Cross-Layer MAC Scheduling for Multiple Antenna Systems

Marc Realp¹ and Ana I. Pérez-Neira²

¹marc.realp@cttc.es; Telecommun. Technological Center of Catalonia (CTTC); Barcelona (Catalonia-Spain)
²anuskagps.tsc.upc.es; Technical University of Catalonia (UPC); Barcelona (Catalonia-Spain)

Abstract—In multiple access wireless channels mechanisms for an efficient management of the channel usage are of special interest. Particularly, in multiple access systems with multiple transmit and receive antennas, many antenna selection algorithms that exploit instantaneous channel information have been presented in the literature. On the other hand, MAC protocols control the access to the channel by evaluating average performance parameters such as throughput and delay. Both approaches, namely PHY layer and MAC layer approaches, present advantages and drawbacks. Far from being competitive, we argue that such approaches are complementary. Therefore, we present a unified view of both, PHY and MAC layer visions for a multiple input multiple output (MIMO) multiple access channel. First, we evaluate the impact in terms of throughput of a MAC scheduling strategy based on throughput maximization and compare it with a PHY scheduling strategy based on capacity maximization. Second, a MAC scheduler with QoS requirements is presented. The particularity of such MAC scheduler is that it combines PHY level instantaneous optimization with MAC level average maximization. We present mathematical expressions to evaluate the schedulers. We also present our design as an extension of Multiuser Diversity scheduler for the multi-user detector case.

I. INTRODUCTION

One of the main problems in multiple access communication systems is the design of mechanisms for an efficient management of the channel usage, for instance, in a multiple access system with multiple antennas, different algorithms might be designed in order to obtain the optimal set of simultaneous transmitting/receiving antennas that maximizes the total information-theoretic channel capacity. Such algorithms are based on the knowledge of the channel state between each pair of transmitter-receiver antennas. A very recent work [1] presents a nice overview on antenna selection algorithms including those in [2]. Other strategies [3] also use power allocation techniques to achieve other Quality of Service (QoS) requirements, as for instance, an Equal Bit Error Rate (BER) or a maximum BER for the worst transmission. However, although at the PHY layer the performance trade-offs as given by information theory are well understood, the impact of such mechanisms on network performance parameters as throughput or delay is much less understood.

On the other hand, the MAC layer manages the access to the channel dealing with traffic characteristics and QoS requirements of many terminals and/or services. Therefore, the MAC layer is forced to work with the long-term evaluation parameters throughput, delay and/or jitter. Unfortunately, average MAC strategies lead to not exploiting the benefits of the instantaneous PHY layer performance and, just recently, new MAC protocols that consider, for instance, the multipacket reception capability of the receiver [5],[4] have appeared.

In this paper, we present a unified view of both, PHY and MAC layer approaches for a multiple input multiple output (MIMO) multiple access channel. First, we present a MAC scheduler based on a Throughput Maximization Criterion that exploits the multipacket reception capability of the receiver and compare it with an scheduler based on a Capacity Maximization Criterion. Second, a MAC scheduler with QoS requirements is presented. The particularity of this MAC scheduler is that it deals with PHY level instantaneous QoS requirements (minimum instantaneous SNIR) as well as it aims to maximize MAC level long term system performance (throughput). In order to manage the channel usage more efficiently, a Cross-Layer (CL) design is presented. The idea behind the CL concept [6] is that in order to achieve optimal performance in one layer of the communication system, it is important that this layer is aware of some parameters or characteristics of the others. As a particular case of interest, we also present our CL design as an extension of the Multiuser Diversity schemes [11] to the multi-user receiver case.

II. MIMO SYSTEM MODEL

Let us consider a MIMO multiple access channel as in figure 1. Many terminals with multiple antennas wish to communicate with an other terminal, namely Receiver Terminal (RT), provided of $N$ antennas. Independent of the number of terminals, a total of $M$ antennas can transmit packets to the RT.

