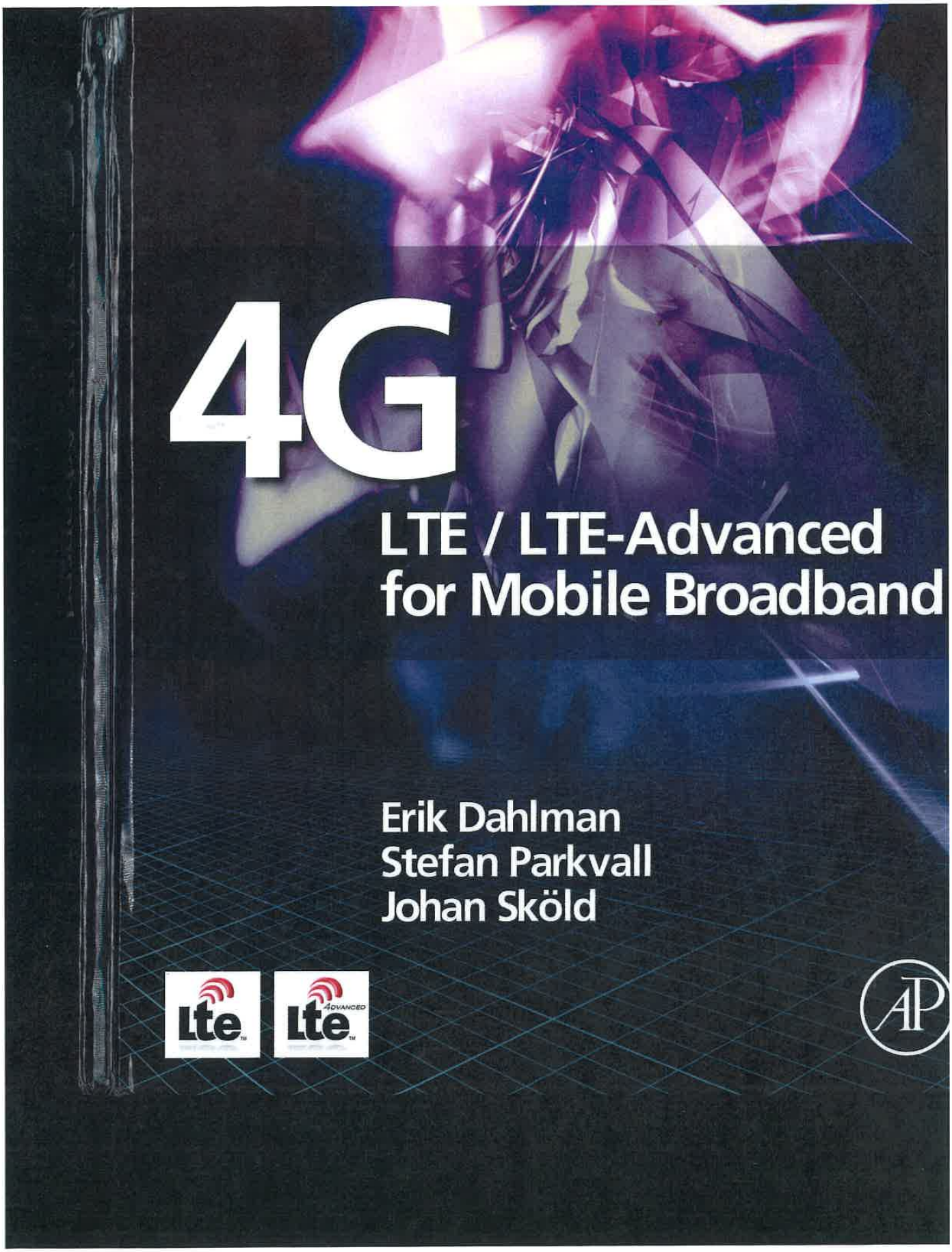

4G
LTE / LTE-Advanced
for Mobile Broadband
Erik Dahlman
Stefan Parkvall
Johan Sköld
Lte
Lte
4Advanced
AP

4G LTE/LTE-Advanced for Mobile Broadband

4G LTE/LTE-Advanced for Mobile Broadband

Erik Dahlman, Stefan Parkvall, and Johan Sköld



ELSEVIER

AMSTERDAM • BOSTON • HEIDELBERG • LONDON • NEW YORK • OXFORD
PARIS • SAN DIEGO • SAN FRANCISCO • SINGAPORE • SYDNEY • TOKYO

Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK
30 Corporate Drive, Suite 400, Burlington, MA 01803, USA

First published 2011

Copyright © 2011 Erik Dahlman, Stefan Parkvall & Johan Sköld. Published by Elsevier Ltd. All rights reserved

The rights of Erik Dahlman, Stefan Parkvall & Johan Sköld to be identified as the authors of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangement with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices
Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Control Number: 2011921244

ISBN: 978-0-12-385489-6

For information on all Academic Press publications
visit our website at www.elsevierdirect.com

Typeset by MPS Limited, a Macmillan Company, Chennai, India
www.macmillansolutions.com

Printed in the United States of America
Transferred to Digital Printing, 2012

Working together to grow
libraries in developing countries
www.elsevier.com | www.bookaid.org | www.sabre.org
ELSEVIER BOOK AID International Sabre Foundation

Contents

Preface xiii
Acknowledgements xv
Abbreviations and Acronyms xvii

CHAPTER 1 Background of LTE 1
1.1 Introduction 1
1.2 Evolution of Mobile Systems Before LTE 2
1.2.1 The First 3G Standardization 3
1.3 ITU Activities 4
1.3.1 IMT-2000 and IMT-Advanced 4
1.3.2 Spectrum for IMT Systems 6
1.4 Drivers for LTE 7
1.5 Standardization of LTE 8
1.5.1 The Standardization Process 8
1.5.2 The 3GPP Process 9
1.5.3 The 3G Evolution to 4G 11

CHAPTER 2 High Data Rates in Mobile Communication 15
2.1 High Data Rates: Fundamental Constraints 15
2.1.1 High Data Rates in Noise-Limited Scenarios 17
2.1.2 Higher Data Rates in Interference-Limited Scenarios 18
2.2 Higher Data Rates Within a Limited Bandwidth: Higher-Order Modulation 19
2.2.1 Higher-Order Modulation in Combination with Channel Coding 20
2.2.2 Variations in Instantaneous Transmit Power 20
2.3 Wider Bandwidth Including Multi-Carrier Transmission 21
2.3.1 Multi-Carrier Transmission 23

CHAPTER 3 OFDM Transmission 27
3.1 Basic Principles of OFDM 27
3.2 OFDM Demodulation 29
3.3 OFDM Implementation Using IFFT/FFT Processing 30
3.4 Cyclic-Prefix Insertion 32
3.5 Frequency-Domain Model of OFDM Transmission 34
3.6 Channel Estimation and Reference Symbols 35
3.7 Frequency Diversity with OFDM: Importance of Channel Coding 36

Samsung Exhibit 1019, Page 5 of 195

3.8 Selection of Basic OFDM Parameters 37
3.8.1 OFDM Subcarrier Spacing 38
3.8.2 Number of Subcarriers 39
3.8.3 Cyclic-Prefix Length 40
3.9 Variations in Instantaneous Transmission Power 40
3.10 OFDM as a User-Multiplexing and Multiple-Access Scheme 41
3.11 Multi-Cell Broadcast/Multicast Transmission and OFDM 43

CHAPTER 4 Wider-Band "Single-Carrier" Transmission 45
4.1 Equalization Against Radio-Channel Frequency Selectivity 45
4.1.1 Time-Domain Linear Equalization 45
4.1.2 Frequency-Domain Equalization 47
4.1.3 Other Equalizer Strategies 49
4.2 Uplink FDMA with Flexible Bandwidth Assignment 50
4.3 DFT-Spread OFDM 52
4.3.1 Basic Principles 52
4.3.2 DFTS-OFDM Receiver 54
4.3.3 User Multiplexing with DFTS-OFDM 55
4.3.4 Distributed DFTS-OFDM 55

CHAPTER 5 Multi-Antenna Techniques 59
5.1 Multi-Antenna Configurations 59
5.2 Benefits of Multi-Antenna Techniques 60
5.3 Multiple Receive Antennas 60
5.4 Multiple Transmit Antennas 65
5.4.1 Transmit-Antenna Diversity 65
5.4.2 Transmitter-Side Beam-Forming 68
5.5 Spatial Multiplexing 71
5.5.1 Basic Principles 71
5.5.2 Precoder-Based Spatial Multiplexing 74
5.5.3 Nonlinear Receiver Processing 75

CHAPTER 6 Scheduling, Link Adaptation, and Hybrid ARQ 79
6.1 Link Adaptation: Power and Rate Control 79
6.2 Channel-Dependent Scheduling 81
6.2.1 Downlink Scheduling 81
6.2.2 Uplink Scheduling 85
6.2.3 Link Adaptation and Channel-Dependent Scheduling in the Frequency Domain 87
6.2.4 Acquiring on Channel-State Information 87
6.2.5 Traffic Behavior and Scheduling 88

6.3 Advanced Retransmission Schemes 89
6.4 Hybrid ARQ with Soft Combining 90

CHAPTER 7 LTE Radio Access: An Overview 95
7.1 Basic Principles 96
7.1.1 Transmission Scheme 96
7.1.2 Channel-Dependent Scheduling and Rate Adaptation 97
7.1.3 Inter-Cell Interference Coordination 99
7.1.4 Hybrid ARQ with Soft Combining 99
7.1.5 Multi-Antenna Transmission 100
7.1.6 Spectrum Flexibility 100
7.2 LTE Release 9 102
7.2.1 Multicast and Broadcast Support 102
7.2.2 Positioning 103
7.2.3 Dual-Layer Beam-Forming 103
7.3 LTE Release 10 and IMT-Advanced 103
7.3.1 Carrier Aggregation 104
7.3.2 Extended Multi-Antenna Transmission 104
7.3.3 Relaying 105
7.3.4 Heterogeneous Deployments 105
7.4 Terminal Capabilities 105

CHAPTER 8 Radio-Interface Architecture 109
8.1 Overall System Architecture 109
8.1.1 Core Network 109
8.1.2 Radio-Access Network 111
8.2 Radio Protocol Architecture 111
8.2.1 Radio-Link Control 113
8.2.2 Medium-Access Control 115
8.2.3 Physical Layer 123
8.3 Control-Plane Protocols 124
8.3.1 State Machine 125

CHAPTER 9 Physical Transmission Resources 127
9.1 Overall Time–Frequency Structure 127
9.2 Normal Subframes and MBSFN Subframes 131
9.3 Carrier Aggregation 132
9.4 Frequency-Domain Location of LTE Carriers 134
9.5 Duplex Schemes 135
9.5.1 Frequency-Division Duplex (FDD) 136
9.5.2 Time-Division Duplex (TDD) 137
9.5.3 LTE and TD-SCDMA Coexistence 140

CHAPTER 10 Downlink Physical-Layer Processing 143
10.1 Transport-Channel Processing 143
10.1.1 Processing Steps 143
10.1.2 Localized and Distributed Resource Mapping 149
10.2 Downlink Reference Signals 152
10.2.1 Cell-Specific Reference Signals 153
10.2.2 Demodulation Reference Signals 156
10.2.3 CSI Reference Signals 158
10.3 Multi-Antenna Transmission 161
10.3.1 Transmit Diversity 163
10.3.2 Codebook-Based Precoding 165
10.3.3 Non-Codebook-Based Precoding 169
10.3.4 Downlink Multi-User MIMO 171
10.4 Downlink L1/L2 Control Signaling 173
10.4.1 Physical Control Format Indicator Channel 174
10.4.2 Physical Hybrid-ARQ Indicator Channel 177
10.4.3 Physical Downlink Control Channel 179
10.4.4 Downlink Scheduling Assignment 181
10.4.5 Uplink Scheduling Grants 189
10.4.6 Carrier Aggregation and Cross-Carrier Scheduling 193
10.4.7 Power-Control Commands 195
10.4.8 PDCCH Processing 195
10.4.9 Blind Decoding of PDCCHs 199

CHAPTER 11 Uplink Physical-Layer Processing 203
11.1 Transport-Channel Processing 203
11.1.1 Processing Steps 203
11.1.2 Mapping to the Physical Resource 205
11.1.3 PUSCH Frequency Hopping 207
11.2 Uplink Reference Signals 210
11.2.1 Uplink Demodulation Reference Signals 210
11.2.2 Uplink Sounding Reference Signals 217
11.3 Uplink Multi-Antenna Transmission 221
11.3.1 Precoder-Based Multi-Antenna Transmission for PUSCH 221
11.3.2 Uplink Multi-User MIMO 224
11.3.3 PUCCH Transmit Diversity 226
11.4 Uplink L1/L2 Control Signaling 226
11.4.1 Uplink L1/L2 Control Signaling on PUCCH 227
11.4.2 Uplink L1/L2 Control Signaling on PUSCH 242
11.5 Uplink Timing Alignment 245

CHAPTER 12 Retransmission Protocols 247
12.1 Hybrid ARQ with Soft Combining 248
12.1.1 Downlink Hybrid ARQ 251
12.1.2 Uplink Hybrid ARQ 251
12.1.3 Hybrid-ARQ Timing 255
12.2 Radio-Link Control 259
12.2.1 Segmentation, Concatenation, and Reassembly of RLC SDUs 260
12.2.2 RLC Retransmission 261
12.2.3 In-Sequence Delivery 262
12.2.4 RLC Operation 262

CHAPTER 13 Power Control, Scheduling, and Interference Handling 265
13.1 Uplink Power Control 265
13.1.1 Uplink Power Control – Some Basic Rules 266
13.1.2 Power Control for PUCCH 267
13.1.3 Power Control for PUSCH 269
13.1.4 Power Control for SRS 271
13.1.5 Power Headroom 271
13.2 Scheduling and Rate Adaptation 272
13.2.1 Downlink Scheduling 273
13.2.2 Uplink Scheduling 275
13.2.3 Semi-Persistent Scheduling 280
13.2.4 Scheduling for Half-Duplex FDD 281
13.2.5 Channel-State Reporting 282
13.2.6 Discontinuous Reception (DRX) and Component Carrier Deactivation 287
13.3 Inter-Cell Interference Coordination 290
13.4 Heterogeneous Network Deployments 293
13.4.1 Interference Handling in a Heterogeneous Deployment 294
13.4.2 Interference Coordination in the Case of Home-eNodeB 299

CHAPTER 14 Access Procedures 301
14.1 Acquisition and Cell Search 301
14.1.1 Overview of LTE Cell Search 301
14.1.2 PSS Structure 303
14.1.3 SSS Structure 304
14.2 System Information 304
14.2.1 MIB and BCH Transmission 305
14.2.2 System-Information Blocks 308

14.3 Random Access 310
14.3.1 Step 1: Random-Access Preamble Transmission 312
14.3.2 Step 2: Random-Access Response 317
14.3.3 Step 3: Terminal Identification 318
14.3.4 Step 4: Contention Resolution 319
14.4 Paging 319

CHAPTER 15 Multimedia Broadcast/Multicast Services 323
15.1 Architecture 324
15.2 Overall Channel Structure and Physical-Layer Processing 325
15.3 Scheduling of MBMS Services 328

CHAPTER 16 Relaying 331
16.1 Relays in LTE 332
16.2 Overall Architecture 333
16.3 Backhaul Design for Inband Relaying 334
16.3.1 Access-Link Hybrid-ARQ Operation 335
16.3.2 Backhaul-Link Hybrid-ARQ Operation 336
16.3.3 Backhaul Downlink Control Signaling 337
16.3.4 Reference Signals for the Backhaul Link 341
16.3.5 Backhaul–Access Link Timing 342

CHAPTER 17 Spectrum and RF Characteristics 347
17.1 Spectrum for LTE 347
17.1.1 Spectrum Defined for IMT Systems by the ITU-R 347
17.1.2 Frequency Bands for LTE 348
17.1.3 New Frequency Bands 351
17.2 Flexible Spectrum Use 352
17.3 Flexible Channel Bandwidth Operation 353
17.4 Carrier Aggregation for LTE 355
17.5 Multi-Standard Radio Base Stations 359
17.6 Overview of RF Requirements for LTE 361
17.6.1 Transmitter Characteristics 362
17.6.2 Receiver Characteristics 363
17.6.3 Regional Requirements 363
17.6.4 Band-Specific UE Requirements Through Network Signaling 364
17.6.5 Base-Station Classes 365
17.7 Output Power Level Requirements 366
17.7.1 Base-Station Output Power and Dynamic Range 366
17.7.2 UE Output Power and Dynamic Range 366

17.8 Transmitted Signal Quality 366
17.8.1 EVM and Frequency Error 367
17.8.2 UE In-Band Emissions 367
17.8.3 Base-Station Time Alignment 367
17.9 Unwanted Emissions Requirements 367
17.9.1 Implementation Aspects 368
17.9.2 Spectrum Emission Mask 368
17.9.3 Adjacent Channel Leakage Ratio 370
17.9.4 Spurious Emissions 372
17.9.5 Occupied Bandwidth 372
17.9.6 Transmitter Intermodulation 372
17.10 Sensitivity and Dynamic Range 373
17.11 Receiver Susceptibility to Interfering Signals 373

CHAPTER 18 Performance 377
18.1 Performance Assessment 377
18.1.1 End-User Perspective of Performance 378
18.1.2 Operator Perspective 379
18.2 Performance in Terms of Peak Data Rates and Latency 379
18.3 Performance Evaluation of LTE-Advanced 380
18.3.1 Models and Assumptions 381
18.3.2 Evaluation Criteria 383
18.3.3 Performance Numbers for FDD 384
18.3.4 Performance Numbers for TDD 385
18.4 Conclusion 387

CHAPTER 19 Other Wireless Communications Systems 389
19.1 HSPA 389
19.1.1 Architecture 391
19.1.2 Channel-Dependent Scheduling 392
19.1.3 Hybrid ARQ with Soft Combining 393
19.1.4 Control-Plane Latency Reductions 394
19.1.5 Spatial Multiplexing 395
19.1.6 Carrier Aggregation 395
19.1.7 UTRA TDD 395
19.2 GSM/EDGE 396
19.2.1 Objectives for GSM/EDGE Evolution 397
19.2.2 Dual-Antenna Terminals 398
19.2.3 Multi-Carrier EDGE 398
19.2.4 Reduced TTI and Fast Feedback 399
19.2.5 Improved Modulation and Coding 400

19.2.6 Higher Symbol Rates 400
19.2.7 Voice Service over Adaptive Multi-User Channels 401
19.3 CDMA2000 and HRPD/1x EV-DO 402
19.3.1 CDMA2000 1x 403
19.3.2 1x EV-DO Rev 0 403
19.3.3 1x EV-DO Rev A 404
19.3.4 1x EV-DO Rev B 405
19.3.5 1x EV-DO Rev C 405
19.4 IEEE 802.16e, Mobile WiMAX and 802.16 m 405
19.4.1 IEEE 802.16e and Mobile WiMAX 406
19.4.2 IEEE 802.16 m – WiMAX for IMT-Advanced 408
19.5 Summary 409

CHAPTER 20 Final Thoughts 411
20.1 Where to Go in the Future? 411
20.1.1 Advanced Multi-Cell Coordination 412
20.1.2 Network Energy Efficiency 412
20.1.3 Machine-Type Communication 413
20.1.4 New ways of Using Spectrum 414
20.1.5 Direct Device-to-Device Communication 414
20.2 Concluding Remarks 415

References 417
Index 423

Preface

During the past years, there has been a quickly rising interest in radio access technologies for providing mobile as well as nomadic and fixed services for voice, video, and data. The difference in design, implementation, and use between telecom and datacom technologies is also becoming more blurred. One example is cellular technologies from the telecom world being used for broadband data and wireless LAN from the datacom world being used for voice-over IP.

Today, the most widespread radio access technology for mobile communication is digital cellular, with the number of users passing 5 billion by 2010, which is more than half of the world's population. It has emerged from early deployments of an expensive voice service for a few car-borne users, to today's widespread use of mobile-communication devices that provide a range of mobile services and often include camera, MP3 player, and PDA functions. With this widespread use and increasing interest in mobile communication, a continuing evolution ahead is foreseen.

This book describes LTE, developed in 3GPP (*Third Generation Partnership Project*) and providing true 4G broadband mobile access, starting from the first version in release 8 and through the continuing evolution to release 10, the latest version of LTE. Release 10, also known as LTE-Advanced, is of particular interest as it is the major technology approved by the ITU as fulfilling the IMT-Advanced requirements. The description in this book is based on LTE release 10 and thus provides a complete description of the LTE-Advanced radio access from the bottom up.

Chapter 1 gives the background to LTE and its evolution, looking also at the different standards bodies and organizations involved in the process of defining 4G. It also gives a discussion of the reasons and driving forces behind the evolution.

Chapters 2–6 provide a deeper insight into some of the technologies that are part of LTE and its evolution. Because of its generic nature, these chapters can be used as a background not only for LTE as described in this book, but also for readers who want to understand the technology behind other systems, such as WCDMA/HSPA, WiMAX, and CDMA2000.

Chapters 7–17 constitute the main part of the book. As a start, an introductory technical overview of LTE is given, where the most important technology components are introduced based on the generic technologies described in previous chapters. The following chapters provide a detailed description of the protocol structure, the downlink and uplink transmission schemes, and the associated mechanisms for scheduling, retransmission and interference handling. Broadcast operation and relaying are also described. This is followed by a discussion of the spectrum flexibility and the associated requirements from an RF perspective.

Finally, in Chapters 18–20, an assessment is made on LTE. Through an overview of similar technologies developed in other standards bodies, it will be clear that the technologies adopted for the evolution in 3GPP are implemented in many other systems as well. Finally, looking into the future, it will be seen that the evolution does not stop with LTE-Advanced but that new features are continuously added to LTE in order to meet future requirements.

Samsung Exhibit 1019, Page 13 of 195

Acknowledgements

We thank all our colleagues at Ericsson for assisting in this project by helping with contributions to the book, giving suggestions and comments on the contents, and taking part in the huge team effort of developing LTE.

The standardization process involves people from all parts of the world, and we acknowledge the efforts of our colleagues in the wireless industry in general and in 3GPP RAN in particular. Without their work and contributions to the standardization, this book would not have been possible.

Finally, we are immensely grateful to our families for bearing with us and supporting us during the long process of writing this book.

Samsung Exhibit 1019, Page 14 of 195

Abbreviations and Acronyms

3GPP	Third Generation Partnership Project
3GPP2	Third Generation Partnership Project 2
ACIR	Adjacent Channel Interference Ratio
ACK	Acknowledgement (in ARQ protocols)
ACLR	Adjacent Channel Leakage Ratio
ACS	Adjacent Channel Selectivity
AM	Acknowledged Mode (RLC configuration)
AMC	Adaptive Modulation and Coding
A-MPR	Additional Maximum Power Reduction
AMPS	Advanced Mobile Phone System
AQPSK	Adaptive QPSK
ARI	Acknowledgement Resource Indicator
ARIB	Association of Radio Industries and Businesses
ARQ	Automatic Repeat-reQuest
AS	Access Stratum
ATIS	Alliance for Telecommunications Industry Solutions
AWGN	Additive White Gaussian Noise
BC	Band Category
BCCH	Broadcast Control Channel
BCH	Broadcast Channel
BER	Bit-Error Rate
BLER	Block-Error Rate
BM-SC	Broadcast Multicast Service Center
BPSK	Binary Phase-Shift Keying
BS	Base Station
BSC	Base Station Controller
BTS	Base Transceiver Station
CA	Carrier Aggregation
CC	Convolutional Code (in the context of coding), or Component Carrier (in the context of carrier aggregation)
CCCH	Common Control Channel
CCE	Control Channel Element
CCSA	China Communications Standards Association
CDD	Cyclic-Delay Diversity
CDF	Cumulative Density Function
CDM	Code-Division Multiplexing
CDMA	Code-Division Multiple Access

Samsung Exhibit 1019, Page 15 of 195

CEPT	European Conference of Postal and Telecommunications Administrations
CN	Core Network
CoMP	Coordinated Multi-Point transmission/reception
CP	Cyclic Prefix
CPC	Continuous Packet Connectivity
CQI	Channel-Quality Indicator
C-RAN	Centralized RAN
CRC	Cyclic Redundancy Check
C-RNTI	Cell Radio-Network Temporary Identifier
CRS	Cell-specific Reference Signal
CS	Circuit Switched (or Cyclic Shift)
CS	Capability Set (for MSR base stations)
CSA	Common Subframe Allocation
CSG	Closed Subscriber Group
CSI	Channel-State Information
CSI-RS	CSI reference signals
CW	Continuous Wave
DAI	Downlink Assignment Index
DCCH	Dedicated Control Channel
DCH	Dedicated Channel
DCI	Downlink Control Information
DFE	Decision-Feedback Equalization
DFT	Discrete Fourier Transform
DFTS-OFDM	DFT-Spread OFDM (DFT-precoded OFDM, see also SC-FDMA)
DL	Downlink
DL-SCH	Downlink Shared Channel
DM-RS	Demodulation Reference Signal
DRX	Discontinuous Reception
DTCH	Dedicated Traffic Channel
DTX	Discontinuous Transmission
DwPTS	The downlink part of the special subframe (for TDD operation).
EDGE	Enhanced Data rates for GSM Evolution, Enhanced Data rates for Global Evolution
EGPRS	Enhanced GPRS
eNB	eNodeB
eNodeB	E-UTRAN NodeB
EPC	Evolved Packet Core
EPS	Evolved Packet System
ETSI	European Telecommunications Standards Institute
E-UTRA	Evolved UTRA
E-UTRAN	Evolved UTRAN
EV-DO	Evolution-Data Only (of CDMA2000 1x)
EV-DV	Evolution-Data and Voice (of CDMA2000 1x)
EVM	Error Vector Magnitude

FACH	Forward Access Channel
FCC	Federal Communications Commission
FDD	Frequency Division Duplex
FDM	Frequency-Division Multiplex
FDMA	Frequency-Division Multiple Access
FEC	Forward Error Correction
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FPLMTS	Future Public Land Mobile Telecommunications Systems
FRAMES	Future Radio Wideband Multiple Access Systems
FSTD	Frequency Switched Transmit Diversity
GERAN	GSM/EDGE Radio Access Network
GGSN	Gateway GPRS Support Node
GP	Guard Period (for TDD operation)
GPRS	General Packet Radio Services
GPS	Global Positioning System
GSM	Global System for Mobile communications
HARQ	Hybrid ARQ
HII	High-Interference Indicator
HLR	Home Location Register
HRPD	High Rate Packet Data
HSDPA	High-Speed Downlink Packet Access
HSPA	High-Speed Packet Access
HSS	Home Subscriber Server
HS-SCCH	High-Speed Shared Control Channel
ICIC	Inter-Cell Interference Coordination
ICS	In-Channel Selectivity
ICT	Information and Communication Technologies
IDFT	Inverse DFT
IEEE	Institute of Electrical and Electronics Engineers
IFDMA	Interleaved FDMA
IFFT	Inverse Fast Fourier Transform
IMT-2000	International Mobile Telecommunications 2000 (ITU's name for the family of 3G standards)
IMT-Advanced	International Mobile Telecommunications Advanced (ITU's name for the family of 4G standards)
IP	Internet Protocol
IR	Incremental Redundancy
IRC	Interference Rejection Combining
ITU	International Telecommunications Union
ITU-R	International Telecommunications Union-Radiocommunications Sector

J-TACS	Japanese Total Access Communication System
LAN	Local Area Network
LCID	Logical Channel Index
LDPC	Low-Density Parity Check Code
LTE	Long-Term Evolution
MAC	Medium Access Control
MAN	Metropolitan Area Network
MBMS	Multimedia Broadcast/Multicast Service
MBMS-GW	MBMS gateway
MBS	Multicast and Broadcast Service
MBSFN	Multicast-Broadcast Single Frequency Network
MC	Multi-Carrier
MCCH	MBMS Control Channel
MCE	MBMS Coordination Entity
MCH	Multicast Channel
MCS	Modulation and Coding Scheme
MDHO	Macro-Diversity Handover
MIB	Master Information Block
MIMO	Multiple-Input Multiple-Output
ML	Maximum Likelihood
MLSE	Maximum-Likelihood Sequence Estimation
MME	Mobility Management Entity
MMS	Multimedia Messaging Service
MMSE	Minimum Mean Square Error
MPR	Maximum Power Reduction
MRC	Maximum Ratio Combining
MSA	MCH Subframe Allocation
MSC	Mobile Switching Center
MSI	MCH Scheduling Information
MSP	MCH Scheduling Period
MSR	Multi-Standard Radio
MSS	Mobile Satellite Service
MTCH	MBMS Traffic Channel
MU-MIMO	Multi-User MIMO
MUX	Multiplexer or Multiplexing
NAK, NACK	Negative Acknowledgement (in ARQ protocols)
NAS	Non-Access Stratum (a functional layer between the core network and the terminal that supports signaling and user data transfer)
NDI	New-data indicator
NSPS	National Security and Public Safety
NMT	Nordisk MobilTelefon (Nordic Mobile Telephony)

NodeB	NodeB, a logical node handling transmission/reception in multiple cells. Commonly, but not necessarily, corresponding to a base station.
NS	Network Signaling
OCC	Orthogonal Cover Code
OFDM	Orthogonal Frequency-Division Multiplexing
OFDMA	Orthogonal Frequency-Division Multiple Access
OI	Overload Indicator
OOB	Out-Of-Band (emissions)
PAPR	Peak-to-Average Power Ratio
PAR	Peak-to-Average Ratio (same as PAPR)
PARC	Per-Antenna Rate Control
PBCH	Physical Broadcast Channel
PCCH	Paging Control Channel
PCFICH	Physical Control Format Indicator Channel
PCG	Project Coordination Group (in 3GPP)
PCH	Paging Channel
PCRF	Policy and Charging Rules Function
PCS	Personal Communications Systems
PDA	Personal Digital Assistant
PDC	Personal Digital Cellular
PDCCH	Physical Downlink Control Channel
PDCP	Packet Data Convergence Protocol
PDSCH	Physical Downlink Shared Channel
PDN	Packet Data Network
PDU	Protocol Data Unit
PF	Proportional Fair (a type of scheduler)
P-GW	Packet-Data Network Gateway (also PDN-GW)
PHICH	Physical Hybrid-ARQ Indicator Channel
PHS	Personal Handy-phone System
PHY	Physical layer
PMCH	Physical Multicast Channel
PMI	Precoding-Matrix Indicator
POTS	Plain Old Telephony Services
PRACH	Physical Random Access Channel
PRB	Physical Resource Block
P-RNTI	Paging RNTI
PS	Packet Switched
PSK	Phase Shift Keying
PSS	Primary Synchronization Signal
PSTN	Public Switched Telephone Networks
PUCCH	Physical Uplink Control Channel

PUSC	Partially Used Subcarriers (for WiMAX)
PUSCH	Physical Uplink Shared Channel
QAM	Quadrature Amplitude Modulation
QoS	Quality-of-Service
QPP	Quadrature Permutation Polynomial
QPSK	Quadrature Phase-Shift Keying
RAB	Radio Access Bearer
RACH	Random Access Channel
RAN	Radio Access Network
RA-RNTI	Random Access RNTI
RAT	Radio Access Technology
RB	Resource Block
RE	Reseource Element
RF	Radio Frequency
RI	Rank Indicator
RIT	Radio Interface Technology
RLC	Radio Link Control
RNC	Radio Network Controller
RNTI	Radio-Network Temporary Identifier
RNTP	Relative Narrowband Transmit Power
ROHC	Robust Header Compression
R-PDCCH	Relay Physical Downlink Control Channel
RR	Round-Robin (a type of scheduler)
RRC	Radio Resource Control
RRM	Radio Resource Management
RS	Reference Symbol
RSPC	IMT-2000 radio interface specifications
RSRP	Reference Signal Received Power
RSRQ	Reference Signal Received Quality
RTP	Real Time Protocol
RTT	Round-Trip Time
RV	Redundancy Version
RX	Receiver
S1	The interface between eNodeB and the Evolved Packet Core.
S1-c	The control-plane part of S1
S1-u	The user-plane part of S1
SAE	System Architecture Evolution
SCM	Spatial Channel Model
SDMA	Spatial Division Multiple Access
SDO	Standards Developing Organization
SDU	Service Data Unit
SEM	Spectrum Emissions Mask

SF	Spreading Factor
SFBC	Space-Frequency Block Coding
SFN	Single-Frequency Network (in general, see also MBSFN) or System Frame Number (in 3GPP)
SFTD	Space–Frequency Time Diversity
SGSN	Serving GPRS Support Node
S-GW	Serving Gateway
SI	System Information message
SIB	System Information Block
SIC	Successive Interference Combining
SIM	Subscriber Identity Module
SINR	Signal-to-Interference-and-Noise Ratio
SIR	Signal-to-Interference Ratio
SI-RNTI	System Information RNTI
SMS	Short Message Service
SNR	Signal-to-Noise Ratio
SOHO	Soft Handover
SORTD	Spatial Orthogonal-Resource Transmit Diversity
SR	Scheduling Request
SRS	Sounding Reference Signal
SSS	Secondary Synchronization Signal
STBC	Space–Time Block Coding
STC	Space–Time Coding
STTD	Space-Time Transmit Diversity
SU-MIMO	Single-User MIMO
TACS	Total Access Communication System
TCP	Transmission Control Protocol
TC-RNTI	Temporary C-RNTI
TD-CDMA	Time-Division Code-Division Multiple Access
TDD	Time-Division Duplex
TDM	Time-Division Multiplexing
TDMA	Time-Division Multiple Access
TD-SCDMA	Time-Division-Synchronous Code-Division Multiple Access
TF	Transport Format
TIA	Telecommunications Industry Association
TM	Transparent Mode (RLC configuration)
TR	Technical Report
TS	Technical Specification
TSG	Technical Specification Group
TTA	Telecommunications Technology Association
TTC	Telecommunications Technology Committee
TTI	Transmission Time Interval
TX	Transmitter

UCI	Uplink Control Information
UE	User Equipment, the 3GPP name for the mobile terminal
UL	Uplink
UL-SCH	Uplink Shared Channel
UM	Unacknowledged Mode (RLC configuration)
UMB	Ultra Mobile Broadband
UMTS	Universal Mobile Telecommunications System
UpPTS	The uplink part of the special subframe (for TDD operation).
US-TDMA	US Time-Division Multiple Access standard
UTRA	Universal Terrestrial Radio Access
UTRAN	Universal Terrestrial Radio Access Network
VAMOS	Voice services over Adaptive Multi-user channels
VoIP	Voice-over-IP
VRB	Virtual Resource Block
WAN	Wide Area Network
WARC	World Administrative Radio Congress
WCDMA	Wideband Code-Division Multiple Access
WG	Working Group
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WMAN	Wireless Metropolitan Area Network
WP5D	Working Party 5D
WRC	World Radiocommunication Conference
X2	The interface between eNodeBs.
ZC	Zadoff-Chu
ZF	Zero Forcing

CHAPTER

Background of LTE

1

1.1 INTRODUCTION

Mobile communications has become an everyday commodity. In the last decades, it has evolved from being an expensive technology for a few selected individuals to today's ubiquitous systems used by a majority of the world's population. From the first experiments with radio communication by Guglielmo Marconi in the 1890s, the road to truly mobile radio communication has been quite long. To understand the complex mobile-communication systems of today, it is important to understand where they came from and how cellular systems have evolved. The task of developing mobile technologies has also changed, from being a national or regional concern, to becoming an increasingly complex task undertaken by global standards-developing organizations such as the *Third Generation Partnership Project* (3GPP) and involving thousands of people.

Mobile communication technologies are often divided into generations, with 1G being the analog mobile radio systems of the 1980s, 2G the first digital mobile systems, and 3G the first mobile systems handling broadband data. The *Long-Term Evolution* (LTE) is often called "4G", but many also claim that LTE release 10, also referred to as *LTE-Advanced*, is the true 4G evolution step, with the first release of LTE (release 8) then being labeled as "3.9G". This continuing race of increasing sequence numbers of mobile system generations is in fact just a matter of labels. What is important is the actual system capabilities and how they have evolved, which is the topic of this chapter.

In this context, it must first be pointed out that LTE and LTE-Advanced is the same technology, with the "Advanced" label primarily being added to highlight the relation between LTE release 10 (LTE-Advanced) and ITU/IMT-Advanced, as discussed later. This does not make LTE-Advanced a different system than LTE and it is not in any way the final evolution step to be taken for LTE. Another important aspect is that the work on developing LTE and LTE-Advanced is performed as a continuing task within 3GPP, the same forum that developed the first 3G system (WCDMA/HSPA).

This chapter describes the background for the development of the LTE system, in terms of events, activities, organizations and other factors that have played an important role. First, the technologies and mobile systems leading up to the starting point for 3G mobile systems will be discussed. Next, international activities in the ITU that were part of shaping 3G and the 3G evolution and the market and technology drivers behind LTE will be discussed. The final part of the chapter describes the standardization process that provided the detailed specification work leading to the LTE systems deployed and in operation today.

4G LTE/LTE-Advanced for Mobile Broadband.
© 2011 Erik Dahlman, Stefan Parkvall & Johan Sköld. Published by Elsevier Ltd. All rights reserved.

Samsung Exhibit 1019, Page 23 of 195

1.2 EVOLUTION OF MOBILE SYSTEMS BEFORE LTE

The US *Federal Communications Commission* (FCC) approved the first commercial car-borne telephony service in 1946, operated by AT&T. In 1947 AT&T also introduced the cellular concept of reusing radio frequencies, which became fundamental to all subsequent mobile-communication systems. Similar systems were operated by several monopoly telephone administrations and wire-line operators during the 1950s and 1960s, using bulky and power-hungry equipment and providing car-borne services for a very limited number of users.

The big uptake of subscribers and usage came when mobile communication became an international concern involving several interested parties, in the beginning mainly the operators. The first international mobile communication systems were started in the early 1980s; the best-known ones are NMT that was started up in the Nordic countries, AMPS in the USA, TACS in Europe, and J-TACS in Japan. Equipment was still bulky, mainly car-borne, and voice quality was often inconsistent, with "cross-talk" between users being a common problem. With NMT came the concept of "roaming", giving a service also for users traveling outside the area of their "home" operator. This also gave a larger market for mobile phones, attracting more companies into the mobile-communication business.

The analog first-generation cellular systems supported "plain old telephony services" (POTS) – that is, voice with some related supplementary services. With the advent of digital communication during the 1980s, the opportunity to develop a second generation of mobile-communication standards and systems, based on digital technology, surfaced. With digital technology came an opportunity to increase the capacity of the systems, to give a more consistent quality of the service, and to develop much more attractive and truly mobile devices.

In Europe, the GSM (originally *Groupe Spécial Mobile*, later *Global System for Mobile communications*) project to develop a pan-European mobile-telephony system was initiated in the mid 1980s by the telecommunication administrations in CEPT[1] and later continued within the new *European Telecommunication Standards Institute* (ETSI). The GSM standard was based on *Time-Division Multiple Access* (TDMA), as were the US-TDMA standard and the Japanese PDC standard that were introduced in the same time frame. A somewhat later development of a *Code-Division Multiple Access* (CDMA) standard called IS-95 was completed in the USA in 1993.

All these standards were "narrowband" in the sense that they targeted "low-bandwidth" services such as voice. With the second-generation digital mobile communications came also the opportunity to provide data services over the mobile-communication networks. The primary data services introduced in 2G were text messaging (*Short Message Services*, SMS) and circuit-switched data services enabling e-mail and other data applications, initially at a modest peak data rate of 9.6 kbit/s. Higher data rates were introduced later in evolved 2G systems by assigning multiple time slots to a user and through modified coding schemes.

Packet data over cellular systems became a reality during the second half of the 1990s, with *General Packet Radio Services* (GPRS) introduced in GSM and packet data also added to other cellular technologies such as the Japanese PDC standard. These technologies are often referred to as 2.5G. The success of the wireless data service iMode in Japan, which included a complete "ecosystem"

[1] The European Conference of Postal and Telecommunications Administrations (CEPT) consists of the telecom administrations from 48 countries.

for service delivery, charging etc., gave a very clear indication of the potential for applications over packet data in mobile systems, in spite of the fairly low data rates supported at the time.

With the advent of 3G and the higher-bandwidth radio interface of UTRA (*Universal Terrestrial Radio Access*) came possibilities for a range of new services that were only hinted at with 2G and 2.5G. The 3G radio access development is today handled in 3GPP. However, the initial steps for 3G were taken in the early 1990s, long before 3GPP was formed.

What also set the stage for 3G was the internationalization of cellular standards. GSM was a pan-European project, but it quickly attracted worldwide interest when the GSM standard was deployed in a number of countries outside Europe. A global standard gains in economy of scale, since the market for products becomes larger. This has driven a much tighter international cooperation around 3G cellular technologies than for the earlier generations.

1.2.1 The First 3G Standardization

Work on a third-generation mobile communication started in ITU (*International Telecommunication Union*) in the 1980s, first under the label *Future Public Land Mobile Telecommunications Systems* (FPLMTS), later changed to IMT-2000 [1]. The World Administrative Radio Congress WARC-92 identified 230 MHz of spectrum for IMT-2000 on a worldwide basis. Of these 230 MHz, 2×60 MHz was identified as paired spectrum for FDD (*Frequency-Division Duplex*) and 35 MHz as unpaired spectrum for TDD (*Time-Division Duplex*), both for terrestrial use. Some spectrum was also set aside for satellite services. With that, the stage was set to specify IMT-2000.

In parallel with the widespread deployment and evolution of 2G mobile-communication systems during the 1990s, substantial efforts were put into 3G research activities worldwide. In Europe, a number of partially EU-funded projects resulted in a multiple access concept that included a Wideband CDMA component that was input to ETSI in 1996. In Japan, the *Association of Radio Industries and Businesses* (ARIB) was at the same time defining a 3G wireless communication technology based on Wideband CDMA and also in the USA a Wideband CDMA concept called WIMS was developed within the T1.P1[2] committee. South Korea also started work on Wideband CDMA at this time.

When the standardization activities for 3G started in ETSI in 1996, there were WCDMA concepts proposed both from a European research project (FRAMES) and from the ARIB standardization in Japan. The Wideband CDMA proposals from Europe and Japan were merged and came out as part of the winning concept in early 1998 in the European work on Universal Mobile Telecommunication Services (UMTS), which was the European name for 3G. Standardization of WCDMA continued in parallel in several standards groups until the end of 1998, when the *Third Generation Partnership Project* (3GPP) was formed by standards-developing organizations from all regions of the world. This solved the problem of trying to maintain parallel development of aligned specifications in multiple regions. The present organizational partners of 3GPP are ARIB (Japan), CCSA (China), ETSI (Europe), ATIS (USA), TTA (South Korea), and TTC (Japan).

At this time, when the standardization bodies were ready to put the details into the 3GPP specifications, work on 3G mobile systems had already been ongoing for some time in the international arena within the ITU-R. That work was influenced by and also provided a broader international framework for the standardization work in 3GPP.

[2]The T1.P1 committee was part of T1, which presently has joined the ATIS standardization organization.

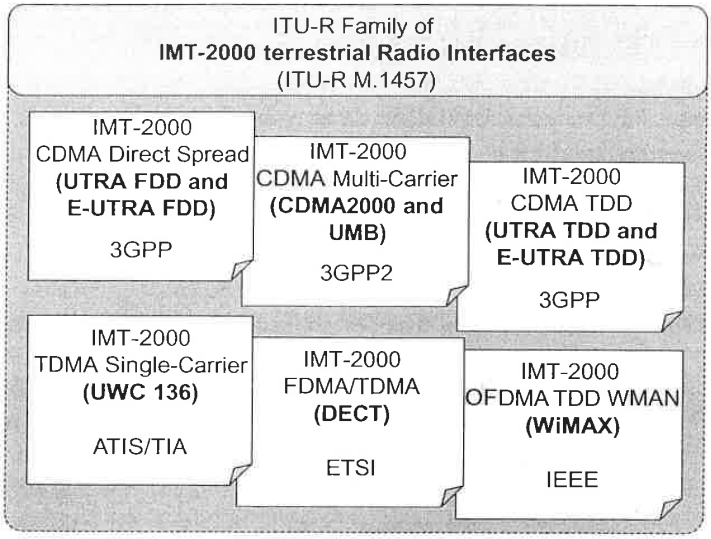


FIGURE 1.1

The definition of IMT-2000 in ITU-R.

1.3 ITU ACTIVITIES

1.3.1 IMT-2000 and IMT-Advanced

ITU-R *Working Party 5D* (WP5D) has the responsibility for IMT systems, which is the umbrella name for 3G (IMT-2000) and 4G (IMT-Advanced). WP5D does not write technical specifications for IMT, but has kept the role of defining IMT in cooperation with the regional standardization bodies and to maintain a set of recommendations for IMT-2000 and IMT-Advanced.

The main IMT-2000 recommendation is ITU-R M.1457 [2], which identifies the IMT-2000 *radio interface specifications* (RSPC). The recommendation contains a "family" of radio interfaces, all included on an equal basis. The family of six terrestrial radio interfaces is illustrated in Figure 1.1, which also shows the *Standards Developing Organizations* (SDO) or Partnership Projects that produce the specifications. In addition, there are several IMT-2000 satellite radio interfaces defined, not illustrated in Figure 1.1.

For each radio interface, M.1457 contains an overview of the radio interface, followed by a list of references to the detailed specifications. The actual specifications are maintained by the individual SDOs and M.1457 provides references to the specifications maintained by each SDO.

With the continuing development of the IMT-2000 radio interfaces, including the evolution of UTRA to Evolved UTRA, the ITU recommendations also need to be updated. ITU-R WP5D continuously revises recommendation M.1457 and at the time of writing it is in its ninth version. Input to the updates is provided by the SDOs and Partnership Projects writing the standards. In the latest revision of ITU-R M.1457, LTE (or E-UTRA) is included in the family through the 3GPP family members for UTRA FDD and TDD, as shown in the figure.

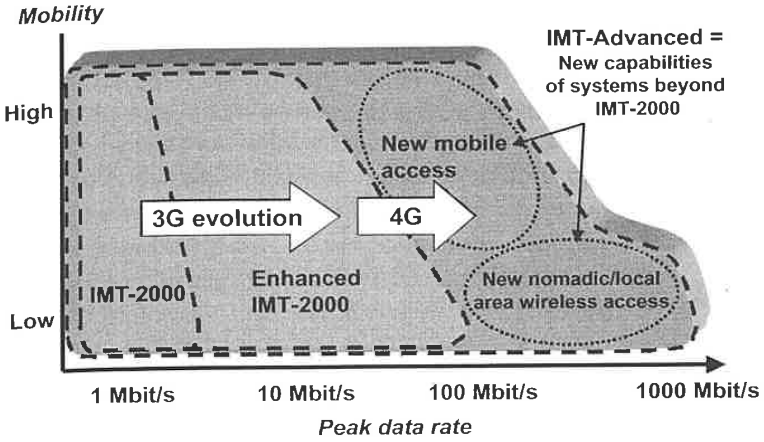


FIGURE 1.2

Illustration of capabilities of IMT-2000 and IMT-Advanced, based on the framework described in ITU -R Recommendation M.1645 [4].

In addition to maintaining the IMT-2000 specifications, a main activity in ITU-R WP5D is the work on systems beyond IMT-2000, now called IMT-Advanced. The term IMT-Advanced is used for systems that include new radio interfaces supporting the new capabilities of systems beyond IMT-2000, as demonstrated with the "van diagram" in Figure 1.2. The step into IMT-Advanced capabilities is seen by ITU-R as the step into 4G, the next generation of mobile technologies after 3G.

The process for defining IMT-Advanced was set by ITU-R WP5D [3] and was quite similar to the process used in developing the IMT-2000 recommendations. ITU-R first concluded studies for IMT-Advanced of services and technologies, market forecasts, principles for standardization, estimation of spectrum needs, and identification of candidate frequency bands [4]. Evaluation criteria were agreed, where proposed technologies were to be evaluated according to a set of minimum technical requirements. All ITU members and other organizations were then invited to the process through a circular letter [5] in March 2008. After submission of six candidate technologies in 2009, an evaluation was performed in cooperation with external bodies such as standards-developing organizations, industry forums, and national groups.

An evolution of LTE as developed by 3GPP was submitted as one candidate to the ITU-R evaluation. While actually being a new release (release 10) of the LTE system and thus an integral part of the continuing LTE development, the candidate was named LTE-Advanced for the purpose of ITU submission. 3GPP also set up its own set of technical requirements for LTE-Advanced, with the ITU-R requirements as a basis. The specifics of LTE-Advanced will be described in more detail as part of the description of LTE later in this book. The performance evaluation of LTE-Advanced for the ITU-R submission is described further in Chapter 18.

The target of the process was always harmonization of the candidates through consensus building. ITU-R determined in October 2010 that two technologies will be included in the first release of IMT-Advanced, those two being LTE release 10 ("LTE-Advanced") and WirelessMAN-Advanced [6] based on the IEEE 802.16m specification. The two can be viewed as the "family" of IMT-Advanced

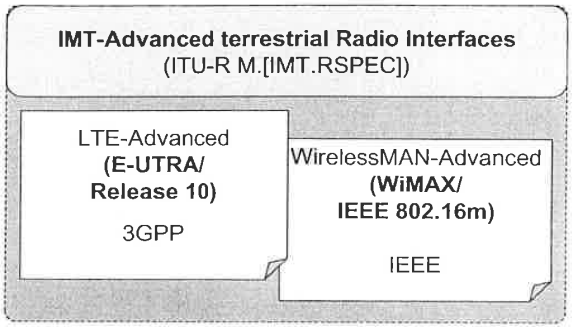


FIGURE 1.3

Radio interface technologies for IMT-Advanced.

technologies as shown in Figure 1.3. The main IMT-Advanced recommendation, identifying the IMT-Advanced radio interface specifications, is presently named ITU-R.[IMT.RSPEC] [7] and will be completed during 2011. As for the corresponding IMT-2000 specification, it will contain an overview of each radio interface, followed by a list of references to the detailed specifications.

1.3.2 Spectrum for IMT Systems

Another major activity within ITU-R concerning IMT-Advanced has been to identify globally available spectrum, suitable for IMT systems. The spectrum work has involved sharing studies between IMT and other technologies in those bands. Adequate spectrum availability and globally harmonized spectrum are identified as essential for IMT-Advanced.

Spectrum for 3G was first identified at the *World Administrative Radio Congress* WARC-92, where 230 MHz was identified as intended for use by national administrations that want to implement IMT-2000. The so-called IMT-2000 "core band" at 2 GHz is in this frequency range and was the first band where 3G systems were deployed.

Additional spectrum was identified for IMT-2000 at later World Radio communication conferences. WRC-2000 identified the existing 2G bands at 800/900 MHz and 1800/1900 MHz plus an additional 190 MHz of spectrum at 2.6 GHz, all for IMT-2000. As additional spectrum for IMT-2000, WRC'07 identified a band at 450 MHz, the so-called "digital dividend" at 698–806 MHz, plus an additional 300 MHz of spectrum at higher frequencies. The applicability of these new bands varies on a regional and national basis.

The worldwide frequency arrangements for IMT-2000 are outlined in ITU-R recommendation M.1036 [8], which is presently being updated with the arrangements for the most recent frequency bands added at WRC'07. The recommendation outlines the regional variations in how the bands are implemented and also identifies which parts of the spectrum are paired and which are unpaired. For the paired spectrum, the bands for uplink (mobile transmit) and downlink (base-station transmit) are identified for *Frequency-Division Duplex* (FDD) operation. The unpaired bands can, for example, be used for *Time-Division Duplex* (TDD) operation. Note that the band that is most globally deployed for 3G is still 2 GHz.

1.4 DRIVERS FOR LTE

The evolution of 3G systems into 4G is driven by the creation and development of new services for mobile devices, and is enabled by advancement of the technology available for mobile systems. There has also been an evolution of the environment in which mobile systems are deployed and operated, in terms of competition between mobile operators, challenges from other mobile technologies, and new regulation of spectrum use and market aspects of mobile systems.

The rapid evolution of the technology used in telecommunication systems, consumer electronics, and specifically mobile devices has been remarkable in the last 20 years. Moore's law illustrates this and indicates a continuing evolution of processor performance and increased memory size, often combined with reduced size, power consumption, and cost for devices. High-resolution color displays and megapixel camera sensors are also coming into all types of mobile devices. Combined with a high-speed internet backbone often based on optical fiber networks, we see that a range of technology enablers are in place to go hand-in-hand with advancement in mobile communications technology such as LTE.

The rapid increase in use of the internet to provide all kinds of services since the 1990s started at the same time as 2G and 3G mobile systems came into widespread use. The natural next step was that those internet-based services also moved to the mobile devices, creating what is today know as *mobile broadband*. Being able to support the same *Internet Protocol* (IP)-based services in a mobile device that people use at home with a fixed broadband connection is a major challenge and a prime driver for the evolution of LTE. A few services were already supported by the evolved 2.5G systems, but it is not until the systems are designed primarily for IP-based services that the real mobile IP revolution can take off. An interesting aspect of the migration of broadband services to mobile devices is that a mobile "flavor" is also added. The mobile position and the mobility and roaming capabilities do in fact create a whole new range of services tailored to the mobile environment.

Fixed telephony (POTS) and earlier generations of mobile technology were built for circuit switched services, primarily voice. The first data services over GSM were circuit switched, with packet-based GPRS coming in as a later addition. This also influenced the first development of 3G, which was based on circuit switched data, with packet-switched services as an add-on. It was not until the 3G evolution into HSPA and later LTE/LTE-Advanced that packet-switched services and IP were made the primary design target. The old circuit-switched services remain, but will on LTE be provided over IP, with Voice-over IP (VoIP) as an example.

IP is in itself service agnostic and thereby enables a range of services with different requirements. The main service-related design parameters for a radio interface supporting a variety of services are:

- **Data rate.** Many services with lower data rates such as voice services are important and still occupy a large part of a mobile network's overall capacity, but it is the higher data rate services that drive the design of the radio interface. The ever increasing demand for higher data rates for web browsing, streaming and file transfer pushes the peak data rates for mobile systems from kbit/s for 2G, to Mbit/s for 3G and getting close to Gbit/s for 4G.
- **Delay.** Interactive services such as real-time gaming, but also web browsing and interactive file transfer, have requirements for very low delay, making it a primary design target. There are, however, many applications such as e-mail and television where the delay requirements are not as strict. The delay for a packet sent from a server to a client and back is called *latency*.

- **Capacity.** From the mobile system operator's point of view, it is not only the peak data rates provided to the end-user that are of importance, but also the total data rate that can be provided on average from each deployed base station site and per hertz of licensed spectrum. This measure of capacity is called *spectral efficiency*. In the case of capacity shortage in a mobile system, the Quality-of-Service (QoS) for the individual end-users may be degraded.

How these three main design parameters influenced the development of LTE is described in more detail in Chapter 7 and an evaluation of what performance is achieved for the design parameters above is presented in Chapter 18.

The demand for new services and for higher peak bit rates and system capacity is not only met by evolution of the technology to 4G. There is also a demand for more spectrum resources to expand systems and the demand also leads to more competition between an increasing number of mobile operators and between alternative technologies to provide mobile broadband services. An overview of some technologies other than LTE is given in Chapter 19.

With more spectrum coming into use for mobile broadband, there is a need to operate mobile systems in a number of different frequency bands, in spectrum allocations of different sizes and sometimes also in fragmented spectrum. This calls for high spectrum flexibility with the possibility for a varying channel bandwidth, which was also a driver and an essential design parameter for LTE.

The demand for new mobile services and the evolution of the radio interface to LTE have served as drivers to evolve the core network. The core network developed for GSM in the 1980s was extended to support GPRS, EDGE, and WCDMA in the 1990s, but was still very much built around the circuit-switched domain. A *System Architecture Evolution* (SAE) was initiated at the same time as LTE development started and has resulted in an *Evolved Packet Core* (EPC), developed to support HSPA and LTE/LTE-Advanced, focusing on the packet-switched domain. For more details on SAE/EPC, please refer to [9].

1.5 STANDARDIZATION OF LTE

With a framework for IMT systems set up by the ITU-R, with spectrum made available by the WRC and with an ever increasing demand for better performance, the task of specifying the LTE system that meets the design targets falls on 3GPP. 3GPP writes specifications for 2G, 3G, and 4G mobile systems, and 3GPP technologies are the most widely deployed in the world, with more than 4.5 billion connections in 2010. In order to understand how 3GPP works, it is important to understand the process of writing standards.

1.5.1 The Standardization Process

Setting a standard for mobile communication is not a one-time job, it is an ongoing process. The standardization forums are constantly evolving their standards trying to meet new demands for services and features. The standardization process is different in the different forums, but typically includes the four phases illustrated in Figure 1.4:

1. *Requirements*, where it is decided what is to be achieved by the standard.
2. *Architecture*, where the main building blocks and interfaces are decided.

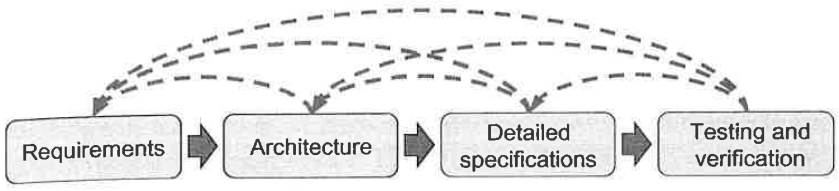


FIGURE 1.4

The standardization phases and iterative process.

3. *Detailed specifications*, where every interface is specified in detail.
4. *Testing and verification*, where the interface specifications are proven to work with real-life equipment.

These phases are overlapping and iterative. As an example, requirements can be added, changed, or dropped during the later phases if the technical solutions call for it. Likewise, the technical solution in the detailed specifications can change due to problems found in the testing and verification phase.

Standardization starts with the *requirements* phase, where the standards body decides what should be achieved with the standard. This phase is usually relatively short.

In the *architecture* phase, the standards body decides about the architecture – that is, the principles of how to meet the requirements. The architecture phase includes decisions about reference points and interfaces to be standardized. This phase is usually quite long and may change the requirements.

After the architecture phase, the *detailed specification* phase starts. It is in this phase the details for each of the identified interfaces are specified. During the detailed specification of the interfaces, the standards body may find that previous decisions in the architecture or even in the requirements phases need to be revisited.

Finally, the *testing and verification* phase starts. It is usually not a part of the actual standardization in the standards bodies, but takes place in parallel through testing by vendors and interoperability testing between vendors. This phase is the final proof of the standard. During the testing and verification phase, errors in the standard may still be found and those errors may change decisions in the detailed standard. Albeit not common, changes may also need to be made to the architecture or the requirements. To verify the standard, products are needed. Hence, the implementation of the products starts after (or during) the detailed specification phase. The testing and verification phase ends when there are stable test specifications that can be used to verify that the equipment is fulfilling the standard.

Normally, it takes one to two years from the time when the standard is completed until commercial products are out on the market. However, if the standard is built from scratch, it may take longer since there are no stable components to build from.

1.5.2 The 3GPP Process

The *Third-Generation Partnership Project* (3GPP) is the standards-developing body that specifies the LTE/LTE-Advanced, as well as 3G UTRA and 2G GSM systems. 3GPP is a partnership project formed by the standards bodies ETSI, ARIB, TTC, TTA, CCSA, and ATIS. 3GPP consists of four *Technical Specifications Groups* (TSGs) – see Figure 1.5.

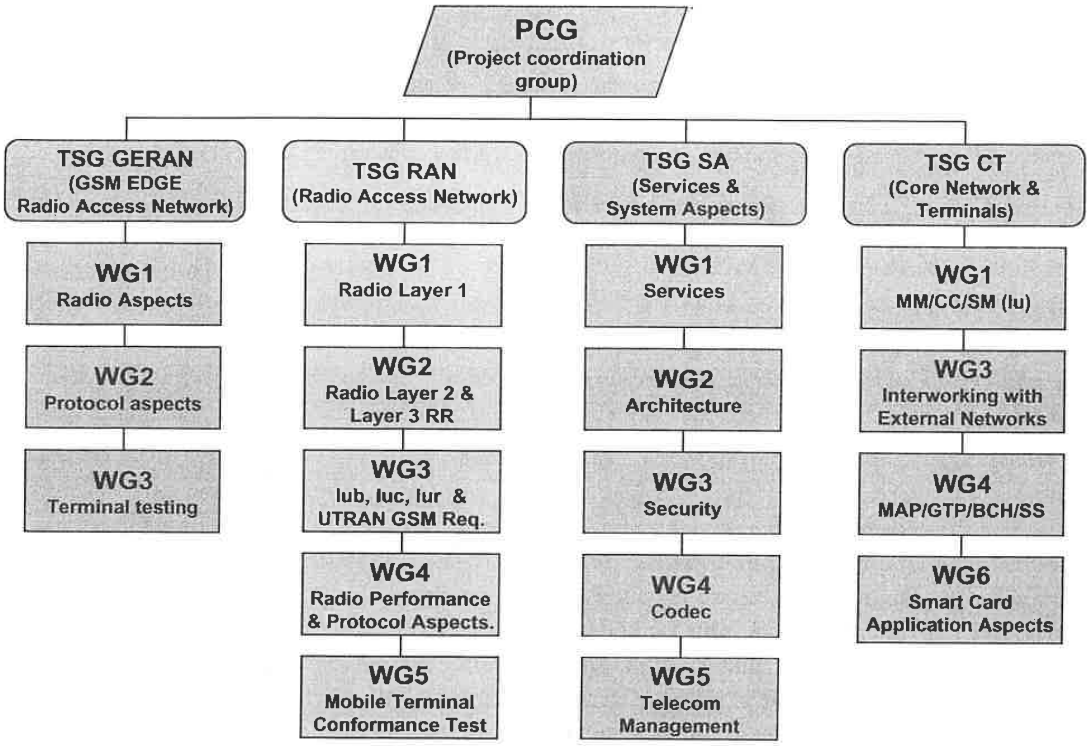


FIGURE 1.5

3GPP organization.

A parallel partnership project called 3GPP2 was formed in 1999. It also develops 3G specifications, but for CDMA2000, which is the 3G technology developed from the 2G CDMA-based standard IS-95. It is also a global project, and the organizational partners are ARIB, CCSA, TIA, TTA, and TTC.

3GPP TSG RAN (*Radio Access Network*) is the technical specification group that has developed WCDMA, its evolution HSPA, as well as LTE/LTE-Advanced, and is in the forefront of the technology. TSG RAN consists of five working groups (WGs):

1. RAN WG1, dealing with the physical layer specifications.
2. RAN WG2, dealing with the layer 2 and layer 3 radio interface specifications.
3. RAN WG3, dealing with the fixed RAN interfaces – for example, interfaces between nodes in the RAN – but also the interface between the RAN and the core network.
4. RAN WG4, dealing with the *radio frequency* (RF) and *radio resource management* (RRM) performance requirements.
5. RAN WG 5, dealing with the terminal conformance testing.

The scope of 3GPP when it was formed in 1998 was to produce global specifications for a 3G mobile system based on an evolved GSM core network, including the WCDMA-based radio access of

the UTRA FDD and the TD-CDMA-based radio access of the UTRA TDD mode. The task to maintain and develop the GSM/EDGE specifications was added to 3GPP at a later stage and the work now also includes LTE (E-UTRA). The UTRA, E-UTRA and GSM/EDGE specifications are developed, maintained, and approved in 3GPP. After approval, the organizational partners transpose them into appropriate deliverables as standards in each region.

In parallel with the initial 3GPP work, a 3G system based on TD-SCDMA was developed in China. TD-SCDMA was eventually merged into release 4 of the 3GPP specifications as an additional TDD mode.

The work in 3GPP is carried out with relevant ITU recommendations in mind and the result of the work is also submitted to ITU. The organizational partners are obliged to identify regional requirements that may lead to options in the standard. Examples are regional frequency bands and special protection requirements local to a region. The specifications are developed with global roaming and circulation of terminals in mind. This implies that many regional requirements in essence will be global requirements for all terminals, since a roaming terminal has to meet the strictest of all regional requirements. Regional options in the specifications are thus more common for base stations than for terminals.

The specifications of all releases can be updated after each set of TSG meetings, which occur four times a year. The 3GPP documents are divided into releases, where each release has a set of features added compared to the previous release. The features are defined in Work Items agreed and undertaken by the TSGs. The releases from release 8 and onwards, with some main features listed for LTE, are shown in Figure 1.6. The date shown for each release is the day the content of the release was frozen. Release 10 of LTE is the version approved by ITU-R as an IMT-Advanced technology and is therefore also named *LTE-Advanced*.

1.5.3 The 3G Evolution to 4G

The first release of WCDMA Radio Access developed in TSG RAN was called release 99[3] and contained all features needed to meet the IMT-2000 requirements as defined by the ITU. This included circuit-switched voice and video services, and data services over both packet-switched and circuit-switched bearers. The first major addition of radio access features to WCDMA was HSPA, which was added in release 5 with *High Speed Downlink Packet Access* (HSDPA) and release 6 with *Enhanced Uplink*. These two are together referred to as HSPA and an overview of HSPA is given in Chapter 19 of this book. With HSPA, UTRA goes beyond the definition of a 3G mobile system and also encompasses broadband mobile data.

The 3G evolution continued in 2004, when a workshop was organized to initiate work on the 3GPP *Long-Term Evolution* (LTE) radio interface. The result of the LTE workshop was that a study item in 3GPP TSG RAN was created in December 2004. The first 6 months were spent on defining the requirements, or design targets, for LTE. These were documented in a 3GPP technical report [10] and approved in June 2005. Most notable are the requirements on high data rate at the cell edge and the importance of low delay, in addition to the normal capacity and peak data rate requirements. Furthermore, spectrum flexibility and maximum commonality between FDD and TDD solutions are pronounced.

[3] For historical reasons, the first 3GPP release is named after the year it was frozen (1999), while the following releases are numbered 4, 5, 6, etc.

During the fall of 2005, 3GPP TSG RAN WG1 made extensive studies of different basic physical layer technologies and in December 2005 the TSG RAN plenary decided that the LTE radio access should be based on OFDM in the downlink and DFT-precoded OFDM in the uplink. TSG RAN and its working groups then worked on the LTE specifications and the specifications were approved in December 2007. Work has since then continued on LTE, with new features added in each release, as shown in Figure 1.6. Chapters 7–17 will go through the details of the LTE radio interface in more detail.

