Tero Jokela

Design and Analysis of Forward Error Control Coding and Signaling for Guaranteeing QoS in Wireless Broadcast Systems

Turku Centre for Computer Science

TUCS Dissertations
No 128, October 2010

Design and Analysis of Forward Error Control Coding and Signaling for Guaranteeing QoS in Wireless Broadcast Systems

Tero Jokela

To be presented, with the permission of the Faculty of Mathematics and Natural Sciences of the University of Turku, for public criticism in Auditorium Delta in ICT building on October 8 2010, at 12 noon.

University of Turku
Department of Information Technology
Communication Systems

2010

Supervisors

Adjunct professor Valery Ipatov
Professor Jouni Isoaho
Dr. Tech. Jarkko Paavola

Department of Information Technology
University of Turku
FIN-20014 University of Turku
Finland

Reviewers

Professor Catherine Douillard
Telecom Bretagne
Technopole Brest-Iroise, CS 83818
29238 Brest Cedex 3
France

Professor Markku Juntti
Centre for Wireless Communications
University of Oulu
P.O. Box 4500, FIN-90014 University of Oulu
Finland

Opponent

Professor Pablo Angueira
Department of Electronics and Telecommunications
Bilbao Faculty of Engineering
University of the Basque Country (UPV/EHU)
Alda. Urkijo S/N 48013 Bilbao
Spain

ISBN 978-952-12-2461-4
ISBN 978-952-12-2462-1 (PDF)
ISSN 1239-1883
Painosalama Oy

Abstract

Broadcasting systems are networks where the transmission is received by several terminals. Generally broadcast receivers are passive devices in the network, meaning that they do not interact with the transmitter. Providing a certain Quality of Service (QoS) for the receivers in heterogeneous reception environment with no feedback is not an easy task. Forward error control coding can be used for protection against transmission errors to enhance the QoS for broadcast services. For good performance in terrestrial wireless networks, diversity should be utilized. The diversity is utilized by application of interleaving together with the forward error correction codes. In this dissertation the design and analysis of forward error control and control signaling for providing QoS in wireless broadcasting systems are studied.

Control signaling is used in broadcasting networks to give the receiver necessary information on how to connect to the network itself and how to receive the services that are being transmitted. Usually control signaling is considered to be transmitted through a dedicated path in the systems. Therefore, the relationship of the signaling and service data paths should be considered early in the design phase. Modeling and simulations are used in the case studies of this dissertation to study this relationship.

This dissertation begins with a survey on the broadcasting environment and mechanisms for providing QoS therein. Then case studies present analysis and design of such mechanisms in real systems. The mechanisms for providing QoS considering signaling and service data paths and their relationship at the DVB-H link layer are analyzed as the first case study. In particular the performance of different service data decoding mechanisms and optimal signaling transmission parameter selection are presented. The second case study investigates the design of signaling and service data paths for the more modern DVB-T2 physical layer. Furthermore, by comparing the performances of the signaling and service data paths by simulations, configuration guidelines for the DVB-T2 physical layer signaling are given. The presented guidelines can prove useful when configuring DVB-T2 transmission networks. Finally, recommendations for the design of data and signaling paths are given based on findings from the case studies. The requirements for the signaling design should be derived from the requirements for the main services. Generally, these requirements for signaling should be more demanding as the signaling is the enabler for service reception.

Acknowledgments

The research for this thesis was funded by the University of Turku and Turku Centre for Computer Science. I also thank Nokia foundation, HPY research foundation and Finnish foundation for technology promotion (Tekniikan Edistämissäätiö) for further financial support.

I am grateful to several persons who have contributed substantially to the completion of this thesis. Dr. Jarkko Paavola has contributed to the thesis greatly through guidance all along the process. Professor Valery Ipatov directed and encouraged in the crucial early stages of the research process. I want to thank Professor Jouni Isoaho for providing me the opportunity to work towards the dissertation at the university of Turku. I thank my reviewers Professor Catherine Douillard and Professor Markku Juntti for prompt reviews and constructive comments that enhanced the quality of the thesis. I also express my gratitude to Professor Pablo Angueira for agreeing to act as the opponent.

Warm thanks go also to Mikko Jalonen, the manager of the DTV-group in which I have had the privilege to work. The spirit of the team has been superb. I wish to thank all my colleagues at the team, in the Department of Information Technology and in Åbo Akademi University I have had the opportunity to work with. Specially I would like present my gratitude to two excellent colleagues and friends Heidi Himmanen and Jussi Poikonen for their friendship over the years. The co-operation of DTV-group with Nokia team in Turku, currently located in Salo, affected the outcome of this thesis remarkably by widening the perspective of the author. Therefore, I wish also to thank the personnel of the Nokia team for fruitful co-operation without which this thesis would have looked very different.

Finally and most importantly, I thank my family and friends. My parents Reija and Hannu have always encouraged and supported me in whatever I have decided to pursue. Specially I am grateful to my dear wife, Sanna for being there for me. Without her support and encouragement this thesis would never have been done.

Turku, September 13, 2010

Tero Jokela

Contents

List of Acronyms

0G	Analog wireless telephone systems before 1G
1G	First Generation cellular mobile phone system
2G	Second Generation cellular mobile phone system
3G	Third Generation cellular mobile phone system
ACE	Active Constellation Extension
ARP	AutoRadioPuhelin
ARQ	Automatic Repeat-reQuest
ATSC	Advanced Television Systems Committee
ATSC-M/H	Advanced Television Systems Committee - Mobile/Handheld
AWGN	Additive White Gaussian Noise
BB	BaseBand
BCH	Bose-Chaudhuri-Hocquenghem
BER	Bit Error Rate
BICM	Bit Interleaved Coding and Modulation
BPSK	Binary Phase Shift Keying
C/N	Carrier-to-Noise ratio
CDS	Carrier Distribution Sequence
CE	CRC Erasure
CFT	Call for Technologies
CM	Commercial Module
CMMB	China Multimedia Mobile Broadcasting
CRC	Cyclic Redundancy Check
CSP	Common Simulation Platform
DAB	Digital Audio Broadcasting
DBPSK	Differential Binary Phase Shift Keying
DFT	Discrete Fourier Transform
DMB-T/H	Digital Multimedia Broadcast-Terrestrial/Handheld
DRM	Digital Radio Mondiale
DTMB	Digital Terrestrial Multimedia Broadcast
DVB-C	Digital Video Broadcasting - Cable
DVB-C2	Digital Video Broadcasting - 2nd generation Cable
DVB-H	Digital Video Broadcasting - Handheld
DVB-IPDC	Digital Video Broadcasting - IP DataCasting
DVB-NGH	Digital Video Broadcasting - Next Generation Handheld
DVB-S	Digital Video Broadcasting - Satellite
DVB-S2	Digital Video Broadcasting - 2nd generation Satellite
DVB-SH	Digital Video Broadcasting - Satellite Handheld
DVB-T	Digital Video Broadcasting - Terrestrial
DVB-T2	Digital Video Broadcasting - 2nd generation Terrestrial

EPG	Electronic Program Guide
ESR	Erroneous Seconds Ratio
ETSI	European Telecommunications Standards Institute
FEC	Forward Error Correction
FEF	Future Extension Frame
FER	Frame Error Rate
FFT	Fast Fourier Transform
FM	Frequency Modulation
GF	Galois Field
GPS	Global Positioning System
GS	Generic Stream
GSE	Generic Stream Encapsulation
GSM	Global System for Mobile communications
HDTV	High Definition TeleVision
INT	IP/MAC Notification Table
IP	Internet Protocol
ISDB-T	Integrated Services Digital Broadcasting - Terrestrial
ISI	InterSymbol Interference
ITU	International Telecommunication Union
L1	Physical layer
L2	Data link layer
LDPC	Low Density Parity Check
LLR	Log-Likelihood Ratio
MBMS	Multimedia Broadcast and Multicast Services
Mbps	Megabits per second
MDS	Maximum Distance Separable
MFER	MPE-FEC Frame Error Rate
MFN	Multiple Frequency Network
MHE	MPE Header Erasure
MIMO	Multiple-Input and Multiple-Output
MISO	Multiple-Input and Single-Output
MPE	Multi-Protocol Encapsulation
MPE-FEC	Multi-Protocol Encapsulation - Forward Error Correction
MPEG	Moving Picture Experts Group
MSS	Modulation Signaling Sequence
NIT	Network Information Table
NMT	Nordic Mobile Telephony
OFDM	Orthogonal Frequency Division Multiplexing
OSI	Open Systems Interconnection
PAPR	Peak-to-Average Power Ratio
PAT	Program Association Table
PE	PID Erasure
PER	Packet Error Rate

PID	Packet IDentifier
PLP	Physical Layer Pipe
PMT	Program Map Table
PP	Pilot Pattern
PRBS	Pseudo Random Binary Sequence
PSI/SI	Program Specific Information / Service Information
QEF	Quasi Error Free
QoS	Quality of Service
QAM	Quadrature amplitude modulation
QPSK	Quadrature phase-shift keying
RF	Radio Frequency
RS	Reed-Solomon
RSSI	Received Signal Strength Indicator
SDTV	Standard Definition TeleVision
SFN	Single Frequency Network
SINR	Signal-to-Interference-plus-Noise Ratio
SNR	Signal-to-noise ratio
SVC	Scalable Video Coding
T-DMB	Terrestrial Digital Multimedia Broadcasting
TDM	Time Division Multiplexing
TDT	Time and Date Table
TEI	Transport Error Indicator
TFS	Time Frequency Slicing
TI	Time Interleaving
TM	Technical Module
TPS	Transmission Parameter Signaling
TR	Tone Reservation
TS	Transport Stream
TSE	Transport Stream Erasure
UMTS	Universal Mobile Telecommunications System
VSB	Vestigial Side Band
WRC	World Radiocommunication Conference

List of Symbols

$\mathbf{c}$	Codeword (vector)
$\mathcal{C}(t_u, t_e)$	Complexity function
d	Minimum Hamming distance
d_e	Euclidian distance
$\mathbf{e}$	Error vector
$\mathbf{f}$	Erasure vector
f	Frequency
f_D	Doppler frequency
Δf_{int}	Estimated integer frequency offset
G	Generator Matrix
$g(z)$	Generator polynomial
H	Parity check matrix
k	Number of information symbols
κ	Actual number of erroneus symbols
$\lambda(x)$	Error-erasure locator polynomial (time domain)
$\Lambda(x)$	Error-erasure locator polynomial (frequency domain)
N_s	Section length
n	Codeword length
p	Bit error probability
p_e	Probility of detected symbol error
p_s	Symbol error probability
p_u	Probability of undetected erroneous symbol
$\psi(x)$	Erasure locator polynomial (time domain)
$\Psi(x)$	Erasure locator polynomial (frequency domain)
q	Symbol alphabet size
R	Code rate
R_I	Information rate
t	Number of correctable errors
T_B	Cyclic extension (B) length
T_c	Coherence time
T_C	Cyclic extension (C) length
t_e	Amount of detected symbol errors
t_u	Amount of undetected symbol errors
v	Velocity
$\mathbf{v}$	Received vector
W_c	Coherence bandwidth
$\mathbf{x}$	Information vector
y	Received symbol

Chapter 1

Introduction

Wireless communication using radio frequencies (RF) plays an important role in today's world by enabling communication from places that would be impossible or at least very expensive to reach otherwise, for example with wires. Thanks to advancements in wireless communication during the last few decades, many things that were thought some 20 years ago to be science fiction are nowadays everyday life. For example, connecting to Internet or receiving video calls while on the move are currently possible in many places on the world. This rather dramatic evolution of services has been enabled mainly by the digitalization of communication systems. Perhaps one of the most well known wireless communication applications is the cellular phone system. The Finnish Autoradiopuhelin (ARP, car radio phone in English) launched in 1971 is considered to be one of the first successful commercial mobile phone systems, sometimes referred to as zero generation (0G) cellular networks. In Finland the network covered 100 % of the geographical area with 140 base stations. Nevertheless, moving between the cells (that are coverage areas of transmitters) of this analog system was not seamless, meaning that the call would be interrupted when moving from one cell to another. The follower of ARP in the nordic countries called Nordic Mobile Telephony (NMT) was opened for service in 1981 and it is referred to as being a first generation (1G) system. The analog NMT system introduced several enhancements over ARP, for example seamless cell switching, support for an increased number of users and smaller terminal size. Common to 0G and 1G systems, the main application was speech. The first modern 2nd generation (2G) digital cellular phone system called Global System for Mobile communications or Groupe Spécial Mobile (GSM) was launched in Europe in the beginning of the 1990s. Since then users all over the world have become accustomed to be able to do the same operations on their mobile devices as they can do on their personal computers and more. The third generation cellular system (3G) also called Universal Mobile Telecommunications System (UMTS) [1] was introduced in Europe in 2005. It enables higher data rates for the users as compared to 2G and thus a wider range of multimedia services is possible. For further increase in data rate, Long Term Evolution (LTE) technology is being developed and several network pilots are running at the time of writing this thesis.

Another important field of communication is broadcasting, whose development has been slightly different from that of cellular systems. Wireless broadcasting is an efficient means to deliver a certain set of services to a large number of customers with no fixed connection. Frequency Modulation (FM) radio or terrestrial television are common examples of wireless broadcasting. Wireless broadcasting dates back to the first half of the 20th century. In the beginning the broadcasts were mainly radio, but in the 1930s television broadcasts began. Since then both have been widely used mechanisms to provide people information on the world around them (for example news) and entertainment. Similarly as cellular systems, also the broadcasting world is moving towards digital systems. Currently, in many countries, the broadcasting of television is being digitalized taking into use digital broadcasting systems and switching off the analog television transmissions. In European countries the digital television is implemented using the Terrestrial Digital Video Broadcasting (DVB-T) standard [2]. For digital radio transmission Digital Audio Broadcasting (DAB) [3] is used in Europe.

The increasing popularity of digital cellular phones has led to the development of the Digital Video Broadcasting for Handheld (DVB-H) [4] system as the users of cellular phones are willing to watch televion on their terminals. DVB-H is being deployed in Europe for the delivery of streaming services such as mobile television. The presence of competitive standards indicates that there is interest towards mobile broadcasting. Examples of other mobile broadcasting standards are Terrestrial Digital Multimedia Broadcasting (T-DMB) in South Korea, China Multimedia Mobile Broadcasting (CMMB) in China, MediaFLO in North America and 1seg in Japan. The delivery of mobile television to cellular terminals can be done over cellular 3G networks also, but the delivery for numerous users is more cost efficient over true broadcast networks [5]. The advantage of digital systems is that also hybrid networks can be designed. In this case the same service could be received either via 3G or a broadcasting system depending on the availability of the signal in the networks.

The latest trend in video transmission systems is High Definition Tele-Vision (HDTV) services. For the HDTV purposes in Europe second generations of systems for cable, satellite and terrestrial transmission have been developed. The second generation Terrestrial Digital Video Broadcasting (DVB-T2) [6] services will be introduced in the near future. DVB-T2 offers a higher throughput as compared to DVB-T and thus will accommodate more HDTV services in one multiplex. Also, DVB-T2 is designed to be a very flexible system and thus it may be that also low data rate services for mobile users in combination with high data rate HDTV services for home television users will be transmitted through common DVB-T2 multiplexes. In this thesis the application focus will be on DVB-H and DVB-T2.

1.1 Quality of Service mechanisms in broadcasting systems

Service in the context of broadcasting can be roughly divided into two types, streaming and data services. The latter is sometimes also referred to as file-casting. A conventional TV channel is a streaming service while teletext can be considered a data service. Quality of Service (QoS) is used to describe the fidelity of the service from user's viewpoint. In general at least the following technical issues are considered to affect the QoS in digital broadcasting systems [7], [8]:

- Throughput (Bit rate)
- Delay or latency
- Delay variation (delay jitter)
- Error rate

QoS requirements vary for different types of services. From the standpoint of streaming video services, the throughput directly affects the audiovisual quality. Having higher throughput, a larger resolution or less lossy video encoding scheme can be utilized resulting in better video quality. For data services the throughput has an effect on the download time of the data package. Delay or latency of the content itself may not be a problem in most situations in a one-way system as it would be for a two-way systems such as for telephony or video conferencing. Still, for example in live events with possible local DVB-H broadcast of the performance, the delay can be an issue. If the user in the audience would like to see some different camera angle of the performance on the mobile television device, it will most certainly be annoying to see a few seconds delay in the video as compared to the live content. The delay variation affects the QoS through halts or frame rate changes in the video. The delay variation within certain limits can be handled by proper buffering in the receiver and in dedicated broadcasting networks the transmitter can control the amount of delay variation. On the other hand for data services the delay variation may not be a problem. The effect of the error rate is direct: the higher the error rate the more visible errors in the video. For the data services the content should be error free before it can be used. Therefore, in broadcasting data downloads repetition or application layer forward error correction (FEC) is necessary.

A broadcasting system is a one-to-all system where feedback from the receivers to the transmitter is not present. In general, in wireless transmission channels transmission errors always occur. The nature of broadcasting imposes certain restrictions on how the QoS can be provided for the users. The absence of the feedback channel directly rules out any kind of repeat request mechanisms such as Automatic Repeat-reQuest (ARQ) [9] or adaptive methods such as beamforming which means transmitting more power to the direction where users suffer from low signal. What is possible for broadcasting systems, is to use forward error control [10] codes. When using FEC codes, the transmitter inserts a certain amount of redundancy to the transmitted data that can be used in the receiver to correct transmission errors that occurred between the transmitter and the receiver, that is to reduce the error rate. The enhancement in the robustness of the service data is obtained at the cost of throughput of the system. To get the most out of the FEC codes in difficult transmission environments, interleavers are used in combination with them. Generally, interleaving means shuffling the code symbols to disperse possible error bursts from the transmission channel. As data needs to be buffered for the interleaving, this imposes delay (or latency) to the system. Therefore, interleaving together with FEC is used to trade-off between the throughput, latency and error rate.

Wireless broadcasting networks are designed to enable certain QoS for a certain share of users, which can be for example 95 % of households. Naturally there can always be some users in such unfortunate locations of the network that the QoS is below acceptable level, but for most of the users the QoS should be acceptable. From the point of view of network design the FEC and interleaving have an effect on the area that can be covered by a transmitter with given power [11]. Thus, the FEC and interleaving chain can also be considered as a measure to trade-off between the coverage area of a transmitter and throughput of the system.

1.2 Role of signaling to provide QoS

In wireless transmission systems control signaling later referred to as *signaling* is used to give the receiver necessary information on how to connect to the network itself and how to receive the services that are transmitted. Usually signaling is considered to be transmitted in its individual “signaling path” in the systems. The most challenging situation for the signaling are so called ad-hoc networks, where devices form a network by themselves with help of signaling. Also broadcasting systems introduce a challenging signaling scenario due to the lack of feedback. In broadcast systems the signaling has an effect on the QoS, as it is the enabler for the service reception. While broadcasting systems must work without a feedback channel, the signaling must be transmitted with certain intervals to enable receivers to connect to the network and to follow the services. This interval affects the delay for the user when for example switching between services. Also, if the error rate of the signaling is high it may hinder the user from receiving services at all.

Commonly, the signaling information is separated into several Open Systems Interconnection (OSI) layers [12] in the systems where each layer signals layer specific information. In this thesis mainly layers 1 (L1) and 2 (L2), namely physical and data link layers are considered with respect to both signaling and service data transmission. The physical layer is responsible for converting the binary data to a format that can be transmitted over the physical medium (that is air in wireless communications). The data link layer provides data transfer through the physical layer. The effects of higher layer signaling in a mobile broadcasting system are analyzed for example in [13].

1.3 Problem formulation for broadcasting system design

The fundamental property of broadcasting is that data transmission is unidirectional. This property emphasizes the importance of the transmitter that runs the whole show. Therefore, for broadcasting systems mainly transmitter and the signal it should produce are standardized and the receivers should be designed so that they can receive the specified signal. Still, it is the receiver user that sets the requirements for transmitter design in form of required quality of service. Therefore, receiver and transmitter should be considered simultaneously during the design process. For this a simulation model of a receiver can be implemented to assist in designing the transmitter. Simulator is an efficient tool when designing broadcasting systems as well as other telecommunication systems. The design process and use of simulations to assist the design process are studied in Chapter 3.

The technical side of system design introduces restrictions for the use of measures for enabling the wanted quality of service. For example the decoder complexity of certain well performing forward error control codes may be too high to lead to receivers of reasonable cost. For this reason, the system designers need to make trade-offs between performance and complexity. Another example of the trade-off is the limitation in the memory (and thus silicon die size) that allows interleaving. The more memory is available, the longer interleaving can be performed resulting in enhanced performance but also increased cost. Taking into account all technical and QoS specific needs and limitations is not a trivial task in the design of relatively complex systems. In this thesis trade-offs considering mainly FEC and interleaving processes for both service data and signaling paths are considered.

In addition to the service data path, the protection of the signaling path should be carefully designed. The concatenation of signaling and data on different layers in a transmission system has a role on providing QoS. This can be approached for example from the point of view of which operations a DVB-T2 receiver needs to perform in order to receive service data. First, the receiver needs to synchronize to the radio signal. Then it needs to receive the physical layer (L1) signaling in order to receive the L1 data carrying data link layer (L2) signaling. Once L2 signaling is obtained, L2 data is received in order to gain access to the next layer signaling. The concatenation of signaling and data on different layers has the effect that for example if L1 signaling is disturbed badly in the transmission channel and the L1 data cannot be received, also L2 signaling is unobtainable. Therefore, the physical layer signaling should be the most robust part of the system after the synchronization. This relationship of data and signaling paths should be carefully taken into account in the system design. How this is done in the design of DVB-T2 system is described in Chapter 6. A further problem is how the signaling and data paths can be compared to each other, that is, what are the requirements for both and using which criteria they should be designed. A commonly used design criterion for the data transmission is Bit Error Rate (BER) which is not the optimal measure of robustness for the signaling path. For L1 signaling in DVB-T2 system, the Frame Error Rate (FER) presents itself as a more representative error measure than BER.

1.4 Objectives and scope of the thesis

The main research topic of the dissertation is the relationship of the service data and signaling paths. More specifically how their transmission can be analyzed and should be designed in order to allow for optimal overall system performance in a terrestrial broadcasting environment. The objectives of the thesis can be crystallized in the following bullets:

- Study and analyze the design of mechanisms for providing required QoS in broadcasting systems. The forward error control and interleaving methods are mainly considered.
- Provide guidelines and ideas for the design and analysis of these mechanisms for future broadcasting systems emphasizing the relationship of the service data and signaling transmission.

To approach the solution to this problem, the properties of the terrestrial transmission environment and mechanisms utilized for provisioning QoS therein are studied. Then the design process for wireless communication systems is examined. To obtain practical insight into the issue the transmission of service data and signaling in the present DVB-H system are analyzed. Further, the design process for the DVB-T2 system is investigated and the performances of signaling and data transmission mechanisms are studied. Based on these studies conclusions are drawn to present ideas and suggest guidelines for the design and analysis of data and signaling paths for future systems.

The thesis is organized in eight chapters: In chapter 2 the mechanisms of guaranteeing QoS in digital broadcasting systems are studied. First of all the specialties of the broadcasting environment are presented. Then methods for utilizing diversity that is available in broadcasting systems, that is, FEC and different interleaving mechanisms are investigated.

In chapter 3 the system design process is studied focusing on the mechanisms for provisioning the QoS. Also the performance measures against which the systems are designed will be studied. Further, the use of simulations in the design process is investigated.

In chapter 4 the current status of the digital broadcasting systems worldwide is presented and the family of DVB standards is investigated in more detail to provide knowledge that helps the reader in following the next chapters dealing with these standards.

In chapter 5 the performance analysis of the DVB-H system link layer is discussed. Both signaling and data path are studied and their performance compared. In addition to theoretical calculations on the performance, simulations and field measurements are presented.

In chapter 6 the DVB-T2 physical layer standard and its design process are investigated. Also simulation results on the performance of the data and signaling are presented and guidelines for the selection of transmission parameters for the signaling are considered.

In chapter 7 the future evolution of the broadcasting systems is discussed. Further, guidelines for designing future systems based on the findings of the studies with DVB-H and DVB-T2 in chapters 5 and 6 are given.

1.5 Author's contributions

This dissertation is based on the following contributions by the author that are further appended and harmonized here. Specially sections 5.2.2, 6.1.2, 6.2.4 and 7.2 present new results. The contribution of the author is indicated for each publication.

- Tero Jokela, Jarkko Paavola, Heidi Himmanen, Valery Ipatov, "Performance Analysis of Different Reed-Solomon Erasure decoding Strategies at the DVB-H Link Layer", in *proc. IEEE International Sympo-*

sium on Personal, Indoor and Mobile Radio Communications, September 2006

Author's contributions: The author performed the theoretical analysis of the performance of different MPE-FEC decoding methods presented by Heidi Himmanen earlier. Other authors supervised and instructed in the preparation of the paper. Section 5.1 of the dissertation bases on this publication.

- Jarkko Paavola, Heidi Himmanen, Tero Jokela, Jussi Poikonen and Valery Ipatov, "The Performance Analysis of MPE-FEC Decoding Methods at the DVB-H Link Layer for Efficient IP Packet Retrieval", *IEEE Transactions on Broadcasting, Special issue on "Mobile Multimedia Broadcasting"*, Volume 53, Issue 1, Part 2, pages 263-275, March 2007

  Author's contributions: The performance comparison of the MPE-FEC decoding schemes theoretically and by simulation was the contribution of the author in this paper. Section 5.1 of the dissertation bases on this publication.

- Tero Jokela and Jani Väre, "Simulations of PSI/SI transmission in DVB-H systems", in *proc. IEEE International Symposium on Broadband Multimedia and Broadcasting*, March 2007

  Author's contributions: The performance evaluation of the PSI/SI transmission in DVB-H system was performed by the author. Jani Väre provided important practical insight on the transmission mechanism for the signaling and real receiver operations. Section 5.2.1 of the dissertation is based on this publication.

- Heidi Himmanen and Tero Jokela, "DVB-H field trials: Studying channel characteristics", in *proc. IEEE International Symposium on Broadband Multimedia and Broadcasting*, March 2007

  Author's contributions: The author performed the measurements together with Heidi Himmanen and further processed the measured raw data. The field measurements used in part of the simulations presented in section 5.2.1 are based on the measurements performed for this publication.

- Tero Jokela and Eero Lehtonen, "Reed-Solomon decoding algorithms and their complexities at the DVB-H link layer", in *proc. IEEE International Symposium on Wireless Communication Systems*, October 2007

  Author's contributions: Main contribution on the paper is on the application of the calculations for the Reed-Solomon decoding methods on the DVB-H MPE-FEC decoding. Eero Lehtonen performed the theoretical complexity calculations for the time and frequency domain Reed-Solomon decoders. Appendix A is based on this publication.

- Heidi Himmanen, Tero Jokela, Jarkko Paavola and Valery Ipatov, "Performance analysis of Reed-Solomon coding combined with error detection as the link layer FEC by computer simulations", appears in *Handbook of Mobile Broadcasting: DVB-H, DMB, ISDB-T, AND MEDIAFLO*, CRC Press, 2008

  Author's contributions: The performance comparison of the MPE-FEC decoding schemes theoretically and by simulation was the contribution of the author in this book chapter. Section 5.1 of the dissertation bases on this publication.

- Tero Jokela, "Performance Analysis of Substituting DVB-S2 LDPC Code for DVB-T Error Control Coding System", in *proc. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, April 2008

  Author's contributions: This paper is an independent contribution by the author, where simulation platform is generated and utilized for studying the channel coding alternative for the existing DVB-T system. The reason for the study was to assist in the design of the second generation terrestrial DVB system. Section 6.1.1 of the dissertation bases on this publication.

- Tero Jokela and Jarkko Paavola, "Robustness analysis of physical layer signalling transmission in DVB-T2", in *proc. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, May 2009

  Author's contributions: The paper is mainly contribution of the author. Jarkko Paavola supervised and assisted in the writing. Paper studies the performance of the signaling in DVB-T2 system with the help of simulator created partly by the author. Sections 6.2.2, 6.2.3 and 6.2.4 base on this publication.

- Tero Jokela, Miika Tupala and Jarkko Paavola, "Analysis of physical layer signaling transmission in DVB-T2", *IEEE Transactions on Broadcasting*, Volume 56, Issue 3, pages 410-417, September 2010

  Author's contributions: Authors contribution is on the analysis of the signaling transmitted in P2 symbols. Miika Tupala contributed on the performance evaluation of the P1 signaling reception. Jarkko Paavola supervised and assisted in the writing process. Sections 6.2.2, 6.2.3 and 6.2.4 base on this publication.

Chapter 2

Mechanisms of guaranteeing QoS in wireless broadcast system

In this chapter the mechanisms that are used to guarantee the performance of a wireless broadcasting system are presented. The mechanisms apply to both data path and signaling. We start with the characteristics of the broadcasting environment. Let us assume that we are given certain bandwidth at some frequency range and transmission power. These are regulated by the governments in most countries. Having bandwidth and power fixed, the most important resource to enable reliable wireless communication is diversity. The concept of diversity is described in section 2.2. Further, the mechanisms for utilizing the diversity in digital communication systems to combat the harmful effects of the wireless environment are presented. The aim of the chapter is to make the reader familiar with the basic concepts that will be discussed later in the dissertation.

2.1 Wireless broadcasting environment

The reception scenarios in wireless broadcasting are usually separated into mobile and static reception. The transmission channel observed by a mobile receiver is more challenging from the receiver point of view than the static one. A mobile receiver can move around in a city center, rural area, inside buildings, inside a train and so on. Thus a rich variety of different usage scenarios and environments is possible. A typical behavior for the mobile wireless channel is that it varies over time and frequency and the variations can be relatively rapid. Static receiver corresponds to for example a normal living room television with a fixed roof top antenna. The transmission received by a static receiver will experience much less changes in the

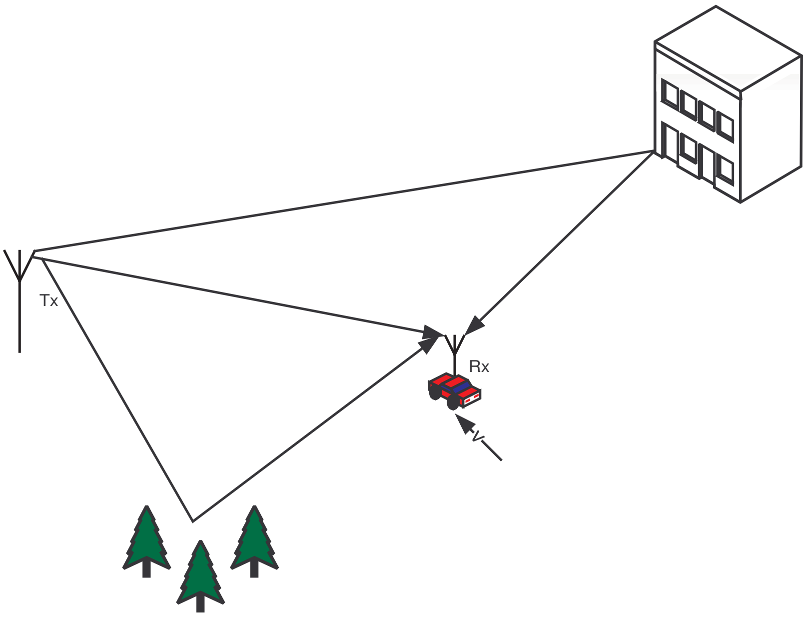

Figure 2.1: An example scenario of multipath propagation.

channel quality as compared to the mobile one. These changes are due to moving objects near the receiving antenna such as cars or swaying trees on windy weather for example. The variations are commonly divided into two categories as defined in [14]:

- *Large scale fading* is due to path loss of the signal as the receiver moves and the signal is shadowed by large objects such as buildings etc.

- *Small scale fading* is due to constructive and destructive interference of the different signal paths from the transmitter to the receiver.

Large scale fading is mostly relevant network planning when placing and directing the transmitters while small scale fading is important for the design of the communication systems [14]. This is why the small scale fading is mostly addressed in this chapter.

An example of a scenario illustrating multipath propagation resulting in small scale fading in a terrestrial environment is shown in Fig. 2.1. It is observed that the different propagation paths can have different lengths and thus different delays. Assuming that the transmitted signal is sinusoidal, the different delays of the paths result in different phases for the signals received through different paths. This effectively causes the signals from different paths to add either constructively or destructively in the receiver, depending on the phase. When most of the paths add constructively, the signal is strong. If most paths add destructively, the signal is weak. Let us further consider that the receiver is moving with velocity v for example to the direction shown by the arrow in the figure. Now the lengths of the paths can vary with different velocities. As the distances of the receiver from the transmitter via different paths change, the paths experience Doppler shifts, that is, shifts in frequency. The Doppler shift of one path can be evaluated by

$$\Delta f = \frac{v_{path}}{c} f_0, \tag{2.1}$$

where v_{path} is the velocity of the change of the current path, c is the speed of the wave in the medium, that is, the speed of light in the case of a radio wave propagating on air and f_0 is the emitted frequency. The change in frequency Δf can be either positive if the receiver is moving towards the transmitter or negative if moving away from the transmitter. The *coherence time* T_c of a channel is the time duration over which two received signals have a strong potential for amplitude correlation [15]. Based on the maximum Doppler shift Δf_{max} among the propagation paths the popular rule of thumb for the coherence time of the channel can be estimated by [15]:

$$T_c = \frac{0.423}{\Delta f_{max}}. \tag{2.2}$$

The coherence time defines how fast the channel changes in time. From the equation it is important to note that the larger the Doppler shift, the smaller the coherence time. This intuitively means that the faster the receiver moves, the faster the channel changes. In addition to coherence time, also coherence bandwidth W_c plays an important role in communications. Coherence bandwidth indicates how quickly the channel changes in frequency. Coherence bandwidth can be evaluated based on the delay spread T_d indicating the difference in propagation time of the shortest and the longest path with significant energy. W_c is estimated by [14]:

$$W_c = \frac{1}{2T_d}. \tag{2.3}$$

In addition to fading, wireless channels often suffer from impulsive noise. Impulsive noise is a short noise pulse of relatively high power. Possible sources of impulsive noise are for example ignition systems of automobiles and electrical switches that can introduce sparks. Usually impulsive noise introduces a burst of errors in the receiver. This kind of noise has been found to have an effect on the transmission of DVB-T signals and receiver algorithms for mitigating it are presented for example in [16].

2.2 Diversity

As the wireless transmission channel varies in time and frequency, there is a significant probability that the channel will be in a deep fade every now and then. When the channel is in a deep fade, the signal is severely distorted by the channel and the receiver is not able to decode the transmission resulting in transmission errors. To minimize the effect of the fades the system should be designed so that the information passes through multiple channels that fade independently. Now the probability that all the channels will be in deep fade simultaneously is made smaller and thus the reliability of the communication system is increased. The concept is called *diversity*. Diversity can be viewed as a form of redundancy [17]. It is widely utilized in present communication systems.

As presented in [14] and [17], different kinds of diversity in telecommunications can be divided into three main categories: *Time*, *Frequency* and *Space* diversity. For time variant channels utilization of time diversity translates into usage of error control coding and interleaving. Information is coded and the coded symbols are dispersed over time in different coherence periods of the channel. This way different parts of the codewords fade independently. An illustration of the utilization of time diversity is shown in Fig. 2.2. Let us consider that codewords (shown with different colors) consist of four symbols and the error control code can correct one erroneous symbol in each codeword. In our example if the time diversity is not utilized, one codeword (blue) is transmitted during a deep fade and all of its symbols are erroneous. Clearly our error control code capable of correcting one erroneous symbol in each codeword is not capable of correcting the errors of this codeword. Next, consider a situation when the codewords are dispersed over a longer time period, say over the duration of four codewords. We can see that with a similar channel only one symbol of each codeword is erroneous and this can be corrected by the error control code. This is the idea for utilizing time diversity in communication systems. In general to exploit time diversity, interleaving and coding over several coherence time periods T_c should be performed [14].

For wideband signals where the signal bandwidth W is greater than the coherence bandwidth W_c of the transmission channel the frequency response is not flat. Thus the channel varies in frequency and resource of frequency diversity is available. For Orthogonal Frequency Division Multiplexing (OFDM) systems (such as the ones discussed in this dissertation) the total signal bandwidth is divided into orthogonal sub-carriers. Selecting the number of sub-carriers in relation to the bandwidth of the signal and coherence bandwidth appropriately, the sub-carriers will experience narrowband flat fading. Therefore, in OFDM systems the frequency diversity can be utilized in the same way as for time diversity by dispersing the coded in-

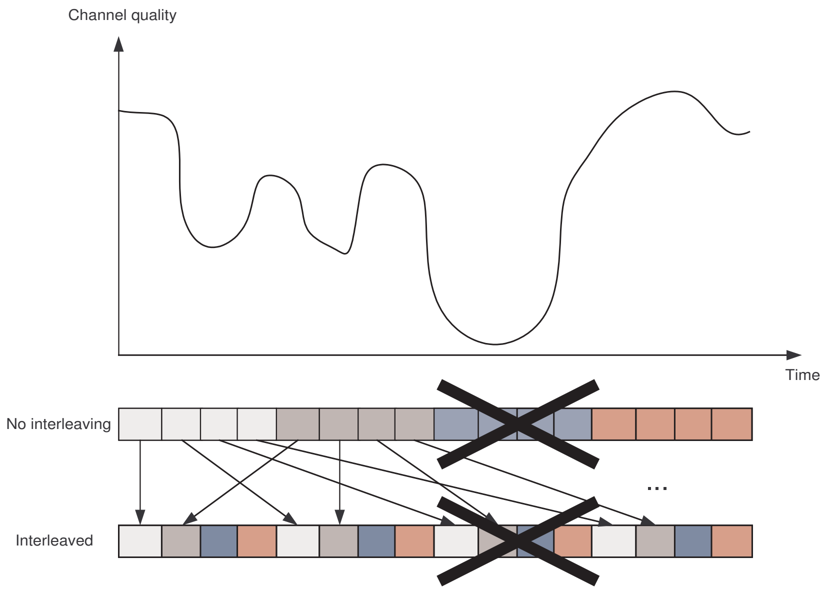

Figure 2.2: Illustration of the utilization of time diversity by error control coding and interleaving.

formation in the frequency direction, that is, over the sub-carriers of several coherence bandwidths of the channel. For information on the mechanisms for utilizing the frequency diversity in single carrier and spread spectrum systems, the interested reader can turn for example to [14].

In a transmission system with multiple transmitting and/or receiving antennas diversity in space can also be obtained. When the antennas within the transmitting or receiving antenna array are placed sufficiently far from each other physically, they introduce signal paths that experience independent fading. The sufficient distance between the antennas in the antenna array depends on the environment and carrier frequency. Typically a separation of half to one carrier wavelength is considered to be enough [14]. Also, different paths can be obtained by using different polarizations in the transmissions. When separate independently fading paths are available, they can be combined to avoid deep fades. Several combining methods have been designed and a good overview of the most important ones can be found in [18]. Further, when the transmitter has more than one antenna, space-time coding can be performed. Space-time coding distributes a code over several antennas and time instants. A description of a popular mechanism for space-time coding that is called *Alamouti scheme* after its inventor is presented in [19].

2.3 Techniques for utilizing diversity

Now that the basics of wireless transmission environments and diversity issues are presented we can focus on the mechanisms that allow the designer to utilize the different forms of diversity. First we go through Forward Error Control (FEC) mechanisms and then introduce different interleaving schemes.

2.3.1 Forward error correction

Forward Error Correction is a mechanism for error control. It is mainly used in digital communication and data storage systems. In communication systems the sender adds redundant data to its messages that allows detection and correction of a certain number of errors in the receiver. The algorithm that is used for calculating this redundancy is commonly called the *error control code* or in short just *code* and its properties set the amount of errors the code can correct. The redundant symbols generally depend on many information symbols. There are several different error control codes known today with different characteristics. One code can be better suited for low complexity decoding while some other may have very good error correction capability for example. Therefore it is an integral part of system design to select the best code for the application.

FEC codes are usually divided into two categories, *block* and *trellis* codes. The difference is in the processing of the input data. Block codes take input as chunks of k information symbols and calculate the redundancy for these symbols ending up in a codeword of length n. Once the redundancy is added to the information symbols, the next chunk is processed. One important subclass of block codes are *linear* block codes. They are widely used because their structure allows for low complexity decoding algorithms [20]. Linear block codes are generally defined by their *generator matrices*.

