WILEY

LTE for UMTS
OFDMA and SC-FDMA Based Radio Access

Harri Holma and Antti Toskala



Exhibit 1008
IPR2024-00310
U.S. Patent 10,932,207

LTE for UMTS –
OFDMA and SC-FDMA Based Radio Access

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6

LTE for UMTS –
OFDMA and SC-FDMA Based Radio Access

Edited by

Harri Holma and Antti Toskala

both of Nokia Siemens Networks, Finland



John Wiley & Sons, Ltd

This edition first published 2009
© 2009 John Wiley & Sons Ltd.

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

The right of the author to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

LTE is a trademark, registered by ETSI for the benefit of the 3GPP Partners

Library of Congress Cataloging-in-Publication Data

LTE for UMTS-OFDMA and SC-FDMA based radio access / edited by Harri Holma, Antti Toskala.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-99401-6 (cloth : alk. paper) 1. Universal Mobile Telecommunications System. 2. Wireless communication systems--Standards. 3. Mobile communication systems--Standards. 4. Global system for mobile communications. I. Holma, Harri, 1970- II. Toskala, Antti.
TK5103.4883.L78 2009
621.3845'6--dc22

2008052792

A catalogue record for this book is available from the British Library.

ISBN 9780470994016 (H/B)

Set in 10/12 pt Times by Sparks, Oxford – www.sparkspublishing.com
Printed and bound in Great Britain by Antony Rowe, Chippenham, UK

Contents

Preface **xiii**

Acknowledgements **xv**

List of Abbreviations **xvii**

1 Introduction **1**
Harri Holma and Antti Toskala
1.1 Mobile Voice Subscriber Growth 1
1.2 Mobile Data Usage Growth 2
1.3 Wireline Technologies Evolution 3
1.4 Motivation and Targets for LTE 4
1.5 Overview of LTE 5
1.6 3GPP Family of Technologies 7
1.7 Wireless Spectrum 8
1.8 New Spectrum Identified by WRC-07 10
1.9 LTE-Advanced 11

2 LTE Standardization **13**
Antti Toskala
2.1 Introduction 13
2.2 Overview of 3GPP Releases and Process 13
2.3 LTE Targets 14
2.4 LTE Standardization Phases 16
2.5 Evolution Beyond Release 8 18
2.6 LTE-Advanced for IMT-Advanced 19
2.7 LTE Specifications and 3GPP Structure 21
References 22

3 System Architecture Based on 3GPP SAE **23**
Atte Länsisalmi and Antti Toskala
3.1 System Architecture Evolution in 3GPP 23
3.2 Basic System Architecture Configuration with only E-UTRAN Access Network 25

3.2.1 Overview of Basic System Architecture Configuration 25
3.2.2 Logical Elements in Basic System Architecture Configuration 26
3.2.3 Self-configuration of S1-MME and X2 interfaces 34
3.2.4 Interfaces and Protocols in Basic System Architecture Configuration 35
3.2.5 Roaming in Basic System Architecture Configuration 39
3.3 System Architecture with E-UTRAN and Legacy 3GPP Access Networks 40
3.3.1 Overview of 3GPP Inter-working System Architecture Configuration 40
3.3.2 Additional and Updated Logical Elements in 3GPP Inter-working System Architecture Configuration 42
3.3.3 Interfaces and Protocols in 3GPP Inter-working System Architecture Configuration 44
3.3.4 Inter-working with Legacy 3GPP CS Infrastructure 44
3.4 System Architecture with E-UTRAN and Non-3GPP Access Networks 45
3.4.1 Overview of 3GPP and Non-3GPP Inter-working System Architecture Configuration 45
3.4.2 Additional and Updated Logical Elements in 3GPP Inter-working System Architecture Configuration 47
3.4.3 Interfaces and Protocols in Non-3GPP Inter-working System Architecture Configuration 50
3.4.4 Roaming in Non-3GPP Inter-working System Architecture Configuration 51
3.5 Inter-working with cdma2000® Access Networks 51
3.5.1 Architecture for cdma2000® HRPD Inter-working 51
3.5.2 Additional and Updated Logical Elements for cdma2000® HRPD Inter-working 54
3.5.3 Protocols and Interfaces in cdma2000® HRPD Inter-working 55
3.5.4 Inter-working with cdma2000® 1xRTT 56
3.6 IMS Architecture 56
3.6.1 Overview 56
3.6.2 Session Management and Routing 58
3.6.3 Databases 59
3.6.4 Services Elements 59
3.6.5 Inter-working Elements 59
3.7 PCC and QoS 60
3.7.1 PCC 60
3.7.2 QoS 63
References 65

4 Introduction to OFDMA and SC-FDMA and to MIMO in LTE 67
Antti Toskala and Timo Lunttila
4.1 Introduction 67
4.2 LTE Multiple Access Background 67
4.3 OFDMA Basics 70
4.4 SC-FDMA Basics 76
4.5 MIMO Basics 80
4.6 Summary 82
References 82

5 Physical Layer **83**
Antti Toskala, Timo Lunttila, Esa Tiirola, Kari Hooli and Juha Korhonen
5.1 Introduction 83
5.2 Transport Channels and Their Mapping to the Physical Channels 83
5.3 Modulation 85
5.4 Uplink User Data Transmission 86
5.5 Downlink User Data Transmission 89
5.6 Uplink Physical Layer Signaling Transmission 93
5.6.1 Physical Uplink Control Channel (PUCCH) 94
5.6.2 PUCCH Configuration 97
5.6.3 Control Signaling on PUSCH 101
5.6.4 Uplink Reference Signals 103
5.7 PRACH Structure 109
5.7.1 Physical Random Access Channel 109
5.7.2 Preamble Sequence 110
5.8 Downlink Physical Layer Signaling Transmission 112
5.8.1 Physical Control Format Indicator Channel (PCFICH) 112
5.8.2 Physical Downlink Control Channel (PDCCH) 113
5.8.3 Physical HARQ Indicator Channel (PHICH) 115
5.8.4 Downlink Transmission Modes 115
5.8.5 Physical Broadcast Channel (PBCH) 116
5.8.6 Synchronization Signal 117
5.9 Physical Layer Procedures 117
5.9.1 HARQ Procedure 118
5.9.2 Timing Advance 119
5.9.3 Power Control 119
5.9.4 Paging 120
5.9.5 Random Access Procedure 120
5.9.6 Channel Feedback Reporting Procedure 123
5.9.7 Multiple Input Multiple Output (MIMO) Antenna Technology 129
5.9.8 Cell Search Procedure 130
5.9.9 Half Duplex Operation 130
5.10 UE Capability Classes and Supported Features 131
5.11 Physical Layer Measurements 132
5.11.1 eNodeB Measurements 132
5.11.2 UE Measurements and Measurement Procedure 133
5.12 Physical Layer Parameter Configuration 133
5.13 Summary 134
References 135

6 LTE Radio Protocols **137**
Antti Toskala and Woonhee Hwang
6.1 Introduction 137
6.2 Protocol Architecture 137
6.3 Medium Access Control 139
6.3.1 Logical Channels 140
6.3.2 Data Flow in MAC Layer 142

6.4 Radio Link Control Layer 143
6.4.1 RLC Modes of Operation 144
6.4.2 Data Flow in RLC Layer 145
6.5 Packet Data Convergence Protocol 145
6.6 Radio Resource Control (RRC) 146
6.6.1 UE States and State Transitions Including Inter-RAT 147
6.6.2 RRC Functions and Signaling Procedures 148
6.7 X2 Interface Protocols 158
6.7.1 Handover on X2 Interface 159
6.7.2 Load Management 160
6.8 Early UE Handling in LTE 162
6.9 Summary 162
References 163

7 Mobility 165
Chris Callender, Harri Holma, Jarkko Koskela and Jussi Reunanen
7.1 Introduction 165
7.2 Mobility Management in Idle State 166
7.2.1 Overview of Idle Mode Mobility 166
7.2.2 Cell Selection and Reselection Process 167
7.2.3 Tracking Area Optimization 169
7.3 Intra-LTE Handovers 170
7.3.1 Procedure 170
7.3.2 Signaling 171
7.3.3 Handover Measurements 174
7.3.4 Automatic Neighbor Relations 174
7.3.5 Handover Frequency 175
7.3.6 Handover Delay 177
7.4 Inter-system Handovers 177
7.5 Differences in E-UTRAN and UTRAN Mobility 178
7.6 Summary 179
References 180

8 Radio Resource Management 181
Harri Holma, Troels Kolding, Daniela Laselva, Klaus Pedersen, Claudio Rosa and Ingo Viering
8.1 Introduction 181
8.2 Overview of RRM Algorithms 181
8.3 Admission Control and QoS Parameters 182
8.4 Downlink Dynamic Scheduling and Link Adaptation 184
8.4.1 Layer 2 Scheduling and Link Adaptation Framework 184
8.4.2 Frequency Domain Packet Scheduling 185
8.4.3 Combined Time and Frequency Domain Scheduling Algorithms 187
8.4.4 Packet Scheduling with MIMO 188
8.4.5 Downlink Packet Scheduling Illustrations 189
8.5 Uplink Dynamic Scheduling and Link Adaptation 192
8.5.1 Signaling to Support Uplink Link Adaptation and Packet Scheduling 196

8.5.2 Uplink Link Adaptation 199
8.5.3 Uplink Packet Scheduling 200
8.6 Interference Management and Power Settings 204
8.6.1 Downlink Transmit Power Settings 205
8.6.2 Uplink Interference Coordination 206
8.7 Discontinuous Transmission and Reception (DTX/DRX) 207
8.8 RRC Connection Maintenance 209
8.9 Summary 209
References 210

9 Performance 213
Harri Holma, Pasi Kinnunen, István Z. Kovács, Kari Pajukoski, Klaus Pedersen and Jussi Reunanen
9.1 Introduction 213
9.2 Layer 1 Peak Bit Rates 213
9.3 Terminal Categories 216
9.4 Link Level Performance 217
9.4.1 Downlink Link Performance 217
9.4.2 Uplink Link Performance 219
9.5 Link Budgets 222
9.6 Spectral Efficiency 224
9.6.1 System Deployment Scenarios 224
9.6.2 Downlink System Performance 228
9.6.3 Uplink System Performance 231
9.6.4 Multi-antenna MIMO Evolution Beyond 2×2 234
9.6.5 Higher Order Sectorization (Six Sectors) 238
9.6.6 Spectral Efficiency as a Function of LTE Bandwidth 240
9.6.7 Spectral Efficiency Evaluation in 3GPP 242
9.6.8 Benchmarking LTE to HSPA 243
9.7 Latency 244
9.7.1 User Plane Latency 244
9.8 LTE Refarming to GSM Spectrum 246
9.9 Dimensioning 247
9.10 Capacity Management Examples from HSPA Networks 249
9.10.1 Data Volume Analysis 250
9.10.2 Cell Performance Analysis 252
9.11 Summary 256
References 257

10 Voice over IP (VoIP) 259
Harri Holma, Juha Kallio, Markku Kuusela, Petteri Lundén, Esa Malkamäki, Jussi Ojala and Haiming Wang
10.1 Introduction 259
10.2 VoIP Codecs 259
10.3 VoIP Requirements 261
10.4 Delay Budget 262
10.5 Scheduling and Control Channels 263

10.6 LTE Voice Capacity 265
10.7 Voice Capacity Evolution 271
10.8 Uplink Coverage 273
10.9 Circuit Switched Fallback for LTE 275
10.10 Single Radio Voice Call Continuity (SR-VCC) 277
10.11 Summary 280
References 281

11 Performance Requirements 283
Andrea Ancora, Iwajlo Angelow, Dominique Brunel, Chris Callender, Harri Holma, Peter Muszynski, Earl McCune and Laurent Noël
11.1 Introduction 283
11.2 Frequency Bands and Channel Arrangements 283
11.2.1 Frequency Bands 283
11.2.2 Channel Bandwidth 285
11.2.3 Channel Arrangements 287
11.3 eNodeB RF Transmitter 288
11.3.1 Operating Band Unwanted Emissions 288
11.3.2 Coexistence with Other Systems on Adjacent Carriers Within the Same Operating Band 290
11.3.3 Coexistence with Other Systems in Adjacent Operating Bands 292
11.3.4 Transmitted Signal Quality 295
11.4 eNodeB RF Receiver 300
11.4.1 Reference Sensitivity Level 300
11.4.2 Dynamic Range 301
11.4.3 In-channel Selectivity 301
11.4.4 Adjacent Channel Selectivity (ACS) and Narrow-band Blocking 303
11.4.5 Blocking 304
11.4.6 Receiver Spurious Emissions 306
11.4.7 Receiver Intermodulation 306
11.5 eNodeB Demodulation Performance 307
11.5.1 PUSCH 307
11.5.2 PUCCH 309
11.5.3 PRACH 310
11.6 UE Design Principles and Challenges 311
11.6.1 Introduction 311
11.6.2 RF Subsystem Design Challenges 311
11.6.3 RF–Baseband Interface Design Challenges 318
11.6.4 LTE vs HSDPA Baseband Design Complexity 324
11.7 UE RF Transmitter 327
11.7.1 LTE UE Transmitter Requirement 327
11.7.2 LTE Transmit Modulation Accuracy, EVM 328
11.7.3 Desensitization for Band and Bandwidth Combinations (Desense) 329
11.7.4 Transmitter Architecture 329
11.8 UE RF Receiver Requirements 331
11.8.1 Reference Sensitivity Level 331
11.8.2 Introduction to UE Self-desensitization Contributors in FDD UEs 336

11.8.3 ACS, Narrowband Blockers and ADC Design Challenges 341
11.8.4 EVM Contributors: A Comparison Between LTE and WCDMA Receivers 348
11.9 UE Demodulation Performance 352
11.9.1 Transmission Modes 352
11.9.2 Channel Modeling and Estimation 354
11.9.3 Demodulation Performance 356
11.10 Requirements for Radio Resource Management 358
11.10.1 Idle State Mobility 360
11.10.2 Connected State Mobility when DRX is Not Active 360
11.10.3 Connected State Mobility when DRX is Active 362
11.10.4 Handover Execution Performance Requirements 363
11.11 Summary 364
References 364

12 LTE TDD Mode 367
Che Xiangguang, Troels Kolding, Peter Skov, Wang Haiming and Antti Toskala
12.1 Introduction 367
12.2 LTE TDD Fundamentals 368
12.2.1 LTE TDD Frame Structure 369
12.2.2 Asymmetric Uplink/Downlink Capacity Allocation 371
12.2.3 Co-existence with TD-SCDMA 371
12.2.4 Channel Reciprocity 372
12.2.5 Multiple Access Schemes 373
12.3 TDD Control Design 374
12.3.1 Common Control Channels 374
12.3.2 Sounding Reference Signal 376
12.3.3 HARQ Process and Timing 376
12.3.4 HARQ Design for UL TTI Bundling 379
12.3.5 UL HARQ-ACK/NACK Transmission 380
12.3.6 DL HARQ-ACK/NACK Transmission 380
12.3.7 DL HARQ-ACK/NACK Transmission with SRI and/or CQI over PUCCH 381
12.4 Semi-persistent Scheduling 381
12.5 MIMO and Dedicated Reference Signals 383
12.6 LTE TDD Performance 385
12.6.1 Link Performance 386
12.6.2 Link Budget and Coverage for TDD System 386
12.6.3 System Level Performance 389
12.6.4 Evolution of LTE TDD 396
12.7 Summary 396
References 397

13 HSPA Evolution 399
Harri Holma, Karri Ranta-aho and Antti Toskala
13.1 Introduction 399
13.2 Discontinuous Transmission and Reception (DTX/DRX) 400

13.3 Circuit Switched Voice on HSPA 401
13.4 Enhanced FACH and RACH 404
13.5 Downlink MIMO and 64QAM 405
13.6 Dual Carrier HSDPA 407
13.7 Uplink 16QAM 409
13.8 Layer 2 Optimization 410
13.9 Single Frequency Network (SFN) MBMS 411
13.10 Architecture Evolution 412
13.11 Summary 414
References 415

Index **417**

Preface

The number of mobile subscribers has increased tremendously in recent years. Voice communication has become mobile in a massive way and the mobile is the preferred way for voice communication. At the same time the data usage has grown fast in those networks where 3GPP High Speed Packet Access (HSPA) was introduced indicating that the users find value in broadband wireless data. The average data consumption exceeds hundreds of Megabytes per subscriber per month. The end users expect data performance similar to the fixed lines. The operators request high data capacity with low cost of data delivery. 3GPP Long Term Evolution (LTE) is designed to meet those targets. This book presents 3GPP LTE standard in Release 8 and describes its expected performance.

The book is structured as follows. Chapter 1 presents an introduction. The standardization background and process is described in Chapter 2. The system architecture evolution (SAE) is presented in Chapter 3, and the basics of air interface modulation choices in Chapter 4. Chapter 5 describes 3GPP LTE physical layer solutions, and Chapter 6 protocol solutions. The mobility

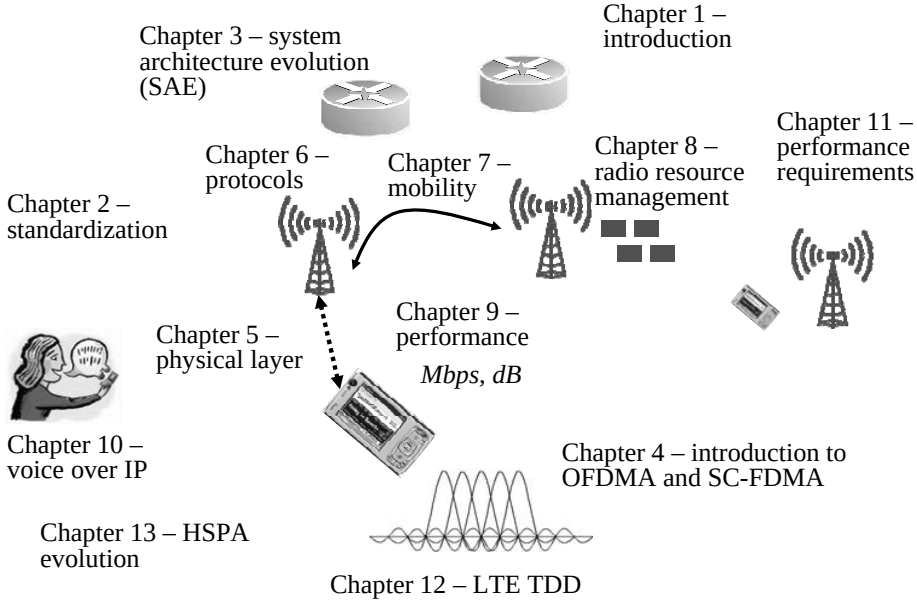


Figure 0.1 Contents of the book

aspects are addressed in Chapter 7, and the radio resource management in Chapter 8. The radio and end-to-end performance is illustrated in Chapter 9. The voice performance is presented in Chapter 10. Chapter 11 explains the 3GPP performance requirements. Chapter 12 presents the main LTE Time Division Duplex (TDD). Chapter 13 describes HSPA evolution in 3GPP Releases 7 and 8.

LTE can access a very large global market – not only GSM/UMTS operators, but also CDMA operators and potentially also fixed network service providers. The potential market can attract a large number of companies to the market place pushing the economies of scale which enable wide scale LTE adoption with lower cost. This book is particularly designed for chip set and mobile vendors, network vendors, network operators, application developers, technology managers and regulators who would like to get a deeper understanding of LTE technology and its capabilities.

Acknowledgements

The editors would like to acknowledge the hard work of the contributors from Nokia Siemens Networks, Nokia, ST-Ericsson and Nomor Research: Andrea Ancora, Iwajlo Angelow, Dominique Brunel, Chris Callender, Kari Hooli, Woonhee Hwang, Juha Kallio, Matti Kiiski, Pasi Kinnunen, Troels Kolding, Juha Korhonen, Jarkko Koskela, Istvan Kovacs, Markku Kuusela, Daniela Laselva, Earl McCune, Peter Muszynski, Petteri Lunden, Timo Lunttila, Atte Länsisalmi, Esa Malkamäki, Laurent Noel, Jussi Ojala, Kari Pajukoski, Klaus Pedersen, Karri Ranta-aho, Jussi Reunanen, Haiming Wang, Peter Skov, Esa Tiirola, Ingo Viering, Haiming Wang and Che Xiangguang.

We also would like to thank the following colleagues for their valuable comments: Asbjörn Grovlen, Jorma Kaikkonen, Michael Koonert, Peter Merz, Preben Mogensen, Sari Nielsen, Gunnar Nitsche, Miikka Poikselkä, Sabine Rössel, Benoist Sebire, Issam Toufik and Helen Waite.

The editors appreciate the fast and smooth editing process provided by Wiley and especially Sarah Tilley, Mark Hammond, Katharine Unwin, Brett Wells, Tom Fryer and Mitch Fitton.

We are grateful to our families, as well as the families of all the authors, for their patience during the late night and weekend editing sessions.

The editors and authors welcome any comments and suggestions for improvements or changes that could be implemented in forthcoming editions of this book. The feedback is welcome to editors' email addresses harri.holma@nsn.com and antti.toskala@nsn.com.

List of Abbreviations

3GPP	Third Generation Partnership Project
AAA	Authentication, Authorization and Accounting
ACF	Analog Channel Filter
ACIR	Adjacent Channel Interference Rejection
ACK	Acknowledgement
ACLR	Adjacent Channel Leakage Ratio
ACS	Adjacent Channel Selectivity
ADC	Analog-to Digital Conversion
ADSL	Asymmetric Digital Subscriber Line
AKA	Authentication and Key Agreement
AM	Acknowledged Mode
AMBR	Aggregate Maximum Bit Rate
AMD	Acknowledged Mode Data
AMR	Adaptive Multi-Rate
AMR-NB	Adaptive Multi-Rate Narrowband
AMR-WB	Adaptive Multi-Rate Wideband
ARP	Allocation Retention Priority
ASN	Abstract Syntax Notation
ASN.1	Abstract Syntax Notation One
ATM	Adaptive Transmission Bandwidth
AWGN	Additive White Gaussian Noise
AWGN	Additive White Gaussian Noise
BB	Baseband
BCCH	Broadcast Control Channel
BCH	Broadcast Channel
BE	Best Effort
BEM	Block Edge Mask
BICC	Bearer Independent Call Control Protocol
BiCMOS	Bipolar CMOS
BLER	Block Error Rate
BO	Backoff
BOM	Bill of Material
BPF	Band Pass Filter
BPSK	Binary Phase Shift Keying

BS Base Station
BSC Base Station Controller
BSR Buffer Status Report
BT Bluetooth
BTS Base Station
BW Bandwidth
CAZAC Constant Amplitude Zero Autocorrelation Codes
CBR Constant Bit Rate
CCE Control Channel Element
CCCH Common Control Channel
CDD Cyclic Delay Diversity
CDF Cumulative Density Function
CDM Code Division Multiplexing
CDMA Code Division Multiple Access
CIR Carrier to Interference Ratio
CLM Closed Loop Mode
CM Cubic Metric
CMOS Complementary Metal Oxide Semiconductor
CoMP Coordinated Multiple Point
CP Cyclic Prefix
CPE Common Phase Error
CPICH Common Pilot Channel
CQI Channel Quality Information
CRC Cyclic Redundancy Check
C-RNTI Cell Radio Network Temporary Identifier
CS Circuit Switched
CSCF Call Session Control Function
CSFB Circuit Switched Fallback
CSI Channel State Information
CT Core and Terminals
CTL Control
CW Continuous Wave
DAC Digital to Analog Conversion
DARP Downlink Advanced Receiver Performance
D-BCH Dynamic Broadcast Channel
DC Direct Current
DCCH Dedicated Control Channel
DCH Dedicated Channel
DC-HSDPA Dual Cell (Dual Carrier) HSDPA
DCI Downlink Control Information
DCR Direct Conversion Receiver
DCXO Digitally-Compensated Crystal Oscillator
DD Duplex Distance
DFCA Dynamic Frequency and Channel Allocation
DFT Discrete Fourier Transform
DG Duplex Gap
DL Downlink

DL-SCH Downlink Shared Channel
DPCCH Dedicated Physical Control Channel
DR Dynamic Range
DRX Discontinuous Reception
DSP Digital Signal Processing
DTCH Dedicated Traffic Channel
DTM Dual Transfer Mode
DTX Discontinuous Transmission
DVB-H Digital Video Broadcast – Handheld
DwPTS Downlink Pilot Time Slot
E-DCH Enhanced DCH
EDGE Enhanced Data Rates for GSM Evolution
EFL Effective Frequency Load
EFR Enhanced Full Rate
EGPRS Enhanced GPRS
E-HRDP Evolved HRPD (High Rate Packet Data) network
EIRP Equivalent Isotropic Radiated Power
EMI Electromagnetic Interference
EPA Extended Pedestrian A
EPC Evolved Packet Core
EPDG Evolved Packet Data Gateway
ETU Extended Typical Urban
E-UTRA Evolved Universal Terrestrial Radio Access
EVA Extended Vehicular A
EVDO Evolution Data Only
EVM Error Vector Magnitude
EVS Error Vector Spectrum
FACH Forward Access Channel
FCC Federal Communications Commission
FD Frequency Domain
FDD Frequency Division Duplex
FDE Frequency Domain Equalizer
FDM Frequency Division Multiplexing
FDPS Frequency Domain Packet Scheduling
FE Front End
FFT Fast Fourier Transform
FM Frequency Modulated
FNS Frequency Non-Selective
FR Full Rate
FRC Fixed Reference Channel
FS Frequency Selective
GB Gigabyte
GBF Guaranteed Bit Rate
GDD Group Delay Distortion
GERAN GSM/EDGE Radio Access Network
GF G-Factor
GGSN Gateway GPRS Support Node

GMSK Gaussian Minimum Shift Keying
GP Guard Period
GPON Gigabit Passive Optical Network
GPRS General packet radio service
GPS Global Positioning System
GRE Generic Routing Encapsulation
GSM Global System for Mobile Communications
GTP GPRS Tunneling Protocol
GTP-C GPRS Tunneling Protocol, Control Plane
GUTI Globally Unique Temporary Identity
GW Gateway
HARQ Hybrid Adaptive Repeat and Request
HB High Band
HD-FDD Half Duplex Frequency Division Duplex
HFN Hyper Frame Number
HII High Interference Indicator
HO Handover
HPBW Half Power Beam Width
HPF High Pass Filter
HPSK Hybrid Phase Shift Keying
HRPD High Rate Packet Data
HSDPA High Speed Downlink Packet Access
HS-DSCH High Speed Downlink Shared Channel
HSGW HRPD Serving Gateway
HSPA High Speed Packet Access
HS-PDSCH High Speed Physical Downlink Shared Channel
HSS Home Subscriber Server
HS-SCCH High Speed Shared Control Channel
HSUPA High Speed Uplink Packet Access
IC Integrated Circuit
IC Interference Cancellation
ICI Inter-carrier Interference
ICIC Inter-cell Interference Control
ICS IMS Centralized Service
ID Identity
IETF Internet Engineering Task Force
IFFT Inverse Fast Fourier Transform
IL Insertion Loss
iLBC Internet Lob Bit Rate Codec
IM Implementation Margin
IMD Intermodulation
IMS IP Multimedia Subsystem
IMT International Mobile Telecommunications
IoT Interference over Thermal
IOT Inter-Operability Testing
IP Internet Protocol
IR Image Rejection

IRC Interference Rejection Combining
ISD Inter-site Distance
ISDN Integrated Services Digital Network
ISI Inter-system Interference
ISTO Industry Standards and Technology Organization
ISUP ISDN User Part
IWF Interworking Funtion
LAI Location Area Identity
LMA Local Mobility Anchor
LB Low Band
LCID Logical Channel Identification
LCS Location Services
LMMSE Linear Mininum Mean Square Error
LNA Low Noise Amplifier
LO Local Oscillator
LOS Line of Sight
LTE Long Term Evolution
MAC Medium Access Control
MAP Maximum a Posteriori
MAP Mobile Application Part
MBMS Multimedia Broadcast Multicast System
MBR Maximum Bit Rate
MCH Multicast Channel
MCL Minimum Coupling Loss
MCS Modulation and Coding Scheme
MGW Media Gateway
MIB Master Information Block
MIMO Multiple Input Multiple Output
MIP Mobile IP
MIPI Mobile Industry Processor Interface
MIPS Million Instructions Per Second
MM Mobility Management
MME Mobility Management Entity
MMSE Minimum Mean Square Error
MPR Maximum Power Reduction
MRC Maximal Ratio Combining
MSC Mobile Switching Center
MSC-S Mobile Switching Center Server
MSD Maximum Sensitivity Degradation
MU Multiuser
NACC Network Assisted Cell Change
NACK Negative Acknowledgement
NAS Non-access Stratum
NAT Network Address Table
NB Narrowband
NF Noise Figure
NMO Network Mode of Operation

NRT	Non-real Time
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
OI	Overload Indicator
OLLA	Outer Loop Link Adaptation
OOB	Out of Band
OOBN	Out-of-Band Noise
O&M	Operation and Maintenance
PA	Power Amplifier
PAPR	Peak to Average Power Ratio
PAR	Peak-to-Average Ratio
PBR	Prioritized Bit Rate
PC	Personal Computer
PC	Power Control
PCC	Policy and Charging Control
PCCC	Parallel Concatenated Convolution Coding
PCCPCH	Primary Common Control Physical Channel
PCFICH	Physical Control Format Indicator Channel
PCH	Paging Channel
PCI	Physical Cell Identity
PCM	Pulse Code Modulation
PCRF	Policy and Charging Resource Function
PCS	Personal Communication Services
PDCCH	Physical Downlink Control Channel
PDCP	Packet Data Convergence Protocol
PDF	Probability Density Function
PDN	Packet Data Network
PDU	Payload Data Unit
PDSCH	Physical Downlink Shared Channel
PF	Proportional Fair
P-GW	Packet Data Network Gateway
PHICH	Physical HARQ Indicator Channel
PHR	Power Headroom Report
PHS	Personal Handyphone System
PHY	Physical Layer
PLL	Phase Locked Loop
PLMN	Public Land Mobile Network
PMI	Precoding Matrix Index
PMIP	Proxy Mobile IP
PN	Phase Noise
PRACH	Physical Random Access Channel
PRB	Physical Resource Block
PS	Packet Switched
PSD	Power Spectral Density
PSS	Primary Synchronization Signal
PUCCH	Physical Uplink Control Channel
PUSCH	Physical Uplink Shared Channel

QAM Quadrature Amplitude Modulation
QCI QoS Class Identifier
QD Quasi Dynamic
QN Quantization Noise
QoS Quality of Service
QPSK Quadrature Phase Shift Keying
RACH Random Access Channel
RAD Required Activity Detection
RAN Radio Access Network
RAR Random Access Response
RAT Radio Access Technology
RB Resource Block
RBG Radio Bearer Group
RF Radio Frequency
RI Rank Indicator
RLC Radio Link Control
RNC Radio Network Controller
RNTP Relative Narrowband Transmit Power
ROHC Robust Header Compression
RR Round Robin
RRC Radio Resource Control
RRM Radio Resource Management
RS Reference Signal
RSCP Received Symbol Code Power
RSRP Reference Symbol Received Power
RSRQ Reference Symbol Received Quality
RSSI Received Signal Strength Indicator
RT Real Time
RTT Round Trip Time
RV Redundancy Version
SA Services and System Aspects
SAE System Architecture Evolution
SAIC Single Antenna Interference Cancellation
S-CCPCH Secondary Common Control Physical Channel
SC-FDMA Single Carrier Frequency Division Multiple Access
SCH Synchronization Channel
SCM Spatial Channel Model
SCTP Stream Control Transmission Protocol
SDQNR Signal to Distortion Quantization Noise Ratio
SDU Service Data Unit
SE Spectral Efficiency
SEM Spectrum Emission Mask
SF Spreading Factor
SFBC Space Frequency Block Coding
SFN System Frame Number
SGSN Serving GPRS Support Node
S-GW Serving Gateway

SIB System Information Block
SID Silence Indicator Frame
SIM Subscriber Identity Module
SIMO Single Input Multiple Output
SINR Signal to Interference and Noise Ratio
SMS Short Message Service
SNR Signal to Noise Ratio
SON Self Optimized Networks
SON Self Organizing Networks
SR Scheduling Request
S-RACH Short Random Access Channel
SRB Signaling Radio Bearer
S-RNC Serving RNC
SRS Sounding Reference Signals
SSS Secondary Synchronization Signal
SR-VCC Single Radio Voice Call Continuity
S-TMSI S-Temporary Mobile Subscriber Identity
SU-MIMO Single User Multiple Input Multiple Output
S1AP S1 Application Protocol
TA Tracking Area
TBS Transport Block Size
TD Time Domain
TDD Time Division Duplex
TD-LTE Time Division Long Term Evolution
TD-SCDMA Time Division Synchronous Code Division Multiple Access
TM Transparent Mode
TPC Transmit Power Control
TRX Transceiver
TSG Technical Specification Group
TTI Transmission Time Interval
TU Typical Urban
UDP Unit Data Protocol
UE User Equipment
UHF Ultra High Frequency
UICC Universal Integrated Circuit Card
UL Uplink
UL-SCH Uplink Shared Channel
UM Unacknowledged Mode
UMD Unacknowledged Mode Data
UMTS Universal Mobile Telecommunications System
UpPTS Uplink Pilot Time Slot
USB Universal Serial Bus
USIM Universal Subscriber Identity Module
USSD Unstructured Supplementary Service Data
UTRA Universal Terrestrial Radio Access
UTRAN Universal Terrestrial Radio Access Network
VCC Voice Call Continuity

VCO	Voltage Controlled Oscillator
VDSL	Very High Data Rate Subscriber Line
VLR	Visitor Location Register
V-MIMO	Virtual MIMO
VoIP	Voice over IP
WCDMA	Wideband Code Division Multiple Access
WG	Working Group
WLAN	Wireless Local Area Network
WRC	World Radio Conference
X1AP	X1 Application Protocol
ZF	Zero Forcing

1

Introduction

Harri Holma and Antti Toskala

1.1 Mobile Voice Subscriber Growth

The number of mobile subscribers has increased tremendously during the last decade: the first billion landmark was exceeded in 2002, the second billion in 2005, the third billion in 2007 and the fourth billion by the end of 2008. More than 1 million new subscribers per day have been added globally, that is more than ten subscribers on average every second. This growth is illustrated in Figure 1.1. Mobile phone penetration worldwide is approaching

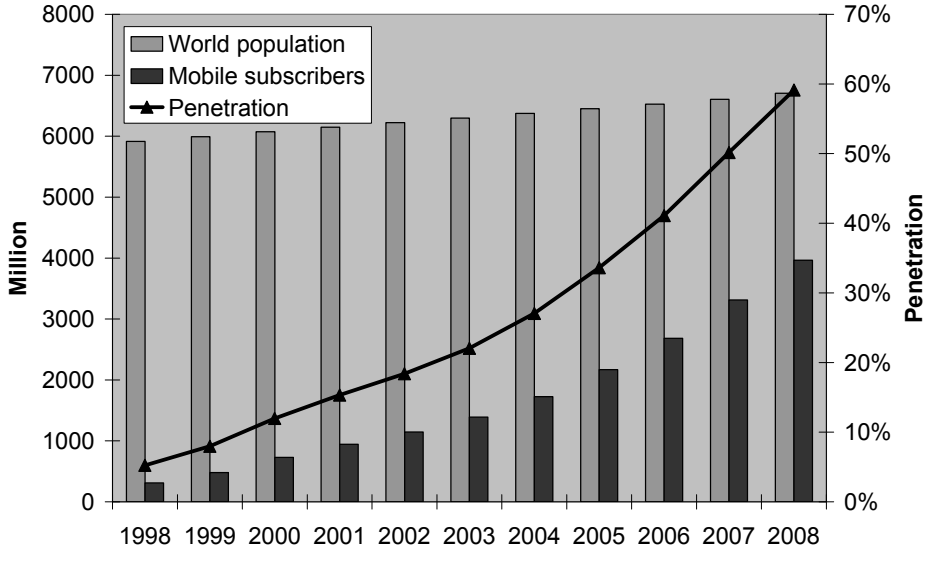


Figure 1.1 Growth of mobile subscribers

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6

60%[1]. Voice communication has become mobile in a massive way. The mobile is the preferred method for voice communication, with mobile networks covering over 90% of the world's population. This voice growth has been fuelled by low cost mobile phones and efficient network coverage and capacity, which is enabled by standardized solutions and by an open ecosystem leading to the economies of scale. Mobile voice is not the privilege of the rich but also brings value for users on low incomes – because of the benefits of being connected, low income users spend a larger part of their income on mobile communications.

