Dynamic Resource Allocation in Heterogeneous Networks

Tania Villa*, Ruben Merz‡, Raymond Knopp*
∗ Mobile Communications Dept., Eurecom, Sophia Antipolis, France
‡ Swisscom, Strategy & Innovation, Bern, Switzerland

***Abstract*—LTE systems do not suffer from intra-cell interference, but they are affected by interference coming from adjacent cells. However, most of the research on resource allocation and repetition protocols has not paid attention to the interference case. In this paper, we consider the problem of dynamic resource allocation for IR-HARQ schemes under the presence of interference. We consider resource allocation by means of rate and physical dimensions adaptation in each HARQ round. We provide a mathematical framework that can be applied for the analysis of heterogeneous networks. Rather than performing extensive simulations, we take an information theoretic approach to derive analytical expressions that represent the long-term throughput of the network and we consider distributed resource allocation policies. Our policies are applicable for both the uplink and downlink channels.**

I. Introduction

LTE performance in terms of spectral efficiency and available data rates is, relatively speaking, more limited by interference from adjacent cells as compared to previous communication standards [1]. Means to reduce or control the inter-cell interference can potentially provide substantial benefits to LTE performance, especially in terms of the quality of service provided to every user.

Traditionally, hybrid automatic repeat and request (HARQ) has been used as a way to recover from errors occurring during the transmission of information. When the retransmission consists of the addition of new parity bits, we refer to it as incremental redundancy (IR) [2]. Another technique to improve system performance is to adapt the coding rate. The code rate can be fine-tuned by puncturing, generating different redundancy versions to match the number of coded bits to the channel. The code rate, rate matching, and the number of resources allocated for one transmission determine the transport-block size [3]. In essence, rate adaptation adapts the modulation and coding scheme (MCS) to the current channel conditions which translates to the data rate or error probability of each link. The MCS, in terms of spectral efficiency, represents the number of information bits per modulation symbol. The use of rate adaptation provides manufacturers an incentive to implement more advanced receivers since it will result in higher end-user data rates than standard receivers [1].

Extensive research has explored variable-rate adaptation techniques. However, very little attention has been paid to the more performance-limited case of interference. The throughput of HARQ has been investigated in the limit of infinite block length for Gaussian input signals [4] over a Gaussian channel with fading. In [5], the long-term throughput analysis of a HARQ protocol under slow-fading channels is presented for fixed-rate, variable-power transmissions under the framework of the renewal-reward theory of [4]. Rate adaptation for HARQ protocols under delay constraints is studied in [6], and for time-correlated channels in [7] and [8]. Power adaptation is presented in [9] to minimize the outage probability and in [10], both power and rate control are derived through dynamic programming without outage constraints. In [11], the optimization of either the packet drop probability or the average transmit power is shown for the case of IR HARQ with a maximum number of retransmissions. In [12], the information theoretic approach of [4] is adapted to variable rate transmissions in the case of HARQ with IR. Finally, in [13], a mathematical framework based on a sum-rate analysis for heterogeneous networks with partial feedback is developed.

We consider dynamic resource allocation for IR-HARQ schemes under the presence of interference. The latter is a real possibility in schedulers for LTE base stations and, to the best of our knowledge, no well-known methodology exists for adapting physical resources across HARQ rounds when subject to time-varying channels. Rather than performing extensive simulations, we take an information theoretic approach to derive analytical expressions that represent the long-term throughput of the network and consider practical cases where there is a constraint on the outage probability representing the latency of the protocol. Our contributions are the following:

- We motivate the use of inter-round resource allocation through a simple but illustrative analysis with Gaussian signals and interference.
- We provide a mathematical framework for the analysis of heterogeneous networks and we derive analytical expressions, based on mutual information modeling, that capture the throughput performance of such networks.
- We develop adaptive resource allocation policies that are applicable both for the uplink (UL) and downlink (DL) channels which are based on dynamic adaptation of the physical dimensions and coding rate used in each HARQ round.

The remainder of this paper is organized as follows. The system model and assumptions are presented in section II. A motivating example for rate adaptation with interference is given in section III. Our resource allocation policies are exposed in section IV. Finally, we conclude in section V.

II. Modeling and Assumptions

We consider a slotted transmission scheme and we take an information theoretic approach to analyze the throughput performance. When there is more than one user, we assume that all transmissions in every slot are synchronized and we randomize the interference process with the use of activity factors. The latter models sporadic interference patterns characteristic of future heterogeneous networking deployments, in particular the interference seen from small cell base stations with bursty traffic in the receiver of a macro cell user. It can also model dual-carrier networks with cross-carrier scheduling. In this type of network, we can talk about clean and dirty carriers. On the one hand, clean carriers are used by the macro cell to carry their data plus signaling for small cells because of their controlled interference property. On the other hand, dirty carriers are interfering carriers where the "cleaning" is done with the use of HARQ.

