Enhanced Multicarrier Techniques for Professional Ad-Hoc and Cell-Based Communications

(EMPhAtiC)

Document Number D5.3

Cross-layer optimization of RRM with Physical and Application layers in PMR networks

Contractual date of delivery to the CEC:	31 December 2014
Actual date of delivery to the CEC:	28 January 2015
Project Number and Acronym:	318362 EMPhAtiC
Editor:	Stephan Pfletschinger (CTTC)
Authors:	Dimitris Tsolkas (CTI), Marius Caus (CTTC) Yao Cheng (ITU), Martin Haardt (ITU), Antonio Cipriano (TCS), Luxmiram Vijayandran (TCS), Mylene Pischella (CNAM)
Participants:	CTI, CNAM, TCS, CTTC, ITU
Workpackage:	WP5
Security:	Public (PU)
Nature:	Report
Version:	1.1
Total Number of Pages:	42

Abstract:

This report presents new progress on cross-layer techniques and Radio Resource Management (RRM) algorithms for broadband PMR networks. Cross-layer adaptation in cell-based PMR communications is first investigated, and a cross-layer scheme is devised to offer an increased number of PMR connections with satisfactory Quality of Experience (QoE), which is crucial in crisis scenarios. Taking into account the inter-cluster interference, an enhanced version of the Distributed bUffer Stabilization RRM algorithm (DUST) is proposed for PMR networks. In addition, with a focus on multi-user MISO downlink scenarios, an efficient margin adaptive scheduling algorithm is tailored for FBMC, contributing to a solution to the challenging joint design of transmit and receive processing, the channel assignment as well as the power allocation.

Table of Contents

1. Introduction

This report considers cross-layer techniques for the optimization of Radio Resource Management (RRM) algorithms in PMR networks and places special emphasis on the interactions between the physical and the application layers. Techniques and results on cross-layer RRM for multiple clusters which have been reported in deliverable D5.2 have been extended to consider cross-cluster interference.

The first section considers cross-layer adaptation in cell-based PMR communications: The interplay between the source coding of critical voice communication, which is located in the application layer, and the number of simultaneous calls which can be accommodated by the system is investigated and an algorithm inspired by the design principle *continuous flow modelling* has been developed.

Section 3 investigates RRM for multiple clusters considering Cyclic Prefix Orthogonal Frequency Division Multiplex (CP-OFDM), filter bank multicarrier (FBMC) and *perfect modulation (PM)*, which is an idealized modulation that does not cause adjacent interference. The design objective is a stable queue for all users and the minimization of the energy consumption taking into account non-negligible inter-cluster interference. An extension of the backpressure algorithm is developed and the achieved performance is evaluated for different scenarios, in particular the differences between OFDM and FBMC are assessed. Signalling between different clusters for interference coordination has been considered.

The last section addresses the Multi-User Single Input Multiple Output (SU-SIMO) margin adaptive problem for the FBMC modulation based on the offset QAM (OQAM), referred to as FBMC. The analysis conducted in that section reveals that due to the specific FBMC transmission format the system has to deal with inter-user, inter-symbol and inter-carrier interference to solve the resource allocation problem. To manage the interference, the successive channel allocation method, originally proposed for the CP-OFDM, has been modified by organizing subcarriers in groups and assigning subcarriers to users in a block-wise fashion. Following this strategy and making some approximations we demonstrate that it is possible to pose a single problem that is valid for OFDM and FBMC. In this case, numerical results show that FBMC is more energy-efficient than OFDM because it requires less power to transmit the same information rate.

2. Increasing PMR capacity through encoding rate adaptation

2.1 Adopted PMR scenario

The adopted scenario refers to cell-based PMR communications, where in each cell, time and frequency synchronized transmissions (including potentially direct transmissions through Direct Mode Operation - DMO) take place under the control of a base station (BS). It is assumed that network dimensioning/planning procedures have taken place, while the BSs are inter-connected through a backbone infrastructure. The Hand Helds (HHs) and the Mobile Stations (MS) are randomly deployed in each cell, while the transmissions of a dynamic set of HHs/MSs may be critical referring to communications for Public Protection and Disaster Relief (PPDR).

This scenario may model situations such as a car accident, train crash, traffic jam, an earthquake event etc., where coordination among police officers/firemen is needed at a specific area inside a cell while all citizens try to communicate especially the first moments after the event (Fig. 2-1). Practically, in such situations specific set of HHs/MSs (e.g., HHs of police officers) requires scheduling priority, since their transmissions carry vital information. The BSs must serve the critical traffic with priority and guarantee the quality requirements of the non-critical traffic as well. An efficient way to deal with this problem has been proposed in D5.2 [1], where an enhanced proportional fair RRM algorithm has been proposed. However, in such critical scenarios, even if the critical traffic is served with priority, it increases rapidly requiring solutions for improving the capacity of the critical traffic. To this end, here we focus on critical voice connections and propose a cross-layer mechanism for optimizing RRM by adjusting the encoding rate in application layer. The proposed optimization targets at maximizing the active HHs in the system, with the cost of a controlled degradation on end users' quality of experience (QoE).

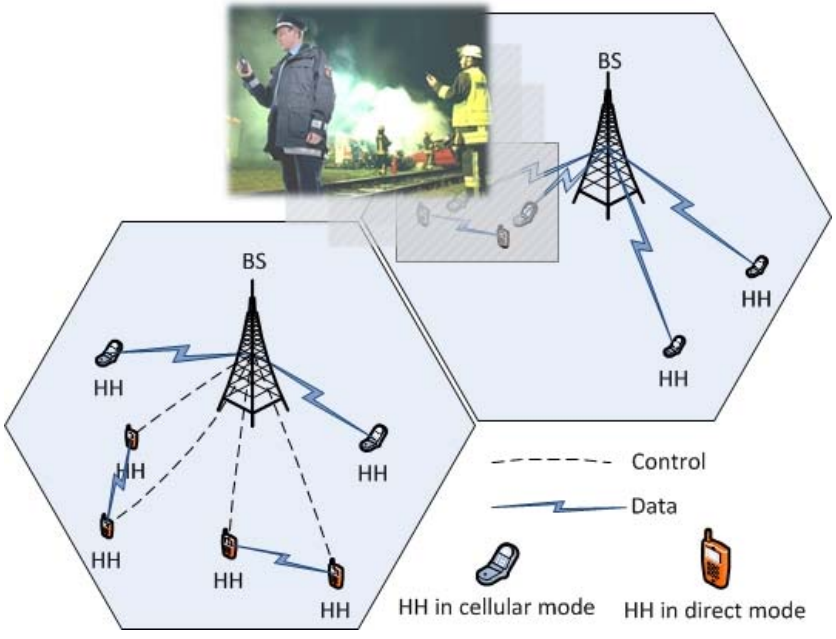


Figure 2-1: Illustration of the cell-based PMR scenario

2.2 The Proposed cross-layer scheme

The proposed cross-layer scheme is split into two parts, namely the BS part and the HH part, residing at the BS and each HH, respectively (Fig. 2-2). Regarding the first part, the BS starts by collecting all the required information regarding the performance status of each one of its active connections. Subsequently, the BS uses the collected information to run a decision algorithm for adjusting the encoding mode in application layer. This decision is taken separately by the HH part after BS's order and aims at counterbalancing the number of PMR voice connections in the system and the achievable QoE at HHs. This is a critical issue in crisis scenarios, since the establishment of a voice connection is of high importance, while the achievable QoE level can be more relaxed. Our main focus is on proposing a cross layer decision algorithm for the BS part. This algorithm takes into account the packet timeout rate and the mean delay of each connection and decides on the adaptation of media encoding rates towards an increased number of PMR connections with acceptable QoE in the system.

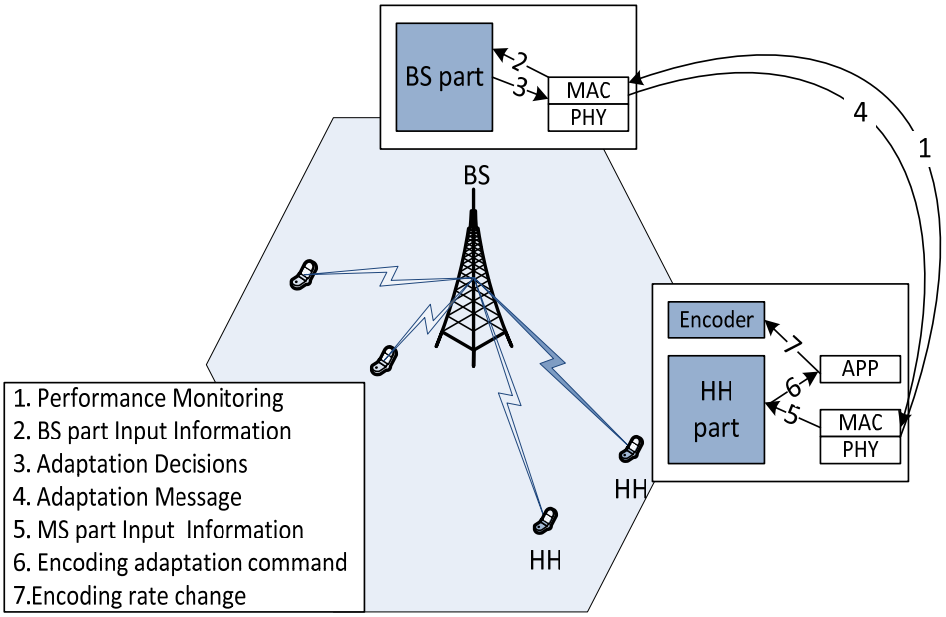


Figure 2-2: The proposed cross layer scheme

Assume N active PMR connections in a cell, the parameters that are used by the decision algorithm are the following:

- $R_{timeout,i}$ is the packet timeout rate of the ith connection, i.e., the percentage of packets that were lost due to deadline expiration (packet delay has exceeded maximum acceptable delay).
- S_i is the mean delay of the ith connection.
- $\mu_{i,d}$ is the encoding mode (of the D available modes) for the ith connection, $d = (1,2,\dots D)$.
- $\varepsilon_{max,i}, S_{max,i}$ are the maximum tolerable timeout rate and maximum acceptable delay of the ith connection, respectively,

- β_i is a threshold indicating a very low delay ratio: ($\frac{S_i}{S_{max,i}}$).

The algorithm is initiated at regular time instants and takes proper adaptation decisions based on the packet timeout rate and delay of each connection. The actions to be taken by the decision algorithm are as follows:

if $R_{timeout,i} > \varepsilon_{max,i}$, a high number of packet timeouts is observed in the system due to high traffic load. The BS part instructs the HH part for a media encoding rate reduction in order to moderate timeouts.

To achieve an efficient performance under all possible conditions, the algorithm makes adaptation decisions also when the QoS and QoE for a specific connection is improved.

if $\frac{S_i}{S_{max,i}} < \beta_i$, the timeout rate decreases significantly, and BS part instructs for a media encoding rate increase.

The encoding rate reduction policy is similar to the slow start TCP congestion control mechanism, meaning that when $R_{timeout,i} > \varepsilon_{max,i}$, the encoding rate is reduced to the lower one and while $R_{timeout,i} \leq \varepsilon_{max,i}$, and $\frac{S_i}{S_{max,i}} < \beta_i$ it grows linearly. Fig. 2-3 illustrates the proposed decision algorithm.

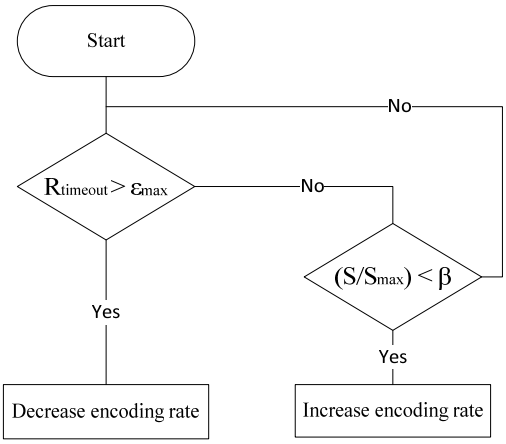


Figure 2-3: Proposed decision algorithm at the BS part of the cross-layer scheme

2.3 Theoretic analysis using Continuous Flow Modelling

In this section the theoretic analysis of the proposed cross-layer scheme is presented. With the use of Continuous Flow Modelling (CFM), the system is considered as a ‘‘fluid’’ queue with inflow and outflow rates representing its traffic generation and service rates, respectively. Each data source (of the N active PMR connections) is modeled as a Markov chain, from the steady-state of which the optimal adaptation thresholds of the cross-layer mechanism are derived.

2.3.1 General Continuous Flow Model

The operation of the system that employs the proposed scheme can be modelled as a CFM [2, 3]. CFMs are widely used in the relevant bibliography for network modelling and analysis

[4–6]. CFM representation of communication networks' traffic and elements is justified from the fact that in a high-speed packet switching network a packet may be considered as a water molecule with a virtually infinitesimal effect on the entire flow. Thus, the adoption of a CFM can significantly reduce the complexity of the representation of the system that employs the proposed scheme, as compared to a standard representation at packet level. The CFM used here consists of a ''fluid'' queue whose inflow and outflow processes are characterized by flow rates, while its content is defined by the volume of the stored fluid. The size of the queue is finite. Thus, in case the queue is full, the excess flow cannot be absorbed, leading to overflow. The basic storage unit of this model is a symbol. Assume that the N active PMR connections in a cell are identical and each one is fed by a data source producing a continuous data flow with maximum tolerable delay S_{max}. The basic parameters of the CFM are the following:

1. $a(t)$inflow rate in symbols/second
2. $\varphi(t)$ Constant service rate in symbols/s (the PHY layer time frame is considered to have a duration of T_f seconds and serve c symbols).
3. $C = c \cdot \frac{S_{max}}{T_f}$ Queue size in symbols. Its value is such that overflow occurs when the delay of a queued symbol exceeds the maximum tolerable value S_{max}.
4. $x(t)$: Queue load in symbols.
5. $\eta(t)$: Outflow rate in symbols/second.
6. $\lambda(t)$: Overflow rate in symbols/second.

Let us define for each data flow: a variable encoding rate equal to $\mu_{i,d}(t)$ bits/second where $d = (1,2, \dots D)$ and D is the number of discrete encoding modes, and a modulation order $b_i(t)$ bits/symbol. The inflow rate of each source is:

$a_i(t) = \frac{\mu_{i,d}(t)}{b_i(t)}$, while the total inflow rate:

$$a(t) = \sum_{i=1}^{N} a_i(t) \tag{2.1}$$

The buffer load $x(t)$ is described by the differential equation:

$$\frac{dx(t)}{dt^+} = \begin{Bmatrix} 0, & if\ x(t) = 0\ and\ a(t) < \varphi(t) \\ 0, & if\ x(t) = C\ and\ a(t) \geq \varphi(t) \\ a(t) - \varphi(t), & else \end{Bmatrix} \tag{2.2}$$

The total outflow rate $\eta(t)$ is defined as:

$$\eta(t) = \begin{Bmatrix} \varphi(t), & if\ x(t) > 0 \\ a(t), & if\ x(t) = 0 \end{Bmatrix} \tag{2.3}$$

The total overflow rate $\lambda(t)$ is defined as:

$$\lambda(t) = \begin{Bmatrix} a(t) - \varphi(t), & if\ x(t) = C\ and\ a(t) \geq \varphi(t) \\ 0, & if\ x(t) = 0 \end{Bmatrix} \tag{2.4}$$

The overflow rate $\lambda_i(t)$ of each source is proportional to the percentage of the source inflow rate with respect to the total system inflow rate.

$$\lambda_i(t) = \frac{a_i(t)}{a(t)} \cdot \lambda(t) \quad (2.5)$$

where

$$\lambda(t) = \sum_{i=1}^{N} \lambda_i(t) \quad (2.6)$$

The total overflow volume during the time interval $[0, T]$ is defined as follows:

$$L = \int_0^T \lambda(\tau) d\tau \quad (2.7)$$

while the overflow of the ith data source is

$$L_i = \int_0^T \lambda_i(\tau) d\tau \quad (2.8)$$

Thus, the total overflow rate during the time interval $[0, T]$ is

$$R_{overflow} = \frac{L}{\int_0^T a(\tau) d\tau} \quad (2.9)$$

while the overflow rate of the ith source is

$$R_{overflow,i} = \frac{L}{\int_0^T a_i(\tau) d\tau} \quad (2.10)$$

This corresponds to the connection packet timeout rate $R_{timeout,i}$, which is the key value used in the proposed decision algorithm.

