tilde{\lambda}_i^*$. The subgradients required to update them are:

$$d_k = \sum_{i=1}^{R}\sum_{j}^{m_i} p_{ij}\delta_{ij}^k - P_{\max}^k, \quad \tilde{\tilde{d}}_i = \alpha_1 \sum_{j=1}^{m_i}\log_2\left(1+s_{ij}^2 p_{ij}\right) - \alpha_{2q} r_{2iq} \tag{14}$$

4. EVALUATION AND RESULTS

The evaluation of the proposed approach is done on a radio access network based on 802.16m specifications [1] at the 2.6 GHz band and 20 MHz bandwidth. Channel models adopted are outdoor-to-outdoor obtained from [12]. We assume LOS conditions for all BS-RS links and distance-dependent LOS/NLOS condition for BS-MS and RS-MS links. 3 BSs and a total of 6 RSs are deployed. On each scenario, 6 MSs are dropped, each attached to a different RS. All RSs are at the same distance d to their associated BS, equal to 60% of the cell radius (experimentally found as the best position for the case with F=1). Transmit powers are 40 dBm and 30 dBm, and antenna gains of 10,6 dBi and 5 dBi, for BS and RS respectively. Noise spectral density is -174 dBm/Hz. The number of antennas is n_B=4 at the BS, n_R=2 at the RS and n_M=1 at the MS, and thus $m_i = 2$.

It has been observed that users close to the BS are not benefited from the RS assistance. While we recognize that a practical scheduling scheme should consider splitting the population between those close-to and those far-from their serving BS, the topic is deferred to a forthcoming study. In order to include only those users benefiting from the presence of relays, in our evaluations users are uniformly placed beyond 35% of the cell radius. The cell arrangement is shown in Figure 1.

Two fundamental measures are adopted: *cellular spectral efficiency* (S_e), as the sum rate of R users averaged over many deployments, and *outage rate* (r_{out}), as the peak achievable rate of the ε-percentile worst users in the cell over many deployments. Both capture most of the benefits offered by coordination of BS and relay-based transmission.

Table I. Algorithm solving $\mathrm{P_{WSR}}$ for B=3 and R=6

1. Initialize: $\lambda_1^{\max}, \lambda_1^{\min}, \lambda_2^{\max}, \lambda_2^{\min}, \lambda_3^{\max}, \lambda_3^{\min}$
2. while $\left|\lambda_1^{\max} - \lambda_1^{\min}\right| \le \varepsilon$ do
3. $\quad \lambda_1 = \frac{1}{2}\left(\lambda_1^{\max} + \lambda_1^{\min}\right)$
4. $\quad$ while $\left|\lambda_2^{\max} - \lambda_2^{\min}\right| \le \varepsilon$ do
5. $\qquad \lambda_2 = \frac{1}{2}\left(\lambda_2^{\max} + \lambda_2^{\min}\right)$
6. $\qquad$ while $\left|\lambda_3^{\max} - \lambda_3^{\min}\right| \le \varepsilon$ do
7. $\qquad\quad \lambda_3 = \frac{1}{2}\left(\lambda_3^{\max} + \lambda_3^{\min}\right)$
8. $\qquad\quad \left[\tilde{\lambda}_1, \ldots, \tilde{\lambda}_6\right] =$ *Ellipsoid method* {
9. $\qquad\qquad$ Initialize $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_6$
10. $\qquad\qquad$ Repeat
11. $\qquad\qquad\quad$ - Compute $p_{ij}\left(\lambda_1, \lambda_2, \lambda_3, \tilde{\lambda}_i\right)$ given by (13)
12. $\qquad\qquad\quad$ - Compute subgradient $\tilde{\tilde{d}}_i$ given by (14)
13. $\qquad\qquad\quad$ - Update $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_6$, [14]
14 $\qquad\qquad$ until convergence}
15. $\qquad$ if $d_3 < 0$, $\lambda_3^{\max} = \lambda_3$, else $\lambda_3^{\min} = \lambda_3$
16. $\qquad$ end while
17. $\quad$ if $d_2 < 0$, $\lambda_2^{\max} = \lambda_2$, else $\lambda_2^{\min} = \lambda_2$
18. $\quad$ end while
19. if $d_1 < 0$, $\lambda_1^{\max} = \lambda_1$, else $\lambda_1^{\min} = \lambda_1$
20. end while