Trellis codes, the most important subclass of which are convolutional codes, on the other hand process a continuous stream of input data and add parity constantly. The division between trellis and block codes is not a very strict one. Any block code can be described in terms of its trellis and also any trellis code can be treated like a very long block code and a special rule of encoding information bits into redundancy [21]. Tail biting technique can be used to obtain shorter block lengths as studied in [22].

Let us go through some important properties of block codes with help of a binary Hamming code of length $n = 7$ and $k = 4$. Hamming codes were discovered by R. W. Hamming [23] around 1950. The generator matrix $\mathbf{G}$

for the $(7,4)$ Hamming code is

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$$

The encoding of the information symbols to the codewords is simply done by multiplying the generator matrix by the information vector (using modulo-2 arithmetic for binary code), that is:

$$\mathbf{c} = \mathbf{x}\mathbf{G}. \tag{2.4}$$

For example with the information sequence $\mathbf{x} = (1,0,0,0)$ the resulting codeword is $\mathbf{c} = (1,0,0,0,1,1,1)$. This code is called *systematic*, as the information sequence is visible in the codeword. As we introduce additional symbols when encoding the messages, the bandwidth required by the transmission of the information is increased. A measure for this bandwidth expansion is *code rate* that is the ratio of the useful information and the total codeword length:

$$R = \frac{k}{n}. \tag{2.5}$$

For our Hamming code the rate is $R = 4/7$. Another important measure for codes is the minimum distance of the code. The minimum distance is defined as the minimum Hamming distance between any two codewords of the code. Hamming distance is the number of differing symbols in two codewords. For the $(7,4)$ Hamming code the minimum distance is known to be $d = 3$. For block codes the error correction capability depends on the minimum distance properties of the code in the following way:

$$d \geq 2t + 1, \tag{2.6}$$

where t is the number of errors that can be corrected. With this formula we can calculate that the $(7,4)$ Hamming code can correct one error in each codeword. The error correction capability t can also be illustrated geometrically as shown in Fig. 2.3. There are three different codewords $\mathbf{u_1}$, $\mathbf{u_2}$ and $\mathbf{u_3}$ that are surrounded by a Hamming sphere of radius t. Clearly, if the spheres do not intersect, the decoder will decode a vector within each sphere to the corresponding closest codeword. This means that t or less errors in each codeword will be corrected. To avoid the intersection of the spheres the minimum hamming distance between the codewords should be exactly as defined in Eq. 2.6. In the figure the received vector $\mathbf{v}$ is within the hamming sphere of codeword $\mathbf{u_3}$ and will thus be decoded as being $\mathbf{u_3}$.

The function of the decoder of the FEC codes in general is to find the codeword that was most likely transmitted knowing the received vector.

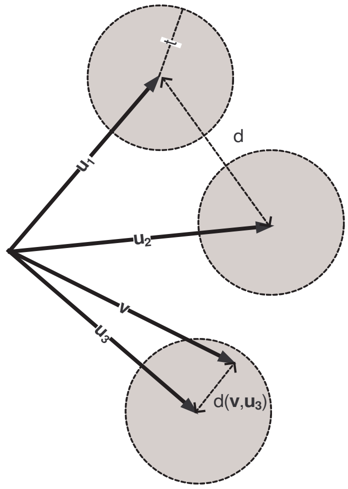

Figure 2.3: Illustration of correction capability and minimum distance of a code.

One measure for evaluating the likelihood is the Hamming distance. The decoder can evaluate the Hamming distance of the received vector to all codewords and decide that the codeword with the smallest Hamming distance was transmitted. This kind of decoding is called *hard decision* decoding. In hard decision decoding the observed symbols are converted back to the original alphabet (binary for example) before the decoding process. If more fine grained information on the received signal is available at the decoder, *soft decision* decoding can be performed. In soft decision decoding symbols are not converted back to the original alphabet before the decoding, instead likelihood for each symbol is used in calculating the distance of the received vector to the codewords. For soft decision decoding *Euclidean distance* instead of Hamming distance can be used as the measure for estimating the likelihood of the whole codeword. The Euclidean distance for the observation $\mathbf{v} = (v_1, v_2, ..., v_n)$ and codeword $\mathbf{c} = (c_1, c_2, ..., c_n)$ is calculated as:

$$d_e(\mathbf{v},\mathbf{c}) = \sqrt{\sum_{i=1}^{n}(v_i - c_i)^2}. \tag{2.7}$$

Obviously for best error correction capability under soft decision decoding, the Euclidean distance between the codewords should be as large as possible. Again, based on the observation $\mathbf{v}$, the codeword closest to the received vector is selected. Generally soft decision decoding is more efficient than hard decision decoding as making hard decisions before the decoding leaves

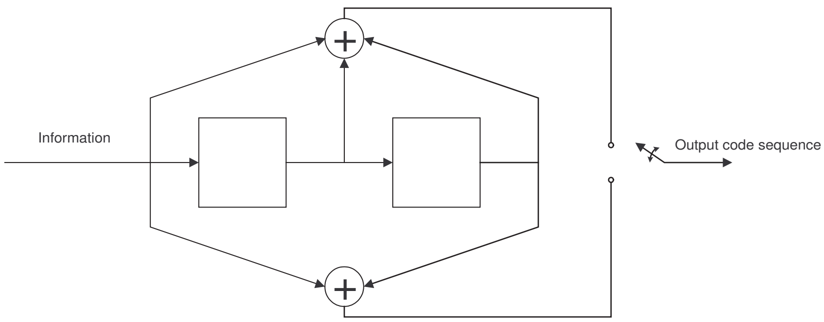

Figure 2.4: Convolutional encoder (Rate 1/2).

the decoder with less information to base its decisions on. Different kinds of decoder algorithms for both soft and hard decision decoding for different codes are widely presented in the literature. For soft decision decoding perhaps the most important are the Viterbi algorithm [24] for convolutional codes and belief-propagation based iterative algorithms used for decoding Low Density Parity Check (LDPC) [25] and turbo codes [26].

Let us next go through the important properties of convolutional codes by studying the encoder of a simple code shown in Fig. 2.4. The figure illustrates a shift register that can be used to encode rate $R = 1/2$ convolutional code. The operation of the encoder is the following. As one information bit is clocked into the register, the outputs of the modulo-2 adders are multiplexed to form two coded bits (thus the rate 1/2). When clocking the next information bit in, the previous information bit moves into the next memory element. It is observed that a certain information bit affects not only the output during its own bit interval, but also during the next two intervals it is stored in the memory elements. A convolutional code is fully defined by the number of memory elements, the number of outputs and the connections between the shift registers and the modulo-2 adders [20]. The connections are commonly presented by *generator polynomials* where the degree of the formal variable z corresponds to the index number of a memory element and its coefficient is 1 only if a connection from the output of the memory element to the modulo-2 adder is present. For the code of Fig. 2.4 the generator polynomial for the upper output path is $g_1(z) = 1 + z + z^2$, since there is a connection from before the first memory element ($z^0 = 1$) and after both memory elements(z^1 and z^2). Similarly for the lower output path the generator polynomial is $g_2(z) = 1 + z^2$. Often, the generator polynomials are presented in the literature in octal form. To obtain this form, all polynomial coefficients are written from left to right as a single binary number and then converted into octal number system. The generator polynomials of the code in our example would be $(g_1(z), g_2(z)) = (7, 5)$. An important parameter that affects the performance of the convolutional code is the *constraint length* that is defined in [20] to be the number of memory elements in the encoder plus one. For the code of our example the constraint length is three. In analogy to the minimum distance of the block codes, the convolutional codes have a measure called *free distance* that describes the distance properties of the code. It does not directly tell the designer how many errors the code can correct as the convolutional codes do not have codewords or they can be considered to be of infinite length. However, the bigger the free distance of the code the better the performance. The error correction capability of a convolutional code within a sequence of a few (3 to 5) constraint lengths can be evaluated using the Eq. 2.6 [18].

One more classification for error control codes is to divide them into *systematic* and *non-systematic*. For systematic codes information and redundancy are separated from each other, so that for block codes the information can be observed from fixed locations of the codewords even before the decoding. Convolutional codes are said to be systematic if one of their generator polynomials equals one, that is, the input bit is directly one of the output bits.

The best codes known today from the error correction performance point of view are the LDPC and turbo codes. Due to high calculation capabilities of the current receivers, the codes can be efficiently decoded using iterative belief-propagation algorithms. Both these codes are reported to perform less than a decibel away from the theoretical capacity of the channel also called the Shannon Limit. For large length LDPC codes, the performance is even reported to be only fractions of a decibel away from the Shannon limit [27]. Similar results are presented for turbo codes in [28].

2.3.2 Interleaving

Most error control codes are designed to correct random independent errors. When the channel has memory, like the multipath channel already described does, the errors occurring in the channel are not independent. Instead they appear usually in error bursts of variable duration. Introducing bursty errors to the decoder of a code designed for independent errors degrades the error correction performance of the system. Coding mechanisms for channels with memory have been presented, but the problem is to match them to the variable nature of for example wireless multipath channels [29]. A more direct approach is to interleave the encoded data to make the channel with memory seem memoryless for the decoder. Therefore the task of the interleaver is to shuffle the code symbols over a span of several code block lengths. The span is determined by the channel characteristics as de-

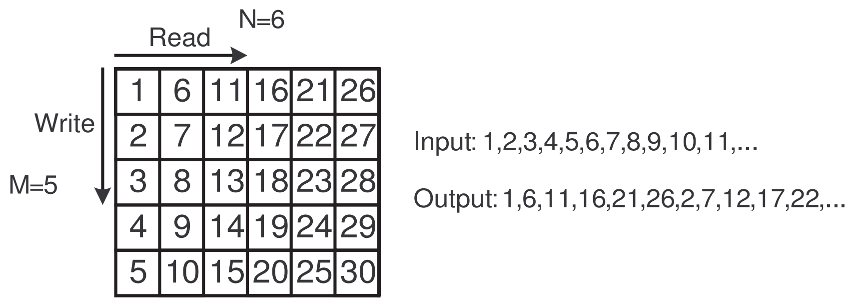

Figure 2.5: 5x6 block interleaver.

scribed in section 2.2. The two most often used procedures for performing the interleaving are called *block* and *convolutional* interleaving.

A block interleaver can be considered as an M-row-by N-column matrix. Usually the permutation of symbols is obtained by filling the matrix column by column. Also other ways to fill the matrix are possible, as is the case for example for the DVB-T2 bit interleaver described in section 6.1. Once the matrix is completely filled, the data is read out from the matrix row by row. The deinterleaver performs similar processing in reverse order, that is, data is written row by row and read out column by column. Fig. 2.5 illustrates the permuting operation of the block interleaver for $M = 5$ and $N = 6$ block size. The entries in the matrix indicate the order in which the symbols are placed in the interleaver. If it is assumed that a code of length $n \leq M$ is applied, then burst of length N in the channel introduces one or zero errors in each codeword. The end-to-end delay for the interleaver deinterleaver pair is $2MN$ symbol times. The amount of memory required in both transmitter and the receiver is MN symbols. Usually, in the receiver the memory needs to be doubled, because in the same time the current interleaver block is being emptied the next may need to be filled already [18].

Convolutional interleavers have been proposed by Ramsey in [30] and Forney in [29] in the early 1970s. The operation of the convolutional interleaver is easiest to describe with help of Fig. 2.6. The symbols are fed sequentially into the bank of N registers. Each successive register is J symbols longer than the previous one [18]. The first register is not actually a register, there is no storage and the symbol is output directly. Every time a new code symbol is read in, the switch turns to the next register and the code symbol is clocked into the register while the oldest one is clocked out. After the last register, the switch turns back to the first one and the process is repeated. The performance of the convolutional interleaver is reported to be very similar to that of the block interleaver [18]. The advantage of the

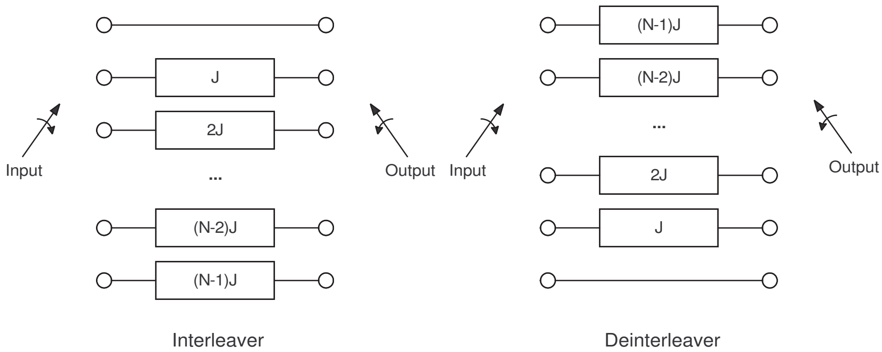

Figure 2.6: Convolutional interleaver.

convolutional interleaver is that its end-to-end delay is $M(N-1)$ and the memory required in the transmitter and the receiver is $M(N-1)/2$, where $M = NJ$ [18]. It is observed that both delay and memory requirement are halved compared to the block interleaver. Naturally the interleaver and deinterleaver need to be carefully synchronized for proper operation.

Chapter 3

Design of digital broadcast systems

When designing a broadcasting system, or as well any communication system, the first thing to consider is whether there is a real need for such a system. It is pointless to expend considerable effort to design a system that will never be used. This analysis task may rather be done by commercially oriented people than engineers themselves. For example, the need to create digital television standards stems from the fact that radio frequencies are an expensive and limited resource. With help of digital signal processing and compression, several television channels can be transmitted in the same frequency bandwidth as one analog television channel. Clearly, there is an advantage for the service providers.

After a need is identified, the requirements for the system to be commercially viable should be found. For example in the DVB organization that generates standards for media delivery, there is a special group called *commercial module* (CM) for defining the commercial requirements for the systems. This group defines the guidelines for the system design in form of requirements the system needs to fulfill from the commercial point of view. These requirements can state for example what kind of services should be carried, what should the maximum cost for the receiver chip of the system be, in which environment it is to be used, are there backwards compatibility requirements and so on. To continue with the digital television introduction example, the customers, that is the television audience in this case, would have to buy new equipment to receive the digital television transmissions. Therefore the maximum cost for the receivers is an important issue. Also, the users accustomed to the services and quality of analog television expect to see at least the same services and quality and more with the new system. Once the commercial requirements are defined, the technical design to fulfill the requirements as will be described in section 3.1 can begin. The process of the development of broadcasting systems is also studied in detail by Himmanen in [7].

The use of modeling and simulation in the design process is studied in section 3.2. The design of coding, interleaving and modulation are examined in more detail in section 3.3. The purpose of this chapter is not to go through all possible choices the designer needs to make in the communication system design. The purpose is rather to go through the design process and the most important trade-offs that are available in the design of a broadcasting system concentrating on the lowest two OSI layers, namely physical and data link layers. These trade-offs and design process are further elaborated with respect to DVB-T2 system in chapter 6.

3.1 System design process

The system design process is an iterative and recursive problem solving process [31]. During the process the needs and requirements for the system are transformed into system and process descriptions. The system design process can be divided into four phases as described in [31]: *Requirements analysis*, *Functional Analysis and Allocation*, *Design Synthesis* and *Verification*. The relationships between the phases are visualized in Fig. 3.1.

In the requirements analysis process the functional and performance requirements are developed for the system. This means translating the customer needs into requirements that define what the system must do and how well. Also design constraints (such as cost or backwards compatibility issues for example) need to be identified and taken into account in defining the requirements. This is the area in which the DVB commercial module operates. The developed requirements should be unambiguous and complete, so that there is no room for misunderstanding them. The requirements lay the foundation for the design process as the main purpose of the system design is to transform the requirements together with constraints into a system.

In the functional analysis and allocation process the functionality of the system is studied and split into smaller functional modules that together perform the overall functionality of the system. The performance requirements for the smaller modules are identified from the performance requirements of the whole system. The process of finding the requirements can be repeated to find architectures of ever increasing level of detail. During the functional analysis the designer's understanding of what the system should do and how it can do it together with the relationship of the smaller modules to one another is clarified. The relationships of the modules and the interfaces between them are studied. For the purposes of functional analysis and allocation, functional block diagrams and timing schedules for the system

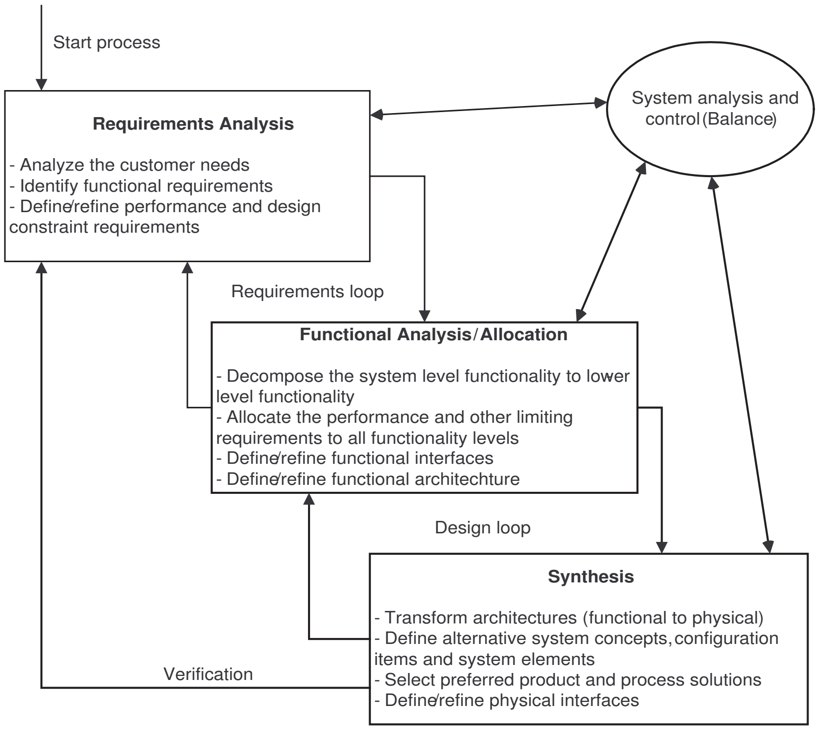

Figure 3.1: The system design process, modified from [31].

are defined. The result is a functional architecture of the system that is a description of the system in terms of functions and performance measures. Once the analysis is complete and the functional architecture is defined the requirement analysis should be reconsidered based on the findings of the functional analysis and it should be verified that each function relates to the requirements.

In the design synthesis the concept or actual physical designs are created based on the functional descriptions. The output of the synthesis is referred to as a physical architecture. After the synthesis process the result will be compared to the functional requirements to check that the required functions are fulfilled and the performance requirements are met by the design. This comparison helps also in optimizing the synthesized design. During the synthesis, such design issues may be encountered that the functional requirements need to be re-evaluated. It is possible that several alternative designs for the system and modules are generated during the synthesis phase. Out of these, the best option should be selected based on the most relevant criteria (cost for example).

Modeling techniques are often used during the synthesis. With help of modeling the performance and behavior of the system can be estimated before making final design decisions. Also, knowledge of the system based on the models may assist the designer in finding better solutions for the independent modules of the system.

Once the synthesis process is ready, the system should be verified. This is done by comparing the design to the requirements and constraints for the system. One important method for verification is *simulation*, where the system is modeled by a computer software. With simulations the performance of the system can be estimated to make a comparison to the performance requirements before physically manufacturing the device. Also, when implementing the simulator, design issues requiring changes in the design may come up. Thus, already implementing a simulator is a form of verification for the designed architecture. Once the design is verified by simulations, a hardware prototype of the design is produced.

3.2 Use of modeling and simulation in design and verification

Modeling and simulation are widely used tools for the communication system design. System modeling represents mapping of a real system or its prototype into a definition that can be used to study the system characteristics. In addition to studying the system characteristics, modeling used in combination with simulations is an effective tool for system design. Models of different abstraction levels and accuracy can be created for the systems based on the purpose of the model. Generally, a system is represented by a chain of models for lower level functional blocks of the system. Ideally, the designer would like to create models that exactly replicate the real system. For most situations the exact models would be if not impossible to implement, at least too costly in simulation complexity or run time [32]. Therefore simplifications to models are often necessary resulting in loss of accuracy for the models compared to the real system. For example in modeling the receiver design in a communication system to study the performance of error control coding in the transmitter-receiver chain it is common to assume that the receiver is perfectly synchronized to the received signal. Therefore, the simulator does not need to implement the receiver synchronization algorithms.

Next, let us consider how the simulations can be used in the design of a communication system. Suppose a system functional block diagram shown in Fig. 3.2 is specified by the functional analysis. It is known from the user requirements that the system is to be used in terrestrial static environment. Therefore the channel-block should present a respective scenario. Let us fur-

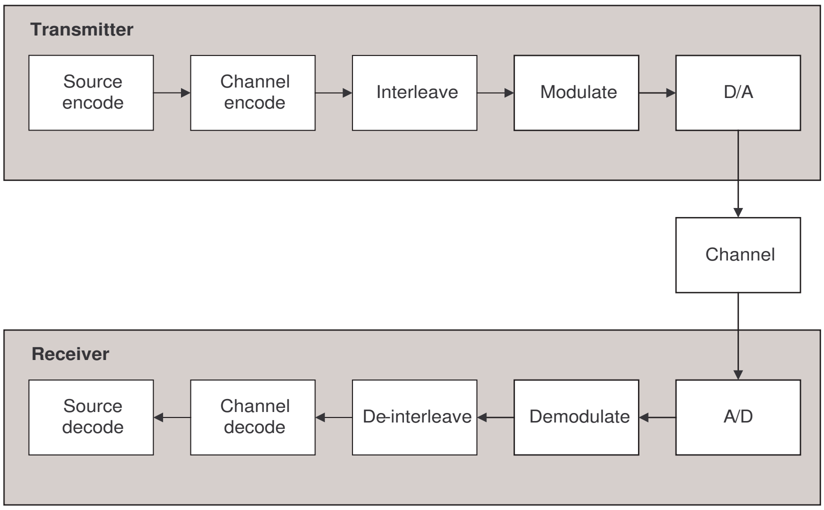

Figure 3.2: Block schema of an exemplary communication system.

ther assume that algorithms for transmitter and receiver blocks other than channel coding are fixed. Now the designer is interested in what channel coding algorithm should be used for the system. The designer simulates the different alternatives for channel coding algorithms (convolutional or turbo for example) and checks whether their performance in the presence of a channel required by the user requirements fulfills the performance requirements and further ranks the algorithms based on their performance. Once the best code is found out on the performance basis, the designer should calculate the complexity of the different schemes and check these against the complexity and timing constraints for the design. Usually, the best algorithm option from the performance point of view is not the best option from the complexity point of view. Therefore, the designer needs to decide what compromises can be made.

The different modules in the system are often designed by different engineers to meet the functional and performance requirements. Once initial designs for the modules are ready, they can be put together in a simulator and the performance of the system can be evaluated. With the help of the simulator the performance of the modules can be optimized based on the performance of the whole system instead of basing the optimization on one module at a time. This is important as a combination of "optimal" modules may not result in an optimal system from the performance or complexity point of view. For example, if different designers have designed modules incorporating interleaving there is a risk that once the modules are put together the different independently optimal interleavers may even counteract each other resulting in non-optimal performance of the whole system.

Suppose now that a receiver design is present for the static user scenario. Next, it is found out for example in studies of consumer habits that the customers would be willing to use the same services with portable receivers. What the designer can do, is to change the channel block in the simulator to one corresponding to the portable scenario. The performance of the receiver in this scenario is studied and possibly necessary changes in the design are identified. These changes can be implemented in the simulator and the performance checked against the requirements for portable reception. Considering costs versus the gains of the changes, the designer judges their viability. This is clearly an advantage of modeling and simulation; the engineer can "prototype" between different ideas before implementing the costly hardware prototype of the design.

In communication system manufacturing, usually both transmitter and receiver need to be available for successful testing of implemented hardware. In telecommunications industry, however, the same manufacturer does not necessarily produce both transmitter and receiver systems. In broadcasting systems it is common that the transmitter is standardized and the receiver can be implemented in various ways to receive the standard accordant transmission. Simulations can be used for the transmitter design to verify that the transmitter produces the correct waveforms at its output. Also, having a reference model for the receiver available at the time of transmitter design guides the designer's choices towards alternatives that assist the receiver performance and implementation. The receiver manufacturer does not necessarily need to implement the transmitter just for receiver testing purposes, it is possible to use recorded test sequences. Also, when an exact (bit-true) simulation model for the transmitter is available, it is possible to generate test streams for the receiver implementation in a cost efficient manner by the simulator.

There are several different ways to perform a simulation. The standard is so called Monte Carlo simulation, in which the simulation is very close to being the software counterpart of the actual system [32]. Monte Carlo simulations are based on repeated random sampling. Monte Carlo simulations are especially useful when it is very hard or even impossible to compute the system performance using deterministic algorithms. For example wireless communication channels are often simulated using statistical channel models. The simulation model for a communication system most importantly contains implementations for the transmitter, transmission channel and receiver. Bit-true models can be created for the transmitter and receiver that mimic the actions of the actual system so that bits in the simulator correspond to the bits that would go through the actual system. The accuracy of the Monte Carlo simulations depends on the accuracy of the models. For

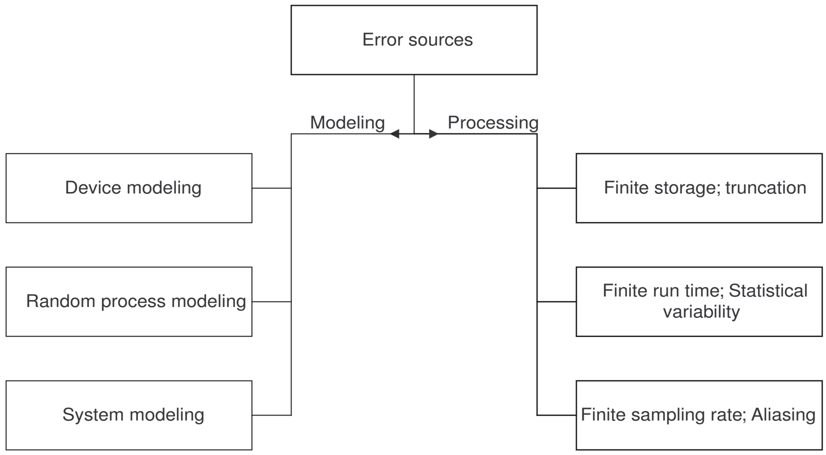

Figure 3.3: Sources of error in simulation [32].

high precision models, the Monte Carlo simulations tend to be highly time-consuming. To reduce the run time for the Monte Carlo simulations some parts of the models can be handled using analytical techniques. This kind of hybrid simulators are referred to as *quasianalytical* simulators [32].

3.2.1 Error sources and limitations in simulations

The results obtained with a simulator should not be taken directly as the final truth. There are several sources of errors and inaccuracy in the simulators that should be considered to have an idea of the accuracy and reliability of the obtained results. The sources of errors or inaccuracies in simulations can be divided into two categories of modeling and processing errors, as shown in Fig. 3.3 presented in [32].

Errors or inaccuracy deriving from system modeling relate to the "intentional" errors caused by the fact that not necessarily all functional blocks of the system are incorporated into the simulation system (such as synchronization for example). Functional blocks may be omitted from the simulations as they are considered to have only a small effect on the overall performance of the system. Of course, there is room for human error here, if the effect of a block on the system performance is greater than expected.

Device modeling errors stem from the fact that the blocks of the simulator are not exactly like the blocks of the actual system. In theory it is possible to create exact simulation blocks for discrete system blocks, but as soon as physical phenomena of the device are simulated the match is not exact any more. Ultimately the goal of modeling and simulation is to generate models that are close enough to the real system generating tolerable amount of error. What is tolerable, depends on the system itself and what should be learned out of the simulation results. Further, the cumulative effect of the error induced by different blocks of the system is hard to predict. Therefore, verifications of the simulator at different abstraction levels is necessary.

The error in random process modeling is an outcome of the differences of the behavior of the random process and the actual process it is modeling. Random processes used in the simulators relate to the signal and noise sources.

Processing errors stem from the restrictions of the simulation platform such as amount of memory, computational power and precision. Simulators running on computers use discrete-time and discrete-value representations of signals due to the discrete nature of computers. The sampling of analog signals for example presents some inaccuracy to the signal. It is well known that if an insufficient sampling frequency is used, frequencies in the signal are aliased to lower frequencies [33]. Further as the values of the signals at discrete time moments are quantized, the precision of presenting the values has an effect on the processing error. The amount of memory limits how much data can be handled by the simulator at a time. The memory limitation demands for example truncating infinite impulse responses in the simulations resulting in processing error. The computational power of the simulation platform limits the length of the simulation that can be performed within reasonable time. To obtain reliable results with Monte Carlo simulations where the measure of interest is often a random variable, it is essential to perform simulations that are long enough to eliminate the statistical variability of the results. Running too short simulations thus results in uncertainty of the results. To obtain reasonable uncertainty in bit error rate simulations for certain error rate p_b, 10 to 100 error events should be recorded [32]. Therefore, the amount of simulated bits should be in the order of $10/p_b$ to $100/p_b$. For example to obtain reliable results at BER 10^{-6}, the simulation should process 10^7 to 10^8 bits total. When simulating packet error rates, where the packets consist usually of several thousands of bits, also 10 to 100 error events (that is packets this time) should be obtained. This leads to increased simulation time as compared to the bit error rate simulations.

3.3 Design of coding, interleaving and modulation at the physical layer

The technical design process of communication systems consists of the strive to create a system that fulfills the requirements set initially by the user. Ideally, the designer tries to create the best possible system taking into account the limitations of information theory. This ideal must be loosened since there are constraints imposed by the system complexity and QoS requirements present. Due to the restrictions that can sometimes contradict the requirements, the technical design process consists of trade-offs between different ideals and restrictions. The technical design of a communication system contains at least the following goals presented in [18].

1. Maximize bit rate

2. Minimize error probability

3. Minimize required power

4. Minimize required bandwidth

5. Maximize system utilization, that is, provide reliable services to maximal number of users with minimum delay and maximal resistance to interference

6. Minimize system complexity, computational load and cost

Naturally all these goals cannot be met at once, which means that trade-offs between them need to be done. Clearly requirements 1 and 2 are in contradiction with 3 and 4 as the bit rate cannot be maximized while minimizing error probability, required transmission power (or signal to noise ratio) and bandwidth. Therefore, an appropriate balance needs to be found. Let us next go through the requirements for a broadcasting system and then study the trade-offs in the technical design phase for coding, interleaving and modulation in more detail as they are important building blocks to balance between the goals (1-6).

3.3.1 Requirements analysis for broadcasting systems

For broadcasting systems there are at least two types of customers, which are the broadcasters and the service consumers (receiver users). In the design of the system, the needs of both customers need to be taken into account. The first consumer need that sets requirements and constraints for the broadcasting system design is the environment of use. The environment is defined here to contain most importantly the transmitter configuration (satellite or terrestrial) and mobility of the receiver (static or mobile). The mobility of the receiver requires more robustness from the transmission system and also the battery powered mobile receivers set constraints on the power consumption. The broadcasters on the other hand require that the network costs should be as small as possible. This translates into as large coverage area for the transmitters as possible (with minimum power), cheap transmitters and possible utilization of existing infrastructure (transmission masts and delivery network for example).

Another requirement from the service consumers is the service types that should be broadcasted and what is the expected quality of service. The most important question is, what quality the user expects when using the system. The expectation of the service quality, on the other hand, depends on the usage scenario. For example, a viewer following a soccer game on HDTV television at home may feel more annoyed if there is a glitch in the video than the viewer following the same game on a bus with a mobile device with a small screen. All in all, it is obvious that the user having no interest in the underlying technology, is willing to see the best quality of service possible. For the design purposes the question should be considered the other way round, that is, what is the worst quality the user could tolerate while still being satisfied enough. This criterion sets the limits for the design as the designer knows that the system should do at least better than this. Considering broadcasting, the main types of service are *streaming* and *data* services. An example of a streaming service is a digital television channel and for data download it could be podcasts, that are downloadable media files. For these two types of services the QoS requirements are different.

Commonly used QoS requirements of the error rate for the streaming HDTV or Standard Definition TeleVision (SDTV) service is [6] “one visible/audible error in an hour”. For data download the requirement could be for example “average 5 minutes time to receive a certain file correctly”. In addition to the amount of tolerable errors the time for channel switching (or zapping) is important for the streaming service consumer. The channel switching time should be in the order of one second [7]. For the mobile viewers the requirement of the tolerable error rate for the streaming could be less strict, for example “one visible/audible error in a minute”. The bit rates required to provide good enough quality for different service types and also different usage scenarios are different. One HDTV service requires 5 Mbps to provide satisfactory audiovisual quality while mobile service destined for terminals with small display requires only 1 Mbps or less. Further, for data services, even lower bit rates may satisfy the required download time.

The designer translates the QoS requirements into performance requirements that can be used as a benchmark in the technical design as described in sections 3.1 and 3.2. Common performance requirements are based on bit rate, delay, delay jitter and error rate. Knowing the bit rate of the service and the effect of bit errors on the video codec the requirement of one visible/audible error in an hour can be translated into a certain bit error rate or maximum burst error length, that can be used when designing the coding and modulation for the system. Further, knowing the tolerable service switching time the maximal delay for the system can be defined assisting in the design of the interleaving and framing structure. In addition to interleaving, signaling has an effect on the channel switching time. If for example the receiver receiving a certain service obtains only the signaling necessary to continue receiving the current service, it takes a while to receive signaling for switching to the target service. Thus the delay requirement affects also the design of signaling transmission. The allowed delay jitter by the codecs guides the design of buffering in the system. In analogy, the specified criteria for data download services can be transformed into error rate, delay, delay-jitter and bit rate requirements.

As an example, the DVB-T2 system [6] defined to be mainly used in static reception conditions was designed against the Quasi Error Free (QEF) criterion of one error in an hour. This criterion corresponds with a 5 Mbps service to a Transport Stream packet error rate of approximately 10^{-7}. It is defined that 5 Mbps is enough to deliver satisfactory audiovisual quality to the customer. Transport stream packets are 188 byte packets carrying data. Assuming independent bit errors, the packet error rate can be translated into a target bit error rate of roughly 10^{-10} to 10^{-11}. The target service switching time at the physical layer used in the design of DVB-T2 is 0.5 seconds. De-jitter buffers are included in the system to minimize the effect of delay-jitter.

As an example of the mobile reception scenario, the performance requirement commonly used for design and measurements of a streaming service of the DVB-H system in a mobile usage scenario is MPE-FEC Frame Error Rate (MFER) smaller than 5% at the link layer. This error criterion is much less strict than the quasi error free criterion for the stationary reception of a streaming service. The MFER criterion is criticized for being too ambiguous [34], as one erroneous frame may contain a variable number of errors. Therefore other criteria such as Erroneous Seconds Ratio (ESR) have been defined for mobile systems.

3.3.2 Design trade-offs

Once the functional and performance requirements for the system have been considered, the more detailed functionality of the system is studied. Common functional elements for all current physical layer designs of broadcasting systems are forward error control, interleaving and modulation. They are essential for trading off between different inherent resources of communication systems. The general idea of trading off between the communication resources using coding and modulation is visualized together with the transmission capacity bound in Fig. 3.4 [18]. Notations G, C and F stand for "Gained", "Cost" and "Fixed" respectively. The parameters that can be traded are p_b (bit error probability or rate), R_I/W (bandwidth efficiency, information rate divided by bandwidth) and P (power or ratio of signal power to noise power). Radio frequencies for broadcasting are allocated in

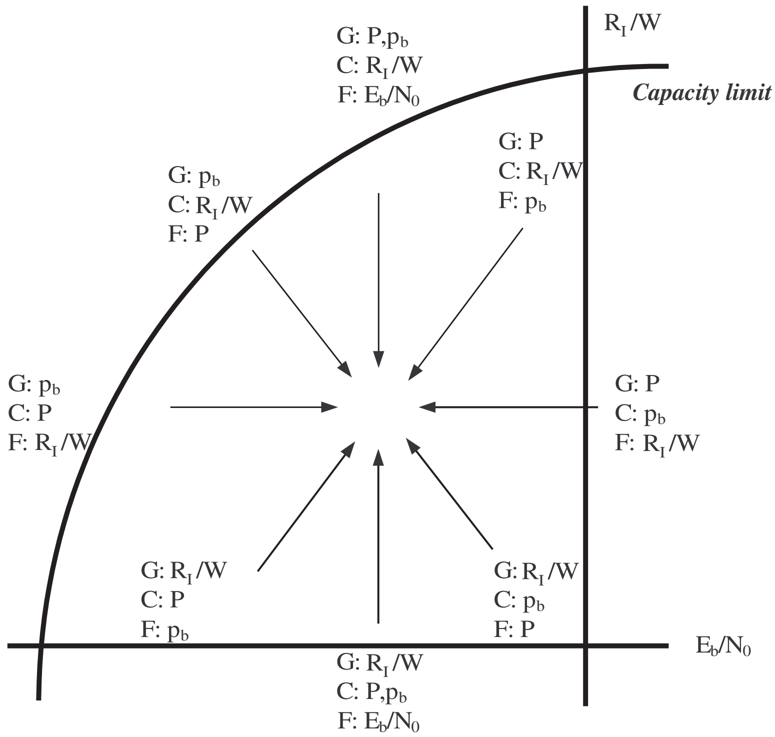

Figure 3.4: Trading off between communication system resources [18].

blocks of frequency space. The use of frequency space is regulated by ITU (International Telecommunication Union) and by governments. ITU arranges World Radio communication Conferences (WRC) every two to three years to revise the radio regulations.

As the radio frequencies are allocated in blocks, we assume here for the design purposes that the bandwidth W for the system is fixed. Also, as we have a desired error rate from the performance requirements, the p_b is considered fixed. Now, there are two arrows in the figure indicating that when p_b is fixed, trade-offs between P and R_I/W can be done. If the main target is minimizing the required power P, then a sacrifice on the bandwidth efficiency is necessary. As the bandwidth is fixed the R_I/W can be altered with coding and modulation. Therefore, to lower the power requirement, a stronger code (of lower rate) or lower order modulation (carrying less bits/symbol) is to be used. On the other hand if a gain on the R_I/W is necessary, a sacrifice on the power needs to be done. That is, if we want to have higher throughput with fixed bandwidth, we need to use higher rate codes, higher order modulation and consequently higher received power is necessary to reach the target p_b. Note that the E_b/N_0 is affected by both power P and used modulation method (number of bits in each modulation symbol).

The selection of the optimal FEC scheme depends on the performance characteristics of the scheme and on its suitability for the particular system approach. The selection between a block or trellis based coding scheme has an effect on the design of other system blocks. Selecting a block-code makes it natural to process data also in other system parts in blocks, while for trellis based coding continuous data processing is more natural. Also, rate-compatibility of the coding schemes should be considered if multiple code rates are to be supported. For example for systematic Reed Solomon codes the codewords can be shortened by setting as many first information symbols as necessary to zero and not transmitting them. Puncturing can be done by not transmitting a selected amount of parity symbols. For the LDPC codes used in DVB-T2 on the other hand, the shortening and puncturing must be done based on the specific structure of the used code to maintain the performance. For trellis codes, the puncturing is done according to selected puncturing patterns. In DVB-T/H the different physical layer code rates are obtained by puncturing the mother convolutional code of rate 1/2. When shortening and puncturing block codes, the codeword length varies with the code rate. If codewords of same length for different code rates are favoured in the system design, individual generator matrices for different rates are necessary, as it is done in the DVB-T2 systems. Further, the performance characteristics of different coding schemes can be simulated as described in 3.2 to find out the scheme with optimal performance matching the design approach.

The selection of modulation schemes depends heavily on the expected usage environments and their common characteristics. An optimal selection for the modulation scheme should provide low bit error rates at low signal-to-noise ratios, should perform well in multipath and fading conditions, should occupy minimum bandwidth and should be cost efficient to implement in the receivers [15]. All of these cannot be simultaneously satisfied. Important issues in modulation scheme selection are *power efficiency* and bandwidth efficiency. Power efficiency describes the ability of the modulation scheme to maintain the fidelity of the digital information at low received power levels. Bandwidth efficiency describes the ability of the modulation scheme to accommodate data to a given bandwidth. The required throughput sets constraints on the modulation order, that is the amount of bits carried on each modulation symbol when the bandwidth, symbol time and number of subcarriers for multicarrier modulation are considered fixed. The receiver cost for the demodulation is an important factor in the modulation scheme selection as the number of receivers in broadcasting networks is large. The final selection of modulation mechanism is in general based on simulations of the modulator and different demodulator implementations in presence of the impairments of the expected usage scenarios. Further, the trade-off between applying a higher order modulation and lower rate channel coding with different channel models can be studied with simulations. That is, if spectral efficiency of 2 bits/symbol is required it can be studied whether it is better to use for example Quadrature Phase-Shift Keying (QPSK) and no channel coding or 16 point Quadrature Amplitude Modulation (16-QAM) with a channel code of rate 1/2.