Time axis is divided into slots, each antenna can transmit up to one packet per slot and transmissions of different antennas are synchronized. Assume $N \geq M$ and that $k$ ($k \leq M$) simultaneous transmissions take place. The following is a model for this multiaccess antenna-array communication link

$$y = Hs + w$$

(1)

In (1), $s = (s_1, s_2, ..., s_M)^T$ is the transmitted symbol vector where $s_i$ is the transmitted symbol of the $i$th antenna which is transmitted with a power $P_i$, $H = (h_{11}, h_{12}, ..., h_{M1})$ is $N \times k$ flat-fading channel matrix where the scalar $h_{ij}$ represents the fading of the $i$th transmitting antenna at the $j$th receiver antenna. The entries of $H$ are independent and identically distributed complex Gaussian random variables with zero mean and unit variance. The vector $w$ is a complex-valued, background Gaussian noise with zero mean and variance $\sigma_w^2$.

We consider that the channel remains constant during one packet transmission and therefore, by means of a training sequence, the channel can be perfectly estimated at the RT. Previous to the transmission of packets, each terminal must have information on the antennas that it is allowed to use. The feedback channel consists of a $M$ bit Feedback Multicast Sequence (FMS) where the $i$th position $(FMS_i) = 1$ or not $(FMS_i = 0)$, Instantaneous perfect feedback channel is assumed. As we will see, information conveyed in the feedback channel might be determined in many different ways.

During transmission of packets, a Zero Forcing (ZF) beamforming is performed at the RT by means of the $N$ antennas. The ZF criterion provides multipacket reception capability in a low-complex closed-form solution. However, its performance can be drastically reduced in the case that not all the expected transmitting antennas do, in fact, transmit a packet. Hence, the ZF would try to null non-existing interferences. This problem is solved in [7] and references therein by adding, previous to the ZF, an additional stage that detects the $k$ active transmissions. We will assume the RT to use this additional stage and that active transmissions are detected without error.
For a given average SNR \( \gamma = \frac{P}{\sigma^2} \) at the reception, the post-detection instantaneous SNIR \( \gamma_i \) for the \( i \)th transmitting antenna can be defined as

\[
\gamma_i = \frac{\gamma}{\sum_j \alpha_j}
\]

where \( \alpha_i \) is a random variable that can be seen as a PHY layer quality measurement that accounts for both, channel fading and Multiple Access Interference (MAI) through receiver implementation. With the ZF receiver, \( \alpha_i \) is equal to \( \frac{\gamma}{\|H_i^H H_i\|_2} \) with \( \|H_i^H H_i\|_2 \) defining the \( i \)th element of the diagonal of \( (H_i^H H_i)^{-1} \). Notice that \( \alpha_i \) strongly depends on the number of columns \( k \) of \( H \). From (2), we can obtain a value for the BER, by using the approximate BER expression presented in [8] \( BER(\gamma_i) \approx C_1 \exp(-C_2 \gamma_i) \) where constants \( C_1 \) and \( C_2 \) are modulation dependant (for instance, for a QPSK, \( C_1 = 0.2 \) and \( C_2 = \frac{\gamma}{\pi} \)).

III. SYSTEM ANALYSIS

In a system as described in section 2, if the packet length is \( P_l \) and \( r \) is the number of correctable errors in a packet, the instantaneous Packet Success Rate (PSR) for the \( i \)th transmission or equivalently, the probability of successfully receive a packet from the \( i \)th transmitting antenna, is defined as

\[
PSR(\alpha_i) = \sum_{m=0}^{r} \binom{P_l}{m}(BER)^m(1-BER)^{P_l-m}
\]