The duration of time slot α used in the simulations is based on the results obtained in [13], where it is shown that the optimum α_1 in terms of spectral efficiency is a random variable that depends on the particular scenario and the target function to be maximized. It was observed that its mean value is lower for F=6 than for F=1. Fixing the position of the RS at 60% of the cell radius, the mean value of the optimum α_1 does not exhibit an appreciable variation with higher cell radius for F=6 (no intereference in the second hop). For F=1 it increases with cell radius due to the fact that interference in the second hop is reduced with higher cell radius while rate losses in the BS-RS link are larger due to increased distance. However the main variation of α values is given by the objective function to be maximized. When using WSR and the weights are inversely proportional to the rates in the second hop (that is $\mu_i = 1/r_{2i}$ in order to avoid unfair service to deprived users) lower values of α_1 are observed than for SR.

In a first study, the optimal solution to the (P_{WSR}) in (12) is evaluated for the SR and the WSR, over 1000 random user deployments with F=1. The objective function used in (9) is the product of rates because it achieves an enhanced r_{out} and allows the design of a more fair service in the second hop, although other criteria could also be used. Figure 2 displays r_{out} vs. S_e for cell radiuses of 500, 750 and 1000 m adopting different transmission strategies: N-MIMO with relay-assistance (Relayed N-MIMO α_1=0.2 and Relayed N-MIMO α_1=0.3 in legend), N-MIMO with relay-assistance and dynamic optimized α (Relayed N-MIMO), N-MIMO direct transmission (N-MIMO) and relayed transmissions with uncoordinated beamforming precoding at the BS (BF-TDMA), both under the SR criterion. In this later case, each BS serves its associated RS under round-robin TDMA. Figure 2 shows significant gains in terms of S_e by N-MIMO strategies as compared to BF-TDMA, remarkably boosted by the use of relays. Moreover, when α is optimized, the gains in terms of spectral efficiency and outage rate are comparable with the fixed α case. This confirms the possibility of having systems gains even if the duration of phases is kept fixed over the time

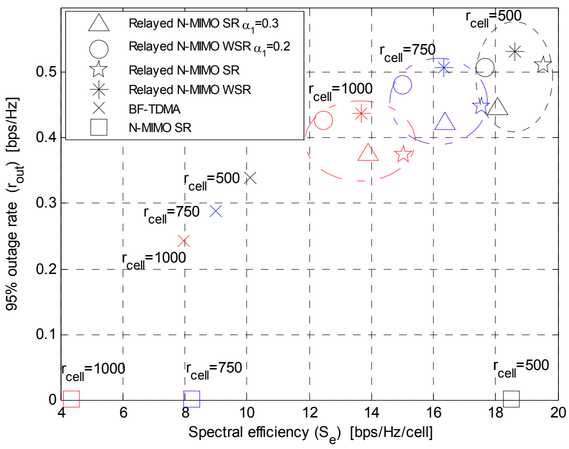

Figure 2 - Outage transmission rate (r_{out}) vs. spectral efficiency (S_e) for different transmission strategies and F=1, using and not using RS. Cell radiuses are 500 m (black), 750 m (red) and 1000 m (blue).

In the following we would like to analyze the total power efficiency of our system for the cases where the transmitted power at the relays is optimized (Relayed N-MIMO), as it was considered in Figure 2, the case where RSs transmit at full power (Relayed N-MIMO P^R_{max}) and finally with the direct transmission case (N-MIMO). Notice that by optimizing the transmitted power at RSs, we transmit less power than the Relayed N-MIMO P^R_{max} case. In order to fairly compare all cases, we define power efficiency as:

$$\xi_{ef} = \tfrac{1}{P}\sum_{i=1}^{R} r_i \quad \text{(bps/Hz/Watt)}$$

where, due to half duplexing of the RS:

$$P = \alpha_1 \sum_{k=1}^{B} P_k + \alpha_2 \sum_{i=1}^{R} P_i^R$$

Figure 3 shows the cumulative density function (cdf) of ξ_{ef} when WSR or SR criteria, both with F=1, are adopted. We can see that by optimizing the transmitting power at RSs, ξ_{ef} is nearly doubled: achievable user's bit rate are higher and the system power consumption is lower. Moreover, when compared to direct N-MIMO transmissions, the power efficiency of relay transmissions is nearly fourfold as a consequence of the enhanced spectral efficiency and the fact that RSs transmit lower power than BSs.