The data successfully served by the queue (i.e., scheduled for transmission) are transmitted over the wireless medium. The data loss due to channel conditions of the connection i is as follows:

$$\zeta_i(t) = \eta_i(t) \cdot (1 - (1 - BER_i)^{b_i(t)}) \quad (2.11)$$

where BER_i is the BER for the ith connection. Consequently, the total loss due to the wireless channel is:

$$\zeta(t) = \sum_{i=1}^{N} \zeta_i(t) \quad (2.12)$$

During the time interval $[0, T]$ is:

$$E = \int_0^T \zeta(\tau) d\tau \quad (2.13)$$

And, thus, the total loss rate due to the wireless channel is

$$R_{err} = \frac{E}{\int_0^T \eta(\tau) d\tau} \quad (2.14)$$

Finally, the total loss rate due to timeouts (overflow) and the wireless channel is

$$R_{loss} = R_{overflow} + R_{err} \cdot (1 - R_{overflow}) \quad (2.15)$$

The proposed algorithm changes its decision in specific time intervals. The BS monitors the network for a period of time (a number of subframes) and then decides on changing or not the encoding rate. To this end, next we provide the calculation of the loss rate and the delay in discrete time intervals.

2.3.2 Calculation of the timeout rate and delay in discrete time intervals

Let us consider a system that consists of a single data source, the fluid queue and the wireless medium. The source generates data with an encoding rate $\mu_{i,d}(t)$ bits/s, where $d = (1,2,\dots D)$. It is assumed that during the observation interval $[0,T]$ the functions $a(t)$ and $\varphi(t)$ are piecewise constant with a finite number of ''jumps''. In this case, the function $x(t)$ is piecewise linear, while the functions $\eta(t)$ and $\lambda(t)$ are piecewise constant as well. The decision algorithm has to determine the appropriate values of $\mu_{i,d}(t)$, thus to adapt the value of $a(t)$ every ΔT seconds (where ΔT is a constant). If the quality of the wireless channel is assumed to be constant during a time frame, then $[0,T] = T_f$. Thus, the $[0,T]$ interval is divided into $K = \frac{T}{T_f}$ subintervals of T_f seconds: $[T_k, T_{k+1})$, where $T_0 = 0$ and $T_K = T$, during each of which the value of the function $a(t)$ is constant. The value of $\varphi(t)$ is by definition constant during all these time intervals.

2.3.2.1 Calculation of the timeout rate

Let a_k, φ_k, η_k be the values of $a(t)$, $\varphi(t)$, $\eta(t)$, respectively, during the time interval $[T_k, T_{k+1})$. Similarly, $x_k = x(T_k)$ is the buffer load at T_k, and $L_k = \int_{T_k}^{T_{k+1}} \lambda(\tau) d\tau$ is the overflow volume during $[T_k, T_{k+1})$.

The buffer load at T_{k+1} is: $x_{k+1} = \min\{\max\{x_k + [a_k - \varphi_k] \cdot T_f, 0\}, C\}$, while the overflow volume during $[T_k, T_{k+1})$, is $L_k = \begin{cases} [a_k - \varphi_k] + x_k - C, & if\ x_{k+1} = C \\ 0, & else \end{cases}$.

Thus, the overflow rate is:

$$R_{overflow,k} = \frac{L_k}{a_k \cdot T_f} \tag{2.16}$$

Additionally the loss rate due to the wireless channel is:

$$R_{err,k} = 1 - (1 - BER_k)^{b_k} \tag{2.17}$$

Consequently,

$$R_{loss,k} = R_{overflow,k} + R_{err,k} \cdot (1 - R_{overflow,k}) \tag{2.18}$$

2.3.2.2 Calculation of the mean delay

The value of the data source mean delay, S_k, during $[T_k, T_{k+1})$ is derived from the mean buffer load (given the constant service rate of the queue). More specifically, in the time interval $[T_k, T_{k+1})$ the mean buffer load x_k is:

$$x_k = \frac{1}{T_f}\int_{T_k}^{T_{k+1}} x(\tau)d\tau \quad (2.19)$$

Thus, for the mean delay $S_k$we have:

$$S_k = x_k \cdot \frac{T_f}{c} \quad (2.20)$$

2.4 Derivation of the optimal adaptation thresholds

Voice is one of the dominant services in crisis scenarios. Taking this into account the main target of the proposed scheme is to adapt the voice encoding rate of HHs at application layer under controlled degradation at the end-user's QoE. QoE is a strongly subjective term and also one of the dominant factors for assessing a provided service. The most common measure of the QoE is the Mean Opinion Score (MOS) scale recommended in [7]. The MOS ranges from 1 to 5, with 5 representing the best quality, and is commonly produced by averaging the results of a subjective test, where end-users are called under a controlled environment to rate their experience with a provided service. However, the subjective methodologies (e.g., use of questionnaires) are cost-demanding and practically inapplicable for real-time monitoring of the service performance.

2.4.1 Objective estimation of QoE

Different objective methods have been proposed to measure the speech quality. These methods can be classified into intrusive and non-intrusive methods. Intrusive methods, such as the Perceptual Speech Quality Measure (PSQM) and the PESQ (Perceptual Evaluation of Speech Quality), estimate the distortion of a reference signal that travels through a network by mapping the quality deterioration of the received signal to MOS values. However, the need for a reference signal is a drawback for using intrusive methods for QoE monitoring. To this end, non-intrusive methods have been defined such as the E-model and the ITU P.563 [8]. Since the ITU P.563 method has increased computational requirements, making it inappropriate for monitoring in real-time basis, we adopt the easier to be applied E-model, described in the next section.

2.4.1.1 The E-model

The E-model [9] has been proposed by the ITU-T for measuring objectively the MOS of voice communications. It is a parametric model that takes into account a variety of transmission impairments producing the so-called Transmission Rating factor (R factor) scaling from 0 (worst) to 100 (best). Then a mathematic formula is used to translate this to MOS values. In the case of the baseline scenario where no network or equipment impairments exist, the R factor is given by:

$$R = R_0 = 93.2 \quad (2.21)$$

However, delays and signal impairments are involved in a practical scenario and hence the R factor is given by:

$$R = R_0 - I_s - I_d - I_{ef} + A \quad (2.22)$$

where:

I_s : the impairments that are generated during the voice traveling,

I_d : the delays introduced from end-to-end signal traveling,

I_{ef} : the impairments introduced by the equipment,

A : allows for compensation of impairment factors when there are other advantages of access to the user (Advantage Factor). It describes the tolerance of a user due to a certain advantage that he/she enjoys, e.g., not paying for the service or being mobile. Typical value for cellular networks: $A = 10$.

Focusing on parameters which depend on the wireless part of the communication it holds that [10]:

$$I_d = 0.024d + 0.11(d - 177.3)H(d - 177.3) \quad (2.23)$$

where:

$$H(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

and d is the average packet delivery delay composed by two parts, the scheduling delay, denoted by d_s, plus the roundtrip delay, denoted by d_τ i.e., $d = d_s + d_\tau$. Low delay values below 150ms have very little impact on call quality or interactivity. As delay values continue to increase above 150ms call quality further degrades, with delays above 400ms making a duplex call extremely difficult due to loss of interactivity. Regarding the proposed scheme the value d_s for a connection i corresponds to the S_i value defined in the CFM description above.

Also, the packet loss rate, R_{loss}, affects the parameter I_{ef} as follows [11]:

$$I_{ef} = I_e + (95 - I_e) \cdot \frac{100 \cdot R_{loss}}{\frac{100 \cdot R_{loss}}{busrtR} + B} \quad (2.24)$$

where I_e is an encoding mode depended equipment impairment factor obtained from [12] and shown in Table 2-1, B is a codec specific loss robustness factor and $busrtR$ is the quotient of the average burst length and is dependent upon the theoretical burst length under random loss conditions. Considering that a single frame is sent per packet giving independent losses we set $busrtR$ = 1. For AMR B has not been defined for all modes. However, it is dependent upon the inter-packet dependencies and the packet loss concealment scheme, and so since all AMR codec modes have a similar structure we utilize the value defined for the 12.2kbps mode of $B = 10$ for all codec modes [12].

The R factor can be transformed to MOS values to retrieve results comparable with results provided by subjective methods. The transformation formula is as follows:

$$MOS = \begin{cases} 1, \ if \ R < 0, \\ 1 + 0.035R + R(R - 60)(100 - R) \cdot 7 \cdot 10^{-6}, \ if \ 0 \leq R \leq 100, \\ 4.5, \ if \ R > 100 \, . \end{cases} \quad (2.25)$$

The proposed scheme increases the PMR capacity in respect to a specific R (or MOS) threshold, which represent the minimum acceptable QoE level for a PMR voice connection. Based on this threshold the $\varepsilon_{max,i}$ and β_i thresholds are calculated as described in the next subsection.

Table 2-1: Encoding modes

Encoding Mode	Application layer Encoding Rate (Kbps)	I_{ef}
1	12.2	5
2	10.2	9
3	7.95	15
4	7.4	16
5	6,7	20
6	5,9	23
7	5,15	27
8	4,75	29

2.4.2 Thresholds calculation

We first depict in Fig. 2-4 the achievable MOS values for different $R_{timeout,i}$,and delay values. For minimum acceptable MOS=3 and MOS=2.5, the threshold values that can be used by the proposed scheme are depicted in Fig. 2-5, and Fig. 2-6, respectively.

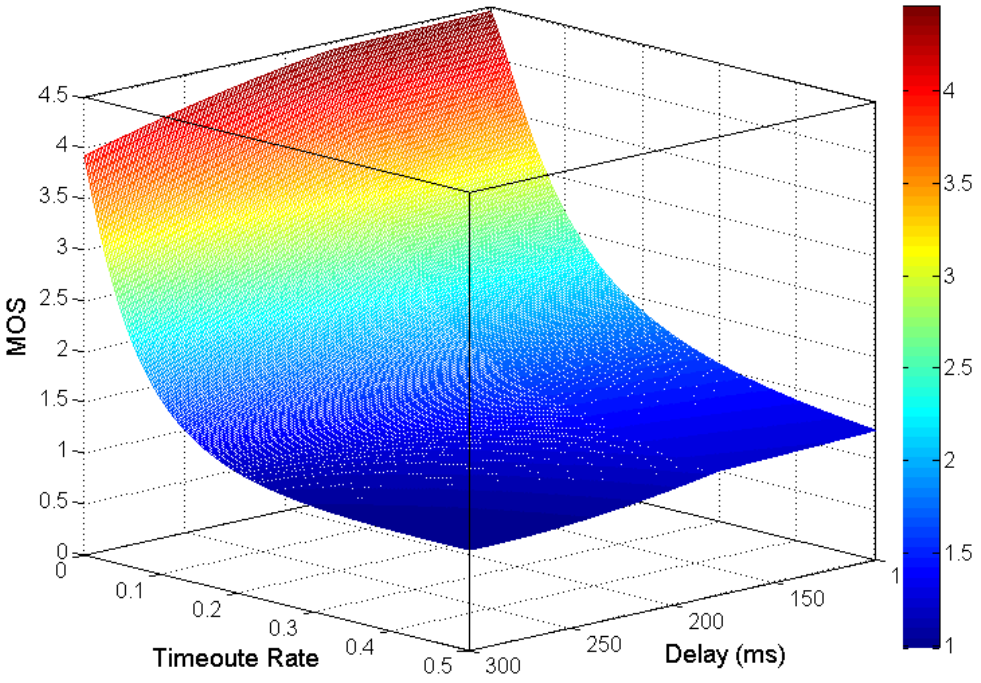


Fig. 2-4: MOS for various delay and timeout rate values

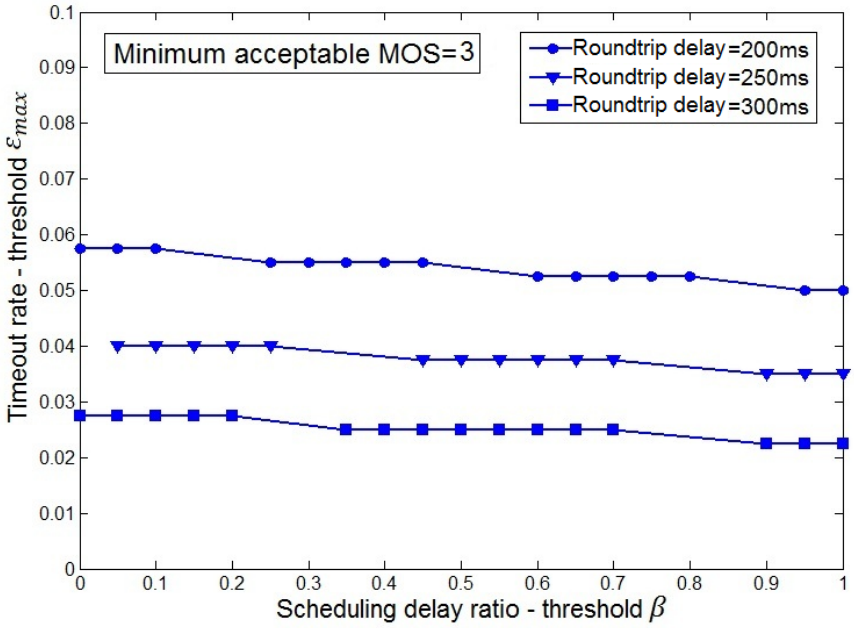


Fig. 2-5: Threshold values that can be used by the proposed scheme for minimum acceptable MOS=3

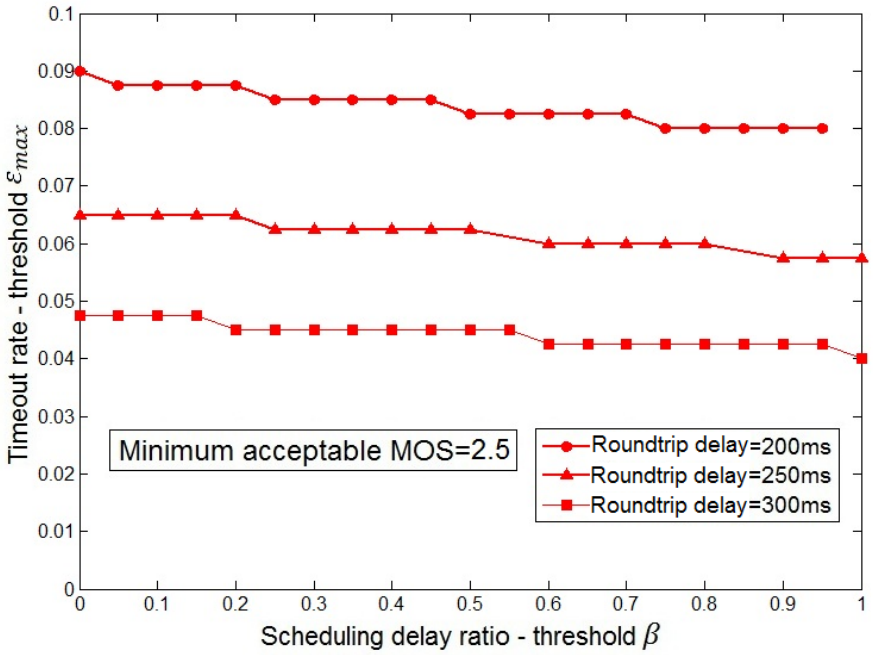


Fig. 2-6: Threshold values that can be used by the proposed scheme for minimum acceptable MOS=2.5

2.5 Performance evaluation

We evaluated the proposed scheme when it is applied on top of OFDMA and FBMC PHY, using the parameters depicted in Table 2-2. The adaptation decision is taken every 10ms, while the timeout and delay thresholds $\varepsilon_{max} = 0.05$, $\beta = 0.3$, are used, respectively. Based on the analysis in previous subsection, these thresholds guarantees QoE of about MOS=2.5 for a

roundtrip delay 300ms. Since the proposed scheme is based on the timeout rate, meaning on the rate of packets that are not scheduled for transmission due to high traffic conditions, the key difference between the OFDMA and FBMC scenario lies on the channel service rate. Using the same bandwidth for both the scenarios, FBMC provides higher service rate due to the absence of CP, i.e., the same amount of data can be sent in shorter frame duration [20].