Interleaving has no effect on the R_I/W for fixed bandwidth but in the case of fixed p_b it enables to trade-off between the system delay and P in time or frequency variant reception conditions. The most widely utilized interleaving mechanisms are block and convolutional based as described in section 2.3.2. Convolutional interleaving is ideally suited for use with trellis (or convolutional) codes [15], while block interleaving is better suited for block coding. The required channel switch time sets the maximum delay for the time interleaver. For the frequency interleaver the utilized bandwidth sets the limits for the interleaving.

It is common to design a broadcasting system so that the same system may be used by both mobile and static users. Therefore broadcasting systems are often *configurable*. In a configurable system a "toolbox" of modes differing for example in modulation scheme, FEC code rate and interleaving configuration is available. For broadcasting systems this leads to configuring the network to transmit using a certain mode (or modes as in DVB-T2, this will be discussed in section 4.2.2), so that desired user coverage for selected user scenarios is achieved. If multiple configurations in the transmission are allowed (as in DVB-T2) the cost for introducing new services to new groups of receivers in the existing network is decreased. For example, the DVB-T2 system may start by transmitting only HDTV services for static receivers, but after a while services using a more robust configuration (meaning lower code rate, lower order modulation and longer time interleaving) may be introduced in the same network for mobile receivers. Due to the configurability of the system there is a need for signaling to inform the receiver on current transmission parameters.

3.3.3 The role of signaling

The QoS requirements for the L1 and L2 signaling need to be found out based on the requirements for the services at respective layers, as the signaling enables the reception of the services and thus affects their QoS. In general, the requirements for the signaling should be more demanding than for the services, due to the "enabler" nature of signaling.

The signaling can be divided into in-band and out-of-band signaling based on its transmission mechanism. In-band signaling is transmitted in the same channel as the information itself. Out-of-band signaling is transmitted in its own dedicated signaling channel. The advantage of the out-of-band signaling is that the signaling information for multiple services and networks can be centralized. Centralizing signaling information is useful in broadcasting networks, as the receiver can obtain necessary information for the reception of all services in the multiplex from one place instead of having to gather information from all the services. Naturally the in-band signaling could contain signaling information for all other services, but this would increase the overhead introduced by the signaling to the system as the same information would have to be repeated in each service. Both in- and out-of-band signaling can be used simultaneously in the same system as in DVB-T2.

The design of coding, modulation and interleaving for the out-of-band signaling is critical to enable robust and fast service discovery and switching in broadcasting networks. In terminals with for example 3G and broadcast receivers, it would in theory be possible to arrange the signaling channel in a hybrid manner through the interactive 3G packet data. In this thesis, however, it is assumed that the signaling channel is also a broadcast channel, arranged in the signal by for example time division multiplexing as in DVB-T2. For in-band signaling the coding, modulation and interleaving are the same as for the service data as the signaling is embedded in the service.

The signaling information at all system layers in broadcasting networks needs to be transmitted repeatedly to enable the receivers to connect to the network and start decoding the services. The modulation and coding design for the L1 signaling channel can be done similarly as for the data path. For the service discovery, the signaling information should be obtainable by the receiver without any prior knowledge except the presence of the signal. In the design this leads to a predetermined transmission mechanism, that is, fixed code rate, modulation, interleaving and mapping to framing structure for the out-of-band signaling. The performance criteria for the signaling are derived from the performance criteria for the data path. To enable the reception of the services the signaling should be at least as robust as the service it is signaling, preferably it should be more robust in the expected usage environment. This corresponds to a more strict error criterion for the signaling, that is, the signaling should reach the required error criterion at smaller Signal-to-Noise Ratio (SNR) values than the data path. For systems with seldom changing signaling information already the repetition enhances the robustness. For systems with rather variable signaling information such as DVB-T2 for example, the robustness gain of repetition is not necessarily present and thus lower order modulation or lower rate code than for the data path are necessary.

The delay limitation for the signaling derived from the service switch and discovery time sets the maximum repetition interval and interleaving length for the signaling. The throughput required by the signaling depends on the amount of signaling that needs to be transmitted for the receiver. The capacity that the signaling requires from the network is taken away from the transmission capacity of the main data, that is, the services. Therefore, compact signaling information conveying only the necessary information is desirable. Also, the repetition rate of the signaling has direct effect on the capacity requirement for the signaling. The design issues for the signaling in the special case of DVB-T2 system are further discussed in chapter 6.2.1.

Chapter 4

Digital wireless broadcasting standards

At the time of writing this dissertation in the beginning of year 2010, there are five families of standards for digital wireless broadcasting of television and other data services. Geographically the standards are roughly divided into European DVB-family, north American ATSC-family, Japanese ISDB-family, Chinese standards and South Korean DMB-family. There are also digital radio broadcasting standards such as DAB and Digital Radio Mondiale (DRM) that are used for broadcasting audio and data services. In addition to terrestrial broadcasting, also satellite systems are wireless, but this chapter concentrates on the terrestrial systems. Thus, when referring to wireless broadcasting, a terrestrial situation is considered. In this chapter, the current status of the systems used globally for terrestrial wireless broadcasting is studied first. Then a closer inspection of the DVB standards family is presented to gain insight into the systems studied in the case-study chapters 5 and 6.

4.1 Global situation of wireless broadcasting standards

The two most widely used wireless broadcasting standards are DVB-T and Advanced Television Systems Committee (ATSC). They are used mainly for the delivery of digital television to homes with fixed antennas. ATSC is used in North America and South Korea, while DVB-T is used in European countries, Australia, many Asian countries and some African countries. The third important broadcasting system is Terrestrial Integrated Services Digital Broadcasting (ISDB-T) that is used in Japan and South America. The official Chinese standard for wireless broadcasting is called Digital Terrestrial Multimedia Broadcast (DTMB) or sometimes Digital Multimedia

Broadcast for Terrestrial and Handheld (DMB-T/H). It is likely to have very large amount of users due to the large population of China.

In addition to transmitting services to fixed receivers, there are currently nine more or less competing standards that are designed for broadcasting services to handheld devices. In Europe the main system used is the DVB-T based DVB-H (Handheld) and probably DVB-SH (Satellite Handheld) in the future. The European Commission has endorsed DVB-H to be the preferred standard for digital TV broadcasting to mobile phones or other handheld devices. There are also some countries in Europe that use other standards than DVB-H. Norway, for example, has decided to use the South-Korean T-DMB for mobile broadcasting.

In north America, the main mobile standards are ATSC-M/H and MediaFLO. ATSC-M/H is a new standard based on the basic ATSC used for terrestrial and cable digital television. Differing from many other mobile and terrestrial standards, it uses a Vestigial Side Band (VSB) transmission scheme instead of OFDM. MediaFLO is a proprietary system, which resembles DVB-H on many system parts.

In Japan, ISDB-T is used for delivery of terrestrial digital television. The ISDB-T transmission is divided into 13 segments. Commonly, television services (HDTV or SDTV) occupy 12 segments and one is left for the mobile services. Therefore, the name of the Japanese mobile broadcasting is 1seg.

In China the CMMB system is used for mobile broadcasting. Its operating philosophy is similar to the DVB-SH system, where terrestrial transmitters are used as "gap-fillers" for satellite network coverage. The CMMB standard has real potential to become the most widely used of the mobile broadcasting standards, due to the large population of China. Also, if the receiver markets become large the prices may fall and also other countries than China may find the CMMB an appealing alternative. Also, DTMB is used in China. It is the official digital television standard in China that is supposed to serve both mobile and fixed terminals.

In South-Korea T-DMB and S-DMB are used. The T-DMB system utilizes the DAB standard for data transmission. The T-DMB services are free to air, while there is a price for satellite S-DMB services. A large share of new cellular phones sold in Korea contain a T-DMB receiver. Therefore, the services are popular and consequently they can be financed by advertisements. The coverage area of the T-DMB services is limited to large cities and the larger coverage area of the satellite based S-DMB is the reason why also S-DMB services are quite popular.

In conclusion, the technology field for fixed and mobile television systems is very fragmented. It is quite unlikely that some standard would "win the war" and become the de facto standard of the whole world for both usage scenarios. Still, from the user point of view this would be a nice feature for at least mobile standards. When traveling from one country to another the same device could be used for watching television everywhere.

4.2 Family of DVB standards

The Digital Video Broadcasting (DVB) project was established in 1993 when broadcasters, equipment manufacturers and regulatory bodies in Europe gathered to prepare the introduction of digital television. Since then, the DVB project has played an important role in developing and standardizing digital broadcasting systems that are used in Europe and also in other parts of the world.

The core principle in the work of the DVB project is that the commercial module figures out the market needs for a certain system. Usually, this work is preceded by a study mission to identify possible new technologies that can be used in a future system. Once market a need is obvious, the CM generates commercial requirements, specifying in more detail what the characteristics of the system should be from the commercial point of view. After this the technical module (TM) works to create a technical specification to meet the commercial requirements. After the technical work is done, the CM reviews the technical specification and then it is sent to the DVB project's steering board for approval. Finally, the specification is sent to formal standardization (for example to the European Telecommunications Standards Institute ETSI). The important rule in DVB work is that the technical work is commercially driven by the requirements of the member companies [35], [36]. The member companies represent industry related to the standards such as transmitter and receiver manufacturers, service providers, measurement system manufacturers and so on.

The first main physical layer standards that the DVB project generated in the 1990s were DVB-S (Satellite) [37], DVB-C (Cable) [38] and DVB-T (Terrestrial) [2]. These three standards are currently in use in the delivery of digital television in many parts of the world. In addition, several related higher layer standards are used on top of these, for example for transmitting subtitles or service information [39] that is called layer 2 (L2) signaling in this thesis. Further, based on DVB-T the DVB-H (Handheld) [4] standard was specified to enable the broadcasting of digital services to handheld devices. Currently, the second generations for the main standards DVB-S2 (Second generation Satellite) [40], DVB-C2 (Second generation Cable) [41] and DVB-T2 (Second generation Terrestrial) [6] are being developed and deployed. The main driver for the second generation has been the introduction of HDTV. Also, the analogue switch-off has released frequencies that should be utilized as efficiently as possible. Advances in technology have enabled the use of such mechanisms in the second generation that were considered in the time of design of the first generation systems to be too costly for consumer products [35]. For example, the LDPC coding to be decoded with iterative belief-propagation algorithms is used in the second generation systems. Had these decoding algorithms been available in the time of specifying the first generation systems in the 1990s, they would have been too computationally complex for products of that time.

The principle with the satellite, cable and terrestrial physical layer standards has been that the standards should have as many commonalities as possible, hence they are called the family of standards. For example, the first generation physical layer standards differ mainly in the modulation method to suit the systems in different transmission environments but for all of them the input is an MPEG-2 transport stream. For DVB-S and DVB-C, single carrier transmission is used, while DVB-T uses OFDM multicarrier transmission that is better suited for the terrestrial environment. Also the second generation systems have clear similarities in the transmitter processing to ease the interoperability between the different standards. In this thesis, the main focus is on the terrestrial standards DVB-T, DVB-H and DVB-T2. They are in general technically the most advanced of the standards, as they need to mitigate the harmful effects of the challenging terrestrial transmission channels. Let us next go through the main characteristics of the terrestrial standards.

4.2.1 First generation: DVB-T and DVB-H

The design of the first generation DVB systems was driven by the need to digitalize the television transmissions for more efficient frequency spectrum utilization than with analog transmission. Further, support for country-wide Single Frequency Network (SFN) operation was desired for optimal utilization of the available frequency spectrum. At the time of the design of the terrestrial system, standards for cable and satellite transmission of digital television were available and the introduction of the terrestrial standard was a natural continuation to make the family complete. This also led to a requirement of maximal level of commonality with the DVB-S and -C standards [42]. The first version of the DVB-T standard [2] was approved in 1995 by ETSI. Since then it has been adopted in many countries as the bearer of SDTV services. Still, thanks to the flexibility of the standard, some countries such as Norway, Singapore, France and Estonia transmit also HDTV services over DVB-T networks [43]. The DVB-T system was originally designed mainly for delivery of digital broadcast for stationary receivers with rooftop antennas. Portable reception was also considered desirable but not mandatory, as it was not mentioned in the commercial requirements [42]. Trials with mobile reception have shown that the flexibility of DVB-T enables reception also in other scenarios than the stationary one.

Due to adequate performance in some mobile scenarios, DVB-T is used also for broadcasting services for mobile receivers for example in cars.

The DVB-H standard designed for mobile handheld receivers was adopted as an ETSI standard in 2004. In the trials of mobile reception of DVB-T it was found that the system should be enhanced to enable mobile reception with handheld devices [44]. The handheld property of the receiver sets requirements that are not directly met by the DVB-T standard. The enhancements were seen to be required to enable low power consumption for battery-powered devices, increase flexibility in the network planning, enable handover when moving around in an Multiple Frequency Network (MFN), improve the performance in mobile channels with a small and possibly not optimally oriented reception antenna and add compatibility with IP networks. DVB-H can be considered as an update for DVB-T to support handheld reception. Most important updates for DVB-H are the addition of time interleaving to combat challenging transmission channels and time slicing (that is bursty transmission) to save the receiver battery. In addition to power saving the time-slicing facilitates handover. The changes were designed in a backwards compatible manner, so that the physical layer would be almost the same as in DVB-T. This enables sharing the transmission networks between the DVB-T and -H.

Physical Layer

The DVB-T as well as DVB-H system physical layer consists of the following processes that are applied to the data stream [2]:

- Transport multiplex adaptation and randomization for energy dispersal
- Outer coding (Reed-Solomon(204,188) code)
- Outer interleaving (convolutional interleaving)
- Inner coding (punctured convolutional code)
- Inner interleaving
- Mapping and modulation (QPSK, 16-QAM and 64-QAM)
- OFDM

The block schema of the DVB-T standard together with the relationship of the DVB-T to the DVB-S standard is illustrated in Fig. 4.1. It is seen that mostly the concatenated and interleaved FEC scheme is inherited from the DVB-S system. The input to the DVB-T or DVB-H system physical layer is fixed length MPEG-2 Transport Stream (TS) packets [45]. The packets

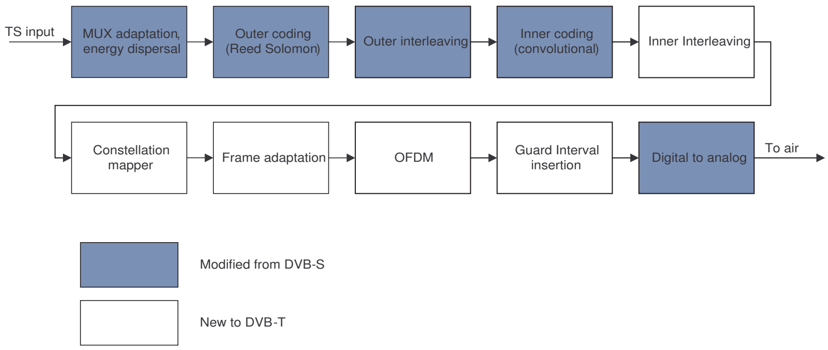

Figure 4.1: Block schema for DVB-T/H with relationship to previous DVB-S standard.

are randomized by a Pseudo Random Binary Sequence (PRBS) for energy dispersal. Outer encoding with the Reed-Solomon (RS) code with length matched to the length of the TS packets is performed. The RS encoded packets are interleaved to randomize the bit stream for the inner convolutional code. The rate 1/2 mother convolutional code is punctured with predetermined puncturing patterns to obtain different code rates (1/2, 2/3, 3/4, 5/6, 7/8) according to the robustness needs. The inner interleaving assures that subsequent bits from the convolutional encoder are not mapped to the same modulation symbols in the mapping and modulation. The modulation schemes defined for the DVB-T system are QPSK, 16-QAM and 64-QAM. The modulation schemes carry 2, 4 and 6 bits in one modulation symbol respectively. Finally the modulated symbols are carried in different subcarriers of the OFDM system. In addition to data, the OFDM subcarriers carry physical layer (L1) signaling information, called Transmission Parameter Signaling (TPS) and pilot symbols that are used for estimating the transmission channel and synchronization in the receiver. Two different OFDM transmission modes named after the Discrete Fourier Transform (DFT) size, namely 2k (2048) and 8k (8192) are defined. The Fourier transform and inverse transform are commonly performed using the widely known Fast Fourier Transform (FFT) algorithm [46]. The size of the FFT defines how many subcarriers there can be in the OFDM symbol. The OFDM symbol is formed in frequency domain and inverse transformed into time domain for transmission. To reduce the effect of intersymbol interference in multipath and SFN environments, a guard interval also known as cyclic prefix is inserted into the OFDM symbols.

The physical layer of DVB-H is relatively similar to DVB-T with only optional additions. Therefore, DVB-T and DVB-H can co-exist in the same

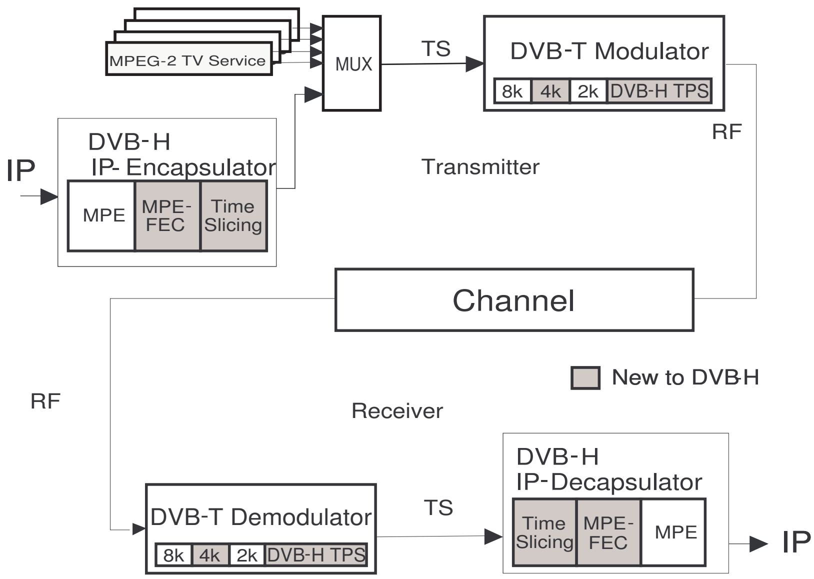

Figure 4.2: A conceptual description of the DVB-H system [4].

transport multiplex if these additions are not used. The relationship between the DVB-T and -H is presented in Fig. 4.2. The additions to the physical layer contain an additional 4k (4096 carriers) transmission mode, an in-depth interleaver that can be used for 2k and 4k modes and an extension of the TPS information to support DVB-H systems. The additional 4k mode is intended to provide more flexibility to network planning being a compromise between the large SFN size of the 8k mode and robustness in mobility of the 2k mode. The in-depth interleaver performs interleaving over 2 and 4 OFDM symbols for 4k and 2k transmission modes increasing the time diversity to the same level with 8k transmission mode without the need for extra interleaving memory. The TPS is extended to carry bits signaling the use of the new 4k transmission mode, in-depth interleaver and the use of Link Layer MPE-FEC and Time slicing discussed in the next section.

Link Layer

The input to the DVB-T physical layer from the link layer is a stream of MPEG-2 Transport Stream (TS) packets of constant length 188 bytes. In DVB-T the services are multiplexed into a continuous stream of TS packets. The TS packet header field Packet IDentifier (PID) is used to identify different elementary streams carrying different services within the transport stream. L2 signaling, also called Program Specific Information and Service

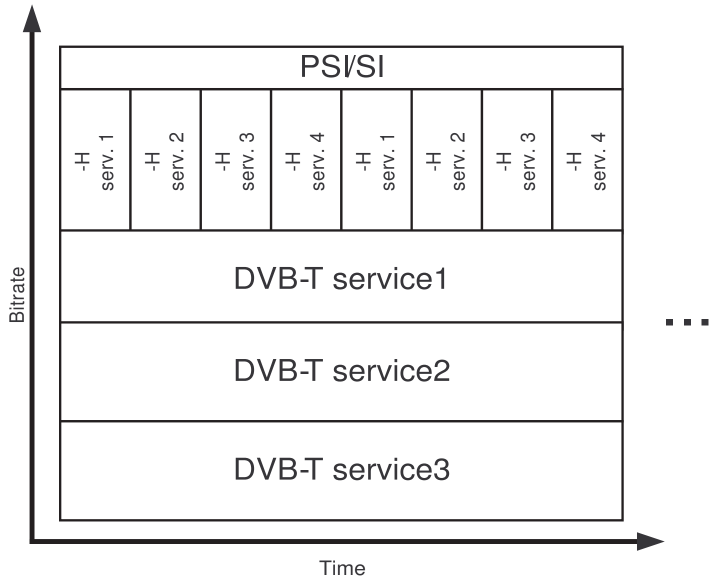

Figure 4.3: Multiplexing of DVB-T and DVB-H services.

Information (PSI/SI) signaling, is used to inform the receiver about the network structure and services, for example, which services are found on the elementary streams with different PID values. Therefore, the PSI/SI plays an important role in service discovery and reception. The performance and transmission mechanism for the PSI/SI is studied in section 5.2. In order to receive a single elementary stream the receiver needs to filter the TS packets based on the PID.

The link layer is the main improvement of the DVB-H over DVB-T. The DVB-H services are IP-packet based. Therefore, the IP packets are encapsulated in Multi-Protocol Encapsulation (MPE) packets for transmission in TS packets as defined in [47]. The additional features at the link layer include time slicing and one more stage of error correction called the Multi-Protocol Encapsulation - Forward Error Correction (MPE-FEC). Time slicing means that the transmission is time division multiplexed, that is, one service is sent in bursts separated in time. Therefore the multiplexing of the services is done in a different way than for the DVB-T. This difference is illustrated in Fig. 4.3 with an exemplary configuration of a signal carrying three DVB-T and four DVB-H services. The power-saving is achieved due to the fact that the DVB-H receiver can switch off some of the power consuming radio components between the bursts of the service it is following. If the receiver is following service 1, it can go to sleep mode for the transmission time of the services 2, 3 and 4 while the DVB-T receiver needs to be on all the time for receiving one service. It is important to note that the DVB-H receiver performing time slicing receives updates for the PSI/SI only during the on time of the burst. The time slicing parameter *delta-t* indicating the time to the beginning of the next burst is embedded in the MPE-packet header in each burst. The MPE-FEC utilizes a RS(255,191) code combined with *virtual* time interleaving to combat channel fading. The term virtual is used because the data itself is not interleaved, only the redundancy is calculated in an interleaved manner. The MPE-FEC and time-slicing are closely related as exactly one MPE-FEC frame is transmitted in one time slicing burst. The properties and performance of the MPE-FEC will be further studied in chapter 5.1.

In addition to power saving, the bursty transmission enables the receiver to perform seamless handover [48]. During the time between the bursts of interest the receiver can probe the neighbouring frequencies transmitting the same service. The neighbouring frequencies are signaled in PSI/SI. If a better signal is found from another frequency, the receiver can switch to receive the service from that frequency without the user necessarily even noticing that a handover has happened.

4.2.2 Second generation: DVB-T2

The development of the second generation terrestrial digital video broadcasting standard DVB-T2 [6] started in 2006 when DVB initialized a study on how the DVB-T standard could be upgraded. Several appealing new technologies had become available after the introduction of the first generation standards and potential for substantial enhancement was identified. The commercial requirements for the DVB-T2 [49] indicated that a certain degree of backwards compatibility to DVB-T was required. It stated that the reception of DVB-T2 should be possible with the current antenna installations for DVB-T. Therefore, the static reception scenario was the main target, but the system should also support services targeted for portable and mobile receivers. Otherwise, the requirements gave the designers rather free hands. From the family of standards point of view, as many existing parts of the standards family as possible should be utilized in the design. Therefore, similarities with the first second generation standard DVB-S2 are visible. The main requirements steering the technical design given by the commercial requirements are as listed in [7]:

- Provide a capacity increase of at least 30 % compared to DVB-T
- Provide improved SFN performance
- Provide service specific robustness

- Provide means to reduce peak-to-average power ratio
- Provide bandwidth and frequency flexibility

Once the commercial requirements were created for DVB-T2 the technical design for the next generation terrestrial system was initialized. First the technical module issued a Call for Technologies (CFT) [50] that asked for proposals for the technologies to be used in the new system. The proposals were to incorporate simulation models and performance evaluation if possible to assist the technology selection process. Further, these models for individual system modules were incorporated in the common simulation platform for the whole system to study the overall performance of the system. This illustrates a practical use of simulation in the design process as described in Chapter 3.2. The first version of the physical layer standard was finalized in 2008. Let us next go through the characteristics of the DVB-T2 system and examine how these requirements are met by the design.

Physical Layer

The DVB-T2 standard [6] as its predecessor DVB-T is based on OFDM as the transmission scheme. In addition to the FFT modes defined in DVB-T (2k and 8k), the DVB-T2 standard defines 1k, 4k, 16k and 32k FFT sizes. Having a larger FFT size increases the number of carriers in each OFDM symbol and consequently lengthens the duration of symbols as the sampling frequency is the same for all FFT sizes in DVB-T2 when the bandwidth is kept constant. On the other hand, the carriers are packed more densely in the spectral domain increasing vulnerability to Doppler effect present in a mobile environment. In DVB-T the overhead introduced by the pilot carriers and guard interval is large, as already the pilots introduce an overhead of order of 10 % [51]. DVB-T2 introduces new pilot patterns and guard intervals that can be used to reduce the overhead and thus increase the useful data rate for the services. The less dense pilot patterns reduce the performance of the channel estimation for demanding reception conditions, but provide good enough performance for static reception conditions. The quard interval has an effect on the SFN performance of the system. For example if a guard interval fraction of 1/4 is required for proper SFN operation with 8k mode, the guard interval fraction to obtain the same absolute duration for the guard interval for 32k mode is smaller. Thus, the overhead introduced by the guard interval is smaller for 32k. Consequently, if the same 1/4 fraction for 32k were used, the SFN performance would be increased as compared to the 8k mode. For further increase in data rate 256-QAM is introduced.

One significant architectural change as compared to DVB-T is the introduction of Physical Layer Pipes (PLPs) that enables service specific robustness. Using the physical layer pipes Time Division Multiplexing (TDM) of the services can be performed and only the wanted PLP can be be received at a time. This is different from DVB-T where the whole multiplex must be received, but resembles the DVB-H transmission with time slicing. The difference to DVB-H is that the time slicing is performed at the physical layer instead of the link layer. Further, in DVB-T2 optional Time Frequency Slicing (TFS) is defined. In TFS one time sliced PLP is transmitted on different radio frequencies in different bursts. Each PLP can have its own modulation mode (QPSK, 16-QAM, 64-QAM or 256-QAM) and code rate (1/2, 3/5, 2/3, 3/4, 4/5 or 5/6). Optionally for each PLP constellation rotation can be used. In constellation rotation the constellation is rotated and imaginary and real parts of the modulation symbols are transmitted on different OFDM subcarriers. Once this is done, iterative demodulation in the receiver can be performed. Constellation rotation is reported to enhance the performance of the system in challenging channel conditions.

Forward error correction coding in DVB-T2 is based on a similar mechanism as in DVB-S2 consisting of concatenated LDPC and BCH (Bose Chaudhuri Hocquenghem) coding. Some of the codes defined in DVB-S2 have been changed to match the terrestrial system better. The LDPC coding mechanism provides robustness gain over the DVB-T mechanism based on concatenated convolutional and Reed-Solomon coding. Also, it has been pointed out that in the DVB-T systems the lack of time interleaving introduces challenges in difficult reception conditions (for example mobile reception) [52]. Therefore, in DVB-T2 physical layer time interleaving is introduced in order to obtain better time diversity and thus better performance in mobile and portable usage scenarios. The block schema of a DVB-T2 transmitter and relationship to other DVB standards is visualized in Fig. 4.4. The input to the DVB-T2 physical layer are Transport or Generic streams. The packets of these streams are transmitted in baseband (BB) frames that are input to the bit interleaved coding and modulation (BICM) module of the system as shown in Fig. 4.4.

In order to enable the reception of PLPs separately, physical layer (L1) signaling indicating how the PLP can be received is necessary. In DVB-T2 the L1 signaling is transmitted in preambles located in the beginning of so called T2 frames. This is quite different from DVB-T where TPS information is transmitted in a continuous manner over 68 OFDM symbols. One T2 frame is a sequence of a configurable number of OFDM symbols. The T2 frame begins with a preamble consisting of P1 and P2 OFDM symbols followed by data symbols. A P1 symbol is a specific preamble symbol that enables quick detection of the presence of a T2 signal on the RF channel and initial synchronization to the signal. P2 symbols carry most of the L1 signaling. The structure of the T2 frame is depicted in Fig. 4.5.

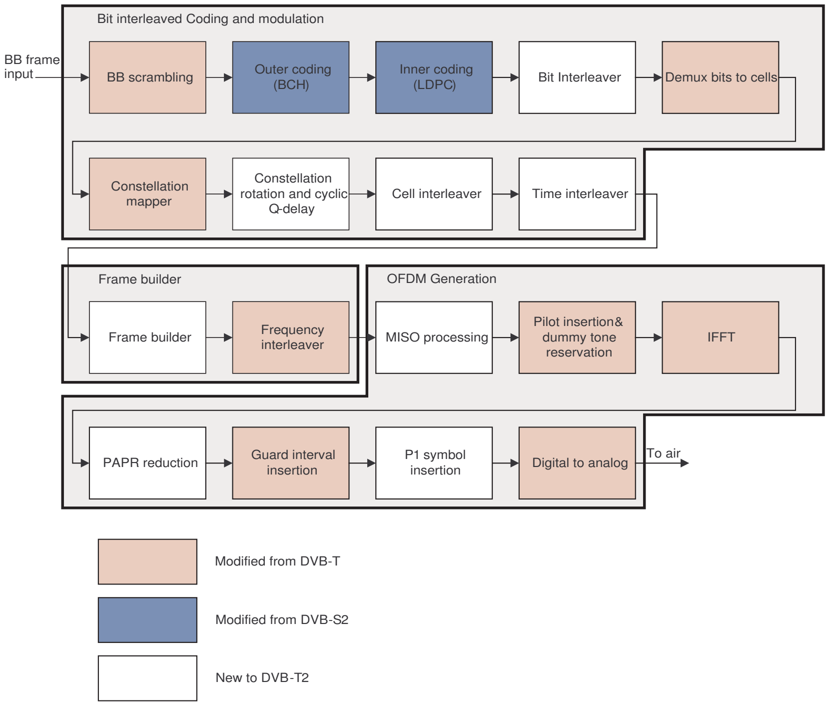

Figure 4.4: Block schema for DVB-T2 transmitter with relationship to previous DVB standards.

The multiplexing is done in the frame builder (Fig. 4.4). The multiplexing of the services to the T2 signal in one RF channel is further presented in Fig. 4.6 by an example of two data PLPs. The L1 signaling is divided into L1-pre and L1-post signaling, both transmitted in P2 symbols. L1-pre signaling provides general information on the system and the means to receive the L1-post signaling. L1-post contains the information such as addressing, modulation order, code rate and so on to give the receiver the means to receive the PLPs. The common PLPs are used for transmitting information that is common to several PLPs, for example link layer (L2) signaling and electronic program guide (EPG). The data PLPs of type 1 consisting of one burst in the T2-frame are mapped directly after the common PLPs. Type 2 PLPs consisting of several bursts also called subslices (not shown in the figure) are mapped after the type 1 PLPs. It is important that the OFDM subcarriers allocated for the PLPs can be different from frame to frame to enable statistical multiplexing over PLPs. Statistical multiplexing enables better usage of the available bandwidth for variable bit rate services than allocating a fixed number of carriers for each PLP. Therefore, the ser-

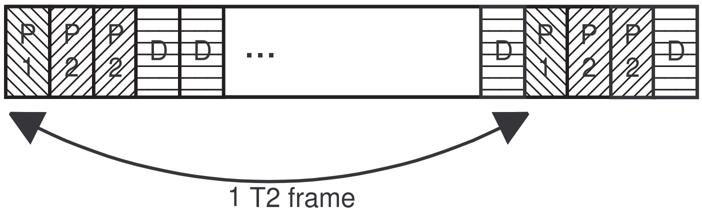

Figure 4.5: Structure of a T2 frame.

vice provider is able to allocate more services into a DVB-T2 signal with statistical multiplexing than would be possible without it.

Peak-to-Average Power Ratio (PAPR) is a problem in OFDM systems. High PAPR results in spectral spreading and distortion when the signal is passed through a nonlinear device, such as the power amplifier in the transmitter [53]. Therefore DVB-T2 specifies two means to reduce the PAPR, namely Active Constellation Extension (ACE) and Tone Reservation (TR).

Bandwidth and frequency flexibility is obtained in the system by allowing several RF signal bandwidths (1.7, 5, 6, 7, 8 and 10 MHz). Several RF signals can further be combined to form a T2 signal by means of TFS. This allows a great deal of freedom for the network designers. Further, when using TFS there is a potential for increased statistical multiplexing gain as there are more services to perform the multiplexing over.

Link Layer

The input to the DVB-T2 physical layer from the link layer is Transport Streams, Generic Streams or both. With TS input, the link layer operations are rather similar to DVB-T with updates in the PSI/SI signaling. Further, TS may be split into different PLPs based on services. Also, part of link layer signaling common to all services transmitted in different PLPs may be transmitted in common PLPs instead of repeating it for each PLP. The option of using TS is available to enable easy transition from the first generation to the second, as the available transmission network equipment for DVB-T supports TS directly.

The GS input provides a more optimized encapsulation mechanism [54] for carrying IP and other network layer packet based content over the second generation DVB physical layers. Generic Stream Encapsulation (GSE) offers a reduced overhead in encapsulation as compared to the MPE used in DVB-IPDC (IP Datacast) on top of the transport stream. Reduction of overhead by a factor 2-3 with respect to MPE over TS is reported in [55]. GSE also

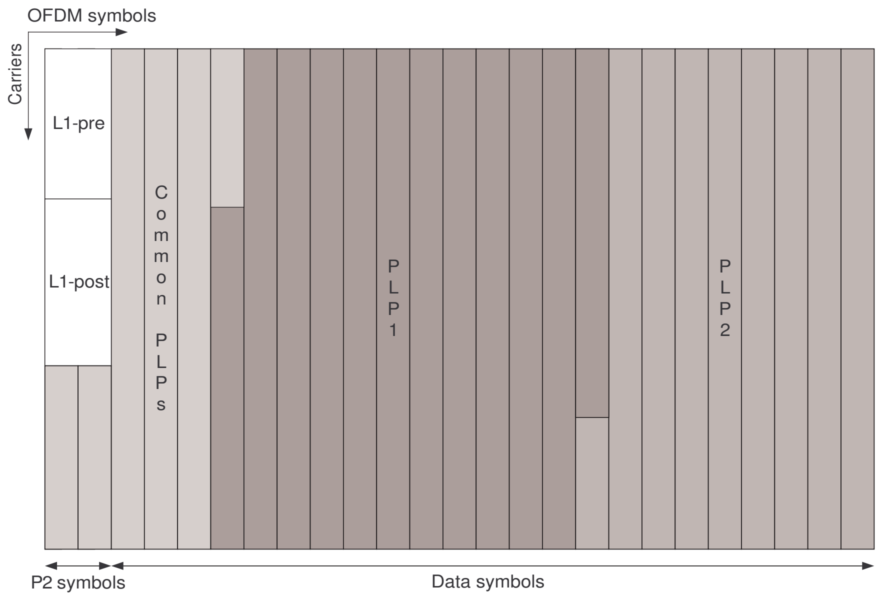

Figure 4.6: Mapping of PLPs to T2 frame.

performs the fragmentation of the network layer packets to fit them into baseband frames. Thus the problems arising from the network layer packet fragmentation can be avoided and the GSE level fragmentation is hidden from the network layer [55].

Chapter 5

Case study: DVB-H link layer analysis

As considered in previous chapters, the main building blocks that affect the QoS at the physical and link layers of a broadcasting system are forward error control coding, interleaving and signaling. In this chapter we study the mechanisms used and their performance at the DVB-H system link layer to obtain a practical view on the effect.

In DVB-H forward error control can be applied at three different system layers: physical, link, and application layers. Here, the FEC at the physical and link layers are mainly considered and analyzed. In addition to FEC, also interleaving processes are investigated, since they are closely related to FEC and they have a great effect on the system performance in mobile transmission environments as described in chapter 2.

Control signaling later referred to as solely signaling affects the QoS observed by the user as described in section 3.3.3. Signaling is used at several system layers of the DVB-H system to give the receiver necessary information to synchronize each layer to the transmission network. At the physical layer (L1) the signaling information consists of for example the modulation order and code rate used so that the physical layer of the receiver is able to extract the service data and to pass it to the link layer. The link layer (L2) signaling information consists of parameters that are required at the link layer for service discovery and reception. With this information the receiver link layer can locate the required service data from the physical layer output. There should be no need for the L2 to know the L1 signaling parameters or vice versa. This way the independence of the different system layers is maintained. Still, interaction between the layers is necessary. In DVB-H the delta-T indicating the time from burst to burst can for example be considered to be L2 information, being carried in the L2 packet (MPE) header, but it is needed in switching the physical layer parts on and off for power saving. Thus, the link layer controls the physical layer. In this chapter the L2 signaling is analyzed to gain knowledge on the robustness of the service following and discovery.

5.1 Error control coding and diversity in DVB-H

The physical layer forward error control coding in DVB-H consists of concatenated RS and convolutional coding as described in section 4.2.1. The length of the RS code is matched to the length of the TS packet, that is, the contents of each TS packet are individually protected by the RS code. Thanks to this individual protection, the physical layer in receiver can detect whether it was able to decode each TS packet correctly or not. This information collected by the receiver is utilized in part of the analyses presented in this chapter. After the RS encoding, the codewords are interleaved by a convolutional interleaver with $N = 12$ and $J = 17$ (see Fig. 2.6). The memory required by such an interleaver is $NJ(N-1)/2 = 1122$ bytes. Therefore, this interleaver is capable of spreading data over a couple of OFDM symbols, the amount depending on the modulation scheme and FFT size used. Synchronization of the deinterleaver in the receiver is enabled by the synchronization bytes of the TS packets that are always passed through the interleaver branch without delay. The synchronization byte (or inverted synchronization byte) should be found from every 204th byte.

The output from the convolutional interleaver is transformed into binary form and fed to convolutional code encoder with rate $R = 1/2$ and generator polynomials $G_1 = 171_{oct}$ and $G_2 = 133_{oct}$. The outputs of the rate 1/2 encoder are punctured to allow code rates 2/3, 3/4, 5/6 and 7/8. The interleaver after the convolutional encoding consists of bit-wise interleaving that is followed by symbol interleaving (the latter can also be called a frequency interleaver as it interleaves the data to the subcarriers of an OFDM symbol). The symbol interleaver counteracts the effect of frequency selective fading that is often present in a terrestrial environment. Both interleavers operate within one OFDM symbol if the optional in-depth interleaving is not utilized. With in-depth interleaving the interleaving is performed over four or two OFDM symbols for 2k and 4k modes respectively. As a conclusion time diversity corresponding to the duration of a few 8k OFDM symbols can be obtained with the whole physical layer interleaving chain. In real time this corresponds to a few milliseconds. For mobile channels, this is not enough and therefore additional MPE-FEC at the link layer is introduced to cover up for the lack of time interleaving at the physical layer. Time interleaving at the physical layer would have been a good feature for DVB-H, and a more effective one than the link layer interleaver, but the backwards compatibility requirements to DVB-T imposed restrictions for the re-design.

