CONFIDENTIAL

FP7-ICT Future Networks
SPECIFIC TARGETTED RESEARCH PROJECT
Project Deliverable

PHYDYAS Doc. Number	PHYDYAS_006
Project Number	ICT - 211887
Project Acronym+Title	PHYDYAS - PHYsical layer for DYnamic AccesS and cognitive radio
Deliverable Nature	Report
Deliverable Number	D6.1
Contractual Delivery Date	January 1st, 2009
Actual Delivery Date	January 1st, 2009
Title of Deliverable	Duplexing and multiple access techniques, software description
Contributing Work Package	WP6
Project starting date; Duration	01/01/2008; 30 months
Dissemination Level	CO
Author(s)	Leonardo Baltar, Qing Bai, Josef Nossek (TUM-WP6 leader); Nikos Passas (RA-CTI); Ismael Gutiérrez, Carlos Bader (CTTC); Michel Terré (CNAM), Tilde Fusco, Mario Tanda (UNINA)

Abstract: In this report we derive the relations between FBMC and duplexing techniques like TDD and FDD. We determine the requirements of the front-end equipment when the concept of In-band FDD is employed. Then we consider the effects of time and frequency offset in the uplink and multiple access interference in the downlink. Finally, we contemplate the cross-layer issues of separate optimal resource allocation and scheduling, and joint optimal resource allocation/scheduling.

Contents

List of Acronyms

3GPP	3rd Generation Partnership Project
ADC	Analog to digital converter
AGC	Automatic Gain Control
ALC	Air Link Control
AMC	Adaptive Modulation and Coding
ARQ	Automatic Repeat Request
AWGN	Additive White Gaussian Noise
b2PFS	Buffer Based PFS
BE	Best Effort
BER	Bit Error Rate
BS	Base Station
BPSK	Binary Phase Shift Keying
CBR	Constant Bit Rate
CC	Convolutional Coding
CF	Continuously Feedback
CFO	Carrier Frequency Offset
CID	Connection Identifiers
CLARA	Cross-layer Assisted Resource Allocation
CNAM	Conservatoire National des Arts et Métiers
CP	Cyclic Prefix
CPE	Customer Premise Equipment
CP-OFDM	Cyclic Prefix based Orthogonal Frequency Division Multiplexing
CQI	Channel Quality Indicators
CQM	channel Quality Metrics
CSI	Channel State Information
CSIT	Channel State Information at the Transmitter
CTC	Convolutional Turbo Coding
DAC	Digital to Analog Converter
DL	Downlink
DL-MAP	Message defines the usage of the downlink intervals for a burst mode PHY
DMC	Discrete Memoryless Channel
DR	Dynamic Range
EDF	Earliest Deadline First
ertPS	extended real-time Polling Service
ESM	Effective SNR Mapping
ESNR	Effective SNR
FB	Filter Bank
FBMC	Filter Bank Multiple Carrier

FCH	Frame Control Header
FEC	Forward Error Correcting Code
FFT	Fast Fourier Transform
FDD	Frequency Division Duplexing
FDMA	Frequency Division Multiple Access
FTP	File Transfer Protocol
FUSC	Full Usage of Subcarriers
FIFO	First In First Out
HARQ	Hybrid Automatic Repeat Request
ICI	Interchannel Interference
IEEE	Institute of Electrical and Electronics Engineers
IF	Intermediate Frequency
IP	Internet Protocol
ISI	Intersymbol Interference
ISR	Interference-to-signal ratio
ITU	International Telecommunication Union
LNA	Low Noise Amplifier
LTE	Long Term Evolution
LO	Local Oscillator
MA	Margin Adaptation
MAC	Medium Access Control
MCS	Modulation and Coding Scheme
MDS	Minimum Detectable Signal
MIMO	Multiple Input Multiple Output
MRU	Minimum Resource Units
MS	Mobile Station
MTBA	Mixed TUSC and Band AMC
nrtPS	non-real Time Polling Service
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OS	Opportunistic Scheduling
PA	Power Amplifier
PA	Power Allocation
PC	Power Control
PDU	Packet Data Unit
PER	Packet Error Rate
PFS	Proportional Fair Scheduling
PHY	Physical Layer
PMP	Point to Multiple Points
PS	Physical Slots
PSD	Power Spectral Density
PUSC	Partial Usage of Subcarriers
QoS	Quality of Service

QPSK	Quadrature Phase Shift Keying
QAM	Quadrature Amplitude Modulation
RA	Resource Allocator
RF	Radio Frequency
RF	Requested Feedback
RCFO	Residual Carrier Frequency Offset
RTD	Round Trip Delay
RTG	Receive-Transmit Transition Gap
RTO	Residual Time Offset
RRA	Radio Resource Allocation
RRC	Radio Resource Controller
RRM	Radio Resource Management
rtPS	Real-Time Polling Service
SA	Subchannel Assignment
SF	Service Flow
SFID	Service Flow Identifiers
SNR	Signal to Noise Ratio
SS	Subscriber Station
TDD	Time Division Duplexing
TO	Time Offset
TSPS	Time Stamped Packets Scheduling
TTG	Transmit-Receive Transition Gap
TTI	Transmission Time Interval
TUSC	Tile Usage Subchannelization
UL	Uplink
UL-MAP	Defines the uplink usage in terms of offset of the burst relative to the allocation start time
UMTS	Universal Mobile Telecommunications System
UGS	Unsolicited Grant Service
VBR	variable bit rate
VoIP	Voice Over IP
WFQ	Weighted Fair Queuing
WiMAX	Worldwide Interoperability for Microwave Access
WP	Work Package
WWW	World Wide Web

Chapter 1

Introduction

1.1 Scope

In this report the ongoing research conducted in work package 6 is described. The different topics in this work package are

- investigation of duplexing techniques: comparison of FDD and TDD and their relative advantages and disadvantages with respect to FBMC and CP-OFDM
- a special FDD approach, i.e. In-band FDD has been investigated
- multiple access interference in the uplink has been investigated with different allocation schemes
- intercell interference tables for the downlink have been derived
- cross-layer issues design have been tackled with two different approaches.

Work on schemes for feeding back channel state information from the receiver to the transmitter has also been start but is not part of this report. It will be reported in the next deliverable of work package 6 together with all the other ongoing investigations.

Chapter 2

FBMC and duplexing schemes

2.1 Time vs. Frequency division duplexing

WiMAX prioritizes the time division duplexing (TDD) scheme [1], although frequency division duplexing (FDD) is also included in the IEEE 802.16 standard [2]. Standards like the LTE for UMTS from the 3GPP consortium include only the FDD scheme. In this section we will make an overview of both schemes considering the advantages and disadvantages of both schemes and in the following section we will point the characteristics of the FBMC system that need to be taken into account when employing both schemes.

2.1.1 Advantages and disadvantages of TDD

Both transmitter and receiver equipment share the same band in the spectrum, allowing a dynamic allocation of transmission rate between up- and downlink. The redistribution of the data rate is achieved by only altering the subframe durations. But two transition gaps need to be included between the subframes: The transmit/receive transition gap (TTG) is the guard interval between down- and uplink and the receive/transmit transition gap (RTG) is the guard interval between up- and downlink. Both intervals have to be long enough to include the equipment switch time and the propagation delay. Those transition gaps reduce the spectral efficiency.

TDD allows a reduction in the cost of the radio frequency (RF) front-end equipment since both transmitter and receiver share the same bandwidth and the same antenna, but in different periods of time [3]. In that way, only one local oscillator (LO) and one RF filter are necessary for both transmitter and receiver. The filter does not need to be very selective, since the out-of-band requirements are not as strict as in the FDD case. In any case, the out-of-band radiation needs to be considered, because it can interfere with nearby radio receivers.

A reduction in the overhead of training data is possible with a consequent increase in the spectrum efficiency, since the same band in the spectrum is used for both up- and downlink allowing the to rely on the channel reciprocity in low mobility scenarios. But a careful calibration procedure is necessary during the operation time to benefit from

that reciprocity, since the RF components have divergent characteristics in different radio equipments and because of synchronization issues, since all the equipments need to know exactly when to start transmitting.

2.1.2 Advantages and disadvantages of FDD

FDD allows a continuous transmission for both up- and downlink simplify any protocol of the upper layers that is dependent on negotiation between the equipments, since the answers to any consultation come timely. The channel state information also arrive quicker when any transmit processing is used. But since the bands stay fixed for both links, no flexible allocation of transmission rate is possible. Some spectral efficiency is lost because of the necessary feedback.

2.2 FBMC in TDD

Figure 2.1 depicts the allocation of time in one TDD frame for a CP-OFDM system. We

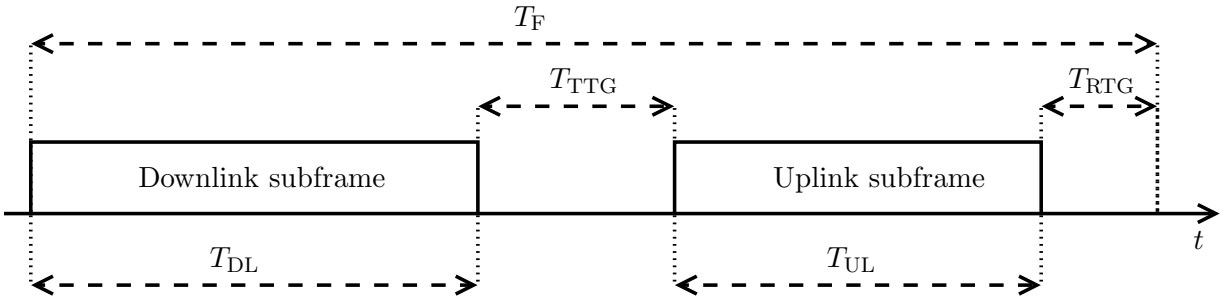


Figure 2.1: Representation of a TDD frame in an OFDM system

define the frame duration as $T_{\mathrm{F}} = T_{\mathrm{DL}} + T_{\mathrm{TTG}} + T_{\mathrm{UL}} + T_{\mathrm{RTG}}$ and it represents the period of time between the beginning of two downlink transmissions. Clearly, not all the duration of both down- and uplink subframes are filled with payload data: a part of the subframes is filled with the cyclic prefix. In other words, the spectral efficiency will be reduced depending on the length of the CP.

WiMAX defines 8 possible durations for T_{F}: from 2 ms to 20 ms, but only a duration of 5 ms is mandatory. In the case of CP-OFDM using a prefix of length L_{cp}, WiMAX also assumes that N_{S}[1] OFDM symbols - each with a duration of $T_{\mathrm{b}} + T_{\mathrm{g}}$, where T_{b} is the duration of one OFDM symbol and T_{g} is the duration of the CP - are included in one frame, leaving the rest of the time available for both TTG and RTG. We define the fraction of the subframes (or the fraction of the useful part of one frame) spent by the cyclic prefix as $N_{\mathrm{S}}T_{\mathrm{g}}/(T_{\mathrm{DL}} + T_{\mathrm{UL}})$.

In the case of the FBMC, besides the TTG and the RTG, a guard time is necessary to include the longer impulse response of the prototype. Those guard times are called pre-tail and post-tails, depending if they appear at the beginning or at the end of the subframe. Figure 2.2 shows an example of an output of the synthesis filter bank. It is clear that an smooth transition exist in opposition to CP-OFDM where the block starts abruptly[2].

Let us take as an example the output of a synthesis FB with M subcarriers and a prototype filter of length KM, where $K = 4$. We call N_{S} here the number of blocks inside one burst. Figure 2.3 depicts how the sequence of blocks overlap and which of them generate the tails. Each box represent a stream of M samples at the output of the synthesis FB. The numbers represent the index of the blocks at the input of the transmitter. We can see in this case, that each input block generates a sequence of $4M$ samples and that one

[1]We will assume here, that the number of symbols N_{S} inside one frame is obtained by the formula $N_{\mathrm{S}} = \left\lfloor \frac{T_{\mathrm{F}} F_{\mathrm{S}}}{M + L_{\mathrm{cp}}} - 1 \right\rfloor$.

[2]Here we are only looking at the output of the digital multicarrier modulation algorithm. Of course that the radiated signal signal in implemented systems using CP-OFDM will also present some kind of transition due to the digital and analog processing blocks

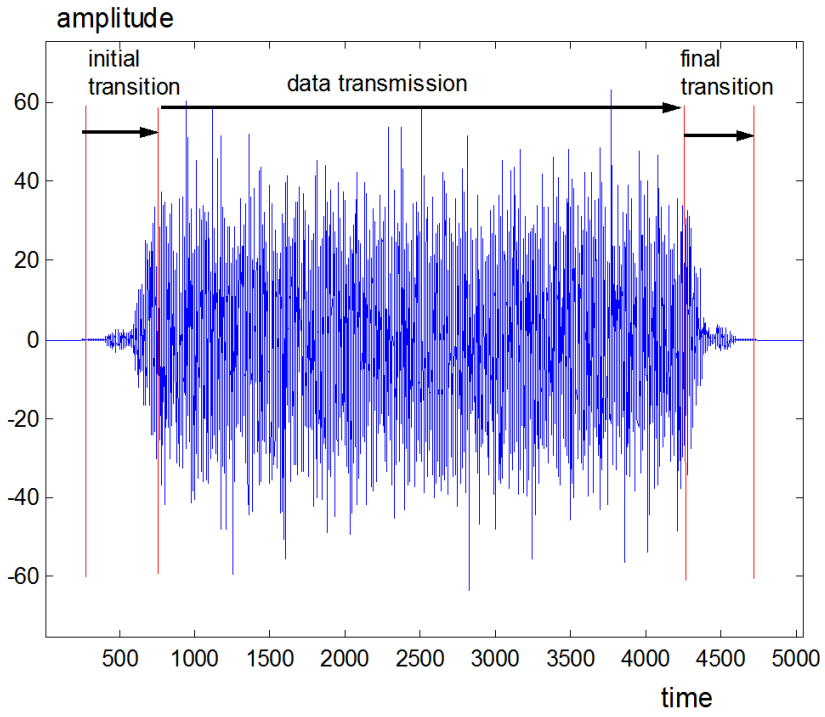


Figure 2.2: Example of an output of a synthesis FB

block will overlap with the following 3 blocks. Only the first two blocks are responsible for the pre-tail and only the last two blocks are responsible for the post-tail.

Figure 2.4 shows a the allocation of time inside an FBMC TDD frame. Now we define the FBMC frame duration as $T_{\mathrm{F}} = T'_{\mathrm{DL}} + T_{\mathrm{TTG}} + T'_{\mathrm{UL}} + T_{\mathrm{RTG}} + 2T_{\mathrm{P}}$, where $T_{\mathrm{P}} = (K-1)M/F_{\mathrm{S}}$ is the total time consumed by the pre- and the post-tails in one subframe. It is clear that the new T'_{DL} and T'_{UL} have to be shorter than in the CP-OFDM case presented before, i.e. $T_{\mathrm{DL}} = T'_{\mathrm{DL}} + T_{\mathrm{P}}$ and $T_{\mathrm{UL}} = T'_{\mathrm{UL}} + T_{\mathrm{P}}$. But again, in the CP-OFDM case a fraction of the subframes is composed by redundant data. We define as $2T_{\mathrm{P}}/(T_{\mathrm{DL}} + T_{\mathrm{UL}})$ the fraction of the subframes consumed by the tails of the filtering operation.

The existence of tails reduces the bandwidth efficiency, since during the pre- and post-tail intervals the analog front-ends from both transmitter and receiver need to be active for a longer time than the duration of the useful data.

2.2.1 Burst truncation

In Section 6 of [4] the effect of FBMC burst[3] truncation is analyzed and some techniques to shorten the pre- and post-tails are presented. As a drawback there is a degradation in some of the received symbols. The truncation of the pre-tail is also called memory pre-loading, because it is equivalent to pre-loading the memory elements of the polyphase components

[3] An FBMC burst in the context of time division duplexing is equivalent to a subframe.

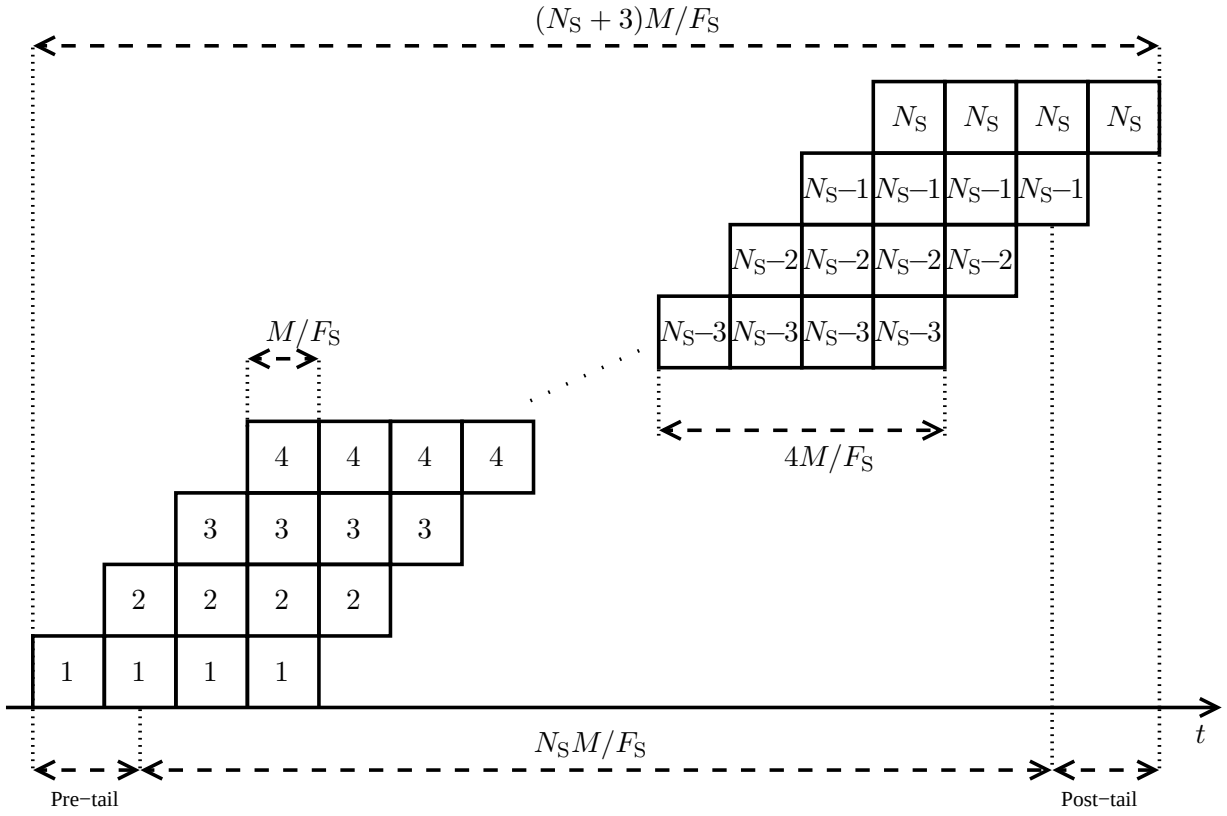


Figure 2.3: Illustration of a burst of a FBMC system for $K = 4$

with the input values, instead of waiting for the clock transitions to propagate them over the tap delay line.

A basic implementation of the burst truncation is through the control of the switching time of the transmitters analog front-end. The transmitted signal could start to be radiated just after all the memory elements are filled and the radiation could stop before all the memory elements are emptied. This would be equivalent to the truncation using a rectangular window. In [4] also a raised cosine window is presented, to smooth the transition after the truncation.

2.2.2 Examples

Let us consider, for example, a bandwidth of 10 MHz or $M = 1024$ subcarriers, which corresponds to a sampling frequency of 11.2 MHz. If we take the mandatory frame length of 5 ms, WiMAX defines N_S=47 as the number of OFDM symbols inside one frame for an $L_{cp} = M/8$ (Those symbols still have to be allocated to the subframes). In this case the total time reserved for the TTG and RTG is $(T_{TTG} + T_{RTG})$=0.7 ms. We will take this value as fixed for all the possible frame lengths to calculate the useful part of the frame $(T_{DL} + T_{UL})$.

Table 2.1 exhibits the fraction (in percents) of the subframes consumed by the pre- and

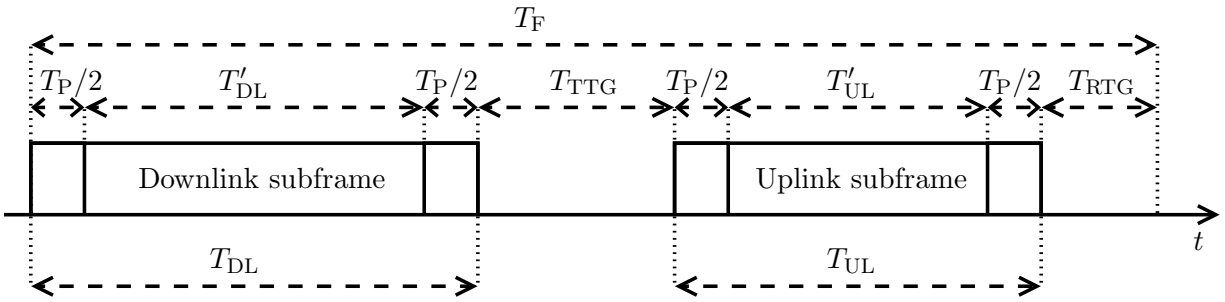


Figure 2.4: Representation of a TDD frame in an FBMC system

post-tails in the FBMC case for the various frame lengths and for two prototype lengths, namely $KM = 4096$ and $KM = 6144$. For each K the values without and with truncation are shown. Based on the results in Section 6 of [4], a reduction to 28 % of the tail length would give satisfactory results. For comparison reasons we also provide the percentage of the subframes used by the CP in the CP-OFDM system. From the Table 2.1 we can note

	FBMC				CP-OFDM			
	$K=4$		$K=6$		$L_{cp} =$			
T_F	full	trunc.	full	trunc.	$M/4$	$M/8$	$M/16$	$M/32$
2 ms	42.2 %	11.8 %	70.3 %	19.7 %	25 %	12.5 %	6.25%	3.125%
2.5 ms	30.5 %	8.5 %	50.8 %	14.2 %				
4 ms	16.6 %	4.6 %	27.7 %	7.7 %				
5 ms	12.7 %	3.5 %	21.3 %	5.9 %				
8 ms	7.5 %	2.1 %	12.5 %	3.5 %				
10 ms	5.9 %	1.6 %	9.8 %	2.7 %				
12 ms	4.8 %	1.3 %	8.1 %	2.3 %				
20 ms	2.8 %	0.8 %	4.7 %	1.3 %				

Table 2.1: Percentage of the frame duration consumed by the pre- and post-tails in the FBMC case and by the cyclic prefix in the CP-OFDM case.

that in the FBMC system the lost in spectral efficiency is dependent on the frame length and on the prototype length, in contrast to CP-OFDM, where the cyclic prefix determines that. We can see that, when a longer prefix is used in CP-OFDM and a shorter prototype is employed in the FBMC, the latter is always more efficient. At the other side, for shorter prefixes and longer prototypes, CP-OFDM is more efficient.

It is worth mentioning that, for a complete analysis of the spectral efficiency, the size of the burst truncation will determine how many blocks can be inserted in the two subframes. Besides that, it has to be taken into account the number of active subcarriers, since FBMC permits a to increase that number in relation to CP-OFDM.

2.3 FBMC in FDD

The FBMC system should allow a reduction in the frequency guard band between up- and downlink and between two adjacent channels and/or operators, since the shaping of the prototype filter significantly reduces the out-of-band radiation in comparison to CP-OFDM.

To illustrate the reduction of the out-of-band radiation, Figure 2.5 depicts the power spectral density (PSD) for both FBMC and CP-OFDM systems with 1024 subcarriers, where 840 are considered active. In WiMAX context, that corresponds to a bandwidth of 10 MHz and a sampling rate of 11.2 MHz. The prototype employed was designed using the frequency sampling method of [5] (CNAM prototype) with length $4M$.

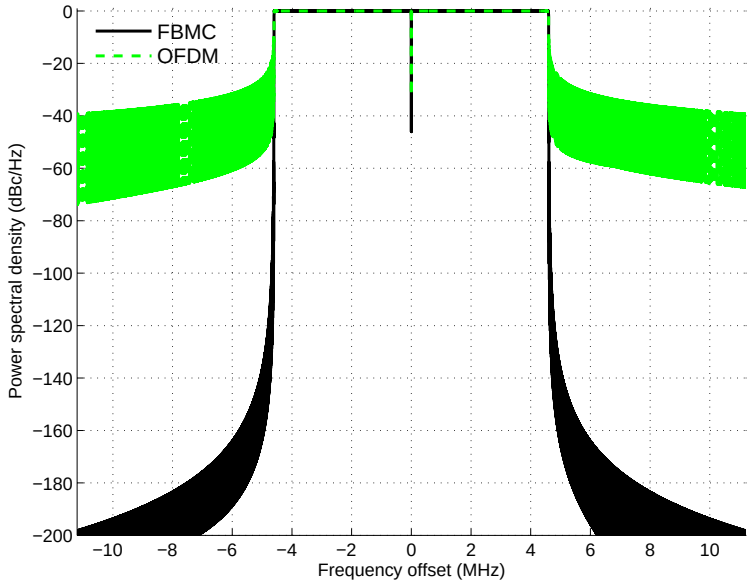


Figure 2.5: PSD of FBMC and CP-OFDM systems for $M = 1024$ and $N_{\text{used}} = 840$.

The PSD represented in Figure 2.5 does not take into account neither the digital post-processing nor the analog processing. Usually, the digital to analog converter and the power amplifier will not work exclusively in the linear region, resulting in a compression of the signal amplitude. The consequence of this non-linearity is the so called spectral regrowth. In other words, the out-of-band radiation will increase and both FBMC and CP-OFDM will show less attenuation in adjacent channels. If the amplitude clipping is modeled by a series expansion, the resulting out-of-band radiation is composed by the attenuation seen in Figure 2.5 plus the higher order elements of the in-band signal. In FBMC the attenuation of the undistorted signal is so low that it will have almost no influence in the final signal. But in CP-OFDM the attenuation of the undistorted signal is already too low, adding to the spectral regrowth.

2.4 In-band FDD

2.4.1 Introduction

In-band FDD (IB-FDD) [6] is a variation of the frequency division duplex (FDD) concept. It is a very attractive scheme for the dynamic allocation of bandwidth between downlink and uplink, when a continuous block of spectrum is available. IB-FDD shares some similarities with TDD, for example, it allows variable data rate for both downlink and uplink and relies only on one LO. The IB-FDD scheme also provides the same advantages in a Hybrid-FDD[4] framework [3].

Given a continuous piece of spectrum for both downlink and uplink, the FBMC system should be able to fill that bandwidth dynamically, depending on the required data rate. In a specific equipment (mobile or base station) the transmitter has a synthesis FB with M subcarriers, where a chunk of N_{Tx} subcarriers are filled with data and pilots. The receiver is composed by an analysis FB also with M subcarriers, but only a chunk of N_{Rx} subcarriers is demodulated and decoded. Obviously the total number of used subcarriers is $N_{\mathrm{Tx}} + N_{\mathrm{Rx}} \leq M$.

Figure 2.6 shows a simplified model of the analog front-end of a general wireless device designed for an IB-FDD scheme. At the transmitter part, after the digital to analog

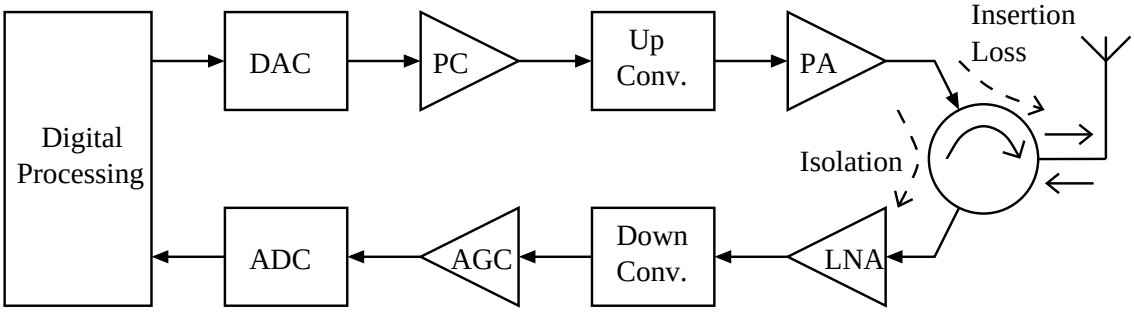


Figure 2.6: Simplified model of the analog front end.

conversion (DAC), the signal passes through the power control (PC) device and is up converted to the RF, power amplified (PA) and fed into the antenna through a duplexer. At the receiver part, the signal coming from the antenna is fed into the low noise amplifier (LNA) through the duplexer, down converted to the baseband, amplified by the automatic gain control (AGC) and analog to digital converted (ADC). Depending on the technology employed, the duplexer receives different names, like circulator, directional coupler, hybrid, etc. In Figure 2.6 the duplexer is represented by a circulator. The up and down conversion can also be executed in stages, i.e. the signal is converted to an intermediate frequency (IF) before being converted to the RF or baseband.

[4]In a Hybrid-FDD scheme with frequency division multiple access, the MSs operate in TDD and the BS operates in FDD. For example, within one cell in a certain instant some MSs are only transmitting while some are only receiving and the BS is transmitting and receiving simultaneously.

In an usual FDD scheme an analog low-pass, high-pass or band-pass filter is placed between the PA/LNA and the duplexer, or the filters are building blocks of the duplexer. But to allow maximum flexibility in the bandwidth allocation between the downlink and the uplink, no fixed analog filters are used and the transmitted signal flows into the receiver attenuated only by the isolation of the duplexer.

Ideally, a duplexer has an insertion loss equal to zero and a perfect isolation, i.e. no leakage is observed. Would this be true, IB-FDD would be simply implementable.

Unfortunately, a commercially available duplexer has a minimum insertion loss of 0.2 dB and a maximum isolation of 25 dB. The insertion loss implies in a lower efficiency for the transmitter, but does not influence the requirements of the receiver. However, the non-ideal isolation has a big impact on the design of the analog front-end components, the ADC and on the admitted level of out-of-band radiation. The objective of this work is to quantify these requirements in a framework of FBMC systems and using the WiMAX specifications.

In a system based on cyclic prefix orthogonal frequency division multiplexing (CP-OFDM) the IB-FDD scheme would not come into consideration, unless a measure is taken to reduce the out-of-band radiation with its respective increase in complexity. We will later show, that the high attenuation in the stop-band of the prototype filter allows an FBMC system to fulfill that requirements. It is worth noting that in cases when analog filters are employed, the FBMC system allows to relax the attenuation specifications.

Section 2.4.2 lists the relevant data extracted from [1] and [2] to derive the requirements of the impacted components of the analog front-end, the ADC and the prototype filter. In Section 2.4.4 we give some examples considering both base station (BS) and mobile station (MS) equipment, and also different bandwidth allocation scenarios. In Section 2.4.5 some conclusions are drawn.