Table 2-2: Evaluation parameters

Parameter	value
Bandwidth	5MHz
frame duration	10ms (OFDMA) 9.33ms (FBMC)
Slots/subframe	2
Data Symbols/slot	7
Data Bits/Symbol	6
Number of RB	25
Data subcarriers/RB	12
Roundtrip delay	300ms
BER	10e-7

Fig. 2-7 depicts the timeout rate in the network when the proposed scheme is disabled ("No adaptation") and enabled ("Proposed adaptation") for both the OFDMA and FBMC cases. The first observation regarding Fig. 2-7 is that for the conventional case, with no rate adaptation, the FBMC leads to about 14% more PMR capacity comparing to OFDMA scenario (FBMC can serve 126 HHs while OFDMA 110 HHs). By enabling the proposed scheme, the total gain for the FBMC and OFDMA scenario is 72% and 50%, respectively.

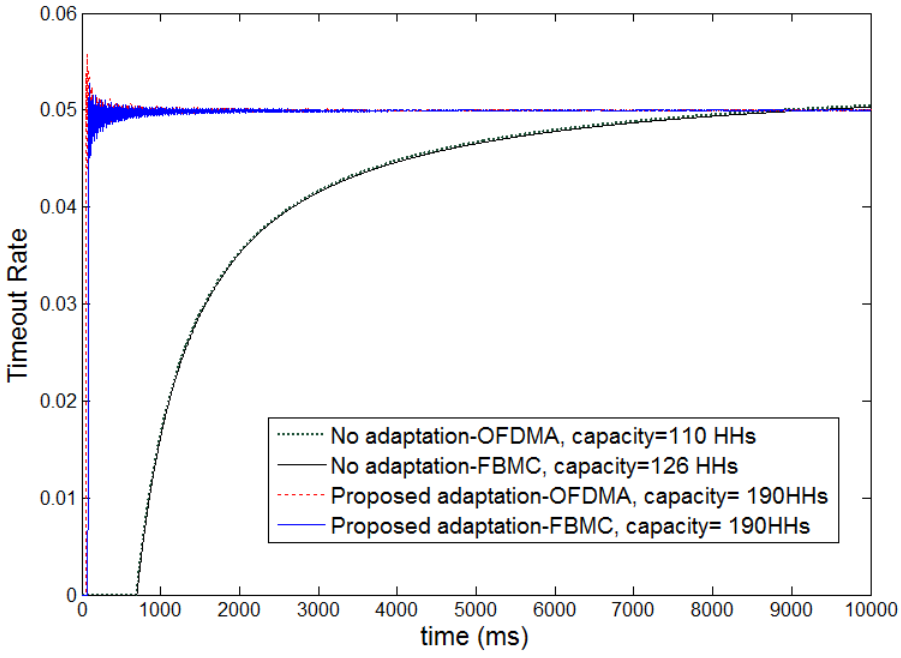


Fig. 2-7: Comparison of the proposed scheme with the conventional scenario (no adaptation) for the OFDMA and FBMC cases

Here is the performance summary of the proposed cross-layer scheme:

- ✓ To involve application layer performance metrics in lower layer procedures, reliable monitoring of these metrics is required. In the case of the QoE metric, the E-model provides an efficient and easy to apply formula, which can be exploited by the BS of the PMR network.
- ✓ Application layer rate adaptation provides sufficient capacity gain for the critical transmissions in a PMR network, using as a criterion the end-user's QoE.
- ✓ The performance of the proposed scheme has negligible variations for the different PHY schemes (OFDMA and FBMC), since it is based on scheduling and application layer procedures. Cross layer schemes that involve physical layer parameters, e.g., the SINR, are expected to lead to higher performance differentiations. Such a scheme is provided in the next section.

3. Evolution of DUST algorithm using interference information

This work is an evolution of the Distributed bUffer Stabilization RRM (DUST) algorithm described in D5.2, Section 7. In particular, the main objective is to address large inter-cluster interference by taking advantage of the aggregate interference cross-layer information $I(t)$ obtained from the physical layer, assuming that it has sensing capability.

3.1 Evolution compared to D5.2

In D5.2, we addressed the resource allocation problem for an asynchronous clusterised PMR ad-hoc network that uses the spectral gaps in between other existing communication networks, , as depicted in Figure 3-1.

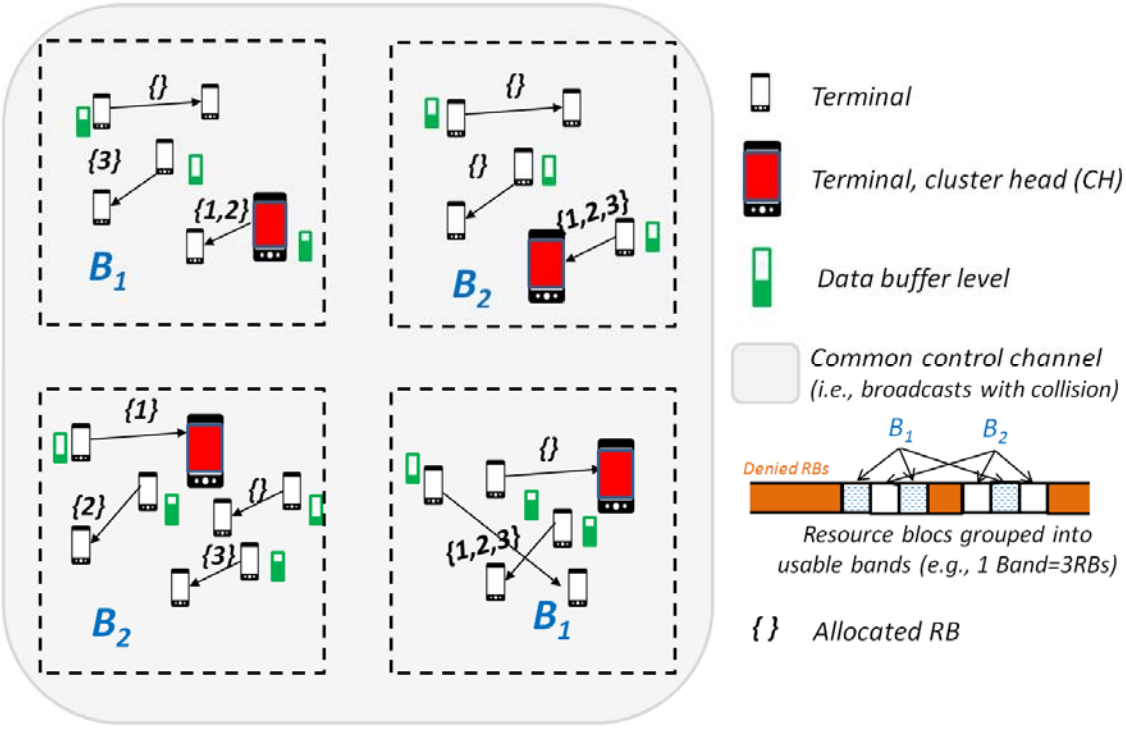


Figure 3-1: Clusterised network scenario

For such heterogeneous networks, besides the direct channel interference, the adjacent ones also need to be considered. As such, we considered both CP-OFDM and FBMC modulations, as well as the so-called Perfect modulation (i.e., no adjacent interference). We sought to stabilize each user data queue while minimizing their average transmission energy. In order to cope with the long term stochastic nature of the problem and the inter-cluster interference in a distributed way an algorithm was proposed adapted from the centralized backpressure approach. It was shown that the algorithm allows to trade between energy efficiency and the buffer delay. Compared to the classical centralized backpressure approach [13], a special dynamic weight was introduced to balance the independent cluster traffic requirements (i.e.,

D5.2-Eq. 7.7), such that the objective function to be maximized at every instance was given by, D5.2-Eq 7.9, i.e., optimize P to maximize

$$\sum_{m\in M_n} \alpha_n(t)\zeta_m(t)\mu_m(t) - \eta_n(t)p_m(t) \tag{3-1}$$

where M_n is the set of all links affected to cluster n, $\mu_{m(t)}$ and $p_{m(t)}$ are the theoretical Shannon capacity and the transmit power, respectively, $\alpha_n(t)$, $\zeta_n(t)$ and $\eta_n(t)$ are dynamic variables controlled by the algorithm. The proposed algorithm was shown to perform well when the inter-cluster interference is not high. However, when the inter-cluster interference is high it may lead to destructive competition. In other words, if a given cluster does not succeed to get the desired throughput (thus steadily increasing data queue), it will keep on increasing its transmission power until reaching the maximum power. Then, links will most probably always interfere each other (especially those near the edge) without any of them achieving there desired rate. Thus, we propose to enhance the DUST algorithm even when the inter-cluster interference can be high (depending also on the transmit power). To do so, we use a key statement that was proved in [14]. It is stated that when the cross-interference is higher than some level, the optimal allocation is an FDMA-type allocation. In other words, transmission should be separated by any mean (e.g., frequency or time, FDMA, TDMA) such that there is no interference among different clusters. "*One intuitive way to impose orthogonal communications between independent clusters' links is to randomly turn-off some of the available Resource Blocks (RBs) for a random period of time"*. The above intuition is the base line of the entire enhancement as compared to D5.2 Section 7. In the sequel, we refer this process of turning-off the RBs as transmission back-off algorithm. It is important to note that the selection of allowed RBs using the back-off algorithm is done before the basic backpressure algorithm. The output of the back-off algorithm is an entry for the main algorithm by providing the pre-selected RBs for the RRM optimization.
Of course, the main and difficult objective is to cleverly control the back-off selection, in particular using side information. Here are some important desired and undesired features to consider:

- As a Cluster Head (CH), turning-off some RBs will allow links from other clusters to experience better SINR. The effect is reciprocal if other clusters follow the same policy and will be beneficial for him.
- Clusters being selfish, back-off is performed if it does not harm their own communication in the long-run.
- If the cluster is willing to back-off, the time and the number of RBs for back-off should depend on the relative ratio between the local throughput requirements and others requirements (which entails signalling).

As for the side-information, we use two main assumptions. First, nodes are enhanced with sensing capability such that they can assess the level of aggregate interference introduced by other links from other clusters. The interference information from the physical layer is then provided to the MAC layer (i.e., cross-layer information). Second, a unique control channel, limited in capacity, exists where any CH can listen or broadcast. This channel is used by the CHs to exchange the required signalling information in order to run the back-off algorithm. Since in practice exchanges of signalling are limited and sensible to losses, we will also investigate the performance degradation when the probability of correctly receiving the broadcasts decreases.

Let us first resume the problem setting detailed in D5.2, Section 7. Consider an asynchronous clustered PMR scenario where the unique CH, elected in each cluster, is in charge to manage the local RRM and can broadcast limited signalling to all neighbour CHs using the limited common channel. The main objective is to minimize the nodes transmission energy while stabilizing their data buffer queue (i.e., for brevity, on the long term the data buffer remains within a constant level, assuming that the system is feasible). In particular, we seek to optimize the long term energy efficiency instead of the instantaneous power (see, D5.2). Assume N clusters composed of M total communication links that form an asynchronous clusterised ad-hoc network. The entire frequency band is composed of K resource blocks (RBs) representing the set K. A cluster n is associated with the sub-set of allowed RBs K_n, $K_n \in K$. The transmitter of link m has a data buffer whose dynamic is simply defined as

$$q(t) = \left[q(t-1) - \mu(t)\right]^+ + a(t)$$

where q(t) represents the data buffer size, a(t) is the stochastic random input bits, $\mu(t)$ is the theoretical throughput, and $\left[x\right]^+ = \max(x,0)$. Each CH asynchronously optimizes and updates the allocation policy for all the internal links following some algorithm, e.g., DUST algorithm. The following assumptions are considered:

- to ease the analysis, we assume a time-slotted system (i.e., discrete time t, representing for example a milli-second),
- the capacity $\mu(t)$ of a given link m using the k^{th} RB is defined by the basic Shannon capacity $\mu_m(t) = B_k \sum_{k \in K_n} \log_2(1 + \frac{p_m^{(k)}(t) g_{mm}^{(k)}(t)}{N_0 + I_m^{(k)}})$, where B_k is the bandwidth of the k^{th} RB.
- the intra-cluster links allocations are assumed orthogonal, i.e., only one link can be scheduled in a given RB, such that there is no intra-cluster interference. Furthermore, we also assume that adjacent channel interference within the same cluster is neglected (e.g., CP large enough with OFDM to account for the maximum delay inside the cluster, and using guard subcarriers with FBMC). However, since each asynchronous CH optimizes their RRM independently, inter-clusters interference cannot be avoided for both the direct and the adjacent channels, which is the main difficulty for the distributed RRM.
- intra-cluster signalling between nodes and their CH is assumed perfect without losses, whereas broadcast signalling is assumed sensible to collisions, e.g., SINR smaller than some threshold.
- the clustering of the M links and the selection of the unique cluster head (CH) in each cluster is not investigated here and is assumed to follow some clustering algorithm (e.g., [17], [32]),
- the pre-allocation of the set K_n for each cluster n is not addressed here

3.2 Algorithm description

We first describe the overall algorithm provided in Algorithm 1 with few modifications compared to the one described in D5.2.