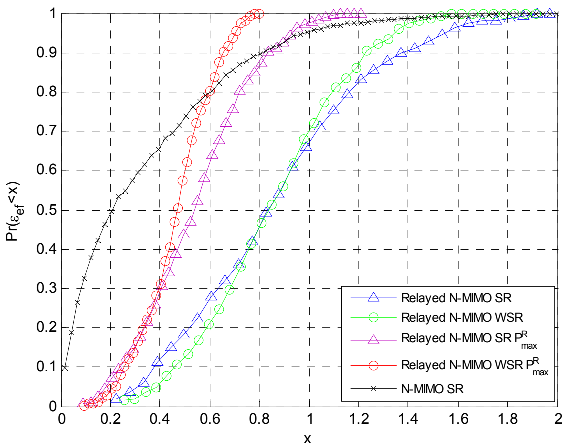

Figure 3 - Cumulative density function (cdf) of power efficiency for a cell radius of 750 m when RS transmits with less or equal to its maximum allocated power.

5. CONCLUSIONS AND OUTLOOK

Optimal resource allocation algorithms for QoS-constrained relay-assisted cellular systems have been proposed, where cooperation between BS is appointed and the duration of the relaying phases is fixed. Results show that optimizing the duration is not critical, and still large gains both in terms of spectral efficiency and outage rate are obtained. This observation facilitates the adoption of relay transmissions in next-generation wireless systems, as this guarantees the predictability of other cell-clusters interference within transmission frames. Further work is oriented to study modulation-constrained resource allocation, user's grouping and scheduling strategies and coordination of transmissions between cell clusters.

6. APPENDIX

The Lagrangian function of (P_{WSR}) in equation (12) becomes:

$$L\left(\{r_i\},\{p_{ij}\},\{\lambda_k\},\{\psi_i,\gamma_i,\eta_i\},\{\varphi_{ij}\}\right) = -\sum_{i=1}^{R} \mu_i r_i +$$
$$+\sum_{k=1}^{B} \lambda_k \left(\sum_{i=1}^{R} \sum_{j}^{m_i} p_{ij} \delta_{ij}^k - P_{\max}^k \right) + \sum_{i=1}^{R} \psi_i \left(r_i - \alpha_1 \sum_{j=1}^{m_i} \log_2\left(1 + s_{ij}^2 p_{ij}\right) \right) \quad (15)$$
$$+\sum_{i=1}^{R} \gamma_i (r_i - \alpha_{2q} r_{2iq}) - \sum_{i=1}^{R} \eta_i r_i - \sum_{i=1}^{R} \sum_{j=1}^{m_i} \varphi_{ij} p_{ij}$$

where $\lambda_k, \psi_i, \gamma_i$ denote the Lagrange multipliers or dual variables associated to the max-power per BS, Shannon bitrate bound and max-rate constraints, respectively. Finally, η_i, φ_{ij} are the Lagrange multipliers needed for having positive values of bitrate and allocated power at BSs. The conditions to minimize the Lagrangian as a function of the bitrate r_i and allocated power p_{ij} become,

$$\begin{cases} \partial L/\partial r_i = 0 & \rightarrow -\mu_i + \psi_i + \gamma_i - \eta_i = 0 \\ \partial L/\partial p_{ij} = 0 & \rightarrow -\dfrac{\alpha_1 \psi_i}{\ln 2} \dfrac{s_{ij}^2}{1 + s_{ij}^2 p_{ij}} + \sum_{k=1}^{B} \lambda_k \delta_{ij}^k - \varphi_{ij} = 0 \end{cases} \tag{16}$$

The previous conditions allow having an expression for p_{ij} but only if there is at least one BS using its max power ($\lambda_k \neq 0$). However, if the second hop is limiting the maximum rate, not all the available power is needed ($\lambda_k = 0\ \forall k$). In such a case, there are multiple power allocation strategies providing maximum weighted sum-rate.