In (3), we have considered the use of hard-decision decoding of perfect linear binary block codes. However, expression (3) could be modified to consider other codification techniques. The ZF receiver allows the reception of multiple packets simultaneously. The multipacket reception performance of the receiver can be modeled as the expected (or the average) number of successfully received packets when \( k \) transmissions take place simultaneously. This is

\[
C_k = \sum_{i=1}^{k} \int_{0}^{\infty} PSR(\alpha_i)p(\alpha_i)d\alpha_i
\]

where \( p(\alpha_i) \) is the p.d.f of the \( i \)th transmission. As mentioned before, access to the channel of the transmitting antennas is controlled by means of the FMS sequence. However, due to QoS requirements or/and due to data traffic issues, not all the antennas with the corresponding FMSi bit set to 1 will, in fact, transmit a packet. Therefore, we define \( K \) as the number of transmitting antennas allowed to access the channel and the probability \( p_{Kk} \) as the probability of having \( k \) active transmissions when access to the channel is given to \( K \) transmitting antennas. Throughput is computed as

\[
\eta_K = \sum_{k=0}^{K} p_{Kk} C_k
\]

It is worth mentioning that expression (5) is an average evaluation among channel statistics and transmissions that gives the average number of packets successfully received per slot. Particularly, (5) is a tool to evaluate the impact of PHY layer characteristics on the throughput of the system.

Since the performance of our system will be evaluated in terms of throughput at the MAC layer, in the following sections we will consider different MAC scheduling designs that aim to construct a FMS sequence with the optimal number of scheduled transmissions \( K_0 \) that maximizes throughput.

IV. MAC SCHEDULER BASED ON THROUGHPUT MAXIMIZATION CRITERION (TMC)

Let us consider a system with a very simple traffic model such that each transmitting antenna transmits (and retransmits) a packet with probability \( q \). Then, the probability \( p_{Kk} \) only depends on \( q \) in the binomial form

\[
p_{Kk} = \binom{K}{k} q^k (1-q)^{K-k}
\]

Throughput is computed as

\[
\eta_K = \sum_{k=0}^{K} \binom{K}{k} q^k (1-q)^{K-k} C_k
\]

Notice that when \( q = 1 \) expression (7) is equivalent to (4). Note that if no other scheduling criteria are considered the p.d.f. defined by \( p(\alpha_i) \) is the probability density function of a Chi-Square distributed random variable with \( n = 2(N-k+1) \) degrees of freedom.

Throughput maximization reduces to find \( K_0 \) such that

\[
K_0 = \arg\max_{K=1..M} \eta_K
\]

Clearly, \( K_0 \) is the optimal number of transmitting antennas that guarantees an optimal trade-off between the number of simultaneous transmissions and the amount of multiple access interference. How to allocate \( K_0 \) transmissions in the FMS sequence would depend on transmission priorities, fairness and other QoS parameters. In our case, transmitting antennas are selected randomly or equivalently, the FMS sequence is constructed by randomly allocating \( K_0 \) ones and \( M - K_0 \) zeros. Figure 2 shows how the optimal number of transmissions \( K_0 \) that maximize throughput depends on the packet probability \( q \). For example, maximum throughput of 10.2 packets/slot is obtained for \( K_0 = 11 \) when \( q = 1 \). We also observe that when \( q \leq 0.6 \), interference from other transmissions is very low and the background noise dominates. Therefore, \( K_0 = 15 \). In between, we obtain a curve for \( q = 0.8 \) with a maximum throughput when \( K_0 = 13 \).
A. Throughput Maximization Criterion Vs. Capacity Maximization Criterion

Instead of schedule transmissions to guarantee a maximum throughput, traditionally, transmitting antennas are selected in order to maximize the capacity of the link. As it is shown in [12], the capacity of a set of $K$ parallel AWGN channels after a $ZF$ receiver is

$$\text{Cap}_{ZF} = \frac{K}{\log_2(1 + \frac{P}{\sigma^2[H^TH]^{-1}})}$$  \hspace{1cm} (9)$$

The optimal scheduling strategy would be to select the set of $K_0$ transmissions that maximize (9). Notice that such Capacity Maximization Criterion (CMC) selects not only the number of simultaneous transmissions (as the TMC does) but a set of transmissions with cardinality $K_0$. Since CMC is based on an instantaneous maximization, needs to be performed once every time the channel changes. Consequently, CMC implies higher computational costs than TMC. Furthermore, no traffic considerations are taken in (9).