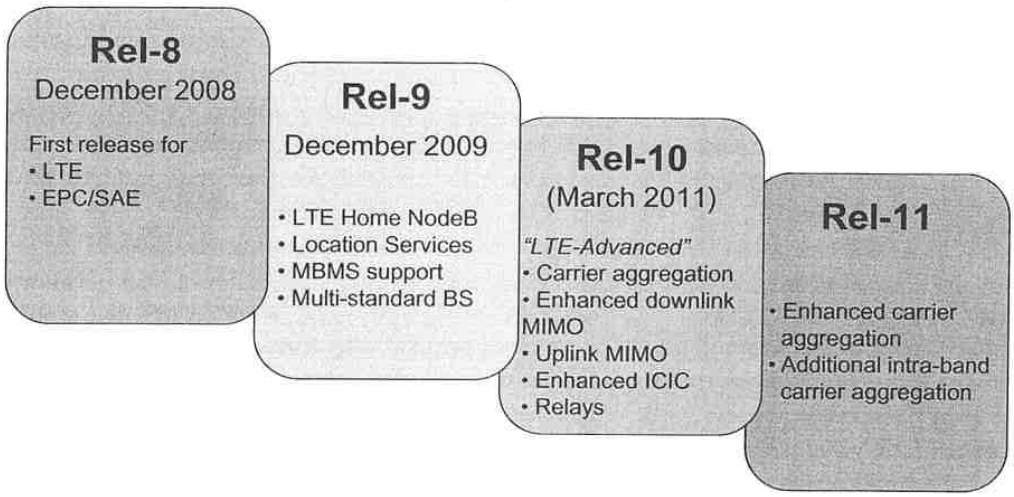


FIGURE 1.6

Releases of 3GPP specifications for LTE.

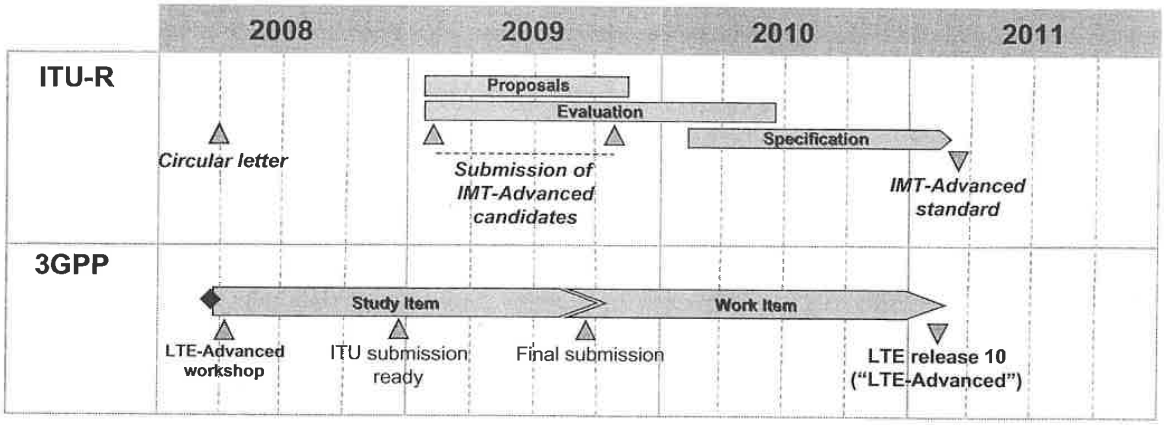


FIGURE 1.7

3GPP time schedule for LTE-Advanced in relation to ITU time-schedule on IMT-Advanced.

The work on IMT-Advanced within ITU-R WP5D came in 2008 into a phase where the detailed requirements and process were announced through a circular letter [5]. Among other things, this triggered activities in 3GPP, where a study item on *LTE-Advanced* was started. The task was to define requirements and investigate and propose technology components to be part of LTE-Advanced. The work was turned into a Work Item in 2009 in order to develop the detailed specifications.

Within 3GPP, LTE-Advanced is seen as the next major step in the evolution of LTE. LTE-Advanced is therefore not a new technology; it is an evolutionary step in the continuing development of LTE. As shown in Figure 1.6, the features that form LTE-Advanced are part of release 10 of 3GPP LTE specifications. Wider bandwidth through aggregation of multiple carriers and evolved use of advanced antenna techniques in both uplink and downlink are the major components added in LTE release 10 to reach the IMT-Advanced targets.

The work on LTE-Advanced within 3GPP is planned with the ITU-R time frame in mind, as shown in Figure 1.7. LTE-Advanced was submitted as a candidate to the ITU-R in 2009 and is now included in the set of radio interface technologies announced by ITU-R [6] in October 2010 to be included as a part of the IMT-Advanced radio interface specifications. This is very much aligned with what was from the start stated as a goal for LTE, namely that *LTE should provide the starting point for a smooth transition to 4G (= IMT-Advanced) radio access.*

Since LTE-Advanced is an integral part of 3GPP LTE release 10, it is described in detail together with the corresponding components of LTE in Chapters 7–17 of this book.

CHAPTER

OFDM Transmission

3

In this chapter, a more detailed overview of OFDM or *Orthogonal Frequency-Division Multiplexing* will be given. OFDM is the transmission scheme used for 3GPP LTE and is also used for several other radio-access technologies, for example WiMAX [14] and the DVB broadcast technologies [15].

3.1 BASIC PRINCIPLES OF OFDM

Transmission by means of OFDM can be seen as a kind of multi-carrier transmission. The basic characteristics of OFDM transmission, which distinguish it from a straightforward multi-carrier extension of a more narrowband transmission scheme as outlined in Figure 2.5 in the previous chapter, are:

- The use of a typically very large number of relatively narrowband subcarriers. In contrast, a straightforward multi-carrier extension as outlined in Figure 2.5 would typically consist of only a few subcarriers, each with a relatively wide bandwidth. As an example, the HSPA multi-carrier evolution to a 20 MHz overall transmission bandwidth consists of four (sub)carriers, each with a bandwidth of the order of 5 MHz. In comparison, OFDM transmission may imply that several hundred subcarriers are transmitted over the same radio link to the same receiver.
- Simple rectangular pulse shaping is illustrated in Figure 3.1a. This corresponds to a sinc-square-shaped per-subcarrier spectrum, as illustrated in Figure 3.1b.
- Tight frequency-domain packing of the subcarriers with a subcarrier spacing $\Delta f = 1/T_u$, where T_u is the per-subcarrier modulation-symbol time (see Figure 3.2). The subcarrier spacing is thus equal to the per-subcarrier modulation rate $1/T_u$.[1]

An illustrative description of a basic OFDM modulator is provided in Figure 3.3. It consists of a bank of N_c complex modulators, where each modulator corresponds to one OFDM subcarrier.

In complex baseband notation, a basic OFDM signal $x(t)$ during the time interval $mT_u \leq t < (m+1)T_u$ can thus be expressed as:

$$x(t) = \sum_{k=0}^{N_c-1} x_k(t) = \sum_{k=0}^{N_c-1} a_k^{(m)} e^{j2\pi k\Delta f t}, \tag{3.1}$$

where $x_k(t)$ is the kth modulated subcarrier with frequency $f_k = k \cdot \Delta f$ and $a_k^{(m)}$ is the, generally complex, modulation symbol applied to the kth subcarrier during the mth OFDM symbol interval – that is, during the time interval $mT_u \leq t < (m+1)T_u$. OFDM transmission is thus block based, implying

[1] This ignores a possible *cyclic prefix*, see Section 3.4.

4G LTE/LTE-Advanced for Mobile Broadband.
© 2011 Erik Dahlman, Stefan Parkvall & Johan Sköld. Published by Elsevier Ltd. All rights reserved.

Samsung Exhibit 1019, Page 36 of 195

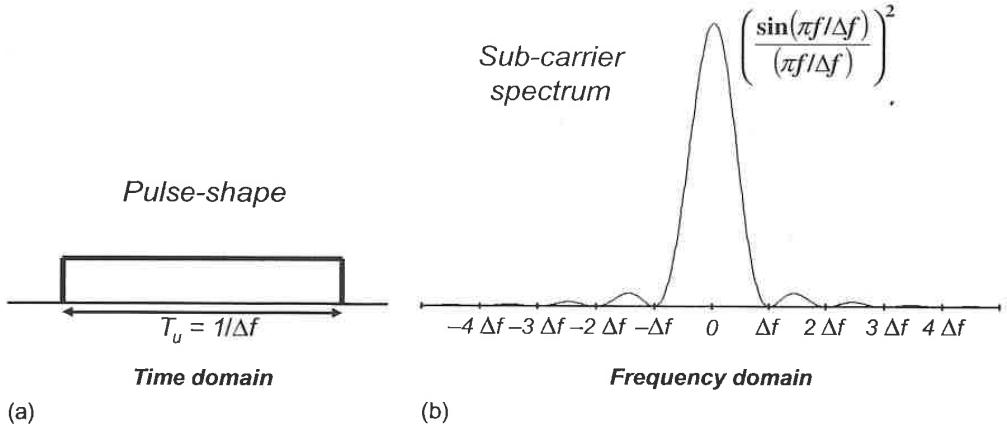


FIGURE 3.1

(a) Per-subcarrier pulse shape. (b) Spectrum for basic OFDM transmission.

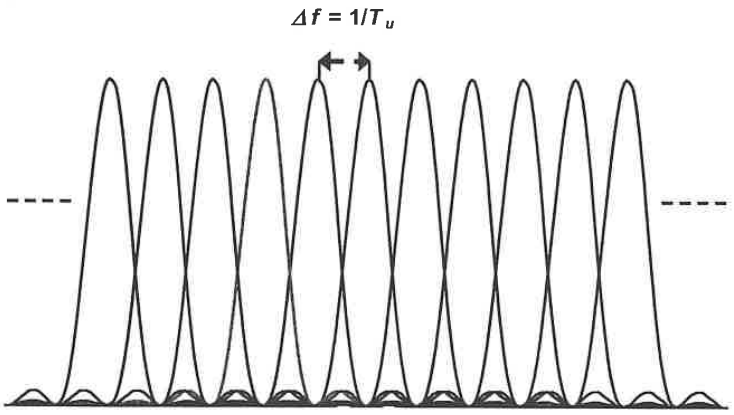


FIGURE 3.2

OFDM subcarrier spacing.

that, during each OFDM symbol interval, N_c modulation symbols are transmitted in parallel. The modulation symbols can be from any modulation alphabet, such as QPSK, 16QAM, or 64QAM.

The number of OFDM subcarriers can range from less than hundred to several thousand, with the subcarrier spacing ranging from several hundred kHz down to a few kHz. What subcarrier spacing to use depends on what types of environments the system is to operate in, including such aspects as the maximum expected radio-channel frequency selectivity (maximum expected time dispersion) and the maximum expected rate of channel variations (maximum expected Doppler spread). Once the subcarrier spacing has been selected, the number of subcarriers can be decided based on the assumed overall transmission bandwidth, taking into account acceptable out-of-band emission, etc.

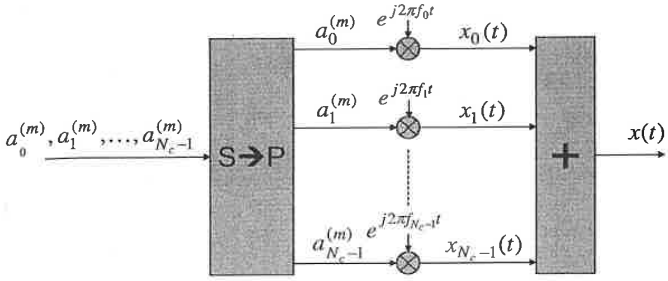


FIGURE 3.3

Principles for OFDM modulation.

The selection of OFDM subcarrier spacing and number of subcarriers is discussed in somewhat more detail in Section 3.8.

As an example, for 3GPP LTE the basic subcarrier spacing equals 15 kHz. On the other hand, the number of subcarriers depends on the transmission bandwidth, with of the order of 600 subcarriers in the case of operation in a 10 MHz spectrum allocation and correspondingly fewer/more subcarriers in the case of smaller/larger overall transmission bandwidths.

The term *Orthogonal Frequency-Division Multiplex* is due to the fact that two modulated OFDM subcarriers x_{k_1} and x_{k_2} are mutually *orthogonal* over the time interval $mT_u \le t < (m+1)T_u$, that is:

$$\int_{mT_u}^{(m+1)T_u} x_{k_1}(t)x_{k_2}^*(t)\,dt = \int_{mT_u}^{(m+1)T_u} a_{k_1}a_{k_2}^* e^{j2\pi k_1 \Delta f t} e^{-j2\pi k_2 \Delta f t}\,dt = 0 \quad \text{for } k_1 \neq k_2. \tag{3.2}$$

Thus, basic OFDM transmission can be seen as the modulation of a set of orthogonal functions $\varphi_k(t)$, where

$$\varphi_k(t) = \begin{cases} e^{j2\pi k \Delta f t} & 0 \le t < T_u \\ 0 & \text{otherwise} \end{cases}. \tag{3.3}$$

The "physical resource" in the case of OFDM transmission is often illustrated as a time–frequency grid according to Figure 3.4, where each "column" corresponds to one OFDM symbol and each "row" corresponds to one OFDM subcarrier.

3.2 OFDM DEMODULATION

Figure 3.5 illustrates the basic principle of OFDM demodulation consisting of a bank of correlators, one for each subcarrier. Taking into account the orthogonality between subcarriers according to Eqn (3.2), it is clear that, in the ideal case, two OFDM subcarriers do not cause any interference to

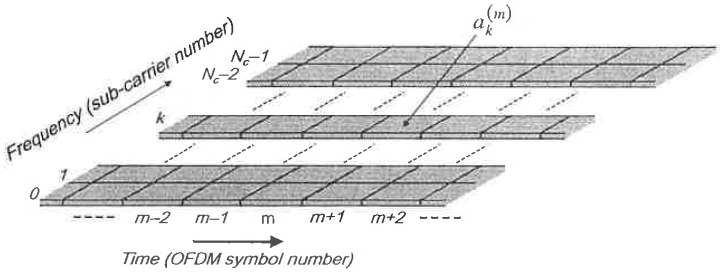


FIGURE 3.4

OFDM time–frequency grid.

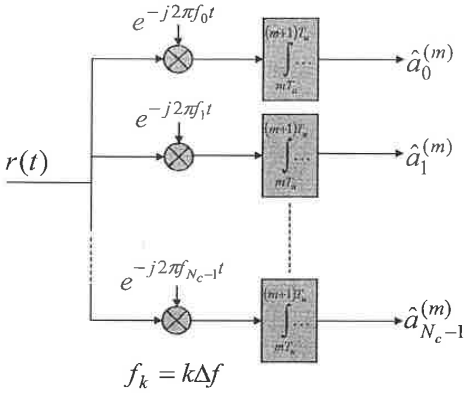


FIGURE 3.5

Basic principle of OFDM demodulation.

each other after demodulation. Note that this is the case despite the fact that the spectrum of neighbor subcarriers clearly overlaps, as can be seen from Figure 3.2. Thus, the avoidance of interference between OFDM subcarriers is not simply due to a subcarrier spectrum separation, which is, for example, the case for the kind of straightforward multi-carrier extension outlined in Figure 2.5 in the previous chapter. Rather, the subcarrier orthogonality is due to the *specific* frequency-domain structure of each subcarrier in combination with the *specific* choice of a subcarrier spacing Δf equal to the per-subcarrier symbol rate $1/T_u$. However, this also implies that, in contrast to the kind of multi-carrier transmission outlined in Section 2.3.1 of the previous chapter, any corruption of the frequency-domain structure of the OFDM subcarriers, for example due to a frequency-selective radio channel, may lead to a loss of inter-subcarrier orthogonality and thus to interference between subcarriers. To handle this and to make an OFDM signal truly robust to radio-channel frequency selectivity, *cyclic-prefix insertion* is typically used, as will be further discussed in Section 3.4.

3.3 OFDM IMPLEMENTATION USING IFFT/FFT PROCESSING

Although a bank of modulators/correlators according to Figures 3.3 and 3.5 can be used to illustrate the basic principles of OFDM modulation and demodulation respectively, these are not the most appropriate modulator/demodulator structures for actual implementation. Actually, due to its specific structure and the selection of a subcarrier spacing Δf equal to the per-subcarrier symbol rate $1/T_u$, OFDM allows for low-complexity implementation by means of computationally efficient *Fast Fourier Transform* (FFT) processing.

To confirm this, consider a time-discrete (sampled) OFDM signal where it is assumed that the sampling rate f_s is a multiple of the subcarrier spacing Δf – that is, $f_s = 1/T_s = N \cdot \Delta f$. The parameter N should be chosen so that the sampling theorem [12] is sufficiently fulfilled.[2] As $N_c \cdot \Delta f$ can be seen

[2]An OFDM signal defined according to Eqn (3.1) in theory has an infinite bandwidth and thus the sampling theorem can never be fulfilled completely.

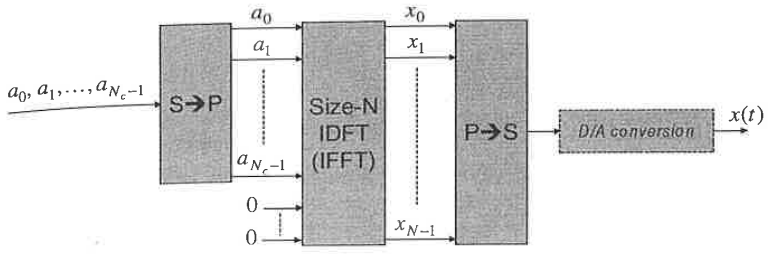


FIGURE 3.6

OFDM modulation by means of IFFT processing.

as the nominal bandwidth of the OFDM signal, this implies that N should exceed N_c with a sufficient margin.

With these assumptions, the time-discrete OFDM signal can be expressed as:[3]

$$x_n = x(nT_s) = \sum_{k=0}^{N_c-1} a_k e^{j2\pi k\Delta f n T_s} = \sum_{k=0}^{N_c-1} a_k e^{j2\pi kn/N} = \sum_{k=0}^{N-1} a'_k e^{j2\pi kn/N}, \tag{3.4}$$

where

$$a'_k = \begin{cases} a_k & 0 \le k < N_c \\ 0 & N_c \le k < N \end{cases}. \tag{3.5}$$

Thus, the sequence x_n, or, in other words, the sampled OFDM signal, is the size-N Inverse Discrete Fourier Transform (IDFT) of the block of modulation symbols a_0, a_1, …, a_{N_c-1} extended with zeros to length N. OFDM modulation can thus be implemented by means of IDFT processing followed by digital-to-analog conversion, as illustrated in Figure 3.6. In particular, by selecting the IDFT size N equal to 2^m for some integer m, the OFDM modulation can be implemented by means of implementation-efficient radix-2 Inverse Fast Fourier Transform (IFFT) processing. It should be noted that the ratio N/N_c, which could be seen as the *over-sampling* of the time-discrete OFDM signal, could very well be, and typically is, a non-integer number. As an example and as already mentioned, for 3GPP LTE the number of subcarriers N_c is approximately 600 in the case of a 10 MHz spectrum allocation. The IFFT size can then, for example, be selected as $N = 1024$. This corresponds to a sampling rate $f_s = N \cdot \Delta f = 15.36$ MHz, where $\Delta f = 15$ kHz is the LTE subcarrier spacing.

It is important to understand that IDFT/IFFT-based implementation of an OFDM modulator, and even more so the exact IDFT/IFFT size, are just transmitter-implementation choices and not something that would be mandated by any radio-access specification. As an example, nothing forbids the implementation of an OFDM modulator as a set of parallel modulators, as illustrated in Figure 3.3. Also, nothing prevents the use of a larger IFFT size, for example a size-2048 IFFT size, even in the case of a smaller number of OFDM subcarriers.

[3] From now on the index m on the modulation symbols, indicating the OFDM symbol number, will be ignored unless especially needed.

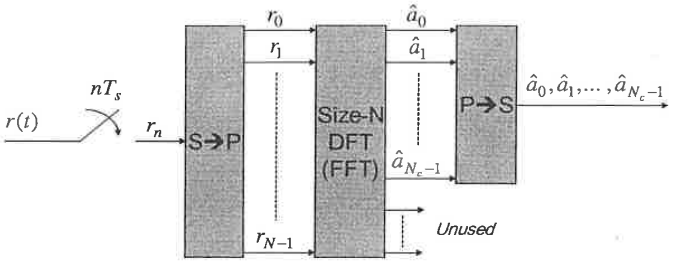


FIGURE 3.7

OFDM demodulation by means of FFT processing.

Similar to OFDM modulation, efficient FFT processing can be used for OFDM demodulation, replacing the bank of N_c parallel demodulators of Figure 3.5 with sampling with some sampling rate $f_s = 1/T_s$, followed by a size-N DFT/FFT, as illustrated in Figure 3.7.

3.4 CYCLIC-PREFIX INSERTION

As described in Section 3.2, an uncorrupted OFDM signal can be demodulated without any interference between subcarriers. One way to understand this *subcarrier orthogonality* is to recognize that a modulated subcarrier $x_k(t)$ in Eqn (3.1) consists of an integer number of periods of complex exponentials during the demodulator integration interval $T_u = 1/\Delta f$.

However, in case of a time-dispersive channel the orthogonality between the subcarriers will, at least partly, be lost. The reason for this loss of subcarrier orthogonality in the case of a time-dispersive channel is that, in this case, the demodulator correlation interval for one path will overlap with the symbol boundary of a different path, as illustrated in Figure 3.8. Thus, the integration interval will not necessarily correspond to an integer number of periods of complex exponentials of that path as the modulation symbols a_k may differ between consecutive symbol intervals. As a consequence, in the case of a time-dispersive channel there will not only be inter-symbol interference within a subcarrier, but also interference between subcarriers.

Another way to explain the interference between subcarriers in the case of a time-dispersive channel is to have in mind that time dispersion on the radio channel is equivalent to a frequency-selective channel frequency response. As clarified in Section 3.2, orthogonality between OFDM subcarriers is not simply due to frequency-domain separation but due to the *specific* frequency-domain structure of each subcarrier. Even if the frequency-domain channel is constant over a bandwidth corresponding to the main lobe of an OFDM subcarrier and only the subcarrier side lobes are corrupted due to the radio-channel frequency selectivity, the orthogonality between subcarriers will be lost with inter-subcarrier interference as a consequence. Due to the relatively large side lobes of each OFDM subcarrier, already a relatively limited amount of time dispersion or, equivalently, a relatively modest radio-channel frequency selectivity may cause substantial interference between subcarriers.

To deal with this problem and to make an OFDM signal truly insensitive to time dispersion on the radio channel, so-called *cyclic-prefix insertion* is typically used in OFDM transmission. As illustrated

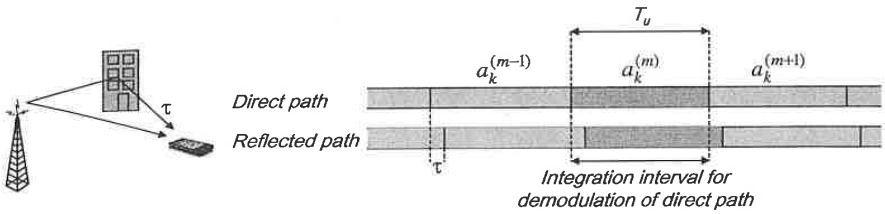


FIGURE 3.8

Time dispersion and corresponding received-signal timing.

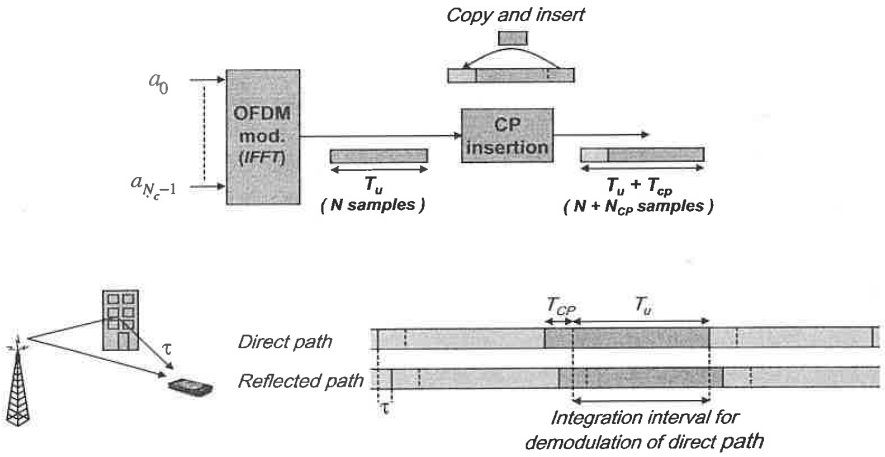


FIGURE 3.9

Cyclic-prefix insertion.

in Figure 3.9, cyclic-prefix insertion implies that the last part of the OFDM symbol is copied and inserted at the beginning of the OFDM symbol. Cyclic-prefix insertion thus increases the length of the OFDM symbol from T_u to $T_u + T_{CP}$, where T_{CP} is the length of the cyclic prefix, with a corresponding reduction in the OFDM symbol rate as a consequence. As illustrated in the lower part of Figure 3.9, if the correlation at the receiver side is still only carried out over a time interval $T_u = 1/\Delta f$, subcarrier orthogonality will then also be preserved in the case of a time-dispersive channel, as long as the span of the time dispersion is shorter than the cyclic-prefix length.

In practice, cyclic-prefix insertion is carried out on the time-discrete output of the transmitter IFFT. Cyclic-prefix insertion then implies that the last N_{CP} samples of the IFFT output block of length N are copied and inserted at the beginning of the block, increasing the block length from N to $N + N_{CP}$. At the receiver side, the corresponding samples are discarded before OFDM demodulation by means of, for example, DFT/FFT processing.

Cyclic-prefix insertion is beneficial in the sense that it makes an OFDM signal insensitive to time dispersion as long as the span of the time dispersion does not exceed the length of the cyclic prefix. The drawback of cyclic-prefix insertion is that only a fraction $T_u/(T_u + T_{CP})$ of the received signal

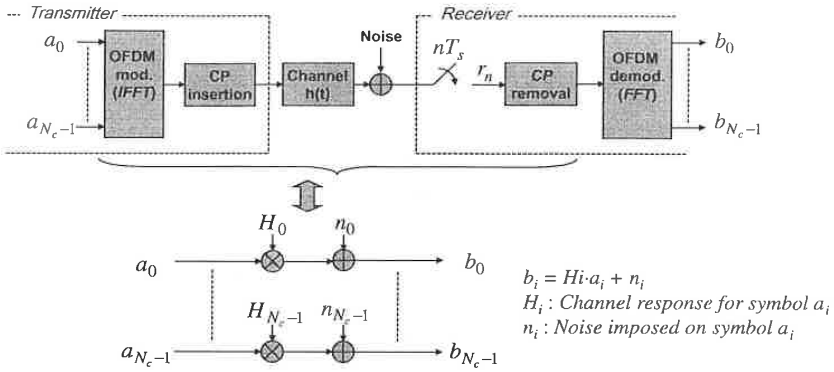


FIGURE 3.10

Frequency-domain model of OFDM transmission/reception.

power is actually utilized by the OFDM demodulator, implying a corresponding power loss in the demodulation. In addition to this power loss, cyclic-prefix insertion also implies a corresponding loss in terms of bandwidth as the OFDM symbol rate is reduced without a corresponding reduction in the overall signal bandwidth.

One way to reduce the relative overhead due to cyclic-prefix insertion is to reduce the subcarrier spacing Δf, with a corresponding increase in the symbol time T_u as a consequence. However, this will increase the sensitivity of the OFDM transmission to fast channel variations – that is, high Doppler spread – as well as different types of frequency errors (see further Section 3.8).

It is also important to understand that the cyclic prefix does not necessarily have to cover the entire length of the channel time dispersion. In general, there is a trade-off between the power loss due to the cyclic prefix and the signal corruption (inter-symbol and inter-subcarrier interference) due to residual time dispersion not covered by the cyclic prefix and, at a certain point, further reduction of the signal corruption due to further increase of the cyclic-prefix length will not justify the corresponding additional power loss. This also means that, although the amount of time dispersion typically increases with the cell size, beyond a certain cell size there is often no reason to increase the cyclic prefix further as the corresponding power loss due to a further increase of the cyclic prefix would have a larger negative impact, compared to the signal corruption due to the residual time dispersion not covered by the cyclic prefix [16].

3.5 FREQUENCY-DOMAIN MODEL OF OFDM TRANSMISSION

Assuming a sufficiently large cyclic prefix, the linear convolution of a time-dispersive radio channel will appear as a circular convolution during the demodulator integration interval T_u. The combination of OFDM modulation (IFFT processing), a time-dispersive radio channel, and OFDM demodulation (*FFT* processing) can then be seen as a *frequency-domain* channel as illustrated in Figure 3.10, where

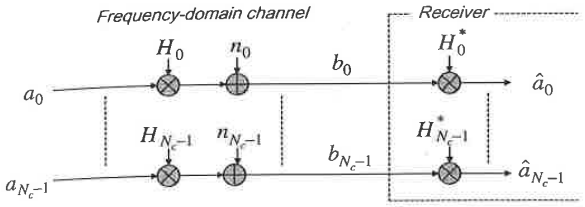


FIGURE 3.11

Frequency-domain model of OFDM transmission/reception with "one-tap equalization" at the receiver.

the frequency-domain channel taps $H_0, \ldots, H_{N_c-1}$ can be directly derived from the channel impulse response.

The demodulator output b_k in Figure 3.10 is the transmitted modulation symbol a_k scaled and phase rotated by the complex frequency-domain channel tap H_k and impaired by noise n_k. To properly recover the transmitted symbol for further processing, for example data demodulation and channel decoding, the receiver should multiply b_k by the complex conjugate of H_k, as illustrated in Figure 3.11. This is often expressed as a *one-tap equalizer* being applied to each received subcarrier.

3.6 CHANNEL ESTIMATION AND REFERENCE SYMBOLS

As described above, to demodulate the transmitted modulation symbol a_k and allow for proper decoding of the transmitted information at the receiver side, scaling with the complex conjugate of the frequency-domain channel tap H_k should be applied after OFDM demodulation (*FFT* processing) (see Figure 3.11). To be able to do this, the receiver obviously needs an estimate of the frequency-domain channel taps $H_0, \ldots, H_{N_c-1}$.

The frequency-domain channel taps can be estimated indirectly by first estimating the channel impulse response and, from that, calculating an estimate of H_k. However, a more straightforward approach is to estimate the frequency-domain channel taps directly. This can be done by inserting known *reference symbols*, sometimes also referred to as *pilot symbols*, at regular intervals within the OFDM time–frequency grid, as illustrated in Figure 3.12. Using knowledge about the reference symbols, the receiver can estimate the frequency-domain channel around the location of the reference symbol. The reference symbols should also have a sufficiently high density in both the time and the frequency domains to be able to provide estimates for the entire time–frequency grid in the case of radio channels subject to high frequency and/or time selectivity.

Different more or less advanced algorithms can be used for the channel estimation, ranging from simple averaging in combination with linear interpolation to Minimum-Mean-Square-Error (MMSE) estimation relying on more detailed knowledge of the channel time/frequency-domain characteristics. Readers are referred to, for example, [17] for a more in-depth discussion on channel estimation for OFDM.

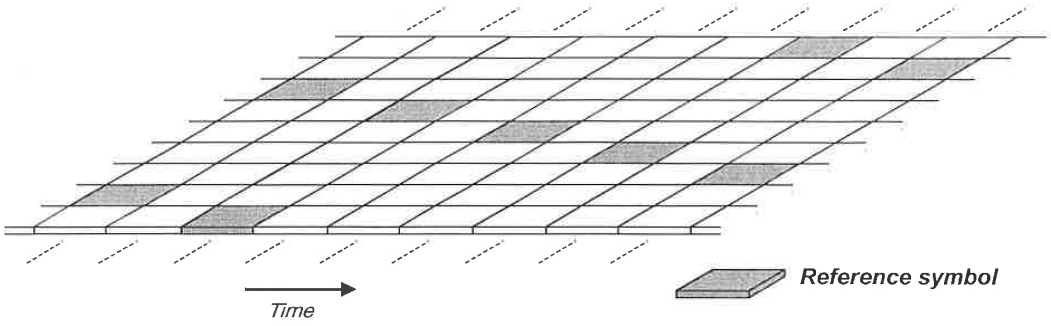


FIGURE 3.12

Time–frequency grid with known reference symbols.

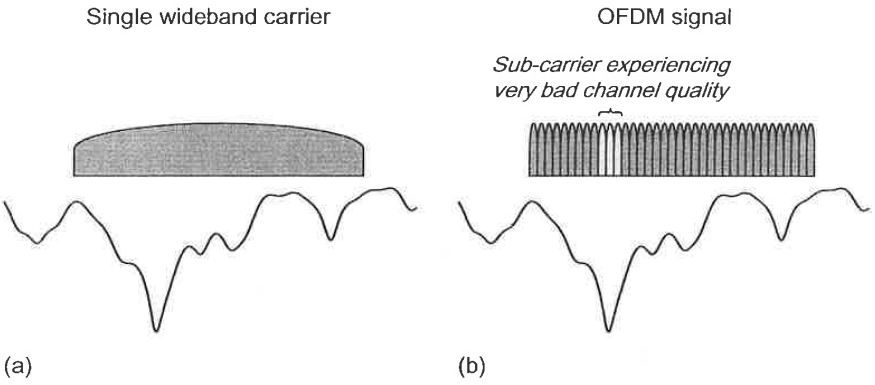


FIGURE 3.13

(a) Transmission of single wideband carrier. (b) OFDM transmission over a frequency-selective channel.

3.7 FREQUENCY DIVERSITY WITH OFDM: IMPORTANCE OF CHANNEL CODING

As discussed in Section 2.3 in the previous chapter, a radio channel is always subject to some degree of frequency selectivity, implying that the channel quality will vary in the frequency domain. In the case of a single wideband carrier, such as a WCDMA carrier, each modulation symbol is transmitted over the entire signal bandwidth. Thus, in the case of the transmission of a single wideband carrier over a highly frequency-selective channel (see Figure 3.13a), each modulation symbol will be transmitted both over frequency bands with relatively good quality (relatively high signal strength) and frequency bands with low quality (low signal strength). Such transmission of information over multiple frequency bands with different instantaneous channel quality is also referred to as *frequency diversity*.

On the other hand, in the case of OFDM transmission each modulation symbol is mainly confined to a relatively narrow bandwidth. Thus, for OFDM transmission over a frequency-selective channel, certain modulation symbols may be fully confined to a frequency band with very low instantaneous

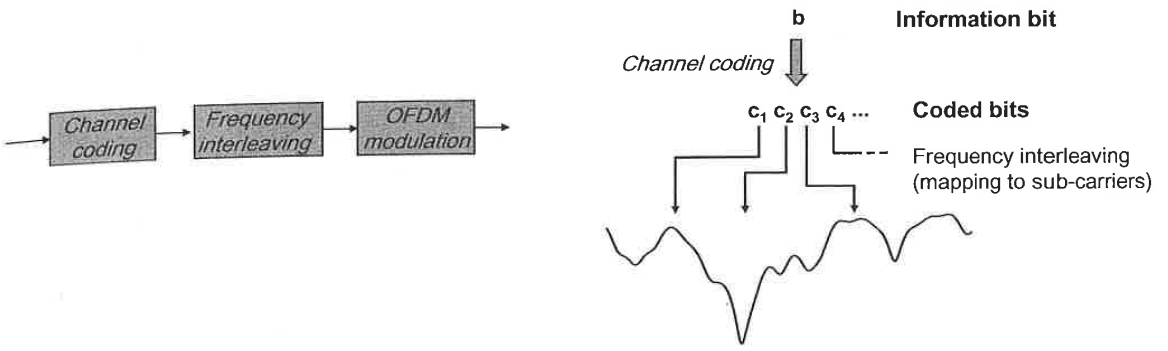


FIGURE 3.14

Channel coding in combination with frequency-domain interleaving to provide frequency diversity in the case of OFDM transmission.

signal strength, as illustrated in Figure 3.13b. Thus, the individual modulation symbols will typically not experience any substantial frequency diversity even if the channel is highly frequency selective over the overall OFDM transmission bandwidth. As a consequence, the basic error-rate performance of OFDM transmission over a frequency-selective channel is relatively poor and especially much worse than the basic error rate in the case of a single wideband carrier.

However, in practice channel coding is used in most cases of digital communication and especially in mobile communication. Channel coding implies that each bit of information to be transmitted is spread over several, often very many, code bits. If these coded bits are then, via modulation symbols, mapped to a set of OFDM subcarriers that are well distributed over the overall transmission bandwidth of the OFDM signal, as illustrated in Figure 3.14, each information bit will experience frequency diversity in the case of transmission over a radio channel that is frequency selective over the transmission bandwidth, despite the fact that the subcarriers, and thus also the code bits, will not experience any frequency diversity. Distributing the code bits in the frequency domain, as illustrated in Figure 3.14, is sometimes referred to as *frequency interleaving*. This is similar to the use of time-domain interleaving to benefit from channel coding in the case of fading that varies in time.

Thus, in contrast to the transmission of a single wideband carrier, channel coding (combined with frequency interleaving) is an essential component in order for OFDM transmission to be able to benefit from frequency diversity on a frequency-selective channel. As channel coding is typically used in most cases of mobile communication this is not a very serious drawback, especially taking into account that a significant part of the available frequency diversity can be captured already with a relatively high code rate.

3.8 SELECTION OF BASIC OFDM PARAMETERS

If OFDM is to be used as the transmission scheme in a mobile-communication system, the following basic OFDM parameters need to be decided upon:

- The subcarrier spacing Δf.
- The number of subcarriers N_c, which, together with the subcarrier spacing, determines the overall transmission bandwidth of the OFDM signal.

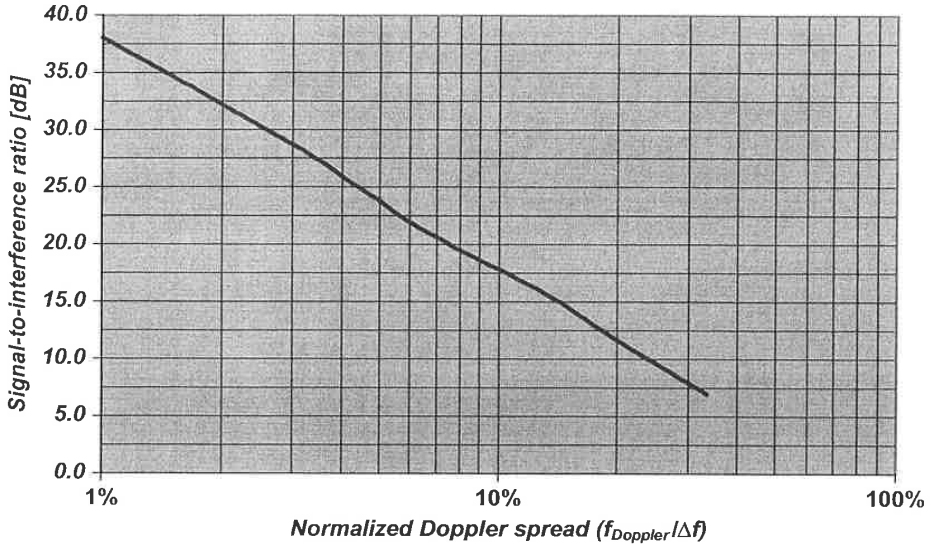


FIGURE 3.15

Subcarrier interference as a function of the normalized Doppler spread $f_{Doppler}/\Delta f$.

- The cyclic-prefix length T_{CP}. Together with the subcarrier spacing $\Delta f = 1/T_u$, the cyclic-prefix length determines the overall OFDM symbol time $T = T_{CP} + T_u$ or, equivalently, the OFDM symbol rate.

3.8.1 OFDM Subcarrier Spacing

There are two factors that constrain the selection of the OFDM subcarrier spacing:

- The OFDM subcarrier spacing should be as small as possible (T_u as large as possible) to minimize the relative cyclic-prefix overhead $T_{CP}/(T_u + T_{CP})$ – see further Section 3.8.3.
- A too small subcarrier spacing increases the sensitivity of the OFDM transmission to Doppler spread and different kinds of frequency inaccuracies.

A requirement for the OFDM subcarrier orthogonality (3.2) to hold at the receiver side – that is, after the transmitted signal has propagated over the radio channel – is that the instantaneous channel does not vary noticeably during the demodulator correlation interval T_u (see Figure 3.5). In the case of such channel variations, for example due to very high Doppler spread, the orthogonality between subcarriers will be lost, with inter-subcarrier interference as a consequence. Figure 3.15 illustrates the subcarrier signal-to-interference ratio due to inter-subcarrier interference between two neighboring subcarriers, as a function of the normalized Doppler spread. When considering Figure 3.15, it should be borne in mind that a subcarrier will be subject to interference from multiple subcarriers on both sides[4] – that is, the overall inter-subcarrier interference from all subcarriers will be higher than what is illustrated in Figure 3.15.

[4]Except for the subcarrier at the edge of the spectrum.

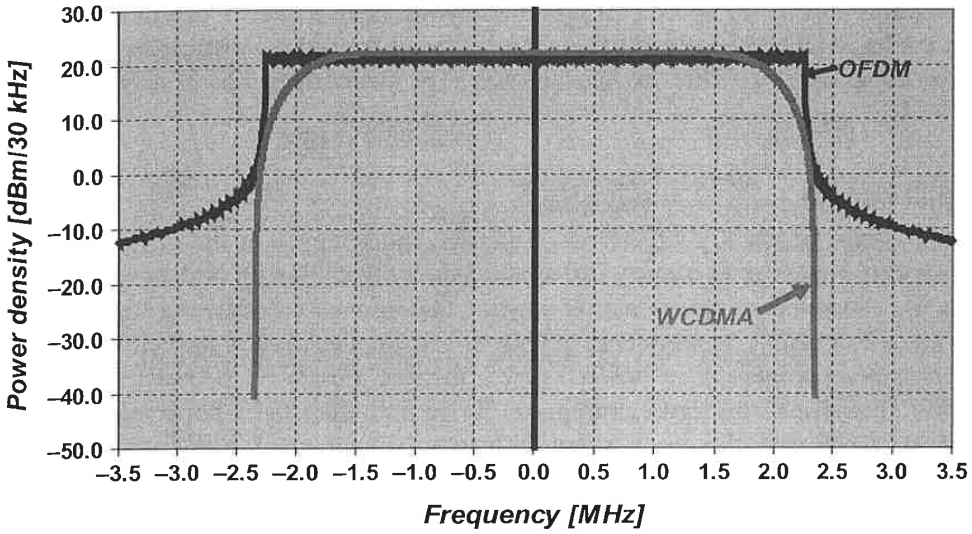


FIGURE 3.16

Spectrum of a basic 5 MHz OFDM signal compared with WCDMA spectrum.

In practice, the amount of inter-subcarrier interference that can be accepted very much depends on the service to be provided and to what extent the received signal is anyway corrupted due to noise and other impairments. As an example, on the cell border of large cells the signal-to-noise/interference ratio will anyway be relatively low, with relatively low achievable data rates as a consequence. A small amount of additional inter-subcarrier interference, for example due to Doppler spread, may then be more or less negligible. At the same time, in high signal-to-noise/interference scenarios, for example in small cells with low traffic or close to the base station, where high data rates are to be provided, the same amount of inter-subcarrier interference may have a much more negative impact.

It should also be noted that, in addition to Doppler spread, inter-subcarrier interference will also be due to different transmitter and receiver inaccuracies, such as frequency errors and phase noise.

3.8.2 Number of Subcarriers

Once the subcarrier spacing has been selected based on environment, expected Doppler spread and time dispersion, etc., the number of subcarriers can be determined based on the amount of spectrum available and the acceptable out-of-band emissions.

The basic bandwidth of an OFDM signal equals $N_c \cdot \Delta f$ – that is, the number of subcarriers multiplied by the subcarrier spacing. However, as can be seen in Figure 3.16, the spectrum of a basic OFDM signal falls off very slowly outside the basic OFDM bandwidth and especially much slower than for a WCDMA signal. The reason for the large *out-of-band emission* of a basic OFDM signal is the use of rectangular pulse shaping (Figure 3.1), leading to per-subcarrier side lobes that fall off relatively slowly. However, in practice, straightforward filtering or *time-domain windowing* [18] will be used to suppress a main part of the OFDM out-of-band emissions. Thus, in practice, typically of the

order of 10% guard-band is needed for an OFDM signal implying, as an example, that in a spectrum allocation of 5 MHz, the basic OFDM bandwidth $N_c \cdot \Delta f$ could be of the order of 4.5 MHz. Assuming, for example, a subcarrier spacing of 15 kHz as selected for LTE, this corresponds to approximately 300 subcarriers in 5 MHz.

3.8.3 Cyclic-Prefix Length

In principle, the cyclic-prefix length T_{CP} should cover the maximum length of the time dispersion expected to be experienced. However, as already discussed, increasing the length of the cyclic prefix, without a corresponding reduction in the subcarrier spacing Δf, implies an additional overhead in terms of power as well as bandwidth. In particular, the power loss implies that, as the cell size grows and the system performance becomes more power limited, there is a trade-off between the loss in power due to the cyclic prefix and the signal corruption due to time dispersion not covered by the cyclic prefix. As already mentioned, this implies that, although the amount of time dispersion typically increases with the cell size, beyond a certain cell size there is often no reason to increase the cyclic prefix further as the corresponding power loss would have a larger negative impact, compared to the signal corruption due to the residual time dispersion not covered by the cyclic prefix [16].

One situation where a longer cyclic prefix may be needed is in the case of multi-cell transmission using SFN (Single-Frequency Network), as further discussed in Section 3.11.

Thus, to be able to optimize performance to different environments, some OFDM-based systems support multiple cyclic-prefix lengths. The different cyclic-prefix lengths can then be used in different transmission scenarios:

- Shorter cyclic prefix in small-cell environments to minimize the cyclic-prefix overhead.
- Longer cyclic prefix in environments with extreme time dispersion and especially in the case of SFN operation.

3.9 VARIATIONS IN INSTANTANEOUS TRANSMISSION POWER

According to Section 2.3.1, one of the drawbacks of multi-carrier transmission is the corresponding large variations in the instantaneous transmit power, implying a reduced power-amplifier efficiency and higher mobile-terminal power consumption; alternatively, that the power-amplifier output power has to be reduced with a reduced range as a consequence. Being a kind of multi-carrier transmission scheme, OFDM is subject to the same drawback.

However, a large number of different methods have been proposed to reduce the large power peaks of an OFDM signal:

- In the case of *tone reservation* [19], a subset of the OFDM subcarriers are not used for data transmission. Instead, these subcarriers are modulated in such a way that the largest peaks of the overall OFDM signal are suppressed, allowing for a reduced power-amplifier back-off. One drawback of tone reservation is the bandwidth loss due to the fact that a number of subcarriers are not available for actual data transmission. The calculation of what modulation to apply to the reserved tones can also be of relatively high complexity.
- In the case of *prefiltering* or *precoding*, linear processing is applied to the sequence of modulation symbols before OFDM modulation. DFTS-spread OFDM (DFTS-OFDM) as described in the next chapter and which is used for the LTE uplink, can be seen as one kind of prefiltering.

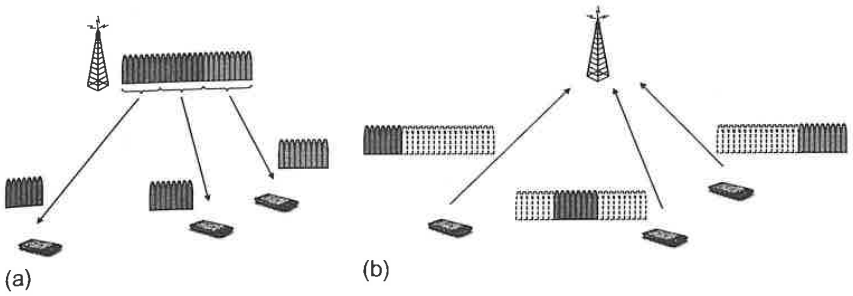


FIGURE 3.17

OFDM as a user-multiplexing/multiple-access scheme: (a) downlink; (b) uplink.

- In the case of *selective scrambling* [20], the coded-bit sequence to be transmitted is scrambled with a number of different scrambling codes. Each scrambled sequence is then OFDM modulated and the signal with the lowest peak power is selected for transmission. After OFDM demodulation at the receiver side, descrambling and subsequent decoding is carried out for all the possible scrambling sequences. Only the decoding carried out for the scrambling code actually used for the transmission will provide a correct decoding result. A drawback of selective scrambling is an increased receiver complexity as multiple decodings need to be carried out in parallel.

Readers are referred to the references above for a more in-depth discussion on different peak-reduction schemes.

3.10 OFDM AS A USER-MULTIPLEXING AND MULTIPLE-ACCESS SCHEME

The discussion has, until now, implicitly assumed that all OFDM subcarriers are transmitted from the same transmitter to a certain receiver, that is:

- downlink transmission of all subcarriers to a *single* terminal;
- uplink transmission of all subcarriers from a *single* terminal.

However, OFDM can also be used as a *user-multiplexing* or *multiple-access scheme*, allowing for simultaneous frequency-separated transmissions to/from multiple terminals (see Figure 3.17).

In the downlink direction, OFDM as a user-multiplexing scheme implies that, in each OFDM symbol interval, different subsets of the overall set of available subcarriers are used for transmission to *different* terminals (see Figure 3.17a).

Similarly, in the uplink direction, OFDM as a *user-multiplexing* or *multiple-access scheme* implies that, in each OFDM symbol interval, different subsets of the overall set of subcarriers are used for data transmission from *different* terminals (see Figure 3.17b). In this case, the term Orthogonal Frequency-Division Multiple Access or OFDMA is also often used.[5]

[5]The term OFDMA is sometimes also used to denote the use of OFDM to multiplex multiple users in the downlink, as illustrated in Figure 3.17a.

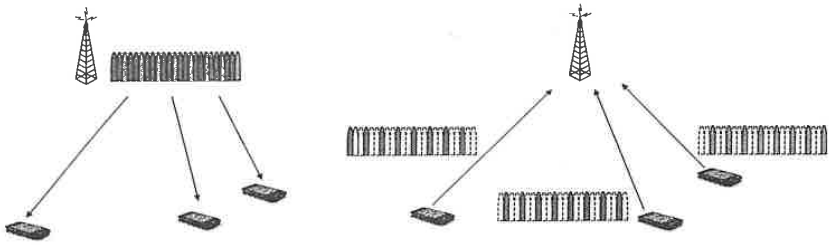


FIGURE 3.18

Distributed user multiplexing.

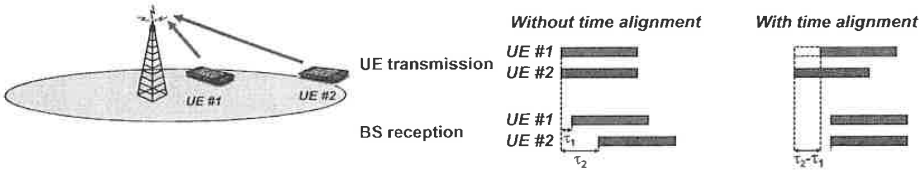


FIGURE 3.19

Uplink transmission-timing control.

Figure 3.17 assumes that *consecutive* subcarriers are used for transmission to/from the same terminal. However, distributing the subcarriers to/from a terminal in the frequency domain is also possible, as illustrated in Figure 3.18. The benefit of such *distributed* user multiplexing or *distributed* multiple access is a possibility for additional frequency diversity as each transmission is spread over a wider bandwidth.

In the case when OFDMA is used as an uplink multiple-access scheme – that is, in the case of frequency multiplexing of OFDM signals from multiple terminals – it is critical that the transmissions from the different terminals arrive approximately time aligned at the base station. More specifically, the transmissions from the different terminals should arrive at the base station with a timing misalignment less than the length of the cyclic prefix to preserve orthogonality between subcarriers received from different terminals and thus avoid inter-user interference.

Due to the differences in distance to the base station for different terminals and the corresponding differences in the propagation time (which may far exceed the length of the cyclic prefix), it is therefore necessary to control the uplink transmission timing of each terminal (see Figure 3.19). Such *transmission-timing control* should adjust the transmit timing of each terminal to ensure that uplink transmissions arrive approximately time aligned at the base station. As the propagation time changes as the terminal is moving within the cell, the transmission-timing control should be an active process, continuously adjusting the exact transmit timing of each terminal.

Furthermore, even in the case of perfect transmission-timing control, there will always be some interference between subcarriers, for example due to frequency errors. Typically this interference is relatively low in the case of reasonable frequency errors, Doppler spread, etc. (see Section 3.8). However, this assumes that the different subcarriers are received with at least approximately the same

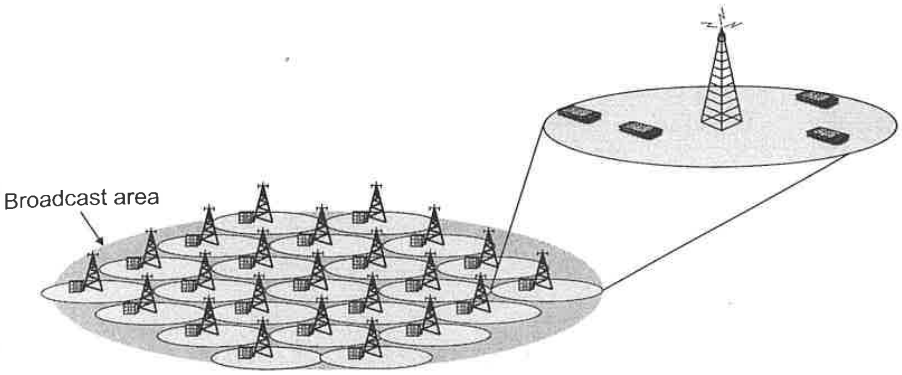


FIGURE 3.20

Broadcast scenario.

power. In the uplink, the propagation distance and thus the path loss of the different mobile-terminal transmissions may differ significantly. If two terminals are transmitting with the same power, the received-signal strengths may thus differ significantly, implying a potentially significant interference from the stronger signal to the weaker signal unless the subcarrier orthogonality is perfectly retained. To avoid this, at least some degree of *uplink transmission-power control* may need to be applied in case of uplink OFDMA, reducing the transmission power of user terminals close to the base station and ensuring that all received signals will be of approximately the same power.

3.11 MULTI-CELL BROADCAST/MULTICAST TRANSMISSION AND OFDM

The provision of broadcast/multicast services in a mobile-communication system implies that the same information is to be *simultaneously* provided to multiple terminals, often dispersed over a large area corresponding to a large number of cells, as shown in Figure 3.20. The broadcast/multicast information may be a TV news clip, information about the local weather conditions, stock-market information, or any other kind of information that, at a given time, may be of interest to a large number of people.

When the same information is to be provided to multiple terminals within a cell it is often beneficial to provide this information as a single "broadcast" radio transmission covering the entire cell and simultaneously being received by all relevant terminals (Figure 3.21a), rather than providing the information by means of individual transmissions to each terminal (unicast transmission; see Figure 3.21b).

As a broadcast transmission according to Figure 3.21 also has to be dimensioned to reach the worst-case terminals, including terminals at the cell border, it will be relative costly in terms of the recourses (base-station transmit power) needed to provide a certain broadcast-service data rate. Alternatively, taking into account the limited signal-to-noise ratio that can be achieved at, for example, the cell edge, the achievable broadcast data rates may be relatively limited, especially in the case of large cells. One way to increase the broadcast data rates would then be to reduce the cell size, thereby increasing the cell-edge receive power. However, this will increase the number of cells to cover a certain area and is thus obviously negative from a cost-of-deployment point of view.

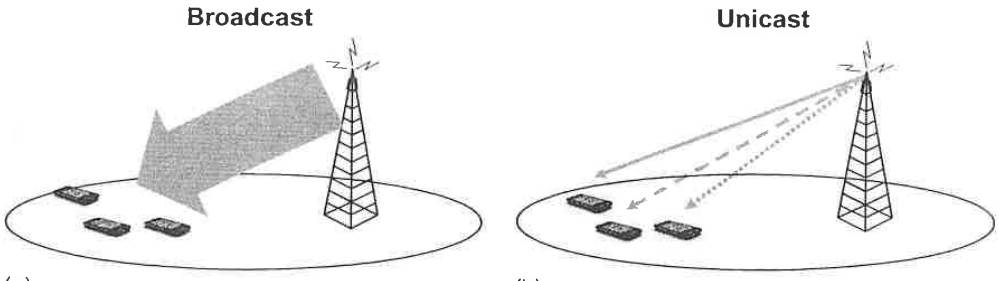


FIGURE 3.21

Broadcast vs. unicast transmission. (a) Broadcast. (b) Unicast.

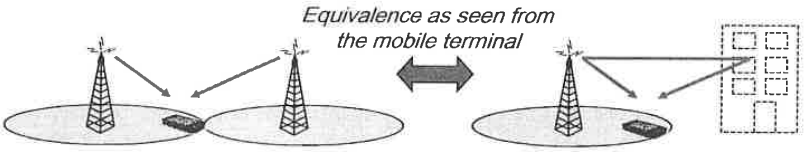


FIGURE 3.22

Equivalence between simulcast transmission and multi-path propagation.

However, as discussed above, the provision of broadcast/multicast services in a mobile-communication network typically implies that identical information is to be provided over a large number of cells. In such a case, the resources (downlink transmit power) needed to provide a certain broadcast data rate can be considerably reduced if terminals at the cell edge can utilize the received power from broadcast transmissions from multiple cells when detecting/decoding the broadcast data.

One way to mitigate this and further improve the provision of broadcast/multicast services in a mobile-communication network is to ensure that the broadcast transmissions from different cells *are truly identical* and *transmitted mutually time aligned*. In this case, the transmissions received from multiple cells will, as seen from the terminal, appear as a single transmission subject to severe multi-path propagation, as illustrated in Figure 3.22. The transmission of identical time-aligned signals from multiple cells, especially in the case of provision of broadcast/multicast services, is sometimes referred to as Single-Frequency Network (SFN) operation [21].

In the case of such identical time-aligned transmissions from multiple cells, the "inter-cell interference" due to transmissions in neighboring cells will, from a terminal point of view, be replaced by signal corruption due to time dispersion. If the broadcast transmission is based on OFDM with a cyclic prefix that covers the main part of this "time dispersion", the achievable broadcast data rates are thus only limited by noise, implying that, especially in smaller cells, very high broadcast data rates can be achieved. Furthermore, the OFDM receiver does not need to explicitly identify the cells to be soft combined. Rather, all transmissions that fall within the cyclic prefix will "automatically" be captured by the receiver.

CHAPTER

Scheduling, Link Adaptation, and Hybrid ARQ

6

One key characteristic of mobile radio communication is the typically rapid and significant variations in the instantaneous channel conditions. There are several reasons for these variations. Frequency-selective fading will result in rapid and random variations in the channel attenuation. Shadow fading and distance-dependent path loss will also significantly affect the average received signal strength. Finally, the interference at the receiver due to transmissions in other cells and by other terminals will also impact the interference level. Hence, to summarize, there will be rapid, and to some extent random, variations in the experienced quality of each radio link in a cell, variations that must be taken into account and preferably exploited.

In this chapter, some of the techniques for handling variations in the instantaneous radio-link quality will be discussed. *Channel-dependent scheduling* in a mobile-communication system deals with the question of how to share, between different users (different terminals), the radio resource(s) available in the system to achieve as efficient resource utilization as possible. Typically, this implies minimizing the amount of resources needed per user and thus allowing for as many users as possible in the system, while still satisfying whatever quality-of-service requirements that may exist. Closely related to scheduling is *link adaptation*, which deals with how to set the transmission parameters of a radio link to handle variations of the radio-link quality.

Both channel-dependent scheduling and link adaptation try to exploit the channel variations through appropriate processing *prior* to transmission of the data. However, due to the random nature of the variations in the radio-link quality, perfect adaptation to the instantaneous radio-link quality is never possible. *Hybrid ARQ,* which requests retransmission of erroneously received data packets, is therefore useful. This can be seen as a mechanism for handling variations in the instantaneous radio-link quality *after* transmission and nicely complements channel-dependent scheduling and link adaptation. Hybrid ARQ also serves the purpose of handling random errors due to, for example, noise in the receiver.

6.1 LINK ADAPTATION: POWER AND RATE CONTROL

Historically, *dynamic transmit-power control* has been used in CDMA-based mobile-communication systems such as WCDMA and CDMA2000 to compensate for variations in the instantaneous channel conditions.[1] As the name suggests, dynamic power control dynamically adjusts the radio-link transmit

[1] Power control is also used in GSM, but operating on a much slower basis than in CDMA-based systems. In CDMA-based systems, power control is also essential to handle the near–far problem and the associated performance impact from non-orthogonal uplink transmissions.

4G LTE/LTE-Advanced for Mobile Broadband.
© 2011 Erik Dahlman, Stefan Parkvall & Johan Sköld. Published by Elsevier Ltd. All rights reserved.

Samsung Exhibit 1019, Page 54 of 195

power to compensate for variations and differences in the instantaneous channel conditions. The aim of these adjustments is to maintain a (near) constant E_b/N_0 at the receiver to successfully transmit data without the error probability becoming too high. In principle, transmit-power control increases the power at the transmitter when the radio link experiences poor radio conditions (and vice versa). Thus, the transmit power is in essence inversely proportional to the channel quality, as illustrated in Figure 6.1a. This results in a basically constant data rate, regardless of the channel variations. For services such as circuit-switched voice, this is a desirable property. Transmit-power control can be seen as one type of link adaptation – that is, the adjustment of transmission parameters, in this case the transmit power – to adapt to differences and variations in the instantaneous channel conditions to maintain the received E_b/N_0 at a desired level.

However, in many cases of mobile communication, especially in the case of packet-data traffic, there is not a strong need to provide a certain constant data rate over a radio link. Rather, from a user perspective, the data rate provided over the radio interface should simply be as "high as possible". Actually, even in case of typical "constant-rate" services such as voice and video, (short-term)

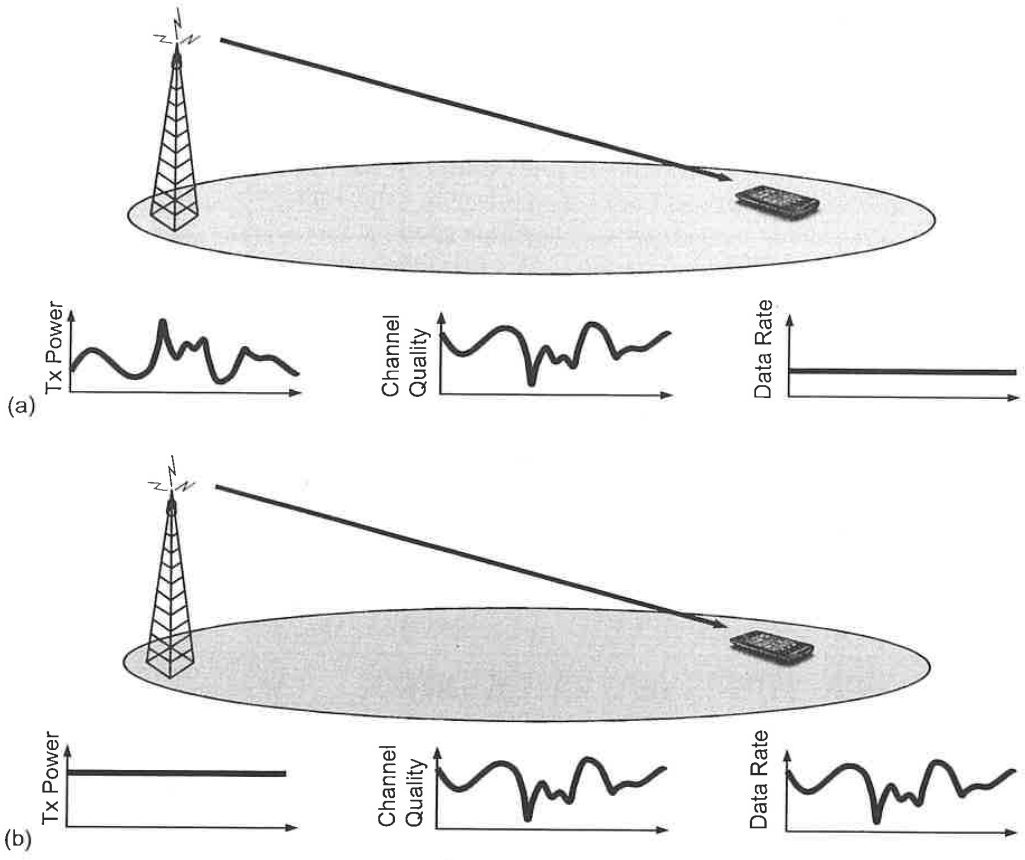


FIGURE 6.1

(a) Power control (top). (b) Rate control (bottom).

variations in the data rate are often not an issue, as long as the average data rate remains constant, assuming averaging over some relatively short time interval. In such cases – that is, when a constant data rate is not required – an alternative to transmit-power control is link adaptation by means of *dynamic rate control*. Rate control does not aim at keeping the instantaneous radio-link data rate constant, regardless of the instantaneous channel conditions. Instead, with rate control, the data rate is dynamically adjusted to compensate for the varying channel conditions. In situations with advantageous channel conditions, the data rate is increased, and vice versa. Thus, rate control maintains the $E_b/N_0 \sim P/R$ at the desired level, not by adjusting the transmission power P, but rather by adjusting the data rate R. This is illustrated in Figure 6.1b.

It can be shown that rate control is more efficient than power control [42,43]. Rate control in principle implies that the power amplifier is always transmitting at full power and therefore efficiently utilized. Power control, on the other hand, results in the power amplifier in most situations not being efficiently utilized as the transmission power is less than its maximum.

In practice, the radio-link data rate is controlled by adjusting the modulation scheme and/or the channel coding rate. In the case of advantageous radio-link conditions, the E_b/N_0 at the receiver is high and the main limitation of the data rate is the bandwidth of the radio link. Hence, in such situations higher-order modulation, for example 16QAM or 64QAM, together with a high code rate, is appropriate as discussed in Chapter 2. Similarly, in the case of poor radio-link conditions, QPSK and low-rate coding is used. For this reason, link adaptation by means of rate control is sometimes also referred to as *Adaptive Modulation and Coding* (AMC).

6.2 CHANNEL-DEPENDENT SCHEDULING

Scheduling controls the allocation of the shared resources among users at each time instant. It is closely related to link adaptation, and often scheduling and link adaptation is seen as one joint function. The scheduling principles, as well as which resources are shared between users, differ depending on the radio-interface characteristics – for example, whether uplink or downlink is considered and whether different users' transmissions are mutually orthogonal or not.

6.2.1 Downlink Scheduling

In the downlink, transmissions to different terminals within a cell are typically mutually orthogonal, implying that, at least in theory, there is no interference between the transmissions (no intra-cell interference). Downlink intra-cell orthogonality can be achieved in the time domain (*Time-Division Multiplexing*, TDM), in the frequency domain (*Frequency-Domain Multiplexing*, FDM), or in the code domain (*Code-Domain Multiplexing*, CDM). In addition, the spatial domain can also be used to separate users, at least in a quasi-orthogonal way, through different antenna arrangements. This is sometimes referred to as *Spatial-Division Multiplexing* (SDM), although in most cases it is used in combination with one or several of the above multiplexing strategies and is not discussed further in this chapter.