1.2 Mobile Data Usage Growth

The second generation mobile networks – like Global System for Mobile Communications (GSM) – were originally designed for carrying voice traffic while the data capability was added later. Data usage has increased but the traffic volume in second generation networks is clearly dominated by voice traffic. The introduction of third generation networks with High Speed Downlink Packet Access (HSDPA) has boosted data usage considerably. Example operator statistics for 12 months are shown in Figure 1.2 where the HSDPA downlink data volumes are several terabytes per day, which correspond to beyond 1 Gbps busy hour network level throughput. Such fast data growth shows that the end users find value in the wireless broadband access.

Data traffic volume has in many cases already exceeded voice traffic volume when voice traffic is converted into terabytes by assuming a voice data rate of 12 kbps. A typical case is illustrated in Figure 1.3. HSDPA data growth is advanced by high speed radio capability, flat rate pricing schemes and simple device installation. In short, the introduction of HSDPA has changed mobile networks from voice dominated to packet data dominated networks.

Data usage is advanced by a number of bandwidth hungry laptop applications including internet and intranet access, file sharing, streaming services to distribute video content and mobile

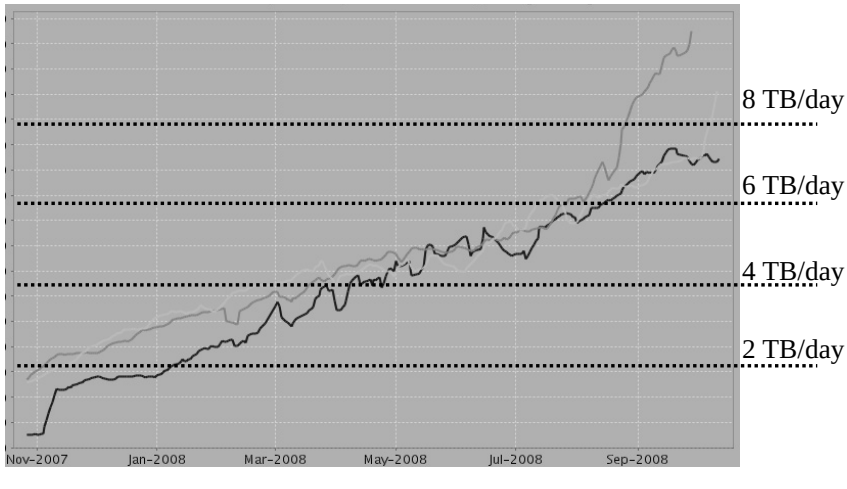


Figure 1.2 Growth of HSDPA data traffic

1 The actual user penetration can be different since some users have multiple subscriptions and some subscriptions are shared by multiple users.

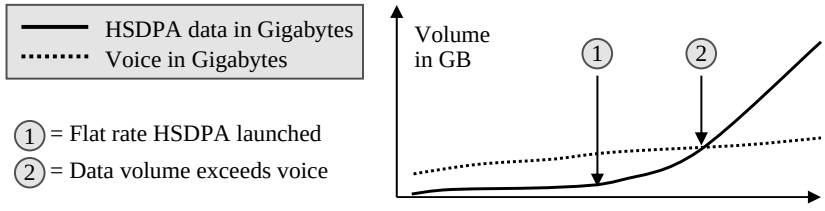


Figure 1.3 HSDPA data volume exceeds voice volume

TV and interactive gaming. In addition, service bundles of video, data and voice – known also as triple play – are entering the mobile market, also replacing the traditional fixed line voice and broadband data services with mobile services both at home and in the office.

A typical voice subscriber uses 300 minutes per month, which is equal to approximately 30 megabyte of data with a voice data rate of 12.2 kbps. A broadband data user can easily consume more than 1000 megabyte (1 gigabyte) of data. Heavy broadband data usage takes 10–100× more capacity than voice usage, which sets high requirements for the capacity and efficiency of network data.

It is expected that by 2015, 5 billion people will be connected to the internet. Broadband internet connections will be available practically anywhere in the world. Already today, the existing wireline installations can reach approximately 1 billion households and the mobile networks connect over 3 billion subscribers. These installations need to evolve into broadband internet access. Further extensive use of wireless access as well as new wireline installations with enhanced capabilities are required to offer true broadband connectivity to the 5 billion customers.

1.3 Wireline Technologies Evolution

Although wide area wireless networks have experienced a fast evolution of data rates, wireline networks still provide the highest data rates. The evolution of the peak user data rate both in wireless and wireline networks is illustrated in Figure 1.4. Interestingly, the shape of the evolution curve is similar in both domains with a relative difference of approximately 30 times. An application of Moore's law predicts that data rates double every 18 months. Currently, copper based wireline solutions with Very High Data Rate Digital Subscriber Line (VDSL2) can offer bit rates of tens of Mbps and the passive optical fibre based solution gives rates in excess of 100 Mbps. Both copper and fibre based solutions will have further data rate evolution in the near future, increasing the data rate offerings to the Gbps range.

Wireless networks must make data rates higher in order to match the user experience provided by wireline networks. When customers are used to wireline performance, they expect the wireless network to offer comparable performance. The applications designed for wireline networks advance the evolution of the wireless data rates.

Wireless technologies, on the other hand, have the huge benefit of being capable of offering personal broadband access independently of user location – in other words, mobility, for nomadic or full mobile use cases. A wireless solution can also provide low cost broadband coverage compared to new wireline installations if there is no existing wireline infrastructure. Therefore, wireless broadband access is an attractive option, especially in new growth markets in urban areas as well as in rural areas in other markets.

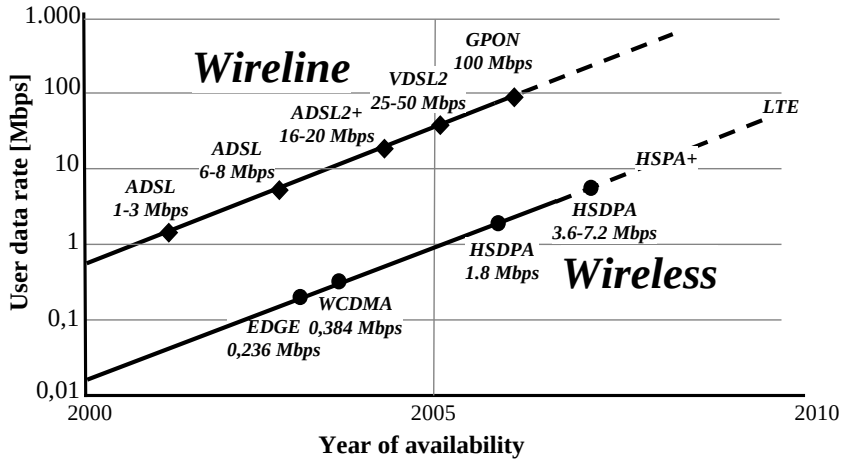


Figure 1.4 Evolution of wireless and wireline user data rates [Broadband Access for All - A Brief Technology Guide, Nokia Siemens Networks white paper (2007)]. GPON = Gigabit Passive Optical Network; VDSL = Very High Data Rate Subscriber Line; ADSL = Asymmetric Digital Subscriber Line

1.4 Motivation and Targets for LTE

The work towards 3rd Generation Partnership Project (3GPP) Long Term Evolution (LTE) started in 2004 with the definition of the targets. Even though HSDPA was not yet deployed at that time, it became evident that work for the next radio system should be started. It takes more than 5 years from setting the system targets to commercial deployment using interoperable standards. Therefore, system standardization must be started early enough to be ready by the time the need is there. A few driving forces can be identified advancing LTE development: wireline capability evolution, the need for additional wireless capacity, the need for lower cost wireless data delivery and the competition of other wireless technologies. As wireline technology keeps improving, a similar evolution is required in the wireless domain to make sure that the applications also work fluently in the wireless domain. There are also other wireless technologies – including IEEE 802.16 – which promise high data capabilities. 3GPP technologies must match and exceed the competition. More capacity is a clear requirement for taking maximum advantage of the available spectrum and base station sites. These reasons are summarized in Figure 1.5.

LTE must be able to deliver superior performance compared to existing 3GPP networks based on High Speed Packet Access (HSPA) technology. The performance targets in 3GPP are defined relative to HSPA in Release 6. The peak user throughput should be minimum 100 Mbps in downlink and 50 Mbps in uplink, which is ten times more than HSPA Release 6. Also the latency must be reduced in order to improve the end user performance. The terminal power consumption must be minimized to enable more usage of the multimedia applications without recharging the battery. The main performance targets are shown in Figure 1.6 and are listed below:

- spectral efficiency two to four times more than with HSPA Release 6;
- peak rates exceed 100 Mbps in downlink and 50 Mbps in uplink;

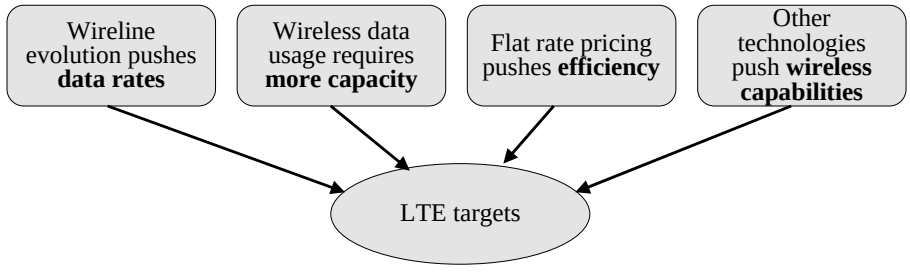


Figure 1.5 Driving forces for LTE development

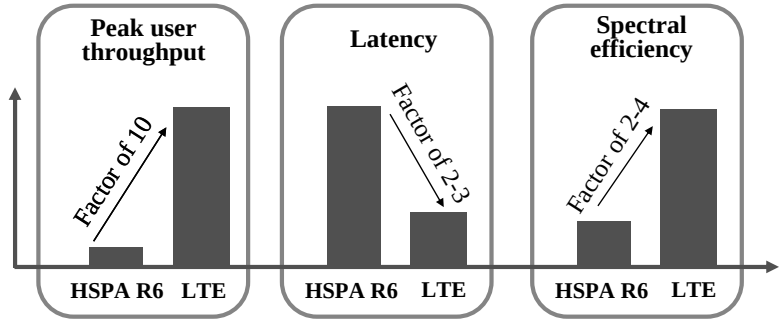


Figure 1.6 Main LTE performance targets

- enables round trip time <10 ms;
- packet switched optimized;
- high level of mobility and security;
- optimized terminal power efficiency;
- frequency flexibility with from below 1.5 MHz up to 20 MHz allocations.

1.5 Overview of LTE

The multiple access scheme in LTE downlink uses Orthogonal Frequency Division Multiple Access (OFDMA) and uplink uses Single Carrier Frequency Division Multiple Access (SC-FDMA). These multiple access solutions provide orthogonality between the users, reducing the interference and improving the network capacity. The resource allocation in the frequency domain takes place with a resolution of 180 kHz resource blocks both in uplink and in downlink. The frequency dimension in the packet scheduling is one reason for the high LTE capacity. The uplink user specific allocation is continuous to enable single carrier transmission while the downlink can use resource blocks freely from different parts of the spectrum. The uplink single carrier solution is also designed to allow efficient terminal power amplifier design, which is relevant for the terminal battery life. The LTE solution enables spectrum flexibility where the transmission bandwidth can be selected between 1.4 MHz and 20 MHz depending on the available spectrum. The 20 MHz bandwidth can provide up to 150 Mbps downlink user data rate with 2×2 MIMO, and 300 Mbps with 4×4 MIMO. The uplink peak data rate is 75 Mbps. The multiple access schemes are illustrated in Figure 1.7.

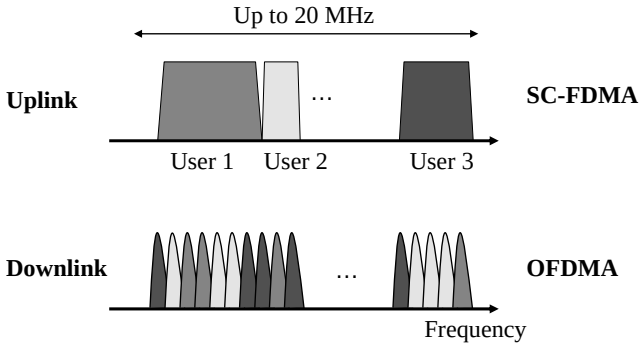


Figure 1.7 LTE multiple access schemes

The high network capacity also requires an efficient network architecture in addition to the advanced radio features. The target in 3GPP Release 8 is to improve the network scalability for traffic increase and to minimize the end-to-end latency by reducing the number of network elements. All radio protocols, mobility management, header compression and all packet retransmissions are located in the base stations called eNodeB. eNodeB includes all those algorithms that are located in Radio Network Controller (RNC) in 3GPP Release 6 architecture. Also the core network is streamlined by separating the user and the control planes. The Mobility Management Entity (MME) is just the control plane element while the user plane bypasses MME directly to System Architecture Evolution (SAE) Gateway (GW). The architecture evolution is illustrated in Figure 1.8. This Release 8 core network is also often referred to as Evolved Packet Core (EPC) while for the whole system the term Evolved Packet System (EPS) can also be used.

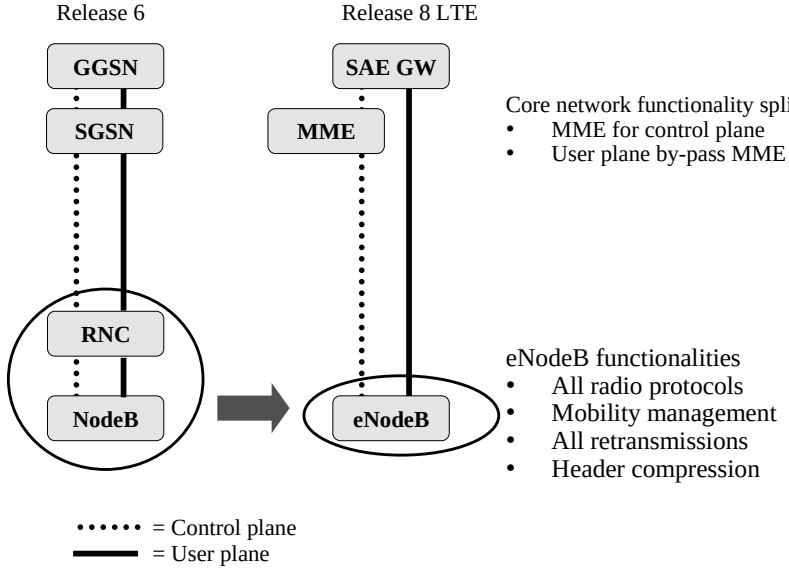


Figure 1.8 LTE network architecture

1.6 3GPP Family of Technologies

3GPP technologies – GSM/EDGE and Wideband Code Division Multiple Access (WCDMA)/ HSPA – are currently serving nearly 90% of the global mobile subscribers. The market share development of 3GPP technologies is illustrated in Figure 1.9. A number of major Code Division Multiple Access (CDMA) operators have already turned or are soon turning to GSM/WCDMA for voice evolution and to HSPA/LTE for data evolution to get access to the benefits of the large and open 3GPP ecosystem and economics of scale for low cost mobile devices. The number of subscribers using 3GPP based technologies is currently more than 3.5 billion. The 3GPP LTE will be built on this large base of 3GPP technologies.

The time schedule of 3GPP specifications and the commercial deployments is illustrated in Figure 1.10. Enhanced Data rates for GSM Evolution (EDGE) was defined in 3GPP in 1997 and WCDMA at the end of 1999. Both systems had their first commercial deployments during

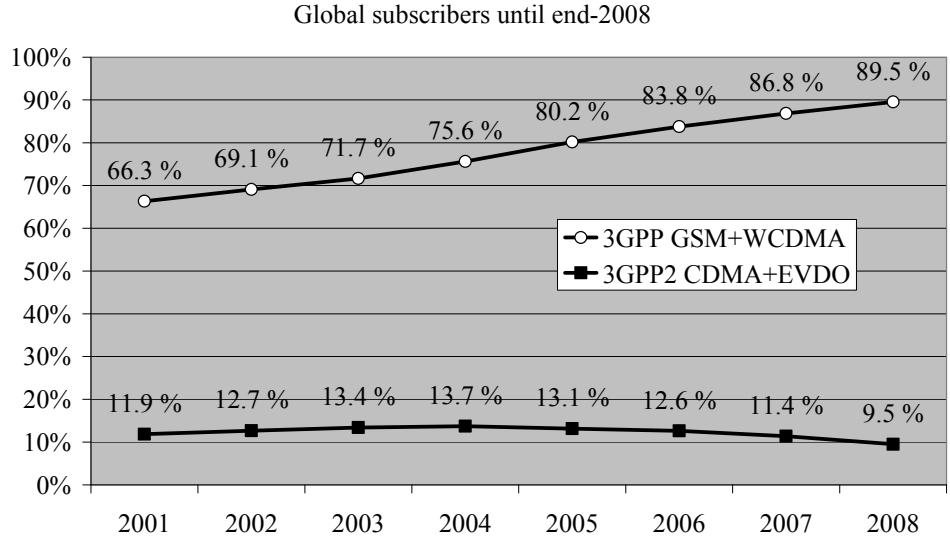


Figure 1.9 Global market share of 3GPP and 3GPP2 technologies. EVDO, evolution data only

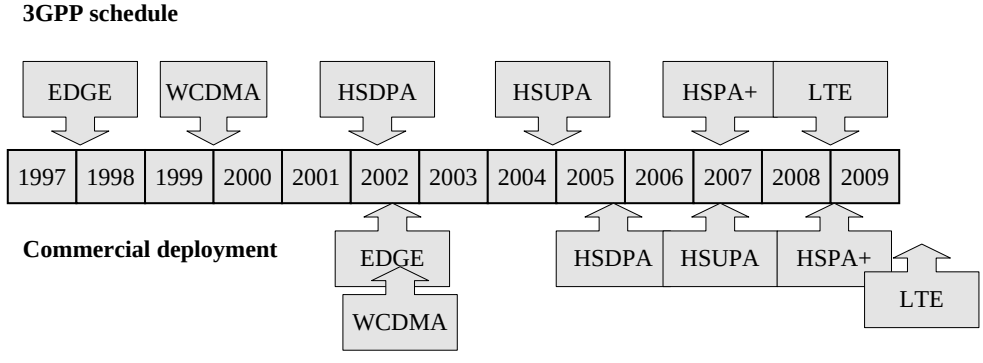


Figure 1.10 Schedule of 3GPP standard and their commercial deployments

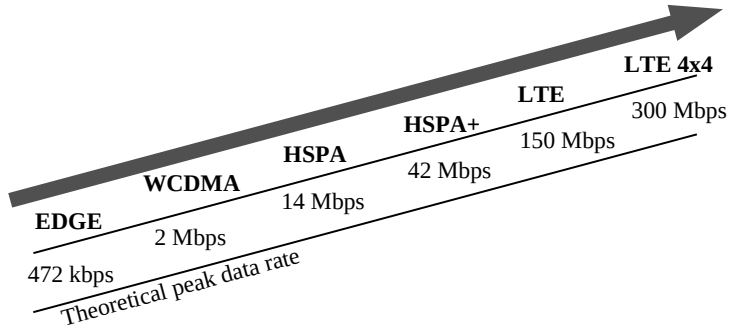


Figure 1.11 Peak data rate evolution of 3GPP technologies

2002. The HSDPA and High Speed Uplink Packet Access (HSUPA) standards were completed in March 2002 and December 2004, and the commercial deployments followed in 2005 and 2007. The first phase of HSPA evolution, also called HSPA+, was completed in June 2007 and the deployments start during 2009. The LTE standard was approved at the end of 2007, the backwards compatibility is expected to start in March 2009 and commercial deployments are expected in 2010.

The new generation of technologies pushes data rates higher. The evolution of the peak user data rates is illustrated in Figure 1.11. The first WCDMA deployments in 2002 offered 384 kbps, current HSDPA networks 7.2–14.4 Mbps, HSPA evolution 21–42 Mbps and LTE 2010 150 Mbps, that is a more than 300 times higher data rate over 8 years.

The 3GPP technologies are designed for smooth inter-working and coexistence. The LTE will support bi-directional handovers between LTE and GSM and between LTE and UMTS. GSM, UMTS and LTE can share a number of network elements including core network elements. It is also expected that some of the 3G network elements can be upgraded to support LTE and there will be single network platforms supporting both HSPA and LTE. The subscriber management and Subscriber Identity Module (SIM) based authentication will be used also in LTE; however, in LTE the system access requires the more modern and more secure Universal SIM (USIM) instead of the older 2G originated SIM card.

1.7 Wireless Spectrum

The LTE frequency bands in 3GPP specifications are shown in Figure 1.12 for paired bands and in Figure 1.13 for unpaired bands. There are 17 paired bands and 8 unpaired bands defined currently and more bands will be added during the standardization process. Some bands are currently used by other technologies and LTE can coexist with the legacy technologies. Similarly, in Europe and in Asia, WCDMA was initially deployed in the new 2100 MHz band while the refarming to the existing 900 MHz started during 2007. LTE will likely start by using the new 2600 MHz band and refarming to 900 and 1800 MHz bands. In the best case in Europe there is in total a 565 MHz spectrum available for the mobile operators when including 900 MHz, 1800 MHz, 2100 MHz Frequency Division Duplex (FDD) and Time Division Duplex (TDD) bands and the new 2600 MHz allocation all together.

Operating band	3GPP name	Total spectrum	Uplink [MHz]	Downlink [MHz]
Band 1	2100	2x60 MHz	1920-1980	2110-2170
Band 2	1900	2x60 MHz	1850-1910	1930-1990
Band 3	1800	2x75 MHz	1710-1785	1805-1880
Band 4	1700/2100	2x45 MHz	1710-1755	2110-2155
Band 5	850	2x25 MHz	824-849	869-894
Band 6	800	2x10 MHz	830-840	875-885
Band 7	2600	2x70 MHz	2500-2570	2620-2690
Band 8	900	2x35 MHz	880-915	925-960
Band 9	1700	2x35 MHz	1750-1785	1845-1880
Band 10	1700/2100	2x60 MHz	1710-1770	2110-2170
Band 11	1500	2x25 MHz	1427.9-1452.9	1475.9-1500.9
Band 12	US700	2x18 MHz	698-716	728-746
Band 13	US700	2x10 MHz	777-787	746-756
Band 14	US700	2x10 MHz	788-798	758-768
Band 17	US700	2x10 MHz	704-716	734-746
Band 18	Japan800	2x30 MHz	815-830	860-875
Band 19	Japan800	2x30 MHz	830-845	875-890

Figure 1.12 Frequency bands for paired bands in 3GPP specifications

Operating band	3GPP name	Total spectrum	Uplink and downlink [MHz]
Band 33	UMTS TDD1	1x20 MHz	1900-1920
Band 34	UMTS TDD2	1x15 MHz	2010-2025
Band 35	US1900 UL	1x60 MHz	1850-1910
Band 36	US1900 DL	1x60 MHz	1930-1990
Band 37	US1900	1x20 MHz	1910-1930
Band 38	2600	1x50 MHz	2570-2620
Band 39	UMTS TDD	1x40 MHz	1880-1920
Band 40	2300	1x50 MHz	2300-2400

Figure 1.13 Frequency bands for unpaired bands in 3GPP specifications

In the USA the WCDMA networks have been refarmed to 850 and 1900 MHz. The new frequencies at 1700/2100 are also used for 3G deployment. LTE will be deployed using 700 and 1700/2100 bands, and later refarmed to the existing bands.

In Japan the LTE deployments start using the 2100 band followed later by 800, 1500 and 1700 bands.

Flexible bandwidth is desirable to take advantage of the diverse spectrum assets: refarming typically requires a narrowband option below 5 MHz, while the new spectrum allocations could

take benefit of a wideband option up to 20 MHz and higher data rates. It is also evident that both FDD and TDD modes are required to take full benefit of the available paired and unpaired spectrum. These requirements are taken into account in the LTE system specification.

1.8 New Spectrum Identified by WRC-07

The ITU-R World Radiocommunication Conference (WRC-07) worked in October and November 2007 to identify the new spectrum for International Mobile Telecommunications (IMT). The following bands were identified for IMT and are illustrated in Figure 1.14. The target was to identify both low bands for coverage and high bands for capacity.

The main additional coverage band will be in UHF frequencies 470–806/862 MHz that are currently used for terrestrial TV broadcasting. The sub-band 790–862 MHz was identified in Europe and Asia-Pacific. The availability of the band depends on the national time schedules of the analogue to digital TV switchover and it can become widely available within the 2012 to 2015 timeframe. The band allows, for example, three operators each running 10 MHz LTE FDD.

The sub-band 698–806 MHz was identified for IMT in the Americas. In the USA, part of the band has already been auctioned.

The main capacity band will be in 3.4–4.2 GHz (C-band). Total 200 MHz in the sub-band 3.4–3.8 GHz was identified for IMT in Europe and in Asia-Pacific. This spectrum can facilitate the deployment of a larger bandwidth of IMT-Advanced to provide the highest bit rates and capacities.

Additionally, the band 2.3–2.4 GHz was identified for IMT, but this band is not expected to be available in Europe or in the Americas. This band was already identified for IMT-2000 in China at the WRC-2000. The sub-band 450–470 MHz was identified for IMT globally, but it is not expected to be widely available in Europe. This spectrum will be narrow with a maximum 2×5 MHz deployment.

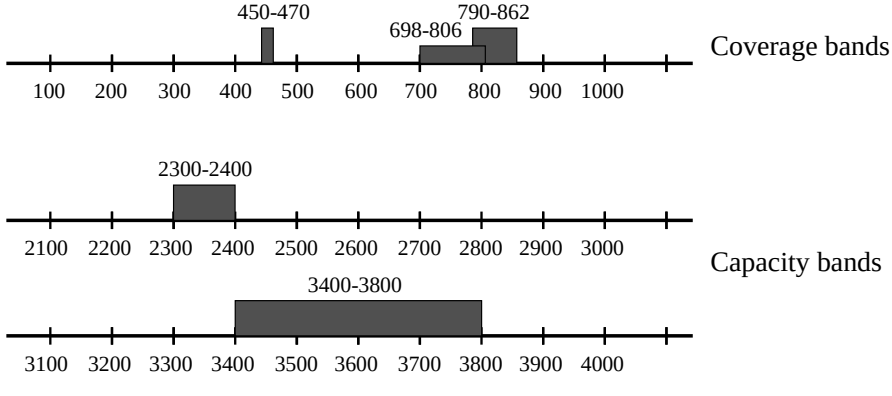


Figure 1.14 Main new frequencies identified for IMT in WRC-07

1.9 LTE-Advanced

International Mobile Telecommunications-Advanced (IMT-Advanced) is a concept for mobile systems with capabilities beyond IMT-2000. IMT-Advanced was previously known as Systems beyond IMT-2000. During 2009, there will be an open call for candidates for IMT-Advanced to be submitted to ITU, as well as the start of assessment activities of candidate technologies and systems. The radio interface submission deadline is expected by October 2009 and the final specifications by 2011.

The new capabilities of these IMT-Advanced systems are envisaged to handle a wide range of supported data rates according to economic and service demands in multi-user environments with target peak data rates of up to approximately 100 Mbps for high mobility and up to 1 Gbps for low mobility such as nomadic/local wireless access. 3GPP has started to work towards IMT-Advanced targets also for the local area radio under the name LTE-Advanced. LTE-Advanced is planned to be part of 3GPP Release 10 and the commercial deployment of IMT-Advanced will be 2013 or later. The high level evolution of 3GPP technologies to meet IMT requirements is shown in Figure 1.15.

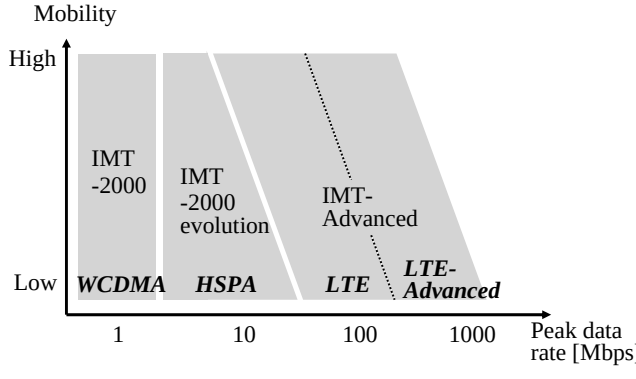


Figure 1.15 Bit rate and mobility evolution to IMT-Advanced

2

LTE Standardization

Antti Toskala

2.1 Introduction

Long Term Evolution (LTE) standardization is being carried out in the 3rd Generation Partnership Project (3GPP), as was also the case for Wideband CDMA (WCDMA), and the later phase of GSM evolution. This chapter introduces first the 3GPP LTE release schedule and the 3GPP standardization process. The requirements set for LTE by the 3GPP community are then reviewed, and the steps foreseen for later LTE Releases, including the LTE-Advanced work for the IMT-Advanced process, are covered. This chapter concludes with the introduction of LTE specifications and 3GPP structure.

2.2 Overview of 3GPP Releases and Process

The 3GPP has a background of 10 years for WCDMA development (or Universal Terrestrial Radio Access, UTRA) since the start of 3GPP in 1998. The major 3GPP releases are shown in Figure 2.1 starting from the first WCDMA release, Release 99, and covering the releases that followed. In Figure 2.1 the releases are shown with the date when the release content was finalized, not

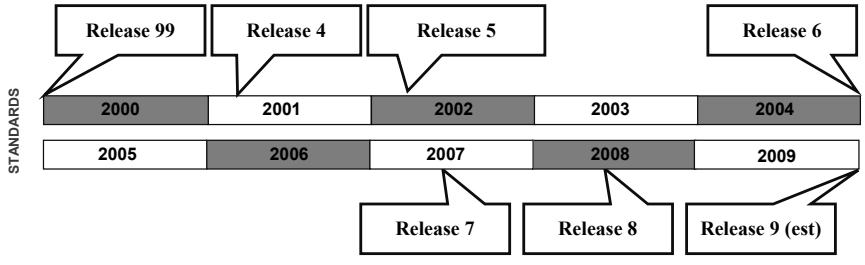


Figure 2.1 3GPP releases schedule with estimated Release 9 closure

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6

the actual protocol freezing date (backwards compatibility start). The first WCDMA release – Release 99 – was published in December 1999 and contained the basic WCDMA features with theoretical data rates up to 2 Mbps, based on the different multiple access for Frequency Division Duplex (FDD) mode and Time Division Duplex (TDD) operation. After that, 3GPP abandoned the yearly release principle and thus release naming was also changed as from Release 4, completed in March 2001. Release 4 did not have many major WCDMA features, but contained the new low chip rate TDD version (TD-SCDMA) for the TDD mode of UTRA. Release 5 followed with High Speed Downlink Packet Access (HSDPA) in March 2002 and Release 6 with High Speed Uplink Packet Access (HSUPA) in December 2004 for WCDMA. Release 7 was completed in June 2007 with the introduction of several HSDPA and HSUPA enhancements. Now 3GPP has just finalized Release 8 (with a few issues pending, for March 2009), which brought along further HSDPA/HSUPA improvements (often referred to jointly as High Speed Packet Access (HSPA) evolution) as well as containing the first LTE Release. The feature content for Release 8 was completed in December 2008. A more detailed description of the WCDMA/HSPA release content can be found in Chapter 13 covering Release 8 and in [1] for the earlier releases.

The earlier 3GPP Releases have a relationship to LTE in Release 8. Several of the novel features adopted – especially with HSDPA and HSUPA – are also used in LTE, such as base station based scheduling with physical layer feedback, physical layer retransmissions and link adaptation. Also, LTE specifications reuse the WCDMA design in areas where it could be carried out without compromising performance, thus facilitating reuse of the design and platforms developed for WCDMA. The first LTE release, Release 8, supports data rates up to 300 Mbps in downlink and up to 75 Mbps in uplink with low latency and flat radio architecture. Release 8 also facilitates radio level inter-working with GSM, WCDMA and cdma2000®.

3GPP is introducing new work items and study items for Release 9, some of them related to features postponed from Release 8 and some of them for new topics raised for Release 9 with the topics introduced in December 2009. Release 9 is scheduled to be completed around the end of 2009. Release 10 is then foreseen to contain further radio capability enhancement in the form of LTE-Advanced, intended to be submitted to ITU-R IMT-Advanced process with data rate capabilities foreseen to range up to 1 Gbps. First Release 10 specifications are expected to be ready at the end of 2010.

The 3GPP process is such that more topics are started than eventually end up in the specifications. Often a study is initially carried out for more complicated issues, as was the case with LTE. Typically during a study, more alternatives are looked at than the small set of features that eventually enter a specification. Sometimes a study is completed with the finding that there is not enough gain to justify the added complexity in the system. A change requested in the work item phase could also be rejected for this same reason. The 3GPP process starting from a study item is shown in Figure 2.2.

2.3 LTE Targets

At the start of the work during the first half 2005, the 3GPP defined the requirements for LTE development. The key elements included in the target setting for the LTE feasibility study work, as defined in [2], were as follows:

- The LTE system should be packet switched domain optimized. This means that circuit switched elements are not really considered, but everything is assumed to be based on a

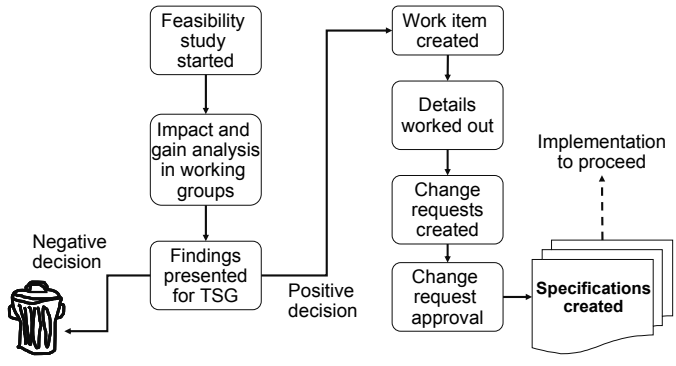


Figure 2.2. 3GPP process for moving from study towards work item and specification creation

packet type of operation. The system was required to support IP Multimedia Sub-system (IMS) and further evolved 3GPP packet core.

- As the data rates increase, latency needs to come down for the data rates to show any practical improvement. Thus the requirement for the LTE radio round trip time was set to be below 10 ms and access delay below 300 ms.
- The requirements for the data rates were defined to ensure sufficient steps in terms of data rates in contrast to HSPA. The peak rate requirements for uplink and downlink were set at 50 Mbps and 100 Mbps respectively.
- As the 3GPP community was used to a good level of security and mobility with the earlier systems – starting from GSM – it was also a natural requirement that these should be sustained. This also included inter-system mobility with GSM and WCMA, as well as cdma2000®, since there was (and is) a major interest in the cdma2000® community to evolve to LTE for next generation networks.
- With WCDMA, one of the topics that had caused challenges – especially in the beginning – was terminal power consumption, thus it was required to improve terminal power efficiency.
- In the 3GPP technology family there were both a narrowband system (GSM with 200 kHz) and a wideband system (WCDMA with 5 MHz). Thus it was now required that the new system facilitate frequency allocation flexibility with 1.25/2.5, 5, 10, 15 and 20 MHz allocations. Later during the course of work, the actual bandwidth values were slightly adjusted for the two smallest bandwidths (to use 1.4 and 3 MHz bandwidths) to give a good match for both GSM and cdma2000® refarming cases. The possibility of using LTE in a deployment with WCDMA or GSM as the system on the adjacent band was also required.
- The 'standard' requirement for any new system is also to have higher capacity. The benchmark level chosen was 3GPP Release 6, which had a stable specification and a known performance at the time. Thus Release 6 was a stable comparison level for running the LTE performance simulations during the feasibility study phase. Depending on the case, 2- to 4-times higher capacity than provided with the Release 6 HSDPA/HSUPA reference case, was required.
- One of the drivers for the work was cost, to ensure that the new system could facilitate lower investment and operating costs compared to the earlier system. This was a natural result of the flat rate charging model appearing at the time for data use and created pressure for the price vs data volume level.