We consider a maximum of $\mathcal{R}$ HARQ transmission rounds and the channel is independent and identically distributed (iid) or constant over all the transmission rounds of the protocol. After each transmission we receive an error-free acknowledgment (ACK or NACK) indicating a successful or unsuccessful transmission. We define the probability of outage as being unsuccessful to correctly receive the information at the end of the HARQ protocol. This probability translates to the latency of the protocol and quality of service in our system.

In general, we define $R_{r,i}$ as the code rate at the rth round. For user i, we define the number of dimensions in time as T and the number of dimensions in frequency as $L_{r,i}$. At each transmission round, the total number of dimensions is $L_{r,i}T$. Assuming the channel does not vary during T time dimensions and for a packet length of M information bits, the rate $R_{r,i}$ at the rth round, in bits/dim is given by:

$$R_{r,i} = \frac{M}{L_{r,i}T} \quad \text{bits/dim.} \tag{1}$$

In IR-HARQ, the retransmission consists of the same set of information bits as the original, however, the set of coded bits are chosen differently and they may contain additional parity bits. In each of the transmission rounds there are $L_{r,i}T$ dimensions (see figure 1).

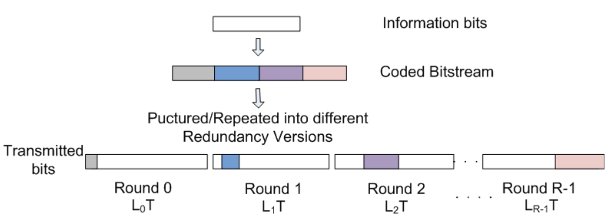

Fig. 1. Coding Model

In the context of LTE, the number of physical dimensions $L_{r,i}T$ refers to the number of resource blocks allocated to one user in one subframe of 1 ms duration (one TTI). There are at most two transport blocks delivered to the physical layer in the case of spatial multiplexing [14]. In a single-user LTE system, there is only one transport block in one TTI, representing only one codeword "in the air" at the same time. Each transport block is carried by an HARQ process, and each process is assigned to a subframe (number of processes is fixed). In our model, if the number of dimensions for user i is less than the maximum number of available resources N_R, $(L_{r,i}T < \max\{N_R\})$, then the rest will not be utilized. Although not possible in the current LTE standard, one could propose to assign the unused resources to transmit multiple codewords in parallel (at the same time), to increase the throughput. In a multiuser system, the remaining dimensions would be allocated to other users and thus the efficiency of the protocol should be chosen to maximize the aggregate spectral efficiency of the cell.

Let P_{succ_0} be the probability of having a successful transmission in the first round, and $P_{succ_r,fail_{r-1}}$ the probability of not having a successful transmission in the $(r-1)$th round, but being successful in the rth round. Finally, let P_{out} represent the probability of outage. The overall throughput can thus be expressed as:

$$T = P_{succ_0} R_0 + \sum_{r=1}^{\mathcal{R}-1} P_{succ_r,fail_{r-1}} \left(\frac{R_r}{r}\right) \quad \text{bits/dim.} \tag{2}$$

where the outage probability is given by $P_{out} = \Pr(r = \mathcal{R}) = 1 - \sum_{r=0}^{\mathcal{R}-1} P_{succ_r}$

We consider N transmitters, where user i is the transmitter of interest, and the remaining $N-1$ transmitters are interferers. We model an OFDMA physical layer with L subcarriers. We let $\mu_{j,k}$ be the activity factor, P_j the transmission power, and $x_{j,k}$ the input signal of the jth user on the kth subcarrier. We assume discrete signals with equal probabilities and size of the constellation $|x_{j,k}| = \mathcal{S}$, $z_{i,k}$ is the zero mean complex Gaussian noise with variance σ^2. Since we assume Rayleigh fading, $h_{j,k}$ is a circularly symmetric complex Gaussian random variable with unit mean. The received signal at node i is y_i and is given by:

$$y_i = \sum_{j=0}^{N-1} \sqrt{P_j} \sum_{k=0}^{L-1} h_{j,k}\mu_{j,k}x_{j,k} + z_{i,k} \tag{3}$$

Variations in the channel are caused at the receiver because of the activity factor plus the frequency shifting from the resource allocation process. For the interfering users, the channel variation depends on whether we consider the UL or DL. In the UL, it is caused by the interference coming from different user terminals. In the DL, the activity factors will introduce variations originated from the fact that the interfering cells are not active the whole time.