For packet data, where the traffic is often very bursty, it can be shown that TDM is preferable from a theoretical point of view [44,45] and is therefore typically the main component in the downlink [46,47]. However, as discussed in Chapter 3, the TDM component is often combined with

sharing of the radio resource also in the frequency domain (FDM) or in the code domain (CDM). For example, in the case of HSPA (see Chapter 19), downlink multiplexing is a combination of TDM and CDM. On the other hand, in the case of LTE, downlink multiplexing is a combination of TDM and FDM. The reasons for sharing the resources not only in the time domain will be elaborated upon later in this section.

When transmissions to multiple users occur in parallel, either by using FDM or CDM, there is also an instantaneous sharing of the total available cell transmit power. In other words, not only are the time/frequency/code resources shared resources, but also the power resource in the base station. In contrast, in the case of sharing only in the time domain there is, by definition, only a single transmission at a time and thus no instantaneous sharing of the total available cell transmit power.

For the purpose of discussion, assume initially a TDM-based downlink with a single user being scheduled at a time. In this case, the utilization of the radio resources is maximized if, at each time instant, all resources are assigned to the user with the best instantaneous channel condition:

- In the case of link adaptation based on power control, this implies that the lowest possible transmit power can be used for a given data rate and thus minimizes the interference to transmissions in other cells for a given link utilization.
- In the case of link adaptation based on rate control, this implies that the highest data rate is achieved for a given transmit power or, in other words, for a given interference to other cells, the highest link utilization is achieved.

However, if applied to the downlink, transmit-power control in combination with TDM scheduling implies that the total available cell transmit power will, in most cases, not be fully utilized. Thus, rate control is generally preferred [42,44,47,48].

The strategy outlined above is an example of channel-dependent scheduling, where the scheduler takes the instantaneous radio-link conditions into account. Scheduling the user with the instantaneously best radio-link conditions is often referred to as *max-C/I* (or *maximum rate*) scheduling. Since the radio conditions for the different radio links within a cell typically vary independently, at each point in time there is almost always a radio link whose channel quality is near its peak (see Figure 6.2). Thus, the channel eventually used for transmission will typically have a high quality and, with rate control, a correspondingly high data rate can be used. This translates into a high system capacity. The gain obtained by transmitting to users with favorable radio-link conditions is commonly known as multi-user diversity; the multi-user diversity gains are larger, the larger the channel variations and the larger the number of users in a cell. Hence, in contrast to the traditional view that fast fading – that is, rapid variations in the radio-link quality – is an undesirable effect that has to be combated, the possibility of channel-dependent scheduling implies that *fading is in fact potentially beneficial and should be exploited.*

Mathematically, the max-C/I (maximum rate) scheduler can be expressed as scheduling user k given by:

$$k = \arg\max_i R_i,$$

where R_i is the instantaneous data rate for user i. Although, from a system capacity perspective, max-C/I scheduling is beneficial, this scheduling principle will not be fair in all situations. If all terminals are, on average, experiencing similar channel conditions and large variations in the instantaneous channel conditions are only due to, for example, fast multi-path fading, all users will experience the

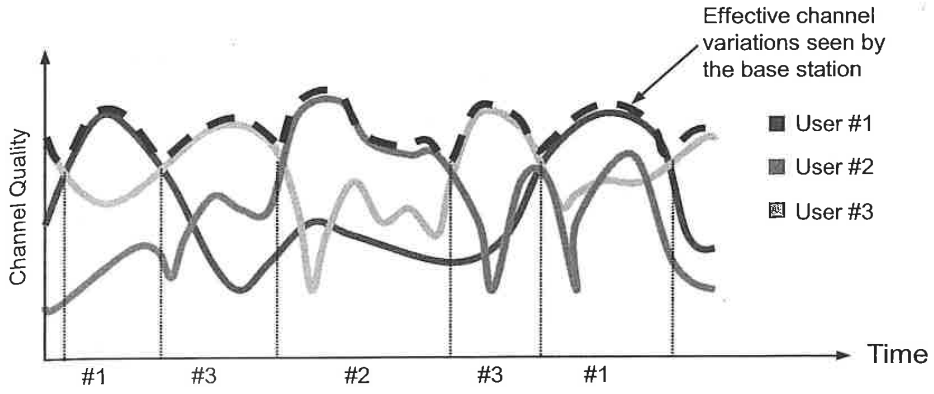


FIGURE 6.2

Channel-dependent scheduling.

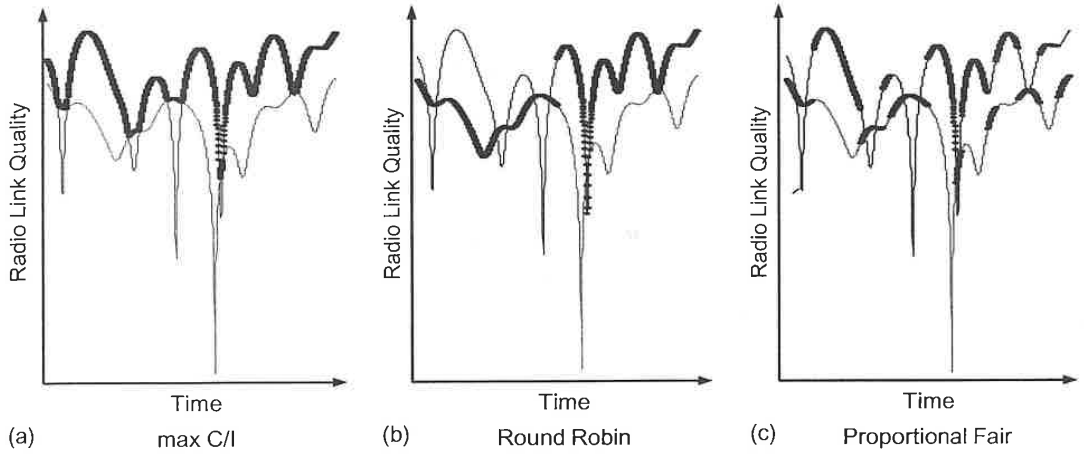


FIGURE 6.3

Examples of three different scheduling behaviors for two users with different average channel quality: (a) max-C/I; (b) round robin; (c) proportional fair. The selected user is shown with bold lines.

same average data rate. Any variations in the instantaneous data rate are rapid and often not even noticeable by the user. However, in practice different terminals will also experience differences in the (short-term) average channel conditions – for example, due to differences in the distance and shadow fading between the base station and the terminal. In this case, the channel conditions experienced by one terminal may, for a relatively long time, be worse than the channel conditions experienced by other terminals. A pure max-C/I-scheduling strategy may then, in essence, "starve" the terminal with the bad channel conditions, and the terminal with bad channel conditions will never be scheduled. This is illustrated in Figure 6.3a, where a max-C/I scheduler is used to schedule between two different users with different average channel quality. The same user is scheduled virtually all the

time. Although resulting in the highest system capacity, this situation is often not acceptable from a quality-of-service point of view.

An alternative to the max-C/I scheduling strategy is so-called *round-robin* scheduling, illustrated in Figure 6.3b. This scheduling strategy lets users take turns in using the shared resources, without taking the instantaneous channel conditions into account. Round-robin scheduling can be seen as fair scheduling in the sense that the same amount of radio resources (the same amount of time) is given to each communication link. However, round-robin scheduling is not fair in the sense of providing the same service quality to all communication links. In that case more radio resources (more time) must be given to communication links with bad channel conditions. Furthermore, as round-robin scheduling does not take the instantaneous channel conditions into account in the scheduling process, it will lead to lower overall system performance but more equal service quality between different communication links, compared to max-C/I scheduling.

Thus, what is needed is a scheduling strategy that is able to utilize the fast channel variations to improve the overall cell throughput while still ensuring the same average user throughput for all users or at least a certain minimum user throughput for all users. When discussing and comparing different scheduling algorithms it is important to distinguish between different types of variations in the service quality:

- Fast variations in the service quality corresponding to, for example, fast multi-path fading and fast variations in the interference level. For many packet-data applications, relatively large short-term variations in service quality are often acceptable or not even noticeable to the user.
- More long-term differences in the service quality between different communication links corresponding to, for example, differences in the distance to the cell site and shadow fading. In many cases there is a need to limit such long-term differences in service quality.

A practical scheduler should thus operate somewhere in between the max-C/I scheduler and the round-robin scheduler – that is, try to utilize fast variations in channel conditions as much as possible while still satisfying some degree of fairness between users.

One example of such a scheduler is the proportional-fair scheduler [49–51], illustrated in Figure 6.3c. In this strategy, the shared resources are assigned to the user with the *relatively* best radio-link conditions – that is, at each time instant, user k is selected for transmission according to:

$$k = \arg\max_i \frac{R_i}{\overline{R_i}},$$

where R_i is the instantaneous data rate for user i and $\overline{R_i}$ is the average data rate for user i. The average is calculated over a certain averaging period T_{PF}. To ensure efficient usage of the short-term channel variations and, at the same time, limit the long-term differences in service quality to an acceptable level, the time constant T_{PF} should be set longer than the time constant for the short-term variations. At the same time, T_{PF} should be sufficiently short so that quality variations within the interval T_{PF} are not strongly noticed by a user. Typically, T_{PF} can be set to be of the order of one second.

In the above discussion, it was assumed that all the radio resources in the downlink were assigned to a single user at a time – that is, scheduling was done purely in the time domain using TDM between users. However, in several situations, TDM is complemented by CDM or FDM. In principle, there are two reasons for not relying solely on TDM in the downlink:

- In the case of insufficient payload – that is, the amount of data to transfer to a user is not sufficiently large to utilize the full channel capacity, and a fraction of resources could be assigned to another user, either through FDM or CDM.
- In the case where channel variations in the frequency domain are exploited through FDM, as discussed further below.

The scheduling strategies in these cases can be seen as generalizations of the schemes discussed for the TDM-only cases above. For example, to handle small payloads, a greedy filling approach can be used, where the scheduled user is selected according to max-C/I (or any other scheduling scheme). Once this user has been assigned resources matching the amount of data awaiting transmission, the second best user according to the scheduling strategy is selected and assigned (a fraction of) the residual resources and so on.

Finally, it should also be noted that the scheduling algorithm typically is a base-station-implementation issue and nothing that is normally specified in any standard. What need to be specified in a standard to support channel-dependent scheduling are channel-quality measurements/reports and the signaling needed for dynamic resource allocation.

6.2.2 Uplink Scheduling

The previous section discussed scheduling from a downlink perspective. However, scheduling is equally applicable to uplink transmissions and to a large extent the same principles can be considered, although there are some differences between the two.

Fundamentally, the uplink power resource is *distributed* among the users, while in the downlink the power resource is *centralized* within the base station. Furthermore, the maximum uplink transmission power of a single terminal is typically significantly lower than the output power of a base station. This has a significant impact on the scheduling strategy. Unlike the downlink, where pure TDMA can often be used, uplink scheduling typically has to rely on sharing in the frequency and/or code domain in addition to the time domain, as a single terminal may not have sufficient power for efficiently utilizing the link capacity.

Similar to the downlink case, channel-dependent scheduling is also beneficial in the uplink case. However, the characteristics of the underlying radio interface, most notably whether the uplink relies on orthogonal or non-orthogonal multiple access and the type of link adaptation scheme used, also have a significant impact on the uplink scheduling strategy.

In the case of a non-orthogonal multiple-access scheme such as CDMA, power control is typically essential for proper operation. As discussed earlier in this chapter, the purpose of power control is to control the received E_b/N_0 such that the received information can be recovered. However, in a non-orthogonal multiple-access setting, power control also serves the purpose of controlling the amount of interference affecting *other* users. This can be expressed as the *maximum tolerable interference level* at the base station in a shared resource. Even if, from a single user's perspective, it would be beneficial to transmit at full power to maximize the data rate, this may not be acceptable from an interference perspective as other terminals in this case may not be able to successfully transfer any data. Thus, with non-orthogonal multiple access, scheduling a terminal when the channel conditions are favorable may not directly translate into a higher data rate as the interference generated to other simultaneously transmitting terminals in the cell must be taken into account. Stated differently, the

received power (and thus the data rate) is, thanks to power control, in principle constant, regardless of the channel conditions at the time of transmission, while the *transmitted* power depends on the channel conditions at the time of transmission. Hence, even though channel-dependent scheduling in this example does not give a direct gain in terms of a higher data rate from the terminal, channel-dependent scheduling will still provide a gain for the system in terms of lower intra-cell interference.

The above discussion on non-orthogonal multiple access was simplified in the sense that no bounds on the terminals' transmission power were assumed. In practice, the transmission power of a terminal is upper bounded, due both to implementation and regulatory reasons, and scheduling a terminal for transmission in favorable channel conditions decreases the probability that the terminal has insufficient power to utilize the channel capacity.

In the case of an orthogonal multiple-access scheme, intra-cell power control is fundamentally not necessary and the benefits of channel-dependent scheduling become more similar to the downlink case. In principle, from an intra-cell perspective, a terminal can transmit at full power and the scheduler assigns a suitable part of the orthogonal resources (in practice a suitable part of the overall bandwidth) to the terminal for transmission. The remaining orthogonal resources can be assigned to other users. However, implementation constraints, for example leakage between the received signals or limited dynamic range in the receiver circuitry, may pose restrictions on the maximum tolerable power difference between the signals from simultaneously transmitting terminals. As a consequence, a certain degree of power control may be necessary, making the situation somewhat similar to the non-orthogonal case.

The discussion on non-orthogonal and orthogonal multiple access mainly considered intra-cell multiple access. However, in many practical systems universal frequency reuse between cells is applied. In this case, the inter-cell multiple access is non-orthogonal, regardless of the intra-cell multiple access, which sets limits on the allowable transmission power from a terminal.

Regardless of whether orthogonal or non-orthogonal multiple access is used, the same basic scheduling principles as for the downlink can be used. A max-C/I scheduler would assign all the uplink resources to the terminal with the best uplink channel conditions. Neglecting any power limitations in the terminal, this would result in the highest capacity (in an isolated cell) [44].

In the case of a non-orthogonal multiple-access scheme, *greedy filling* is one possible scheduling strategy [48]. With greedy filling, the terminal with the best radio conditions is assigned as high a data rate as possible. If the interference level at the receiver is smaller than the maximum tolerable level, the terminal with the second best channel conditions is allowed to transmit as well, continuing with more and more terminals until the maximum tolerable interference level at the receiver is reached. This strategy maximizes the air interface utilization but is achieved at the cost of potentially large differences in data rates between users. In the extreme case, a user at the cell border with poor channel conditions may not be allowed to transmit at all.

Strategies between greedy filling and max-C/I can also be envisioned – for example, different proportional-fair strategies. This can be achieved by including a weighting factor for each user, proportional to the ratio between the instantaneous and average data rates, into the greedy filling algorithm.

The schedulers above all assume knowledge of the instantaneous radio-link conditions, knowledge that can be hard to obtain in the uplink scenario, as discussed in Section 7.2.4. In situations when no information about the uplink radio-link quality is available at the scheduler, *round-robin* scheduling can be used. Similar to the downlink, round-robin implies terminals taking turns in transmitting, thus creating a TDMA-like operation with inter-user orthogonality in the time domain. Although the round-robin scheduler is simple, it is far from the optimal scheduling strategy.

However, as already discussed in Chapter 4, the transmission power in a terminal is limited and therefore additional sharing of the uplink resources in the frequency and/or code domain is required. This also impacts the scheduling decisions. For example, terminals far from the base station typically operate in the power-limited region, in contrast to terminals close to the base stations, which often are in the bandwidth-limited region (for a discussion on power-limited vs. bandwidth-limited operation, see Chapter 2). Thus, for a terminal far from the base station, increasing the bandwidth will not result in an increased data rate and it is better to only assign a small amount of the bandwidth to this terminal and assign the remaining bandwidth to other terminals. On the other hand, for terminals close to the base station, an increase in the assigned bandwidth will provide a higher data rate.

6.2.3 Link Adaptation and Channel-Dependent Scheduling in the Frequency Domain

In the previous section, TDM-based scheduling was assumed and it was explained how, in this case, channel variations in the time domain could be utilized to improve system performance by applying channel-dependent scheduling, especially in combination with dynamic rate control. However, if the scheduler has access to the frequency domain, for example through the use of OFDM transmission, scheduling and link adaptation can also take place in the frequency domain.

Link adaptation in the frequency domain implies that, based on knowledge about the instantaneous channel conditions also in the frequency domain – that is, knowledge about the attenuation as well as the noise/interference level of, in the extreme case, every OFDM subcarrier – the power and/or the data rate of each OFDM carrier can be individually adjusted for optimal utilization.

Similarly, channel-dependent scheduling in the frequency domain implies that, based on knowledge about the instantaneous channel conditions also in the frequency domain, different subcarriers are used for transmission to or from different terminals. The scheduling gains from exploiting variations in the frequency domain are similar to those obtained from time-domain variations. Obviously, in situations where the channel quality varies significantly with the frequency while the channel quality only varies slowly with time, channel-dependent scheduling in the frequency domain can enhance system capacity. An example of such a situation is a wideband indoor system with low mobility, where the quality only varies slowly with time.

6.2.4 Acquiring on Channel-State Information

To select a suitable data rate, in practice a suitable modulation scheme and channel-coding rate, the transmitter needs information about the radio-link channel conditions. Such information is also required for the purpose of channel-dependent scheduling. In the case of a system based on frequency-division duplex (FDD), only the receiver can accurately estimate the radio-link channel conditions.

For the downlink, most systems provide a downlink signal of a predetermined structure, known as the downlink pilot or the downlink reference signal. This reference signal is transmitted from the base station with constant power and can be used by the terminal to estimate the instantaneous downlink channel conditions. Information about the instantaneous downlink conditions can then be reported to the base station.

Basically, what is relevant for the transmitter is an estimate reflecting the channel conditions *at the time of transmission*. Hence, in principle, the terminal could apply a predictor, trying to predict the future channel conditions and report this predicted value to the base station. However, as this would

require specification of prediction algorithms and how they would operate when the terminal is moving at different speeds, most practical systems simply report the measured channel conditions to the base station. This can be seen as a very simple predictor, basically assuming the conditions in the near future will be similar to the current conditions. Thus, the more rapid the time-domain channel variations, the less efficient the link adaptation.

As there inevitably will be a delay between the point in time when the terminal measured the channel conditions and the application of the reported value in the transmitter, channel-dependent scheduling and link adaptation typically operates at its best at low terminal mobility. If the terminal starts to move at high speed, the measurement reports will be outdated once arriving at the base station. In such cases, it is often preferable to perform link adaptation on the long-term average channel quality and rely on hybrid ARQ with soft combining for the rapid adaptation.

For the uplink, estimation of the uplink channel conditions is not as straightforward, as there is typically no reference signal transmitted with constant power from each terminal. Means to estimate uplink channel conditions for LTE are discussed in Chapter 11.

In the case of a system with time-division duplex (TDD), where uplink and downlink communication are time multiplexed within the same frequency band, the instantaneous uplink signal quality attenuation could be estimated from downlink measurements of the terminal, due to the reciprocity of the multi-path fading in the case of TDD. However, it should then be noted that this may not provide full knowledge of the downlink channel conditions. As an example, the interference situations at the terminal and the base station are different also in the case of TDD.

6.2.5 Traffic Behavior and Scheduling

It should be noted that there is little difference between different scheduling algorithms at low system load – that is, when only one or, in some cases, a few users have data waiting for transmission at the base station at each scheduling instant. The differences between different scheduling algorithms are primarily visible at high load. However, not only the load, but also the traffic behavior affects the overall scheduling performance.

As discussed above, channel-dependent scheduling tries to exploit short-term variations in radio quality. Generally, a certain degree of long-term fairness in service quality is desirable, which should be accounted for in the scheduler design. However, since system throughput decreases the more fairness is enforced, a trade-off between fairness and system throughput is necessary. In this trade-off, it is important to take traffic characteristics into account as they have a significant influence on the trade-off between system throughput and service quality.

To illustrate this, consider three different downlink schedulers:

1. *Round-robin (RR) scheduler,* where channel conditions are not taken into account.
2. *Proportional-fair (PF) scheduler,* where short-term channel variations are exploited while maintaining the long-term average user data rate.
3. *Max-C/I scheduler,* where the user with the best instantaneous channel quality in absolute terms is scheduled.

For a full buffer scenario when there is always data available at the base station for all terminals in the cell, a max-C/I scheduler will result in no, or very low, user throughput for users at the cell edge with a low average channel quality. The reason is the fundamental strategy of the max-C/I

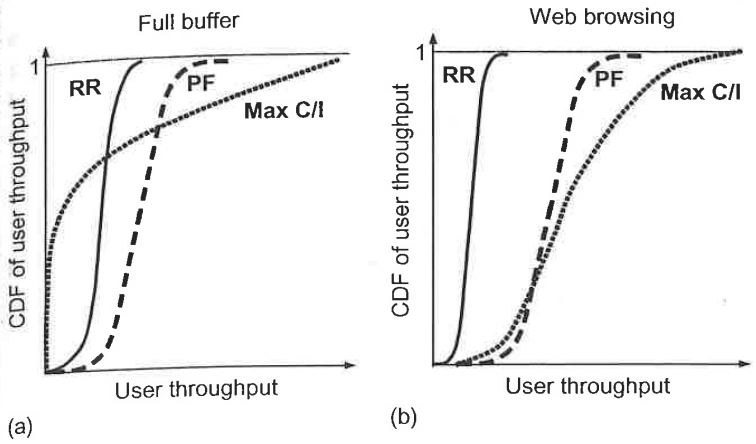


FIGURE 6.4

Illustration of the principle behavior of different scheduling strategies: (a) for full buffers; (b) for web-browsing traffic model.

scheduler – all resources are allocated for transmission to the terminal whose channel conditions support the highest data rate. Only in the rare, not to say unlikely, case of a cell-edge user having better conditions than a cell-center user, for example due to a deep fading dip for the cell-center user, will the cell-edge user be scheduled. A proportional-fair scheduler, on the other hand, will ensure some degree of fairness by selecting the user supporting the highest data rate relative to its average data rate. Hence, users tend to be scheduled on their fading peaks, regardless of the absolute quality. Thus, users on the cell edge will also be scheduled, thereby resulting in some degree of fairness between users.

For a scenario with bursty packet data, the situation is different. In this case, the users' buffers will be finite and in many cases also empty. For example, a web page has a certain size and, after transmitting the page, there is no more data to be transmitted to the terminal in question until the users requests a new page by clicking on a link. In this case, a max-C/I scheduler can still provide a certain degree of fairness. Once the buffer for the user with the highest C/I has been emptied, another user with non-empty buffers will have the highest C/I and be scheduled and so on. This is the reason for the difference between full buffer and web-browsing traffic illustrated in Figure 6.4. The proportional-fair scheduler has similar performance in both scenarios.

Clearly, the degree of fairness introduced by the traffic properties depends heavily on the actual traffic; a design made with certain assumptions may be less desirable in an actual network where the traffic pattern may be different from the assumptions made during the design. Therefore, relying solely on the traffic properties for fairness is not a good strategy, but the discussion above emphasizes the need to design the scheduler not only for the full buffer case.

6.3 ADVANCED RETRANSMISSION SCHEMES

Transmissions over wireless channels are subject to errors, for example due to variations in the received signal quality. To some degree, such variations can be counteracted through link adaptation,

as discussed above. However, receiver noise and unpredictable interference variations cannot be counteracted. Therefore, virtually all wireless communications systems employ some form of *Forward Error Correction* (FEC), tracing its roots to the pioneering work of Claude Shannon in 1948 [11]. There is a rich literature in the area of error-correction coding (see, for example, [52,53] and references therein), and a detailed description is beyond the scope of this book. In short, the basic principle beyond forward error-correction coding is to introduce redundancy in the transmitted signal. This is achieved by adding *parity bits* to the information bits prior to transmission (alternatively, the transmission could consists of parity bits alone, depending on the coding scheme used). The parity bits are computed from the information bits using a method given by the coding structure used. Thus, the number of bits transmitted over the channel is larger than the number of original information bits and a certain amount of *redundancy* has been introduced in the transmitted signal.

Another approach to handle transmissions errors is to use *Automatic Repeat Request* (ARQ). In an ARQ scheme, the receiver uses an error-detection code, typically a Cyclic Redundancy Check (CRC), to detect if the received packet is in error or not. If no error is detected in the received data packet, the received data is declared error-free and the transmitter is notified by sending a positive acknowledgement (ACK). On the other hand, if an error is detected, the receiver discards the received data and notifies the transmitter via a return channel by sending a negative acknowledgement (NAK). In response to an NAK, the transmitter retransmits the same information.

Virtually all modern communication systems, including LTE, employ a combination of forward error-correction coding and ARQ, known as *hybrid ARQ*. Hybrid ARQ uses forward error-correction codes to correct a subset of all errors and relies on error detection to detect uncorrectable errors. Erroneously received packets are discarded and the receiver requests retransmissions of corrupted packets. Thus, it is a combination of FEC and ARQ as described above. Hybrid ARQ was first proposed in [54] and numerous publications on hybrid ARQ have appeared since (see [52] and references therein). Most practical hybrid ARQ schemes are built around a CRC code for error detection and convolutional or Turbo codes for error correction, but in principle any error-detection and error-correction code can be used.

6.4 HYBRID ARQ WITH SOFT COMBINING

The hybrid-ARQ operation described above discards erroneously received packets and requests retransmission. However, despite it not being possible to decode the packet, the received signal still contains information, which is lost by discarding erroneously received packets. This shortcoming is addressed by *hybrid ARQ with soft combining*. In hybrid ARQ with soft combining, the erroneously received packet is stored in a buffer memory and later combined with the retransmission to obtain a single, combined packet that is more reliable than its constituents. Decoding of the error-correction code operates on the combined signal. If the decoding fails (typically a CRC code is used to detect this event), a retransmission is requested.

Retransmission in any hybrid-ARQ scheme must, by definition, represent the same set of information bits as the original transmission. However, the set of coded bits transmitted in each retransmission may be selected differently as long as they represent the same set of information bits. Hybrid ARQ with soft combining is therefore usually categorized into *Chase combining* and *incremental redundancy*, depending on whether the retransmitted bits are required to be identical to the original transmission or not.

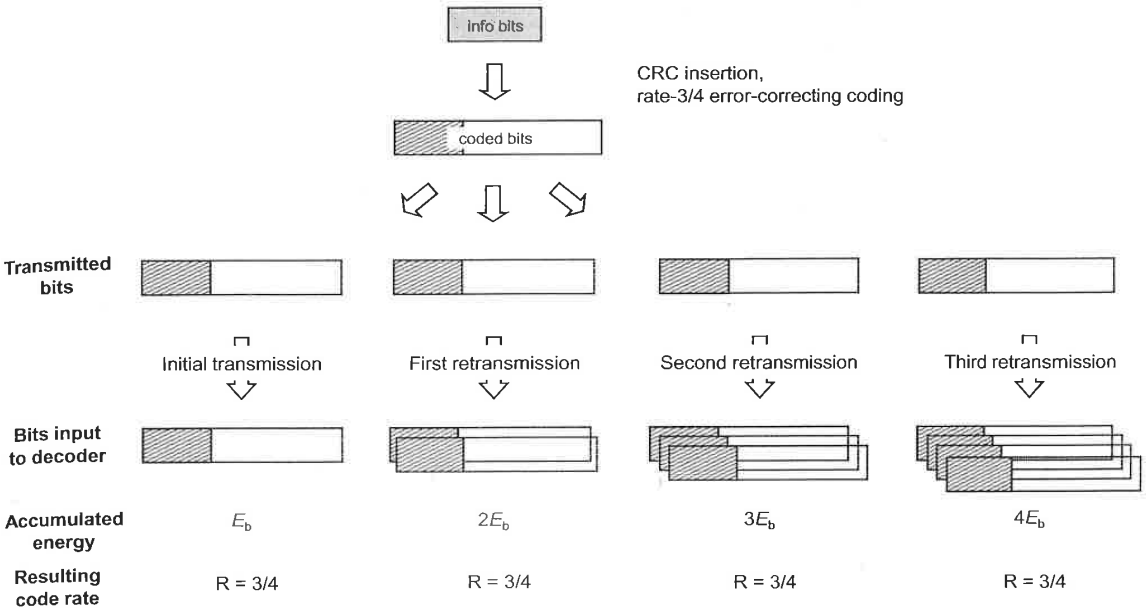


FIGURE 6.5

Example of Chase combining.

Chase combining, where the retransmissions consist of the same set of coded bits as the original transmission, was first proposed in [55]. After each retransmission, the receiver uses maximum-ratio combination to combine each received channel bit with any previous transmissions of the same bit, and the combined signal is fed to the decoder. As each retransmission is an identical copy of the original transmission, retransmissions with Chase combining can be seen as additional repetition coding. Therefore, as no new redundancy is transmitted, Chase combining does not give any additional coding gain but only increases the accumulated received E_b/N_0 for each retransmission (Figure 6.5).

Several variants of Chase combining exist. For example, only a subset of the bits transmitted in the original transmission might be retransmitted, so-called partial Chase combining. Furthermore, although combination is often done after demodulation but before channel decoding, combination can also be carried out at the modulation symbol level before demodulation, as long as the modulation scheme is unchanged between transmission and retransmission.

With *Incremental Redundancy* (IR), each retransmission does not have to be identical to the original transmission. Instead, *multiple sets* of coded bits are generated, each representing the same set of information bits [56,57]. Whenever a retransmission is required, the retransmission typically uses a different set of coded bits than the previous transmission. The receiver combines the retransmission with previous transmission attempts of the same packet. As the retransmission may contain additional parity bits not included in the previous transmission attempts, the resulting code rate is generally lowered by a retransmission. Furthermore, each retransmission does not necessarily have to consist of the same number of coded bits as the original and, in general, the modulation scheme can also be different for different

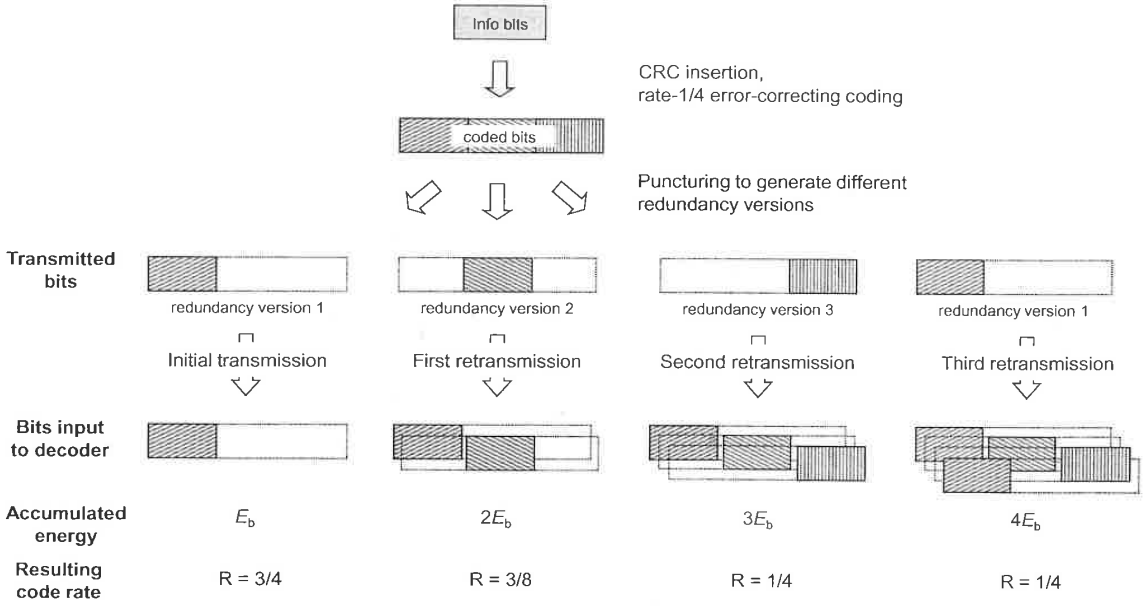


FIGURE 6.6

Example of incremental redundancy.

retransmissions. Hence, incremental redundancy can be seen as a generalization of Chase combining or, stated differently, Chase combining is a special case of incremental redundancy.

Typically, incremental redundancy is based on a low-rate code and the different redundancy versions are generated by puncturing the output of the encoder. In the first transmission only a limited number of coded bits are transmitted, effectively leading to a high-rate code. In the retransmissions, additional coded bits are transmitted. As an example, assume a basic rate-1/4 code. In the first transmission, only every third coded bit is transmitted, effectively giving a rate-3/4 code, as illustrated in Figure 6.6. In the case of a decoding error and a subsequent request for a retransmission, additional bits are transmitted, effectively leading to a rate-3/8 code. After a second retransmission the code rate is 1/4. In the case of more than two retransmissions, transmitted coded bits would already be repeated. In addition to a gain in accumulated received E_b/N_0, incremental redundancy also results in a coding gain for each retransmission. The gain with IR compared to Chase is larger for high initial code rates, while at lower initial coding rates Chase combining is almost as good as IR [58]. Furthermore, as shown in [59], the performance gain of incremental redundancy compared to Chase combining can also depend on the relative power difference between the transmission attempts.

With incremental redundancy, the code used for the first transmission should provide good performance not only when used alone, but also when used in combination with the code for the second transmission. The same holds for subsequent retransmissions. Thus, as the different redundancy versions typically are generated through puncturing of a low-rate mother code, the puncturing patterns should be defined such that all the code bits used by a high-rate code should also be part of any lower-rate codes. In other words, the resulting code rate R_i after transmission attempt i, consisting of

the coded bits from redundancy versions RV_k, $k = 1, \ldots, i$, should have similar performance as a good code designed directly for rate R_i. Examples of this for convolutional codes are the so-called rate-compatible convolutional codes [60].

In the discussion so far, it has been assumed that the receiver has received all the previously transmitted redundancy versions. If all redundancy versions provide the same amount of information about the data packet, the order of the redundancy versions is not critical. However, for some code structures, not all redundancy versions are of equal importance. One example here is Turbo codes, where the systematic bits are of higher importance than the parity bits. Hence, the initial transmission should at least include all the systematic bits and some parity bits. In the retransmission(s), parity bits not in the initial transmission can be included. However, if the initial transmission was received with poor quality or not at all, a retransmission with only parity bits is not appropriate as a retransmission of (some of) the systematic bits provides better performance. Incremental redundancy with Turbo codes can therefore benefit from multiple levels of feedback, for example by using two different negative acknowledgements – NAK to request additional parity bits and LOST to request a retransmission of the systematic bits. In general, the problem of determining the amount of systematic and parity bits in a retransmission based on the signal quality of previous transmission attempts is non-trivial.

Hybrid ARQ with soft combining, regardless of whether Chase or incremental redundancy is used, leads to an implicit reduction of the data rate by means of retransmissions and can thus be seen as *implicit* link adaptation. However, in contrast to link adaptation based on explicit estimates of the instantaneous channel conditions, hybrid ARQ with soft combining implicitly adjusts the coding rate based on the result of the decoding. In terms of overall throughput this kind of implicit link adaptation can be superior to explicit link adaptation, as additional redundancy is only added *when needed* – that is, when previous higher-rate transmissions were not possible to decode correctly. Furthermore, as it does not try to predict any channel variations, it works equally well, regardless of the speed at which the terminal is moving. Since implicit link adaptation can provide a gain in system throughput, a valid question is why explicit link adaptation is necessary at all. One major reason for having explicit link adaptation is the reduced delay. Although relying on implicit link adaptation alone is sufficient from a system throughput perspective, the end-user service quality may not be acceptable from a delay perspective.

CHAPTER

LTE Radio Access: An Overview

7

The work on LTE was initiated in late 2004 with the overall aim of providing a new radio-access technology focusing on packet-switched data only. The first phase of the 3GPP work on LTE was to define a set of performance and capability targets for LTE [10]. This included targets on peak data rates, user/system throughput, spectral efficiency, and control/user-plane latency. In addition, requirements were also set on spectrum flexibility, as well as on interaction/compatibility with other 3GPP radio-access technologies (GSM, WCDMA/HSPA, and TD-SCDMA).

Once the targets were set, 3GPP studies on the feasibility of different technical solutions considered for LTE were followed by development of the detailed specifications. The first release of the LTE specifications, release 8, was completed in the spring of 2008 and commercial network operation began in late 2009. Release 8 has then been followed by additional LTE releases, introducing additional functionality and capabilities in different areas, as illustrated in Figure 7.1.

In parallel to the development of LTE, there has also been an evolution of the overall 3GPP network architecture, termed *System Architecture Evolution* (SAE), including both the radio-access network and the core network. Requirements were also set on the architecture evolution, leading to a new flat radio-access-network architecture with a single type of node, the *eNodeB*[1], as well as a new core-network architecture. An excellent description of the LTE-associated core-network architecture, the *Evolved Packet Core* (EPC), can be found in [9].

The remaining part of this chapter provides an overview of LTE up to and including release 10. The most important technologies used by LTE release 8 – including transmission schemes, scheduling, multi-antenna support, and spectrum flexibility – are covered, as well as the additional features introduced in LTE releases 9 and 10. The chapter can either be read on its own to get a high-level overview of LTE, or as an introduction to the subsequent chapters.

The following chapters, Chapters 8–18, then contain a detailed description of the LTE radio-access technology. Chapter 8 provides an overview of the LTE protocol structure, including RLC, MAC, and the physical layer, explaining the logical and physical channels, and the related data flow. The time–frequency structure on which LTE is based is covered in Chapter 9, followed by a detailed description of the physical layer for downlink and uplink transmission in Chapters 10 and 11 respectively. Chapter 12 contains a description of the retransmission mechanisms used in LTE, followed by a discussion on power control, scheduling, and interference management in Chapter 13. Access procedures, necessary for a terminal to connect to the network, are the topic of Chapter 14. Chapter 15 covers the multicast/broadcast functionality of LTE and Chapter 16 describes relaying operation. Chapter 17 addresses how radio-frequency (RF) requirements are defined in LTE, taking into account the spectrum flexibility. Finally, Chapter 18 contains an assessment of the system performance of LTE.

[1] eNodeB is a 3GPP term that can roughly be seen as being equivalent to a base station, see further in Chapter 8.

4G LTE/LTE-Advanced for Mobile Broadband.
© 2011 Erik Dahlman, Stefan Parkvall & Johan Sköld. Published by Elsevier Ltd. All rights reserved.

Samsung Exhibit 1019, Page 69 of 195

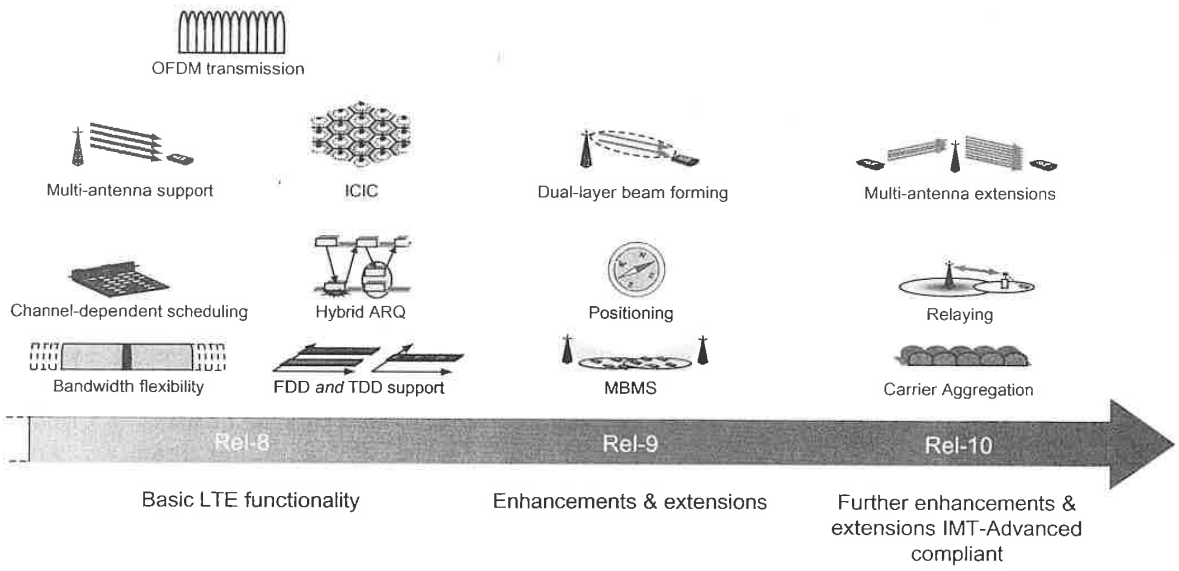


FIGURE 7.1

LTE and its evolution.

7.1 BASIC PRINCIPLES

Building upon the basic technologies described in the previous chapters, the main principles behind LTE will be described in the following.

7.1.1 Transmission Scheme

The LTE downlink transmission scheme is based on conventional OFDM. As discussed in Chapter 3, OFDM is an attractive transmission scheme for several reasons. Due to the relatively long OFDM symbol time in combination with a cyclic prefix, OFDM provides a high degree of robustness against channel frequency selectivity. Although signal corruption due to a frequency-selective channel can, in principle, be handled by equalization at the receiver side, the complexity of such equalization starts to become unattractively high for implementation in a terminal at larger bandwidths and especially in combination with advanced multi-antenna transmission schemes such as spatial multiplexing, thereby making OFDM an attractive choice for LTE for which a wide bandwidth and support for advanced multi-antenna transmission were key requirements.

OFDM also provides some additional benefits relevant for LTE:

- OFDM provides access to the frequency domain, thereby enabling an additional degree of freedom to the channel-dependent scheduler compared to time-domain-only scheduling used in major 3G systems.
- Flexible transmission bandwidth to support operation in spectrum allocations of different size is straightforward with OFDM, at least from a baseband perspective, by varying the number of OFDM subcarriers used for transmission. Note, however, that support of a flexible transmission

bandwidth also requires flexible RF filtering, etc., for which the exact transmission scheme is irrelevant. Nevertheless, maintaining the same baseband-processing structure, regardless of the bandwidth, eases terminal development and implementation.

- Broadcast/multicast transmission, where the same information is transmitted from multiple base stations, is straightforward with OFDM as already described in Chapter 3 and elaborated upon from an LTE perspective in Chapter 15.

Also, the LTE uplink is based on OFDM transmission. However, different means are taken to reduce the *cubic metric* of the uplink transmission, thereby enabling higher terminal power-amplifier efficiency. More specifically, for uplink data transmission, the OFDM modulator is preceded by a DFT precoder, leading to *DFT-spread OFDM* (DFTS-OFDM) as described in Chapter 4. Often, the term DFTS-OFDM is used to describe the LTE uplink transmission scheme in general. However, it should be understood that DFTS-OFDM is only applicable to uplink data transmission. Other means are used to achieve a low cubic metric for other types of uplink transmissions. Thus, the LTE uplink transmission scheme should be described as OFDM with different techniques, including DFT precoding for data transmission, being used to reduce the cubic metric of the transmitted signal.

The use of DFTS-OFDM on the LTE uplink allows for orthogonal separation of uplink transmissions also in the frequency domain. Orthogonal separation is in many cases beneficial as it avoids interference between uplink transmissions from different terminals within the cell (*intra-cell interference*). As discussed in Chapter 4, allocating a very large instantaneous bandwidth for transmission from a single terminal is not an efficient strategy in situations where the data rate is mainly limited by the available terminal transmit power rather than the bandwidth. In such situations, a terminal can instead be allocated only a part of the total available bandwidth and other terminals within the cell can be scheduled to transmit in parallel on the remaining part of the spectrum. In other words, the LTE uplink transmission scheme allows for both *time division* (TDMA) and *frequency division* (FDMA) between users.

7.1.2 Channel-Dependent Scheduling and Rate Adaptation

At the core of the LTE transmission scheme is the use of *shared-channel transmission* with the overall time–frequency resource dynamically shared between users. The use of shared-channel transmission is well matched to the rapidly varying resource requirements posed by packet-data communication and also enables several of the other key technologies on which LTE is based.

The scheduler controls, for each time instant, to which users the different parts of the shared resource should be assigned. The scheduler also determines the data rate to be used for each transmission. Thus, *rate adaptation* can be seen as a part of the scheduling functionality. The scheduler is thus a key element and to a large extent determines the overall system performance, especially in a highly loaded network. Both downlink and uplink transmissions are subject to tight scheduling in LTE. From Chapter 6 it is well known that a substantial gain in system capacity can be achieved if the channel conditions are taken into account in the scheduling decision, so-called *channel-dependent scheduling*. Due to the use of OFDM in both the downlink and uplink transmission directions, the scheduler has access to both the time and frequency domains. In other words, the scheduler can, for each time instant and frequency region, select the user with the best channel conditions, as illustrated in Figure 7.2.

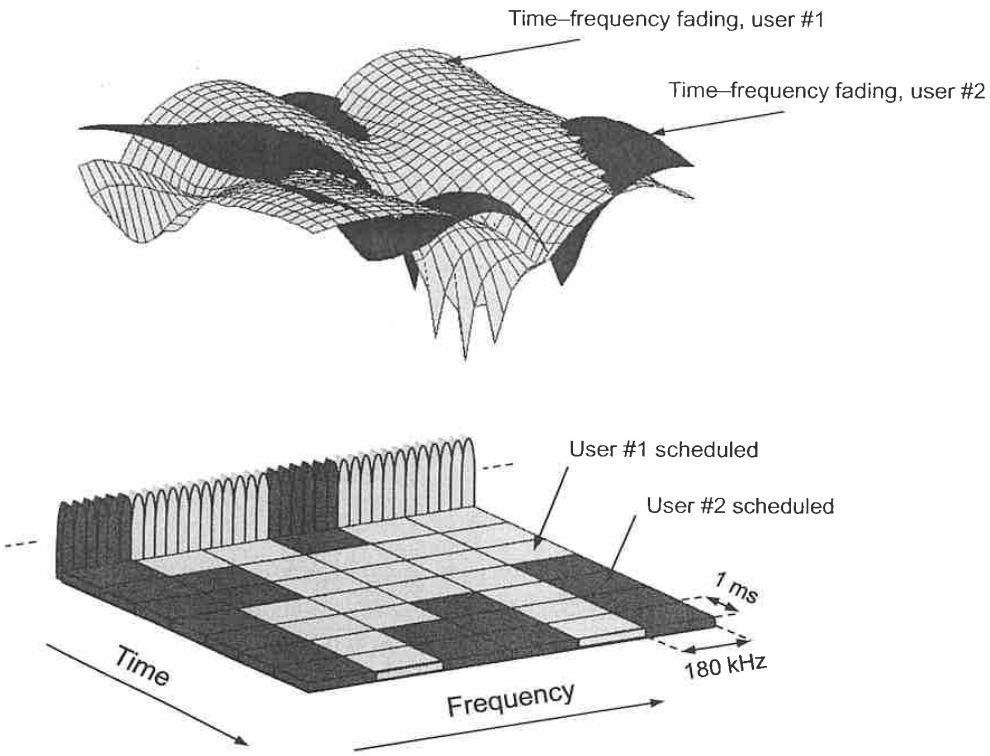


FIGURE 7.2

Downlink channel-dependent scheduling in the time and frequency domains.

The possibility of channel-dependent scheduling in the frequency domain is particularly useful at low terminal speeds – in other words, when the channel is varying slowly in time. As discussed in Chapter 6, channel-dependent scheduling relies on channel-quality variations between users to obtain a gain in system capacity. For delay-sensitive services, a time-domain-only scheduler may, due to the delay constraints, be forced to schedule a particular user, despite the channel quality not being at its peak. In such situations, exploiting channel-quality variations also in the frequency domain will help to improve the overall performance of the system. For LTE, scheduling decisions can be taken as often as once every 1 ms and the granularity in the frequency domain is 180 kHz. This allows for relatively rapid channel variations in both the time and frequency domains to be tracked and utilized by the scheduler.

To support downlink scheduling, a terminal may provide the network with *channel-state* reports indicating the instantaneous downlink channel quality in both the time and frequency domains. The channel state is typically obtained by measuring on *reference signals* transmitted in the downlink. Based on the channel-state reports, also referred to as *channel-state information* (CSI), the downlink scheduler can assign resources for downlink transmission to different terminals, taking the channel quality into account in the scheduling decision. In principle, a scheduled terminal can be assigned an arbitrary combination of 180 kHz wide *resource blocks* in each 1 ms scheduling interval.

As already mentioned, the LTE uplink is based on orthogonal separation of different uplink transmissions and it is the task of the uplink scheduler to assign resources in both the time and frequency domains to different terminals. Scheduling decisions, taken once per 1 ms, control what set of terminals are allowed to transmit within a cell during a given time interval and, for each terminal, on what frequency resources the transmission is to take place and what transmission parameters, including the data rate, to use.

Channel conditions can also be taken into account in the uplink scheduling process, similar to the downlink scheduling. However, as will be discussed in more detail in subsequent chapters, obtaining information about the uplink channel conditions may not be feasible or desirable in all situations. Therefore, different means to obtain *uplink diversity* are important as a complement in situations where uplink channel-dependent scheduling is not suitable.

7.1.3 Inter-Cell Interference Coordination

LTE is designed to operate with a one-cell frequency reuse, implying that the same time–frequency resources can be used in neighboring cells. In particular, the basic control channels are designed to operate properly also with the relatively low signal-to-interference ratio that may be experienced in a reuse-one deployment.

From an overall system-efficiency point-of-view, having access to the entire available spectrum in each cell and operating with one-cell reuse is always beneficial. However, it may also lead to relatively large variations in the signal-to-interference ratio, and thus also in the achievable data rates, over the cell area with potentially only relatively low data rates being available at the cell border. Thus, system performance, and especially the cell-edge user quality, can be further enhanced by allowing for some coordination in the scheduling between cells. The basic aim of such *inter-cell interference coordination* (ICIC) is to, if possible, avoid scheduling transmissions to/from terminals at the cell border simultaneously in neighboring cells, thereby avoiding the worst-case interference situations.

To support such interference coordination, the LTE specification includes several messages that can be communicated between eNodeBs using the so-called *X2 interface*, see Chapter 8. These messages provide information about the interference situation and/or scheduling strategies of the eNodeB issuing the message and can be used by an eNodeB receiving the message as input to its scheduling process. LTE interference coordination is discussed in more detail in Chapter 13.

An even more complicated interference situation may occur in so-called heterogeneous network deployments consisting of overlapping cell layers with large differences in the cell output power. This will be briefly discussed in Section 7.3.4 as part of the discussion on LTE release-10 features. The interference handling in heterogeneous network deployments will then be discussed in more detail in Chapter 13.

7.1.4 Hybrid ARQ with Soft Combining

Fast hybrid ARQ with soft combining is used in LTE to allow the terminal to rapidly request retransmissions of erroneously received transport blocks and to provide a tool for implicit rate adaptation. Retransmissions can be rapidly requested after each packet transmission, thereby minimizing the impact on end-user performance from erroneously received packets. Incremental redundancy is used as the soft combining strategy and the receiver buffers the soft bits to be able to perform soft combining between transmission attempts.

7.1.5 Multi-Antenna Transmission

Already from its first release, LTE included support for different multi-antenna transmission techniques as an integral part of the radio-interface specifications. In many respects, the use of multiple antennas is the key technology to reach many of the aggressive LTE performance targets. As discussed in Chapter 5, multiple antennas can be used in different ways for different purposes:

- Multiple receive antennas can be used for receive diversity. For uplink transmissions, this has been used in many cellular systems for several years. However, as dual receive antennas are the baseline for all LTE terminals, the downlink performance is also improved. The simplest way of using multiple receive antennas is classical receive diversity to collect additional energy and suppress fading, but additional gains can be achieved in interference-limited scenarios if the antennas are used not only to provide diversity, but also to suppress interference, as discussed in Chapter 5.
- Multiple transmit antennas at the base station can be used for transmit diversity and different types of beam-forming. The main goal of beam-forming is to improve the received SINR and, eventually, improve system capacity and coverage.
- *Spatial multiplexing*, sometimes referred to as MIMO, using multiple antennas at both the transmitter and receiver is supported by LTE. Spatial multiplexing results in an increased data rate, channel conditions permitting, in bandwidth-limited scenarios by creating several parallel "channels", as described in Chapter 5. Alternatively, by combining the spatial properties with the appropriate interference-suppressing receiver processing, multiple terminals can transmit on the same time–frequency resource in order to improve the overall cell capacity. This is sometimes referred to as *multi-user MIMO*.

In general, the different multi-antenna techniques are beneficial in different scenarios. As an example, at relatively low SINR, such as at high load or at the cell edge, spatial multiplexing provides relatively limited benefits. Instead, in such scenarios multiple antennas at the transmitter side should be used to raise the SINR by means of beam-forming. On the other hand, in scenarios where there already is a relatively high SINR, for example in small cells, raising the signal quality further provides relatively minor gains as the achievable data rates are then mainly bandwidth limited rather than SINR limited. In such scenarios, spatial multiplexing should instead be used to fully exploit the good channel conditions. The multi-antenna scheme used is under control of the base station, which therefore can select a suitable scheme for each transmission.

Up to four layers can be spatially multiplexed in release 8. Later releases further enhance the multi-antenna support, as described later.

7.1.6 Spectrum Flexibility

A high degree of spectrum flexibility is one of the main characteristics of the LTE radio-access technology. The aim of this spectrum flexibility is to allow for the deployment of LTE radio access in difference frequency bands with different characteristics, including different duplex arrangements and different sizes of the available spectrum. Chapter 17 outlines further details of how spectrum flexibility is achieved in LTE.

7.1.6.1 Flexibility in Duplex Arrangement

One important part of the LTE requirements in terms of spectrum flexibility is the possibility to deploy LTE-based radio access in both paired *and* unpaired spectrum. Therefore, LTE supports both

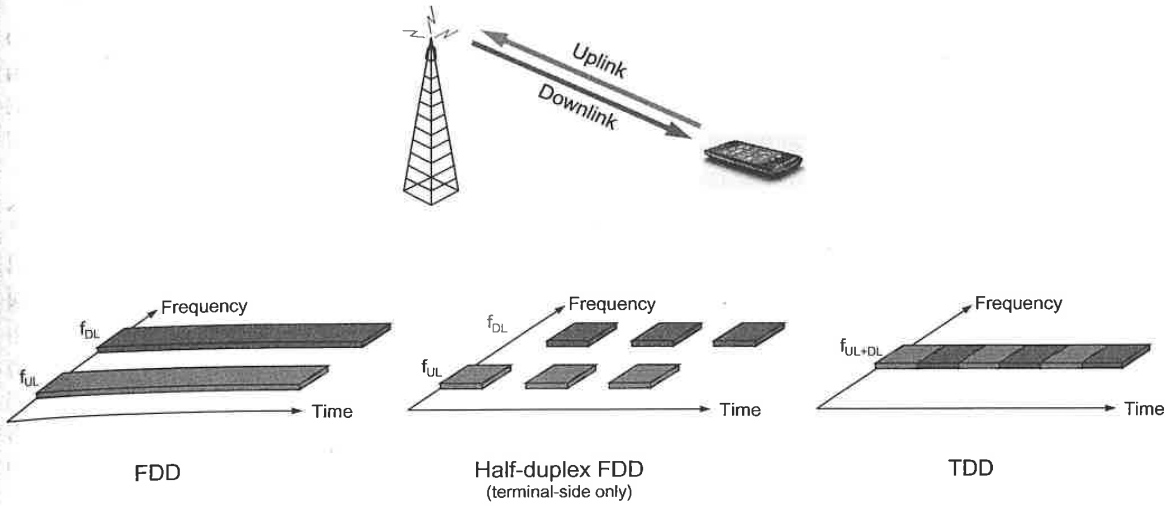


FIGURE 7.3

Frequency- and time-division duplex.

frequency- and time-division-based duplex arrangements. *Frequency-Division Duplex* (FDD), as illustrated on the left in Figure 7.3, implies that downlink and uplink transmission take place in different, sufficiently separated, frequency bands. *Time-Division Duplex* (TDD), as illustrated on the right in Figure 7.3, implies that downlink and uplink transmission take place in different, non-overlapping time slots. Thus, TDD can operate in unpaired spectrum, whereas FDD requires paired spectrum. The required flexibility and resulting requirements to support LTE operation in different paired and unpaired frequency arrangements are further discussed in Chapter 17.

Operation in both paired and unpaired spectrum has been supported by 3GPP radio-access technologies even before the introduction of LTE by means of FDD-based WCDMA/HSPA in combination with TDD-based TD-SCDMA radio. However, this was then achieved by means of, at least in the details, relatively different radio-access technologies leading to additional effort and complexity when developing and implementing dual-mode terminals capable of both FDD and TDD operation. LTE, on the other hand, supports both FDD and TDD *within a single radio-access technology*, leading to a minimum of deviation between FDD and TDD for LTE-based radio access. As a consequence of this, the overview of the LTE radio access provided in the following chapters is, to a large extent, valid for both FDD and TDD. In the case of differences between FDD and TDD, these differences will be explicitly indicated. Furthermore, the TDD mode, also known as TD-LTE, is designed with coexistence between TD-LTE and TD-SCDMA in mind to simplify a gradual migration from TD-SCDMA to TD-LTE.

LTE also supports *half-duplex* FDD at the terminal (illustrated in the middle of Figure 7.3). In half-duplex FDD, transmission and reception *at a specific terminal* are separated in both frequency and time. The base station still uses full-duplex FDD as it simultaneously may schedule *different* terminals in uplink and downlink; this is similar to, for example, GSM operation. The main benefit with half-duplex FDD is the reduced terminal complexity as no duplex filter is needed in the terminal. This is especially beneficial in the case of multi-band terminals which otherwise would need multiple sets of duplex filters.

7.1.6.2 Bandwidth Flexibility

An important characteristic of LTE is the possibility for different transmission bandwidths on both downlink and uplink. The main reason for this is that the amount of spectrum available for LTE deployment may vary significantly between different frequency bands and also depending on the exact situation of the operator. Furthermore, the possibility of operating in different spectrum allocations gives the possibility for gradual migration of spectrum from other radio-access technologies to LTE.

LTE supports operation in a wide range of spectrum allocations, achieved by a flexible transmission bandwidth being part of the LTE specifications. To efficiently support very high data rates when spectrum is available, a wide transmission bandwidth is necessary, as discussed in Chapter 2. However, a sufficiently large amount of spectrum may not always be available, either due to the band of operation or due to a gradual migration from another radio-access technology, in which case LTE can be operated with a more narrow transmission bandwidth. Obviously, in such cases, the maximum achievable data rates will be reduced accordingly. As discussed below, the spectrum flexibility is further improved in later releases of LTE.

The LTE physical-layer specifications [61–64] are bandwidth agnostic and do not make any particular assumption on the supported transmission bandwidths beyond a minimum value. As will be seen in the following, the basic radio-access specification, including the physical-layer and protocol specifications, allows for any transmission bandwidth ranging from roughly 1 MHz up to around 20 MHz. At the same time, at an initial stage, radio-frequency requirements are only specified for a limited subset of transmission bandwidths, corresponding to what is predicted to be relevant spectrum-allocation sizes and relevant migration scenarios. Thus, in practice the LTE radio-access technology supports a limited set of transmission bandwidths, but additional transmission bandwidths can easily be introduced by updating only the RF specifications.

7.2 LTE RELEASE 9

After completing the first release of LTE, work continued in 3GPP with introducing additional functionality in the second release of the LTE specifications, release 9. The main enhancements seen in release 9, completed in late 2009, were support for multicast transmission, support for network-assisted positioning services, and enhancements to beam-forming in the downlink.

7.2.1 Multicast and Broadcast Support

Multi-cell broadcast implies transmission of the same information from multiple cells, as described in Chapter 3. By exploiting this at the terminal, effectively using signal power from multiple cell sites at the detection, a substantial improvement in coverage (or higher broadcast data rates) can be achieved. By transmitting not only identical signals from multiple cell sites (with identical coding and modulation), but also synchronizing the transmission timing between the cells, the signal at the terminal will appear exactly as a signal transmitted from a single cell site and subject to multi-path propagation. Due to the OFDM robustness to multi-path propagation, such multi-cell transmission, in 3GPP also referred to as *Multicast/Broadcast Single-Frequency Network* (MBSFN) transmission, will then not only improve the received signal strength, but also eliminate the inter-cell interference as described in

Chapter 3. Thus, with OFDM, multi-cell broadcast/multicast throughput may eventually be limited by noise only and can then, in the case of small cells, reach extremely high values.

It should be noted that the use of MBSFN transmission for multi-cell broadcast/multicast assumes the use of tight synchronization and time alignment of the signals transmitted from different cell sites.

7.2.2 Positioning

Positioning, as the name implies, refers to functionality in the radio-access network to determine the location of individual terminals. Determining the position of a terminal can, in principle, be done by including a GPS receiver in the terminal. Although this is a quite common feature, not all terminals include the necessary GSP receiver and there may also be cases when the GPS service is not available. LTE release 9 therefore introduces positioning support inherent in the radio-access network. By measuring on special reference signals transmitted regularly from different cell sites, the location of the terminal can be determined.

7.2.3 Dual-Layer Beam-Forming

Release 9 enhances the support for combining spatial multiplexing with beam-forming. Although the combination of beam-forming and spatial multiplexing was already possible in release 8, this was then restricted to so-called *codebook-based precoding* (see Chapter 10). In release 9, spatial multiplexing can be combined with so-called *non-codebook-based precoding*, thereby improving the flexibility in deploying various multi-antenna schemes. The enhancements can also be seen as the basis for a further improvement in the area in LTE release10.

7.3 LTE RELEASE 10 AND IMT-ADVANCED

As described in Chapter 1, *IMT-Advanced* is the term used by the ITU for radio-access technologies beyond IMT-2000. As a first step in defining IMT-Advanced, the ITU defined a set of requirements that any IMT-Advanced compliant technology should fulfill [65]. Examples of these requirements are support for at least 40 MHz bandwidth, peak spectral efficiencies of 15 bit/s/Hz in downlink and 6.75 bit/s/Hz in uplink (corresponding to peak rates of at least 600 and 270 Mbit/s respectively), and control and user plane latency of less than 100 and 10 ms respectively.

One of the main targets of LTE release 10 was to ensure that the LTE radio-access technology would be fully compliant with the IMT-Advanced requirements, thus the name *LTE-Advanced* is often used for LTE release 10. However, in addition to the ITU requirements, 3GPP also defined its own targets and requirements [66] for LTE release 10 (LTE-Advanced). These targets/requirements extended the ITU requirements both in terms of being more aggressive as well as including additional requirements. One important requirement was *backwards compatibility*. Essentially this means that an earlier-release LTE terminal should always be able to access a carrier supporting LTE release-10 functionality, although obviously not being able to utilize all the release-10 features of that carrier.

LTE release 10 was completed in late 2010 and enhances LTE spectrum flexibility through *carrier aggregation*, further extends multi-antenna transmission, introduces support for *relaying*, and provides improvements in the area of inter-cell interference coordination in *heterogeneous network deployments*.

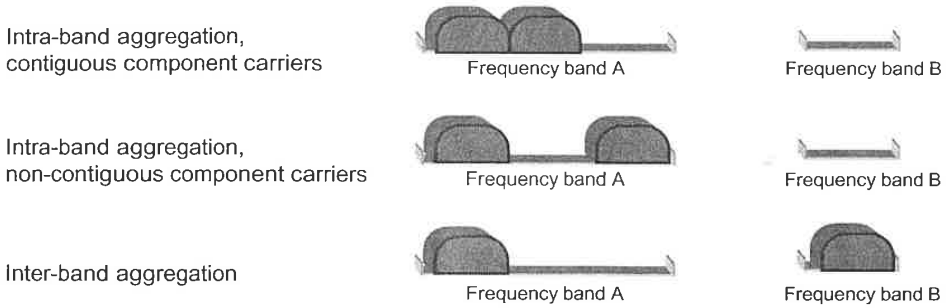


FIGURE 7.4

Carrier aggregation.

7.3.1 Carrier Aggregation

As mentioned earlier, the first release of LTE already provided extensive support for deployment in spectrum allocations of various characteristics, with bandwidths ranging from roughly 1 MHz up to 20 MHz in both paired and unpaired bands. In LTE release 10 the transmission bandwidth can be further extended by means of so-called *carrier aggregation* (CA), where multiple *component carriers* are aggregated and jointly used for transmission to/from a single terminal. Up to five component carriers, possibly each of different bandwidth, can be aggregated, allowing for transmission bandwidths up to 100 MHz. Backwards compatibility is catered for as each component carrier uses the release-8 structure. Hence, to a release-8/9 terminal each component carrier will appear as an LTE release-8 carrier, while a carrier-aggregation-capable terminal can exploit the total aggregated bandwidth, enabling higher data rates. In the general case, a different number of component carriers can be aggregated for the downlink and uplink.

Component carriers do not have to be contiguous in frequency, which enables exploitation of *fragmented spectrum*; operators with a fragmented spectrum can provide high-data-rate services based on the availability of a wide overall bandwidth even though they do not posses a single wideband spectrum allocation. From a baseband perspective, there is no difference between the cases in Figure 7.4 and they are all supported by LTE release 10. However, the RF-implementation complexity is vastly different with the first case being the least complex. Thus, although spectrum aggregation is supported by the basic specifications, the actual release-10 RF requirements will be strongly constrained, including specification of only a limited number of aggregation scenarios and including support of inter-band aggregation only for the most advanced terminals, but excluding non-contiguous intra-band aggregation.

7.3.2 Extended Multi-Antenna Transmission

In release 10, downlink spatial multiplexing is expanded to support up to eight transmission layers. Along with this, and as discussed in more detail in Chapter 10 an enhanced *reference-signal structure* is introduced to improve the support of various beam-forming solutions. This can be seen as an extension of the release-9 dual-layer beam-forming to support up to eight antenna ports and eight corresponding layers. Together with the support for carrier aggregation, this enables downlink data rates up to 3 Gbit/s or 30 bit/s/Hz.

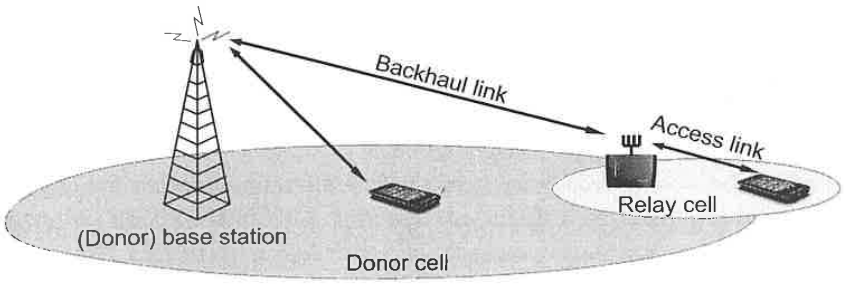


FIGURE 7.5

Example of relaying.

Uplink spatial multiplexing of up to four layers is also part of release 10. It consists of a codebook-based scheme under control of the base station, which means that the structure can also be used for uplink transmit-side beam-forming. Together with the possibility for uplink carrier aggregations, this allows for uplink data rates up to 1.5 Gbit/s or 15 bit/s/Hz.