Since we are interested in attaining the maximum rate but using the minimum required power, we transform the bitrate inequalities in the second hop into max-power constraints. In this regard, let us define the maximum power used in the first hop for the i-th RS as the solution of the optimization problem,

$$(\mathrm{P_{SP}}): \quad \begin{array}{ll} \min\limits_{\{p_{ij}\}} & \dfrac{1}{\mu_i} \sum\limits_{j=1}^{m_i} p_{ij} \\ s.t. & \begin{cases} \alpha_{2q} r_{2iq} - \alpha_1 \sum_{j=1}^{m_i} \log_2 \left(1 + s_{ij}^2 p_{ij}\right) \le 0 \\ -p_{ij} \le 0 \qquad j = 1, \ldots, m_i \end{cases} \end{array} \tag{17}$$

The power allocation turns out to be,

$$p_{ij}(\omega_i) = \left[\frac{\alpha_1 \mu_i}{\ln 2} \omega_i - \frac{1}{s_{ij}^2} \right]^+, \quad \omega_i^* = 2^{\frac{1}{m_i}\left(\alpha_2 r_{2iq} - \sum\limits_{j=1}^{m_i} \log_2 \left(\frac{\alpha_1 \mu_i s_{ij}^2}{\ln 2} \right) \right)} \tag{18}$$

and the total power employed for the transmission to the i-th RS is

$$P_{2iq} = \sum_{j=1}^{m_i} p_{ij}\left(\omega_i^*\right) \tag{19}$$

Now, we reformulate the problem ($\mathrm{P_{WSR}}$) in (12) taking into account the max-power per stream when we are limited by the bitrate of the second hop (19) :

$$\left(\tilde{\mathrm{P}}_{\mathrm{WSR}}\right): \quad \begin{array}{ll} \min\limits_{\{r_i\},\{p_{ij}\}} & -\sum\limits_{i=1}^{R} \mu_i r_i \qquad \{ i = 1, \ldots, R \}, \{ k = 1, \ldots, B \} \\ s.t. & \begin{cases} (\psi_i): \ r_i - \alpha_1 \sum_{j=1}^{m_i} \log_2 \left(1 + s_{ij}^2 p_{ij}\right) \le 0 \\ (\lambda_k): \ \sum_{i=1}^{R} \sum_{j=1}^{m_i} p_{ij} \delta_{ij}^k - P_{\max}^k \le 0 \\ (\tilde{\lambda}_i): \ \sum_{j=1}^{m_i} p_{ij} - P_{2iq} \le 0 \\ (\eta_i): \ -r_i \le 0 \\ (\varphi_{ij}): \ -p_{ij} \le 0 \qquad j = 1, \ldots, m_i \end{cases} \end{array} \tag{20}$$

Notice that we have included the Lagrange multipliers $\tilde{\lambda}_i$ instead of γ_i in (15) with the proper power constraint P_{2iq} as obtained from (19). The conditions to optimize (20) become,

$$\begin{cases} \psi_i = \mu_i, \quad \eta_i = 0, \varphi_{ij} = 0 \\ p_{ij}\left(\lambda_k, \tilde{\lambda}_i\right) = \left[\dfrac{\alpha_1}{\ln 2} \dfrac{\mu_i}{\sum_{k=1}^{B} \lambda_k \delta_{ij}^k + \tilde{\lambda}_i} - \dfrac{1}{s_{ij}^2} \right]^+ \end{cases} \tag{21}$$

Let us define the *dual function* of $\left(\tilde{\mathrm{P}}_{\mathrm{WSR}}\right)$ taking into account its Lagrangian and the solution in (21),

$$\begin{gathered} g\left(\{\lambda_k\}, \{\tilde{\lambda}_i\}\right) = -\sum_{i=1}^{R} \mu_i \alpha_1 \sum_{j=1}^{m_i} \log_2 \left(1 + s_{ij}^2 p_{ij}\right) + \sum_{k=1}^{B} \lambda_k d_k + \sum_{i=1}^{R} \tilde{\lambda}_i \tilde{d}_i \\ d_k = \sum_{i=1}^{R} \sum_{j=1}^{m_i} p_{ij} \delta_{ij}^k - P_{\max}^k, \quad \tilde{d}_i = \sum_{j=1}^{m_i} p_{ij} - P_{2iq} \end{gathered} \tag{22}$$