2.4.2 IEEE 802.16 and WiMAX requirements

2.4.2.1 Transmitter requirements

2.4.2.1.1 Mobile station (MS) transmitter output power

- Customer premise equipment (CPE) outdoor system with LOS: $P_{\mathrm{Tx}} = 20$ dBm (100 mW)

- Indoor systems inside buildings: 24 dBm (250 mW) $\leq P_{\mathrm{Tx}} \leq$ 27 dBm(0.5 W)

2.4.2.1.2 Base station (BS) transmitter output power

- 30 dBm (1 W) $\leq P_{\mathrm{Tx}} \leq$ 43 dBm (20 W)

2.4.2.2 Receiver requirements

2.4.2.2.1 Receiver sensitivity in dBm

$$R_{\mathrm{S}} = -114 \text{ dBm} + \mathrm{SNR}_{\mathrm{Rx}} - 10\log_{10}(R) + 10\log_{10}\left(\frac{F_{\mathrm{S}} N_{\mathrm{Rx}}}{M}\right) + \rho + \alpha \qquad (2.1)$$

where

- $\mathrm{SNR}_{\mathrm{Rx}}$ in dB is the minimum receiver signal-to-noise ratio requirement for a coded BER of 10^{-6} as in Table 2.2 for convolutional coding (CC) in Table 2.3 for convolutional turbo coding (CTC).
- R is the repetition factor, as described in Section 8.4.9 of [2]
- F_{S} is the sampling frequency in MHz as defined in Section 8.4.2.4 of [2]
- N_{Rx} is the number of used subcarriers in the received signal.
- M is the total number of subcarriers
- ρ is the implementation loss in dB, which includes non-ideal receiver effects such as channel estimation errors, tracking errors, quantization errors and phase noise. The assumed value is 5 dB.
- α is the receiver noise figure. The assumed value is 8 dB.

Obs.: Classically the noise floor (or minimum detectable signal (MDS)) γ is defined in dBm as

$$\gamma = 10\log_{10}\left(\frac{kTB_{\mathrm{Rx}}}{1 \text{ mW}}\right) + \alpha = -174 + 10\log_{10} B_{\mathrm{Rx}} + \alpha, \qquad (2.2)$$

where $k = 1.38\times10^{-23}$ W/K is the Boltzmann´s constant, $T = 290$ K (16.85 $^{\mathrm{o}}$C) is the temperature and with B_{Rx} given in Hz. The noise figure α is defined as $\alpha = 10\log_{10}\left(\frac{\mathrm{SNR}_{\mathrm{in}}}{\mathrm{SNR}_{\mathrm{out}}}\right)$ dB, where $\mathrm{SNR}_{\mathrm{in}}$ and $\mathrm{SNR}_{\mathrm{out}}$ are the signal-to-noise ratios at the input and output of the analog signal chain. Since in the IEEE 802.16 standard B_{Rx} is given in MHz, 60 dB are added resulting in -114 dBm. Using (2.2) the receiver sensitivity can be rewritten as

$$R_{\mathrm{S}} = \gamma + \mathrm{SNR}_{\mathrm{Rx}} - 10\log_{10}(R) + \rho. \qquad (2.3)$$

2.4.2.2.2 BS receiver maximum input signal The BS receiver shall be capable of decoding a maximum on-channel signal of -45 dBm (30 nW).

2.4.2.2.3 MS receiver maximum input signal The MS receiver shall be capable of decoding a maximum on-channel signal of -30 dBm (1 μW).

2.4.2.2.4 BS receiver maximum tolerable signal The BS receiver shall tolerate a maximum signal of -10 dBm (0.1 mW) without damage.

2.4.2.2.5 MS receiver maximum tolerable signal The BS receiver shall tolerate a maximum signal of 0 dBm (1 mW) without damage.

2.4.2.2.6 Minimum receiver signal-to-noise ratio The minimum receiver signal-to-noise ratio requirements for a coded bit error rate of 10^{-6} are given in Table 2.2 for convolutional coding and in Table 2.3 for convolutional turbo coding. They are reproductions from Tables 4.2.3 in [1] and 338 in [2].

Modulation	**Coding rate**	$\mathrm{SNR_{Rx}}$ **in dB**
QPSK	1/2	5
	3/4	8
16-QAM	1/2	10.5
	3/4	14
64-QAM	1/2	16
	2/3	18
	3/4	20

Table 2.2: Minimum $\mathrm{SNR_{Rx}}$ for CC

Modulation	**Coding rate**	$\mathrm{SNR_{Rx}}$ **in dB**
QPSK	1/2	2.9
	3/4	6.3
16-QAM	1/2	8.6
	3/4	12.7
64-QAM	1/2	13.8
	2/3	16.9
	3/4	18
	5/6	19.9

Table 2.3: Minimum $\mathrm{SNR_{Rx}}$ for CTC

2.4.3 Complementary definitions

2.4.3.1 Transmit and receive bandwidth

- $B_{\mathrm{Rx}} = \frac{F_{\mathrm{S}} N_{\mathrm{Rx}}}{M}$ is the bandwidth of the receive signal in MHz.
- $B_{\mathrm{Tx}} = \frac{F_{\mathrm{S}} N_{\mathrm{Tx}}}{M}$ is the bandwidth of the transmit signal in MHz.

Since $N_{\mathrm{Tx}} + N_{\mathrm{Rx}} \leq M$ it turns out that $B_{\mathrm{Tx}} + B_{\mathrm{Rx}} \leq F_{\mathrm{S}}$.

2.4.3.2 ADC parametrization

Let us assume that the quantization noise is white, uniformly distributed, has zero mean and variance $\sigma_{\mathrm{q}}^2 = \frac{\Delta^2 N_{\mathrm{Rx}}}{12M}$, since the received signal is confined into the bandwidth B_{Rx} and Δ is the step-size [7].

If we assume that the received signal has a Gaussian power density function (PDF), it can be said that 99.993% of the signal amplitude swings inside the interval $[-4\sigma_{\mathrm{s}}, 4\sigma_{\mathrm{s}}]$, where σ_{s}^2 is the signal variance. Setting the full-scale level of the ADC to $4\sigma_{\mathrm{s}}$, the step-size is given by $\Delta = \frac{8\sigma_{\mathrm{s}}}{2^\beta}$, where β is the number of bits of the ADC. Consequently, the ratio between the variance of the received signal and the variance of the quantization noise is given by

$$\frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{q}}^2} = \frac{2^{2\beta} 3M}{16 N_{\mathrm{Rx}}}. \tag{2.4}$$

The dynamic range (DR) at the input of the LNA is defined as $\mathrm{DR} = 10\log_{10}\frac{\sigma_{\mathrm{s}}^2}{\sigma_{\mathrm{n}}^2}$, where $\sigma_{\mathrm{n}}^2 = \sigma_{\mathrm{LNA}}^2 + \sigma_{\mathrm{q}}^2$ and σ_{LNA}^2 is the variance of the noise portion that does not include the quantization noise. The signal variance is then

$$\sigma_{\mathrm{s}}^2 = \sigma_{\mathrm{n}}^2 10^{\frac{\mathrm{DR}}{10}}. \tag{2.5}$$

If we assume that the implementation loss is caused only by the quantization noise inserted by the ADC, in other words, that the ADC may reduce the dynamic range by ρ dB, we have that $\rho = 10\log_{10}\frac{\sigma_{\mathrm{n}}^2}{\sigma_{\mathrm{LNA}}^2}$ or that

$$\sigma_{\mathrm{q}}^2 = \sigma_{\mathrm{LNA}}^2 (10^{\frac{\rho}{10}} - 1). \tag{2.6}$$

Using (2.5) and (2.6) in (2.4) and solving for β, the number of bits of the ADC is given by

$$\beta = \left\lceil \frac{\mathrm{DR} + \rho - 10\log_{10}\left(10^{\frac{\rho}{10}} - 1\right) - 10\log_{10}\left(\frac{3M}{16 N_{\mathrm{Rx}}}\right)}{20\log_{10} 2} \right\rceil \tag{2.7}$$

where the operator $\lceil x \rceil$ is the smallest integer not less than x.

2.4.4 Examples

The examples are divided in two sections: In Section 2.4.4.1 a symmetric bandwidth allocation is made between uplink and downlink, in other words, both links get the same bandwidth. In Section 2.4.4.2 an asymmetric bandwidth is considered. In that case 90 % of the bandwidth is given to one link and 10 % to the other.

Each section has four examples, where two involve the base station and two the mobile station. For each station a best case and a worst case are considered. The worst case would represent a pessimistic situation where the stations are so far from each other, or the channel has a such a deep fading and/or shadowing, that the received signal arrives

with the minimum power necessary for a fixed BER and the transmit power is the highest possible. In the other extreme are the best case examples, that would represent a scenario where the stations are so close to each other, that the signal arrives with the maximum minimum power for a fixed BER and the transmit power is the lowest possible.

In fact, the optimum situation from the receiver perspective is when one link gets the whole bandwidth, viz. the transmitter at the same equipment is simply turned off, and the signal arrives with the maximum allowed power without damage to the equipment. But in that case we could not classify it as a frequency duplexing scheme, since the data flow is unidirectional in a certain time instant.

2.4.4.1 Symmetric bandwidth

In the following examples we assume a total available bandwidth of 10 MHz for both downlink and uplink. In this case, according to WiMAX, the sampling frequency should be $F_{\mathrm{S}} = 11.2$ MHz and the total number of subcarriers should be $M = 1024$, from which $N_{\mathrm{Tx}} = 420$ subcarriers are used by the transmitter and $N_{\mathrm{Rx}} = 420$ are used by the receiver. That results in $B_{\mathrm{Rx}} = B_{\mathrm{Tx}} = 4.6$ MHz. We do not consider here the application of a repetition code, i.e. $R = 1$, since they are not used for data transmission.

2.4.4.1.1 BS worst case: minimum receive power and maximum transmit power Let us consider a BS with full power downlink transmission and a minimum power uplink signal reception. In this case the BS is transmitting with $P_{\mathrm{Tx}} = 43$ dBm. We can see that the signal leakage from the transmitter to the receiver has a power of 18 dBm. It is clear that this value is even higher than the maximum input power allowed without damaging the receiver specified in WiMAX. To guarantee the detection, an even lower input power is specified.

If we assume that the MS is using a CTC with rate 1/2 and QPSK modulation in each subcarrier, for a coded BER of 10^{-6}, Table 4.2.3 of [1] specifies an $\mathrm{SNR}_{\mathrm{Rx}} = 2.9$ dB. Inserting that values in (2.1), we end up with a minimum received power of $R_{\mathrm{S}} = -91.5$ dBm. The dynamic range of the LNA input signal specified for that received power is of 46.5 dB. But considering the suggested setup, the necessary dynamic range would be of 109.5 dB.

It could be said that, without any analog filtering in the path between the PA and the LNA, either a duplexer with better isolation or/and a LNA with a higher allowed input power and input dynamic range would be necessary. Another alternative is to reduce the maximum cell radius and consequently the transmit power.

Assuming that the duplexer and the LNA fulfills the requisites we can see what happens with the interference between Tx and Rx. In Figure 2.7 is depicted the PSD of the signal at the input of the LNA. Both leaked Tx signal and low power Rx are represented. The prototype employed is designed with the frequency sampling method of [5] (CNAM prototype) with length $4M + 1$. It is evident that the interference level of the Tx signal is too high. To keep a interference level below 10 dB some subcarriers need to be left

empty between Tx and Rx. In this example, 40 subcarriers would be enough to keep that requisite as depicted in Figure 2.8.

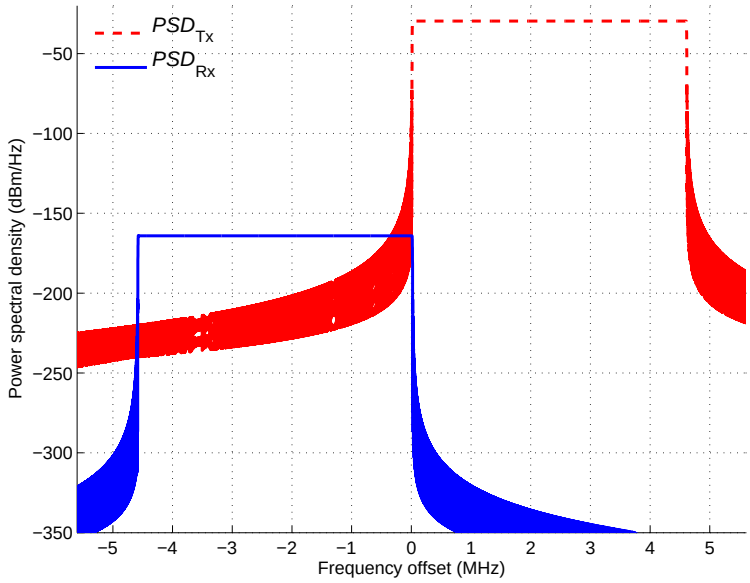


Figure 2.7: PSD of the LNA input signal. BS worst case, symmetric bandwidth, 1 free subcarrier between Tx and Rx and CNAM prototype.

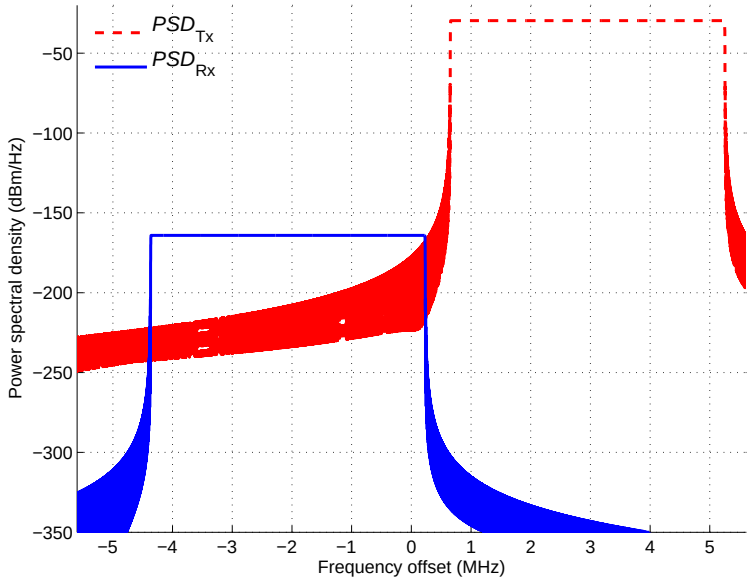


Figure 2.8: PSD of the LNA input signal. BS worst case, symmetric bandwidth, 40 free subcarriers between Tx and Rx and CNAM prototype.

Let us now look how many bits the ADC needs assuming that the LNA fulfills the

requirements of dynamic range. We consider that the only contribution to the implementation loss comes from the ADC, i.e. all other imperfections are negligible. In that case we can calculate the number of bits necessary for the dynamic range calculated and assuming $\rho = 5$ dB and by using (2.7). The specified dynamic range of 46.5 dB gives a total of 9 bits and the dynamic range of 109.5 dB requires 20 bits.

2.4.4.1.2 MS worst case: minimum receive power and maximum transmit power Now we consider a MS with full power uplink transmission and a minimum downlink signal reception. In that case the MS is transmitting with $P_{\mathrm{Tx}} = 25$ dBm. Now the signal leaking from the transmitter to the receiver has a power in the of 0 dBm. This value is still higher than the maximum input power allowed for signal detection specified in WiMAX.

If we assume that the BS is using a CTC with rate 1/2 and QPSK modulation in each subcarrier, we have again a $\mathrm{SNR}_{\mathrm{Rx}} = 2.9$ dB for a coded BER of 10^{-6}. That gives again an $R_{\mathrm{S}} = -91.5$ dBm. Now the dynamic range of the LNA input signal specified for that received power is of 61.5 dB. But in our setup, the necessary dynamic range would be of 91.5 dB.

Again either a duplexer with better isolation or/and a LNA with a higher allowed input power and input dynamic range would be necessary, but in the MS case this is not so extreme as in the BS, because of the lower transmit power.

In this case, the guard band between up- and downlink comprises 23 subcarriers and the ADC needs 12 bits for the WiMAX specifications and 17 bits for the IB-FDD specifications.

2.4.4.1.3 BS best case: maximum received power and minimum transmit power In an optimistic scenario, the BS will transmit with a power of $P_{\mathrm{Tx}} = 30$ dBm, i. e. with the lowest transmit power, and the received signal has the maximum minimum power for a BER of 10^{-6}. In this case, the leakage is of 5 dBm.

For this configuration the MS uses at the uplink a CC with rate 3/4 and 64-QAM modulation and the signal to noise ratio is $\mathrm{SNR}_{\mathrm{Rx}} = 20$ dB. This would result in a received signal with power $R_{\mathrm{S}} = -74.4$ dBm. The WiMAX dynamic range in this case is 29.4 dB and the DR necessary for IB-FDD would be 79.4 dB.

Here 13 subcarriers would be enough to isolate the links and the ADC would need 6 bits in a WiMAX equipment and 15 bits in an IB-FDD ready equipment.

2.4.4.1.4 MS best case: maximum received power and minimum transmit power Again, in an optimistic scenario, the MS is transmitting with a power of $P_{\mathrm{Tx}} = 20$ dBm , resulting in a leakage of -5 dBm.

The BS would be using a CC with rate 3/4 and 64-QAM modulation, resulting in an $\mathrm{SNR}_{\mathrm{Rx}} = 20$ dB for a coded BER of 10^{-6} and in an $R_{\mathrm{S}} = -74.4$ dBm. The dynamic range for WiMAX would be 44.4 dB and for IB-FDD 69.5 dB.

Figure 2.9 shows the PSD for this example. It can be seen that 9 subcarriers are needed between transmit and receive signal. The ADC resolution would be 9 bits for WiMAX and

13 bits to employ IB-FDD.

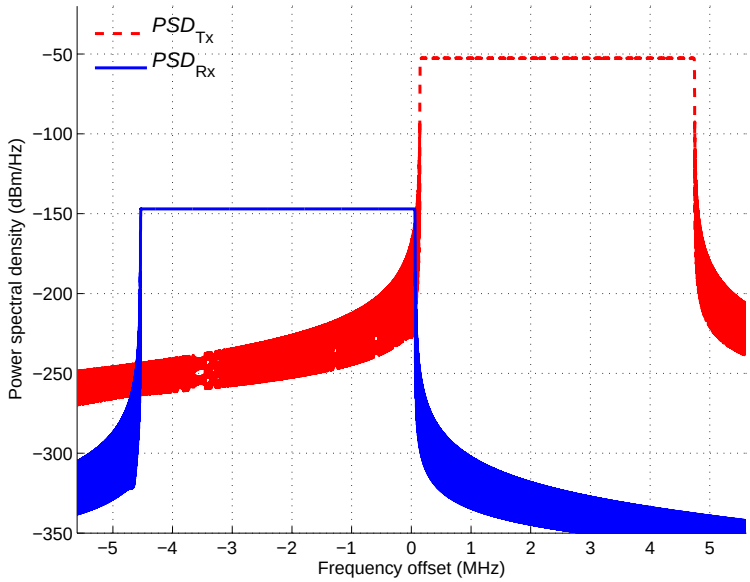


Figure 2.9: PSD of the LNA input signal. MS best case, symmetric bandwidth, 9 free subcarriers between Tx and Rx and CNAM prototype.

A summary of the examples for a symmetric bandwidth allocation is given in Table 2.4.

Parameter\Case	BS worst	MS worst	BS best	MS best
P_{Tx}	43 dBm	25 dBm	30 dBm	20 dBm
$\mathrm{SNR}_{\mathrm{Rx}}$	2.9 dB	2.9dB	20 dB	20 dB
R_{s}	-91.5 dBm	-91.5 dBm	-74.4 dBm	-74.4 dBm
DR WiMAX	46.5 dB	61.5 dB	29.4 dB	44.4 dB
DR IB-FDD	109.5 dB	91.5 dB	79.4 dB	69.5 dB
Guard Band	40 subch.	23 subch.	13 subch.	9 subch.
ADC WiMAX	9 bits	12 bits	6 bits	9 bits
ADC IB-FDD	20 bits	17 bits	15 bits	13 bits

Table 2.4: Summary of the examples for symmetric bandwidth allocation

2.4.4.2 Asymmetric bandwidth

We assume again a total available bandwidth of 10 MHz for both downlink and uplink. The sampling frequency remains $F_{\mathrm{S}} = 11.2$ MHz and the total number of subcarriers $M = 1024$. But now the bandwidth of uplink and downlink are asymmetric. Again, we do not consider the use of a repetition code, i.e. $R = 1$.

2.4.4.2.1 BS worst case: minimum receive power/bandwidth and maximum transmit power/bandwidth Here 90% of the subcarriers are used in the downlink and 10% are used in the uplink, i.e. $N_{\mathrm{Tx}} = 756$ and $N_{\mathrm{Rx}} = 84$, corresponding to $B_{\mathrm{Tx}} = 8.75$ MHz and $B_{\mathrm{Rx}} = 920$ kHz. The transmit power is $P_{\mathrm{Tx}} = 43$ dBm and the uplink signal arrives with $R_{\mathrm{S}} = -98.5$ dBm. Considering this, the dynamic range for WiMAX would be 53.5 dB and for IB-FDD 116.5 dB.

The PSD for this example is represented in Figure 2.10, where 60 subcarriers are left unused between the links. For this scenario, the ADC needs 10 bits for the WiMAX

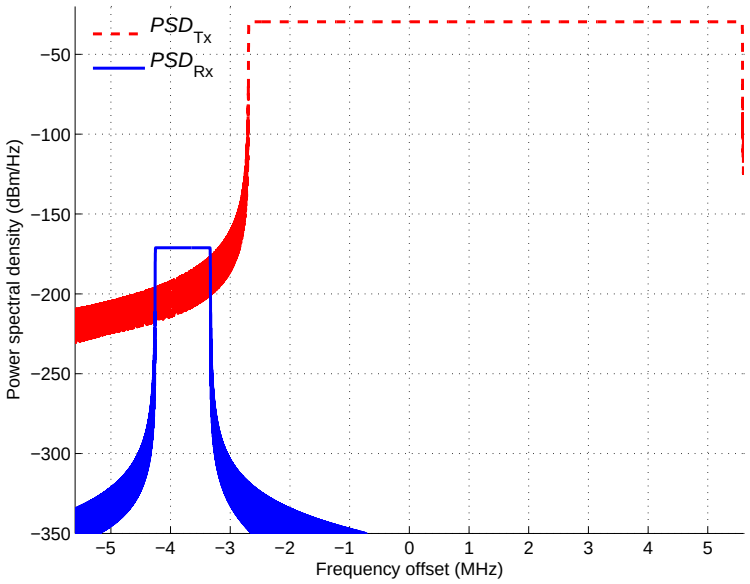


Figure 2.10: PSD of the LNA input signal. BS worst case, asymmetric bandwidth, 60 free subcarrier between Tx and Rx and CNAM prototype.

specifications and 21 bits for the deployment of IB-FDD.

2.4.4.2.2 MS worst case: minimum receive power/bandwidth and maximum transmit power/bandwidth Now 90% of the subcarriers are used in the uplink and 10% are used in the downlink, i.e. $N_{\mathrm{Tx}} = 756$ and $N_{\mathrm{Rx}} = 84$, corresponding to $B_{\mathrm{Tx}} = 8.75$ MHz and $B_{\mathrm{Rx}} = 920$ kHz, and the transmit power is $P_{\mathrm{Tx}} = 25$ dBm and the downlik signal arrives with $R_{\mathrm{S}} = -98.5$ dBm. Then, the dynamic range for WiMAX would be 68.5 dB and for IB-FDD 98.5 dB.

At least 30 empty subcarriers are necessary between the links and the resolution of the ADC would be of 13 bits for WiMAX and 18 bits for the deployment of IB-FDD.

2.4.4.2.3 BS best case: maximum receive power/bandwidth and minimum transmit power/bandwidth In this example, 90% of the subcarriers are used in the uplink and 10% are used in the downlink, i.e. $N_{\mathrm{Rx}} = 756$ and $N_{\mathrm{Tx}} = 84$, corresponding

to $B_{\mathrm{Rx}} = 8.75$ MHz and $B_{\mathrm{Tx}} = 920$ kHz. The transmit power is $P_{\mathrm{Tx}} = 30$ dBm and the receive power is $R_{\mathrm{S}} = -71.8$ dBm. This gives a dynamic range for WiMAX of 26.8 dB and for IB-FDD 76.8 dB.

In this example 13 empty subcarriers suffice to isolate the links and an ADC with 6 bits for WiMAX and 14 bits for IB-FDD would tolerate the necessary dynamic range.

2.4.4.2.4 MS best case: maximum receive power/bandwidth and minimum transmit power/bandwidth Here 90% of the subcarriers are used in the downlink and 10% are used in the uplink, i.e. $N_{\mathrm{Rx}} = 756$ and $N_{\mathrm{Tx}} = 84$, corresponding to $B_{\mathrm{Rx}} = 8.75$ MHz and $B_{\mathrm{Tx}} = 920$ kHz. The transmit power is again $P_{\mathrm{Tx}} = 30$ dBm and the receive power is again $R_{\mathrm{S}} = -71.8$ dBm. This implies in a dynamic range for WiMAX of 41.8 dB and for IB-FDD 66.8 dB.

Figure 2.11 shows the PSD for this example. It can be seen that 8 subcarriers between the links would be enough. An ADC with 8 bits for WiMAX and with 12 bits for IB-FDD would be sufficient.

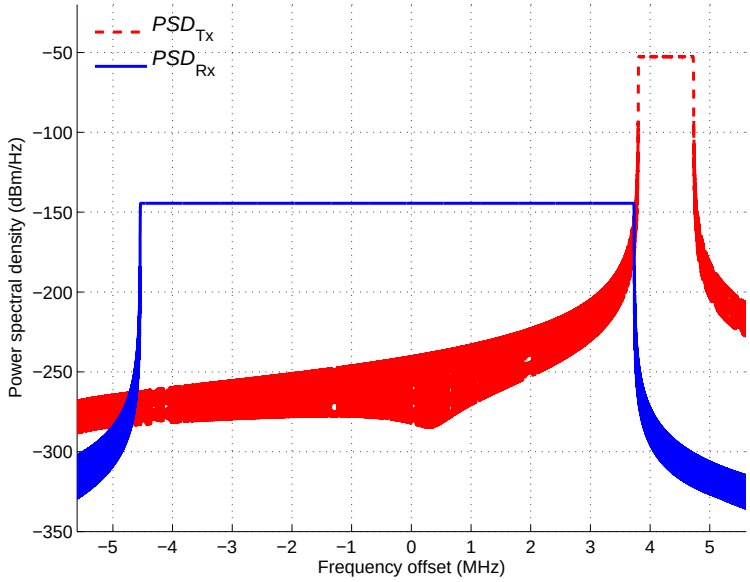


Figure 2.11: PSD of the LNA input signal. MS best case, asymmetric bandwidth, 8 free subcarriers between Tx and Rx and CNAM prototype.

A summary of the examples for an asymmetric bandwidth allocation is given in Table 2.5.

2.4.5 Additional comments and conclusions

In this study we did not consider multiple antenna systems. In that case, we should assume that the sum of the power used on each antenna is equal to the power used in the single antenna case. Depending on which multiple input multiple (MIMO) scheme is applied,

Parameter\Case	BS worst	MS worst	BS best	MS best
P_{Tx}	43 dBm	25 dBm	30 dBm	20 dBm
$\mathrm{SNR}_{\mathrm{Rx}}$	2.9 dB	2.9dB	20 dB	20 dB
R_{S}	-98.5 dbm	-98.5 dBm	-71.8 dBm	-71.8 dBm
DR WiMAX	53.5 dB	68.5 dB	26.8 dB	41.8 dB
DR IB-FDD	116.5 dB	98.5 dB	76.8 dB	66.8 dB
Guard band	60 subch.	30 subch.	13 subch.	8 subch.
ADC WiMAX	10 bits	13 bits	6 bits	8 bits
ADC IB-FDD	21 bits	18 bits	14 bits	12 bits

Table 2.5: Summary of the examples for asymmetric bandwidth allocation.

the distribution of power will not be uniform. In the limit, one transmit antenna gets all the power. In this case the receiver sharing the same antenna picks all the transmit power similar to the worst case in a single antenna system. However, how much transmit signal enters into the receivers connected to adjacent and nearby antennas depends on many other factors, like distance, antenna gain, antenna geometry, for example.

We also did not take into account non-linearities of the DAC and PA. That is, we considered them always working in the linear region. If those devices clip the signal to be transmitted, some spectral regrowth will occur, worsening the situation for the receiver in the same equipment. As a consequence, more interference will leak from the transmit band to the receive band. In other words, the transmitter has also to fulfill severe requirements.

It should be noted that the problem of interference between Tx and Rx after the analog equipment fulfill the requisites, depends only on the design criteria of the prototype. A minimax design, for example, would result in a more flat out-of-band radiation, possibly reducing the number of necessary guard subcarriers. The design of the prototype will also have a crucial role if the number of subcarriers and the latency are lowered, i. e. the relation between the length of the prototype and the number of subcarriers is kept, and the same transmission bandwidth is maintained, since in that case there is an increase in the interference between Tx and Rx signal.

In this case study we showed through examples the requirements of the analog receive front end if an IB-FDD scheme is implemented the FBMC modulation is employed. The parameters were extracted from the WiMAX and IEEE 802.16 standards. The examples show that these requisites are very demanding and that the design of the prototype could take into account the interference between downlink and uplink.

Nevertheless, the concept of the IB-FDD scheme is very attractive since the dynamic allocation between up- and downlink will play a crucial role in future wireless communication systems. In the examples studied here we considered a contiguous band for each link, but this does not need to be the rule. Each link between two stations could occupy different regions of the spectrum, maximizing the efficiency according to the quality of the transmission channel in each region of the given band.

2.5 Conclusions

From the viewpoint of spectral efficiency, FDD is to be preferred in an FBMC system, because here FBMC benefits from not needing a CP, loading more subcarriers with data, because of the much lower out-of-band radiation and pre- and post-tails are not harmful because of continuous data transmission or at least long blocks of transmission.

The desirable flexibility of IB-FDD will not be achievable because of the extreme requirements for both transmit and receive front end.

Chapter 3

Multiple access techniques

3.1 Introduction

The multiple access technique used in a multicarrier system naturally is Frequency Division Multiple access (FDMA). First a multiuser scenario with single antenna at both MS and BS is considered. In the downlink, there is no carrier frequency offset (CFO) and no timing offset (TO) between the subcarriers dedicated to different users. Therefore, orthogonality between users is not destroyed. In the uplink the situation is quite different. Between users, there is CFO and TO and the BS cannot synchronize to all users simultaneously. The performance degradation due to these offsets is investigated in Section 3.2 for different allocation schemes and for both AWGN and for multipath channels. FBMC and CP-OFDM are compared in the basis of the same bandwidth and same symbol rate on the channel. Therefore, the data rate for the FBMC system is higher, because there is no CP and more subcarriers can be loaded because of the much better reduction of the out-of-band emissions. The FBMC system needs an equalizer per sub-carrier to mitigate multipath induced intersymbol interference (ISI) and intercarrier interference (ICI). Such an equalizer will need several taps and will also reduce the degradation due to CFO and TO.

In Section 3.3 the effect of intercell interference is investigated. In a cellular system it is desirable to work with frequency reuse one. In such a case users at the cell edge will suffer from strong interference. First steps to quantify these effect are taken by computing tables for the interference-to-signal ratio (ISR) as a basis for further investigation and comparison of FBMC and CP-OFDM.