The core of the optimization problem to be solved is given in line 5. Compared to the initial DUST algorithm described in D5.2, Section 7.5.1, where the objective was to maximize $\sum_{m \in M_n} \alpha_n(t)\zeta_m(t)\mu_m(t) - \eta_n(t)p_m(t)$, we now discard the use of the parameters $\alpha_n(t)\zeta_n(t)$ and the related time varying trade-off parameter $\eta(t)$. The main reason is that this approach slowly balances the power transmission from one link to another one in a continuous motion. However, recall the key statement from [14] which says that when the cross-talk interference (i.e., inter-cluster interference) is important, the best allocation is an orthogonal one. Hence, this continuous balancing of power instead of an ON/OFF style allocation can yield important wastes. Thus, instead of this continuous power balancing approach thanks to the variation of parameters $\alpha_n(t)\zeta_n(t)$ and $\eta(t)$, we decide to keep unchanged the basic centralized back-pressure approach (see e.g., [13] [15]) in each cluster, i.e., we simply maximize $\sum_{m \in M_n} q_m(t)\mu_m(t) - \eta p_m(t)$ where the data queue level $q_m(t)$.is the only weight for the throughput and $\eta$is a constant that does not vary along time However, we employ the idea of RBs back-off, which directly impacts the computation of $\mu_m(t) = \sum_{k \in K_n(t)} \mu_m(t)$, by restricting the allowed set $K_n(t)$. This restriction is crucial and is handled by the proposed back-off Algorithm 2, applied in line 5 of Algorithm 1. In order to run Algorithm 2, each local CH needs information about the other clusters, in our case through ψ (as explained in more details after Algorithm 1). Line 10 to 16 represents the preparation of that local information at each cluster before broadcasting, whereas Line 17 to 22 represent the update after successful or failed broadcast reception by the clusters receiving the broadcast.

Algorithm 1 Main algorithm

Init: $\mathcal{K} = [1:K], \mathcal{M} = [1:M], \mathcal{N} = [1:N]$
$\forall n : \mathcal{M}_n \in \mathcal{M}, \mathcal{K}_n \in \mathcal{K}$
$\forall m : q_m(1) = Q_{\text{INIT}}$,
$\forall n : \psi_{n,n'}(1) = 0, \forall n'$, $\mathcal{K}_n(1) = \mathcal{K}_n, \forall n$
Key constants: $\eta > 0$, $\zeta_1 \in [0,1], \zeta_2 > 0, \zeta_3, \zeta_4 > 0$

1 : For t = 1, 2, ...
2 : For $\forall n \in \mathcal{N}$ {*optimize cluster internal resource based on local info*)}
3 : $C_{\text{TimeOff},n} = [C_{\text{TimeOff},n} - 1, 0]^+$
4 : If $n \in \mathcal{N}_{\text{Optz}}(t)$
5 : Get $\mathcal{K}_n(t) \in \mathcal{K}_n$ from back-off **Algorithm 2** using $\psi_{n,n'}(t), \forall n'$, $C_{\text{TimeOff},n}$, and $\{\zeta_1, \zeta_2, \zeta_3\}$
6 : Update $\mathbf{P}_n^*(t) = \text{argmax} \sum_{m \in \mathcal{M}_n} \sum_{k \in \mathcal{K}_n(t)} \left(s_m^k q_m(t) \log_2 \left(1 + \frac{p_m^{(k)}(t) g_{m,m}^{(k)}(t)}{N_m^{(k)} + I_m^{(k)}(t-1)}\right) - \eta p_m(t) \right)$
s.t. $\sum_{m \in \mathcal{M}_n} s_m^k \leq 1$, $s_m^k \in \{0,1\}$, $\sum_{k \in \mathcal{K}_n(t)} p_m^k(t) \leq P_{\text{MAX}}$

7 : Update all rate $\mu_m(t)$, $\forall m$, using $\mathbf{P}^*(t)$. Record $I_m^{(k)}(t), \forall m \in \mathcal{M}, \forall k \in \mathcal{K}$
8 : Update $q_m(t+1) = [q_m(t) - \mu_m(t), 0]^+ + a_m(t)$, $\forall m \in \mathcal{M}$
9 : Update $\mathbf{P}_n^*(t+1) = \mathbf{P}_n^*(t), \forall n \in \mathcal{N}$

10 : For $\forall n \in \mathcal{N}$ {*prepare local info for broadcast*)}
11: $\psi_n(t+1) = 0$
12: For $\forall m \in \mathcal{M}_n$
13: if $\exists P_{m,\Delta_T}(t) > 0$ and $\frac{\bar{P}_{m,\Delta_T}^{(>0)}(t) \bar{g}_{k,\Delta T}(t)}{I_m^k(t-1)} < \zeta_4, \forall k \in \mathcal{K}_n$ OR $\nexists \mu_{m,\Delta_T}(\tau) > 0$ and $\bar{Q}_{m,\Delta_T}(t) > 0$
14: $\psi_n(t+1) = \psi_n(t+1) + q_m(t)$
15: If $n \in \mathcal{N}_{\text{Br}}(t)$
16: Broadcast $\psi_n(t+1)$

17: For $\forall n' \in \mathcal{N}$ {*update locally info received from broadcast*)}
18: For $\forall n \in \mathcal{N}_{\text{Br}}(t), n' \neq n$
19: If n' correctly decodes broadcast signal from n
20: $\psi_{n',n}(t+1) = \psi_n(t+1)$
21: else
22: $\psi_{n',n}(t+1) = \psi_{n',n}(t)$

- $\mathcal{N}_{\text{Br}}(t) \in \mathcal{N}$ represents the set of clusters that transmit a broadcast at time t.
- $\bar{x}_{k,\Delta T}(t) \stackrel{\text{def}}{=} \frac{1}{T} \sum_{\tau = max(t - \Delta_T + 1, 0)}^{t} x(\tau)$.
- $\bar{P}_{m,\Delta_T}^{(>0)}(t) \stackrel{\text{def}}{=} \frac{1}{\sum \mathbb{1}_{p_m(t)>0}} \sum_{\tau = max(t-\Delta_T+1,0)}^{t} p_m(t) > 0$ i.e., average power considering only positive power.
- ζ_1: Interfered buffer ratio threshold, ζ_2: Max queue variation within a time window, ζ_3: Maximum time for random back-off, ζ_4: SIR threshold considered as inter-cluster interference.

Note that in line 6, when performing the optimization for the current asynchronous allocation, the interference is based on the previous time epoch, i.e., a CH cannot know the new allocation in advance. However, when computing the theoretical throughput in line 7, the interference is from the current time (to be recorded for the next time epoch optimization).

Let us now focus on the back-off algorithm. In order to decide how many random RBs are to be turned-off as well as the random back-off time, we define information, called ψ_n that is to be broadcasted by the n^{th} CH. That information simply reflects the interfered links cumulated queue. Assume for example 3 links in a cluster with current buffer size, 1Kb, 2Kb, 5Kb, and such that only user 2 and 3 are interfered (e.g., because there are located at the cluster edge). Thus, CH n needs to broadcast $\psi_n = 7$ Kb, and all other CHs, record that information. However, since broadcast can be lost, and all CHs may not correctly receive the broadcast, it is necessary that each CH maintains its own specific knowledge about others, (up-to-date, or not). Thus,

instead of a unique ψ_n known by all CHs, we define $\psi_{n,n'}(t)$ as the information known by cluster n about cluster n'. Of course $n = n'$ means the cluster's local information which is always assumed up-to-date (i.e., $\psi_n = \psi_{n,n}$).

To create ψ_n at each CH, it is first important to define when a link is "considered as interfered or not". Following an intuitive (heuristic) approach, the interfered state is assumed as true for one of the following cases:

- the average SINR is lower than some threshold ζ_4. To compute the average SINR, we use the average non-zero transmit power within the last Δ_T epoch, and the corresponding average channel strength and aggregate interference.
- there is no transmission within Δ_T, and the average queue size within Δ_T is larger than some threshold.

The reason behind the first point is related to the fact that the decision whether or not a link is in an interference state, is through the SINR and not only the interference $I_m(t-1)$ itself. Using a threshold directly on the interference is not meaningful without considering the level of the utile signal. Thus, we consider the past non-null power to compute the average transmit power within Δ_T. However, recall that the basic back-pressure algorithm may decide to not transmit when the interference of a given link is too high compared to the level of its queue, such that it may postpone for a long period before deciding to transmit. Therefore, we need to assume to be in interference state if there is no transmission for a period of time but having the mean queue larger than some value (i.e., second criteria). Of course Δ_T should be selected not too small. Assume for example a link with very low traffic requirement and good channels that do not often need to transmit but in burst. It would be wrong to assume that it is experiencing interference.

We have just defined the information that is computed by each CH then broadcasted. Based on $\psi_{nn'}(t), \forall n'$, each cluster n then needs to decide how many random RBs will be turned-off and for how long. As, such we propose the following heuristic approach, with the following intuitive (heuristic) objectives.

- the number of RBs to be turned off should be larger as the ratio between the number of total interfered links, i.e., $\sum_{n' \neq n} \psi_n(n')$, becomes larger than the internal cluster links total queue plus $\sum_{n' \neq n} \psi_n(n')$. This means that if a cluster n is experiencing lots of interfered links, it will expects other clusters to apply some back-offs. In the other hand, if cluster n is informed that others are experiencing some bad interferences, it will apply some back offs. . Note that the choice of the ratio threshold, ζ_1 should be related to the system, e.g., to the number of clusters (or estimation).
- the clusters being selfish, most probable in practice, the back-off would be accepted only under certain conditions e.g., cluster n does not experience too much delay in helping others,
- the RBs to be turned-off by cluster n is randomly selected among K_n,

- the time for back-off is randomly selected, given a maximum time,
- since each cluster optimize its RRM based on the previous interference state, doing a back-off for only 1 time cannot be useful for others (otherwise when the other CH decides to transmit believing there is no interference, the local cluster would already start again to transmit, thus interfering). Thus, back-off should be activated if and only if the time for back-off is at least equal to 2 and the number of RBs to be turned off is at least equal to 1.

Based on the intuitive objectives described above, the detailed algorithm for the back-off is described hereafter.

Algorithm 2 Distributed back-off at CH n: RBs and back-off time decition algorithm

Input: $\psi_{n,n'}(t), \forall n'$ and $q_m(t) \forall m \in \mathcal{M}_n$
$\{\zeta_1, \zeta_2, \zeta_3\}$
$\mathcal{M}_n, \mathcal{K}_n$
$C_{\text{TimeOff},n}$

If $C_{\text{TimeOff},n} = 0$
  ratio $= \frac{\sum_{n' \neq n} \psi(n,n,t)}{\sum_{m \in \mathcal{M}_n} q_m(t) + \sum_{n' \neq n} \psi(n,n,t)}$
  If ratio $> \zeta_1$ and $q_m(t) - q_m(t - \Delta_T) < \zeta_2, \forall m \in \mathcal{M}_n$
    $c_1 = \min(\lceil \text{rand}() \times (2(\text{ratio} - 0.5))^2 \times |\mathcal{K}_n| \rceil, |\mathcal{K}_n|)$
    $c_2 = \lceil \text{rand}() \times \zeta_4 \rceil$
    $\mathcal{K}_{\text{Off},n} = \{\}$
    If $c_1 > 0$ and $c_2 > 1$
      $\mathcal{K}_{\text{Off},n} =$ randomly choose c_1 RBs $\in \mathcal{K}_n$
      $C_{\text{TimeOff},n} = c_2$
    End
    $\mathcal{K}_n(t) \stackrel{\text{def}}{=} \mathcal{K}_n - \mathcal{K}_{\text{Off},n}$
  End
Else
  $\mathcal{K}_n(t) = \mathcal{K}_n(t-1)$
End

Output: $\mathcal{K}_n(t), C_{\text{TimeOff},n}$

3.3 Numerical Experiments

The performance of the Algorithm 1 is demonstrated using the static spatial setting depicted in the following Figure 3-2.

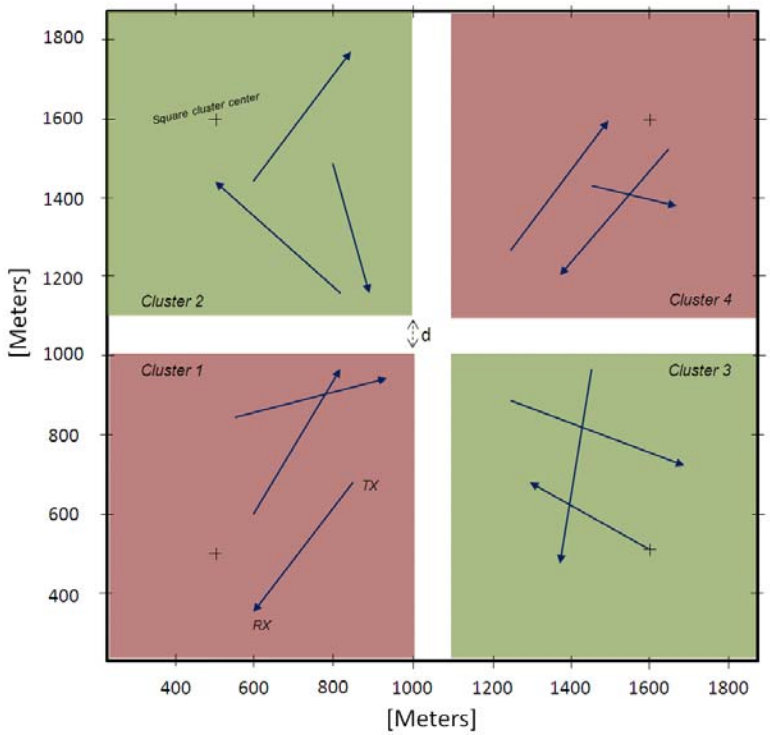


Figure 3-2: Simulation scenario: 4 squared-clusters with cluster center represented by the cross; 3 random links per cluster; each cluster uses one of the 2 sets of RBs (pink or green) with 3RBs in each set.

There are 4 clusters ($\mathrm{N} = 4$) and 3 random links in each cluster ($\mathrm{M} = 12$). The size of the square-clusters is 1 km wide. It is interesting to assess the algorithm for two main different situations, low and high inter-cluster interference. To do so we use two different spreadings of the 4 clusters by varying d in Figure 3-2 such that $\mathrm{d} = 2$ km and $\mathrm{d} = 100$ m to model low and high inter-cell interference, respectively. Note that, as opposed to D5.2, we do not use the ITU model with 9dB shadowing, in order to easily assess the relation between the different observed performances and the spatial static setting. Thus, the channel gains are obtained using only d^{-4}. The noise level is defined by 174 dBm/Hz, and $\mathrm{P\_max}$=30 dBm (note that in [18] the BS are provided with 43-47 dBm transmission capability, whereas 21-23 dBm for handsets).

The Phy settings are the following: F_c equals to 700 MHz and total bandwidth of 1.4 Mhz, where we assume that each RB is constrained by a minimum of 12 subcarriers (as in LTE) each of size 15 kHz , i.e., one RB equals 0.18 Mhz. Thus, for a total bandwidth of 1.4 MHz, we have in total 6 RBs available. Based on this setting, evaluating the aggregate interference ($I_m^{(k)}(t) = \sum_{\forall j \neq m, \forall m \in M_n} \sum_{k'=1}^{K} p_j^{(k')}(t) v_{k-k'} g_{jm}^{(k')}(t)$) of all the subcarriers, inside or outside a given RB, the coefficients were already provided in D5.2- Table 7.2, *i.e.*,

Modulation type	Vector of interference coefficient […, v(-2), v(-1), v(0), v(1), v(2), v(3), …]
Perfect	[0, 1, 0]
OFDM	[0.0201, 1.2324, 0.0201]
FBMC	[0.0073, 1.1615, 0.0073]

Note that all other coefficients are null. We assume that the interferences are symmetric in terms of RB, i.e., $v_{k-k'} = v_{k'-k}$. We recall that the v has been derived in [19] using Eq. 4.6-4.7 by assuming that the clusters are asynchronous with a time delay which is uniform over the duration of the OFDM symbol.