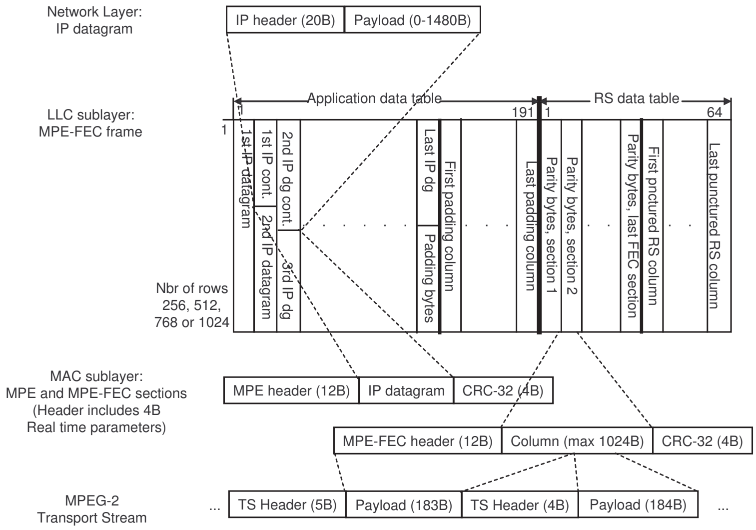

Figure 5.1: The link layer operations of DVB-H [61].

The input to the DVB-H system link layer in the transmitter side is IP datagrams. When MPE-FEC is used, the datagrams are inserted into the MPE-FEC frame column-wise (Fig. 5.1) for the row-wise calculation of redundancy bytes with the RS(255,191) code. The number of rows in the frame can be 256, 512, 768 or 1024. The number of rows in the frame defines how much data there can be in one frame and consequently affects the duration of the time-slice and delta-t for a given service bit rate. Time-slicing and MPE-FEC are closely related to one another, since exactly one MPE-FEC frame is transmitted during a time-slicing burst. The number of data columns is one to 191 and the number of redundancy columns is zero to 64. As the RS code parity is calculated horizontally, using less than 191 of the columns for the data translates into effectively shortening the code. Respectively, omitting redundancy columns translates into puncturing the code. A combination of code shortening and puncturing by using different amounts of columns for the data and parity is used to achieve different MPE-FEC code rates. The different code rates are used to find the trade-off between good enough mobile performance and overhead introduced by the parity when configuring the network. The code rate is close to 3/4 if all 191 data columns are used and all 64 redundancy columns are transmitted.

For the transmission over the physical layer the MPE-FEC frame is divided into sections so that an IP datagram forms the payload of a MPE section and a redundancy column forms the payload of a MPE-FEC section. Important for the receiver operation, information to which address in the frame the contents of the section belongs to and real time parameters such as delta-T are inserted into the section headers. After the section header is attached, the four CRC-32 redundancy bytes are calculated for the section. The sections are transmitted in MPEG-2 transport stream format as defined in [45] and [47]. The device performing most of the DVB-H specific operations as compared to DVB-T (packet encapsulation, time slicing and MPE-FEC encoding) is called an IP encapsulator [56]. The operations performed by the link layer are illustrated in Fig. 5.1.

In the receiver after demodulating the information bits the convolutional code is decoded. Commonly convolutional codes are decoded using the well known Viterbi-algorithm [24]. The implementation of soft decision decoding with the Viterbi-algorithm is of reasonable complexity and therefore often used. For the RS code on the other hand, hard decision decoding is the main alternative, as the complexity of soft decision decoding RS codes is high.

The output of the DVB-H physical layer at the receiver consists of TS packets that are the input to the DVB-H link layer for IP decapsulation. The receiver reconstructs the MPE and MPE-FEC sections from the TS packets. The IP datagrams and redundancy columns carried inside the sections are inserted to correct locations in the MPE-FEC frame based on the knowledge carried in the section header. The RS decoding is then performed row-wise. As the FEC redundancy is calculated row-wise and the transmission occurs column-wise, there is "virtual" time interleaving for the MPE-FEC frame. The term virtual here is used because only the redundancy is transmitted in an interleaved manner and data is transmitted without interleaving. This link layer interleaving is introduced to compensate for the lack of physical layer time interleaving. Again, leaving the data without interleaving was a design decision to allow compatibility with receivers that are not capable of performing the MPE-FEC decoding at all. After the MPE-FEC decoding, the IP packets are forwarded to the application layer for processing.

The application layer FEC for DVB-H is defined in [57]. It is mainly designed for use with file delivery services, but also studies on its applicability for streaming services have been presented for example in [58]. The main advantage of the application layer FEC is that it can provide great time diversity for the file download, where the delay constraint for the content and service switching is very loose as compared to streaming services. Also, the memory at the application layer is "cheaper" than at link or physical layers, as general purpose memory can be used at the application layer instead of dedicated memory. Therefore longer interleaving can be implemented at the same cost. On the other hand, application layer FEC mechanisms usually have to perform less effective hard decision decoding instead of soft decision, as conveying soft information from physical layer to application layer would be very complex to implement in the receivers.

5.1.1 MPE-FEC performance analysis

The standard [47] defines the Reed-Solomon code used in the MPE-FEC and how to puncture or shorten it. The decoding method is, however, left open for each receiver manufacturer to decide. For Reed-Solomon codes decoding with erasures is possible and recommended in [59] among primary options. Erasure in this context stands for an unknown symbol value in a known location in a codeword. The advantage of the erasure decoder is that it is capable of correcting more erroneous code symbols than the non-erasure decoder. Further, the error correction performance of the code could be enhanced by soft decoding using algorithms such as the one by Koetter and Vardy (KV) [60]. At the link layer, soft information obtained from the demodulator is unfortunately not easily accessible and therefore soft decision decoding performance is not considered here. To convey the soft information from the demodulator it would require that the convolutional and Reed-Solomon codes of the physical layer were decoded so that soft information output could be provided.

In DVB-H, different options exist to obtain the erasure information for MPE-FEC decoding. It is suggested in [59] that the erasure information could be obtained from the Cyclic Redundancy Check (CRC) error detection mechanism embedded in MPE and MPE-FEC sections in the encapsulation process. The erasure decoding based on CRC check is referred to in the following as CRC-erasure (CE) decoding. Another option is to apply error information provided in the transport stream (TS) packet headers. This erasure decoding method is referred to as transport stream erasure (TSE). In [61, 62] two decoding methods based on correcting both errors and erasures were proposed. These decoding methods were called hierarchical section erasure (HSE) decoding and hierarchical transport stream (HTS) erasure decoding. Here these methods are called MPE-header-erasure decoding (MHE) and PID-erasure decoding (PE) with respect to the way the erasure information is obtained. The main difference between the decoding methods from the performance point of view is the way the erasure information is obtained and how it is utilized in the decoding process.

In the following, the error correction capability of the RS-based MPE-FEC at the DVB-H link layer is analyzed theoretically based on decoding error probabilities for the four different decoding methods. A stationary memoryless channel for the bit stream arriving at the link layer is assumed. This assumption is justified when the interleaving procedures preceding the link layer decoding stage are effective enough to disperse the error bursts occurring in the transmission channel. The following analysis treats the physical layer bit stream as an output of a binary symmetric channel with the bit error (crossover) probability p. Then the probability of error p_s for one eight bit RS symbol (byte) is

$$p_s = 1 - (1 - p)^8 \approx 8p, \tag{5.1}$$

approximation being valid whenever $p << 1$.

The criterion to compare the performance of the different decoding methods is MPE-FEC frame error rate (MFER). A frame is considered erroneous whenever the decoding of the frame is not successful (that is when the decoder was unable to decode at least one row). The payloads of the sections belonging to the RS data table are always of the same length N_s coinciding with the number of rows in the frame (Fig. 5.1). For the sake of simplicity we assume here that the length of the IP packets also coincides with the number of rows in the MPE-FEC frame. For this analysis $N_s = 535$ is chosen to enable simple and unified analysis of the different decoding methods. This assumption makes it possible to have an integer number of TS packets in a column of the MPE-FEC frame. Although this option is not defined in the standard, it gives results that are very close to the defined case with $N_s = 512$.

CRC-erasure decoding (CE)

Let us now turn our attention to the erasure decoding. Introduce designations p_u, p_e, t_u and t_e for the probability of undetected corrupted RS symbol (byte in this case), the probability of an erased symbol, number of corrupted symbols in a RS codeword that were not detected and number of erased symbols in a codeword respectively. According to [63] and [64] any code of distance d corrects t_e erasures and t_u errors whenever

$$t_e + 2t_u < d. \tag{5.2}$$

Since we only consider decoding within the code distance and lose an MPE-FEC frame any time the correction of errors and erasures fails, every violation of (5.2) is treated as a decoding error. Now, for a code of length n there are $\begin{pmatrix} n \\ t_e \end{pmatrix}$ equiprobable patterns of t_e erasures and for each of them $\begin{pmatrix} n - t_e \\ t_u \end{pmatrix}$ equiprobable placements of t_u undetected symbol errors on the $n - t_e$ positions left. Since the probability of any fixed pattern of t_e erasures and t_u undetected errors is $p_e^{t_e} p_u^{t_u} (1 - p_e - p_u)^{n - t_e - t_u}$, the joint probability distribution of t_e, t_u is [65]:

$$\begin{aligned} p(t_e, t_u) = \begin{pmatrix} n \\ t_e \end{pmatrix} \begin{pmatrix} n - t_e \\ t_u \end{pmatrix} p_e^{t_e} p_u^{t_u} \cdot \\ \cdot (1 - p_e - p_u)^{n - t_e - t_u}. \end{aligned} \tag{5.3}$$

As a result the probability of correct decoding of one codeword for erasure decoding is evaluated:

$$\begin{aligned} P_{cE} &= P(t_e + 2t_u < d) = \\ &= \sum_{t_e=0}^{d-1} \sum_{t_u=0}^{\frac{d-1-t_e}{2}} p(t_e, t_u). \end{aligned} \tag{5.4}$$

In the course of CRC processing at the DVB-H link layer all the MPE and MPE-FEC sections undergo testing on whether they are corrupted by bit crossovers or not and a whole section is discarded if a CRC error is discovered. Thus, under the assumption of the section length $N_s = 535$ coinciding with the number of rows in the MPE-FEC frame, every detected error within a section erases precisely one symbol (byte) in every RS codeword in the frame.

The error detection capability of CRC-32 is rather high and is not nearly exhausted by only detecting all errors of weight up to three [66] as the distance properties of the CRC-32 code would indicate. Like any other binary linear code used for error detection it may miss only fraction 2^{-r} of all possible error patterns, r being the number of redundant bits [63]. For the CRC-32 $r = 32$, and the share of undetectable corrupted section patterns does not go beyond $2^{-32} < 3 \cdot 10^{-10}$. Besides, the probability of an undetected corrupted symbol in a RS codeword appears to be much smaller against the probability of CRC fault, since in a missed corrupted section not all bytes are necessarily wrong. Therefore $p_u << 1 - p_e$ and we may neglect p_u and put $t_u = 0$ in (5.3). Assuming an absolute reliability of CRC, we can use the probability of an error occurring in the MPE(-FEC) section as the erasure probability for the RS codewords in (5.3). This probability can be calculated as

$$p_e \approx 1 - (1 - p_s)^{N_s}. \tag{5.5}$$

Since the erasure patterns for each RS codeword in the MPE-FEC table are exactly the same, the probability of erroneous CE decoding is simply calculated by

$$P_{eCE} = 1 - P_{cE}. \tag{5.6}$$

Transport stream packet erasure decoding (TSE)

For the analysis of TSE decoding (described in [67]) let us assume that when the TS packet is declared correct by the physical layer, the data carried inside the packet can always be decapsulated into the MPE-FEC frame. This could be accomplished for example with the help of the continuity counter in the TS header. The information on the correctness of the TS packets is obtained from a one bit flag in the TS header. This flag is set by the physical layer to

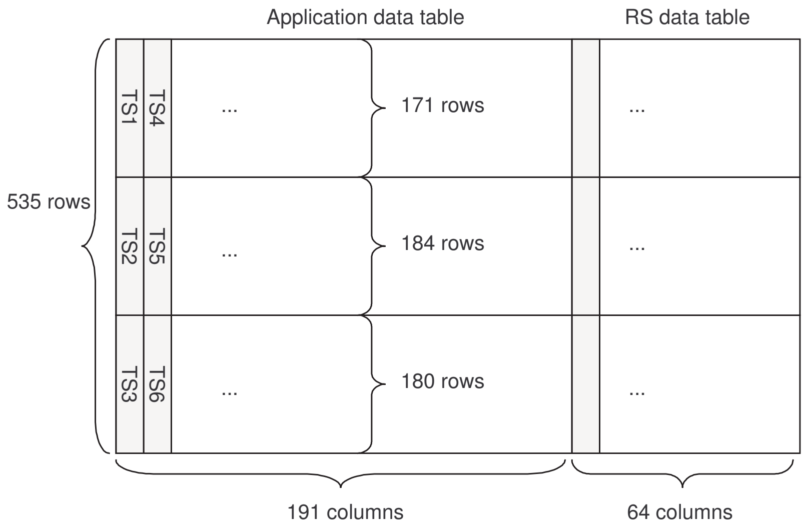

Figure 5.2: Structure of the MPE-FEC frame used in calculations.

indicate the situation where an uncorrectable error pattern in the TS packet is recognized by the physical layer RS(204,188) decoder. When the flag is set the data carried inside the TS packet in question is marked as erased for the MPE-FEC decoder. To enable simple calculations the MPE-FEC frame is considered to consist of three subframes having 535 rows in total (Fig. 5.2). The sizes of the subframes are 171, 184 and 180 rows, since following our assumptions the first TS packet carries the 12 byte MPE header and one byte payload unit start pointer and the third contains the 4 CRC-32 bytes. In this decoding scheme the information provided by the CRC-32 decoding is totally ignored. For our analysis the reliability of the erasure information obtained from the physical layer RS decoder needs to be evaluated first.

One way to estimate the probability of undetected error pattern in MDS (Maximum Distance Separable) codes is studied in [68], where results support the intuitive idea that the probability in question (if small enough) may well be approximated by the share of undetectable error patterns:

$$\begin{aligned} P_{decError} &\cong \frac{\text{number of decodable words}}{\text{number of words}} = \\ &= \frac{(q^k - 1)V_n(t)}{q^n} \approx q^{-(n-k)}V_n(t), \end{aligned} \tag{5.7}$$

where k is the number of information symbols in a MDS codeword and $V_n(t)$ is the volume of a Hamming sphere of radius t, t being code correction capability. This result can be used for any MDS code, including shortened RS codes [69] such as the DVB-H physical layer RS(204,188) code. Since for the RS codes $n - k = 2t$, we have

$$P_{decError} \approx q^{-2t} V_n(t). \tag{5.8}$$

For $V_n(t)$ the following estimate holds:

$$\begin{aligned} V_n(t) &= \sum_{i=0}^{t} \binom{n}{i} (q-1)^i \\ &< q^t \sum_{i=0}^{t} \binom{n}{i} < q^t 2^{nh(\frac{t}{n})}, \end{aligned} \tag{5.9}$$

where $h(x)$ is binary entropy [63]. Then from (5.8) and (5.9) together we obtain

$$P_{decError} < q^{-t} 2^{nh(\frac{t}{n})}. \tag{5.10}$$

For $t = 8$, $n = 204$ and $q = 256$: $nh(\frac{t}{n}) - tlog_2 q \approx 204 \cdot 0.24 - 64 \approx -15$ so that $P_{decError} \approx 2^{-15} \approx 3 \cdot 10^{-5}$. This shows that any error pattern of weight greater than t will be almost for sure (that is with probability $\geq 1 - 3 \cdot 10^{-5}$) detected in the course of RS decoding so that in (5.3) p_u may be neglected and t_u put to zero.

Now the probability of erasure of one code symbol in every codeword in one of the three subframes is evaluated:

$$p_e \approx 1 - (1 - p_s)^{188}. \tag{5.11}$$

The probability of correct decoding of one subframe is evaluated by (5.4). The whole frame will be correctly decoded whenever decoding of all three subframes is successful leading to decoding error probability

$$P_{eTSE} = 1 - P_{cE}^3. \tag{5.12}$$

MPE-header-erasure decoding (MHE)

The MPE-header-erasure decoding scheme is presented in [61] as Hierarchical Section decoding. The main idea is that on the contrary to what is suggested in [59] the data carried inside an unreliable section (or TS packet) having detected errors is inserted into the MPE-FEC frame for decoding whenever possible. In the case of MHE decoding it is assumed that the payload of the section can be inserted into the frame for decoding whenever the MPE header and thus the address of the section in the MPE-FEC frame is not corrupted. Thus from the decoder point of view the section is erased only when an error hits the 12 byte section header leading to the probability of erased symbol in a codeword:

$$p_e \approx 1 - (1 - p_s)^{12}. \tag{5.13}$$

As described in [61] the decoding procedure has several stages, the first one being the decoding using both erased and unreliable sections as erasures. The performance of this first stage coincides with that of CE decoding. From the error correction capability point of view (and thus for this analysis) the second stage when all the data that has been put to the frame is considered reliable and only lost sections having errors in the MPE header are considered erased is more interesting. The probability of undetected symbol error in this situation is just the symbol error probability from the physical layer ($p_u = p_s$), since the information on payload errors within the section is discarded (that is CRC-32 information is not used) in this step of the decoding scheme. Using an MPE-FEC frame with 535 rows the probability of erroneous decoding for MHE can be calculated after substituting p_e from (5.13) and $p_u = p_s$ in (5.3) by

$$P_{eMHE} = 1 - P_{cE}^{535}. \tag{5.14}$$

PID-erasure decoding (PE)

The idea of the PE decoding scheme is presented in [61] as Hierarchical TS decoding. PE is rather similar to MHE, except that TS packet based information rather than section based is used. In the analysis of the PE decoding it is assumed that the payload of a TS packet can be put into the frame when the TS packet can be received. This could be accomplished for example with the help of the continuity counter in the TS packet header. A TS packet can be received whenever the two byte (actually 13 bits, but approximated here to be two bytes) PID in the TS header is correct. If the PID was not correct, the receiving equipment link layer would not be able to recognize the packet to be a part of the received stream and thus would filter it out. When the packet is not received it is considered erased by the receiver. With this observation, the probability of an erasure in each symbol of a MPE-FEC codeword in one sub-frame is approximated by the probability of the situation when an error hits the PID:

$$p_e \approx 1 - (1 - p_s)^2. \tag{5.15}$$

Again the interesting situation takes place when only completely lost TS packets (that is having errors in PID) are considered erased and knowledge of detected payload errors (error information from the physical layer RS

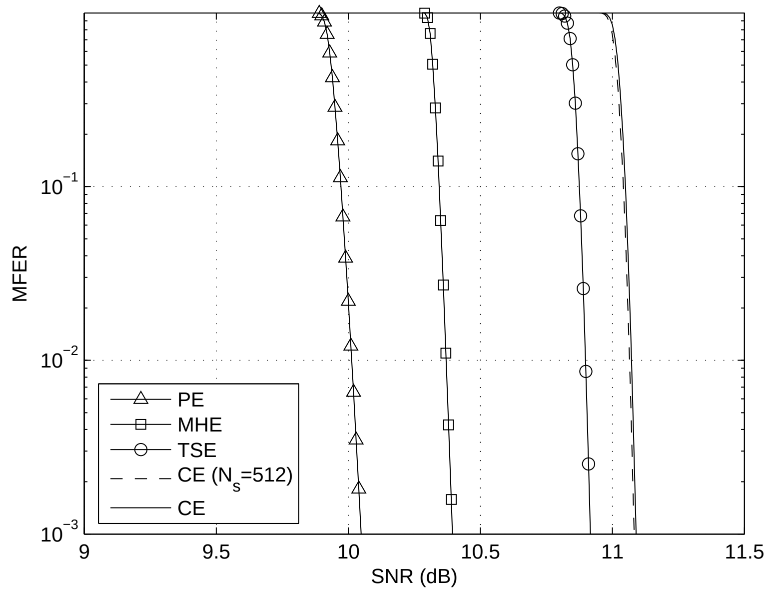

Figure 5.3: Comparison of different decoding methods (AWGN channel, 16-QAM, convolutional code rate 3/4, MPE-FEC code rate 3/4).

decoder and link layer CRC-32) is discarded. The probability of undetected error is now $p_u = p_s$. Using p_e (5.15) and $p_u = p_s$ in (5.3) the probability of erroneous decoding of one frame having three subframes (see Fig. 5.2) is evaluated by

$$P_{ePE} = 1 - P_{cE}^{171} P_{cE}^{184} P_{cE}^{180} = 1 - P_{cE}^{535}. \tag{5.16}$$

Comparison of the decoding methods

To compare the performance of the different decoding schemes for MPE-FEC with code rate 3/4 in terms of required SNR (E_s/N_0), MPE-FEC frame error rates are calculated from (5.6), (5.12), (5.14) and (5.16) using physical layer output byte error probabilities p_s related to SNR. The physical layer output error probabilities are simulated by a DVB-T physical layer simulator. The physical layer parameters used in the simulations are: 16-QAM modulation, convolutional code rate 3/4, 8K OFDM mode and guard interval duration equal to 1/4 of "pure" OFDM symbol duration. The channel model used in the simulator is the simple Additive White Gaussian Noise (AWGN). The results are shown in Fig. 5.3. The curve for CE with $N_s = 512$ (dash line practically coinciding with a solid line for CE decoding with $N_s = 535$) is included to show the negligible effect of deviation of $N_s = 535$ from the

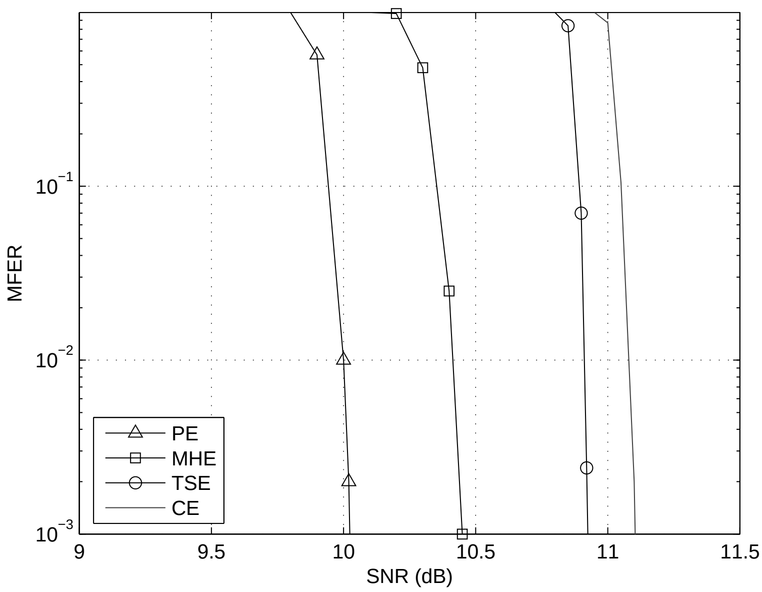

Figure 5.4: Link layer simulation results for different decoding schemes (AWGN channel, 16-QAM, convolutional code rate 3/4, MPE-FEC code rate 3/4).

standard frame size. The ranking of the compared decoding methods follows from the order of frame error probabilities: $P_{ePE} < P_{eMHE} < P_{eTSE} < P_{eCE}$. The coding gains of the decoding methods over the CE decoding suggested in the implementation guidelines [59] are approximately PE=1, MHE=0.7 and TSE=0.2 dB at $MFER = 10^{-3}$ in the AWGN channel.

Example

Take for example SNR 10 dB for PE decoding. From physical layer simulations it is known that the symbol error probability for this SNR is $p_s = 0.033946$. Now we can calculate from (5.15) that $p_e = 1-(1-0.033946)^2 = 0.066740$ and set $p_u = p_s = 0.033946$. Substituting these into (5.3) and further in (5.4) for calculations results in $P_{cE} = 0.999959$ for MPE-FEC coderate 3/4. Then, MPE-FEC frame error probability can be evaluated to be $P_{ePE} = 1 - 0.999959^{535} \approx 0.022$. This value can be observed in Fig. 5.3.

Link layer simulation results using the same physical layer parameters are shown in Fig. 5.4. The link layer simulator generates the MPE-FEC frames, introduces errors according to the physical layer simulation, calculates the numbers of errors and erasures in rows of the frame and decides whether the frame in question can be decoded correctly by the RS code or not. As a result frame error rates are obtained. These simulations show similar results as the theoretical calculations in the previous section thus supporting them.

Simulation results for the presented decoding methods in a mobile multipath channel can be found from [61]. The results of those simulations at different Doppler frequencies indicate similar ranking of the decoding methods as shown here. The gain for the PE (HTS in [61]) over CE (SE in [61]) is also around 1 dB.

In the previous analysis it turned out that the PE decoding is the best of these four methods while CE decoding arranged as suggested in [59] has the worst decoding capability. The significant result is that all other decoding methods including the TSE decoding that ignore the CRC-32 information perform better than the CE decoding. Therefore it would be of interest to find some other more benefitical use for the 4 byte overhead caused by the CRC-32 in MPE(-FEC) sections giving the system performance gain. For example, an error correcting code with 4 byte redundancy over each MPE and MPE-FEC section header could be used to additionally protect the real time parameters that are necessary in the reception. Of course, this kind of change is not possible as the standards for the DVB-H system are stabilized, but the issue should be taken into account in the design of the future systems. The main reason for the poor performance of the CE decoding is that using the CRC-erasure information over rather long MPE(-FEC) sections erases many sound bytes along with the erroneous ones. If, for example, there is one real byte error in a section of length 512 bytes, 511 correct bytes are erased from the MPE-FEC frame when the CE decoding is performed.

As it is clear that PE obtains the best error correction capability for the MPE-FEC, it is important to be aware of the effect of its utilization on the total complexity of the MPE-FEC decoding. Study on the effect of the MPE-FEC decoding methods on the complexity of Reed-Solomon decoding is presented in Appendix A. The significant outcome of the study is that of the decoding methods studied here the one having the best error correction capability also minimizes the amount of operations necessary in the Reed-Solomon decoding.

5.2 Signaling

Let us now turn the attention from the data path to signaling and its transmission in DVB-H systems. The physical layer (L1) signaling is called Transmission Parameter Signaling (TPS). The TPS consists of total 53 information bits. These bits signal the number of the current frame in the superframe, cell identification, transmission mode (2k, 4k, 8k), inner physical

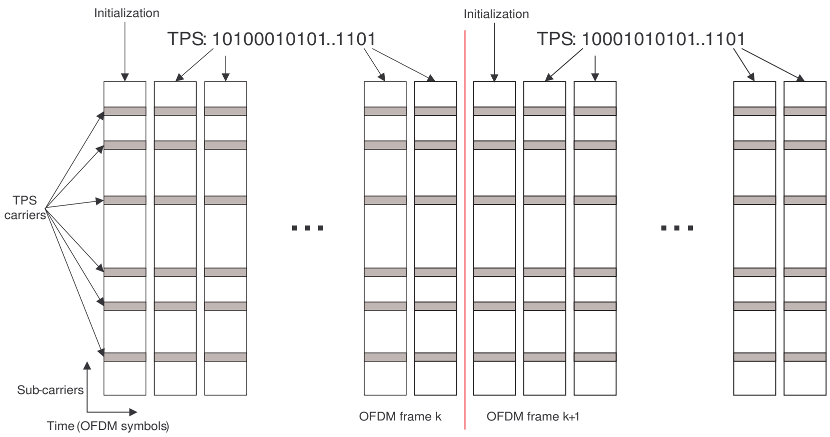

Figure 5.5: Transmission mechanism for TPS in DVB-H.

layer (convolutional) code rate, guard interval, constellation and hierarchy information, usage of MPE-FEC and time slicing and usage of in-depth interleaver. A BCH(67,53) code capable of correcting 2 bit errors is applied over the TPS information, resulting in codeword of length 67. TPS is transmitted on dedicated carriers in OFDM symbols. In 2k mode there are 17, in 4k mode 34 and in 8k mode 68 dedicated carriers. The same data (differentially encoded) is transmitted in parallel on each dedicated carrier of one OFDM symbol (one OFDM symbol carries exactly one bit of TPS information) to ensure robust transmission of TPS bits. Therefore, one TPS codeword is transmitted in one OFDM frame consisting 68 OFDM symbols. One superframe consists of four frames. The first OFDM symbol must carry the initialization for the differential modulation of the following 67 information symbols. The transmission is illustrated in Fig. 5.5. An interesting observation on TPS is that to receive the TPS signaling, some of the parameters it is signaling are already needed, such as guard interval and transmission mode (that is the size of FFT) for example. This means that having this information in TPS is intended mainly for making physical layer configuration changes smooth. This is possible because the TPS information carried in one frame refers to the next frame. The receiver visiting a physical channel for the first time must detect the guard interval and FFT size before it can receive the TPS information signaling the rest of the data transmission parameters.

Once the physical layer parameters are known, the receiver link layer (L2) can obtain the signaling from the known PID values. The PSI/SI (Program Specific Information/Service Information) [70–72] is an essential part of service discovery and following in DVB-H and DVB-T systems. The

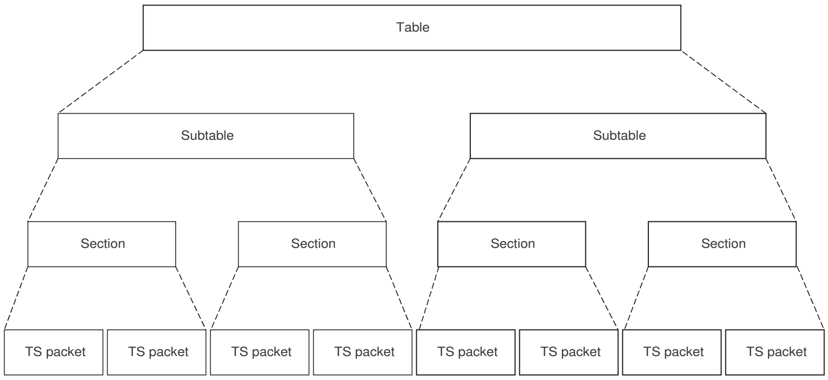

Figure 5.6: Mapping of PSI/SI tables to TS packets.

PSI/SI access time has a direct effect on the total latency in service access. From the end user point of view, fast service access time is preferred and hence the PSI/SI access time should be minimized. From the network operator point of view on the other hand, the capacity consumed by the signaling should be minimized to leave as much capacity for the payload services as possible. In DVB-H the PSI/SI information is carried in MPEG-2 private table structures [70]. These structures are tables that generally consist of subtables. The subtables are transmitted in sections that are carried inside transport stream (TS) packets. The mapping of a table to the TS packets is illustrated in Fig. 5.6. PSI/SI tables are transmitted with a certain interval to enable tapping into the network when the receiver is powered on. The transmitted sections contain CRC information allowing the receiver to check the correctness of the received sections. If the receiver detects that some sections carrying one subtable during one transmission are in error it has to wait for the next transmission to receive the sections. The PSI/SI tables needed for service discovery in DVB-H networks are [59]: IP/MAC Notification Table (INT), Network Information Table (NIT), Program Association Table (PAT), Program Map Table (PMT) and Time and Date Table (TDT).

Although the application data is protected at the link layer by the MPE-FEC to combat the harmful effects of mobility, the PSI/SI information is not. Also, the PSI/SI information is transmitted continuously in the DVB-H multiplex on the contrary to DVB-H services that are time sliced as illustrated in Fig. 4.3. The updates to the PSI/SI information are therefore received only during the on-time of the receiver and during some transmission of the table the receiver may not be on. When receiving the PSI/SI for the first time, the receiver needs to stay constantly on to obtain PSI/SI instructing how to receive the time sliced services. In the following analysis

Table 5.1: Simulated Configurations

Configuration	LC	T	MM
Mode	8k	8k	8k
Constellation	QPSK	16-QAM	16-QAM
Conv. Coderate	1/2	1/2	2/3
Guard Interval	1/4	1/4	1/4
Bandwidth	8 MHz	8 MHz	8MHz

mainly the initial scan use case where the tables need to be received from scratch is considered.

5.2.1 Performance analysis of L2 signaling

As the L2 signaling path is not protected at the link layer by FEC while the data path is, the question arises, whether the transmission of the PSI/SI information in its present form is robust enough for mobile receivers. The issue of robustness is briefly touched upon in [59] and further studied in [73], but needs still further research for the mobile environment. Here, the performance of the PSI/SI transmission in a mobile environment is studied by computer simulations and measurements. The effect of configuring the network with different repetition intervals and section sizes for the tables is studied. Also, the effect of different configurations on the capacity reserved by the PSI/SI signaling is considered. The simulations cover the transmission and reception of the PSI/SI. The physical channel model between the transmitter and receiver in the simulated scenarios is the commonly used TU6 [74] mobile multipath channel model. The measurements from two vehicular and two pedestrian use cases are also studied. Three different network configurations are considered. These configurations are named "Low Cost" (LC), "Typical" (T) and "Maximum Mobile" (MM) and their network parameters are given in Table 5.1. LC corresponds to a low cost network where the coverage area of the transmitters is maximized at the cost of reduced capacity available for the services. MM on the other hand represents a network with high throughput obtained at the cost of coverage area of the transmitters. T corresponds to a compromise between the LC and MM network configurations. These configurations are the same as the ones considered in [73] except for the different Guard Interval (1/4 instead of 1/8 of pure OFDM symbol duration).

Three different decoding methods for the PSI/SI tables were simulated. The first one is called an "intelligent" decoder. It is capable of keeping all correctly received sections in memory until all the sections are correctly received and the subtable can be reconstructed. It is assumed that the contents of the subtables do not change between the transmissions, thus enabling

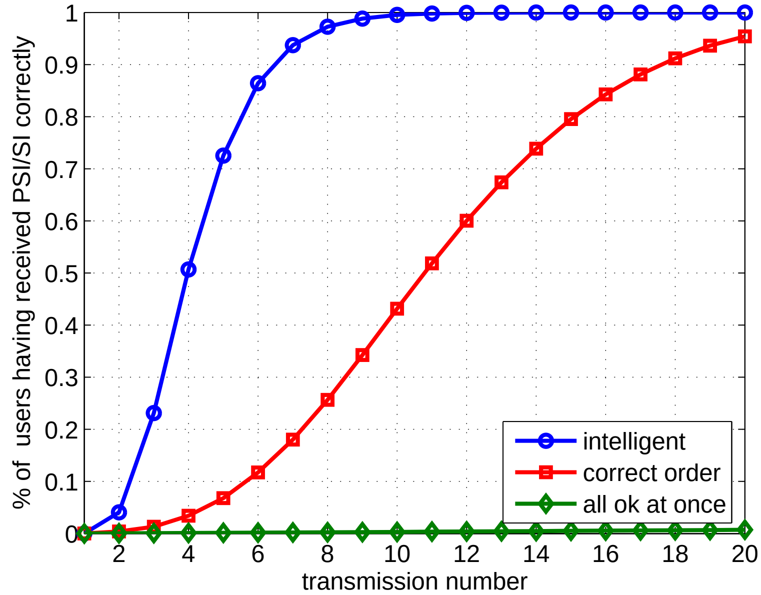

Figure 5.7: Performance of different decoding methods for PSI/SI in the receiver (TS packet error rate 15%).

combining sections from different repeated transmissions. The changes in the information can be detected from the version number in the section header; every time the contents change the version number is incremented. The proper order of the sections can be obtained with the help of section numbering also carried in the section header. The second simulated decoding method is called "correct order" and it requires the correct sections to be received in correct order but not necessarily from the same transmission cycle. The third decoding method is the least intelligent and called "all ok at once". It requires that all sections of a subtable must be received correctly during one transmission for correct reception. The output of the simulations is the percentage of users that have received the table correctly with a certain number of retransmissions (in the following this is called coverage). The performance of the different decoders for a table of size 8192B (Byte) transmitted in 512B sections is illustrated in Fig. 5.7 for TS PER (TS Packet Error Rate) 15% and uniform TS packet error distribution. It is seen that to obtain user coverage of 95%, eight transmissions are needed for the "intelligent" decoder. For the "correct order" decoder this number is 20 and for "all ok at once" it is very large. It is easy to see and deduce that in any situation the intelligent decoder is the best of these three as it loses the least information and its use is therefore recommendable. If the contents of the sections (or subtables) change very often, use of an "all ok at once" type of decoder could be justified. Still, it is probable that the large subtables transmitted in several sections have relatively static contents. This is the reason why our main focus in the following sections is on the intelligent decoder, unless otherwise stated.

The link layer TU6 PSI/SI simulations were performed with recorded TS packet error traces at the output of the physical layer. As a real hardware receiver was used, these traces are inherently receiver dependent. To obtain the error traces the transmitted signal was generated, modulated and ran through a channel simulator using a TU6 channel profile. Then the signal was received by a DVB-T receiver supporting mobile reception (8-tap channel estimation in time direction) and the transport error indicators (TEI) were collected. The receiver was constantly on (it did not perform time slicing) and thus TEI for every packet of every service was collected. Transport error indicators directly inform us whether each TS packet was decoded correctly by the physical layer Reed-Solomon decoder or not, that is whether there are errors left in the packet. All the TS packets that were not decoded correctly were considered to be discarded by each studied PSI/SI decoding method. With TU6 error traces the effect of the motion of the receiver on the PSI/SI transmission can be studied. Error traces were available for Doppler frequencies $f_D = 2$ Hz, 10 Hz, 30 Hz and 80 Hz and Carrier to Noise ratio (C/N) values corresponding to TS PER from around 40 % to nearly error free transmission with 1 dB resolution. Simulated table and section sizes together with repetition intervals for all three network configurations and decoding schemes are given in Table 5.2. The simulation matrix is large to reveal the general behaviour and to cover the possible table and section sizes of the link layer signaling transmission with different possible system configurations. In real DVB-H networks using shorter sections with some subtables increases the total size of the subtable. In these simulations the hypothetic table and section sizes given in Table 5.2 are used meaning that the variation in section size doesn't introduce any changes in the total table size. This is done to obtain more general insight into the behavior of the PSI/SI transmission in mobile multipath channel. 10 000 users were considered to be enough for each simulation to provide reliable results.

The effect of repetition interval

The repetition interval has only a slight effect on the C/N requirement to obtain certain reception coverage with different numbers of transmissions. As an example, the percentage of receivers receiving a table of size 8192B with 512B sections correctly with three transmissions is presented in Fig. 5.8 for typical (T) configuration and Doppler frequency $f_D = 10$ Hz. It is evident that the curves differ very little from each other. The natural rea-

Table 5.2: Simulated table and section sizes and repetition intervals

Table (B)	Section lengths (B)	Rep. Intervals (s)
16384	64, 128, 256, 512, 1024, 2048, 4096	1, 5, 10, 15, 30
8192	64, 128, 256, 512, 1024, 2048, 4096	1, 5, 10, 15, 30
4096	64, 128, 256, 512, 1024, 2048, 4096	1, 5, 10, 15, 30
1024	64, 128, 256, 512, 1024	1, 5, 10, 15, 30
768	64, 128, 256, 512	1, 5, 10, 15, 30
512	64, 128, 256, 512	1, 5, 10, 15, 30
256	64, 128, 256	1, 5, 10, 15, 30
16	8, 16	.025, .05, .075, .1
8	8	.025, .05, .075, .1

soning for this is that already one second repetition interval is much longer than the coherence time of the channel. Therefore, it does not make anymore difference from the performance point of view whether the repetition interval is one or 30 seconds. The network capacity required by the signaling transmission and the time for the receiver to obtain the PSI/SI table, on the other hand, are directly affected by the repetition interval. Sending tables more frequently naturally increases the consumed network capacity and shortens the time a receiver has to wait before acquiring all the sections correctly.

The effect of section size

The effect of section size on the performance can be illustrated by considering first two Doppler frequencies $f_D = 10$ Hz and 80 Hz with the typical (T) network configuration. Consider that the intelligent decoder is used and 95% user coverage is necessary. The results for transmission of the table of total size 16384B with sections of sizes 64B, 1024B and 4096B are shown in Fig. 5.9. The figure shows the required C/N to obtain 95% user coverage with certain amount of transmissions and section size. It can be observed that using 64B sections instead of 4096B ones gives us approximately 2-3 dB gain in TU6 channel with $f_D = 10$ Hz in T network configuration. For $f_D =$ 80 Hz the gain is even 4 dB. According to the simulations the performance of the section sizes between these extremes that are not shown is between the ones shown. Similar order with different gains is observed also with other considered network configurations (LC and MM). As a conclusion, the shorter the section the better the performance. That is, the required user coverage is obtained with smaller C/N using shorter sections than with the longer ones.