3.2 Effects of timing and frequency offset in the uplink

In the presence of timing offset per user τ_i, carrier frequency offset (CFO) per user normalized to the subcarrier spacing ε_i and carrier phase offset per user ϕ_i, the received

continuous-time signal in the uplink multiuser FBMC system is

$$r(t) = \sum_{i=1}^{U} e^{j2\pi\left(\frac{\varepsilon_i}{T}t+\varphi_i\right)} |c_i|\, s_i\left(t-\tau_i\right) + \eta(t) \tag{3.1}$$

where

$$s_i(t) = \sum_{k\in\mathcal{M}_u^i} \sum_{n=-\infty}^{+\infty} d_{k,n}^i \beta_{k,n}\theta_{k,n} p\left(t - n\frac{T}{2}\right) e^{j\frac{2\pi}{T}kt}. \tag{3.2}$$

Taking into account the analysis filter bank (see D2.1), the received symbol for the i-th user in the subcarrier k is given by

$$\begin{aligned}
\hat{d}_{k,n}^i &= d_{k,n}^i |c_i| \,\mathrm{Re}\left\{ b_{k,k}^{i,\varepsilon_i}\left(-\tau_i\right) e^{j2\pi\left(\varphi_i - \frac{\tau_i}{T}k\right)}\right\} \\
&+ \sum_{\substack{k'\in\mathcal{M}_u^i \\ k'\neq k}} d_{k',n}^i |c_i| \,\mathrm{Re}\left\{ b_{k',k}^{i,\varepsilon_i}\left(-\tau_i\right) e^{j2\pi\left(\varphi_i - \frac{\tau_i}{T}k'\right)}\right\} \\
&+ \sum_{k'\in\mathcal{M}_u^i} \sum_{\substack{n'=-\infty \\ n'\neq n}}^{+\infty} d_{k',n'}^i |c_i| \,\mathrm{Re}\left\{ b_{k',k}^{i,\varepsilon_i}\left(\left(n'-n\right)T/2 - \tau_i\right) e^{j2\pi\left(\varphi_i - \frac{\tau_i}{T}k'\right)}\right\} \\
&+ \sum_{\substack{i'=1 \\ i'\neq i}}^{U} \sum_{k'\in\mathcal{M}_u^i} \sum_{n'=-\infty}^{+\infty} d_{k',n'}^{i'} |c_i| \,\mathrm{Re}\left\{ b_{k',k}^{i',\varepsilon_{i'}}\left(\left(n'-n\right)T/2 - \tau_{i'}\right) e^{j2\pi\left(\varphi_{i'} - \frac{\tau_i}{T}k'\right)}\right\} \\
&+ \mathrm{Re}\left\{ \eta\left(qT_s\right) e^{-j\frac{2\pi}{T}kqT_s} \beta_{k,n}^*\theta_{k,n}^* \otimes \tilde{p}\left(qT_s\right)\Big|_{q=nN/2}\right\}
\end{aligned} \tag{3.3}$$

where $T_s = T/M$ and

$$b_{k',k}^{i,\varepsilon_i}\left(\left(n'-n\right)T/2 - \tau_i\right) = \beta_{k,n}^*\theta_{k,n}^*\beta_{k',n'}\theta_{k',n'} e^{j\pi\varepsilon_i} e^{j\pi n(k'-k)} p\left(\left(n'-n\right)T/2 - \tau_i\right) \otimes \tilde{p}\left(nT/2\right) \tag{3.4}$$

and $\tilde{p}(t)$ is the matched filter to $p(t)$.

In the presence of synchronization errors and non dispersive channel, at the receiver side the useful term is subject to an attenuation and a phase rotation related to the subcarrier index k, the timing offset τ_i, the phase offset ϕ_i, the CFO ε_i and the index of information symbol n. Furthermore, intercarrier interference, intersymbol interference and multiple access interference are present. In our simulation analysis, we have considered three different allocations schemes: blockwise, interleaved and interleaved b assignment schemes. In particular, as illustrated in Figure 3.1, in the blockwise and interleaved b allocation schemes group of adjacent subcarriers are allocated to the same user or different users, respectively, while in the interleaved allocation scheme one subcarrier is dropped between two adjacent users.

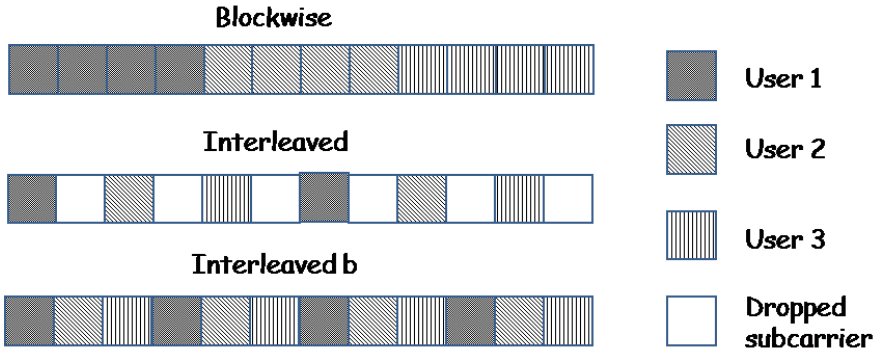


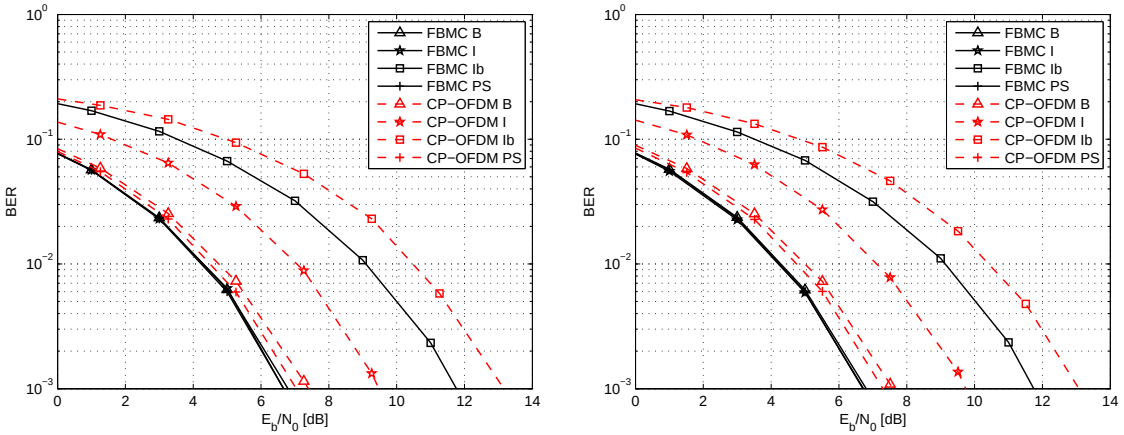
Figure 3.1: Allocation schemes.

The modulation format on all subcarriers and in all the following simulation results is QPSK. The number of subcarriers is $M = 1024$. In the FBMC case a prototype filter of length $K = 4$ with a roll-off factor of one has been used. The number of used subcarriers is evenly split between the users.

In Figure 3.2 we show that, in the case of a multi-user system in which the user of interest is perfectly synchronized to the base station (BS), the FBMC-MA systems are much more robust with respect to misalignments between different users than CP-OFDMA systems, assuring in the case of the blockwise assignment a performance practically coincident with that obtained in the case of perfect synchronization. We have considered the case of an asynchronous scenario with $U = 4$ users, where three of them have normalized frequency offsets uniformly distributed in the range ± 0.125 and timing offsets uniformly distributed within $\{-T/2, ..., T/2\}$ [8]. The BER depicted in Figure 3.2 corresponds to the perfect synchronized user, who is suffering from the multiple access interference of the three asynchronous users. Figures 3.2(a),(b) and (c) shows always the same performance for the FBMC system while comparing it with CP-OFDM with different CP-length. It is obvious, that FBMC with blockwise allocation of subcarriers per user is always the best choice and the performance advantage over CP-OFDM is higher as longer is the CP.

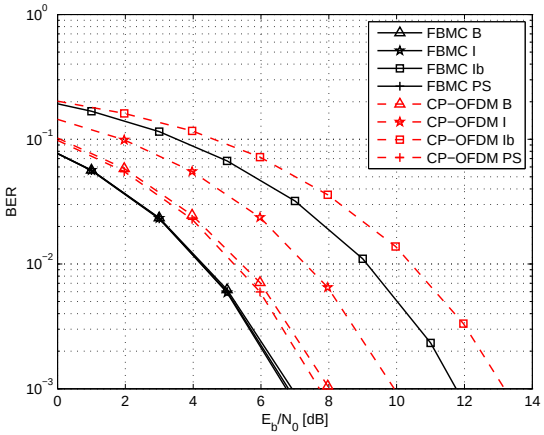
Furthermore, we have studied the sensitivity of the FBMC system in terms of BER to timing inaccuracy for the user of interest, under the hypothesis of ideal carrier frequency recovery. The other users are supposed asynchronous with normalized frequency offsets uniformly distributed in the range ± 0.5 and a timing offset uniformly distributed within $\{-T/2, ..., T/2\}$. In Figure 3.3, we report the BER of the user of interest with $E_s/N_0 =$ 20dB as a function of the residual timing offset (RTO) normalized to the sampling interval $(\hat{\tau} - \tau)/T_s$ in AWGN channel (solid line) and ITU-Vehicular A channel (dashed line) and $U = 4$ users. The E_s/N_0 of the other users is equal to $(E_s/N_0)_i = 20\text{dB}, \ i = 2, 3, 4$.

Note that the effect of the RTO on the digital data at the output of each subcarrier is an attenuation and a phase rotation proportional to the RTO and to the subcarrier index (3.3). This phase rotation incorporated in the channel gain should be compensated by the subcarrier equalizer. It is assumed that on each subcarrier a one-tap equalizer with perfect knowledge of the channel and of the RTO is used. The results show that the



(a) CP length $M/16$

(b) CP length $M/8$



(c) CP length $M/4$

Figure 3.2: Comparisons between FBMC-MA and CP-OFDMA.

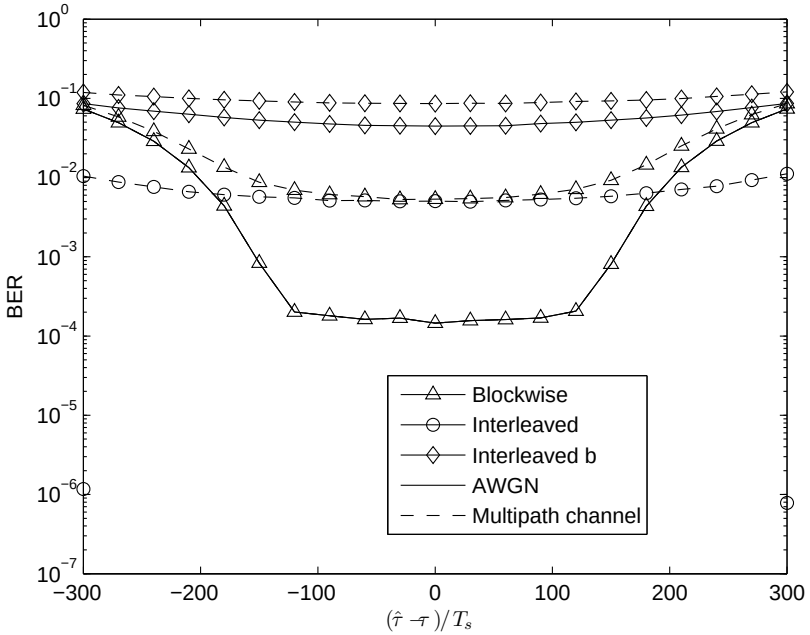


Figure 3.3: BER as a function of the RTO normalized to the sampling interval for $E_s/N_0 =$ 20 dB.

interleaved allocation scheme assures the lowest sensitivity to the presence of a RTO, both in AWGN and multipath channel, but on the other hand, it presents a loss of spectral efficiency, since one subcarrier is dropped between two adjacent users. Instead, with the blockwise assignment scheme, an accuracy of ± 100 samples ($\cong 10\%$ of the FBMC symbol interval T) is sufficient to assure an acceptable performance degradation with respect to the case of perfect synchronization (RTO=0). (Note that the BER corresponding to the interleaved scheme in AWGN is lower than 10^{-5}.) Since the FBMC system needs a multitap equalizer anyway, this will additionally increase the robustness against RTO. This will be investigated and compared to the CP-OFDM system. An improved robustness for FBMC over CP-OFDM is expected.

We have also analyzed the effect of the presence of a residual carrier frequency offset (RCFO) for the user of interest, for whom RTO is zero, with the other users completely asynchronous. In Figure 3.4, we report the BER of the user of interest with $E_s/N_0 = 20$ dB as a function of the RCFO normalized to the subcarrier spacing $\left(\hat{\Delta}f - \Delta f\right) T$ in AWGN channel (solid line) and ITU-Vehicular A channel (dashed line) for an FBMC-MA system with $U = 4$ users. The E_s/N_0 of the other users is also 20 dB.

Note that the effect of the RCFO on the digital data at the output of each subcarrier is an attenuation and a phase rotation proportional to the RTO and to the symbol index (3.3). This phase rotation incorporated in the channel gain should be compensated by the subcarrier equalizer. It is assumed that on each subcarrier a one-tap equalizer with perfect

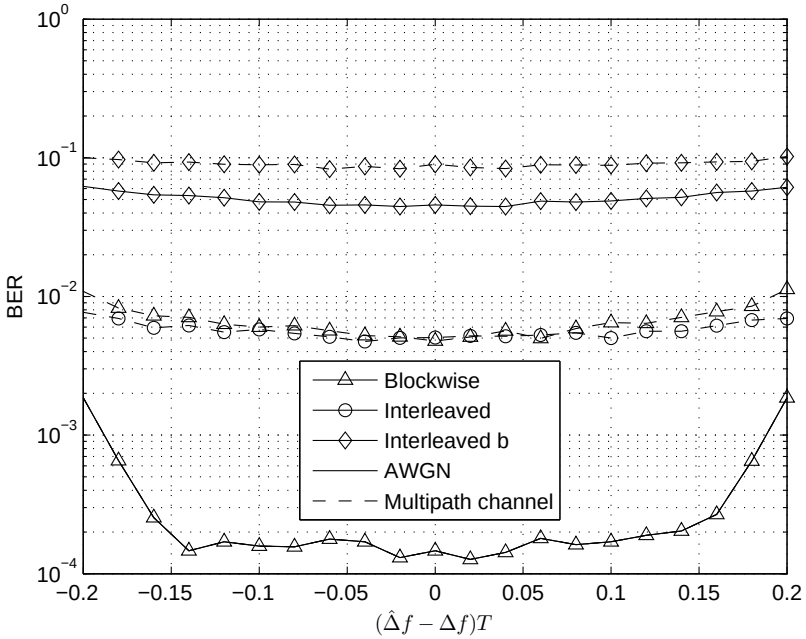


Figure 3.4: as a function of the RCFO normalized to the subcarrier spacing for $E_s/N_0 = 20$ dB.

knowledge of the channel and of the RCFO is used.

For blockwise and interleaved allocation schemes an accuracy of $\pm 15\%$ of the subcarrier spacing can provide an acceptable performance degradation with respect to the case of perfect synchronization (RCFO=0). The degradation of the FBMC system to RCFO again will be reduced by a multitap equalizer and is expected to be much less than with CP-OFDM.

In addition to the sensitivity of an FBMC-MA system to the presence of synchronization errors, we have studied the performance of an asynchronous FBMC-MA system in which, after a first timing and frequency offset compensation, the user of interest has a residual synchronization error.

Let us analyze the case of the presence of a RTO and ideal carrier frequency recovery. For the initial synchronization at the BS, we exploit the preamble-based estimator derived in [9], whose expression is the following (see D2.1 for further details):

$$\tilde{\tau}_i = \arg\max_{\tilde{\tau}_i} \left[\left| \sum_{k=0}^{\eta M-1} r[m] z_i^*[m - \tilde{\tau}_i] \right| \right], \quad i = 1, \ldots, U, \tag{3.5}$$

where

$$z_i[m] = \sum_{k \in M_u^i} \sum_{n=0}^{2N_{TR}-1} d_{k,n} \theta_{k,n} \beta_{k,n} p\left[m - nM/2\right] e^{j\frac{2\pi}{M}km} \tag{3.6}$$

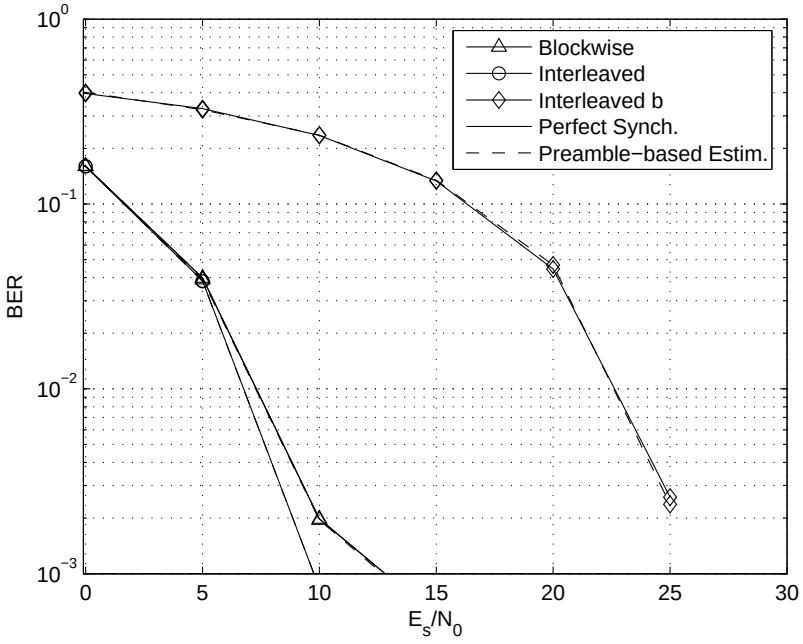


Figure 3.5: BER in the presence of a RTO (AWGN channel).

is the training sequence of the i-th user and the notation of the type $\tilde{x}$ indicates trial value of x.

In Figure 3.5 and Figure 3.6, the BER obtained by exploiting the proposed algorithm is compared with that observed in the case of perfect synchronization. Both in AWGN and multipath channel, the adoption of the preamble-based estimator can assure nearly the same performance of one tap channel equalization with perfect knowledge of the channel and perfect synchronization for all the considered allocation schemes.

In the presence of a RCFO and ideal timing recovery, we can exploit the preamble based estimator proposed for the initial synchronization at the BS, (see [4] and [10] for further details)

$$\hat{\varepsilon}_i = \arg\max_{\tilde{\varepsilon}_i} \left[\left| \sum_{k \in M_u^i} \sum_{n=0}^{2N_{TR}-1} d_{k,n} \theta_{k,n}^* \beta_{k,n}^* w_n^{(k)} \left(\tilde{\varepsilon}_i \right) \right| \right], \quad i = 1, \ldots, U, \tag{3.7}$$

where

$$w_n^{(k)} \left(\tilde{\varepsilon}_i \right) = \sum_{m=0}^{\eta M-1} r[m] p \left[m - nM/2 \right] e^{-j \frac{2\pi}{M} (k + \tilde{\varepsilon}_i) m}. \tag{3.8}$$

In Figure 3.7 and Figure 3.8 we report the BER obtained by exploiting the proposed CFO estimator. The considered algorithm assures, in the whole range of considered E_s/N_0 values, a negligible performance degradation with respect to the case of one-tap equalizer with perfect knowledge of the channel and perfect synchronization.

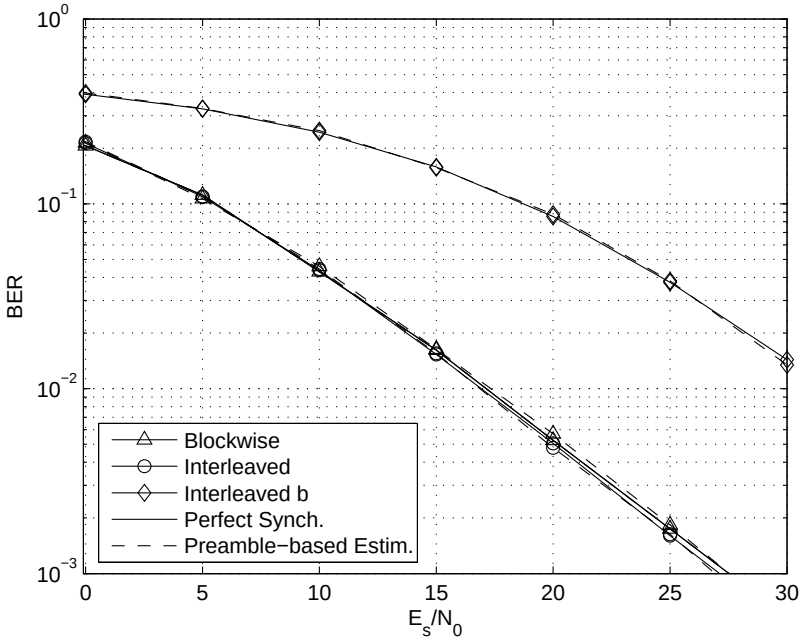


Figure 3.6: BER in the presence of a RTO (multipath channel).

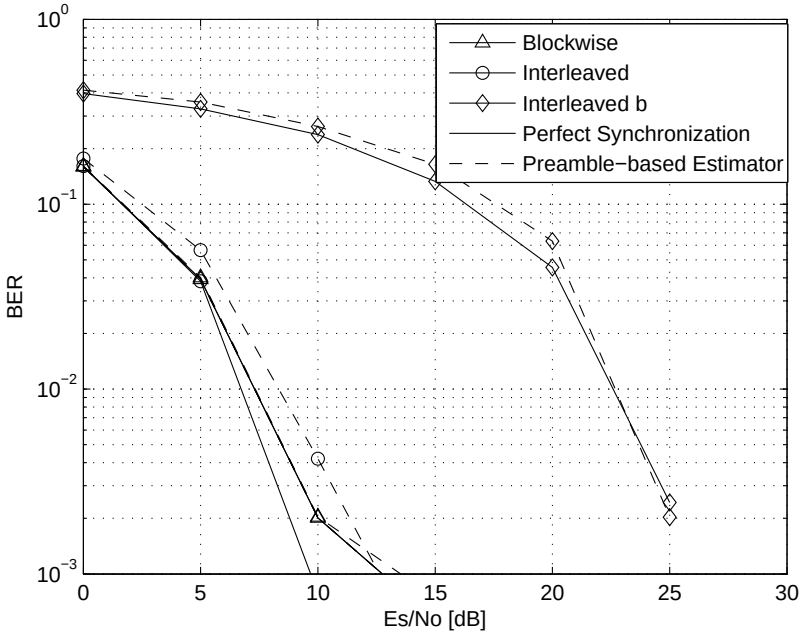


Figure 3.7: BER in the presence of a RCFO (AWGN channel).

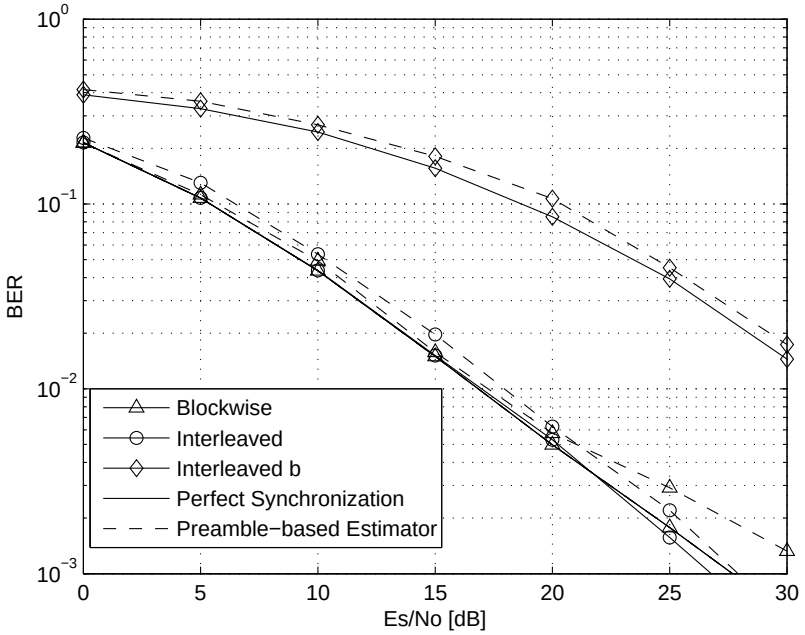


Figure 3.8: BER in the presence of a RCFO (multipath channel).

Finally, we consider the case of the presence of timing and frequency offsets. For the initial synchronization at the BS, we can use the joint carrier frequency offset and symbol timing estimator (see D2.1 and D2.2 for further details)

$$\hat{\tau}_i = \arg\max_{\tilde{\tau}_i} \left[|A(\tilde{\tau}_i)| + |B(\tilde{\tau}_i)| \right], \quad i = 1, \ldots, U, \tag{3.9}$$

$$\hat{\varepsilon}_i(\hat{\tau}_i) = \frac{1}{\pi} \angle \left\{ A^*(\hat{\tau}_i) B(\hat{\tau}_i) \right\}, \quad i = 1, \ldots, U, \tag{3.10}$$

where

$$A(\tilde{\tau}_i) = \sum_{k \in M_u^i} d_{k,0} \theta_{k,0}^* \beta_{k,0}^* w_0^{(k)} (0, \tilde{\tau}_i) \tag{3.11}$$

$$B(\tilde{\tau}_i) = \sum_{k \in M_u^i} d_{k,1} \theta_{k,1}^* \beta_{k,1}^* w_1^{(k)} (0, \tilde{\tau}_i) \tag{3.12}$$

and

$$w_n^{(k)} (\tilde{\varepsilon}_i, \tilde{\tau}_i) = \sum_{m=0}^{\eta M - 1} r[m] p[m - nM/2 - \tilde{\tau}_i / T_s] e^{j \frac{2\pi}{T} k (\tilde{\tau}_i - m)} e^{-j \frac{2\pi \tilde{\varepsilon}_i}{M} m}. \tag{3.13}$$

The derived joint symbol timing and CFO estimator in (3.9) and (3.10) results to be particularly attractive since it provides a closed form solution for the CFO estimate and then it requires only a one-dimensional maximization with respect to the continuous parameter $\tilde{\tau}_i$.

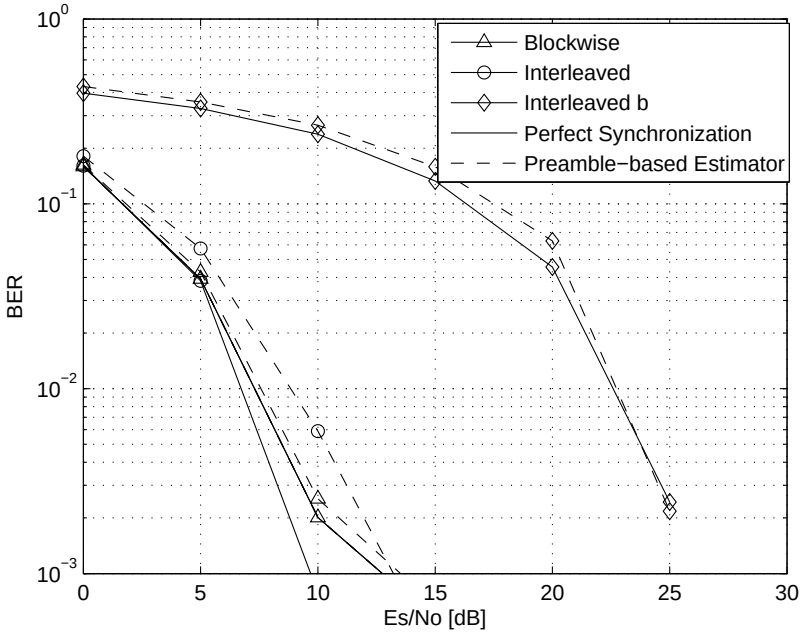


Figure 3.9: BER in the presence of a RTO and RCFO (AWGN channel).

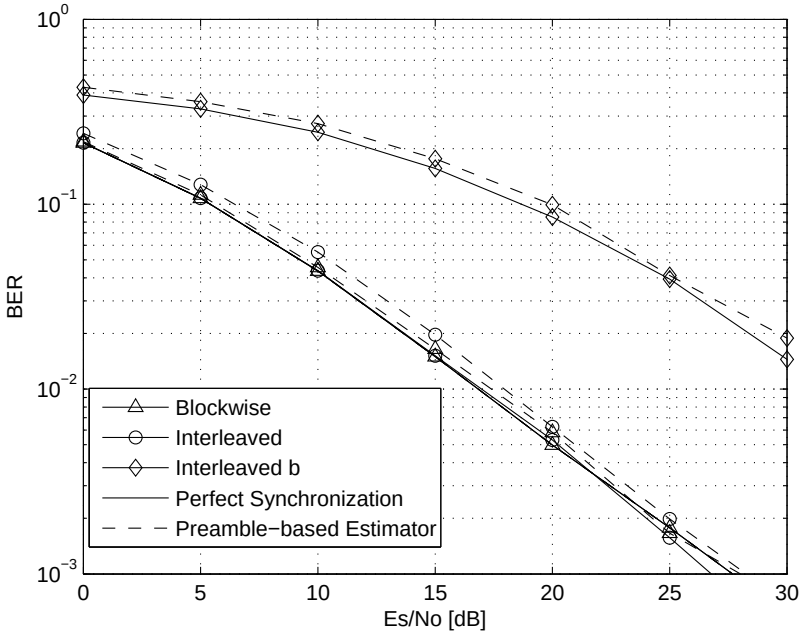


Figure 3.10: BER in the presence of a RTO and RCFO (multipath channel).

In Figure 3.9 and Figure 3.10, the BER obtained by exploiting the proposed joint symbol timing and CFO algorithm and that observed in the case of perfect synchronization. The results show that, both in AWGN and multipath channel, and for all the considered allocation schemes, the performance of the proposed estimator is practically coincident with that obtained the case of one-tap equalization with perfect knowledge of the channel and perfect synchronization.

Concluding the section on multiple access intracell interference in the uplink it is important to note that, the FBMC approach relaxes the required synchronization accuracy compared to CP-OFDM. This will be quantified further especially when a multitap equalizer [11, 12] is utilized. From the different allocation schemes, the blockwise assignment of subcarriers to user is the most efficient: it combines acceptable sensitivity with high spectral efficiency. Its sensitivity can be further reduced by dropping a single subcarrier between blocks assigned to different user with only marginally reducing spectral efficiency.

3.3 Analysis of multiple access interference in the downlink

In the previous section the multiple access interference within an isolated cell has been considered. Here we will investigate the interference, which comes from the adjacent cells in a multicellular network.

3.3.1 Interferences tables

In this section we compare CP-OFDM and FBMC in the downlink. We focus on the impact of inter-cell interference in an unsynchronized FDD context.

The reuse factor is one, meaning that all the cells present in the network use the same frequency band. If we consider one user of a given cell, he will receive both the desired signal $s_1(t)$ from its own base station and the interference $s_2(t)$ from neighboring cell in the same frequency band, as illustrated in Figure 3.11.

We assume a perfect frequency and time synchronization between the user and his base station. Hence, the interference $s_2(t)$ will only come from the signal of the other cell. Our aim is to calculate the inter-cell interference power in FBMC and CP-OFDM systems and to carry out a performance comparison of these two technical solutions.

In our analysis, frequency offset and phase offset will be disregarded. Nevertheless, as two cells are not time synchronized, we will be interested in the timing offset τ between the user and the base station of the other cell. For example $\tau = 0$ means that frames received from the two base stations are, at the receiver location, perfectly synchronized.

Under this assumption, we will analyze the interference created by each subcarrier of the cell B on the k-th subcarrier of the cell A.

For the sequel of the paragraph, a subcarrier is called "*frequency slot*" and the CP-OFDM symbol duration is called "*time slot*". We consider then the problem in 2D (time and frequency) space.