It was also expected that further development of WCDMA would continue in parallel with the LTE activity, and this has been also carried out with the Release 8 HSPA improvements, as covered in Chapter 13.

2.4 LTE Standardization Phases

LTE work was started as a study in the 3GPP, with the first workshop held in November 2004 in Canada. In the workshop the first presentations were on both the expected requirements for the work and the expected technologies to be adopted. Contributions were made from both operator and vendor viewpoints.

Following the workshop, 3GPP TSG RAN approved the start of the study for LTE in December 2004, with the work first running in the RAN plenary level to define the requirements and then moving to the working groups for detailed technical discussions for multiple access, protocol solutions and architecture. The first key issues to be resolved were what the requirements are, as discussed in section 2.3, and these were mainly settled during the first half of 2005, visible in [2], with the first approved version in June 2005. Then the work focused on solving two key questions:

- What should be the LTE radio technology in terms of multiple access?
- What should be the system architecture?

The multiple access discussion was soon concluded with the finding that something new was needed instead of just extending WCDMA. This conclusion was the result of the large range of requirements for covering different bandwidths and data rates with reasonable complexity. The use of Orthogonal Frequency Division Multiple Access (OFDMA) in the downlink was obvious early on, and had already been reflected in many of the presentations in the original LTE workshop in 2004. For the uplink multiple access, the Single Carrier Frequency Division Multiple Access (SC-FDMA) soon emerged as the most favourable choice that was supported by many key vendors and operators, as could be seen, for example, in [3]. A noticeable improvement from the WCDMA was that both FDD and TDD modes had the same multiple access solution, as addressed for the FDD and TDD differences in Chapter 12. The OFDMA and SC-FDMA principles and motivational aspects are further covered in Chapter 4. The multiple access decision was officially endorsed at the end of 2005 and after that the LTE radio work focused on the chosen technologies. The LTE milestones are shown in Figure 2.3. The FDD/TDD alignment refers to the agreement on

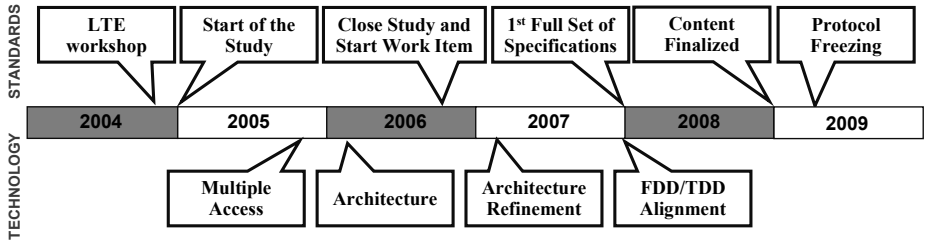


Figure 2.3 LTE milestones in 3GPP

the adjustment of the frame structure to minimize the differences between FDD and TDD modes of operation.

In the area of LTE architecture, after some debate it was decided to aim for a single node RAN, resulting in all radio related functionality being placed in the base station. This time the term used in 3GPP became eNodeB, with 'e' standing for evolved. The original architecture split, as shown in Figure 2.4, was endorsed in March 2006 with a slight adjustment carried out in early 2007 (with the Packet Data Convergence Protocol (PDCP) shifted from core network side to eNodeB). The fundamental difference to the WCDMA network was the lack of the Radio Network Controller (RNC) type of an element. The architecture is further described in Chapter 3.

The study also evaluated the resulting LTE capacity, and the studies reported in [4] and further refined studies summarized in [5] show that the requirements could be reached.

The study item was closed formally in September 2006 and detailed work items started to make the LTE part of the 3GPP Release 8 specifications.

The LTE specification work produced the first set of approved physical layer specifications in September 2007 and the first full set of approved LTE specifications in December 2007. Clearly there were open issues in the specifications at that time, especially in the protocol specifications and in the area of performance requirements. The remaining specification freezing process could be divided into three different steps:

1 Freezing the functional content of the LTE specifications in terms of what will be finalized in Release 8. This has meant leaving out some of the originally planned functionality such as support for broadcast use (point to multipoint data broadcasting). Functional freeze thus means that no new functionality can be introduced anymore but the agreed content will be finalized. In LTE the introduction of new functionality was basically over after June 2008 and during the rest of 2008 the work focused on completing the missing pieces (and correcting detected errors), especially in the protocol specifications, mainly completed for December 2008.

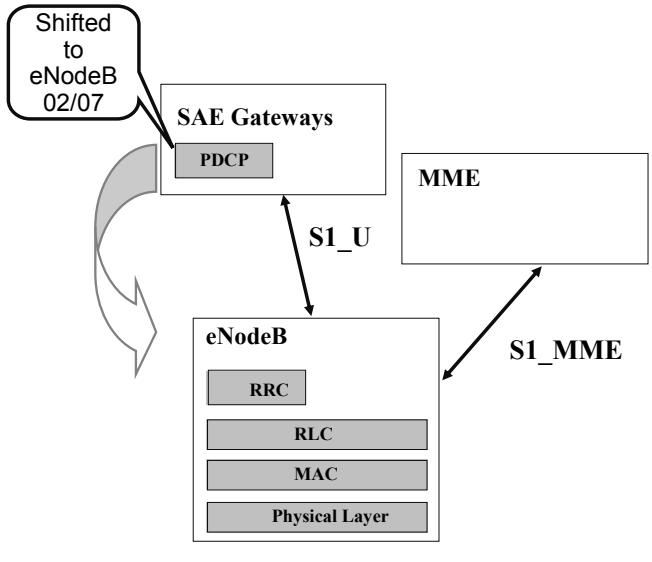


Figure 2.4 Original network architecture for LTE radio protocols

2 Once all the content is ready, the next step is to freeze the protocol specifications in terms of starting backwards compatibility. Backwards compatibility defines for a protocol the first version which can be the commercial implementation baseline. Until backwards compatibility is started, the protocol specifications are corrected by deleting any information elements that are not working as intended and replacing them with new ones. Once the start of backwards compatibility is reached, older information elements are no longer removed but extensions are used. This allows the equipment based on the older version to work based on the old information elements (though not necessarily 100% optimally), while equipment with newer software can read the improved/corrected information element after noticing the extension bit being set. Obviously, core functionality needs to work properly so that the start of backwards compatibility makes sense, as if something is totally wrong, fixing it with backwards compatible correction does not help older software versions if there is no operational functionality. It is planned to reach this step with 3GPP Release 8 protocol specifications in March 2009 when the protocol language, Abstract Syntax Notation One (ASN.1), related review for debugging all the errors is completed. With Release 7 specifications (containing HSPA improvements) this phase was reached in December 2007 following the content completion in June 2007.

3 The last phase is 'deep' freeze of the specifications, when any changes to the specifications will no longer be allowed. This is something that is valid for a release that is already rolled out in the field, such as Release 5 with HSDPA and Release 6 with HSUPA. With the devices out in the field, core functionality has been proven and tested and there is no point in any further changes to those releases, but potential improvement would need to be carried out in a later release. This kind of problem may arise when some feature has not been implemented (and thus no testing with the network has been possible) and the problem is only detected later. Then the resulting outcome could be to correct it in a later release and also recommend that the network activates it only for devices which are based on this later release. For LTE specifications, this phase is expected to be just before the actual roll-out, in a 2010 time frame, as typically some errors are detected in the implementation and trialling phase.

2.5 Evolution Beyond Release 8

The work in 3GPP during 2008 focused on Release 8 finalization, but work was started for issues beyond Release 8, including the first Release 9 topics as well as LTE-Advanced for IMT-Advanced. The following topics have been decided in 3GPP to be considered beyond Release 8:

- LTE MBMS, which is expected to cover the operation of broadcast type data both for a dedicated MBMS carrier and for a shared carrier. When synchronized properly, an OFDMA based broadcast signal can be sent in the same resource space from different base stations (with identical content) and then the signal from multiple base stations can be combined in the devices. This principle is already in use in, for example, Digital Video Broadcasting for Handhelds (DVB-H) devices in the market. DVB-H is also an OFDMA based system but only intended for broadcast use.
- Self Optimized Networks (SON) enhancements. 3GPP has worked on the self optimization/configuration aspects of LTE and that work is expected to continue in Release 9.
- Further improvements for enhanced VoIP support in LTE. In the discussions in 3GPP, it has been identified that VoIP could be further optimized to improve the maximum number

of VoIP users that could be supported simultaneously. The current capability is rather high already, as shown in Chapter 10.
- The requirements for the multi-bandwidth and multi-radio access technology base stations. The scope of this work is to define the requirements for the operation so that the same Radio Frequency (RF) part is used for transmitting, for example, LTE and GSM or LTE and WCDMA signals. Currently the requirements for the emissions on the adjacent frequencies, for example, take only a single Radio Access Technology (RAT) into account, while the requirements will now be developed for different combinations, including running multiple LTE bandwidths in parallel in addition to the multi-RAT case.

2.6 LTE-Advanced for IMT-Advanced

In parallel to the work for LTE corrections and further optimization in Release 9, the 3GPP is also targeting the creation of the input for the IMT-Advanced process in ITU-R. The ITU-R is developing the framework for next generation wireless networks. The following are the requirements from the ITU-R side for the IMT-Advanced candidate technologies, as reflected in details in the information available from ITU-R, accessible via the links given in [6]:

- support for peak data rates up to 1 Gbps for nomadic (low mobility case) and 100 Mbps for the high mobility case;
- support for larger bandwidths, and thus also 3GPP is considering specifying up to 100 MHz bandwidth support for LTE-Advanced;
- requirements for the expected spectral efficiency in different environments. In ITU-R requirements these are defined as minimum requirements and are thus different from the target type of value setting in 3GPP.

3GPP thus also has its own requirements, with the first version of the requirements approved in May 2008 as reflected in [7]. One of the 3GPP specific requirements is the backwards compatibility from the 3GPP Release 8 LTE. The requirement is defined so that a Release 8 based LTE device can operate in the LTE-Advanced system and, respectively, the Release 10 LTE Advanced device can access the Release 8 LTE networks. Obviously a Release 9 terminal would also be similarly accommodated. This could be covered, for example, with the multicarrier type of alternative as shown in Figure 2.5. The mobility between LTE-Advanced needs to work with LTE as well as GSM/EDGE, HSPA and cdma2000®.

The ITU-R process, as shown in Figure 2.6, aims for early 2011 completion of the ITU-R specifications, which requires 3GPP to submit the first full set of specifications around the end of 2010. This is one of the factors shaping the Release 10 finalization schedule, though officially the Release 10 schedule has not yet been defined in 3GPP, but will be discussed further once Release 9 work has progressed further.

3GPP has held a number of discussions on LTE-Advanced during 2008, and the technologies to be investigated include:

- Relay nodes. These are targeted for extending coverage by allowing User Equipment (UE) further away from the base station to send their data via relay nodes that can hear the eNodeB better than, for example, UE located indoors.

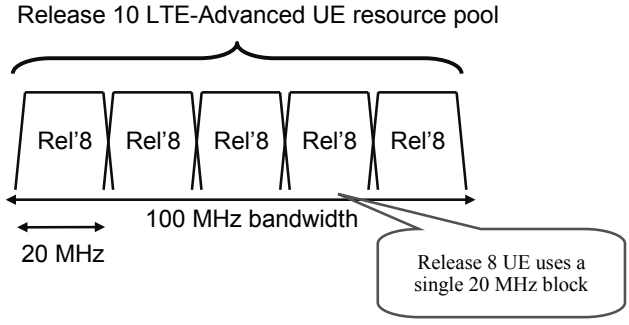


Figure 2.5 Resource sharing between LTE and LTE-Advanced

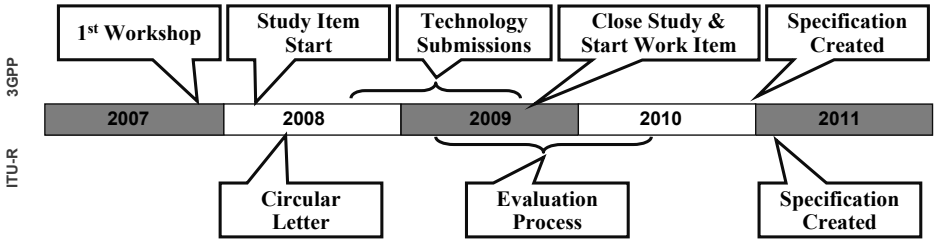


Figure 2.6 3GPP LTE-Advanced and ITU-R IMT-Advanced schedules

- UE dual transmit antenna solutions for uplink Single User MIMO (SU-MIMO) and diversity MIMO.
- Scalable system bandwidth exceeding 20 MHz, potentially up to 100 MHz. In connection with this the study has been investigating aspects related to multiple access technology with up to 100 MHz system bandwidth, and it is foreseen to be based strongly on the existing LTE solutions with extensions to larger bandwidths. How to extend the bandwidth (and how that is reflected in the multiple access) is the first topic where conclusions are expected in LTE-Advanced studies.
- Nomadic/Local Area network and mobility solutions.
- Flexible Spectrum Usage.
- Automatic and autonomous network configuration and operation.
- Coordinated Multiple Point (CoMP) transmission and reception, which is referring to MIMO transmission coordinated between different transmitters (in different sectors or even different sites in an extreme case).

It is worth noting that even though some technology is being studied, it does not necessarily mean that it will be included in the Release 10 specifications. It may be decided that some issues are already needed for Release 9 (scheduled for the end of 2009), while other issues may not be necessary at all due to low gain and/or high complexity. The 3GPP study will be completed in the second half of 2009 and then work towards the actual specifications of Release 10 will start. Some of the items from the LTE-Advanced studies are also expected to be postponed to beyond Release 10.

The process in ITU-R is open for other RAT submissions as well. Similarly, as was the case in the original IMT-2000 process, multiple RAT submissions are expected to be made available for the evaluation phase. Assuming those submissions can meet the IMT-Advanced minimum requirements, the RATs submitted are then expected to be part of the IMT-Advanced family.

2.7 LTE Specifications and 3GPP Structure

The LTE specifications mostly follow similar notation to that of the WCDMA specifications, just using the 36-series numbering. For example, when WCDMA RRC is 25.331, the corresponding LTE spec is 36.331. The LTE specifications use the term Evolved Universal Terrestrial Radio Access (E-UTRA) while the WCDMA specifications use the UTRA term (and UTRAN with N standing for Network). The terms LTE and E-UTRAN, as well as WCDMA and UTRA, are used interchangeably in the book. In the physical layer there are some differences, e.g. the specification on spreading and modulation was not needed, such as WCDMA specification 25.213. Now due to use of the same multiple access, the FDD and TDD modes are covered in the same physical layer specification series. In Figure 2.7 the specification numbers are shown for the physical layer and different protocols over the radio or internal interfaces. Note that not all the performance related specifications are shown. The following chapters will introduce the functionality in each of the interfaces shown in Figure 2.7. All the specifications listed are available from the 3GPP website [8]. When using a 3GPP specification it is always recommended that the latest version of the release in question is used. For example, version 8.0.0 is always the first approved version and versions with the number 8.4.0 (or higher) are normally more stable with fewer errors.

Inside 3GPP, the 3GPP TSG RAN is responsible for LTE specification development. The detailed specification work is covered in the Working Groups (WGs) under each TSG. TSG RAN has a total of five working groups, as shown in Figure 2.8, where working groups under other TSGs are not shown. The specifications for the Evolved Packet Core (EPC) are covered in TSA SA and in TSG CT and are also needed for an end-to-end functioning system. The TSG GERAN is responsible for the necessary Release 8 changes in GSM/EDGE specifications to facilitate the LTE-GERAN inter-working from the GERAN perspective.

From the RAN working groups the physical layer specifications 36.2 series are developed by WG1, as shown in Figure 2.8. Respectively, the Layer 2 (L2) and Layer 3 (L3) specifica-

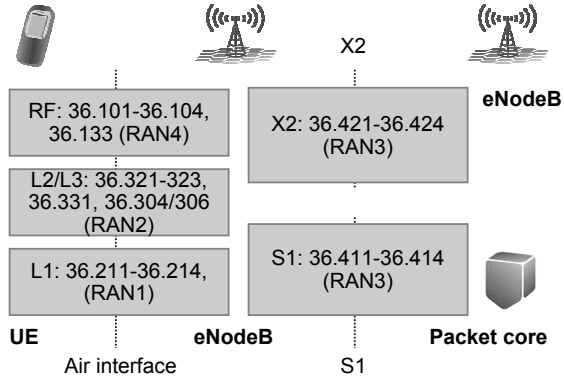


Figure 2.7. Specifications with responsible working groups for different LTE interfaces

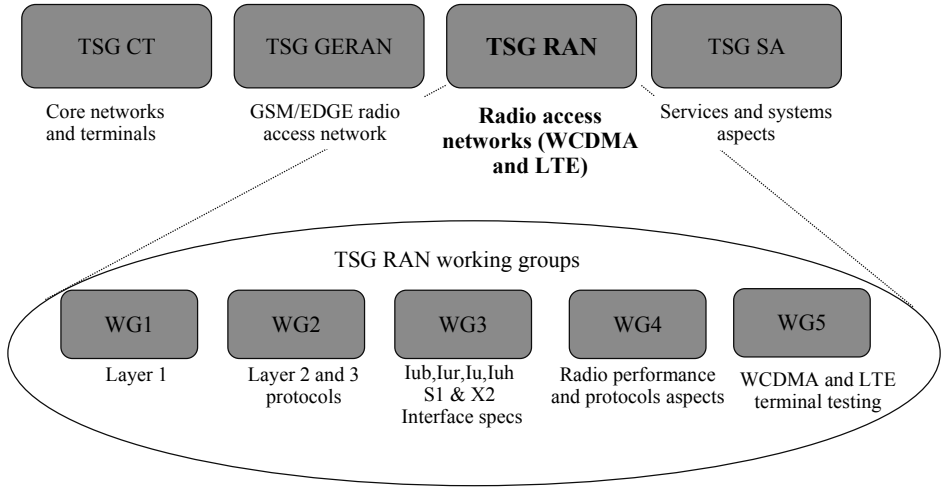


Figure 2.8 3GPP structure

tions are in the 36.3 series from WG2, internal interfaces in the 36.4 series from WG3, and radio performance requirements in the 36.1 series from WG4. Outside Figure 2.7 can be noted the LTE terminal test specifications from WG5. All groups cover the respective areas also for WCDMA/HSPA further releases as well.

References

[1] H.Holma, A.Toskala, 'WCDMA for UMTS', 4th edition, Wiley 2007.
[2] 3GPP Technical Report, TR 25.913, 'Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN)' version 7.0.0, June 2005.
[3] 3GPP Tdoc, RP-050758, LS on UTRAN LTE Multiple Access Selection, 3GPP TSG RAN WG1, November 2005.
[4] 3GPP Technical Report, TR 25.814, 'Physical layer aspects for evolved Universal Terrestrial Radio Access (UTRA)', 3GPP TSG RAN, September 2006,
[5] 3GPP Tdoc, RP-060535, LS on LTE SI Conclusions, 3GPP TSG RAN WG1, September 2006.
[6] 3GPP Tdoc, RP-080448, 'Receipt of ITU-R Circular Letter 5/LCCE/2 on IMT-Advanced', May 2008.
[7] 3GPP Technical Report, TR 36.913, 'Requirements for Further Advancements for E-UTRA (LTE-Advanced)', 3GPP TSG RAN, version 8.0.0, May 2008.
[8] www.3gpp.org

3

System Architecture Based on 3GPP SAE

Atte Länsisalmi and Antti Toskala

3.1 System Architecture Evolution in 3GPP

When the evolution of the radio interface started, it soon became clear that the system architecture would also need to be evolved. The general drive towards optimizing the system only for packet switched services is one reason that alone would have set the need for evolution, but some of the radio interface design goals – such as removal of soft handover – opened up new opportunities in the architecture design. Also, since it had been shown by High Speed Packet Access (HSPA) that all radio functionality can be efficiently co-located in the NodeB, the door was left open for discussions of flatter overall architecture.

Discussions for System Architecture Evolution (SAE) then soon followed the radio interface development, and it was agreed to schedule the completion of the work in Release 8. There had been several reasons for starting this work, and there were also many targets. The following lists some of the targets that possibly shaped the outcome the most:

- optimization for packet switched services in general, when there is no longer a need to support the circuit switched mode of operation;
- optimized support for higher throughput required for higher end user bit rates;
- improvement in the response times for activation and bearer set-up;
- improvement in the packet delivery delays;
- overall simplification of the system compared to the existing 3GPP and other cellular systems;
- optimized inter-working with other 3GPP access networks;
- optimized inter-working with other wireless access networks.

Many of the targets implied that a flat architecture would need to be developed. Flat architecture with less involved nodes reduces latencies and improves performance. Development

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6

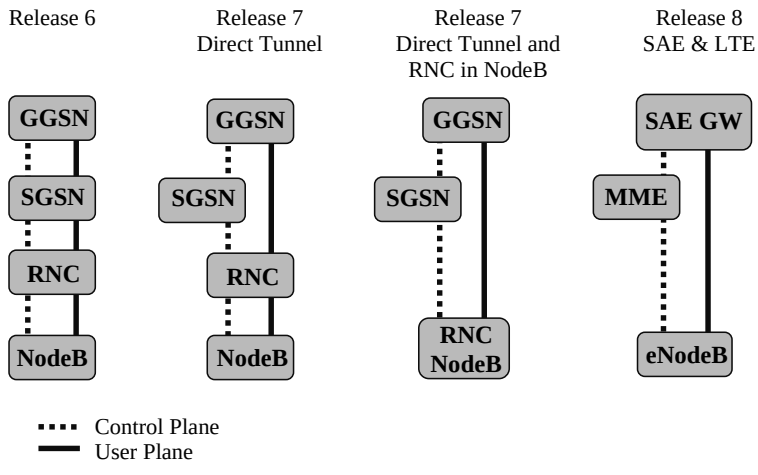


Figure 3.1 3GPP architecture evolution towards flat architecture

towards this direction had already started in Release 7 where the Direct Tunnel concept allows User Plane (UP) to bypass the SGSN, and the placement of RNC functions to HSPA NodeB was made possible. Figure 3.1 shows these evolution steps and how this aspect was captured at a high level in SAE architecture.

Some of the targets seem to drive the architecture development in completely different directions. For example, optimized inter-working with several wireless access networks (ANs) indicates the need to introduce a set of new functions and maybe even new interfaces to support specific protocols separately for each one of them. This works against the target of keeping the architecture simple. Therefore, since it is likely that that none of the actual deployments of the architecture would need to support all of the potential inter-working scenarios, the 3GPP architecture specifications were split into two tracks:

- GPRS enhancements for E-UTRAN access [1]: This document describes the architecture and its functions in its native 3GPP environment with E-UTRAN and all the other 3GPP ANs, and defines the inter-working procedures between them. The common nominator for these ANs is the use of GTP (GPRS Tunnelling Protocol) as the network mobility protocol.
- Architecture enhancements for non-3GPP accesses [2]: This document describes the architecture and functions when inter-working with non-3GPP ANs, such as cdma2000® High Rate Packet Data (HRPD), is needed. The mobility functionality in this document is based on IETF protocols, such as MIP (Mobile Internet Protocol) and PMIP (Proxy MIP), and the document also describes E-UTRAN in that protocol environment.

This chapter further describes the 3GPP system architecture in some likely deployment scenarios: basic scenario with only E-UTRAN, legacy 3GPP operator scenario with existing 3GPP ANs and E-UTRAN, and finally E-UTRAN with non-3GPP ANs, where inter-working with cdma2000® is shown as a specific example.

3.2 Basic System Architecture Configuration with only E-UTRAN Access Network

3.2.1 Overview of Basic System Architecture Configuration

Figure 3.2 describes the architecture and network elements in the architecture configuration where only the E-UTRAN AN is involved. The logical nodes and connections shown in this figure represent the basic system architecture configuration. These elements and functions are needed in all cases when E-UTRAN is involved. The other system architecture configurations described in the next sections also include some additional functions.

This figure also shows the division of the architecture into four main high level domains: User Equipment (UE), Evolved UTRAN (E-UTRAN), Evolved Packet Core Network (EPC), and the Services domain.

The high level architectural domains are functionally equivalent to those in the existing 3GPP systems. The new architectural development is limited to Radio Access and Core Networks, the E-UTRAN and the EPC respectively. UE and Services domains remain architecturally intact, but functional evolution has also continued in those areas.

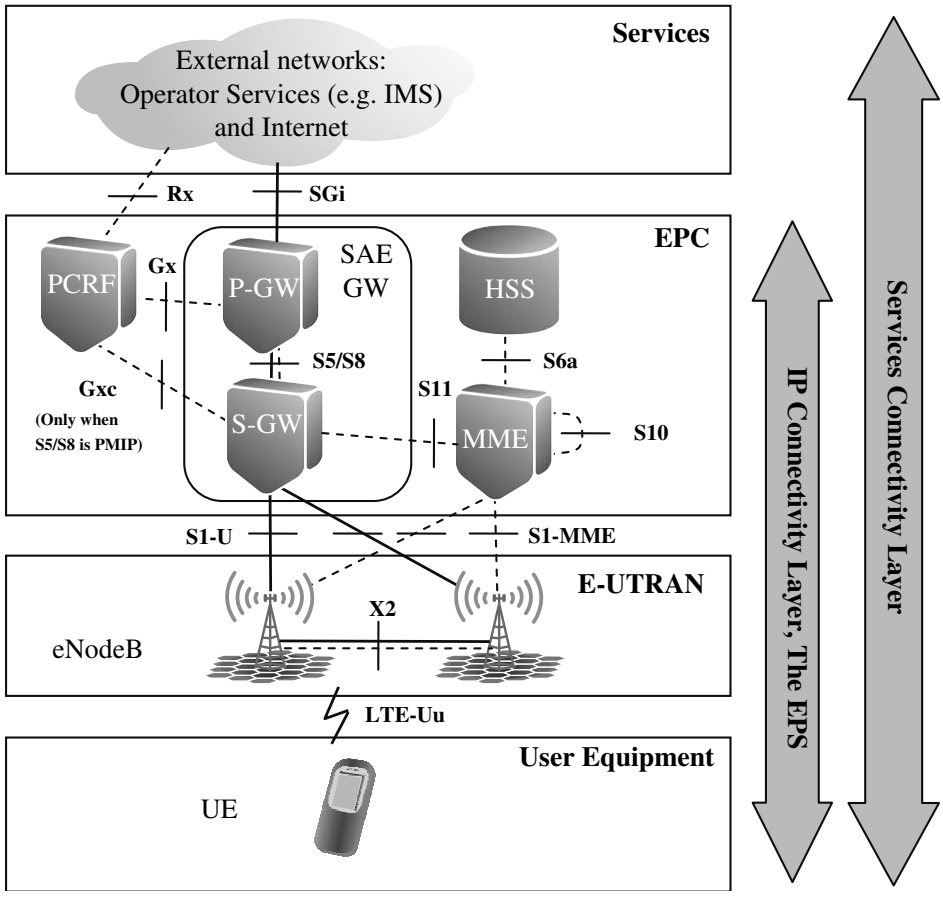


Figure 3.2 System architecture for E-UTRAN only network

UE, E-UTRAN and EPC together represent the Internet Protocol (IP) Connectivity Layer. This part of the system is also called the Evolved Packet System (EPS). The main function of this layer is to provide IP based connectivity, and it is highly optimized for that purpose only. All services will be offered on top of IP, and circuit switched nodes and interfaces seen in earlier 3GPP architectures are not present in E-UTRAN and EPC at all. IP technologies are also dominant in the transport, where everything is designed to be operated on top of IP transport.

The IP Multimedia Sub-System (IMS) [3] is a good example of service machinery that can be used in the Services Connectivity Layer to provide services on top of the IP connectivity provided by the lower layers. For example, to support the voice service, IMS can provide Voice over IP (VoIP) and interconnectivity to legacy circuit switched networks PSTN and ISDN through Media Gateways it controls.

The development in E-UTRAN is concentrated on one node, the evolved Node B (eNodeB). All radio functionality is collapsed there, i.e. the eNodeB is the termination point for all radio related protocols. As a network, E-UTRAN is simply a mesh of eNodeBs connected to neighbouring eNodeBs with the X2 interface.

One of the big architectural changes in the core network area is that the EPC does not contain a circuit switched domain, and no direct connectivity to traditional circuit switched networks such as ISDN or PSTN is needed in this layer. Functionally the EPC is equivalent to the packet switched domain of the existing 3GPP networks. There are, however, significant changes in the arrangement of functions and most nodes and the architecture in this part should be considered to be completely new.

Both Figure 3.1 and Figure 3.2 show an element called SAE GW. As the latter figure indicates, this represents the combination of the two gateways, Serving Gateway (S-GW) and Packet Data Network Gateway (P-GW) defined for the UP handling in EPC. Implementing them together as the SAE GW represents one possible deployment scenario, but the standards define the interface between them, and all operations have also been specified for when they are separate. The same approach is followed in this chapter of the book.

The Basic System Architecture Configuration and its functionality are documented in 3GPP TS 23.401 [1]. This document shows the operation when the S5/S8 interface uses the GTP protocol. However, when the S5/S8 interface uses PMIP, the functionality for these interfaces is slightly different, and the Gxc interface also is needed between the Policy and Charging Resource Function (PCRF) and S-GW. The appropriate places are clearly marked in [1] and the additional functions are described in detail in 3GPP TS 23.402 [2]. In the following sections the functions are described together for all cases that involve E-UTRAN.

3.2.2 Logical Elements in Basic System Architecture Configuration

This section introduces the logical network elements for the Basic System Architecture configuration.

3.2.2.1 User Equipment (UE)

UE is the device that the end user uses for communication. Typically it is a hand held device such as a smart phone or a data card such as those used currently in 2G and 3G, or it could be embedded, e.g. to a laptop. UE also contains the Universal Subscriber Identity Module (USIM)

that is a separate module from the rest of the UE, which is often called the Terminal Equipment (TE). USIM is an application placed into a removable smart card called the Universal Integrated Circuit Card (UICC). USIM is used to identify and authenticate the user and to derive security keys for protecting the radio interface transmission.

Functionally the UE is a platform for communication applications, which signal with the network for setting up, maintaining and removing the communication links the end user needs. This includes mobility management functions such as handovers and reporting the terminals location, and in these the UE performs as instructed by the network. Maybe most importantly, the UE provides the user interface to the end user so that applications such as a VoIP client can be used to set up a voice call.

3.2.2.2 E-UTRAN Node B (eNodeB)

The only node in the E-UTRAN is the E-UTRAN Node B (eNodeB). Simply put, the eNodeB is a radio base station that is in control of all radio related functions in the fixed part of the system. Base stations such as eNodeB are typically distributed throughout the networks coverage area, each eNodeB residing near the actual radio antennas.

Functionally eNodeB acts as a layer 2 bridge between UE and the EPC, by being the termination point of all the radio protocols towards the UE, and relaying data between the radio connection and the corresponding IP based connectivity towards the EPC. In this role, the eNodeB performs ciphering/deciphering of the UP data, and also IP header compression/decompression, which means avoiding repeatedly sending the same or sequential data in IP header.

The eNodeB is also responsible for many Control Plane (CP) functions. The eNodeB is responsible for the Radio Resource Management (RRM), i.e. controlling the usage of the radio interface, which includes, for example, allocating resources based on requests, prioritizing and scheduling traffic according to required Quality of Service (QoS), and constant monitoring of the resource usage situation.

In addition, the eNodeB has an important role in Mobility Management (MM). The eNodeB controls and analyses radio signal level measurements carried out by the UE, makes similar measurements itself, and based on those makes decisions to handover UEs between cells. This includes exchanging handover signalling between other eNodeBs and the MME. When a new UE activates under eNodeB and requests connection to the network, the eNodeB is also responsible for routing this request to the MME that previously served that UE, or selecting a new MME, if a route to the previous MME is not available or routing information is absent.

Details of these and other E-UTRAN radio interface functions are described extensively elsewhere in this book. The eNodeB has a central role in many of these functions.

Figure 3.3 shows the connections that eNodeB has to the surrounding logical nodes, and summarizes the main functions in these interfaces. In all the connections the eNodeB may be in a one-to-many or a many-to-many relationship. The eNodeB may be serving multiple UEs at its coverage area, but each UE is connected to only one eNodeB at a time. The eNodeB will need to be connected to those of its neighbouring eNodeBs with which a handover may need to be made.

Both MMEs and S-GWs may be pooled, which means that a set of those nodes is assigned to serve a particular set of eNodeBs. From a single eNodeB perspective this means that it may need to connect to many MMEs and S-GWs. However, each UE will be served by only one MME and S-GW at a time, and the eNodeB has to keep track of this association. This associa-

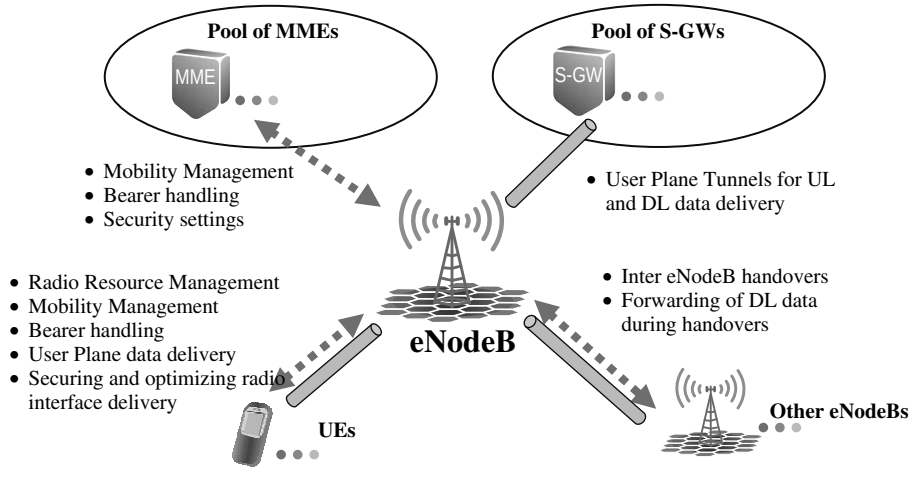


Figure 3.3 eNodeB connections to other logical nodes and main functions

tion will never change from a single eNodeB point of view, because MME or S-GW can only change in association with inter-eNodeB handover.

3.2.2.3 Mobility Management Entity (MME)

Mobility Management Entity (MME) is the main control element in the EPC. Typically the MME would be a server in a secure location in the operator's premises. It operates only in the CP, and is not involved in the path of UP data.

In addition to interfaces that terminate to MME in the architecture as shown in Figure 3.2, the MME also has a logically direct CP connection to the UE, and this connection is used as the primary control channel between the UE and the network. The following lists the main MME functions in the basic System Architecture Configuration:

- Authentication and Security: When a UE registers to the network for the first time, the MME initiates the authentication, by performing the following: it finds out the UE's permanent identity either from the previously visited network or the UE itself; requests from the Home Subscription Server (HSS) in UE's home network the authentication vectors which contain the authentication challenge – response parameter pairs; sends the challenge to the UE; and compares the response received from the UE to the one received from the home network. This function is needed to assure that the UE is who it claims to be. The details of EPS-AKA authentication are defined in [4]. The MME may repeat authentication when needed or periodically. The MME will calculate UEs ciphering and integrity protection keys from the master key received in the authentication vector from the home network, and it controls the related settings in E-UTRAN for UP and CP separately. These functions are used to protect the communication from eavesdropping and from alteration by unauthorized third parties respectively. To protect the UE privacy, MME also allocates each UE a temporary identity called the Globally Unique Temporary Identity (GUTI), so that the need to send the permanent UE identity – International Mobile Subscriber Identity (IMSI) – over the

radio interface is minimized. The GUTI may be re-allocated, e.g. periodically to prevent unauthorized UE tracking.