To model heterogeneous networks where the interference is not constant, we let the user of interest to be active all the time (macro cell user) and we let the small interfering cells to transmit with a certain probability. Our downlink case in

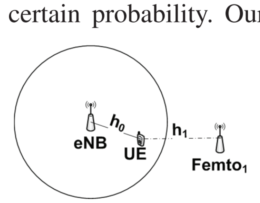

Fig. 2. Downlink of a macrocell with a femtocell interfering.

section IV-A, models the downlink of a macro user with an interfering femtocell active with probability p (see figure 2). In the next section, we provide a motivating example on the importance of adapting the rate and physical dimensions across the rounds of the HARQ protocol.

III. Simple interference analysis in zero-outage

For the sake of analytical tractability, we start by looking at the special case of (3) where the $h_{j,k}$ are fixed (AWGN channel) and we assume Gaussian signals. We consider one interferer and we model it with an activity factor, which means that the interferer could be active or inactive. The activity factor is bernoulli distributed with probability p.

The rate with Gaussian codebooks that can be achieved by the protocol depends on the interference state (interference active or inactive). Let R_H be the capacity that can be achieved without interference, and R_L the corresponding capacity with interference, which are given by:

$$R_H = \log_2(1 + \text{SNR}_1) \tag{4}$$

$$R_L = \log_2\left(1 + \frac{\text{SNR}_1}{1 + \text{SNR}_2}\right) \tag{5}$$

where SNR_1 is the signal-to-noise ratio (SNR) for the user of interest, SNR_2 is the corresponding SNR for the interferer and we assume unitary noise variance.

If we consider a HARQ protocol with two rounds, then in the first round we transmit with ρN_R dimensions and $(1-\rho)N_R$ dimensions in the second round. For a packet of length M bits, the rate in the first round is $R_1 = \frac{1}{\rho N_R}\log_2 M = R_H$ and in the second round $R_2 = \frac{1}{N_R}\log_2 M = R_L$. Therefore, $R_L = \rho R_H$, and $\rho = \frac{R_L}{R_H}$.

In the remainder of this section, we derive the zero-outage throughput with and without feedback and we consider also the case of a residual outage at the end of the protocol with feedback.

A. Zero-outage throughput without feedback and no delay constraint

We now look at the case of no feedback. Let μ define the state of the interference, if $\mu = 0$ there is no interference and $\mu = 1$ means that the interference is active and it happens with probability p. Then the throughput R with zero outage (without delay) is given by:

$$\begin{aligned} R &= \mathbb{E}_\mu I(X;Y|\mu) && (6)\\ &= (1-p)R_H + pR_L && (7)\end{aligned}$$

It is interesting to note that (7) is the ergodic capacity (average over all possible states). In the next section we explore the case when feedback becomes available and we look at the case of more than two transmission rounds.

B. Zero-outage throughput with feedback

In this case, we assume that we have feedback from the HARQ protocol and we vary the tolerable latency by fixing the maximum number of transmission rounds $\mathcal{R}$, but still assume zero-outage probability. Then, given that we want zero-outage at round $\mathcal{R}$, we choose the rate that guarantees successful decoding (i.e. R_L). We choose the rate in the first round to be as high as possible, and the intermediate rates are at the optimal value between R_L and R_H. Therefore, the rate at the r^{th} round is given by:

$$R_1 = \frac{\log_2 M}{\rho_1 N_R} = R_H \qquad r = 1 \tag{8}$$

$$R_r = \frac{\log_2 M}{(\sum_{j=1}^{r}\rho_j)N_R} = \left(\frac{\rho_1}{\sum_{j=1}^{r}\rho_j}\right)R_H \qquad 2 \le r < \mathcal{R} \tag{9}$$

$$R_{\mathcal{R}} = \frac{\log_2 M}{N_R} = R_L = \rho_1 R_H \;\Rightarrow \rho_1 = \frac{R_L}{R_H} \qquad r = \mathcal{R} \tag{10}$$

In this case, the throughput expression for $\mathcal{R}$ rounds is given by:

$$R = (1-p)R_H + \sum_{r=2}^{\mathcal{R}-1} p^{r-1}(1-p)\left(\frac{\rho_1}{\sum_{j=1}^{r}\rho_j}\right)R_H + p^{(\mathcal{R}-1)}R_L \tag{11}$$