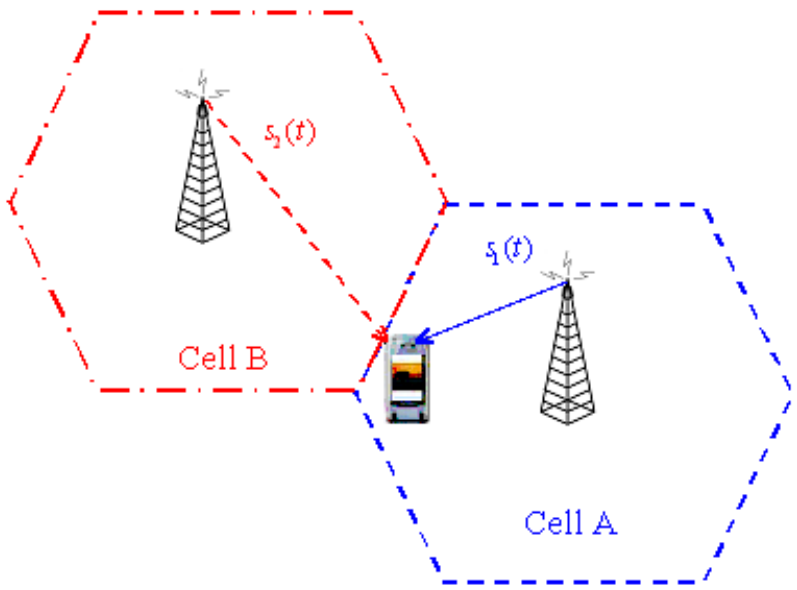


Figure 3.11: Interference in multicellular (2 cells) system

In the CP-OFDM case, we note that the interference comes from two time slots and a high number of frequency slots used in Cell B. We can then define a set Ω_{OFDM} of time/frequency slots of cell B that are interfering with the analyzed time/frequency slot of cell A, so called "victim slot" in the Figure 3.12 hereafter.

On the other hand, in the case of FBMC transmission, the interference is much more localized in frequency (It does not exceed the adjacent subcarriers) but is spread out over K CP-OFDM symbols in time (depending of the length $KM + 1$ of the prototype filter).

Finally, the main differences between the two systems is the size of the two sets Ω_{OFDM} and Ω_{FBMC}.

Thanks to FBMC characteristics, we have: $\text{Card}(\Omega_{\text{FBMC}}) < \text{Card}(\Omega_{\text{OFDM}})$, nevertheless we have to calculate precisely the power coupling coefficients between slots of Ω_{OFDM} (resp. Ω_{FBMC}) and the victim slot.

3.3.2 Theoretical derivation leading to interference tables

In our derivation, we consider that we transmit an isolated communication symbol on each time / frequency slot and we evaluate the impact on victim slot.

3.3.2.1 CP-OFDM system

As is shown in Figure 3.14, we assume that the single transmit symbol from base station A – the desired signal – arrives at the receiver at $t = 0$. The receiver therefore evaluates the signal at the receive filter output at the end of the symbol, *i.e.*, at $t = T + \Delta$. The

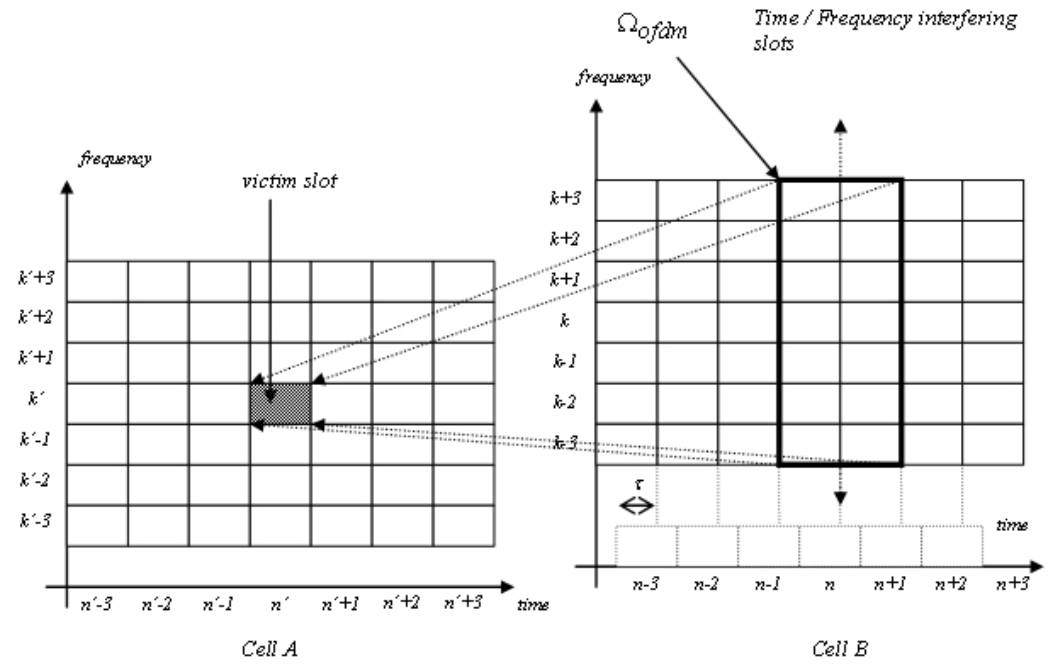


Figure 3.12: Set of time/frequency slots of cell B that interferes with the (n, k) slot of cell A with an CP-OFDM transmission

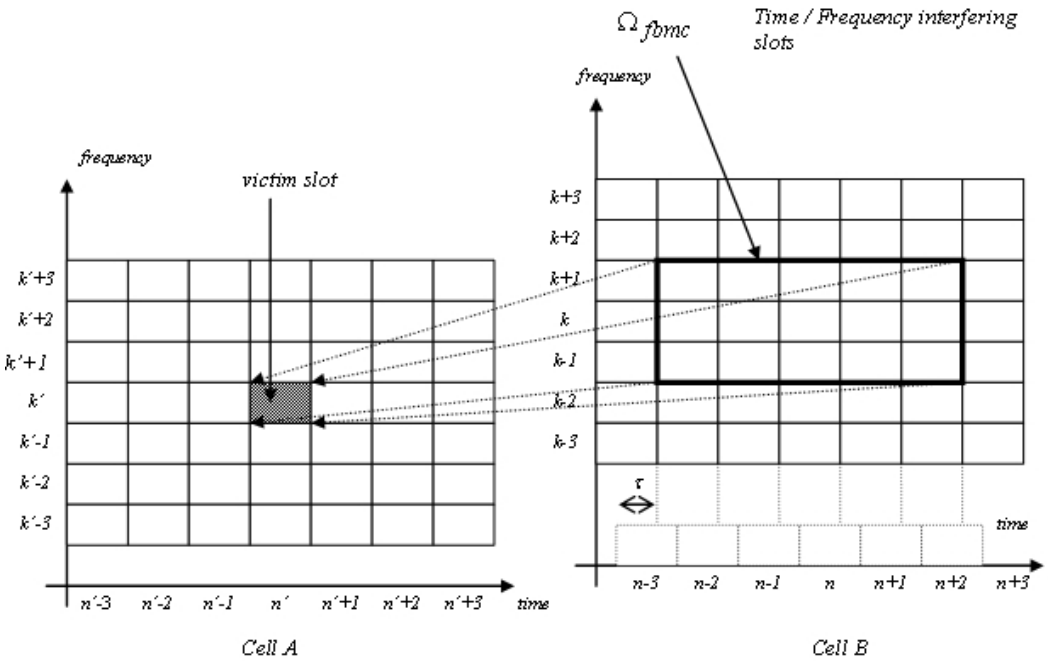


Figure 3.13: Set of time/frequency slots of cell B that interferes with the (n, k) slot of cell A for a FBMC context

interfering symbol from base station B arrives at an offset of $-\tau$. Therefore we have

$$s(t) = g(t+\tau)e^{j\frac{2\pi}{T}k(t+\tau)}, \tag{3.14}$$

where

$$g(t) = \begin{cases} \frac{1}{\sqrt{T}} & t \in [0, T+\Delta], \\ 0 & \text{else.} \end{cases} \tag{3.15}$$

Since the receive filter impulse response is

$$f(t) = \begin{cases} \frac{1}{\sqrt{T}} & t \in [0, T], \\ 0 & \text{else,} \end{cases} \tag{3.16}$$

we get for the receive filter output

$$y(t) = s(t) \cdot e^{-j\frac{2\pi}{T}k't} * f(t), \tag{3.17}$$

where k is the subcarrier number of the transmitting base station B while k' is the number of the receiving victim subcarrier and $l = k - k'$ denotes the offset between them. Plugging in $s(t)$ we have

$$\begin{aligned} y(t) &= e^{j\frac{2\pi}{T}k\tau} \cdot e^{j\frac{2\pi}{T}lt} g(t+\tau) * f(t) \\ &= e^{j\frac{2\pi}{T}k\tau} \int_{-\infty}^{+\infty} g(t'+\tau) f(t-t') e^{j\frac{2\pi}{T}lt'} \mathrm{d}t' |_{t=T+\Delta}. \end{aligned}$$

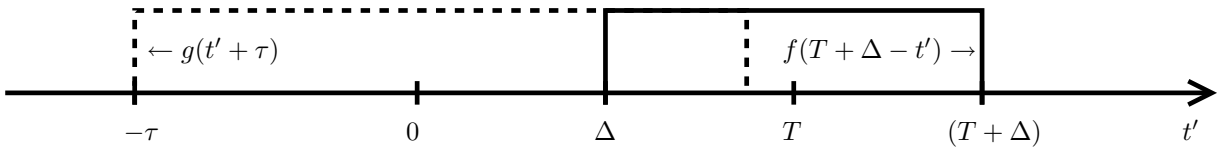


Figure 3.14: Displaced transmitter and receiver pulse shape

We see that y can be nonzero only for values of the offset $-\tau \in [-T, T+\Delta]$, which we have to split into three subintervals

- Case 1: $\tau \in [0, T]$,
- Case 2: $\tau \in [-\Delta, 0]$,
- Case 3: $\tau \in [-\Delta - T, -\Delta]$.

In general we have

$$y(t = T + \Delta, k, k, \tau) = e^{j\frac{2\pi}{T}k\tau} \int_{-\infty}^{+\infty} g(t' + \tau) f(T + \Delta - t') e^{j\frac{2\pi}{T}lt'} \mathrm{d}t', \tag{3.18}$$

which for <u>case 1</u> ($\tau \in [0, T]$) leads to

$$|y(t = T + \Delta, k, k, \tau)| = (1 - \frac{\tau}{T}) \left| \mathrm{sinc}\left(\pi l(1 - \frac{\tau}{T})\right)\right|. \tag{3.19}$$

The corresponding interference power is

$$I(l, \tau) = (1 - \frac{\tau}{T})^2 \left(\mathrm{sinc}\left(\pi l(1 - \frac{\tau}{T})\right)\right)^2. \tag{3.20}$$

For <u>case 2</u> ($\tau \in [-\Delta, 0]$) we have

$$I(l, \tau) = \begin{cases} 1, & l = 0, \\ 0, & \text{else.} \end{cases} \tag{3.21}$$

For <u>case 3</u> ($\tau \in [-\Delta - T, -\Delta]$) we have

$$I(l, \tau) = (1 + \frac{\Delta + \tau}{T})^2 \left(\mathrm{sinc}\left(\pi l(1 + \frac{\Delta + \tau}{T})\right)\right)^2. \tag{3.22}$$

For calculating an average interference power an interval $\tau \in [-\Delta/2, T + \Delta/2]$ has been chosen and leads to the following result

$$\bar{I}(l) = \frac{1}{T + \Delta} \int_{-\Delta/2}^{T+\Delta/2} I(l, \tau) d\tau. \tag{3.23}$$

For $l = 0$ we get

$$\bar{I}(0) = \frac{2T + 3\Delta}{T + \Delta} \tag{3.24}$$

and for $l \neq 0$

$$\bar{I}(l) = \frac{T}{2(T + \Delta)} \cdot \frac{1}{(\pi l)^2}. \tag{3.25}$$

It is worth to consider this choice for the mean interference due to periodicity of $T_1 = T + \Delta$ and the symmetry. The theoretical results for this case are given in Table 3.1.

Table 3.2 represents the simulated interference power. It corresponds to a transmitted power that equals 1 with channel gains of 1 between the two base stations and the victim user. As expected, it appears that the interference power is spread over a high number of time/frequency slots.

0	***n***	***n+1***
k+15	2.0014E-004	2.0014E-004
k+14	2.2975E-004	2.2975E-004
k+13	2.6646E-004	2.6646E-004
k+12	3.1272E-004	3.1272E-004
k+11	3.7216E-004	3.7216E-004
k+10	4.5032E-004	4.5032E-004
k+9	5.5595E-004	5.5595E-004
k+8	7.0362E-004	7.0362E-004
k+7	9.1901E-004	9.1901E-004
k+6	1.2509E-003	1.2509E-003
k+5	1.8013E-003	1.8013E-003
k+4	2.8145E-003	2.8145E-003
k+3	5.0035E-003	5.0035E-003
k+2	1.1258E-002	1.1258E-002
k+1	4.5032E-002	4.5032E-002
k	**3.5185-001**	**3.5185E-001**
k-1	4.5032E-002	4.5032E-002
k-2	1.1258E-002	1.1258E-002
k-3	5.0035E-003	5.0035E-003
k-4	2.8145E-003	2.8145E-003
k-5	1.8013E-003	1.8013E-003
k-6	1.2509E-003	1.2509E-003
k-7	9.1901E-004	9.1901E-004
k-8	7.0362E-004	7.0362E-004
k-9	5.5595E-004	5.5595E-004
k-10	4.5032E-004	4.5032E-004
k-11	3.7216E-004	3.7216E-004
k-12	3.1272E-004	3.1272E-004
k-13	2.6646E-004	2.6646E-004
k-14	2.2975E-004	2.2975E-004
k-15	2.0014E-004	2.0014E-004

Table 3.1: Theoretical mean power interference table T^{OFDM}

0	***n***	***n+1***
k+15	2.0172E-004	2.0172E-004
k+14	2.3122E-004	2.3122E-004
k+13	2.6780E-004	2.6780E-004
k+12	3.1390E-004	3.1390E-004
k+11	3.7314E-004	3.7314E-004
k+10	4.5102E-004	4.5102E-004
k+9	5.5628E-004	5.5628E-004
k+8	7.0344E-004	7.0344E-004
k+7	9.1809E-004	9.1809E-004
k+6	1.2488E-003	1.2488E-003
k+5	1.7973E-003	1.7973E-003
k+4	2.8070E-003	2.8070E-003
k+3	4.9885E-003	4.9885E-003
k+2	1.1221E-002	1.1221E-002
k+1	4.4878E-002	4.4878E-002
k	**3.5237E-001**	**3.5237E-001**
k-1	4.4878E-002	4.4878E-002
k-2	1.1221E-002	1.1221E-002
k-3	4.9885E-003	4.9885E-003
k-4	2.8070E-003	2.8070E-003
k-5	1.7973E-003	1.7973E-003
k-6	1.2488E-003	1.2488E-003
k-7	9.1809E-004	9.1809E-004
k-8	7.0344E-004	7.0344E-004
k-9	5.5628E-004	5.5628E-004
k-10	4.5102E-004	4.5102E-004
k-11	3.7314E-004	3.7314E-004
k-12	3.1390E-004	3.1390E-004
k-13	2.6780E-004	2.6780E-004
k-14	2.3122E-004	2.3122E-004
k-15	2.0172E-004	2.0172E-004

Table 3.2: Simulated mean power interference table T^{OFDM}

3.3.2.2 FBMC system

The transmit signal is given by

$$s(t) = [a[n]g(t - nT - \tau) + jb[n]g(t - nT - \tau - T/2)]\, e^{j\left(\frac{2\pi}{T}(t-\tau)+\frac{\pi}{2}\right)k}, \tag{3.26}$$

$$\text{with } a^2[n] + b^2[n] = 1. \tag{3.27}$$

The k-th output of the receiver reads

$$y_{k'}(t) = \left(s(t)e^{-j\left(\frac{2\pi}{T}t+\frac{\pi}{2}\right)k'}\right) * f(t) \tag{3.28}$$

$$= e^{-j\frac{2\pi}{T}k\tau}\left\{a[\ell]\int_{-\infty}^{+\infty} g(t - \ell T - \tau - \alpha)f(\tau)e^{j\left(\frac{2\pi}{T}(t-\alpha)+\frac{\pi}{2}\right)(k-k')}d\alpha\right. \tag{3.29}$$

$$\left. + jb[\ell]\int_{-\infty}^{+\infty} g(t - \ell T - T/2 - \tau - \alpha)f(\tau)e^{j\left(\frac{2\pi}{T}(t-\alpha)+\frac{\pi}{2}\right)(k-k')}d\alpha\right\} \tag{3.30}$$

After the decision, the interference power can be written as

$$\text{Interference}(n, k') = \left|\Re[y_{k'}(t)]_{t=nT} + \Im[y_{k'}(t)]_{t=(n+1/2)T}\right|^2 \tag{3.31}$$

The mean FBMC interference matrix for values of $\tau \in [0, T]$ is given in Table 3.3. The prototype filter used for this table was a root raised cosine filter with a roll-off equal to 1. The overlapping factor was $K = 4$.

	n-2	***n-1***	***n***	***n+1***	***n+2***	***n+3***
k+5	3.57E-07	4.30E-09	1.07E-10	1.07E-10	4.30E-09	3.57E-07
k+4	5.91E-07	5.69E-09	1.81E-10	1.81E-10	5.69E-09	5.91E-07
k+3	1.22E-06	6.68E-09	3.80E-10	3.80E-10	6.68E-09	1.22E-06
k+2	5.10E-06	1.09E-07	9.91E-09	9.91E-09	1.09E-07	5.10E-06
k+1	7.34E-05	1.78E-03	6.06E-02	6.06E-02	1.78E-03	7.34E-05
k	1.48E-05	1.87E-04	**3.77E-01**	**3.77E-01**	1.87E-04	1.48E-05
k-1	7.34E-05	1.78E-03	6.06E-02	6.06E-02	1.78E-03	7.34E-05
k-2	5.10E-06	1.09E-07	9.91E-09	9.91E-09	1.09E-07	5.10E-06
k-3	1.22E-06	6.68E-09	3.80E-10	3.80E-10	6.68E-09	1.22E-06
k-4	5.91E-07	5.69E-09	1.81E-10	1.81E-10	5.69E-09	5.91E-07
k-5	3.57E-07	4.30E-09	1.07E-10	1.07E-10	4.30E-09	3.57E-07

Table 3.3: Simulated mean power interference table T^{FBMC} (root square raised cosine, roll-off=1, K=4)

The mean FBMC interference matrix for values of $\tau \in [0, T]$ for the reference prototype filter is shown in Table 3.4.

Finally, if we consider only main interfering slots, with a hard threshold $T_{i,j} > 10^{-3}$, we obtain 12 interfering slots for the FBMC case: Card(Ω_{FBMC})=12 and 30 slots for the CP-OFDM case: Card(Ω_{OFDM})=30.

	n-2	***n-1***	***n***	***n+1***	***n+2***	***n+3***
k+5	1.37E-10	1.37E-10	1.09E-12	1.09E-12	1.37E-10	1.37E-10
k+4	6.13E-10	6.13E-10	5.54E-12	5.54E-12	6.13E-10	6.13E-10
k+3	5.06E-09	5.06E-09	5.54E-11	5.54E-11	5.06E-09	5.06E-09
k+2	2.65E-07	2.65E-07	1.58E-09	1.58E-09	2.65E-07	2.65E-07
k+1	4.11E-06	3.95E-03	4.03E-02	4.03E-02	3.95E-03	4.11E-06
k	8.43E-06	2.96E-03	**4.10E-01**	**4.10E-01**	2.96E-03	8.43E-06
k-1	4,11E-06	3.95E-03	4.03E-02	4.03E-02	3.95E-03	4.11E-06
k-2	2.65E-07	2.65E-07	1.58E-09	1.58E-09	2.65E-07	2.65E-07
k-3	5.06E-09	5.06E-09	5.54E-11	5.54E-11	5.06E-09	5.06E-09
k-4	6.13E-10	6.13E-10	5.54E-12	5.54E-12	6.13E-10	6.13E-10
k-5	1.37E-10	1.37E-10	1.09E-12	1.09E-12	1.37E-10	1.37E-10

Table 3.4: Simulated mean power interference table T^{FBMC} (with the reference filter)

3.3.3 Rules for using the interference tables

The interference tables can easily be used for calculating the mean interference power due to a set of active slots in the interfering cell.

Let us consider a simple case where two slots (p, q) and (u, v) of the interfering cell are switched on, other slots being switched off. These two slots interfere with the (k, n) victim slot of the analyzed cell. We can then calculate the instantaneous interference (complex value), for a particular value of τ (time delay representing the non synchronization of the two cells). We have then:

$$I_{n,k} = g_q c_{p,q}(\tau) s_{p,q} + g_v c_{u,v}(\tau) s_{u,v} \tag{3.32}$$

where:

- $I_{n,k}$ represents the complex interference value received in the (n, k) victim slot.
- g_p (resp. g_v) represents the channel gain at frequency p (resp. u), between the cell B base station and the victim.
- $c_{p,q}(\tau)$ represents the complex amplitude coupling coefficient between the active slot (p, q) and the victim slot (n, k).
- $s_{p,q}$ represents the modulation symbol transmitted on slot (p, q) in the interfering cell B.

Evaluating the mean value of the square of the interference, we obtain:

$$\mathsf{E}\left[|I_{n,k}|^2\right] = \mathsf{E}\left[|g_q c_{p,q}(\tau) s_{p,q} + g_v c_{u,v}(\tau) s_{u,v}|^2\right] \tag{3.33}$$

The statistical expectation E is calculated over:

- all channel gains: (g_p)
- all time delays: (τ)
- all communication symbols: $(s_{p,q})$

As g_p, $c_{p,q}(\tau)$ and $s_{p,q}$ are independent variables, the expectation can be simplified. Moreover, as communication symbols are zero mean uncorrelated variables, we have:

$$\mathsf{E}\left[|I_{n,k}|^2\right] = \mathsf{E}_g\left[|g_q|^2\right]\mathsf{E}_\tau\left[|c_{p,q}(\tau)|^2\right]\mathsf{E}_s\left[|s_{p,q}|^2\right] + \mathsf{E}_g\left[|g_v|^2\right]\mathsf{E}_\tau\left[|c_{u,v}(\tau)|^2\right]\mathsf{E}_s\left[|s_{u,v}|^2\right], \tag{3.34}$$

where:

- $\mathsf{E}_g\left[|g_q|^2\right]$ is the mean "power" of the channel gain.
- $\mathsf{E}_\tau\left[|c_{p,q}(\tau)|^2\right] = T_{p,q}$is directly given by the interfering table.
- $\mathsf{E}_s\left[|s_{u,v}|^2\right] = p_S$ is the mean power of the transmitted communication symbol.

Then:

$$\mathsf{E}\left[|I_{n,k}|^2\right] = P_S\left(\mathsf{E}_g\left[|g_q|^2\right]T_{p,q} + \mathsf{E}_g\left[|g_v|^2\right]T_{u,v}\right). \tag{3.35}$$

If we consider $\mathsf{E}_g\left[|g_q|^2\right] = \mathsf{E}_g\left[|g_v|^2\right] = g_2$, we obtain:

$$\mathsf{E}\left[|I_{n,k}|^2\right] = g_2 P_S\left(T_{p,q} + T_{u,v}\right). \tag{3.36}$$

Now, for the general case with a set Ω_a of switched on slots, we will obtain:

$$\mathsf{E}\left[|I_{n,k}|^2\right] = g_2 P_S \sum_{(i,j)\in\Omega_a} T_{i,j}. \tag{3.37}$$

If we consider that the mean useful received power (from the base station of Cell A) is equal to $g_1 P_s$, we have the following expression for the mean Signal to Interference power Ratio (SIR):

$$\mathrm{SIR} = \frac{g_1}{g_2 \sum_{(i,j)\in\Omega_a} T_{i,j}}. \tag{3.38}$$

Finally it appears that inner values of the interference tables can easily be linearly added in order to calculate the interference power due to a set of active slots. Interference tables can then be used as main parameters of the physical layer for resource allocation algorithms. Moreover, a resource allocation algorithm, knowing which time/frequency slots are currently used in the neighbor cells, could then more easily avoid interfering slots in a FBMC context than in an OFDM context (due to sizes of the two sets Ω_{OFDM} and Ω_{FBMC}).

3.3.4 Interferences probability

Power interference tables introduced in the previous section are important for estimating the probability law of the interference power. If we introduce the probability p representing the probability of having an active transmission in a frequency/time slot, then it appears that we can calculate the probability of having k active frequency/time slots in the Ω_{OFDM} (resp. Ω_{FBMC}) set:

$$P_k = C_N^k p^k (1-p)^{N-k} \tag{3.39}$$

where N stands for Card(Ω_{OFDM}) (resp. Card(Ω_{FBMC})).

Having the power interference tables, the discrete set of available interference power values (2^N values) can be calculated. Then, having available values and their probabilities, we have finally an estimation of the discrete interference probability law. Figures 3.15-3.18 correspond to a traffic load: $p = 0.3$. Probabilities were obtained with an exhaustive enumeration for the case of FBMC transmission while they have been obtained via Monte Carlo simulation for CP-OFDM transmission (due to the very high value of Card(Ω_{OFDM})).

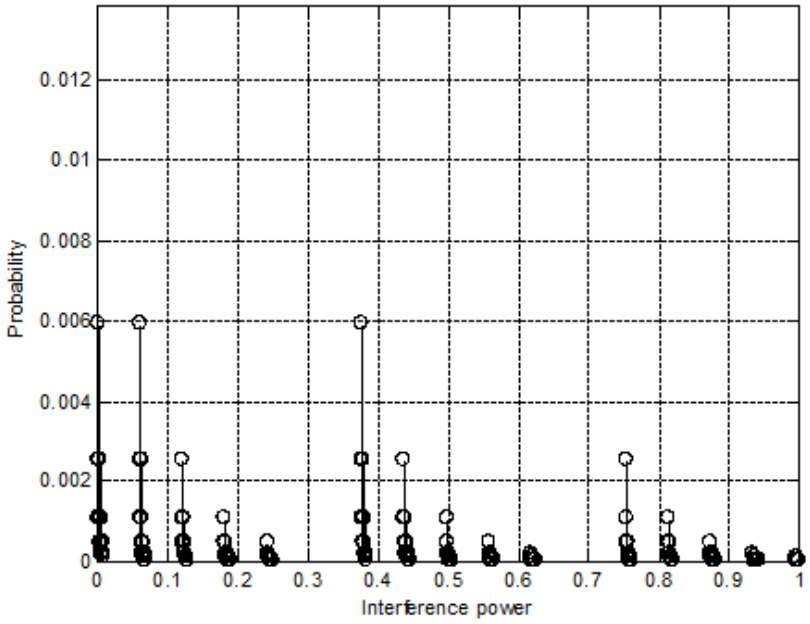


Figure 3.15: Available interference power and associated probabilities for FBMC transmission

It appears finally that the number of potentially interfering slots, having a high interfering power, is weaker in the FBMC context, than in the CP-OFDM context. This could be a key point in the comparison of the two waveforms. A resource allocation algorithm, knowing which time/frequency slots are currently used in the neighbor cells, could then more easily avoid interfering slots in a FBMC context than in an CP-OFDM context.

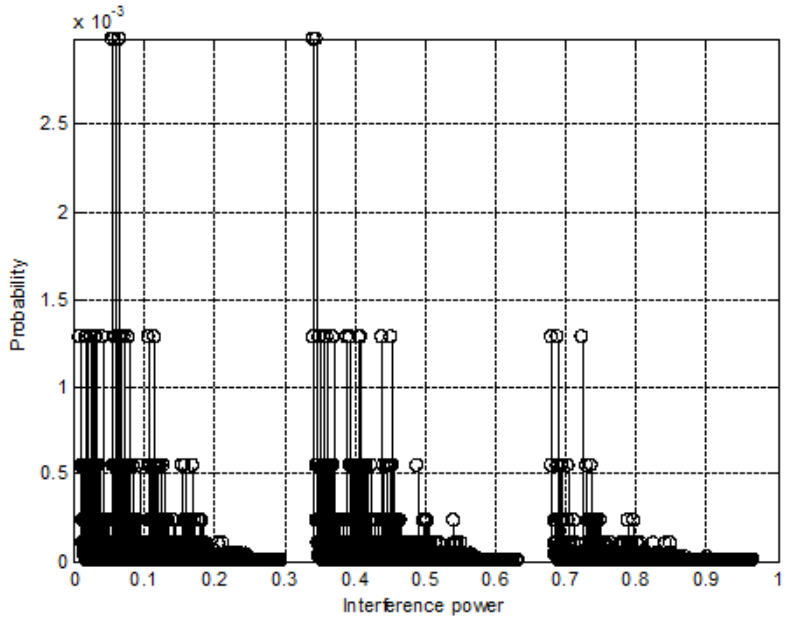


Figure 3.16: Available interference power and associated probabilities for CP-OFDM transmission (curve obtained by Monte Carlo simulation)

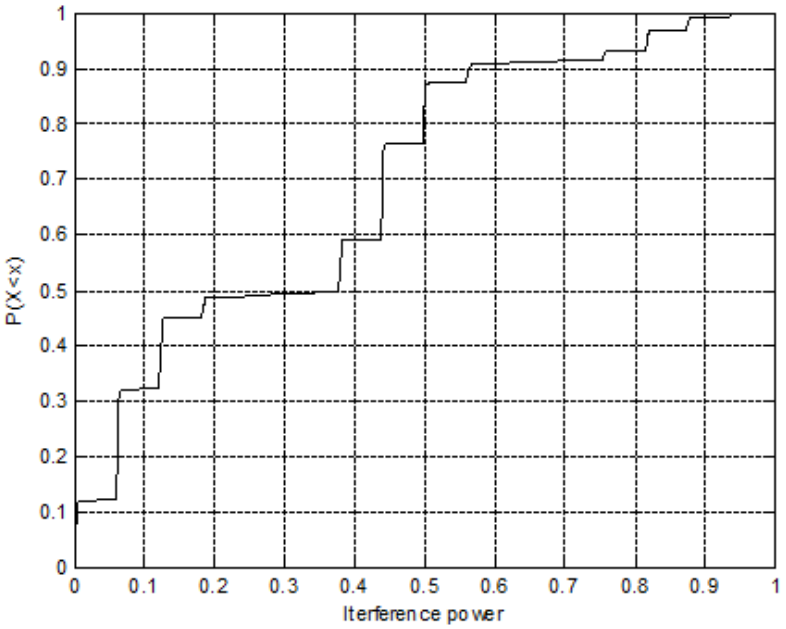


Figure 3.17: Cumulative probability law of interference power for FBMC transmission

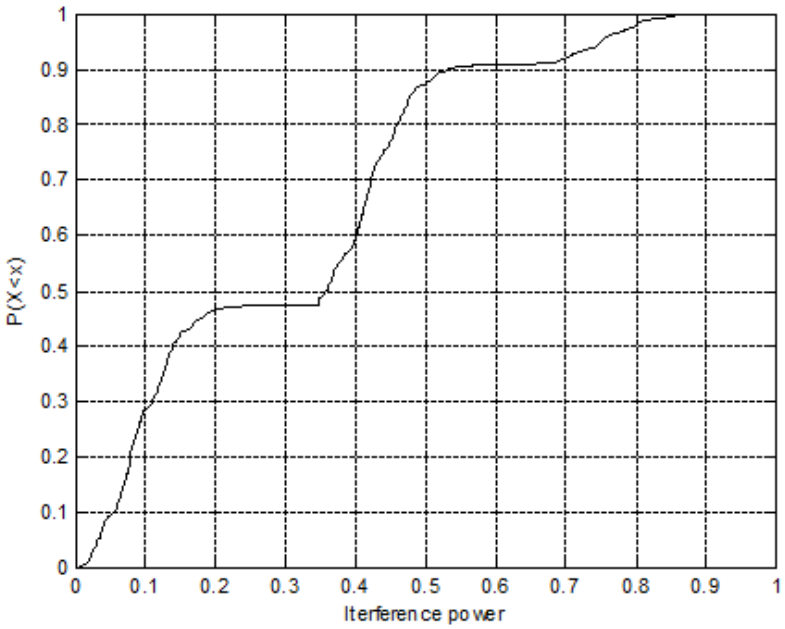


Figure 3.18: Cumulative probability law of interference power for CP-OFDM transmission (curve obtained by Monte Carlo simulation)

We see from the tables, that in the CP-OFDM system the interference is spread over many subcarriers, which would be even worse, if in addition to the timing offset a carrier frequency offset between base stations A and B has been assumed. Therefore, the results shown will degrade by taking a CFO into account. In the FBMC system only immediate adjacent subcarriers are causing interference, but not only two subsequent symbols in time may interfere, but as many as the length of the prototype filter. But still, the number of interfering time/frequency slots is greatly reduced as compared with CP-OFDM and therefore interference coordination should be much easier. This will be exploited in future investigations.