- Mobility Management: The MME keeps track of the location of all UEs in its service area. When a UE makes its first registration to the network, the MME will create an entry for the UE, and signal the location to the HSS in the UE's home network. The MME requests the appropriate resources to be set up in the eNodeB, as well as in the S-GW which it selects for the UE. The MME will then keep tracking the UE's location either on the level of eNodeB, if the UE remains connected, i.e. is in active communication, or at the level of Tracking Area (TA), which is a group of eNodeBs in case the UE goes to idle mode, and maintaining a through connected data path is not needed. The MME controls the setting up and releasing of resources based on the UE's activity mode changes. The MME also participates in control signalling for handover of an active mode UE between eNodeBs, S-GWs or MMEs. MME is involved in every eNodeB change, since there is no separate Radio Network Controller to hide most of these events. An idle UE will report its location either periodically, or when it moves to another Tracking Area. If data are received from the external networks for an idle UE, the MME will be notified, and it requests the eNodeBs in the TA that is stored for the UE to page the UE.
- Managing Subscription Profile and Service Connectivity: At the time of a UE registering to the network, the MME will be responsible for retrieving its subscription profile from the home network. The MME will store this information for the duration it is serving the UE. This profile determines what Packet Data Network connections should be allocated to the UE at network attachment. The MME will automatically set up the default bearer, which gives the UE the basic IP connectivity. This includes CP signalling with the eNodeB, and the S-GW. At any point later on, the MME may need to be involved in setting up dedicated bearers for services that benefit from higher treatment. The MME may receive the request to set up a dedicated bearer either from the S-GW if the request originates from the operator service domain, or directly from the UE, if the UE requires a connection for a service that is not known by the operator service domain, and therefore cannot be initiated from there.

Figure 3.4 shows the connections MME has to the surrounding logical nodes, and summarizes the main functions in these interfaces. In principle the MME may be connected to any other MME in the system, but typically the connectivity is limited to one operator network only. The remote connectivity between MMEs may be used when a UE that has travelled far away while powered down registers to a new MME, which then retrieves the UE's permanent identity, the International Mobile Subscriber Identity (IMSI), from the previously visited MME. The inter-MME connection with neighbouring MMEs is used in handovers.

Connectivity to a number of HSSs will also need to be supported. The HSS is located in each user's home network, and a route to that can be found based on the IMSI. Each MME will be configured to control a set of S-GWs and eNodeBs. Both the S-GWs and eNodeBs may also be connected to other MMEs. The MME may serve a number of UEs at the same time, while each UE will only connect to one MME at a time.

3.2.2.4 Serving Gateway (S-GW)

In the Basic System Architecture configuration, the high level function of S-GW is UP tunnel management and switching. The S-GW is part of the network infrastructure maintained centrally in operation premises.

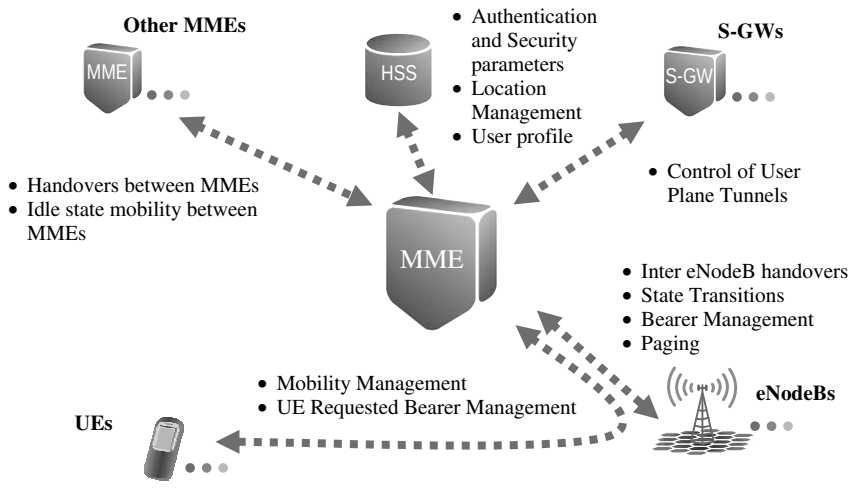


Figure 3.4 MME connections to other logical nodes and main functions

When the S5/S8 interface is based on GTP, the S-GW will have GTP tunnels on all its UP interfaces. Mapping between IP service flows and GTP tunnels is done in P-GW, and the S-GW does not need to be connected to PCRF. All control is related to the GTP tunnels, and comes from either MME or P-GW. When the S5/S8 interface uses PMIP, the S-GW will perform the mapping between IP service flows in S5/S8 and GTP tunnels in S1-U interfaces, and will connect to PCRF to receive the mapping information.

The S-GW has a very minor role in control functions. It is only responsible for its own resources, and it allocates them based on requests from MME, P-GW or PCRF, which in turn are acting on the need to set up, modify or clear bearers for the UE. If the request was received from P-GW or PCRF, the S-GW will also relay the command on to the MME so that it can control the tunnel to eNodeB. Similarly, when the MME initiated the request, the S-GW will signal on to either the P-GW or the PCRF, depending on whether S5/S8 is based on GTP or PMIP respectively. If the S5/S8 interface is based on PMIP, the data in that interface will be IP flows in one GRE tunnel for each UE, whereas in the GTP based S5/S8 interface each bearer will have its own GTP tunnel. Therefore S-GW supporting PMIP S5/S8 is responsible for bearer binding, i.e. mapping the IP flows in S5/S8 interface to bearers in the S1 interface. This function in S-GW is called Bearer Binding and Event Reporting Function (BBERF). Irrespective of where the bearer signalling started, the BBERF always receives the bearer binding information from PCRF.

During mobility between eNodeBs, the S-GW acts as the local mobility anchor. The MME commands the S-GW to switch the tunnel from one eNodeB to another. The MME may also request the S-GW to provide tunnelling resources for data forwarding, when there is a need to forward data from source eNodeB to target eNodeB during the time UE makes the radio handover. The mobility scenarios also include changing from one S-GW to another, and the MME controls this change accordingly, by removing tunnels in the old S-GW and setting them up in a new S-GW.

For all data flows belonging to a UE in connected mode, the S-GW relays the data between eNodeB and P-GW. However, when a UE is in idle mode, the resources in eNodeB are released, and the data path terminates in the S-GW. If S-GW receives data packets from P-GW on any such

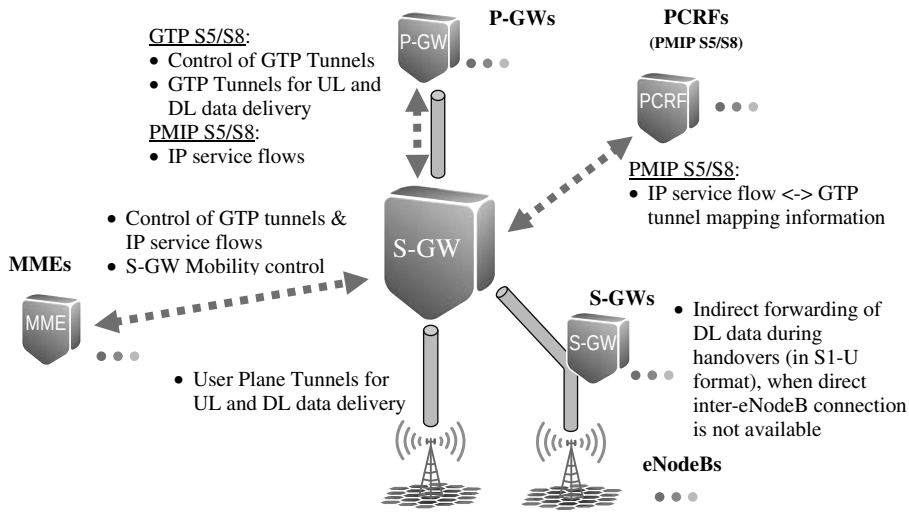


Figure 3.5 S-GW connections to other logical nodes and main functions

tunnel, it will buffer the packets, and request the MME to initiate paging of the UE. Paging will cause the UE to re-connect, and when the tunnels are re-connected, the buffered packets will be sent on. The S-GW will monitor data in the tunnels, and may also collect data needed for accounting and user charging. The S-GW also includes functionality for Lawful Interception, which means the capability to deliver the monitored user's data to authorities for further inspection.

Figure 3.5 shows how the S-GW is connected to other logical nodes, and lists the main functions in these interfaces. All interfaces have to be configured in a one-to-many fashion from the S-GW point of view. One S-GW may be serving only a particular geographical area with a limited set of eNodeBs, and likewise there may be a limited set of MMEs that control that area. The S-GW should be able to connect to any P-GW in the whole network, because P-GW will not change during mobility, while the S-GW may be relocated, when the UE moves. For connections related to one UE, the S-GW will always signal with only one MME, and the UP points to one eNodeB at a time (indirect data forwarding is the exception, see next paragraph). If one UE is allowed to connect to multiple PDNs through different P-GWs, then the S-GW needs to connect to those separately. If the S5/S8 interface is based on PMIP, the S-GW connects to one PCRF for each separate P-GW the UE is using.

Figure 3.5 also shows the indirect data forwarding case where UP data is forwarded between eNodeBs through the S-GWs. There is no specific interface name associated to the interface between S-GWs, since the format is exactly the same as in the S1-U interface, and the involved S-GWs may consider that they are communicating directly with an eNodeB. This would be the case if indirect data forwarding takes place via only one S-GW, i.e. both eNodeBs can be connected to the same S-GW.

3.2.2.5 Packet Data Network Gateway (P-GW)

Packet Data Network Gateway (P-GW, also often abbreviated as PDN-GW) is the edge router between the EPS and external packet data networks. It is the highest level mobility anchor

in the system, and usually it acts as the IP point of attachment for the UE. It performs traffic gating and filtering functions as required by the service in question. Similarly to the S-GW, the P-GWs are maintained in operator premises in a centralized location.

Typically the P-GW allocates the IP address to the UE, and the UE uses that to communicate with other IP hosts in external networks, e.g. the internet. It is also possible that the external PDN to which the UE is connected allocates the address that is to be used by the UE, and the P-GW tunnels all traffic to that network. The IP address is always allocated when the UE requests a PDN connection, which happens at least when the UE attaches to the network, and it may happen subsequently when a new PDN connectivity is needed. The P-GW performs the required Dynamic Host Configuration Protocol (DHCP) functionality, or queries an external DHCP server, and delivers the address to the UE. Also dynamic auto-configuration is supported by the standards. Only IPv4, only IPv6 or both addresses may be allocated depending on the need, and the UE may signal whether it wants to receive the address(es) in the Attach signalling, or if it wishes to perform address configuration after the link layer is connected.

The P-GW includes the PCEF, which means that it performs gating and filtering functions as required by the policies set for the UE and the service in question, and it collects and reports the related charging information.

The UP traffic between P-GW and external networks is in the form of IP packets that belong to various IP service flows. If the S5/S8 interface towards S-GW is based on GTP, the P-GW performs the mapping between the IP data flows to GTP tunnels, which represent the bearers. The P-GW sets up bearers based on request either through the PCRF or from the S-GW, which relays information from the MME. In the latter case, the P-GW may also need to interact with the PCRF to receive the appropriate policy control information, if that is not configured in the P-GW locally. If the S5/S8 interface is based on PMIP, the P-GW maps all the IP Service flows from external networks that belong to one UE to a single GRE tunnel, and all control information is exchanged with PCRF only. The P-GW also has functionality for monitoring the data flow for accounting purposes, as well as for Lawful Interception.

P-GW is the highest level mobility anchor in the system. When a UE moves from one S-GW to another, the bearers have to be switched in the P-GW. The P-GW will receive an indication to switch the flows from the new S-GW.

Figure 3.6 shows the connections P-GW has to the surrounding logical nodes, and lists the main functions in these interfaces. Each P-GW may be connected to one or more PCRF, S-GW and external network. For a given UE that is associated with the P-GW, there is only one S-GW, but connections to many external networks and respectively to many PCRFs may need to be supported, if connectivity to multiple PDNs is supported through one P-GW.

3.2.2.6 Policy and Charging Resource Function (PCRF)

Policy and Charging Resource Function (PCRF) is the network element that is responsible for Policy and Charging Control (PCC). It makes decisions on how to handle the services in terms of QoS, and provides information to the PCEF located in the P-GW, and if applicable also to the BBERF located in the S-GW, so that appropriate bearers and policing can be set up. PCRF is part of the PCC framework defined in [5]. PCRF is a server usually located with other CN elements in operator switching centres.

The information the PCRF provides to the PCEF is called the PCC rules. The PCRF will send the PCC rules whenever a new bearer is to be set up. Bearer set-up is required, for exam-

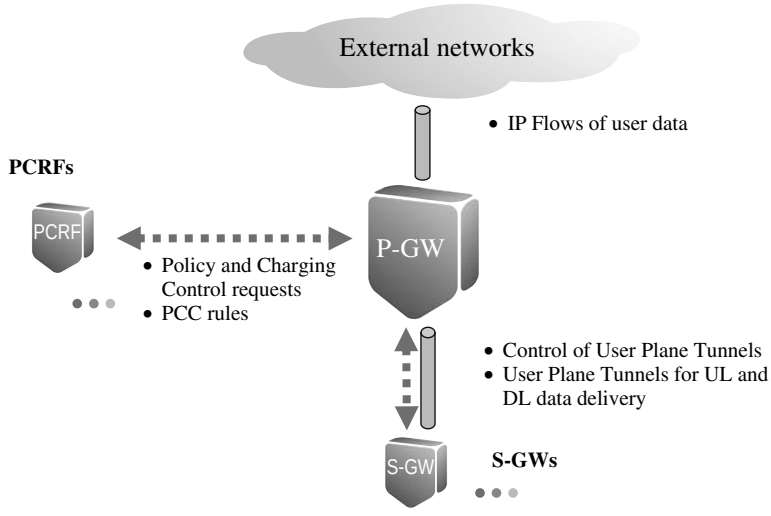


Figure 3.6 P-GW connections to other logical nodes and main functions

ple, when the UE initially attaches to the network and the default bearer will be set up, and subsequently when one or more dedicated bearers are set up. The PCRF will be able to provide PCC rules based on request either from the P-GW and also the S-GW in PMIP case, like in the attach case, and also based on request from the Application Function (AF) that resides in the Services Domain. In this scenario the UE has signalled directly with the Services Domain, e.g. with the IMS, and the AF pushes the service QoS information to PCRF, which makes a PCC decision, and pushes the PCC rules to the P-GW, and bearer mapping information to S-GW in PMIP S5/S8 case. The EPC bearers are then set up based on those.

The connections between the PCRF and the other nodes are shown in Figure 3.7. Each PCRF may be associated with one or more AF, P-GW and S-GW. There is only one PCRF associated with each PDN connection that a single UE has.

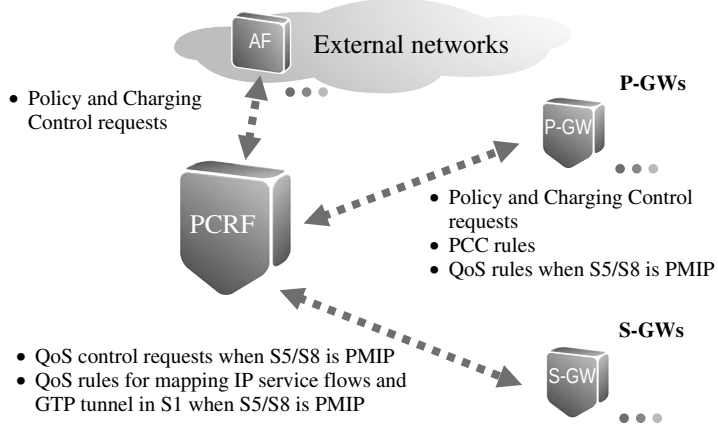


Figure 3.7 PCRF connections to other logical nodes and main functions

3.2.2.7 Home Subscription Server (HSS)

Home Subscription Server (HSS) is the subscription data repository for all permanent user data. It also records the location of the user in the level of visited network control node, such as MME. It is a database server maintained centrally in the home operator's premises.

The HSS stores the master copy of the subscriber profile, which contains information about the services that are applicable to the user, including information about the allowed PDN connections, and whether roaming to a particular visited network is allowed or not. For supporting mobility between non-3GPP ANs, the HSS also stores the Identities of those P-GWs that are in use. The permanent key, which is used to calculate the authentication vectors that are sent to a visited network for user authentication and deriving subsequent keys for encryption and integrity protection, is stored in the Authentication Center (AuC), which is typically part of the HSS. In all signalling related to these functions, the HSS interacts with the MME. The HSS will need to be able to connect with every MME in the whole network, where its UEs are allowed to move. For each UE, the HSS records will point to one serving MME at a time, and as soon as a new MME reports that it is serving the UE, the HSS will cancel the location from the previous MME.

3.2.2.8 Services Domain

The Services domain may include various sub-systems, which in turn may contain several logical nodes. The following is a categorization of the types of services that will be made available, and a short description of what kind of infrastructure would be needed to provide them:

- IMS based operator services: The IP Multimedia Sub-system (IMS) is service machinery that the operator may use to provide services using the Session Initiation Protocol (SIP). IMS has 3GPP defined architecture of its own, and is described in section 3.6, and more thoroughly, e.g. in [3].
- Non-IMS based operator services: The architecture for non-IMS based operator services is not defined in the standards. The operator may simply place a server into their network, and the UEs connect to that via some agreed protocol that is supported by an application in the UE. A video streaming service provided from a streaming server is one such example.
- Other services not provided by the mobile network operator, e.g. services provided through the internet: This architecture is not addressed by the 3GPP standards, and the architecture depends on the service in question. The typical configuration would be that the UE connects to a server in the internet, e.g. to a web-server for web browsing services, or to a SIP server for internet telephony service (i.e. VoIP).

3.2.3 Self-configuration of S1-MME and X2 interfaces

In 3GPP Release 8 development it has been agreed to define the support for self-configuration of the S1-MME and X2 interfaces. The basic process is as presented in Figure 3.8, where the eNodeB once turned on (and given that the IP connection exists) will connect to the O&M (based on the known IP address) to obtain then further parameters in terms of which other network elements to connect (and also for eNodeB software download) as well as initial parameters for the operation, such as in which part of the frequency band to operate and what kind of parameters to include for the broadcast channels.

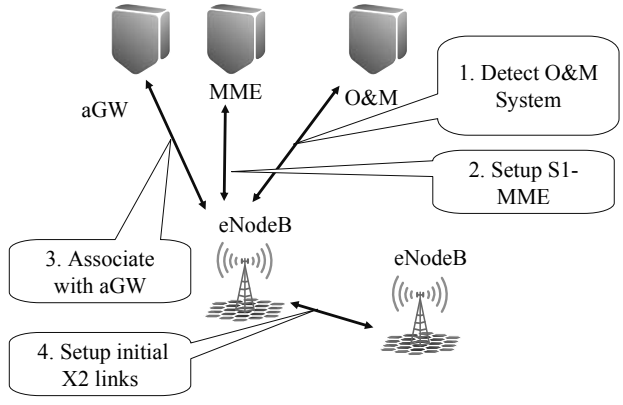


Figure 3.8 eNodeB self-configuration steps

This is expected to include setting the S1-MME connection by first setting up the SCTP association with at least one MME, and once that is connected to continue with application level information exchange to make S1-MME interface operational. Once the link to MME exists, there needs to be then association with S-GW created for UP data transfer.

To enable functionalities such as mobility and inter-cell interference control, the X2 interface configuration follows similar principles to the S1-MME interface. The difference here is that initially the eNodeB will set up the X2 connection for those eNodeBs indicated from the O&M and it may then later adapt more to the environment based on the Automatic Neighbour Relationship (ANR) functionality – as covered in Chapter 7 – to further optimize the X2 connectivity domain based on actual handover needs. The parameters that are exchanged over the X2 interface include:

- global eNodeB ID;
- information of the cell specific parameters such as Physical Cell ID (PCI), uplink/downlink frequency used, bandwidth in use;
- MMEs connected (MME Pool).

For the PCI there is also support for auto-configuration in the Release 8 specifications as covered in Chapter 5, other parameters then coming from the O&M direction with procedures that can be automated to limit the need for on-site configuration by installation personnel.

3.2.4 Interfaces and Protocols in Basic System Architecture Configuration

Figure 3.9 shows the CP protocols related to a UE's connection to a PDN. The interfaces from a single MME are shown in two parts, the one on top showing protocols towards the E-UTRAN and UE, and the bottom one showing protocols towards the gateways. Those protocols that are shown in white background are developed by 3GPP, while the protocols with light grey background are developed in IETF, and represent standard internet technologies that are used for transport in EPS. 3GPP has only defined the specific ways of how these protocols are used.

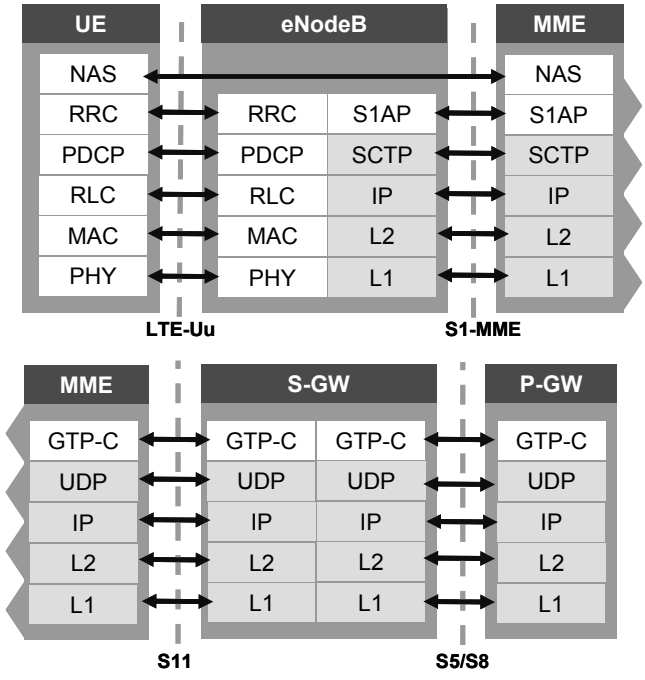


Figure 3.9 Control plane protocol stack in EPS

The topmost layer in the CP is the Non-Access Stratum (NAS), which consists of two separate protocols that are carried on direct signalling transport between the UE and the MME. The content of the NAS layer protocols is not visible to the eNodeB, and the eNodeB is not involved in these transactions by any other means, besides transporting the messages, and providing some additional transport layer indications along with the messages in some cases. The NAS layer protocols are:

- EPS Mobility Management (EMM): The EMM protocol is responsible for handling the UE mobility within the system. It includes functions for attaching to and detaching from the network, and performing location updating in between. This is called Tracking Area Updating (TAU), and it happens in idle mode. Note that the handovers in connected mode are handled by the lower layer protocols, but the EMM layer does include functions for re-activating the UE from idle mode. The UE initiated case is called Service Request, while Paging represents the network initiated case. Authentication and protecting the UE identity, i.e. allocating the temporary identity GUTI to the UE are also part of the EMM layer, as well as the control of NAS layer security functions, encryption and integrity protection.
- EPS Session Management (ESM): This protocol may be used to handle the bearer management between the UE and MME, and it is used in addition for E-UTRAN bearer management procedures. Note that the intention is not to use the ESM procedures if the bearer contexts are already available in the network and E-UTRAN procedures can be run immediately. This would be the case, for example, when the UE has already signalled with an operator affiliated Application Function in the network, and the relevant information has been made available through the PCRF.

The radio interface protocols are (only short descriptions are included here, since these functions are described extensively in other sections of this book):

- Radio Resource Control (RRC): This protocol is in control of the radio resource usage. It manages UE's signalling and data connections, and includes functions for handover.
- Packet Data Convergence Protocol (PDCP): The main functions of PDCP are IP header compression (UP), encryption and integrity protection (CP only).
- Radio Link Control (RLC): The RLC protocol is responsible for segmenting and concatenation of the PDCP-PDUs for radio interface transmission. It also performs error correction with the Automatic Repeat Request (ARQ) method.
- Medium Access Control (MAC): The MAC layer is responsible for scheduling the data according to priorities, and multiplexing data to Layer 1 transport blocks. The MAC layer also provides error correction with Hybrid ARQ.
- Physical Layer (PHY): This is the Layer 1 of LTE-Uu radio interface that takes care of DS-CDMA Layer functions.

The S1 interface connects the E-UTRAN to the EPC, and involves the following protocols:

- S1 Application Protocol (S1AP): S1AP handles the UE's CP and UP connections between the E-UTRAN and EPC, including participating in the handover when EPC is involved.
- SCTP/IP signalling transport: The Stream Control Transmission Protocol (SCTP) and Internet Protocol (IP) represent standard IP transport suitable for signalling messages. SCTP provides the reliable transport and sequenced delivery functions. IP itself can be run on a variety of data link and physical layer technologies (L2 and L1), which may be selected based on availability.

In the EPC, there are two alternative protocols for the S5/S8 interface. The following protocols are involved, when GTP is used in S5/S8:

- GPRS Tunnelling Protocol, Control Plane (GTP-C): It manages the UP connections in the EPC. This includes signalling the QoS and other parameters. If GTP is used in the S5/S8 interface it also manages the GTP-U tunnels. GTP-C also performs the mobility management functions within the EPC, e.g. when the GTP-U tunnels of a UE need to be switched from one node to the other.
- UDP/IP transport. The Unit Data Protocol (UDP) and IP are used as the standard and basic IP transport. UDP is used instead of Transmission Control Protocol (TCP) because the higher layers already provide reliable transport with error recovery and re-transmission. IP packets in EPC may be transported on top of a variety of L2 and L1 technologies. Ethernet and ATM are some examples.

The following protocols are used, when S5/S8 is based on PMIP:

- Proxy Mobile IP (PMIP): PMIP is the alternative protocol for the S5/S8 interface. It takes care of mobility management, but does not include bearer management functions as such. All traffic belonging to a UE's connection to a particular PDN is handled together.
- IP: PMIP runs directly on top of IP, and it is used as the standard IP transport.

Figure 3.10 illustrates the UP protocol structure for UE connecting to P-GW.

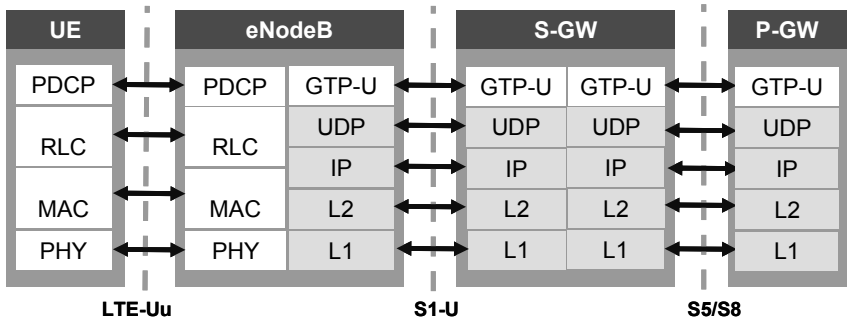


Figure 3.10 User plane protocol stack in EPS

The UP shown in Figure 3.10 includes the layers below the end user IP, i.e. these protocols form the Layer 2 used for carrying the end user IP packets. The protocol structure is very similar to the CP. This highlights the fact that the whole system is designed for generic packet data transport, and both CP signalling and UP data are ultimately packet data. Only the volumes are different. Most of the protocols have been introduced already above, with the exception of the following two that follow the selection of protocol suite in S5/S8 interface:

- GPRS Tunnelling Protocol, User Plane (GTP-U): GTP-U is used when S5/S8 is GTP based. GTP-U forms the GTP-U tunnel that is used to send End user IP packets belonging to one EPS bearer. It is used in S1-U interface, and is used in S5/S8 if the CP uses GTP-C.
- Generic Routing Encapsulation (GRE): GRE is used in the S5/S8 interface in conjunction with PMIP. GRE forms an IP in IP tunnel for transporting all data belonging to one UE's connection to a particular PDN. GRE is directly on top of IP, and UDP is not used.

Figure 3.11 illustrates the X2 interface protocol structure, which resembles that of the S1 interface. Only the CP Application Protocol is different. X2 interface is used in mobility between the eNodeBs, and the X2AP includes functions for handover preparation, and overall maintenance of the relation between neighbouring eNodeBs. The UP in the X2 interface is used for forwarding data in a transient state during handover, when the radio interface is already

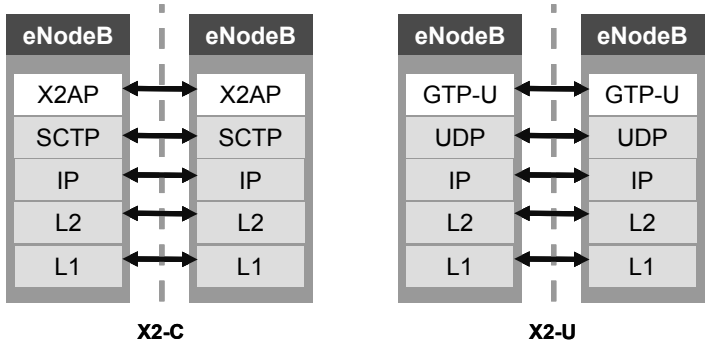


Figure 3.11 Control and user plane protocol stacks for X2 interface

Table 3.1 Summary of interfaces and protocols in Basic System Architecture configuration

Interface	Protocols	Specification
LTE-Uu	CP: RRC/PDCP/RLC/MAC/PHY	36.300 [6]
	UP: PDCP/RLC/MAC/PHY	(stage 2)
X2	CP: X2AP/SCTP/IP	36.423 [7]
	UP: GTP-U/UDP/IP	29.274 [8]
S1-MME	S1AP/SCTP/UDP/IP	36.413 [9]
S1-U	GTP-U/UDP/IP	29.274 [8]
S10	GTP-C/UDP/IP	29.274 [8]
S11	GTP-C/UDP/IP	29.274 [8]
S5/S8 (GTP)	GTP/UDP/IP	29.274 [8]
S5/S8 (PMIP)	CP: PMIP/IP	29.275 [10]
	UP: GRE/IP	
SGi	IP (also Diameter & Radius)	29.061 [11]
S6a	Diameter/SCTP/IP	29.272 [12]
Gx	Diameter/SCTP/IP	29.212 [13]
Gxc	Diameter/SCTP/IP	29.212 [13]
Rx	Diameter/SCTP/IP	29.214 [14]
UE – MME	EMM, ESM	24.301 [15]

disconnected on the source side, and has not yet resumed on the target side. Data forwarding is done for the DL data, since the UL data can be throttled effectively by the UE.

Table 3.1 summarizes the protocols and interfaces in Basic System Architecture configuration.

3.2.5 Roaming in Basic System Architecture Configuration

Roaming is an important functionality, where operators share their networks with each other's subscribers. Typically roaming happens between operators serving different areas, such as different countries, since this does not cause conflicts in the competition between the operators, and the combined larger service area benefits them as well as the subscribers. The words *home* and *visited* are used as prefixes to many other architectural terms to describe where the subscriber originates from and where it roams to respectively.

3GPP SAE specifications define which interfaces can be used between operators, and what additional considerations are needed if an operator boundary is crossed. In addition to the connectivity between the networks, roaming requires that the operators agree on many things at the service level, e.g. what services are available, how they are realized, and how accounting and charging is handled. This agreement is called the *Roaming Agreement*, and it can be made directly between the operators, or through a broker. The 3GPP specifications do not cover these items, and operators using 3GPP technologies discuss roaming related general questions in a private forum called the GSM Association, which has published recommendations to cover these additional requirements.

Roaming defined for SAE follows quite similar principles to the earlier 3GPP architectures. The E-UTRAN is always locally in the visited network, but the data may be routed either to the home network, or can break out to external networks directly from the visited network. This aspect differentiates the two roaming models supported for SAE, which are defined as follows:

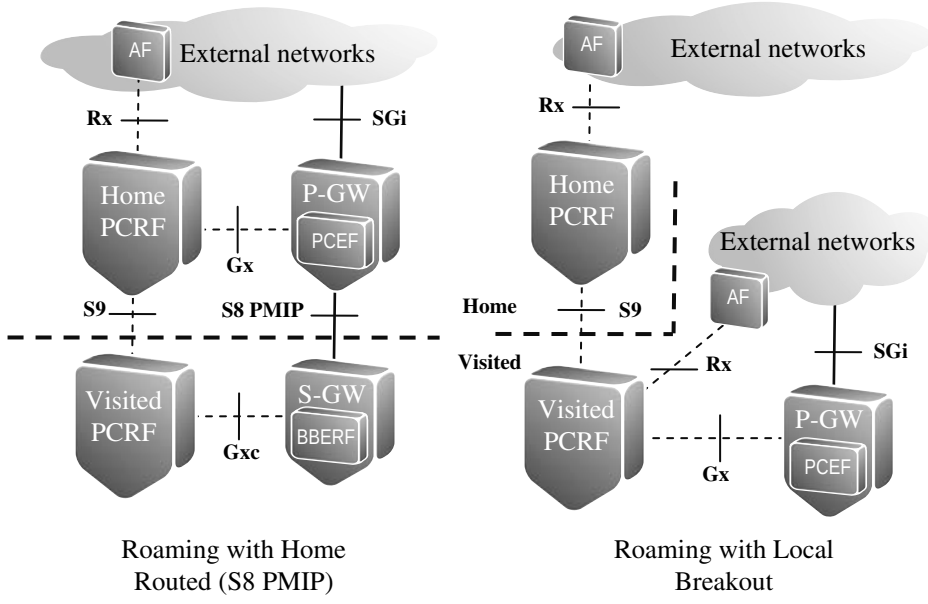


Figure 3.12 Home routed and local breakout roaming

- Home Routed model: The P-GW, HSS and PCRF reside in the home operator network, and the S-GW, MME and the radio networks reside in the visited operator network. In this roaming configuration the interface between P-GW and S-GW is called S8, whereas the same interface is called S5 when S-GW and P-GW are in the same operator's network. S5 and S8 are technically equivalent. When the S8 interface is based in GTP, the roaming architecture is as shown in Figure 3.2 (Gxa does not apply with GTP). When the S8 interface uses PMIP, the PCRF will also be divided into home and visited nodes with the S9 interface between them. This is the scenario shown in Figure 3.12 on the left, and is explained with more detail in section 3.7.1. The Home Routed roaming model applies to legacy 3GPP ANs in the same way, the additional detail being that the SGSN introduced in the next chapter and shown in Figure 3.12 resides in the visited network.
- Local Breakout model: In this model, shown in the right side of Figure 3.12, the P-GW will be located in the visited network, and the HSS is in the home network. If dynamic policy control is used, there will again be two PCRFs involved, one in the home network, and the other in the visited network. Depending on which operator's services are used, the PCRF in that operator's network is also connected to the AF. Also this scenario is explained with more detail in section 3.7.1. With these constraints the Local Breakout model also works with the legacy 3GPP ANs.

3.3 System Architecture with E-UTRAN and Legacy 3GPP Access Networks

3.3.1 Overview of 3GPP Inter-working System Architecture Configuration

Figure 3.13 describes the architecture and network elements in the architecture configuration where all 3GPP defined ANs, E-UTRAN, UTRAN and GERAN, are connected to the EPC.

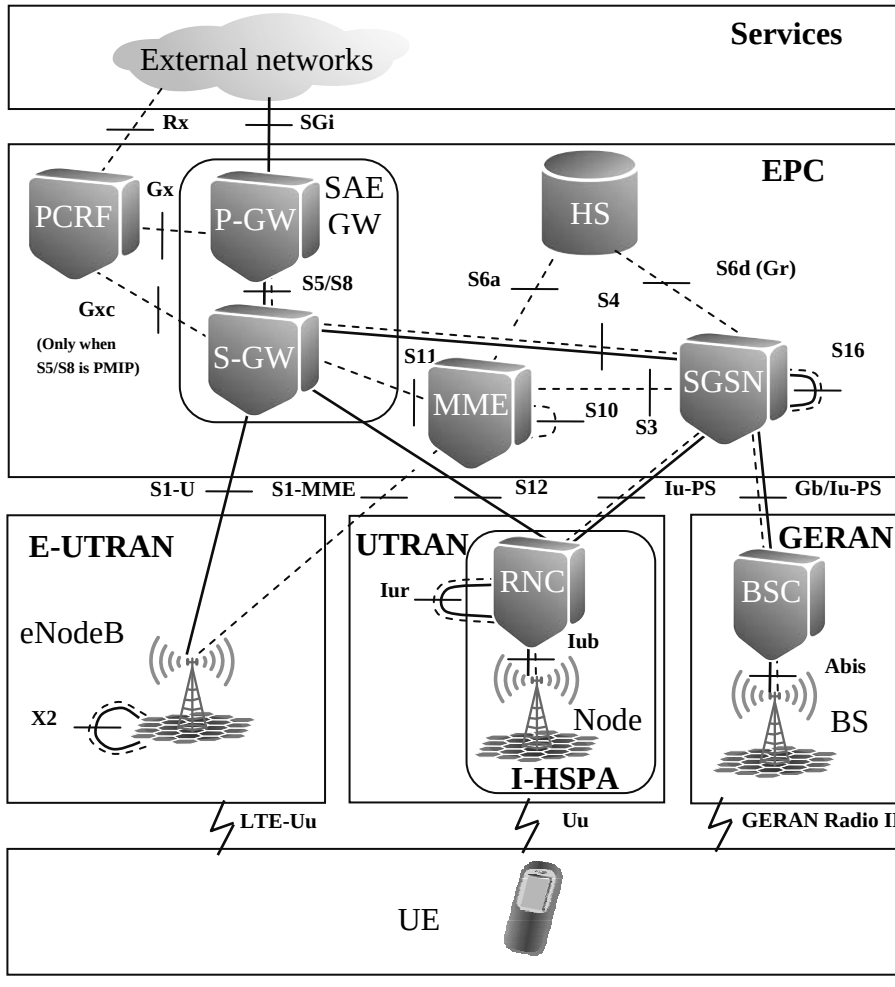


Figure 3.13 System architecture for 3GPP access networks

This is called here the 3GPP Inter-working System Architecture Configuration, and it allows optimized inter-working between the mentioned accesses.