Figure 3 shows a comparison in terms of throughput between TMC and CMC. We observe that when $q = 1$, the CMC outperforms the TMC because maximizing the instantaneous capacity as defined in (9) also implies maximizing the number of successfully received packets per slot. However, when $q \leq 1$, it might happen that some of the antennas selected by the CMC do not transmit and hence, throughput is not maximized. Such traffic considerations are taken in the TMC leading to better throughput results.

V. MAC SCHEDULER UNDER QoS REQUIREMENTS

As mentioned before, previous to the transmission of data, the transmitting terminals must receive the Feedback Multicast Sequence (FMS) determining through which antennas each terminal is allowed transmitting terminals must receive the Feedback Multicast Sequence (FMS). Particularly, it implements an algorithm that guarantees all the scheduled transmissions to be over an instantaneous Signal to Noise-Interference Ratio (SNIR) or equivalently, all the scheduled transmissions to be below an instantaneous BER. Notice though, we do not consider either power allocation techniques or fairness or computational cost issues, only allocation of transmissions in time and space. We define the minimum required instantaneous BER as $BER_{th}$. Hence, we would like to schedule a set of transmissions defined in the FMS sequence such that

$$FMS_i = 1 \text{ for } 0 < i \leq M \text{ s.t. } BER_i \leq BER_{th}$$  \hspace{1cm} (10)$$

If an instantaneous SNIR threshold $\gamma_{th}$ corresponding to a $BER_{th}$ is defined, expression (10) can be rewritten as

$$FMS_i = 1 \text{ for } 0 < i \leq M \text{ s.t. } \gamma_i \geq \gamma_{th}$$  \hspace{1cm} (11)$$

Since, in general, there may exist many different algorithms that could be applied to construct the FMS sequence as in (11), let us define our IBC algorithm:

1) Set $FMS = \{1, ..., 1\}$
2) Obtain $H = [h_1, ..., h_N]$.
3) According to $H$, compute $\gamma_i$ as in (2) for $i \in M$.
4) If minimum $\gamma_i$ is over $\gamma_{th}$, Go to step 6.
5) Else, set $FMS_i = 0$ being $i$ the index for the minimum $\gamma_i$ and remove the corresponding column from $H$. Go to step 3.
6) Send $FMS$.

Notice that the IBC algorithm is an iterative algorithm which, by means of steps 3-4-5, discards the "poorest" transmission and hence, reduces the MAI of the other transmissions and improves their quality. Steps 3-4-5 are repeated until all transmissions are guaranteed to be over an instantaneous SNIR threshold. The main drawback of iterative algorithms is to analyze them analytically. For instance, the well known V-Blast iterative receiver was first proposed in 1996 and since then, its analytical analysis has been always carried out from a lower bound point of view [9]. Very recently though, some studies have appeared that propose analytical expressions for the outage probability of the V-Blast receiver. In our case, a similar approach to that in [10] could be followed, but such analysis is not straightforward and is out of the scope of this work. We keep such studies for further work. In this paper, we use simulated curves instead. Figure 4 shows the simulated Cumulative Distribution Function (c.d.f) of the minimum SNIR at step 4 of the IBC algorithm from the first iteration (left) to the last iteration (right).
\[ p_{kk}(\alpha_k) = \Pr(\alpha_1 < \alpha_k)\ldots \Pr(\alpha_{K-k} < \alpha_k)(1 - \Pr(\alpha_{K-k+1} < \alpha_k))\ldots (1 - \Pr(\alpha_K < \alpha_k)) \]  
\[ p_{kk}(\alpha_k) = \sum_{t=k}^{K} \Pr(\alpha_1 < \alpha_k)\ldots \Pr(\alpha_{K-t} < \alpha_k)(1 - \Pr(\alpha_{K-t+1} < \alpha_k))\ldots (1 - \Pr(\alpha_K < \alpha_k)) \left( \frac{t}{k} \right) a^{t-k} \]  
\[ C_k(K, \alpha_k) = \sum_{i=K-k+1}^{K} \frac{\int_{\alpha_i}^{\infty} PSR(\gamma_i)\mu(\gamma_i) d\gamma_i}{\int_{\alpha_i}^{\infty} \mu(\gamma_i) d\gamma_i} \]  
and maximization reduces to find \( \alpha_{th} \) such that
\[ \alpha_{th} = \arg \max_{\alpha_{th} \in [0, \infty)} \eta_{\alpha_{th}} \]  