In the sequel, the following curve specification is used in all the figures (unless redefined): solid-lines represent the case where there is no back-off procedure, whereas dash-lines represent the case where the back-off is activated in the algorithm. The diamond, circle, and plus, refer to PM, OFDM, and FBMC, respectively. The black coloured curves are used for the left-hand side Y-axis metric, and the blue coloured curves for the right-hand side metric.

We first start by testing the overall algorithm on a scenario where the inter-cluster interference is almost inexistent or very low. To do so, we spread the clusters using $\mathrm{d} = 2$ km. Figure 3-3- displays the performance of Algorithm 1 with and without back-off (i.e., without back-off always yields $K_n(t) = K_n$ in Line 5 of Algorithm 1), and for the 3 types of modulations. The average power consumption and the average data queue size are provided as η varies. An interesting metric to compare the mean queue size with is the average input traffic that is provided by the straight dashed-line (also provided in Figure 3-4 and Figure 3-5). It is important to notice that there is almost no difference between using back-off or not and between the 3 modulations, i.e., the 6 curves are superposed. This outcome is quite intuitive as we have largely spread clusters with low inter-cluster interference such that the back-off algorithm simply does not need to turn-off the RBs. Moreover, as the adjacent interference due to neighbouring clusters is also low for the same reason, there is no impact whether we use OFDM or FBMC, as compared to PM.

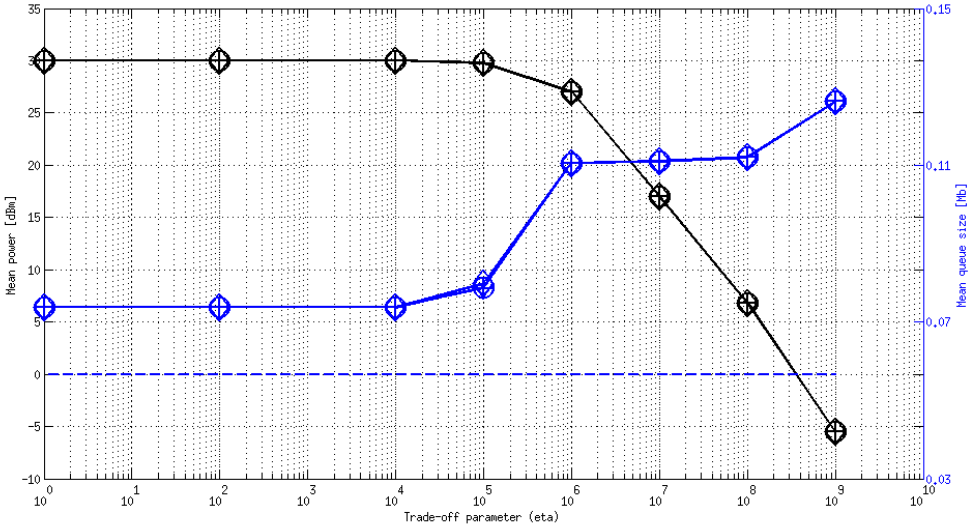
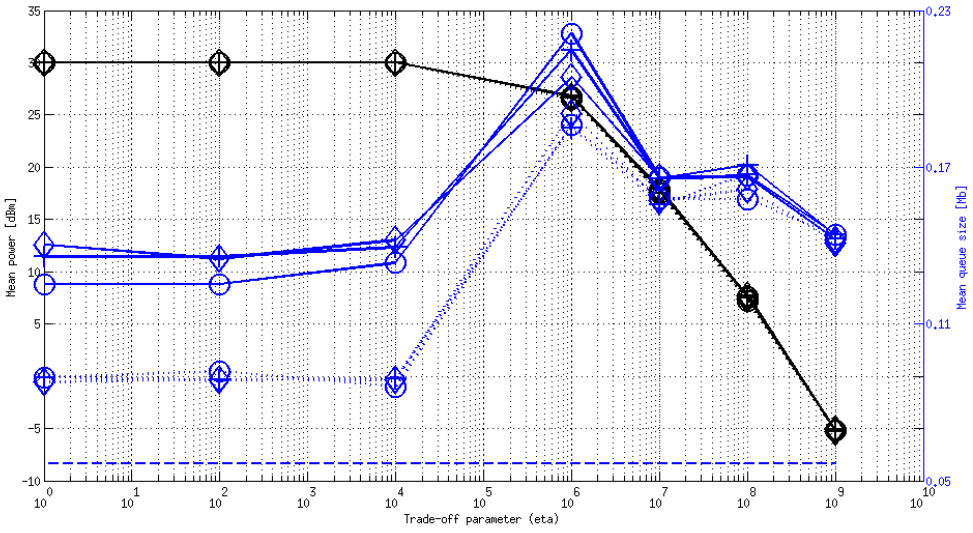


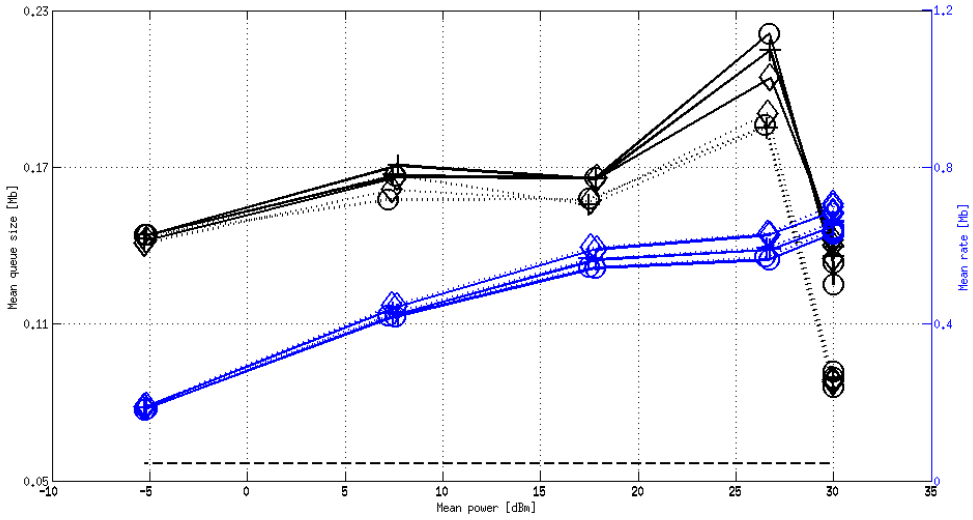
Figure 3-3: Low inter-cluster interference scenario: Average power consumption and delay for varying η, for the 3 modulations, with (dotted-lines) and without back-off (solid-lines); diamond, circle, and plus, refer to PM, OFDM, and FBMC, respectively.

Let us now assess the algorithm performance when the clusters are close to each other such that the interference may be high (depending on the transmit power). It is important to note that even links from nearby clusters may succeed to transmit their data, thanks to relative distances that may yield good SINR. Taking for example the links depicted in Figure 3-2, link 11 may certainly get bad SINR if link 9 transmits, but can succeed if instead of link 9 it is for example link 7 which uses the same RB. But of course, in the long run, the RRM algorithm not using the back-off algorithm should yield larger delay. Figure 3-4 a) presents the same comparison as in Figure 3-3 with now $d = 100$ m to model possible high inter-cluster interference. In particular, it is important to observe the average queue size (delay). We first note, that for any η, the delay is always lower when using the back-off algorithm compared to not using it. This is of course a desired behaviour. This is especially true for η small. This is because as η becomes small, the average transmit power increases (close to P_{max}), such that links from nearby clusters are more likely to always interfere each other. Thus, the use of the RBs back-off becomes very relevant. Now, when η increases, and the transmission power diminishes, the maximum interfering region of each cluster diminishes, such that two simultaneously transmitting links get better and better SINR (of course the transmit power cannot be arbitrarily small since the SINR even with vanishing interference is limited by the receiver noise level, N_0). Thus, back-off becomes less necessary such that the performances between PM, FBMC and OFDM become approximately the same.

It is important to note that, as for the basic backpressure algorithm, increasing infinitely the η parameter, does not continuously decrease the transmit power. This is because, as the transmit power becomes close to the optimal minimum (cannot go lower with queue instability), the data queues will start increasing steadily, and thanks to the counter effect of the queues in the optimization problem (i.e., $\sum_{m \in M_n} q_m(t)\mu_m(t) - \eta p_m(t)$), any high η will always be compensated at some time by the increasing queue. However, this may yield undesirable delays (i.e. time for the queues to be large enough to compensate for η).



a) Average power consumption and delay for varying η.

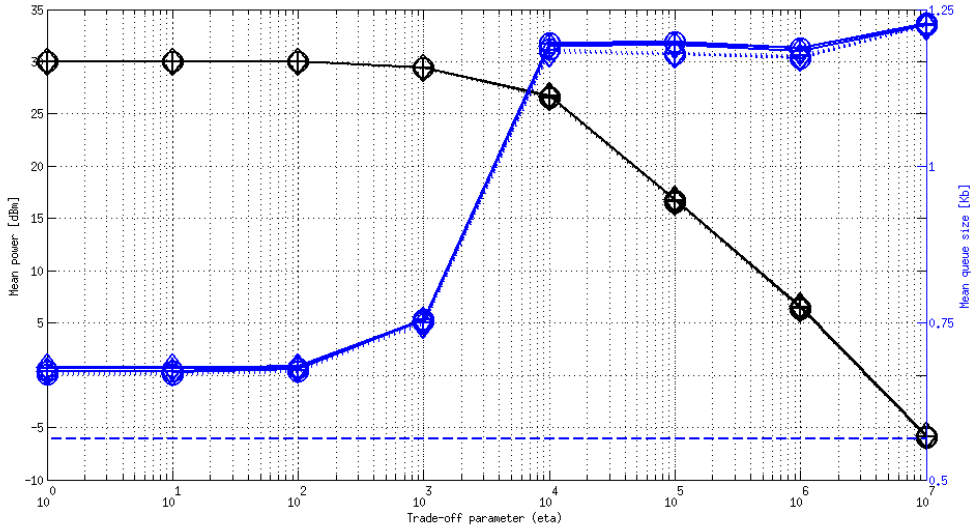


b) Average delay and theoretical throughput for the associated power consumption

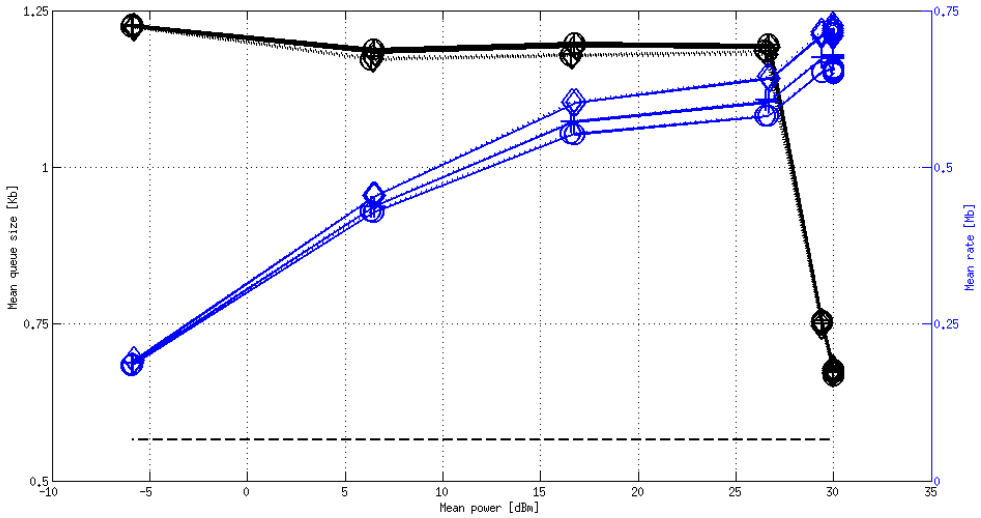
Figure 3-4: High inter-cluster interference scenario: for the 3 modulations, with (doted-lines) and without back-off (solid-lines); diamond, circle, and plus, refer to PM, OFDM, and FBMC, respectively.

Another point to notice is that one would always expect that the delay performance for the PM is always lower than for the FBMC and itself lower than OFDM. However, Figure 3-4 a) does not always reflect that intuitive relation. Although the differences quite vanish using the back-off algorithm, they are well pronounced when not using the back-off. In order to further investigate this effect, we plot Figure 3-4 b) for the same performance as in Figure 3-4 a) but using different axis. This figure better reflects the performance (delay and theoretical throughput) in terms of the effective depleted power instead of η (since for a given η different settings can yield different average transmit power). In the second figure, the difference of performance between the 3 modulations is not significant in terms of the average queue. However, when observing the theoretical throughput, $\mu(t)$, we note that the expected performance ranking between the 3 modulations is always respected, i.e., PM higher than FBMC and itself better than OFDM.

In Figure 3-4 we assessed the algorithm performance for some high rate traffic. We provide the same results in Figure 3-5 using lower traffic input (simply dividing by 1000 the input bits). It is very clear that there is no much difference between using or not the back-off (or quite small). As explained earlier, even links from nearby clusters can well communicate simultaneously depending on the geographical location (thus their channel quality), as far as the SINR is good enough. Thus, since the traffic is much smaller than in Figure 3-4, the few times when link 11 and link 7 are luckily scheduled at the same time (i.e., no much self interference) is enough to deplete the small data queues. Thus, back-off is not very crucial, as opposed to the high traffic scenario.



a) Average power consumption and delay for varying η

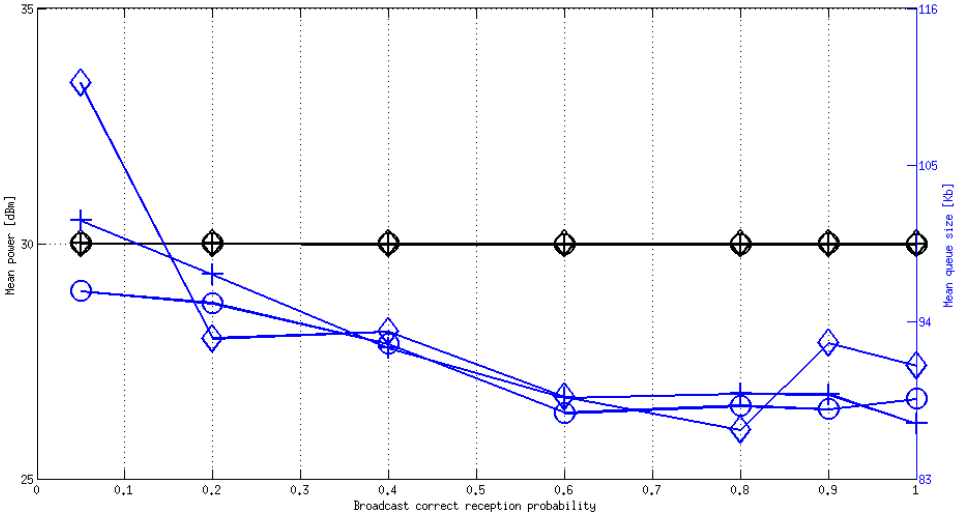


b) Average delay and theoretical throughput for the associated power consumption

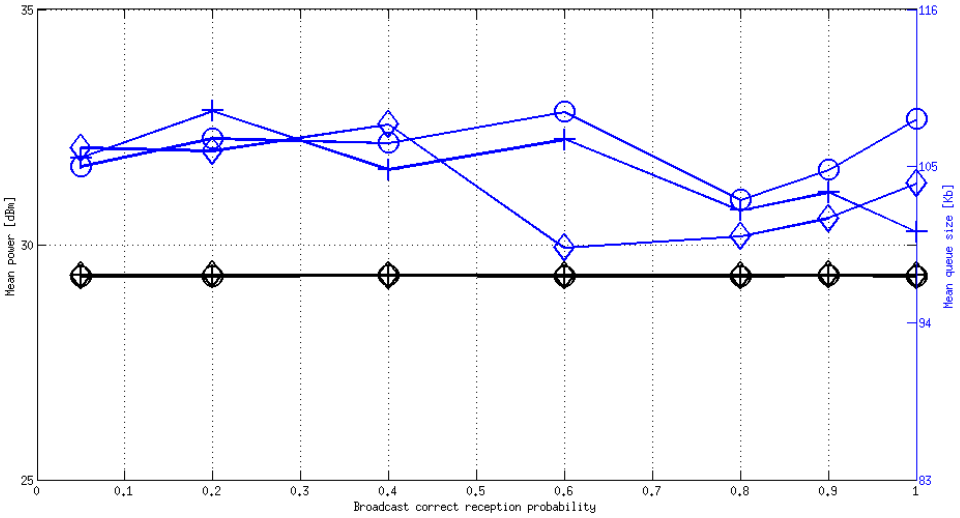
Figure 3-5: High inter-cluster interference scenario but lower traffic input compared to Figure 3-4 (divided by 1e3): for the 3 modulations, with (dotted-lines) and without back-off (solid-lines); diamond, circle, and plus, refer to PM, OFDM, and FBMC, respectively.