Functionally the E-UTRAN, UTRAN and GERAN all provide very similar connectivity services, especially when looking at the situation from the end user point of view, where the only difference may be the different data rates and improved performance, but architecturally these ANs are quite different, and many things are carried out differently. There are, for example, big differences in how the bearers are managed in the EPS compared to the existing networks with UTRAN or GERAN access. However, when UTRAN or GERAN is connected to EPC, they may still operate as before from this perspective, and for this purpose the S-GW simply assumes the role of the Gateway GPRS Support Node (GGSN). Also in optimized inter-working with the E-UTRAN, the GERAN and UTRAN ANs behave almost the same way as they behave when inter-working between themselves. The differences become more visible in the EPC, because what used to be the fixed GGSN is now the S-GW that may be changed along with the SGSN change during UE mobility.

All nodes and functions described in the previous section for the Basic System Architecture Configuration are needed here also. The EPC needs the addition of a few new interfaces and functions to connect and inter-work with UTRAN and GERAN. The corresponding functions will also be required from GERAN and UTRAN. The new interfaces are S3, S4 and S12 as shown in Figure 3.12. The interface from SGSN to HSS can also be updated to Diameter based S6d, but the use of the legacy MAP based Gr is also possible.

Keeping E-UTRAN, i.e. the eNodeB design as focused to, and as optimized for the requirements of the new OFDMA radio interface, and as clean of inter-working functionality as possible, was an important guideline for the inter-working design. Consequently, the eNodeB does not interface directly with the other 3GPP ANs, and the interaction towards the EPC is the same as in other mobility cases that involve EPC. However, optimized inter-working means that the network is in control of mobility events, such as handovers, and provides functionality to hand the communication over with minimum interruption to services. This means that an eNodeB must be able to coordinate UE measuring UTRAN and GERAN cells, and perform handover decisions based on measurement results, and thus E-UTRAN radio interface protocols have been appended to support the corresponding new functions. Similar additions will be required from UTRAN and GERAN to support handover to E-UTRAN.

3.3.2 Additional and Updated Logical Elements in 3GPP Inter-working System Architecture Configuration

3.3.2.1 User Equipment

From the UE point of view, inter-working means that it needs to support the radio technologies in question, and the mobility operations defined for moving between them. The optimized inter-working means that the network controls the usage of radio transmitter and receiver in the UE in a way that only one set of them needs to be operating at the same time. This is called single radio operation, and allows UE implementations where only one pair of physical radio transmitter and receiver is implemented.

The standard does not preclude implementing multiple radio transmitters and receivers, and operating them simultaneously in dual radio operation. However, single radio operation is an important mode, because the different ANs often operate in frequencies that are so close to each other that dual radio operation would cause too much interference within the terminal. That, together with the additional power consumption, will decrease the overall performance.

3.3.2.2 E-UTRAN

The only addition to E-UTRAN eNodeB compared to the Basic System Architecture Configuration is the mobility to and from other 3GPP ANs. From the eNodeB perspective the functions are very similar irrespective of whether the other 3GPP AN is UTRAN or GERAN.

For the purpose of handover from E-UTRAN to UTRAN or GERAN, the neighbouring cells from those networks need to be configured into the eNodeB. The eNodeB may then consider handover for those UEs that indicate corresponding radio capability. The eNodeB requests the UE to measure the signal level of the UTRAN or GERAN cells, and analyses the measurement reports. If the eNodeB decides to start the handover, it signals the need to the MME in the

same way that it would signal inter-eNodeB handover when the X2 interface is not available. Subsequently, the eNodeB will receive the information needed for the Handover Command from the target Access System via the MME. The eNodeB will send the Handover Command to the UE without the need for interpreting the content of this information.

In the case of handover from UTRAN or GERAN to E-UTRAN, the eNodeB does not need to make any specific preparations compared to other handovers where the handover preparation request comes through the MME. The eNodeB will allocate the requested resources, and prepare the information for handover command, which it sends to the MME, from where it is delivered to the UE through the other 3GPP Access System that originated the handover.

3.3.2.3 UTRAN

In UTRAN, the radio control functionality is handled by the Radio Network Controller (RNC), and under its control the Node B performs Layer 2 bridging between the Uu and Iub interfaces. UTRAN functionality is described extensively in [16].

UTRAN has evolved from its initial introduction in Release 99 in many ways, including the evolution of architectural aspects. The first such item is Iu flex, where the RNC may be connected to many Serving GPRS Support Nodes (SGSNs) instead of just one. Another such concept is I-HSPA, where the essential set of packet data related RNC functions is included with the Node B, and that connects to Iu-PS as a single node. Figure 3.13 also shows the direct UP connection from RNC to S-GW, which is introduced to 3G CN by the Direct Tunnel concept, where the SGSN is bypassed in UP.

Inter-working with E-UTRAN requires that UTRAN performs the same measurement control and analysis functions as well as the transparent handover information delivery in Handover Command that were described for eNodeB in the earlier section. Also the UTRAN performs similar logic that it already uses with Relocation between RNCs, when the Iur interface is not used.

3.3.2.4 GERAN

GSM EDGE Radio AN (GERAN) is the evolved version of GSM AN, which can also be connected to 3G Core Network. It consists of the Base Station Controller (BSC) and the Base Station (BS), and the radio interface functionalities are divided between them. An overview of GERAN functionality and the whole GSM system can be found in [17].

The GERAN is always connected to the SGSN in both Control and UPs, and this connection is used for all the inter-working functionality. Also the GERAN uses logic similar to that described above for E-UTRAN and UTRAN for inter-working handover.

3.3.2.5 EPC

The EPC has a central role for the inter-working system architecture by anchoring the ANs together. In addition to what has been described earlier, the MME and S-GW will support connectivity and functions for inter-working. Also the SGSN, which supports the UTRAN and GERAN access networks, will need to support these functions, and when these additions are supported, it can be considered to belong to the EPC.

The S-GW is the mobility anchor for all 3GPP access systems. In the basic bearer operations and mobility between SGSNs, it behaves like a GGSN towards the SGSN, and also towards the RNC if UP tunnels are set up in Direct Tunnel fashion bypassing the SGSN. Many of the GGSN functions are actually performed in the P-GW, but this is not visible to the SGSN. The S-GW retains its role as a UP Gateway, which is controlled by either the MME or the SGSN depending on which AN the UE is being served by.

To support the inter-working mobility, the MME will need to signal with the SGSN. These operations are essentially the same as between those two MMEs, and have been described earlier in section 3.2. An additional aspect of the MME is that it may need to combine the change of S-GW and the inter-working mobility with SGSN.

The SGSN maintains its role as the controlling node in core network for both UTRAN and GERAN. These functions are defined in [18]. The SGSN has a role very similar to that of the MME. The SGSN needs to be updated to support for S-GW change during mobility between SGSNs or RNCs, because from the legacy SGSN point of view this case looks like GGSN changing, which is not supported. As discussed earlier, the SGSN may direct the UP to be routed directly between the S-GW and UTRAN RNC, or it may remain involved in the UP handling. From the S-GW point of view this does not really make a difference, since it does not need to know which type of node terminates the far end of the UP tunnel.

3.3.3 Interfaces and Protocols in 3GPP Inter-working System Architecture Configuration

Table 3.2 summarizes the interfaces in the 3GPP Inter-working System Architecture Configuration and the protocols used in them. Interfaces and protocols in legacy 3GPP networks are not listed. Interfaces and protocols listed for Basic System Architecture Configuration are needed in addition to these.

3.3.4 Inter-working with Legacy 3GPP CS Infrastructure

While the EPS is purely a Packet Switched (PS) only system without a specific Circuit Switched (CS) domain with support for VoIP, the legacy 3GPP systems treat CS services such as voice calls with a specific CS infrastructure. IMS VoIP may not be ubiquitously available, and therefore the SAE design includes two special solutions that address inter-working with circuit switched voice. A description of how inter-working between E-UTRAN and the legacy 3GPP CS domain can be

Table 3.2 Summary of additional interfaces and protocols in 3GPP Inter-working System Architecture configuration

Interface	Protocols	Specification
S3	GTP-C/UDP/IP	29.274 [8]
S4	GTP/UDP/IP	29.274 [8]
S12	GTP-U/UDP/IP	29.274 [8]
S16	GTP/UDP/IP	29.274 [8]
S6d	Diameter/SCTP/IP	29.272 [12]

Table 3.3 Summary of additional interfaces and protocols for inter-working with legacy 3GPP CS infrastructure

Interface	Protocols	Specification
SGs	SGsAP/SCTP/IP	29.118 [21]
Sv	GTP-C(subset)/UDP/IP	29.280 [22]

arranged is given in Chapter 10 on VoIP. Two specific functions have been defined for that purpose, Circuit Switched Fall Back (CSFB) and Single Radio Voice Call Continuity (SR-VCC).

CSFB [19] is a solution for networks that do not have support for IMS VoIP. Instead, the voice calls are handled by the CS domain, and the UE is handed over there at the time of a voice call. The SGs interface between the MME and MSC Server is used for related control signalling, as shown with more detail in Chapter 10.

SR-VCC [20] is a solution for converting and handing over an IMS VoIP call to a CS voice call in the legacy CS domain. This functionality would be needed when the coverage of an IMS VoIP capable network is smaller than that of the legacy CS networks. SR-VCC allows a UE entering the edge of the VoIP coverage area with an ongoing VoIP call to be handed over to the CS network without interrupting the call. SR-VCC is a one way handover from the PS network with VoIP to the CS network. If E-UTRAN coverage becomes available again, the UE may return there when the call ends and the UE becomes idle. The solution relies on running only one radio at a time, i.e. the UE does not need to communicate simultaneously with both systems. In this solution the MME is connected to the MSC Server in the CS domain via a Sv interface, which is used for control signalling in the SR-VCC handover. The details of the solution are presented in Chapter 10. A summary of additional interfaces and protocols for inter-working with legacy 3GPP CS infrastructure is given in Table 3.3.

3.4 System Architecture with E-UTRAN and Non-3GPP Access Networks

3.4.1 Overview of 3GPP and Non-3GPP Inter-working System Architecture Configuration

Inter-working with non-3GPP ANs was one of the key design goals for SAE, and to support it, a completely separate architecture specification [2] was developed in 3GPP. The non-3GPP Inter-working System Architecture includes a set of solutions in two categories. The first category contains a set of generic and loose inter-working solutions that can be used with any other non-3GPP AN. Mobility solutions defined in this category are also called Handovers without Optimizations, and the same procedures are applicable in both connected and idle mode. The second category includes a specific and tighter inter-working solution with one selected AN, the cdma2000® HRPD. This solution category is also called Handovers with Optimizations, and it specifies separate procedures for connected and idle mode.

The generic non-3GPP Inter-working System Architecture is shown in Figure 3.14. The specific application of the architecture for cdma2000® HRPD inter-working and the required additional interfaces are described with more detail in section 3.5.

Figure 3.14 describes the generic inter-working solution that relies only on loose coupling with generic interfacing means, and without AN level interfaces. Since there are so many different kinds of ANs, they have been categorized to two groups, the trusted and un-trusted non-3GPP ANs, depending on whether it can be safely assumed that 3GPP defined authentication can be run by the network, which makes it trusted, or if authentication has to be done in overlay fashion and the AN is un-trusted. The P-GW will maintain the role of mobility anchor, and the non-3GPP ANs are connected to it either via the S2a or the S2b interface, depending on whether the non-3GPP AN functions as a Trusted or Un-trusted non-3GPP AN. Both use network controlled IP layer mobility with the PMIP protocol. For networks that do not support PMIP, Client MIPv4 Foreign Agent mode is available as an option in S2a. In addition to mobility functions, the architecture includes interfaces for authenticating the UE within and through the non-3GPP ANs, and also allows PCC functionality in them via the Gxa and Gxb interfaces. Note that the detailed functions and protocols for Gxb are not specified in Release 8.

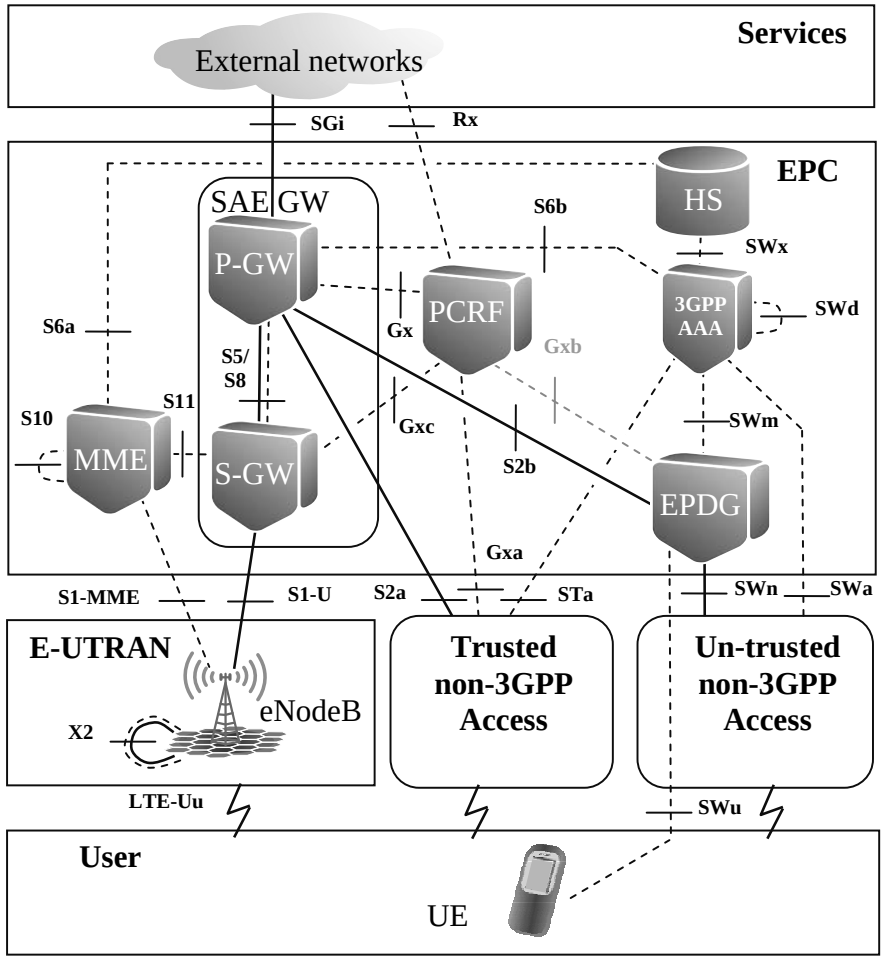


Figure 3.14 System architecture for 3GPP and non-3GPP access networks

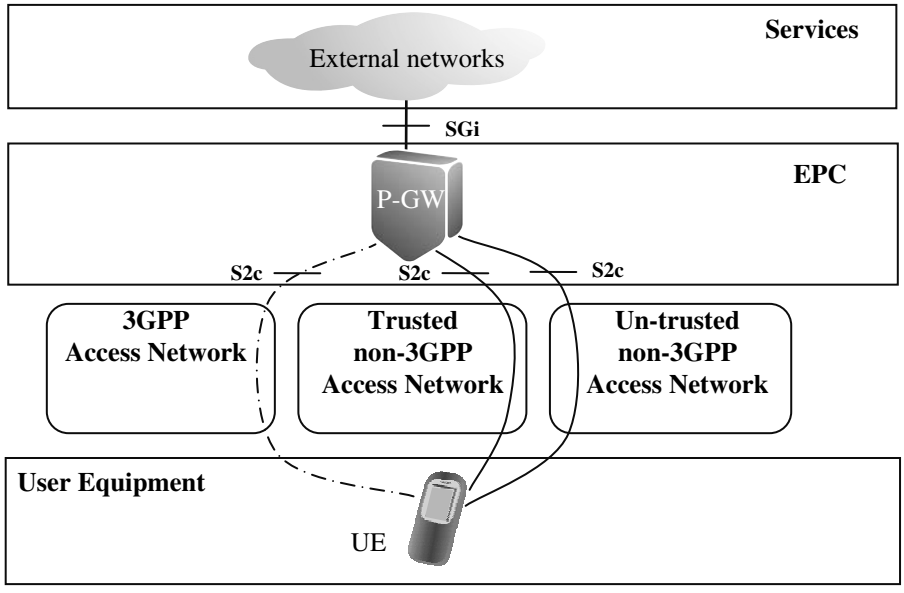


Figure 3.15 Simplified system architecture showing only S2c

In addition to the network controlled mobility solutions, a completely UE centric solution with DSMIPv6 is also included in the inter-working solutions. This scenario is depicted in Figure 3.15.

In this configuration the UE may register in any non-3GPP AN, receive an IP address from there, and register that to the Home Agent in P-GW. This solution addresses the mobility as an overlay function. While the UE is served by one of the 3GPP ANs, the UE is considered to be in home link, and thus the overhead caused by additional MIP headers is avoided.

Another inter-working scenario that brings additional flexibility is called the chained S8 and S2a/S2b scenario. In that scenario the non-3GPP AN is connected to S-GW in the visited Public Land Mobile Network (PLMN) through the S2a or S2b interface, while the P-GW is in the home PLMN. This enables the visited network to offer a roaming subscriber the use of non-3GPP ANs that might not be associated with the home operator at all, even in the case where P-GW is in the home PLMN. This scenario requires that S-GW performs functions that normally belong to P-GW in order to behave as the termination point for the S2a or S2b interfaces. In Release 8, this scenario does not support dynamic policies through the PCC infrastructure, i.e. the Gxc interface will not be used. Also, chaining with GTP based S5/S8 is not supported. All other interfaces related to non-3GPP ANs are used normally as shown in Figure 3.14.

3.4.2 Additional and Updated Logical Elements in 3GPP Inter-working System Architecture Configuration

3.4.2.1 User Equipment

Inter-working between the non-3GPP ANs requires that the UE supports the corresponding radio technologies, and the specified mobility procedures. The mobility procedures and required

radio capabilities vary depending on whether optimizations are in place or not. The procedures defined for Handovers without Optimizations do not make any assumption about the UE's capability to use the radio transmitters and receivers simultaneously, and both single radio and dual radio configurations can use the procedures. However, the handover gap time is expected to be shorter, if preparing the connections towards the target side can start already while data are still flowing through the source side. This is caused by the fact that Handovers without Optimizations do not have procedures in the network side to assist in handover preparations, and the procedures follow the principle where UE registers to the target network according to the method defined for that network, and then the network switches the flow to the target network. This may be time consuming, since it normally includes procedures such as authentication. Also, the decision to make these handovers is the responsibility of the UE.

The Handovers with Optimizations, i.e. inter-working with cdma2000® HRPD, assume that they do include network control for connected mode, so the handovers are decided by the network, while the idle mode mobility relies on UE decision making, which may use cdma2000® HRPD related information in the LTE-Uu broadcast. Furthermore, the procedures are designed with the assumption that single radio configuration is enough for the UE.

3.4.2.2 Trusted Non-3GPP Access Networks

The term trusted non-3GPP AN refers to networks that can be trusted to run 3GPP defined authentication. 3GPP Release 8 security architecture specification for non-3GPP ANs [23] mandates that the Improved Extensible Authentication Protocol Method for 3rd Generation Authentication and Key Agreement (EAP-AKA') [24] is performed. The related procedures are performed over the STa interface.

The trusted non-3GPP ANs are typically other mobile networks, such as the cdma2000® HRPD. The STa interface supports also delivery of subscription profile information from Authentication, Authorization and Accounting (AAA)/HSS to the AN, and charging information from the AN to AAA Server, which are typical functions needed in mobile networks. It can also be assumed that such ANs may benefit from connecting to the PCC infrastructure, and therefore the Gxc interface may be used to exchange related information with the PCRF.

The trusted non-3GPP AN connects to the P-GW with the S2a interface, with either PMIP or MIPv4 Foreign Agent mode. The switching of UP flows in P-GW is therefore the responsibility of the trusted non-3GPP AN when UE moves into the AN's service area.

3.4.2.3 Un-trusted Non-3GPP Access Networks

To a large extent, the architectural concepts that apply for un-trusted non-3GPP ANs are inherited from the Wireless Local Area Network Inter-Working (WLAN IW) defined originally in Release 6 [25]. The Release 8 functionality for connecting un-trusted non-3GPP ANs to EPC is specified fully in [2] with references to the earlier WLAN IW specifications when applicable.

The main principle is that the AN is not assumed to perform any other functions besides delivery of packets. A secure tunnel is established between UE and a special node called the Enhanced Packet Data Gateway (EPDG) via the SWu interface, and the data delivery takes place through that tunnel. Furthermore, the P-GW has a trust relationship with the EPDG con-

nected to it via the S2b interface, and neither node needs to have secure association with the un-trusted non-3GPP AN itself.

As an optional feature, the un-trusted non-3GPP AN may be connected to the AAA Server with the SWa interface, and this interface may be used to authenticate the UE already in the non-3GPP AN level. This can be done only in addition to authentication and authorization with the EPDG.

3.4.2.4 EPC

The EPC includes quite a few additional functions for the support of non-3GPP ANs, when compared to the previously introduced architecture configurations. The main changes are in the P-GW, PCRF and HSS, and also in S-GW for the chained S8 and S2a/S2b scenario. In addition, completely new elements, such as the EPDG (Evolved Packet Data Gateway) and the AAA are introduced. The AAA infrastructure contains the AAA Server, and it may also contain separate AAA proxies in roaming situations. Figure 3.16 highlights the AAA connections and functions for non-3GPP ANs.

The P-GW is the mobility anchor for the non-3GPP ANs. For PMIP based S2a and S2b interfaces, the P-GW hosts the Local Mobility Anchor (LMA) function in a manner similar to that for the S5/S8 PMIP interfaces. Also the Home Agent (HA) function for the Client MIPv4 Foreign Agent mode in S2a is located in P-GW. The relation between P-GWs and non-3GPP ANs is many to many. The P-GW will also interface with the AAA Server, which subsequently connects to HSS. This interface is used for reporting the selected P-GW to the HSS so that it is available in mobility between non-3GPP ANs, and to authenticate and authorize users connecting with S2c mode. Each P-GW may connect to more than one AAA server.

The PCRF supports PCC interfaces for non-3GPP ANs. The Gxa is used towards trusted non-3GPP ANs, and Gxb towards un-trusted non-3GPP ANs. Only Gxa is specified in detail level

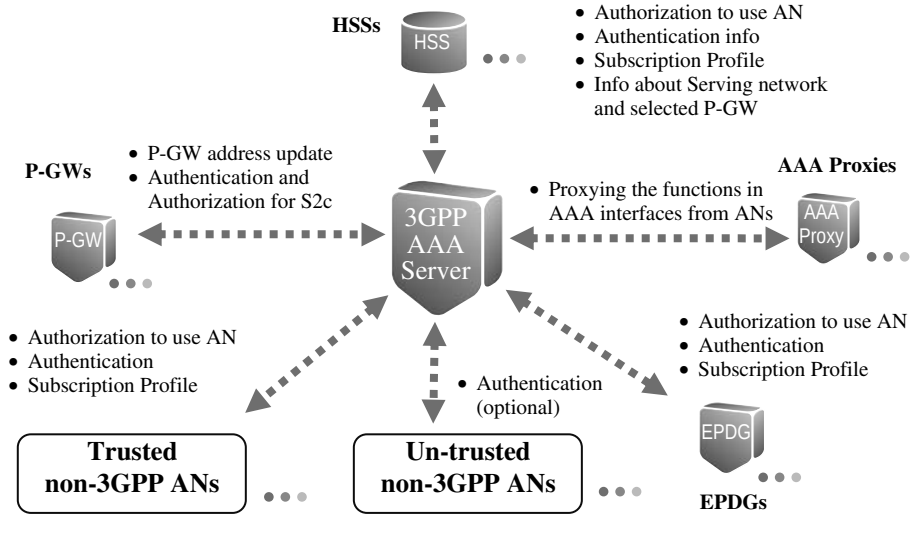


Figure 3.16 3GPP AAA server interfaces and main functions

in Release 8. The Gxa interface functions in a fashion similar to that of the Gxc interface. In this case the BBERF function will be in the non-3GPP AN, and it will receive instructions from the PCRF on how to handle the bearer level functions for the IP flows in the S2a interface. The further bearer functions internal to non-3GPP ANs are not addressed in 3GPP specifications.

The EPDG is a dedicated node for controlling the UE and inter-network connection, when an un-trusted non-3GPP AN is connected to EPC. Since the AN is not trusted, the main function is to secure the connection, as defined in [23]. The EPDG establishes an IPsec tunnel to the UE through the un-trusted non-3GPP AN with IKEv2 signalling [26] over the SWu interface. During the same signalling transaction the EAP-AKA authentication is run, and for that the EPDG signals with the AAA Server through the SWm interface. While the SWm interface is logically between UE and the EPDG, the SWn interface represents the interface on a lower layer between the EPDG and the un-trusted non-3GPP AN. The Release 8 specifications do not assume that EPDG would signal with PCRF for any PCC functions, but the architecture already contains the Gxb interface for that purpose.

The 3GPP AAA Server, and possibly a AAA Proxy in the visited network, performs a 3GPP defined set of AAA functions. These functions are a subset of what the standard IETF defined AAA infrastructure includes, and do not necessarily map with the way other networks use AAA infrastructure. The AAA Server acts between the ANs and the HSS, and in doing so it creates a context for the UEs it serves, and may store some of their information for further use. Thus, the 3GPP AAA Server consolidates the signalling from different types of ANs into a single SWx interface towards the HSS, and terminates the access specific interfaces S6b, STa, SWm and SWa. Most importantly the AAA Server performs as the authenticator for the EAP-AKA authentication through the non-3GPP ANs. It checks the authenticity of the user, and informs the AN about the outcome. The authorization to use the AN in question will also be performed during this step. Depending on the AN type in question, the AAA Server may also relay subscription profile information to the AN, which the AN may further use to better serve the UE. When the UE is no longer served by a given non-3GPP AN, the AAA Server participates in removing the UE's association from the HSS. Figure 3.16 summarizes the AAA Server main functions in relation to other nodes.

The HSS performs functions similar to those for the 3GPP ANs. It stores the main copy of the subscription profile as well as the secret security key in the AuC portion of it, and when requested, it provides the profile data and authentication vectors to be used in UEs connecting through non-3GPP ANs. One addition compared to 3GPP ANs is that since the non-3GPP ANs do not interface on the AN level, the selected P-GW needs to be stored in the HSS, and retrieved from there when the UE mobility involves a non-3GPP AN. The variety of different AN types are mostly hidden from the HSS, since the AAA Server terminates the interfaces that are specific to them, and HSS only sees a single SWx interface. On the other hand, the subscription profile stored in the HSS must reflect the needs of all the different types of ANs that are valid for that operator.

3.4.3 Interfaces and Protocols in Non-3GPP Inter-working System Architecture Configuration

Connecting the non-3GPP ANs to EPC and operating them with it requires additional interfaces to those introduced in earlier sections. Table 3.4 lists the new interfaces.

Table 3.4 Summary of additional interfaces and protocols in non-3GPP Inter-working System Architecture configuration

Interface	Protocols	Specification
S2a	PMIP/IP, or MIPv4/UDP/IP	29.275 [10]
S2b	PMIP/IP	29.275 [10]
S2c	DSMIPv6, IKEv2	24.303 [27]
S6b	Diameter/SCTP/IP	29.273 [28]
Gxa	Diameter/SCTP/IP	29.212 [13]
Gxb	*Not defined in Release 8*	*N.A.*
STa	Diameter/SCTP/IP	29.273 [28]
SWa	Diameter/SCTP/IP	29.273 [28]
SWd	Diameter/SCTP/IP	29.273 [28]
SWm	Diameter/SCTP/IP	29.273 [28]
SWn	PMIP	29.275 [10]
SWu	IKEv2, MOBIKE	24.302 [29]
SWx	Diameter/SCTP/IP	29.273 [28]
UE – foreign agent in trusted non-3GPP Access	MIPv4	24.304 [30]
UE – Trusted or Un-trusted non-3GPP access	EAP-AKA	24.302 [29]

3.4.4 Roaming in Non-3GPP Inter-working System Architecture Configuration

The principles for roaming with non-3GPP accesses are equivalent to those described in section 3.2.4 for 3GPP ANs. Both home routed and local breakout scenarios are supported and the main variations in the architecture relate to the PCC arrangement, which depends on where the services are consumed. This aspect is highlighted more in section 3.7.1.

The additional consideration that non-3GPP ANs bring to roaming is related to the case where the user is roaming to a visited 3GPP network in Home Routed model, and it would be beneficial to use a local non-3GPP AN that is affiliated with the visited network, but there is no association between that network and the home operator. For this scenario, the 3GPP Release 8 includes a so-called *chained case*, where the S-GW may behave as the anchor for the non-3GPP ANs also, i.e. it terminates the S2a or S2b interface, and routes the traffic via the S8 interface to the P-GW in the home network.

3.5 Inter-working with cdma2000® Access Networks

3.5.1 Architecture for cdma2000® HRPD Inter-working

The best inter-working performance in terms of handover gap time is achieved by specifying the networks to inter-operate very tightly to exchange critical information. This creates a specific solution that is valid for only the ANs in question. With the limited time and resources available for specification work, the number of such solutions in 3GPP Release 8 could only be limited. A tight inter-working solution also requires changes in the other ANs, and by definition the development of non-3GPP ANs is not within the control of 3GPP. Achieving a well designed solution requires special attention to coordination between the developments in

different standardization bodies. With these difficulties at hand, 3GPP Release 8 only includes an optimized inter-working solution with cdma2000® HRPD AN.

Figure 3.17 highlights the architecture for cdma2000® HRPD inter-working. It shows the Evolved HRPD (E-HRPD) network, where a number of modifications have been applied to make it suitable for connecting to the EPC. Due to these modifications it will be called E-HRPD in this chapter to distinguish it from legacy HRPD systems that do not support these functions. The radio interface and the Radio Access Network have been kept as similar as possible, but the HRPD Serving Gateway (HSGW) is a completely new node inheriting many of its functions from S-GW.

The E-HRPD is generally treated as a trusted non-3GPP AN, and it is therefore connected to the EPC via S2a, Gxa and STa interfaces. These interfaces operate as described earlier. Since the inter-working solution is optimized, and does not rely on UE performing the attach

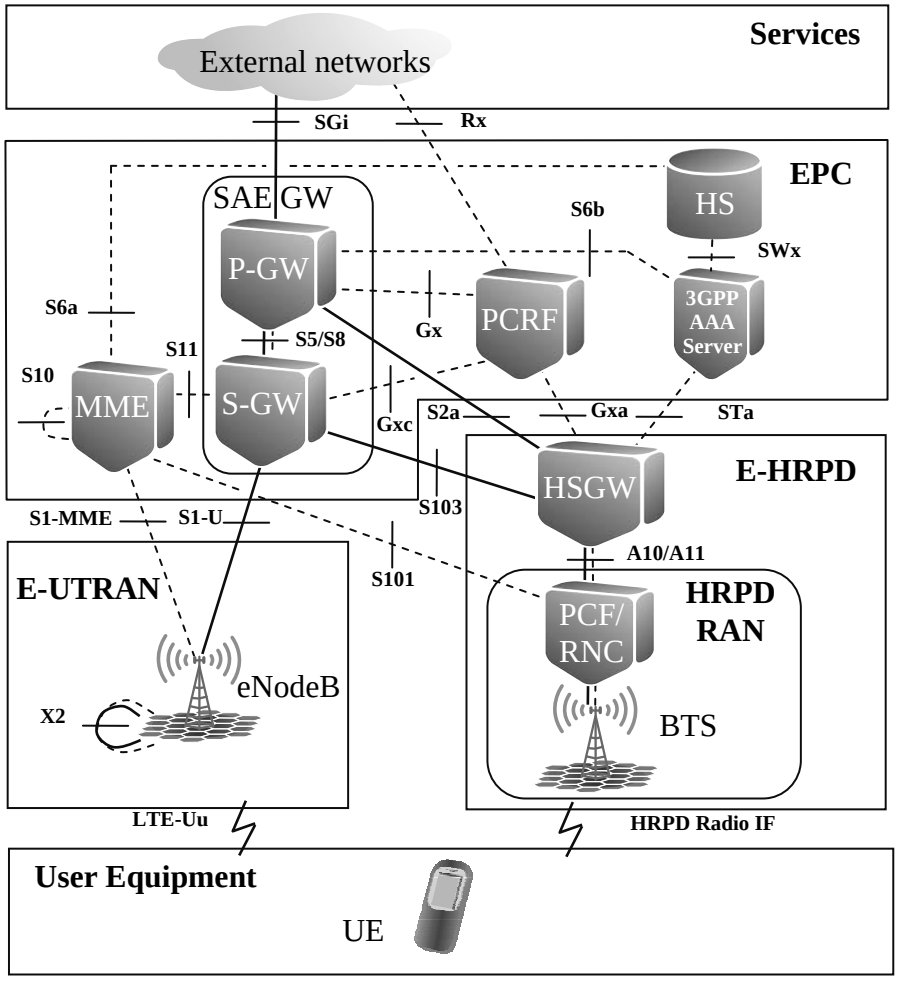


Figure 3.17 System architecture for 3GPP and cdma2000® HRPD inter-working

procedure directly to the target network, two new interfaces, S101 and S103, were defined for the CP and UP interactions respectively.

3GPP ANs and the 3GPP2-defined cdma2000® ANs share many things in common, but many things are also different. Both systems use a prepared handover, where the source system signals to the target system to give it essential parameters to be able to serve the terminal there, and the target system gives the source system parameters that can be further given to the terminal to guide it to make the access to the target radio. While there are similarities in these methods, the parameters themselves do not match well at all, and this method could not be used by applying a simple protocol conversion. To ease up on the need to align every information element that would need to be exchanged in handover, it was decided to use a transparent signalling transport method.

Figure 3.18 shows how the transparent tunnel is used in mobility from E-UTRAN to E-HRPD on the left, and the opposite direction is shown on the right. The thick black arrow indicates the signalling which is carried transparently through the source access system and over the S101 interface to the target system. In this method the source access system gives guidance to the UE to register to the target system through the tunnel. This creates the UE context in the target system without the source system having to convert its information to the target system format. This is called pre-registration, and the purpose is to take the time consuming registration/attach function away from the time critical path of handover. The transparent tunnel may also be used to build the bearer context in the target system so that when the time to make the handover is at hand, everything will be ready and waiting at the target side. The actual handover is decided based on radio interface conditions, and this solution requires that both systems are able to handle measurements from the other system. The following inter-working scenarios are supported between E-UTRAN and E-HRPD:

- E-UTRAN → E-HRPD handover: The pre-registration may be performed well before the actual handover takes place, and also all bearers are set up in the E-HRPD side. The UE remains in a dormant state (equal to idle mode) from the E-HRPD system point of view before handover, and this state may be long lived. When the radio conditions indicate the need for handover, the eNodeB commands the UE to start requesting traffic channel from E-HRPD. This takes place through the transparent tunnel, and once it is completed, the eNodeB commands the UE to make the actual handover. The S103 interface is used only

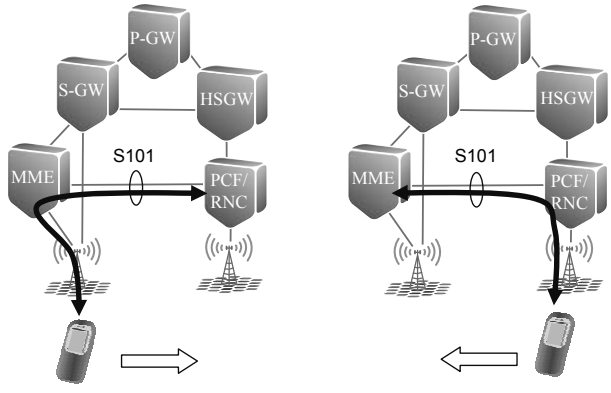


Figure 3.18 Tunnelled pre-registration to eHRPD and to E-UTRAN

in this handover scenario to forward DL data during the time when the UE is making the switch between the radios.