The optimal values of the Lagrange multipliers are obtained by maximizing the dual function,

$$(\mathrm{P_{DP}}): \quad \begin{array}{ll} \max & g\left(\{\lambda_k\}, \{\tilde{\lambda}_i\}\right) \\ s.t. & \lambda_k \ge 0\ \forall k, \quad \tilde{\lambda}_i \ge 0\ \forall i \end{array} \tag{23}$$

Since ($\mathrm{P_{DP}}$) is convex, gradient-type search is guaranteed to converge to the global optimum of (23). Search directions given by $d_k, \tilde{d}_i$ in (22) coincide with the subgradient, [14]. This suggests that if a given constraint is exceeded the associated Lagrange multiplier should be increased, or decreased otherwise. In this respect, we can avoid calculating P_{2iq} in (19) by defining the gradient in equation (14) that also accounts for the maximum rate.

The algorithm presented in Table I compiles the method, and it is able to provide the optimal values for λ_k^*, $\tilde{\lambda}_i^*$ with a polynomial complexity.

7. REFERENCES

[1] IEEE 802.16 Broadband Wireless Access Group, "IEEE 802.16m system requirements", IEEE 802.16m-07/002r4, October 2007.

[2] A. Host-Madsen, J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans on Information Theory,* vol.51, no.6, pp.2020-2040, June 2005.

[3] A. Agustin, J. Vidal, O. Munoz, "Protocols and Resource Allocation for the Two-Way Relay Channel with Half-Duplex Terminals," *IEEE Int. Conference on Communications, 2009*, 14-18 June 2009.

[4] R. S. Blum, "MIMO Capacity with Interference", *IEEE J. Select. Areas Commun.,* vol. 21, no. 5, June 2003.

[5] H. Takahashi, *et al*, "Field Experimental Results on E-UTRA Downlink Throughput Using Polarization Diversity Antennas", *IEEE Vehicular Technology Conference,* April 2009.

[6] G. J. Foschini, K. Karakayali, R. A. Valenzuela, "Network Coordination for Spectrally Efficient Communications in Cellular Systems", *IEE Proc.-Commun.,* Vol. 153, No. 4, August 2006.

[7] G. Kramer, M. Gastpar, P. Gupta, "Cooperative Strategies and Capacity Theorems for Relay Networks," *IEEE Trans on Information Theory*, vol.51, no.9, pp. 3037- 3063, Sept. 2005.

[8] M. Kieβling, "Unifying Analysis of Ergodic MIMO Capacity in Correlated Rayleigh Fading Environments", *European Transactions on Telecommunications*, vol. 16, issue 1, 2005.

[9] Q.H. Spencer, A.L. Swindlehurst, M. Haardt, "Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels", *IEEE Trans on Signal Processing*, vol. 52, no. 2, February 2004.

[10] S. Shi, M. Schubert, N. Vucic, H. Boche, "MMSE Optimization with Per-Base-Station Power Constraints for Network MIMO Systems", *IEEE Int. Conference on Communications (ICC)*, Beijing, May 2008.

[11] S. Boyd, L. Vandenberghe, *Convex Optimization*, Cambridge University Press, USA, 2004.

[12] WINNER II consortium, "Channel Models Part II: Radio Channel Measurements and Analysis Results", Deliverable 1.1.2, IST-4-027756 WINNER II, September 2007.

[13] J. Vidal, A. Agustín, S. Lagén, E. Valera, O. Muñoz, A. García Armada, M. Sánchez Fernández, "Network-MIMO backhauling for QoS-constrained relay tranmission", *IEEE Int Conf on Acoustics, Speech and Signal Processing,* ICASSP 2011, Prague, May 2011.

[14] W. Yu, R. Lui, "Dual Methods for Nonconvex Spectrum Optimization of Multicarrier systems", *IEEE Trans. on Communications*, vol. 54, no.7, July 2006.

[15] R. Zhang, "Cooperative Multi-Cell Block Diagonalization with Per-Base-Station Power Constraints", *IEEE J. Select. Areas Commun.,* vol.28, no. 9, December 2010.