For the rates to be achievable, we observe that there is a restriction on the ratio of dimensions after the second round ρ_r, $r > 1$. This restriction comes from the fact that the rate after round r is $\left(\sum_{j=1}^{r-1}\rho_j\right)N_R R_L + \rho_r N_R R_H$ which means that after round r we decode if:

$$\left(\sum_{j=1}^{r}\rho_j\right)N_R R_r < \left(\sum_{j=1}^{r-1}\rho_j\right)N_R R_L + \rho_r N_R R_H \tag{12}$$

$$R_r = \left(\frac{\rho_1}{\sum_{j=1}^{r}\rho_j}\right)R_H < \frac{\left(\sum_{j=1}^{r-1}\rho_j\right)R_L + \rho_r R_H}{\sum_{j=1}^{r}\rho_j}$$

$$\rho_r > \rho_1\left(1 - \sum_{j=1}^{r}\rho_j\right) \tag{13}$$

If we look at figure 3, the solid lines and the right axis show the zero-outage throughput for the HARQ protocol with a maximum number of rounds $\mathcal{R} = \{1, 2, 3, 4\}$. We can see that there is a high gain when going from one to two rounds and after three rounds there is only a marginal gain. The dashed lines and left axis show how the dimensions are being distributed across the rounds of the protocol. We illustrate the case of three rounds (i.e. $\mathcal{R} = 3$) and we look at the proportion of physical dimensions used in each round (ρ_r). In both cases, the interference strength is the same as the user of interest ($\text{SNR}_1 = \text{SNR}_2$), the channel is AWGN and we assume Gaussian signals with one interferer active with probability 50%.

C. Throughput with outage and feedback

In this case, we allow the protocol to have a residual outage probability which is overcome by an upper layer ARQ process on top of the IR-HARQ [15], and we assume that we have feedback. For two rounds, the throughput is now given by:

$$\begin{aligned} R = &(1 - P_{out,2}(\rho, R_2))\Big[(1 - P_{out,1}(\rho, R_2))\frac{R_2}{\rho} && (14)\\ &+ P_{out,1}\left(1 - P_{out,2}(\rho, R_2|Out_1)\right)R_2\Big] && (15)\end{aligned}$$

$I(\mu_r)$ is the mutual information as a function of the state of the interference at round r, and it is defined by μ_r:

$$I(\mu_r) = \begin{cases}\log_2(1+\text{SNR}_1) & \mu_r = 0\\ \log_2\left(1 + \frac{\text{SNR}_1}{1+\text{SNR}_2}\right) & \mu_r = 1\end{cases} \tag{16}$$

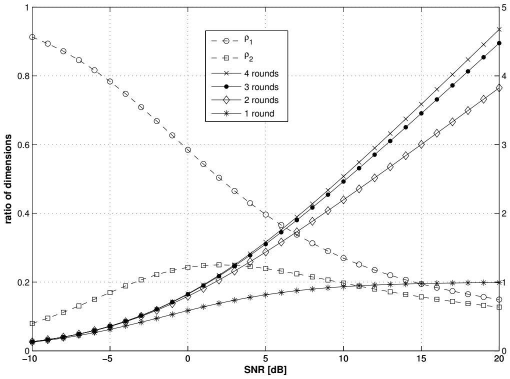

Fig. 3. The axis on the right (solid lines) shows the zero-outage throughput for the HARQ protocol with different number of rounds, while the axis on the left (dashed lines) shows the ratio of dimensions per round for the three rounds, zero-outage HARQ protocol. In both cases the channel is AWGN with Gaussian signals and there is one interferer with probability $p = 0.5$. The interference strength is the same as the user of interest ($\mathrm{SNR}_1 = \mathrm{SNR}_2$).

Now, we can define the probabilities in (14) as follows: $P_{out,1}(\rho, R_2)$ is the outage probability at the first round and it is given by:

$$\begin{aligned} P_{out,1}(\rho, R_2) &= \Pr(R_2 > \rho I(\mu_1)) \\ &= \begin{cases} 1 & \text{if } R_2 > \rho I(0) \\ 0 & \text{if } R_2 < \rho I(1) \\ \Pr(\mu_1 = 1) = p & \text{if } \rho I(1) < R_2 < \rho I(0) \end{cases} \end{aligned} \tag{17}$$