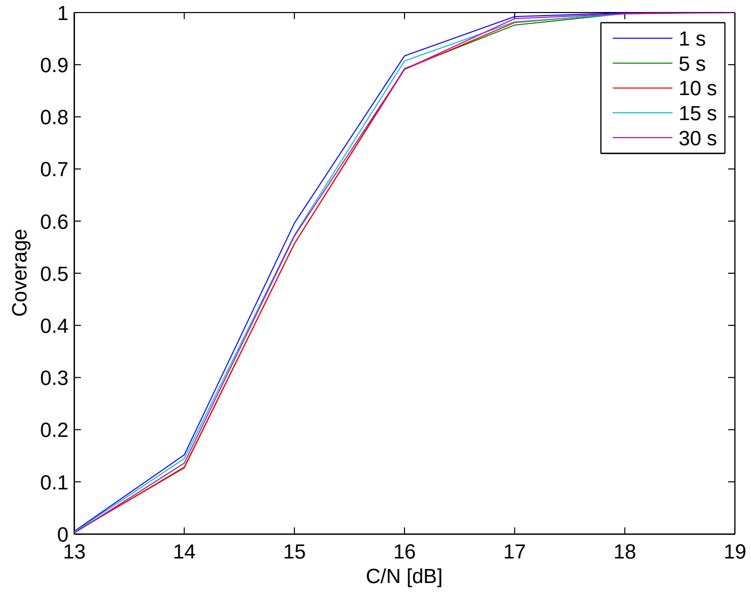

Figure 5.8: The effect of repetition interval for 8192B table transmitted with 512B sections (TU6 $f_D = 10$Hz, T network configuration).

Comparison of the signaling to the data path

Next let us investigate the effect of Doppler frequency on the transmission of PSI/SI as compared to the transmission of the data protected by MPE-FEC with 512 row frame and code rate 3/4. Comparison to the data path is very informative, since at least when the data path is operating above an accepted error criterion, the PSI/SI transmission should work as well to enable service discovery. In fact, it is desirable that the signaling is more robust against channel impairments than the data path as described in section 3.3.3. For all PSI/SI results presented in the following the total table size used is 16384B and it is required to obtain the coverage using 6 transmissions with a 5 second repetition interval. This could be a realistic size for the largest table INT used in the DVB-H systems being in the same time the most error prone. The maximum repetition interval for INT is defined to be 30 seconds [47]. Let us decide for the analysis purposes that 95 % user coverage should be obtained during this time. So, if a repetition interval of 5 seconds is used, six transmissions should be enough to obtain the wanted coverage for the signaling. If the network is planned according to the MFER=5% criterion, the PSI/SI transmission requiring higher C/N value than the data path to reach its error criterion would not be considered viable. For example, in Fig.

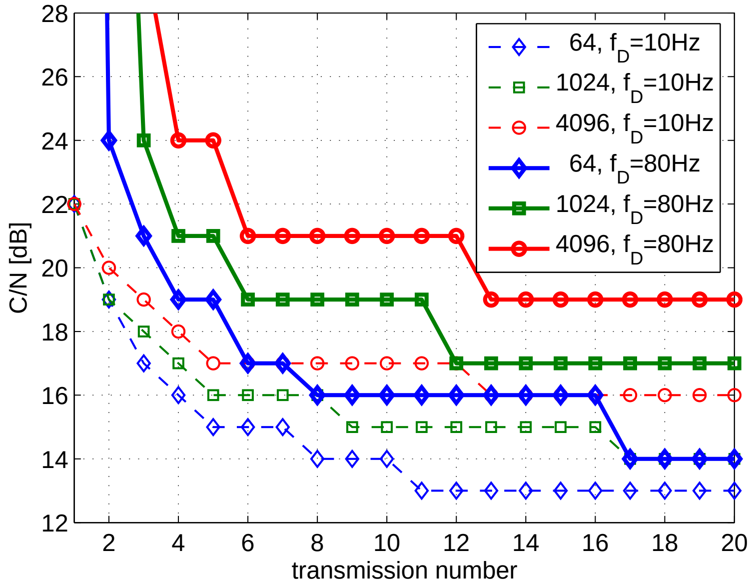

Figure 5.9: Required C/N to obtain 95% user coverage for 16384B table transmitted with different section sizes and number of transmissions (TU6 f_D = 10Hz, T network configuration).

5.10 for LC configuration correct order and all-ok decoders require higher C/N to obtain 95 % user coverage than what is designed for the data path and only the intelligent decoder is a working option. For T configuration the effect of section size on the Doppler performance is presented in Fig. 5.11. The gain of using the smallest sections instead of the largest ones is observed to be approximately 3 dB in T network configuration. The values for Doppler frequencies f_D = 10 Hz and 80 Hz can also be seen in Fig. 5.9. With the 64 byte sections it is also observed that obtaining the required coverage is possible with smaller a C/N value than for obtaining MFER=5% in the datapath, so the PSI/SI transmission is more robust than the datapath in this case and it can be considered to be working. Similarly the use of a 4096B section is not recommended. Considering the situation the other way round, obtaining the required user coverage with 4096B sections would take a longer time than the 30 seconds specified here. From the user point of view, for many users the delay for beginning to receive services would be longer than the 30 seconds.

The Doppler performance for the MM configuration is shown in Fig. 5.12. For this configuration it is observed that the effect of Doppler at f_D = 80 Hz is more than the receiver can handle as the data path requires very

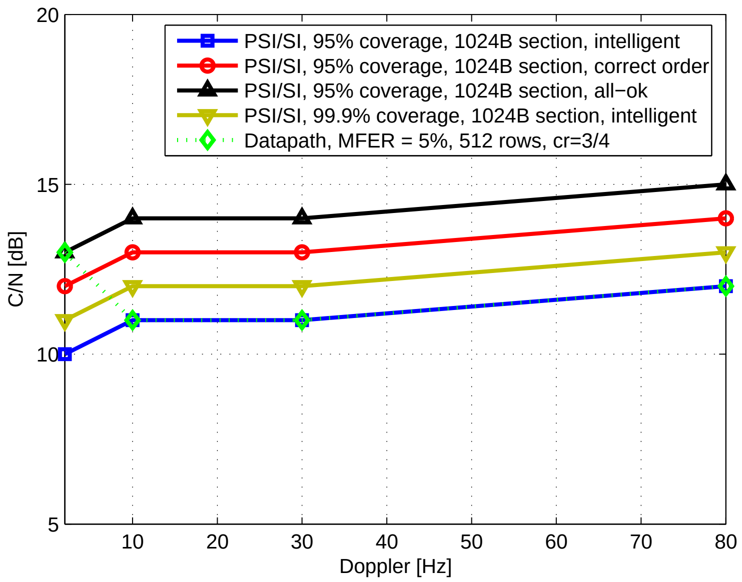

Figure 5.10: Comparison of decoding methods and required coverage percentage for 16384B table (TU6, LC network configuration, MPE-FEC code rate 3/4 for data path).

high C/N to reach the MFER = 5%. Still, when using short sections for the signaling transmission, the signaling path reaches the error criterion of 95% coverage. The curves for section sizes between 64B and 4096B again lie between the curves for these section sizes. It must be noted that the exact shape of the Doppler curves could not be obtained, since only error traces for $f_D = 2$, 10, 30 and 80 Hz were available [75].

It is evident from the simulations that to obtain similar coverage for large tables with longest sections as with the shortest ones, several dB higher C/N is required. For the small tables (8B and 16B) no visible difference between the different section sizes in the simulation results were observed, because they already are smaller than the TS packet. Of the four measured Doppler frequencies $f_D = 80$ Hz is the most challenging, that is highest C/N is required to obtain 95% coverage and respectively $f_D = 2$ Hz is the least demanding.

Network capacity considerations

When the effect of varying section size on the total table size is taken into account (as is done in [73]), the shortest sections are not necessarily the

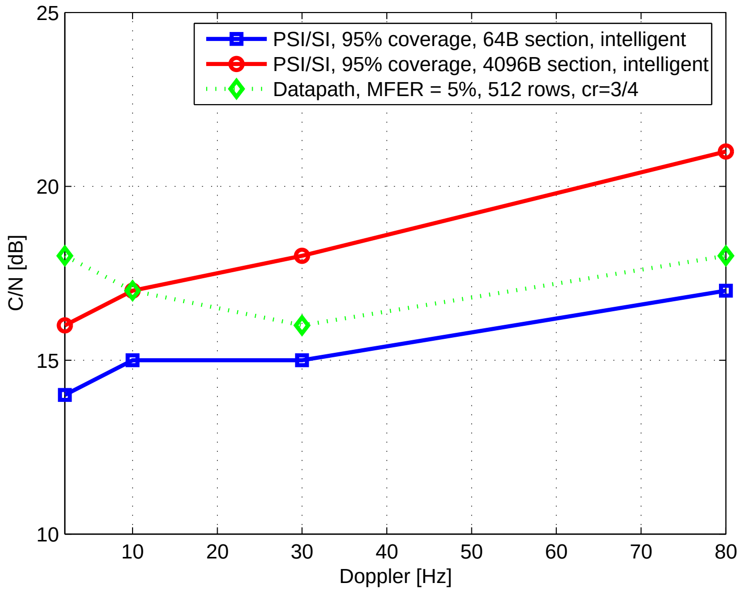

Figure 5.11: Comparison of different section sizes for 16384B table (TU6, T network configuration, MPE-FEC code rate 3/4 for data path).

best from the reserved network capacity point of view. The increase in total table size with decreasing section size is caused by the additional overhead induced by the increased number of sections. The total table size for each section size used in the analysis are calculated for the network configuration studied in [73] and shown for INT in Table 5.3. For NIT the simulated section sizes are 96, 128, 256, 512 and 756 bytes and corresponding total table sizes are 1330, 1166, 920, 838 and 756 bytes respectively.

Let us consider here again the T configuration. The numbers of required transmissions to obtain the 95% coverage in the T network configuration for

Table 5.3: Size of INT table with different section sizes (in bytes)

	Total size of the table		
Section size	LC	T	MM
203	6857	13843	18963
512	4921	9971	13507
1024	4217	8739	11747
2048	3865	8035	10867
4096	3689	7683	10515

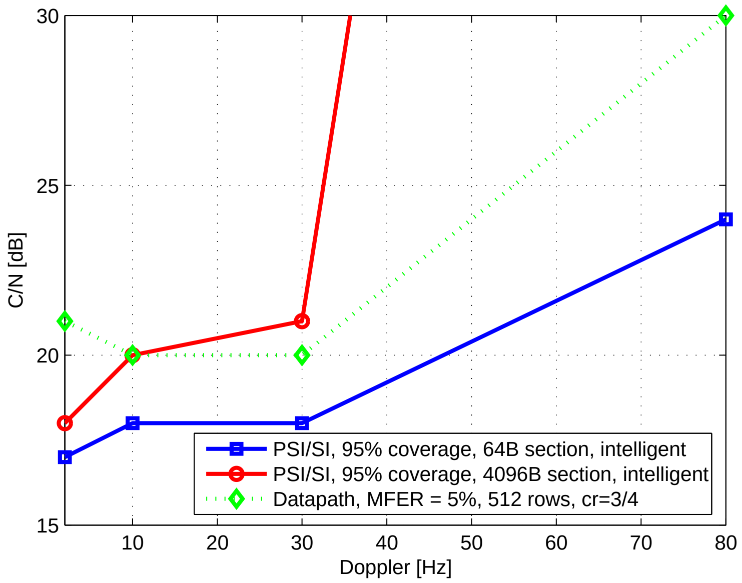

Figure 5.12: Comparison of different section sizes for 16384B table (TU6, MM network configuration, MPE-FEC code rate 3/4 for data path).

the PSI/SI at the C/N where the MFER=5%, are shown in Table 5.4. For different section lengths the worst case of the Doppler frequencies needs to be considered (marked with bold face font in Table 5.4), that is the largest amount of transmissions is the limiting factor for each section size. When planning the network for mobile users, the worst case acceptable velocities of the receivers should be taken into account. For example, INT with 512B sections needs to be transmitted 5 times to obtain the required coverage. The required network capacity for the transmission of the tables can then be calculated if we assume that it is necessary to reach the coverage of 95 % of users within the maximal repetition intervals given in Table 5.5 (collected from the standards [47], [59] and [71]) at any Doppler frequency. Fig. 5.13 presents the reserved network capacity as a function of the section length for INT and NIT. From the network capacity point of view, the optimal section length for INT is 512B and for NIT it is 756B.

The total sizes for the tables using sections of these sizes are calculated to be: INT 9971B, NIT 756B, PMT 394B, PAT 16B and TDT 8B. The overall network capacity reserved by the five tables using these optimal section sizes is calculated to be: $\frac{9971\times8}{(30s/5)} + \frac{756\times8}{(10s/3)} + \frac{394\times8}{(0.1s/2)} + \frac{16\times8}{0.1s} + \frac{8\times8}{30s} \approx 77.6kbps$, that is less than 1 % of the total network capacity of 9.95 Mbps in T net-

Table 5.4: Necessary number of transmissions to reach 95% coverage at C/N corresponding to MFER = 5% (T) for different section lengths

Table				
INT	f_D=2Hz	f_D=10Hz	f_D=30Hz	f_D=80Hz
203B	2	3	**5**	4
512B	2	3	**5**	4
1024B	2	3	**6**	4
2048B	2	4	**9**	6
4096B	2	4	**19**	10
NIT				
96B	1	2	**3**	2
128B	1	2	**3**	2
256B	1	2	**3**	2
512B	1	2	**3**	2
756B	1	2	**3**	2
PAT				
16B	**1**	**1**	**1**	**1**
PMT				
394B	1	**2**	**2**	**2**
TDT				
8B	**1**	**1**	**1**	**1**

work configuration. Most of this network capacity is reserved by PMT that requires 61.6 kbps due to the short repetition interval.

Simulations based on field measurements

The field measurements were performed in the city of Turku, Finland on a non-hierarchical DVB-H signal with the center frequency 498 MHz in two transmitter Single Frequency Network (SFN), with the transmitters located approximately 4 km from each other [76]. Physical layer parameters measured were the ones for "Typical" configuration: 16-QAM modulation

Table 5.5: Repetition interval ranges for the PSI/SI tables used in DVB-H

Table	**Min**	**Max**
NIT	25 ms	10 s
PAT	25 ms	100 ms
PMT	25 ms	100 ms
INT	25 ms	30 s
TDT	25 ms	30 s

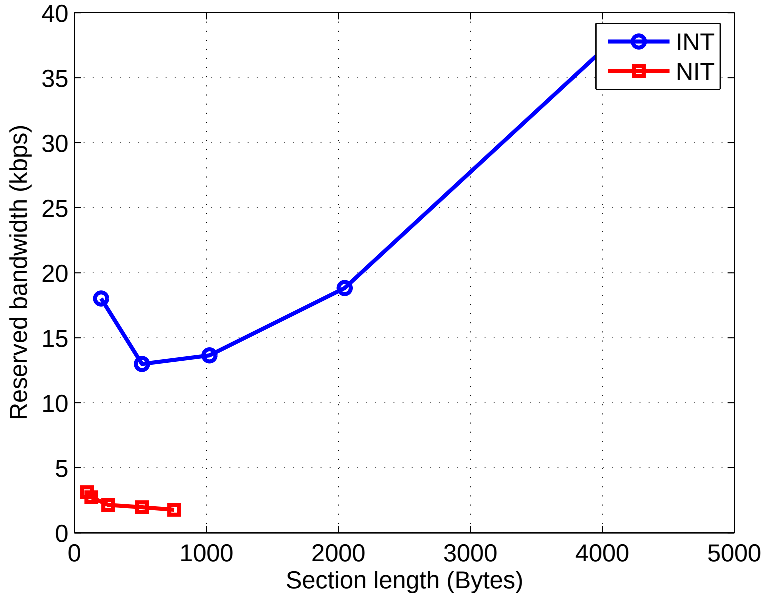

Figure 5.13: Network capacities reserved by INT and NIT (TU6, T network configuration).

with convolutional code rate 1/2, 8k OFDM mode and guard interval of 1/4 of the OFDM symbol duration. The receivers were the same as were used for obtaining the TU6 error traces used in the previous analysis. In the measurements four use cases were considered: Pedestrian outdoor (3 km/h), Pedestrian indoor (3 km/h), Vehicular urban (30 km/h) and Motorway (100 km/h). The block diagram for the measurement setup is shown in Fig. 5.14. The received signal from the antenna is fed to two identical receivers through a signal splitter. The receivers output Received Signal Strength Indicators (RSSI) and a TS packet error trace that are stored together with GPS information on a laptop for post processing. The recorded transport stream packet error information is then used for simulating the PSI/SI performance.

In the pedestrian use cases the measurement system was carried in a backpack with the antenna outside the pack and in the vehicular measurements the antenna was located on the rooftop of the car. The indoor measurements were performed in a shopping center with a variation of open squares with glass roofs and narrow passageways. Measurements in both pedestrian use cases were performed close to the transmitter in the city center; the vehicular urban use case was measured between the two transmitters in an urban environment and the motorway use case in the coverage area of the farther transmitter in rural surroundings.

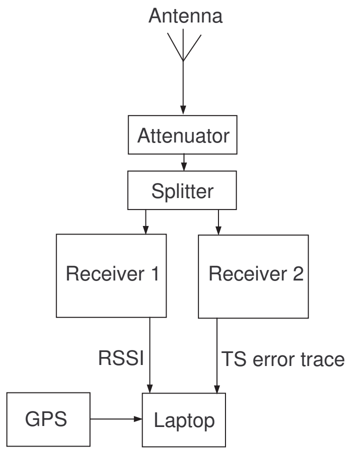

Figure 5.14: Measurement setup.

The results of simulations on PSI/SI transmission based on the measured TS packet error information in the four usage scenarios are presented in Fig. 5.15. The INT table of total size 7683B in T network configuration transmitted with 4096B sections is shown as an example. An attenuator was used to obtain such a signal level that errors do occur in the reception. During one measurement the attenuation was kept constant. The MFER values for the measurements used in the simulations shown in Fig. 5.15 are: Pedestrian indoor = 53%, Pedestrian outdoor = 10%, Vehicular urban = 5% and Motorway = 35%. An important notion of the figures is that even when the reception of the application data can be considered to be impossible, for example in the pedestrian indoor case with MFER=53%, 95% coverage for the INT table can be reached already with 6 transmissions. The number of necessary retransmissions could be reduced by using shorter sections as was learned with TU6 simulations, but the result with 4096B sections serves as the worst situation. It is clear that the distribution of TS errors is different in the use cases. For example, in pedestrian indoor measurements there are error bursts of duration reaching up to several seconds due to shadowing and sometimes the reception can be perfectly clear resulting in quite good overall PSI/SI transmission. As a contradictory example, in the motorway measurements TS packet errors are rather evenly distributed due to shorter shadows and Doppler effect and the performance of PSI/SI transmission is quite poor. The curves in the figure are not directly comparable to one another since there are fluctuations in signal strength, but the important issue is that it can be verified that PSI/SI transmission actually works in a real mobile environment also.

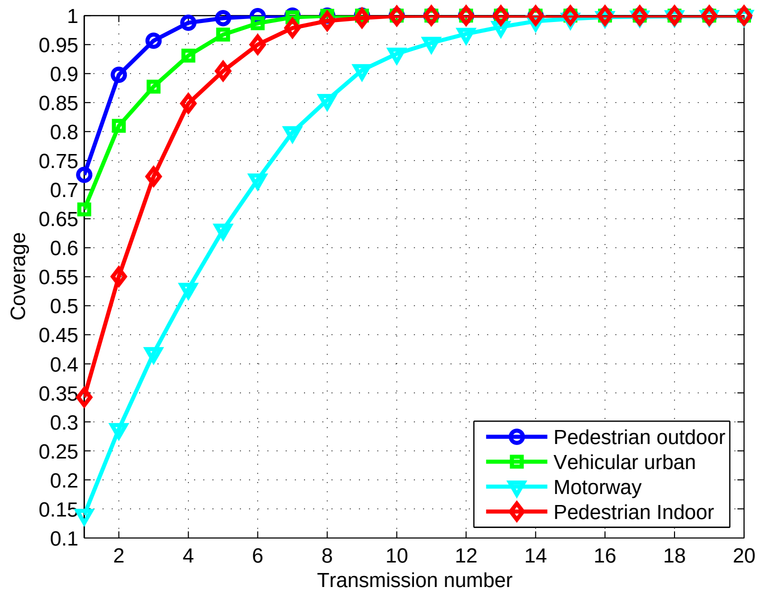

Figure 5.15: INT table coverage in measured environments as a function number of transmissions (T network configuration).

In this section simulations on the performance of PSI/SI transmission in DVB-H systems in a mobile environment were presented. The effects of repetition interval, section size and decoding method on different channel conditions and network configurations were presented. First of all, the intelligent decoder should be used due to its superior performance over the other two decoding methods. Based on the simulation results on the robustness of the signaling it is advisable to use as short sections in the transmission of PSI/SI as possible. If the transmission of PSI/SI is further optimized with respect to the used network capacity and the impact of varying section size on the total table size is taken into account, the smallest section size is not necessarily the best option. Still, an optimal section size can be found for example with the help of simulations as described in this section. It seems according to the simulations with TU6 channel and field measurements that the PSI/SI transmission as it is organized presently is able to provide robust enough transmission when network design effort is put on the selection of the PSI/SI transmission parameters.

5.2.2 Improving the performance for the L2 signaling

Consider now that the large tables (that is INT and NIT) could be transmitted within the scope of MPE-FEC. This kind of approach is not consistent with the present standards, but studying it will provide valuable knowledge for design of future systems. Let us assume that one service with MPE-FEC frame of size 256 rows is allocated for these tables. This frame is capable of carrying in total $191 \times 256 = 48896$ bytes of information so that in most situations the INT and NIT tables will fit in it. Let us contemplate a scenario of transmitting only INT within MPE-FEC and where the size of the INT is 16384B. This table fills exactly 64 columns of the MPE-FEC frame resulting in MPE-FEC code rate 1/2 if all 64 redundancy columns are transmitted. Further the CRC-32 information of sections could be used as erasure information for the MPE-FEC decoder. The addressing of the sections in the MPE-FEC frame could be calculated from the section header information such as section length and number. For the T network configuration the MPE-FEC frame error rate for this kind of frame using sections of 4096 bytes at $C/N = 16$ dB and $f_D = 10$ Hz is simulated to be approximately 5 %. Thus the coverage of 95 % is reached with just one transmission. If the INT were transmitted conventionally in the same channel conditions using also 4096B sections it would require 13 transmissions to obtain the corresponding coverage (Fig. 5.9). If the coverage is to be reached within 30s, the bandwidth reserved by the proposed enhancement is $(16384 \times 8 + 64 \times 256 \times 8)/30 \approx 8.53$ kbps (the additional data is the redundancy information of MPE-FEC). For the conventional transmission the necessary repetition interval is $30/13 \approx 2.3$s, leading to bandwidth $16384 \times 8/2.3 \approx 55.7$ kbps. It is clear that including the INT to the scope of MPE-FEC would benefit in addition to the error rate also in the bandwidth used for PSI/SI transmission. The problem with the long repetition interval is that if the PSI/SI data is not stored in non-volatile memory in the receiver or the receiver is turned on for the first time in the network, there is potentially a long delay for every user before the reception of the service can be started. For shorter repetition intervals the users in the good conditions can receive the signaling more quickly and the unfortunate users in harder reception conditions have to wait for several retransmissions. Therefore a compromise between the repetition interval and the reserved bandwidth must be made. In Fig. 5.16 the error rate of a table of size 16384B transmitted conventionally with 4096B sections and within a MPE-FEC frame of size 256 rows (code rate 1/2) are compared at $f_D = 10$ Hz and clear advantage in terms of decibels for the considered enhanced scheme is visible. At error rate 0.05 gain for MPE-FEC utilization of approximately 5 dB is observed.

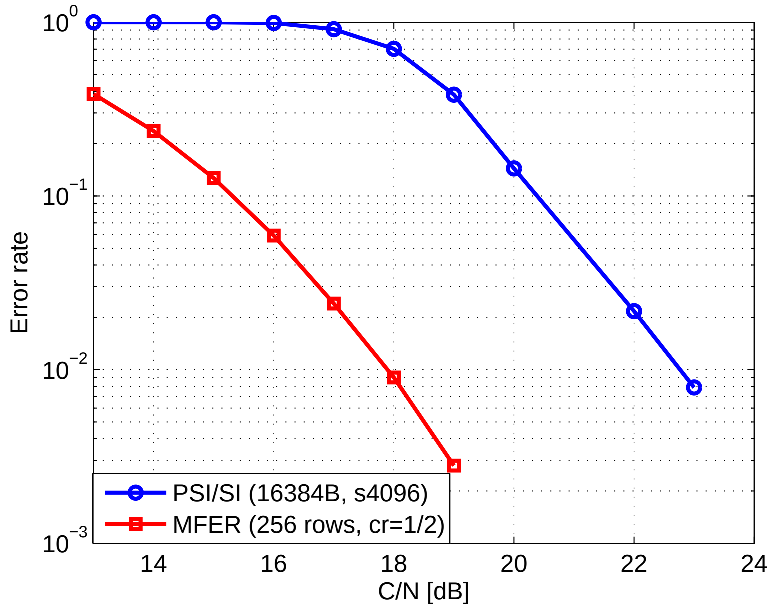

Figure 5.16: Error rates of INT transmission and considered enhancement (TU6 $f_D = 10$Hz, T network configuration, MPE-FEC code rate 1/2 for the studied enhancement).

For the DVB-H receiver, the introduction of L2 "signaling bursts" would slightly increase the power consumption of the terminal as reception of additional bursts would be necessary. Also, additional memory for the signaling MPE-FEC table would be necessary. Therefore, such changes in the already stable DVB-H system definition are not imaginable. Still, currently something very similar is implemented in the second generation DVB-T2, where common parts of L2 signaling can be transmitted in a common PLP that practically is a L2 signaling burst transmitted in the beginning of T2 frames. This kind of approach is natural when the transmission system utilizes a time domain framing structure. The signaling and its transmission in DVB-T2 systems will be studied in detail in chapter 6.

Chapter 6

Case study: DVB-T2 physical layer design and analysis

The design process of the mechanisms for providing QoS in the DVB-T2 system physical layer is studied in this chapter. The selected mechanisms and possible configurations allowed by the standard are analyzed by system simulations. The view on the design process bases on participation in the DVB organization technical module defining the standard. Therefore, the focus is not on design of all aspects of the standard but rather on selected items, namely utilization of diversity, physical layer signaling design and use of computer simulations for design and analysis purposes. The observations from the design process assist in future system design and these issues are further discussed in chapter 7. First, the mechanisms for obtaining time and frequency diversity for the DVB-T2 transmission are studied. Then the utilization of simulations for the design of the diversity mechanisms is presented. Further, the design process for the physical layer signaling is studied and the performance of the signaling and service data in a terrestrial environment is investigated. Finally, a mechanism and guidelines for selecting the transmission parameters for the signaling in combination with the parameters for the service data are presented.

6.1 Mechanisms for diversity in the service data path

In DVB-T2 the error control coding scheme consists of concatenated LDPC and BCH coding. The BCH code is required mainly to remove the error floor that is typical for LDPC codes decoded with belief propagation iterative decoders. One BCH codeword represents exactly the contents of one LDPC codeword. The error correction capability of 10-12 erroneous bits within each codeword is considered to be enough to remove the error floor from the LDPC decoder.

LDPC codes were discovered in 1962 by Robert Gallager in his Doctoral Thesis [77]. At that time these codes were not considered to be practical due to unavailable computational resources in the technologies of the time. It was only after the success of the turbo codes in the 1990s that the LDPC codes were re-discovered by MacKay and Neal [25]. Since then, there has been extensive research on LDPC codes. As opposed to turbo codes, the decoding of the LDPC codes can be easily parallelized and the degree of the parallelism can be adjusted in the design phase [78]. The more parallel is the decoder implementation used, the more silicon area it consumes but respectively the decoding process requires less clock cycles.

LDPC codes are linear codes that are specified, as already their name indicates, by sparse (Low Density) parity check matrices. The sparse matrices contain a small number of ones as compared to zeros. An example of a parity check matrix as shown in [79] (not sparse though) could be:

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

This matrix shows that 1st, 4th, 5th and 8th bit of the codeword contribute to the first parity check and 2nd, 3rd, 5th and 7th contribute to the second parity check etc. For a valid codeword the modulo-2 sums of all bits contributing to parity checks are zero, that is

$$\mathbf{H}\mathbf{c}^T = \mathbf{0}, \tag{6.1}$$

where $\mathbf{c} = (c_1, c_2, ..., c_n)$ is codeword vector and n is the codeword length.

For the encoding of the linear block codes, the parity check matrix can be transformed to a generator matrix by for example Gaussian elimination [80] and then matrix multiplication is performed to obtain the encoded codeword. Unfortunately, the generator matrix of a sparse parity check matrix is not necessarily sparse [78]. Storing the dense generator matrices instead of non-dense parity check matrices for usually very long LDPC codes needs a significant amount of memory. To save the memory and to make the encoding computationally less expensive, the structure of the LDPC codes used in DVB-T2 is designed so that encoding basing directly on the parity check matrix is possible. The structure enabling linear encoding complexity for the T2 LDPC codes is following: $\mathbf{H}_{(n-k)\times n} = [\mathbf{A}_{(n-k)\times k}\mathbf{B}_{(n-k)\times(n-k)}]$ where $\mathbf{B}$ is a staircase lower triangular matrix. The general structure of the parity check matrices for T2 LDPC codes is visualized below. Both parts **A** and **B** for the matrix **H** are sparse.

$$\mathbf{H} = \begin{bmatrix} 0 & 0 & 0 & 1 & . & . & . & 1 & 0 & 0 & . & . & . & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & . & . & . & 1 & 1 & 0 & . & . & . & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & . & . & . & 0 & 1 & 1 & . & . & . & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & . & . & . & 0 & 0 & 1 & . & . & . & 0 & 0 & 0 \\ . & . & . & . & . & & & . & . & . & . & & & . & . & . \\ . & . & . & . & & . & & . & . & . & & . & & . & . & . \\ . & . & . & . & & & . & . & . & . & & & . & . & . & . \\ 0 & 0 & 0 & 0 & . & . & . & 0 & 0 & 0 & . & . & . & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & . & . & . & 0 & 0 & 0 & . & . & . & 0 & 1 & 1 \end{bmatrix}$$

Now, the information block $\mathbf{i} = (i_0, i_1, ..., i_{k-1})$ can be encoded recursively to systematic codeword $\mathbf{c} = (i_0, i_1, ..., i_{k-1}, p_0, p_1, ...p_{n-k-1})$ where the information is in the beginning and the redundancy is in the end as presented in [78]:

$a_{0,0}i_0 + a_{0,1}i_1 + ... + a_{0,k-1}i_{k-1} + p_0 = 0 \Rightarrow$ solve p_0

$a_{1,0}i_0 + a_{1,1}i_1 + ... + a_{1,k-1}i_{k-1} + p_0 + p_1 = 0 \Rightarrow$ solve p_1

$\vdots$

$a_{n-k-1,0}i_0 + a_{n-k-1,1}i_1 + ... + a_{n-k-1,k-1}i_{k-1} + p_{n-k-2} + p_{n-k-1} = 0 \Rightarrow$ solve p_{n-k-1}

The LDPC codes used in DVB-T2 are irregular meaning that different number of parity checks correspond to different bits, that means the bits have different *degree*. In general, irregular LDPC codes have better performance under belief propagation decoding than regular ones. The reason for this is that the higher degree bits obtain more information from the several parity checks and they get corrected with a small number of decoding iterations. Once they are correct they help to recover the bits with smaller degree. Codewords of lengths 16200 and 64800 bits are used in DVB-T2 and individual parity check matrices are defined for all supported code rates. The DVB-T2 codes are reported to perform 0.7 dB to 0.8 dB from the Shannon limit in ideal memoryless channels [78]. This would indicate that there is not much space for further improvements on the error control coding mechanism itself. But of course it is the overall performance of the system in challenging reception conditions that the designers are interested in, not only the performance of individual system blocks such as error control coding in ideal channels. Therefore the utilization of diversity plays an important role in the overall system performance for a terrestrial transmission system.

To utilize both time and frequency diversity, there are several physical layer interleavers in the DVB-T2 system. Bit, cell, time and frequency interleavers are specified. The distribution of the information in time and frequency by the interleaver chain is illustrated in Fig. 6.1. In the figure one time interleaver block is considered to contain only four LDPC codewords for simple illustration. The contents of the codewords are shuffled by bit and cell interleavers. The bit interleaver consists of parity and block interleaving. First the LDPC parity bits are interleaved and then the codeword interleaved by a block interleaver to which the data is inserted column by column applying circular shift for each column. This is called in the standard *column twist interleaving*. The cell interleaver is a pseudo random interleaver that is applied over the cells (constellation points) of LDPC block. It is noteworthy that both bit and cell interleavers shuffle information only within the codewords.

The time interleaver is the main tool for obtaining time diversity for DVB-T2 transmissions. The time interleaver is a block interleaver, where each LDPC codeword fills exactly five columns of the interleaver matrix. The frame builder maps the output of the time interleaver to the subcarriers of the OFDM symbols. Further, the frequency interleaver spreads the data in frequency (subcarrier) direction. Thus, as illustrated in Fig. 6.1 both time and frequency diversity is obtained by the interleaving chain. The amount of time diversity can be adjusted by selecting the size of the time interleaving (TI) block. The maximum size of the time interleaver block is set by the size of the time interleaver memory defined in the DVB-T2 standard to be $2^{19} = 524288$ cells. The frame builder extends the operations of the interleavers when subslicing or interleaving over several T2 frames is used. When performing the subslicing, the TI block is divided to several bursts in one T2 frame. When additional time diversity is required the TI block can be split to be transmitted over several T2 frames. Also, if TFS (Time Frequency Slicing) is used, frequency diversity over several RF channels can be obtained. In TFS the transmission multiplex consists of up to six RF channels that are transmitted in a synchronized manner and the received PLP is received from different RF channels in successive T2 frames. The selection of time interleaving parameters has an effect on the QoS. For example having longer TI block gives better error rate performance but increases also the latency due to buffering need. These configuration trade-offs are discussed further in section 6.1.2.

6.1.1 Using simulations for design of modulation, coding and interleaving

It was stated in the DVB-T2 call for technologies [50] issued by the DVB project, that the DVB-S2 LDPC code was the working assumption for the error control code to be used in the system. It was further stated that if the DVB-S2 LDPC code were not suitable for terrestrial channels, other channel coding schemes could be considered. Let us next describe a straightforward way for investigating the suitability issue.

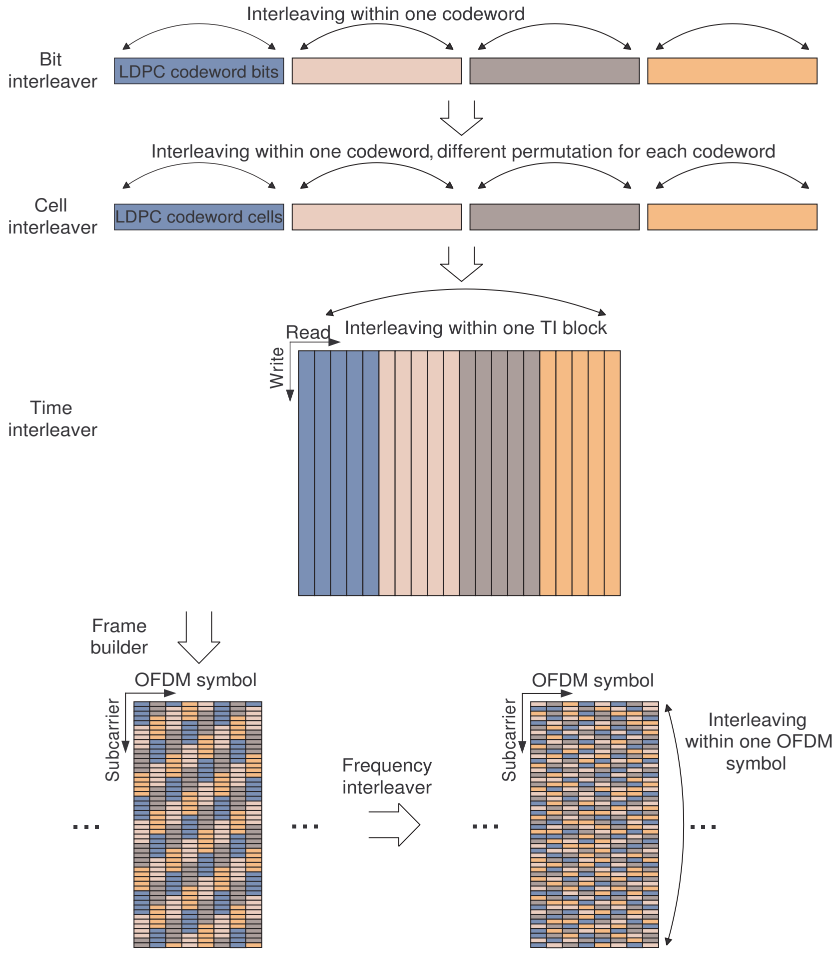

Figure 6.1: Illustration of the interleaving operations in DVB-T2.

The suitability of the DVB-S2 LDPC code to terrestrial environments can be studied with simulations based on a DVB-T system simulator, assuming that there will be significant similarities between DVB-T and -T2 in system parts other than the error control coding. Such a study is presented by the author in [81], where the error control coding scheme of the DVB-T system is substituted with the LDPC code from DVB-S2. The BCH code of the DVB-S2 is not implemented in the simulator, as its effect on the performance is very small when the error floor is not reached by the simulations. In the study a multipath channel model is used to represent a terrestrial mobile environment. A mobile environment is studied because it is considered

Table 6.1: Simulated system parameters

Parameter	Value
Guard interval	1/4
FFT size	8k
Modulation	16-QAM
LDPC code length	64800 bits
LDPC max. Iterations	50
Coderates	1/2, 2/3, 3/4, 5/6

to be more challenging for the error control coding than a static terrestrial environment. The simulated system parameters are shown in Table 6.1.

In order to allow for soft decision decoding of the LDPC codes in the receiver the demapper of the DVB-T receiver is modified to output bit level LLRs (Log-Likelihood Ratios). Calculating exact LLR values involves calculating distances of the received signal point to all the constellation points. For high order modulations this is a tedious job. In the simulations approximate LLRs were calculated. In the calculation of approximate LLRs only the nearest constellation points with bit values 1 and 0 in corresponding bit location are considered rather than all constellation points. For Gray coded QAM signals as in DVB-T, the LLR approximation can be based on the decision regions due to the bit ordering of Gray coding. A mechanism for this kind of LLR calculation is presented in [82]. LDPC decoding is performed by the belief propagation algorithm. Decoding is an iterative process where the reliability (soft) information of the received bits is refined iteration by iteration. Iteration consists of computing parity-check sums and updating reliability information based on the results of the parity-checks. The output of iteration is used as an input for the next one. This process of updating the reliability continues until all parity checks are fulfilled or maximum iteration count is reached. Maximum iteration count of 50 was used in the simulations of this section as it was observed that the decoding performance did not increase significantly when further increasing the iteration count. After this, hard decisions on the symbols are made to come to an estimation of what was sent. For more detailed information on the decoding algorithms as well as the LDPC codes themselves, interested reader can turn to [10].

The gain of using the LDPC code instead of the DVB-T coding in presence of an AWGN channel is visualized in Fig. 6.2. AWGN is commonly used for performance evaluation of the coding schemes to have a reference comparison. It should be noted that the true code rate of the DVB-T concatenated scheme is lower than that of LDPC because of the concatenation of the RS code (true code rate $(1/2) \times (188/204) = 0.46$ instead of code rate 1/2 for example). LDPC code rate 2/3 is compared to the DVB-T

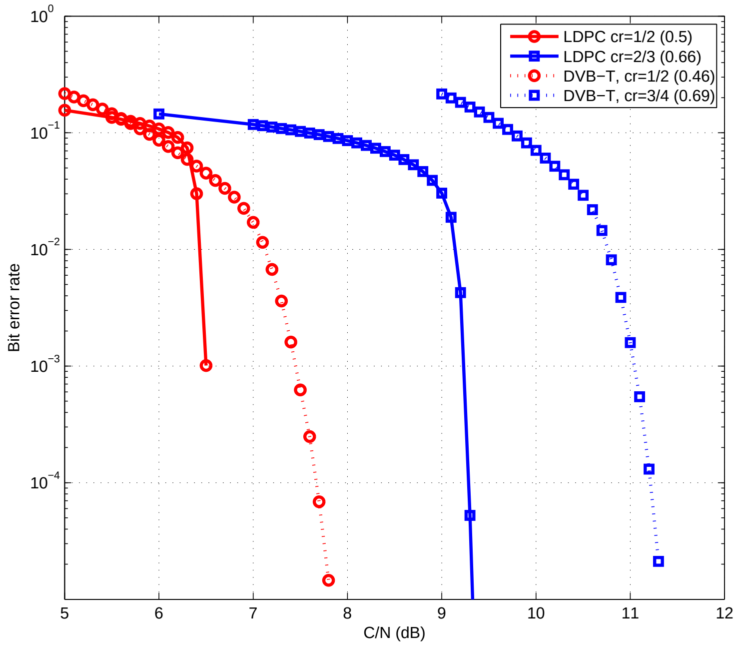

Figure 6.2: Comparison of LDPC and DVB-T coding (AWGN channel, 16-QAM, 8k mode, guard interval 1/4).