In all the previous performance assessments, we assumed that all signalling broadcasts where always successfully received by all neighbour clusters. However, in practice the broadcast can be lost due to collisions. In order to model this important real constraint, we randomly impose some broadcasts losses at each neighbour cluster. As such we analyse the delay performance as the probability of correctly receiving a broadcast varies. In particular, we do this test for two η parameters, $\eta=10^{2}$ and $\eta=10^{5}$, i.e. yielding high and low transmission power. We

observe in Figure 3-6 a), for $\eta=10^{2}$, that as the probability of correctly receiving the broadcast diminishes, the delay increases. However, when observing Figure 3-6 b), for $\eta=10^{5}$, that the deterioration becomes less obvious. These are expected results, since as η increases, the transmission power diminishes along with the inter-cluster interferences such that the back-off algorithm becomes less important. Thus, the frequency update of the signalling message ψ_n also becomes less important.



$\eta=10^{2}$



$\eta=10^{5}$

Figure 3-6: Delay performance for varying probability of correctly receiving the broadcast, for the 3 modulations; diamond, circle, and plus, refer to PM, OFDM, and FBMC, respectively.

It is also important to emphasize the nature of the broadcasted information, Ψ_n, which is the data quantity. When transmitting that quantity in a real system, it is most probably required to quantize that information. However, for such type of information quantization loss may not necessarily impact the performance of the algorithm. This of course depends on the overall queues' size. In other word, having 1 Mb difference for an average traffic of the order of few Mb is more critical than for traffics of the order of tens or hundreds of Mb.

3.4 Performance summary

Here is the performance summary of the proposed new algorithm taking advantage of the SINR information:

- ✓ As for the classical back-pressure algorithm, increasing the trade-off parameter η decreases the energy consumption. This helps neighbour clusters to interfere less among each other such that can have acceptable SINR when communicating at the same time. However, using arbitrarily large η will yield unreasonable delays.
- ✓ The performance deterioration with lower successful broadcast is more visible when clusters interfere more each other, i.e., nearby clusters use high transmit powers, such that information update about other clusters state becomes more important.
- ✓ The difference in performance between the three types of modulations becomes small when using the back-off algorithm since the direct channel interference is better controlled. This is thanks to the orthogonality in transmission obtained using the back-off algorithm when inter-cluster interference is too high.
- ✓ Although, the difference in delay performance is not significant, the better performances of PM over FBMC, itself over OFDM is clear in terms of throughput. This is especially true when the power transmission is high. In practice, some system can only work in an ON/OFF transmission fashion or with restricted set of power transmission. In such a case the use of Back-off is very useful.

To summarize, the back-pressure algorithm provides a trade-off between long-term energy reduction and system latency, thanks to the tunable η parameter. As a simple example, right after a disaster, much voice traffic with stringent quality may be necessary. This would require low η parameter for the algorithm, accepting the fact the energy depletion is not the most important. Then, few days later,, high quality photos or video of the area may be needed to be transmitted for a long period. In such a case, one would use higher η parameter, accepting the fact that the communication latency can be higher but with more efficient usage of the energy storage. In addition to the salient feature of the back-pressure algorithm, the back-off feature is important in a distributed clusterised network where inter-cluster interference may be high. The idea for the back-off on top of the back-pressure algorithm is based on the important fact that when the cross-talk interference is higher than some threshold, the best communication is an orthogonal-type communication like FDMA [14]. Of course, to back-off or not depends on many criteria for each independent and selfish (most probably) cluster The back-off is done without agreement between clusters but independently based on broadcast signaling sent by the different CHs. That signaling is limited in capacity and the overall algorithm is quite robust to some loss.

4. Scheduling algorithms for FBMC systems

4.1 Introduction

It is well-known that the merits of OFDM come with the cost of transmitting redundancy in the form of a cyclic prefix (CP) and shaping the subcarrier signals with the rectangular window. A viable alternative is the FBMC modulation scheme [20], which does not transmit the CP and shapes subcarriers with well-frequency-localized waveforms. The price that is paid to refrain from transmitting a CP is a weak orthogonality that it is only satisfied in the real domain. Thus, multipath fading ruins the orthogonality of FBMC. This is the main obstacle to extend FBMC to multiple-input-multiple-output (MIMO) architectures [21], which is crucial to enhance the performance. Knowing that the solutions devised for OFDM cannot be applied to FBMC in general, several techniques have been proposed to combine MIMO and FBMC [22].

One of the few works that addresses the resource allocation problem in the FBMC context is presented in [23], where the rate in the multiple access channel is maximized given power and users' rate constraints. In [24] the authors seek for maximizing the downlink capacity of cognitive radio systems. By contrast, the work presented in this section aims at minimizing the transmit power in the downlink subject to users' rate constraints. The propagation conditions considered in previous works are such that the interference can be neglected [24], or cancelled out by applying the same signal processing techniques as in OFDM [23]. This is not the case in this section, where the channel is more frequency selective revealing that the signal received by each user is affected by inter-user interference if the user allocation is different in adjacent subcarriers, i.e. when the active users is not the same. To overcome this issue we propose to assign subcarriers to users in a block-wise fashion. Then, we demonstrate that FBMC systems can benefit from the algorithms proposed in [25] to solve the margin adaptive problem. It is important to remark that this section focuses on multi-user multiple-input-single-output (MU-MISO) communication systems, because in this configuration FBMC exhibits robustness against the frequency selectivity, while it achieves the same spatial channels gains as OFDM [26].

The main contribution of this section consists in modifying the scheduling algorithm presented in [25] according to the FBMC transmission scheme, by proposing a specific subcarrier grouping and taking into account the existence of inter-user, inter-symbol and inter-carrier interference.

4.2 System model

Consider the downlink of a MU-MISO communication system where the terminals and the base station (BS) are equipped with a single and N_T antennas, respectively. The BS uses the spatial dimension to simultaneously serve up to N_U users in the same frequency resources. When the FBMC transceiver is considered, the signal received by the *l*th user after demodulating the *q*th subcarrier is [26]:

$$y_{lq}[k] = \sum_{u=1}^{N_U} \sum_{m=q-1}^{q+1} \sum_{\tau=-3}^{3} \alpha_{qm}[\tau] \boldsymbol{H}_{lm} \boldsymbol{B}_{um} \times \theta_m[k-\tau] d_{um}[k-\tau] + w_{lq}[k] \quad (4\text{-}1)$$

$$\boldsymbol{H}_{lm} = [H_{l1}(m) \quad \dots \quad H_{lN_T}(m)] \quad (4\text{-}2)$$

$$\theta_m[k] = \begin{cases} 1 & k+m \text{ even} \\ j & k+m \text{ odd} \end{cases}, \quad (4\text{-}3)$$

for $0 \leq q \leq M-1$ and $1 \leq l \leq N_U$. Let $w_{lq}[k]$ be the noise that contaminates the reception of the *l*th user on the *q*th subcarrier, which follows this distribution $w_{lq}[k] \sim \mathcal{CN}(0, N_0)$. The channel frequency response is assumed flat at the subcarrier level. In this sense, the term $\boldsymbol{H}_{lm} \in \mathbb{C}^{1 \times N_T}$ denotes the MISO channel frequency response seen by the *l*th user on the *m*th subcarrier. The sequence $d_{um}[k]$ is frequency multiplexed on the *m*th subcarrier and it contains real-valued PAM symbols that are intended for the *u*th user. Note that $d_{um}[k]$ is linearly precoded with $\boldsymbol{B}_{um} \in \mathbb{C}^{N_T \times 1}$. The coefficients $\{\alpha_{qm}[k]\}$ denote the intrinsic interference and they are defined as

$$\alpha_{qm}[k] = \left(f_m[n] * f_q^*[-n]\right)_{\downarrow \frac{M}{2}} \quad (4\text{-}4)$$

$$f_m[n] = p[n] e^{j\frac{2\pi}{M} m\left(n - \frac{L-1}{2}\right)}, \quad (4\text{-}5)$$

where $*$ denotes convolution. Note that $f_m[n]$ is the subband pulse that is obtained by frequency shifting the low-pass prototype pulse $p[n]$ the length of which is L. In this section we opt for the design described in [27] with an overlapping factor equal to four. The operation $(.)_{\downarrow x}$ performs a decimation by a factor of x. In the q even case $\{\alpha_{qm}[k]\}$ takes the values gathered in Table 1. The same magnitudes hold when q is odd but the sign may vary.

To get rid of the interferences $y_{lq}[k]$ is post-processed as follows:

$$\check{d}_{lq}[k] = \Re\left(\theta_q^*[k] a_{lq} y_{lq}[k]\right) = a_{lq} \Re\left(\boldsymbol{H}_{lq} \boldsymbol{B}_{lq}\right) d_{lq}[k] + \sum_{\substack{(m,\tau,u) \neq \\ (q,0,l)}} a_{lq} \Re\left(\theta_q^*[k] \theta_m[k-\tau] \alpha_{qm}[\tau] \boldsymbol{H}_{lm} \boldsymbol{B}_{um}\right) d_{um}[k-\tau] + a_{lq} \Re\left(\theta_q^*[k] w_{lq}[k]\right). \quad (4\text{-}6)$$

Note that the equalizer a_{lq} is constrained to be real-valued [26].

	k=-3	k=-2	k=-1	k=0	k=1	k=2	k=3
$\alpha_{qq-1}[k]$	-j0.0429	-0.1250	j0.2058	0.2393	-j0.2058	-0.1250	j0.0429
$\alpha_{qq}[k]$	-0.0668	0	0.5644	1	0.5644	0	-0.0668
$\alpha_{qq+1}[k]$	j0.0429	-0.1250	-j0.2058	0.2393	j0.2058	-0.1250	-j0.0429

Table 1: Intrinsic interferences under ideal propagation conditions

4.3 SDMA with block diagonalization

One solution to achieve spatial division multiple access (SDMA) consists in following the block diagonalization (BD) approach [28]. Although the BD technique succeeds in achieving interference-free data multiplexing, it imposes $N_T \geq N_U$ in MU-MISO communication systems. It can be checked that the subband processing proposed in [26] guarantees that (4-6)

is free of inter-symbol interference (ISI), inter-carrier interference (ICI) and inter-user interference (IUI), as long as the users served on a given subcarrier are the same on adjacent subcarriers. Then, the maximum achievable rate becomes

$$r_{lq} = \frac{1}{2}\log_2\left(1 + SINR_{lq}\right) \tag{4-7}$$

$$SINR_{lq} = p_{lq}\left\|\boldsymbol{H}_{lq}\breve{\boldsymbol{V}}_{lq}^{0}\right\|_2^2 / 0.5N_0 = p_{lq}\lambda_{lq}, \tag{4-8}$$

where $\breve{\boldsymbol{V}}_{lq}^{0} \in \mathbb{C}^{N_T \times N_T - N_U + 1}$ spans the null space of

$$\breve{\boldsymbol{H}}_{lq} = \left[\boldsymbol{H}_{1q}^T \cdots \boldsymbol{H}_{l-1q}^T \ \boldsymbol{H}_{l+1q}^T \cdots \boldsymbol{H}_{N_U q}^T\right]^T. \tag{4-9}$$

Note that the power allocated to $d_{lq}[k]$ is given by p_{lq}. The factor $\frac{1}{2}$ in (4-7) has to do with the fact that the variables in (4-6) are real-valued. To get (4-7) we assume a continuous transmission, so that the tailsof the pulse have no impact on the rate. This does not hold true in burst-like transmission, but we leave this case for future work. In the OFDM counterpart the rate is given by (4-7) but the factor $\frac{1}{2}$ is dropped because the information is conveyed in both the in-phase and quadrature components, i.e. the rate is formulated as

$$r_{lq} = \log_2\left(1 + 2p_{lq}\left\|\boldsymbol{H}_{lq}\breve{\boldsymbol{V}}_{lq}^{0}\right\|_2^2 / N_0\right). \tag{4-10}$$

The SINR is not modified but there are two aspects that have to be taken into account. The first one is that the power of the desired signal is multiplied by two since the symbols transmitted in OFDM are drawn from the QAM scheme and the real-valued symbols $\{d_{lq}[k]\}$ are obtained from either the real or the imaginary parts of the QAM constellation points. The second relevant aspect has to do with the fact that the power of the noise is not halved because detection is performed in the complex domain. Let us stress that when the power coefficients, the bandwidth and the sampling frequency are kept unchanged in both modulations, then the rate expressed in bps is the same in FBMC and OFDM without CP.

To ease the comparison between FBMC and OFDM we consider the rate that corresponds to one OFDM symbol period. Thus, the rate that will be used from this point on unless otherwise stated is expressed in this form

$$r_{lq} = \log_2\left(1 + SINR_{lq}\right), \tag{4-11}$$

which is valid for both modulations. Bearing in mind that (4-11) is achieved in FBMC systems by transmitting two multicarrier symbols, it follows that the total transmitted power in one OFDM symbol period is $\sum_{l=1}^{N_U}\sum_{q=0}^{M-1} 2\,p_{lq}$. By contrast, taking into account that the symbols in OFDM are obtained from a complex-valued constellation diagram, the transmitted power in OFDM is equal to $\sum_{l=1}^{N_U}\sum_{q=0}^{M-1}\left(1 + \frac{L_{CP}}{M}\right)2p_{lq}$, where L_{CP} denotes the CP length. For the sake of clarity the CP will be only considered when the transmitted power is computed and will be neglected when the rate is evaluated.