$P_{out,2}(\rho, R_2|Out_1)$ is the outage probability at the second round, given that there was an outage in the first round and it is given by:

$$\begin{aligned} P_{out,2}(\rho, R_2|Out_1) &= \Pr(R_2 > \rho I(\mu_1) + (1-\rho)I(\mu_2)\dots \\ &\quad \dots|R_2 > \rho I(\mu_1)) \\ &= \begin{cases} 1 & R_2 > \rho I(0) \\ 1 & R_2 < \rho I(1) \\ \Pr((1-\rho)I(\mu_2) < R_2 - \rho I(1)) & \rho I(1) < R_2 < \rho I(0) \end{cases} \end{aligned}$$

where $\Pr((1-\rho)I(\mu_2) + \rho I(1) < R_2) =$

$$\begin{cases} p & \text{if } I(1) < R_2 < \rho I(1) + (1-\rho)I(0) \\ 0 & \text{if } I(1) > R_2 \\ 1 & R_2 > \rho I(1) + (1-\rho)I(0) \end{cases} \tag{18}$$

Finally, $P_{out,2}(\rho, R_2)$ is the probability of outage after the second round, independently of the interference state at the first round and it is given by:

$$P_{out,2}(\rho, R_2) = \begin{cases} p & R_2 > \rho I(1) + (1-\rho)I(0) \\ p^2 & R_2 < \rho I(1) \end{cases} \tag{19}$$

Figure 4 shows the throughput of the HARQ protocol with two transmission rounds. There is one interferer with probability $p = \{0.05, 0.5\}$. We compare the zero-outage throughput against the throughput that allows an outage at the end of the protocol. We also plot the maximum capacity achieved with one round and no interference R_H and the corresponding capacity for interference R_L. If we look at the case of 50% probability of interference, we can see that the zero-outage throughput is higher for all SNR values, however, if we look at a case with a lower probability of having interference ($p = 0.05$, or 5%), we have almost the same throughput, except at high SNR, where the throughput with an outage is slightly higher. In this case, we also see that the capacity that can be achieved by adapting the rate and dimensions gets close to the capacity achieved without interference.

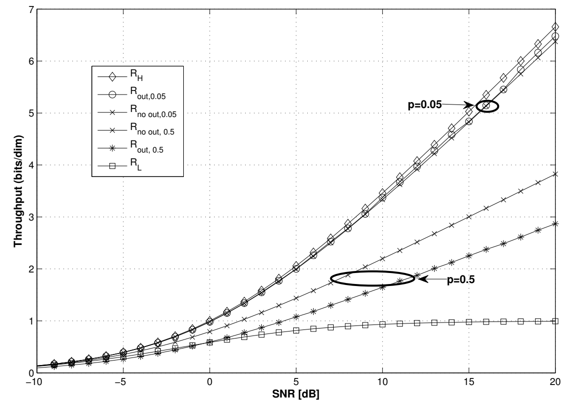

Fig. 4. Throughput of the two rounds HARQ protocol in an AWGN channel with Gaussian signals. There is one interferer with probability $p = 0.05, 0.5$.

D. Discussion

If we consider the case with the ergodic capacity and no feedback, we transmit N_R dimensions per channel realization. Therefore, we have the average capacity:

$$\mathbb{E}_\mu = \begin{cases} \log_2(1 + \mathrm{SNR}_1) & \mu = 0 \\ \log_2\left(1 + \frac{\mathrm{SNR}_1}{1+\mathrm{SNR}_2}\right) & \mu = 1 \end{cases} \tag{20}$$

where μ is the state of the interference. Now, if we consider a channel with feedback of the state of the interference (non-causal feedback). Then at round r, the transmit signal is a function of the message W and the interference state μ:

$$\begin{cases} x_r = f(W, \mu) & r > 0 \\ x_0 = f(W) \end{cases} \tag{21}$$

To get an insight into how a rate-adaptive scheme performs when changing the number of dimensions across rounds, we focus on the case of the HARQ protocol with two transmission rounds. At round r, if $\mu = 1$, then there is no transmission, and it happens with probability $\Pr(\mu_r = 1) = p$. However, if there is no interference, $\mu_r = 0$, it transmits with $\frac{N_R}{1-p}$ dimensions, and in this case we get a throughput= $(1-p)\left(\frac{\log_2(1+\mathrm{SNR}_1)\frac{N_R}{(1-p)}}{N_R}\right) = \log_2(1 + \mathrm{SNR}_1)$ which is the maximum achievable spectral efficiency. When feedback becomes available, it allows the scheme to perform better than the ergodic capacity. The latter is in contrast to the work in [4] where in the infinite delay case, the authors conclude that the maximum that can be achieved is the ergodic capacity. The difference comes from the fact that in [4] there is always a fixed bandwidth allocation for each user, regardless of the state of the channel. In our case, we dynamically adapt the bandwidth for each user depending on the interference conditions of past transmissions for the same codeword. From the perspective of the scheduler, the bandwidth is better distributed. From our initial analysis, we can conclude that the highest spectral efficiency that can be achieved happens in the case of the zero-outage protocol where increasing the delay becomes beneficial to a certain point and brings only a marginal gain after this point. In the next section, we look at practical interference scenarios where having zero-outage throughput is not possible. However, a constraint on the outage probability can be imposed.