Chapter 4

Resource Allocation and Scheduling Methods

4.1 Introduction

Resource allocation in multi-carrier systems such as FBMC, refers mainly to the allocation of time and frequency per user, in order to transmit an amount of data which will maintain the required QoS level per case. A resource allocation method in wireless multi-user systems consists of two main components, namely the Scheduler and the Resource Allocator. The Scheduler decides on the order of packet transmissions for all transmitting users, while the Resource Allocator makes the real allocation of exact time slots and frequencies for each transmission. In this chapter, two different approaches of the general resource allocation problem are described, that aim to take advantage of the good frequency selectivity of FBMC to result in considerably improved performance compared to CP-OFDMA. The first approach considers separated scheduling and resource allocation procedures, where the Scheduler provides the Resource Allocator with a prioritized list of packets for transmission. In the second approach the two procedures are operating jointly to investigate if this case improve performance. Besides possible improvements of the two schemes on their own, comparative studies on the two will be done which involve not only system performance and robustness, but also their prospective to benefit from FBMC systems.

4.2 Separated resource allocation and scheduling approach

4.2.1 An Overview

4.2.1.1 Introduction

The resource allocation problem at the downlink of a multicarrier system is addressed, where scheduling of packets, subchannel assignment, power allocation and adaptive mod-

ulation and coding (AMC) are jointly taken into consideration to optimize system performance based on the prerequisite that the *quality of service* (QoS) requirements of individual users are satisfied. Since parameters from the MAC and PHY layers are jointly optimized, we term the proposed scheme as a *Cross-Layer Assisted Resource Allocation*, or *CLARA* in short.

To be more specific, consider the base station employing a multicarrier system as the downlink transmitter in an isolated cell. There are a number of users, say U users in the cell, each having one of the five types of service flows as defined in IEEE 802.16, namely Unsolicited grant service (UGS), Real-time polling service (rtPS), Extended real-time polling service (ertPS), Non-real-time polling service (nrtPS) and Best effort service (BE). Determined by the characteristics of the service flow, the traffic from each user has its own density and delay tolerance. Thus at the data link layer of the BS arrive different amount of packets with various latency requirements from the U users. The problem we explore then is how to serve as many packets as possible, by using the available bandwidth and transmit power.

In the second section of this document, the basic resource allocation procedure and system structure are introduced. After that the machinery of the *resource allocator* (RA), which is one of the system's two main components, is explained in detail. Due to the lack of appropriate performance metrics and a comparative benchmark, simulation results are absent and will be presented in the coming reports.

4.2.1.2 Basic Procedure and System Structure

The resource allocation procedure is done on a per *slot* basis, where a *slot* is a short time period of length T during which the wireless channel is assumed to stay constant. As information bits loaded onto consecutive slots are independently modulated and coded, a slot can formally be referred to as a *Transmission Time Interval* (TTI). For each TTI, the scheduler, located at the MAC layer, receives a number of packets passed down from higher layers. Depending on the QoS requirements and previous statistics, the scheduler decides which packets are to be served and in which priority order, and provides the prioritized list of packets to the resource allocator. The RA then looks for the specific subchannel assignment, power allocation and modulation and coding scheme (MCS) that could serve the list of packets best, under the current channel realization.

From the procedure description above, it can be seen that beside the two central components scheduler and RA, there are three auxiliary components needed in the system to make simulations and evaluations possible: a traffic modeler, a channel generator, and a statistics module. Moreover, a control unit is necessary for scenario setup and system initialization. The basic structure of the system and the interconnections between the components are shown in Figure 4.1, and explained in the following.

- Control unit
  As the initialization step, the control unit creates U user objects, the properties of each describe the physical status (*e.g.*, downlink or uplink, distance from the BS)

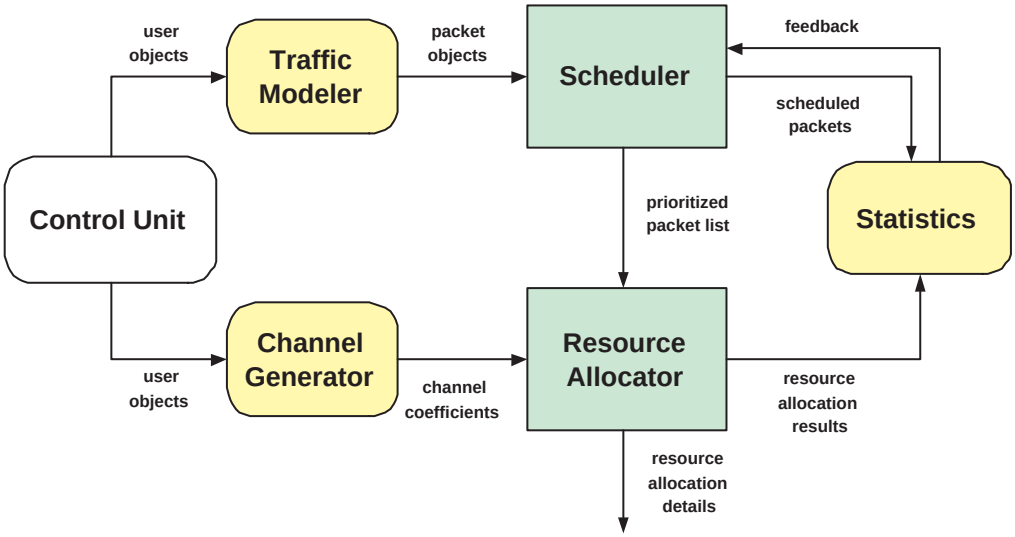


Figure 4.1: System Structure

as well as the QoS requirements (*e.g.*, service flow type, minimum sustainable data rate, maximum latency) of the particular user. These user objects are assumed to be static over a large number of TTI's.

- Traffic modeler
  The traffic modeler simulates the data traffic of the U users as various numbers of packets with different lengths and latency requirements, such that each user gets satisfactory QoS. It then passes these packet objects to the scheduler, as the upper layers in a communication system do.

- Scheduler
  The input packets are scheduled and put into a prioritized list. The decision is informed to the statistics component.

- Channel generator
  The channel generator randomly generates the path amplitudes the users are experiencing, and computes the channel gains of the users with path losses taken into account.

- Resource allocator
  With a prioritized packet list as input for every TTI, the RA outputs not only the resource allocation details such as a specific subchannel assignment and power allocation to certain physical layer modules, but also the servabilities of the packets from the original list.

- Statistics
  The resource allocation results for consecutive TTI's are stored and processed at the

statistics component, which helps the scheduler in providing the packet list and the evaluation of the system performance.

4.2.2 The Resource Allocator Design

The main task of the RA is to find the resource allocation that could serve the maximum number of packets from the input packet list. Such a performance optimization under the total available resources constraint falls into the category of *bottom-up* optimizations, since the problem starts with the physical resources from the bottom of the communication system, finds the mapping from resources to certain QoS parameters, and optimizes these QoS parameters which are critical to the applications supported by upper layers in the system. Due to the complexity of the maximization, we propose an RA algorithm that gives a suboptimal solution by solving a sequence of *top-down* optimizations, which to the contrary of bottom-up optimizations, start with the QoS the applications ask for, find the mapping from these QoS parameters to physical resources, and minimize the amount of resources needed to support such QoS. In our case, the minimum amount of resources needed to serve a certain number of packets is computed and compared to the total amount available. If what is needed is more than what is available, then at least one of the packets has to be dropped or replaced by some other packets. Thus another top-down optimization is formed, and the procedure repeats until the resources we have are enough to serve the left packets. The strict priority order of the input packets given by the scheduler helps the RA choose the subset of packets to serve and avoid the enumeration of all possible packet combinations, which is evidently impossible when the number of packets is large.

In this section we first introduce the system model which adopts the cross-layer framework from [13], where adjustments are done to accommodate the special features of multicarrier systems. Then the bottom-up capacity maximization problem and the top-down power minimization problem are formally formulated. The algorithm we have developed is explained afterwards.

4.2.2.1 System Model

We consider the downlink scenario of an isolated single-cell multicarrier system with U active users each having one service flow that generates packets. As the detailed characteristics of the service flows are dealt with by the scheduler and made transparent to the RA, the RA only needs to care for the *latency* requirement of each packet which is defined as:

Definition: The latency time τ_k is the delay packet k experiences until it is received correctly with an outage probability of no more than the predefined value $\pi^{(\text{out})}$. Let $f_k[m]$ be the probability that it takes exactly m TTI's to transmit packet k error-free, and M_k be the minimum number of transmissions needed to guarantee that in a fraction of $1-\pi^{(\text{out})}$

out of all cases the packet transmission is successful, *i.e.*,

$$M_k = \min_M M \quad \text{s.t.} \quad \sum_{m=1}^{M} f_k[m] \geq 1 - \pi^{(\text{out})}. \tag{4.1}$$

Then $\tau_k = (M_k - 1)(\text{RTD} + T) + T$ where RTD represents *round trip delay*.

In the following subsections, the mathematical descriptions of the regarded system components are derived which lay the basis for cross-layer optimization.

4.2.2.1.1 Channel Model The downlink broadcast channel is modeled as frequency-selective fading over the total system bandwidth and frequency-flat fading over each *subchannel*, which is consist of N_c adjacent subcarriers. FDMA is employed meaning the assignment of every subchannel is exclusive to one packet, and *intercarrier interference* (ICI) is not taken into account. Note that although multiple packets from the same user could share one subchannel, we forbid such cases since those packets could have different latency requirements and it is much simpler to keep the modulation and coding scheme consistent over one TTI. Moreover, we restrict ourselves here to the single-antenna case both at the base station (BS) and at the mobile stations (MS).

Let $H_{k,n}$ and $\sigma^2_{k,n}$ be the channel coefficient and Gaussian noise variance at the reception of packet k to user u_k on the nth subchannel, and p_n be the amount of power allocated on subchannel n. When assigned to packet k, the *signal-to-noise-ratio* (SNR) on subchannel n can be computed as

$$\gamma_{k,n} = \frac{|H_{k,n}|^2}{\sigma^2_{k,n}} \cdot p_n. \tag{4.2}$$

Note that throughout this work the index k refers to packets and index n refers to subchannels.

We choose the TTI to be of length $T = 5$ ms and assume that the channel condition is constant during one TTI. The WiMAX standard suggests a symbol duration of 102.9 μs in a system with 10 MHz bandwidth and an FFT size of 1024. Based on this number we assume that one TTI contains $N_\text{s} = 40$ symbols for data transmission.

4.2.2.1.2 FEC coding and modulation We assume that modulation and coding is done on a per subchannel basis, and with reference to the WiMAX standard 7 modulation and coding schemes (MCS) are chosen as candidates, which are listed in Table 4.1.

Since with the help of cyclic prefix or an equalizer, intersymbol interference is not present in the system, each subchannel can be modeled as a *discrete memoryless channel* (DMC) over which the *noisy channel coding theorem* [14] can be applied. Let the modulation alphabet and the coding rate on the nth subchannel be $\mathcal{A}_n = \{a_1, \ldots, a_{A_n}\}$ and R_n respectively. The *cutoff rate* of subchannel n with SNR $\gamma_{k,n}$ can be expressed as

$$R_0(\gamma_{k,n}, A_n) = \log_2 A_n - \log_2 \left[1 + \frac{2}{A_n} \sum_{m=1}^{A_n - 1} \sum_{l=m+1}^{A_n} e^{-\frac{1}{4}|a_l - a_m|^2 \gamma_{k,n}} \right]. \tag{4.3}$$

Table 4.1: Modulation and Coding Schemes (MCS)

Index	Modulation Type	Alphabet Size A	Code Rate R	$R\log_2 A$
1	BPSK	2	1/2	0.5
2	QPSK	4	1/2	1
3	QPSK	4	3/4	1.5
4	16-QAM	16	1/2	2
5	16-QAM	16	3/4	3
6	64-QAM	64	2/3	4
7	64-QAM	64	3/4	4.5

Note that the cutoff rate is monotonically increasing with SNR when the modulation alphabet is fixed, yet it is not monotonic with varying modulation levels when SNR is fixed, which is rather an unfavorable property regarding optimizations.

According to the noisy channel coding theorem, there always exists a block code with block length l and binary code rate $R_n \log_2 A_n \leq R_0(\gamma_{k,n}, A_n)$ in bits per subchannel use, such that with maximum likelihood decoding the error probability $\tilde{\pi}_{k,n}$ of a code word satisfies

$$\tilde{\pi}_{k,n} \leq 2^{-l\left(R_0(\gamma_{k,n},A_n)-R_n\log_2 A_n\right)}. \tag{4.4}$$

In order to apply this upper bound on code word error probability to the extensively used turbo decoded convolutional code, quantitative investigations have been done in [13] and an expression for the *equivalent block length* is derived based on link level simulations. The result from [13] shows that the performance of a turbo decoded convolutional code of rate R_n applied to B_n information bits in a very good approximation equals the performance of a block code with block length

$$n_{\text{eq}} = \beta_{n_{\text{eq}}} \ln L_n, \tag{4.5}$$

where $L_n = B_n/R_n$ and parameter $\beta_{n_{\text{eq}}}$ is used to adapt this model to the specifics of the employed turbo code. In our situation L_n can be plugged in with the number of bits contained in one TTI on subchannel n, *i.e.*, $L_n = N_{\text{s}} N_{\text{c}} \log_2 A_n$. Consequently, the transmission of L_n bits is equivalent to the sequential transmission of L_n/n_{eq} blocks of length n_{eq} and has an error rate of

$$\begin{aligned}\pi_{k,n} &= 1-(1-\tilde{\pi}_{k,n})^{\frac{L_n}{n_{\text{eq}}}} \\ &\leq 1-\left(1-2^{-n_{\text{eq}}\left(R_0(\gamma_{k,n},A_n)-R_n\log_2 A_n\right)}\right)^{\frac{L_n}{n_{\text{eq}}}}.\end{aligned} \tag{4.6}$$

4.2.2.1.3 Protocol At the MAC layer an *automatic repeat request* (ARQ) protocol is employed. The data sequence transmitted in one TTI over one subchannel, which will be referred to as a *subpacket*, is used as the retransmission unit since it is independently

decodable. We set no limit on the maximum number of retransmissions and consider the case where the corrupted subpackets are simply abandoned at the receiver, *i.e.*, no Hybrid ARQ.

Denote the set of subchannels assigned to packet k as $\mathcal{S}_k$, the number of information bits in packet k as b_k, *i.e.*, length of the packet uncoded, and the length of its part loaded on subchannel n as $b_{k,n}$. The completeness of the transmission of packet k requires

$$\sum_{n\in\mathcal{S}_k} b_{k,n} = b_k. \tag{4.7}$$

On the other hand, the latency time τ_k is obviously determined by the largest subpacket error rate of packet k, denoted by $\pi_k = \max_{n\in\mathcal{S}_k} \pi_{k,n}$. Assuming that the subpacket error probability of a retransmitted subpacket is the same as that of its original transmission, then $f_k[m] = \pi_k^{m-1}(1-\pi_k)$ becomes a geometric series with ratio π_k, and

$$\sum_{m=1}^{M} f_k[m] = 1 - \pi_k^M \geq 1 - \pi^{(\mathrm{out})} \tag{4.8}$$

implies that for packet k, the number of transmissions needed to keep the outage probability below $\pi^{(\mathrm{out})}$ is

$$M_k = \left\lceil \frac{\ln \pi^{(\mathrm{out})}}{\ln \pi_k} \right\rceil. \tag{4.9}$$

The latency time τ_k then follows from its definition.

The quantities mentioned in this section, their notations, as well as their simulation values are summarized in Table 4.2.

Table 4.2: System Parameters

Total bandwidth		10 MHz
Center frequency	f_c	2.5 GHz
FFT size		1024
Number of data subcarriers		720
Number of subchannels	N	30
Number of subcarriers per subchannel	N_c	$720/30 = 24$
Transmission Time Interval (TTI)	T	5 ms
Number of data symbols per TTI	N_s	40
Round Trip Delay (RTD)	RTD	10 ms
Turbo code dependent parameter	β_{n_eq}	32
Outage probability	$\pi^{(\mathrm{out})}$	0.01
Number of users in the cell	U	8

4.2.2.2 Cross-layer Optimization

Recall that for each TTI of length T, the scheduler provides the RA with a prioritized list of K_{tot} packets, each in the format of

User ID u_k	Latency requirement $\tau_k^{(\text{rq})}$	Length in bits b_k.

Ideally, the capacity maximizing resource allocation is given by the solution to the optimization

$$\begin{aligned}
\max_{K,\{\mathcal{S}_k\},\boldsymbol{p},\boldsymbol{A},\boldsymbol{R}} \quad & K \\
\text{s.t.} \quad K &\in \{0,\ldots,K_{\text{tot}}\} \\
\mathbf{1}^{\text{T}}\boldsymbol{p} &\leq P_{\text{tot}}, \\
\mathcal{S}_i \cap \mathcal{S}_j &= \emptyset, && i,j=1,\ldots,K, i\neq j \\
\cup_{k=1}^{K}\mathcal{S}_k &\subseteq \mathcal{N}, \\
(A_n,R_n) &\in \mathcal{M}, && n=1,\ldots,N \\
\textstyle\sum_{n\in\mathcal{S}_k} R_n \log_2 A_n &\geq \tfrac{b_k}{N_{\text{s}}N_{\text{c}}}, && k=1,\ldots,K \\
\tau_k &\leq \tau_k^{(\text{rq})}, && k=1,\ldots,K,
\end{aligned} \tag{4.10}$$

where K is the number of packets getting served, $\boldsymbol{p}, \boldsymbol{A}, \boldsymbol{R}$ are $N \times 1$ vectors of power values, modulation alphabet size and coding rate on each subchannel, P_{tot} is the total available transmit power, $\mathcal{N}$ is the set of all subchannels, and $\mathcal{M}$ is the set of all available MCS. The objective of the optimization is to maximize the number of packets that could be served, and the constrains include total transmit power, FDMA, available MCS, and latency requirements of the packets.

On the other hand, the top-down optimization problem of minimizing transmit power on serving K chosen packets is formulated as

$$\begin{aligned}
\min_{\{\mathcal{S}_k\},\boldsymbol{p},\boldsymbol{A},\boldsymbol{R}} \quad & \mathbf{1}^{\text{T}}\boldsymbol{p} \\
\text{s.t.} \quad \mathcal{S}_i \cap \mathcal{S}_j &= \emptyset, && i,j=1,\ldots,K, i\neq j \\
\cup_{k=1}^{K}\mathcal{S}_k &\subseteq \mathcal{N}, \\
(A_n,R_n) &\in \mathcal{M}, && n=1,\ldots,N \\
\textstyle\sum_{n\in\mathcal{S}_k} R_n \log_2 A_n &\geq \tfrac{b_k}{N_{\text{s}}N_{\text{c}}}, && k=1,\ldots,K \\
\tau_k &\leq \tau_k^{(\text{rq})}, && k=1,\ldots,K,
\end{aligned} \tag{4.11}$$

and its optimum value is denoted by $P_{\min}$.

It can be observed that optimization (4.10) is discrete, non-convex, and of a combinatorial nature. If all packets are treated equally, the complexity of the optimization grows exponentially with K_{tot} and is thus intractable. Fortunately with the priority order of packets given by the scheduler, we propose a feasible way to suboptimally tackle the problem, wherein problem (4.11) plays the role of feasibility tests.

Our principle is to serve as many packets as possible according to the given priority order. Starting with the first one, the packets from the list are sequentially added to the set of servable packets $\mathcal{P}_{\text{S}}$, as long as the current expansion to set $\mathcal{P}_{\text{S}}$ passes the feasibility test, *i.e.*, the optimal value $P_{\min}$ to problem (4.11) satisfies $P_{\min} \leq P_{\text{tot}}$. When the current

test fails, the packet under consideration is marked as unservable and the program moves on to the next prioritized packet. Following this routine, problem (4.11) needs to be solved K_{tot} times before we find $\mathcal{P}_{\text{S}}$, which still takes much computational effort.

We exploit two means to further reduce the number of iterations for the RA. As $R_n \log_2 A_n$ is within the range from 0.5 to 4.5, we set both an upper and a lower bound on the number of subchannels that can be assigned to packet k as

$$N_k^{(\text{l})} = \left\lceil \frac{b_k}{4.5 N_{\text{s}} N_{\text{c}}} \right\rceil, \quad N_k^{(\text{u})} = \left\lceil \frac{b_k}{0.5 N_{\text{s}} N_{\text{c}}} \right\rceil. \tag{4.12}$$

In the multicarrier system context, the lower bound $N_k^{(\text{l})}$ is well suited for a preliminary check whether there are enough subchannels to support the chosen subset of packets, and the upper bound $N_k^{(\text{u})}$ is useful in subchannel assignment, which will become clear in the next section. Let the current set of servable packets be $\mathcal{P}_{\text{S}}$, if the upcoming packet k' yields $N_{k'}^{(\text{l})} + \sum_{k \in \mathcal{P}_{\text{S}}} N_k^{(\text{l})} > N$ where N is the total number of subchannels, it can be abandoned directly without putting to the feasibility test which is bound to fail.

The second usage of $N_k^{(\text{l})}$ is to roughly estimate how many consecutive packets from the highest priority can be served, *i.e.*, solution to problem

$$K = \max_{K'} K' \quad \text{s.t.} \quad \sum_{k=1}^{K'} N_k^{(\text{l})} \leq N \tag{4.13}$$

provides an initial $\mathcal{P}_{\text{S}}$ as $\{1, \ldots, K\}$. Therefore instead of creating and expanding $\mathcal{P}_{\text{S}}$ one packet after another, we could start with a $\mathcal{P}_{\text{S}}$ of reasonable size and delete the member packets that turn out unservable, and then follow the principal routine to add packets with relatively lower priorities. This modification to the original method reduces the number of iterations only on average.

The basic RA algorithm is summarized in Algorithm 1.

4.2.2.3 The Three-step Approach

The non-convex top-down optimization problem has intrinsically a complicated structure in that the optimization variables are discrete and closely related to each other causing a direct decomposition of the original problem impossible. Regarding the complexity of finding the optimal solution, searching for the best subchannel assignment alone has reached a complexity of $O(K^N)$, which is obviously computationally intractable. In this work we employ a three-step approach, *i.e.*, first the subchannel assignment (SA) is determined, then power is allocated to each subchannel, at last the subchannel assignment and power allocation (PA) are adjusted based on the result of the first two steps. The basic idea is that at each step, some variables are kept fixed while some others are being optimized: in the SA step, the MCS on every subchannel is fixed and different assignments are compared by computing the power needed on each subchannel to achieve the required packet error rate; in the PA step, the subchannel assignment is fixed and the best combination of MCS on each subset of subchannels is found.

Algorithm 1 Resource Allocation Procedure by RA

Require: A prioritized list of K_{tot} packets with QoS requirements
Ensure: The set of packets $\mathcal{P}_{\text{S}}$ that could be served

```
Compute N_k^(l), N_k^(u)
K ← max K'  s.t.  Σ_{k=1}^{K'} N_k^(l) ≤ N
P_S ← {1, ..., K}
Solve the top-down optimization (4.11) with P_S
while P_min > P_tot do
  P_S ← P_S \ {K}
  K ← K − 1
  Solve the top-down optimization (4.11) with P_S
end while
if K < K_tot − 1 & Σ_{k=1}^{K} N_k^(l) < N then
  k' ← K + 2
  while k' ≤ K_tot do
    while k' ≤ K_tot & N_{k'}^(l) + Σ_{k∈P_S} N_k^(l) > N do
      k' ← k' + 1
    end while
    if k' ≤ K_tot then
      Solve the top-down optimization (4.11) with P_S ∪ {k'}
      if P_min ≤ P_tot then
        P_S ← P_S ∪ {k'}
        K ← K + 1
      end if
      k' ← k' + 1
    end if
  end while
end if
```

4.2.2.3.1 Subchannel Assignment (SA) From the derivations in Section 4.2.2.1 it can be observed that with fixed MCS, the PER decreases with increasing allocated power onto the subchannel. Note that as the latency time a packet has to experience is given by

$$\tau_k = (\left\lceil \frac{\ln \pi^{(\text{out})}}{\ln \pi_k} \right\rceil - 1)(\text{RTD} + T) + T \quad \text{where} \quad \pi_k = \max_{n \in \mathcal{S}_k} \pi_{k,n}, \tag{4.14}$$

letting the PER on each subchannel assigned to packet k be of the same value requires the minimum power, for otherwise the total required power can be reduced without increasing the latency time. Therefore we fix the PER on subchannels assigned to packet k to its maximum allowable value, *i.e.*,

$$\pi_{k,n} = \pi_k = \pi_k^{(\text{rq})} = \sqrt[M_k]{\pi^{(\text{out})}}, \quad \forall n \in \mathcal{S}_k, \tag{4.15}$$

where

$$M_k = \left\lfloor \frac{\tau_k^{(\text{rq})} - T}{\text{RTD} + T} + 1 \right\rfloor. \tag{4.16}$$

We assume that the MCS $(R, A) = (0.75, 64)$ is used on every subchannel. According to (4.3), (4.5), (4.6), the minimum power required to achieve $\pi_k^{(\text{rq})}$ on any subchannel can be computed by using a bisection method, and the results are recorded in matrix $\boldsymbol{P} \in \mathbb{R}_+^{K \times N}$. The subchannel assignment problem is formulated as

$$\begin{aligned} \min_{\{\mathcal{S}_k\}} & \sum_{k=1}^{K} \sum_{n \in \mathcal{S}_k} p_{k,n} \\ \text{s.t.} \quad & \mathcal{S}_i \cap \mathcal{S}_j = \emptyset, \qquad i, j = 1, \ldots, K, i \neq j \\ & \cup_{k=1}^{K} \mathcal{S}_k \subseteq \mathcal{N}, \\ & |\mathcal{S}_k| \geq N_k^{(\text{l})}, \quad k = 1, \ldots, K \\ & |\mathcal{S}_k| \leq N_k^{(\text{u})}, \quad k = 1, \ldots, K, \end{aligned} \tag{4.17}$$

i.e., from each column of $\boldsymbol{P}$ one entry is picked such that the kth row has between $N_k^{(\text{l})}$ and $N_k^{(\text{u})}$ picked entries, and the sum of all picked entries is minimized.

Intuitively we will first pick up the minimum entry from every column in $\boldsymbol{P}$. In the ideal case, if such a choice happens to fulfill $N_k^{(\text{l})} \leq |\mathcal{S}_k| \leq N_k^{(\text{u})}, k = 1, \ldots, K$, then solution to (4.17) is found. Normally a set of unsatisfied packets with $|\mathcal{S}_k| < N_k^{(\text{l})}$, and at the same time a set of oversatisfied packets with $|\mathcal{S}_k| > N_k^{(\text{l})}$ are obtained, which are denoted by $\mathcal{K}_\text{u}$ and $\mathcal{K}_\text{o}$ respectively. We deprive from the oversatisfied packets the subchannels with the least advantage assigning to them, until $\forall k \in \mathcal{K}_\text{o}, |\mathcal{S}_k| = N_k^{(\text{l})}$. By "the least advantage" it is meant that the difference between the absolute minimum entry and minimum entry among unsatisfied packets in a column of $\boldsymbol{P}$ is the smallest among all columns assigned to the oversatisfied packet, *i.e.*, assigning such a column to an unsatisfied packet would probably not hurt much. All the deprived subchannels form a set of extra subchannels $\mathcal{N}_\text{e}$. The same assigning procedure is then repeated on $\mathcal{K}_\text{u}$ and $\mathcal{N}_\text{e}$. The recursion stops

when there are no more unsatisfied packets. If some subchannels are left unassigned, *i.e.*, some oversatisfied packets are over-deprived, theses subchannels are assigned back to those packets, until they reach the upper bound $N_k^{(\mathrm{u})}$.

The SA algorithm is summarized in Algorithm 2. It is a greedy algorithm in the sense that during the assignment, the unsatisfied and exactly-satisfied packets never give up the subchannels already assigned to them, unless the unsatisfied packets turn into oversatisfied packets eventually.