- E-UTRAN → E-HRPD idle mode mobility: The pre-registration state works as described above for the handover. The UE is in idle mode in the E-UTRAN also, and it moves within the system, selecting the cells on its own, and when it selects an E-HRPD cell, it briefly connects to the E-HRPD network to get the mobility pointers updated to the E-HRPD side.
- E-HRPD → E-UTRAN handover: The E-HRPD AN will request the UE to make tunnelled pre-registration (attach) only at the time the handover is needed, and the UE will immediately proceed to requesting connection directly from the E-UTRAN cell after the registration is complete. The bearers are set up in embedded fashion with the registration and connection request procedures.
- E-HRPD → E-UTRAN idle mode mobility: The idle mode procedure follows the same guidelines as the handover for the tunnelled registration (attach), but the UE accesses the E-UTRAN radio only by reporting its new location (Tracking Area), since there is no need to set up bearers in E-UTRAN for UE in idle mode.

3.5.2 Additional and Updated Logical Elements for cdma2000® HRPD Inter-working

Inter-working with eHRPD in an optimized manner brings a lot of new features in the basic SAE network elements, and introduces few totally new elements in the HRPD side. The UE, eNodeB, MME and S-GW will all be modified to support new functions, and MME and S-GW will also deal with new interfaces. The eHRPD is a new network of its own, and it consists of elements such as Base Station, Radio Network Controller (RNC), Packet Control Function (PCF) and HRPD Serving Gateway (HSGW).

The UE will need to support both radio interfaces. The design of the procedure assumes that UE is capable of single mode operation only. On the other hand, the integration is kept loose enough so that it would be possible to implement terminal with separate chip sets for E-UTRAN and E-HRPD. This means that the UE is not required to make measurements of cells in the other technology in as tightly a timewise controlled manner as is normally seen within a single radio technology. The UE will also need to support the tunnelled signalling operation. The tunnelled signalling itself is the same signalling as the UE would use directly with the other system.

The main new requirement for the eNodeB is that it also needs to be able to control mobility towards the eHRPD access. From the radio interface perspective it does this much in the same manner as with the other 3GPP accesses, by instructing the UE to make measurements of the neighbouring eHRPD cells, and making the handover decision based on this information. On the other hand, the eNodeB does not signal the handover preparation towards the eHRPD, like it would for other handovers in S1 interface. Instead the handover preparation works so that the UE sends traffic channel requests to the eHRPD AN through the transparent tunnel, and the eNodeB is only responsible for marking the uplink messages with appropriate routing information, so that the MME can select the right node in the eHRPD AN, and noting the progress of the handover from the headers of the S1 messages carrying the eHRPD signalling.

The MME implements the new S101 interface towards the eHRPD RAN. For UE originated messages, it needs to be able to route them to the right eHRPD RAN node based on a reference given by the eNodeB. In the reverse direction the messages are identified by the IMSI of

the UE, and the basis that the MME can route them to the right S1 signalling connection. The MME does not need to interpret the eHRPD signalling message contents, but the status of the HO progress is indicated along with those messages that require special action from the MME. For example, at a given point during E-UTRAN → E-HRPD handover, the MME will set up the data forwarding tunnels in the S-GW. The MME also needs to memorize the identity of the E-HRPD AN node that a UE has been signalling with, so that if MME change takes place, the MME can update the S101 context in the HRPD AN node.

The S-GW supports the new S103 interface, which is used for forwarding DL data during the time in handover, when the radio link cannot be used. The forwarding function is similar to the function S-GW has for the E-UTRAN handovers. The difference is that S103 is based on a GRE tunnel, and there will be only one tunnel for each UE in handover, so the S-GW needs to map all GTP tunnels from the S1-U interface to a single GRE tunnel in the S103 interface.

The E-HRPD network is a completely new way to use the existing HRPD radio technology with the SAE, by connecting it to the EPC. Compared to the original HRPD, many changes are caused by the inter-working, and connecting to the EPD requires some new functions, e.g. the support of EAP-AKA authentication. The HSGW is taking the S-GW role for E-HRPD access, and performs much like a S-GW towards the P-GW. The HSGW also includes many CP functions. Towards the eHRPD AN, it behaves like the Packet Data Serving Node (PDSN) in a legacy HRPD network. It also signals with the 3GPP AAA Server to authenticate the UE, and to receive its service profile. The CN aspects of the E-HRPD are specified in [31] and the evolved RAN is documented in [32].

3.5.3 Protocols and Interfaces in cdma2000® HRPD Inter-working

The optimized inter-working introduces two new interfaces – S101 and S103 – to the architecture (see Table 3.5). The S2a, Gxc and STa are as described earlier. The following summarizes the new interfaces:

- S101 is a CP interface that in principle forms a signalling tunnel for the eHRPD messages. The CP protocol is S101AP, which is specified in [33]. The S101AP uses the same message structure and coding as the newest version of GTP. The main function is to carry the signalling messages, with the IMSI as a reference and with an additional handover status parameter that is set by either the UE or either one of the networks it signals with. In addition, when the data forwarding tunnel needs to be set up, the address information is also included in S101AP. S101AP also includes a procedure to switch the interface from one MME to another if handover in E-UTRAN causes MME change.
- S103 is a simple GRE tunnel for UP data forwarding in handover. It is only used for DL data in handover from E-UTRAN to E-HRPD. S103 is a UP interface only, and all control information to set up the GRE tunnel is carried in other interfaces. It is specified with S101AP in [33].

Table 3.5 Additional interfaces and protocols for inter-working with cdma2000® eHRPD

Interface	Protocols	Specification
S101	S101AP/UDP/IP	29.276 [33]
S103	GRE/IP	29.276 [33]

Table 3.6 Additional interfaces and protocols for inter-working with cdma2000® 1xRTT

Interface	Protocols	Specification
S102	S102 protocol	29.277 [34]
A21	A21 protocol	A.S0008-C [35]

3.5.4 Inter-working with cdma2000® 1xRTT

The cdma2000® 1xRTT is a system supporting CS bearers, and is primarily used for voice calls. In this respect it is functionally equivalent to the legacy 3GPP CS infrastructure such as the MSC and the CS bearer capabilities of GERAN and UTRAN. As described in Chapter 10, the 3GPP standard includes two functions to support inter-working between the E-UTRAN and the legacy CS infrastructure. These are the CSFB [19] and SR-VCC [20]. These functions have been extended to cover inter-working with cdma2000® 1xRTT also, and at a high level they work in the same way as described in Chapter 10. In the 1xRTT case, the interface between MME and the cdma2000® 1xRTT infrastructure is called S102. S102 carries a protocol specified in 3GPP2 for the A21 interface, which is used in cdma2000® systems for voice call continuity (see Table 3.6).

3.6 IMS Architecture

3.6.1 Overview

The IP Multimedia Services Sub-System (IMS) is the preferred service machinery for LTE/SAE. IMS was first introduced in Release 5, and with the well defined inter-working with existing networks and services that have been introduced since, the Rel-8 IMS can now be used to provide services over fixed and wireless accesses alike. The IMS architecture is defined in [36], and the functionality is defined in [37]. A comprehensive description of IMS can also be found in [3]. For the purpose of this book, the functional architecture of IMS is presented in Figure 3.19, and a short description of the main functions follows below.

IMS is an overlay service layer on top of the IP connectivity layer that the EPS provides. Figure 3.19 shows a thick grey line from UE to P-GW that represents the UE's IP connectivity to IMS and other external networks through the RAN and EPC. The signalling interfaces Gm and Ut run on top of this connection, which typically use the default bearer a UE will always have in LTE/SAE. The services may further require that dedicated bearers are set up through EPC, and the service data flows may need to be handled by one of the Inter-working or Services Elements.

In principle the IMS is independent of the connectivity layer, which requires its own registration and session management procedures, but it has also been specifically designed to operate over the 3GPP-defined ANs, and it works seamlessly with the PCC described in section 3.7. IMS uses SIP protocol for registration and for controlling the service sessions. SIP is used both between the terminal and the IMS (Gm Interface) and between various IMS nodes (ISC, Mw, Mg, Mr, Mi, Mj, Mx, Mk and Mm Interfaces). The SIP usage in IMS is defined in [38]. Diameter (Cx, Dx, Dh and Sh Interfaces) and H.248 (Mp) are the other protocols used in IMS.

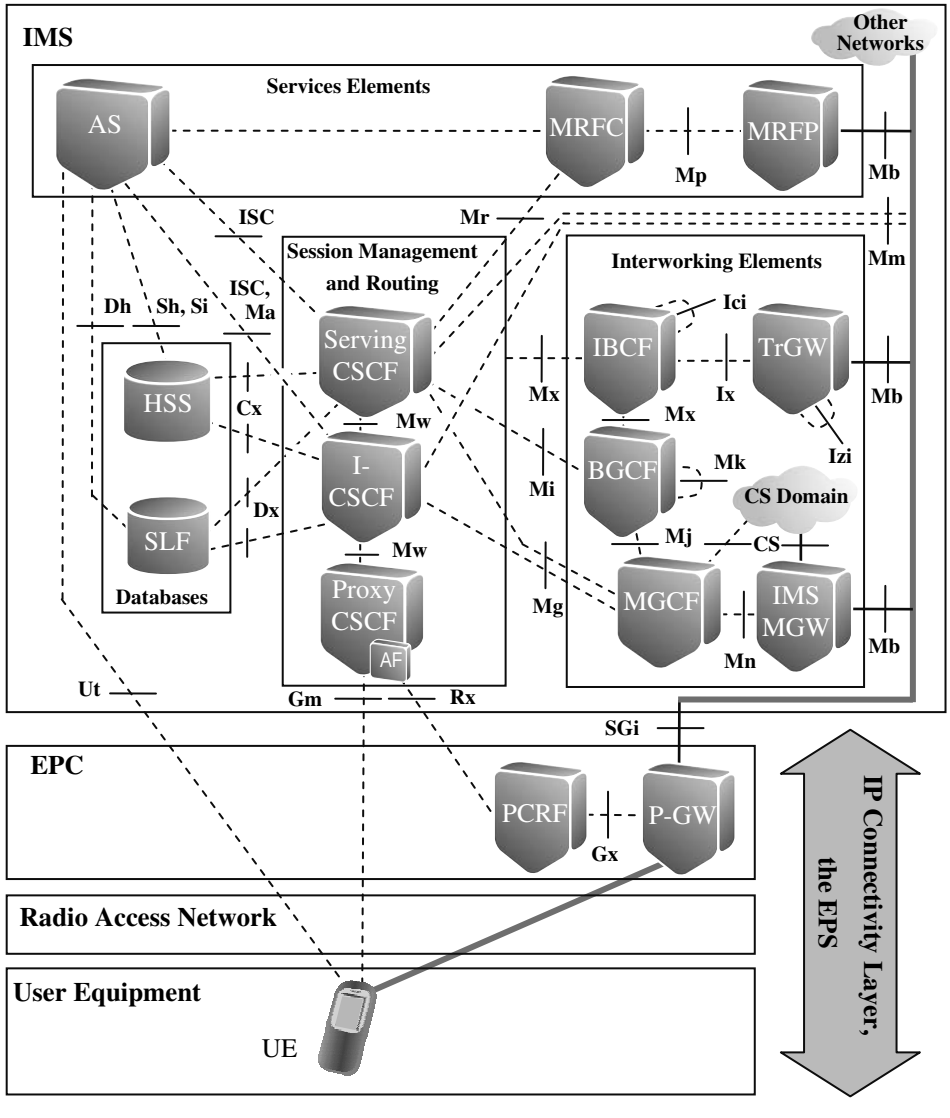


Figure 3.19 IMS architecture

The UE primarily signals with the CSCFs for the services it wishes to use, and in addition some service specific signalling may be run directly with the Application Servers. The signalling may also be network originated for terminating services. The *Session Management and Routing* functions are handled by the CSCFs that are in control of the UE's registration in IMS. For that purpose they signal with the *Databases* to get the appropriate information. The CSCFs are also in control of the UE's service sessions in IMS, and for that purpose they may need to signal with one or more of the *Services Elements* to know what kind of connectivity is needed for the service in question, and then with the connectivity layer through the Rx interface to make corresponding requests to bearer resources. Finally, the CSCFs may need to signal with

one or more of the *Inter-working Elements* to control the interconnection between networks. Whenever the UP flow is routed through one of the IMS elements, it is done through the Mb interface that connects IMS to IP networks. The following sections introduce the main functions in the functional groups highlighted in Figure 3.19.

Most IMS elements responsible for session management and routing or inter-working are involved in collecting charging information. Rf and Ro interfaces are the main IMS charging interfaces (see section 3.7.1). For simplicity, charging related nodes and interfaces are not shown in Figure 3.19.

3.6.2 Session Management and Routing

The Call State Control Function (CSCF) is the central element in SIP signalling between the UE and the IMS, and it takes care of the UE's registration to the IMS, and service session management. The registration includes authentication. The primary authentication method is IMS-AKA [39], but other methods such as http digest [40] may also be used. CSCF is defined to have three different roles that may reside in the same node, or separate nodes connected through the Mw interface, and all are involved in the UE-related SIP signalling transactions:

- The Serving CSCF (S-CSCF) locates in the user's home network, and it will maintain the user's registration and session state. At registration, it interfaces with the HSS to receive the subscription profile, including authentication information, and it will authenticate the UE. For the service sessions, the S-CSCF signals with the UE through the other CSCFs, and may also interact with the Application Servers (ASs) or the MRFCs for setting up the service session properly. It also carries the main responsibility for controlling the Inter-working Elements. The S-CSCF may also need to interact with MGCF for inter-working with CS networks, or with other multimedia networks for UE requested services.
- The Interrogating CSCF (I-SCSF) is located at the border of the home network, and it is responsible for finding out the UE's registration status, and either assigning a new S-CSCF or routing to the right existing S-CSCF. The request may come from Proxy CSCF (P-CSCF), from other multimedia networks, or from CS networks through the Media Gateway Control Function (MGCF). Also I-CSCF may need to interact with the ASs for service handling. The Ma interface is used for this when Public Service Identity (PSI) is used to identify the service, and the I-CSCF can route the request directly to the proper AS.
- The (P-CSCF) is the closest IMS node the UE interacts with, and it is responsible for all functions related to controlling the IP connectivity layer, i.e. the EPS. For this purpose the P-CSCF contains the Application Function (AF) that is a logical element for the PCC concept, which is described in section 3.7.1. The P-CSCF is typically located in the same network as the EPS, but the Rel-8 includes a so-called Local Breakout concept that allows P-CSCF to remain in the home network, while PCRF in the visited network may still be used.

In addition to the above-mentioned three CSCF roles, a fourth role, the Emergency CSCF (E-CSCF), has been defined. As the name indicates, the E-CSCF is dedicated to handling the emergency call service in IMS. The E-CSCF connects to the P-CSCF via the Mw interface, and these nodes must always be in the same network. In addition, the E-CSCF is also connected to a Location Retrieval Function (LRF) through the Mi Interface. The LRF can provide

the location of the UE, and routing information to route the emergency call appropriately. The E-CSCF and LRF are not shown in Figure 3.19 for simplicity

The CSCFs are connected to each other with the Mw interface, and to other multimedia networks through the Mm interface. Interconnection between CSCFs in different operators' networks may be routed through a common point called the Interconnection Border Control Function (IBCF). See section 3.6.5.

3.6.3 Databases

The Home Subscriber Server (HSS) is the main database used by the IMS. The HSS contains the master copy of subscription data, and it is used in much the same way as with the IP connectivity layer. It provides the location and authentication information based on requests from the I- or S-CSCF, or the AS. The interface between the HSS and the Services Elements will be either Sh or Si depending on the type of services elements. The Sh interface is used in case of SIP or OSA service capability server and the Si when CAMEL based AS is in question.

When there is more than one addressable HSS, another database called the Subscription Locator Function (SLF) may be used to find the right HSS.

3.6.4 Services Elements

The actual service logic is located in the Application Servers (AS). A variety of different services may be provided with different ASs, and the standards do not aim to cover all possible services. Some of the main services are covered in order to facilitate easier inter-working with operators in roaming, and to provide for consistent user experience. One example of a standardized AS is the Telephony Application Server (TAS), which may be used to provide the IMS VoIP service.

The media component of the service can be handled by the Multimedia Resource Function (MRF), which is defined as a separate controller (MRFC) and processor (MRFP). The UP may be routed through MRFP for playing announcements as well as for conferencing and transcoding. For coordination purposes, the MRFC may also be connected to the related AS.

3.6.5 Inter-working Elements

The Inter-working Elements are needed when the IMS interoperates with other networks, such as other IMS networks, or CS networks. The following are the main functions of the standardized inter-working elements:

- The Breakout Gateway Control Function (BGCF) is used when inter-working with CS networks is needed, and it is responsible for selecting where the interconnection will take place. It may select the Media Gateway Control Function (MGCF) if the breakout is to happen in the same network, or it may forward the request to another BGCF in another network. This interaction may be routed through the Interconnection Border Control Function (IBCF).
- The Interconnection Border Control Function (IBCF) is used when interconnection between operators is desired to be routed through defined points, which hide the topology inside the network. The IBCF may be used in interconnection between CSCFs or BGCFs and it is in control of Transition Gateway (TrGW), which is used for the same function in the UP.

Note that the IBCF–TrGW interface is not fully specified in Release 8. The IBCFs and the TrGWs in different operators' networks may be interconnected to each other via the Ici and Izi interfaces respectively, and together they comprise the Inter IMS Network to Network Interface (II-NNI).

- The Media Gateway Control Function (MGCF) and IMS-Media Gateway (IMS-MGW) are the CP and UP nodes for inter-working with the CS networks such as the legacy 3GPP networks with CS domain for GERAN or UTRAN, or for PSTN/ISDN. Both incoming and outgoing IMS VoIP calls are supported with the required signalling inter-working and transcoding between different voice coding schemes. The MGCF works in the control of either the CSCF or BGCF.

3.7 PCC and QoS

3.7.1 PCC

Policy and Charging Control (PCC) has a key role in the way users' services are handled in the Release 8 LTE/SAE system. It provides a way to manage the service related connections in a consistent and controlled way. It determines how bearer resources are allocated for a given service, including how the service flows are partitioned to bearers, what QoS characteristics those bearers will have, and finally, what kind of accounting and charging will be applied. If an operator uses only a very simple QoS model, then a static configuration of these parameters may be sufficient, but Release 8 PCC allows the operator to set these parameters dynamically for each service and even each user separately.

The PCC functions are defined in [5] and the PCC signalling transactions as well as the QoS parameter mapping are defined in [41]. Figure 3.20 shows the PCC functions and interfaces in the basic configuration when PCC is applied in one operator's network.

The primary way to set up service flows in Release 8 is one where the UE first signals the request for the service in the Service Layer, and the Application Function (AF) residing in that layer contacts the Policy and Charging Resource Function (PCRF) for appropriate bearer

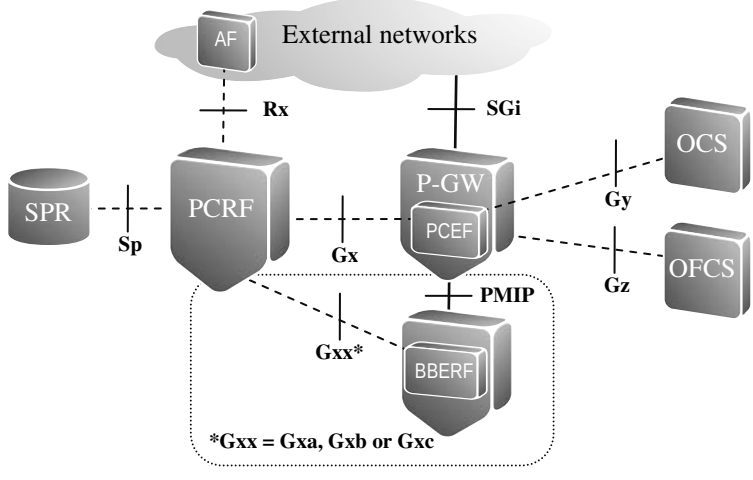


Figure 3.20 Basic PCC functions

resources. The PCRF is in charge for making the decisions on what PCC to use for the service in question. If subscriber specific policies are used, then the PCRF may enquire subscription related policies from the Subscription Profile Repository (SPR). Further details about SPR structure and, for example, its relation to HSS, and the Sp interface are not specified in Release 8. Based on the decision, the PCRF creates the appropriate PCC rules that determine the handling in the EPS.

If the interface from P-GW to the S-GW is based on GTP, the PCRF pushes the PCC rules to the Policy and Charging Enforcement Function (PCEF) residing in the P-GW, and it alone will be responsible for enforcing the PCC rules, e.g. setting up the corresponding dedicated bearers, or modifying the existing bearers so that the new IP service flows can be mapped to them, and by ensuring that only authorized service flows are allowed and QoS limits are not exceeded. In this case the Gxx interface shown in Figure 3.20 does not apply.

If the interface from P-GW towards the AN is based on PMIP, i.e. if it is S5 PMIP, S2a or S2b, there is no means to signal the bearer level information onwards from the P-GW, and the PCRF will create a separate set of QoS rules, and those are first sent to the BBERF, which will handle the mapping between IP service flows and bearers over the AN. Depending on the AN type, the BBERF may reside in S-GW (S5 PMIP), trusted non-3GPP AN, e.g. in HSGW (S2a), or in the EPDG (S2b) for the un-trusted non-3GPP AN (S2b is not supported in Release 8). Also in this case the PCC rules are also sent to PCEF in the P-GW, and it performs the service flow and QoS enforcement.

Release 8 also supports UE initiated bearer activation within the EPS, which is applicable to the case when there is no defined service that both the UE and the serving network could address. In this case the UE signals with the AN and the BBERF requests the service resources from the PCRF. The PCRF makes the PCC decision, and the logic then continues as described above.

The PCC standard [5] defines two charging interfaces, Gy and Gz, which are used for online and offline charging respectively. The Gy interface connects the PCEF to the Online Charging System (OCS), which is used for flow based charging information transfer and control in an online fashion. The Gz interface is used between the P-GW and the Offline Charging System (OFCS), and it is applied when charging records are consolidated in an offline fashion. The charging specifications [42] and [43] further define that the Gy interface is functionally equivalent to the Ro interface that uses Diameter Credit-Control Application as defined in [44]. The Gz interface may be based on either the Rf interface, which relies on the mentioned Diameter Credit-Control Application, or the Ga interface, which uses the 3GPP defined GTP protocol. The Ro and Rf interfaces are also used for charging in IMS, and were originally specified for that purpose.

The PCRF in control of the PCEF/P-GW and the BBERF typically reside in the same operator's network. In the case of roaming, they may reside in different networks, and the S9 interface between PCRFs is used to enable the use of a local PCRF. The S9 interface is defined in [9], and it re-uses the applicable parts from Rx, Gx and Gxx interfaces to convey the information between the PCRFs.

There are two different cases when the S9 interface is used. The first case, which is shown in Figure 3.21, applies when the roaming interface is based on PMIP, and the PCEF and BBERF are in different networks. In this scenario traffic is routed to the home network. In the second case, shown in Figure 3.22, the Local Breakout model is applied, and the P-GW resides in the visited network. The AF and Rx interface will be used from the same network that provides the service in question. The OCS will reside in the home network. As described

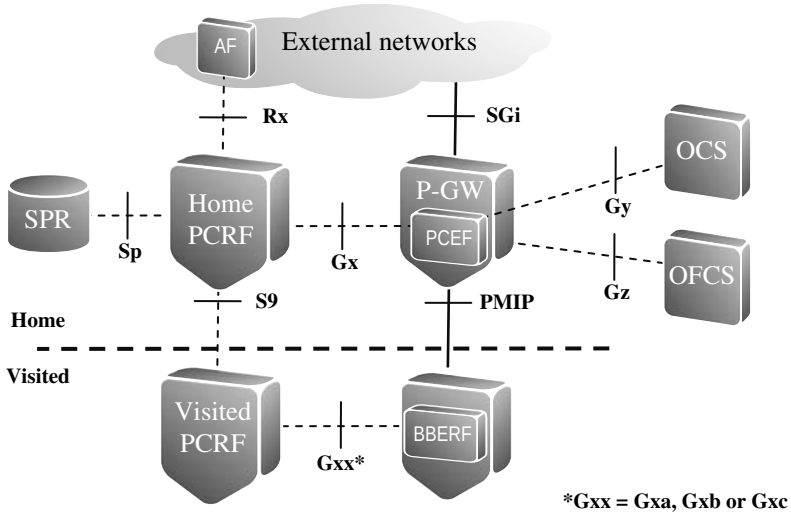


Figure 3.21 PCC functions in roaming with PMIP, home routed model

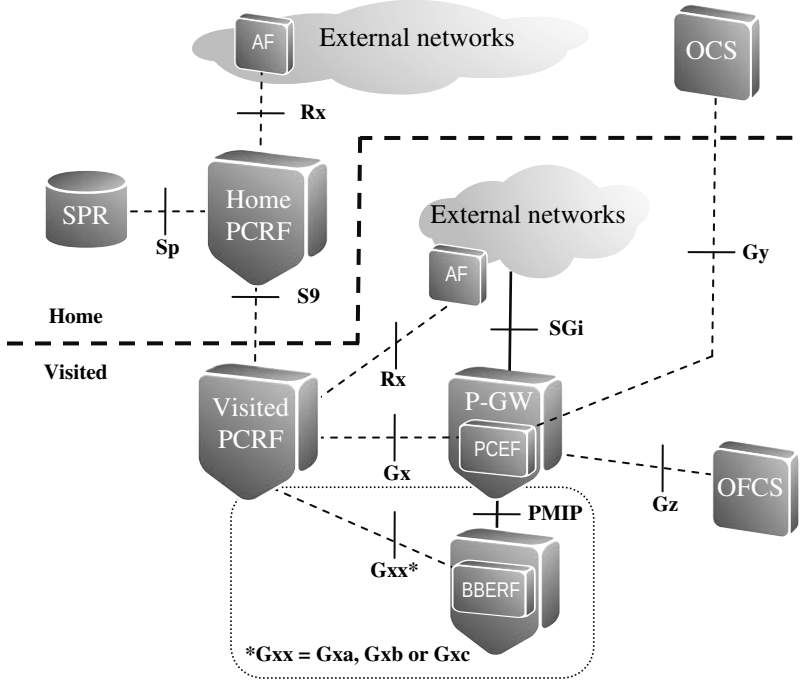


Figure 3.22 PCC functions in roaming, local breakout model

above, the separate BBERF and Gxx interfaces apply only if PMIP is used from the P-GW in the visited network.

Table 3.7 lists the PCC related interfaces and the protocols, and the standards where they are specified.

Table 3.7 Summary of PCC interfaces

Interface	Protocols	Specification
Gx	Diameter/SCTP/IP	29.212 [13]
Gxx (Gxa or Gxc)	Diameter/SCTP/IP	29.212 [13]
Rx	Diameter/SCTP/IP	29.214 [14]
S9	Diameter/SCTP/IP	29.215 [45]
Sp	*Not defined in Release 8*	*N.A.*
Gy =		32.240 [42]
Ro	Diameter/SCTP/IP	32.299 [46]
Gz =		32.251 [43]
Rf or	Diameter/SCTP/IP or	32.295 [47] or
Ga	GTP'/UDP or TCP/IP	32.299 [46]

3.7.2 QoS

The development of the SAE bearer model and the QoS concept started with the assumption that improvements compared to the existing 3GPP systems with, e.g. UTRAN access, should be made, and the existing model should not be taken for granted. Some potential areas had already been identified. It had not been easy for the operators to use QoS in the legacy 3GPP systems. An extensive set of QoS attributes was available, but it was to some extent disconnected from the application layer, and thus it had not been easy to configure the attributes in the correct way. This problem was emphasized by the fact that the UE was responsible for setting the QoS attributes for a bearer. Also, the bearer model had many layers, each signalling just about the same information. It was therefore agreed that for SAE, only a reduced set of QoS parameters and standardized characteristics would be specified. Also it was decided to turn the bearer set-up logic so that the network resource management is solely network controlled, and the network decides how the parameters are set, and the main bearer set-up logic consists of only one signalling transaction from the network to the UE and all interim network elements.

The resulting SAE bearer model is shown in Figure 3.23. The bearer model itself is very similar to the GPRS bearer model, but it has fewer layers. EPS supports the always-on concept.

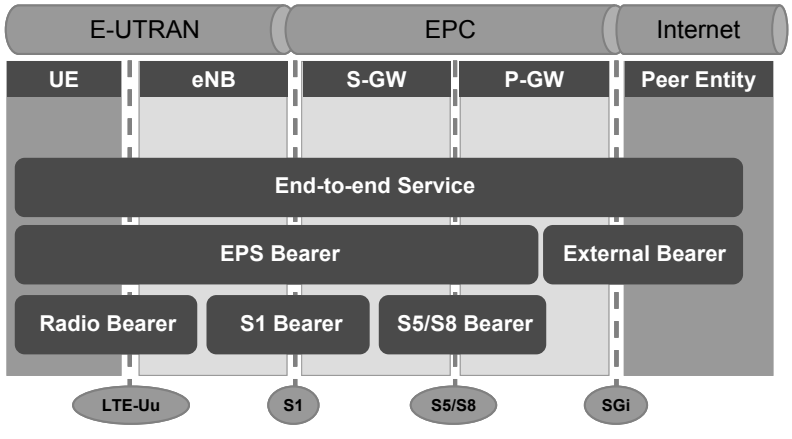


Figure 3.23 SAE Bearer model

Each UE that is registered to the system has at least one bearer called the default bearer available, so that continuous IP connectivity is provided. The default bearer may have quite basic QoS capabilities, but additional bearers may be set up on demand for services that need more stringent QoS. These are called dedicated bearers. The network may also map several IP flows that have matching QoS characteristics to the same EPS bearer.

The bearer set-up logic works so that the UE first signals on the application layer, on top of the default bearer, to an Application Server (AS) in the operator service cloud, e.g. with IMS, to set up the End-to-end Service. This signalling may include QoS parameters, or simply indication to a known service. The AS will then request the set-up of the corresponding EPS bearer through the PCC infrastructure. There is no separate signalling transaction for the EPS bearer layer, but the EPS bearer is set up together with the signalling for the lower layers, i.e. S5/S8 bearer, S1 Bearer and Radio Bearer. Furthermore, since the eNodeB is responsible for controlling the radio interface transmission in the uplink as well, the UE can operate based on very basic QoS information. The overall goal for network orientation in bearer set-up is to minimize the need for QoS knowledge and configuration in the UE.

Also the QoS parameters were optimized for SAE. Only a limited set of signalled QoS parameters are included in the specifications. They are:

- QoS Class Identifier (QCI): It is an index that identifies a set of locally configured values for three QoS attributes: Priority, Delay and Loss Rate. QCI is signalled instead of the values of these parameters. Ten pre-configured classes have been specified in two categories of bearers, Guaranteed Bit Rate (GBR) and Non-Guaranteed Bit-Rate (Non-GBR) bearers. In addition operators can create their own classes that apply within their network. The standard QCI classes and the values for the parameters within the class are shown in Table 3.8.
- Allocation and Retention Priority (ARP): Indicates the priority of the bearer compared to other bearers. This provides the basis for admission control in bearer set-up, and further in a congestion situation if bearers need to be dropped.
- Maximum Bit Rate (MBR): Identifies the maximum bit rate for the bearer. Note that a Release 8 network is not required to support differentiation between the MBR and GBR, and the MBR value is always set to equal to the GBR.
- Guaranteed Bit Rate (GBR): Identifies the bit rate that will be guaranteed to the bearer.
- Aggregate Maximum Bit Rate (AMBR): Many IP flows may be mapped to the same bearer, and this parameter indicates the total maximum bit rate a UE may have for all bearers in the same PDN connection.

Table 3.8 QoS parameters for QCI

QCI	Resource type	Priority	Delay budget	Loss rate	Example application
1	GBR	2	100 ms	1e-2	VoIP
2	GBR	4	150 ms	1e-3	Video call
3	GBR	5	300 ms	1e-6	Streaming
4	GBR	3	50 ms	1e-3	Real time gaming
5	Non-GBR	1	100 ms	1e-6	IMS signalling
6	Non-GBR	7	100 ms	1e-3	Interactive gaming
7	Non-GBR	6	300 ms	1e-6	Application with TCP:
8	Non-GBR	8			browsing, email, file
9	Non-GBR	9			download, etc.

Table 3.8 shows the QoS parameters that are part of the QCI class, and the nine standardized classes. The QoS parameters are:

- Resource Type: Indicates which classes will have GBR associated to them.
- Priority: Used to define the priority for the packet scheduling of the radio interface.
- Delay Budget: Helps the packet scheduler to maintain sufficient scheduling rate to meet the delay requirements for the bearer.
- Loss Rate: Helps to use appropriate RLC settings, e.g. number of re-transmissions.

References

[1] 3GPP TS 23.401, 'General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access (Release 8)'.
[2] 3GPP TS 23.402, 'Architecture enhancements for non-3GPP accesses (Release 8)'.
[3] M. Poikselkä et al., 'The IMS: IP Multimedia Concepts and Services', 2nd edition, Wiley, 2006.
[4] 3GPP TS 33.401, 'Security Architecture (Release 8)'.
[5] 3GPP TS 23.203, 'Policy and charging control architecture (Release 8)'.
[6] 3GPP TS 36.413, 'Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description (Release 8)'.
[7] 3GPP TS 36.423, 'Evolved Universal Terrestrial Radio Access Network (E-UTRAN); X2 Application Protocol (X2AP) (Release 8)'.
[8] 3GPP TS 29.274, 'Evolved GPRS Tunnelling Protocol (eGTP) for EPS (Release 8)'.
[9] 3GPP TS 36.413, 'Evolved Universal Terrestrial Radio Access (E-UTRA); S1 Application Protocol (S1AP) (Release 8)'.
[10] 3GPP TS 29.275, 'PMIP based Mobility and Tunnelling protocols (Release 8)'.
[11] 3GPP TS 29.061, 'Inter-working between the Public Land Mobile Network (PLMN) supporting packet based services and Packet Data Networks (PDN) (Release 8)'.
[12] 3GPP TS 29.272, 'MME Related Interfaces Based on Diameter Protocol (Release 8)'.
[13] 3GPP TS 29.212, 'Policy and charging control over Gx reference point (Release 8)'.
[14] 3GPP TS 29.214, 'Policy and charging control over Rx reference point (Release 8)'.
[15] 3GPP TS 24.301, 'Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS) (Release 8)'.
[16] H. Holma, A. Toskala, 'WCDMA for UMTS – HSPA Evolution and LTE', 4th edition, Wiley, 2007.
[17] T. Halonen, J. Romero, J. Melero, 'GSM, GPRS and EDGE Performance: Evolution Towards 3G/UMTS', 2nd edition, Wiley, 2003.
[18] 3GPP TS 23.060, 'General Packet Radio Service (GPRS); Service description; Stage 2 (Release 8)'.
[19] 3GPP TS 23.272, 'Circuit Switched (CS) fallback in Evolved Packet System (EPS); Stage 2 (Release 8)'.
[20] 3GPP TS 23.216, 'Single Radio Voice Call Continuity (SRVCC); Stage 2 (Release 8)'.
[21] 3GPP TS 29.118, 'Mobility Management Entity (MME) – Visitor Location Register (VLR) SGs interface specification (Release 8)'.
[22] 3GPP TS 29.280, '3GPP EPS Sv interface (MME to MSC) for SRVCC (Release 8)'.
[23] 3GPP TS 33.402, 'Security aspects of non-3GPP accesses (Release 8)'.
[24] IETF Internet-Draft, draft-arkko-eap-aka-kdf, 'Improved Extensible Authentication Protocol Method for 3rd Generation Authentication and Key Agreement (EAP-AKA)'. (J. Arkko,V. Lehtovirta, P. Eronen) 2008.
[25] 3GPP TS 23.234, '3GPP system to Wireless Local Area Network (WLAN) interworking; System description (Release 7)'.
[26] IETF RFC 4306, 'Internet Key Exchange (IKEv2) Protocol.' C. Kaufman, Editor, 2005.
[27] 3GPP TS 24.303, 'Mobility management based on Dual-Stack Mobile IPv6 (Release 8)'.
[28] 3GPP TS 29.273, 'Evolved Packet System (EPS); 3GPP EPS AAA interfaces (Release 8)'.
[29] 3GPP TS 24.302, 'Access to the Evolved Packet Core (EPC) via non-3GPP access networks (Release 8)'.
[30] 3GPP TS 24.304, 'Mobility management based on Mobile IPv4; User Equipment (UE) – foreign agent interface (Release 8)'.
[31] 3GPP2 Specification X.P0057, 'E-UTRAN – HRPD Connectivity and Interworking: Core Network Aspects (2008)'.