IV. Practical interference scenarios

In this section, we derive the throughput in terms of the mutual information expressions when the signals come from discrete constellations.

Let H_r denote the vector of channel realizations in the rth round, then $I_r(H) = I_r(Y;X|H)$ denotes the corresponding instantaneous mutual information at round r. For user i, we define the mutual information at round r, in bits, as:

$$I_{r,i}(H) = T \sum_{l=0}^{r-1} \sum_{k=0}^{L_{l,i}-1} I_{k,i}(H) \tag{22}$$

where $L_{r,i}$ is the number of dimensions up to round r, $(\sum L_{r,i} = L_i)$, and $I_{k,i}(H)$ is the mutual information for user i at a particular subcarrier k, and it is given by:

$$\begin{aligned} I(Y;X|H=h) = \frac{1}{\mathcal{S}_1 \mathcal{S}_2} \sum_{x_1} \sum_{x_2} \int_y f(y|x_1, x_2, H) \\ \times \log_2 \left[\frac{\sum_{x_2'} f(y|x_1, x_2', H)}{\frac{1}{\mathcal{S}_1} \sum_{x_1'} \sum_{x_2'} f(y|x_1', x_2', H)} \right] dy \end{aligned} \tag{23}$$

In the rest of this section, we refer to the mutual information in bits/dim. For this purpose, we define $I'_{r,i}(H)$ as the mutual information in bits/dim as: $I'_{r,i}(H) = \frac{1}{L_{r,i} T} I_{r,i}(H)$. We can relate the generic throughput expression to the mutual information by defining the probabilities in (2) as:

$$P_{succ_r} = \Pr(I'_r(H) > R_r) \tag{24}$$

$$P_{succ_r, fail_{r-1}} = \Pr(I'_r(H) > R_r, I'_{r-1}(H) < R_{r-1}) \tag{25}$$

If we focus on user i, for a given channel realization $h_{r,i}$ and a particular value of SNR, the maximum rate of reliable communication supported by the channel at round r is $I_{r,i}(H)$ bits/s/Hz, which is a function of the random channel gain $h_{r,i}$ and is therefore random. If the transmitter encodes data at a rate R_r bits/s/Hz, then at round r, if the channel realization $h_{r,i}$ is such that $I_{r,i}(H) < R_r$, the transmission is called unsuccessful and this happens with probability $\Pr(I'_{r,i}(H) < R_r)$.

For IR-HARQ, mutual information is accumulated over rounds. Let $R_1 = \frac{M}{L_1 T}$ be the rate at the first round, and $R_2 = \frac{M}{L_2 T}$ the rate at the second round. If L_1 is the number of dimensions used in the first round and L_2 for the second round, the overall throughput expression is:

$$T_{HARQ} = \left(\Pr(I'_{1,i}(H)) > R_1\right) R_1 + \left(\Pr(I'_{2,i}(H)) > R_2\right) R_2 \tag{26}$$

where the outage probability is $P_{out}^{HARQ} = \Pr(I'_{2,i}(H)) < R_2)$.

In this case, zero-outage is impossible since power control and channel state feedback are not assumed [9]. However, we assume an outage constraint at the end of the HARQ protocol. To model this constraint, we consider an IR-HARQ protocol with a maximum of $\mathcal{R}$ rounds, and we say that the constraint is met whenever the packet error probability after $\mathcal{R}$ rounds is smaller than a predefined threshold P_{out}.

To find the operating rates of the protocol, we start by choosing the rate in the second round that satisfies the outage constraint, i.e. we solve $\Pr(I'_{2,i}(H) < R_2)$ for R_2. The rate in the first round (R_1), is choosen as the one that maximizes the throughput expression in (26) while satisfying the given constraint. In this case, we also optimize the number of dimensions used in each of the retransmission rounds.

Since there is no closed-form expression for the probability of outage of discrete signals, we notice that $\Pr(I(H_2) < R_2)$ represents the cumulative distribution function (cdf) of the mutual information evaluated at R_2, i.e. $F_I(R_2)$. With the help of the inversion formula in [16], we use the characteristic function of the mutual information $\Phi_I(\omega)$ to find the cdf as:

$$F_I(R_2) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{\Im\left\{\exp(-j\omega R_2)\Phi_I(\omega)\right\}}{\omega} d\omega \tag{27}$$

Finally, we can use (27) and the outage constraint to find R_2.