7.3.3 Relaying

Relaying implies that the terminal communicates with the network via a *relay node* that is *wirelessly connected* to a *donor cell* using the LTE radio-interface technology (Figure 7.5). From a terminal point of view, the relay node will appear as an ordinary cell. This has the important advantage of simplifying the terminal implementation and making the relay node backwards compatible – that is, also accessible to LTE release-8/9 terminals. In essence, the relay is a low-power base station wirelessly connected to the remaining part of the network. One of the attractive features of a relay is the LTE-based wireless backhaul, as this could provide a simple way of improving coverage, for example in indoor environments, by simply placing relays at the problematic locations. At a later stage, if motivated by the traffic situation, the wireless donor-relay link could be replaced by an optical fiber in order to use the precious radio resources in the donor cell for terminal communication instead of serving the relay.

7.3.4 Heterogeneous Deployments

Heterogeneous deployments refer to deployments with a mixture of cells with different downlink transmission power, operating on (partially) the same set of frequencies and with overlapping geographical coverage (Figure 7.6). A typical example is a pico cell placed within the coverage area of a macro cell. Although such deployments were already supported in release 8, release 10 introduced improved inter-cell interference handling focusing on scenarios with large power differences between overlapping cells.

7.4 TERMINAL CAPABILITIES

To support different scenarios, which may call for different terminal capabilities in terms of data rates, as well as to allow for market differentiation in terms of low- and high-end terminals with a

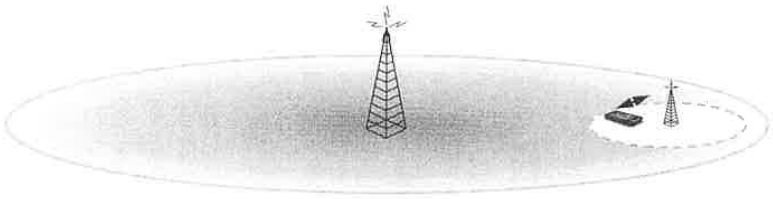


FIGURE 7.6

Example of heterogeneous deployment with a pico cell inside a macro cell.

Table 7.1 UE Categories

	Category							
	Release 8/9/10					Release 10 only		
	1	2	3	4	5	6	7	8
Downlink peak rate (Mbit/s)	10	50	100	150	300	300	300	3000
Uplink peak rate (Mbit/s)	5	25	50	50	75	50	150	1500
Maximum downlink modulation	64QAM							
Maximum uplink modulation	16QAM				64QAM	16QAM	64QAM	
Max. number of layers for downlink spatial multiplexing	1	2			4	Signaled separately		

corresponding difference in price, not all terminals support all capabilities. Furthermore, terminals from an earlier release of the standard will not support features introduced in later versions of LTE. For example, a release-8 terminal will obviously not support carrier aggregation as this feature was introduced in release 10. Therefore, as part of the connection setup, the terminal indicates not only which release of LTE it supports, but also its capabilities within the release.

In principle, the different parameters could be specified separately, but to limit the number of combinations and avoid a parameter combination that does not make sense, a set of physical-layer capabilities are lumped together to form a UE category (UE, *User Equipment*, is the term used in 3GPP to denote a terminal). In total five different UE categories have been specified for LTE release 8/9, ranging from the low-end category 1 not supporting spatial multiplexing to the high-end category 5 supporting the full set of features in the release-8/9 physical layer specifications. The categories are summarized in Table 7.1 (in simplified form; for the full set of details, see [67]). Note that, regardless

of the category, a terminal is always capable of receiving transmissions from up to four antenna ports. This is necessary as the system information can be transmitted on up to four antenna ports.

In LTE release 10, features such as carrier aggregation and uplink spatial multiplexing are introduced, which calls for additional capability signaling compared to release 8/9, either in the form of additional UE categories or as separate capabilities. Defining new categories for each foreseen combination of the maximum number of component carriers and maximum degree of spatial multiplexing could be done in principle, although the number of categories might become very large and which categories a terminal support may be frequency-band dependent. Therefore, in release 10, three additional UE categories were defined as seen in Table 7.1, and the maximum number of component carriers and degree of spatial multiplexing supported, both in uplink and downlink, are signaled separately from the category number. A release-10 terminal may therefore declare itself as, for example, category 4 but capable of uplink spatial multiplexing. Hence, categories 1–5 may have a slightly different meaning for a release-8/9 and a release-10 terminal, depending on the value of the separately declared capabilities. Furthermore, in order to be able to operate in release-8/9 networks, a release-10 UE has to be able to declare both release-8/9 and release-10 categories.

In addition to the capabilities mentioned in the different UE categories, there are some capabilities specified outside the categories. The duplexing schemes supported is one such example, the support of UE-specific reference signals for FDD in release 8 is another. Whether the terminal supports other radio-access technologies, for example GSM and WCDMA, is also declared separately.

CHAPTER

Radio-Interface Architecture

8

This chapter contains a brief overview of the overall architecture of an LTE radio-access network and the associated core network, followed by descriptions of the radio-access network user-plane and control-plane protocols.

8.1 OVERALL SYSTEM ARCHITECTURE

In parallel to the work on the LTE radio-access technology in 3GPP, the overall system architecture of both the *Radio-Access Network* (RAN) and the *Core Network* (CN) was revisited, including the split of functionality between the two network parts. This work was known as the *System Architecture Evolution* (SAE) and resulted in a flat RAN architecture, as well as a new core network architecture referred to as the *Evolved Packet Core* (EPC). Together, the LTE RAN and the EPC can be referred to as the *Evolved Packet System* (EPS).[1]

The RAN is responsible for all radio-related functionality of the overall network including, for example, scheduling, radio-resource handling, retransmission protocols, coding and various multi-antenna schemes. These functions will be discussed in detail in the subsequent chapters.

The EPC is responsible for functions not related to the radio interface but needed for providing a complete mobile-broadband network. This includes, for example, authentication, charging functionality, and setup of end-to-end connections. Handling these functions separately, instead of integrating them into the RAN, is beneficial as it allows for several radio-access technologies to be served by the same core network.

Although this book focuses on the LTE RAN, a brief overview of the EPC, as well as how it connects to the RAN, is useful. For an excellent in-depth discussion of EPC, the reader is referred to [9].

8.1.1 Core Network

The EPC is a radical evolution from the GSM/GPRS core network used for GSM and WCDMA/HSPA. EPC supports access to the *packet-switched domain* only, with no access to *the circuit-switched domain*. It consists of several different types of nodes, some of which are briefly described below and illustrated in Figure 8.1.

[1] UTRAN, the WCDMA/HSPA radio-access network, is also part of the EPS.

4G LTE/LTE-Advanced for Mobile Broadband.
© 2011 Erik Dahlman, Stefan Parkvall & Johan Sköld. Published by Elsevier Ltd. All rights reserved.

Samsung Exhibit 1019, Page 82 of 195

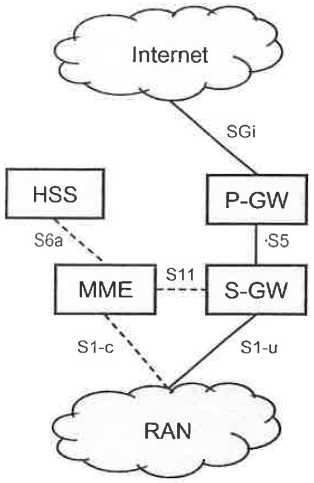


FIGURE 8.1

Core-network (EPC) architecture.

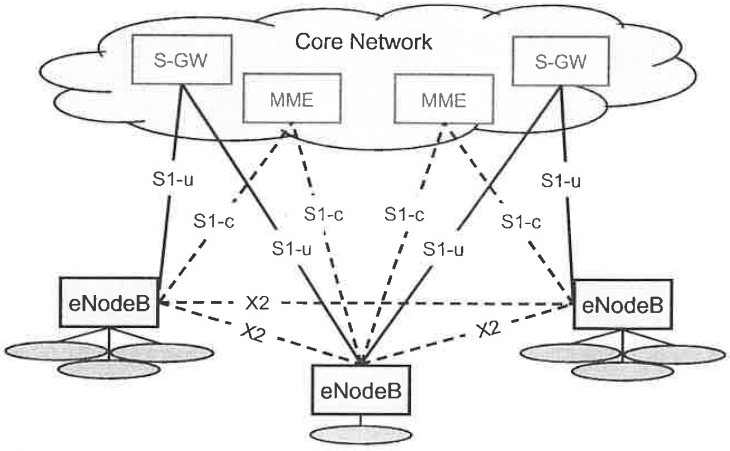


FIGURE 8.2

Radio-access-network interfaces.

The *Mobility Management Entity* (MME) is the control-plane node of the EPC. Its responsibilities include connection/release of bearers to a terminal, handling of IDLE to ACTIVE transitions, and handling of security keys. The functionality operating between the EPC and the terminal is sometimes referred to as the *Non-Access Stratum* (NAS), to separate it from the *Access Stratum* (AS) which handles functionality operating between the terminal and the radio-access network.

The *Serving Gateway* (S-GW) is the user-plane node connecting the EPC to the LTE RAN. The S-GW acts as a mobility anchor when terminals move between eNodeBs (see next section), as well as a mobility anchor for other 3GPP technologies (GSM/GPRS and HSPA). Collection of information and statistics necessary for charging is also handled by the S-GW.

The *Packet Data Network Gateway* (PDN Gateway, P-GW) connects the EPC to the internet. Allocation of the IP address for a specific terminal is handled by the P-GW, as well as quality-of-service enforcement according to the policy controlled by the PCRF (see below). The P-GW is also the mobility anchor for non-3GPP radio-access technologies, such as CDMA2000, connected to the EPC.

In addition, the EPC also contains other types of nodes such as *Policy and Charging Rules Function* (PCRF) responsible for quality-of-service (QoS) handling and charging, and the *Home Subscriber Service* (HSS) node, a database containing subscriber information. There are also some additional nodes present as regards network support of *Multimedia Broadcast Multicast Services* (MBMS) (see Chapter 15 for a more detailed description of MBMS, including the related architecture aspects).

It should be noted that the nodes discussed above are *logical* nodes. In an actual physical implementation, several of them may very well be combined. For example, the MME, P-GW, and S-GW could very well be combined into a single physical node.

8.1.2 Radio-Access Network

The LTE radio-access network uses a flat architecture with a single type of node[2] – the *eNodeB*. The eNodeB is responsible for all radio-related functions in one or several cells. It is important to note that an eNodeB is a *logical* node and not a physical implementation. One common implementation of an eNodeB is a three-sector site, where a base station is handling transmissions in three cells, although other implementations can be found as well, such as one baseband processing unit to which a number of remote radio heads are connected. One example of the latter is a large number of indoor cells, or several cells along a highway, belonging to the same eNodeB. Thus, a base station is a *possible* implementation of, but not *the same* as, an eNodeB.

As can be seen in Figure 8.2, the eNodeB is connected to the EPC by means of the *S1 interface*, more specifically to the S-GW by means of the *S1 user-plane part*, S1-u, and to the MME by means of the *S1 control-plane part*, S1-c. One eNodeB can be connected to multiple MMEs/S-GWs for the purpose of load sharing and redundancy.

The *X2 interface*, connecting eNodeBs to each other, is mainly used to support active-mode mobility. This interface may also be used for multi-cell *Radio Resource Management* (RRM) functions such as *Inter-Cell Interference Coordination* (ICIC) discussed in Chapter 13. The X2 interface is also used to support lossless mobility between neighboring cells by means of packet forwarding.

8.2 RADIO PROTOCOL ARCHITECTURE

With the overall network architecture in mind, the RAN protocol architecture for the user as well as the control planes can be discussed. Figure 8.3 illustrates the RAN protocol architecture (the MME is, as discussed in the previous section, not part of the RAN but is included in the figure for completeness). As seen in the figure, many of the protocol entities are common to the user and control planes. Therefore, although this section mainly describes the protocol architecture from a user-plane perspective, the description is in many respects also applicable to the control plane. Control-plane-specific aspects are discussed in Section 8.3.

The LTE radio-access network provides one or more *Radio Bearers* to which IP packets are mapped according to their Quality-of-Service requirements. A general overview of the LTE (user-plane) protocol architecture for the downlink is illustrated in Figure 8.4. As will become clear in the subsequent discussion, not all the entities illustrated in Figure 8.4 are applicable in all situations. For example, neither MAC scheduling nor hybrid ARQ with soft combining is used for broadcast of the basic system information. The LTE protocol structure related to uplink transmissions is similar to the downlink structure in Figure 8.4, although there are some differences with respect to, for example, transport-format selection.

The different protocol entities of the radio-access network are summarized below and described in more detail in the following sections.

- *Packet Data Convergence Protocol* (PDCP) performs IP header compression to reduce the number of bits to transmit over the radio interface. The header-compression mechanism is based

[2]The introduction of MBMS (see Chapter 15) in release 9 and relaying (see Chapter 16) in release 10 brings additional node types to the RAN.

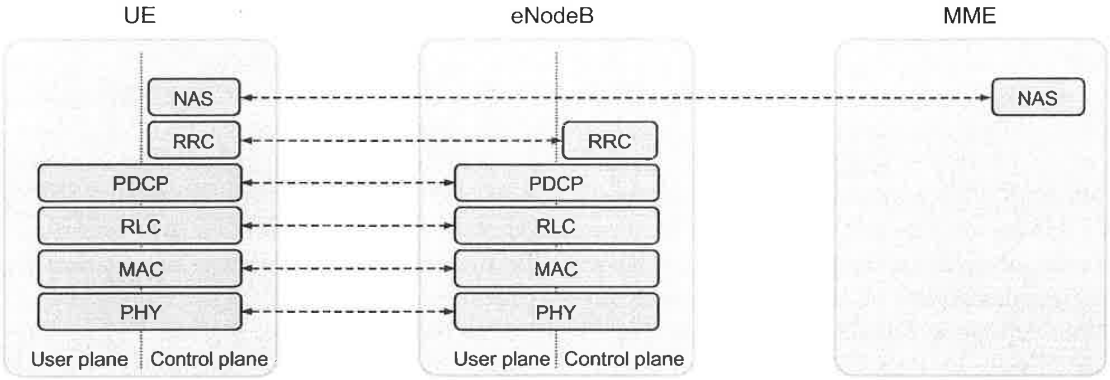


FIGURE 8.3

Overall RAN protocol architecture.

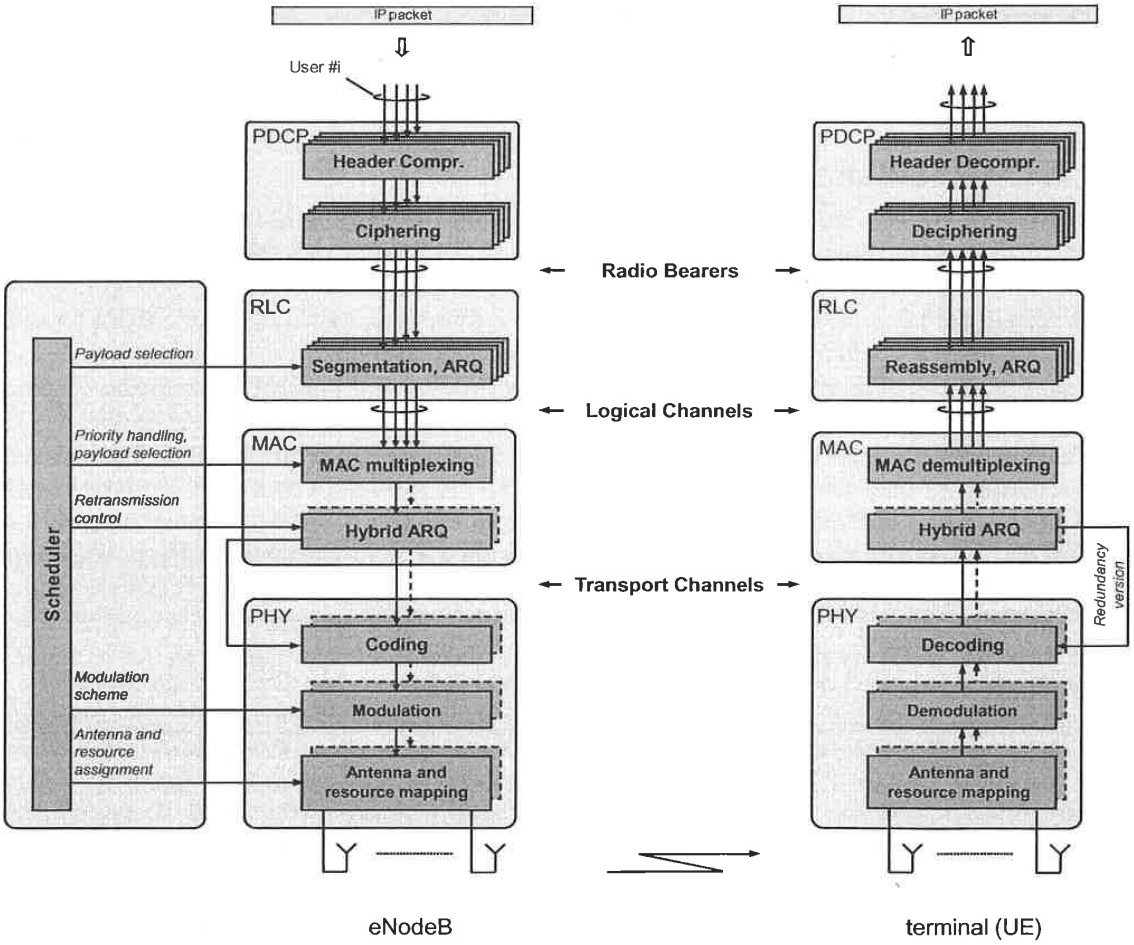


FIGURE 8.4

LTE protocol architecture (downlink).

on Robust Header Compression (ROHC) [68], a standardized header-compression algorithm also used for several mobile-communication technologies. PDCP is also responsible for ciphering and, for the control plane, integrity protection of the transmitted data, as well as in-sequence delivery and duplicate removal for handover. At the receiver side, the PDCP protocol performs the corresponding deciphering and decompression operations. There is one PDCP entity per radio bearer configured for a terminal.

- *Radio-Link Control* (RLC) is responsible for segmentation/concatenation, retransmission handling, duplicate detection, and in-sequence delivery to higher layers. The RLC provides services to the PDCP in the form of *radio bearers*. There is one RLC entity per radio bearer configured for a terminal.
- *Medium-Access Control* (MAC) handles multiplexing of logical channels, hybrid-ARQ retransmissions, and uplink and downlink scheduling. The scheduling functionality is located in the eNodeB for both uplink and downlink. The hybrid-ARQ protocol part is present in both the transmitting and receiving ends of the MAC protocol. The MAC provides services to the RLC in the form of *logical channels*.
- *Physical Layer* (PHY) handles coding/decoding, modulation/demodulation, multi-antenna mapping, and other typical physical-layer functions. The physical layer offers services to the MAC layer in the form of *transport channels*.

To summarize the flow of downlink data through all the protocol layers, an example illustration for a case with three IP packets, two on one radio bearer and one on another radio bearer, is given in Figure 8.5. The data flow in the case of uplink transmission is similar. The PDCP performs (optional) IP-header compression, followed by ciphering. A PDCP header is added, carrying information required for deciphering in the terminal. The output from the PDCP is forwarded to the RLC.

The RLC protocol performs concatenation and/or segmentation of the PDCP SDUs[3] and adds an RLC header. The header is used for in-sequence delivery (per logical channel) in the terminal and for identification of RLC PDUs in the case of retransmissions. The RLC PDUs are forwarded to the MAC layer, which multiplexes a number of RLC PDUs and attaches a MAC header to form a transport block. The transport-block size depends on the instantaneous data rate selected by the link-adaptation mechanism. Thus, the link adaptation affects both the MAC and RLC processing. Finally, the physical layer attaches a CRC to the transport block for error-detection purposes, performs coding and modulation, and transmits the resulting signal, possibly using multiple transmit antennas.

The remainder of the chapter contains an overview of the RLC, MAC, and physical layers. A more detailed description of the LTE physical-layer processing is given in Chapters 10 (downlink) and 11 (uplink), followed by descriptions of some specific uplink and downlink radio-interface functions and procedures in the subsequent chapters.

8.2.1 Radio-Link Control

The RLC protocol is responsible for segmentation/concatenation of (header-compressed) IP packets, also known as RLC SDUs, from the PDCP into suitably sized RLC PDUs. It also handles

[3] In general, the data entity from/to a higher protocol layer is known as a Service Data Unit (SDU) and the corresponding entity to/from a lower protocol layer entity is called a Protocol Data Unit (PDU).

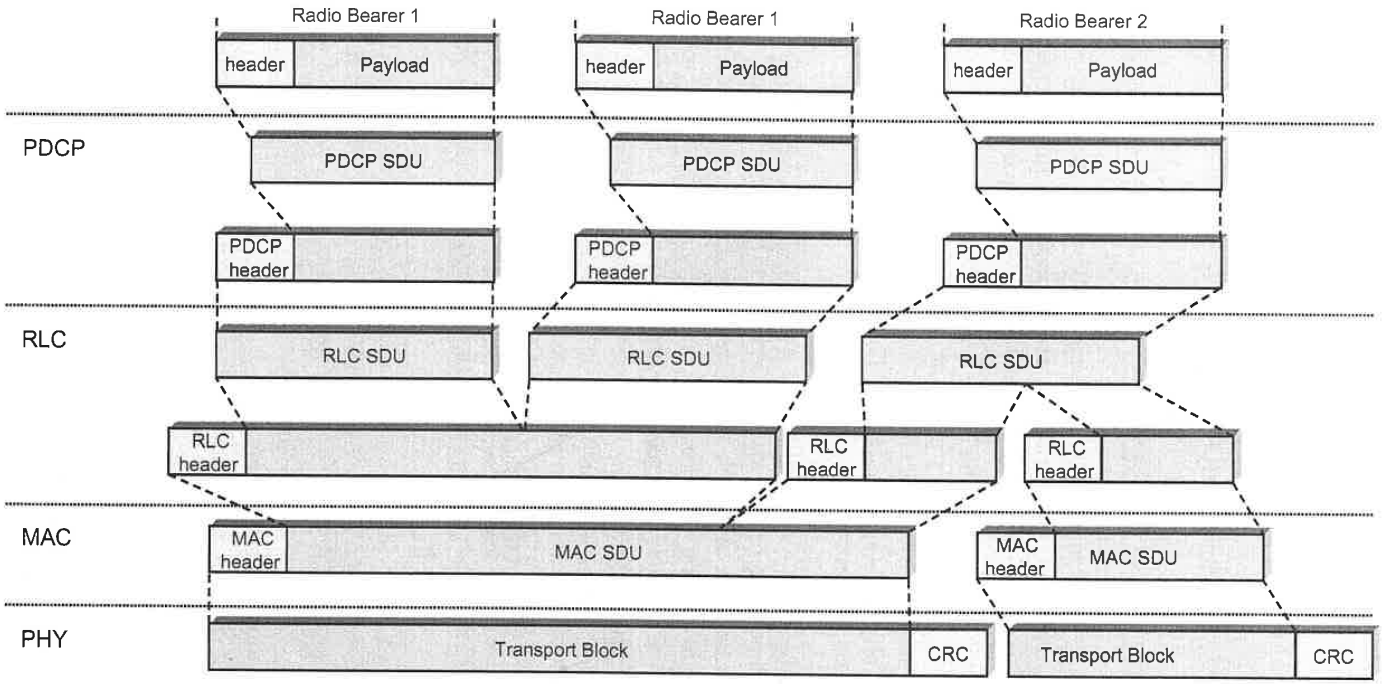


FIGURE 8.5

Example of LTE data flow.

Samsung Exhibit 1019, Page 87 of 195

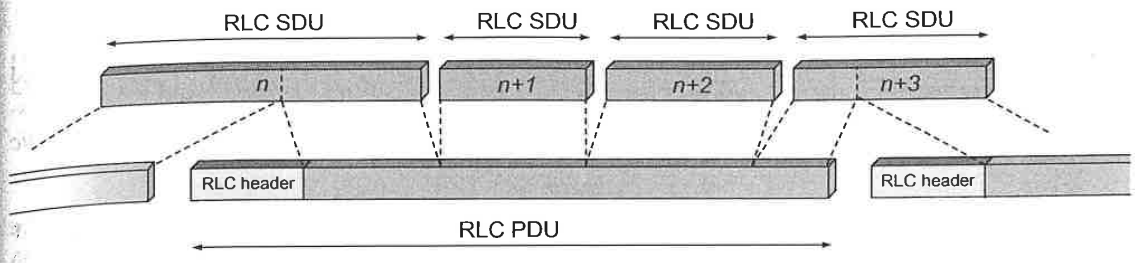


FIGURE 8.6

RLC segmentation and concatenation.

retransmission of erroneously received PDUs, as well as removal of duplicated PDUs. Finally, the RLC ensures in-sequence delivery of SDUs to upper layers. Depending on the type of service, the RLC can be configured in different modes to perform some or all of these functions.

Segmentation and concatenation, one of the main RLC functions, is illustrated in Figure 8.6. Depending on the scheduler decision, a certain amount of data is selected for transmission from the RLC SDU buffer and the SDUs are segmented/concatenated to create the RLC PDU. Thus, for LTE the RLC PDU size varies *dynamically*. For high data rates, a large PDU size results in a smaller relative overhead, while for low data rates, a small PDU size is required as the payload would otherwise be too large. Hence, as the LTE data rates may range from a few kbit/s up to 3 Gbit/s, dynamic PDU sizes are motivated for LTE in contrast to earlier mobile-communication technologies, which typically use a fixed PDU size. Since the RLC, scheduler, and rate adaptation mechanisms are all located in the eNodeB, dynamic PDU sizes are easily supported for LTE. In each RLC PDU, a header is included, containing, among other things, a sequence number used for in-sequence delivery and by the retransmission mechanism.

The RLC retransmission mechanism is also responsible for providing error-free delivery of data to higher layers. To accomplish this, a retransmission protocol operates between the RLC entities in the receiver and transmitter. By monitoring the sequence numbers of the incoming PDUs, the receiving RLC can identify missing PDUs. Status reports are then fed back to the transmitting RLC entity, requesting retransmission of missing PDUs. Based on the received status report, the RLC entity at the transmitter can take the appropriate action and retransmit the missing PDUs if needed.

Although the RLC is capable of handling transmission errors due to noise, unpredictable channel variations, etc., error-free delivery is in most cases handled by the MAC-based hybrid-ARQ protocol. The use of a retransmission mechanism in the RLC may therefore seem superfluous at first. However, as will be discussed in Section 8.2.2.3, this is not the case and the use of both RLC- and MAC-based retransmission mechanisms is in fact well motivated by the differences in the feedback signaling.

The details of RLC are further described in Chapter 12.

8.2.2 Medium-Access Control

The MAC layer handles logical-channel multiplexing, hybrid-ARQ retransmissions, and uplink and downlink scheduling. It is also responsible for multiplexing/demultiplexing data across multiple component carriers when carrier aggregation is used.

8.2.2.1 Logical Channels and Transport Channels

The MAC provides services to the RLC in the form of *logical channels*. A logical channel is defined by the *type* of information it carries and is generally classified as a *control channel*, used for transmission of control and configuration information necessary for operating an LTE system, or as a *traffic channel*, used for the user data. The set of logical-channel types specified for LTE includes:

- *The Broadcast Control Channel* (BCCH), used for transmission of *system information* from the network to all terminals in a cell. Prior to accessing the system, a terminal needs to acquire the system information to find out how the system is configured and, in general, how to behave properly within a cell.
- The *Paging Control Channel* (PCCH), used for paging of terminals whose location on a cell level is not known to the network. The paging message therefore needs to be transmitted in multiple cells.
- The *Common Control Channel* (CCCH), used for transmission of control information in conjunction with random access.
- The *Dedicated Control Channel* (DCCH), used for transmission of control information to/from a terminal. This channel is used for individual configuration of terminals such as different handover messages.
- The *Multicast Control Channel* (MCCH), used for transmission of control information required for reception of the MTCH (see below).
- The *Dedicated Traffic Channel* (DTCH), used for transmission of user data to/from a terminal. This is the logical channel type used for transmission of all uplink and non-MBSFN downlink user data.
- The *Multicast Traffic Channel* (MTCH), used for downlink transmission of MBMS services.

From the physical layer, the MAC layer uses services in the form of *transport channels*. A transport channel is defined by *how* and *with what characteristics* the information is transmitted over the radio interface. Data on a transport channel is organized into *transport blocks*. In each *Transmission Time Interval* (TTI), at most one transport block of dynamic size is transmitted over the radio interface to/from a terminal in the absence of spatial multiplexing. In the case of spatial multiplexing (MIMO), there can be up to two transport blocks per TTI.

Associated with each transport block is a *Transport Format* (TF), specifying *how* the transport block is to be transmitted over the radio interface. The transport format includes information about the transport-block size, the modulation-and-coding scheme, and the antenna mapping. By varying the transport format, the MAC layer can thus realize different data rates. Rate control is therefore also known as *transport-format selection*.

The following transport-channel types are defined for LTE:

- The *Broadcast Channel* (BCH) has a fixed transport format, provided by the specifications. It is used for transmission of parts of the BCCH system information, more specifically the so-called *Master Information Block* (MIB), as described in Chapter 14.
- The *Paging Channel* (PCH) is used for transmission of paging information from the PCCH logical channel. The PCH supports *discontinuous reception* (DRX) to allow the terminal to save battery power by waking up to receive the PCH only at predefined time instants. The LTE paging mechanism is also described in Chapter 14.

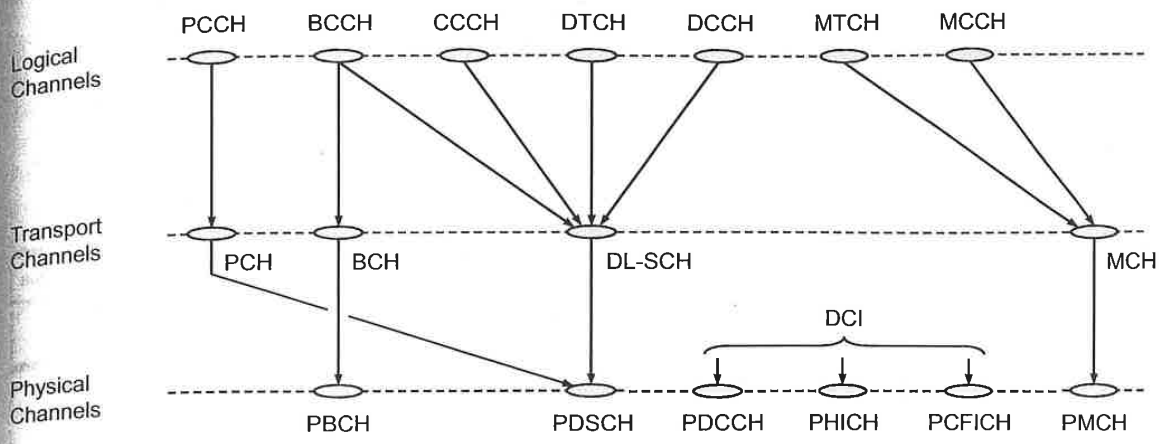


FIGURE 8.7

Downlink channel mapping.

- The *Downlink Shared Channel* (DL-SCH) is the main transport channel used for transmission of downlink data in LTE. It supports key LTE features such as dynamic rate adaptation and channel-dependent scheduling in the time and frequency domains, hybrid ARQ with soft combining, and spatial multiplexing. It also supports DRX to reduce terminal power consumption while still providing an always-on experience. The DL-SCH is also used for transmission of the parts of the BCCH system information not mapped to the BCH. There can be multiple DL-SCHs in a cell, one per terminal[4] scheduled in this TTI, and, in some subframes, one DL-SCH carrying system information.
- The *Multicast Channel* (MCH) is used to support MBMS. It is characterized by a semi-static transport format and semi-static scheduling. In the case of multi-cell transmission using MBSFN, the scheduling and transport format configuration is coordinated among the transmission points involved in the MBSFN transmission. MBSFN transmission is described in Chapter 15.
- The Uplink Shared Channel (UL-SCH) is the uplink counterpart to the DL-SCH – that is, the uplink transport channel used for transmission of uplink data.

In addition, the *Random-Access Channel* (RACH) is also defined as a transport channel, although it does not carry transport blocks.

Part of the MAC functionality is multiplexing of different logical channels and mapping of the logical channels to the appropriate transport channels. The supported mappings between logical-channel types and transport-channel types are given in Figure 8.7 for the downlink and Figure 8.8 for the uplink. The figures clearly indicate how DL-SCH and UL-SCH are the main downlink and uplink transport channels respectively. In the figures, the corresponding physical channels, described further below, are also included and the mapping between transport channels and physical channels is illustrated.

To support priority handling, multiple logical channels, where each logical channel has its own RLC entity, can be multiplexed into one transport channel by the MAC layer. At the receiver, the

[4]For carrier aggregation, a UE may receive multiple DL-SCHs, one per component carrier.

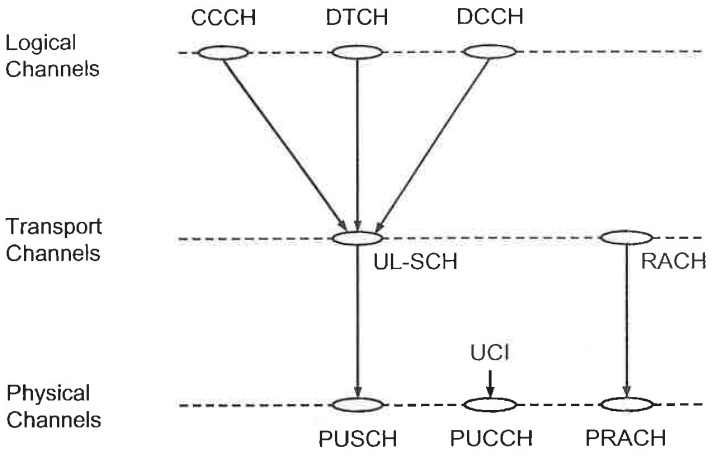


FIGURE 8.8

Uplink channel mapping.

MAC layer handles the corresponding demultiplexing and forwards the RLC PDUs to their respective RLC entity for in-sequence delivery and the other functions handled by the RLC. To support the demultiplexing at the receiver, a MAC header, shown in Figure 8.9, is used. To each RLC PDU, there is an associated sub-header in the MAC header. The sub-header contains the identity of the logical channel (LCID) from which the RLC PDU originated and the length of the PDU in bytes. There is also a flag indicating whether this is the last sub-header or not. One or several RLC PDUs, together with the MAC header and, if necessary, padding to meet the scheduled transport-block size, form one transport block which is forwarded to the physical layer.

In addition to multiplexing of different logical channels, the MAC layer can also insert the so-called *MAC control elements* into the transport blocks to be transmitted over the transport channels. A MAC control element is used for inband control signaling – for example, timing-advance commands and random-access response, as described in Sections 11.5 and 14.3 respectively. Control elements are identified with reserved values in the LCID field, where the LCID value indicates the type of control information. Furthermore, the length field in the sub-header is removed for control elements with a fixed length.

The MAC multiplexing functionality is also responsible for handling of multiple component carriers in the case of carrier aggregation. The basic principle for carrier aggregation is independent processing of the component carriers in the physical layer, including control signaling, scheduling and hybrid-ARQ retransmissions, while carrier aggregation is invisible to RLC and PDCP. Carrier aggregation is therefore mainly seen in the MAC layer, as illustrated in Figure 8.10, where logical channels, including any MAC control elements, are multiplexed to form one (two in the case of spatial multiplexing) transport block(s) per component carrier with each component carrier having its own hybrid-ARQ entity.

8.2.2.2 Scheduling

One of the basic principles of LTE radio access is shared-channel transmission – that is, time–frequency resources are dynamically shared between users. The *scheduler* is part of the MAC layer

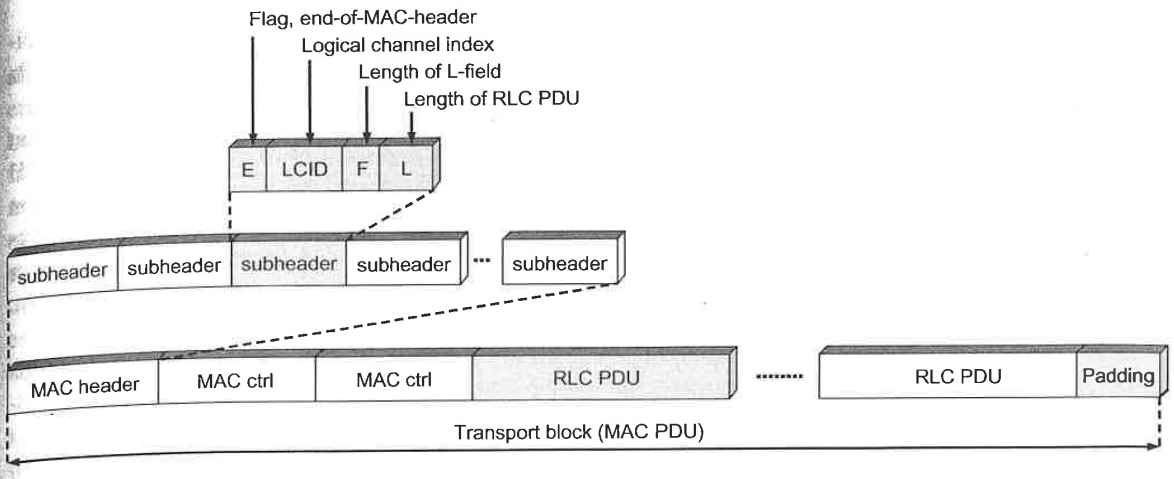


FIGURE 8.9

MAC header and SDU multiplexing.

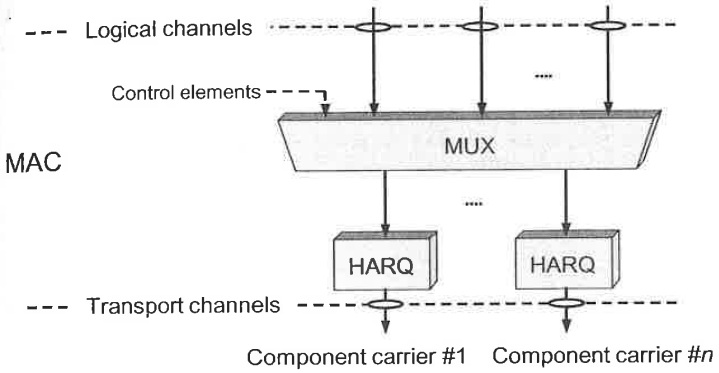


FIGURE 8.10

MAC operation in the case of multiple component carriers.

(although often better viewed as a separate entity as illustrated in Figure 8.4) and controls the assignment of uplink and downlink resources in terms of so-called *resource-block pairs*. Resource blocks correspond to a time–frequency unit of 1 ms times 180 kHz, as described in more detail in Chapter 9.

The basic operation of the scheduler is so-called *dynamic* scheduling, where the eNodeB in each 1 ms interval takes a scheduling decision and sends scheduling information to the selected set of terminals. However, there is also a possibility for semi-persistent scheduling where a semi-static scheduling pattern is signaled in advance to reduce the control-signaling overhead. Coordination of scheduling decisions across multiple cells residing in different eNodeBs is supported using signaling over the X2 interface.

Uplink and downlink scheduling are separated in LTE, and uplink and downlink scheduling decisions can be taken independently of each other (within the limits set by the uplink/downlink split in the case of half-duplex FDD operation).

The downlink scheduler is responsible for (dynamically) controlling which terminal(s) to transmit to and, for each of these terminals, the set of resource blocks upon which the terminal's DL-SCH should be transmitted. Transport-format selection (selection of transport-block size, modulation scheme, and antenna mapping) and logical-channel multiplexing for downlink transmissions are controlled by the eNodeB, as illustrated in the left part of Figure 8.11. As a consequence of the scheduler controlling the data rate, the RLC segmentation and MAC multiplexing will also be affected by the scheduling decision. The outputs from the downlink scheduler can be seen in Figure 8.4.

The uplink scheduler serves a similar purpose, namely to (dynamically) control which terminals are to transmit on their respective UL-SCH and on which uplink time–frequency resources (including component carrier). Despite the fact that the eNodeB scheduler determines the transport format for the terminal, it is important to point out that the uplink scheduling decision is taken *per terminal* and not per radio bearer. Thus, although the eNodeB scheduler controls the payload of a scheduled terminal, the terminal is still responsible for selecting *from which radio bearer(s)* the data is taken. The terminal autonomously handles logical-channel multiplexing according to rules, the parameters of which can be configured by the eNodeB. This is illustrated in the right part of Figure 8.11, where the eNodeB scheduler controls the transport format and the terminal controls the logical-channel multiplexing.

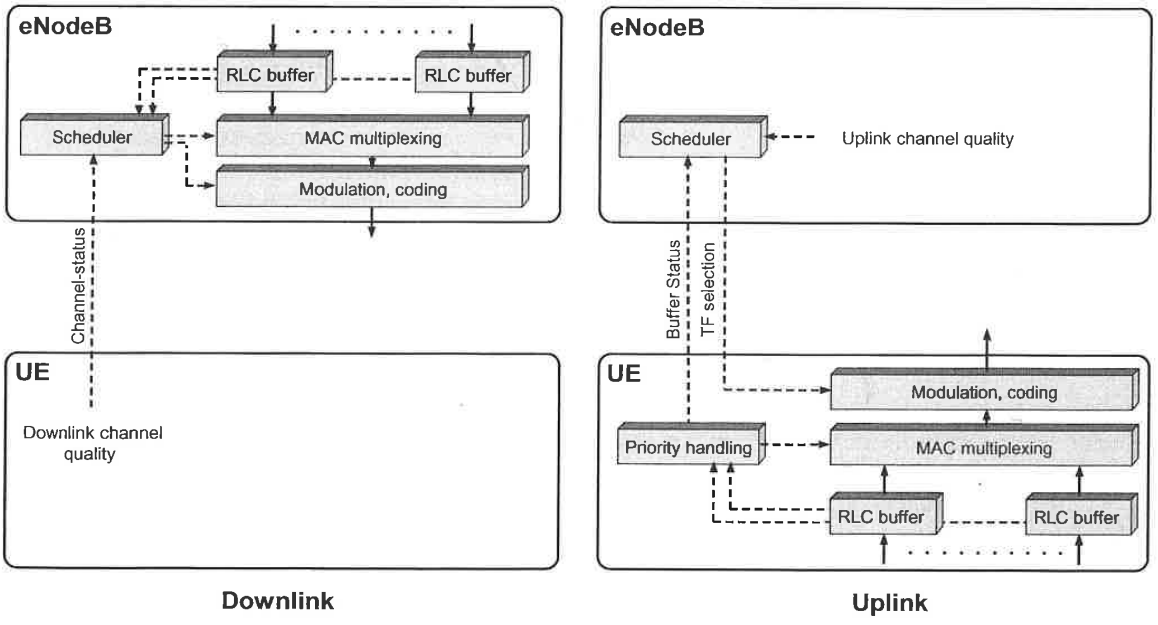


FIGURE 8.11

Transport-format selection in downlink and uplink.

Although the scheduling strategy is implementation specific and not specified by 3GPP, the overall goal of most schedulers is to take advantage of the channel variations between terminals and preferably schedule transmissions to a terminal on resources with advantageous channel conditions. A benefit of the use of OFDM in LTE is the possibility to exploit channel variations in both time and frequency domains through channel-dependent scheduling. This was mentioned earlier in Chapter 7 and illustrated in Figure 7.2. For the larger bandwidths supported by LTE, where a significant amount of frequency-selective fading will often be experienced, the possibility for the scheduler to also exploit frequency-domain channel variations becomes increasingly important compared to exploiting time-domain variations only. The possibility to also exploit frequency-domain variations is beneficial especially at low speeds, where the variations in the time domain are relatively slow compared to the delay requirements set by many services.

Downlink channel-dependent scheduling is supported through *channel-state reports*, transmitted by the UE and reflecting the instantaneous channel quality in the time and frequency domains, as well as information necessary to determine the appropriate antenna processing in the case of spatial multiplexing. In the uplink, the channel-state information necessary for uplink channel-dependent scheduling can be based on a *sounding reference signal* transmitted from each terminal for which the eNodeB wants to estimate the uplink channel quality. To aid the uplink scheduler in its decisions, the terminal can transmit buffer-status information to the eNodeB using a MAC message. Obviously, this information can only be transmitted if the terminal has been given a valid scheduling grant. For situations when this is not the case, an indicator that the terminal needs uplink resources is provided as part of the uplink L1/L2 control-signaling structure (see Chapter 11).

Interference coordination, which tries to control the inter-cell interference on a slow basis, as mentioned in Chapter 7, is also part of the scheduler. As the scheduling strategy is not mandated by the specifications, the interference-coordination scheme (if used) is vendor specific and may range from simple higher-order reuse deployments to more advanced schemes. The mechanisms supporting inter-cell interference coordination are discussed in Chapter 13.

8.2.2.3 Hybrid ARQ with Soft Combining

Hybrid ARQ with soft combining provides robustness against transmission errors. As hybrid-ARQ retransmissions are fast, many services allow for one or multiple retransmissions, thereby forming an implicit (closed loop) rate-control mechanism. The hybrid-ARQ protocol is part of the MAC layer, while the actual soft combining is handled by the physical layer.[5]

Obviously, hybrid ARQ is not applicable for all types of traffic. For example, broadcast transmissions, where the same information is intended for multiple terminals, typically do not rely on hybrid ARQ. Hence, hybrid ARQ is only supported for the DL-SCH and the UL-SCH, although its usage is optional.

The LTE hybrid-ARQ protocol uses multiple parallel stop-and-wait processes. Upon reception of a transport block, the receiver makes an attempt to decode the transport block and informs the transmitter about the outcome of the decoding operation through a single acknowledgement bit indicating whether the decoding was successful or if a retransmission of the transport block is required. Further details on transmission of hybrid-ARQ acknowledgements are found in Chapters 10–12. Clearly, the receiver must know to which hybrid-ARQ process a received acknowledgement is associated. This is

[5]The soft combining is done before or as part of the channel decoding, which clearly is a physical-layer functionality.

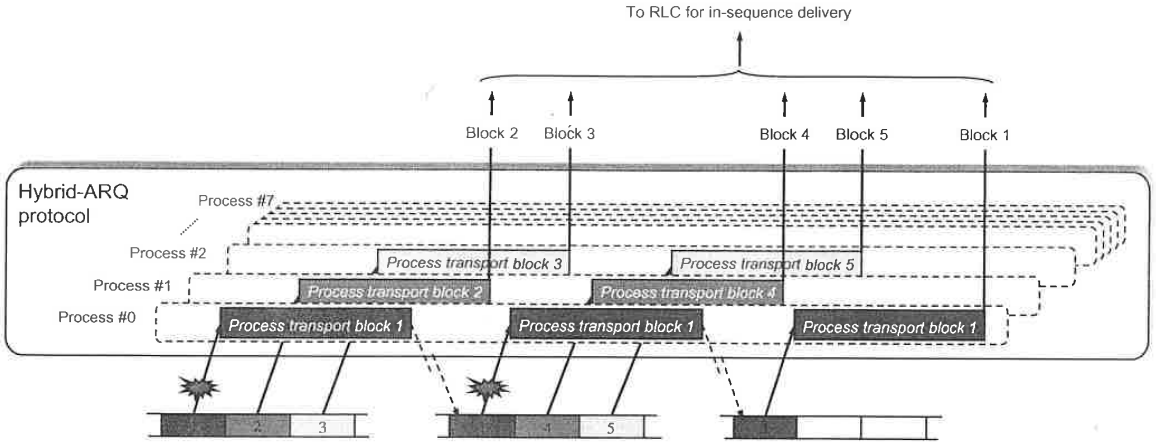


FIGURE 8.12

Multiple parallel hybrid-ARQ processes.

solved by using the timing of the acknowledgement for association with a certain hybrid-ARQ process. Note that, in the case of TDD operation, the time relation between the reception of data in a certain hybrid-ARQ process and the transmission of the acknowledgement is also affected by the uplink/downlink allocation.

The use of multiple parallel hybrid-ARQ processes, illustrated in Figure 8.12, for each user can result in data being delivered from the hybrid-ARQ mechanism out of sequence. For example, transport block 5 in the figure was successfully decoded before transport block 1 which required retransmissions. In-sequence delivery of data is therefore ensured by the RLC layer. The same reordering mechanism in the RLC is also used to handle in-sequence delivery across multiple component carriers in the case of carrier aggregation. As mentioned above, the hybrid-ARQ retransmissions are handled independently per component carrier, which may result in out-of-sequence delivery in a similar way as within a component carrier.

Downlink retransmissions may occur at any time after the initial transmission – that is, the protocol is asynchronous – and an explicit hybrid-ARQ process number is used to indicate which process is being addressed. In an asynchronous hybrid-ARQ protocol, the retransmissions are in principle scheduled similarly to the initial transmissions. Uplink retransmissions, on the other hand, are based on a synchronous protocol, the retransmission occurs at a predefined time after the initial transmission and the process number can be implicitly derived. In a synchronous protocol the time instant for the retransmissions is fixed once the initial transmission has been scheduled, which must be accounted for in the scheduling operation. However, note that the scheduler knows from the hybrid-ARQ entity in the eNodeB whether a terminal will perform a retransmission or not.

The hybrid-ARQ mechanism will rapidly correct transmission errors due to noise or unpredictable channel variations. As discussed above, the RLC is also capable of requesting retransmissions, which at first sight may seem unnecessary. However, the reason for having two retransmission mechanisms on top of each other can be seen in the feedback signaling – hybrid ARQ provides fast retransmissions but due to errors in the feedback the residual error rate is typically too high for, for example,

good TCP performance, while RLC ensures (almost) error-free data delivery but slower retransmissions than the hybrid-ARQ protocol. Hence, the combination of hybrid ARQ and RLC provides an attractive combination of small round-trip time and reliable data delivery. Furthermore, as the RLC and hybrid ARQ are located in the same node, tight interaction between the two is possible, as discussed in Chapter 12.

8.2.3 Physical Layer

The physical layer is responsible for coding, physical-layer hybrid-ARQ processing, modulation, multi-antenna processing, and mapping of the signal to the appropriate physical time–frequency resources. It also handles mapping of transport channels to physical channels, as shown in Figures 8.7 and 8.8.

As mentioned in the introduction, the physical layer provides services to the MAC layer in the form of transport channels. Data transmission in downlink and uplink use the DL-SCH and UL-SCH transport-channel types respectively. There is at most one or, in the case of spatial multiplexing, two transport blocks per TTI on a DL-SCH or UL-SCH. In the case of carrier aggregation, there is one DL-SCH (or UL-SCH) per component carrier.

A *physical channel* corresponds to the set of time–frequency resources used for transmission of a particular transport channel and each transport channel is mapped to a corresponding physical channel, as shown in Figures 8.7 and 8.8. In addition to the physical channels with a corresponding transport channel, there are also physical channels without a corresponding transport channel. These channels, known as L1/L2 control channels, are used for *downlink control information* (DCI), providing the terminal with the necessary information for proper reception and decoding of the downlink data transmission, and *uplink control information* (UCI) used for providing the scheduler and the hybrid-ARQ protocol with information about the situation at the terminal.

The physical-channel types defined in LTE include the following:

- The *Physical Downlink Shared Channel* (PDSCH) is the main physical channel used for unicast data transmission, but also for transmission of paging information.
- The *Physical Broadcast Channel* (PBCH) carries part of the system information, required by the terminal in order to access the network.
- The *Physical Multicast Channel* (PMCH) is used for MBSFN operation.
- The *Physical Downlink Control Channel* (PDCCH) is used for downlink control information, mainly scheduling decisions, required for reception of PDSCH, and for scheduling grants enabling transmission on the PUSCH.
- The *Physical Hybrid-ARQ Indicator Channel* (PHICH) carries the hybrid-ARQ acknowledgement to indicate to the terminal whether a transport block should be retransmitted or not.
- The *Physical Control Format Indicator Channel* (PCFICH) is a channel providing the terminals with information necessary to decode the set of PDCCHs. There is only one PCFICH per component carrier.
- The *Physical Uplink Shared Channel* (PUSCH) is the uplink counterpart to the PDSCH. There is at most one PUSCH per uplink component carrier per terminal.
- The *Physical Uplink Control Channel* (PUCCH) is used by the terminal to send hybrid-ARQ acknowledgements, indicating to the eNodeB whether the downlink transport block(s) was successfully received or not, to send channel-state reports aiding downlink channel-dependent scheduling, and for requesting resources to transmit uplink data upon. There is at most one PUCCH per terminal.

- The *Physical Random-Access Channel* (PRACH) is used for random access, as described in Chapter 14.

Note that some of the physical channels, more specifically the channels used for downlink control information (PCFICH, PDCCH, PHICH) and uplink control information (PUCCH), do not have a corresponding transport channel.

The remaining downlink transport channels are based on the same general physical-layer processing as the DL-SCH, although with some restrictions in the set of features used. This is especially true for PCH and MCH transport channels. For the broadcast of system information on the BCH, a terminal must be able to receive this information channel as one of the first steps prior to accessing the system. Consequently, the transmission format must be known to the terminals a priori and there is no dynamic control of any of the transmission parameters from the MAC layer in this case. The BCH is also mapped to the physical resource (the OFDM time–frequency grid) in a different way, as described in more detail in Chapter 14.

For transmission of paging messages on the PCH, dynamic adaptation of the transmission parameters can, to some extent, be used. In general, the processing in this case is similar to the generic DL-SCH processing. The MAC can control modulation, the amount of resources, and the antenna mapping. However, as an uplink has not yet been established when a terminal is paged, hybrid ARQ cannot be used as there is no possibility for the terminal to transmit a hybrid-ARQ acknowledgement.

The MCH is used for MBMS transmissions, typically with single-frequency network operation, as described in Chapter 3, by transmitting from multiple cells on the same resources with the same format at the same time. Hence, the scheduling of MCH transmissions must be coordinated between the cells involved and dynamic selection of transmission parameters by the MAC is not possible.

8.3 CONTROL-PLANE PROTOCOLS

The control-plane protocols are, among other things, responsible for connection setup, mobility, and security. Control messages transmitted from the network to the terminals can originate either from the MME, located in the core network, or from the *Radio Resource Control* (RRC), located in the eNodeB.

NAS control-plane functionality, handled by the MME, includes EPS bearer management, authentication, security, and different idle-mode procedures such as paging. It is also responsible for assigning an IP address to a terminal. For a detailed discussion about the NAS control-plane functionality, see [9].

The RRC is located in the eNodeB and is responsible for handling the RAN-related procedures, including:

- Broadcast of system information necessary for the terminal to be able to communicate with a cell. Acquisition of system information is described in Chapter 14.
- Transmission of paging messages originating from the MME to notify the terminal about incoming connection requests. Paging, discussed further in Chapter 14, is used in the RRC_IDLE state (described further below) when the terminal is not connected to a particular cell. Indication of system-information update is another use of the paging mechanism, as is public warning systems.

- Connection management, including setting up bearers and mobility within LTE. This includes establishing an RRC context – that is, configuring the parameters necessary for communication between the terminal and the radio-access network.
- Mobility functions such as cell (re)selection.
- Measurement configuration and reporting.
- Handling of UE capabilities; when connection is established the terminal will announce its capabilities as all terminals are not capable of supporting all the functionality described in the LTE specifications, as briefly discussed in Chapter 7.

RRC messages are transmitted to the terminal using *signaling radio bearers* (SRBs), using the same set of protocol layers (PDCP, RLC, MAC and PHY) as described in Section 8.2. The SRB is mapped to the common control channel (CCCH) during establishment of connection and, once a connection is established, to the dedicated control channel (DCCH). Control-plane and user-plane data can be multiplexed in the MAC layer and transmitted to the terminal in the same TTI. The aforementioned MAC control elements can also be used for control of radio resources in some specific cases where low latency is more important than ciphering, integrity protection, and reliable transfer.

8.3.1 State Machine

In LTE, a terminal can be in two different states, as illustrated in Figure 8.13, RRC_CONNECTED and RRC_IDLE.

In RRC_CONNECTED, there is an RRC context established – that is, the parameters necessary for communication between the terminal and the radio-access network are known to both entities. The cell to which the terminal is belongs is known and an identity of the terminal, the *Cell Radio-Network Temporary Identifier* (C-RNTI), used for signaling purposes between the terminal and the network, has been configured. RRC_CONNECTED is intended for data transfer to/from the terminal, but *discontinuous reception* (DRX) can be configured in order to reduce terminal power consumption (DRX is described in further detail in Chapter 13). Since there is an RRC context established in the eNodeB in RRC_CONNECTED, leaving DRX and starting to receive/transmit data is relatively fast as no connection with its associated signaling is needed.

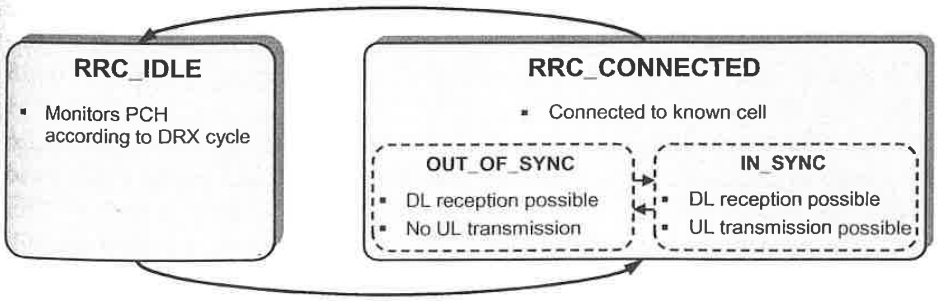


FIGURE 8.13

LTE states.

Although expressed differently in the specifications, RRC_CONNECTED can be thought of as having two substates, IN_SYNC and OUT_OF_SYNC, depending on whether the uplink is synchronized to the network or not. Since LTE uses an orthogonal FDMA/TDMA-based uplink, it is necessary to synchronize the uplink transmission from different terminals such that they arrive at the receiver at (approximately) the same time. The procedure for obtaining and maintaining uplink synchronization is described in Chapter 11, but in short the receiver measures the arrival time of the transmissions from each actively transmitting terminal and sends timing-correction commands in the downlink. As long as the uplink is synchronized, uplink transmission of user data and L1/L2 control signaling is possible. If no uplink transmission has taken place within a configurable time window, timing alignment is obviously not possible and the uplink is declared to be non-synchronized. In this case, the terminal needs to perform a random-access procedure to restore uplink synchronization prior to transmission of uplink data or control information.

In RRC_IDLE, there is no RRC context in the radio-access network and the terminal does not belong to a specific cell. No data transfer may take place as the terminal sleeps most of the time in order to reduce battery consumption. Uplink synchronization is not maintained and hence the only uplink transmission activity that may take place is random access, discussed in Chapter 14, to move to RRC_CONNECTED. When moving to RRC_CONNECTED the RRC context needs to be established in both the radio-access network and the terminal. Compared to leaving DRX this takes a somewhat longer time. In the downlink, terminals in RRC_IDLE periodically wake up in order to receive paging messages, if any, from the network, as described in Chapter 14.

CHAPTER

9 Physical Transmission Resources

In Chapter 8, the overall LTE architecture was discussed, including an overview of the different protocol layers. Prior to discussing the detailed LTE downlink and uplink transmission schemes, a description of the basic time- and frequency-domain structure of LTE will be provided in this chapter.

9.1 OVERALL TIME–FREQUENCY STRUCTURE

OFDM is the basic transmission scheme for both the downlink and uplink transmission directions in LTE although, for the uplink, specific means are taken to reduce the cubic metric of the transmitted signal, thereby allowing for improved efficiency for the terminal transmitter power amplifier. Thus, for data transmission, DFT precoding is applied before OFDM modulation, leading to *DFT-spread OFDM* or *DFTS-OFDM*, as described in Chapter 4. It should be noted though that DFTS-OFDM is only applied to uplink *data* transmission – that is, for the transmission of the UL-SCH transport channel. As will be seen in Chapter 11, for other uplink transmissions, such as the transmission of *L1/L2 control signaling* and different types of *reference-signal transmissions*, other means are taken to limit the cubic metric of the transmitted signal.

The LTE OFDM subcarrier spacing equals 15 kHz for both downlink and uplink. As discussed in Chapter 3, the selection of the subcarrier spacing in an OFDM-based system needs to carefully balance overhead from the cyclic prefix against sensitivity to Doppler spread/shift and other types of frequency errors and inaccuracies. The choice of 15 kHz for the LTE subcarrier spacing was found to offer a good balance between these two constraints.

Assuming an FFT-based transmitter/receiver implementation, 15 kHz subcarrier spacing corresponds to a sampling rate $f_s = 15000 \cdot N_{FFT}$, where N_{FFT} is the FFT size. It is important to understand though that the LTE specifications do not in any way mandate the use of FFT-based transmitter/receiver implementations and even less so a particular FFT size or sampling rate. Nevertheless, FFT-based implementations of OFDM are common practice and an FFT size of 2048, with a corresponding sampling rate of 30.72 MHz, is suitable for the wider LTE carrier bandwidths, such as bandwidths of the order of 15 MHz and above. However, for smaller carrier bandwidths, a smaller FFT size and a correspondingly lower sampling rate can very well be used. The sampling rate above illustrates another factor influencing the choice of the LTE subcarrier spacing, namely a desire to simplify implementation of dual-mode LTE/HSPA terminals. Assuming a power-of-two FFT size and a subcarrier spacing of 15 kHz, the sampling rate $\Delta f \cdot N_{FFT}$ will be a multiple or sub-multiple of the HSPA chip rate of 3.84 Mchip/s.

4G LTE/LTE-Advanced for Mobile Broadband.
© 2011 Erik Dahlman, Stefan Parkvall & Johan Sköld. Published by Elsevier Ltd. All rights reserved.

Samsung Exhibit 1019, Page 100 of 195

In addition to the 15 kHz subcarrier spacing, a *reduced subcarrier spacing* of 7.5 kHz, with a corresponding OFDM symbol time that is twice as long, is also defined for LTE. The introduction of the reduced subcarrier spacing specifically targeted MBSFN-based multicast/broadcast transmissions (see Chapter 15). However, currently the 7.5 kHz subcarrier numerology is only partly implemented in the LTE specifications. Thus, at least for LTE up to and including release 10, only the 15 kHz subcarrier spacing is fully supported. The remaining discussions within this and the following chapters will assume the 15 kHz subcarrier spacing unless explicitly stated otherwise.

In the time domain, LTE transmissions are organized into *(radio) frames* of length 10 ms, each of which is divided into ten equally sized *subframes* of length 1 ms, as illustrated in Figure 9.1. Each subframe consists of two equally sized *slots* of length $T_{\text{slot}} = 0.5$ ms, with each slot consisting of a number of OFDM symbols including cyclic prefix.[1]

To provide consistent and exact timing definitions, different time intervals within the LTE specifications are defined as multiples of a basic time unit $T_s = 1/(15000 \cdot 2048)$. The basic time unit T_s can thus be seen as the sampling time of an FFT-based transmitter/receiver implementation with an FFT size equal to 2048. The time intervals outlined in Figure 9.1 can thus also be expressed as $T_{\text{frame}} = 307200 \cdot T_s$, $T_{\text{subframe}} = 30720 \cdot T_s$, and $T_{\text{slot}} = 15360 \cdot T_s$ for the frame, subframe, and slot durations respectively.

On a higher level, each frame is identified by a *System Frame Number* (SFN). The SFN is used to control different transmission cycles that may have a period longer than one frame, such as paging sleep-mode cycles and periods for channel-status reporting. The SFN period equals 1024, thus the SFN repeats itself after 1024 frames or roughly 10 seconds.

The 15 kHz LTE subcarrier spacing corresponds to a useful symbol time $T_u = 2048 \cdot T_s$ or approximately 66.7 μs. The overall OFDM symbol time is then the sum of the useful symbol time and the cyclic-prefix length T_{CP}. As illustrated in Figure 9.1, LTE defines two cyclic-prefix lengths, the *normal* cyclic prefix and an *extended* cyclic prefix, corresponding to seven and six OFDM symbols per slot respectively. The exact cyclic-prefix lengths, expressed in the basic time unit T_s, are given in Figure 9.1. It can be noted that, in the case of the normal cyclic prefix, the cyclic-prefix length for the first OFDM symbol of a slot is somewhat larger compared to the remaining OFDM symbols. The reason for this is simply to fill the entire 0.5 ms slot, as the number of basic time units T_s per slot (15360) is not divisible by seven.

The reasons for defining two cyclic-prefix lengths for LTE are twofold:

- A longer cyclic prefix, although less efficient from a cyclic-prefix-overhead point of view, may be beneficial in specific environments with extensive delay spread, for example in very large cells. It is important to have in mind, though, that a longer cyclic prefix is not necessarily beneficial in the case of large cells, even if the delay spread is very extensive in such cases. If, in large cells, link performance is limited by noise rather than by signal corruption due to residual time dispersion not covered by the cyclic prefix, the additional robustness to radio-channel time dispersion, due to the use of a longer cyclic prefix, may not justify the corresponding additional energy overhead of a longer cyclic prefix.

[1] This is valid for the "normal" downlink subframes. As described in Section 9.5.2, the *special subframe* present in the case of TDD operation is not divided into two slots but rather into three fields. However, the number of OFDM symbols in the special subframe is the same as for the normal subframes illustrated in Figure 9.1.