Algorithm 2 Subchannel assignment

Require: $\boldsymbol{P} = (p_{k,n})$, $N_k^{(\mathrm{u})}$ and $N_k^{(\mathrm{l})}$
  $\mathcal{S}_k \leftarrow \emptyset, k = 1, \ldots, K$
  for $n = 1, \ldots, N$ **do**
    $k \leftarrow \mathrm{argmin}_k\, p_{k,n}$, $\mathcal{S}_k \leftarrow \mathcal{S}_k \cup \{n\}$
  end for
  $\mathcal{K}_{\mathrm{u}} \leftarrow \{k : |\mathcal{S}_k| < N_k^{(\mathrm{l})}\}$, $\mathcal{N}_{\mathrm{e}} \leftarrow \emptyset$
  while $\mathcal{K}_{\mathrm{u}} \neq \emptyset$ **do**
    for each $k \in \{k : |\mathcal{S}_k| > N_k^{(\mathrm{l})}\}$ **do**
      while $|\mathcal{S}_k| > N_k^{(\mathrm{l})}$ **do**
        $n \leftarrow \mathrm{argmin}_{n \in \mathcal{S}_k} \left(\min_{k' \in \mathcal{K}_{\mathrm{u}}} (p_{k',n} - p_{k,n})\right)$
        $\mathcal{N}_{\mathrm{e}} \leftarrow \mathcal{N}_{\mathrm{e}} \cup \{n\}$, $\mathcal{S}_k \leftarrow \mathcal{S}_k \setminus \{n\}$
      end while
    end for
    for each $n \in \mathcal{N}_{\mathrm{e}}$ **do**
      $k = \mathrm{argmin}_{k \in \mathcal{K}_{\mathrm{u}}}\, p_{k,n}$, $\mathcal{S}_k \leftarrow \mathcal{S}_k \cup \{n\}$
    end for
    $\mathcal{K}_{\mathrm{u}} \leftarrow \{k : |\mathcal{S}_k| < N_k^{(\mathrm{l})}\}$, $\mathcal{N}_{\mathrm{e}} \leftarrow \emptyset$
  end while
  $\mathcal{N}_{\mathrm{e}} \leftarrow \{n :$ extra subchannels assigned to the last $k \in \mathcal{K}_{\mathrm{u}}\}$
  $\mathcal{K}_{\mathrm{a}} \leftarrow \{k : |\mathcal{S}_k| < N_k^{(\mathrm{u})}\}$
  for each $n \in \mathcal{N}_{\mathrm{e}}$ **do**
    $k = \mathrm{argmin}_{k \in \mathcal{K}_{\mathrm{a}}}\, p_{k,n}$, $\mathcal{S}_k \leftarrow \mathcal{S}_k \cup \{n\}$, update $\mathcal{K}_{\mathrm{a}}$
  end for

4.2.2.3.2 Power Allocation (PA) With the SA result as input, power allocation is no longer coupled among the packets when intercarrier interference (ICI) is neglected. The top-down optimization (4.11) can therefore be decomposed into K independent optimizations as

$$\min \sum_{n \in \mathcal{S}_k} p_n \quad \text{s.t.} \quad \sum_{n \in \mathcal{S}_k} R_n \log_2 A_n \geq \frac{b_k}{N_{\mathrm{s}} N_{\mathrm{c}}}, \tau_k \leq \tau_k^{(\mathrm{rq})}. \tag{4.18}$$

Firstly we look for all *efficient* MCS combinations on $\mathcal{S}_k$ such that

$$\sum_{n\in\mathcal{S}_k} R_n \log_2 A_n \geq \frac{b_k}{N_\mathrm{s} N_\mathrm{c}} \tag{4.19}$$

is fulfilled, which form a set denoted by $\mathcal{M}_\mathrm{c}$. An MCS combination M_c is said to be *efficient* if $\nexists M_\mathrm{c}' \neq M_\mathrm{c}$ such that $\forall n \in \mathcal{S}_k, (R_n \log_2 A_n)_{M_\mathrm{c}'} \leq (R_n \log_2 A_n)_{M_\mathrm{c}}$ and still fulfills (4.19). Apparently, to achieve the same PER, the most power-saving MCS combination is always an efficient one.

For each MCS combination, the minimum power required to achieve $\pi_k^{(\mathrm{rq})}$ on every subchannel can be computed using a bisection method. Note that for each MCS, the SNR that is required to achieve $\pi_k^{(\mathrm{rq})}$ is actually independent from channel realizations and can thus be computed and stored off-line. When tracking the power values, only a simple additional calculation is needed, *i.e.*,

$$p_{k,n} = \frac{\sigma_{k,n}^2}{|H_{k,n}|^2} \cdot \gamma_{k,n}^{(\mathrm{rq})}. \tag{4.20}$$

Therefore, even if a packet is large and takes many subchannels which gives a large volumed $\mathcal{M}_\mathrm{c}$, the sum power comparison between each MCS combination is not computationally expensive.

The power allocation algorithm is outlined in Algorithm 3.

Algorithm 3 Power allocation for packet k

Require: $\mathcal{S}_k$, $\pi_k^{(\mathrm{rq})}$ and b_k
  $\mathcal{M}_\mathrm{c} \leftarrow \{\text{efficient } \{(R_n, A_n) : n \in \mathcal{S}_k\} \text{ satisfying (4.19)}\}$
  $P_{\min,k} \leftarrow \infty$
  for each $M_\mathrm{c} \in \mathcal{M}_\mathrm{c}$ **do**
    $P_k \leftarrow \sum_{n\in\mathcal{S}_k} p_n \quad \text{s.t.} \quad \pi_{k,n} = \pi_k^{(\mathrm{rq})}, \forall n \in \mathcal{S}_k$
    $P_{\min,k} \leftarrow \min\{P_{\min,k}, P_k\}$
  end for

4.2.2.4 Adjustment

The outcome of PA might indicate zero MCS on some subchannels, which means these subchannels are released from occupation, and can be assigned again to other packets. The reassignment is also based on $\boldsymbol{P}$, and the packets getting new subchannels have to go through the PA procedure again, after which the total consumed power is compared with the previously obtained value. We set the limit that each subchannel has only one chance to be reassigned so that the program does not trap in a dead loop. The adjustment phase is in fact an amendment to SA.

4.2.2.5 Outlooks

At the moment the integration of all components of the system is successful, and with the test setups we have validated the system functionality. Yet to come up with more meaningful and representative simulation scenarios, and more importantly, to find a comparative benchmark to the system, joint discussions with our partners are necessary.

On the other hand, to demonstrate the benefits of filter bank based multicarrier systems, we plan to investigate the resource allocation problem in the uplink scenario, where synchronization among the users is hardly perfect. Also possible is to include intercell interference into the system model. Last but not the least, the case that multiple antennas are employed both at the transmitter and the receiver should be explored, which, though rendering the problem more complicated, is expected to give a considerable performance improvement.

4.2.3 The Scheduler design

4.2.3.1 QoS provision mechanisms

The system architecture of WiMAX consists of Base Stations (BSs), each one responsible for a specific area cell, and stationary Subscriber Stations (SSs). Two operation modes are defined: Point-to-Multipoint (PMP), for communication between the BS and the SSs of its cell, and Mesh mode for direct SS-to-SS communications without the need of a BS. At the PMP mode, each BS regulates all the communication in its cell. The communication path between SS and BS has two directions: uplink (from SS to BS) and downlink (from BS to SS), multiplexed either with Time Division Duplex (TDD) or Frequency Division Duplex (FDD). Transmission parameters, including the modulation and coding schemes, may be adjusted individually for each SS on a frame-by-frame basis. A TDD frame has a fixed duration which may take various values: 0.5, 1, 2, 2.5, 4, 5, 8, 10, 12.5 or 20 msec. In any case, the frame is divided into a downlink subframe, and an uplink subframe. The TDD framing is adaptive in that the bandwidth allocated to the downlink versus the uplink direction may vary. Each subframe consists of an integer number of Physical Slots (PSs) representing the minimum portion of allocated bandwidth. The downlink subframe begins with information necessary for frame synchronization and control. A Frame Start Preamble is used for synchronization and equalization. This is followed by the frame control section, containing the DL-MAP and UL-MAP fields, that state the PSs at which bursts begin in both directions. The following frame portion carries the data, organized into bursts with different burst profiles and therefore different levels of transmission robustness. Each SS receives and decodes the control information of the downlink direction contained in the DL-MAP and looks for MAC headers indicating data for that SS in the remainder of the downlink subframe. Through the UL-MAP, the BS determines the transmission opportunities of its subordinates SSs, based on the bandwidth requests of each SS. Bandwidth requests are transmitted through special purpose information elements referred as BW-Requests. Each SS having decoded the corresponding control information contained in UL-MAP, knows exactly during which PSs of the uplink subframe it is allowed to transmit

and what kind of transmission it can make. In both directions, data bits are randomized, FEC encoded, and mapped to one of the five mandatory (Spread BPSK, BPSK, QPSK, 16-QAM, 64-QAM) or optional 256-QAM signal constellation.

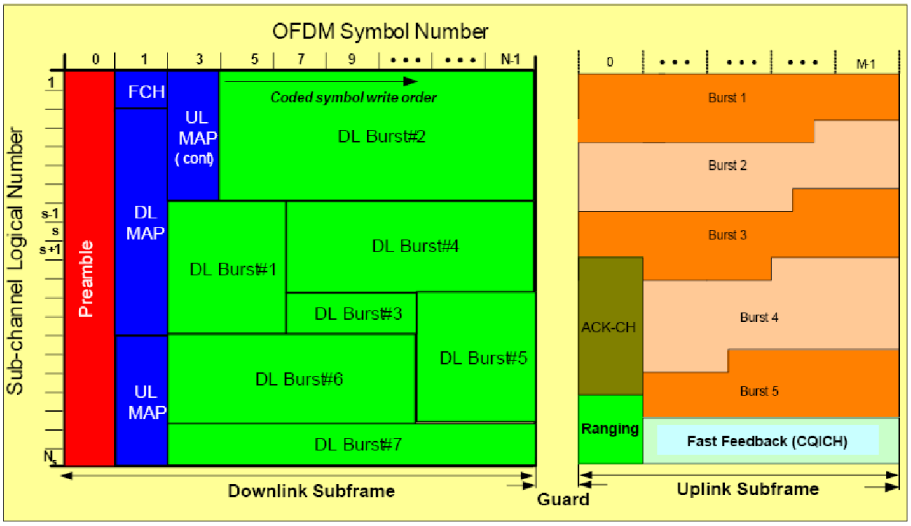


Figure 4.2: WiMAX TDD frame structure

IEEE 802.16 can support multiple communication services (data, voice, video, etc.) with different QoS requirements organized into different connections. Each connection is associated with a single service flow and specifies a set of traffic and QoS parameters that quantify its traffic behavior and QoS expectations. This set includes:

- minimum reserved traffic rate (in bits/sec),
- maximum sustained traffic rate (in bits/sec),
- maximum latency (in ms),
- tolerated jitter (maximum delay variation in ms),
- traffic priority (values 0-7, with 7 the highest),
- etc.

The standard defines four different services:

- *Unsolicited Grant Service (UGS)*: This service supports real-time data streams consisting of fixed-size data packets transmitted at periodic intervals, such as Voice over IP without silence suppression. These applications require constant bandwidth allocation, so bandwidth requests are not required.

QoS Category	Application	Specification
UGS Unsolicited Grant Service	VOIP	•Maximum Sustained Rate •Maximum Latency Tolerance •Jitter Tolerance
rtPS Real-Time Polling Service	Streaming Audio or Video	•Maximum Sustained Rate •Minimum Reserved Rate •Maximum Latency Tolerance •Traffic Priority
ErtPS Extend Real-Time Polling Service	Voice with Activity Detection	•Maximum Sustained Rate •Minimum Reserved Rate •Maximum Latency Tolerance •Jitter Tolerance •Traffic Priority
nrtPS Non-Real-Time polling service	File Transfer Protocol	•Maximum Sustained Rate •Minimum Reserved Rate •Traffic Priority
BE Best-Effort Service	Data Transfer, Web Browsing	•Maximum Sustained Rate •Traffic Priority

Figure 4.3: QoS

- *Real-time Polling Service (rtPS)*: This service supports data streams consisting of variable-sized data packets that are transmitted at fixed intervals, such as MPEG video. These applications have specific bandwidth requirements, as well as a maximum acceptable latency. Late packets that miss the deadline are considered useless.

- *Non-real-time Polling Service (nrtPS)*: This service is for non-real-time connections that require better than best effort service, e.g., bandwidth intensive file transfer. These applications are time-insensitive but require a minimum bandwidth allocation.

- *Best Effort service (BE)*: This service is for best effort traffic with no QoS guarantee. The applications of this kind of service share the remaining bandwidth after allocation to the rest of the services is completed. BE uses only contention mode.

In [15], a new service, referred to as enhanced rtPS (ertPS), is defined to better support real-time service flows that generate variable size data packets on a periodic basis, e.g., VoIP with silence suppression. Unsolicited Grant Service (UGS) is designed to support real-time applications, with strict delay requirements, which generate fixed-size data packets at periodic intervals, such as T1/E1. Therefore, UGS is defined so as to closely follow the packet arrival pattern. Grants occur on a periodic basis. The base period and the grant size are specified during the connection setup phase. After that, SSs never request bandwidth for UGS connections. For these reasons, we did not find this scheduling service interesting from a MAC standpoint, and so its performance is not assessed in this paper. Real-time Polling Service (rtPS) is designed to support real-time applications with less stringent

delay requirements, which generate variable-size data packets at periodic intervals, such as Moving Pictures Expert Group (MPEG) video and VoIP with silence suppression. Unlike UGS tailored applications, the size of arriving packets with rtPS is not fixed, thus SSs are required to explicitly make a request for bandwidth from the BS. The standard provides that the BS periodically sends unicast polls to rtPS connections. The base period can be specified during the connection setup. Specifically, it is possible to set the polling period to the interval at which packets are expected to be generated by the application. A unicast poll consists of an uplink allocation from the BS to the polled SS of the bandwidth needed to transmit a bandwidth request PDU. Unlike UGS and rtPS scheduling services, non-real-time Polling Service (nrtPS) and Best Effort (BE) are designed for applications that do not have specific delay requirements. The main difference between them is that nrtPS connections are reserved a minimum amount of bandwidth (by means of the Minimum Reserved Traffic Rate parameter). Additionally, the BS grants unicast polls to nrtPS connections on a large time-scale. The IEEE 802.16 standard specifies this scale to be one second or less. Both nrtPS and BE uplink connections typically use contention-based bandwidth requests. Such requests are sent in response to broadcast/multicast polls, which are advertised by the BS in the ULMAP. The BS is free to use any algorithm to decide which uplink subframe portion is reserved for broadcast/multicast contention slots on a frame-by-frame basis. The main drawback of this mechanism is that a collision occurs whenever two or more SSs access the medium in the same contention slot to send a bandwidth request. A bandwidth request is considered lost (i.e., a collision occurred) if the transmitting SS does not receive the related data grant within a specified timeout (50 ms, in our analysis). To reduce the likelihood of this event, a collision avoidance scheme is used. SSs randomly select a number in the backoff window which indicates the number of contention slots the SSs must defer before transmitting. When collisions occur, a truncated binary exponential backoff algorithm is employed to increase the backoff window. Consequently, this polling mechanism is tailored to serve traffic with no specific delay requirements, such as bursty Web traffic. In addition, an SS can issue an unsolicited bandwidth request for one of its non-UGS backlogged connections by consuming part of the grant that it was allocated for the transmission of data. Optionally, incremental unsolicited bandwidth requests can be piggybacked to PDUs by means of a specific 2 bytes MAC subheader [16].

4.2.3.2 The scheduling algorithm

The traffic scheduler located at the Base Station (BS) decides on the allocation of the physical slots in each time frame. Uplink scheduling is performed by the BS with the aim of providing each Subscriber Station (SS) with enough bandwidth for uplink transmissions or opportunities for extra transmission requests. When additional bandwidth is needed, the SS utilizes its transmission opportunities during contention periods or when it is polled by the BS, depending on its agreed QoS characteristics, to pass its transmission requests. Downlink scheduling on the other hand, considers packets waiting for transmission at the BS as implicit requests for bandwidth allocation. A logical representation of the scheduler's operation is shown in the following figure.

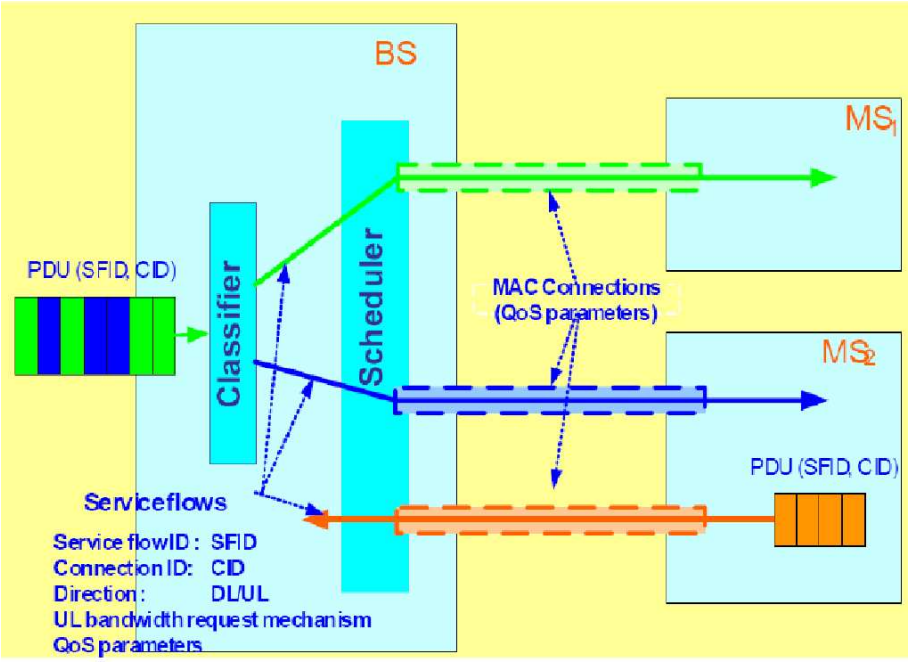


Figure 4.4: Representation of the scheduler's operation

Based on well-accepted studies (e.g., [17]), the scheduler has to combine the following properties:

- **Fast Data Scheduling**: The MAC scheduler must efficiently allocate available resources in response to bursty data traffic and time-varying channel conditions. The scheduler is located at each BS to enable rapid response to traffic requirements and channel conditions. The data packets are associated to traffic connections with well defined QoS parameters in the MAC layer so that the scheduler can correctly determine the packet transmission ordering over the air interface.

- **Scheduling for both DL and UL**: The scheduling service is provided for both DL and UL traffic. In order for the MAC scheduler to make an efficient resource allocation and provide the desired QoS in the UL, the UL must feedback accurate and timely information as to the traffic conditions and QoS requirements. Multiple uplink bandwidth request mechanisms, such as bandwidth request through ranging channel, piggyback request and polling are designed to support UL bandwidth requests. The UL service flow defines the feedback mechanism for each uplink connection to ensure predictable UL scheduler behavior. Furthermore, with orthogonal UL sub-channels, there is no intra-cell interference. UL scheduling can allocate resource more efficiently and better enforce QoS.

- **Dynamic Resource Allocation**: The MAC supports frequency-time resource allocation in both DL and UL on a per-frame basis. The resource allocation is delivered

in MAP messages at the beginning of each frame. Therefore, the resource allocation can be changed frame-by-frame in response to traffic and channel conditions. Additionally, the amount of resource in each allocation can range from one slot to the entire frame. The fast and fine granular resource allocation allows superior QoS for data traffic.

- **QoS Oriented**: The MAC scheduler handles data transport on a connection-by-connection basis. Each connection is associated with a single data service with a set of QoS parameters that quantify the aspects of its behavior. With the ability to dynamically allocate resources in both DL and UL, the scheduler can provide superior QoS for both DL and UL traffic. Particularly with uplink scheduling - the uplink resource is more efficiently allocated, performance is more predictable, and QoS is better enforced.

To efficiently support all types of connections (UGS, rtPS, ertPS, nrtPS and BE) as they provided by the standard, the scheduler designed for PHYDYAS is based on ideas found in [17] and uses a combination of strict priority service discipline, earliest deadline first (EDF) [15] and weight fair queue (WFQ) [16] algorithms. The hierarchical structure of the bandwidth allocation is shown in the following figure.

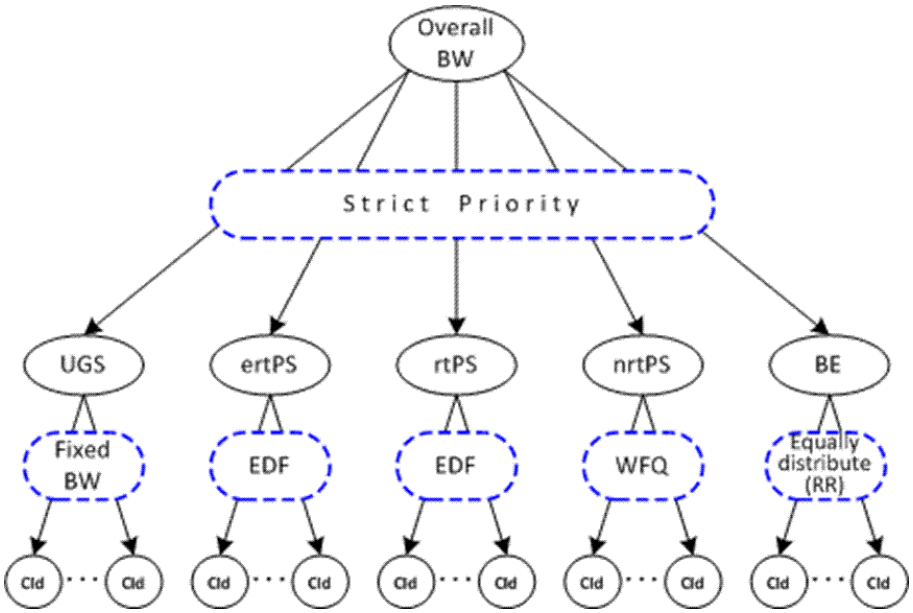


Figure 4.5: Representation of the scheduler's operation

The basic scheduling principles of the algorithms are as follows:

1. Overall bandwidth allocation: The bandwidth allocation per traffic class follows strict priority, from highest to lowest: UGS, ertPS, rtPS, nrtPS and BE. One disadvantage of the strict priority service is that higher priority connections may lead

lower priority connections to bandwidth starvation. To overcome this problem, a traffic policing module is included in each terminal, which forces the connection's bandwidth demands to stay within the traffic contract, as agreed during connection setup. This prevents the higher priority connections from using bandwidth more than their allocation, and allows for fair treatment of all traffic.

2. Bandwidth allocation within UGS connections: The scheduler allocates fixed bandwidth to UGS connections based on their fixed bandwidth requirements. This policy is determined clearly by the IEEE 802.16 standard, without the need for real-time transmission requests.

3. Bandwidth allocation within ertPS and rtPS connections: The earliest deadline first (EDF) service is adopted for these connections, to allow packets with the earliest deadline to be scheduled first. In case two packets belonging to two different service types (one of ertPS and one of rtPS) expiry at exactly the same time, the scheduler will give priority to the ertPS packet, considering this packet of higher priority. Bandwidth needs are constantly updated through real-time transmission requests.

4. Bandwidth allocation within nrtPS connections: The weighted fair queue (WFQ) service is applied for this traffic class. For each nrtPS connection, the ratio of its average data to the total nrtPS average data rates is computed, and resources being left from the higher priority classes (UGS,ertPS and rtPS) are distributed according to the computed weights of the connections. No transmission requests are required on this case.

5. Bandwidth allocation within BE connections: The remaining bandwidth is equally allocated to each BE connection following the Round Robin model, without transmission requests.

The scheduler described above combines both simplicity and efficiency, as can be easily implemented without the need for complex calculations, while it can provide service differentiation and QoS guarantees to all traffic classes. Simplicity was a critical requirement in our case, as the algorithm has to operate in real-time on a frame-by-frame basis. Nevertheless, it is expected that this will not sacrifice the algorithm's capability of operate under different traffic and channel conditions. Moreover, the scheduler can take advantage of the improved performance of FBMC compared to CP-OFDM, by fairly supporting a larger number of connections.

4.3 Joint resource allocation and scheduling

4.3.1 Scheduling and Resource allocation

The Radio Resource Management (RRM) is performed at the BS by the Radio Resource Agent (RRA), and by the Radio Resource Controller (RRC) which can be implemented

apart from the Base Station (BS). The tasks performed by both the RRA and the RRC include the channel estimation, the channel quality indicators (CQI) management, and the control of the radio resources assigned to the BS. Only the Medium Access Controller (MAC) layer and the Physical (PHY) layer are defined within the IEEE 802.16 standard [2]. This work will be carried out at the MAC layer blocks which perform the resource allocation (RA), the Packet Data Unit (PDU) management, the fragmentation, and the burst mapping. Therefore, as it is depicted in Figure 4.6 all the blocks within the dotted lines (the yellow box) are affected by the current work. On the other hand, the Air Link Control (ALC) is in charge of recollecting the MS's channel state information which is later used by the scheduling and RA processes, as well as other procedures such as the power control or the ranging among others.

Each incoming data stream is classified according to its class of service and mapped to a single Service Flow (SF). Without loss of generality, in this work it is considered that each Mobile Station (MS) has only one active SF. The packets from each SF are then independently buffered and each incoming packet is time stamped. Four service classes are defined in [2]: the *Unsolicited Grant Service* (UGS), the *Real-Time Polling Service* (rtPS), the *Non-real Time Polling Service* (nrtPS), and the *Best Effort* (BE) service.

As it is depicted in Figure 4.6, the data from the input buffers is monitored by the *Scheduling and RA* block. During each frame all the input packets are evaluated for transmission, and according to the channel state from each user and the scheduling policy some of the packets are scheduled (and maybe fragmented) to be transmitted in the subsequent frame. The scheduling process is strictly connected to the resource allocation process since the latter is who determines how many resources are assigned to each SF in every frame.

4.3.1.1 Resource allocation and MCS selection problem formulation

The main goal is to maximize the system throughput (i.e. the spectral efficiency) while guaranteeing the QoS constraints for each SF. Most of these constraints are defined by: the average bit rate, the peak bit rate, the minimum bit rate, the maximum tolerated delay per packet (and jitter), and the average bit error rate (or packet error rate). Nevertheless, one key issue for any resource allocation scheme is to minimize the signalling that is required to inform the receivers about the frame structure.

The optimum shape and position of each burst (with its respective MCS) is explored while the QoS requirements are fulfilled for each user. To reduce the algorithm complexity, uniform power allocation across subcarriers is considered and each SF is allocated with a single burst per frame. According to these premises and considering that there are M active SFs, the resource allocation and the rate adaptation problem that guarantees the different QoS requirements while maximizing the spectral efficiency can be mathematically

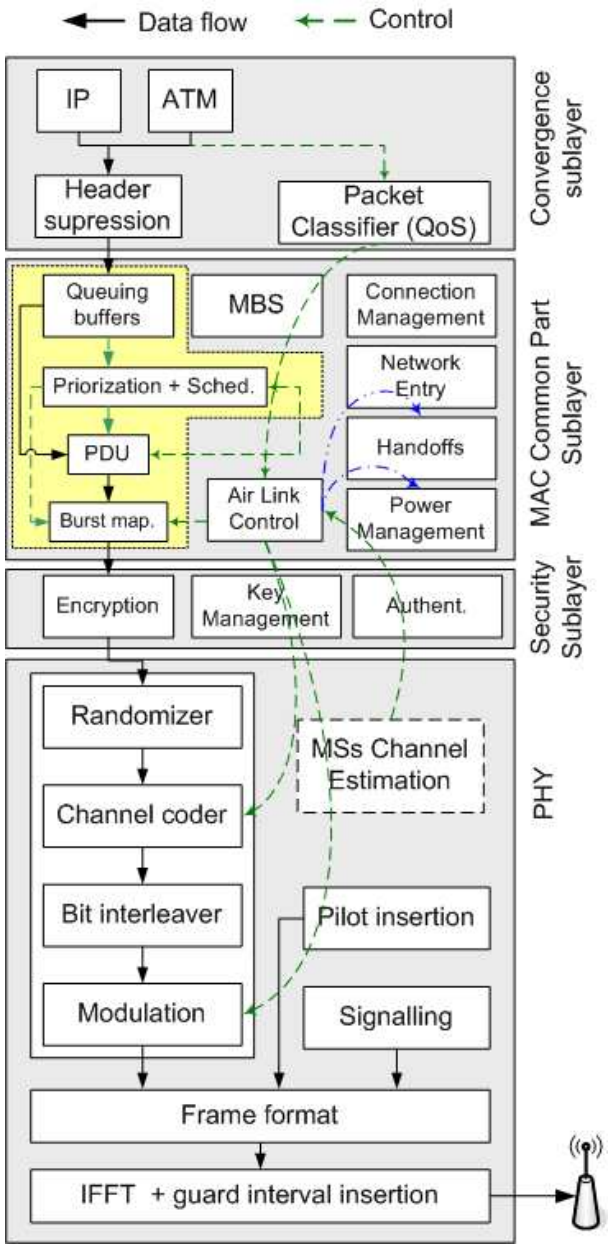


Figure 4.6: Protocol stack at the BS and Cross-layer interactions

expressed as

$$\underset{\xi}{\text{argmax}} \left\{ \sum_{i=1}^{M} \sum_{n=1}^{Q} \sum_{k=1}^{T} \eta_i(n,k) - M I_{\text{CC}} \right\} \tag{4.21}$$

$$\text{s.t. } b_i = T_{\text{frame}} \sum_{p=1}^{P_i} \frac{L_{i,p}}{(\tau_{max,i} - \tau_{i,p})}, \tag{4.22}$$

with

$$\xi_i(n,k)\xi_j(n,k) = 0, \text{ for } i \neq j \text{ and } n \in [0, Q-1], k \in [0, T-1], \tag{4.23}$$

$$\eta_i|_{\text{BER} \leq \mu} = \psi(\text{SNR}_{\text{eff},i}), \tag{4.24}$$

$$R_i = \sum_{n=1}^{Q} \sum_{k=1}^{T} \eta_i \xi_i(n,k) \geq b_i. \tag{4.25}$$

In (4.21) the term I_{CC} means the number of the required signaling bits transmitted within the control channel for each burst. The minimum required bits per frame b_i for the i-th SF is obtained in (4.22), where $L_{i,p}$ is the p-th packet size in bits from the i-th SF, $\tau_{i,p}$ is the packet delay (time the packet has been queued in the buffer), $\tau_{\text{max},i}$ is the maximum allowed delay per packet for the i-th SF, and P_i the total number of the queued packets. ξ_i is a binary $(Q \times T)$ matrix which points out which MRUs are allocated for the i-th SF (i.e. $\xi_i(n,k)$=1 means the (n,k) MRU has been assigned to the i-th SF). Equation (4.23) guarantees that the different bursts do not overlap (as it is later depicted in Figure 4.7). Finally, (4.24) and (4.25) determine the current number of bits transmitted within the i-th burst R_i. The term η_i represents the upper layer throughput (in bits) per MRU, and it is obtained as a function of the calculated effective SNR (SNR_{eff}) per each burst, the available MCS and the upper bound BER.

4.3.1.2 Proposed Joint Packet scheduling and Resource Allocation

The solution of (4.21) up to (4.25) could be obtained using non-linear programming techniques. However, such techniques are not feasible for practical systems due to the very high computational complexity. Furthermore, the problem as defined above is very rigid since it forces the number of bursts to be equal to the number of service flows, and in consequence all the service flows are scheduled during each frame. However, the number of bursts per frame B should be adapted according to the different channel conditions (a MS may experience deep fading during certain frames). In addition, using a unique burst per user may decrease the spectral efficiency when the burst spans over a large bandwidth due to the effect of frequency selective fading. To overcome these limitations, a low complexity iterative algorithm is proposed which adapts the number of bursts for user scheduling and resource allocation purposes. In order to maximize the spectral efficiency and undertaking the service flows QoS requirements, the resource allocation and the rate adaptation problem described in this subsection is divided into two stages:

i) the minimum requirements fulfillment,

ii) and the spectral efficiency maximization.

For each stage a different prioritization function is applied.

Service Flows prioritization:

In order to select which resources will be assigned to each SF (and thus to each MS), each i-th service has a priority value assigned over each n-th subchannel (we assume a constant channel for each subcarrier during the whole frame). The number of subcarriers per subchannel is subject to the shape of the minimum resource unit. Taking as example the well-known Proportional Fair Scheduling function [18], the priority $\varphi_i(n)$ assigned to each i-th SF in each n-th subchannel is given by

$$\varphi_i(n)|_{\text{PFS}} = \begin{cases} \dfrac{1}{\overline{Th_i}(t)} \dfrac{\eta_i(n)}{\eta_{\max}}, & \text{if } \sum_{p=1}^{P} L_{i,p} > 0 \\ 0, & \text{otherwise,} \end{cases} \tag{4.26}$$

where $\eta_i(n)$ is the spectral efficiency achieved by the highest MCS that can be applied on the n-th subchannel given an instantaneous BER lower than a certain upper bound $\text{BER}_{\max}$. Thus, $\eta_i(n)$=0 denotes a deep fading in the n-th subchannel for the i-th MS, and clearly in this case the priority becomes zero. $\eta_{\max}$ is the spectral efficiency achieved by the highest MCS. $\overline{Th_i}(t)$ is the average throughput obtained by a moving average window with α as the latency scale and $Th_i(t)$ the instantaneous throughput, thus

$$\overline{Th_i}(t) = \frac{1}{\alpha} Th_i(t) + \left(1 - \frac{1}{\alpha}\right) \overline{Th_i}(t-1), \quad \text{with } Th_i(t) \geq 0. \tag{4.27}$$

On the other hand, fairness might be also achieved by means of adhoc user satisfaction indicators as proposed in [19, 20, 21]. However, most of these algorithms have been designed based on the average bit rate requirements, without considering the buffer's state nor the variable bit rate (VBR) nature of the traffic. To overcome these restrictions, a Time Stamped Packets Scheduling (TSPS) function, proposed in [22], is applied. This function uses the input buffers status, the time stamp from each packet, and the channel metrics in order to determine the priority for each service flow. The allocation of the resources is divided into two phases:

i) QoS fulfillment (*phase* 1),

ii) and the spectral efficiency optimization (*phase* 2).