RS-convolutional coding scheme with convolutional code rate 3/4, because the true code rate of the DVB-T coding with convolutional code rate 3/4 (shown in parentheses in the figure) is closer to that of LDPC 2/3 than for the DVB-T coding with convolutional code rate 2/3. Gains of 1 dB for rate 1/2 and 2 dB for rate 2/3 LDPC code are visible in the AWGN channel at error rate 10^{-4} when slightly extrapolating the curve for LDPC 1/2. It is also observed that the slopes of the BER curves for LDPC codes are steeper than those of the RS-convolutional coding.

Performance comparison of the coding schemes in TU6 channel with f_D=10Hz is shown in Fig. 6.3. This Doppler frequency corresponds to relatively slow receiver movement assuming that the center frequency of the transmitted signal is around 500 MHz. The true code rates are again given in parentheses in the figure. The coding schemes with rather similar overall code rates are compared. Coding gain for the LDPC code over the DVB-T coding is observed. At bit error rate level 10^{-5}, for true code rates 1/2, 2/3, 3/4 and 5/6 gains of 0.6 dB, 2.5 dB, 3 dB and 2.7 dB respectively are observed. On the other hand, it can be seen, that for example for LDPC code

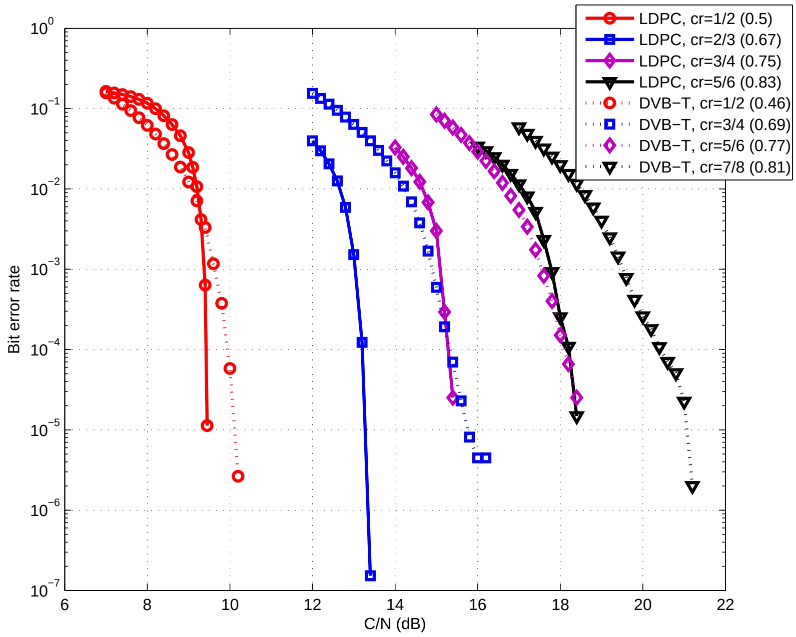

Figure 6.3: Comparison of LDPC and DVB-T coding (TU6 f_D = 10Hz, 16-QAM, 8k mode, guard interval 1/4).

rate 3/4 and DVB-T coding with convolutional code rate 3/4 almost similar BER is obtained with similar C/N values. Thus, with LDPC code higher data throughput is obtained with similar C/N value since true code rate for LDPC code is 0.75 and for DVB-T coding it is 0.69. In this exemplary case, the gain of LDPC code in data throughput is nearly 9%.

As a conclusion, using a DVB-T simulator it is observed that there is potential gain in using the DVB-S2 LDPC codes in a terrestrial system. Further the LDPC code is likely to show better better performance than what is shown by the simulations when using a more optimal LLR calculation scheme than the approximate calculation used in the simulator.

It is known from the measurement studies and the introduction of the DVB-H system for mobile services that the DVB-T system as it is does not have enough time diversity to combat the challenges of fading and channels with impulsive noise [16]. To have additional robustness against these kinds of interference in the second generation DVB-T2, longer time interleaving than what is in the present DVB-T system is considered necessary. Also the effect of adding an interleaver can be studied simply by adding a block interleaver to the DVB-T simulator with the LDPC code. A block interleaver

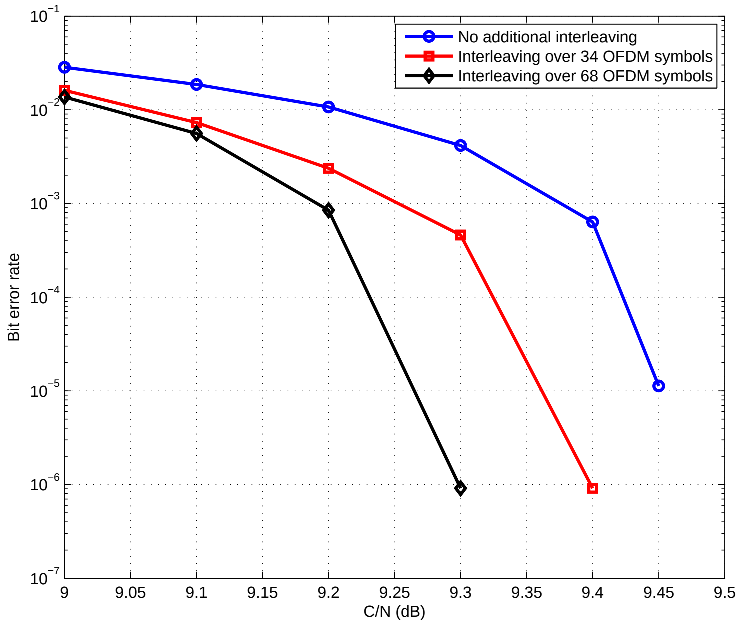

Figure 6.4: The effect of additional time interleaver (TU6 $f_D = 10$Hz, 16-QAM, 8k mode, guard interval 1/4).

is selected as it is a natural choice with a block error control code. Let us further study the effect of adding a time interleaver over 34 and 68 OFDM symbols (that is over half and one DVB-T OFDM frame) to our simulation system. Interleaving depths of about 13 and 25 LDPC codewords are obtained with this exemplary interleaver. With the used system parameters this corresponds to approximately 38 and 76 milliseconds. The results with a code rate 1/2 LDPC code in the TU6 channel with f_D=10Hz using the additional time interleaver are shown in Fig. 6.4. It is observed that some gain (0.1dB - 0.2dB) can be obtained already by adding this rather un-optimized interleaver to the system. A time interleaver providing better results in this particular system should be found by careful system design. Also, the interleaving length provided by the interleaver is not enough for the simulated channel as the coherence time of the channel is in the order of 40 milliseconds calculated with equation (2.2). Preferably the interleaving length should be several coherence times to allow for the channel averaging effect. Therefore, longer time interleaving should be incorporated into the DVB-T2 system if such a mobile scenario is to be supported.

As a more efficient FEC was defined to be used in DVB-T2, higher order modulation schemes can be considered. The FEC compensates for the high SNR required by high order modulation schemes. Therefore, up to 256-QAM is specified in DVB-T2 for maximal system throughput. Also, for services destined for stationary receivers with rooftop antennas, the signal is not often severely distorted and the performance with high order modulation is sufficiently good.

Already the call for technologies for DVB-T2 requested for simulation results to support the proposed technologies. In addition to the initial study on the suggested technologies, the DVB-T2 system was simulated extensively during the design phase. For this purpose, a Common Simulation Platform (CSP) based on DVB-T was used throughout the design process. The CSP was agreed to be used and updated by the organizations developing new parts of the DVB-T2 standard. The simulation platform evolved alongside with the specification, thus allowing the evaluation of the effect of the design selections on the overall performance of the system under development. Such application of simulations is very useful and allows in the best case to avoid bad design decisions and justify the good ones thus speeding up the design process. Further, the simulation platform has been used as a reference in the verification process for real DVB-T2 implementations.

6.1.2 Time interleaving configuration analysis

Let us consider here two different PLP interleaving configurations to illustrate the flexibility of PLP allocation provided by the DVB-T2 standard. What mainly can be configured in DVB-T2 is the time interleaving length and mapping of the service to T2 frames. The two selected configurations are illustrated in Fig. 6.5. Configuration A represents what is called in the standard *Mode A*, when only one PLP is used for the transmission and the data rate of the whole multiplex is allocated for this PLP. As the time interleaver memory is limited, three time interleaver blocks are required in each T2 frame when the duration of the T2 frame is configured to be 250 ms. Configuration B represents a situation where there are several PLPs transmitted in the DVB-T2 system. This is called *Mode B* in the standard. In configuration B one TI block of the considered PLP is transmitted in two successive T2 frames. In mode B the time interleaver configuration as well as other transmission parameters for each PLP can be individually configured.

Mode A is designed for transmitting several services multiplexed in one transport stream as is done in DVB-T. Using the mode A transmission the distribution network for the DVB-T transmitters can be easily utilized for introducing DVB-T2 networks. Also, statistical multiplexing can be performed inside the PLP similarly as it is done in DVB-T [83]. The drawbacks of mode A as compared to mode B are that all the services have the same

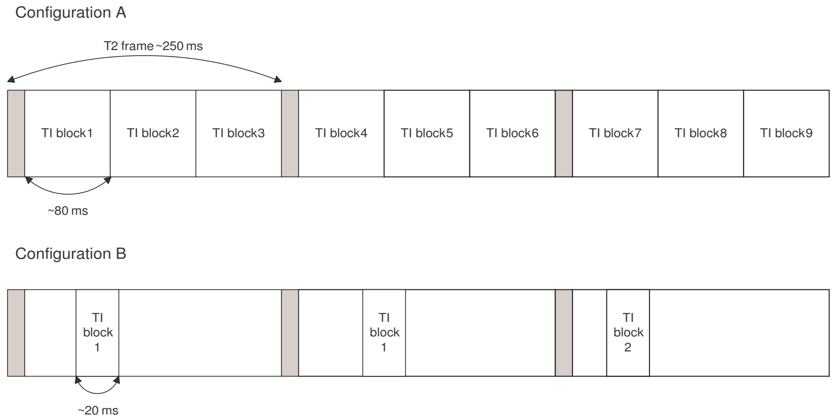

Figure 6.5: The compared time interleaving configurations.

robustness level (same modulation and coding) and the time diversity that can be obtained for the services is limited by the time interleaver memory being shared by all the services. In addition, the receiver physical layer needs to decode all the services in the multiplex and the link layer would filter only the desired service from the transport stream.

Mode B is designed for transmitting services in the same network with different robustness levels. One PLP can carry one or several services. In mode B services for mobile receivers can be transmitted with low code rate and low order modulation in some PLPs and for example HDTV channels destined for stationary receivers can be transmitted with high code rate and high order modulation in other PLPs. Further, mode B may enable use of scalable video coding (SVC) [84] in the future in a way that different layers of the video are transmitted with different robustness levels in different PLPs. This would enable graceful degradation for the video and also sharing video components between mobile and stationary service consumers. Mode B enables also mapping the time interleaver blocks over several T2 frames and use of subslicing and therefore provides means for further time diversity for the services as compared to the mode A. For mode B the statistical multiplexing can be done over the PLPs to obtain maximal utilization of the available bandwidth. In general, very long TI lengths can be obtained for mode B, but the channel switching time increases correspondingly. Therefore the network operator needs to consider trade-offs between the robustness and the channel switching time when designing the PLP configurations to be used in the network.

The performance for both configurations A and B in a mobile multipath TU6 channel with f_D=80 Hz with respect to SNR (E_s/N_0) are shown in

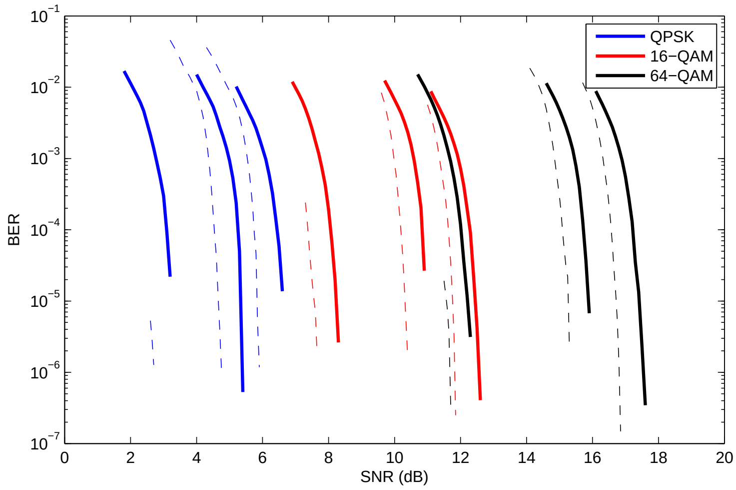

Figure 6.6: The performance DVB-T2 mode A (dashed line) and B (continuous line) PLPs (TU6 $f_D = 80$Hz, 8k mode, guard interval 1/4). Code rates 1/2, 2/3 and 3/5 from left to right.

Fig. 6.6. Mobile varying channel model is used to observe differences for the interleaving configurations, as this is the case for which the interleaving is mainly intended for. For configuration A the whole time interleaver memory is used while for the configuration B only half of the memory is used. Due to mapping and interleaver configuration the data rate of the PLP in configuration B is 1/12 of the data rate for configuration A PLP. The curves for code rates shown from left to right for each modulation scheme and for both interleaver configurations are 1/2, 2/3 and 3/5. Dashed lines represent configuration A and solid lines configuration B. It is observed that the performance for configuration A in this channel is better as the whole time interleaver memory is used. For slower varying channels (for example coherence time of around 100 ms) configuration B could outperform configuration A, because one deep fade in the channel could corrupt the whole TI block for configuration A while for configuration B only half of the TI block would be severely faded. All in all, the selection of a suitable time interleaving configuration can well be assisted by simulations.

6.2 Signaling path

The transmission scheme of the physical layer signaling carried in the preambles of the DVB-T2 system differs from that of the data path. As the signaling enables the reception of the actual data, it should naturally be more robust against channel impairments than the data itself. This is to be taken into account already in the system design process when defining the performance requirements for the signaling. Therefore, the robustness of the transmission of the L1 (physical layer) signaling needs to be studied and compared to the robustness of the data path. Here the L1 signaling and its design process are studied. Further, a method for selection of L1 signaling transmission parameters based on the simulation studies is presented.

6.2.1 Design process for signaling

The transmission scheme selected already in early design phase for DVB-T2 is OFDM with time division multiplexing for the services. Also, it was selected that the physical layer signaling is to be transmitted in preambles. This system approach facilitates the use of time frequency slicing that is enabled by the DVB-T2 specification. In TFS the signal is transmitted on several RF channels in a synchronized manner and the signaling should be available on all the channels. This decision sets limitations but also possibilities for the signaling transmission solution. The duration of the preamble should be adequate to accommodate necessary signaling information and on the other hand the overhead introduced by the preamble should be minimized. The signaling information should be retransmitted at suitable intervals to allow the receivers to tap into the network when powering on. As the preamble based system approach was taken, the duration of the frame to which the preamble is attached sets the repetition interval for the signaling. Further, the duration of the preamble together with the requirement for the initial scan duration for the receiver set limitations to the time diversity that can be achieved for the signaling.

Translating the QoS requirements for the whole system to functional and performance requirements for the signaling is not as straightforward as it is for the data path. The relationship of for example error rate and the delay of the signaling and the corresponding parameters for the services themselves need to be studied. This requires knowledge on the receiver operations for both service data and signaling reception. In general, the requirements for the signaling should be more strict than for the data due to the enabler nature of the signaling. As an example, if the service switching time for the system should be smaller than one second, the physical layer signaling should allow for switching to receive a different PLP in a much shorter time period as there are delays also on other system layers (sychronizing the video codecs, receiving the upper layer signaling and so on). The same applies for error rate. The signaling should be error free at a signal quality where the performance requirement of the error rate for the data is reached. This way no additional errors are introduced by the signaling to the service reception and signaling and data paths can be designed individually taking into account the designed robustness for the service data.

As the purpose of the signaling is to aid the receiver in receiving the services, the starting point for the signaling design is on the receiver operations and reception environment. The main use cases regarding broadcast receiver operations for the static receiver are *initial scan*, *channel switching* and *normal data reception*. The signaling designed for DVB-T2 should support at least these. Further, if mobile reception of the broadcast is to be considered the signaling should support *handover* from one transmission area to another. When beginning the initial scan, the receiver has no knowledge of the transmission network and has to start from scratch to resolve the transmission parameters to begin the service reception. This is the most time consuming of the considered use cases. In the channel switching use case the receiver is already consuming services on the network and the user is willing to switch to another service. In the normal data reception use case, the receiver keeps receiving the current service. To perform the handover in an efficient way, the signaling content should provide knowledge about other networks as well. Based on these use cases and the selection of the preamble based solution the required signaling content and the transmission mechanism can be designed.

For the initial scan purposes the preamble structure must present an entry point for the receiver. In DVB-T2 this entry point was designed to be P1 symbol and L1-pre signaling in the P2 symbols. The P1 symbol allows the synchronization and signals parameters that are fundamental, such as the FFT size. After the P1 symbol is received and synchronization performed the receiver obtains the P2 symbols. The number of preamble P2 symbols was designed to be known based on FFT size and this is known by the receiver as the P1 signaling is decoded. Now that the receiver is provided with the means to receive the correct amount P2 symbols, the design question arises, how should the L1-pre information be protected and mapped to the subcarriers of P2 symbols to give the required performance. Also, the contents of the L1-pre signaling are to be selected. As the reception of services is not possible without obtaining the L1-pre first (it being the entry point), the L1-pre is encoded by the rate 1/5 LDPC code and BPSK modulated to make it very robust against interference. Further, the L1-pre is interleaved over all P2 symbols to obtain best possible time diversity within the preamble to combat impulsive noise and fading.

Naturally, the whole physical layer signaling package could have been coded and modulated with the strongest code rate and lowest order modulation. This would have resulted in high overhead introduced by the L1 signaling in relation to data. If the transmission network is designed for example according to such signal strength that the services using 256-QAM are received correctly, using BPSK modulation for the L1-post signaling would lead to unnecessary overhead. Therefore, in DVB-T2 the L1-post containing most of the signaling information required to receive data from selected PLPs was designed to have configurable modulation order (BPSK, QPSK, 16-QAM and 64-QAM) encoded with an LDPC code with rate 4/9. The signaling content was designed so that the L1-pre contains only the most important system parameters and information that allows the receiver to extract the L1-post signaling. L1-post signaling then carries the rest of the physical layer information. Now, the network operator is able to maximize the throughput of the system by configuring a L1-post modulation that suits the needs for the data PLP configuration of the network. The selection of suitable L1-post modulation order is further studied in section 6.2.4. The coded and modulated L1-post is mapped to the subcarriers of the P2 symbol in a similar manner as L1-pre to obtain time diversity over the whole duration of the preamble. The subcarriers left over after L1 signaling in the P2 symbols are allocated for the data PLPs. It is intended that the first PLP (PLP0) would be a common PLP that carries information that is common for all the services, such as parts of link layer (L2) signaling for example. It must be noted that with low order modulations for the L1-post signaling the capacity available in the P2 symbols of the preamble limits the number of PLPs that can be signaled in one DVB-T2 multiplex. The signaling system could also have been designed to support a preamble of configurable length and not to limit the capacity that is available for the signaling, but this would potentially have imposed difficulties for the receiver design.

As the statistical multiplexing was to be supported by the DVB-T2 standard as indicated in the Call for Technologies [50], the signaling must allow the scheduling of the PLP contents in the T2 frame to change frame to frame. This will allow performing statistical multiplexing over the PLPs. Consequently part of the L1 signaling information needs to change every T2 frame. Therefore, part of the L1-post signaling is called *dynamic signaling* as it contains most importantly the addressing of the PLP data in the T2 frame. This dynamic signaling is critical for all the three receiver operations as if it is lost, the receiver is not able to extract the contents of the PLP from the T2 frame. Therefore, repetition of the dynamic L1-post signaling information was designed. For the repetition, the dynamic parameters for both current and the following frame are transmitted in the preamble. When using the repetition, loss of one preamble can be tolerated assuming that there is no configuration change between the two frames and the reception of the PLP can continue without interruption.

In normal data reception the receiver would receive the L1-pre and L1-post information from the preamble of every frame. Based on the signaling information it would receive the PLP data from the signaled locations in the frame. Taking into account the receiver power consumption, it was considered whether there should be a method to allow the receiver to follow services without needing to wake up for each preamble. The solution for this is the *in-band signaling* that contains part of dynamic signaling being multiplexed with the service data. The in-band signaling contains at least the dynamic signaling for the service being followed but it can contain also dynamic signaling for other services as well. In addition to allowing the service reception without receiving every preamble, in-band signaling provides an additional path for the signaling and thus increases the probability of receiving the signaling correctly.

For service switching there are actually two different sub-use cases, namely switching to a service in the same or in a different multiplex. If the service to be switched to is in the same multiplex, the receiver already knows the scheduling (dynamic signaling) of the target service if it is following the preambles every frame or the dynamic signaling of the desired service is also carried in the in-band signaling. If the dynamic parameters for the target service are not transmitted in-band, the preamble of the next T2 frame is received to obtain all L1 signaling. The reception of the service can begin already from the next T2 frame and the service switching is fast. When the desired service is transmitted in another multiplex, the receiver has to synchronize to the new multiplex and to receive the L1 signaling starting from P1 as in the initial scan. The delay for the channel switching will be longer than in the first situation.

The simulations can be used in the design of signaling transmission as well as for the design of error control coding, interleaving and modulation for the data path. This enables the comparison of the performance of the designed signaling mechanism and data path early in the design phase to avoid unwanted design flaws such as having less robust signaling than the data path for example. This would lead to the fact that the signaling is the critical point for the system operation.

6.2.2 Description of the signaling preamble

The preamble symbol P1 is mainly intended for fast identification of available T2 signals. The challenge in the current DVB-T/H system is the large number of parameters that need to be blindly tested during the initial scan (that is when the receiver is turned on for the first time). In DVB-T2, the P1 symbol can be received without knowing the data transmission parameters, which makes the initial scan much faster. A somewhat similar preamble symbol is utilized in the WiMAX standard [85] for synchronization.

A DVB-T2 transmitter may produce a frequency offset to the signal in order to avoid adjacent channel interference. This offset is at most ± 0.5 MHz compared to the nominal center frequency. The P1 structure however allows detection at the nominal center frequency and the receiver can be later tuned according to the offset. In addition, the P1 symbol provides

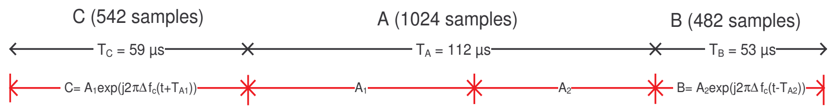

Figure 6.7: P1 symbol in time domain.

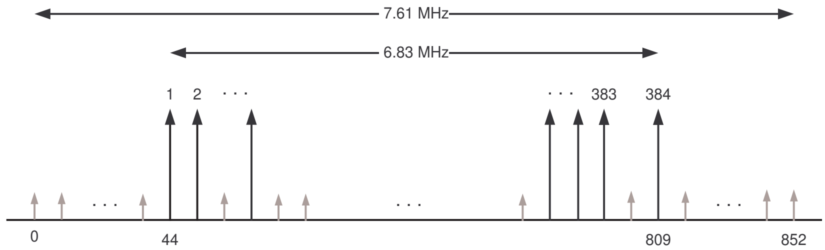

Figure 6.8: P1 symbol in frequency domain.

seven signaling bits to declare some transmission parameters. To conclude, the P1 symbol is used in the initial scan for 1) detecting the presence of DVB-T2 signal on the current frequency, 2) estimating the frequency offset, and 3) signaling.

The structure of the P1 symbol in time domain is shown in Fig. 6.7. The P1 symbol consists of a 1k OFDM symbol (A), which is Differential Binary Phase Shift Keying (DBPSK) modulated in frequency direction by pseudo random binary modulation signaling sequences (MSS) called S1 and S2. In addition to the main symbol part, P1 includes two frequency shifted cyclic extensions C and B. These extensions have an offset of 30 samples such that C has 542 samples and B has 482 samples. Thus, the total length of P1 is 2048 samples (that is the same as for 2k OFDM symbol). The cyclic prefix C is a frequency shifted version of the first part of A, and the cyclic suffix B is similarly a frequency shifted version of the latter part of A. The cyclic extensions thus contain together the same information as part A. The frequency shift is one subcarrier upwards for both C and B.

The frequency domain structure of P1 (A) is shown in Fig. 6.8. The P1 symbol uses 384 of the 853 subcarriers that are available in a normal 1k symbol. The pseudo random allocation of the used carriers is determined in the standard by the carrier distribution sequence (CDS). This carrier pattern has been designed to have good auto-correlation properties for finding the integer frequency offset by studying the locations of the used carriers. The subcarriers have been selected from the middle of the band so that most of

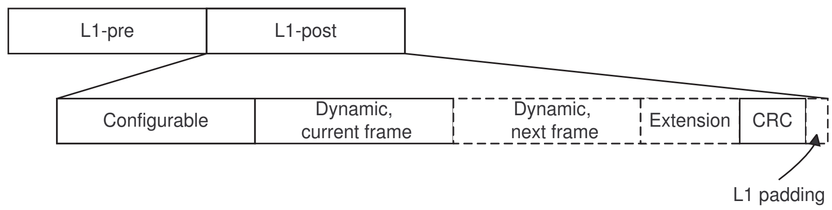

Figure 6.9: Division of L1 signaling.

them stay within the nominal band of 7.61 MHz (for 8 MHz system) even when the largest frequency offset is used in the transmitter. The P1 carriers are boosted to normalize the power between the P1 and data symbols. This boost also enhances the reception of the carriers due to higher power.

P1 symbol provides seven signaling bits that are carried by the modulation signaling sequences S1 and S2. The length of S1 sequence is 64 and there are eight possible sequences, that is, S1 carries three bits. The S2 sequence occupies 256 subcarriers and provides four signaling bits. The S1 bits define the type of the system, or more precisely the type of the following P2 symbols. The S1 signaling field is common for all DVB-T2 based systems including possible future extensions for example to a system dedicated for handheld reception. In DVB-T2, S2 bits announce the FFT size and also the guard interval group is signaled for the FFT modes, that are considered to be the most usual ones. With this information the receiver is able to receive the P2 preamble symbols.

All of the L1 signaling (also repeating information carried by S1 and S2 in P1 symbol) and PLP data if there is space, are transmitted in the P2 OFDM symbols located in the beginning of the T2 frames (see Fig. 4.5). As the L1 signaling is transmitted in P2 symbols, the T2 frame length sets the minimum transmission interval for the L1 signaling. The number of the P2 symbols in T2 frame depends on the FFT size used in the system. For 1k, 2k, 4k, 8k, 16k and 32k modes 16, 8, 4, 2, 1 and 1 P2 symbols are used, respectively. P2 symbols use the same FFT size as the data symbols. In P2 symbols every 3rd carrier is a pilot to enable accurate channel estimation already from the P2 symbols. The amount of available carriers in the P2 symbols sets the maximum capacity for L1 signaling and thus is often the limiting factor for the amount of PLPs that can be signaled in the DVB-T2 system.

The L1 signaling transmitted in P2 symbols is divided in the DVB-T2 standard into two parts, L1-pre and L1-post signaling as illustrated in Fig. 6.9. L1-pre signaling acts as a key to receive L1-post signaling. The L1-post signals the parameters necessary for the reception of the PLP data. The

L1-post signaling parameters are further divided into two parts based on their frequency of change: configurable and dynamic (Fig. 6.9). The configurable parameters change only when the network transmission parameters are changed, for example PLPs added or removed. The changes in the configurable parameters are applied only on the border of a T2 super frame. T2 super frame consists of configurable number of T2 frames. The number of frames in the superframe is signaled in L1-pre signaling. Dynamic parameters on the other hand can change from frame to frame. These parameters most importantly contain information on the mapping of PLPs to the frame (to which carriers in which OFDM symbols the PLP is mapped to) and frame index number. To obtain further robustness for the dynamic L1-post parameters, the dynamic parameters of the next frame together with the ones for the current frame can be transmitted in each preamble. This is called repetition of L1-post signaling. These dynamic parameters or part of them can also be transmitted in the in-band signaling. Extension field in the L1-post signaling can be used in the future if some extensions in the signaling are necessary. L1 padding bits are inserted to make the amount of encoded L1-post data divisible by the number of P2 symbols. The divisibility is important when mapping the information to the P2 symbols.

Transmission of L1 signaling in P2 symbols

As L1-pre signaling enables the reception of the L1-post signaling it needs to be as robust as possible and therefore it is always BPSK (Binary Phase Shift Keying) modulated and protected by the rate 1/5 and length 16200 LDPC code concatenated with the BCH code. The rate 1/5 LDPC code is inherited from the selection of DVB-S2 LDPC codes. In DVB-T2 it is used only for L1-pre transmission. The amount of L1-pre signaling data is 200 bits and the amount of BCH redundancy is 168 bits. Thus, the LDPC codeword needs to be shortened and punctured to be able to transmit this small amount of data using the given LDPC with the original amount of information bits $k = 3240$. The LDPC code is shortened and punctured to maintain the original rate. Therefore, the amount of encoded L1-pre bits is $(200 + 168) \times 5 = 1840$. The overall code rate including BCH for the L1-pre signaling is $R = \frac{200}{1840} \approx \frac{1}{9}$.

The L1-post signaling data is protected by rate 4/9 (called rate 1/2 in the standard) length 16200 LDPC code concatenated with the BCH code. Shortening and puncturing are used. The modulation schemes used with the L1-post signaling are BPSK, QPSK, 16-QAM and 64-QAM. For the data path, the most robust transmission scheme is QPSK with code rate 1/2. It is designed considering static transmission environments that L1-post can always be configured to be more robust than the data path by selecting lower order modulation (for example BPSK in the case of QPSK code rate 1/2 for

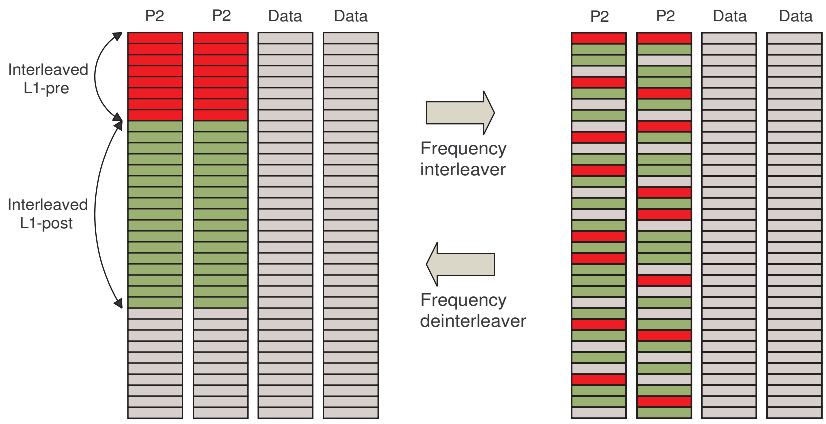

Figure 6.10: Mapping of L1 signaling to preamble symbols.

the data path). The amount of the L1-post signaling data depends on the transmission system parameters, that is on the amount of PLPs used in the system, usage of L1 post repetition, use of Future Extension Frames and so on. Therefore, the puncturing and shortening scheme must be flexible to allow for different amounts of L1-post signaling data. The shortening and puncturing schemes specified in the standard are based on the structure of the LDPC codes to ensure good performance with all possible codeword lengths.

The coded and modulated signaling cells are inserted to the carriers of the P2 symbols so that L1-pre and L1-post data are evenly distributed over all P2 symbols of one T2 frame. This is done to obtain as much time diversity for the signaling data within the preamble as possible. The mapping of signaling to P2 symbols is illustrated in Fig. 6.10. PLP data is inserted to the carriers available in P2 symbols after the insertion of the L1-pre and -post signaling. This ensures that there is no other overhead introduced by the P2 symbols than what is imposed by the pilot pattern that is more dense than that of the data OFDM symbols.

6.2.3 Performance analysis

First, the performance of P1 symbol is studied, and later the analysis is extended to include the L1 signaling in the P2 symbols and also service data. The performance has been evaluated in three different channel conditions: AWGN, DVB-T P1 (static Rayleigh) [86] and TU6 (mobile 6-tap Typical Urban) [74]. The SNR in all figures stands for E_s/N_0.

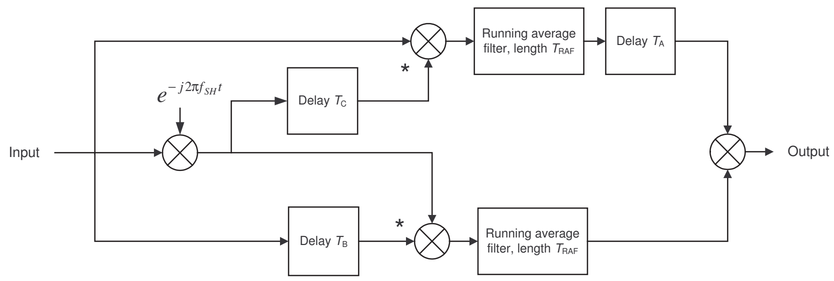

Figure 6.11: Time domain P1 correlator.

P1 detection algorithms

P1 detection begins with time domain correlation, which is based on the basic guard interval correlation as presented in [87]. The correlator structure, which has been modified to take into account both cyclic extensions and frequency shifts, is presented in Fig. 6.11. There are two running average filters. The lengths of the filters can be selected to be equal to the cyclic extensions, T_C and T_B, which gives a clean correlation pulse as shown in Fig. 6.12. This option has been chosen for the simulations. Selecting the running average filter lengths equal to part A for both filters results in a wider pulse (Fig. 6.13) but this has merit of eliminating possible continuous wave interference [86]. Note that guard interval correlation is immune to frequency offset which is vital since the offset is unknown at the signal detection stage. Also, fractional frequency offset can be estimated from the time domain correlation.

Guard interval correlation is intended for situations where the delay spread of the channel stays within the guard interval, which may not be the case with P1 symbol in large scale SFNs, for example with 32k OFDM mode. In this case, the delays longer than the guard interval introduce Intersymbol Interference (ISI). This is however not a problem because P1 symbol has been designed to cope with very low Signal-to-Interference-plus-Noise Ratio (SINR).

After coarse time and fractional frequency synchronization, the received P1 symbol can be translated to frequency domain by FFT. The output of the time domain correlator may be noisy due to ISI and low SNR, and this may result in a false alarm. Therefore, the presence of a P1 symbol is validated in frequency domain.

Next, the receiver estimates the integer frequency offset, that is, the frequency offset in terms of subcarrier spacings. This is done by comparing the received power spectrum to the presumed subcarrier locations (CDS). The power correlation is calculated over all possible frequency shifts and

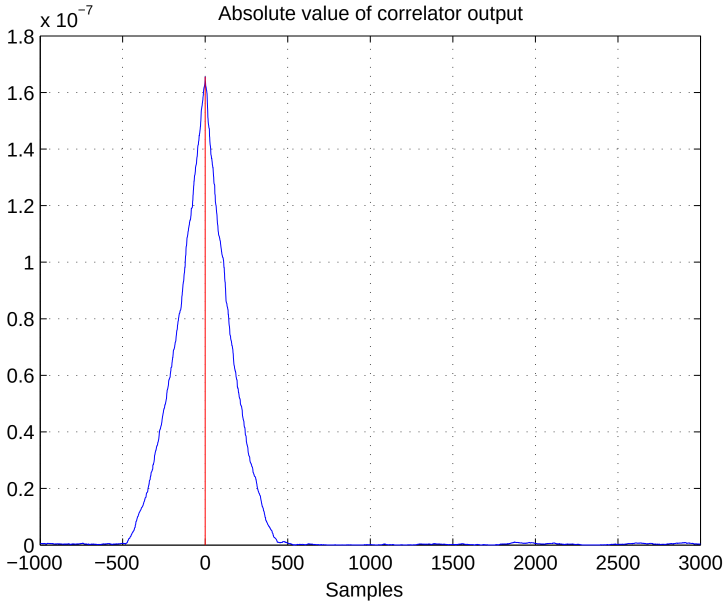

Figure 6.12: Ideal correlator output with running average filter lengths equal to the cyclic extension.

the estimated integer frequency offset Δf_{int} is obtained by maximizing the correlation

$$\Delta f_{int} = \max_k \left[\sum_{p \in P} y_{p+k} y^*_{p+k} \right], \tag{6.2}$$

where P is the set of P1 pilot subcarriers without frequency offset, y_{p+k} is the received symbol at carrier $p + k$, and k is an integer limited to the possible integer frequency offset range. In practice, it is possible to reduce the offset range by pre-processing but in the simulations the range has been the whole allowed ± 0.5 MHz. P1 validation is done by setting a threshold for this power correlation.

Finally, MSS sequences are obtained by soft DBPSK demodulation, and the S1 and S2 bits are decoded by selecting the MSS sequence that has the highest correlation to the received sequence.

P1 simulation model

The P1 simulation model represents a simplified version of the DVB-T2 system, where the transmitted signal consists only of the DVB-T2 preamble

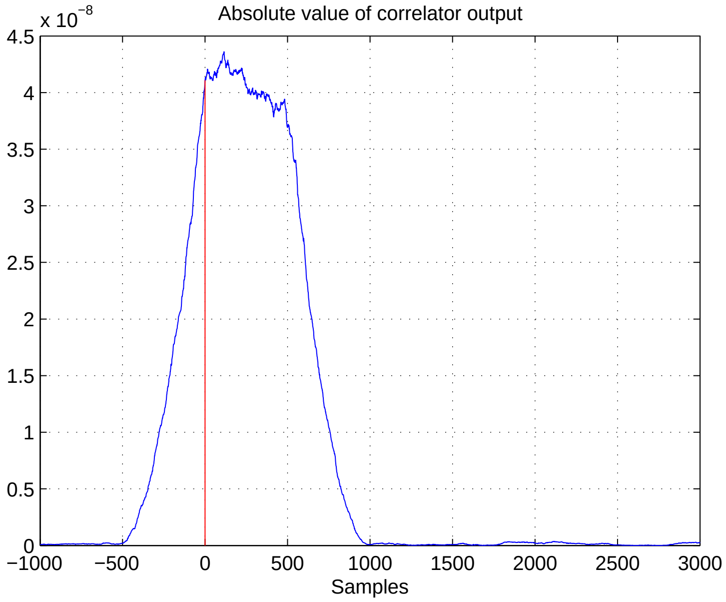

Figure 6.13: Ideal correlator output with running average filter lengths equal to the length of main symbol part A.

and random OFDM symbols representing the data. The simplified model is used to reduce the required simulation time. The transmitted time domain signal is passed through multipath channel model and noise is added, after which the signal is fed to the receiver. The receiver performs time and fractional frequency synchronization, integer frequency offset estimation by power correlation, P1 validation, and decodes the S1 and S2 signaling fields.

The simulation results for detecting the P1 symbol and decoding the S1 and S2 bits are presented in Fig. 6.14 and 6.15. In the figures there are four curves: T2 detection gives the probability of detecting a valid P1 symbol by power correlation, false alarm means that a P1 symbol is detected in power correlation although there was no P1 symbol, S1 and S2 give the probability of detecting the signaling correctly.

Fig. 6.14 shows the performance in AWGN channel. It can be seen that the P1 symbol can be correctly received in adverse SNR conditions. It is interesting that the performance for T2 detection and signaling is almost the same. This can be explained by the high threshold that has been set for power correlation to indicate correctly detected P1. A high threshold is necessary to avoid false alarms, but the consequence is that once the threshold is exceeded, the signal is good enough also for detecting the signaling.