In figure 5, we show the case without interference. To obtain the maximum throughput, we remove the constraint on the modulation. To do this, we choose the rate in the first round according to the mutual information expression for Gaussian inputs, which is not bounded by a particular modulation order, and we choose the modulation that allows us to achieve this rate. We define threshold values for changing modulation between QPSK, 16-QAM and 64-QAM according to the maximum rate achieved with each modulation for a particular SNR value. The dashed lines represent the results for an outage constraint of 10%, and the solid lines correspond to an outage constraint of 1%. We observe a gain of around

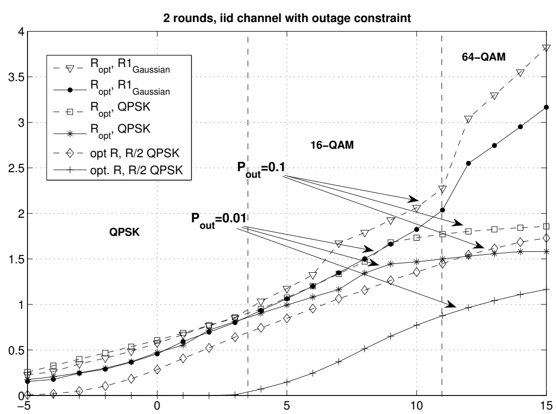

Fig. 5. Throughput of the HARQ protocol in a Rayleigh fading channel with discrete signals and without interference.

7dB for the 1% case and around 2dB for the 10% case. When we change the modulation with respect to the SNR, we observe a higher throughput in the high SNR region. This is caused by allowing the protocol to use higher modulation orders. It can be noticed that the throughput with rate optimization and an outage constraint of 1% is only lower by a small quantity as compared to the outage constraint of 10%. This tells us that optimizing the rate and dimensions can, indirectly, minimize the outage constraint (achieving almost the same throughput in the 1% case as with the more relaxed constraint of 10%). In the remainder of this section, we analyze the case for UL and DL channels with an outage constraint and interference.

A. Downlink

If we focus on the DL, since the interference is coming from adjacent cells, we can model it as non-iid with an activity factor. Figure 6 shows the throughput optimization results for

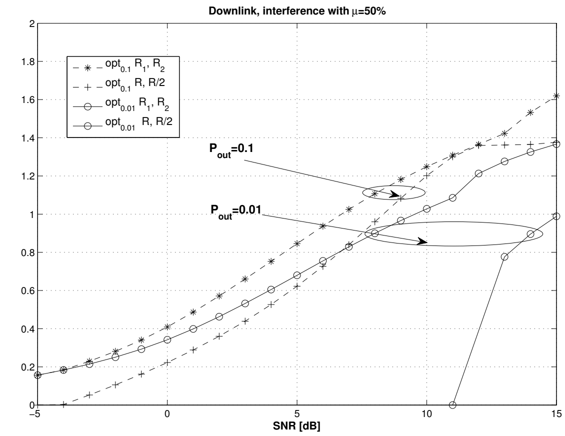

Fig. 6. Downlink optimization for QPSK signals under Rayleigh fading. There is one dominant interferer with an activity factor of 50%. The solid lines represent an outage constraint of 10% and the dashed lines correspond to the outage constraint of 1%

the DL with an activity factor of 50%, and an outage constraint of 1% and 10%. We compare it to the case of using equal number of dimensions across rounds. Although we can not have zero-outage, there is a clear gain from optimizing the dimensions across rounds for all SNR values with a maximum gain of almost 3dB for the 10% outage constraint and more than 10dB for the 1% case. The optimization of rate and dimensions has a significant impact for the lower latency case.

B. Uplink

For the UL channel, the interference is coming from users in the vicinity, therefore, we model it as iid and we consider an outage constraint of 10%. In figure 7, we show the results for the throughput optimization with an equal splitting of the dimensions across rounds and we compare it to the optimized case. For the case of the UL, results are in agreement with the DL. With a maximum gain of more than 2dB, we benefit from adapting the number of dimensions across the HARQ rounds.