In the first stage a scheme similar to the PFS is proposed where a priority level is assigned to each service flow according to the amount of instantaneous required resources (instead of the historical average data rate used in the PFS), and its instantaneous CQM. The number of the instantaneous required resources is obtained assuming that each packet p at the input buffers of the q-th service flow is time stamped.

Thus, in order to send every packet before its maximum delay, we propose that each

service flow should be allocated in the next frame with b_q bits such that

$$b_q = \begin{cases} T_{\text{frame}} \sum_{p=1}^{P} \dfrac{L_{i,p}}{\tau_{\max,q} - \Delta\tau - \tau_{p,q}}, & \text{if } \forall p' \rightarrow \tau_{p',q} < (\tau_{\max,q} - \Delta\tau) \\ T_{\text{frame}} \sum\limits_{\substack{p=1 \\ p \neq p'}}^{P} \dfrac{L_{i,p}}{\tau_{\max,q} - \Delta\tau - \tau_{p,q}} + \sum\limits_{p'} L_{p',q}, & \text{otherwise} \end{cases}, \quad (4.28)$$

where $L_{p,q}$ are the bits that still have to be sent from the p-th packet of the q-th service flow, and $\tau_{p,q}$ are the seconds that such a packet has been already in the queue. In case where a packet (referred as p') exceeding its maximum delay all its packets will be transmitted within b_q bits. The priority assigned to each service flow is then a function of the instantaneous required resources normalized by the resources required by all the active service flows, and the achievable spectral efficiency R_q expressed in bits/sec/Hz normalized by $R_{\max}$ that represent the maximum spectral efficiency that can be achieved according to the set of available MCS. The priority for each q-th service flow in each n-th subchannel is

$$\varphi_{\text{PRE},q}(n) = \frac{R_q(n)}{R_{\max}} \times \frac{b_q}{\sum_{u=1}^{M} b_u}. \quad (4.29)$$

Note that we assume that the effective channel from each subchannel is practically constant during the frame time, thus the time index is omitted.

Furthermore, in case all the service flows allocated their required b_q bits within the frame, the remaining resources are then allocated following a greedy algorithm, where the first required resources are assigned to the users with better CQMs. According to this, the priority assigned to each service flow for this second stage is

$$\varphi_{\text{POST},q}(n) = \frac{R_q(n)}{R_{\max}}, \quad (4.30)$$

which it is not affected by any QoS constraint.

Since the proposed prioritization scheme aims to obtain packet delays lower than $\tau_{\max}$, the offered data rate will be equal to the average data rate. However, there is no guarantee that all the packets get to the receiver before $\tau_{\max}$. Then, for real time applications where the packet cannot exceed its maximum delay (otherwise the packet could be dropped from the queue), the priority level is further modified in order to avoid excessive packet drops. We propose to apply an urgency factor as in [20], thus in case a packet is close to being discarded, the related service flow receives a higher priority forcing the packet to be transmitted within the following frame. Thus we reformulate (4.29) as,

$$\varphi_{\text{PRE},q}(n) = \begin{cases} P_{\text{urgency}} \dfrac{R_q(n)}{R_{\max}}, & \text{if } \tau_{p',q} \geq (\tau_{\max,q} - \Delta\tau) \\ \dfrac{R_q(n)}{R_{\max}} \times \dfrac{b_q}{\sum_{u=1}^{M} b_u}, & \text{otherwise,} \end{cases} \quad (4.31)$$

where the index p' means any value in the range $[1, P_q]$ such that P_q is the number of packets queued from any q-th service flow that belongs to the real time service class. P_{urgency} is a fixed constant such that $P_{\text{urgency}} > (R_{\min}/R_{\max})^{-1}$ ($R_{\min}$ is the minimum spectral efficiency achieved by the more robust MCS), thus the urgent packets are scheduled prior to other packets despite the CQM. Moreover, in case where a service flow received a higher priority to flush the urgent packets, b_q is updated such that the required resources are equal to the size of the urgent packets. Finally, in case where the channel is in a deep fade (i.e. $R_q(n)$=0), the priority assigned to the service flow is then equal to zero despite of the input buffer status.

4.3.2 Burst Allocation in WiMAX and IEEE 802.16e

4.3.2.1 Introduction

In multicarrier based schemes, the subcarriers are grouped into larger units referred to as subchannels. Then, these subchannels are grouped across the frequency and time domains into bursts, where each burst is mapped to one user (in unicast) or a group of users (in broadcast). The burst allocation and the Modulation and Coding Scheme (MCS) applied to each burst are adapted on a frame basis. This will allow the Base Station (BS) to dynamically adjust the bandwidth usage per user according to the users' service requirements.

One of the most popular scheduling policies, currently used in the 3G networks, is the Proportional Fair Scheduling (PFS) [23][24][18][25]. In each radio resource unit PFS assigns each user a priority that is proportional to the channel quality and inversely proportional to the offered data rate. However, the main drawback of PFS comes from the fact that it considers full buffers and Constant Bit Rate (CBR) streams. Clearly, multimedia networks have to deal with different traffic types, e.g. Variable Bit Rate (VBR) streams with very strict packet delay requirements. Recent trends in packet scheduling consider cross-layer implementations such as those proposed in [19, 20, 21]. Liu et al. in [19] and Soo et al. in [20] proposed a scheduling algorithm where a priority is assigned to each user according to its instantaneous channel and service status, the CSI is directly obtained from the average received Signal to Noise Ratio (SNR), and the service status is obtained from the delay of the head-of-line packet. The same principle is extended to an OFDMA system in [21].

However, none of the above mentioned algorithms considered the effects of the resource allocation with respect to the required signaling and its payload neither the need of rectangular shaped bursts. In this work we define one burst as a set of continuous Minimum Resource Units (MRU) (logical or physical) in both time and frequency domains following a rectangular shape containing data from one Service Flow (SF). Each SF is a unidirectional flow of packets with a particular set of QoS parameters [2]. The resource allocation approaches in [26] and [27] follow the "raster approach" but they have been conceived considering that the channel within each subchannel is uncorrelated among subcarriers (thus a subcarrier permutation algorithm is assumed), hence the number of MRUs required for each user can be determined a priori according to the average SNR independently of the

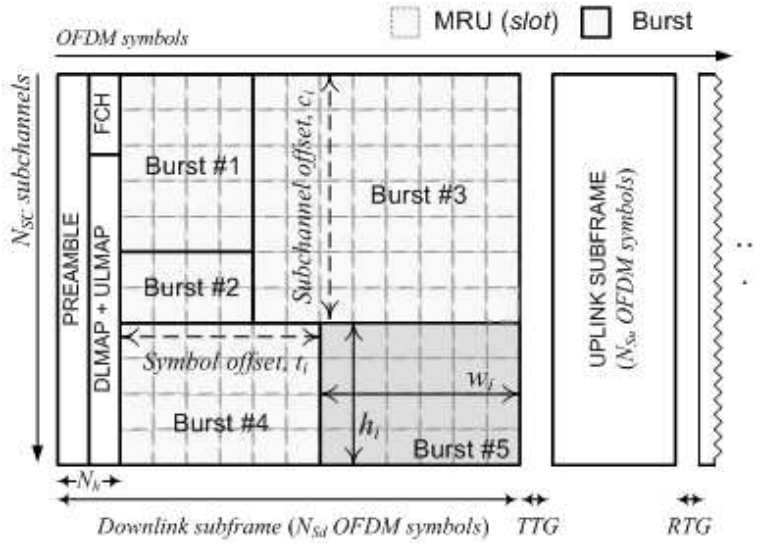


Figure 4.7: IEEE 802.16e OFDMA frame in TDD mode and burst structure

instantaneous channel. Though these proposals a good trade off between complexity and spectral efficiency may be achieved, as the gain from frequency scheduling is minimized since the channel effects have been averaged through all the bandwidth.

To overcome the above limitations, a new dynamic two-dimensional (frequency and time) resource allocation and scheduling scheme considering the rectangular burst shape is presented and hereafter analyzed. Moreover, the resource allocation algorithm is able to reduce the number of bursts per frame by allocating continuous MRUs, hence it reduces the required signaling per frame[1].

Furthermore, since the user's data are in almost all the cases packed together in the time and/or the frequency domain, the Mobile Stations (MS) power consumption is also reduced due of the reduced number of active symbols (shorter connections) or the reduced number of active sub-channels (lower computational cost at the receiver) [28].

4.3.2.2 System description

We consider one single cell with a total of K MSs within its cell area with no interference sources. The Time Division Duplexing (TDD) scheme is considered, thus channel reciprocity can be assumed between uplink and downlink. The whole TDD frame is formed by a total of N_{s} symbols with T_{frame} duration. The number of downlink and uplink symbols usually follows the ratio 2:1 or 3:1, however it can be adjusted by the BS according to users' demand [2].The whole transmission bandwidth BW is formed by a total of N_{c} subcarriers where only N_{used} are active.

[1]Each burst is signaled at least by its position in the frame (starting subcarrier and symbol, c_i and t_i in Figure 4.7), the number of allocated MRUs in frequency and time (h_i and w_i), the MCS and (optionally) the associated service flow or connection identifier (SFID/ CID) [19]. Table 4.3 resumes the fields that are transmitted for each burst.

In the FUSC and PUSC, the subcarriers assigned to each subchannel are distant in frequency, whereas for the Band AMC the subcarriers from one subchannel are adjacent. Note that the FUSC and PUSC increase the frequency diversity and average the interference, whereas the Band AMC mapping mode is more convenient for loading and beamforming where multiuser diversity is increased [21].

As it is depicted in Figure 4.7, the MRUs allocated to any data stream within a frame have a two dimensional shape constructed by at least one subchannel and one OFDM symbol (i.e. 2 FBMC symbols). In IEEE 802.16 standard the specific size of the MRU varies according to the permutation scheme, concretely for the Band AMC it may take the shapes 9×6, 18×3 or 27×2 (subcarriers $\times$ time symbols, respectively), where 1/9 of the subcarriers are dedicated to pilots. We define a MRU as a resource unit formed by a set of $N_{\mathrm{sc}} \times N_{\mathrm{st}}$ symbols in frequency and time domains respectively. Once the size of the MRUs is defined, we can obtain the total number of MRUs per frame $Q \times T$, where $Q = N_{\mathrm{c}}/N_{\mathrm{sc}}$ is the number of subchannels and $T = N_{\mathrm{s}}/N_{\mathrm{st}}$ defines the number of the time slots.

Field	**Size in bits**
Number of CIDs, J	8
CIDs (optional)	$J \times 16$
MCS	4
OFDMA symbol offset, t_i	8
Subchannel offset, c_i	6
Number of OFDMA symbols, w_i	7
Number of subchannels, h_i	6
Boosting	3

Table 4.3: Signaling data per burst used in the DL-MAP

Several MRUs can be grouped into a *data region* or *burst* (see Figure 4.7), formed by successive MRUs in both frequency and time directions. Both the MRU and the data region always follow a rectangular shape structure. Since the MS receiver needs to know how the downlink frame is organized in order to properly decode the data, the downlink control channel includes the number of bursts transmitted as well as the signaling for each burst.

The burst is signaled by the parameters indicated in Table 4.3 [2], the multicast transmission is addressed by mapping different Connection Identifiers (CID) to each burst, where the BS is responsible for issuing the Service Flow Identifiers (SFID) and mapping it over a single CID.

4.3.2.3 Radio Resource Management

Once the resources per SF have been resolved, the Packet Data Unit (PDU) block represented in Figure 4.6 prepares the data that will be mapped into each burst at the PHY layer. The *Burst Mapping* block breaks the packet data units in order to map each fragment into one physical burst. Each physical burst may apply a different MCS. The MCS

for each burst is obtained according to the effective SNR (SNR_{eff}) of the channel over the MRUs assigned to the burst. For low mobility scenarios, we can consider the channel for each subcarrier quasi-constant during the whole frame, thus the SNR_{eff} is an arbitrary function of the different post-processing SNR per subcarrier (SNR_i) such that

$$\text{SNR}_{\text{eff}} = f(\text{SNR}_1, \text{SNR}_2, ..., \text{SNR}_n, \text{MCS}), \tag{4.32}$$

where SNR_{eff} would be the SNR that in case of an additive white Gaussian noise (AWGN) channel it would achieve the same Bit Error Rate (BER). Several metrics as the; Exponentially Effective SNR (EESM), the Mean Instantaneous Capacity (MIC), or others based on the Mutual Information per Bit can be applied to obtain the SNR_{eff} [29][30]. In the presented work the harmonic mean of the channel values has been used (for more detailed information see [31]), which gives a tight lower bound of the BER and it is independent of the MCS.

4.3.2.4 Iterative Resource Allocation algorithm

In order to allocate the resource while the QoS from each SF is guaranteed, the resources are allocated following the TSPS function developed in section 4.3.1.2. Then once the priority for each SF over each subchannel $\varphi_i(n)$ and the minimum bits per frame b_i have been obtained, the MRUs are allocated iteratively in order to guarantee the QoS of all SFs (their minimum required bits per frame). Moreover, the resources are allocated following the rectangular burst shape thus the signaling required is minimized. Since each i-th burst must follow a rectangular shape, and considering the burst starts at n_i and k_i with h_i and w_i the number of the MRUs in frequency and time respectively, ξ_i in (4.21) is given by

$$\xi_i(n,k) = \begin{cases} 1, & \text{if}(n_i \leq n \leq n_i + h_i - 1) \text{ and } (k_i \leq k \leq k_i + w_i - 1), \\ 0, & \text{others.} \end{cases} \tag{4.33}$$

Although heuristic, an efficient solution to the resource allocation and scheduling problem is proposed based on iterative burst increments. As it can be observed in Figure 4.8, each burst may be increased towards four directions, i.e. top, bottom, left and right w.r.t. its position in the frame. In order to determine in which direction the increase is more advantageous or suitable, an equivalent priority D_x ($x \in \{\text{top, bottom, left, right }\}$) is assigned to each direction as indicated in Figure 4.8, where D_x is obtained by averaging the priority values $\varphi_i(n)$ of the MRU that are covered by the enlarged burst. Whether in the x direction there is any occupied MRU or the burst is at the frame boundary then D_x is forced to "0". An example of the iterative burst enlargement principle is proposed and depicted in Figure 4.8 where the number inside each rectangle indicates the order in which the resources have been allocated to each burst.

In this example three bursts have been created after 15 iterations, where the number indicated inside each MRU indicates the order in which the MRUs have been allocated. On each iteration one MRU is assigned to one service flow according to the priorities previously computed and updated after each iteration.

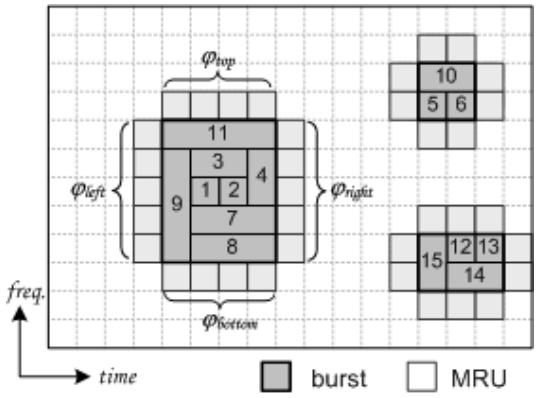


Figure 4.8: Contiguous MRU allocation scheme for localized bursts.

Two cases are considered for every iteration:

i) a new burst might be created and,

ii) a burst might be increased by allocating a MRU (or a group of MRU) to the burst.

While increasing the number of bursts per frame also increases the signaling, when one MRU is allocated to an existing burst no extra signaling is required.

The flowchart of the proposed iterative resource allocation algorithm is depicted in Figure 4.9. The algorithm starts without any allocated burst (B=0). For the first burst, the (n, k)-th MRU is allocated according to the i-th service flow and the n-th subchannel combination that maximizes the value of $\varphi_i(n)$. The position on the time axis of the MRU allocated to the first burst is forced such that k=0. Once the first burst is created, the iterative process starts checking the possible increments of the already existing bursts while at the same time it tries to generate new bursts.

Iteratively, the option with highest priority has a new is allocated MRU (in case of creating a new burst) or a group of MRUs (in case of enlarging an existing burst). Furthermore, the equivalent priorities associated to each burst increment are multiplied by a P_{burst} factor (e.g. P_{burst}=10) to push forward the enlargement of the existing bursts instead of generating new bursts.

In case a new burst is created, since $\varphi_i(n)$ is time independent, its position over the time axis is determined by that time slot whose distance to other already allocated MRU is maximized. This fact assures that the new created burst has higher chances to be increased than if it is placed near to the other already allocated bursts. The algorithm is iterated until all the requirements are fulfilled or when all the resources have been allocated. The number of bursts is then variable and may change from frame to frame depending on the buffers state, the QoS requirements and the channel state conditions. Moreover, since each SF may have more than one burst, another auxiliary matrix θ with size $Q \times T$ is defined. Each value of θ indicates to which burst the MRU is allocated. Both matrices ξ and θ are updated each time a new MRU is allocated.

Considering the MCS applied in each burst, we can obtain how many bits from each buffer are going to be transmitted and thus checking if the minimum requirements are meet.

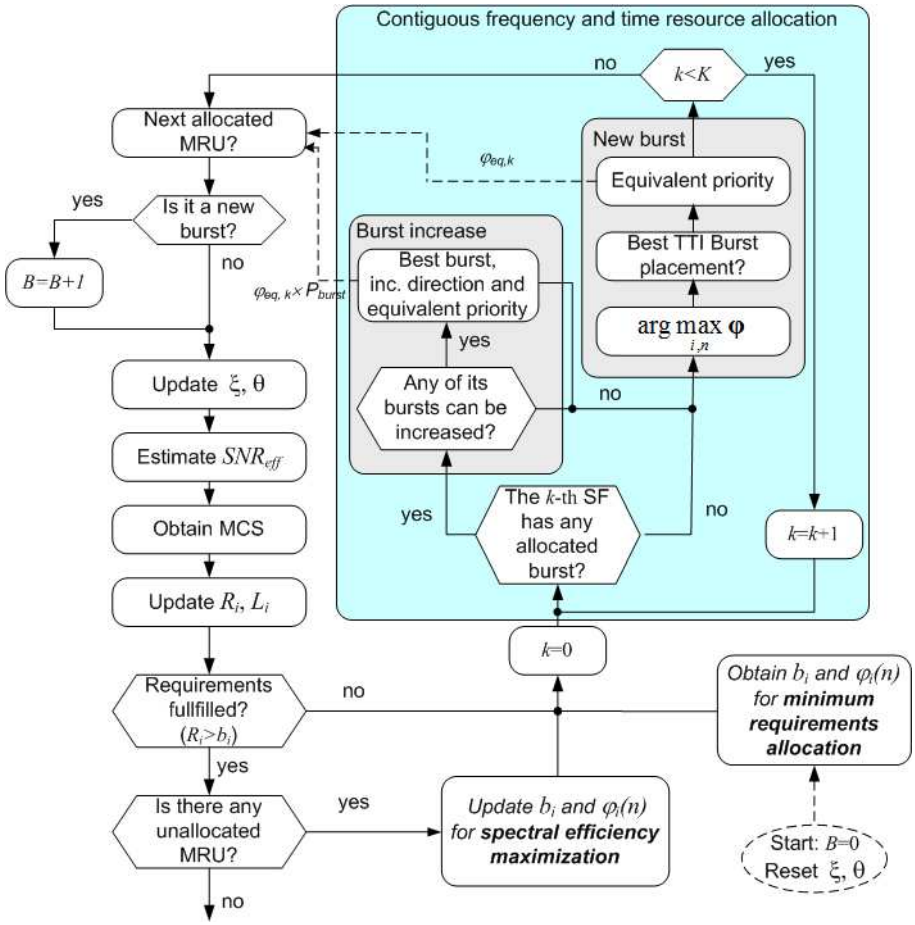


Figure 4.9: Resource Allocation and Scheduling algorithm flowchart

When the minimum requirements are satisfied (i.e. $R_i \geq b_i$ for $i = \{1..K\}$), in case there are still unassigned MRUs these resources should be used to flush the input buffers. Since the minimum requirements for the SF have been already allocated, the spectral efficiency can be maximized by transmitting the data from those SFs associated to the best channel conditions.

Finally, the end of the joint RRA and packet scheduling process may be achieved due to two main indicators: i) all the MRU have been allocated, or ii) the input buffers have been emptied. The packet delay is considered as the time from the packet is queued at the buffer until the instant where all the bits from the packet have been transmitted.

Class of Service	**rtPS**	**nrtPS**	**UGS**	**BE**	
Application	Video call	Streaming	Voice	HTTP	FTP
Average bit rate	380Kbps	2Mbps	15Kbps	N/A	N/A
Peak bit rate	2Mbps	10Mbps	15Kbps	2Mbps	10Mbps
Packet rate	10 packets/s	10 packets/s	10 packets/s	Variable	Variable
Max. delay	50ms	300ms	75ms	N/A	N/A
Max. BER	10^{-4}	10^{-4}	10^{-4}	10^{-6}	10^{-6}
Packet dropping ratio	1%	0-1%	1%	0	0

Table 4.4: Example of QoS requirements for applications belonging to each Service Class

In addition to the PFS and in order to validate and to check the performance of the proposed TSPS scheduling scheme, a modified version of the PFS named Buffer Based PFS (b^2PFS) is also introduced, where, instead of balancing the throughput of the different users the scheduler equalizes the number of buffered bits from each user, therefore the VBR streams can be managed. The prioritization function of the b^2PFS scheduler is

$$\varphi_i(n)\big|_{\mathrm{b^2PFS}} = \begin{cases} \dfrac{\overline{L_i}(t)}{\sum_i \overline{L_i}(t)} \dfrac{\eta_i(n)}{\eta_{\max}}, & \text{if } \sum_{p=1}^{P} L_{i,p} > 0 \\ 0, & \text{otherwise,} \end{cases} \tag{4.34}$$

where

$$\overline{L_i}(t) = \frac{1}{\alpha} L_i(t) + \left(1 - \frac{1}{\alpha}\right) \overline{L_i}(t-1), \text{ with } L_i(t) = \sum_p L_{i,p}. \tag{4.35}$$

The performance of the proposed TSPS prioritization function is evaluated and compared to the PFS and the b^2PFS. For the PFS and b^2PFS scheduling functions, the number of bits per frame b_i that should be transmitted is assumed equal to the number of buffered bits ($b_i = L_i(t)$). The latency scale is fixed to 10 frames (i.e. $\alpha = 10$).

4.3.2.5 Conclusions from preliminary results

The proposed resource allocation and data scheduling algorithm for multicarrier systems such as OFDMA and FBMC might be applied for Band AMC subcarrier permutation and

CP-OFDMA Air Interface and System Level parameters	
Carrier Freq	3.5GHz
Bandwidth	20MHz
Sampling Frequency	22.857Msps
Subcarrier Permutation	Band AMC
CP	12.5 %
FFT length	2048
# of used subcarriers	1728
# of subcarriers per MRU	18
# of symbols per MRU	3
# of data symbols per MRU	48 (efficiency = 8/9)
Modulation	M-QAM, M={4, 16, 64}
Channel coding	Punctured Convolutional
Bit Error Rate (BER)	$< 10^{-6}$
Channel Model	Pedestrian B
MS velocity	10 Km/h
Channel estimation and feedback	Ideal
Shadowing standard deviation	5 dB
BS Tx power	49dBm
BS antenna gain and pattern	14 dB (sectorial antenna), 70º
MS antenna gain and pattern	0dB, Omnidirectional
Other Link budget parameters	BS height = 30m MS height =1.5m, MS Noise Figure = 7dB, Connectors Loss = 2dB
Path loss, urban environment	$139.57 + 28 * \log 10(R)$, R = distance BS to MS in Km.
Thermal noise	-174dBm/Hz
# of sectors simulated	1
Frame duration, T_{frame}	5ms
DL/UL rate	2:1
# of symbols in the DL subframe	30

Table 4.5: System parameters

the PUSC subcarrier permutation indistinctly. Moreover, the proposed TSPS scheme has shown in preliminary results which will be extended and validated in deliverable 6.3 its ability to handle sensitive delay applications (i.e. rtPS and nrtPS), while high spectral efficiency is reached by exploiting the multiuser diversity on those unallocated resources. On the other hand, the proposed algorithm, which is based on iterative burst increments, aims to decrease the number of required bursts per frame, which leads to an important reduction of the overload signaling as well as a reduction on the receiver complexity and power consumption.

Another advantage of this RA scheme is its lower computational complexity compared to the case where each MRU is independently evaluated. Actually, since in many cases several MRUs might be allocated in a single iteration, the number of required iterations is much lower compared to the case where one MRU is allocated per iteration.

4.3.3 Mixed TUSC and Band AMC subcarrier permutation zone proposal for efficient Resource Allocation and Scheduling in multicarrier systems with Limited Feedback

4.3.3.1 Introduction

In the Orthogonal Frequency Division Multiple Access (OFDMA) scheme as well as in the Filter Bank Multicarrier (FBMC) scheme, the whole bandwidth is first divided into subcarriers, which are later grouped into subchannels. The subcarriers assigned to each subchannel may belong to distant regions of the spectrum (hence frequency diversity is enhanced), or may be adjacent subcarriers from a specific part of the spectrum (thus the channel is highly correlated between subcarriers within the subchannel). Both schemes have been defined in the CP-OFDMA interface of the IEEE 802.16 standard under the names of Partial Usage of the Subchannels (PUSC) and Band Adaptive Modulation and Coding (Band AMC) respectively [2][32]. The Band AMC scheme is preferred in case Channel State Information (CSI) is available at the Transmitter (CSIT). In such case the optimum power, Modulation and Coding Scheme (MCS) might be selected for each allocated resource in order to maximize the throughput (named Rate Adaptation - RA) or to minimize the transmission power (named Margin Adaptation - MA) [33]. Furthermore, when in a multiuser environment the Base Station (BS) knows the CSI from each user in each subchannel, the frequency scheduling can be combined with multiuser diversity bringing out large system gains whether the Opportunistic Scheduling (OS) is applied in the downlink [34][35]. Channel reciprocity has been assumed by many researches [30] in case of Time Division Duplexing (TDD) to obtain the CSI at the BS, however, this implies that every Mobile Station (MS) must transmit pilots over all the bandwidth during the uplink interval or report the CSI measured by the MS. As a result, the Band AMC scheme implies a large signaling load in the uplink to know the CSI from all the users, and in the downlink to indicate how the resources have been allocated to the users (i.e. the Downlink-MAP in [2]). In order to reduce such signaling two common techniques are employed: i) the use of limited-rate feedback techniques to minimize the uplink feedback requirements,

and ii) two allocate contiguous sets of Minimum Resource Units (MRU) leading to a unique Data Region (also referred as burst) which also preserves the rectangular shape.

In order to reduce the feedback requirements, Han et al. proposed in [36] to send the channel quality metrics (CQM) of the n-best subchannels, as well as sending the CQMs of n-listed subchannels which are requested by the BS. Moreover, selective feedback where only the CQM of those subchannels above a threshold are reported, is also a quite spread technique [37, 38, 39, 40, 41]. In this case, only the index of the subchannel needs to be sent in the feedback channel. In case of opportunistic feedback, a trade-off between dedicated feedback resources and number of collisions arises, thus, the system must be tuned in order to achieve the best performances. Other researchers have focused on determining the CQMs that best model the channel effects and how much they can be compressed [42]. Most of the CQMs proposed are a modified version of the instantaneous Signal to Noise Ratio (SNR) measured at the receiver side, where in most the cases it can be transmitted over 2-5 bits according to the number of available MCS.

On the other hand, the problem of allocating the bursts following a rectangle shape into the frames has been previously studied in different contexts (i.e. floor planning, job scheduling, etc.) and referred as the two-dimensional bin packing problem. Hung et al. showed in [43] that the burst allocation structure is NP-complete. Several heuristic solutions have been proposed to solve it in real-time applications [26]. Similarly, Erta et al. proposed in [27] to proceed backwards in column-wise order, thus there is no need to book in advance resources for the transmission within the DL-MAP. However, the schemes proposed in [26, 27, 44] have assumed a total independence between the burst position over the frame (hence distributed OFDMA is considered where only the average SNR is needed).

In this subsection, a new scheme is proposed where both techniques (selective feedback and two-dimensional bin packing) can be combined efficiently in case when distributed and localized CP-OFDMA are considered jointly. Moreover, in order to further reduce the feedback requirements, we propose that the BS signals which MS have packets pending to be scheduled thus only these MS report the CQM of their n-best subchannels. Then, the resource allocation algorithm allocates first those subchannels where selective CQMs are available. For these subchannels, localized CP-OFDMA is considered and consequently the best MCS is selected considering power transmission uniformly distributed in the overall bandwidth. Furthermore, the assigned resources to each user are given following a new iterative allocation algorithm proposed leads to a reduction on the number of bursts assigned to each user. Afterwards, in case there are still subchannels to allocate, the remainder subchannels can be allocated following the already proposed raster approach in [27]. During this second allocation process only the average SNR value is necessary, which is feedback periodically from all the Mobile Stations at a very low rate (each 15 ms). In order to be able to signal the bursts that have been created during the allocation process, an overlapping signaling scheme similar to the one proposed in [45] is used, where the signaling order is considered to determine the remaining resources.

4.3.3.2 System model

The system studied is a Point to Multi-Point scheme as represented in Figure 4.10 with a single BS serving K MSs, and any interfering source is considered. The transmission bandwidth BW is formed by a total of N_c subcarriers, and the subcarrier spacing considered is much smaller than the channel coherence bandwidth, thus a frequency flat fading can be assumed. The whole set of subcarriers can be divided into data subcarriers, pilot subcarriers, and null subcarriers where the average transmitting power P_T is distributed uniformly over the N_c active subcarriers (non null sub-carriers). All the active subcarriers are then grouped into Q subchannels where each subchannel consists on a set of both pilots and data subcarriers. Therefore, the number of subcarriers assigned to each subchannel becomes $N_{sc} = N_c/Q$. Two subcarrier permutations are employed: the PUSC, and the Band AMC [2]. In the PUSC case, the subcarriers assigned to each subchannel are placed at distant positions over the transmission bandwidth. On the other hand, the Band AMC maps all the subcarriers within a subchannel at adjacent positions in the spectrum. The PUSC scheme is used to increase frequency diversity and averaging the interference. Whereas, having the CSI at the transmitter, the Band AMC scheme is used to obtain the MCS that best fits to that subchannel, hence increasing the spectral efficiency.