Looking thoroughly into (16), one might observe some similarities with the Multiuser Diversity strategies used in single-user detector environments [11]. Multiuser Diversity is based on the idea that in fading channels, access to the channel should be given to the terminal (or antenna, in this case) whose SNR is greater. Particularly, (16) is an extension of Multiuser Diversity scheduling to the multi-user receiver case. Hence, similar to [11], it is found out that optimality is achieved when transmissions are scheduled according to their channel state (\( \alpha_i \)). Nevertheless, the main difference is that, in our case, many simultaneous transmissions might occur. Furthermore, optimization is performed by throughput maximization rather than capacity maximization. Hence, a CL approach. An example is shown in figure 5. The maximum throughput for the IBC algorithm is obtained when \( SNIR_0 = 10dB \) (\( \alpha_{th}^0 = 6dB \)). In that case, the average number of active transmissions per slot is 12.59.

D. Design of optimal set of transmissions

In many different situations we might find out that \( \alpha_{th} \) is a parameter fixed due to system requirements. For instance, to give upper layers QoS or to determine a Received Signal Strength Indicator (RSSI) which decides whether demodulate a signal or not. In any case, the aim is to obtain the optimal \( K_0 \) knowing the value of \( \alpha_{th} \). One can follow two different approaches. First, what we call the solution with priority, where a set of transmitting antennas with cardinality \( K_0 \) is chosen first, and then, IBC is performed among these selected antennas resulting on a FMS sequence with \( K < K_0 \) scheduled transmissions that all are over \( \alpha_{th} \). This solution, reduces the number of computations to be performed by the IBC algorithm because limits the set of possible transmissions to \( K_0 \) and allows to more easily control fairness and priority issues by deciding where to allocate the \( K_0 \) transmissions. In the case with priority, throughput is expressed as
\[ \eta_{K, \alpha_{th}} = \sum_{k=0}^{K} p_{kk}(\alpha_k)C_k(K, \alpha_k) \]  
and maximum throughput is obtained for
\[ K_0 = \arg \max_{K \in [1,M]} \eta_{K, \alpha_{th}} \]  

The second alternative, namely the solution without priority, is such that the IBC selects, among \( M \), all the transmissions that are over \( \alpha_{th} \). Then, the FMS sequence is constructed with the \( K \leq K_0 \) best transmissions that are over \( \alpha_{th} \). In this second alternative, at the expenses of higher computational costs instantaneous channel information is much more exploited. However, no priority among transmissions can be imposed because transmissions are strictly selected according to their \( \alpha_i \). Throughput and \( K_0 \) in the without priority case are
constructed with packets/slot is desired. Then, we observe that these QoS requirements for the multi-user detector case.

Second, a MAC scheduler with QoS requirements was presented. The particularity of such MAC scheduler is that it combines PHY level Maximization Criterion (TMC) and compared it with a PHY scheduler for the case with priority and without priority approaches are considered. For a given value of $SNIR_{th}$ the value $K_0$ for the case with priority and without priority that maximize throughput are obtained. Hence, for instance, $SNIR_{th} = 4dB$, $K_0$ is 11 for the case with priority and $K_0$ is 12 for the case without priority. As a matter of example, let us consider, that for QoS reasons, the minimum instantaneous $SNIR$ is set to $6dB$ and that a normalized throughput (throughput/q) over 10.2 packets/slot is desired. Then, we observe that these QoS requirements would be only achieved if the without priority approach was used.