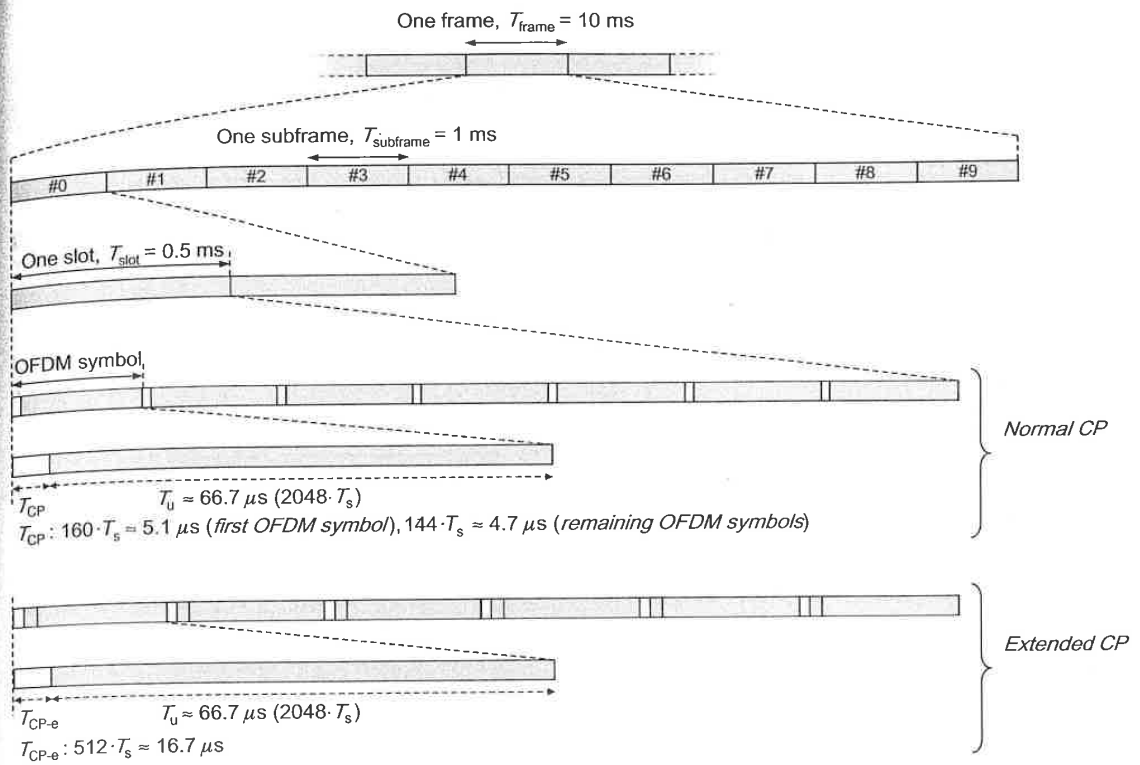


FIGURE 9.1

LTE time-domain structure.

- As already discussed in Chapter 3, in the case of MBSFN-based multicast/broadcast transmission, the cyclic prefix should not only cover the main part of the actual channel time dispersion, but also the timing difference between the transmissions received from the cells involved in the MBSFN transmission. In the case of MBSFN operation, the extended cyclic prefix is therefore used.

It should be noted that different cyclic-prefix lengths may be used for different subframes within a frame. As an example, to be discussed further in Chapter 15, MBSFN-based multicast/broadcast transmission is always confined to a limited set of subframes, in which case the use of the extended cyclic prefix, with its associated additional cyclic-prefix overhead, may only be applied to these subframes.[2]

A *resource element*, consisting of one subcarrier during one OFDM symbol, is the smallest physical resource in LTE. Furthermore, as illustrated in Figure 9.2, resource elements are grouped into *resource blocks*, where each resource block consists of 12 consecutive subcarriers in the frequency domain and one 0.5 ms slot in the time domain. Each resource block thus consists of $7 \cdot 12 = 84$

[2]The extended cyclic prefix is then actually applied only to the so-called MBSFN part of the MBSFN subframes (see Section 9.2).

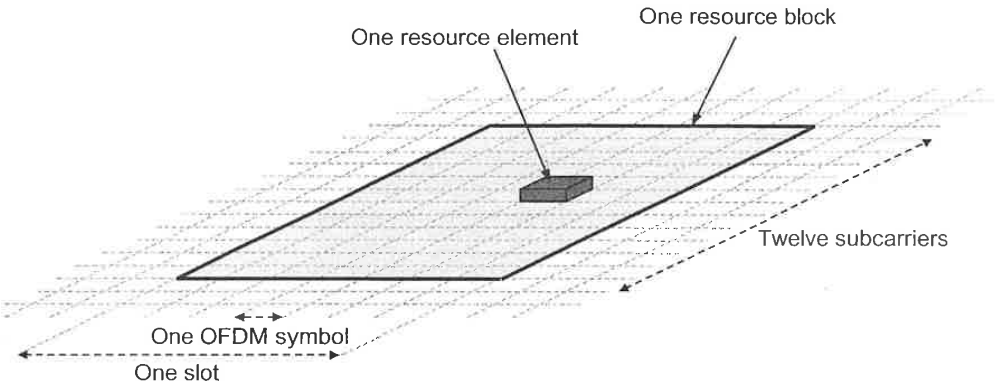


FIGURE 9.2

The LTE physical time–frequency resource.

resource elements in the case of a normal cyclic prefix and $6 \cdot 12 = 72$ resource elements in the case of an extended cyclic prefix.

Although resource blocks are defined over one slot, the basic time-domain unit for dynamic scheduling in LTE is one subframe, consisting of two consecutive slots. The reason for defining the resource blocks over one slot is that *distributed downlink transmission*, as described in Chapter 10 and *uplink frequency hopping* (described in Chapter 11) are defined on a slot or resource-block basis. The minimum scheduling unit, consisting of two time-consecutive resource blocks within one subframe (one resource block per slot), can be referred to as a *resource-block pair*.

The LTE physical-layer specification allows for a carrier to consist of any number of resource blocks in the frequency domain, ranging from a minimum of six resource blocks up to a maximum of 110 resource blocks. This corresponds to an overall transmission bandwidth ranging from roughly 1 MHz up to in the order of 20 MHz with very fine granularity and thus allows for a very high degree of LTE bandwidth flexibility, at least from a physical-layer-specification point of view. However, as mentioned in Chapter 7, LTE radio-frequency requirements are, at least initially, only specified for a limited set of transmission bandwidths, corresponding to a limited set of possible values for the number of resource blocks within a carrier. Also note that, in LTE release 10, the total bandwidth of the transmitted signal can be as large as 100 MHz by aggregating multiple carriers, as mentioned in Chapter 7 and further described in Section 9.3.

The resource-block definition above applies to both the downlink and uplink transmission directions. However, there is a minor difference between the downlink and uplink in terms of where the carrier center frequency is located in relation to the subcarriers.

In the downlink (upper part of Figure 9.3), there is an unused *DC-subcarrier* that coincides with the carrier center frequency. The reason why the DC-subcarrier is not used for downlink transmission is that it may be subject to disproportionately high interference due, for example, to local-oscillator leakage.

On the other hand, in the uplink (lower part of Figure 9.3), no unused DC-subcarrier is defined and the center frequency of an uplink carrier is located *between* two uplink subcarriers. The presence of an unused DC-carrier in the center of the spectrum would have prevented the assignment of the entire cell bandwidth to a single terminal and still retain the assumption of mapping to consecutive

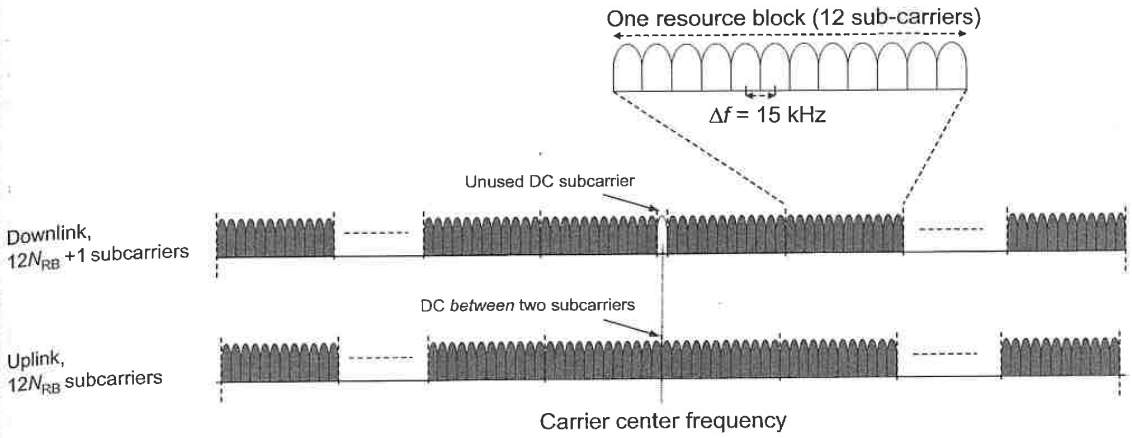


FIGURE 9.3

Frequency-domain structure for LTE.

inputs of the OFDM modulator, something that is needed to retain the low-cubic-metric property of the DFTS-OFDM modulation used for uplink data transmission.

9.2 NORMAL SUBFRAMES AND MBSFN SUBFRAMES

In LTE, each downlink subframe (and the DwPTS in the case of TDD; see Section 9.5.2 for a discussion of the TDD frame structure) is normally divided into a *control region*, consisting of the first few OFDM symbols, and a *data region*, consisting of the remaining part of the subframe. The detailed usage of the resource elements in the two regions will be discussed in detail in Chapter 10; at this stage all we need to know is that the control region carries L1/L2 signaling necessary to control uplink and downlink data transmissions.

Additionally, already from the first release of LTE, so-called *MBSFN subframes* have also been defined. The original intention with MBSFN subframes was, as indicated by the name, to support MBSFN transmission, as described in Chapter 15. However, MBSFN subframes have also been found to be useful in other contexts, for example as part of relaying functionality, as discussed in Chapter 16. Hence, MBSFN subframes are therefore better seen as a generic tool and not related only to MBSFN transmission.

An MBSFN subframe, illustrated in Figure 9.4, consists of a control region of length one or two OFDM symbols, which is in essence identical to its counterpart in a normal subframe, followed by an *MBSFN region* whose contents depend on the usage of the MBSFN subframe. The reason for keeping the control region also in MBSFN subframes is, for example, to transmit control signaling necessary for uplink transmissions. All terminals, from LTE release 8 and onwards, are capable of receiving the control region of an MBSFN subframe. This is the reason why MBSFN subframes have been found useful as a generic tool to introduce, in a backwards-compatible way, new types of signaling and transmission not part of an earlier release of the LTE radio-access specification. Terminals not capable of receiving transmissions in the MBSFN region will simply ignore those transmissions.

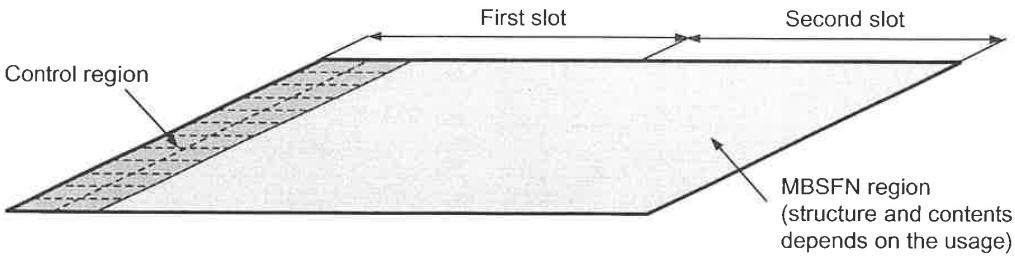


FIGURE 9.4

Resource-block structure for MBSFN subframes, assuming normal cyclic prefix for the control region and extended cyclic prefix for the MBSFN region.

Information about the set of subframes that are configured as MBSFN subframes in a cell is provided as part of the system information. In principle, an arbitrary pattern of MBSFN subframes can be configured with the pattern repeating after 40 ms.[3] However, as information necessary to operate the system (to be more specific, synchronization signals, system information and paging, all of which will be discussed in detail in later chapters) needs to be transmitted in order for terminals to find and connect to a cell, subframes where such information is provided cannot be configured as MBSFN subframes. Therefore, subframes 0, 4, 5, and 9 for FDD and subframes 0, 1, 5, and 6 for TDD cannot be configured as MBSFN subframes, leaving the remaining six subframes as candidates for MBSFN subframes.

9.3 CARRIER AGGREGATION

The possibility for *carrier aggregation* was introduced in LTE release 10. In the case of carrier aggregation, multiple LTE carriers, each with a bandwidth up to 20 MHz, can be transmitted in parallel to/from the same terminal, thereby allowing for an overall wider bandwidth and correspondingly higher per-link data rates. In the context of carrier aggregation, each carrier is referred to as a *component carrier*[4] as, from an RF point-of-view, the entire set of aggregated carriers can be seen as a single (RF) carrier.

Up to five component carriers, possibly of different bandwidths up to 20 MHz, can be aggregated allowing for overall transmission bandwidths up to 100 MHz. A terminal capable of carrier aggregation may receive or transmit simultaneously on multiple component carriers. Each component carrier can also be accessed by an LTE terminal from earlier releases – that is, component carriers are *backwards compatible*. Thus, in most respects and unless otherwise mentioned, the physical-layer description in the following chapters applies to each component carrier separately in the case of carrier aggregation.

[3]The reason for this value is that both the 10 ms frame length and the 8 ms hybrid-ARQ round-trip time are factors in the 40 ms, which is important for relaying operation (see also Chapter 16).

[4]In the specifications, the term *cell* is used instead of component carrier, but as the term cell is something of a misnomer in the uplink case, the term component carrier is used herein.

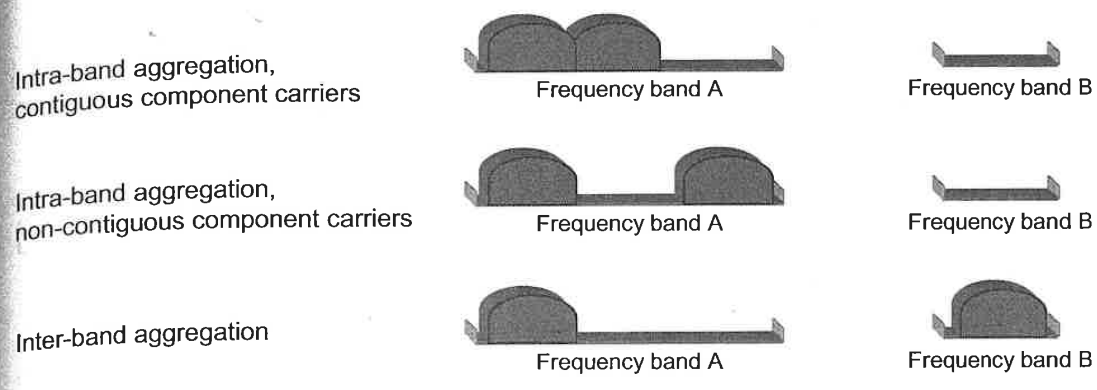


FIGURE 9.5

Different types of carrier aggregation.

It should be noted that aggregated component carriers do not need to be contiguous in the frequency domain. Rather, with respect to the frequency location of the different component carriers, three different cases can be identified (see also Figure 9.5):

- Intra-band aggregation with frequency-contiguous component carriers
- Intra-band aggregation with non-contiguous component carriers
- Inter-band aggregation with non-contiguous component carriers.

The possibility to aggregate non-adjacent component carriers allows for exploitation of a fragmented spectrum; operators with a fragmented spectrum can provide high-data-rate services based on the availability of a wide overall bandwidth even though they do not posses a single wideband spectrum allocation. Except from an RF point of view there is no difference between the three different cases outlined in Figure 9.5 and they are all supported by the basic LTE release-10 specification. However, the complexity of RF implementation is vastly different, with the first case being the least complex. Thus, although spectrum aggregation is supported by the physical-layer and protocol specifications, the actual implementation will be strongly constrained, including specification of only a limited number of aggregation scenarios and aggregation over a dispersed spectrum only being supported by the most advanced terminals.

A terminal capable of carrier aggregation has one downlink *primary* component carrier and an associated uplink primary component carrier. In addition, it may have one or several *secondary* component carriers in each direction. Different terminals may have different carriers as their primary component carrier – that is, the configuration of the primary component carrier is terminal specific. The association between the downlink primary carrier and the corresponding uplink primary carrier is cell specific and signaled as part of the system information. This is similar to the case without carrier aggregation, although in the latter case the association is trivial. The reason for such an association is, for example, to determine to which uplink component carrier a certain scheduling grant transmitted on the downlink relates without having to explicitly signal the component-carrier number.

All idle mode procedures apply to the primary component carrier only or, expressed differently, carrier aggregation with additional secondary carriers configured only applies to terminals in the RRC_CONNECTED state. Upon connection to the network, the terminal performs the related procedures such

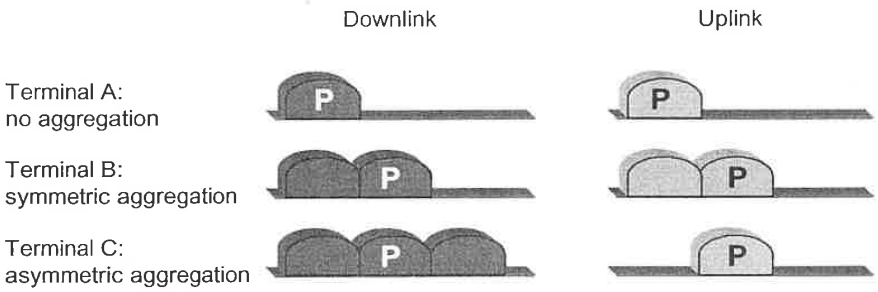


FIGURE 9.6

Examples of carrier aggregation ("P" denotes the primary component carrier).

as cell search and random access (see Chapter 14 for a detailed description of these procedures) following the same steps as in the absence of carrier aggregation. Once the communication between the network and the terminal is established, additional secondary component carriers can be configured.

The fact that carrier aggregation is terminal specific – that is, different terminals may be configured to use different sets of component carriers – is useful not only from a network perspective to balance the load across component carriers, but also to handle different capabilities between terminals. Some terminals may be able to transmit/receive on multiple component carriers, while other terminals may do so on only a single carrier. This is an obvious consequence of being able to serve terminals from earlier releases at the same time as a carrier-aggregation-capable release-10 terminal, but it also allows for different capabilities in terms of carrier aggregation for different terminals as well as a differentiation between downlink and uplink carrier-aggregation capability. For example, a terminal may be capable of two component carriers in the downlink but of only a single component carrier – that is, no carrier aggregation – in the uplink, as is the case for terminal C in Figure 9.6. Note also that the primary component-carrier configuration can differ between terminals. Asymmetric carrier aggregation can also be useful to handle different spectrum allocations, for example if an operator has more spectrum available for downlink transmissions than uplink transmissions.

In release 10, only downlink-heavy asymmetries are supported – that is, the number of uplink component carriers configured for a terminal is always equal to or smaller than the number of configured downlink component carriers. Uplink-heavy asymmetries are less likely to be of practical interest and would also complicate the overall control signaling structure, as in such a case multiple uplink component carriers would need to be associated with the same downlink component carrier.

Carrier aggregation is supported for both FDD and TDD, although all component carriers need to have the same duplex scheme. Furthermore, in the case of TDD, the uplink–downlink configuration should be the same across component carriers. The special subframe configuration can be different for the different components carriers though, as long as the resulting downlink–uplink switch time is sufficiently large.

9.4 FREQUENCY-DOMAIN LOCATION OF LTE CARRIERS

In principle, an LTE carrier could be positioned anywhere within the spectrum and, actually, the basic LTE physical-layer specification does not say anything about the exact frequency location of an LTE

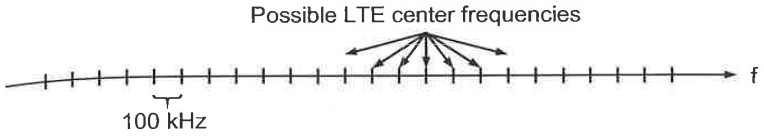


FIGURE 9.7

LTE carrier raster.

carrier, including the frequency band. However, in practice, there is obviously a need for restrictions on where an LTE carrier can be positioned in the frequency domain.

- In the end, an LTE terminal must be implemented and RF-wise such a terminal can obviously only support certain frequency bands. The frequency bands for which LTE is specified to operate are discussed in Chapter 17.
- After being activated, an LTE terminal has to search for a network-transmitted carrier within the frequency bands supported by the terminal. In order for that carrier search not to take an unreasonably long time, there is a need to limit the set of frequencies to be searched.

For this reason it is assumed that, within each supported frequency band, LTE carriers may exist on a 100 kHz *carrier raster* or *carrier grid* – that is, the carrier center frequency can be expressed as $m \cdot 100$ kHz, where m is an integer (see Figure 9.7).

In the case of carrier aggregation, multiple carriers can be transmitted to/from the same terminal. In order for the different component carriers to be accessible also by earlier-release terminals, each component carrier should fall on the 100 kHz carrier grid. However, in the case of carrier aggregation, there is an additional constraint that the carrier spacing between adjacent component carriers should be a multiple of the 15 kHz subcarrier spacing to allow transmission/reception with a single FFT.[5] Thus, in the case of carrier aggregation, the carrier spacing between the different component carriers should be a multiple of 300 kHz, the smallest carrier spacing being a multiple of both 100 kHz (the raster grid) and 15 kHz (the subcarrier spacing). A consequence of this is that there will always be a small number of unused subcarriers between two component carriers, even when they are located as close as possible to each other, as illustrated in Figure 9.8.

9.5 DUPLEX SCHEMES

Spectrum flexibility is one of the key features of LTE. In addition to the flexibility in transmission bandwidth, LTE also supports operation in both paired and unpaired spectrum by supporting both FDD- and TDD-based duplex operation with the time–frequency structures illustrated in Figure 9.9. Although the time-domain structure is, in most respects, the same for FDD and TDD, there are some differences, most notably the presence of a *special subframe* in the case of TDD. The special subframe is used to provide the necessary guard time for downlink–uplink switching, as discussed below.

[5]This is obviously only relevant for component carriers that are contiguous in the frequency domain. Furthermore, in the case of independent frequency errors between component carriers, separate FFTs may be needed at the receiver.

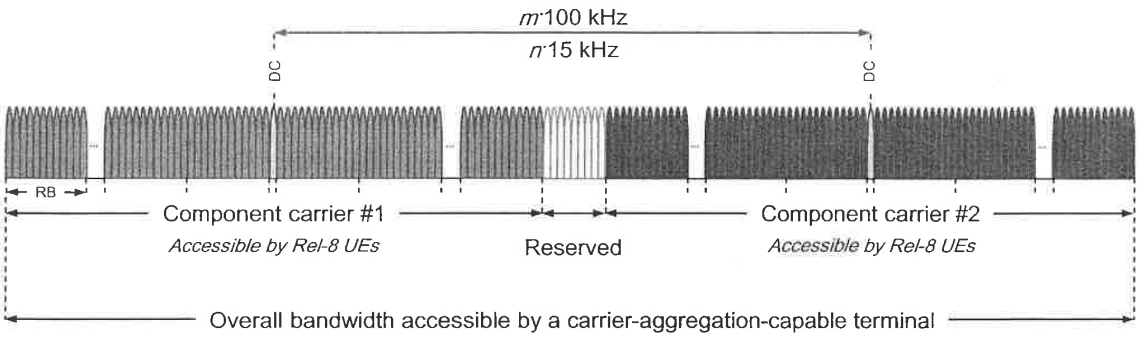


FIGURE 9.8

LTE carrier raster and carrier aggregation.

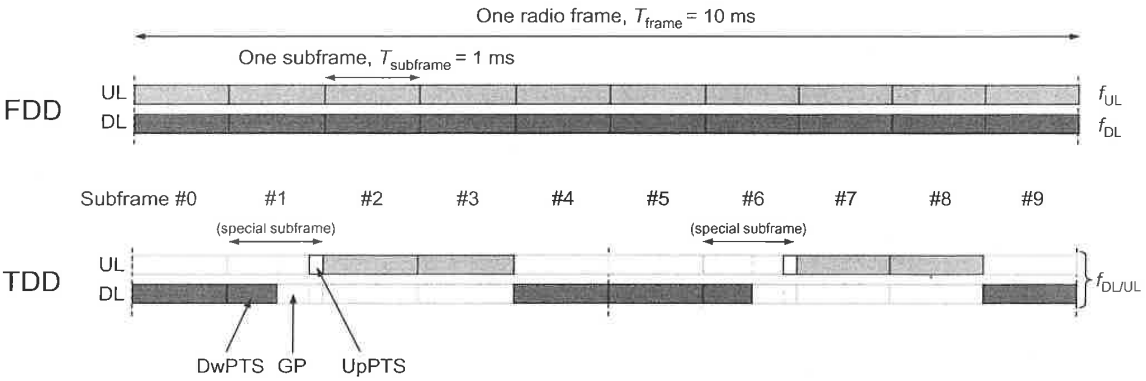


FIGURE 9.9

Uplink/downlink time–frequency structure for FDD and TDD.

9.5.1 Frequency-Division Duplex (FDD)

In the case of FDD operation (upper part of Figure 9.9), there are two carrier frequencies, one for uplink transmission (f_{UL}) and one for downlink transmission (f_{DL}). During each frame, there are thus ten uplink subframes and ten downlink subframes, and uplink and downlink transmission can occur simultaneously within a cell. Isolation between downlink and uplink transmissions is achieved by transmission/reception filters, known as duplex filters, and a sufficiently large *duplex separation* in the frequency domain.

Even if uplink and downlink transmission can occur simultaneously within a cell in the case of FDD operation, a terminal may be capable of *full-duplex* operation or only *half-duplex* operation for a certain frequency band, depending on whether or not it is capable of simultaneous transmission/reception. In the case of full-duplex capability, transmission and reception may also occur simultaneously at a terminal, whereas a terminal capable of only half-duplex operation cannot transmit and receive simultaneously. As mentioned in Chapter 7, supporting only half-duplex operation allows for simplified terminal implementation due to relaxed duplex-filter requirements. This applies especially

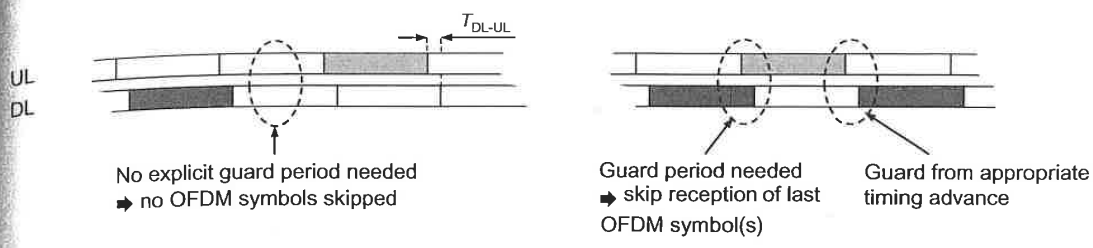


FIGURE 9.10

Guard time at the terminal for half-duplex FDD.

for certain frequency bands with a narrow duplex gap. Hence, full duplex support is *frequency-band dependent* such that a terminal may support only half-duplex operation in certain frequency bands while being capable of full-duplex operation in the remaining supported bands. It should be noted that full/half-duplex capability is a property of the *terminal*; the base station is operating in full duplex irrespective of the terminal capabilities. Hence, as the relevant transmission structures and timing relations are identical between full-duplex and half-duplex FDD, a single cell may simultaneously support a mixture of full-duplex and half-duplex FDD terminals. Half-duplex operation has an impact on the sustained data rates that can be provided to/from a single mobile terminal as it cannot transmit in all uplink subframes, but the cell capacity is hardly affected as typically it is possible to schedule different terminals in uplink and downlink in a given subframe.

Since a half-duplex terminal is not capable of simultaneous transmission and reception, the scheduling decisions must take this into account and half-duplex operation can be seen as a scheduling restriction, as will be discussed in more detail in Chapter 13. If a terminal is scheduled such that downlink reception in one subframe immediately precedes a subframe of uplink transmission, a guard time is necessary for the terminal to switch from reception to transmission. This is created in such cases by allowing the terminal to skip receiving[6] the last OFDM symbol(s) in the downlink subframe, as illustrated in Figure 9.10.

9.5.2 Time-Division Duplex (TDD)

In the case of TDD operation (lower part of Figure 9.9), there is a single carrier frequency only and uplink and downlink transmissions are separated in the time domain on a cell basis. As seen in the figure, some subframes are allocated for uplink transmissions and some subframes for downlink transmission, with the switch between downlink and uplink occurring in the *special subframe* (subframe 1 and, in some cases, subframe 6). Different asymmetries in terms of the amount of resources – that is, subframes – allocated for uplink and downlink transmission respectively are provided through the seven different downlink/uplink configurations illustrated in Figure 9.11. As seen in the figure, subframes 0 and 5 are always allocated for downlink transmission while subframe 2 is always allocated for uplink transmissions. The remaining subframes (except the special subframe; see below)

[6]The impact on the decoding performance can be accounted for by the rate adaptation mechanism. For very high data rates, the performance impact may be somewhat larger due to the downlink interleaving structure, in which case it may be preferable not to schedule an uplink transmission immediately after a downlink reception.

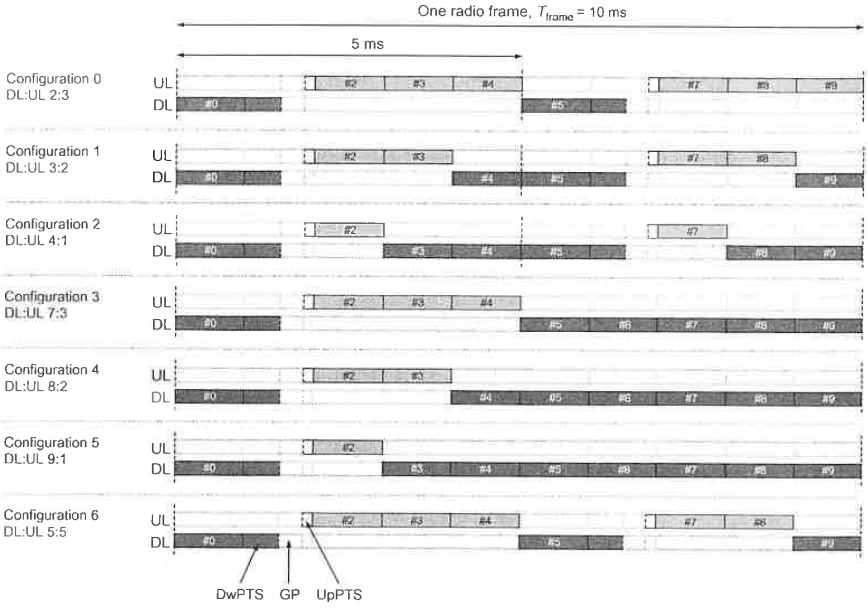


FIGURE 9.11

Different downlink/uplink configurations in the case of TDD.

can then be flexibly allocated for downlink or uplink transmission depending on the configured downlink/uplink configuration. To avoid severe interference between downlink and uplink transmissions in different cells, neighboring cells typically have the same downlink/uplink configuration. This makes it difficult to change the downlink/uplink configuration dynamically, for example on a frame-by-frame basis. Therefore, the current LTE specification assumes that the downlink/uplink configuration is relatively static. It can be changed on a very slow basis though, in order to adapt to changing traffic patterns. It could also, in principle, be different in different areas, for example to match different traffic patterns, although inter-cell interference needs to be carefully addressed in this case.

As the same carrier frequency is used for uplink and downlink transmission, both the base station and the terminal need to switch from transmission to reception and vice versa. The switch between downlink and uplink occurs in the special subframe, which is split into three parts: a downlink part (DwPTS), a guard period (GP), and an uplink part (UpPTS).

The DwPTS is in essence treated as a normal downlink subframe,[7] although the amount of data that can be transmitted is smaller due to the reduced length of the DwPTS compared to a normal subframe. The UpPTS, however, is not used for data transmission due to the very short duration. Instead, it can be used for channel sounding or random access. It can also be left empty, in which case it serves as extra guard period.

An essential aspect of any TDD system is the possibility to provide a sufficiently large *guard period* (or guard time), where neither downlink nor uplink transmissions occur. This guard period is necessary for switching from downlink to uplink transmission and vice versa and, as already mentioned, is

[7] For the shortest DwPTS duration of three OFDM symbols, DwPTS cannot be used for PDSCH transmission.

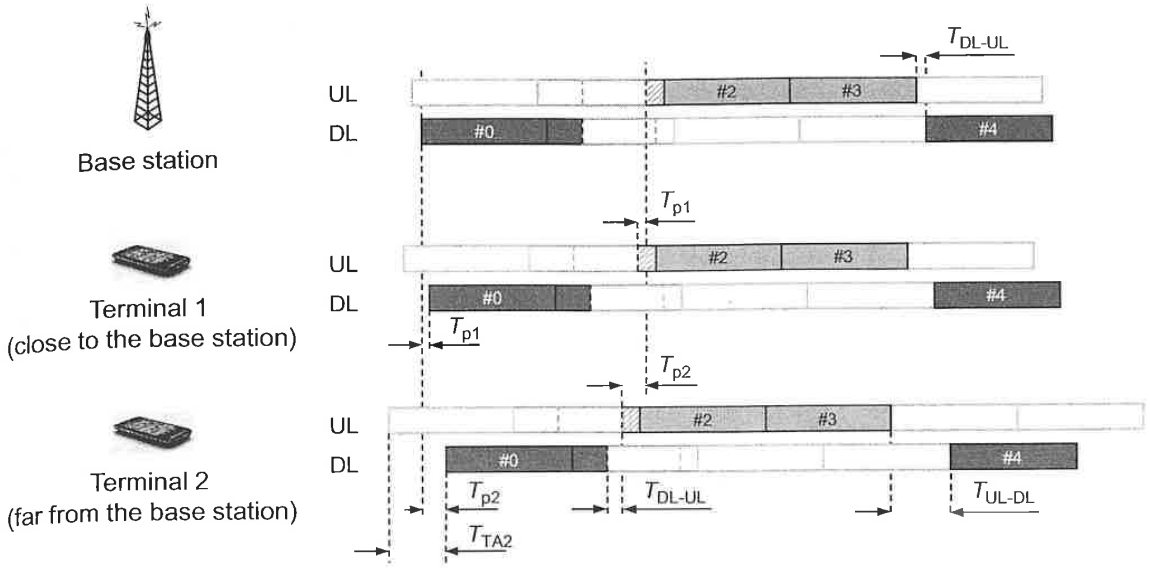


FIGURE 9.12

Timing relation for TDD operation.

obtained from the special subframe. The required length of the guard period depends on several factors. First, it should be sufficiently large to provide the necessary time for the circuitry in base stations and the terminals to switch from downlink to uplink. Switching is typically relatively fast, of the order of 20 microseconds, and in most deployments does not significantly contribute to the required guard time.

Secondly, the guard time should also ensure that uplink and downlink transmissions do not interfere at the base station. This is handled by advancing the uplink timing at the terminals such that, at the base station, the last uplink subframe before the uplink-to-downlink switch ends before the start of the first downlink subframe. The uplink timing of each terminal can be controlled by the base station by using the timing advance mechanism, as will be elaborated upon in Chapter 11. Obviously, the guard period must be large enough to allow the terminal to receive the downlink transmission and switch to transmission before it starts the (timing-advanced) uplink transmission. In essence, some of the guard period of the special subframe is "moved" from the downlink-to-uplink switch to the uplink-to-downlink switch by the timing-advance mechanism. This is illustrated in Figure 9.12. As the timing advance is proportional to the distance to the base station, a larger guard period is required when operating in large cells compared to small cells.

Finally, the selection of the guard period also needs to take interference between base stations into account. In a multi-cell network, inter-cell interference from downlink transmissions in neighboring cells must decay to a sufficiently low level before the base station can start to receive uplink transmissions. Hence, a larger guard period than is motivated by the cell size itself may be required as the last part of the downlink transmissions from distant base stations otherwise may interfere with uplink reception. The amount of guard period depends on the propagation environments, but in some cases the inter-base-station interference is a non-negligible factor when determining the guard period.

Table 9.1 Resulting Guard Period in OFDM Symbols for Different DwPTS and UpPTS Lengths (Normal Cyclic Prefix)

	DwPTS								
	12	11		10		9		3	
GP	1	1	2	2	3	3	4	9	10
UpPTS	1	2	1	2	1	2	1	2	1

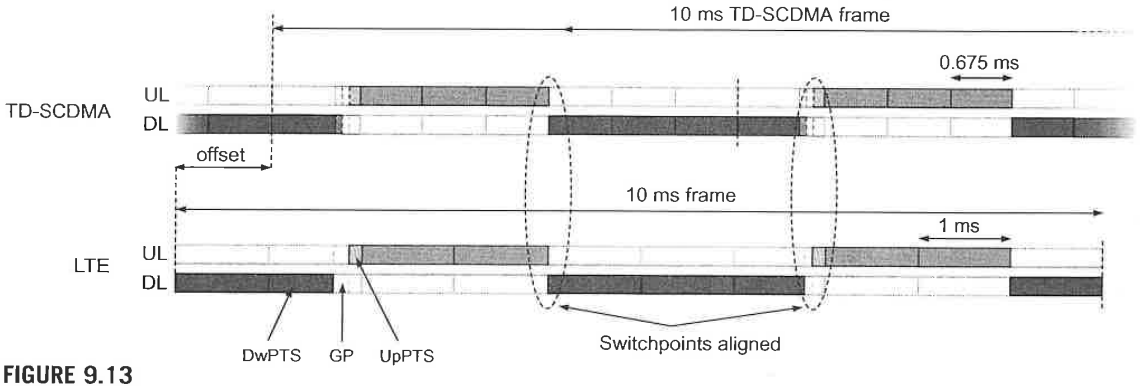


FIGURE 9.13

Coexistence between TD-SCDMA and LTE.

From the above discussion, it is clear that a sufficient amount of configurability of the guard period is needed to meet different deployment scenarios. Therefore, a set of DwPTS/GP/UpPTS configurations is supported as shown in Table 9.1, where each configuration corresponds to a given length of the three fields in the special subframes. The DwPTS/GP/UpPTS configuration used in the cell is signaled as part of the system information.

9.5.3 LTE and TD-SCDMA Coexistence

In addition to supporting a wide range of different guard periods, an important aspect in the design of TDD in LTE was to simplify coexistence with and migration from systems based on the 3GPP TD-SCDMA standard (see Chapter 19 for details on TD-SCDMA). Basically, to handle inter-system interference from two different but co-sited TDD systems operating close in frequency, it is necessary to align the switch-points between the two systems. Since LTE supports configurable lengths of the DwPTS field, the switch-points of LTE and TD-SCDMA (or any other TDD system) can be aligned, despite the different subframe lengths used in the two systems. Aligning the switch-points between TD-SCDMA and LTE is the technical reason for splitting the special subframe into the three fields DwPTS/GP/UpPTS instead of locating the switch-point at the subframe boundary. An example of LTE/TD-SCDMA coexistence is given in Figure 9.13.

The set of possible lengths of DwPTS/GP/UpPTS is selected to support common coexistence deployments, as well as to provide a high degree of guard-period flexibility for the reasons discussed earlier. The UpPTS length is one or two OFDM symbols and the DwPTS length can vary from three[8] to twelve OFDM symbols, resulting in guard periods ranging from one to ten OFDM symbols. The resulting guard period for the different DwPTS and UpPTS configurations supported is summarized in Table 9.1 for the case of normal cyclic prefix. As discussed earlier, the DwPTS can be used for downlink data transmission, while the UpPTS can be used for sounding or random access only, due to its short duration.

[8]The smallest DwPTS length is motivated by the location of the primary synchronization signal in the DwPTS (see Chapter 14).

CHAPTER

Downlink Physical-Layer Processing

10

In Chapter 8, the LTE radio-interface architecture was discussed with an overview of the functions and characteristics of the different protocol layers. Chapter 9 then gave an overview of the basic time–frequency structure of LTE transmissions, including the structure of the OFDM time–frequency grid being the fundamental physical resource on both uplink and downlink.

This chapter will provide a more detailed description of the downlink physical-layer functionality, including the transport-channel processing (Section 10.1), downlink reference signals (Section 10.2), details on downlink multi-antenna transmission (Section 10.3), and downlink L1/L2 control signaling (Section 10.4). Chapter 11 will provide a corresponding description for the *uplink* transmission direction. The later chapters will then go further into the details of some specific uplink and downlink functions and procedures.

10.1 TRANSPORT-CHANNEL PROCESSING

As described in Chapter 8, transport channels provide the interface between the MAC layer and the physical layer. As also described, for the LTE downlink there are four different types of transport channels defined, the Downlink Shared Channel (DL-SCH), the Multicast Channel (MCH), the Paging Channel (PCH), and the Broadcast Channel (BCH). This section provides a description of the physical-layer processing applied to DL-SCH transport channels, including the mapping to the physical resource – that is, to the resource elements of the OFDM time–frequency grid. DL-SCH is the main downlink transport-channel type in LTE and is used for transmission of user-specific higher-layer information, both user data and dedicated control information, as well as part of the downlink system information. The DL-SCH physical-layer processing is to a large extent applicable also to MCH and PCH transport channels, although with some additional constraints. On the other hand, the physical-layer processing, and the transmission structure in general, for the BCH is quite different. BCH transmission is described in Chapter 14 as part of the discussion on LTE system information.

10.1.1 Processing Steps

The different steps of the DL-SCH physical layer processing are outlined in Figure 10.1. In the case of carrier aggregation – that is, transmission on multiple component carriers in parallel to the same terminal – the transmissions on the different carriers correspond to separate transport channels with separate and more or less independent physical-layer processing. The transport-channel processing outlined in Figure 10.1 and the discussion below is thus valid also in the case of carrier aggregation.

4G LTE/LTE-Advanced for Mobile Broadband.
© 2011 Erik Dahlman, Stefan Parkvall & Johan Sköld. Published by Elsevier Ltd. All rights reserved.

Samsung Exhibit 1019, Page 115 of 195

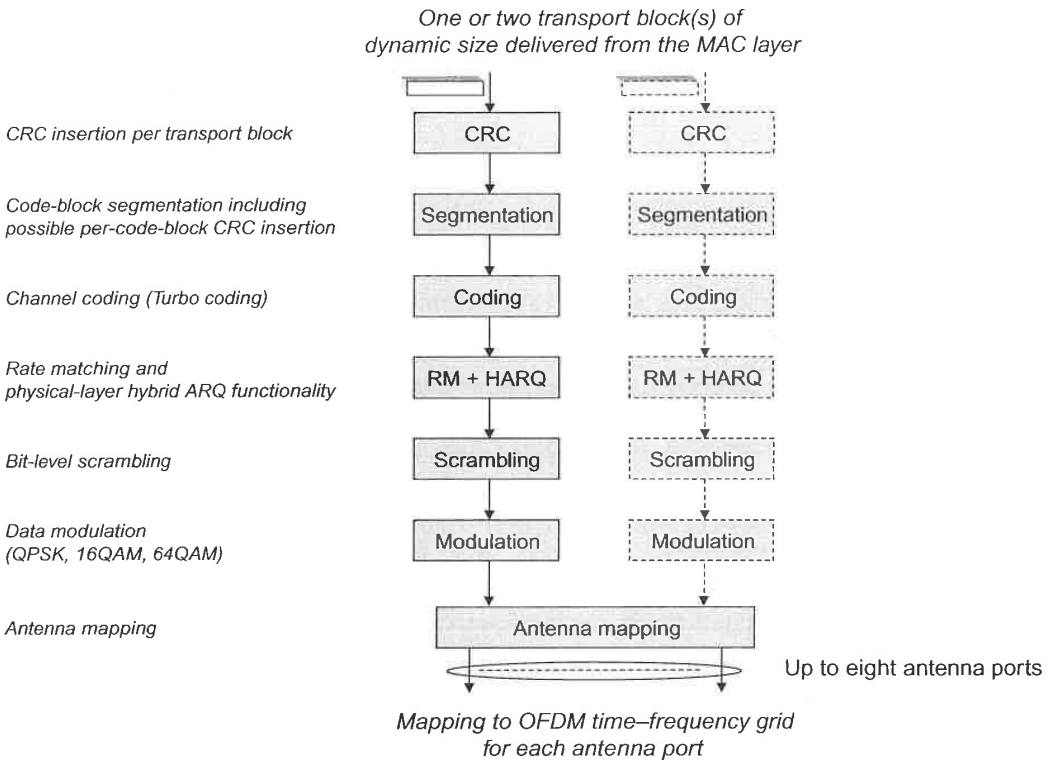


FIGURE 10.1

Physical-layer processing for DL-SCH.

Within each *Transmission Time Interval* (TTI), corresponding to one subframe of length 1 ms, up to two transport blocks of dynamic size are delivered to the physical layer and transmitted over the radio interface for each component carrier. The number of transport blocks transmitted within a TTI depends on the configuration of the multi-antenna transmission scheme, as described in Section 10.3.

- In the case of no *spatial multiplexing* there is at most a single transport block in a TTI.
- In the case of spatial multiplexing, with transmission on multiple *layers* in parallel to the same terminal, there are two transport blocks within a TTI.[1]

10.1.1.1 CRC Insertion Per Transport Block

In the first step of the physical-layer processing, a 24-bit CRC is calculated for and appended to each transport block. The CRC allows for receiver-side detection of errors in the decoded transport block. The corresponding error indication can, for example, be used by the downlink hybrid-ARQ protocol as a trigger for requesting retransmissions.

[1]This is true for initial transmissions. In the case of hybrid-ARQ retransmissions there may also be cases when a single transport block is transmitted over multiple layers, as discussed, for example, in Section 10.3.2.

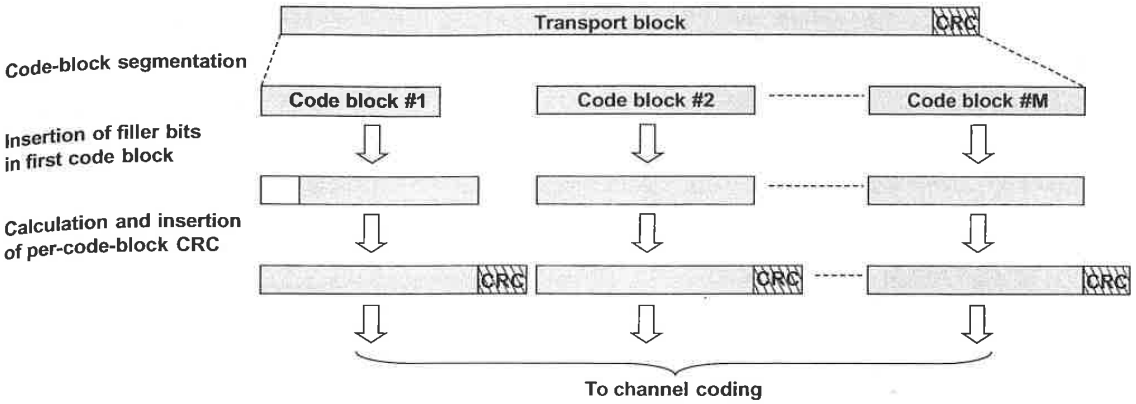


FIGURE 10.2

Code-block segmentation and per-code-block CRC insertion.

10.1.1.2 Code-Block Segmentation and Per-Code-Block CRC Insertion

The LTE Turbo-coder internal interleaver is only defined for a limited number of code-block sizes, with a maximum block size of 6144 bits. If the transport block, including the transport-block CRC, exceeds this maximum code-block size, *code-block segmentation*, illustrated in Figure 10.2, is applied before the Turbo coding. Code-block segmentation implies that the transport block is segmented into smaller *code blocks*, the sizes of which should match the set of code-block sizes supported by the Turbo coder.

In order to ensure that a transport block of arbitrary size can be segmented into code blocks that match the set of available code-block sizes, the specification includes the possibility to insert "dummy" *filler bits* at the head of the first code block. However, the set of transport-block sizes currently defined for LTE has been selected so that filler bits are not needed.

As can be seen in Figure 10.2, code-block segmentation also implies that an additional CRC (also of length 24 bits, but different compared to the transport-block CRC described above) is calculated for and appended to each code block. Having a CRC per code block allows for early detection of correctly decoded code blocks and correspondingly early termination of the iterative decoding of that code block. This can be used to reduce the terminal processing effort and corresponding energy consumption. In the case of a single code block no additional code-block CRC is applied.

One could argue that, in case of code-block segmentation, the transport-block CRC is redundant and implies unnecessary overhead as the set of code-block CRCs should indirectly provide information about the correctness of the complete transport block. However, code-block segmentation is only applied to large transport blocks for which the relative extra overhead due to the additional transport-block CRC is small. The transport-block CRC also adds additional error-detection capabilities and thus further reduces the risk for undetected errors in the decoded transport block.

Information about the transport-block size is provided to the terminal as part of the scheduling assignment transmitted on the PDCCH control channel, as described in Section 10.4.4. Based on this information, the terminal can determine the code-block size and number of code blocks. The terminal receiver can thus, based on the information provided in the scheduling assignment, straightforwardly undo the code-block segmentation and recover the decoded transport blocks.

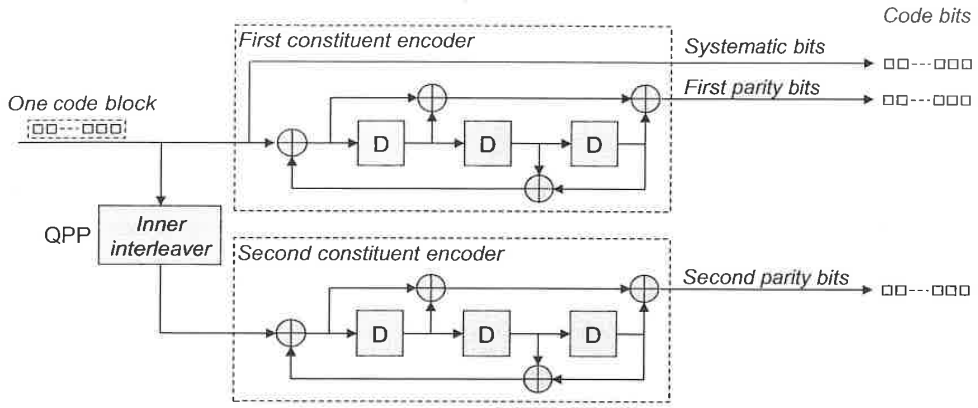


FIGURE 10.3

LTE Turbo encoder.

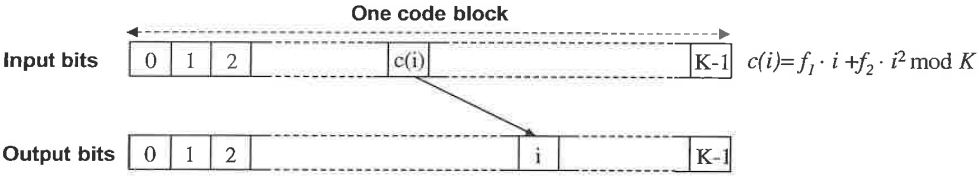


FIGURE 10.4

Principles of QPP-based interleaving.

10.1.1.3 Channel Coding

Channel coding for DL-SCH (as well as for PCH and MCH) is based on Turbo coding [53], with encoding according to Figure 10.3. The encoding consists of two rate-1/2, eight-state constituent encoders, implying an overall code rate of 1/3, in combination with QPP-based[2] interleaving [69]. As illustrated in Figure 10.4, the QPP interleaver provides a mapping from the input (non-interleaved) bits to the output (interleaved) bits according to the function:

$$c(i) = f_1 \cdot i + f_2 \cdot i^2 \bmod K,$$

where i is the index of the bit at the output of the interleaver, $c(i)$ is the index of the same bit at the input of the interleaver, and K is the code-block/interleaver size. The values of the parameters f_1 and f_2 depend on the code-block size K. The LTE specification lists all supported code-block sizes, ranging from a minimum of 40 bits to a maximum of 6144 bits, together with the associated values for the parameters f_1 and f_2. Thus, once the code-block size is known, the Turbo-coder inner interleaving, as well as the corresponding de-interleaving at the receiver side, can straightforwardly be carried out.

A QPP-based interleaver is *maximum contention free* [70], implying that the decoding can be parallelized without the risk for contention when the different parallel processes are accessing the

[2]QPP = Quadrature Permutation Polynomial.

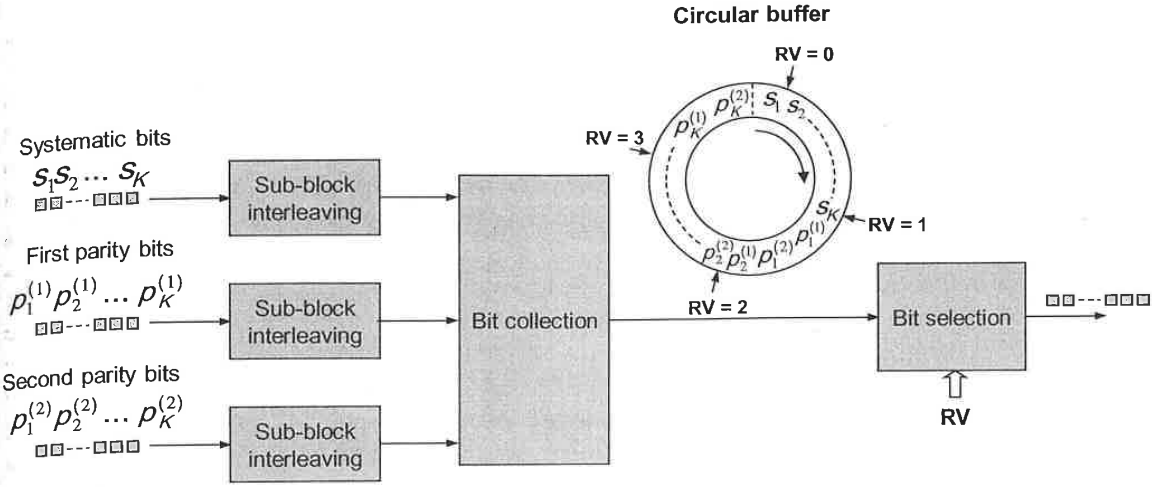


FIGURE 10.5

Rate matching and hybrid-ARQ functionality.

interleaver memory. For the very high data rates supported by LTE, the improved possibilities for parallel processing offered by QPP-based interleaving can substantially simplify the Turbo-encoder/decoder implementation.

10.1.1.4 Rate Matching and Physical-Layer Hybrid-ARQ Functionality

The task of the rate-matching and physical-layer hybrid-ARQ functionality is to extract, from the blocks of code bits delivered by the channel encoder, the exact set of code bits to be transmitted within a given TTI/subframe.

As illustrated in Figure 10.5, the outputs of the Turbo encoder (systematic bits, first parity bits, and second parity bits) are first separately interleaved. The interleaved bits are then inserted into what can be described as a circular buffer with the systematic bits inserted first, followed by alternating insertion of the first and second parity bits.

The bit selection then extracts consecutive bits from the circular buffer to an extent that matches the number of available resource elements in the resource blocks assigned for the transmission. The exact set of bits to extract depends on the *redundancy version* (RV) corresponding to different starting points for the extraction of coded bits from the circular buffer. As can be seen, there are four different alternatives for the redundancy version. The transmitter/scheduler selects the redundancy version and provides information about the selection as part of the scheduling assignment (see Section 10.4.4).

Note that the rate-matching and hybrid-ARQ functionality operates on the full set of code bits corresponding to one transport block and not separately on the code bits corresponding to a single code block.

10.1.1.5 Bit-Level Scrambling

LTE downlink scrambling implies that the block of code bits delivered by the hybrid-ARQ functionality is multiplied (*exclusive-or* operation) by a bit-level *scrambling sequence*. Without downlink scrambling,

the channel decoder at the terminal could, at least in principle, be equally matched to an interfering signal as to the target signal, thus being unable to properly suppress the interference. By applying different scrambling sequences for neighboring cells, the interfering signal(s) after descrambling is (are) randomized, ensuring full utilization of the processing gain provided by the channel code. Thus, the bit scrambling essentially serves the same purpose as the scrambling applied at chip level after the direct-sequence spreading in DS-CDMA-based systems such as WCDMA/HSPA. Fundamentally, channel coding can be seen as "advanced" spreading providing processing gain similar to direct-sequence spreading but also additional coding gain.

In LTE, downlink scrambling is applied to all transport channels as well as to the downlink L1/L2 control signaling. For all downlink transport-channel types except MCH, as well as for the L1/L2 control signaling, the scrambling sequences differ between neighboring cells (*cell-specific scrambling*) to ensure interference randomization between the cells. This is achieved by having the scrambling sequences depend on the physical-layer cell identity (Chapter 14). In contrast, in the case of MBSFN-based transmission using MCH, the same scrambling should be applied to all cells taking part in the MBSFN transmission – that is, all cells within the so-called *MBSFN area* (see Chapter 15).

10.1.1.6 Data Modulation

The downlink data modulation transforms the block of scrambled bits to a corresponding block of complex modulation symbols. The set of modulation schemes supported for the LTE downlink includes QPSK, 16QAM, and 64QAM, corresponding to two, four, and six bits per modulation symbol respectively.

10.1.1.7 Antenna Mapping

The antenna mapping jointly processes the modulation symbols corresponding to the one or two transport blocks and maps the result to different *antenna ports*. The antenna mapping can be configured in different ways corresponding to different multi-antenna transmission schemes, including transmit diversity, beam-forming, and spatial multiplexing. As indicated in Figure 10.1, LTE supports transmission using up to eight antenna ports depending on the exact multi-antenna transmission scheme. More details about LTE downlink multi-antenna transmission are provided in Section 10.3.

Note that the antenna ports referred to above do not necessarily correspond to specific physical antennas. Rather, an antenna port is a more general concept introduced, for example, to allow for beam-forming using multiple physical antennas without the terminal being aware of the beam-forming carried out at the transmitter side.

At least for the downlink, an antenna port can be seen as corresponding to the transmission of a *reference signal* (Section 10.2). Any data transmission from the antenna port can then rely on that reference signal for channel estimation for coherent demodulation. Thus, if the same reference signal is transmitted from multiple physical antennas, these physical antennas correspond to a *single* antenna port. Similarly, if two different reference signals are transmitted from the same set of physical antennas, this corresponds to *two separate antenna ports*.

It should be noted that the LTE specification actually has a somewhat more general definition of an antenna port, essentially just saying that two received signals can be assumed to have experienced the same overall channel, including any joint processing at the transmitter side, if and only if they have been transmitted on the same antenna port.

10.1.1.8 Resource-Block Mapping

The resource-block mapping takes the symbols to be transmitted on each antenna port and maps them to the resource elements of the set of resource blocks assigned by the MAC scheduler for the transmission. As described in Chapter 9, each resource block consists of 84 resource elements (twelve subcarriers during seven OFDM symbols).[3] However, some of the resource elements within a resource block will not be available for the transport-channel transmission as they are occupied by:

- different types of downlink reference signals, as described in Section 10.2;
- downlink L1/L2 control signaling (one, two, or three OFDM symbols at the head of each subframe), as described in Section 10.4[4].

Furthermore, as will be described in Chapter 14, within some resource blocks, additional resource elements are reserved for the transmission of *synchronization signals* as well as for the transmission of the BCH transport channel.

In the TDD special subframe (Section 9.5.2), mapping is limited to the DwPTS.

10.1.2 Localized and Distributed Resource Mapping

As already discussed in Chapter 7, when deciding what set of resource blocks to use for transmission to a specific terminal, the network may take the downlink channel conditions in both the time and frequency domains into account. Such time/frequency-domain channel-dependent scheduling, taking channel variations – for example, due to frequency-selective fading – into account, may significantly improve system performance in terms of achievable data rates and overall cell throughput.

However, in some cases downlink channel-dependent scheduling is not suitable to use or is not practically possible:

- For low-rate services such as voice, the feedback signaling associated with channel-dependent scheduling may lead to extensive relative overhead.
- At high mobility (high terminal speed), it may be difficult or even practically impossible to track the instantaneous channel conditions to the accuracy required for channel-dependent scheduling to be efficient.

In such situations, an alternative means to handle radio-channel frequency selectivity is to achieve frequency diversity by distributing a downlink transmission in the frequency domain.

One way to distribute a downlink transmission in the frequency domain, and thereby achieve frequency diversity, is to assign multiple non-frequency-contiguous resource blocks for the transmission to a terminal. LTE allows for such *distributed resource-block allocation* by means of *resource allocation types 0 and 1* (see Section 10.4.4). However, although sufficient in many cases, distributed resource-block allocation by means of these resource-allocation types has certain drawbacks:

- For both types of resource allocations, the minimum size of the allocated resource can be as large as four resource-block pairs and may thus not be suitable when resource allocations of smaller sizes are needed.
- In general, both these resource-allocation methods are associated with a relatively large PDCCH payload.

[3] 72 resource elements in the case of extended cyclic prefix.
[4] In MBSFN subframes, the control region is limited to a maximum of two OFDM symbols.

In contrast, *resource-allocation type 2* (Section 10.4.4) always allows for the allocation of a single resource-block pair and is also associated with a relatively small PDCCH payload size. However, resource allocation type 2 only allows for the allocation of resource blocks that are contiguous in the frequency domain. In addition, regardless of the type of resource allocation, frequency diversity by means of distributed resource-block allocation will only be achieved in the case of resource allocations larger than one resource-block pair.

In order to provide the possibility for distributed resource-block allocation in the case of resource-allocation type 2, as well as to allow for distributing the transmission of a single resource-block pair in the frequency domain, the notion of a *Virtual Resource Block* (VRB) has been introduced for LTE.

What is being provided in the resource allocation is the resource allocation in terms of VRB pairs. The key to distributed transmission then lies in the mapping from VRB pairs to *Physical Resource Block* (PRB) pairs – that is, to the actual physical resource used for transmission.

The LTE specification defines two types of VRBs: *localized* VRBs and *distributed* VRBs. In the case of localized VRBs, there is a direct mapping from VRB pairs to PRB pairs, as illustrated in Figure 10.6. However, in the case of distributed VRBs, the mapping from VRB pairs to PRB pairs is more elaborate in the sense that:

- consecutive VRBs are not mapped to PRBs that are consecutive in the frequency domain; and
- even a single VRB pair is distributed in the frequency domain.

The basic principle of distributed transmission is outlined in Figure 10.7 and consists of two steps:

- A mapping from VRB pairs to PRB pairs such that consecutive VRB pairs are not mapped to frequency-consecutive PRB pairs (first step of Figure 10.7). This provides frequency diversity between consecutive VRB pairs. The spreading in the frequency domain is done by means of a block-based "interleaver" operating on resource-block pairs.

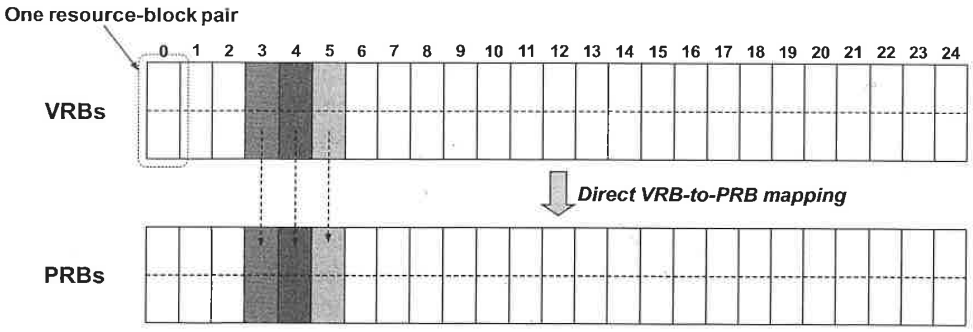


FIGURE 10.6

VRB-to-PRB mapping in the case of localized VRBs. A cell bandwidth corresponding to 25 resource blocks is assumed.

- A split of each resource-block pair such that the two resource blocks of the resource-block pair are transmitted with a certain frequency gap in between (second step of Figure 10.7). This also provides frequency diversity for a single VRB pair. This step can be seen as the introduction of frequency hopping on a slot basis.

Whether the VRBs are localized (and thus mapped according to Figure 10.6) or distributed (mapped according to Figure 10.7) is indicated on the associated PDCCH in the case of type 2 resource allocation. Thus, it is possible to dynamically switch between distributed and localized transmission and also mix distributed and localized transmission for different terminals within the same subframe.

The exact size of the frequency gap in Figure 10.7 depends on the overall downlink cell bandwidth according to Table 10.1. These gaps have been chosen based on two criteria:

1. The gap should be of the order of half the downlink cell bandwidth in order to provide good frequency diversity also in the case of a single VRB pair.
2. The gap should be a multiple of P^2, where P is the size of a *resource-block group* as defined in Section 10.4.4 and used for resource allocation types 0 and 1. The reason for this constraint is to ensure a smooth coexistence in the same subframe between distributed transmission as described above and transmissions based on downlink allocation types 0 and 1.

Due to the constraint that the gap size should be a multiple of P^2, the gap size will in most cases deviate from exactly half the cell bandwidth. In these cases, not all resource blocks within the cell bandwidth can be used for distributed transmission. As an example, for a cell bandwidth corresponding to 25 resource blocks (the example in Figure 10.7) and a corresponding gap size equal to 12 according to Table 10.1, the 25th resource-block pair cannot be used for distributed transmission.

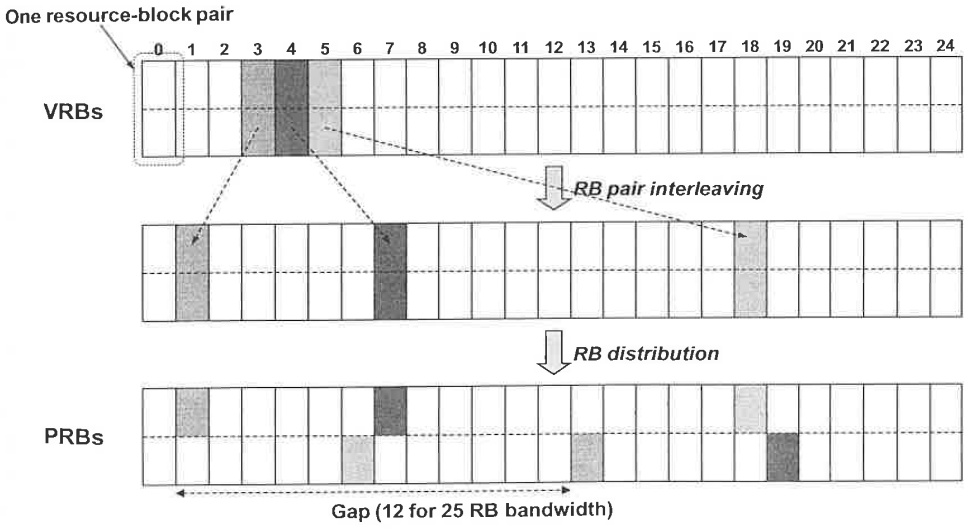


FIGURE 10.7

VRB-to-PRB mapping in the case of distributed VRBs. A cell bandwidth corresponding to 25 resource blocks is assumed.

Table 10.1 Gap Size for Different Cell Bandwidths (Number of Resource Blocks)

	Bandwidth									
	6	**7–8**	**9–10**	**11**	**12–19**	**20–26**	**27–44**	**45–63**	**64–79**	**80–110**
P	1	1	1	2	2	2	3	3	4	4
Gap size	3	4	5	4	8	12	18	27	32	48

Table 10.2 Second Gap Size for Different Cell Bandwidths (Only Applicable for Cell Bandwidths of 50 RBs and Beyond)

	Bandwidth	
	50–63	**64–110**
Gap size	9	16

As another example, for a cell bandwidth corresponding to 50 resource blocks (gap size equal to 27 according to Table 10.1), only 46 resource blocks would be available for distributed transmission.

In addition to the gap size outlined in Table 10.1, for wider cell bandwidths (50 RBs and beyond), there is a possibility to use a second, smaller frequency gap with a size of the order of one-fourth of the cell bandwidth (see Table 10.2). The use of the smaller gap enables restriction of the distributed transmission to only a part of the overall cell bandwidth. Selection between the larger gap according to Table 10.1 and the smaller gap according to Table 10.2 is indicated by an additional bit in the resource allocation on PDCCH.