4.4 Margin adaptive scheduling algorithm

It is worth emphasizing that when $N_T < N_U$ the BS selects N_T out of N_U users over each subcarrier. However, the user assignment has to remain constant over all subbands so that the BD approach is able to remove the ICI in FBMC. This can be easily proved as follows. Consider a toy example where 2 users are allocated in each subcarrier. Imagine that subcarriers $\{q-1, q+1\}$ are assigned to the users u_1 and u_2, while the users l_1 and l_2 are allocated in the qth subcarrier. If $\{a_{lq}\}$ and $\{\boldsymbol{B}_{lq}\}$ are designed according to [26], we get

$$\check{d}_{l_iq}[k] = a_{l_iq}\Re\left(\boldsymbol{H}_{l_iq}\boldsymbol{B}_{l_iq}\right)d_{l_iq}[k] + \sum_{\tau=-3}^{3}\sum_{\substack{m=q-1\\ m\neq q}}^{q+1}\sum_{j=1}^{2} a_{l_iq}\,\Re\left(\theta_q^*[k]\theta_m[k-\tau]\alpha_{qm}[\tau]\boldsymbol{H}_{l_im}\boldsymbol{B}_{u_jm}\right) \times d_{u_jm}[k-\tau] + a_{l_iq}\Re\left(\theta_q^*[k]w_{l_iq}[k]\right), \quad (4\text{-}12)$$

for $i = 1,2$. Note that in (4-12) there is no contribution from user l_j for $j \neq i$. On the negative side, we cannot neglect the ICI because the precoding vectors do not satisfy $\boldsymbol{H}_{l_im}\boldsymbol{B}_{u_jm} = \boldsymbol{0}$. Hence, (4-12) shows the situations that should be avoided to get rid of the interference. To guarantee that all users achieve a certain rate in the absence of interference, we propose to assign subcarriers to users in a block-wise fashion. In other words, the band is partitioned into X subsets, where the subset S_i encompasses these subcarrier indexes $\left[(i-1)\frac{M}{X},\ldots,i\frac{M}{X}-1\right]$ assuming that $\frac{M}{X}$ is an integer number. Taking into account that the roll-off factor of the prototype pulse is close to one, the first carrier of each set is left empty to isolate the blocks. Based on that we define $\bar{S}_i = S_i - \left\{(i-1)\frac{M}{X}\right\}$ for all i.

In the light of the above discussion, the optimization problem is posed as follows:

$$\begin{aligned} &\underset{\{p_{lq}\},\{\rho_{li}\}}{\operatorname{argmin}} \quad \sum_{l=1}^{N_U}\sum_{i=1}^{X}\sum_{q\in\bar{S}_i} p_{lq} \\ &s.t. \sum_{i=1}^{X}\rho_{li}\sum_{q\in\bar{S}_i}\log_2\left(1+p_{lq}\lambda_{lq}\right) \geq R_l, \qquad 1\leq l\leq N_U \\ &\sum_{l=1}^{N_U}\rho_{li} \leq N_T, \qquad \rho_{li}\in\{0,1\}, \qquad 1\leq i\leq X. \end{aligned} \quad (4\text{-}13)$$

By solving (4-13) we find the optimal user selection and power allocation, so that the users' rate constraints are guaranteed with the minimum transmit power. Note that the channel gains $\{\lambda_{lq}\}$ defined in (4-8) depend on the channel interference matrix, which in turns depend on the subcarrier assignment.

With the aim of substantially reducing the complexity, we assume that constant power is used on each block. Then the number of variables to optimize is significantly reduced and the problem can be simplified as

$$\begin{aligned} &\underset{\{p_{li}\},\{\rho_{li}\}}{\operatorname{argmin}} \quad \sum_{l=1}^{N_U}\sum_{i=1}^{X}|\bar{S}_i|\,p_{li} \\ &s.t. \sum_{i=1}^{X}\rho_{li}\sum_{q\in\bar{S}_i}\log_2\left(1+p_{li}\lambda_{lq}\right) \geq R_l, \qquad 1\leq l\leq N_U \\ &\sum_{l=1}^{N_U}\rho_{li} \leq N_T, \qquad \rho_{li}\in\{0,1\}, \qquad 1\leq i\leq X, \end{aligned} \quad (4\text{-}14)$$

where $|\bar{S}_i| = \frac{M}{X} - 1$ and p_{li} denotes the power that user l assigns to all subcarriers that belong to $\bar{S}_i$. By mapping the sum of rates into a single metric we get more tractable expressions. In this regard, we consider this inequality

$$\sum_{q\in\bar{S}_i} \log_2\left(1+p_{li}\lambda_{lq}\right) \geq |\bar{S}_i|\log_2\left(1+p_{li}\hat{\lambda}_{li}\right), \tag{4-15}$$

where

$$\hat{\lambda}_{li} = \left(\prod_{m\in\bar{S}_i} \lambda_{lm}\right)^{1/|\bar{S}_i|} \tag{4-16}$$

is the geometric mean [29]. By substituting the sum of rates by the right hand side of (4-15), we can reformulate (4-14) as

$$\begin{aligned} &\underset{\{p_{li}\},\{\rho_{li}\}}{\operatorname{argmin}} \quad \sum_{l=1}^{N_U}\sum_{i=1}^{X} |\bar{S}_i|\, p_{li} \\ &s.t. \ \sum_{i=1}^{X} \rho_{li}\, |\bar{S}_i|\log_2\left(1+p_{li}\hat{\lambda}_{li}\right) \geq R_l, \qquad 1\leq l\leq N_U \\ &\sum_{l=1}^{N_U} \rho_{li} \leq N_T, \qquad \rho_{li}\in\{0,1\}, \qquad 1\leq i\leq X. \end{aligned} \tag{4-17}$$

Now the blocks play the same role as subcarriers and, thus, we can take advantage of existing low-complexity algorithms, see e.g. [25]. The solution of (4-17) guarantees that the original rate constraints in (4-13) are satisfied because of the inequality established in (4-15).

4.5 Successive channel allocation

This section focuses on solving (4-17) by using the linear programming based successive channel allocation (LPSCA) algorithm described in [25]. Among the possible choices we have favored the LPSCA because it exhibits an excellent tradeoff between complexity and performance. The LPSCA was initially though for OFDM systems, which highlights that it may not be easily tailored to the FBMC scheme. To cast some light into the applicability of the LPSCA to FBMC systems this section briefly describes the algorithm to be employed and the necessary modifications due to the characteristics of the transmitted signal.

The main asset of the LPSCA stems from the reduction of the complexity that is achieved by grouping users into N_T disjoint sets, so that $\{1,\ldots,N_U\} = K_1 \cup \ldots \cup K_{N_T}$. The partitioning is made according to the average channel quality assessment [25]. Then, (4-17) is also partitioned into N_T independent problems that are sequentially optimized. First we allocate the most distant users to the BS, which means that we start with subset K_1 and we terminate with the subset K_{N_T}. When subset K_r is addressed, the users that belong to K_j, for $j<r$ have already been allocated. To guarantee that the rate constraints are not violated, the subband processing is designed to prevent signals that are intended to users in K_r from leaking through users that belong to K_j, for $j<r$. In the rest of the section we focus on the subset K_r without loss of generality. In this sense, the signal received by the user $l\in K_r$ is

$$\begin{aligned} \check{d}_{lq}[k] = a_{lq}\Re\left(\boldsymbol{H}_{lq}\boldsymbol{B}_{lq}\right)d_{lq}[k] \\ + a_{lq}\left(i_{lq}[k]+\Re\left(\theta_q^*[k]w_{lq}[k]\right)\right), \end{aligned} \tag{4-18}$$

where

$$i_{lq}[k] = \sum_{m=q-1}^{q+1} \sum_{u\in\mathcal{A}_m^r} \sum_{\tau=-3}^{3} \Re\left(\theta_q^*[k]\theta_m[k-\tau] \times \alpha_{qm}[\tau]\boldsymbol{H}_{lm}\boldsymbol{B}_{um}\right)d_{um}[k-\tau] + \sum_{(m,\tau)\neq(q,0)} \Re\left(\theta_q^*[k]\theta_m[k-\tau] \times \alpha_{qm}[\tau]\boldsymbol{H}_{lm}\boldsymbol{B}_{lm}\right)d_{lm}[k-\tau]. \quad (4\text{-}19)$$

The already allocated users in the mth subcarrier when the subset K_r is addressed, are included in this set $\mathcal{A}_m^r$ and, thus, $\mathcal{A}_m^r \subseteq \bigcup_{j=1}^{r-1} K_j$. Now the transmit processing is designed to project the channel onto the null space of

$$\breve{\boldsymbol{H}}_{lq} = \left[\boldsymbol{H}_{\mathcal{A}_q^r(1)q}^T \cdots \boldsymbol{H}_{\mathcal{A}_q^r(r-1)q}^T\right]^T. \quad (4\text{-}20)$$

Let $\mathcal{A}_q^r(j)$ denote the jth element of subset $\mathcal{A}_q^r$. Bearing (4-20) in mind, $\boldsymbol{B}_{lq}$ and a_{lq} are designed as [26] proposes in order to remove the interference, yielding

$$\boldsymbol{B}_{lq} = \sqrt{p_{lq}} \frac{\breve{\boldsymbol{U}}_{lq}^0\left(\boldsymbol{H}_{lq}\breve{\boldsymbol{U}}_{lq}^0\right)^H}{\left\|\breve{\boldsymbol{U}}_{lq}^0\left(\boldsymbol{H}_{lq}\breve{\boldsymbol{U}}_{lq}^0\right)^H\right\|_2} \quad (4\text{-}21)$$

$$a_{lq} = \frac{\Re\left(\boldsymbol{H}_{lq}\boldsymbol{B}_{lq}\right)}{\left|\Re\left(\boldsymbol{H}_{lq}\boldsymbol{B}_{lq}\right)\right|^2 + \sigma_{lq}^2}, \quad (4\text{-}22)$$

where $\breve{\boldsymbol{U}}_{lq}^0 \in \mathbb{C}^{N_T\times N_T-r+1}$ spans the null space of (4-20). Then, $\Re\left(\theta_q^*[k]\theta_m[k-\tau]\alpha_{qm}[\tau]\boldsymbol{H}_{lm}\boldsymbol{B}_{lm}\right) = 0$, for $(m,\tau) \neq (q,0)$. Since the users in K_r cannot claim protection against the users in K_j, for $j < r$, the first term of (4-19) is not removed and the variance of the noise plus the interference becomes

$$\sigma_{lq}^2 = \sum_{m=q-1}^{q+1} \sum_{u\in\mathcal{A}_m^r} \sum_{\tau=-3}^{3} \left|\Re\left(\theta_q^*[k]\theta_m[k-\tau] \times \alpha_{qm}[\tau]\boldsymbol{H}_{lm}\boldsymbol{B}_{um}\right)\right|^2 + 0.5N_0. \quad (4\text{-}23)$$

Considering the values of Table 1 along with the fact that the same users are allocated in adjacent subcarriers, i.e. $\mathcal{A}_{q-1}^r$=$\mathcal{A}_q^r$=$\mathcal{A}_{q+1}^r$, (4-23) can be expressed as

$$\sigma_{lq}^2 = 0.5N_0 + \sum_{u\in\mathcal{A}_q^r} \left(\left|\Re\left(\boldsymbol{H}_{lq}\boldsymbol{B}_{uq}\right)\right|^2 + 0.646\left|\Im\left(\boldsymbol{H}_{lq}\boldsymbol{B}_{uq}\right)\right|^2 + 0.1769\left|\Im\left(\boldsymbol{H}_{lq-1}\boldsymbol{B}_{uq-1}\right)\right|^2 + 0.1769\left|\Im\left(\boldsymbol{H}_{lq+1}\boldsymbol{B}_{uq+1}\right)\right|^2\right). \quad (4\text{-}24)$$

If the channel frequency selectivity is not severe we can assume that

$$\left|\Im\left(\boldsymbol{H}_{lq-1}\boldsymbol{B}_{uq-1}\right)\right|^2, \left|\Im\left(\boldsymbol{H}_{lq+1}\boldsymbol{B}_{uq+1}\right)\right|^2 \approx \left|\Im\left(\boldsymbol{H}_{lq}\boldsymbol{B}_{uq}\right)\right|^2. \quad (4\text{-}25)$$

Then, it follows that (4-24) can be approximated by

$$\sigma_{lq}^2 \approx 0.5N_0 + \sum_{u\in\mathcal{A}_q^r} \left|\boldsymbol{H}_{lq}\boldsymbol{B}_{uq}\right|^2. \quad (4\text{-}26)$$

When the subset K_1 is addressed $\sigma_{lq}^2 = 0.5N_0$. In the general form the rate is given by

$$r_{lq} = \log_2\left(1 + SINR_{lq}\right) \quad (4\text{-}27)$$

$$SINR_{lq} = p_{lq}\left\|\boldsymbol{H}_{lq}\breve{\boldsymbol{U}}_{lq}^0\right\|_2^2 / \sigma_{lq}^2 = p_{lq}\beta_{lq}. \quad (4\text{-}28)$$

Once the interference is updated, the problem associated to the subset K_r is

$$\begin{aligned} P_r\colon & \underset{\{p_{li}\},\{\rho_{li}\}}{\operatorname{argmin}} \quad \sum_{l\in K_r}\sum_{i=1}^{X}|\bar{S}_i|\,p_{li} \\ & s.t. \sum_{i=1}^{X}\rho_{li}\,|\bar{S}_i|\log_2\left(1+p_{li}\hat{\beta}_{li}\right)\geq R_l, \qquad l\in K_r \\ & \sum_{l\in K_r}\rho_{li}\leq 1, \qquad \rho_{li}\in\{0,1\}, \qquad 1\leq i\leq X, \end{aligned} \tag{4-29}$$

where

$$\hat{\beta}_{li}=\left(\prod_{m\in\bar{S}_i}\beta_{lm}\right)^{1/|\bar{S}_i|}. \tag{4-30}$$

Similarly to problem (4-17), the users' rate constraints and the objective function have been defined taking into account that the power within each block is constant. The approximation made in (4-25) results in (4-28), which exactly coincides with the SINR that is obtained in OFDM when the sequential channel assignment approach is implemented [25]. Hence, the solution of (4-29) can be indistinctly used in OFDM and FBMC systems. As it has been mentioned we propose to solve (4-29) by executing the LPSCA algorithm that is described in [25].

4.6 Numerical results

This section is devoted to evaluating the user selection and the power allocation algorithm proposed in Section 4.5. Regarding the system parameters, the bandwidth is 10 MHz, the number of subcarriers is $M=1024$ and the sampling frequency is set to 15.36 MHz so that the subcarrier spacing is $\Delta_f=15$ kHz. . Similar settings have been selecting in Section 3 (e.g. $\Delta_f=15$ kHz and B=1.4MHz). The number of users that are connected to the BS is equal to $N_U=10$ and they are uniformly distributed in a cell of radius $R=500$ m. Since the number of transmit antennas is set to $N_T=2$, only two users can be associated with each subcarrier. The thermal noise density is -174 dBm/Hz and the channel is modeled as a Rayleigh fading process with a power delay profile that follows the extended pedestrian A (EPA) channel model [30]. The path loss exponent is $\gamma=4$. It should be mentioned that the frequency selectivity of the EPA channel is such that the system model described in Section 4.2 is valid. In other words, the channel frequency response can be assumed flat at the subcarrier level. Concerning the air-interface, we consider FBMC and OFDM with a CP that encompasses $L_{CP}=M/14$ samples. To comply with the recommendations proposed by the Technical Specification Group for Radio Access Network of the 3GPP, only 600 out 1024 subcarriers are used to transmit data. Furthermore, subcarriers are gathered in RBs of 12 and, thus, there are $X=50$ RBs. In notation terms, let S_a denote the set whose elements are the indexes of those subcarriers that are active and $S_a(i)$ indicates the *i*th active subcarrier. Borrowing the notation from Section 4.5, the elements of subset $\bar{S}_i$ when OFDM is implemented are given by $\{S_a(1+12(i-1))\ldots S_a(12i)\}$, for $i=1,\ldots,50$. To prevent ICI from degrading the FBMC system performance, the blocks are separated by one subcarrier that is intentionally left empty. Then in FBMC the number of subcarriers that are able to convey data is extended from 600 to 650 and the subsets are generated as $\bar{S}_i=\{S_a(2+13(i-1))\ldots S_a(13i)\}$, for $i=1,\ldots,50$. Although inserting one guard band between blocks increases the out of band

radiation, the transmitted signal in FBMC still fits into the spectrum mask of the Universal Mobile Telecommunication System (UMTS) [31]. Actually, it is shown that the occupied subcarriers can be increased in 10 % without violating the spectrum mask.