V. CONCLUSIONS

In this paper, we have demonstrated the benefits of adapting the rate and physical dimensions across transmission rounds of HARQ protocols. We obtained a throughput higher than the ergodic capacity in the case of zero-outage throughput and we have showed that having an upper layer ARQ in case of a residual outage probability results in a lower throughput. In practical scenarios without power control and channel state information, it is not possible to get zero-outage throughput. However, we benefit from the dynamic resource allocation and by imposing a constraint on the outage probability, we

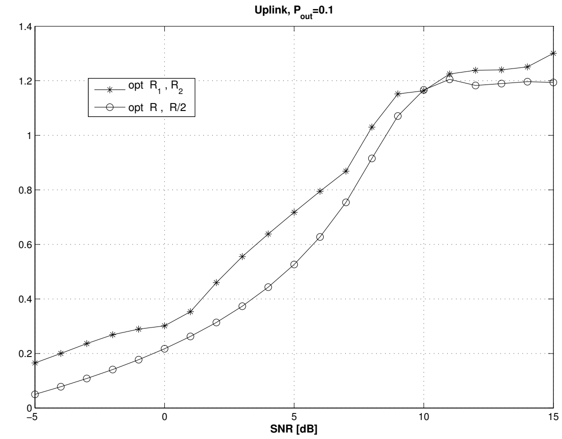

Fig. 7. Rate optimization for QPSK signals under Rayleigh fading on the uplink channel. There is one dominant iid interferer. The outage constraint is 10%

can improve the throughput by varying the latency of the protocol. We are currently investigating the performance of such resource allocation strategies on practical LTE MODEMs in an interference environment under the constraints of LTE coded-modulation.

REFERENCES

[1] E. Dahlman, S. Parkvall, J. Sköld, and P. Beming, *3G Evolution HSPA and LTE for Mobile Broadband*, 1st ed. Academic Press, 2007.

[2] S. Lin, D. Costello, and M. Miller, "Automatic-repeat-request error-control schemes," *Communications Magazine, IEEE*, vol. 22, no. 12, pp. 5 –17, december 1984.

[3] S. Sesia, I. Toufik, and M. Baker, *LTE, The UMTS Long Term Evolution: From Theory to Practice*. Wiley Publishing, 2009.

[4] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the Gaussian collision channel," *Information Theory, IEEE Transactions on*, vol. 47, no. 5, pp. 1971 –1988, july 2001.

[5] D. Tuninetti, "Transmitter channel state information and repetition protocols in block fading channels," in *Information Theory Workshop, 2007. ITW '07. IEEE*, sept. 2007, pp. 505 –510.

[6] P. Wu and N. Jindal, "Performance of hybrid-arq in block-fading channels: A fixed outage probability analysis," *Communications, IEEE Transactions on*, vol. 58, no. 4, April 2010.

[7] N. Gopalakrishnan and S. Gelfand, "Rate selection algorithms for ir hybrid arq," in *Sarnoff Symposium, 2008 IEEE*, april 2008, pp. 1 –6.

[8] S. M. Kim, W. Choi, T. W. Ban, and D. K. Sung, "Optimal rate adaptation for hybrid arq in time-correlated rayleigh fading channels," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 3, pp. 968 –979, march 2011.

[9] G. Caire, G. Taricco, and E. Biglieri, "Optimum power control over fading channels," *Information Theory, IEEE Transactions on*, vol. 45, no. 5, pp. 1468 –1489, jul 1999.

[10] I. Bettesh and S. Shamai, "Optimal power and rate control for minimal average delay: The single-user case," *Information Theory, IEEE Transactions on*, vol. 52, no. 9, pp. 4115 –4141, sept. 2006.

[11] T. Chaitanya and E. Larsson, "Outage-optimal power allocation for hybrid arq with incremental redundancy," *Wireless Communications, IEEE Transactions on*, vol. 10, no. 7, pp. 2069 –2074, july 2011.

[12] L. Szczecinski, C. Correa, and L. Ahumada, "Variable-rate retransmissions for incremental redundancy hybrid arq," *CoRR*, vol. abs/1207.0229, 2012.

[13] Y. Huang and B. Rao, "An analytical framework for heterogeneous partial feedback design in heterogeneous multicell ofdma networks," *Signal Processing, IEEE Transactions on*, vol. 61, no. 3, pp. 753–769, 2013.

[14] E. Dahlman, S. Parkvall, and J. Sköld, *4G LTE/LTE-Advanced for Mobile Broadband*, 1st ed. Academic Press, 2011.

[15] A. Larmo, M. Lindström, M. Meyer, G. Pelletier, J. Torsner, and H. Wiemann, "The LTE link-layer design," *IEEE Communications Magazine*, pp. 52–59, April 2009.

[16] N. G. Shephard, "From characteristic function to distribution function: A simple framework for the theory," *Econometric Theory*, no. 7, pp. 519–529, 1991.