[32] 3GPP2 Specification X.P0022, 'E-UTRAN – HRPD Connectivity and Interworking: Access Network Aspects (E-UTRAN – HRPD IOS) (2008)'.
[33] 3GPP TS 29.276, 'Optimized Handover Procedures and Protocols between EUTRAN Access and cdma2000 HRPD Access (Release 8)'.
[34] 3GPP TS 29.277, 'Optimized Handover Procedures and Protocols between EUTRAN Access and 1xRTT Access (Release 8)'.
[35] 3GPP2 Specification A.S0008-C, 'Interoperability Specification (IOS) for High Rate Packet Data (HRPD) Radio Access Network Interfaces with Session Control in the Access Network (2007)'.
[36] 3GPP TS 23.002, 'Network architecture (Release 8)'.
[37] 3GPP TS 23.228, 'IP Multimedia Subsystem (IMS); Stage 2 (Release 8)'.
[38] 3GPP TS 24.229, 'IP multimedia call control protocol based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3 (Release 8)'.
[39] 3GPP TS 33.203, '3G security; Access security for IP-based services (Release 8)'.
[40] IETF RFC 2617, 'HTTP Authentication: Basic and Digest Access Authentication', J. Franks, P. Hallam-Baker, J. Hostetler, S. Lawrence, P. Leach, A. Luotonen, L. Stewart (1999).
[41] 3GPP TS 29.213, 'Policy and Charging Control signalling flows and QoS parameter mapping (Release 8)'.
[42] 3GPP TS 32.240, 'Telecommunication management; Charging management; Charging architecture and principles (Release 8)'.
[43] 3GPP TS 32.251, 'Telecommunication management; Charging management; Packet Switched (PS) domain charging (Release 8)'.
[44] IETF RFC 4006, 'Diameter Credit-Control Application', H. Hakala, L. Mattila, J.-P. Koskinen, M. Stura, J. Loughney (2005).
[45] 3GPP TS 29.215, 'Policy and Charging Control (PCC) over S9 reference point (Release 8)'.
[46] 3GPP TS 32.299, 'Telecommunication management; Charging management; Diameter charging applications (Release 8)'.
[47] 3GPP TS 32.295, 'Telecommunication management; Charging management; Charging Data Record (CDR) transfer (Release 8)'.

4

Introduction to OFDMA and SC-FDMA and to MIMO in LTE

Antti Toskala and Timo Lunttila

4.1 Introduction

As discussed in Chapter 1, LTE multiple access is different to that of WCDMA. In LTE the downlink multiple access is based on the Orthogonal Frequency Division Multiple Access (OFDMA) and the uplink multiple access is based on the Single Carrier Frequency Division Multiple Access (SC-FDMA). This chapter will introduce the selection background and the basis for both SC-FDMA and OFDMA operation. The basic principles behind the multi-antenna transmission in LTE, using Multiple Input Multiple Output (MIMO) technology, is also introduced. The intention of this chapter is to illustrate the multiple access principles in a descriptive way without too much mathematics. For those interested in the detailed mathematical notation, two selected references are given that provide a mathematical treatment of the different multiple access technologies, covering both OFDMA and SC-FDMA.

4.2 LTE Multiple Access Background

A Single Carrier (SC) transmission means that information is modulated only to one carrier, adjusting the phase or amplitude of the carrier or both. Frequency could also be adjusted, but in LTE this is not effected. The higher the data rate, the higher the symbol rate in a digital system and thus the bandwidth is higher. With the use of simple Quadrature Amplitude Modulation (QAM), with the principles explained, for example in [1], the transmitter adjusts the signal to carry the desired number of bits per modulation symbol. The resulting spectrum waveform is a single carrier spectrum, as shown in Figure 4.1, with the spectrum mask influenced (after filtering) by the pulse shape used.

With the Frequency Division Multiple Access (FDMA) principle, different users would then be using different carriers or sub-carriers, as shown in Figure 4.2, to access the system simultaneously having their data modulation around a different center frequency. Care must be now

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6

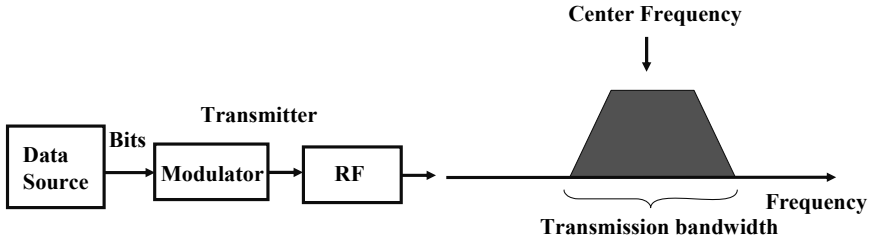


Figure 4.1 Single carrier transmitter

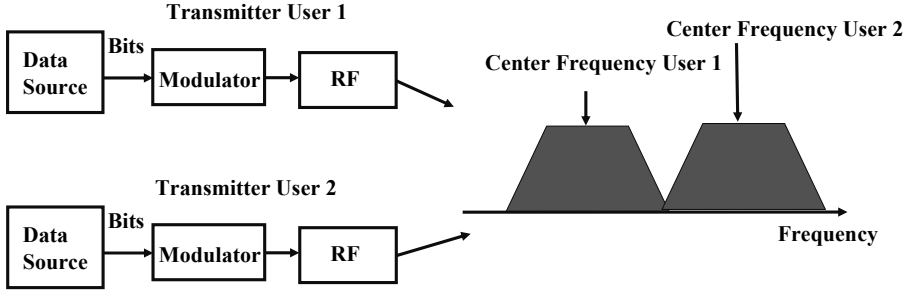


Figure 4.2 FDMA principle

taken to create the waveform in such a way that there is no excessive interference between the carriers, nor should one be required to use extensive guard bands between users.

The use of the multi-carrier principle is shown in Figure 4.3, where data are divided on the different sub-carriers of one transmitter. The example in Figure 4.3 has a filter bank which for practical solutions (such as the ones presented later) is usually replaced with Inverse Fast Fourier Transform (IFFT) for applications where the number of sub-carriers is high. There is a constant spacing between neighboring sub-carriers. One of the approaches to multi-carrier is also the dual carrier WCDMA (dual cell HSDPA, as covered in Chapter 13), which sends two WCDMA next to each other but does not use the principles explained later in this section for high spectrum utilization.

To address the resulting inefficiency from the possible guard band requirements, the approach is to choose the system parameters in such a way as to achieve orthogonality between the dif-

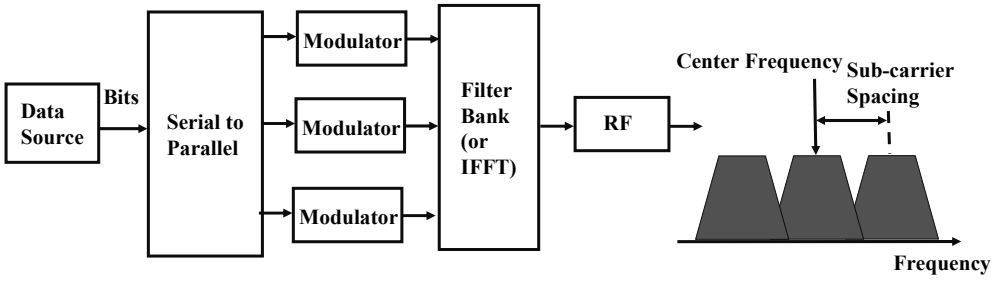


Figure 4.3 Multi-carrier principle

ferent transmissions, and to create the sub-carriers so that they do not interfere with each other but their spectrums could still overlap in the frequency domain. This is what is achieved with the Orthogonal Frequency Division Multiplexing (OFDMA) principle, where each of the center frequencies for the sub-carriers is selected from the set that has such a difference in the frequency domain that the neighboring sub-carriers have zero value at the sampling instant of the desired sub-carrier, as shown in Figure 4.4. For LTE, the constant frequency difference between the sub-carriers has been chosen to be 15 kHz in Release 8 (an alternative of 7.5 kHz is planned to be supported in later releases in connection with broadcast applications such as mobile TV).

The basic principle of OFDMA was already known in the 1950s, at a time when systems were using analog technology, and making the sub-carriers stay orthogonal as a function of component variations and temperature ranges was not a trivial issue. Since the widespread use of digital technology for communications, OFDMA also became more feasible and affordable for consumer use. During recent years OFDMA technology has been widely adopted in many areas such as in digital TV (DVB-T and DVB-H) as well as in Wireless Local Area Network (WLAN) applications.

OFDMA principles have been used in the uplink part of LTE multiple access just as the SC-FDMA uses many of the OFDMA principles in the uplink direction to achieve high spectral efficiency, as described in the next section. The SC-FDMA in the current form, covered in a later section of this chapter, is more novel technology with publications from the late 1990s, such as those presented in [2] and the references therein.

The overall motivation for OFDMA in LTE and in other systems has been due to the following properties:

- good performance in frequency selective fading channels;
- low complexity of base-band receiver;
- good spectral properties and handling of multiple bandwidths;
- link adaptation and frequency domain scheduling;
- compatibility with advanced receiver and antenna technologies.

Many of these benefits (with more explanation provided in the following sections) could only be achieved following the recent developments in the radio access network architecture, meaning setting the radio related control in the base station (or NodeB in 3GPP terms for

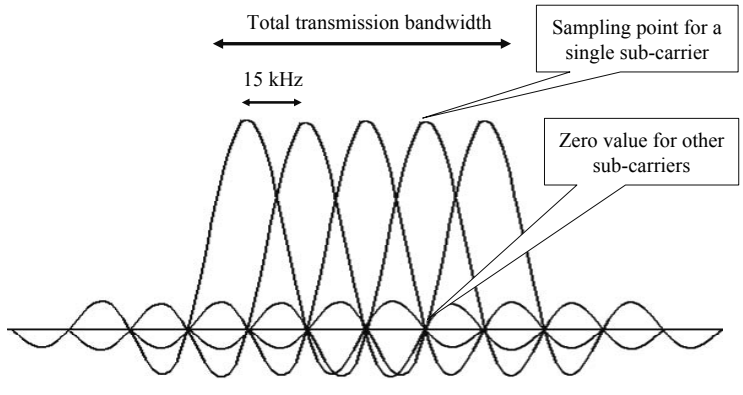


Figure 4.4 Maintaining the sub-carriers' orthogonality

WCDMA), and as the system bandwidths are getting larger, beyond 5 MHz, receiver complexity also becomes more of an issue.

The OFDMA also has challenges, such as:

- Tolerance to frequency offset. This was tackled in LTE design by choosing a sub-carrier spacing of 15 kHz, which gives a large enough tolerance for Doppler shift due to velocity and implementation imperfections.
- The high Peak-to-Average Ratio (PAR) of the transmitted signal, which requires high linearity in the transmitter. The linear amplifiers have a low power conversion efficiency and therefore are not ideal for mobile uplinks. In LTE this was solved by using the SC-FDMA, which enables better power amplifier efficiency.

When looking back, the technology selections carried out for the 3rd generation system in the late 1990s, the lack of a sensible uplink solution, the need for advanced antenna solutions (with more than a single antenna) and having radio resource control centralized in the Radio Network Controller (RNC) were the key factors not to justify the use of OFDMA technology earlier. There were studies to look at the OFDMA together with CDMA in connection with the 3rd generation radio access studies, such as are covered in [3]. The key enabling technologies that make OFDMA work better, such as base station based scheduling (Release 5 and 6) and Multiple Input Multiple Output (MIMO) (Release 7), have been introduced only in the later phase of WCDMA evolution. These enhancements, which were introduced in WCDMA between 2002 and 2007, allowed the OFDMA technology to be better used than would have been the case for the simple use of OFDMA only as a modulation method based on a traditional 2nd generation cellular network without advanced features.

4.3 OFDMA Basics

The practical implementation of an OFDMA system is based on digital technology and more specifically on the use of Discrete Fourier Transform (DFT) and the inverse operation (IDFT) to move between time and frequency domain representation. The resulting signal feeding a sinusoidal wave to the Fast Fourier Transform (FFT) block is illustrated in Figure 4.5. The

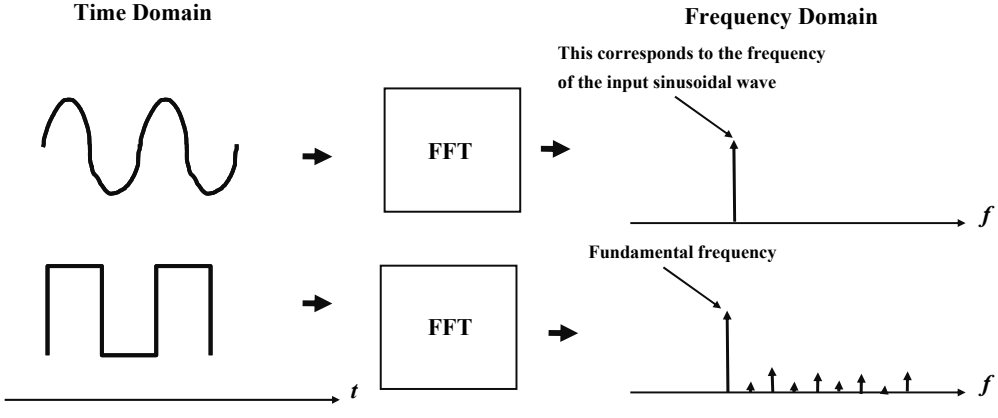


Figure 4.5 Results of the FFT operation with different inputs

practical implementations use the FFT. The FFT operation moves the signal from time domain representation to frequency domain representation. The Inverse Fast Fourier Transform (IFFT) does the operation in the opposite direction. For the sinusoidal wave, the FFT operation's output will have a peak at the corresponding frequency and zero output elsewhere. If the input is a square wave, then the frequency domain output contains peaks at multiple frequencies as such a wave contains several frequencies covered by the FFT operation. An impulse as an input to FFT would have a peak on all frequencies. As the square wave has a regular interval T, there is a bigger peak at the frequency $1/T$ representing the fundamental frequency of the waveform, and a smaller peak at odd harmonics of the fundamental frequency. The FFT operation can be carried out back and forth without losing any of the original information, assuming that the classical requirements for digital signal processing in terms of minimum sampling rates and word lengths (for the numerics) are fulfilled.

The implementation of the FFT is well researched and optimized (low amount of multiplications) when one can stay with power of lengths. Thus for LTE the necessary FFT lengths also tend to be powers of two, such as 512, 1024, etc. From the implementation point of view it is better to have, for example, a FFT size of 1024 even if only 600 outputs are used (see later the discussion on sub-carriers), than try to have another length for FFT between 600 and 1024.

The transmitter principle in any OFDMA system is to use narrow, mutually orthogonal sub-carriers. In LTE the sub-carrier spacing is 15 kHz regardless of the total transmission bandwidth. Different sub-carriers are orthogonal to each other, as at the sampling instant of a single sub-carrier the other sub-carriers have a zero value, as was shown in Figure 4.4. The transmitter of an OFDMA system uses IFFT block to create the signal. The data source feeds to the serial-to-parallel conversion and further to the IFFT block. Each input for the IFFT block corresponds to the input representing a particular sub-carrier (or particular frequency component of the time domain signal) and can be modulated independently of the other sub-carriers. The IFFT block is followed by adding the cyclic extension (cyclix prefix), as shown in Figure 4.6.

The motivation for adding the cyclic extension is to avoid inter-symbol interference. When the transmitter adds a cyclic extension longer than the channel impulse response, the effect of the previous symbol can be avoided by ignoring (removing) the cyclic extension at the receiver.

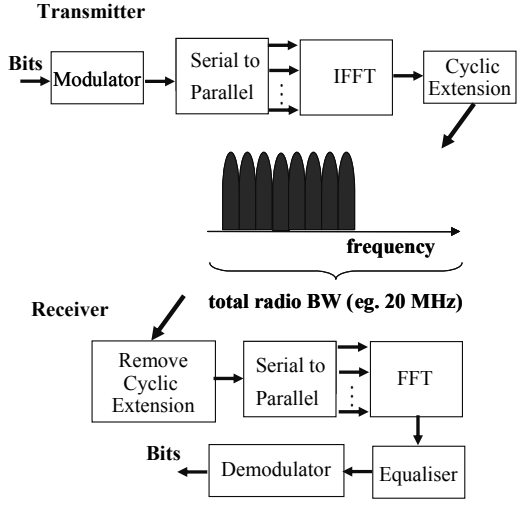


Figure 4.6 OFDMA transmitter and receiver

The cyclic prefix is added by copying part of the symbol at the end and attaching it to the beginning of the symbol, as shown in Figure 4.7. The use of cyclic extension is preferable to simply a break in the transmission (guard interval) as the OFDM symbol then seems to be periodic. When the OFDMA symbol now appears as periodic due to cyclic extension, the impact of the channel ends up corresponding to a multiplication by a scalar, assuming that the cyclic extension is sufficiently long. The periodic nature of the signals also allows for a discrete Fourier spectrum enabling the use of DFT and IDFT in the receiver and transmitter respectively.

Typically the guard interval is designed to be such that it exceeds the delay spread in the environment where the system is intended to be operated. In addition to the channel delay spread, the impact of transmitter and receiver filtering needs to be accounted for in the guard interval design. The OFDMA receiver sees the OFDMA symbol coming as through a FIR filter, without separating individual frequency components like the RAKE receiver as described in [4]. Thus, similar to the channel delay spread, the length of the filter applied to the signal in the receiver and transmitter side will also make this overall 'filtering' effect longer than just the delay spread.

While the receiver does not deal with the inter-symbol interference, it still has to deal with the channel impact for the individual sub-carriers that have experienced frequency dependent phase and amplitude changes. This channel estimation is facilitated by having part of the symbols as known reference or pilot symbols. With the proper placement of these symbols in both the time and frequency domains, the receiver can interpolate the effect of the channel to the different sub-carriers from this time and frequency domain reference symbol 'grid'. An example is shown in Figure 4.8.

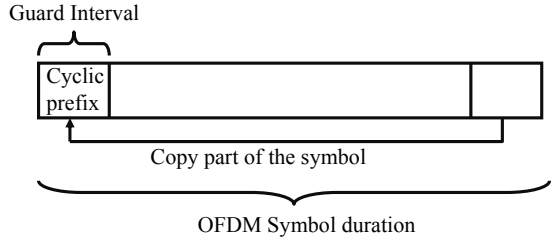


Figure 4.7 Creation of the guard interval for the OFDM symbol

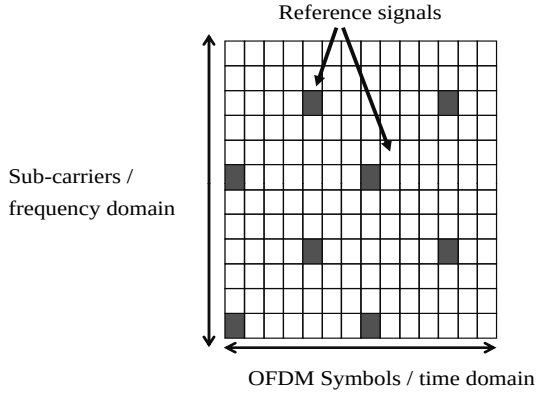


Figure 4.8 Reference symbols spread over OFDMA sub-carriers and symbols

A typical type of receiver solution is the frequency domain equalizer, which basically reverts the channel impact for each sub-carrier. The frequency domain equalizer in OFDMA simply multiplies each sub-carrier (with the complex-valued multiplication) based on the estimated channel frequency response (the phase and amplitude adjustment each sub-carrier has experienced) of the channel. This is clearly a simpler operation compared with WCDMA and is not dependent on channel length (length of multipath in chips) as is the WCDMA equalizer. For WCDMA the challenge would be also to increase the chip rate from the current value of 3.84 Mcps, as then the amount of multi-path components separated would increase (depending on the environment) resulting in the need for more RAKE fingers and contributing heavily to equalizer complexity.

In WCDMA the channel estimation in the downlink is based on the Common Pilot Channel (CPICH) and then on pilot symbols on the Dedicated Channel (DCH), which are transmitted with the spread over the whole transmission bandwidth, and different cells separated by different spreading codes. As in the OFDMA system there is no spreading available, other means must be used to separate the reference symbols between cells or between different antennas. In the multi-antenna transmission, as discussed in further detail later in this chapter, the pilot symbols have different positions. A particular position used for a pilot symbol for one antenna is left unused for other antenna in the same cell. Between different cells this blanking is not used, but different pilot symbol patterns and symbol locations can be used.

The additional tasks that the OFDMA receiver needs to cover are time and frequency synchronization. Synchronization allows the correct frame and OFDMA symbol timing to be obtained so that the correct part of the received signal is dropped (cyclic prefix removal). Time synchronization is typically obtained by correlation with known data samples – based on, for example, the reference symbols – and the actual received data. The frequency synchronization estimates the frequency offset between the transmitter and the receiver and with a good estimate of the frequency offset between the device and base station, the impact can be then compensated both for receiver and transmitter parts. The device locks to the frequency obtained from the base station, as the device oscillator is not as accurate (and expensive) as the one in the base station. The related 3GPP requirements for frequency accuracy are covered in Chapter 11.

Even if in theory the OFDMA transmission has rather good spectral properties, the real transmitter will cause some spreading of the spectrum due to imperfections such as the clipping in the transmitter. Thus the actual OFDMA transmitter needs to have filtering similar to the pulse shape filtering in WCDMA. In the literature this filtering is often referred as windowing, as in the example transmitter shown in Figure 4.9.

An important aspect of the use of OFDMA in a base station transmitter is that users can be allocated basically to any of the sub-carriers in the frequency domain. This is an additional element to the HSDPA scheduler operation, where the allocations were only in the time domain

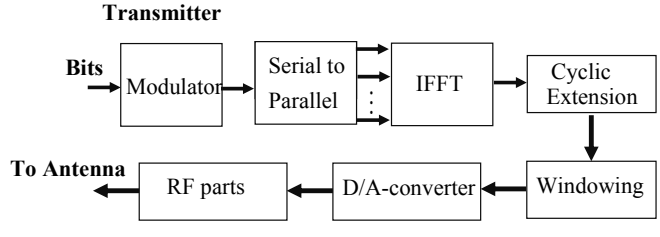


Figure 4.9 OFDMA transmitter with windowing for shaping the spectral mask

and code domain but always occupied the full bandwidth. The possibility of having different sub-carriers to allocated users enables the scheduler to benefit from the diversity in the frequency domain, this diversity being due to the momentary interference and fading differences in different parts of the system bandwidth. The practical limitation is that the signaling resolution due to the resulting overhead has meant that allocation is not done on an individual sub-carrier basis but is based on resource blocks, each consisting of 12 sub-carriers, thus resulting in the minimum bandwidth allocation being 180 kHz. When the respective allocation resolution in the time domain is 1 ms, the downlink transmission resource allocation thus means filling the resource pool with 180 kHz blocks at 1 ms resolution, as shown in Figure 4.10. Note that the resource block in the specifications refers to the 0.5 ms slot, but the resource allocation is done anyway with the 1 ms resolution in the time domain. This element of allocating resources dynamically in the frequency domain is often referred to as frequency domain scheduling or frequency domain diversity. Different sub-carriers could ideally have different modulations if one could adapt the channel without restrictions. For practical reasons it would be far too inefficient to try either to obtain feedback with 15 kHz sub-carrier resolution or to signal the modulation applied on a individual sub-carrier basis. Thus parameters such as modulation are fixed on the resource block basis.

The OFDMA transmission in the frequency domain thus consists of several parallel sub-carriers, which in the time domain correspond to multiple sinusoidal waves with different frequencies filling the system bandwidth with steps of 15 kHz. This causes the signal envelope to vary strongly, as shown in Figure 4.11, compared to a normal QAM modulator, which is only sending one symbol at a time (in the time domain). The momentary sum of sinusoids leads to the Gaussian distribution of different peak amplitude values.

This causes some challenges to the amplifier design as, in a cellular system, one should aim for maximum power amplifier efficiency to achieve minimum power consumption. Figure 4.12 illustrates how a signal with a higher envelope variation (such as the OFDMA signal in the time domain in Figure 4.11) requires the amplifier to use additional back-off compared to a regular single carrier signal. The amplifier must stay in the linear area with the use of extra power back-off in order to prevent problems to the output signal and spectrum mask. The use of additional back-off leads to a reduced amplifier power efficiency or a smaller output power. This either causes the uplink range to be shorter or, when the same average output power level

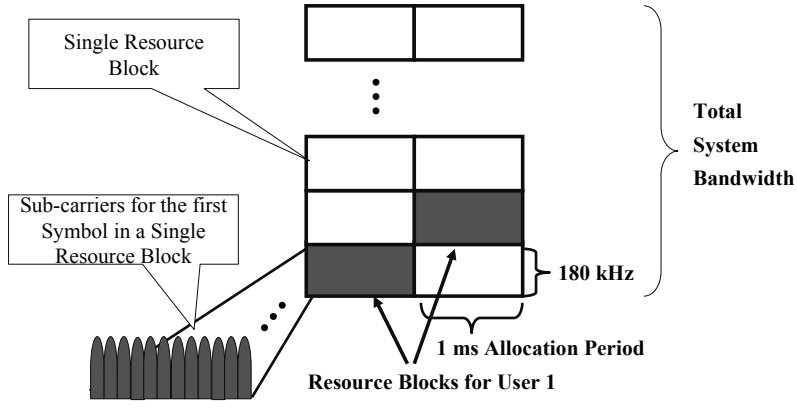


Figure 4.10 OFDMA resource allocation in LTE

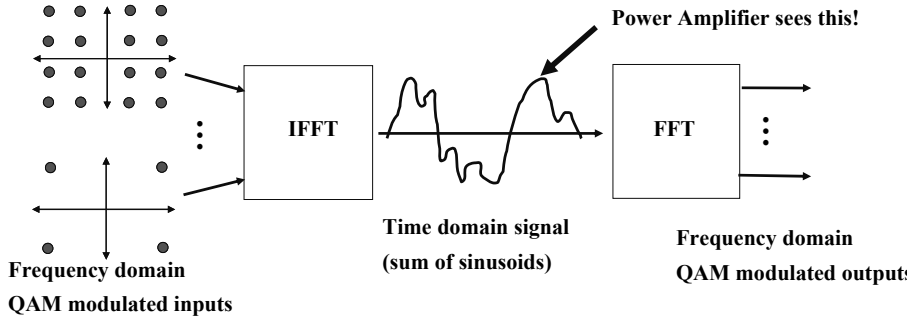


Figure 4.11 OFDMA signal envelope characteristics

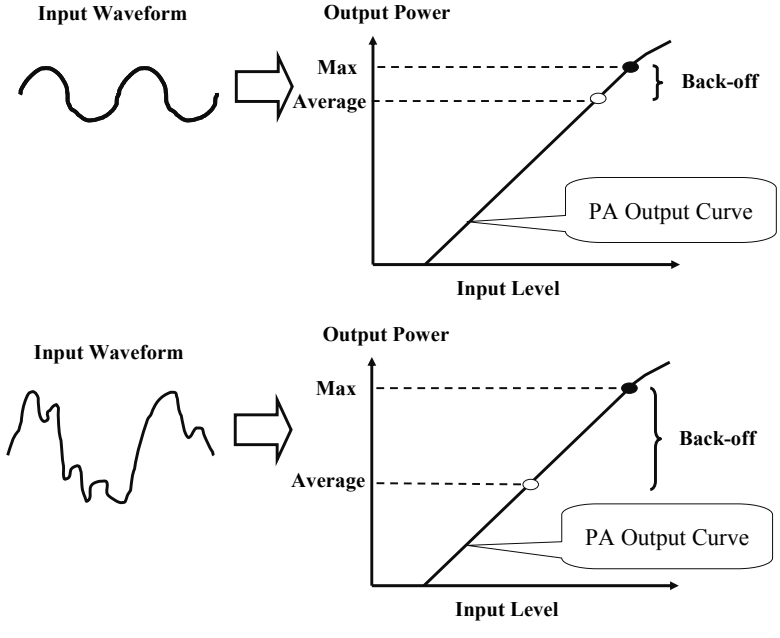


Figure 4.12 Power amplifier back-off requirements for different input waveforms

is maintained, the battery energy is consumed faster due to higher amplifier power consumption. The latter is not considered a problem in fixed applications where the device has a large volume and is connected to the mains, but for small mobile devices running on their own batteries it creates more challenges.

This was the key reason why 3GPP decided to use OFDMA in the downlink direction but to use the power efficient SC-FDMA in the uplink direction. Further principles of SC-FDMA are presented in the next section. From the research several methods are known to reduce the PAR, but of more significance – particularly for the amplifier – is the Cubic Metric (CM), which was introduced in 3GPP to better describe the impact to the amplifier. The exact definition of CM can be found from [5].

An OFDMA system is also sensitive to frequency errors as previously mentioned in section 4.2 . The basic LTE sub-carrier spacing of 15 kHz facilitates enough tolerance for the effects of implementation errors and Doppler effect without too much degradation in the sub-carrier orthogonality. 3GPP has agreed that for broadcast only (on a dedicated carrier) an optional 7.5 kHz sub-carrier spacing can also be used, but full support of the broadcast only carrier is not part of the LTE Release 8. The physical layer details for the 7.5 kHz case principles can already be found in the 36.2 series specifications of Release 8, but details – especially for the higher layer operation – are only to be completed for releases after Release 8.

4.4 SC-FDMA Basics

In the uplink direction 3GPP uses SC-FDMA for multiple access, valid for both FDD and TDD modes of operation. The basic form of SC-FDMA could be seen as equal to the QAM modulation, where each symbol is sent one at a time similarly to Time Division Multiple Access (TDMA) systems such as GSM. Frequency domain generation of the signal, as shown in Figure 4.13, adds the OFDMA property of good spectral waveform in contrast to time domain signal generation with a regular QAM modulator. Thus the need for guard bands between different users can be avoided, similar to the downlink OFDMA principle. As in an OFDMA system, a cyclic prefix is also added periodically – but not after each symbol as the symbol rate is faster in the time domain than in OFDMA – to the transmission to prevent inter-symbol interference and to simplify the receiver design. The receiver still needs to deal with inter-symbol interference as the cyclic prefix now prevents inter-symbol interference between a block of symbols, and thus there will still be inter-symbol interference between the cyclic prefixes. The receiver will thus run the equalizer for a block of symbols until reaching the cyclic prefix that prevents further propagation of the inter-symbol interference.

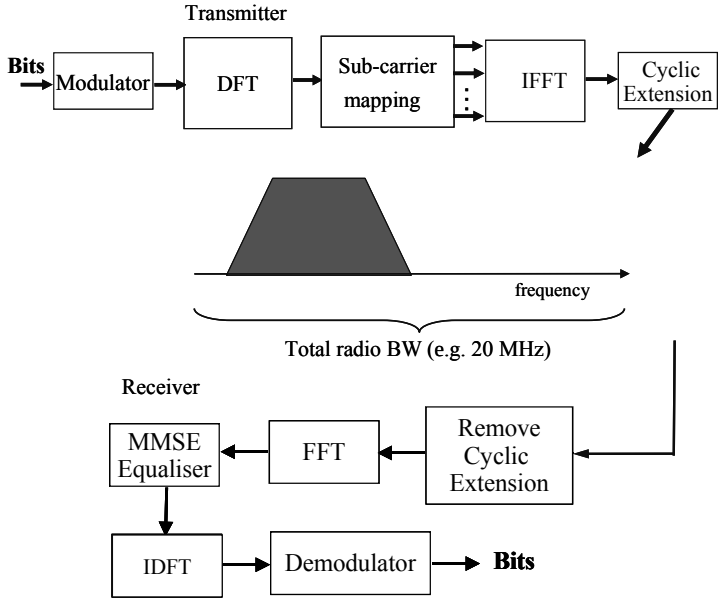


Figure 4.13 SC-FDMA transmitter and receiver with frequency domain signal generation

The transmission occupies the continuous part of the spectrum allocated to the user, and for LTE the system facilitates a 1 ms resolution allocation rate. When the resource allocation in the frequency domain is doubled, so is the data rate, assuming the same level of overhead. The individual transmission (with modulation) is now shorter in time but wider in the frequency domain, as shown in Figure 4.14. The example in Figure 4.14 assumes that in the new resource allocation the existing frequency resource is retained and the same amount of additional transmission spectrum is allocated, thus doubling the transmission capacity. In reality the allocations do not need to have frequency domain continuity, but can take any set of continuous allocation of frequency domain resources. The practical signaling constraints define the allowed amount of 180 kHz resource blocks that can be allocated. The maximum allocated bandwidth depends on the system bandwidth used, which can be up to 20 MHz. The resulting maximum allocation bandwidth is somewhat smaller as the system bandwidth definition includes a guard towards the neighboring operator. For example, with a 10 MHz system channel bandwidth the maximum resource allocation is equal to 50 resource blocks thus having a transmission bandwidth of 9 MHz. The relationship between the *Channel bandwidth* ($BW_{Channel}$) and *Transmission bandwidth configuration* (N_{RB}) is covered in more detail in Chapter 11.

The SC-FDMA resource block for frequency domain signal generation is defined using the same values used in the OFDMA downlink, based on the 15 kHz sub-carrier spacing. Thus even if the actual transmission by name is a single carrier, the signal generation phase uses a sub-carrier term. In the simplest form the minimum resource allocated uses 12 sub-carriers, and is thus equal to 180 kHz. The complex valued modulation symbols with data are allocated to the resource elements not needed for reference symbols (or control information) in the resource block, as shown in Figure 4.15. After the resource mapping has been done the signal is fed to the time domain signal generation that creates the SC-FDMA signal, including the selected length of the cyclic prefix. The example in Figure 4.15 assumes a particular length of cyclic prefix with the two different options introduced in Chapter 5.

As shown in Figure 4.15, reference symbols are located in the middle of the slot. These are used by the receiver to perform the channel estimation. There are different options for the reference symbols to be used; sometimes a reference symbol hopping pattern is also used, as covered in more detail in Chapter 5. Also specifically covered further in Chapter 5 are the sounding reference signals, which are momentarily sent over a larger bandwidth than needed,

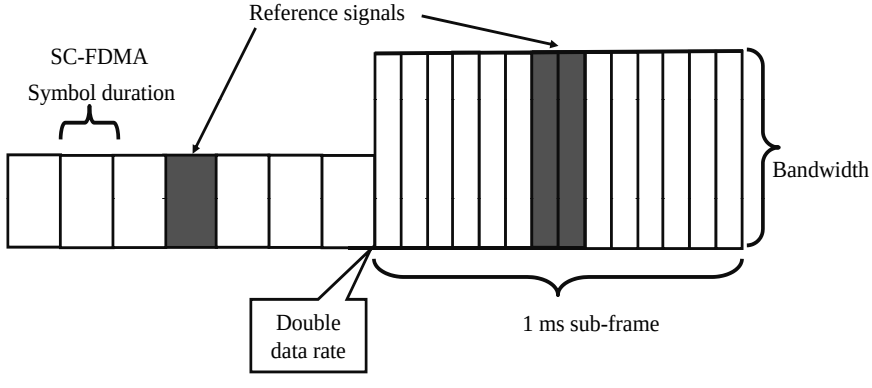


Figure 4.14 Adjusting data rate in a SC-FDMA system

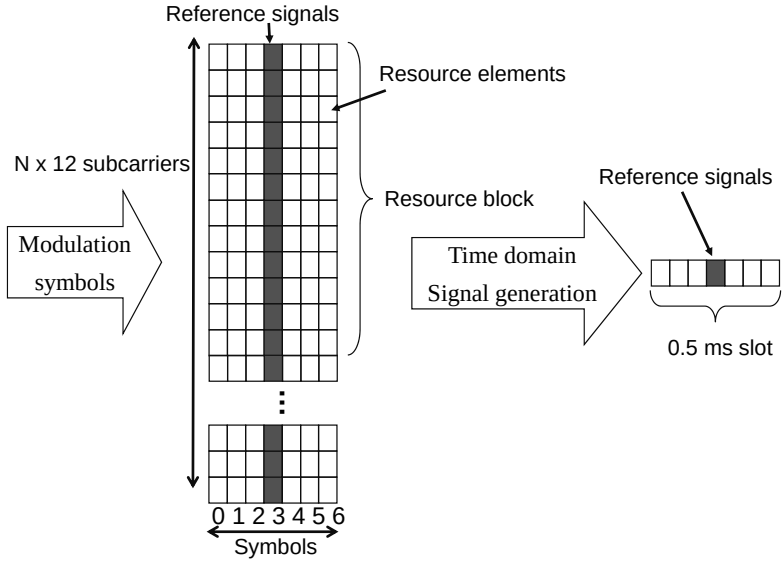


Figure 4.15 Resource mapping in SC-FDMA

for the data to give the base station receiver information of a larger portion of the frequency spectrum to facilitate frequency domain scheduling in the uplink direction.