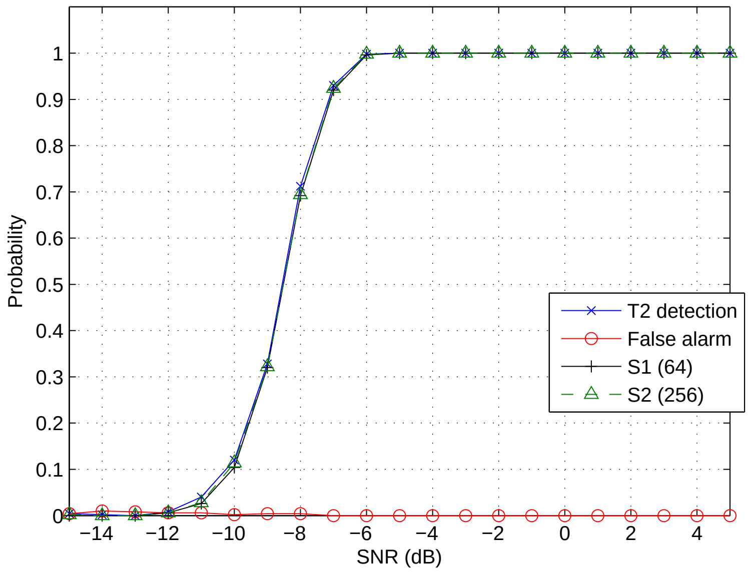

Figure 6.14: P1 performance in AWGN channel.

Fig. 6.15 shows the performance in mobile TU6 channel with Doppler frequency $f_D = 200$ Hz. Because of the low FFT size (1k), the P1 symbol has a good tolerance against Doppler effect. Although there is no channel estimate available, the frequency selectivity of the channel does not degrade the reception because of the differential modulation.

P2 simulations

Simulation results for the L1 signaling carried in P2 symbols are presented next. Performance curves for L1-pre, L1-post and data with code rate 1/2 are presented. FFT size of 8k is used in the simulations resulting in two P2 symbols in each T2 frame. The 8k mode is a possible selection for a network in which the transmitted services are destined for mobile users. Code rate 1/2 is shown for the data as it presents the most robust data path with given modulation. Also, constellation rotation is used for the data path. Other system parameters used in the simulations are: guard interval duration is 1/4 of the "pure" OFDM symbol duration and pilot pattern 1 (PP1 as defined in the standard [6])for the data symbols. Thousand T2 frames are simulated for AWGN and static Rayleigh channels while 2000 T2 frames are simulated for the TU6 channel to obtain statistically more reliable results.

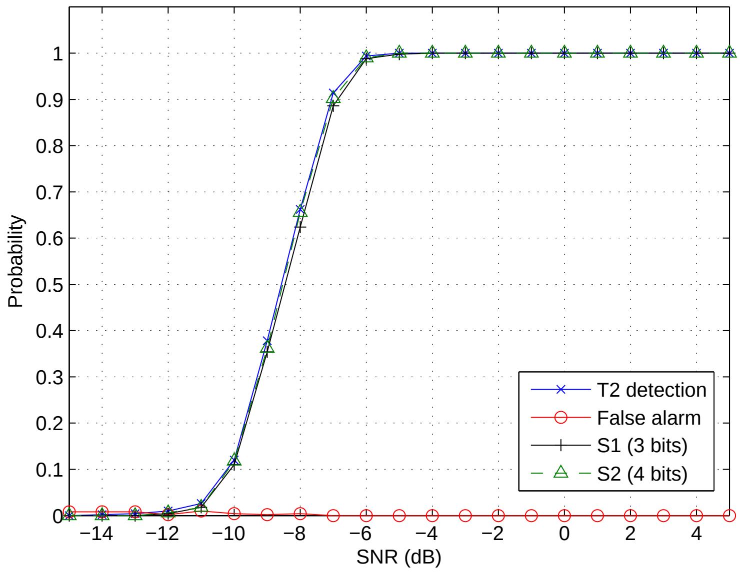

Figure 6.15: P1 performance in TU6 $f_D = 200$Hz channel.

P2 simulation model

The P2 simulations are performed with full DVB-T2 simulator. T2 frames consist of P1, P2 and data OFDM symbols. The L1 signaling data is generated, encoded and mapped to the P2 symbols. Data symbols are inserted after the P2 symbols. The transmitted time domain signal is passed through the same channel simulation blocks as used for the P1 simulations. The receiver extracts the L1 signaling from the P2 symbols and decodes L1-pre and L1-post information. As only the performance characteristics of the system are studied, the actual data transmitted by the simulator is random. The receiver mainly uses receiver algorithms presented in [86]. For example, "Genie Aided" constellation demapping for the data is used, as it is considered to represent the optimal performance that can be achieved by iterative demapping and decoding [86]. Therefore, the results may be optimistic if compared directly to results with real hardware receivers. The constellation rotation for L1 signaling is not allowed by the standard. Further, in the simulations the receiver is synchronized perfectly to the transmitted signal. As ideal channel estimation is used, the effect of pilot pattern on the performance results is negligible, but it would have an effect on the throughput of the system. Short FEC block length (16200 bits) for the data path is

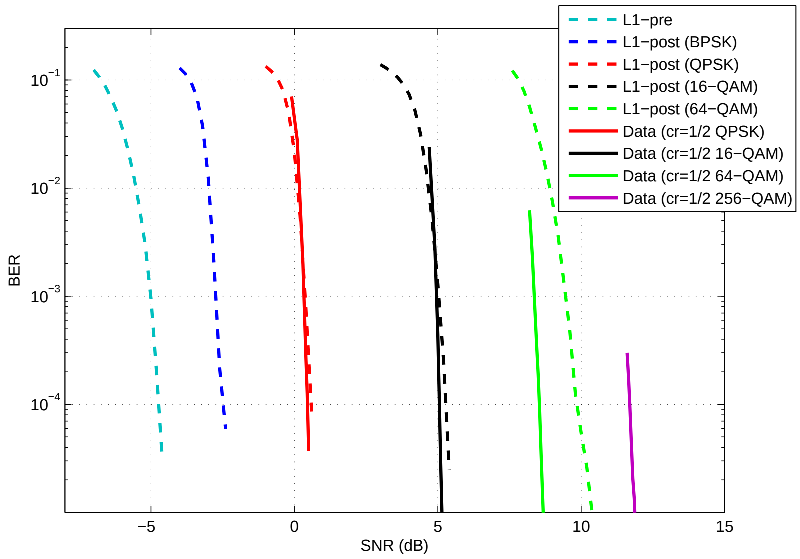

Figure 6.16: L1 signaling performance in AWGN channel.

used. The results for the long FEC block are quite close to the short one as indicated also in the simulation results presented in [86].

First, Bit Error Rate (BER) curves for the L1 signaling and data path in commonly used AWGN channel are presented in Fig. 6.16. It is observed that the L1-pre signal is about 2 dB more robust than the L1-post signal using BPSK modulation. This is as designed, since L1-pre acts as a key for receiving the rest of L1 signaling carried in L1-post. Also, it is seen by comparing to the Fig. 6.14 that the robustness of the P1 reception is much higher than reception of the L1-pre, meaning that P1 is received correctly with lower SNR values. Also this is as designed, since the P1 is the key for receiving L1-pre and L1-post. The performance of the L1-post is rather similar to the data path with similar modulation. It is also seen from Fig. 6.16 that the signaling can be configured to be more robust than the data path by selecting a lower order modulation for the signaling. The selection of the modulation for the signaling is further studied in chapter 6.2.4. The results for the static Rayleigh channel are almost similar as for the AWGN channel but shifted some dBs higher in SNR. This is visualized in Fig. 6.17.

It is possible that DVB-T2 or its variant will be used for broadcasting data services for mobile receivers. Therefore it is important to determine its potential and performance also in mobile channels. The simulation results for the TU6 channel with 80Hz of Doppler frequency are presented in Fig. 6.18. For TU6 the performance of the L1 signaling is significantly worse

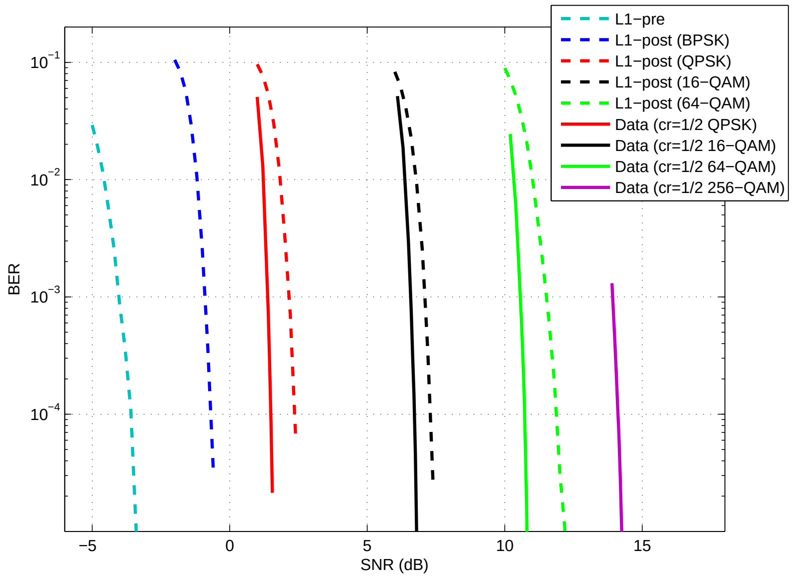

Figure 6.17: L1 signaling performance in DVB-T P1 Rayleigh channel.

than for the static channels. The weak performance is due to the lack of time diversity for L1 signaling as L1 signaling is spread over only the two P2 symbols in 8k case. The time diversity obtained for the L1 signaling in DVB-T2 is 2-4 ms if the L1-post repetition is not taken into account. The data on the other hand can have much greater time diversity (up to several seconds) thanks to the large amount of time interleaving allowed by the standard. In the simulations for the data path interleaving depth of about 80 ms was used as it is the maximum interleaving length that can be obtained with the selected system parameters and a single PLP. This kind of PLP resembles the one illustrated in Fig. 6.5 as configuration A.

Let us now consider another error criterion for the L1 signaling, namely T2 frame error rate (FER) indicating which fraction of the preambles was erroneous. Now for example FER 10^{-3} corresponds to one error in 1000 frames on the average, that translates to an error in 250 seconds (roughly 4 minutes) if 250 ms T2 frames were used. Luckily, the performance of the L1 signaling is enhanced by the optional L1-post repetition. When using repetition, it is assumed that any time two successive preambles are not lost, the reception of PLPs can continue without interruptions. This assumption is justified, because the L1-post dynamic information that is repeated is the most important for the receiver to be correct as it can change from T2 frame to T2 frame, while the configurable information changes rather seldom and

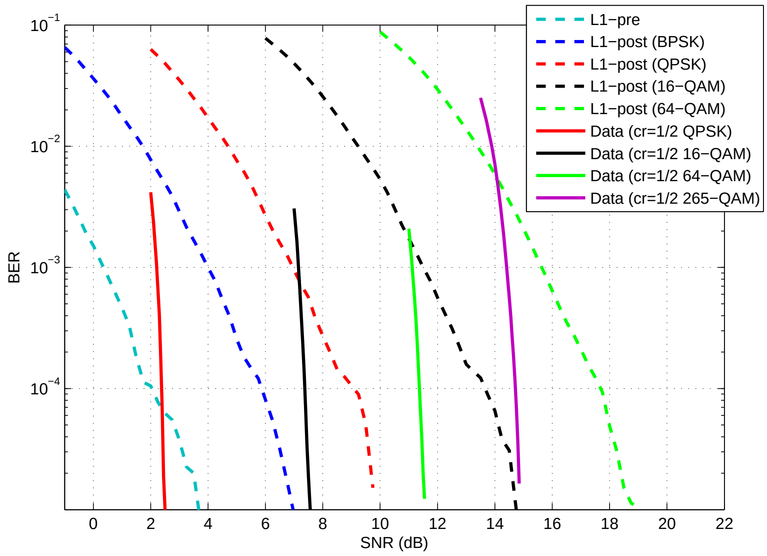

Figure 6.18: L1 signaling performance in TU6 $f_D = 80$Hz channel.

thus does not necessarily need to be correctly received from every T2 frame. The interval for the successive preambles is the duration of the T2-frame. The T2 frame length of 250 ms is used in the simulations. The interval is larger than the coherence time of the used TU6 channel with $f_D = 80$ Hz that is in the order of 5 ms. Therefore it is likely that two successive preambles observe independent fading. Obviously, with very slowly changing channels, when the coherence time of the channel is longer than the T2 frame duration the improvement of the repetition is small as two successive preambles may suffer from the same channel fade. The gain obtained by the L1-post repetition in TU6 ($f_D = 80$ Hz) channel is presented with respect to FER in Fig. 6.19. It is observed that about 3 dB gain can be obtained by the repetition. This is intuitive, as the fading during the preambles are independent. Still, it seems that the the lack of time diversity makes it impossible to make the L1 signaling transmitted in P2 symbols more robust than the most robust data path with QPSK and code rate 1/2. Studying the figures 6.18 and 6.19 it is seen that this most robust data path reaches the bit error rate of 10^{-4} at approximately the same SNR value of 2.4 dB as BPSK modulated L1-post with repetition reaches frame error rate of 10^{-3} (one error in 4 roughly minutes). Thus the use of L1-post repetition brings the performance of the L1-post dynamic signaling close to the performance of the data path at least in the presence of the selected error criteria and system parameters (time interleaving length, FFT size and so on).

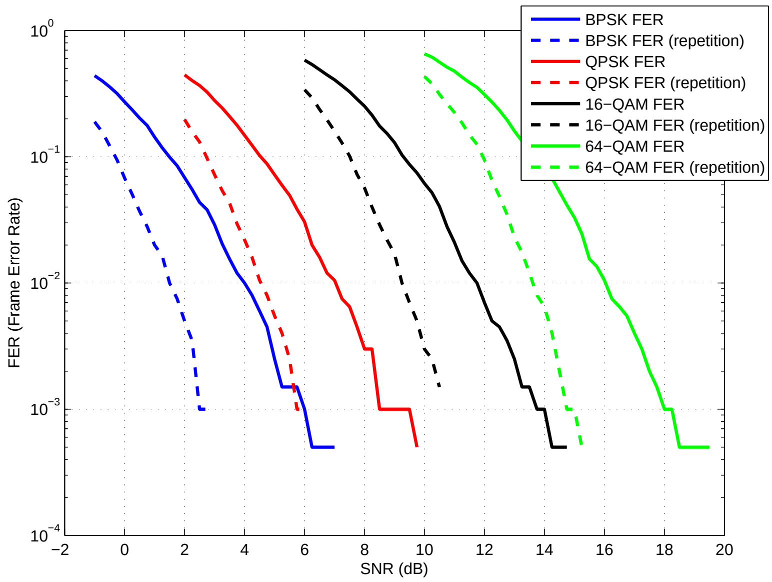

Figure 6.19: L1 signaling frame error rate comparison in TU6 $f_D = 80$Hz channel.

The in-band signaling may further help in severe receiving conditions, as its robustness is exactly the same as the robustness of the data path. Still, in-band signaling is an additional signaling mechanism, as if data frame carrying in-band signaling is lost, so is in-band signaling. In this case the receiver must receive the L1 signaling from the preamble to be able to continue service reception.

As a conclusion, it was observed that the transmission of the most important system parameters as well as detection on the presence of the DVB-T2 signal from the P1 symbol are very robust. Also, the transmission of the rest of physical layer signaling in the P2 symbols can be configured sufficiently robust in static reception conditions. In fast mobile reception conditions on the other hand, the robustness of the signaling may not be high enough due to lack of time diversity for the signaling. Therefore, some improvements in the physical layer signaling transmission may be necessary if the mobile use case is planned during network deployment.

6.2.4 Configuration selection study

Errors in service data may introduce some errors in for example audio or video, but errors in the signaling may destroy the reception of the whole service. Here the robustness of the signaling transmission and the robustness of the data path are compared to come up with guidelines for selecting the modulation parameters in the DVB-T2 network for the L1 signaling. The comparison is based on simulation results for both signaling and data paths.

There are several possible combinations of transmission parameters for L1-post modulation with respect to the parameters used for the data path. The modulation scheme for L1-post can be configured to be BPSK, QPSK, 16-QAM or 64-QAM. Further the use of L1-post repetition and in-band signaling is optional. To have an idea of suitable configurations, we need to define how large the margin in the robustness of signaling and service data is considered adequate. Let us decide for the following considerations that L1-post signaling should be at least 3 dB more robust than the service data path.

Let us begin with the simplest case when no L1-post repetition nor in-band signaling is transmitted. In this situation the receiver needs to continuously follow the preambles for normal service reception to obtain the L1-post dynamic information. The modulation for L1-post signaling that should be used in combination with each data path code rate and modulation can be decided based on the 3 dB criterion at a certain error rate level. For example at bit error rate of around 10^{-4} in AWGN channel data path with 64-QAM and code rate 1/2 requires SNR of 8.5 dB. The L1-post signaling with 16-QAM modulation requires SNR 5.3 dB to reach the same error rate criterion (Fig. 6.16). It is clear that 16-QAM for the L1-post is sufficient according to our selected 3 dB criterion in this case. Similar comparison can be done for all other data path code rates and modulations. The results of this comparison in AWGN and TU6 $f_D = 80$ Hz channels are shown in Table 6.2. Value "1" indicates that the selected 3 dB criterion is fulfilled with the L1 signaling modulation of the column while "0" indicates that it is not. The first value in each column is for AWGN and the second for TU6 without L1-post repetition or in-band signaling.

In general for AWGN it is observed that for high code rates for data (3/4 and above) similar modulation for L1-post can be used. For low data code rates (below 3/4) one order lower modulation for L1-post should be used. For code rate 1/2 256-QAM data, using 64-QAM for L1 signaling is not robust enough (data 11.8 dB and L1 signaling 9.9 dB). Therefore, for this mode 16-QAM for signaling should be used. Of course lower order modulations could also be used for L1-post, if a more strict criterion for the robustness difference for signaling and data is desired (for example 6 dB). Selecting a lower order modulation for the signaling naturally increases the overhead introduced by the signaling. When multiple PLPs are present, the signaling should be matched to the most robust PLP as the signaling is common for all PLPs.

For TU6 it is observed that if QPSK with any code rate or 16-QAM modulation with code rate 1/2 is used for data PLPs, the L1 signaling

Table 6.2: Modulation selected for L1-post for each data configuration based on 3 decibel criterion (AWGN,TU6,TU6+rep,TU6+rep+ib)

Data	BPSK	QPSK	16-QAM	64-QAM
QPSK 1/2	1,0,0,1	0,0,0,0	0,0,0,0	0,0,0,0
3/5	1,0,0,1	0,0,0,0	0,0,0,0	0,0,0,0
2/3	1,0,1,1	0,0,0,1	0,0,0,0	0,0,0,0
3/4	1,0,1,1	1,0,0,1	0,0,0,0	0,0,0,0
4/5	1,0,1,1	1,0,0,1	0,0,0,0	0,0,0,0
5/6	1,0,1,1	1,0,1,1	0,0,0,0	0,0,0,0
16-QAM 1/2	1,0,1,1	1,0,0,1	0,0,0,0	0,0,0,0
3/5	1,1,1,1	1,0,1,1	0,0,0,1	0,0,0,0
2/3	1,1,1,1	1,0,1,1	1,0,0,1	0,0,0,0
3/4	1,1,1,1	1,1,1,1	1,0,0,1	0,0,0,0
4/5	1,1,1,1	1,1,1,1	1,0,1,1	0,0,0,1
5/6	1,1,1,1	1,1,1,1	1,0,1,1	0,0,0,1
64-QAM, 1/2	1,1,1,1	1,0,1,1	1,0,0,1	0,0,0,0
3/5	1,1,1,1	1,1,1,1	1,0,1,1	0,0,0,1
2/3	1,1,1,1	1,1,1,1	1,0,1,1	1,0,0,1
3/4	1,1,1,1	1,1,1,1	1,1,1,1	1,0,1,1
4/5	1,1,1,1	1,1,1,1	1,1,1,1	1,0,1,1
5/6	1,1,1,1	1,1,1,1	1,1,1,1	1,1,1,1
256-QAM 1/2	1,1,1,1	1,1,1,1	1,0,1,1	0,0,0,1
3/5	1,1,1,1	1,1,1,1	1,1,1,1	1,0,1,1
2/3	1,1,1,1	1,1,1,1	1,1,1,1	1,1,1,1
3/4	1,1,1,1	1,1,1,1	1,1,1,1	1,1,1,1
4/5	1,1,1,1	1,1,1,1	1,1,1,1	1,1,1,1
5/6	1,1,1,1	1,1,1,1	1,1,1,1	1,1,1,1

transmitted in the preambles cannot be configured to be 3 dB more robust than the data when the repetition is not used. In this situation the signaling may introduce a bottleneck for the performance of the system.

When configuring to use the L1-post repetition, there is a gain in frame error rate for the signaling as illustrated in Fig. 6.19. Let us consider a situation when there is no in-band signaling and the receiver relies on the L1 signaling in the preambles for normal service reception. Considering the frame error rate of 10^{-3} (one error in the service due to signaling in roughly four minutes) to be enough, a comparison to the data path is made in TU6 channel. The suitable modulation and coding combinations are shown in the third value for each modulation and code rate in Table 6.2. It is observed that with repetition it is possible to decrease the amount of modes when the signaling cannot be made more robust than the data path. Still, there are two modes with which the signaling robustness introduces limitations. These are the most robust data PLP configurations using QPSK and code rates 1/2 and 3/5.

Further, if the in-band signaling is used the requirement for the frame error rate for the L1 signaling transmitted in preambles can be loosened, since the receiver does not need to follow the preambles continuously for normal service reception. The reception of the preamble correctly is still required for performing initial scan or channel switching. The frame error rate requirement for these use cases is lower than for the normal service reception. For example if frame error rate is 20 % for the L1-post signaling in the preamble the probability of receiving the L1-post information correctly from the first frame is 80 %. To receive the signaling correctly from the first or the second frame the probability becomes $0.8 + 0.2 \times 0.8 = 0.96$. The preamble error occurences are assumed to be independent. As the probability that the receiver has been able to obtain full L1 signaling correctly already after two frames (maximum of 500 ms delay) is as high as 96 %, the 20 % is considered a reasonable error rate requirement for the preamble signaling that is used for the initial scan and service switching. Suitable modulation and coding combinations for the preamble signaling are found out based on the 3 dB criterion. The suitable combinations are shown in the fourth column of Table 6.2. It is observed that when using the in-band signaling the transmission of the preamble signaling can be configured to be robust enough according to the selected criterion for the frame error rate thanks to the loosened robustness requirement.

The presented L1-post configuration guidelines can prove to be useful, as a DVB-T2 network can be configured to perform very poorly by choosing unsuitable combination of transmission parameters for L1 signaling and data path. From the AWGN simulations it was seen that the physical layer signaling can be configured to be more robust than the data path as designed. It was also observed that the performance of the signaling with respect to data path in mobile scenario is not as good as in the static environment. Therefore, also the sufficiency of the L1-post repetition and in-band signaling for mobile environment was considered. In general from the performance point of view the L1-post repetition and in-band signaling should be always used if mobile reception is to be enabled by the network. Still, selecting to use repetition and in-band signaling increases the overhead of the signaling and reduces the capacity available for the services. Also, the amount of PLPs that can be signaled in the preamble decreases when the L1-post repetition is used. Therefore the network operator needs to carefully consider the configuration of the signaling as well as the configuration for the services for providing required QoS for the receivers in the network.

Chapter 7

Future evolution of broadcasting systems

The pace of development for microcomputers has been observed to follow what is known as Moore's Law. Gordon Moore noticed already in the 1960s that the number of transistors on a silicon chip had doubled every year. From this he assumed that the growth would not go on with this high pace, but exponentially regardless. Moore assumed that the doubling would happen every 24 months [88]. Later the figure was adjusted to doubling every 18 months by David House. The development of microprocessors has followed this prediction accurately. An important follow up of Moore's Law is that because more transistors can be fit to a single chip, for a fixed cost, chip performance is doubled as the area of the chip is the most important factor in the manufacturing costs.

Broadcast systems have not been able to follow directly the development of digital technology as broadcasters have had a great emphasis on the backwards compatibility of the user equipment when updating transmission systems [89]. This is in contrast to the computer industry where no attention is paid to the compatibility of old computers with new software. Actually, new software can be a marketing mechanism for the consumers to buy new computers to be able to run the latest applications. Due to this different approach, the broadcasting equipment cannot experience such a high pace of development as the computer industry. Mainly, the broadcasting industry can benefit from Moore's Law only when changing technologies such as converting from analogue transmissions to digital ones [89] or upgrading from DVB-T to DVB-T2. As observed in previous chapters, the DVB-T2 system applies much more computationally intensive mechanisms than the first generation system. Also, it was observed during the analogue transmission shut down that the willingness of the users to upgrade for example their television sets or set top boxes is lower than to upgrade their computers [90].

One reason for this is that digital broadcasting is still mainly plain television and switching from analogue to digital transmissions did not introduce many additional features to the service. Also, when switching from DVB-T to DVB-T2, the main update from the user perspective will be the enhanced audiovisual quality. The broadcaster on the other hand has the advantage of increased throughput to accommodate more services in the multiplexes.

7.1 Directions of system design

The current technologies allow the use of mechanisms that result in system performances close to what is considered theoretically achievable. This is observed for mechanisms used in the DVB-T2 system. The heterogeneous nature of terrestrial reception environments assures that there is still room for improvement as the technology advances. The increase in popularity of mobile reception of broadcasting services will certainly keep the engineers busy also in the future. Technologies such as MIMO can be utilized for improved capacity or robustness. The introduction of MIMO was discussed already at the design phase of DVB-T2, but due to the requirement of backwards compatibility for antenna installations the introduction of MIMO into the DVB-T2 system was impossible. The advantages of MIMO were anyway identified and to allow later possible updates of the standard to be carried in the same radio frequency as the DVB-T2, the system was designed to incorporate Future Extension Frames (FEFs). The FEF is seen by the basic DVB-T2 receiver as a "black hole" where any kind of signal can be transmitted. It is possible that the next generation handheld (DVB-NGH) system would be carried in the FEFs.

The user experience for mobile television can be enhanced by utilizing mechanisms allowing graceful degradation [91]. In digital systems such as DVB-T for example, the video quality remains perfect up to a certain received signal level and degrades quickly with lower signal levels. Graceful degradation in broadcasting television services translates into for example lower resolution or smaller frame rate obtained by the receivers in poor reception conditions. This in a way resembles the performance of the previous analogue systems where the picture quality degrades gradually. For digital broadcasting systems the graceful degradation can be obtained with help of scalable video codecs. This is studied for a Multimedia Broadcast and Multicast Services (MBMS) network in [92]. Also, the scalable video codecs would in theory allow sharing service components between HDTV and mobile services for example.

Two directions in broadcasting systems design are currently observed, the toolbox-of-modes approach and the dedicated system approach. The toolbox-of-modes approach is utilized in DVB systems such as DVB-T2.

There, many possible code rates, modulation modes, guard interval lengths and so on are defined at the physical layer to make the system utilization in a wide variety of network scenarios possible. An example of a more dedicated system would be CMMB [93], where a very limited number of configurations is specified. Selection between these two design approaches has an effect on the cost of the devices. The toolbox-of-modes approach requires that all the modes are implemented in the receiver if the receiver is to operate on all networks therefore resulting in higher receiver cost as compared to the dedicated system approach. On the other hand, a dedicated system supports a more limited set of network scenarios.

A clear trend in the mobile industry is the convergence of the different transmission systems in the user device. For example the current cell phones can comprise in addition to GSM and 3G, WLAN, Bluetooth and DVB-H for mobile streaming service reception. Therefore future broadcast receivers need not be devices without any kind of feedback mechanism. The cellular system can be utilized for the feedback in the cell phone example. If the presence of the interactive channel can be assumed, different directions in broadcasting system design can be taken. In this case for example the signaling could be obtained on request over the interactive channel or lost service packets could be requested assuming long enough buffering for the streaming services and low delay on the interactive channel.

The evolution of the convergence of broadcast and telecommunication networks is studied in [94]. It is pointed out that various digital radio systems have been developed and optimized for specific services, such as 3G for mobile communications and DVB-H for mobile broadcasting. For example, provision of services that are of interest for large masses are most efficiently transmitted over a broadcasting system, while interactive services are more suitable for the cellular network. For optimal utilization of the available radio resources, these systems should interact. As currently several receivers are present in many cellular phones, convergence in the receiver side is possible. Naturally this requires also interaction of broadcast and cellular operators, which may be challenging as they usually represent different companies.

For introduction of the standards for new systems the timing is a crucial factor [89]. Choosing to use safe and conservative technologies in the design leads to a shorter life time expectancy than with the most cutting edge technologies. On the other hand using most advanced technologies at the moment, leads to devices with high price tags. The price lowers as time goes according to Moore's Law, but it may be that new more appealing products have already become available. Also, introducing a new generation of the system at the same time the current older system is being deployed may harm the prospects of the current system and make the consumers wait for the next generation instead of choosing to use the current system.

Let us next consider the outcomes of the case studies presented in chapters 5 and 6 and specially what can be learned from the presented analyses for future broadcasting system design.

7.2 Guidelines for the technical design of data and signaling paths based on findings of the case studies

The signaling and data paths at each layer of a broadcasting system should be designed in close relation to each other. This close relation means that in the design of the signaling content as well as the transmission mechanism, the requirements set by the QoS needs should be converted into requirements for the signaling. As the signaling is something that the user does not need to know about, direct requirements on its performance or functionality are not set by the user. The process for finding the performance requirements for the signaling is illustrated in Fig. 7.1. The general system architecture decisions based on functional requirements set restrictions on the signaling content and how it can be transmitted. An example of this is the selection of a time division multiplexing transmission scheme combined with OFDM in the DVB-T2 physical layer. The multiplexing thus requires that the signaling conveys information on the addressing of the multiplexed services to enable receiver processing. Also, in a TDM based system arranging the signaling channel in time domain bursts in form of transmission frame preambles is a natural approach. In general the signaling content should be designed to enable as easy receiver operations as possible. Trying to make things easy for the receivers is a prevailing idea in broadcast systems design. This is natural, as there may be thousands of receivers for one transmitter. Therefore it is more economical to have a more complex transmitter if the receiver cost is reduced accordingly. By studying the receiver operations, the performance requirements for the signaling can be defined as observed in chapter 6.2.4. It was for example observed that the error rate requirement for the signaling in the preamble may be relaxed when in-band signaling is used. Moreover, the requirements for the signaling have an effect on the system architecture. Continuing the DVB-T2 example, the required channel switching time sets requirements for the delay of the service data. This together with the multiplexing scheme and receiver processing for the signaling information sets the delay requirements for the signaling. The delay requirement for the signaling defines how often the signaling should be repeated and thus affects the transmission frame duration.

The simulator proves useful in designing the signaling and data paths in parallel. The availability of the simulator makes it possible to compare the performance of both paths in different usage environments already at

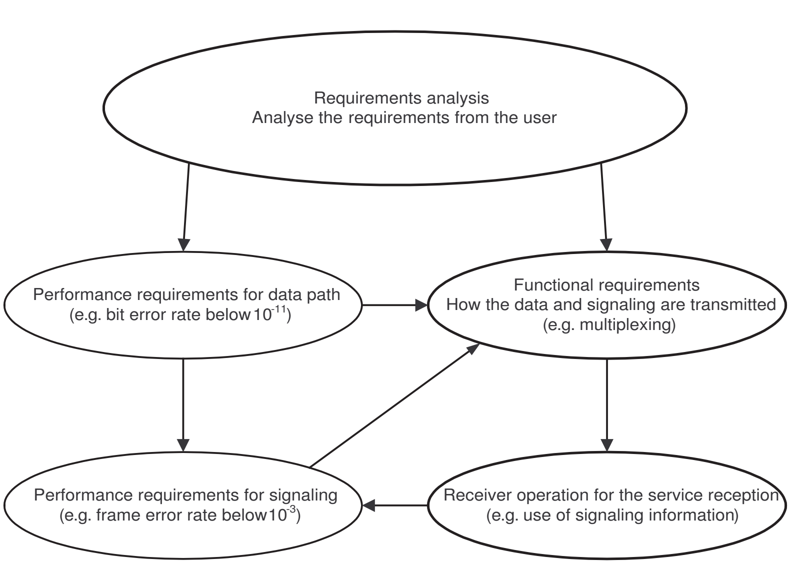

Figure 7.1: Relationship of the performance requirements for the signaling and data paths.

the design phase. Similar comparisons as presented in chapter 6.2.3 can be made. As it was observed, selection of comparable error criteria for data and signaling paths should be carefully considered. As there may be several ways to obtain the signaling (preamble and in-band in DVB-T2) directly the bit error rate comparison between the paths is not representative enough. Already considering frame error rate for the signaling proves to be more representative. There is no sense in designing the system so that the signaling introduces the bottleneck for the system performance. Therefore the signaling should be designed according to more strict error criteria than the data path.

When the signaling is designed for a mobile broadcasting system, diversity plays an important role. As it was observed in chapter 6.2.3 the performance of the physical layer signaling in the mobile environment was severely degraded due to small time diversity in the preamble of the DVB-T2 system. As the DVB-T2 system was mainly designed for stationary reception, the design trade-off of having a short preamble for centralized signaling was made. Therefore, the diversity that can be obtained for the signaling in the preamble is small compared to the data. If the system were destined mainly for mobile reception, the signaling should possess diversity at least of the same order of magnitude as for the data to enable robustness gain for signaling over the data path by using stronger FEC coding. Repetition of the signaling is natural for broadcasting systems. Therefore if the signaling content remains constant, time diversity is obtained already from the repetitive nature of the signaling.

As there is not much to gain by enhancing the forward error control codes, the next promising step towards increased capacity and robustness in broadcasting systems is MIMO. Having multiple antennas at both transmitting and receiving antenna arrays allows for increased throughput or increased robustness or balance between both. One promising scheme for utilizing MIMO in a SFN broadcasting network is presented in [95]. The presented MIMO coding is based on a two encoding layer structure where the encoded information is distributed to different transmission masts. Naturally, introduction of MIMO increases the complexity of the receivers as several antennas and decoding logic for MIMO are necessary. Also, the gain that can be obtained with MIMO on a broadcast system differs from the bi-directional system because the transmitter cannot obtain channel information and only the receivers have access to noisy channel estimates [96]. For mobile receivers of small physical dimensions obtaining the necessary separation for the antennas to obtain independent fading paths requires utilization of high frequencies in the transmissions. Utilization of a cross polarized MIMO transmission scheme as presented in [97] eases the antenna design for future mobile MIMO receivers [98]. From the design point of view, when the data path utilizes MIMO, also the signaling path should utilize it to obtain similar diversity order and thus appropriate robustness for the signaling is obtainable.

If scalable video codes are used for sharing the service components between HDTV and mobile services, the main system should be the HDTV system and the mobile receivers would use only a subset of service components provided in the networks. As the scalable video codecs evolve, sharing the service components potentially leads to more efficient use of the available resources than simulcasting the same service for both types of receivers with for example different video resolution. The use of SVC and simulcasting is studied in [99]. As this together with the graceful degradation are rather remarkable advantages, in future broadcasting systems the use of SVC seems probable. To enable the use of SVC the physical layer data path should enable applying individual robustness levels for the service components and the signaling information should support signaling services comprising of several layers.

Chapter 8

Summary

The goal of this thesis was to study how the signaling and data paths are designed and analyzed in wireless broadcasting systems to provide certain QoS for the users. For broadcasting systems mainly diversity and forward error control coding can be used for QoS provisioning. Therefore the FEC and diversity mechanisms were studied in detail. Also, the design process for the broadcasting systems was studied. For existing systems, the DVB-H link layer performance for signaling and data paths was studied and their effect on QoS considered. For DVB-T2, the design process was studied and performance of signaling and data paths compared ending up in configuration guidelines for DVB-T2 networks. Finally the future evolution of broadcasting systems was contemplated based on the studies on DVB-H and DVB-T2 systems.

The performance of the DVB-H link layer MPE-FEC with different decoding methods was studied by theoretical calculations. Both the error correction capability and complexity were analyzed. The performance of the link layer signaling was studied using simulations. Physical layer error masks from both laboratory and field measurements were used to evaluate the performance of the signaling. Field measurements enabled the comparison to the performance of the data path in a realistic environment. It was observed that the MPE-FEC decoding method utilizing the most accurate erasure information had the best performance. By calculating the complexity of the Reed-Solomon decoding it was further observed that the number of multiplications required by the MPE-FEC decoder was minimized with the decoding method that had the best performance of the considered ones. From the signaling studies the observation was that the link layer signaling can be configured to be more robust than the data path. This requires that some effort is made by the network operator to select appropriate section sizes for the signaling transmission. Mechanism for the selection was presented. Having more robust signaling transmission in broadcasting systems is necessary for service consumption and discovery.

The design of the DVB-T2 system physical layer was studied in detail, focusing on the design of signaling and data paths and their relationship. The utilization of simulations in the course of the design process was stressed. Further the performance of the signaling and data paths was compared by system simulations. It was observed that a simulator is an excellent tool to have available when designing a system. With simulator the robustness of the different paths can be effectively compared already at the design phase. Simulating the performance of the DVB-T2 system with a mobile channel model, it became clear that the performance of signaling with short time diversity is degraded severely. Therefore, when designing future systems for mobile reception, also the diversity for the signaling transmission should be carefully considered. Guidelines for configuring and a mechanism for selecting suitable transmission parameters for the DVB-T2 signaling were given based on the results of the simulations.

Ideas on how the signaling and data paths should be designed together in future systems specially for mobile reception to avoid complications were presented. The performance requirements for the signaling design should be derived from those set by the user for the service. In general, the requirements should be more demanding as the signaling is the enabler for service reception. To obtain more reliable signaling than the data path in a mobile reception environment similar or greater order of diversity should be provided for the signaling than for the service data path, and modulation and coding parameters chosen accordingly. To summarize the thesis:

In Chapter 2 the methods for guaranteeing QoS in broadcasting systems were studied. First, the specialties of the broadcasting environment were described. Then the concept of diversity was discussed. Further, the mechanisms for enhancing the user experience were depicted. These mechanisms consist of FEC codes and interleaving schemes for utilizing the available diversity in the broadcasting environment.

In Chapter 3 the system design process was studied. The use of simulations in the design process was also described. Then the design of coding, interleaving and modulation in broadcasting systems was studied in more detail. More specifically, what the requirements for the broadcasting system design are and what are the trade-offs that can be made between different resources in the design phase. The design aspects of the signaling were also considered.

In Chapter 4 the current status of the digital broadcasting systems worldwide was presented. Also main characteristics of the terrestrial DVB-T, DVB-H and DVB-T2 were presented. It was observed that there are several different standards for terrestrial broadcasting of digital services. More emphasis was put on the DVB-family standards as they are the topics of the case studies in this thesis. The similarities and differences of the first and second generation systems were indicated.

In Chapter 5 performance analysis of the DVB-H system link layer was discussed. Both signaling and data paths were studied and their performance compared. For the data path the performance of the MPE-FEC was studied with different decoding schemes and the complexities for the decoding schemes were calculated in Appendix A. For the link layer signaling the performance was evaluated by simulations and also measured data from the field was used to study the performance of the signaling in a real life environment. The inclusion of the link layer signaling in the scope of the MPE-FEC for increased robustness was briefly discussed.

In Chapter 6 the DVB-T2 standard and its design process were investigated. Use of simulations for the design of modulation, coding and interleaving was illustrated. The performances of the data and signaling paths in the DVB-T2 system were simulated and guidelines for the selection of transmission parameters for the signaling were presented based on the simulation results. From the simulations it was observed that the signaling preamble does not necessarily possess long enough time diversity for use in the mobile reception environment therefore possibly limiting the use of DVB-T2 for mobile broadcasting.

In Chapter 7 future evolution of broadcasting systems was contemplated. The possible relationship of future broadcasting systems to other telecommunication networks such as cellular networks was considered. It is possible that the future will show some amount of convergence between these networks. Further, the findings of the studies with DVB-H and DVB-T2 in chapters 5 and 6 affecting the technical design in future broadcasting systems were pointed out. It was observed that though the modern error control coding schemes utilized in the state of the art broadcasting systems perform close to the theoretical capacity limit, there is still room for overall system performance enhancement on other system parts. One currently hot topic is the utilization of MIMO schemes.