The Figure 4-1 depicts the power that is required to schedule $N_U = 10$ users in 50 blocks of 12 subcarriers each. As it has been pointed out in Section 4.4, one subcarrier is left empty between blocks when the FBMC modulation is employed. It should be emphasizing that this does not mean that interference is removed, but the RBs do not interfere each other. The metric that is evaluated is $\sum_{l=1}^{N_U}\sum_{i=1}^{50} 1\, 2(1 + L_{CP}/M)2p_{li}$, where the power coefficients $\{p_{li}\}$ have been obtained after solving (4-29). The rate constraints have been set equal for all users as follows: $R_l = \alpha 600/N_U$, where α is the target spectral efficiency. It is important to recall that $L_{CP} = 0$ in FBMC systems, and because of that the transmitted power is reduced when compared to OFDM, as Figure 4-1 highlights.

In order to verify that both modulations are able to achieve similar rates we have depicted in Figure 4-2 the overall rate, which is defined as $\sum_{l=1}^{N_U}\sum_{q\in S_a} r_{lq}$. To this end we have used the exact rate definition given by (4-27). Nevertheless, the power of the residual interference plus noise in OFDM and FBMC systems is characterized by (4-26) and (4-24), respectively. In the light of the results of Figure 4-2 we can conclude that the assumption made in (4-25) does not have any negative impact as the overall rates achieved in both modulation schemes practically coincide. Another important aspect that is worth highlighting is that the relative difference between the exact sum-rate and the lower bound based on the geometric mean, which is formulated in (4-15), does not exceed 0.25% after solving (4-29). This holds true for OFDM and FBMC, which supports the simplifications proposed in this section. For the sake of the clarity in the presentation of the results the lower bound has not been included in Figure 4-2.

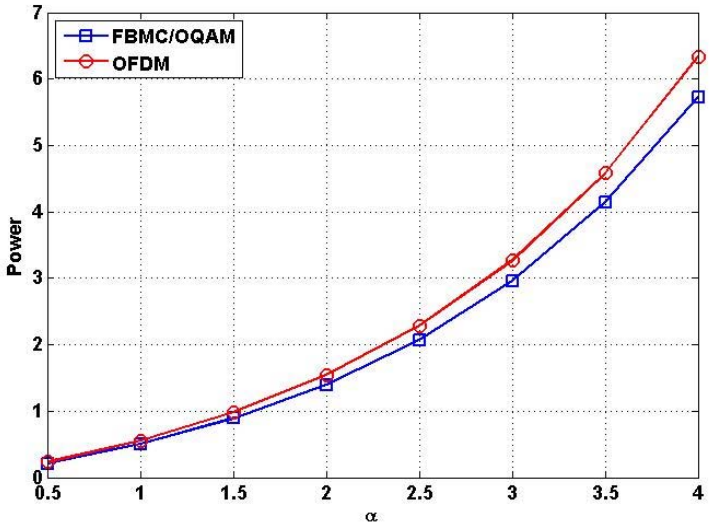


Figure 4-1: Power vs. spectral efficiency in OFDM and FBMC.

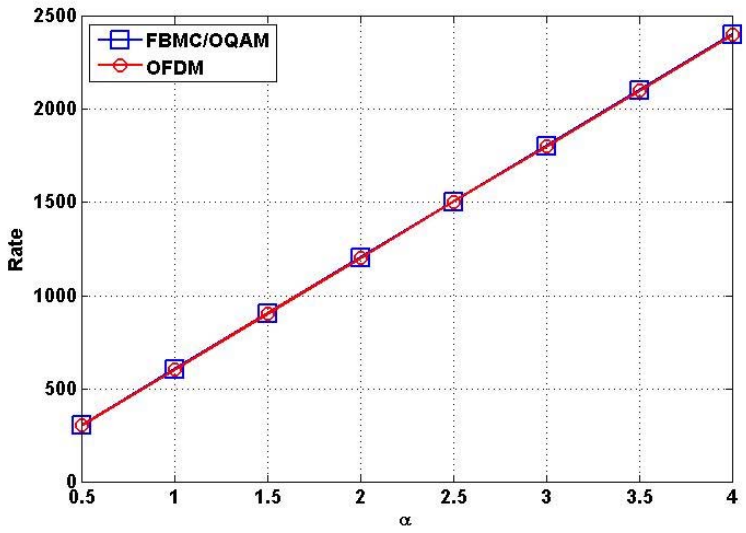


Figure 4-2: Overall rate vs. spectral efficiency in OFDM and FBMC.

4.7 Conclusions

This section shows that the joint optimization of transmit and receive beamforming, the channel assignment and the power allocation is very challenging in the FBMC context. The main reason stems from the fact that the received signals are subject to inter-user, inter-carrier and inter-symbol interference. By grouping subcarriers and keeping the same user selection and power allocation in each group we can exploit spatial diversity to allocate several users in the same frequency resources. It has been demonstrated that the margin adaptive problem in OFDM and FBMC systems can be sub-optimally solved resorting to the LPSCA method. Since no energy is wasted in the FBMC modulation scheme, this modulation is able to transmit the same amount of information as OFDM but using less power.

5. Conclusions

In this report, Radio Resource Management for PMR communications based on FBMC has been investigated. The different techniques presented are further improvements of D5.2, that take into account cross-layer interactions between RRM and both physical and application layers.

The problems considered in this deliverable are all very relevant to PMR communications: voice communications, buffer queue stabilization and energy consumption minimization, power consumption minimization in multiple antennas context.

First, the critical problem of voice communications for PMR communications in cell-based scenario has been studied in Section 2. The E-model has been chosen to model the MOS of voice communications. It provides an efficient way of involving application layer performance metrics in RRM algorithms. The proposed algorithm adapts the encoding rate of voice service depending on the end user's QoE. It provides a high capacity gain compared to non-adaptive algorithms.

In Section 3, the DUST algorithm, initially introduced in D5.2, has been further improved. It implies cross-layer optimization between RRM and the physical layer, with the objective to achieve a stable buffer queue for all users and to minimize energy consumption. The clustered scenario implies inter-cluster interference due to lack of synchronization. In this context, FBMC is known to achieve better performances than CP-OFDM. The achieved performances in terms of delay depend on the multi-carrier modulation and on the back-off algorithm. As the back-off algorithm leads to an orthogonal transmission, the performances are almost the same with all multi-carrier modulations (FBMC, CP-OFDM and Perfect Modulation). If the back-off algorithm is not used, the delay differences between the modulations increase.

Finally, Section 4 has introduced a scheduling algorithm for solving the margin adaptive problem MU SIMO with CP-OFDM and FBMC. In SIMO, inter-user, inter-carrier and inter-symbol interferences must be dealt with. In order to allocate several users in the same frequencies and thus exploit spatial diversity, it is necessary to group subcarriers and allocating the same users with the same power allocation within each subcarriers group. Then the margin adaptive problem can be solved sub-optimally using the LPSCA (linear programming based successive channel allocation) method. FBMC leads to a power consumption decrease compared to OFDM.

6. References

[1] Deliverable D5.2, “Novel algorithms description and performance evaluation for RRM in cell-based and ad-hoc PMR networks”, FP7 EMPHATIC project.

[2] Y. Wardi, B. Melamed, “Variational bounds and sensitivity analysis of traffic processes in continuous flow models”, *Springer’s Discrete Event Dynamic Systems*, vol. 11, no. 3, pp. 249–282, 2001.

[3] Y. Wardi, B. Melamed, “Loss volume in continuous flow models: Fast simulation and sensitivity analysis via IPA”, *Proceedings of 8th IEEE Mediterranean Conference on Control and Automation (MED 2000)*, July 2000.

[4] B. Liu,Y. Guo, J. Kurose.,D. Towsley. W.B. Gong, “Fluid simulation of large scale networks: Issues and tradeoffs”, *Proceedings of International Conference on Parallel and Distributed Processing Techniques and Applications*, June 1999.

[5] K. Kumaran, D. Mitra, “Performance and fluid simulations of a novel shared buffer management systems”, *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, vol. 11, no. 1, pp. 43–75, 2001.

[6] A. Yan, W.B. Gong, “Fluid simulation for high speed networks with flow-based routing”, *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1588–1599, 1999.

[7] ITU-T: *Methods for Subjective Determination of Transmission Quality*. Rec. P.800, 1996.

[8] ITU-T: *Single-ended method for objective speech quality assessment in narrow-band telephony applications*. Rec. P.563, 2005.

[9] ITU-T: The E-Model*: A Computational Model for Use in Transmission Planning*. Rec. G.107, 2012.

[10] S. Jadhav, H. Zhang, Z. Huang, “MOS-based Handover Protocol for Next Generation Wireless Networks”, *IEEE 26th International Conference on Advanced Information Networking and Applications*, pp. 479–486, 2012.

[11] J. Fitzpatrick, "An E-Model based adaptation algorithm for AMR voice calls," *Wireless Days (WD), IFIP* , vol., no., pp.1-6, 10-12 Oct. 2011.

[12] F. Mertz and P. Vary, “Efficient voice communication in wireless packet networks,” in *Proceedings of Sprachkommunikation*, ITG-Fachtagung, 2008, pp. 92–99. [Online]. Available: http://www.vdeverlag.de/proceedings-de/453120011.html

[13] L. Georgiadis, M. J. Neely, L. Tassiulas, ”Resource allocation and cross-layer control in wireless networks”, *Foundations and Trends in Networking*, vol.1, no. 1, 1144, 2006.

[14] S. Hayashi, ZQ Luo, “Spectrum Management for Interference-Limited Multiuser Communication Systems”, *IEEE Trans. Inform. Theory*, vol 55, no. 3, pp. 1153 – 1175, March 2009.

[15] M. J. Neely, *Dynamic Power Allocation and Routing for Satellite and Wireless Networks with Time Varying Channels*. PhD thesis, LIDS, MIT, Cambridge, MA, 2003.

[16] X. Tan, Z. Xiong, and Y. He, “Signal Attenuation-aware Clustering in Wireless Mobile Ad Hoc Networks”, *Journal of Networks*, vol. 8, no. 4, pp.796-803, Apr. 2013.

[17] R. Ghosh and S.Basagni, “Mitigating the impact of node mobility on ad hoc clustering,” *Wireless Communications and Mobile Computing*, vol. 8, no. 3, pp 295-308, Mar. 2008.

[18] Deliverable D9.1, “Definition and specification of hardware demonstrator and software simulator”, FP7 EMPHATICc project.

[19] Deliverable D5.1, “Study of the advantages of FBMC in RRM for PMR”, FP7 EMPHATIC project.

[20] P. Siohan, C. Siclet, N. Lacaille, “Analysis and design of OFDM/OQAM systems based on filterbank theory”, *IEEE Transactions on Signal Processing*, vol. 50, no. 5, pp. 1170 –1183, May 2002.

[21] B. Farhang-Boroujeny, “OFDM Versus Filter Bank Multicarrier,” *Signal Processing Magazine, IEEE*, vol. 28, no. 3, pp. 92 –112, May 2011.

[22] M. Caus, *MIMO Designs for filter bank multicarrier and multiantenna systems based on OQAM*, Ph.D. dissertation, Universitat Politecnica de Catalunya (UPC), 2013. [Online]. Available: http://theses.eurasip.org

[23] M. Payaro, A. Pascual-Iserte, A. Garcia-Armada, M. Sanchez-Fernandez, “Resource Allocation in Multi-Antenna MAC Networks: FBMC vs OFDM,” in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, May 2011.

[24] M. Shaat, F. Bader, “Computationally Efficient Power Allocation Algorithm in Multicarrier-Based Cognitive Radio Networks: OFDM and FBMC Systems,” *EURASIP JASP*, vol. 2010.

[25] M. Moretti and A. I. Perez-Neira., “Efficient Margin Adaptive Scheduling for MIMO-OFDMA Systems,” *Wireless Communications, IEEE Transactions on*, vol. 12, no. 1, pp. 278–287, 2013.

[26] M. Caus, A. I. Perez-Neira, M. Moretti, “SDMA for FBMC with block diagonalization,” in *Signal Processing Advances in Wireless Communications (SPAWC), IEEE 14th International Workshop on*, 2013.

[27] M. Bellanger, “Specification and design of a prototype filter for filter bank based multicarrier transmission.” *ICASSP*, 2001, pp. 2417–2420.

[28] Q. H. Spencer, A. L. Swindlehurst, M. Haardt, “Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels,” *Signal Processing, IEEE Transactions on*, vol. 52, no. 2, Feb. 2004.

[29] J. Huang, V. Subramanian, R. Agrawal, and R. Berry, “Downlink scheduling and resource allocation for OFDM systems,” *Wireless Communications, IEEE Transactions on*, vol. 8, no. 1, pp. 288–296, 2009.

[30] 3GPP TS 36.101 V11.8.0, “Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception (Release 11),” 2014.

[31] L. G. Baltar, D. S. Waldhauser, J. A. Nossek, “Out-of-band radiation in multicarrier systems: a comparison,” in *Multi-Carrier Spread Spectrum 2007*, ser. Lecture Notes Electrical Engineering. Springer Netherlands, 2007, vol. 1, pp. 107–116.

[32] X. Tan, Z. Xiong, and Y. He, “Signal Attenuation-aware Clustering in Wireless Mobile Ad Hoc Networks,” *Journal of Networks*, vol. 8, no. 4, pp. 796-803, 2013.

Glossary and Definitions

Acronym	Meaning
3GPP	3rd Generation Partnership Project
AMR	Adaptive Multi-Rate
BS	Base Station
CFM	Continuous Flow Modelling
CH	Cluster Head
CP	Cyclic Prefix
DMO	Direct Mode Operation
DUST	Distributed bUffer Stabilization RRM algorithm
FBMC	Filter Bank Multicarrier
HH	Hand Held
ICI	Inter-Channel Interference
LPSCA	Linear Programming based Successive Channel Allocation
MAC	Medium access control
MCS	Modulation and coding scheme
MISO	Multiple Input Single Output
MOS	Mean Opinion Score
MS	Mobile Station
MU	Multi-User
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OQAM	Offset Quadrature Amplitude Modulation
PESQ	Perceptual Evaluation of Speech Quality
PSQM	Perceptual Speech Quality Measure
PHY	Physical layer
PMR	Professional Mobile Radio
PM	Perfect Modulation
QoE	Quality of Experience
RRM	Radio Resource Management
RS	Relay Station
RB	Resource Block
SC-FDMA	Single Carrier – Frequency Division Multiple Access
SDMA	Space Division Multiple Access
SIMO	Single Input Multiple Output
TDMA	Time division multiple access