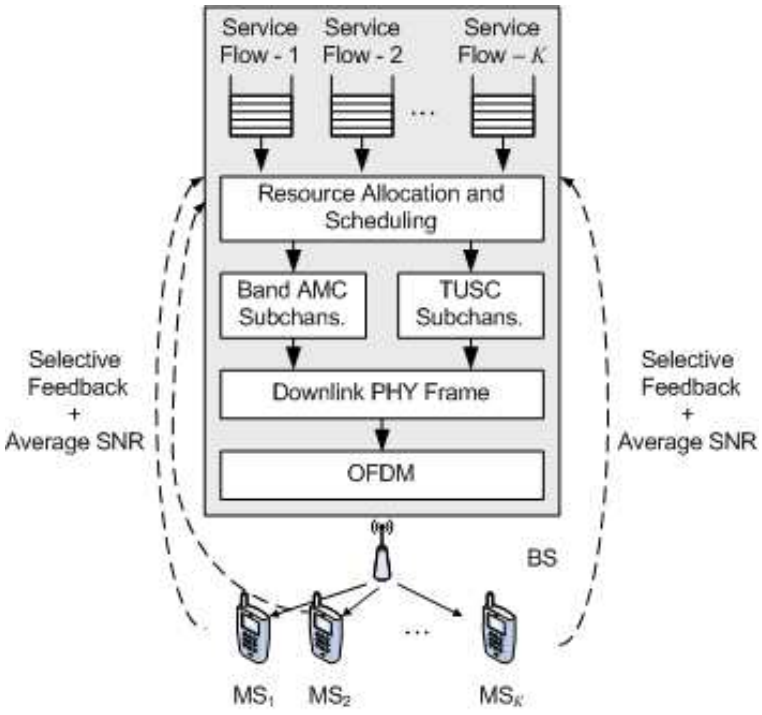


Figure 4.10: Dynamic resource allocation and scheduling schematic

One of the main tasks of the Medium Access Layer (MAC) at the BS is to allocate resources in the downlink and uplink to each MS. We consider that every MS has a single active connection with the BS, which implies a continuous flow of data (referred as service

flow) between the BS and the MS. Furthermore, every service flow also requires a feedback channel where the MS acknowledges the received packets as well as the CSI. Throughout this work we assume that the signaling and the feedback channels are error-free.

In order to reduce the amount of feedback signaling, two types of feedback are considered. First a periodic signaling sent by each MS at a fixed rate (e.g. each 15ms) where the average SNR is transmitted. And the second is a selective feedback sent by those MSs with data pending to be transmitted. To further reduce the signaling, those MS asked to report their CSI will simply report the coefficients of its n-best subchannels. As a result, the BS might take the selective feedback in order to select which user should be allocated in each band (i.e. Band AMC is applied), or could use the average SNR in combination with the PUSC structure.

4.3.3.3 Downlink Sub-Frame Structure and in band signaling

Since most of the actual deployed WiMAX based networks employing time division multiplexing (TDD), the description of the system is based on the TDD frame structure (see Figure 4.7). Each TDD frame lasts for T_{frame} seconds (e.g. $T_{\text{frame}} = 5$ms) where the part dedicated to downlink is formed by a total of N_{Sd} OFDM symbols (e.g. $N_{\text{Sd}} = 30$ for a 2:1 downlink to uplink ratio). The downlink and the uplink subframes are separated by the BS's turnaround times TTG (Transmit-Receive Transition Gap) and RTG (Receive-Transmit Transition Gap). As depicted in Figure 4.7, the first symbols of each frame are occupied by a preamble which is required for the MS synchronization and channel estimation. Following the standard preamble structure in [2], the Frame Control Header (FCH) is transmitted where the DL-MAP's length and the coding scheme are mapped. Straight afterwards, the DL-MAP and UL-MAP messages are transmitted within the downlink and uplink specified subframes structures.

The Minimum Resource Unit (MRU) that the BS is able to allocate to any MS, (also named slot) is defined as a two-dimensional structure formed by at least one subchannel and one OFDM symbol (i.e. 2 FBMC symbols). Without loss of generality, we define the MRU as a resource unit formed by a set of $N_{\text{sc}} \times N_{\text{st}}$ symbols in the frequency and the time domains respectively. Once the size of the MRUs is defined, we obtain the total number of MRUs that could be allocated into the downlink subframe $Q \times T$, having $T = (N_{\text{Sd}} - N_h)/N_{\text{st}}$, where N_h means the number of CP-OFDM symbols occupied by the preamble, the FCH, the DL-MAP and the UL-MAP. Furthermore, in order to reduce the DL-MAP length, several adjacent MRUs can be grouped into one Data Region or burst (note that the burst follows a rectangular shape too as depicted in Figure 4.7). The data conveyed into each DL-burst is then associated to the Connection Identifier (CID) which at the same time has a Service Flow Identifier (SFID) mapped to it. Each DL-burst is signaled within the DL-MAP by means of its frequency and time offset positions, the shape form (number of MRU in frequency and time), the MCS, and the CID among others values. Following the WiMAX standard, around 60 uncoded bits must be transmitted at least within the DL-MAP for every burst [2], hence the number of DL-bursts must be kept as low as possible in order to avoid the loss in spectral efficiency due to the required

signaling.

4.3.3.4 Mixed TUSC and Band AMC zone

According to the IEEE 802.16e, each downlink or uplink sub-frame may be divided into different permutation zones. Each permutation zone is characterized by its specific sub-carrier permutation scheme (e.g. PUSC, Band AMC, Full Usage of Subchannels - FUSC, Tile Usage of Subchannels - TUSC, etc.). The transition between zones is indicated in the DL-Map, whereas any burst can span over multiple zones.

As previously mentioned, the PUSC scheme as well as the TUSC scheme are the schemes that give higher frequency diversity, thus they are preferred in case the CSI information is not available at the transmitter. Furthermore, in case of fast moving MSs or unreliable channel estimations the PUSC is able to cope well with those channel uncertainties. On the other hand, for the full CSIT case, the Band AMC scheme is preferred since the best MCS can be applied to each subchannel thus maximizing the spectral efficiency. However, having the CSI of all the MS in each subchannel may become unpractical. As a result, we propose to combine both permutation schemes into a unique zone named Mixed TUSC and Band AMC (MTBA). As it is illustrated in Figure 4.11, the proposed MTBA zone is constituted by two types of bursts: the *localized bursts* and the *distributed bursts*. For the localized bursts, the MRU within the bursts are constructed with N_{sc} adjacent subcarriers during N_{Sd} symbols (e.g. 18×3, 9×6, 27×2, etc.) in addition the MRUs belonging to each burst are placed in adjacent frequencies and time positions. The channel experienced by each MRU is then obtained thanks to the pilot's subcarriers allocated inside each MRU.

Having the bursts allocated, the remaining subcarriers are divided into tiles. A "tile", is a small structure of adjacent subcarriers in frequency and time domains (e.g. 4×3 or 3×3) where the pilot's subcarriers are allocated. The number of data subcarriers in each tile is fixed and equal to N_{st} (e.g. N_{st}=8). The whole set of tiles are indexed logically thus consecutive tiles are placed at distant positions in frequency and time domain. Afterwards, the MRUs for the distributed bursts are obtained by taking a fixed number N_t of logically consecutive tiles (e.g. N_t=6). Actually, the TUSC permutation scheme in the IEEE-802.16e standard [2] applies the same resource division where the frame is partitioned into tiles, while the subcarriers and symbols belonging to each tile are adjacent in time and frequency. Note that, the degree of diversity obtained using the TUSC scheme is slighter lower than that with the PUSC scheme, since in the later the affected channel coefficients over the subcarriers within each MRU are all uncorrelated. But we have to note that the way on how the PUSC scheme manages the pilots, the data subcarriers and their mapping in each MRU is much more complex than in the TUSC, making much more difficult to allocate localized and distributed bursts within the same permutation zone.

We consider that the TUSC scheme offers a good compromise between diversity and computational complexity. Nevertheless, it is worth mentioning that despite the shape of the MRUs may be changed from the localized bursts to the distributed bursts, the amount of data symbols that can be conveyed into any MRU is fixed and equal to 48. This fact facilitates the resource allocation process since the number of bits that are mapped into

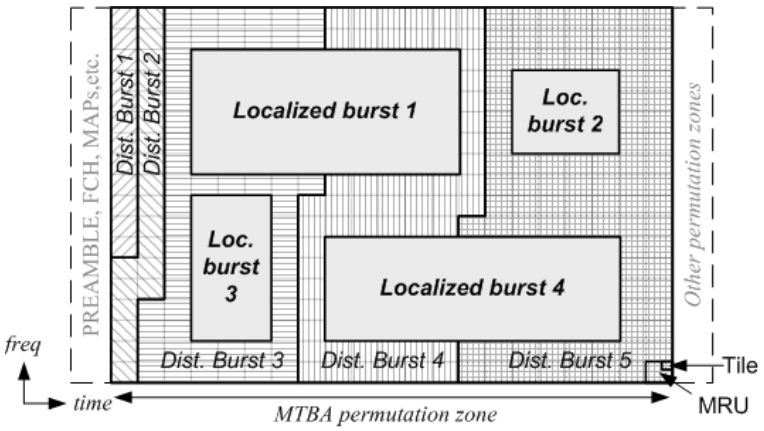


Figure 4.11: Proposed mixed TUSC and Band AMC (MTBA) zone

each MRU can be obtained directly as a function of the MCS applied.

Regarding the signaling associated to the proposed MTBA zone, the information that must be transmitted within the DL-MAP can be also reduced for the MTBA zone following a similar scheme to that in [45]. The signaling required for each localized burst remains unchanged (i.e. the position, shape, MCS, etc. are sent in the same way), however if we assume that for the distributed bursts the signaling associated to the localized burst is decoded before, only the length (in terms of number of MRU) of each distributed burst is required to find its position in the current frame (i.e. the subchannel and CP-OFDM symbol offset, as well as number of MRU in frequency and time are anymore required). Thus, if we assume that 10 bits are enough to indicate the burst length, around 30% of the signaling associated to each distributed burst can be reduced using this approach compared to the signaling proposed in [2]. However, since the information about the distributed burst is accumulative, in order to avoid propagation errors in the decoding of the DL-MAP, the length of the distributed burst should be coded with lower coding rates.

4.3.3.5 Adaptive resource allocation and scheduling with Partial CSI at the Transmitter

One of the main goals of the radio resource management as well as the scheduling is to determine/reach the optimum resource assignment such that (for instance) the total system throughput is maximized while every user's QoS requirements are fulfilled. On the other hand, the highest MCS scheme that can be applied in each MRU depends of the experienced CSI (i.e. the SNR) information, the permutation scheme, and the maximum tolerated Bit Error Rate (BER). In order to obtain the SNR thresholds where each MCS gives a BER lower than a certain value, a link level simulation has been carried out where the BER has been computed through several MRUs. The considered MCSs are M-QAM with M=$\{2,4,16,64\}$ and their corresponding convolutional code with a puncturing r=$\{1/2, 2/3, 3/4, 5/6\}$ coding rates set.

Two types of channel models have been considered. For the localized bursts we have

obtained the SNR thresholds assuming an Additive White Gaussian Noise (AWGN) channel. Since the channel experienced within a burst might be not constant within the burst, an effective channel coefficient $|H_{\text{eff}}|^2$ is obtained by computing the geometric mean of the channel coefficients. This is the more straightforward Effective SNR Mapping (ESM) technique. However, the validity of this approximation has been checked for different channel models where having a total of 16 adjacent MRUs in frequency domain very accurate results were observed (these are here omitted due to page restrictions). The Effective SNR (ESNR) for a MRU placed inside a localized burst is then given by

$$\text{ESNR} = \frac{P_{\text{r}}}{\text{BW}_{\text{active}} \times N_0} \left(\prod_{i=\Delta_{\text{sc}}}^{\Delta_{\text{sc}}+N_{\text{sc}}-1} \prod_{i=\Delta_{\text{st}}}^{\Delta_{\text{st}}+N_{\text{st}}-1} |H(i,j)|^2 \right)^{1/(N_{\text{sc}} \times N_{\text{st}})} \tag{4.36}$$

$$= \frac{P_{\text{r}}}{\text{BW}_{\text{active}} \times N_0} |H_{\text{eff}}|^2, \tag{4.37}$$

where P_{r} is the total received power, which includes the effects of free space channel attenuation and shadowing, and N_0 means the noise power spectral density. The indexes i, and j mean the subchannel and OFDM symbol index respectively, while Δ_{sc}, and Δ_{st} are the burst subchannel and OFDM symbol offsets. $\text{BW}_{\text{active}}$ is the bandwidth occupied by all the active subcarriers. The same principle can be applied to obtain the ESNR for the whole burst. Then, given the ESNR of each burst and having a look up table as the one given in Table 4.6, the BS can decide which is the optimum MCS that should be applied to every localized burst. These values have been obtained using a soft decoder.

On the other hand, for the distributed bursts, only the average SNR is used (as in others open-loop adaptation mechanisms). We assume that the channel experienced by the distributed bursts is an uncorrelated Rayleigh channel. Actually, this imply that the channel transfer function $H(n,k)$ can be modeled as a random variable following a zero mean complex Gaussian distribution with unitary standard deviation. This can be obtained by means of sufficiently large and deep interleaving blocks with subcarrier permutations. In that case, the optimum MCS is obtained considering that the ESNR is equal to the average SNR, and modifying the SNR thresholds to those depicted 4.6. Those values marked in italic with (*) super index should not be used since a MCS with a higher spectral efficiency can be used over the same SNR range. We can observe in 4.6 that those MCS where high puncturing rates are applied fail to obtain a low BER for a Rayleigh channel, which indicates that the diversity order is severely affected by the puncturing.

4.3.3.6 Resource allocation and scheduling algorithm for the proposed MTBA zone

As already introduced in previous subsections, the use of localized bursts instead of distributed bursts improves the spectral efficiency of the system since selective fadings can be often avoided. Consequently, localized bursts must be placed first and in case no more localized bursts can be allocated, the remainder resources are assigned to distributed bursts. Furthermore, we are considering that the BS only knows the n-best CQM of every user, as

Min. SNR_{eff}	**BER=10^{-4} AWGN**	**BER=10^{-4} Rayleigh**	**BER=10^{-6} AWGN**	**BER=10^{-6} Rayleigh**
2PSK,r=1/2	0.5	3.2	1.75	5.5
2PSK, r=2/3	2.2	6.6*	3.3	9.4*
2PSK, r=3/4	3.2	8.8*	4.5	12.2*
2PSK, r=5/6	4.1*	11.5*	5.4*	15.3*
4PSK,r=1/2	3.4	6.6	4.8	9.3
4PSK, r=2/3	5.2	10.1	6.4	13.3
4PSK, r=3/4	6.2	12.5*	7.6	16.8*
4PSK, r=5/6	7.2	15.4*	8.6	20.3*
16QAM,r=1/2	9.6	12.4	11.2	14.8
16QAM,r=2/3	11.5	15.8	13.2	18.6
16QAM,r=3/4	12.6	18.4*	14.1	21.7*
16QAM,r=5/6	13.6	21.5*	15.3	26.3*
64QAM,r=1/2	15.9*	18.5	17.5*	21.2
64QAM,r=2/3	17.7	21.7	19.2	24.8
64QAM,r=3/4	18.8	23.9	20.2	27.5
64QAM,r=5/6	20.1	26.5	21.8	30.7

Table 4.6: Minimum SNR for each MCS using a Soft Decoder

well as the average SNR of each user, thus the channel is only known in certain parts of the spectrum.

So, to allocate the localized bursts, the downlink frame is first segmented into MRUs of adjacent subcarriers and symbols, having $Q \times T$ available MRUs in frequency and time domain respectively. Next, the BS calculates the priorities associated to each user on those known subchannels according to the scheduling function proposed in Section 4.3.1.1. For the unknown subchannels a priority equal to zero is fixed. Then the MRUs are iteratively allocated according to users' priority $\varphi_{\mathrm{PRE},q}(n)$ until every service flow is allocated at least R_q bits or any other localized burst can be created (whatever happens first). After the allocation of the localized bursts, the users' priorities must be updated taking into account the new effective SNR thresholds. A schematic of the proposed algorithm is depicted in Figure 4.12.

Afterwards, when all $\varphi_{\mathrm{PRE},q}(n)$ (or $\varphi_{\mathrm{POST},q}(n)$ in case the minimum resources have been satisfied) become zero, it means that the localized bursts cannot be increased any more since the BS does not know the channel in the remaining MRUs. Those remaining resources, as described in Section 4.3.3.4, are divided into tiles, physically indexed and mapped over N_{dMRU} MRUs and the resources distributed across the entire downlink frame (both in time and frequency domains). Moreover, the N_{dMRU} MRUs are logically indexed such that the tiles of two consecutive MRUs also belong to different parts of the downlink frame. Then, the unallocated N_{dMRU} MRUs are iteratively assigned (irrespective of its index) to the different distributed bursts until all the MRUs have been assigned or all the

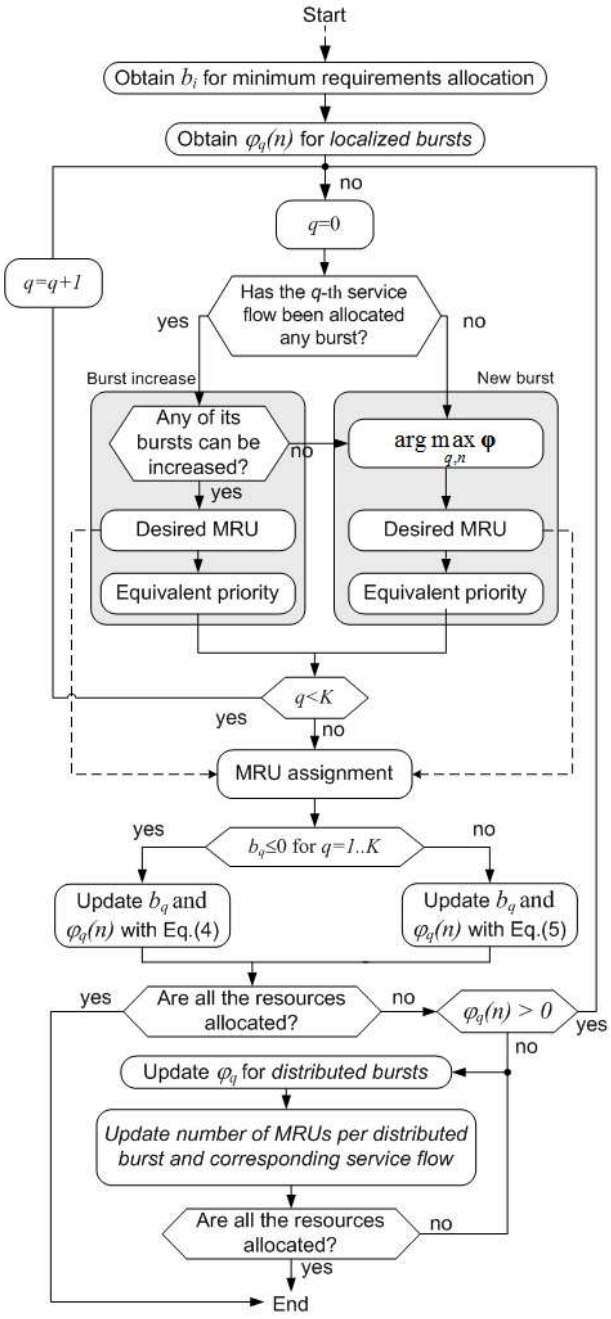


Figure 4.12: *Localized* and *distributed* burst allocation algorithm for the proposed MTBA zone

data in the buffers have been allocated.

The priorities of each service flow are updated every iteration considering the ESNR is equal to the average SNR and the channel is Rayleigh (hence the priorities don't depend on the MRU position). The maximum number of distributed bursts (according to the signaling scheme defined in section 4.3.3.4) is equal to the number of service flows, K, in case all the buffers are non-empty. At the end of this second allocation process, each service flow might be allocated one distributed burst with $N_{\mathrm{dMRU},q}$ consecutive (logically indexed) MRUs.

4.3.3.7 Conclusions from preliminary results

As in subsection 4.3.3, after preliminary simulations and obtained results which will be deeply analysed in deliverable D6.3, it is expected that with the combination of both the TUSC and AMC permutation types in the same MTBA zone, the system parameters as the spectral efficiency, the delay, and the signalling load, can be improved compared with the TUSC and Band AMC zones defined in IEEE 802.16. The use of localized bursts guarantees that unless the system is very saturated, every user will have at least one subchannel assigned. This brings up the idea of adapting the number of feedback subchannels according to user requirements or certain fixed priorities. This can be also exploited in scenarios where several users with different mobility have to be served. When a MTBA zone is used, the low mobility users can be assigned to the localized bursts, whereas the high mobility users can be assigned to the distributed bursts.

4.4 Conclusions

Both approaches described in 4.2 and 4.3 are aiming at the downlink and have to be extended to the uplink. In the uplink with asynchronous user signals arriving at the BS, the advantage gained by the high selectivity of the filter bank will be even more enhanced.

Further studies on this topic are necessary, among which are the following:

- coexistence of low and high mobility users;
- size and shape of bursts assigned per user within one subframe in the context of FBMC with pre- and posttails due to the prototype filter length;
- coordination of intercell interference especially for cell edge users.

Also the results from 3.2 and 3.3 have to be taken into account in Chapter 4.

Chapter 5

Summary

The main results of this report are briefly summarized:

- The increased spectral efficiency of FBMC in comparison to CP-OFDM is best exploited with FDD. TDD is possible, but then a great frame length is beneficial for FBMC.
- FBMC is more robust against synchronization errors especially in the uplink. Less multiple access interference is produced.
- Intercell interference can be better coordinated in an FBMC system and will lead to more efficient frequency reuse in a cellular environment.
- Scheduling and resource allocation can benefit from the greater robustness with respect to synchronization errors and simpler interference coordination of FBMC.

Bibliography

[1] WiMAX Forum Technical Working Group. WiMAX forum mobile system profile release 1.0 approved specification Rev. 1.7.0: 2008-09-18.

[2] IEEE Standard 802.16e 2005. IEEE standard for local and metropolitan area networks part 16: Air interface for fixed and mobile broadband wireless access systems Amendment 2: Physical and medium access control layers for combined fixed and mobile operation in licensed bands and Corrigendum 1.

[3] B. Bisla, R. Eline, and L. M. Franca-Neto. RF system and circuit challenges for WiMAX. *Intel Technology Journal*, 08(03):189–200, August 2004.

[4] Deliverable D2.1 PHYDYAS ICT-211887. Data-aided synchronization and initialization (single antenna). July 2008.

[5] M.G. Bellanger. Specification and design of a prototype filter for filter bank based multicarrier transmission. *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, vol. 4:2417–2420, 2001.

[6] M. Bellanger and M. Terré. Note on the TDD-FDD issue in FBMC. September 2008.

[7] Alan V. Oppenheim, Ronald W. Schafer, and John R. Buck. *Discrete-time signal processing (2nd ed.)*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1999.

[8] T. Fusco, A. Petrella, and M. Tanda. Sensitivity of multi-user filter-bank multicarrier systems to synchronization errors. In *Proc. 3rd International Symposium on Communications, Control and Signal Processing ISCCSP 2008*, pages 393–398, 12–14 March 2008.

[9] T. Fusco, A. Petrella, and M. Tanda. Data-aided symbol timing estimation for multiple access OFDM/OQAM systems. *accepted for presentation at the IEEE International Conference on Communications (ICC 2009)*, 2009.

[10] Deliverable D2.2 PHYDYAS ICT-211887. Blind techniques and strategy for synchronization and initialization. July 2008.

[11] Deliverable D3.1 PHYDYAS ICT-211887. Transmit/receive processing (single antenna). July 2008.

[12] D.S. Waldhauser, L. G. Baltar, and J. A. Nossek. MMSE subcarrier equalization for filter bank based multicarrier systems. In *Signal Processing Advances Wireless Communications, 2008. SPAWC 2008. IEEE 9th Workshop*, Recife, Brazil, July 6-9 2008.

[13] B. Zerlin, M. Ivrlac, W. Utschick, J. Nossek, I. Viering, and A. Klein. Joint optimization of radio parameters in hsdpa. In *Proc. VTC 2005-Spring Vehicular Technology Conference 2005 IEEE 61st*, volume 1, pages 295–299, 30 May–1 June 2005.

[14] R. G. Gallager. *Information Theory and Reliable Communication*. John Wiley and Son, 1968.

[15] A. Demers, S. Keshav, and S. Shenker. Analysis and simulation of a fair queueing algorithm. *ACM SIGCOMM '89*, Sep 1989.

[16] C. Cicconetti, A. Erta, L. Lenzini, and E. Mingozzi. Performance evaluation of the IEEE 802.16 MAC for QoS support. 6(1):26–38, Jan. 2007.

[17] WiMAX Forum. Mobile WiMAX - part I: A technical overview and performance evaluation. Aug 2006.

[18] Hoon Kim and Youngnam Han. A proportional fair scheduling for multicarrier transmission systems. 9(3):210–212, March 2005.

[19] Qingwen Liu, Xin Wang, and G. B. Giannakis. A cross-layer scheduling algorithm with qos support in wireless networks. 55(3):839–847, May 2006.

[20] Sang Soo Jeong, Dong Geun Jeong, and Wha Sook Jeon. Cross-layer design of packet scheduling and resource allocation in ofdma wireless multimedia networks. In *Proc. VTC 2006-Spring Vehicular Technology Conference IEEE 63rd*, volume 1, pages 309–313, 2006.

[21] Lihua Wan, Wenchao Ma, and Zihua Guo. A cross-layer packet scheduling and subchannel allocation scheme in 802.16e ofdma system. In *Proc. IEEE Wireless Communications and Networking Conference WCNC 2007*, pages 1865–1870, 11–15 March 2007.

[22] I. Gutiérrez, F. Bader, and J.L. Pijoan. New prioritization function for user scheduling in ofdma systems. In *Proc. IEEE Radio Wireless Symposium*, San Diego (USA), January 2009.

[23] V. Bharghavan, Songwu Lu, and T. Nandagopal. Fair queuing in wireless networks: issues and approaches. *IEEE Personal Communications*, 6(1):44–53, Feb. 1999.

[24] H. J. Kushner and P. A. Whiting. Convergence of proportional-fair sharing algorithms under general conditions. 3(4):1250–1259, July 2004.

[25] B. Classon, K. Baum, V. Nangia, R. Love, Yakun Sun, R. Nory, K. Stewart, A. Ghosh, R. Ratasuk, Weimin Xiao, and Jun Tan. Overview of UMTS air-interface evolution. In *Proc. VTC-2006 Fall Vehicular Technology Conference 2006 IEEE 64th*, pages 1–5, 25–28 Sept. 2006.

[26] Y. Ben-Shimol, I. Kitroser, and Y. Dinitz. Two-dimensional mapping for wireless OFDMA systems. 52(3):388–396, Sept. 2006.

[27] A. Erta, C. Cicconetti, and L. Lenzini. A downlink data region allocation algorithm for IEEE 802.16e OFDMA. In *Proc. 6th International Conference on Information, Communications & Signal Processing*, pages 1–5, 10–13 Dec. 2007.

[28] C. Desset, E. B. de Lima Filho, and G. Lenoir. WiMAX downlink OFDMA burst placement for optimized receiver duty-cycling. In *Proc. IEEE International Conference on Communications ICC '07*, pages 5149–5154, 24–28 June 2007.

[29] A. Mourad. *On the system Level Performance of MC-CDMA Systems in the Downlink*. PhD thesis, Ecole National Superior Des Telecommunications de Bretagne, June 2006.

[30] IEEE 802.16m 07/002r4. TGm system requirements document (SRD), October 2007.

[31] M. O. Hasna and M. S. Alouini. Application of the harmonic mean statistics to the end-to-end performance of transmission systems with relays. In *Proc. IEEE Global Telecommunications Conference GLOBECOM '02*, volume 2, pages 1310–1314, 17–21 Nov. 2002.

[32] 3GPP TS 36.300 version 8.4.0 Release 8. E-UTRA and E-UTRAN overall description, April 2008.

[33] Cheong Yui Wong, R. S. Cheng, K. B. Lataief, and R. D. Murch. Multiuser ofdm with adaptive subcarrier, bit, and power allocation. 17(10):1747–1758, Oct. 1999.

[34] Xin Wang, G. B. Giannakis, and Yingqun Yu. Channel-adaptive optimal ofdma scheduling. In *Proc. 41st Annual Conference on Information Sciences and Systems CISS '07*, pages 536–541, 14–16 March 2007.

[35] M. Einhaus and O. Klein. Performance evaluation of a basic OFDMA scheduling algorithm for packet data transmissions. In *Proc. 11th IEEE Symposium on Computers and Communications ISCC '06*, pages 695–702, 26–29 June 2006.

[36] Zhong-Hai Han and Yong-Hwan Lee. Opportunistic scheduling with partial channel information in ofdma/fdd systems. In *Proc. VTC2004-Fall Vehicular Technology Conference 2004 IEEE 60th*, volume 1, pages 511–514, 26–29 Sept. 2004.

[37] Seokhyun Yoon, O. Somekh, O. Simeone, and Y. Bar-Ness. A comparison of opportunistic transmission schemes with reduced channel information feedback in ofdma downlink. In *Proc. IEEE 18th International Symposium on Personal, Indoor and Mobile Radio Communications PIMRC 2007*, pages 1–5, 3–7 Sept. 2007.

[38] Y. J. Choi and Saewoong Bahk. Selective channel feedback mechanisms for wireless multichannel scheduling. In *Proc. International Symposium on a World of Wireless, Mobile and Multimedia Networks WoWMoM 2006*, page 10pp., 26–29 June 2006.

[39] Jung Hyoung Kwon, Duho Rhee, Il Mu Byun, Kwang Soon Kim, and Keum Chan Whang. Efficient adaptive transmission technique for multiuser ofdma systems with reduced feedback rate. In *Proc. Wireless Telecommunications Symposium WTS '06*, pages 1–5, April 2006.

[40] K. I. Pedersen, G. Monghal, I. Z. Kovacs, T. E. Kolding, A. Pokhariyal, F. Frederiksen, and P. Mogensen. Frequency domain scheduling for ofdma with limited and noisy channel feedback. In *Proc. VTC-2007 Fall Vehicular Technology Conference 2007 IEEE 66th*, pages 1792–1796, Sept. 30 2007–Oct. 3 2007.

[41] Jieying Chen, Randall A. Berry, and Michael L. Honig. Performance of limited feedback schemes for downlink OFDMA with finite coherence time. In *Proc. IEEE International Symposium on Information Theory ISIT 2007*, pages 2751–2755, 24–29 June 2007.

[42] A. G. Marques, G. B. Giannakis, F. F. Digham, and F. J. Ramos. Power-efficient wireless ofdma using limited-rate feedback. 7(2):685–696, February 2008.

[43] Joseph Y.-T. Leung, Tommy W. Tam, C. S. Wong, Gilbert H. Young, and Francis Y.L. Chin. Packing squares into a square. *J. Parallel Distrib. Comput.*, 10(3):271–275, 1990.

[44] T. Ohseki, M. Morita, and T. Inoue. Burst construction and packet mapping scheme for ofdma downlinks in ieee 802.16 systems. In *Proc. IEEE Global Telecommunications Conference GLOBECOM '07*, pages 4307–4311, 26–30 Nov. 2007.

[45] Ki-Ho Lee, Dong-Ho Cho, Taesoo Kwon, Howon Lee, Sik Choi, and Juyeop Kim. Overlapped two-dimensional resource allocation scheme in ofdma dl-map. *IEEE C802.16e-04/352r2*, August 2004.