10.2 DOWNLINK REFERENCE SIGNALS

Downlink reference signals are predefined signals occupying specific resource elements within the downlink time–frequency grid. The LTE specification includes several types of downlink reference signals that are transmitted in different ways and used for different purposes by the receiving terminal:

- *Cell-specific reference signals* (CRS) are transmitted in every downlink subframe and in every resource block in the frequency domain, thus covering the entire cell bandwidth. The cell-specific reference signals can be used by the terminal for channel estimation for coherent demodulation of any downlink physical channel except for PMCH and for PDSCH in the case of *transmission modes 7, 8, or 9*. As described in Section 10.3, these transmission modes correspond to so-called *non-codebook-based precoding*. The cell-specific reference signals can also be used by the terminal to acquire *channel-state information* (CSI). Finally, terminal measurements on cell-specific reference signals are used as the basis for cell-selection and handover decisions.

- *Demodulation reference signals* (DM-RS), also sometimes referred to as *UE-specific reference signals*, are specifically intended to be used by terminals for channel estimation for PDSCH in the case of transmission modes 7, 8, or 9. The label "*UE-specific*" relates to the fact that each demodulation reference signal is intended for channel estimation by a single terminal. That specific reference signal is then only transmitted within the resource blocks assigned for PDSCH transmission to that terminal.
- *CSI reference signals* (CSI-RS) are specifically intended to be used by terminals to acquire *channel-state information* (CSI) in the case when demodulation reference signals are used for channel estimation.[5] CSI-RS have a significantly lower time/frequency density, thus implying less overhead, compared to the cell-specific reference signals.
- *MBSFN reference signals* are intended to be used for channel estimation for coherent demodulation in the case of MCH transmission using so-called *MBSFN* (see Chapter 15 for more details on MCH transmission).
- *Positioning reference signals* were introduced in LTE release 9 to enhance *LTE positioning functionality*, more specifically to support the use of terminal measurements on multiple LTE cells to estimate the geographical position of the terminal. The positioning reference symbols of a certain cell can be configured to correspond to empty resource elements in neighboring cells, thus enabling high-SIR conditions when receiving neighbor-cell positioning reference signals.

10.2.1 Cell-Specific Reference Signals

Cell-specific reference signals, introduced in the first release of LTE (release 8), are the most basic downlink reference signals in LTE. There can be one, two, or four cell-specific reference signals in a cell, defining one, two, or four corresponding antenna ports.

10.2.1.1 Structure of a Single Reference Signal

Figure 10.8 illustrates the structure of a single cell-specific reference signal. As can be seen, it consists of *reference symbols* of predefined values inserted within the first and third last[6] OFDM symbol of each slot and with a frequency-domain spacing of six subcarriers. Furthermore, there is a frequency-domain staggering of three subcarriers for the reference symbols within the third last OFDM symbol. Within each resource-block pair, consisting of 12 subcarriers during one 1 ms subframe, there are thus eight reference symbols.

In general, the values of the reference symbols vary between different reference-symbol positions and also between different cells. Thus, a cell-specific reference signal can be seen as a *two-dimensional cell-specific sequence*. The period of this sequence equals one 10 ms frame. Furthermore, regardless of the cell bandwidth, the reference-signal sequence is defined assuming the maximum possible LTE carrier bandwidth corresponding to 110 resource blocks in the frequency domain. Thus, the basic reference-signal sequence has a length of 8880 symbols.[7] For cell bandwidths less than the maximum possible value, only the reference symbols within that bandwidth are actually transmitted. The reference symbols in the center part of the band will therefore be the same, regardless of the actual cell bandwidth. This allows for the terminal to estimate the channel corresponding to the center

[5] More specifically, CSI-RS are only used in the case of transmission mode 9.
[6] This corresponds to the fifth and fourth OFDM symbols of the slot for normal and extended cyclic prefixes respectively.
[7] Four reference symbols per resource block, 110 resource blocks per slot, and 20 slots per frame.

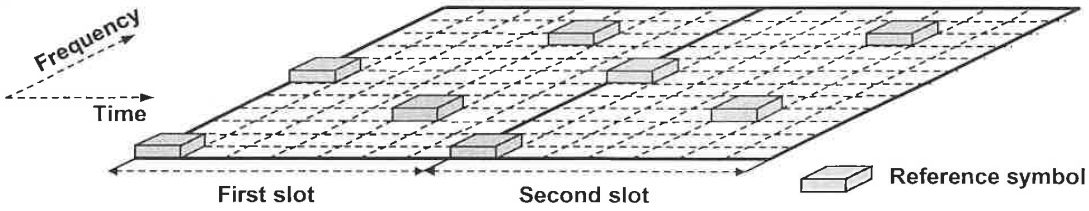


FIGURE 10.8

Structure of cell-specific reference signal within a pair of resource blocks.

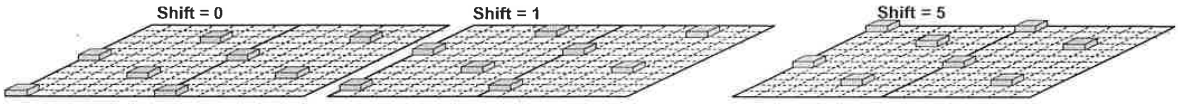


FIGURE 10.9

Different reference-signal frequency shifts.

part of the carrier, where, for example, the basic system information of the cell is transmitted on the BCH transport channel, without knowing the cell bandwidth. Information about the actual cell bandwidth, measured as number of resource blocks, is then provided on the BCH.

There are 504 different reference-signal sequences defined for LTE, where each sequence corresponds to one of 504 different *physical-layer cell identities*. As will be described in more detail in Chapter 14, during the so-called *cell-search procedure* the terminal detects the physical-layer identity of the cell as well as the cell frame timing. Thus, from the cell-search procedure, the terminal knows the reference-signal sequence of the cell (given by the physical-layer cell identity) as well as the start of the reference-signal sequence (given by the frame timing).

The set of reference-symbol positions outlined in Figure 10.8 is only one of six possible *frequency shifts* of the reference symbols, as illustrated in Figure 10.9. The frequency shift to use in a cell depends on the physical-layer identity of the cell such that each shift corresponds to 84 different cell identities. Thus, the six different frequency shifts jointly cover all 504 different cell identities. By properly assigning physical-layer cell identities to different cells, different reference-signal frequency shifts may be used in neighboring cells. This can be beneficial, for example, if the reference symbols are transmitted with higher energy compared to other resource elements, also referred to as *reference-signal power boosting*, in order to improve the reference-signal SIR. If reference signals of neighboring cells were transmitted using the same time/frequency resource, the boosted reference symbols of one cell would be interfered by equally boosted reference symbols of all neighboring cells,[8] implying no gain in the reference-signal SIR. However, if different frequency shifts are used for the reference-signal transmissions of neighboring cells, the reference symbols of one cell will at least partly be interfered by non-reference symbols of neighboring cells, implying an improved reference-signal SIR in the case of reference-signal boosting.

[8]This assumes that the cell transmissions are frame-timing aligned.

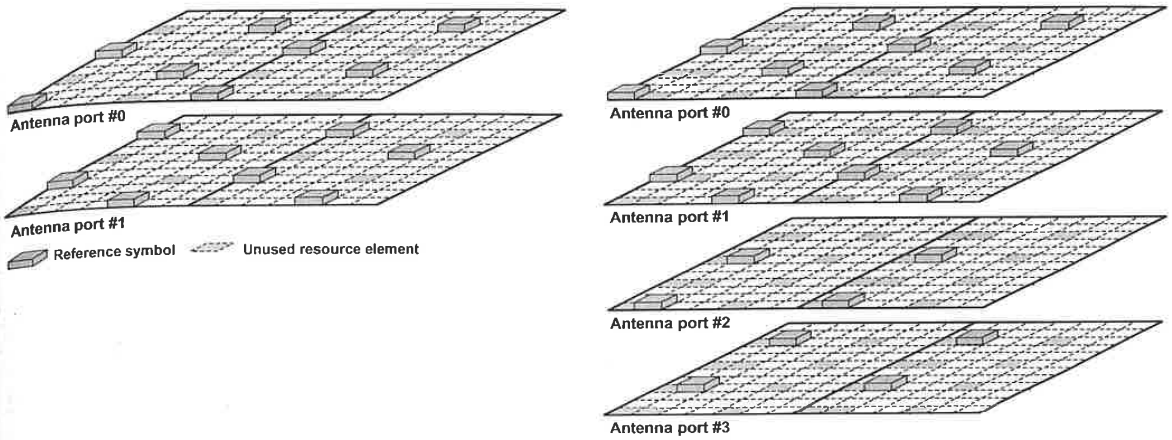


FIGURE 10.10

Structure of cell-specific reference signals in the case of multiple reference signals. Two reference signals corresponding to two antenna ports (left). Four reference signals corresponding to four antenna ports (right).

10.2.1.2 Multiple Reference Signals

Figure 10.10 illustrates the reference-signal structure in the case of multiple, more specifically two and four, cell-specific reference signals, and corresponding multiple antenna ports, within a cell:[9]

- In the case of two reference signals within a cell (left part of Figure 10.10), the second reference signal is frequency multiplexed with the first reference signal, with a frequency-domain offset of three subcarriers.
- In the case of four reference signals (right part of Figure 10.10), the third and fourth reference signals are frequency multiplexed and transmitted within the *second* OFDM symbol of each slot, thus being time multiplexed with the first and second reference signals.

Obviously, the reference-symbol density for the third and fourth reference signals is lower, compared to the density of the first and second reference signals. The reason for this is to reduce the reference-signal overhead in the case of four reference signals. More specifically, while the first and second reference signals each correspond to a relative overhead of approximately 5% (four reference symbols within a resource block consisting of a total of 84 resource elements), the relative overhead of the third and fourth reference signals is only half of that or approximately 2.5%. This obviously has an impact on the potential of the terminal to track very fast channel variations. However, this can be justified based on an expectation that, for example, high-order spatial multiplexing will mainly be applied to scenarios with low mobility.

It can also be noted that in a resource element carrying reference signals for a certain transmission port, nothing is being transmitted on the antenna ports corresponding to the other reference signals. Thus, a cell-specific reference signal is not interfered by transmissions on other antenna ports. Multi-antenna transmission schemes, such as spatial multiplexing, to a large extent rely on good channel

[9] It is not possible to configure a cell with three cell-specific reference signals.

estimates to suppress interference between the different layers at the receiver side. However, in the channel estimation itself there is obviously no such suppression. Reducing the interference to the reference signals of an antenna port is therefore important in order to allow for good channel estimation, and corresponding good interference suppression, at the receiver side.

Note that, in MBSFN subframes, only the reference signals in the two first OFDM symbols of the subframe, corresponding to the control region of the MBSFN subframe, are actually transmitted. Thus, there is no transmission of cell-specific reference signals within the MBSFN part of the MBSFM subframe.

10.2.2 Demodulation Reference Signals

In contrast to cell-specific reference signals, a demodulation reference signal (DM-RS) is intended for a specific terminal and is only transmitted in the resource blocks assigned for transmission to that terminal. Demodulation reference signals are intended to be used for channel estimation for PDSCH transmissions for the case when cell-specific reference signals are not to be used – that is, for non-codebook-based precoding (see Section 10.3.3).

Demodulation reference signals were introduced in the first release of LTE (release 8). However, use of demodulation reference signals was then limited to single-layer transmission – that is, no spatial multiplexing – with at most one reference signal for each terminal. In LTE release 9, transmission based on demodulation reference signals was extended to support also dual-layer transmission, requiring up to two simultaneous reference signals (one for each layer). Transmission based on demodulation reference signals was then further extended in LTE release 10 to support up to eight-layer transmission, corresponding to up to eight reference signals.

Actually, the DM-RS structure introduced in LTE release 9 was not a straightforward extension of the release-8 demodulation reference signals but rather a new structure, supporting both single-layer and dual-layer transmission. Already at the time of LTE release 9, it was relatively clear that the LTE radio-access technology should be further extended to support up to eight-layer spatial multiplexing in release 10. It was also quite clear that this extension would be difficult to achieve based on the release-8 DM-RS structure. Rather than extending the release-8 structure to support two reference signals and then introduce a completely new structure for release 10, it was decided instead to introduce a new, more future-proof structure already in release 9. Here we will focus on the DM-RS structure introduced in LTE release 9, including the release-10 extension to support up to eight simultaneous reference signals.

Figure 10.11 illustrates the DM-RS structure for the case of one or two reference signals.[10] As can be seen, there are 12 reference symbols within a resource-block pair. In contrast to cell-specific reference signals, for which the reference symbols of one antenna port correspond to unused resource elements for the other antenna ports (see Figure 10.10), in the case of two demodulation reference signals all 12 reference symbols in Figure 10.11 are transmitted *for both reference signals* – that is, *on both antenna ports*. Interference between the reference signals is instead handled by applying mutually orthogonal patterns, also referred to as *orthogonal cover codes* (OCC), to pairs of consecutive reference symbols, as illustrated in the lower right corner of the figure.

[10] In the case of TDD, the DM-RS structure is slightly modified in the DwPTS due to the shorter duration of the DwPTS compared with normal downlink subframes.

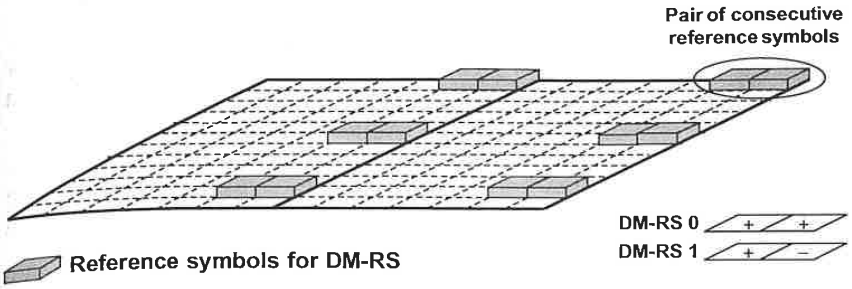


FIGURE 10.11

Structure of demodulation reference signals (DM-RS) for the case of one or two reference signals, including orthogonal cover codes to separate the two reference signals.

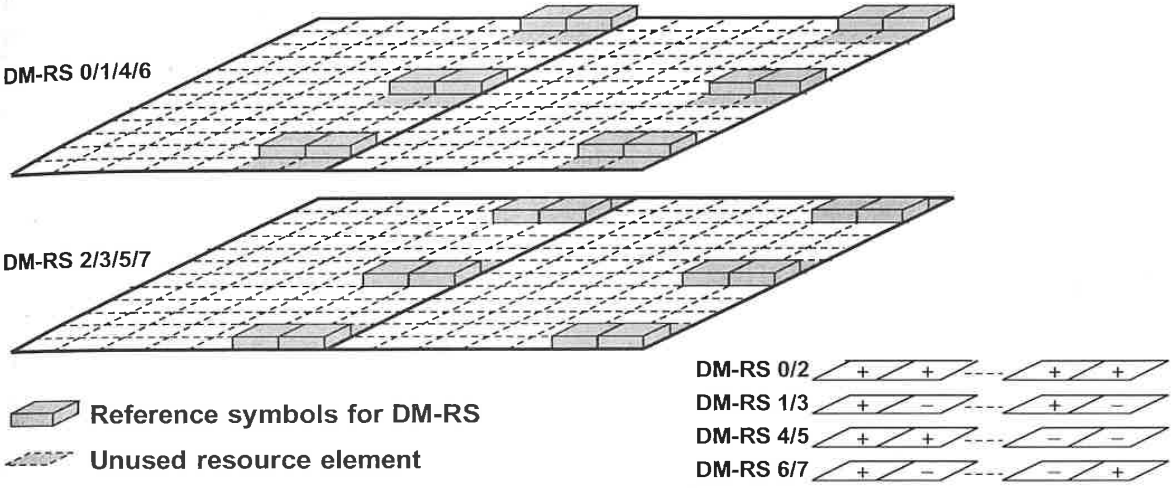


FIGURE 10.12

Demodulation-reference-signal structure for the case of more than two reference signals.

In addition to the mutually orthogonal patterns, one can also apply a pseudo-random sequence to the reference symbols. This sequence is the same for both reference signals and thus does not impact the orthogonality between the transmitted reference signals. Rather, the pseudo-random sequence is intended to separate different demodulation reference signals targeting different terminals in the case of so-called MU-MIMO transmission, as described in Section 10.3.4.1.

Figure 10.12 illustrates the extended DM-RS structure introduced in LTE release 10 to support more than two reference signals. In this case, there are 24 reference symbols within a resource-block pair. The reference signals are frequency multiplexed in groups of four reference signals while, within each group, the reference signals are separated by means of mutually orthogonal patterns covering four reference symbols (two pairs of consecutive reference symbols). It should be noted that orthogonality between the full set of eight reference signals requires that the channel does not vary over the set of reference symbols over which the orthogonal pattern spans. As the four reference symbols are

not consecutive in time, this implies somewhat stronger limitations on the amount of channel variations that can be tolerated without losing the reference-signal orthogonality. However, more than four demodulation reference signals is only applicable to the case of spatial multiplexing with more than four layers, a transmission configuration that is typically anyway only applicable to low-mobility scenarios. Also note that the orthogonal patterns are selected such that, for four or less reference signals, orthogonality is achieved already over pairs of reference symbols. Thus, for three and four reference signals, the constraints on channel variations are the same as for two reference signals according to Figure 10.11.

When demodulation reference signals are transmitted within a resource block, PDSCH mapping to the time–frequency grid of the resource block will be modified to avoid the resource elements in which the reference signals are transmitted. Although this modified mapping is not "understood" by earlier-release terminals not supporting demodulation reference signals, this is not a problem as demodulation reference signals will only be transmitted in resource blocks that are scheduled for PDSCH transmission to terminals of later releases supporting demodulation reference signals and thus "understanding" the modified PDSCH mapping. As will be seen below, the situation is somewhat different in the case of CSI-RS.

As the number of transmitted layers may vary dynamically, the number of transmitted demodulation reference signals may also vary. Thus, the transmission may dynamically change between the DM-RS structures in Figures 10.11 and 10.12. The terminal is informed about the number of transmitted layers (the "transmission rank") as part of the scheduling assignment and will thus know the DM-RS structure and associated PDSCH mapping for each subframe.

10.2.3 CSI Reference Signals

Support for CSI reference signals (CSI-RS) was introduced in LTE release 10. CSI-RS is intended to be used by terminals to acquire channel-state information when demodulation reference signals are used for channel estimation. More specifically, CSI-RS are to be used in transmission mode 9.[11]

As described in Section 10.2.1, the cell-specific reference signals, already available in the first release of LTE, can also be used to acquire channel-state information. There are at least two reasons why CSI-RS was introduced in release 10 as a complement to the already supported cell-specific reference signals:

- In LTE release 8 there could be at most four cell-specific reference signals in a cell. As already mentioned, multi-antenna transmission was extended in LTE release 10 to support downlink spatial multiplexing with up to eight layers, corresponding to eight transmit antennas at the base station. Thus, there was a need to extend the LTE CSI capabilities beyond what was possible with the release 8 cell-specific reference signals.
- The time-frequency density of the cell-specific reference signals was selected to allow for channel estimation for coherent demodulation even with the most extreme channel conditions, including very fast channel variations in both the time and frequency domains. As a consequence, each

[11] The reason why CSI-RS are not used for transmission modes 7 and 8 is simply that these transmission modes were introduced in LTE releases 8 and 9 respectively, and CSI-RS was not introduced until LTE release 10, together with the introduction of transmission mode 9.

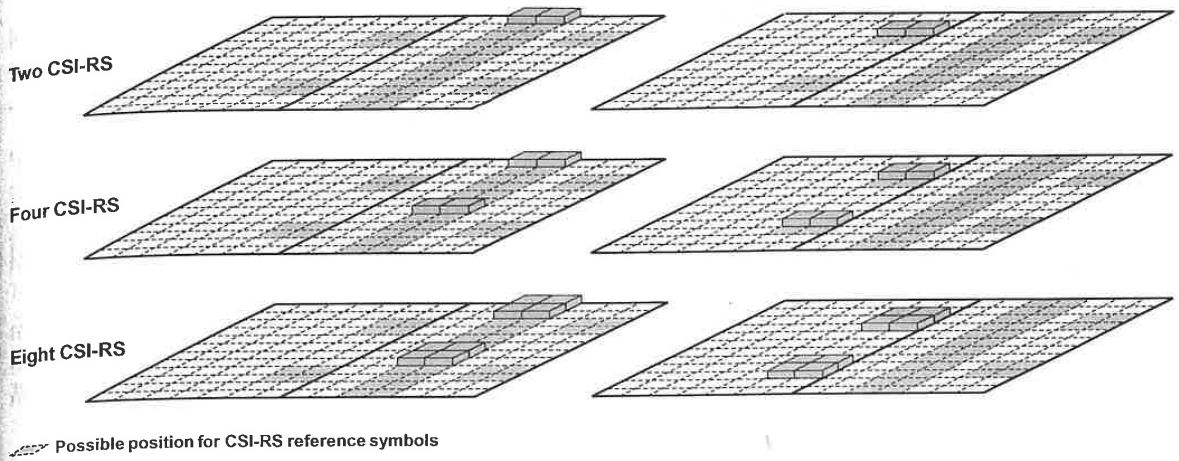


FIGURE 10.13

Examples of reference-signal positions for different number of CSI-RS within a cell. In the case of a single CSI-RS, the same structure as for two CSI-RS is used.

cell-specific reference signal implies a relatively high overhead (approximately 5% and 2.5% for the first/second and third/fourth reference signals respectively). On the other hand, in the case of a reference signal only targeting CSI, such high time-frequency density, and correspondingly relatively large overhead, is not necessary. Thus, rather than extending the cell-specific reference signals it was found to be more efficient to introduce a new type of reference signals (CSI-RS) only targeting CSI, with a flexible and, in general, lower time/frequency density and corresponding lower overhead per reference signal.

10.2.3.1 CSI-RS Structure

A cell can be configured with one, two, four, or eight CSI-RS. The exact CSI-RS structure, including the exact set of resource elements used for CSI-RS in a resource block, depends on the number of CSI-RS configured within the cell and may also be different for different cells. More specifically, within a resource-block pair there are 40 *possible* positions for the reference symbols of CSI-RS and, in a given cell, a subset of the corresponding resource elements is used for CSI-RS transmission, as described below and exemplified in Figure 10.13.[12]

- In the case of two CSI-RS within a cell, the CSI-RS consists of two consecutive reference symbols per resource-block pair, as illustrated in the upper part of Figure 10.13. The two CSI-RS are separated by applying mutually *orthogonal cover codes* (OCC) to the two reference symbols, similar to demodulation reference signals described in Section 10.2.2. Thus, there is a possibility for 20 different CSI-RS configurations, two of which are shown in Figure 10.13, based on the 40 *possible* reference-symbol positions.

[12]For TDD, there are additional *possible* locations for reference symbols for CSI-RS beyond what is illustrated in Figure 10.13.

- In the case of four/eight CSI-RS, the CSI-RS are pair-wise frequency multiplexed, as illustrated in the middle/lower part of Figure 10.13. Thus, there is a possibility for ten/five different CSI-RS configurations, each supporting four/eight CSI-RS.

In the case of a single CSI-RS, the same structure as for two CSI-RS (upper part of Figure 10.13) is used, although with only one of the orthogonal cover codes. The overhead of a single CSI-RS, in terms of occupied resource elements, is thus the same as for two CSI-RS.

In the time domain, the CSI-RS can be transmitted with different periods, ranging from a period of 5 ms (every fifth subframe) to 80 ms (every eighth frame). With a 5 ms period, the overhead per CSI-RS is thus roughly 0.12% per CSI-RS (one resource element per CSI-RS per resource-block pair, but only in every fifth subframe), with longer periods implying correspondingly less overhead.

The exact subframe in which CSI-RS is transmitted in a cell can also be configured, allowing for separation of CSI-RS transmissions between cells also in the time domain, in addition to using different sets of resource elements within the same subframe, as illustrated in Figure 10.13.

In subframes in which CSI-RS is transmitted, it is transmitted in every resource block in the frequency domain. In other words, a CSI-RS transmission covers the entire cell bandwidth.

As mentioned above, when demodulation reference signals are transmitted within a resource block, the corresponding resource elements on which the reference signals are transmitted are explicitly avoided when mapping PDSCH symbols to the resource block. This "modified" PDSCH mapping, which is obviously not "understood" by earlier-release terminals, is possible as demodulation reference signals can be assumed to be transmitted only in resource blocks in which terminals supporting such reference signals are scheduled – that is, terminals of release 10 or later.[13] Expressed alternatively, an earlier-release terminal can be assumed never to be scheduled in a resource block in which demodulation reference signals are transmitted and thus in which the modified PDSCH mapping is used.

The situation is somewhat different for CSI-RS. As CSI-RS is transmitted within all resource blocks in the frequency domain, it would imply a strong scheduler constraint to assume that release-8/9 terminals could never be scheduled in a resource block in which CSI-RS is transmitted. If the PDSCH mapping were modified to explicitly avoid the resource elements in which CSI-RS is transmitted, the mapping would not be recognized by a release-8/9 terminal. Instead, in the case of resource blocks scheduled to release-8/9 terminals, the PDSCH is mapped exactly according to release 8 – that is, the mapping is not modified to avoid the resource elements on which CSI-RS is to be transmitted. The CSI-RS is then simply transmitted on top of the corresponding PDSCH symbols.[14] This will obviously impact the PDSCH demodulation performance, as some PDSCH symbols will be highly corrupted. However, the remaining PDSCH symbols will not be impacted and the PDSCH will still be decodable, although with somewhat reduced performance.

On the other hand, if a release-10 terminal is scheduled in a resource block in which CSI-RS is transmitted, the PDSCH mapping is modified to explicitly avoid the resource elements on which the CSI-RS is transmitted, similar to demodulation reference signals. Thus, if CSI-RS is transmitted in a resource block, the PDSCH mapping to that resource block will be somewhat different depending on the release of the terminal being scheduled in the resource block.

[13] Partly also for terminals of release 9, but then only for a maximum of two DM-RS.

[14] In practice, the base station would probably not transmit PDSCH in these resource elements in order to avoid interference to the CSI-RS transmission. The key thing is that the mapping of the remaining PDSCH symbols is in line with release 8.

It should be noted that release-8 mapping also has to be used for transmission of, for example, system information and paging messages, as such transmissions must be possible to receive also by release-8/9 terminals.

10.2.3.2 Muted CSI-RS

Figure 10.13 illustrated the set of resource elements that can *potentially* be used for CSI-RS, a subset of which is used to create the different CSI-RS within a cell. In normal operation, the remaining resource elements highlighted in Figure 10.13 are used for PDSCH transmission within the cell.

However, it is also possible to additionally configure one or several subsets of the resource elements outlined in Figure 10.13 as *muted CSI-RS*. A muted CSI-RS has the same structure as a normal (non-muted) CSI-RS except that nothing is actually transmitted on the corresponding resource elements. A muted CSI-RS can thus be seen as a normal (non-muted) CSI-RS *with zero power*.

The intention with muted CSI-RS is to be able to create "transmission holes" corresponding to actual CSI-RS transmissions in other (neighboring) cells. This serves two purposes:

1. To make it possible for a terminal to receive CSI-RS of neighboring cells, without being severely interfered by transmissions in its own cell. CSI estimation on neighboring cells would be of interest if support for different multi-cell-transmission techniques such as *CoMP* (see Chapter 20) were introduced in future LTE releases.
2. To reduce the interference to CSI-RS transmissions in other cells. This is especially applicable to so-called *heterogeneous network deployments* where overlapping cell layers with cells of substantially different power are to coexist in the same spectrum. Interference handling in heterogeneous deployments is further discussed in Chapter 13.

There can be one or multiple sets of muted CSI-RS in a cell. In the case of interference avoidance to neighboring cells (corresponding to point 2 above), there is typically sufficient with a single set of muted CSI-RS as one should typically avoid interference to a single lower-power cell layer. When muted CSI-RS are used to remove own-cell interference in order to be able to better receive CSI-RS from neighboring cells (point 1 above), multiple sets of muted CSI-RS are typically needed as different neighboring cells can be assumed to use mutually different CSI-RS configurations.

10.3 MULTI-ANTENNA TRANSMISSION

As illustrated in Figure 10.14, multi-antenna transmission in LTE can, in general, be described as a mapping from the output of the data modulation to the different antennas ports. The input to the antenna mapping thus consists of the modulation symbols (QPSK, 16QAM, 64QAM) corresponding to the one or two transport blocks. To be more specific, there is one transport block per TTI except for spatial multiplexing, in which case there may be two transport blocks per TTI (see also Section 10.3.2).

The output of the antenna mapping is a set of symbols for each antenna port. The symbols of each antenna port are subsequently applied to the OFDM modulator – that is, mapped to the basic OFDM time–frequency grid corresponding to that antenna port.

The different multi-antenna transmission schemes correspond to different so-called *transmission modes*. There are currently nine different transmission modes defined for LTE. They differ in terms of the specific structure of the antenna mapping of Figure 10.14 but also in terms of what reference

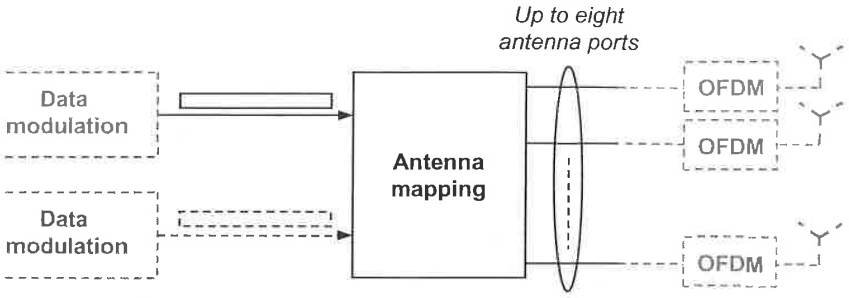


FIGURE 10.14

General structure for LTE downlink multi-antenna transmission. Modulation symbols corresponding to one or two transport blocks mapped to up to eight antenna ports.

signals are assumed to be used for demodulation (cell-specific reference signals or demodulation reference signals respectively) and the type of CSI feedback they rely on. Transmission mode 1 corresponds to single-antenna transmission while the remaining transmission modes correspond to different multi-antenna transmission schemes, including transmit diversity, beam-forming, and spatial multiplexing. Actually, LTE supports both beam-forming and spatial multiplexing as part of more general *antenna precoding*. There are two approaches to downlink antenna precoding, which differ depending on whether cell-specific reference signals or demodulation reference signals are used for channel estimation. These two approaches can also be referred to as *codebook-based precoding* and *non-codebook-based precoding* respectively, the reason for these specific names being further clarified below.

The list below summarizes the currently defined transmission modes and the associated multi-antenna transmission schemes. The different multi-antenna transmission schemes are described in more detail in the sections below.

- **Transmission mode 1:** Single-antenna transmission.
- **Transmission mode 2:** Transmit diversity.
- **Transmission mode 3:** *Open-loop* codebook-based precoding in the case of more than one layer, transmit diversity in the case of rank-one transmission.
- **Transmission mode 4:** *Closed-loop* codebook-based precoding.
- **Transmission mode 5:** Multi-user-MIMO version of transmission mode 4.
- **Transmission mode 6:** Special case of closed-loop codebook-based precoding limited to single-layer transmission.
- **Transmission mode 7:** Release-8 non-codebook-based precoding supporting only single-layer transmission.
- **Transmission mode 8:** Release-9 non-codebook-based precoding supporting up to two layers.
- **Transmission mode 9:** Release-10 non-codebook-based precoding supporting up to eight layers.

Strictly speaking, transmission modes are only applicable for DL-SCH transmission. Thus, a certain transmission mode should not be seen as identical to a certain multi-antenna transmission configuration. Rather, a certain multi-antenna transmission scheme is applied to DL-SCH transmission

according to the list above when the corresponding transmission mode is configured for DL-SCH transmission. However, the same multi-antenna transmission scheme may also be applied to other types of transmissions, such as transmission of BCH and L1/L2 control signaling.[15] However, this does not mean that the corresponding transmission mode is applied to such transmissions.

It should also be mentioned that, although a certain multi-antenna transmission scheme can be seen as being associated with a certain transmission mode, for transmission modes 3–9 there is a possibility for fall-back to transmit diversity without implying that the configured transmission mode is changed.

10.3.1 Transmit Diversity

Transmit diversity can be applied to any downlink physical channel. However, it is especially applicable to transmissions that cannot be adapted to varying channel conditions by means of link adaptation and/or channel-dependent scheduling, and thus for which diversity is more important. This includes transmission of the BCH and PCH transport channels, as well as L1/L2 control signaling. Actually, as already mentioned, transmit diversity is the *only* multi-antenna transmission scheme applicable to these channels. Transmit diversity is also used for transmission of DL-SCH when transmission mode 2 is configured. Furthermore, as also already mentioned, transmit diversity is a fall-back "mode" for transmission mode 3 and higher. More specifically, a scheduling assignment using DCI format 1A (see Section 10.4.4) implies the use of transmit diversity regardless of the configured transmission mode.

Transmit diversity assumes the use of cell-specific reference signals for channel estimation. Thus, a transmit-diversity signal is always transmitted on the same antenna ports as the cell-specific reference signals. Actually, if a cell is configured with two cell-specific reference signals, transmit diversity for two antenna ports *must be used* for BCH and PCH, as well as for the L1/L2 control signaling. Similarly, if four cell-specific reference signals are configured for the cell, transmit diversity for four antenna ports has to be used for the transmission of these channels. In this way, a terminal does not have to be explicitly informed about what multi-antenna transmission scheme is used for these channels. Rather, this is given implicitly from the number of cell-specific reference signals configured for a cell.[16]

10.3.1.1 Transmit Diversity for Two Antenna Ports

In the case of two antenna ports, LTE transmit diversity is based on *Space-Frequency Block Coding* (SFBC), as described in Chapter 5. As can be seen from Figure 10.15, SFBC implies that two consecutive modulation symbols S_i and S_{i+1} are mapped directly to frequency-adjacent resource elements on the first antenna port. On the second antenna port the frequency-swapped and transformed symbols $-S_{i+1}^*$ and S_i^* are mapped to the corresponding resource elements, where "*" denotes complex conjugate.

Figure 10.15 also indicates how the antenna ports on which a transmit-diversity signal is being transmitted correspond to the cell-specific reference signals, more specifically CRS 0 and CRS 1 in

[15] Actually, only single-antenna transmission and transmit diversity is applicable to BCH and L1/L2 control signaling.

[16] Actually, the situation is partly the opposite, i.e. the terminal blindly detects the number of antenna ports used for BCH transmission and, from that, decides on the number of cell-specific reference signals configured within the cell.

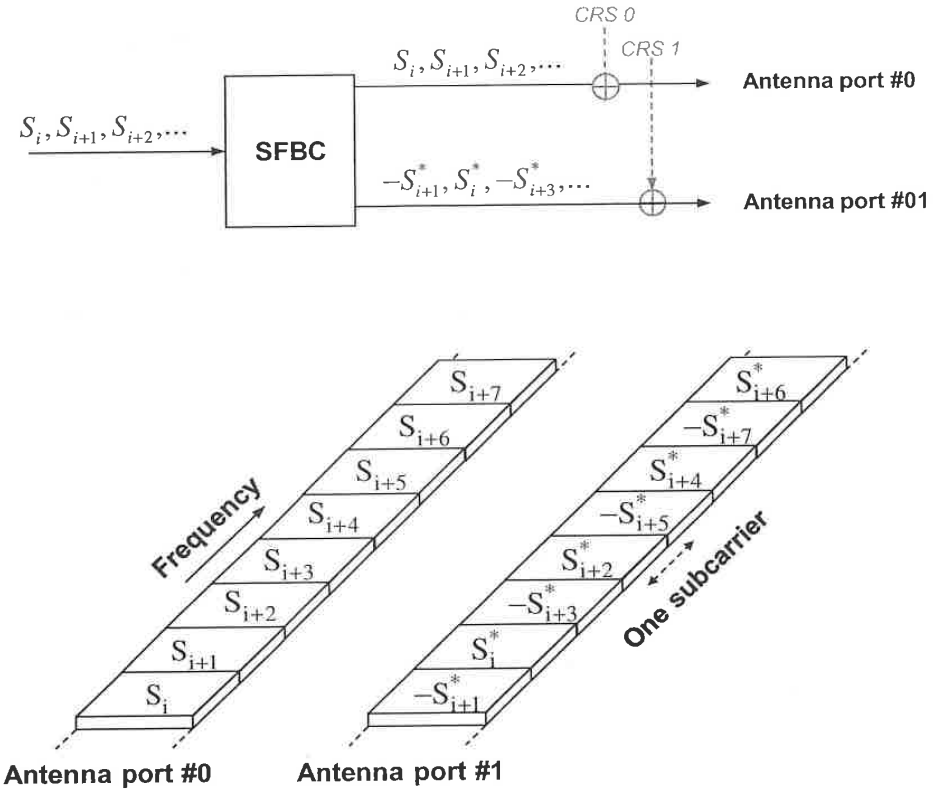


FIGURE 10.15

Transmit diversity for two antenna ports – SFBC.

the case of two antenna ports. Note that one should not interpret this such that the CRS is specifically transmitted for this transmit-diversity signal. The cell-specific reference signals are always transmitted and can be jointly used by multiple transmissions within a cell.

10.3.1.2 Transmit Diversity for Four Antenna Ports

In the case of four antenna ports, LTE transmit diversity is based on a combination of SFBC and *Frequency-Switched Transmit Diversity* (FSTD). As can be seen in Figure 10.16, combined SFBC/FSTD implies that pairs of modulation symbols are transmitted by means of SFBC with transmission alternating between pairs of antenna ports (antenna ports 0 and 2 and antenna ports 1 and 3 respectively). For the resource elements where transmission is on one pair of antenna ports, there is no transmission on the other pair of antenna ports. Thus, combined SFBC/FSTD in some sense operates on groups of four modulation symbols and corresponding groups of four frequency-consecutive resource elements on each antenna port. As mentioned in Section 10.4.1, this is the reason for the use of *resource-element groups*, each consisting of four resource elements, when defining the mapping of the L1/L2 control signaling to the physical resource.

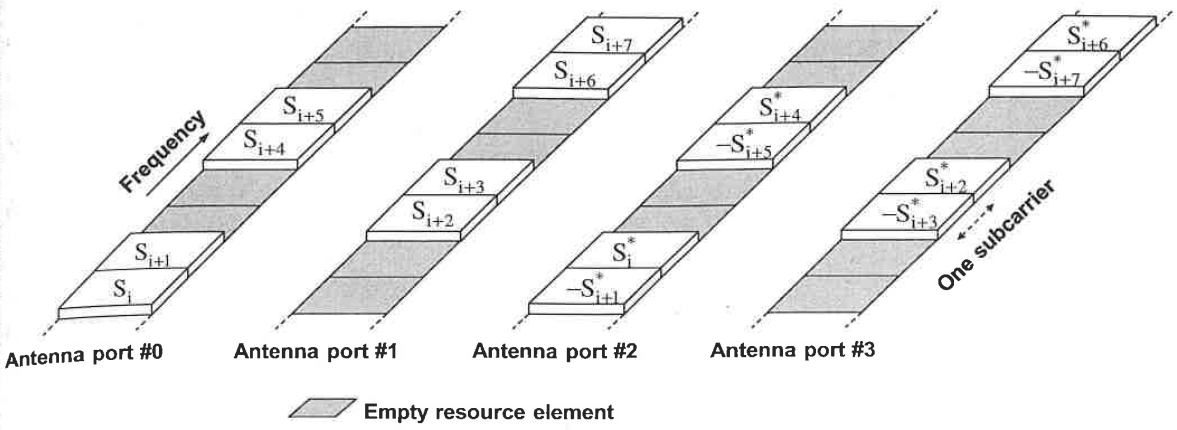


FIGURE 10.16

Transmit diversity for four antenna ports – combined SFBC/FSTD.

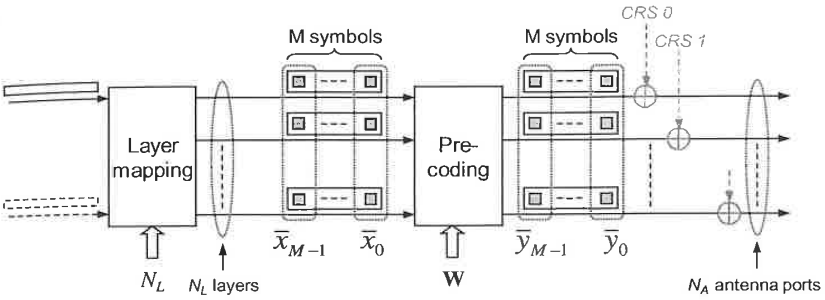


FIGURE 10.17

The basic structure of LTE codebook-based antenna precoding. The figure also indicates how cell-specific reference signals (CRS) are applied after precoding.

10.3.2 Codebook-Based Precoding

The basic processing for codebook-based precoding is illustrated in Figure 10.17. The modulation symbols corresponding to one or two transport blocks are first mapped to N_L *layers*. The number of layers may range from a minimum of one layer up to a maximum number of layers equal to the number of antenna ports.[17] The layers are then mapped to the antenna ports by means of the *precoder functionality*. As codebook-based precoding relies on the cell-specific reference signals for channel estimation, and there are at most four cell-specific reference signals in a cell, codebook-based precoding allows for a maximum of four antenna ports and, as a consequence, a maximum of four layers.

The mapping to layers is outlined in Figure 10.18 for the case of an initial transmission. There is one transport block in the case of a single layer ($N_L = 1$) and two transport blocks for two or more

[17] In practice, the number of layers is also limited by, and should not exceed, the number of receive antennas available at the terminal.

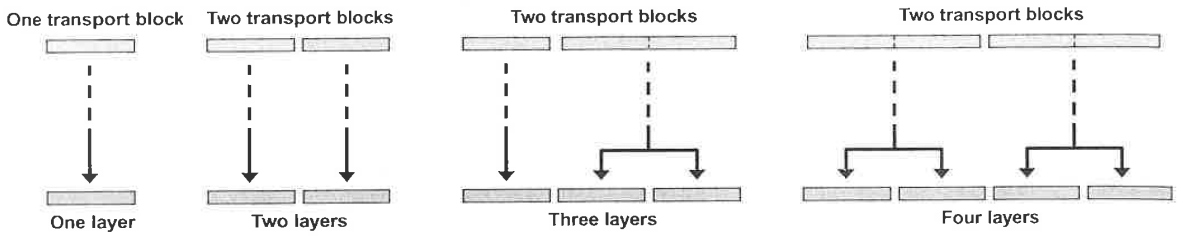


FIGURE 10.18

Transport-block-to-layer mapping for codebook-based antenna precoding (initial transmission).

layers ($N_L > 1$). In the case of a hybrid-ARQ retransmission, if only one of two transport blocks needs to be retransmitted and that transport block was mapped to two layers for the initial transmission, the retransmission may also be carried out on two layers. Thus, in the case of a retransmission, a single transport block may also be transmitted using two layers.

The mapping to layers is such that the number of modulation symbols on each layer is the same and equal to the number of symbols to be transmitted on each antenna port. Thus, in the case of three layers, there should be twice as many modulation symbols corresponding to the second transport block (mapped to the second and third layers) compared to the first transport block (mapped to the first layer). This is ensured by the selection of an appropriate transport-block size in combination with the rate-matching functionality. In the case of four layers, the first transport block is mapped to the first and second layers while the second transport block is mapped to the third and fourth layers. In this case, the number of modulation symbols should thus be the same for the two transport blocks. For one transport block being mapped to two layers, the modulated symbols corresponding to the transport block are mapped to the layers in an alternating fashion – that is, every second modulation symbol is mapped to the first and second layer respectively.

In the case of multi-antenna precoding the number of layers is also often referred to as the *transmission rank*.[18] The transmission rank can vary dynamically, for example based on the number of layers that can be supported by the channel. The latter is sometimes also referred to as the *channel rank*.

After layer mapping, a set of N_L symbols (one symbol from each layer) is linearly combined and mapped to the antenna ports. This combination/mapping can be described by a *precoder matrix* $\mathbf{W}$ of size $N_A \times N_L$, where N_A is the number of antenna ports. More specifically, the vector $\bar{y}_i$ of size N_A, consisting of one symbol for each antenna port, is given by $\bar{y}_i = \mathbf{W} \cdot \bar{x}_i$, where the vector $\bar{x}_i$ of size N_L consists of one symbol from each layer. As the number of layers can vary dynamically, also the number of columns of the precoder matrix will vary dynamically. Specifically, in the case of a single layer, the precoder matrix $\mathbf{W}$ is a vector of size $N_A \times 1$ that provides beam-forming for a single modulation symbol.

Figure 10.17 also indicates how the cell-specific reference signals are applied after antenna precoding. Channel estimation based on the cell-specific reference signals will thus reflect the channel for each antenna port *not including the precoding*. As a consequence, the terminal receiver must have

[18] In the LTE specification, transmit diversity is actually also described as transmission using *multiple layers*. However, transmit diversity is still a *single-rank* transmission scheme.

explicit knowledge about what precoding has been applied at the transmitter side in order to properly process the received signal and recover the different layers. Once again, the figure should not be interpreted such that cell-specific reference signals are inserted specifically for a given PDSCH transmission.

There are two operational modes for codebook-based precoding, *closed-loop operation* and *open-loop operation*. These two modes differ in terms of the exact structure of the precoder matrix and how the matrix is selected by the network and made known to the terminal.

10.3.2.1 Closed-Loop Operation

In case of closed-loop precoding it is assumed that the network selects the precoder matrix based on feedback from the terminal.

Based on measurements on the cell-specific reference signals, the terminal selects a suitable transmission rank and corresponding precoder matrix. Information about the selected rank and precoder matrix is then reported to the network in the form of a *Rank Indication* (RI) and a *Precoder-Matrix Indication* (PMI), as described in Section 11.4. It is important to understand though that the RI and PMI are only recommendations and the network does not need to follow the RI/PMI provided by the terminal when selecting the actual transmission rank and precoder matrix to be used for transmission to the terminal. When not following the terminal recommendation, the network must explicitly inform the terminal what precoder matrix is used for the downlink transmission. On the other hand, if the network uses the precoder matrix recommended by the terminal, only a confirmation that the network is using the recommended matrix is signaled.

To limit the signaling on both uplink and downlink only a limited set of precoder matrices, also referred to as the *codebook*, is defined for each transmission rank for a given number of antenna ports. Both the terminal (when reporting PMI) and the network (when selecting the actual precoder matrix to use for the subsequent downlink transmission to the terminal) should select a precoder matrix from the corresponding codebook. Thus, for terminal PMI reporting, as well as when the network informs the terminal about the actual precoder matrix used for the downlink transmission, only the index of the selected matrix needs to be signaled.

As LTE supports multi-antenna transmission using two and four antenna ports, codebooks are defined for:

- Two antenna ports and one and two layers, corresponding to precoder matrices of size 2×1 and 2×2 respectively.
- Four antenna ports and one, two, three, and four layers, corresponding to precoder matrices of size 4×1, 4×2, 4×3, and 4×4 respectively.

As an example, the precoder matrices specified for the case of two antenna ports are illustrated in Table 10.3. As can be seen, there are four 2×1 precoder matrices for single-layer transmission and three 2×2 precoder matrices for two-layer transmission. In the same way, sets of 4×1, 4×2, 4×3, and 4×4 matrices are defined for the case of four antenna ports and one, two, three, and four layers respectively. It should be pointed out that the first rank-2 (2×2) matrix in Table 10.3 is not used in closed-loop operation but only for *open-loop precoding*, as described in the next section.

Even if the network is following the precoder-matrix recommendation provided by the terminal, the network may, for different reasons, decide to use a lower rank for the transmission, so-called *rank*

Table 10.3 Precoder Matrices for Two Antenna Ports and One and Two Layers (the First 2 × 2 Matrix Is Only Used for Open-Loop Precoding)

One layer	$\frac{1}{\sqrt{2}}\begin{bmatrix}+1\\+1\end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix}+1\\-1\end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix}+1\\+j\end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix}+1\\-j\end{bmatrix}$
Two layers	$\frac{1}{\sqrt{2}}\begin{bmatrix}+1 & 0\\0 & +1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}+1 & +1\\+1 & -1\end{bmatrix}$	$\frac{1}{2}\begin{bmatrix}+1 & +1\\+j & -j\end{bmatrix}$	

override. In that case the network will use a subset of the columns of the recommended precoder matrix. The network precoder confirmation will then include explicit information about the set of columns being used or, equivalently, about the set of layers being transmitted.

There is also a possibility to apply closed-loop precoding strictly limited to single-layer (rank-1) transmission. This kind of multi-antenna transmission is associated with *transmission mode 6*. The reason for defining an additional transmission mode limited to single-layer transmission rather than relying on the general closed-loop precoding associated with transmission mode 4 is that, by strictly limiting to single-layer transmission, the signaling overhead on both downlink and uplink can be reduced. Transmission mode 6 can, for example, be configured for terminals with low SINR for which multi-layer transmission would anyway not apply.

10.3.2.2 Open-Loop Operation

Open-loop precoding does not rely on any detailed precoder recommendation being reported by the terminal and does not require any explicit network signaling of the actual precoder used for the downlink transmission. Instead, the precoder matrix is selected in a predefined and deterministic way known to the terminal in advance. One use of open-loop precoding is in high-mobility scenarios where accurate feedback is difficult to achieve due to the latency in the PMI reporting.

The basic transmission structure for open-loop precoding is aligned with the general codebook-based precoding outlined in Figure 10.17 and only differs from closed-loop precoding in the structure of the precoding matrix $\mathbf{W}$.

In the case of open-loop precoding, the precoder matrix can be described as the product of two matrices $\mathbf{W}'$ and $\mathbf{P}$, where $\mathbf{W}'$ and $\mathbf{P}$ are of size $N_A \times N_L$ and $N_L \times N_L$ respectively:

$$\mathbf{W} = \mathbf{W}' \cdot \mathbf{P}. \tag{10.1}$$

In the case of two antenna ports, the matrix $\mathbf{W}'$ is the normalized 2×2 identity matrix:[19]

$$\mathbf{W}' = \frac{1}{\sqrt{2}}\begin{bmatrix}+1 & 0\\0 & +1\end{bmatrix}. \tag{10.2}$$

[19] As non-codebook-based precoding is not used for rank-1 transmission (see below), there is no need for any matrix $\mathbf{W}'$ of size 2×1.

In the case of four antenna ports, $\mathbf{W}'$ is given by cycling through four of the defined $4 \times N_L$ precoder matrices and is different for consecutive resource elements.

The matrix $\mathbf{P}$ can be expressed as $\mathbf{P} = \mathbf{D}_i \cdot \mathbf{U}$, where $\mathbf{U}$ is a constant matrix of size $N_L \times N_L$ and $\mathbf{D}_i$ is a matrix of size $N_L \times N_L$ that varies between subcarriers (indicated by the index i). As an example, the matrices $\mathbf{U}$ and $\mathbf{D}_i$ for the case of two layers ($N_L = 2$) are given by:

$$\mathbf{U} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & e^{-j2\pi/2} \end{bmatrix}, \quad \mathbf{D}_i = \begin{bmatrix} 1 & 0 \\ 0 & e^{-j2\pi i/2} \end{bmatrix}. \tag{10.3}$$

The basic idea with the matrix $\mathbf{P}$ is to average out any differences in the channel conditions as seen by the different layers.

Similar to closed-loop precoding, the transmission rank for open-loop precoding can also vary dynamically down to a minimum of two layers. Transmission mode 3, associated with open-loop precoding, also allows for rank-1 transmission. In that case, transmit diversity as described in Section 10.3.1 is used – that is, SFBC for two antenna ports and combined SFBC/FSTD for four antenna ports.

10.3.3 Non-Codebook-Based Precoding

Similar to codebook-based precoding, non-codebook-based precoding is only applicable to DL-SCH transmission. Non-codebook-based precoding was introduced in LTE release 9 but was then limited to a maximum of two layers. The extension to eight layers was then introduced as part of release 10. The release-9 scheme, associated with transmission mode 8, is a subset of the extended release-10 scheme (transmission mode 9).

There is also a release-8 non-codebook-based precoding defined, associated with transmission mode 7. Transmission mode 7 relies on the release-8 demodulation reference signals mentioned but not described in detail in Section 10.2.2 and only supports single-layer transmission. In this description we will focus on non-codebook-based precoding corresponding to transmission modes 8 and 9.

The basic principles for non-codebook-based precoding can be explained based on Figure 10.19 (where the precoder is intentionally shaded; see below). As can be seen, this figure is very similar to the corresponding figure illustrating codebook-based precoding (Figure 10.17), with layer mapping of modulation symbols corresponding to one or two transport blocks followed by precoding. The layer mapping also follows the same principles as that of codebook-based precoding (see Figure 10.18) but is extended to support up to eight layers. In particular, at least for an initial transmission, there are two transport blocks per TTI except for the case of a single layer, in which case there is only one transport block within the TTI. Similar to codebook-based precoding, for hybrid-ARQ retransmissions there may in some cases be a single transport block also in the case of multi-layer transmission.

The main difference in Figure 10.19 compared to Figure 10.17 (codebook-based precoding) is the presence of demodulation reference signals (DM-RS) before the precoding. The transmission of precoded reference signals allows for demodulation and recovery of the transmitted layers at the receiver side *without explicit receiver knowledge of the precoding applied at the transmitter side*. Put simply, channel estimation based on precoded demodulation reference signals will reflect the channel experienced by the layers, *including the precoding*, and can thus be used directly for coherent demodulation of the different layers. There is thus no need to signal any precoder-matrix information

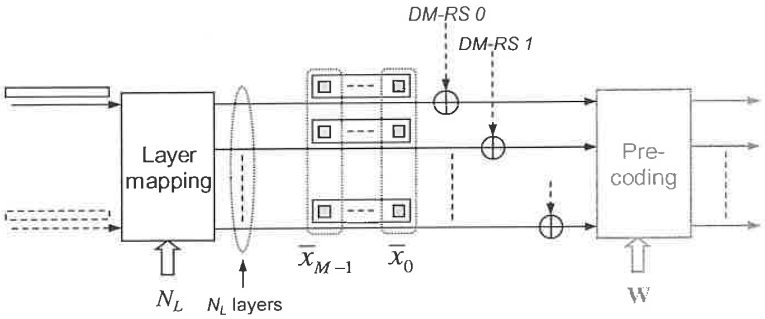


FIGURE 10.19

Basic principles for non-codebook-based antenna precoding.

to the terminal, which only needs to know the number of layers – that is, the transmission rank. As a consequence, the network can select an arbitrary precoder and there is no need for any explicit codebook to select from. This is the reason for the term *non-codebook-based* precoding. It should be noted though that non-codebook-based precoding may still rely on codebooks for the terminal feedback, as described below.

The possibility to select an arbitrary precoder matrix for the transmission is also the reason why the precoder is shaded in Figure 10.19. The precoder part of Figure 10.19 is not visible in the LTE specification and, strictly speaking, in the case of non-codebook-based precoding the antenna mapping defined according to Figure 10.14 consists of only the layer mapping. This also means that the antenna ports defined in Figure 10.14 correspond to the different layers in Figure 10.19 or, expressed differently, precoding occurs *after* the antenna ports.

Still, there must be a way for the network to select a suitable precoder matrix for the transmission. There are essentially two ways by which this can be done in the case of non-codebook-based precoding.

The network may estimate the uplink channel state, for example based on transmission of uplink sounding reference signals as described in the next chapter, and rely on uplink/downlink channel reciprocity when selecting the precoder matrix to use for the downlink transmission. This is especially of interest for TDD operation for which the use of the same frequency for uplink and downlink transmission typically leads to a higher degree of downlink/uplink channel reciprocity. It should be noted though that if the terminal uses multiple receive antennas, it also has to transmit on multiple antennas in order for uplink measurements to fully reflect the downlink channel state.

Alternatively, the network may rely on terminal feedback for precoder-matrix selection. For transmission modes 8 and 9, this feedback is actually very similar to the corresponding feedback for closed-loop codebook-based precoding. In particular, the feedback is based on the same precoder matrices as for codebook-based precoding (see Table 10.3 for an example) but extended to handle up to eight antenna ports. Furthermore, for transmission mode 9, the terminal measurements should be based on CSI-RS, as described in Section 10.2.3, rather than the cell-specific reference signals.

Thus, despite the name, non-codebook-based precoding may also use defined codebooks. However, in contrast to codebook-based precoding, the codebooks are then only used for the terminal PMI reporting and not for the actual downlink transmission.

10.3.4 Downlink Multi-User MIMO

Spatial multiplexing implies the transmission of multiple layers – that is, multiple parallel transmissions on the same time–frequency resource, to the same terminal. The presence of multiple antennas at both the transmitter and receiver sides in combination with transmitter and/or receiver signal processing is then used to suppress interference between the different layers.

Spatial multiplexing is also often referred to as *Multi-Input-Multi-Output* (MIMO) transmission, reflecting the fact that the channel in the case of spatial multiplexing can be seen as having *multiple inputs*, corresponding to the multiple transmit antennas, and *multiple outputs*, corresponding to the multiple receive antennas. The more specific term *Single-User MIMO* (SU-MIMO) is also often used for reasons that will become obvious below.

The term *Multi-User* MIMO (MU-MIMO) is, in 3GPP, used to denote transmission to *different* terminals using the same time–frequency resource, in practice relying on multiple antennas at the transmitter (network) side to separate the two transmissions.

In principle, one could realize MU-MIMO as a direct extension to spatial multiplexing, with the different layers simply being intended for different terminals. The set of terminals would demodulate and perhaps also decode the full set of layers. The data on a layer not intended for a specific terminal would then just be discarded by that terminal after demodulation/decoding.

However, such an approach would imply that all terminals involved in the MU-MIMO transmission would need to know about the full set of layers being transmitted. It would also imply that one would need to have exactly the same resource assignment for all terminals involved in the MU-MIMO transmission. All terminals would also need to include the full set of receive antennas necessary to receive the overall multi-layer transmission.

Instead, 3GPP has chosen an MU-MIMO approach that does not require terminal knowledge about the presence of the other transmissions, allows for only partly overlapping resource assignments, and, at least in principle, does not require the presence of multiple receive antennas at the mobile terminal.[20] There are two approaches to MU-MIMO specified in 3GPP, one being an integrated part of transmission modes 8 and 9 corresponding to non-codebook-based precoding, and one being based on codebook-based precoding but associated with a special transmission mode, *transmission mode 5*.

10.3.4.1 MU-MIMO within Transmission Modes 8/9

The only thing that is needed to support MU-MIMO as part of transmission modes 8 and 9 – that is, within the concept of non-codebook precoding – is to be able to assign different demodulation reference signals for the transmissions to the different terminals. As discussed in Section 10.2.2, for transmission mode 8 it is possible to define two orthogonal reference signals by means of two mutually orthogonal patterns, or *orthogonal cover codes* (OCC). In the case of spatial multiplexing, these OCC and corresponding reference signals are used for the transmission of the two layers. However, they can also be assigned to different users to enable MU-MIMO transmission within transmission mode 8.

In addition, as also mentioned in Section 10.2.2, a pseudo-random sequence is applied to the demodulation reference signals. This sequence is not terminal specific in the sense that each terminal

[20] Note, though, that the LTE performance requirements in general assume the presence of at least two receive antennas at the mobile terminal.

has its own sequence. Rather, there are only two sequences available and information on what pseudo-random sequence is used for a certain demodulation reference signal to a certain terminal is provided in the scheduling assignment together with information about the OCC.

In total, this allows for four different reference signals that can be assigned for four different transmissions being transmitted in parallel to different terminals using MU-MIMO. This assumes no spatial multiplexing on a per-link basis – that is, the transmission to each terminal consists of a single layer. One can also combine MU-MIMO with spatial multiplexing by transmitting two layers to a single terminal, in which case the number of multiplexed terminals is reduced correspondingly as the OCC are then used to separate the two layers transmitted to a single terminal. Obviously, the transmission of multiple layers to the same terminal requires the use of multiple receive antennas at the terminal.

Finally, although MU-MIMO does not rely on any terminal knowledge of the other, parallel, transmissions, nothing prevents a terminal from assuming that there are transmissions on the other "layers", corresponding to the other demodulation reference signals, and applying receiver-side signal processing to suppress the potential interference from those *possibly* present transmissions. Such suppression would then, once again, be based on the presence of multiple receive antennas at the mobile terminal.

10.3.4.2 MU-MIMO Based on CRS

The above-described MU-MIMO transmission is part of transmission modes 8 and 9 and thus became available in LTE release 9, with further extension in release 10. However, already in LTE release 8, MU-MIMO was possible by a minor modification of transmission mode 4 – that is, closed-loop codebook-based beam-forming, leading to *transmission mode 5*. The only difference between transmission modes 4 and 5 is the signaling of an additional power offset between PDSCH and the cell-specific reference signals (CRS).

In general, for transmission modes relying on CRS (as well as when relying on DM-RS) for channel estimation the terminal will use the reference signal as a phase reference but also as a power/amplitude reference for the demodulation of signals transmitted by means of higher-order modulation (16QAM and 64QAM). Thus, for proper demodulation of higher-order modulation, the terminal needs to know the power offset between the CRS and the PDSCH.

The terminal is informed about this power offset by means of higher-layer signaling. However, what is then provided is the offset between CRS power and the overall PDSCH power, including all layers. In the case of spatial multiplexing, the overall PDSCH power has to be divided between the different layers, and it is the relation between the CRS power and the per-layer PDSCH power that is relevant for demodulation.

In the case of pure spatial multiplexing (no MU-MIMO) – that is, transmission modes 3 and 4 – the terminal knows about the number of layers and thus indirectly about the offset between the CRS power and the per-layer PDSCH power.

In the case of MU-MIMO, the total available power will typically also be divided between the transmissions to the different terminals, with less PDSCH power being available for each transmission. However, terminals are not aware of the presence of parallel transmissions to other terminals and are thus not aware of any per-PDSCH power reduction. For this reason, transmission mode 5 includes the explicit signaling of an *additional power offset* of -3 dB to be used by the terminal in addition to the CRS/PDSCH power offset signaled by higher layers.

Transmission mode 5 is limited to single-rank transmission and, in practice, limited to two users being scheduled in parallel, as there is only a single -3 dB offset defined.

Note that the above power-offset signaling is not needed for MU-MIMO based on demodulation reference signals (transmission modes 8 and 9) as, in this case, each transmission has its own set of reference signals. The power per reference signal will thus scale with the number of layers and transmissions, similar to the PDSCH power and the reference-signal to per-layer PDSCH power ratio will remain constant.

10.4 DOWNLINK L1/L2 CONTROL SIGNALING

To support the transmission of downlink and uplink transport channels, there is a need for certain *associated downlink control signaling*. This control signaling is often referred to as *downlink L1/L2 control signaling*, indicating that the corresponding information partly originates from the physical layer (Layer 1) and partly from Layer 2 MAC. Downlink L1/L2 control signaling consists of downlink scheduling assignments, including information required for the terminal to be able to properly receive, demodulate, and decode the DL-SCH[21] on a component carrier, uplink scheduling grants informing the terminal about the resources and transport format to use for uplink (UL-SCH) transmission, and hybrid-ARQ acknowledgements in response to UL-SCH transmissions. In addition, the downlink control signaling can also be used for the transmission of power-control commands for power control of uplink physical channels, as well as for certain special purposes such as MBSFN notifications.

As illustrated in Figure 10.20, the downlink L1/L2 control signaling is transmitted within the first part of each subframe. Thus, each subframe can be said to be divided into a *control region* followed by a *data region*, where the control region corresponds to the part of the subframe in which the L1/L2 control signaling is transmitted. To simplify the overall design, the control region always occupies an integer number of OFDM symbols, more specifically one, two, or three OFDM symbols (for narrow cell bandwidths, 10 resource blocks or less, the control region consists of two, three, or four OFDM symbols to allow for a sufficient amount of control signaling). In the case of carrier aggregation, there is one control region per component carrier.

The size of the control region expressed in number of OFDM symbols, or, equivalently, the start of the data region, can be dynamically varied on a per-subframe basis independently for each component carrier.[22] Thus, the amount of radio resources used for control signaling can be dynamically adjusted to match the instantaneous traffic situation. For a small number of users being scheduled in a subframe, the required amount of control signaling is small and a larger part of the subframe can be used for data transmission (larger data region).

The reason for transmitting the control signaling at the beginning of the subframe is to allow for terminals to decode downlink scheduling assignments as early as possible. Processing of the data region – that is, demodulation and decoding of the DL-SCH transmission – can then begin before the end of the subframe. This reduces the delay in the DL-SCH decoding and thus the overall downlink transmission delay. Furthermore, by transmitting the L1/L2 control channel at the beginning of the subframe – that is, by allowing for early decoding of the L1/L2 control information – terminals that

[21] L1/L2 control signaling is also needed for the reception, demodulation, and decoding of the PCH transport channel.
[22] The start of the data region is semi-statically configured when using cross-carrier scheduling (see Section 10.4.1).

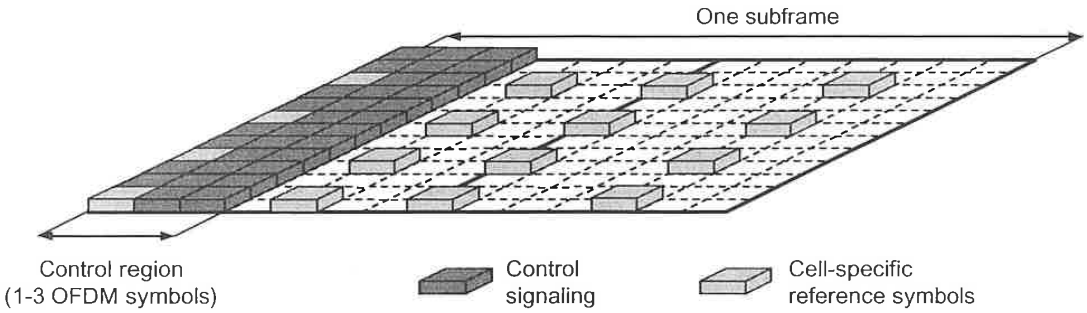


FIGURE 10.20

LTE time–frequency grid illustrating the split of the subframe into (variable-sized) control and data regions.

are not scheduled in the subframe may power down the receiver circuitry for a part of the subframe, allowing for a reduced terminal power consumption.