It is also worth considering the effect on the throughput when $K_0$ is obtained independently of $\alpha_{th}$. Therefore, consider the case when $K_0$ is obtained following a TMC criteria only. However, the IBC algorithm is also performed among the $K_0$ transmissions that are computed according to (8). Hence, the final FMS sequence is constructed with $K \leq K_0$ scheduled transmissions. The impact on the throughput is evaluated in the simulations. For instance, in figure 6 we observe that if the IBC is performed with $SNIR_{th} = 10dB$, independent designs do not achieve maximum throughput because the number of selected transmissions is limited to $K_0$ that has been computed according to equation (8) instead of considering the IBC algorithm as in (19) or (21).

\[ \eta_{K_0, \alpha_{th}} = \sum_{k=0}^{K_0} p_{MK}(\alpha_{th})C_k(M, \alpha_{th}) + \sum_{k=K_0+1}^{M} p_{MK}(\alpha_{th})C_k(M, \alpha_{th}) \]

\[ K_0 = \arg \max_{K=1..M} \eta_{K_0, \alpha_{th}} \quad (21) \]

In either (19) or (21), $K_0$ depends on $\alpha_{th}$ which is fixed. In figure 6, we can observe the throughput, as a function of $\alpha_{th}$, when priority and without priority approaches are considered. For a given value of $SNIR_{th}$ the value $K_0$ for the case with priority and without priority that maximize throughput are obtained. Hence, for instance, $SNIR_{th} = 4dB$, $K_0$ is 11 for the case with priority and $K_0$ is 12 for the case without priority. As a matter of example, let us consider, that for QoS reasons, the minimum instantaneous $SNIR$ is set to $6dB$ and that a normalized throughput (throughput/q) over 10.2 packets/slot is desired. Then, we observe that these QoS requirements would be only achieved if the without priority approach was used.

It is also worth considering the effect on the throughput when $K_0$ is obtained independently of $\alpha_{th}$. Therefore, consider the case when $K_0$ is obtained following a TMC criteria only. However, the IBC algorithm is also performed among the $K_0$ transmissions that are computed according to (8). Hence, the final FMS sequence is constructed with $K \leq K_0$ scheduled transmissions. The impact on the throughput is evaluated in the simulations. For instance, in figure 6 we observe that if the IBC is performed with $SNIR_{th} = 10dB$, independent designs do not achieve maximum throughput because the number of selected transmissions is limited to $K_0$ that has been computed according to equation (8) instead of considering the IBC algorithm as in (19) or (21).

VI. CONCLUSIONS

In this paper, we presented a unified view of both, PHY and MAC layer approaches for a multiple input multiple output (MIMO) multiple access channel. First, we evaluated the impact in terms of throughput of a MAC scheduling strategy based on Throughput Maximization Criterion (TMC) and compared it with a PHY scheduling strategy based on Capacity Maximization Criterion (CMC). Second, a MAC scheduler with QoS requirements was presented. The particularity of such MAC scheduler is that it combines PHY level instantaneous optimization with MAC level average maximization. We presented mathematical expressions to evaluate the schedulers. We also presented our design as an extension of Multis-user Diversity scheduler for the multi-user detector case.

To main issues are foreseen in future work. First, the need to develop analytical expressions for the statistics of $\alpha$ and $C_k$ to fully exploit the idea of CL design. And second, it is envisaged to continue this work accounting for more complex traffic models and hence, considering queueing theory issues.

VII. ACKNOWLEDGEMENTS

This work is supported by the Spanish government under TIC2002-04594-C02 and FIT-070000-2003-257 and by the IST European Commission under FP6-507572 and jointly financed by FEDER.

REFERENCES