In future the signaling on all the system layers could be studied to enable cross-layer optimization of signaling content and transmission. Also, the convergence of telecommunication and broadcasting networks provides interesting research topics for system design due to the availability of a feedback channel through the telecommunication network.

Bibliography

[1] "The 3rd generation partnership project (3GPP), http://www.3gpp.org."

[2] ETSI, "Digital Video Broadcasting (DVB); Framing Structure, channel coding, and modulation for digital terrestrial television", European Telecommunication Standard, EN 300-744, January 2004.

[3] ETSI, "Radio broadcasting systems; Digital Audio Broadcasting (DAB) to mobile, portable and fixed receivers", European Telecommunication Standard, ETS 300 401, May 1997

[4] ETSI, "Digital Video Broadcasting (DVB); Transmission System for Handheld Terminals (DVB-H)", European Telecommunication Standard, EN 302-304, November 2004.

[5] D. Gómez-Barquero, *Cost Efficient Provisioning of Mass Mobile Multimedia Services in Hybrid Cellular and Broadcasting Systems*, PhD thesis, Universidad Politecnica de Valencia, Spain, 2009

[6] ETSI, "Digital Video Broadcasting (DVB); Frame structure channel coding and modulation for a second generation digital terrestrial television broadcasting system (DVB-T2)", European Telecommunication Standard, EN 302-755 (V1.1.1), September 2009.

[7] H. Himmanen, *On Transmission System Design for Wireless Broadcasting*, PhD thesis, University of Turku, Turku 2009

[8] A. Ganz, Z. Ganz and K. Wongthavarawat, *Multimedia Wireless Networks Technologies, Standards and QoS*, Prentice Hall, 2004

[9] G. Fairhurst, "Request for Comments (RFC) 3366 Best Current Practise: Advice to link designers on link Automatic Repeat reQuest (ARQ)", IETF RFC, August 2002

[10] S. Lin and D. J. Costello, *Error Control Coding*, 2nd Edition, Prentice Hall, 2004

[11] P. Unger and D. Gómez-Barquero, "Coverage gain estimation for multi-burst forward error correction in DVB-H networks", in *Proc. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, May 2009

[12] ITU-T recommendation X.200, Information Technology - Open Systems Interconnection - Basic Reference Model: The Basic Model, July 1994

[13] T. Koski et. al., "Five-level DVB-H Protocol Analysis for Interoperability Testing Purposes", in *Proc. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting 2009*, May 2009

[14] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*, Cambridge University Press, 2005.

[15] T. S. Rappaport, *Wireless Communications, Principles & Practise*, Prentice Hall, 1996.

[16] A. Hazmi, *Studies in Digital TV Signal Processing: Impulse Noise Mitigation, Repeater Loop Interference Cancellation, and DVB-T Transmission in CATV Networks*, PhD thesis, Tampere University of Technology, Tampere 2007.

[17] S. Haykin, *Communication Systems*, 4th edition. John Wiley & Sons, 2001.

[18] B. Sklar, *Digital Communications*, Second Edition. Prentice Hall, 2001

[19] S. M. Alamouti, "A simple transmit diversity technique for wireless communications", *IEEE Journal on Selected Areas in Communications*, Vol. 16, Issue 8, Pages 1451 - 1458, October 1998

[20] R. E. Ziemer and R. L. Peterson, *Introduction to Digital Communication*, Second Edition. Prentice Hall, 2001

[21] E. Biglieri, *Coding for Wireless Channels*, Springer US, 2005

[22] H. Ma and J. Wolf, "On Tail Biting Convolutional Codes", *IEEE Transactions on Communications*, Vol.34, No.2, Pages 104 - 111, February 1986

[23] R. W. Hamming, "Error Detecting and Error Correcting Codes", *Bell System Technical Journal*, April 1950

[24] A. J. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm", *IEEE Transactions on Information Theory*, Vol. 13, Pages 260-269, April 1967

[25] D. J. C. MacKay and R. M. Neal, "Near Shannon Limit Performance of Low Density Parity Check Codes", *IEEE Electronics Letters*, Vol. 33, Issue 6, Pages 457-458, March 1997

[26] C. Berrou, A. Glavieux and P. Thitimajshima, "Near Shannon Limit Error-Correcting Coding and Decoding: Turbo-codes", in *Proc. IEEE International Communications Conference*, Geneva Switzerland, May 1993

[27] S. Y. Chung, G.D. Forney Jr., T. J. Richardson and R. Urbanke, "On the Design of Low-Density Parity-Check Codes within 0.0045 dB of the Shannon Limit", *IEEE Communications Letters*, Vol. 5, Issue 2, Pages 58-60, February 2001

[28] J. Boutros et al., "Turbo code at 0.03 dB from capacity limit", in *Proc. IEEE International Symposium on Information Theory*, Lausanne Switzerland, July 2002

[29] G. D. Forney Jr., "Burst-Correcting Codes for the Classic Bursty Channel", *IEEE Transactions on Communication Technology*, vol. COM-19, Pages 772-781, October 1971

[30] J. L. Ramsey, "Realization of Optimum Interleavers", *IEEE Transactions on Information Theory*, vol. IT-16, no. 3, Pages 338-345, May 1970

[31] Department of Defence, *Systems Engineering Fundamentals*, Defence Acquisition University Press, Fort Belvoir, Virginia (USA), January 2001

[32] M. C. Jeruchim, P. Balaban and K. S. Shanmugan, *Simulation of Communication Systems: Modeling, Methodology and Techniques*, Second Edition. Kluwer Academic/Plenum Publishers, 2000.

[33] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing, Principles, Algorithms, and Applications*, Third Edition. Prentice Hall, 1996.

[34] H. Himmanen, M. M. Hannuksela, T. Kurki and J. Isoaho, “Objectives for New Error Criteria for Mobile Broadcastind of Streaming Audiovisual Services”, *EURASIP Journal on Advances in Signal Processing*, Hindawi Publishing, Vol. 2008, Article ID 518219

[35] U. H. Reimers, “DVB - The Family of International Standards for Digital Video Broadcasting”, *Proceedings of the IEEE*, vol. 94, no. 1, Pages 173-182, January 2006

[36] U. H. Reimers, *DVB - The Family of International Standards for Digital Video Broadcasting*, Second Edition. Springer, 2005

[37] ETSI, “Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for 11/12 GHz satellite services”, European Telecommunication Standard, EN 300-421, August 1997.

[38] ETSI, “Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for cable systems”, European Telecommunication Standard, EN 300-429, April 1998.

[39] ETSI, “Digital Video Broadcasting (DVB); Specification for Service Information (SI) in DVB systems”, European Telecommunication Standard, EN 300-468, June 2004.

[40] ETSI, “Digital Video Broadcasting (DVB): Second generation framing structure, channel coding and modulation for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications”, European Telecommunication Standard, EN 302-307 (V1.1.1), June 2004.

[41] DVB Document A138, Frame structure channel coding and modulation for a second generation digital transmission system for cable systems (DVB-C2), April 2009.

[42] U.Ladebusch and C. A. Liss, “Terrestrial DVB (DVB-T): A Broadcast Technology for Stationary Portable and Mobile Use”, *Proceedings of the IEEE*, vol. 94, No. 1, Pages 183-193, January 2006

[43] DVB Fact Sheet: Digital Terrestrial Television - The worlds most flexible and most successful DTT standard, available online (www.dvb.org), cited October 2009

[44] B. Furht and S. Ahson, *Handbook of Mobile Broadcasting - DVB-H, DMB, ISDB-T, and MEDIAFLO*, CRC Press, 2008

[45] ISO/IEC 13818-1, "Information Technology - Generic Coding of Moving Pictures and Associated Audio Information. Part 1: Systems", Second Edition, 2000-12-01.

[46] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series", *Math. Comput.* 19:297-301, 1965

[47] ETSI, "Digital Video Broadcasting (DVB): DVB specification for data broadcasting", Standard, EN 301 192 (V1.4.1), November 2004

[48] J. Väre and M. Puputti, "Soft Handover in terrestrial broadcast networks", in *Proc. IEEE International Conference on Mobile Data Management*, pages 236-242, 2004

[49] DVB bluebook: Commerical Requirement for DVB-T2, April 2007. Available online: *www.dvb.org*, cited 23.11.2009.

[50] DVB document: DVB-T2 Call for Technologies, April 2007. Available online: *www.dvb.org*, cited 23.11.2009.

[51] DVB Fact Sheet: 2nd Generation Terrestrial - The worlds most advanced Digital Terrestrial TV system, available online (www.dvb.org), cited November 2009

[52] R. Burow et al, "On the Performance of the DVB-T System in Mobile Environments", *Multimedia Applications, Services and Techniques, Lecture Notes in Computer Science*, vol. 1425, Pages 467-480, Springer Berlin, 1998

[53] Qing-Song Wen et al, "The Implement of Low-PAPR OFDM System", in *Proc. International Conference on Communications, Circuits and Systems Proceedings*, vol.2, pp.1226-1229, June 2006

[54] ETSI, "Digital Video Broadcasting (DVB); Generic Stream Encapsulation (GSE) Protocol", Technical Specification, TS 102 606 (V1.1.1), October 2007

[55] DVB Fact Sheet: Generic Stream Encapsulation - Enabling the carriage of IP directly over DVB networks, available online (www.dvb.org), cited December 2009

[56] G. Faria, J.A. Henriksson, E. Stare and P. Talmola, "DVB-H: Digital Broadcast Services to Handheld Devices", *Proceedings of the IEEE*, vol. 94, no. 1, Pages 194-209, Jan. 2006.

[57] ETSI, "Digital Video Broadcasting (DVB); IP Datacast over DVB-H: Content Delivery Protocols", Technical Specification, TS 102 472 (V1.2.1), December 2006

[58] D. Gómez-Barquero, A. Bria, "Application Layer FEC for Improved Mobile Reception of DVB-H Streaming Services", in *Proc. Vehicular Technology Conference*, September 2006

[59] ETSI, "Digital Video Broadcasting (DVB); DVB-H Implementation Guidelines", Technical report, TR 102 377 (V0.1.0), Jan. 2005

[60] R. Koetter and A. Vardy, "Algebraic soft-decision decoding of Reed-Solomon codes", *IEEE Transactions on Information Theory*, vol. 49, no. 11, Pages 2809-2825, Nov. 2003.

[61] H. Joki and J. Paavola, "A Novel Algorithm for Decapsulation and Decoding of DVB-H Link Layer Forward Error Correction", in *Proc. IEEE International Conference on Communications*, June 2006

[62] H. Joki, *Modeling of DVB-H link layer*, M.Sc. thesis, Helsinki University of Technology, Helsinki, Finland, May 2005

[63] R.E. Blahut, *Algebraic Codes for Data Transmission*, Cambridge University Press, 2003.

[64] M. Bossert, *Channel Coding for Telecommunications*. John Wiley & Sons, Chichester, 1999.

[65] A. Dür, "On the computation of the performance probabilities for block codes with a bounded-distance decoding rule", *IEEE Transactions on Information Theory*, vol. 34, Pages 70-78, January 1988.

[66] P. Koopman, "Cyclic redundancy codes for Internet applications", in *Proc. International Conference on Dependable Systems and Networks*, June 2002

[67] H. Joki and J. Poikonen, "Analysis and Simulation of DVB-H Link Layer", TUCS Technical Report, No 695, June 2005

[68] Kar-Ming Cheung and R.J. McEliece, "The undetected error probability for Reed-Solomon codes", in *Proc. Military Communications Conference*, vol. 1, October 1988

[69] T. Kasami and S. Lin, "On the Probability of Undetected Error for the Maximum Distance Separable Codes", *IEEE Transactions on Communications*, Vol. COM-32, No. 9, September 1984

[70] ETSI, "IP Datacast over DVB-H: Program Specific Information (PSI)/Service Information (SI)", Technical Specification, TS 102 470 (V1.1.1), April 2006

[71] ETSI, "Digital Video Broadcasting (DVB): Guidelines on implementation and usage of Service Information (SI)", Technical Report, TR 101 211 (V1.6.1), May 2004

[72] ETSI, "Digital video broadcasting (DVB); Specification for Service Information (SI) in DVB Systems.", Standard, EN 300 468 (V1.6.1) , November 2004.

[73] J. Väre, J. Alamaunu, H. Pekonen, T. Auranen, "Optimization of PSI/SI Transmission in IPDC over DVB-H Networks", in *Proc. 56th Annual IEEE Broadcast Symposium*, September 2006

[74] 3GPP Technical Specification, Digital Cellular Telecommunications System (Phase 2+); Radio Transmission and Reception, Nov. 2005, TS 05.05 v.8.20

[75] Wing TV project (EUREKA/Celtic), "Deliverable 15: Simulation report", August 2006

[76] Wing TV project (EUREKA/Celtic), "Deliverable 7: Country field trial report", November 2006

[77] R. G. Gallager, *Low-Density Parity-Check Codes*, Monograph, M.I.T. Press, 1963

[78] Mustafa Eroz, Feng-Wen Sun and Lin-Nan Lee, "DVB-S2 low density parity check codes with near Shannon limit performance", *International Journal of Satellite Communications and Networking*, Vol. 22, Issue 3, Pages 269-279, June 2004

[79] ETSI, "Digital Video Broadcasting (DVB): User guidelines for the second generation system for Broadcasting, Interactive Services, News Gathering and other broadband satellite applications (DVB-S2)", Technical Report, TR 102 376 (V1.1.1), February 2005

[80] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd edition, Johns Hopkins University Press, Baltimore, Maryland 1996

[81] T. Jokela, "Performance Analysis of Substituting DVB-S2 LDPC Code for DVB-T Error Control Coding System", in *Proc. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, April 2008

[82] Ki Seol Kim et al, "General Log-Likelihood Ratio Expression and Its Implementation Algorithm for Gray-Coded QAM Signals", *ETRI journal*, Volume 28, Number 3, Pages 291-300, June 2006

[83] M. Rezaei, I. Bouazizi and M.Gabbouj, "Statistical Time-Frequency Multiplexing of HD Video Traffic in DVB-T2", *International Journal of Digital Multimedia Broadcasting*, Hindawi Publishing, Vol. 2009, Article ID 186960

[84] Lukasz Kondrad et al, "Cross-Layer Optimized Transmission of H.264/SVC streams over DVB-T2 Broadcast System", in *Proc. IEEE International Symposium on Broadband Multimedia Systems and Broadcasting 2009*, May 2009

[85] IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Broadband Wireless Access Systems, 802.16-2009

[86] DVB document A133, Implementation guidelines for a second generation digital terrestrial television broadcasting system (DVB-T2), February 2009

[87] J.J. van de Beek, M. Sandell, P.O. Borjesson, "ML estimation of time and frequency offset in OFDM systems", *IEEE Transactions on Signal Processing*, vol. 45, issue 7, Pages 1800-1805, July 1997

[88] G. E. Moore, "Progress in digital integrated electronics", International Electron Devices Meeting, 1975

[89] P. A. Laven, "Predicting the future of broadcasting", EBU Technical Review, Summer 1998

[90] S. Rainisto, "Digiboksi herättää hylkimisreaktion", Talouselämä, (in English: The customers do not like to buy set-top boxes), 19th April 2006

[91] H. Schwarz, D. Marpe and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 17, No. 9, Pages 1103-1120, September 2007

[92] C. Hellge et al, "3GPP MBMS Mobile-TV Service Using H.264/AVC Temporal Scalability and Layered Transmission", in *Proc. 15th IEEE International Conference on Image Processing*, October 2008

[93] SARFT P.R. China, Mobile Multimedia Broadcasting Part 1: Framing Structure, Channel Coding and Modulation for Broadcasting Channel", GY/T 220.1, 2006

[94] R. Keller et al, "Covergence of Broadcast and New Telecom Networks", *Wireless Personal Communications*, no. 17, Pages 269-282, Kluwer Academic Publishers, 2001

[95] Y. Nasser, J.-F. Hélard and M. Crussière, "3D MIMO scheme for broadcasting future digital TV in single-frequency networks", *IEEE Electronics Letters*, Vol. 44, No. 13, June 2008

[96] P. Piantanida and P. Duhamel,"On the Capacity of the Fading MIMO Broadcast Channel Without Channel Information at the Transmitter and Imperfect Estimation at the Receivers", in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, April 2007

[97] J.D. Mitchell, P.N. Moss and M.J. Thorp, "A dual polarisation MIMO broadcast TV system", WHP 144, BBC Research

[98] R. Serrano, S. Blanch, L. Jofre, "Small antenna fundamentals and technologies: Future trends", in *Proc. IEEE First European Conference on Antennas and Propagation*, November 2006

[99] Z. Avramova, D. De Vleeschauwer, P. Debevere,"Performance of Scalable Video Coding for a TV Broadcast Network with Constant Video Quality and Heterogeneous Receivers", in *Proc. IEEE 10th International Conference on Telecommunications*, June 2009

[100] T. Jokela, *Study on the efficiency of error control coding in DVB-H systems*, Master's thesis, University of Turku, Finland, September 2005

[101] W. G. Horner, "A New Method of Solving Numerical Equations of All Orders, by Continuous Approximation", *Philosophical Transactions of the Royal Society of London*, Vol. 109, pages 308-335, 1819

Appendix A

MPE-FEC complexity analysis

Let us study the complexities of different implementations of RS decoder presented in [63] based on number of Galois Field (GF) operations. Two known bounded distance decoding algorithms for Reed-Solomon codes investigated. To begin with frequency domain decoder is studied. Then the algorithm in the time domain is also investigated to compare their complexity. Both decoding algorithms are based on the Berlekamp-Massey algorithm. A brief study on the complexities of both algorithms based on simulation run-time was performed in [100] and here a more analytical approach is taken. Further, the effect of the decoding schemes presented in previous section on the complexity of the total MPE-FEC decoding is investigated.

A.1 Frequency domain RS decoding algorithm description

Broadly speaking, decoding in frequency domain means Fourier transforming the received vector and performing the decoding tasks on the transformed vector. The Fourier transform in GF is described for example in [63]. In the end of decoding the inverse Fourier transform is taken and a codeword is obtained. This way of decoding is natural for Reed-Solomon codes due to their definition.

The number or erasures in a received vector v is denoted by t_e and the number of undetected errors in the vector is denoted by t_u. The codeword is uniquely decodable under bounded distance decoding whenever t_e and t_u obey the equation (5.2), or slightly rewritten,

$$d \geq 2t_u + t_e + 1,$$

where d is code minimum distance.

We need one more definition before going to the decoding algorithm. Let us define that vectors $\mathbf{e}$ and $\mathbf{f}$ correspond to the error- and erasure-vector of a received vector $\mathbf{v}$, that is, for some codeword $\mathbf{c}$:

$$\mathbf{c} = \mathbf{v} - (\mathbf{e} + \mathbf{f}).$$

Then the *error-erasure-locator polynomial* Λ is the polynomial, which has minimal degree and satisfies the (cyclic) convolutional equation

$$\mathbf{\Lambda} * (\mathbf{E} + \mathbf{F}) = \mathbf{0}.$$

Here we identify the vector

$$\mathbf{\Lambda} = (\Lambda_0, \ldots, \Lambda_{n-1})$$

components of which are the coefficients of the polynomial

$$\Lambda(x) = \Lambda_0 + \ldots + \Lambda_{n-1}x^{n-1}.$$

The next decoding algorithm is represented as in [63].

Algorithm A.1.1 (Frequency domain decoder)**.**

1. Let us assume, that $\mathbf{v}$ is the received vector having $t_e < d$ erasures in the locations $i_1, \ldots, i_{t_e}$, where d is the distance of the code. First, these erasures are filled with zeros. Next, the erasure-locator polynomial $\Psi(x)$ is determined from the formula

   $$\Psi(x) = \prod_{l=1}^{t_e}(1 - x\omega^{i_l}),$$

   where ω is a primitive element of the Galois field $GF(q)$. Also Fourier transform of $\mathbf{v}$ is taken, thus obtaining the vector $\mathbf{V}$.

2. The Berlekamp-Massey algorithm explained below is executed on the vector $\mathbf{V}$ with Ψ to obtain the error-erasure-locator polynomial $\Lambda(x)$.

3. Next, the known $d-1$ values of $\mathbf{E} + \mathbf{F}$ are extended recursively to n values by using the convolutional equation

   $$\mathbf{\Lambda} * (\mathbf{E} + \mathbf{F}) = \mathbf{0}$$

   and the known values of Λ.

4. The decoding is completed by taking the inverse Fourier transform of the vector

   $$\mathbf{C} = \mathbf{V} - (\mathbf{E} + \mathbf{F}).$$

Next the step 2 of the algorithm is studied more carefully. The Berlekamp-Massey algorithm may be written in the following form:

Algorithm A.1.2 (Berlekamp-Massey in frequency domain [63]).

1. Initialization: $\Lambda^{(t_e)}(x) = B^{(t_e)}(x) = \Psi(x), L_{t_e} = 0, r = t_e$, where $\Psi(x)$ is the erasure-locator polynomial.

2. For $r = t_e + 1, \ldots, d-1$:

$$\Delta_r = \sum_{j=0}^{L_{r-1}+t_e} \{\Lambda^{(r-1)}(x)\}_j V_{r-1-j},$$

$$L_r = t_r(r - t_e - L_{r-1}) + (1 - t_r)L_{r-1},$$

$$\begin{pmatrix} \Lambda^{(r)}(x) \\ B^{(r)}(x) \end{pmatrix} = \begin{pmatrix} 1 & -\Delta_r x \\ t_r \Delta_r^{-1} & (1-t_r)x \end{pmatrix} \cdot$$
$$\cdot \begin{pmatrix} \Lambda^{(r-1)}(x) \\ B^{(r-1)}(x) \end{pmatrix},$$

   where $t_r = 1$ if both $\Delta_r \neq 0$ and $2L_{r-1} \leq r - t_e - 1$, otherwise $t_r = 0$.

3. Then $\Lambda^{(d-1)}(x)$ is the locator polynomial for both errors and erasures for the sequence $V_0, \ldots, V_{d-2}$.

In the step 3 of the algorithm A.1.1 the convolutional equation needs to be solved:

$$\mathbf{\Lambda} * (\mathbf{E} + \mathbf{F}) = \mathbf{0}. \tag{A.1}$$

This is done iteratively by extending the known values of $\mathbf{E} + \mathbf{F}$ with the help of the vector $\mathbf{\Lambda}$ from the previous step. The equation (A.1) can be written in the following form:

$$(\mathbf{E} + \mathbf{F})_{((k))} = \sum_{j=1}^{t_u+t_e} \mathbf{\Lambda}_j (\mathbf{E} + \mathbf{F})_{((k-j))}, \quad k = d-1, \ldots, n-1 \tag{A.2}$$

Now the right-hand side uses only known values of the vector $\mathbf{E}+\mathbf{F}$, and hence all of its components can be iteratively obtained.

A.2 Time-domain decoding algorithm description

In time-domain decoding, the operations are performed on the received vector itself rather than taking the Fourier transform of it in the very beginning of the algorithm. However, with Reed-Solomon codes the natural approach is to use the Fourier transform at least implicitly and as we shall see, the algorithm in time domain is closely related to the one already discussed. The most evident way of decoding in time-domain is to take the inverse Fourier transform of each step of the algorithm A.1.1. The algorithm presented here is derived in this manner in [63].

Algorithm A.2.1 (Time-domain decoder)**.**

1. First, the erasure-locator polynomial $\Psi(x)$ is determined as in step 1) in algorithm A.1.1. Next, the inverse Fourier transform of the vector $\mathbf{\Psi}$ is taken to obtain a time-domain vector ψ.
2. The Berlekamp-Massey algorithm in time-domain explained below is executed on the vectors $\mathbf{v}$ and ψ. As a result, a time-domain erasure-error locator λ is obtained.
3. Let us denote $\mathbf{u}^{(d-2)} = \mathbf{v}$. Now, for $r = d-1, \ldots, n-1$, let
$$\begin{aligned} \Delta_r &= \sum_{i=0}^{n-1} \omega^{ir} \lambda_i u_i^{(r-1)} \\ u_i^{(r)} &= u_i^{(r-1)} - \Delta_r \omega^{-ir} \quad \text{for all } i. \end{aligned}$$
4. The code vector $\mathbf{c}$ can now be calculated as $\mathbf{c} = \mathbf{v} - \mathbf{u}^{(n-1)}$.

Again, it is time to concentrate on the Berlekamp-Massey algorithm, now in time-domain.

Algorithm A.2.2 (Berlekamp-Massey in time-domain [63])**.**

1. Initialization: $\lambda^{(t_e)} = b^{(t_e)} = \psi, L_{t_e} = 0, r = t_e$, where ψ is as above.
2. For $r = t_e + 1, \ldots, d-1$,
$$\begin{aligned} \Delta_r &= \sum_{j=0}^{n-1} \omega^{j(r-1)} \lambda_j^{(r-1)} v_j \\ L_r &= t_r(r - t_e - L_{r-1}) + (1 - t_r) L_{r-1} \\ \begin{pmatrix} \lambda_i^{(r)} \\ b_i^{(r)} \end{pmatrix} &= \begin{pmatrix} 1 & -\Delta_r \omega^{-i} \\ t_r \Delta_r^{-1} & (1-t_r)\omega^{-i} \end{pmatrix} \begin{pmatrix} \lambda_i^{(r-1)} \\ b_i^{(r-1)} \end{pmatrix} \end{aligned}$$
for all i, where $t_r = 1$ if both $\Delta_r \neq 0$ and $2L_{r-1} \leq r - t_e - 1$, otherwise $t_r = 0$.
3. Then $\lambda^{(d-1)}$ is the error-erasure locator.

A.3 Complexity analysis of the RS decoding algorithms

The complexities are mainly calculated with respect to the number of multiplications needed in the Galois field. Since the frequency domain decoder is based on Fourier transform, a few words concerning the Fourier transform's complexity are in order.

Fast Fourier transform (FFT) is a well-known, time-wise efficient recursive algorithm. Its complexity is of the order $O(n \log n)$, although this complexity does depend on how the exponentiation of the primitive element is organized. In the following, we assume that a look-up table (that is a table of logarithms and anti-logarithms with respect to the primitive element) is constructed, so that exponentiation of the primitive element ω can be seen as a constant-time operation.

Let us next study the complexity of some polynomial operations. Let us assume, that we have two polynomials in Galois field,

$$\begin{aligned} a(x) &= a_0 + a_1 x + \ldots + a_m x^m \quad \text{and} \\ b(x) &= b_0 + b_1 x + \ldots + b_n x^n, \end{aligned} .$$

To obtain their product (in monomial form) $c(x) = a(x)b(x)$, clearly $(m+1)(n+1)$ multiplications in the Galois field are needed. This number of multiplications may be decreased if more sophisticated methods are used. However, here we assume that the school algorithm for polynomial multiplications is used. If $a_0 = b_0 = 1$, the number of needed multiplications reduces to mn, since multiplying by 1 is a trivial operation. Let now α be some element of the Galois field. To calculate the substitution $a(\alpha)$, m multiplications are necessary, which can be seen by writing the polynomial in different form (Horner scheme [101]),

$$a(\alpha) = (\ldots((a_m\alpha + a_{m-1})\alpha + a_{m-2})\alpha \cdots)\alpha + a_0.$$

A.4 Frequency domain decoder complexity

In the first step of the algorithm A.1.1 the erasure-locator polynomial

$$\Psi(x) = \prod_{l=1}^{t_e} (1 - x\omega^{i_l})$$

is constructed. With the previous considerations it can be calculated, that to obtain $\Psi(x)$ in monomial form, $(1/2)(t_e - 1)t_e$ multiplications in the Galois field are needed. In this step, also Fourier transform is taken hence adding the term $n \log n$ to the overall complexity.

The second step requires calculating the complexity of the Berlekamp-Massey algorithm. The first observation is, that in algorithm A.1.2 always $L_r \leq r - t_e$. For every value of r the computation of Δ_r takes at most

$$L_{r-1} + t_e + 1 \leq (r-1) - t_e + t_e + 1 = r$$

multiplications. No multiplications in the Galois field are needed for computing L_r, but in the matrix equation at most $2(r+1)$ multiplications are performed. Altogether, for every value of r, at most

$$r + 2(r+1) = 3r + 2$$

multiplications in the Galois field are necessary. Counting over all values of r, the number of multiplications is at most

$$\begin{aligned}
\sum_{r=t_e+1}^{d-1} 3r + 2 &= 2(d - t_e - 1) + 3\big(\sum_{r=1}^{d-1} r - \sum_{r=1}^{t_e} r\big) \\
&= 2\big(d - t_e - 1\big) + \big(3/2\big)\big((d-1)d - t_e(t_e+1)\big) \\
&= 2(d - t_e - 1) + (3/2)(d + t_e)(d - t_e - 1) \\
&= \big(d - t_e - 1\big)\big((3/2)(d + t_e) + 2\big).
\end{aligned}$$

The complexity of the third step in the algorithm A.1.1 can be derived from the equation (A.2). The number of needed iterations of this equation is

$$(n-1) - (d-1) + 1 = n - d + 1.$$

For every iteration element, at most $t_u + t_e$ multiplications in the Galois field are performed. Thus, the complexity of the third step is at most $(t_u + t_e)(n - d + 1)$.

The fourth step of the algorithm introduces another Fourier transform and hence is of complexity $n \log n$.

The overall complexity of the decoding algorithm A.1.1 with respect to the number of multiplications in the Galois field $GF(q)$ is at most

$$\begin{gathered}
\mathcal{C}(t_u, t_e) = \tfrac{(t_e-1)t_e}{2} + 2n\log n + \\
\big(d - t_e - 1\big)\big((3/2)(d + t_e) + 2\big) + (t_u + t_e)(n - d + 1).
\end{gathered} \tag{A.3}$$

Note that one easily verifies that the number of additions needed in the algorithm A.1.1 is of the same order of magnitude as the number of multiplications. Hence the overall complexity of the algorithm is of the order mentioned in the equation (A.3).

A.5 Time-domain decoder complexity

Next, let us study the complexity of the algorithm A.2.1. It begins with nearly the same operations as in the previous case, the overall complexity of the first step being $(1/2)(t_e - 1)t_e + n \log n$.

The second step introduces the Berlekamp-Massey algorithm in time-domain. For each step of the iteration, calculating Δ_r takes at most $2n$ multiplications. The matrix multiplications take additional $4n$ multiplications, so the overall complexity for each iteration step is $6n$. Since there are $d - t_e - 1$ of these steps, the complexity of the Berlekamp-Massey algorithm is at most $6n(d - t_e - 1)$.

In the third step of the algorithm A.2.1 at most $3n$ multiplications are performed in each iteration. There are $n - d - 1$ iterations in total, and thus the complexity of this third step is at most $3n(n - d - 1)$.

Thus the complexity of the algorithm A.2.1 is at most

$$\mathcal{C}(t_u, t_e) = \frac{(t_e - 1)t_e}{2} + n \log n + 6n(d - t_e - 1) + 3n(n - d - 1). \tag{A.4}$$

This complexity has no dependence on the number of errors. This is mainly due to two reasons: First of all, the complexity analysis was rather coarse since the number of multiplications may be diminished by not counting the trivial multiplications. Secondly, this algorithm was constructed in such manner, that most of the calculations are done *for all* i, that is, for all components of a vector of size n.

A.6 Complexity of the decoding schemes in the DVB-H system

All the rows of the MPE-FEC frame are decoded with RS decoder that can utilize the erasure information. The link layer operations and the structure of the MPE-FEC frame is shown in Fig. 5.1. From the complexity point of view the thing that makes the difference when the code and decoding algorithm are fixed, is the numbers of undetected errors and erasures hitting RS codewords. The dependence of the complexity on numbers of errors and erasures and code parameters for frequency and time domain decoders for the RS decoding are given in equations (A.3) and (A.4). In MPE-FEC RS(255,191) code is used. As examples of possible scenarios for MPE-FEC decoder errors only, erasures only and combination of both are presented in Fig. A.1. The complexities are calculated from the expressions (A.3) and (A.4). From the complexity analysis of the RS decoding algorithms and from Fig. A.1 it is found out that the most efficient algorithm is the one operating in frequency domain. Its complexity is given in equation (A.3).

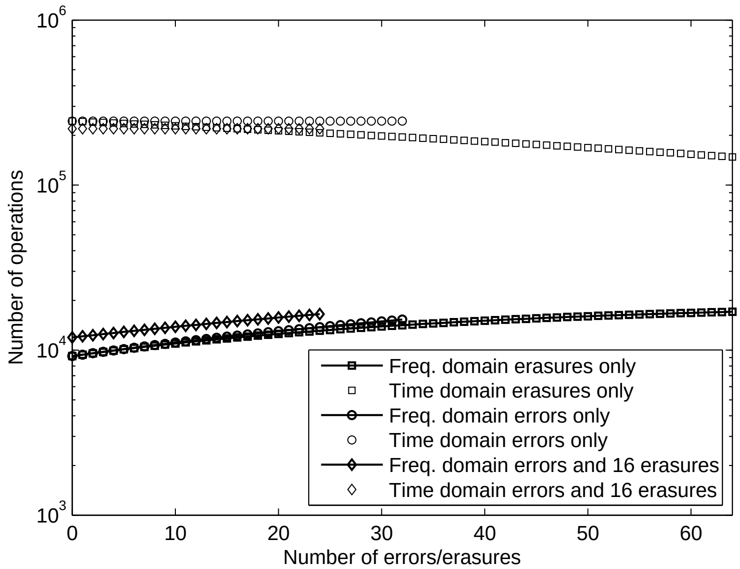

Figure A.1: Dependence of complexity on number of errors and erasures for RS(255,191), $d = 65$ code.

Let us further analyze the expression. First of all, it is noticed that the rightmost component of the sum,

$$(t_u + t_e)(n - d + 1), \tag{A.5}$$

is linear with respect to the sum of t_u and t_e. Let us now denote the actual number of erroneous symbols in the received code vector by κ. Then always

$$\kappa \leq t_u + t_e,$$

equality holding when no correct symbols are erased. Thus to minimize the linear term, there should be no erasures corresponding to correct symbols in the code vector.

Let us next study the term

$$\frac{(t_e - 1)t_e}{2} + \big(d - t_e - 1\big)\big((3/2)(d + t_e) + 2\big). \tag{A.6}$$

Let us denote it by $g(t_e)$ and then minimize it. The first derivative of this function is

$$\begin{aligned} g'(t_e) &= \frac{2t_e - 1}{2} - \frac{3(d + t_e) + 4}{2} + \frac{3(d - t_e - 1)}{2} \\ &= \frac{-4t_e - 8}{2} = 0, \end{aligned}$$

when $t_e = -2$. From this, and the form of expression (A.6) it is observed, that $g(t_e)$ is a decreasing function for $t_e \geq 0$. Thus to minimize the term (A.6), there should be as many erased symbols as possible. This reduces the overall complexity of the frequency domain decoder, since it reduces the complexity of Berlekamp-Massey algorithm in frequency domain. In fact, if it is known that decoding using only erasures can be done, the Berlekamp-Massey algorithm does not need to be executed at all, since the locations of the errors are already known.

Next the minimizing scheme for the whole algorithm A.1.1 needs to be found. For this expressions (A.5) and (A.6) need to be compared. Comparing the first derivatives of these terms with respect to the variable t_e it is found that

$$\begin{aligned}(n-d+1)+g'(t_e) &= n-d+1-2t_e-4 \\ &\geq n-d+1-2(d-1)-4 \\ &= n-3d-1 \geq 0\end{aligned}$$

when $3d \leq n-1$. This is the case for $n = 255$ and $d = 65$, which are the parameters in MPE-FEC Reed-Solomon code. From this it is deduced that for this code, taking extra erasures will only increase the complexity of the frequency domain decoding if those erasures correspond to correct symbols. Hence, to minimize the complexity of the decoding using frequency domain RS decoding, the most beneficial way is to have as reliable erasure information as possible. This is the main idea of the decoding schemes suggested in [61] and denoted in this dissertation as MHE and PE decoding schemes. The performance analysis of the section 5.1 shows that the PE decoding is the best of the discussed decoding schemes with respect to the probability of successful decoding. Here it is observed, that it also minimizes the complexity of the frequency domain RS decoder utilized in MPE-FEC.

The conclusion of the complexity analysis is that from the compared Reed-Solomon decoding algorithms, the one operating in frequency domain is better from the complexity point of view. Further, the significant result is that the MPE-FEC decoding methods suggested in [61] and studied in 5.1 having better error correction capabilities also minimize the amount of operations necessary in the Reed-Solomon decoding.

Turku Centre for Computer Science
TUCS Dissertations

94. **Dubravka Ilić**, Formal Reasoning about Dependability in Model-Driven Development
95. **Kim Solin**, Abstract Algebra of Program Refinement
96. **Tomi Westerlund**, Time Aware Modelling and Analysis of Systems-on-Chip
97. **Kalle Saari**, On the Frequency and Periodicity of Infinite Words
98. **Tomi Kärki**, Similarity Relations on Words: Relational Codes and Periods
99. **Markus M. Mäkelä**, Essays on Software Product Development: A Strategic Management Viewpoint
100. **Roope Vehkalahti**, Class Field Theoretic Methods in the Design of Lattice Signal Constellations
101. **Anne-Maria Ernvall-Hytönen**, On Short Exponential Sums Involving Fourier Coefficients of Holomorphic Cusp Forms
102. **Chang Li**, Parallelism and Complexity in Gene Assembly
103. **Tapio Pahikkala**, New Kernel Functions and Learning Methods for Text and Data Mining
104. **Denis Shestakov**, Search Interfaces on the Web: Querying and Characterizing
105. **Sampo Pyysalo**, A Dependency Parsing Approach to Biomedical Text Mining
106. **Anna Sell**, Mobile Digital Calendars in Knowledge Work
107. **Dorina Marghescu**, Evaluating Multidimensional Visualization Techniques in Data Mining Tasks
108. **Tero Säntti**, A Co-Processor Approach for Efficient Java Execution in Embedded Systems
109. **Kari Salonen**, Setup Optimization in High-Mix Surface Mount PCB Assembly
110. **Pontus Boström**, Formal Design and Verification of Systems Using Domain-Specific Languages
111. **Camilla J. Hollanti**, Order-Theoretic Mehtods for Space-Time Coding: Symmetric and Asymmetric Designs
112. **Heidi Himmanen**, On Transmission System Design for Wireless Broadcasting
113. **Sébastien Lafond**, Simulation of Embedded Systems for Energy Consumption Estimation
114. **Evgeni Tsivtsivadze**, Learning Preferences with Kernel-Based Methods
115. **Petri Salmela**, On Commutation and Conjugacy of Rational Languages and the Fixed Point Method
116. **Siamak Taati**, Conservation Laws in Cellular Automata
117. **Vladimir Rogojin**, Gene Assembly in Stichotrichous Ciliates: Elementary Operations, Parallelism and Computation
118. **Alexey Dudkov**, Chip and Signature Interleaving in DS CDMA Systems
119. **Janne Savela**, Role of Selected Spectral Attributes in the Perception of Synthetic Vowels
120. **Kristian Nybom**, Low-Density Parity-Check Codes for Wireless Datacast Networks
121. **Johanna Tuominen**, Formal Power Analysis of Systems-on-Chip
122. **Teijo Lehtonen**, On Fault Tolerance Methods for Networks-on-Chip
123. **Eeva Suvitie**, On Inner Products Involving Holomorphic Cusp Forms and Maass Forms
124. **Linda Mannila**, Teaching Mathematics and Programming – New Approaches with Empirical Evaluation
125. **Hanna Suominen**, Machine Learning and Clinical Text: Supporting Health Information Flow
126. **Tuomo Saarni**, Segmental Durations of Speech
127. **Johannes Eriksson**, Tool-Supported Invariant-Based Programming
128. **Tero Jokela**, Design and Analysis of Forward Error Control Coding and Signaling for Guaranteeing QoS in Wireless Broadcast Systems

Turku Centre *for* Computer Science

Joukahaisenkatu 3-5 B, 20520 Turku, Finland | www.tucs.fi

University of Turku
- Department of Information Technology
- Department of Mathematics

Åbo Akademi University
- Department of Information Technologies

Turku School of Economics
- Institute of Information Systems Sciences

ISBN 978-952-12-2462-1
ISSN 1239-1883

Tero Jokela

Design and Analysis of Forward Error Control Coding and Signaling for Guaranteeing QoS in Wireless Broadcast Systems

Tero Jokela