The downlink L1/L2 control signaling consists of four different physical-channel types:

- The *Physical Control Format Indicator Channel* (PCFICH), informing the terminal about the size of the control region (one, two, or three OFDM symbols). There is one and only one PCFICH on each component carrier or, equivalently, in each cell.
- The *Physical Downlink Control Channel* (PDCCH), used to signal downlink scheduling assignments and uplink scheduling grants. Each PDCCH typically carries signaling for a single terminal, but can also be used to address a group of terminals. Multiple PDCCHs can exist in each cell.
- The *Physical Hybrid-ARQ Indicator Channel* (PHICH), used to signal hybrid-ARQ acknowledgements in response to uplink UL-SCH transmissions. Multiple PHICHs can exist in each cell.
- The *Relay Physical Downlink Control Channel* (R-PDCCH), used for relaying. A detailed discussion can be found in Chapter 16 in conjunction with the overall description of relays; at this stage it suffices to note that the R-PDCCH is not transmitted in the control region.

The maximum size of the control region is normally three OFDM symbols (four in the case of narrow cell bandwidths), as mentioned above. However, there are a few exceptions to this rule. When operating in TDD mode, the control region in subframes 1 and 6 is restricted to at most two OFDM symbols since, for TDD, the primary synchronization signal (see Chapter 14) occupies the third OFDM symbol in those subframes. Similarly, for MBSFN subframes (see Chapter 9), the control region is restricted to a maximum of two OFDM symbols.

10.4.1 Physical Control Format Indicator Channel

The PCFICH indicates the size of the control region in terms of the number of OFDM symbols – that is, indirectly where in the subframe the data region starts. Correct decoding of the PCFICH information is thus essential. If the PCFICH is incorrectly decoded, the terminal will neither know how to process the control channels nor where the data region starts for the corresponding subframe.[23]

[23]Theoretically, the terminal could blindly try to decode all possible control channel formats and, from which format that was correctly decoded, deduce the starting position of the data region, but this can be a very complex procedure.

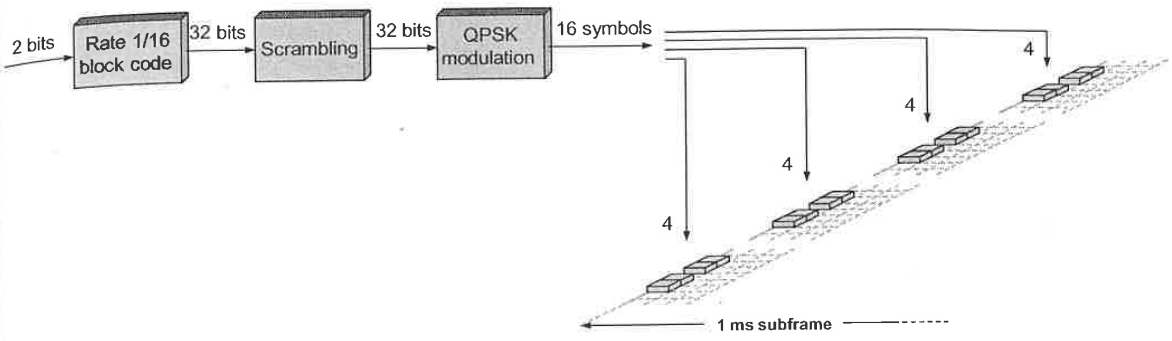


FIGURE 10.21

Overview of the PCFICH processing.

The PCFICH consists of two bits of information, corresponding to the three[24] control-region sizes of one, two, or three OFDM symbols (two, three or four for narrow bandwidths), which are coded into a 32-bit codeword. The coded bits are scrambled with a cell- and subframe-specific scrambling code to randomize inter-cell interference, QPSK modulated, and mapped to 16 resource elements. As the size of the control region is unknown until the PCFICH is decoded, the PCFICH is always mapped to the first OFDM symbol of each subframe.

The mapping of the PCFICH to resource elements in the first OFDM symbol in the subframe is done in groups of four resource elements, with the four groups being well separated in frequency to obtain good diversity. Furthermore, to avoid collisions between PCFICH transmissions in neighboring cells, the location of the four groups in the frequency domain depends on the physical-layer cell identity.

The transmission power of the PCFICH is under control of the eNodeB. If necessary for coverage in a certain cell, the power of the PCFICH can be set higher than for other channels by "borrowing" power from, for example, simultaneously transmitted PDCCHs. Obviously, increasing the power of the PCFICH to improve the performance in an interference-limited system depends on the neighboring cells not increasing their transmit power on the interfering resource elements. Otherwise, the interference would increase as much as the signal power, implying no gain in received SIR. However, as the PCFICH-to-resource-element mapping depends on the cell identity, the probability of (partial) collisions with PCFICH in neighboring cells in synchronized networks is reduced, thereby improving the performance of PCFICH power boosting as a tool to control the error rate.

The overall PCFICH processing is illustrated in Figure 10.21.

To describe the mapping of the PCFICH, and L1/L2 control signaling in general, to resource elements, some terminology is required. As mentioned above, the mapping is specified in terms of groups of four resource elements, so-called *resource-element groups*. To each resource-element group, a *symbol quadruplet* consisting of four (QPSK) symbols is mapped. The main motivation behind this, instead of simply mapping the symbols one by one, is the support of transmit diversity. As discussed in Section 10.3, transmit diversity with up to four antenna ports is specified for L1/L2

[24]The fourth combination is reserved for future use.

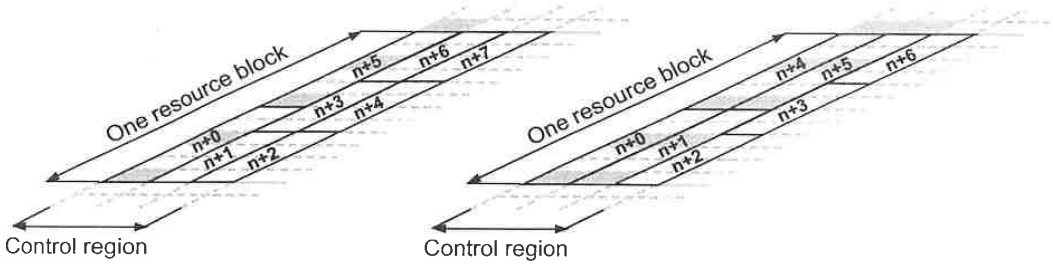


FIGURE 10.22

Numbering of resource-element groups in the control region (assuming a size of three OFDM symbols).

control signaling. Transmit diversity for four antenna ports is specified in terms of groups of four symbols (resource elements) and, consequently, the L1/L2 control-channel processing is also defined in terms of symbol quadruplets.

The definition of the resource-element groups assumes that reference symbols corresponding to two antenna ports are present in the first OFDM symbol, regardless of the actual number of antenna ports configured in the cell. This simplifies the definition and reduces the number of different structures to handle. Thus, as illustrated in Figure 10.22, in the first OFDM symbol there are two resource-element groups per resource block, as every third resource element is reserved for reference signals (or non-used resource elements corresponding to reference symbols on the other antenna port). As also illustrated in Figure 10.22, in the second OFDM symbol (if part of the control region) there are two or three resource-element groups depending on the number of antenna ports configured. Finally, in the third OFDM symbol (if part of the control region) there are always three resource-element groups per resource block. Figure 10.22 also illustrates how resource-element groups are numbered in a time-first manner within the size of the control region.

Returning to the PCFICH, four resource-element groups are used for the transmission of the 16 QPSK symbols. To obtain good frequency diversity the resource-element groups should be well spread in frequency and cover the full downlink cell bandwidth. Therefore, the four resource-element groups are separated by one-fourth of the downlink cell bandwidth in the frequency domain, with the starting position given by physical-layer cell identity. This is illustrated in Figure 10.23, where the PCFICH mapping to the first OFDM symbol in a subframe is shown for three different physical-layer cell identities in the case of a downlink cell bandwidth of eight resource blocks. As seen in the figure, the PCFICH mapping depends on the physical-layer cell identity to reduce the risk of inter-cell PCFICH collisions. The cell-specific shifts of the reference symbols, described in Section 10.2.1, are also seen in the figure.

For terminals capable of carrier aggregation, there is one PCFICH per component carrier. Independent signaling of the control-region size on the different component carriers is used, implying that the control region may be of different size on different component carriers. Hence, in principle the terminal needs to receive the PCFICH on each of the component carriers it is scheduled upon. Furthermore, as different component carriers may have different physical-layer cell identities, the location and scrambling may differ across component carriers.

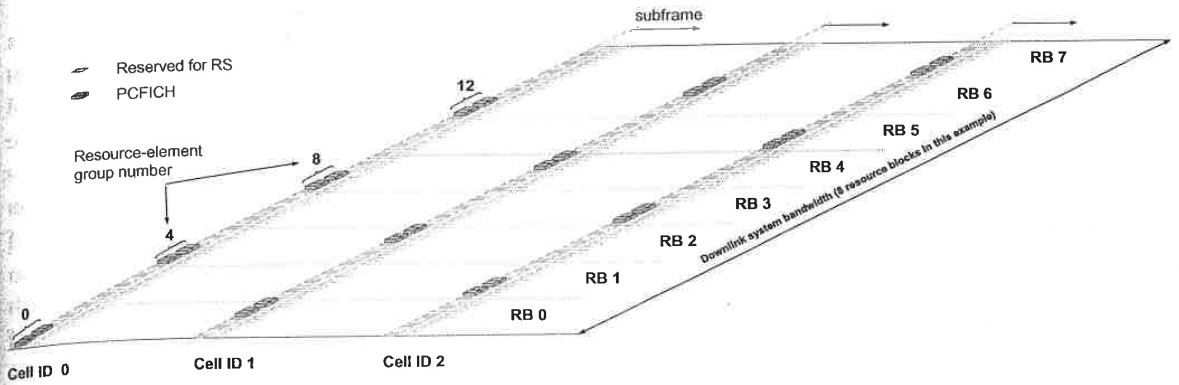


FIGURE 10.23

Example of PCFICH mapping in the first OFDM symbol for three different physical-layer cell identities.

If cross-carrier scheduling is used (see Section 10.4.6) – that is, control signaling related to a certain PDSCH transmission is transmitted on a component carrier other than the PDSCH itself – the terminal needs to know the starting position for the data region on the carrier upon which the PDSCH is transmitted. Using the PCFICH on the component carrier carrying the PDSCH would be possible in principle, although it would increase the probability of incorrectly decoding the PDSCH since there are two PCFICH instances, one for the PDCCH decoding on one component carrier and one for PDSCH reception on the other component carrier. This would be problematic, especially since the main use of cross-carrier scheduling is enhanced support of heterogeneous deployments (see Chapter 13), where some of the component carriers may be subject to strong interference. Therefore, for cross-carrier scheduled transmissions, the start of the data region is not obtained from the PCFICH on that component carrier, but is configured on a semi-static basis. The semi-statically configured value may differ from the value signaled on the PCFICH on the component carrier carrying the PDSCH transmission.

10.4.2 Physical Hybrid-ARQ Indicator Channel

The PHICH is used for transmission of hybrid-ARQ acknowledgements in response to UL-SCH transmission. There is one PHICH transmitted per received transport block and TTI – that is, when uplink spatial multiplexing is used on a component carrier, two PHICHs are used to acknowledge the transmission, one per transport block.

For proper operation of the hybrid-ARQ protocol, as discussed in Chapter 12, the error rate of the PHICH should be sufficiently low. The operating point of the PHICH is not specified and is up to the network operator to decide upon, but typically ACK-to-NAK and NAK-to-ACK error rates of the order of 10^{-2} and 10^{-3}–10^{-4} respectively are targeted. The reason for the asymmetric error rates is that an NAK-to-ACK error would imply loss of a transport block at the MAC level, a loss that has to be recovered by RLC retransmissions with the associated delays, while an ACK-to-NAK error only implies an unnecessary retransmission of an already correctly decoded transport block. To meet these error-rate targets without excessive power, it is beneficial to control the PHICH transmission power as a function of the radio-channel quality of the terminal to which the PHICH is directed. This has influenced the design of the PHICH structure.

In principle, a PHICH could be mapped to a set of resource elements exclusively used by this PHICH. However, taking the dynamic PHICH power setting into account, this could result in significant variations in transmission power between resource elements, which can be challenging from an RF implementation perspective. Therefore, it is preferable to spread each PHICH on multiple resource elements to reduce the power differences while at the same time providing the energy necessary for accurate reception. To fulfill this, a structure where several PHICHs are code multiplexed on to a set of resource elements is used in LTE. The hybrid-ARQ acknowledgement (one single bit of information per transport block) is repeated three times, followed by BPSK modulation on either the I or the Q branch and spreading with a length-four orthogonal sequence. A set of PHICHs transmitted on the same set of resource elements is called a PHICH group, where a PHICH group consists of eight PHICHs in the case of normal cyclic prefix. An individual PHICH can thus be uniquely represented by a single number from which the number of the PHICH group, the number of the orthogonal sequence within the group, and the branch, I or Q, can be derived.

For extended cyclic prefix, which is typically used in time-dispersive environments, the radio channel may not be flat over the frequency spanned by a length-four sequence. A non-flat channel would negatively impact the orthogonality between the sequences. Hence, for extended cyclic prefix, orthogonal sequences of length two are used for spreading, implying only four PHICHs per PHICH group. However, the general structure remains the same as for the normal cyclic prefix.

After forming the composite signal representing the PHICHs in a group, cell-specific scrambling is applied and the 12 scrambled symbols are mapped to three resource-element groups. Similarly to the other L1/L2 control channels, the mapping is described using resource-element groups to be compatible with the transmit diversity schemes defined for LTE.

The overall PHICH processing is illustrated in Figure 10.24.

The requirements on the mapping of PHICH groups to resource elements are similar to those for the PCFICH, namely to obtain good frequency diversity and to avoid collisions between neighboring cells in synchronized networks. Hence, each PHICH group is mapped to three resource-element groups, separated by approximately one-third of the downlink cell bandwidth. In the first OFDM

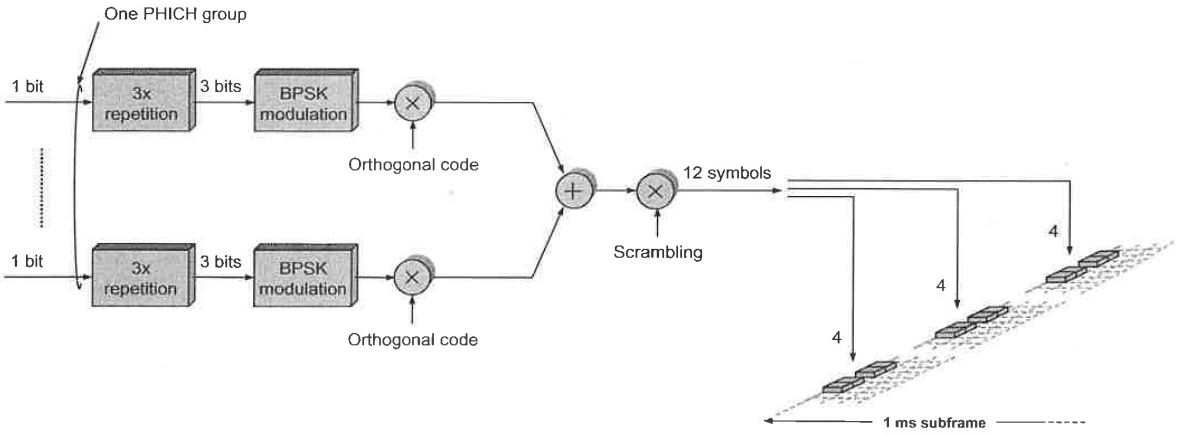


FIGURE 10.24

PHICH structure.

symbol in the control region, resources are first allocated to the PCFICH, the PHICHs are mapped to resource elements not used by the PCFICH, and finally, as will be discussed further below, the PDCCHs are mapped to the remaining resource elements.

Typically, the PHICH is transmitted in the first OFDM symbol only, which allows the terminal to attempt to decode the PHICH even if it failed decoding of the PCFICH. This is advantageous as the error requirements on the PHICH typically are stricter than for PCFICH. However, in some propagation environments, having a PHICH duration of a single OFDM symbol would unnecessarily restrict the coverage. To alleviate this, it is possible to semi-statically configure a PHICH duration of three OFDM symbols. In this case, the control region is three OFDM symbols long in all subframes to fulfill the general principle of separating the control region from data in the time domain only. Obviously, the value transmitted on the PCFICH will be fixed (and can be ignored) in this case. For narrow bandwidths, where the control region can be up to four OFDM symbols long, there is still a need to use the PCFICH to differentiate between a size-three and size-four control region.

As a general principle, LTE transmits the PHICH on the same component carrier that was used for the scheduling grant for the corresponding uplink data transmission. Not only is this principle general in the sense that it can handle symmetric as well as asymmetric carrier aggregation scenarios, it is also beneficial from a terminal power consumption perspective as the terminal only need monitor the component carriers it monitors for uplink scheduling grants (especially as the PDCCH may override the PHICH to support adaptive retransmissions, as discussed in Chapter 12).

The association between PUSCH transmission in an uplink subframe and the corresponding downlink PHICH resource is discussed in Chapter 12. At this stage, it is sufficient to note that, for TDD configuration zeros, uplink transmissions in multiple subframes may need to be acknowledged in one downlink subframe. Consequently, the amount of resources for PHICH may differ between subframes in the case of TDD.

The PHICH configuration is part of the system information transmitted on the PBCH; one bit indicates whether the duration is one or three OFDM symbols and two bits indicate the amount of resources in the control region reserved for PHICHs, expressed as a fraction of the downlink cell bandwidth in terms of resource blocks. Having the possibility to configure the amount of PHICH resources is useful as the PHICH capacity depends on, for example, whether the network uses multi-user MIMO or not. The PHICH configuration must reside on the PBCH, as it needs to be known in order to properly process the PDCCHs for reception of the part of the system information on the DL-SCH. For TDD, the PHICH information provided on the PBCH is not sufficient for the terminal to know the exact set of resources used by PHICH, as there is also a dependency on the uplink/downlink allocation, provided as part of the system information transmitted on the PDSCH. In order to receive the system information on the DL-SCH, which contains the uplink/downlink allocation, the terminal therefore has to blindly process the PDCCHs under different PHICH configuration hypotheses.

10.4.3 Physical Downlink Control Channel

The PDCCH is used to carry downlink control information (DCI) such as scheduling decisions and power-control commands. More specifically, the DCI includes:

- Downlink scheduling assignments, including PDSCH resource indication, transport format, hybrid-ARQ information, and control information related to spatial multiplexing (if applicable). A downlink scheduling assignment also includes a command for power control of the PUCCH used for transmission of hybrid-ARQ acknowledgements in response to downlink scheduling assignments.

- Uplink scheduling grants, including PUSCH resource indication, transport format, and hybrid-ARQ-related information. An uplink scheduling grant also includes a command for power control of the PUSCH.
- Power-control commands for a set of terminals as a complement to the commands included in the scheduling assignments/grants.

The different types of control information, both between the groups above as well as within the groups, correspond to different DCI message sizes. For example, supporting spatial multiplexing with non-contiguous allocation of resource blocks in the frequency domain requires a larger scheduling message in comparison with an uplink grant allowing for frequency-contiguous allocations only. The DCI is therefore categorized into different *DCI formats*, where a format corresponds to a certain message size and usage. The DCI formats are summarized in Table 10.4, sorted in approximate order of the relative size. The actual message size depends on, among other factors, the cell bandwidth as, for larger bandwidths, a larger number of bits is required to indicate the resource-block allocation. The number of cell-specific reference signals in the cell and whether cross-carrier scheduling is configured or not will also affect the absolute size of most DCI formats. Hence, a given DCI format may have different sizes depending on the overall configuration of the cell. This will be discussed further below; at this stage it suffices to note that formats 0, 1A, 3, and 3A have the same message size.[25]

Table 10.4 DCI Formats

Size	Usage				
	Uplink Grant		Downlink Assignment		Power Control
Small	–		1C	Special purpose compact assignment	–
	0	Single layer	1A	Contiguous allocations only	3, 3A
...	–		1B	Codebook-based beam-forming using CRS	–
	–		1D	Multi-user MIMO using CRS	–
	4	Spatial multiplexing	–		–
	–		1	Flexible allocations	–
	–		2A	Open-loop spatial multiplexing using CRS	–
	–		2B	Dual-layer transmission using DM-RS	–
	–		2C	Multi-layer transmission using DM-RS	–
Large	–		2	Closed-loop spatial multiplexing using CRS	–

[25]The smaller of DCI formats 0 and 1A is padded to ensure the same payload size. Which of 0 and 1A is padded depends on the uplink and downlink cell bandwidths; in the case of identical uplink and downlink bandwidths in a cell there is a single bit of padding in format 0.

One PDCCH carries one DCI message with one of the formats above. As multiple terminals can be scheduled simultaneously, on both downlink and uplink, there must be a possibility to transmit multiple scheduling messages within each subframe. Each scheduling message is transmitted on a separate PDCCH, and consequently there are typically multiple simultaneous PDCCH transmissions within each cell. Furthermore, to support different radio-channel conditions, link adaptation can be used, where the code rate of the PDCCH is selected to match the radio-channel conditions.

For carrier aggregation, scheduling assignments/grants are transmitted individually per component carrier. PDCCH transmission for carrier aggregation is further elaborated upon in Section 10.4.6.

10.4.4 Downlink Scheduling Assignment

Downlink scheduling assignments are valid for the same subframe in which they are transmitted. The scheduling assignments use one of the DCI formats 1, 1A, 1B, 1C, 1D, 2, 2A, 2B, or 2C and the DCI formats used depend on the transmission mode configured (the relation between transmission modes and DCI formats is further described in Section 10.4.9). The reason for supporting multiple formats with different message sizes for the same purpose is to allow for a trade-off in control-signaling overhead and scheduling flexibility. Parts of the contents are the same for the different DCI formats, as seen in Table 10.5, but obviously there are also differences due to the different capabilities.

DCI format 1 is the basic downlink assignment format in the absence of spatial multiplexing (transmission modes 1, 2, and 7). It supports non-contiguous allocations of resource blocks and the full range of modulation-and-coding schemes.

DCI format 1A, also known as the "compact" downlink assignment, supports allocation of frequency-contiguous resource blocks only and can be used in all transmission modes. Contiguous allocations reduce the payload size of the control information with a somewhat reduced flexibility in resource allocations. The full range of modulation-and-coding schemes is supported.

DCI format 1B is used to support codebook-based beam-forming described in Section 10.3.2, with a low control-signaling overhead (transmission mode 6). The content is similar to DCI format 1A with the addition of bits for signaling of the precoding matrix. As codebook-based beam-forming can be used to improve the data rates for cell-edge terminals, it is important to keep the related DCI message size small so as not to unnecessarily limit the coverage.

DCI format 1C is used for various special purposes such as random-access response, paging, and transmission of system information. Common for these applications is simultaneous reception of a relatively small amount of information by *multiple* users. Hence, DCI format 1C supports QPSK only, has no support for hybrid-ARQ retransmissions, and does not support closed-loop spatial multiplexing. Consequently, the message size for DCI format 1C is very small, which is beneficial for coverage and efficient transmission of the type of system messages for which it is intended. Furthermore, as only a small number of resource blocks can be indicated, the size of the corresponding indication field in DCI format 1C is independent of the cell bandwidth.

DCI format 1D is used to support multi-user MIMO (transmission mode 5) scheduling of one codeword with precoder information. To support dynamic sharing of the transmission power between the terminals sharing the same resource block in multi-user MIMO, one bit of power offset information is included in DCI format 1D, as described in Section 10.3.4.2.

DCI format 2 is an extension for DCI format 1 to support closed-loop spatial multiplexing (transmission mode 4). Thus, information about the number of transmission layers and the index of the

Table 10.5 DCI Formats Used for Downlink Scheduling

Field		DCI Format								
		1	1A	1B	1C	1D	2	2A	2B	2C
Resource information	Carrier indicator	•	•	•		•	•	•	•	•
	Resource block assignment type	0/1	2	2	2′	2	0/1	0/1	0/1	0/1
HARQ process number		•	•	•		•	•	•	•	•
First transport block	MCS	•	•	•	•	•	•	•	•	•
	RV	•	•	•		•	•	•	•	•
	NDI	•	•	•		•	•	•	•	•
Second transport block	MCS						•	•	•	•
	RV						•	•	•	•
	NDI						•	•	•	•
Multi-antenna information	PMI confirmation			•						
	Precoding information			•		•	•	•		
	Transport block swap flag						•	•		
	Power offset					•				
	DM-RS scrambling								•	
	# Layers/DM-RS scrambling									•
Downlink assignment index		•	•	•		•	•	•	•	•
PUCCH power control		•	•	•		•	•	•	•	•
Flag for 0/1A differentiation			•							
Padding (only if needed)		(•)	(•)	(•)		(•)	(•)	(•)	(•)	(•)
Identity		•	•	•	•	•	•	•	•	•

Samsung Exhibit 1019, Page 154 of 195

precoder matrix used are jointly encoded in the precoding information field. Some of the fields in DCI format 1 have been duplicated to handle the two transport blocks transmitted in parallel in the case of spatial multiplexing.

DCI format 2A is similar to DCI format 2 except that it supports open-loop spatial multiplexing (transmission mode 3) instead of closed-loop spatial multiplexing. The precoder information field is used to signal the number of transmission layers only, hence the field has a smaller size than in DCI format 2. Furthermore, since DCI format 2A is used for scheduling of multi-layer transmissions only, the precoder information field is only necessary in the case of four transmit antenna ports (a two-antenna-port setup cannot be used for more than two transmission layers).

DCI format 2B was introduced in release 9 in order to support dual-layer spatial multiplexing in combination with beam-forming using DM-RS (transmission mode 8). Since scheduling with DCI format 2B relies on DM-RS, precoding/beam-forming is transparent to the terminal and there is no need to signal a precoder index. The number of layers can be controlled by disabling one of the transport blocks. Two different scrambling sequences for the DM-RS can be used, as described in Section 10.2.2.

DCI format 2C was introduced in release 10 and is used to support spatial multiplexing using DM-RS (transmission mode 9). To some extent, it can be seen as a generalization of format 2B to support spatial multiplexing of up to eight layers. DM-RS scrambling and the number of layers are jointly signaled by a single three-bit field.

Many information fields in the different DCI formats are, as already mentioned, common among several of the formats, while some types of information exist only in certain formats. The information in the DCI formats used for downlink scheduling can be organized into different groups, as shown in Table 10.5, with the fields present varying between the DCI formats. The contents of the different DCI formats are explained in more detail below:

- Resource information, consisting of:
  - Carrier indicator (0 or 3 bit). This field is only present if cross-carrier scheduling is enabled via RRC signaling and is used to indicate the component carrier the downlink control information relates to (see Section 10.4.6).
  - Resource-block allocation. This field indicates the resource blocks on one component carrier upon which the terminal should receive the PDSCH. The size of the field depends on the cell bandwidth and on the DCI format, more specifically on the resource indication type, as discussed in Section 10.4.4.1. Resource allocation types 0 and 1, which are the same size, support non-contiguous resource-block allocations, while resource allocation type 2 has a smaller size but supports contiguous allocations only. DCI format 1C uses a restricted version of type 2 in order to further reduce control signaling overhead.
- Hybrid-ARQ process number (3 bit for FDD, 4 bit for TDD), informing the terminal about the hybrid-ARQ process to use for soft combining. Not present in DCI format 1C.
- For the first (or only) transport block:[26]
  - Modulation-and-coding scheme (5 bit), used to provide the terminal with information about the modulation scheme, the code rate, and the transport-block size, as described further below. DCI format 1C has a restricted size of this field as only QPSK is supported.

[26] A transport block can be disabled by setting the modulation-and-coding scheme to zero and the redundancy version to one in the DCI.

  - New-data indicator (1 bit), used to clear the soft buffer for initial transmissions. Not present in DCI format 1C as this format does not support hybrid ARQ.
  - Redundancy version (2 bit).
- For the second transport block (only present in DCI format supporting spatial multiplexing):
  - Modulation-and-coding scheme (5 bit).
  - New-data indicator (1 bit).
  - Redundancy version (2 bit).
- Multi-antenna information. The different DCI formats are intended for different multi-antenna schemes and which of the fields below that are included depends on the DCI format shown in Table 10.5.
  - PMI confirmation (1 bit), present in format 1B only. Indicates whether the eNodeB uses the (frequency-selective) precoding matrix recommendation from the terminal or if the recommendation is overridden by the information in the PMI field.
  - Precoding information, provides information about the index of the precoding matrix used for the downlink transmission and, indirectly, about the number of transmission layers.
  - Transport block swap flag (1 bit), indicating whether the two codewords should be swapped prior to being fed to the hybrid-ARQ processes. Used for averaging the channel quality between the codewords.
  - Power offset between the PDSCH and cell-specific reference signals used to support dynamic power sharing between multiple terminals for multi-user MIMO.
  - Reference-signal scrambling sequence, used to control the generation of quasi-orthogonal reference-signal sequences, as discussed in Section 10.2.2.
  - Number of layers and reference-signal scrambling sequence (jointly encoded information in release 10).
- Downlink assignment index (2 bit), informing the terminal about the number of downlink transmissions for which a single hybrid-ARQ acknowledgement should be generated according to Section 12.1.3. Present for TDD only.
- Transmit-power control for PUCCH (2 bit). For scheduling of a secondary component carrier in the case of carrier aggregation, these bits are reused as *acknowledgement resource indicator* (ARI) – see Chapter 11.
- DCI format 0/1A indication (1 bit), used to differentiate between DCI formats 1A and 0 as the two formats have the same message size. This field is present in DCI formats 0 and 1A only. DCI formats 3 and 3A, which have the same size, are separated from DCI formats 0 and 1A through the use of a different RNTI.
- Padding; the smaller of DCI formats 0 and 1A is padded to ensure the same payload size irrespective of the uplink and downlink cell bandwidths. Padding is also used to ensure that the DCI size is different for different DCI formats (this is rarely required in practice as the payload sizes are different due to the different amounts of information). Finally, padding is used to avoid certain DCI sizes that may cause ambiguous decoding.[27]

[27] For a small set of specific payload sizes, the control signaling may be correctly decoded at an aggregation level other than the one used by the transmitter. Since the PHICH resource is derived from the first CCE used for the PDCCH, this may result in incorrect PHICH being monitored by the terminal. To overcome this, padding is used if necessary to avoid the problematic payload sizes.

- Identity (RNTI) of the terminal for which the PDSCH transmission is intended (16 bit). As described in Section 10.4.8, the identity is not explicitly transmitted but implicitly included in the CRC calculation. There are different RNTIs defined depending on the type of transmission (unicast data transmission, paging, power-control commands, etc.).

10.4.4.1 Signaling of Downlink Resource-Block Allocations

Focusing on the signaling of resource-block allocations, there are three different possibilities, types 0, 1, and 2, as indicated in Table 10.5. Resource-block allocation types 0 and 1 both support non-contiguous allocations of resource blocks in the frequency domain, whereas type 2 supports contiguous allocations only. A natural question is why multiple ways of signaling the resource-block allocations are supported, a question whose answer lies in the number of bits required for the signaling. The most flexible way of indicating the set of resource blocks the terminal is supposed to receive the downlink transmission upon is to include a bitmap with size equal to the number of resource blocks in the cell bandwidth. This would allow for an arbitrary combination of resource blocks to be scheduled for transmission to the terminal but would, unfortunately, also result in a very large bitmap for the larger cell bandwidths. For example, in the case of a downlink cell bandwidth corresponding to 100 resource blocks, the downlink PDCCH would require 100 bits for the bitmap alone, to which the other pieces of information need to be added. Not only would this result in a large control-signaling overhead, but it could also result in downlink coverage problems as more than 100 bits in one OFDM symbol correspond to a data rate exceeding 1.4 Mbit/s. Consequently, there is a need for a resource allocation scheme requiring a smaller number of bits while keeping sufficient allocation flexibility.

In resource allocation type 0, the size of the bitmap has been reduced by pointing not to individual resource blocks in the frequency domain, but to groups of contiguous resource blocks, as shown at the top of Figure 10.25. The size of such a group is determined by the downlink cell bandwidth; for the smallest bandwidths there is only a single resource block in a group, implying that an arbitrary set of resource blocks can be scheduled, whereas for the largest cell bandwidths groups of four resource blocks are used (in the example in Figure 10.25, the cell bandwidth is 25 resource blocks, implying a group size of two resource blocks). Thus, the bitmap for the system with a downlink cell bandwidth of 100 resource blocks is reduced from 100 to 25 bits. A drawback is that the scheduling granularity is reduced; single resource blocks cannot be scheduled for the largest cell bandwidths using allocation type 0.

However, also in large cell bandwidths, frequency resolution of a single resource block is sometimes useful, for example to support small payloads. Resource allocation type 1 addresses this by dividing the total number of resource blocks in the frequency domain into dispersed subsets, as shown in the middle of Figure 10.25. The number of subsets is given from the cell bandwidth with the number of subsets in type 1 being equal to the group size in type 0. Thus, in Figure 10.25, there are two subsets, whereas for a cell bandwidth of 100 resource blocks there would have been four different subsets. Within a subset, a bitmap indicates the resource blocks in the frequency domain upon which the downlink transmission occurs.

To inform the terminal whether resource allocation type 0 or 1 is used, the resource allocation field includes a flag for this purpose, denoted "type" to the left in Figure 10.25. For type 0, the only additional information is the bitmap discussed above. For type 1, on the other hand, in addition to the bitmap

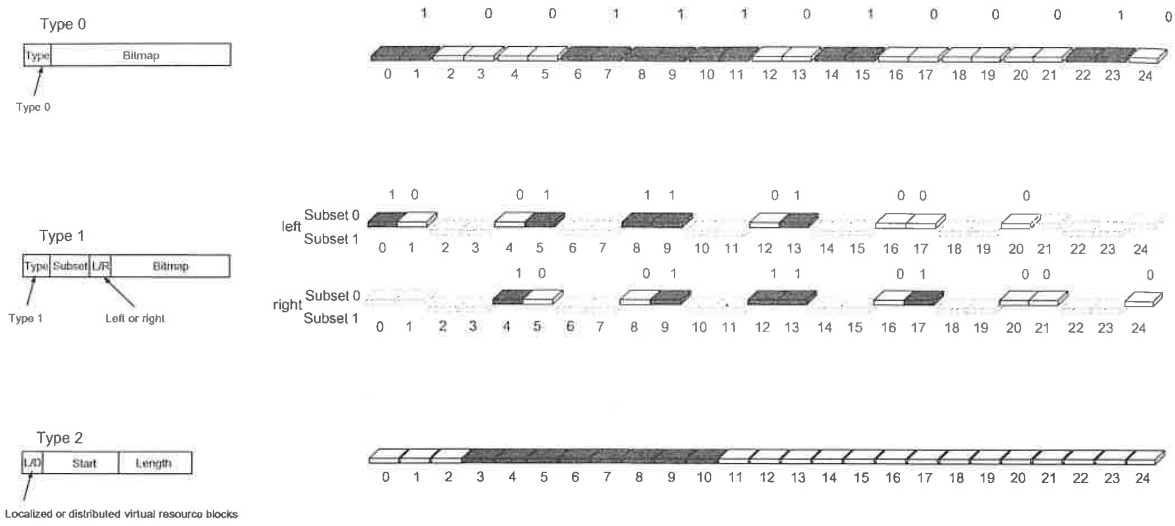


FIGURE 10.25

Illustration of resource-block allocation types (cell bandwidth corresponding to 25 resource blocks used in this example).

itself, information about the subset for which the bitmap relates is also required. As one of the requirements in the design of resource allocation type 1 was to maintain the same number of bits in the allocation as for type 0 without adding unnecessary overhead,[28] the bitmap in resource allocation type 1 is smaller than in type 0 to allow for the signaling of the subset number. However, a consequence of a smaller bitmap is that not all resource blocks in the subset can be addressed simultaneously. To be able to address all resources with the bitmap, there is a flag indicating whether the bitmap relates to the "left" or "right" part of the resource blocks, as depicted in the middle part of Figure 10.25.

Unlike the other two types of resource-block allocation signaling, type 2 does not rely on a bitmap. Instead, it encodes the resource allocation as a start position and length of the resource-block allocation. Thus, it does not support arbitrary allocations of resource blocks but only frequency-contiguous allocations, thereby reducing the number of bits required for signaling the resource-block allocation. The number of bits required for resource-signaling type 2 compared to type 0 or 1 is shown in Figure 10.26 and, as shown, the difference is fairly large for the larger cell bandwidths.

All three resource-allocation types refer to *virtual* resource blocks (see Section 10.1.1.8 for a discussion of resource-block types). For resource-allocation types 0 and 1, the virtual resource blocks are of localized type and the virtual resource blocks are directly mapped to physical resource blocks. For resource-allocation type 2, on the other hand, both localized and distributed virtual resource blocks are supported. One bit in the resource allocation field indicates whether the allocation signaling refers to localized or distributed resource blocks.

[28] Allowing different sizes would result in an increase in the number of blind decoding attempts required in the terminal.

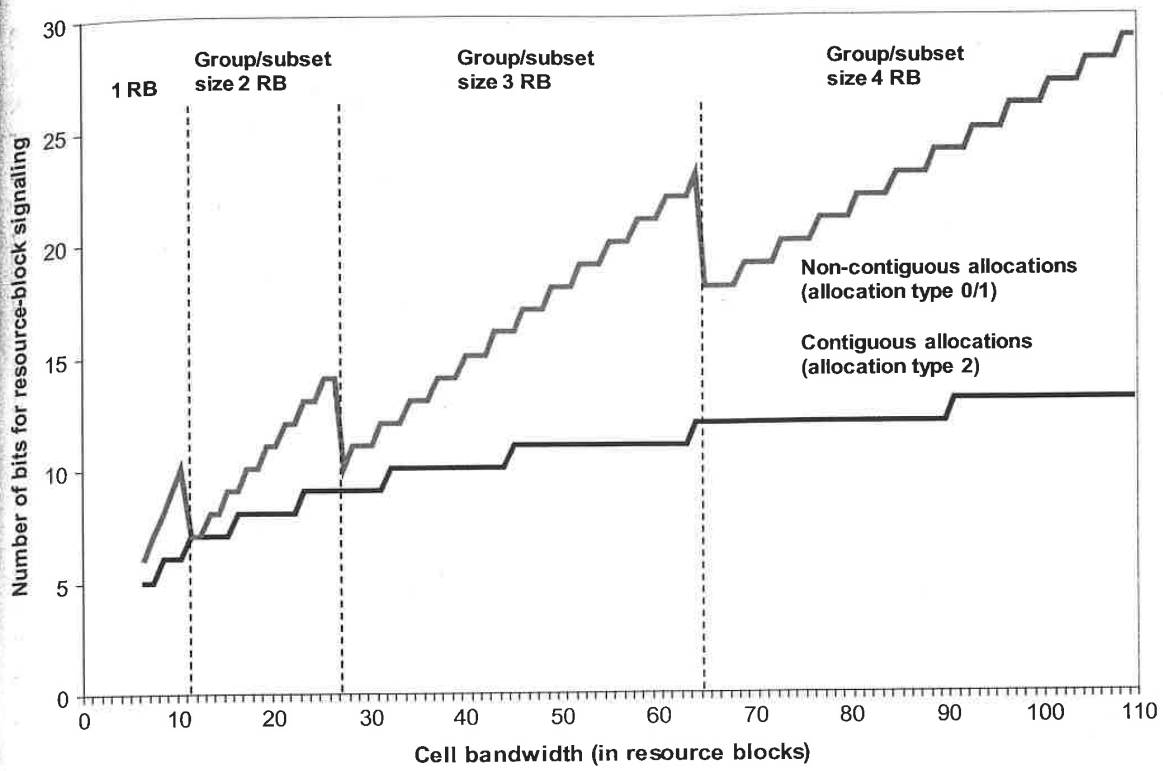


FIGURE 10.26

Number of bits used for downlink resource-allocation signaling for downlink allocation types 0/1 and 2.

10.4.4.2 Signaling of Transport-Block Sizes

Proper reception of a downlink transmission requires, in addition to the set of resource blocks, knowledge about the modulation scheme and the transport-block size, information (indirectly) provided by a five-bit field in the different DCI formats. Of the 32 combinations, 29 are used to signal the modulation-and-coding scheme whereas three are reserved, the purpose of which is described later. Together, the modulation-and-coding scheme and the number of resource blocks assigned provide the transport-block size on the DL-SCH. Thus, the possible transport-block sizes can be described as a table with 29 rows and 110 columns, one column for each number of resource blocks possible to transmit upon (the number of columns follows from the maximum downlink component carrier bandwidth of 110 resource blocks).

Each modulation-and-coding scheme represents a particular combination of modulation scheme and channel-coding rate or, equivalently, a certain spectral efficiency measured in the number of information bits per modulation symbol. Although the 29-by-110 table of transport-block sizes in principle could be filled directly from the modulation-and-coding scheme and the number of resource blocks, this would result in arbitrary transport-block sizes, which is not desirable. First, as all the higher-layer protocol layers are byte aligned, the resulting transport-block sizes should be an integer

number of bytes. Secondly, common payloads, for example RRC signaling messages and VoIP, should be possible to transmit without padding. Aligning with the QPP interleaver sizes is also beneficial, as this would avoid the use of filler bits (see Section 10.1.1.3). Finally, the same transport-block size should ideally appear for several different resource-block allocations, as this allows the number of resource blocks to be changed between retransmission attempts, providing increased scheduling flexibility. Therefore, a "mother table" of transport-block sizes is first defined, fulfilling the requirements above. Each entry in the 29-by-110 table is picked from the mother table such that the resulting spectral efficiency is as close as possible to the spectral efficiency of the signaled modulation-and-coding scheme. The mother table spans the full range of transport-block sizes possible, with an approximately constant worst-case padding.

From a simplicity perspective, it is desirable if the transport-block sizes do not vary with the configuration of the system. The set of transport-block sizes is therefore independent of the actual number of antenna ports and the size of the control region.[29] The design of the table assumes a control region of three OFDM symbols and two antenna ports, the "reference configuration". If the actual configuration is different, the resulting code rate for the DL-SCH will be slightly different as a result of the rate-matching procedure. However, the difference is small and of no practical concern. Also, if the actual size of the control region is smaller than the three-symbol assumption in the reference configuration, the spectral efficiencies will be somewhat smaller than the range indicated by the modulation-and-coding scheme signaled as part of the DCI. Thus, information about the modulation scheme used is obtained directly from the modulation-and-coding scheme, whereas the exact code rate and rate matching is obtained from the implicitly signaled transport-block size together with the number of resource elements used for DL-SCH transmission.

For bandwidths smaller than the maximum of 110 resource blocks, a subset of the table is used. More specifically, in case of a cell bandwidth of N resource blocks, the first N columns of the table are used. Also, in the case of spatial multiplexing, a single transport block can be mapped to up to four layers. To support the higher data rates this facilitates, the set of supported transport-block sizes needs to be extended beyond what is possible in the absence of spatial multiplexing. The additional entries are in principle obtained by multiplying the sizes with the number of layers to which a transport block is mapped and adjusting the result to match the QPP interleaver size.

The 29 combinations of modulation-and-coding schemes each represent a reference spectral efficiency in the approximate range of 0.1–5.9 bits/s/symbol.[30] There is some overlap in the combinations in the sense that some of the 29 combinations represent the same spectral efficiency. The reason is that the best combination for realizing a specific spectral efficiency depends on the channel properties; sometimes higher-order modulation with a low code rate is preferable over lower-order modulation with a higher code rate, sometimes the opposite is true. With the overlap, the eNodeB can select the best combination, given the propagation scenario. As a consequence of the overlap, two of the rows in the 29-by-110 table are duplicates and result in the same spectral efficiency but with different modulation schemes, and there are only 27 unique rows of transport-block sizes.

Returning to the three reserved combinations in the modulation-and-coding field mentioned at the beginning, those entries can be used for retransmissions only. In the case of a retransmission, the

[29] For DwPTS, the transport-block size is scaled by a factor of 0.75 compared to the values found in the table, motivated by the DwPTS having a shorter duration than a normal subframe.

[30] The exact values vary slightly with the number of resource blocks allocated due to rounding.

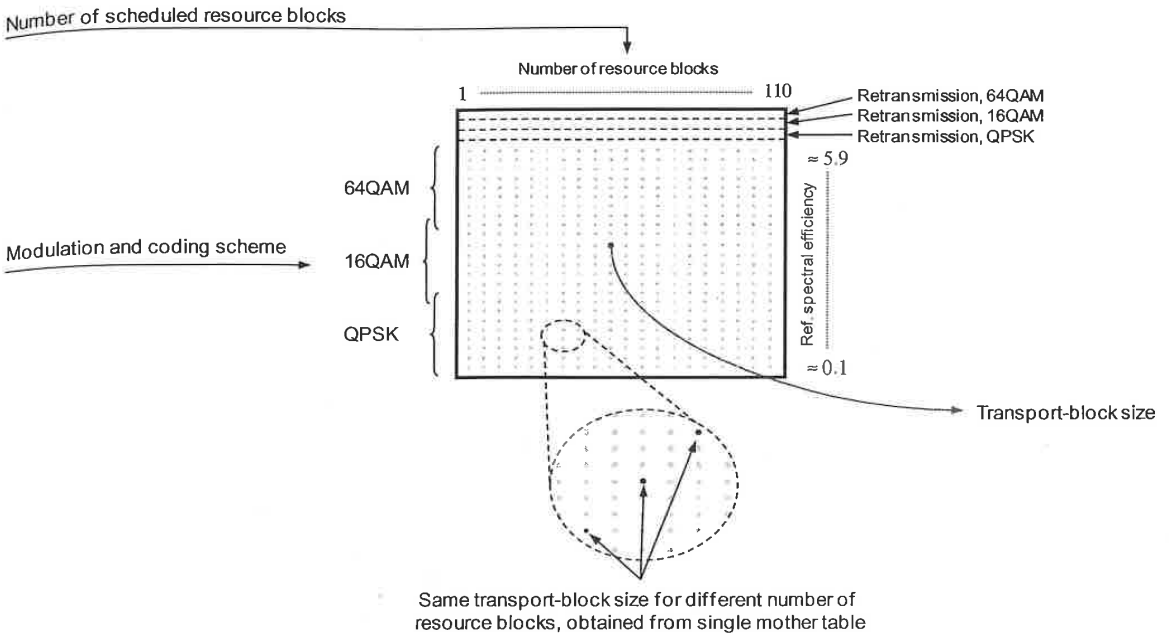


FIGURE 10.27

Computing the transport-block size.

transport-block size is, by definition, unchanged and fundamentally there is no need to signal this piece of information. Instead, the three reserved values represent the modulation scheme, QPSK, 16QAM or 64QAM, which allows the scheduler to use an (almost) arbitrary combination of resource blocks for the retransmission. Obviously, using any of the three reserved combinations assumes that the terminal properly received the control signaling for the initial transmission; if this is not the case, the retransmission should explicitly indicate the transport-block size.

The derivation of the transport-block size from the modulation-and-coding scheme and the number of scheduled resource blocks is illustrated in Figure 10.27.

10.4.5 Uplink Scheduling Grants

Uplink scheduling grants use one of DCI formats 0 or 4, where DCI format 4 was added in release 10 to support uplink spatial multiplexing. The basic resource-allocation scheme for the uplink is single-cluster allocations where the resource blocks are contiguous in the frequency domain, although release 10 added support for multi-cluster transmissions of up to two clusters on a single component carrier.

DCI format 0 is used for scheduling uplink transmissions not using spatial multiplexing on one component carrier. It has the same size control-signaling message as the "compact" downlink assignment (DCI format 1A). A flag in the message is used to inform the terminal whether the message is an uplink scheduling grant (DCI format 0) or a downlink scheduling assignment (DCI format 1A).

Table 10.6 DCI Formats for Uplink Scheduling Grants

Field		DCI Format	
		0	4
Resource information	Carrier indicator	•	•
	Multi-cluster flag	•	•
	Resource block assignment	0	1
First transport block	MCS/RV	•	•
	NDI	•	•
Second transport block	MCS/RV		•
	NDI		•
DMRS phase rotation		•	•
Precoding information			•
CSI request		•	•
SRS request			•
Uplink index/DAI		•	•
PUSCH power control		•	•
Flag for 0/1A differentiation		•	
Padding (only if needed)		(•)	(•)
Identity		•	•

DCI format 4 is used for uplink transmissions using spatial multiplexing on one component carrier. Consequently, the size of DCI format 4 is larger than that of DCI format 0, as additional information fields are required.

Many information fields are common to the two DCI formats, but there are also differences, as shown in Table 10.6. The contents of the different DCI formats are explained in more detail below:

- Resource information, consisting of:
  - Carrier indicator (0 or 3 bit). This field is only present if cross-carrier scheduling is enabled via RRC signaling and is used to indicate the uplink component carrier the grant relates to (see Section 10.4.6).
  - Multi-cluster flag (1 bit), indicating whether one or two clusters of resource blocks are used for the uplink transmission. This flag is not present in releases prior to release 10. In previous releases, the downlink bandwidth is, in practice, always at least as large as the uplink bandwidth, implying that one padding bit is used for DCI format 0 in those releases to align with the size of format 1A. The padding bit could therefore be replaced by the multi-cluster flag in release 10 without sacrificing backwards compatibility. In DCI format 4, supported in release 10 only, the multi-cluster flag is always present.
  - Resource-block allocation, including hopping indication. This field indicates the resource blocks upon which the terminal should transmit the PUSCH using uplink resource-allocation type 0 (DCI format 0) or type 1 (DCI format 4), as described in Section 10.4.5.1. The size of

the field depends on the cell bandwidth. For single-cluster allocations, uplink frequency hopping, as described in Chapter 11, can be applied to the uplink PUSCH transmission.

- For the first (or only) transport block:
  - Modulation-and-coding scheme including redundancy version (5 bit), used to provide the terminal with information about the modulation scheme, the code rate, and the transport-block size. The signaling of the transport-block size uses the same transport-block table as for the downlink – that is, the modulation-and-coding scheme together with the number of scheduled resource blocks provides the transport-block size. However, as the support of 64QAM in the uplink is not mandatory for all terminals, terminals not capable of 64QAM use 16QAM when 64QAM is indicated in the modulation-and-coding field. The use of the three reserved combinations is slightly different than for the downlink; the three reserved values are used for implicit signaling of the redundancy version, as described below. A transport block can be disabled by signaling a specific combination of modulation-and-coding scheme and number of resource blocks. Disabling one transport block is used when retransmitting a single transport block only.
  - New-data indicator (1 bit), used to indicate to the terminal whether transmission of a new transport block or retransmission of the previous transport block is granted.
- For the second transport block (DCI format 4 only):
  - Modulation-and-coding scheme including redundancy version (5 bit).
  - New-data indicator (1 bit).
- Phase rotation of the uplink demodulation reference signal[31] (3 bit), used to support multi-user MIMO, as described in Chapter 11. By assigning different reference-signal phase rotations to terminals scheduled on the same time–frequency resources, the eNodeB can estimate the uplink channel response from each terminal and suppress the inter-terminal interference by the appropriate processing.
- Precoding information, used to signal the precoder to use for the uplink transmission.
- Channel-state request flag (1 or 2 bit). The network can explicitly request an aperiodic channel-state report to be transmitted on the UL-SCH by setting this bit(s) in the uplink grant. In the case of carrier aggregation, two bits are used to indicate which downlink component carrier the CSI should be reported for (see Chapter 12).
- SRS request (2 bit), used to trigger aperiodic sounding using one of up to three preconfigured settings, as discussed in Chapter 11.
- Uplink index/DAI (2 bit). This field is present only when operating in TDD. For uplink–downlink configuration 0 (uplink-heavy configuration), it is used as an uplink index to signal for which uplink subframe(s) the grant is valid, as described in Chapter 13. For other uplink–downlink configurations, it is used as downlink assignment index to indicate the number of downlink transmissions the eNodeB expects hybrid-ARQ acknowledgement for.
- Transmit-power control for PUSCH (2 bit).
- DCI format 0/1A indication (1 bit), used to differentiate between DCI formats 1A and 0 as the two formats have the same message size. This field is present in DCI formats 0 and 1A only.

[31] In release 10, this field can also be used to select the orthogonal cover sequence used for uplink reference signals (see Chapter 11 for details).

- Padding; the smaller of DCI formats 0 and 1A is padded to ensure the same payload size irrespective of the uplink and downlink cell bandwidths. Padding is also used to ensure that the DCI size is different for different DCI formats (this is rarely required in practice as the payload sizes are different due to the different amounts of information). Finally, padding is used to avoid certain DCI sizes that may cause ambiguous decoding.
- Identity (RNTI) of the terminal for which the grant is intended (16 bit). As described in Section 10.4.8, the identity is not explicitly transmitted but implicitly included in the CRC calculation.

There is no explicit signaling of the redundancy version in the uplink scheduling grants. This is motivated by the use of a synchronous hybrid-ARQ protocol in the uplink; retransmissions are normally triggered by a negative acknowledgement on the PHICH and not explicitly scheduled as for downlink data transmissions. However, as described in Chapter 12, there is a possibility to explicitly schedule retransmissions. This is useful in a situation where the network will explicitly move the retransmission in the frequency domain by using the PDCCH instead of the PHICH. Three values of the modulation-and-coding field are reserved to mean redundancy versions one, two, and three. If one of those values is signaled, the terminal should assume that the same modulation and coding as the original transmission is used. The remaining entries are used to signal the modulation-and-coding scheme to use and also imply that redundancy version zero should be used. The difference in usage of the reserved values compared to the downlink scheduling assignments implies that the modulation scheme, unlike the downlink case, cannot change between uplink (re)transmission attempts.

10.4.5.1 Signaling of Uplink Resource-Block Allocations

The basic uplink resource-allocation scheme is single-cluster allocations – that is, allocations contiguous in the frequency domain – but release 10 also provides the possibility for multi-cluster uplink transmissions.

Single-cluster allocations use uplink resource-allocation type 0, which is identical to downlink resource allocation type 2 described in Section 10.4.4.1 except that the single-bit flag indicating localized/distributed transmission is replaced by a single-bit hopping flag. The resource allocation field in the DCI provides the set of virtual resource blocks to use for uplink transmission. The set of physical resource blocks to use in the two slots of a subframe is controlled by the hopping flag, as described in Chapter 11.

Multi-cluster allocations with up to two clusters were introduced in release 10, using uplink resource-allocation type 1. In resource-allocation type 1, the starting and ending positions of two clusters of frequency-contiguous resource blocks are encoded into an index. Uplink resource-allocation type 1 does not support frequency hopping (diversity is achieved through the use of two clusters instead). Indicating two clusters of resources naturally requires additional bits compared to the single-cluster case. At the same time, the total number of bits used for resource-allocation type 1 should be identical to that of type 0 (this is similar to the situation for the downlink allocation types 0 and 1; without aligning the sizes a new DCI format with a corresponding negative impact on the number of blind decodings is necessary). Since frequency hopping is not supported for allocation type 1, the bit otherwise used for the hopping flag can be reused for extending the resource-allocation field. However, despite the extension of the resource-allocation field by one bit, the number of bits is not sufficient to provide a single-resource-block resolution in the two clusters for all bandwidths. Instead, similar to downlink resource allocation type 0, groups of resource blocks are used and the starting and ending positions of the two clusters are

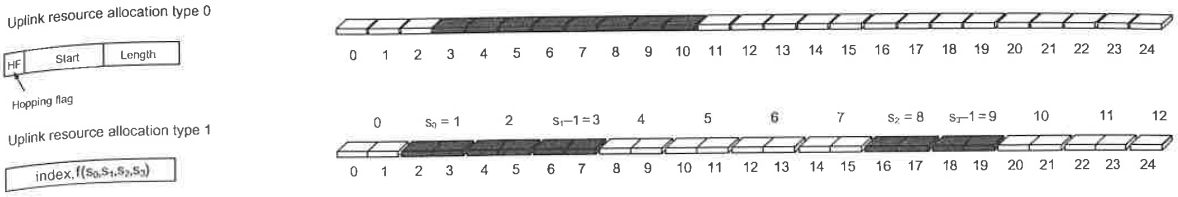


FIGURE 10.28

Illustration of uplink resource-block allocation types (uplink bandwidth corresponding to 25 resource blocks used in this example).

given in terms of group numbers. The size of such a group is determined by the uplink carrier bandwidth in a similar way as for the downlink; for the smallest bandwidths there is only a single resource block in a group, implying that an arbitrary (as long as the limit of at most two clusters is observed) set of resource blocks can be scheduled, whereas for the largest cell bandwidths groups of four resource blocks are used (in the example in Figure 10.28, the cell bandwidth is 25 resource blocks, implying a group size of two resource blocks).

10.4.6 Carrier Aggregation and Cross-Carrier Scheduling

Carrier aggregation, where a terminal receives or transmits on multiple component carriers, is an integral part of LTE from release 10 onwards. Clearly, for a terminal supporting multiple component carriers, the terminal needs to know to which component carrier a certain DCI relates. This information can either be implicit or explicit, depending on whether cross-carrier scheduling is used or not, as illustrated in Figure 10.29. Enabling cross-carrier scheduling is done individually via RRC signaling on a per-terminal and per-component-carrier basis.

In the absence of cross-carrier scheduling, downlink scheduling assignments are valid for the component carrier upon which they are transmitted. Similarly, for uplink grants, there is an association between downlink and uplink component carriers such that each uplink component carrier has an associated downlink component carrier. The association is provided as part of the system information. Thus, from the uplink–downlink association, the terminal will know to which uplink component carrier the downlink control information relates to.

In the presence of cross-carrier scheduling, where downlink PDSCH or uplink PUSCH is transmitted on an (associated) component carrier other than the PDCCH, the carrier indicator in the PDCCH provides information about the component carrier used for the PDSCH or PUSCH.

Whether cross-carrier scheduling is used or not is, as already mentioned, configured using higher-layer signaling; if cross-carrier scheduling is not configured then no carrier indication field is included in the DCI. Thus, most of the DCI formats come in two "flavors", with and without the carrier indication field, and which "flavor" the terminal is supposed to monitor is determined by enabling/disabling support for cross-carrier scheduling.

To signal which component carrier a grant relates to, the component carriers are numbered. The primary component carrier is always given the number zero, while the different secondary component carriers are assigned a unique number each through UE-specific RRC signaling. Hence, even if the

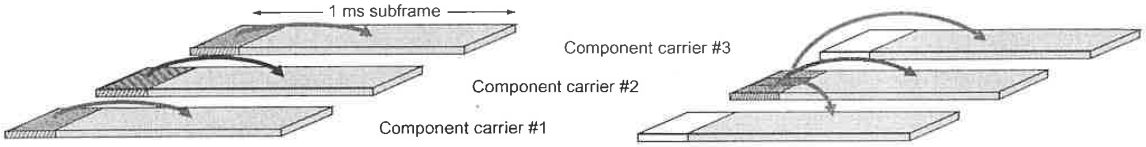


FIGURE 10.29

Cross-carrier scheduling.

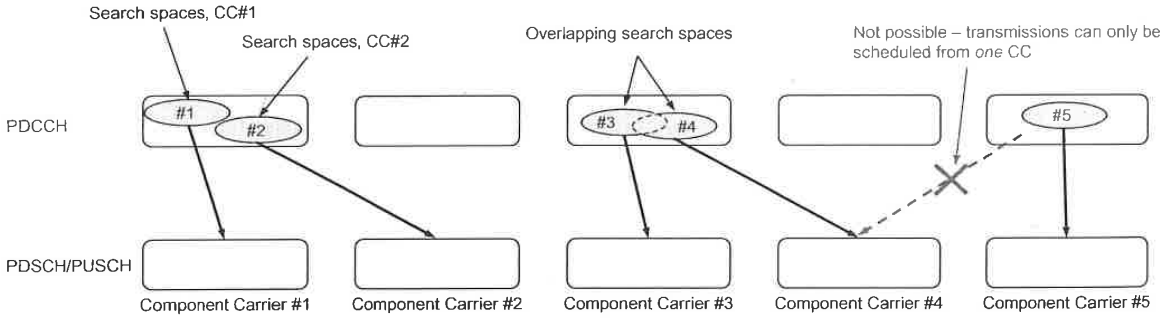


FIGURE 10.30

Example of scheduling of multiple component carriers (search spaces are discussed in Section 10.4.9).

terminal and the eNodeB may have different understandings of the component carrier numbering during a brief period of reconfiguration, at least transmissions on the primary component carrier can be scheduled.

Irrespective of whether cross-carrier scheduling is used or not, PDSCH/PUSCH on a component carrier can only be scheduled from *one* component. Thus, for each PDSCH/PUSCH component carrier there is an associated component carrier, configured via UE-specific RRC signaling, where the corresponding DCI can be transmitted. Figure 10.30 illustrates one example hereof, where PDSCH/PUSCH transmissions on component carrier 1 are scheduled using PDCCHs transmitted on component carrier 1. In this case, as cross-carrier scheduling is not used, there is no carrier indicator in the corresponding DCI formats. PDSCH/PUSCH transmissions on component carrier 2 are cross-carrier scheduled from PDCCHs transmitted on component carrier 1. Hence, the DCI formats in the UE-specific search space for component carrier 2 include the carrier indicator.

Note also, as transmissions on a component carrier can be scheduled by PDCCHs on *one* component carrier only, component carrier 4 cannot be scheduled by PDCCHs on component carrier 5 as the semi-static association between the component carriers used for PDCCH transmission and the actual data transmission has associated data on component carrier 4, with PDCCHs on component carrier 3 in this example.

10.4.7 Power-Control Commands

As a complement to the power-control commands provided as part of the downlink scheduling assignments and the uplink scheduling grants, there is the potential to transmit a power-control command using DCI formats 3 (single-bit command per terminal) or 3A (two-bit command per terminal). The main motivation for DCI format 3/3A is to support power control for semi-persistent scheduling. The power-control message is directed to a group of terminals using an RNTI specific for that group. Each terminal can be allocated two power-control RNTIs, one for PUCCH power control and the other for PUSCH power control. Although the power-control RNTIs are common to a group of terminals, each terminal is informed through RRC signaling which bit(s) in the DCI message it should follow. No carrier indicator is used for formats 3 and 3A.

10.4.8 PDCCH Processing

Having discussed the contents of L1/L2 control signaling in terms of the different DCI formats, the transmission of the DCI message on a PDCCH can be described. The processing of downlink control signaling is illustrated in Figure 10.31. A CRC is attached to each DCI message payload. The identity of the terminal (or terminals) addressed – that is, the RNTI – is included in the CRC calculation and not explicitly transmitted. Depending on the purpose of the DCI message (unicast data transmission, power-control command, random-access response, etc.), different RNTIs are used; for normal unicast data transmission, the terminal-specific C-RNTI is used.

Upon reception of DCI, the terminal will check the CRC using its set of assigned RNTIs. If the CRC checks, the message is declared to be correctly received and intended for the terminal. Thus, the identity of the terminal that is supposed to receive the DCI message is implicitly encoded in the CRC and not explicitly transmitted. This reduces the amount of bits necessary to transmit on the PDCCH as, from a terminal point of view, there is no difference between a corrupt message whose CRC will not check and a message intended for another terminal.

After CRC attachment, the bits are coded with a rate-1/3 tail-biting convolutional code and rate-matched to fit the amount of resources used for PDCCH transmission. Tail-biting convolutional coding is similar to conventional convolutional coding with the exception that no tail bits are used. Instead, the convolutional encoder is initialized with the last bits of the message prior to the encoding process. Thus, the starting and ending states in the trellis in an MLSE (Viterbi) decoder are identical.

After the PDCCHs to be transmitted in a given subframe have been allocated to the desired resource elements (the details of which are given below), the sequence of bits corresponding to all the PDCCH resource elements to be transmitted in the subframe, including the unused resource elements, is scrambled by a cell- and subframe-specific scrambling sequence to randomize inter-cell interference, followed by QPSK modulation and mapping to resource elements.

To allow for simple yet efficient processing of the control channels in the terminal, the mapping of PDCCHs to resource elements is subject to a certain structure. This structure is based on so-called *Control-Channel Elements* (CCEs), which in essence is a convenient name for a set of 36 useful resource elements (nine resource-element groups as defined in Section 10.4.1). The number of CCEs, one, two, four, or eight, required for a certain PDCCH depends on the payload size of the control information (DCI payload) and the channel-coding rate. This is used to realize link adaptation for the PDCCH; if the channel conditions for the terminal to which the PDCCH is intended are

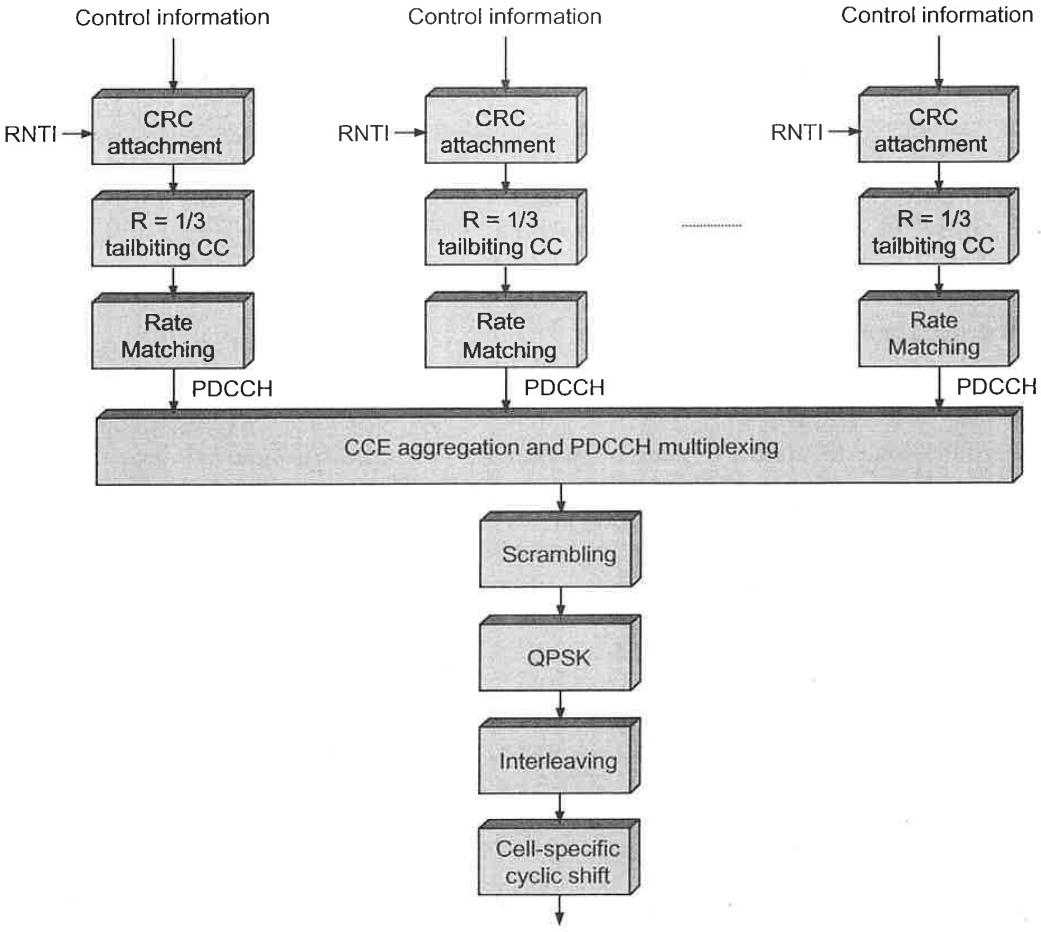


FIGURE 10.31

Processing of L1/L2 control signaling.

disadvantageous, a larger number of CCEs needs to be used compared to the case of advantageous channel conditions. The number of CCEs used for a PDCCH is also referred to as the aggregation level.

The number of CCEs available for PDCCHs depends on the size of the control region, the cell bandwidth, the number of downlink antenna ports, and the amount of resources occupied by PHICH. The size of the control region can vary dynamically from subframe to subframe as indicated by the PCFICH, whereas the other quantities are semi-statically configured. The CCEs available for PDCCH transmission can be numbered from zero and upward, as illustrated in Figure 10.32. A specific PDCCH can thus be identified by the numbers of the corresponding CCEs in the control region.

As the number of CCEs for each of the PDCCHs may vary and is not signaled, the terminal has to blindly determine the number of CCEs used for the PDCCH it is addressed upon. To reduce the complexity of this process somewhat, certain restrictions on the aggregation of contiguous CCEs have been specified. For example, an aggregation of eight CCEs can only start on CCE numbers evenly

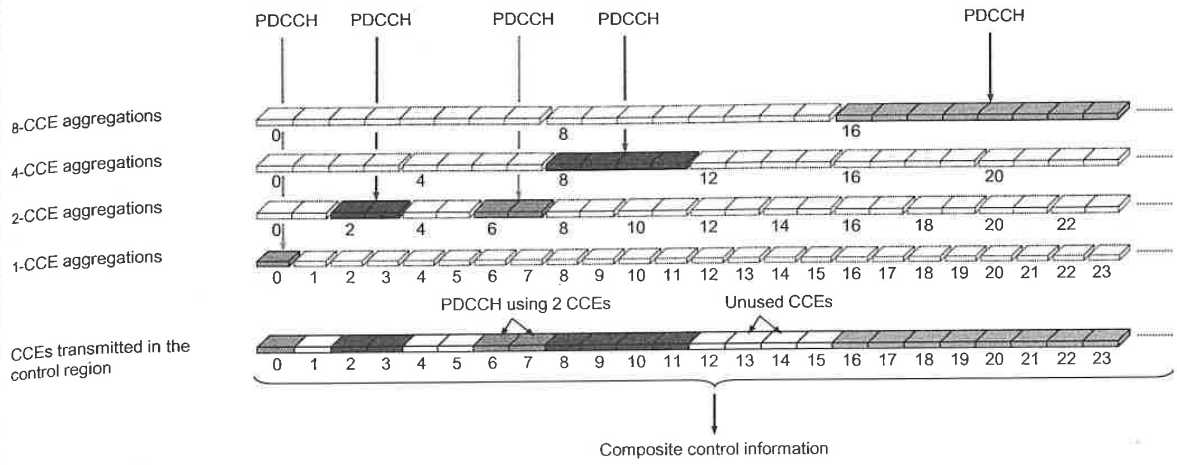


FIGURE 10.32

CCE aggregation and PDCCH multiplexing.

divisible by 8, as illustrated in Figure 10.32. The same principle is applied to the other aggregation levels. Furthermore, some combinations of DCI formats and CCE aggregations that result in excessively high channel-coding rates are removed.

The sequence of CCEs should match the amount of resources available for PDCCH transmission in a given subframe – that is, the number of CCEs varies according to the value transmitted on the PCFICH. In many cases, not all the PDCCHs that can be transmitted in the control region are used. Nevertheless, unused PDCCHs are part of the interleaving and mapping process in the same way as any other PDCCH. At the terminal, the CRC will not check for those "dummy" PDCCHs. Preferably, the transmission power is set to zero for those unused PDCCHs; the power can be used by other control channels.

The mapping of the modulated composite control information is, for the same reason as for the other control channels, described in terms of symbol quadruplets being mapped to resource-element groups. Thus, the first step of the mapping stage is to group the QPSK symbols into symbol quadruplets, each consisting of four consecutive QPSK symbols. In principle, the sequence of quadruplets could be mapped directly to the resource elements in sequential order. However, this would not exploit all the frequency diversity available in the channel and diversity is important for good performance. Furthermore, if the same CCE-to-resource-element mapping is used in all neighboring cells, a given PDCCH will persistently collide with one and the same PDCCH in the neighboring cells assuming a fixed PDCCH format and inter-cell synchronization. In practice, the number of CCEs per PDCCH varies in the cell as a function of the scheduling decisions, which gives some randomization to the interference, but further randomization is desirable to obtain a robust control-channel design. Therefore, the sequence of quadruplets is first interleaved using a block interleaver to allow exploitation of the frequency diversity, followed by a cell-specific cyclic shift to randomize the interference between neighboring cells. The output from the cell-specific shift is mapped to resource-element groups in a time-first manner, as illustrated in Figure 10.22, skipping resource-element groups used for PCFICH and PHICH. Time-first mapping

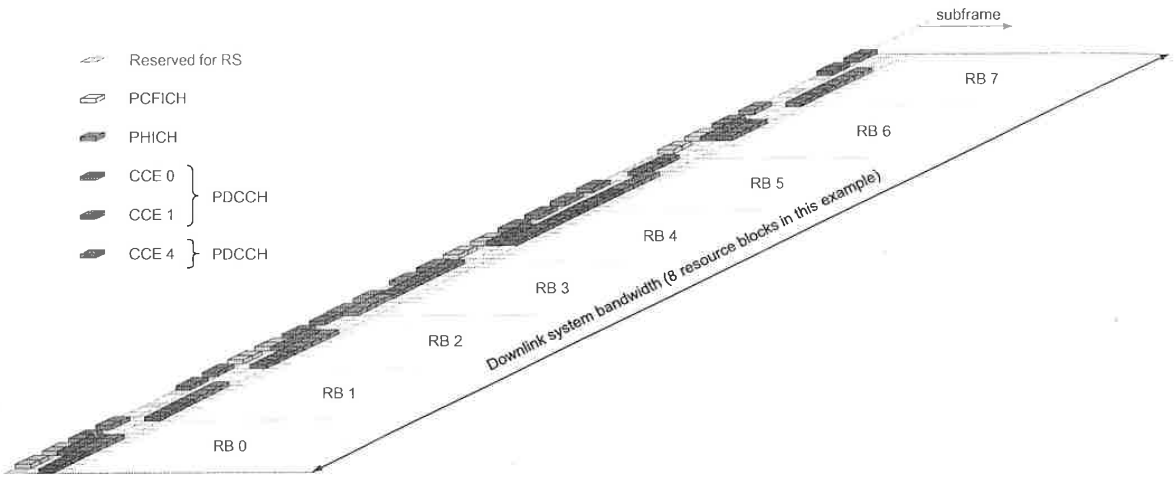


FIGURE 10.33

Example of mapping of PCFICH, PHICH, and PDCCH.

preserves the interleaving properties; with frequency-first over multiple OFDM symbols, resource-element groups that are spread far apart after the interleaving process may end up close in frequency, although on different OFDM symbols.

The interleaving operation described above, in addition to enabling exploitation of the frequency diversity and randomizing the inter-cell interference, also serves the purpose of ensuring that each CCE spans virtually all the OFDM symbols in the control region. This is beneficial for coverage as it allows flexible power balancing between the PDCCHs to ensure good performance for each of the terminals addressed. In principle, the energy available in the OFDM symbols in the control region can be balanced arbitrarily between the PDCCHs. The alternative of restricting each PDCCH to a single OFDM symbol would imply that power cannot be shared between PDCCHs in different OFDM symbols.

Similarly to the PCFICH, the transmission power of each PDCCH is under the control of the eNodeB. Power adjustments can therefore be used as a complementary link adaptation mechanism in addition to adjusting the code rate. Relying on power adjustments alone might seem a tempting solution but, although possible in principle, it can result in relatively large power differences between resource elements. This may have implications on the RF implementation and may violate the out-of-band emission masks specified. Hence, to keep the power differences between the resource elements reasonable, link adaptation through adjusting the channel code rate, or equivalently the number of CCEs aggregated for a PDCCH, is necessary. The two mechanisms for link adaptation, power adjustments and different code rates, complement each other.

To summarize and to illustrate the mapping of PDCCHs to resource elements in the control region, consider the example shown in Figure 10.33. In this example, the size of the control region in the subframe considered equals three OFDM symbols. Two downlink antenna ports are configured (but, as explained above, the mapping would be identical in the case of a single antenna port). One PHICH group is configured and three resource-element groups are therefore used by the PHICHs. The cell identity is assumed to be identical to zero in this case.

The mapping can then be understood as follows: First, the PCFICH is mapped to four resource-element groups, followed by allocating the resource-element groups required for the PHICH. The resource-element groups left after the PCFICH and PHICH are used for the different PDCCHs in the system. In this particular example, one PDCCH is using CCE numbers 0 and 1, while another PDCCH is using CCE number 4. Consequently, there is a relatively large amount of unused resource-element groups in this example; either they can be used for additional PDCCHs or the power otherwise used for the unused CCEs could be allocated to the PDCCHs in use (as long as the power difference between resource elements is kept within the limits set by the RF requirements). Furthermore, depending on the inter-cell interference situation, fractional loading of the control region may be desirable, implying that some CCEs are left unused to reduce the average inter-cell interference.

10.4.9 Blind Decoding of PDCCHs

As described above, each PDCCH supports multiple formats and the format used is a priori unknown to the terminal. Therefore, the terminal needs to blindly detect the format of the PDCCHs. The CCE structure described in the previous section helps in reducing the number of blind decoding attempts, but is not sufficient. Hence, it is required to have mechanisms to limit the CCE aggregations that the terminal is supposed to monitor. Clearly, from a scheduling point of view, restrictions in the allowed CCE aggregations are undesirable as they may influence the scheduling flexibility and require additional processing at the transmitter side. At the same time, requiring the terminal to monitor all possible CCE aggregations, also for the larger cell bandwidths, is not attractive from a terminal-complexity point of view. To impose as few restrictions as possible on the scheduler while at the same time limit the maximum number of blind decoding attempts in the terminal,[32] LTE defines so-called *search spaces*, which describe the set of CCEs the terminal is supposed to monitor for scheduling assignments/grants relating to a certain component carrier. Search spaces are discussed below, starting with the case of a single component carrier and later extended to the case of multiple component carriers.

A search space is a set of candidate control channels formed by CCEs on a given aggregation level, which the terminal is supposed to attempt to decode. As there are multiple aggregation levels, corresponding to one, two, four, and eight CCEs, a terminal has multiple search spaces. In each subframe, the terminals will attempt to decode all the PDCCHs that can be formed from the CCEs in each of its search spaces. If the CRC checks, the content of the control channel is declared as valid for this terminal and the terminal processes the information (scheduling assignment, scheduling grants, etc.). Clearly, the network can only address a terminal if the control information is transmitted on a PDCCH formed by the CCEs in one of the terminal's search spaces. For example, terminal A in Figure 10.34 cannot be addressed on a PDCCH starting at CCE number 20, whereas terminal B can. Furthermore, if terminal A is using CCEs 16–23, terminal B cannot be addressed on aggregation level 4 as all CCEs in its level-4 search space are blocked by the use for the other terminals. From this it can be intuitively understood that for efficient utilization of the CCEs in the system, the search spaces should differ between terminals. Each terminal in the system therefore has a *terminal-specific* search space at each aggregation level.

[32] In release 8/9, the number of blind decoding attempts is 44 per subframe, while for release 10 with uplink spatial multiplexing the number is 60 (assuming a single component carrier).

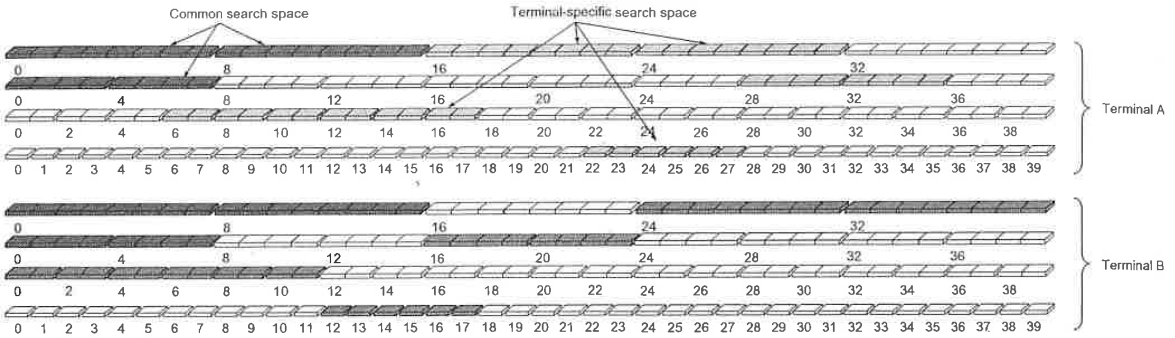


FIGURE 10.34

Principal illustration of search spaces in two terminals.

As the terminal-specific search space is typically smaller than the number of PDCCHs the network could transmit at the corresponding aggregation level, there must be a mechanism determining the set of CCEs in the terminal-specific search space for each aggregation level. One possibility would be to let the network configure the terminal-specific search space in each terminal. However, this would require explicit signaling to each of the terminals and possibly reconfiguration at handover. Instead, the terminal-specific search spaces in LTE are defined without explicit signaling through a function of the terminal identity and implicitly the subframe number. Dependence on the subframe number results in the terminal-specific search spaces being time varying, which helps resolve blocking between terminals. If a given terminal cannot be scheduled in a subframe as all the CCEs that the terminal is monitoring have already been used for scheduling other terminals in the same subframe, the time-varying definition of the terminal-specific search spaces is likely to resolve the blocking in the next subframe.

In several situations, there is a need to address a group of, or all, terminals in the system. One example hereof is dynamic scheduling of system information, another is transmission of paging messages, both described in Chapter 14. Transmission of explicit power-control commands to a group of terminals is a third example. To allow all terminals to be addressed at the same time, LTE has defined *common search spaces* in addition to the terminal-specific search spaces. A common search space is, as the name implies, common, and all terminals in the cell monitor the CCEs in the common search spaces for control information. Although the motivation for the common search space is primarily transmission of various system messages, it can be used to schedule individual terminals as well. Thus, it can be used to resolve situations where scheduling of one terminal is blocked due to lack of available resources in the terminal-specific search space. Unlike unicast transmissions, where the transmission parameters of the control signaling can be tuned to match the channel conditions of a specific terminal, system messages typically need to reach the cell border. Furthermore, the data rate of the associated DL-SCH transmission is typically modest. Consequently, the common search spaces are only defined for aggregation levels of four and eight CCEs and only for the smallest DCI formats, 0/1A/3/3A and 1C. There is no support for DCI formats with spatial multiplexing in the common search space. This helps to reduce the number of blind decoding attempts in the terminal used for monitoring the common search space.

Table 10.7 Downlink DCI Formats Monitored in Different Search Spaces for C-RNTI

Mode	Monitored DCI Formats in Search Space			Description
	Common	UE-Specific	UE-Specific	
1	1A	1A	1	Single antenna transmission
2	1A	1A	1	Transmit diversity
3	1A	1A	2A	Open-loop spatial multiplexing
4	1A	1A	2	Closed-loop spatial multiplexing
5	1A	1A	1D	Multi-user MIMO
6	1A	1A	1B	Single layer codebook-based precoding
7	1A	1A	1	Single-layer transmission using DM-RS
8	1A	1A	2B	Dual-layer transmission using DM-RS
9	1A	1A	2C	Multi-layer transmission using DM-RS

Figure 10.34 illustrates the terminal-specific and common search spaces in two terminals in a certain subframe. The terminal-specific search spaces are different in the two terminals and will, as described above, vary from subframe to subframe. Furthermore, the terminal-specific search spaces partially overlap between the two terminals in this subframe (CCEs 24–31 on aggregation level 8) but, as the terminal-specific search space varies between subframes, the overlap in the next subframe is most likely different.

The downlink DCI formats to decode in the terminal-specific search space depend on the *transmission mode* configured for the terminal. Transmission modes are described in Section 10.3 and, in principle, correspond to different multi-antenna configurations. As an example, there is no need to attempt to decode DCI format 2 when the terminal has not been configured for spatial multiplexing, which helps to reduce the number of blind decoding attempts. Configuration of the transmission mode is done via RRC signaling. As the exact subframe number when this configuration takes effect in the terminal is not specified and may vary depending on, for example, RLC retransmissions, there is a (short) period when the network and the terminal may have different understandings of which transmission mode is configured. Therefore, in order not to lose the possibility of communicating with the terminal, it is necessary to have at least one DCI format that is decoded irrespective of the transmission mode. For downlink transmissions, DCI format 1A serves this purpose and the network can therefore always transmit data to the terminal using this DCI format. Another function of format 1A is to reduce overhead for transmissions when full flexibility in resource-block assignment is not needed.

In Table 10.7, the DCI monitoring is described for the case of a single component carrier. For multiple component carriers, the above procedures are applied in principle to each of the activated[33] downlink component carriers. Hence, in principle there is one UE-specific search space per aggregation level and per (activated) component carrier upon which PDSCH can be received (or PUSCH transmitted), although there are some carrier-aggregation-specific modifications. Clearly, for terminals configured to use carrier aggregation, this results in an increase in the number of blind decoding

[33] Individual component carriers can be activated/deactivated as discussed in Chapter 12.

attempts compared to a terminal not using carrier aggregation, as scheduling assignments/grants for each of the component carriers need to be monitored.

The common search space is only defined for transmissions on the primary component carrier. As the main function of the common search space is to handle scheduling of system information intended for multiple terminals, and such information must be receivable by all terminals in the cell, scheduling in this case uses the common search space. For this reason, the carrier indication field is never present in DCI formats monitored in the common search space.

As mentioned before, there is one UE-specific search space per aggregation level and component carrier[34] used for the PDSCH/PUSCH. This is illustrated in Figure 10.30, where PDSCH/PUSCH transmissions on component carrier 1 are scheduled using PDCCHs transmitted on component carrier 1. No carrier indicator is assumed in the UE-specific search space for component carrier 1 as cross-carrier scheduling is not used. For component carrier 2, on the other hand, a carrier indicator is assumed in the UE-specific search space as component carrier 2 is cross-carrier scheduled from PDCCHs transmitted on component carrier 1.

Search spaces for different component carriers may overlap in some subframes. In Figure 10.30, this happens for the UE-specific search spaces for component carriers 3 and 4. The terminal will handle the two search spaces independently, assuming (in this example) a carrier indicator for component carrier 4 but not for component carrier 3. If the UE-specific and common search spaces relating to different component carriers happen to overlap for some aggregation level when cross-carrier scheduling is configured, the terminal only needs to monitor the common search space. The reason for this is to avoid ambiguities; if the component carriers have different bandwidths a DCI format in the common search space may have the same payload size as another DCI format in the UE-specific search space relating to another component carrier.

[34] If the DCI formats for two different component carriers have the same size, then search spaces can be shared such that, in the example in Figure 10.30, component carrier 2 can be scheduled from search spaces 2 *and* 1. This comes "for free" as the DCI formats in search space 1 need to be monitored anyway and the carrier-indicator field points to the component carrier for which the grant/assignment is intended.

CHAPTER

Final Thoughts 20

The focus of this book has been on the LTE radio-access technology, from the first version in release 8, via release 9 to release 10. LTE release 10 is the latest version of the LTE radio-access specifications, sometimes also referred to as *LTE-Advanced* and approved by the ITU as the first *IMT-Advanced* technology.

The basic technologies described in Chapters 2–6 served as a foundation for the detailed discussion and explanation of the LTE radio-access technology in the subsequent chapters, providing information about the LTE radio-network architecture, the basic transmission structure, and all the protocols and procedures needed for a complete radio-access network.

The capabilities and system performance of LTE were illustrated in Chapter 18 and it was shown that all the requirements defined by the ITU for IMT-Advanced were fulfilled, in many cases by a significant margin.

At the same time as ITU-R is completing the work on the formal IMT-Advanced recommendations, activities are initiated to look at what will come beyond IMT-Advanced in the 2012–2020 time frame. Several regional workshops on *"IMT for the 2010–2020 Decade"* are planned for 2011 and a new report looking at global broadband wireless services and marketplace for IMT has been initiated within ITU-R WP5D. The basis for the continued work is that future mobile broadband consumers will expect similar user experience and access to similar services and applications in the mobile environment as for fixed broadband, with the added convenience of mobility. Studies within ITU-R will include long-term market forecasts and take into account new user behavior, specifically considering personal computers and smart phones.

20.1 WHERE TO GO IN THE FUTURE?

The evolution of LTE does not end with LTE release 10. Rather LTE will continue to evolve into release 11, release 12, and so on, with each new release bringing additional capabilities and further enhanced system performance into the LTE radio-access technology. Not only will the additional capabilities provide better performance in existing applications, they may also open up for, or even be motivated by, new application areas. Examples hereof are home automation, smart transportation, security, and e-books, but the list is continuously growing as additional applications benefiting from mobile connectivity are emerging. Similarly, new ways of deploying cellular networks, for example more extensive use of massive beam-forming or ubiquitous access to optical fibers in the backhaul, may call for enhancements in the radio interface.

4G LTE/LTE-Advanced for Mobile Broadband.
© 2011 Erik Dahlman, Stefan Parkvall & Johan Sköld. Published by Elsevier Ltd. All rights reserved.

Samsung Exhibit 1019, Page 175 of 195

Initial 3GPP discussions on release 11 have already started, with actual work expected to be initiated in early 2011, aiming for completion of this release in mid-2012. In the following, some possible areas for the future evolution of LTE will be briefly discussed. Some of these areas may happen in release 11 or 12, while others may be more futuristic.

20.1.1 Advanced Multi-Cell Coordination

Coordinated Multi-Point transmission and reception (CoMP) refers to a wide range of different techniques with the common denominator being the dynamic coordination of transmission and/or reception at multiple geographically separated sites with the aim to enhance system performance and end-user service quality. Many different coordination schemes of very different characteristics fall under the joint umbrella of CoMP, ranging from dynamic inter-cell scheduling coordination to joint transmission/reception at multiple sites. In the former case, CoMP can, to a large extent, be seen as an extension of the inter-cell interference coordination that is already part of LTE, discussed in Chapter 13.

Joint reception means that the signals received at multiple sites are jointly processed for enhanced reception performance. Maximum-ratio combining and interference-rejection combining are examples of schemes that can be used to combine the uplink transmission received at multiple points. This is, in many respects, similar to softer handover used within a site, for example in WCDMA/HSPA-based systems, but extended to multiple sites.

Joint transmission implies that data is transmitted to a mobile terminal jointly from several sites, thereby not only reducing the interference but also increasing the received power. The transmission from the sites can also take the instantaneous channel conditions at the different terminals into account to enhance the received signal strength, while at the same time reducing the interference between different transmissions.

In general, both joint reception and transmission pose high requirements on low latency in the communication between the network node involved in the joint processing and the different antennas involved in the reception/transmission. Hence, in practice, it can be expected that the different sites may be connected in the form of a *centralized RAN* (C-RAN) deployment for example, as briefly discussed in Chapter 13.

20.1.2 Network Energy Efficiency

Low energy consumption for mobile terminals has been an important requirement ever since the emergence of hand-held terminals roughly 20 years ago. The driving force has been the reduction in battery size and improved battery time. Today, reduced energy consumption *also in the radio-access network* is receiving increased attention for several reasons:

- The cost of energy is a far from negligible part of the overall operational cost for the operator. Thus, reduced energy consumption is one component in the everlasting quest for reduced operating costs for network operators.
- In some rural areas, it may not even be possible to connect the base station to the electrical grid. With sufficiently low energy consumption, reasonably sized solar panels could be used as power source, instead of the diesel generators commonly used today.

- In today's world, where energy consumption and the related climate impact is seen as one of the great challenges for the future, the mobile industry should lead the way for reduced energy consumption.

Regarding the third point, it is important to bear in mind though that the entire ICT (*Information and Communication Technologies*) sector today (2010) contributes roughly 2% of the overall world energy consumption and the contribution of the mobile-communication sector is just a fraction of that. At the same time, the emergence of mobile broadband communication opens up tremendous opportunities for reducing the global energy consumption in general by, for example, virtual meetings replacing traveling to physical meetings.

Still, low energy consumption in mobile networks is important. In particular, taking into account large future traffic increases, where traffic growth of several hundred times up to perhaps 1000 times can be expected in a longer-term perspective, low energy consumption of the mobile networks will most likely be as important a performance metric as capacity, data rates and latency, and must be treated as such.

Reducing the energy consumption of mobile networks is, to a large extent, an implementation issue. However, it is important to ensure that the basic principles of the radio-access technology allows for low energy consumption. One key characteristic is that the energy consumption should scale with the use of the radio resources, not only on average but also on a short-term basis. That is, during points in time where there is no traffic, the energy consumption should be reduced to an absolute minimum, for example by avoiding all but absolutely necessary transmissions. LTE already provides several tools that can be used for this. Nevertheless, the future evolution of LTE should further strive for minimizing transmission of signals strictly not needed.

This will be even more important as the network becomes more and more dense with more and more network nodes. In such network deployments, which may typically be heterogeneous networks as discussed in Chapter 13, pico cells may often be deployed to provide high data rates, rather then being needed for capacity reasons. Thus, the load per cell may be relatively low and each pico cell may often be more or less "empty", further stressing the importance of very low energy consumption when idle.

20.1.3 Machine-Type Communication

Similar to earlier mobile-communication systems, LTE has been designed with data services in mind and much effort has been put into developing techniques for providing high data rates and low latencies for services such as file downloading and web browsing. However, with the increased availability of mobile broadband, connectivity has also become a realistic option for machine-type communication. Machine-type communication spans a wide range of applications, from massive deployment of low-cost battery-powered sensors to remote-controlled utility meters, to surveillance cameras. Many of these applications can be handled by LTE already; communicating with a surveillance camera, for example, is not significantly different from uploading a file, in which case high data rates are paramount. However, other applications may not require transmission of large amounts of data or low latency, but rather pose challenges in terms of a vast amount of devices connecting to the network.

According to some sources [107], 50 billion connections, most of them machine-type communications, can be expected in the year 2020. Handling such a large number of devices is likely to be a challenge mainly for the core network, but improvements in the area of connection setup and power-efficient handling of control signaling in the radio-access network may be of interest.

20.1.4 New ways of Using Spectrum

Spectrum has been, is, and will continue to be a scarce resource for the mobile-communication industry. In particular, in light of the continuous increase in the data rates requested and the corresponding need for wider bandwidths, it is expected that spectrum will remain a scarce resource.

Historically, up until now, the mobile industry has relied on spectrum dedicated for mobile communication and licensed to a certain operator. This will also clearly be the main track for the future. However, in situations where licensed spectrum is not available, other possibilities for increasing the spectrum availability are of interest. This could include the use of unlicensed spectrum, or secondary spectrum primarily used for other communication services, as a complement to operation in the licensed spectrum. Broadcast spectrum not used (in some areas) is often referred to as "white space" [108]. Related to this is the concept of *cognitive radio* [109,110] – that is, radio-access-related functionality related to "smart" selection of spectrum usage. However, the applicability of cognitive radio to cellular communication is a relatively new area and further studies are required to assess the feasibility and impact of such usage.

20.1.5 Direct Device-to-Device Communication

One possible longer-term evolution path would be to extend the LTE radio-access technology with support for direct device-to-device communication – that is, make it possible for two mobile terminals to communicate directly with each other without going via the network. This has been studied in academia for some years, for example as part of the European research project WINNER [111]. Scenarios where the possibility for direct device-to-device communication could be of interest include situations where no network infrastructure is available, for example national security and public safety (NSPS), and situations where network infrastructure is available but when communication directly between the terminals could be more efficient, either in terms of requiring a smaller amount of radio resources for the same communication quality or allowing for improved communication quality. In the latter case, one could envision different degrees of network interaction in the device-to-device communication. However, this area is still in its infancy and scenarios where device-to-device communication offers benefits and how to interact with infrastructure-based communication need to be better understood.

Related to device-to-device communication is the possibility to also use LTE as a radio-access technology in the home, for example a digital camera connecting to the TV. Although such connectivity is possible to some extent today by using (semi-proprietary) technology, basing this type of communication on LTE could be attractive. When outside the home, the devices can communicate with the cellular LTE network, but while at home they can seamlessly connect to the local LTE network covering the home. Furthermore, for devices that are to be used at home only, operation without a SIM card and interaction with an external core network should be possible. Adding this type of functionality to LTE is most likely a relatively simple task.

20.2 CONCLUDING REMARKS

The list of technology areas above should be seen as examples. Some of these technologies will most likely be part of future releases of LTE, while others may not happen at all. There may also be other technologies, not listed above and maybe yet to be discovered, that could be of interest for LTE evolution. What is clear though is that LTE is a very flexible platform that can evolve in different directions to meet the future needs of wireless communication. Given the size of the LTE ecosystem, such evolution is a very attractive path for future wireless communication.

References

[1] ITU-R, International mobile telecommunications-2000 (IMT-2000), Recommendation ITU-R M.687-2, February 1997.

[2] ITU-R, Detailed specifications of the radio interfaces of international mobile telecommunications-2000 (IMT-2000), Recommendation ITU-R M.1457-9, May 2010.

[3] ITU-R, Principles for the process of development of IMT-advanced, Resolution ITU-R 57, October 2007.

[4] ITU-R, Framework and overall objectives of the future development of IMT-2000 and systems beyond IMT-2000, Recommendation ITU-R M.1645, June 2003.

[5] ITU-R, Invitation for submission of proposals for candidate radio interface technologies for the terrestrial components of the radio interface(s) for IMT-advanced and invitation to participate in their subsequent evaluation, ITU-R SG5, Circular Letter 5/LCCE/2, March 2008.

[6] ITU-R, ITU paves way for next-generation 4G mobile technologies; ITU-R IMT-advanced 4G standards to usher new era of mobile broadband communications, ITU Press Release, 21 October 2010.

[7] ITU-R, Working document towards a preliminary draft new recommendation ITU-R M.[IMT. RSPEC] ITU-R WP5D, Contribution 870, Attachment 5.11.

[8] ITU-R, Frequency arrangements for implementation of the terrestrial component of international mobile telecommunications-2000 (IMT-2000) in the bands 806–960 MHz, 1710–2025 MHz, 2110–2200 MHz and 2500–2690 MHz, Recommendation ITU-R M.1036-3, July 2007.

[9] M. Olsson, S. Sultana, S. Rommer, L. Frid, C. Mulligan, SAE and the Evolved Packet Core – Driving the Mobile Broadband Revolution, Academic Press, 2009.

[10] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN) (Release 7), 3GPP TR 25.913.

[11] C.E. Shannon, A mathematical theory of communication, Bell System Tech. J. 27 (July and October 1948) 379–423, 623–656.

[12] J.G. Proakis, Digital Communications, McGraw-Hill, New York, 2001.

[13] G. Bottomley, T. Ottosson, Y.-P. Eric Wang, A generalized RAKE receiver for interference suppression, IEEE J. Sel. Area Comm. 18 (8) (August 2000) 1536–1545.

[14] WiMAX Forum, Mobile WiMAX – Part I: A technical overview and performance evaluation, White Paper, August 2006.

[15] ETSI, Digital video broadcasting (DVB): Framing structure, channel coding and modulation for digital terrestrial television, ETS EN 300 744 v. 1.1.2.

[16] NTT DoCoMo et al., Views on OFDM parameter set for evolved UTRA downlink, Tdoc R1-050386, 3GPP TSG-RAN WG 1, Athens, Greece, 9–13 May 2005.

[17] L. Hanzo, M. Munster, B.J. Choi, and T. Keller, OFDM and MC-CDMA for broadband multi-user communications, WLANs, and broadcasting, Wiley-IEEE Press, Chichester, UK, ISBN 0-48085879.

Samsung Exhibit 1019, Page 180 of 195

[18] R. van Nee, R. Prasad, OFDM for Wireless Multimedia Communications, Artech House Publishers, London, January 2000.
[19] J. Tellado and J.M. Cioffi, PAR reduction in multi-carrier transmission systems, ANSI T1E1.4/97-367.
[20] P.V. Etvalt, Peak to average power reduction for OFDM schemes by selective scrambling, Electron. Lett. 32 (21) (October 1996) 1963–1964.
[21] W. Zirwas, Single frequency network concepts for cellular OFDM radio systems, International OFDM Workshop, Hamburg, Germany, September 2000.
[22] A. Oppenheim, R.W. Schafer, Digital Signal Processing, Prentice-Hall International, ISBN 0-13-214107-8 01.
[23] S. Haykin, Adaptive Filter Theory, Prentice-Hall International, NJ, USA, 1986. ISBN 0-13-004052-5 025.
[24] D. Falconer, S.L. Ariyavisitakul, A. Benyamin-Seeyar, B. Eidson, Frequency domain equalization for single-carrier broadband wireless systems, IEEE Commun. Mag. 40 (4) (April 2002) 58–66.
[25] G. Forney, Maximum likelihood sequence estimation of digital sequences in the presence of intersymbol interference, IEEE T. Inform. Theory IT-18 (May 1972) 363–378.
[26] G. Forney, The viterbi algorithm, Proceedings of the IEEE, Piscataway, NJ, USA, Vol. 61, No. 3, March, 1973, 268–278.
[27] Motorola, Comparison of PAR and Cubic Metric for Power De-rating, Tdoc R1-040642, 3GPP TSG-RAN WG1, May 2004.
[28] U. Sorger, I. De Broeck, M. Schnell, IFDMA – A new spread-spectrum multiple-access scheme, Proceedings of the ICC'98, Atlanta, GA, USA, June 1998, pp. 1267–1272.
[29] W. Lee, Mobile Communications Engineering, McGraw-Hill, New York, USA, ISBN 0-07-037039-7.
[30] J. Karlsson, J. Heinegard, Interference rejection combining for GSM, Proceedings of the 5th IEEE International Conference on Universal Personal Communications, Cambridge, MA, USA, 1996, pp. 433–437.
[31] L.C. Godara, Applications of antenna arrays to mobile communications, Part I: Beam-forming and direction-of-arrival considerations, Proceedings of the IEEE, Vol. 85, No. 7, July, 1997, 1029–1030.
[32] L.C. Godara, Applications of antenna arrays to mobile communications, Part II: beam-forming and direction-of-arrival considerations, Proceedings of the IEEE, Piscataway, Vol. 85, No. 7, July, 1997, 1031–1060.
[33] A. Huebner, F. Schuehlein, M. Bossert, E. Costa, H. Haas, A simple space–frequency coding scheme with cyclic delay diversity for OFDM, Proceedings of the 5th European Personal Mobile Communications Conference, Glasgow, Scotland, April 2000, pp. 106–110.
[34] 3GPP, 3rd generation partnership project; technical specification group radio access network; Physical Channels and Mapping of Transport Channels onto Physical Channels (FDD), 3GP TS 25.211.
[35] V. Tarokh, N. Seshadri, A. Calderbank, Space–time block codes from orthogonal design, IEEE T. Inform. Theory 45 (5) (July 1999) 1456–1467.
[36] A. Hottinen, O. Tirkkonen, and R. Wichman, Multi-Antenna Transceiver Techniques for 3G and Beyond, John Wiley, Chichester, UK, ISBN 0470 84542 2.
[37] C. Kambiz, L. Krasny, Capacity-achieving transmitter and receiver pairs for dispersive MISO channels, IEEE T. Commun. 42 (April 1994) 1431–1440.

[38] R. Horn and C. Johnson, Matrix analysis, Proceedings of the 36th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, November 2002.
[39] K.J. Kim, Channel estimation and data detection algorithms for MIMO–OFDM systems, Proceedings of the 36th Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, November 2002.
[40] M.K. Varanasi and T. Guess, Optimum decision feedback multi-user equalization with successive decoding achieves the total capacity of the gaussian multiple-access channel, Proceedings of the Asilomar conference on Signals, Systems, and Computers, Monterey, CA, November 1997.
[41] S. Grant, J.-F. Cheng, L. Krasny, K. Molnar, Y.-P.E Wang, Per-antenna-rate-control (PARC) in frequency selective fading with SIC-GRAKE receiver, 60th IEEE Vehicular Technology Conference, Los Angeles, CA, USA 2, September 2004, 1458–1462.
[42] S.T. Chung, A.J. Goldsmith, Degrees of freedom in adaptive modulation: A unified view, IEEE T. Commun. 49 (9) (September 2001) 1561–1571.
[43] A.J. Goldsmith, P. Varaiya, Capacity of fading channels with channel side information, IEEE T. Inform. Theory 43 (November 1997) 1986–1992.
[44] R. Knopp, P.A. Humblet, Information capacity and power control in single-cell multi-user communications, Proceedings of the IEEE International Conference on Communications, Seattle, WA, USA, Vol. 1, 1995, 331–335.
[45] D. Tse, Optimal power allocation over parallel gaussian broadcast channels, Proceedings of the International Symposium on Information Theory, Ulm, Germany, June 1997, p. 7.
[46] M.L. Honig and U. Madhow, Hybrid intra-cell TDMA/inter-cell CDMA with inter-cell interference suppression for wireless networks, Proceedings of the IEEE Vehicular Technology Conference, Secaucus, NJ, USA, 1993, pp. 309–312.
[47] S. Ramakrishna, J.M. Holtzman, A scheme for throughput maximization in a dual-class CDMA system, IEEE J. Sel. Area Comm. 16 (6) (1998) 830–844.
[48] S.-J. Oh, K.M. Wasserman, Optimality of greedy power control and variable spreading gain in multi-class CDMA mobile networks, Proceedings of the AMC/IEEE MobiComp, Seattle, WA, USA, 1999, pp. 102–112.
[49] J.M. Holtzman, CDMA forward link waterfilling power control, Proceedings of the IEEE Vehicular Technology Conference, Tokyo, Japan, Vol. 3, May, 2000, 1663–1667.
[50] J.M. Holtzman, Asymptotic analysis of proportional fair algorithm, Proceedings of the IEEE Conference on Personal Indoor and Mobile Radio Communications, San Diego, CA, USA, Vol. 2, 2001, 33–37.
[51] P. Viswanath, D. Tse, R. Laroia, Opportunistic beamforming using dumb antennas, IEEE T. Inform. Theory 48 (6) (2002) 1277–1294.
[52] S. Lin and D. Costello, Error Control Coding, Prentice-Hall, Upper Saddle River, NJ, USA.
[53] C. Schlegel, Trellis and Turbo Coding, Wiley–IEEE Press, Chichester, UK, March 2004.
[54] J.M. Wozencraft, M. Horstein, Digitalised Communication Over Two-way Channels, Fourth London Symposium on Information Theory, London, UK, September 1960.
[55] D. Chase, Code combining – a maximum-likelihood decoding approach for combining and arbitrary number of noisy packets, IEEE T. Commun. 33 (May 1985) 385–393.
[56] M.B. Pursley, S.D. Sandberg, Incremental-redundancy transmission for meteor-burst communications, IEEE T. Commun. 39 (May 1991) 689–702.
[57] S.B. Wicker, M. Bartz, Type-I hybrid ARQ protocols using punctured MDS codes, IEEE T. Commun. 42 (April 1994) 1431–1440.

Samsung Exhibit 1019, Page 182 of 195

[58] J.-F. Cheng, Coding performance of hybrid ARQ schemes, IEEE T. Commun. 54 (June 2006) 1017–1029.
[59] P. Frenger, S. Parkvall, and E. Dahlman, Performance comparison of HARQ with chase combining and incremental redundancy for HSDPA, Proceedings of the IEEE Vehicular Technology Conference, Atlantic City, NJ, USA, October 2001, pp. 1829–1833.
[60] J. Hagenauer, Rate-compatible punctured convolutional codes (RCPC codes) and their applications, IEEE T. Commun. 36 (April 1988) 389–400.
[61] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Physical Channels and Modulation (Release 8), 3GPP TS 36.211.
[62] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Multiplexing and Channel Coding (Release 8), 3GPP TS 36.212.
[63] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Physical Layer Procedures (Release 8), 3GPP TS 36.213.
[64] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Physical Layer – Measurements (Release 8), 3GPP TS 36.214.
[65] ITU-R, Requirements related to technical performance for IMT-Advanced radio interface(s), Report ITU-R M.2134, 2008.
[66] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Requirements for further advancements for Evolved Universal Terrestrial Radio Access (E-UTRA) (LTE-Advanced) (Release 9), 3GPP TR 36.913.
[67] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Evolved universal terrestrial radio access (E-UTRA); User Equipment (UE) Radio Access Capabilities, 3GPP TS 36.306.
[68] IETF, Robust header compression (ROHC): Framework and four profiles: RTP, UDP, ESP, and Uncompressed, RFC 3095.
[69] J. Sun, O.Y. Takeshita, Interleavers for turbo codes using permutation polynomials over integer rings, IEEE T. Inform. Theory 51 (1) (January 2005) 101–119.
[70] O.Y. Takeshita, On maximum contention-free interleavers and permutation polynomials over integer rings, IEEE T. Inform. Theory 52 (3) (March 2006) 1249–1253.
[71] D.C. Chu, Polyphase codes with good periodic correlation properties, IEEE T. Inform. Theory 18 (4) (July 1972) 531–532.
[72] J. Padhye, V. Firoiu, D.F. Towsley, J.F. Kurose, Modelling TCP reno performance: A simple model and its empirical validation, ACM/IEEE T. Network. 8 (2) (2000) 133–145.
[73] CRAN International Workshop, CRAN – Road Towards Green Radio Access Network, China Mobile, 23 April 2010.
[74] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN); User equipment (UE) conformance specification; Radio Transmission and Reception (Part 1, 2 and 3), 3GPP TS 36.521.
[75] 3GPP, Evolved universal terrestrial radio access (E-UTRA); Physical layer for relaying operation, 3GPP TS 36.216.
[76] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Evolved universal terrestrial radio access (E-UTRA); User Equipment (UE) radio transmission and reception, 3GPP TS 36.101.

[77] APT Wireless Forum, APT Report on Harmonised Frequency Arrangements for The Band 698–806 MHz, APT/AWF/REP-14, September 2010 Edition.
[78] 3GPP, 3rd generation partnership project; Technical specification group radio access network; UMTS-LTE 3500 MHz Work Item Technical Report (Release 10), 3GPP TR 37.801.
[79] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Feasibility Study for Evolved Universal Terrestrial Radio Access (UTRA) and Universal Terrestrial Radio Access Network (UTRAN) (Release 7), 3GPP TR 25.912.
[80] 3GPP, E-UTRA, UTRA and GSM/EDGE; Multi-Standard Radio (MSR) Base Station (BS) Radio Transmission and Reception, 3GPP TR 37.104.
[81] 3GPP, E-UTRA, UTRA and GSM/EDGE; Multi-Standard Radio (MSR) Base Station (BS) Conformance Testing, 3GPP TR 37.141.
[82] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Evolved universal terrestrial radio access (E-UTRA); User Equipment (UE) Radio Transmission and Reception, 3GPP TR 36.803.
[83] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Evolved universal terrestrial radio access (E-UTRA); Base Station (BS) Radio Transmission and Reception, 3GPP TR 36.804.
[84] 3GPP, Evolved universal terrestrial radio access (E-UTRA); User Equipment (UE) Radio transmission and reception, 3GPP TR 36.807.
[85] 3GPP, Evolved universal terrestrial radio access (E-UTRA); Carrier Aggregation Base Station (BS) Radio transmission and reception, 3GPP TR 36.808.
[86] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Evolved universal terrestrial radio access (E-UTRA); Base Station (BS) Radio transmission and reception, 3GPP TS 36.104.
[87] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Evolved universal terrestrial radio access (E-UTRA); Base Station (BS) conformance testing, 3GPP TS 36.141.
[88] FCC, Title 47 of the Code of Federal Regulations (CFR), Federal Communications Commission.
[89] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Evolved universal terrestrial radio access (E-UTRA); Radio Frequency (RF) system scenarios, 3GPP TR 36.942.
[90] ITU-R, Unwanted Emissions in the Spurious Domain, Recommendation ITU-R SM.329-10, February 2003.
[91] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Radio Frequency (RF) system scenarios, 3GPP TR 25.942.
[92] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Feasibility study for further advancements for E-UTRA (LTE-Advanced) (Release 9), 3GPP TR 36.912.
[93] ITU-R, Guidelines for Evaluation of Radio Interface Technologies for IMT-Advanced, Report ITU-R M.2135-1, December 2009.
[94] A. Furuskär, Performance evaluations of LTE-advanced – the 3GPP ITU proposal, 12th International Symposium on Wireless Personal Multimedia Communications (WPMC 2009), Invited Paper.

[95] E. Dahlman, H. Ekström, A. Furuskär, Y. Jading, J. Karlsson, M. Lundevall, S. Parkvall, The 3G long-term evolution – radio interface concepts and performance evaluation, 63rd Vehicular Technology Conference, VTC 2006 – Spring, Vol. 1, pp. 137–141, IEEE, 2006.
[96] E. Dahlman, S. Parkvall, J. Sköld, P. Beming, 3G Evolution – HSPA and LTE for Mobile Broadband, second ed., Academic Press, 2008.
[97] WINNER, IST-WINNER II Deliverable 1.1.2 v. 1.2. WINNER II Channel Models, IST-WINNER2, Tech. Rep., 2007.
[98] 3GPP, 3rd generation partnership project; Technical specification group radio access network; Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations (Release 9), 3GPP TR 25.996.
[99] Alcatel Shanghai Bell, CATT, CMCC, RITT, Spectrum Communications, TD-TECH, ZTE, Introduce TD-SCDMA Industry Standard in CCSA to 3GPP, Document R4-071394, 3GPP TSG-RAN WG4 meeting #44, Athens, Greece, August 2007.
[100] 3GPP, 3rd generation partnership project; Technical specification group GSM/EDGE radio access network; Feasibility Study for Evolved GSM/EDGE Radio Access Network (GERAN) (Release 7), 3GPP TR 45.912.
[101] H. Axelsson, P. Björkén, P. de Bruin, S. Eriksson, H. Persson, GSM/EDGE continued evolution, Ericsson Rev. (01) (2006) 20–29 Telefonaktiebolaget LM Ericsson, Stockholm, Sweden.
[102] 3GPP, Updated New WID on Higher Uplink Performance for GERAN Evolution (HUGE), Tdoc GP-061901, 3GPP TSG GERAN #31, Denver, USA, 4–8 September 2006.
[103] 3GPP, Circuit Switched Voice Capacity Evolution for GSM/EDGE Radio Access Network (GERAN), 3GPP TR 45.914.
[104] N. Bhushan, C. Lott, P. Black, R. Attar, Y.-C. Jou, M. Fan, et al., CDMA2000 1xEV-DO revision A: A physical layer and MAC layer overview, IEEE Commun. Mag. (February 2006) 75–87.
[105] M.W. Thelander, The 3G Evolution: Taking CDMA2000 into the Next Decade, Signal Research Group, LLC, White Paper developed for the CDMA Development Group, October 2005.
[106] R. Attar, Evolution of CDMA2000 cellular networks: Multicarrier EV-DO, IEEE Commun. Mag. (March 2006) 46–53.
[107] M. Alendal, Operators need an ecosystem to support 50 billion connections, Ericsson Bus. Rev. (3) (2010).
[108] M. Nekovee, A survey of cognitive radio access to TV white spaces, Int. J. Digi. Multimed. Broadcast. (2010) 1–11.
[109] J. Mitola, Cognitive Radio – An Integrated Agent Architecture for Software Defined Radio, Royal Institute of Technology (KTH), Sweden, May 2000.
[110] S. Haykin, Cognitive radio: Brain-empowered wireless communications, IEEE J. Sel. Area. Comm. 23 (2) (February 2005) 201–220.
[111] Wireless World Initiative New Radio, Eurescom, 2006, https://www.ist-winner.org.

Index

1.28 Mcps TDD, 395
16QAM modulation, 19–21, 399–400
3G, 2–11
3GPP, 9–11
3GPP2, 9, 402–403
32QAM modulation, 399–400
64QAM modulation, 19–21
7.68 Mcps TDD, 395
8PSK modulation, 399–400

A

Absolute grant, 393
Access Stratum (AS), 110
ACIR, 370
Acknowledged mode (for RLC), 261
Acknowledged Resource Indicator (ARI), 184, 238
ACLR, 370–371
ACS, 370–371, 374
Adaptive Modulation and Coding (AMC), 81
Antenna port, 148
Automatic Repeat Request (ARQ), 90–93
  in LTE, 247–259
  for LTE relaying, 335–337

B

Bandwidth utilization, 16
Base-station classes, 365
BCCH *see* Broadcast Control Channel
BCH *see* Broadcast Channel
  transmission of, 305–307
Beam forming:
  classical, 68–70
  in LTE *see* Pre-coding, LTE
Blind decoding of PDCCH, 199–202
Blocking, receiver:
  general, 362, 374
  narrowband, 374
BM-SC *see* Broadcast Multicast Service Center
Broadcast, 43–44
  see also MBMS
Broadcast Control Channel (BCCH), 116
Broadcast Channel (BCH), 116
Broadcast Multicast Service Center (BM-SC), 325
Buffer status report, 277–280

C

Carrier aggregation:
  control signalling for, 193–199
  overview, 104
  physical resources, 132–134
  RF characteristics, 355–358
Carrier Indicator, 183, 190
Carrier raster, 136
CCCH *see* Common Control Channel
CCE *see* Control Channel Element
CDD *see* Cyclic Delay Diversity
CDMA2000, 10, 402–405
Cell identity, 301
Cell-identity group, 302
Cell search, 301–304
Centralized RAN (C-RAN), 292
Channel capacity, 15
Channel coding, LTE:
  downlink, 146–148
  uplink, 204
Channel rank, 166
Channel State Information (CSI), 152

Samsung Exhibit 1019, Page 186 of 195

Channel-state report, 282–287
Channel-Quality Indicator (CQI), 242, 283
Charging, 109–110
Chase combining, 90–92
Closed Subscriber Group (CSG), 294
CN *see* Core Network
Code block, 146
Coexistence
  requirements, 352
  with TD-SCDMA, 140–141
Cognitive radio, 414
Co-location requirements, 352
Common Control Channel (CCCH), 116
Common Subframe Allocation (CSA), 328
CoMP *see* Coordinated Multipoint transmission
Component carrier, 104, 132
Contention resolution, 312, 319
Continuous Packet Connectivity (CPC), 390
Control-channel element, 195
Controlling RNC *see* RNC
Control region, 131, 173–174
Coordinated multipoint transmission, 412
Core Network (CN), 109–110
CQI *see* Channel-Quality Indicator
CRC *see* Cyclic Redundancy Check
C-RNTI, 125, 319
Cross-carrier scheduling, 193–194, 254
CSG *see* Closed Subscriber Group
CSI *see* Channel-State Information
Cubic metric, 54
Cyclic-Delay Diversity (CDD), 66
Cyclic prefix:
  extended, 127–128
  MBSFN, 131–132
  normal, 127–128
  OFDM, 32–34, 40
  Random-access preamble, 315–317
  single-carrier, 40
Cyclic Redundancy Check (CRC), 90
  for DL-SCH, 144–145
  for PDCCH, 195–196
  for UL-SCH, 203–204

D

Data region, 131, 173
DC-subcarrier, 130–131
DCCH *see* Dedicated Control Channel
DCI *see* Downlink Control Information
DCI format, 180–185
Dedicated Control Channel (DCCH), 116
Delay diversity, 65–66
Device-to-device communication, 414
DFT-spread OFDM (DFTS-OFDM), 52–57
  localized vs. distributed, 55–57
  receiver, 54–55
  user multiplexing, 55
  in LTE, 96–97, 205
Discontinuous Reception (DRX), 287–290
Discontinuous Transmission (DTX), 394
  DL-SCH *see* Downlink Shared Channel
Donor cell, 332
Downlink Control Information (DCI), 123, 179
Downlink Shared Channel (DL-SCH), 117
  processing, 143
Downlink Traffic Channel (DTCH), 116
DRX *see* Discontinuous Reception
DTCH *see* Downlink Traffic Channel
DTX *see* Discontinuous Transmission
DwPTS, 138–140

E

EDGE, 396–401
eNB *see* eNodeB
eNodeB, 111
EPC *see* Evolved Packet Core

EPS *see* Evolved Packet System
Equalization, 23, 45–49
  Decision-feedback, 49–50
  Frequency-domain, 47–49
  Time-domain, 45–47
EV-DO, 402–405
EV-DV, 402
EVM requirement, 367
Evolved Packet Core, 109–110
Evolved Packet System, 109

F

FDD *see* Frequency Division Duplex
Femto base station *see* Home-eNodeB
Frame, LTE, 127, 131
Frequency bands, 347–352
  numbering of, 349
Frequency Division Duplex, 3, 6, 101, 136–137, 281–282, 384
Frequency hopping, uplink, 207–209
Frequency-Switched Transmit Diversity, 164
FSTD *see* Frequency-Switched Transmit Diversity

G

GERAN, 10, 397–401
GGSN, 397
GMSK modulation, 399
GSM, 2–3, 7–8, 396–401
  migration from, 352–354
  multi-standard radio for, 359–361
Guard period, 139

H

Half-duplex FDD, 137–138
  in relation to scheduling, 281–282
HARQ *see* Hybrid ARQ
Heterogeneous network deployments, 293–299
Higher-order modulation, 19–21
High Interference Indicator (HII), 292
High Speed Packet Access (HSPA), 389–395
  in TDD, 395–396
HII *see* High Interference Indicator
Home-eNodeB, 293, 299
Home Subscriber Service (HSS), 110
HRPD, 402
HSPA *see* High Speed Packet Access
HSS *see* Home Subscriber Service
HS-SCCH-less operation, 394
Hybrid ARQ, 89–93
  adaptive, 250–251
  control signalling for, 177–179, 229–234
  in LTE, 99, 121–123, 247–259
  process, 248
  relaying, 335–337
  synchronous, 250–251
  with soft combining, 120–123

I

ICIC *see* Inter-cell interference coordination
ICS, 373–374
IEEE 802.16, 405–408
IMT-2000, 4–6
IMT-Advanced, 4–6, 11–13, 103, 408
Incremental redundancy (IR), 90–93, 249, 253
In-sequence delivery, 262
Inter-cell interference coordination (ICIC), 99, 290–299
  in heterogeneous deployments, 294–299
  in home-eNodeB, 299
Interference cancellation, 75

Interference coordination *see Inter-cell Interference Coordination*
Interference Rejection Combining (IRC), 62
Intermodulation:
  receiver, 374–375
  transmitter, 372
Internet Protocol (IP), 7
IR *see* Incremental Redundancy
ITU, 3–6
  see also IMT, IMT-Advanced

L

Latency, 378–379
Layer mapping, LTE, 166, 222
Link adaptation, 79–81, 113, 389
  see also Rate adaptation
Logical channels, 113, 116
LTE-Advanced, 103
  see also Release 10, IMT-Advanced

M

MAC, 115
  control element, 118
Machine-type Communication, 413–414
Master Information Block (MIB), 305–307
MBMS, 323–330
MBMS Control Channel (MCCH), 116, 325–329
MBMS Traffic Channel (MTCH);, 116, 325, 329
MBSFN, 102, 128, 131–132, 324–328
MBSFN area, 324
MBSFN subframe, 131–132, 326–327, 335
MCCH *see* MBMS Control Channel
MCH *see* Multicast Channel
MCH Scheduling Information (MSI), 329
MCH Scheduling Period (MSP), 328
MCH subframe allocation (MSA), 329
MIB *see* Master Information Block
MIMO *see* spatial multiplexing
MME *see* Mobility Management Entity
Mobility Management Entity (MME), 110
MSR *see* Multi-Standard Radio
MTCH *see* MBMS Traffic Channel
Multi-carrier transmission, 21–25
  for CDMA2000, 405
  for EDGE, 398
  for HSPA, 395
  for LTE *see carrier aggregation*
  for TD-SCDMA, 396
Multicast, 43–44
  see also MBMS
Multicast/Broadcast over Single–Frequency Network *see* MBSFN
Multicast Channel (MCH), 117, 325–328
Multi-cell/multicast Coordination Entity (MCE), 324
Multi-Standard Radio (MSR) base stations, 359–361

N

NAS *see* Non-Access Stratum
New-data indicator, 184, 191, 249–253
NodeB, 391
Non-Access Stratum (NAS), 110

O

OCC *see* Orthogonal Cover Codes
OFDM *see* Orthogonal Frequency Division Multiplex
OFDMA, 41
OI *see* Overload Indicator
Operating bands, 348–352
Operating band unwanted emissions, 363, 368–369

Orthogonal Cover Codes (OCC)
for downlink demodulation reference signals, 156
for CSI reference signals, 159
for uplink reference signals, 216
Orthogonal Frequency Division Multiplex (OFDM), 27–44
demodulation, 29
FFT/IFFT implementation, 30–32
in LTE, 127–131
Out-of-band
blocking, 374
emissions, 367–368
Overload Indicator (OI), 292

P

Packet Data Network Gateway (P-GW), 110
Packet Data Convergence Protocol (PDCP), 111
Paging, 319–321
Paging Channel (PCH), 116
Paging Control Channel (PCCH), 116
Partial path-loss compensation *see* path-loss compensation
Path-loss compensation, 270
PBCH *see* Physical Broadcast Channel
PCCH *see* Paging Control Channel
PCFICH *see* Physical Control Format Indicator Channel
PCH *see* Paging Channel
PCRF *see* Policy and Charging Rules Function
PDCCH *see* Physical Downlink Control Channel
PDCP *see* Packet Data Convergence Protocol
PDSCH *see* Physical Downlink Shared Channel
PDU *see* Protocol Data Unit
P-GW *see* Packet Data Network Gateway
PHICH *see* Physical Hybrid-ARQ Indicator Channel
Physical Broadcast Channel (PBCH), 123
Physical channel, 123
Physical Control Format Indicator Channel (PCFICH), 123, 174–177
Physical Downlink Control Channel (PDCCH), 123, 179–181, 195–202
Physical Downlink Shared Channel (PDSCH), 123, 143–152
Physical Hybrid-ARQ Indicator Channel, 123, 177–179
Physical-layer cell identity, 301
Physical Multicast Channel (PMCH), 123, 326–327
Physical Random Access Channel (PRACH), 124, 312
Physical Resource Block (PRB):
downlink, 129–130, 149–152
uplink, 129–130, 205–207
Physical Uplink Control Channel (PUCCH), 123, 227–242
transmit diversity for
Physical Uplink Shared Channel (PUSCH), 123, 203–210
PMCH *see* Physical Multicast Channel
PMI *see* Precoding-Matrix Indicator
Policy and Charging Rules Function (PCRF), 110
Positioning, 103
Power control, 79–81
for PRACH, 315
for PUCCH, 267–269
for PUSCH, 269–271
for SRS, 271–272
in LTE uplink, 265–272
Power Headroom, 271–272, 280
PRACH *see* Physical Random Access Channel

PRB *see* Physical Resource Block
Pre-coding, LTE, 161–173
  codebook-based, 165
  closed-loop, 167
  non-codebook-based, 169
  open-loop, 168
Precoding-Matrix Indicator (PMI), 167, 242, 283
Protocol Data Unit (PDU), 113
PUCCH *see* Physical Uplink Control Channel
PUSCH *see* Physical Uplink Shared Channel

Q

QoS, 110
QPP, 146
QPSK, 19–20

R

RACH *see* Random Access Channel
Radio access network, 95, 111
Radio bearer, 111
Radio frame *see* Frame
Radio Link Control (RLC), 113–115, 247, 257–264
Radio Network Controller (RNC), 391
Radio Resource Control (RRC), 124–125
Radio Resource Management (RRM), 111, 265
RAKE, 45
RAN *see* Radio access network
Random access, 310–319
  preamble, 312–317
Random Access Channel (RACH), 117
Rank indicator, 242, 283
Rate adaptation, 97–99, 115, 272
  see also link adaptation
Rate control, 106–7
  in LTE, 304
Rate matching, 147, 204
Receive diversity, 60, 100
Redundancy version, 92–93, 147, 184, 191, 204, 251, 253–254
Reference sensitivity:
  for LTE BS, 373
  for LTE terminal, 373
Reference signal:
  Cell-specific, 153–156
  CRS *see reference signal, cell specific*
  CSI *see reference signal, channel-state information*
  Demodulation (DM-RS, in downlink) *see reference signal, UE-specific*
  for channel-state information, 158–161
  for MBSFN, 153, 327–328
  for positioning, 153
  for sounding *see* sounding reference signal
  for uplink demodulation, 211–217
  power-boosted, 154UE-specific, 153
Reference-signal sequence:
  phase-rotated/"cyclically shifted", 213–214
  frequency-domain, 210–211
Reference symbol, 35–36, 153
  frequency shifted, 154
Relay, 331–345
  access link, 332
  amplify-and-forward *see repeater*
  architecture, 333–334
  backhaul link, 332
  control signalling *see* Relay Physical Downlink Control Channel
  decode-and-forward, 331
  hybrid-ARQ operation, 336–337
  inband, 332
  outband, 332
Relay Physical Downlink Control Channel (R-PDCCH), 337–342

Relative Narrowband Transmit Power (RNTP), 292
Release 8 (LTE), 95–102
Release 9 (LTE), 95, 102–103
Release 10 (LTE), 103–105
see also LTE-Advanced
Reordering, 262
Repeater, 331
Resource-allocation type, 185–186, 192–193
Resource block:
in downlink, 129–130, 149–153
in uplink, 129–130, 205–210
signaling of, 185–187
Resource-block pair, 130
Resource element, 129
Resource element group, 175
RI *see* Rank Indicator
RLC *see* Radio Link Control
Roaming, 2, 7, 11
RNC *see* Radio Network Controller
RNTP *see* Relative Narrowband Transmit Power
R-PDCCH *see* Relay Physical Downlink Control Channel
RRC *see* Radio Resource Control
RRC_IDLE, 125–126
RRC_CONNECTED, 125–126
RRM *see* Radio Resource Management

S

S1 interface, 111
SAE, 8, 95
Scheduling:
assignment, 179, 181–189
channel-dependent, 79, 81–89, 97–99
downlink, 81–85, 273–275
frequency-domain, 87, 97–98
grant, 180, 189–93
greedy filling, 85, 86
half-duplex FDD, 281–82
max-C/I, 82–85
proportional fair, 83–84
request, 229, 234, 238–239, 277–280
round robin, 83–84
semi-persistent, 119, 280–281
uplink, 85–87, 275–280
SDU *see* Service Data Unit
Search space, 199
Selectivity:
adjacent channel *see* ACS
in-channel *see* ICS
Self-backhauling *see Relay*
Served traffic, 383
Service Data Unit (SDU), 113
Serving cell, 279
Serving Gateway (S-GW), 110
SFBC *see* Space-Frequency (Block) Coding
SFN *see* Single–frequency network
SFN *see* System Frame Number
S-GW *see* Serving Gateway
Shared-channel transmission, 97, 119
SI *see* System Information message
SIB *see* System Information Block
SIC *see* Successive Interference cancellation
Single-frequency network (SFN), 44
see also MBSFN
SI-RNTI *see* System Information RNTI
Slot, 128–30
Soft handover, 391
Sounding reference signal, 217–221
Space-Frequency Block Coding (SFBC), 68, 163–165, 306
Space Time (Block) Coding (ST(B)C), 66–68

Spatial multiplexing, 71–79, 100, 105
multi-codeword based, 75–77
pre-coder based, 74–75
single-codeword based, 75–77
in LTE *see* Pre-coding (LTE)
Special subframe, 138–40
Spectral efficiency, 337
performance, 380–387
requirement, 380, 384
Spectrum flexibility, 100–102, 354
Spurious emissions, 363, 367–369, 372
Standardization, 8–13
Subcarrier spacing:
in general, 38
LTE, 127
reduced, 128
Subframe, 128
Successive Interference Cancellation (SIC)
see Interference cancellation
Symbol quadruplet, 175
Synchronization signal, 302–304
System Frame Number (SFN), 128, 305
System Information, 305
System Information Block (SIB), 305, 308–310
System Information message (SI), 308–310
System Information RNTI (SI-RNTI), 308
System performance, 377–379
System throughput, 379, 383

T

TDD *see* Time Division Duplex
TD-SCDMA, 140, 395–396
Time Division Duplex, 3, 6, 101, 138–141, 385
Timing advance, 245–246
for relaying, 343–344
Transmission mode, 162
Transmission rank, 166
Transmission Time Interval (TTI), 116
Transmit Diversity, 67–69, 163–165
for PUCCH, 226
Transparent mode (for RLC), 261
Transport block, 116
signaling of size, 187–189
Transport channels, 116–118
Transport format, 116
Transport format selection, 116, 120, 273
TTI *see* Transmission Time Interval

U

UCI *see* Uplink Control Information
UE categories, 105–106
UL-SCH *see* Uplink Shared Channel
UMB, 403
Unacknowledged mode (for RLC), 261
Unwanted emissions, 362, 367–372
Uplink Control Information (UCI), 123
Uplink Shared Channel (UL-SCH), 117
UpPTS, 138–140
User Equipment (UE), 106
categories, 106–107
User throughput, 378
UTRA FDD *see* WCDMA
UTRA TDD, 11, 395–396

V

VAMOS, 401
Virtual Resource Block (VRB), 150
distributed vs. localized, 150–151
downlink, 150
uplink, 208
VRB *see* Virtual Resource Block

W
WCDMA, 3, 389–395
White space, 414
WiMAX, 405–408

X
X2 Interface, 111

Z
Zadoff-Chu sequences, 212

TECHNOLOGY/TELECOMMUNICATIONS

4G

LTE / LTE-Advanced for Mobile Broadband

Based on the very successful book *3G Evolution – HSPA and LTE for Mobile Broadband*, but now focusing purely on LTE and LTE-Advanced (Release 10), this book gives an integrated treatment of the underlying technologies of LTE and the LTE standard, explaining both the 'why' and the 'how'. The book gives an understanding of each element on a higher level as well as a detailed insight to the specific solutions, thereby providing an understanding of how to implement and deploy LTE in mobile broadband products and systems.

- Includes full details of the latest addition to the LTE radio-access standards and technologies up to and including 3GPP Release 10
- Clear explanations of the role of the underlying technologies for LTE, including OFDM and MIMO
- Full coverage of LTE-Advanced as submitted to the ITU-R, including LTE carrier aggregation, extended multi-antenna transmission, relaying functionality and heterogeneous networks
- LTE radio-interface architecture, physical layer, access procedures, broadcasting, RF characteristics and system performance are covered in detail

Erik Dahlman, Stefan Parkvall and **Johan Sköld** all work at Ericsson Research and have been deeply involved in 3G and 4G development and standardization since the early days of 3G research. They are leading experts in the field and are today still actively contributing to the standardization of LTE within 3GPP.





ACADEMIC PRESS
An imprint of Elsevier
elsevierdirect.com