Different users are thus sharing the resources in the time as well as in the frequency domain. In the time domain the allocation granularity is 1 ms and in the frequency domain it is 180 kHz. The base station needs to control each transmission so that they do not overlap in the resources. Also to avoid lengthy guard times, timing advance needs to be used, as presented in Chapter 5. By modifying the IFFT inputs, the transmitter can place the transmission in the desired part of the frequency, as shown in Figure 4.16. The base station receiver can detect the

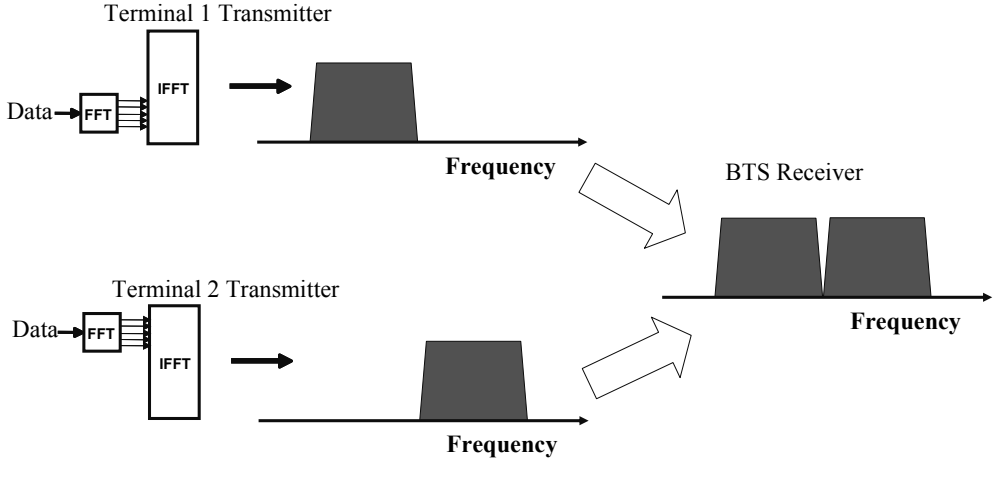


Figure 4.16 Multiple access with resource sharing in the frequency domain with SC-FDMA and frequency domain signal generation

transmission from the correct frequency/time resource. As all the uplink utilization is based on the base station scheduling, with the exception of the random access channel, the base station always knows which user to expect in which resource.

Since we are now transmitting in the time domain only a single modulation symbol at a time, the system retains its good envelope properties and the waveform characteristics are now dominated by the modulation method applied. This allows the SC-FDMA to reach a very low signal PAR or, even more importantly, CM facilitating efficient power amplifiers in the devices. The value of CM as a function modulation applied is shown in Figure 4.17. The use of a low CM modulation method such as Quadrature Phase Shift Keying (QPSK) allows a low CM value and thus the amplifier can operate close to the maximum power level with minimum back-off (Figure 4.17). This allows a good power conversion efficiency of the power amplifier and thus lowers the device power consumption. Note that the pi/2-Binary Phase Shift Keying (BPSK) was originally considered in 3GPP, but as 3GPP performance requirements are such that the full (23 dBm) power level needs to be reached with QPSK, there are no extra benefits for the use of pi/2-BPSK; thus, this was eventually not included in the specifications for user data. The modulation methods in LTE vary depending on whether the symbols are for physical layer control information or for higher layer data (user data or higher layer control signaling) transmission purposes (details in Chapter 5).

The base station receiver for SC-FDMA is slightly more complicated than the corresponding ODFMA receiver on the device side, especially when considering receivers (equalizers) that can reach a performance corresponding to that of an OFDMA receiver. This is the obvious consequence of the receiver having to deal with the inter-symbol interference that is terminated only after a block of symbols and not after every (long) symbol as in OFDMA. This increased need for processing power is, however, not foreseen to be an issue in the base station when compared to the device design constraints and was clearly considered to be outweighed by the benefits of the uplink range and device battery life with SC-FDMA. The benefits of a dynamic resource usage with a 1 ms resolution is also that there is no base-band receiver per UE on standby

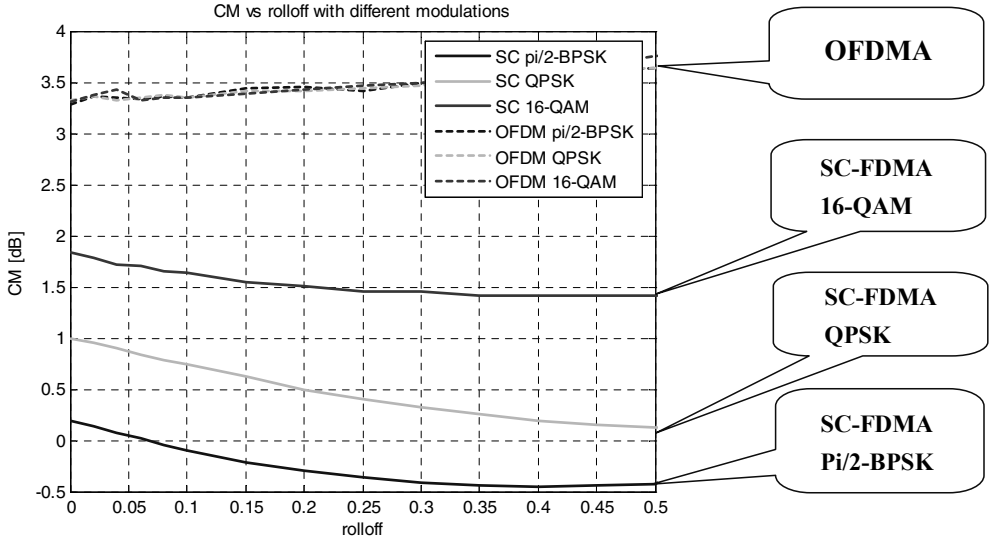


Figure 4.17 CM with OFDMA and SC-FDMA [6]. © 2006 IEEE

but the base station receiver is dynamically used for those users that have data to transmit. In any case the most resource consuming part both in uplink and downlink receiver chains is the channel decoding (turbo decoding) with the increased data rates.

4.5 MIMO Basics

One of the fundamental technologies introduced together with the first LTE Release is the Multiple Input Multiple Output (MIMO) operation including spatial multiplexing as well as pre-coding and transmit diversity. The basic principle in spatial multiplexing is sending signals from two or more different antennas with different data streams and by signal processing means in the receiver separating the data streams, hence increasing the peak data rates by a factor of 2 (or 4 with 4-by-4 antenna configuration). In pre-coding the signals transmitted from the different antennas are weighted in order to maximize the received Signal to Noise Ratio (SNR). Transmit diversity relies on sending the same signal from multiple antennas with some coding in order to exploit the gains from independent fading between the antennas. The use of MIMO has been included earlier in WCDMA specifications as covered in [4], but operating slightly differently than in LTE as a spreading operation is involved. The OFDMA nature is well suited for MIMO operation. As the successful MIMO operation requires reasonably high SNR, with an OFDMA system it can benefit from the locally (in the frequency/time domain) high SNR that is achievable. The basic principle of MIMO is presented in Figure 4.18, where the different data streams are fed to the pre-coding operation and then onwards to signal mapping and OFDMA signal generation.

The reference symbols enable the receiver to separate different antennas from each other. To avoid transmission from another antenna corrupting the channel estimation needed for separating the MIMO streams, one needs to have each reference symbol resource used by a single transmit antenna only. This principle is illustrated in Figure 4.19, where the reference symbols and empty resource elements are mapped to alternate between antennas. This principle can also be extended to cover more than two antennas, with the first LTE Release covering up to four antennas. As the number of antennas increases, the required SNR also increases the resulting transmitter/receiver complexity and the reference symbol overhead.

Even LTE uplink supports the use of MIMO technology. While the device is using only one transmit antenna, the single user data rate cannot be increased with MIMO. The cell level maximum data rate can be doubled, however, by the allocation of two devices with orthogonal reference signals. Thus the transmission in the base station is treated like a MIMO transmission, as shown in Figure 4.20, and the data stream separated with MIMO receiver processing. This

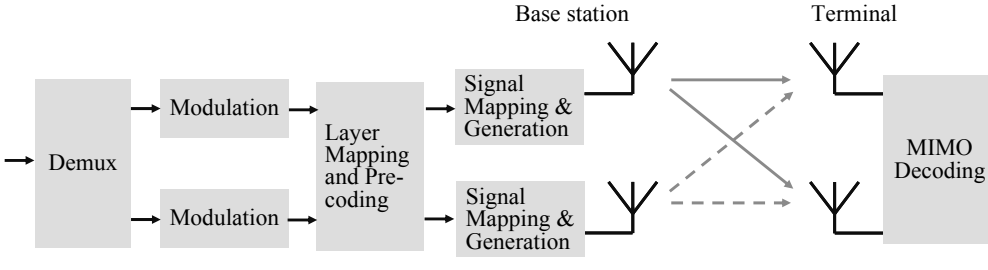


Figure 4.18 MIMO principle with two-by-two antenna configuration

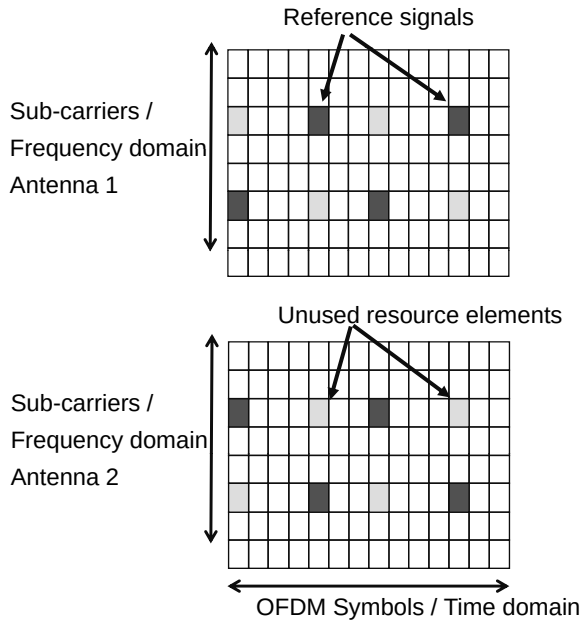


Figure 4.19 OFDMA reference symbols to support two eNodeB transmit antennas

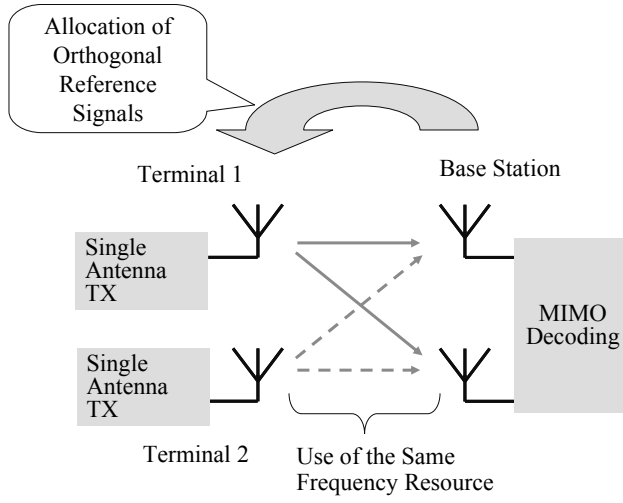


Figure 4.20 Multi-user MIMO principle with single transmit antenna devices

kind of 'virtual' or 'Multi-user' MIMO is supported in LTE Release 8 and does not represent any major implementation complexity from the device perspective as only the reference signal sequence is modified. From the network side, additional processing is needed to separate the users from each other. The use of 'classical' two antenna MIMO transmission is not particularly attractive due to the resulting device impacts, thus discussions on the device support of multi-antenna device transmission are expected to take place for later 3GPP Releases. The SC-FDMA

is well-suited for MIMO use as users are orthogonal (inside the cell) and thus the local SNR may be very high for users close to the base station.

4.6 Summary

Both OFDMA and SC-FDMA are very much related in terms of technical implementation and rely on the use of FFT/IFFT in the transmitter and receiver chain implementation. The SC-FDMA is used to optimize the range and power consumption in the uplink while the OFDMA is used in the downlink direction to minimize receiver complexity, especially with large bandwidths, and to enable frequency domain scheduling with flexibility in resource allocation. Multiple antenna operation with spatial multiplexing has been a fundamental technology of LTE from the outset, and is well suited for LTE multiple access solutions. The mathematical principles of OFDMA and SC-FDMA were not included in this chapter, but can be found from different text books, some of which are included in the references, e.g. [7] for OFDMA and [8] for SC-FDMA.

References

[1] Proakis, J.G., 'Digital Communications', 3rd edition, McGraw-Hill Book Co., 1995.
[2] Czylwik, A., 'Comparison between adaptive OFDM and single carrier modulation with frequency domain equalisation', IEEE Vehicular Technology Conference 1997, VTC-97, Phoenix, USA, pp. 863–869.
[3] Toskala, A., Castro, J., Chalard, L., Hämäläinen, S., Kalliojärvi, K., 'Cellular OFDM/CDMA Downlink Performance in the link and system level', IEE Vehicular Technology Conference 1997, VTC-97, Phoenix, USA, pp. 855–859.
[4] Holma, H., Toskala, A., 'WCDMA for UMTS', 4th edition, Wiley, 2007.
[5] Holma, H., Toskala, A., 'HSDPA/HSUPA for UMTS', Wiley, 2006.
[6] Toskala, A, Holma, H., Pajukoski, K., Tiirola, E., 'UTRAN Long Term Evolution in 3GPP'. The 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC '06), in Proceedings, 2006, Helsinki, Finland.
[7] Schulze, H., Luders, C., 'Theory and Applications of OFDMA and CDMA', Wiley, 2005.
[8] Myung, H.G., Goodman, D.J., 'Single Carrier FDMA: A New Air Interface for Long Term Evolution', Wiley, 2008.

5

Physical Layer

Antti Toskala, Timo Lunttila, Esa Tiirola, Kari Hooli and Juha Korhonen

5.1 Introduction

In this chapter the physical layer of LTE is described, based on the use of OFDMA and SC-FDMA principles as covered in Chapter 4. The LTE physical layer is characterized by the design principle of resource usage based solely on dynamically allocated shared resources rather than having dedicated resources reserved for a single user. This has an analogy with the resource usage in the internet, which is packet based without user specific resource allocation. The physical layer of a radio access system has a key role in defining the resulting capacity and becomes a focal point when comparing different systems for expected performance. Of course a competitive system requires an efficient protocol layer to ensure good performance through to both the application layer and the end user. The flat architecture adopted, as covered in Chapter 3, also enables the dynamic nature of the radio interface because all radio resource control is located close to the radio in the base station site. The 3GPP term for the base station used in this chapter is eNodeB (different to the WCDMA BTS term, which is Node B; e stands for 'evolved'). This chapter first covers the physical channel structures and then introduces the channel coding and physical layer procedures. It concludes with a description of physical layer measurements and device capabilities as well as a brief look at aspects of the parameter configuration of the physical layer. In 3GPP specifications the physical layer was covered in 36.2 series, with the four key physical layer specifications being [1–4]. Many of the issues in this chapter are valid to both FDD and TDD, but in some areas TDD has special solutions because the frame is divided between uplink and downlink. The resulting differences needed for a TDD implementation are covered in Chapter 12.

5.2 Transport Channels and Their Mapping to the Physical Channels

By the nature of the design already discussed, the LTE contains only common transport channels; a dedicated transport channel (Dedicated Channel, DCH, as in WCDMA) does not exist. The transport channels are the 'interface' between the Medium Access Control (MAC) layer

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6

and the physical layer. Each transport channel is characterized by the related physical layer processing applied to the corresponding physical channels used to carry the transport channel in question. The physical layer needs to be able to provide dynamic resource assignment both for data rate variance and for resource division between different users. This section presents the transport channels and their mapping to the physical channels.

- Broadcast Channel (BCH) is a downlink broadcast channel that is used to broadcast the necessary system parameters to enable devices accessing the system (and to identify the operator). Such parameters include, for example, random access related parameters that inform the device about which resource elements are reserved for random access operation.
- Downlink Shared Channel (DL-SCH) carries the user data for point-to-point connections in the downlink direction. All the information (either user data or higher layer control information) intended for only one user or UE is transmitted on the DL-SCH, assuming the UE is already in the RRC_CONNECTED state. As in LTE, however, the role of BCH is mainly for informing the device of the scheduling of the system information; control information intended for multiple devices is carried on DL-SCH as well. In case data on DL-SCH are intended for a single UE only, then dynamic link adaptation and physical layer retransmissions can be used.
- Paging Channel (PCH) is used for carrying the paging information for the device in the downlink direction to move the device from a RRC_IDLE state to a RRC_CONNECTED state.
- Multicast Channel (MCH) is used to transfer multicast service content to the UE in the downlink direction. 3GPP has decided to postpone the full support beyond Release 8.
- Uplink Shared Channel (uplink-SCH) carries the user data as well as device originated control information in the uplink direction in the RRC_CONNECTED state. Similar to the DL-SCH, dynamic link adaptation and retransmissions are available.
- Random Access Channel (RACH) is used in the uplink to respond to the paging message or to initiate the move from/to the RRC_CONNECTED state according to UE data transmission needs. There is no higher layer data or user data transmitted on RACH (such as can be done with WCDMA) but it is used just to enable uplink-SCH transmission where, for example, actual connection set-up with authentication, etc. will take place.

In the uplink direction the uplink-SCH is carried by the Physical Uplink Shared Channel (PUSCH). Correspondingly, the RACH is carried by the Physical Random Access Channel (PRACH). An additional physical channel exists but it is used only for physical layer control information transfer (as covered in connection with section 5.6 on control information). Uplink transport channel mapping to physical channels is illustrated in Figure 5.1.

In the downlink direction the PCH is mapped to the Physical Downlink Shared Channel (PDSCH). The BCH is mapped to Physical Broadcast Channel (PBCH), but as shown in Chapter 6 for the mapping of logical channels to transport channels, only part of the broadcasted param-

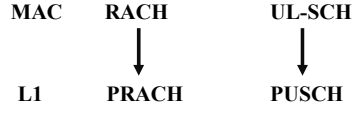


Figure 5.1 Mapping of the uplink transport channels to the physical channels



Figure 5.2 Mapping of the downlink transport channels to the physical channels

eters are on BCH while the actual System Information Blocks (SIBs) are then on DL-SCH. The DL-SCH is mapped to the PDSCH and MCH is mapped to the Physical Multicast Channel, as shown in Figure 5.2.

5.3 Modulation

In the uplink direction the modulation is the more traditional Quadrature Amplitude Modulation (QAM) modulator, as was explained in Chapter 4. The modulation methods available (for user data) are Quadrature Phase Shift Keying (QPSK), 16QAM and 64QAM. The first two are available in all devices while the support for 64QAM in the uplink direction is a UE capability, as covered in section 5.10. The different constellations are shown in Figure 5.3.

The PRACH modulation is phase modulation as the sequences used are generated from Zadoff–Chu sequences with phase differences between different symbols of the sequences (see section 5.7 for further details). Depending on the sequence chosen, the resulting Peak-to-Average Ratio (PAR) or the more practical Cubic Metric (CM) value is somewhat lower or higher compared to the QPSK value. In the uplink direction, the CM signal was discussed in Chapter 4 with SC-FDMA.

The use of QPSK modulation allows good transmitter power efficiency when operating at full transmission power as modulation determines the resulting CM (for SC-FDMA) and thus also the required device amplifier back-off. The devices will use lower maximum transmitter power when operating with 16QAM or 64QAM modulation.

In the downlink direction, the modulation methods for user data are the same as in the uplink direction. In theory an OFDM system could use different modulations for each sub-carrier. To have channel quality information (and signaling) with such a granularity would not be feasible due to the resulting excessive overhead. If modulation was sub-carrier specific, there would be too many bits in the downlink for informing the receiver of parameters for each sub-carrier and in the uplink the Channel Quality Indicator (CQI) feedback would need to be too detailed to achieve sub-carrier level granularity in the adaptation.

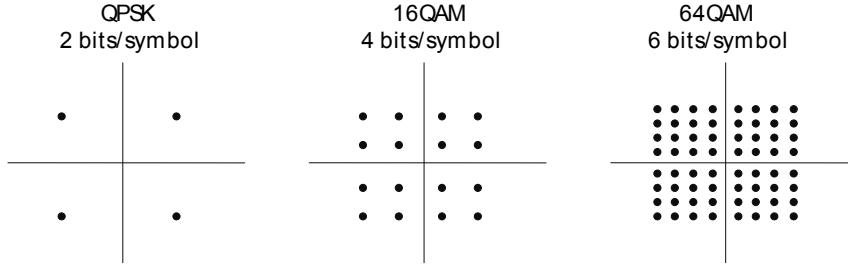


Figure 5.3 LTE modulation constellations

Also Binary Phase Shift Keying (BPSK) has been specified for control channels, which use either BPSK or QPSK for control information transmission. For a control channel, the modulation cannot be freely adapted as one needs to be able to receive them, and a single signaling error must not prevent detecting later control channel messages. This is similar to HSDPA/HSUPA where the control channels have fixed parameterization to prevent error propagation due to frame loss events. The exception is the uplink control data when multiplexed together with the user data – there modulation for data and control is the same – even if 16QAM or 64QAM would be used. This allows the multiplexing rules to be kept simpler.

5.4 Uplink User Data Transmission

The user data in the uplink direction is carried on the PUSCH, which has a 10 ms frame structure and is based on the allocation of time and frequency domain resources with 1 ms and 180 kHz resolution. The resource allocation comes from a scheduler located in the eNodeB, as illustrated in Figure 5.4. Thus there are no fixed resources for the devices, and without prior signaling from the eNodeB only random access resources may be used. For this purpose the device needs to provide information for the uplink scheduler of the transmission requirements (buffer status) it has as well as on the available transmission power resources. This signaling is MAC layer signaling and is covered in detail in Chapter 6.

The frame structure adopts the 0.5 ms slot structure and uses the 2 slot (1 subframe) allocation period. The shorter 0.5 ms allocation period (as initially planned in 3GPP to minimize the round trip time) would have been too signal intensive especially with a large number of users. The 10 ms frame structure is illustrated in Figure 5.5. The frame structure is basically valid for both for FDD and TDD, but TDD mode has additional fields for the uplink/downlink transition point(s) in the frame, as covered in Chapter 12.

Within the 0.5 ms slot there are both reference symbols and user data symbols, in addition to the signaling, covered in a later section. The momentary user data rate thus varies as a function of the uplink resource allocation depending on the allocated momentary bandwidth. The

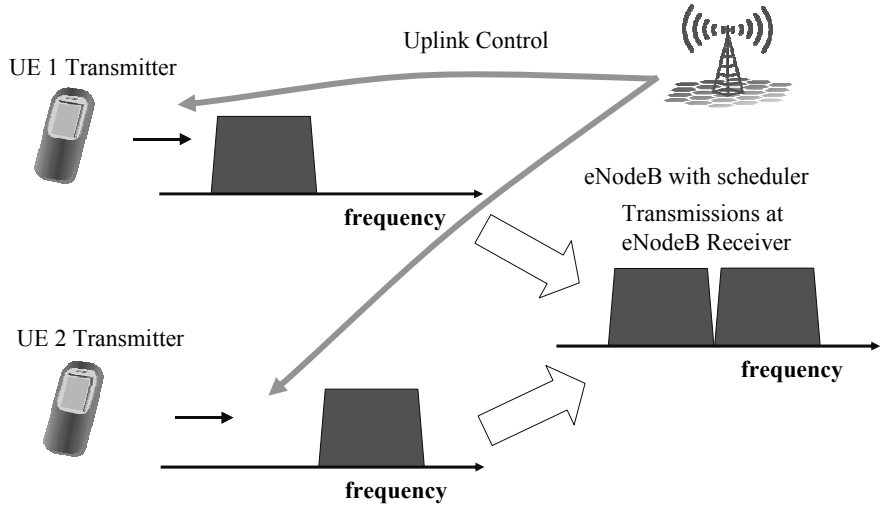


Figure 5.4 Uplink resource allocation controlled by eNodeB scheduler

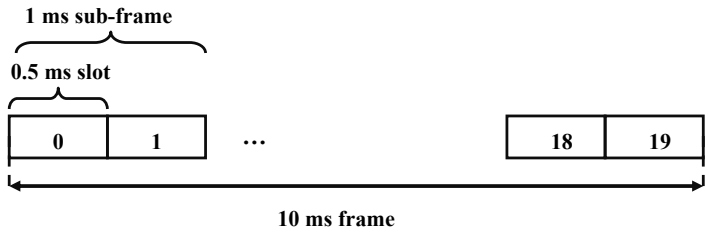


Figure 5.5 LTE FDD frame structure

allocation bandwidth may be between 0 and 20 MHz in the steps of 180 kHz. The allocation is continuous as uplink transmission is FDMA modulated with only one symbol being transmitted at a time. The slot bandwidth adjustment between consecutive TTIs is illustrated in Figure 5.6, where doubling the data rate results in double the bandwidth being used. The reference symbols always occupy the same space in the time domain and thus a higher data rate results in a corresponding increase for the reference symbol data rate.

The cyclic prefix used in uplink has two possible values depending on whether a short or extended cyclic prefix is applied. Other parameters stay unchanged and thus the 0.5 ms slot can accommodate either six or seven symbols as indicated in Figure 5.7. The data payload is reduced if an extended cyclic prefix is used, but it is not used frequently as usually the performance benefit in having seven symbols is far greater than the possible degradation from inter-symbol interference due to channel delay spread longer than the cyclic prefix.

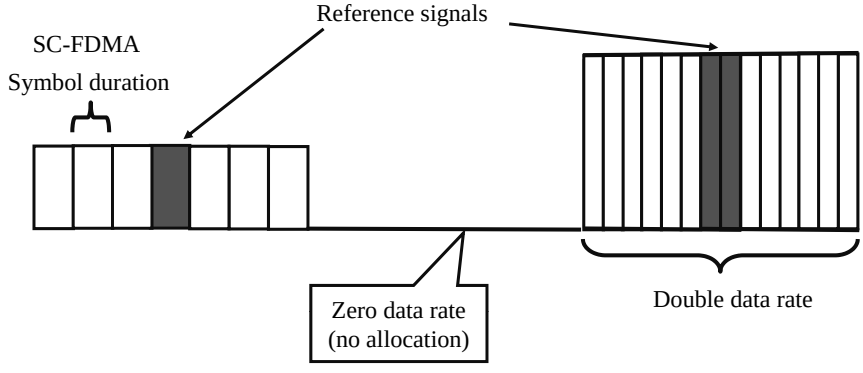


Figure 5.6 Data rate between TTIs in the uplink direction

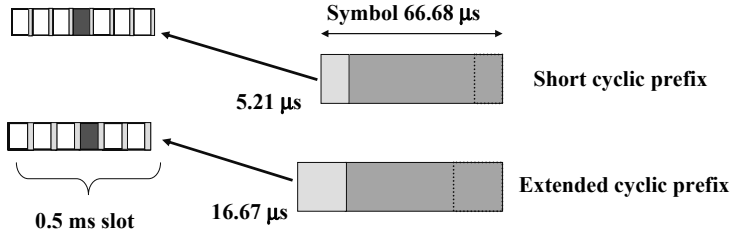


Figure 5.7 Uplink Slot structure with short and extended cyclic prefix

The resulting instantaneous uplink data rate over a 1 ms subframe is a function of the modulation, the number of resource blocks allocated, and the amount of control information overhead as well as of the rate of channel coding applied. The range of the instantaneous uplink peak data rate when calculated from the physical layer resources is between 700 kbps and 86 Mbps. There is no multi-antenna uplink transmission specified in Release 8, as using more than one transmitter branch in a UE is not seen as that attractive from the cost and complexity perspective. The instantaneous data rate for one UE depends on the LTE uplink from:

- Modulation method applied, with 2, 4 or 6 bits per modulations symbol depending on the modulation order for QPSK, 16QAM and 64QAM respectively.
- Bandwidth applied. For 1.4 MHz, the overhead is the largest due to the common channels and synchronization signals. The momentary bandwidth may of course vary between the minimum allocation of 12 sub-carriers (one resource block of 180 kHz) and the system bandwidth, up to 1200 sub-carriers with a 20 MHz bandwidth.
- Channel coding rate applied.
- The average data rate then depends on the time domain resource allocation as well.

The cell or sector specific maximum total data throughput can be increased with the Virtual Multiple Input Multiple Output (V-MIMO). In V-MIMO the eNodeB will treat transmission from two different UEs (with a single transmit antenna each) as one MIMO transmission and separate the data streams from each other based on the UE specific uplink reference symbol sequences. Thus V-MIMO does not contribute to the single user maximum data rate. The maximum data rates taking into account the UE categories are presented in section 5.10, while the maximum data rates for each bandwidth are covered in Chapter 9.

The channel coding chosen for LTE user data was turbo coding. The encoder is Parallel Concatenated Convolution Coding (PCCC) type turbo encoder, exactly the same as in WCDMA/HSPA, as explained in [5]. The turbo interleaver of WCDMA was modified to better fit LTE properties and slot structures and also to allow more flexibility for implementation of parallel signal processing with increased data rates.

LTE also uses physical layer retransmission combining, often referred to as Hybrid Adaptive Repeat and Request (HARQ). In a physical layer HARQ operation the receiver also stores the packets with failed CRC checks and combines the received packet when a retransmission is received. Both soft combining with identical retransmissions and combining with incremental redundancy are facilitated.

The channel coding chain for uplink is shown in Figure 5.8, where the data and control information are separately coded and then mapped to separate symbols for transmission. As the control information has specific locations around the reference symbols, the physical layer control information is separately coded and placed in a predefined set of modulation symbols (but with the same modulation as if the data were transmitted together). Thus the channel interleaver in Figure 5.8 does not refer to truly joint interleaving between control and data.

The data and control information are time multiplexed in the resource element level. Control is not evenly distributed but intended to be either closest for the reference symbols in time domain or then filled in the top rows of Figure 5.9, depending on the type of control information (covered in section 5.6). Data are modulated independently of the control information, but modulation during the 1 ms TTI is the same.

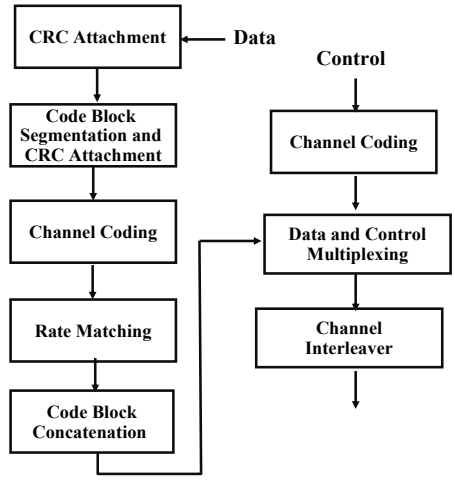


Figure 5.8 PUSCH Channel Coding Chain

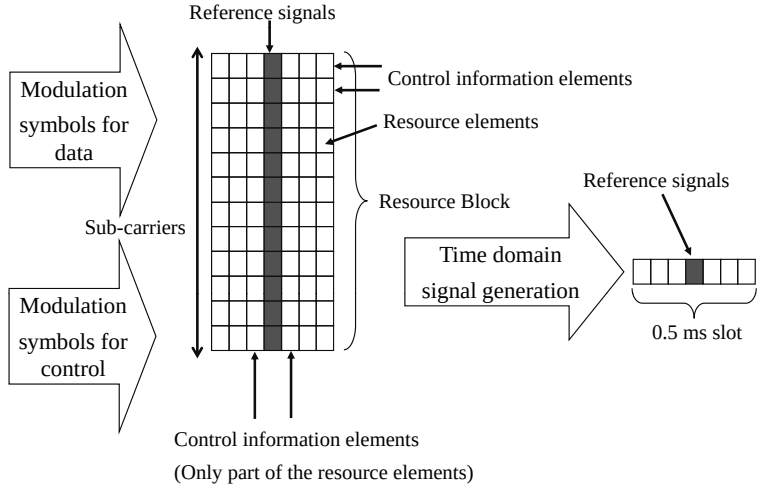


Figure 5.9 Multiplexing of uplink control and data

5.5 Downlink User Data Transmission

The user data rate in the downlink direction is carried on the Physical Downlink Shared Channel (PDSCH). The same 1 ms resource allocation is also valid in the downlink direction. The sub-carriers are allocated to resource units of 12 sub-carriers resulting in 180 kHz allocation units (Physical Resource Blocks, PRBs). With PDSCH, however, as the multiple access is OFDMA, each sub-carrier is transmitted as a parallel 15 kHz sub-carrier and thus the user data rate is dependent on the number of allocated sub-carriers (or resource blocks in practice) for a given user. The eNodeB carries out the resource allocation based on the Channel Quality Indicator (CQI) from the terminal. Similarly to the uplink, the resources are allocated in both the time and the frequency domain, as illustrated in Figure 5.10.

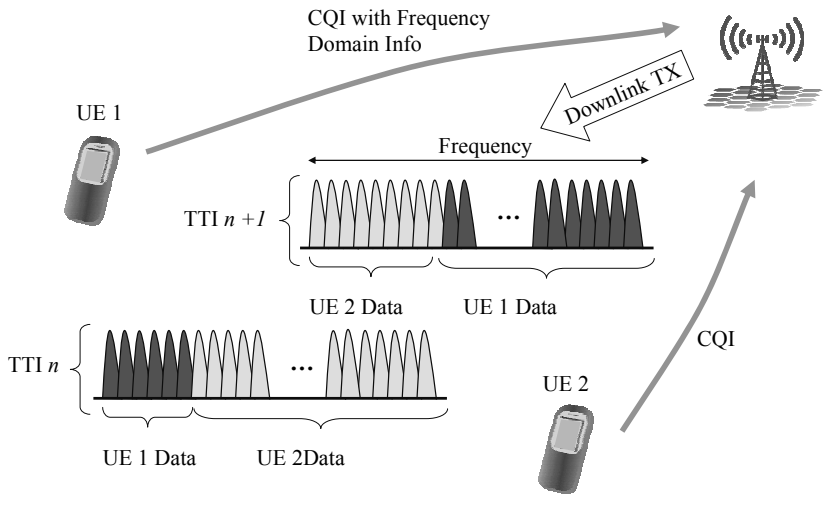


Figure 5.10 Downlink resource allocation at eNodeB

The Physical Downlink Control Channel (PDCCH) informs the device which resource blocks are allocated to it, dynamically with 1 ms allocation granularity. PDSCH data occupy between 3 and 6 symbols per 0.5 ms slot depending on the allocation for PDCCH and depending whether a short or extended cyclic prefix is used. Within the 1 ms subframe, only the first 0.5 ms slot contains PDCCH while the second 0.5 ms slot is purely for data (for PDSCH). For an extended cyclic prefix, 6 symbols are accommodated in the 0.5 ms slot, while with a short cyclic prefix 7 symbols can be fitted, as shown in Figure 5.11. The example in Figure 5.11 assumes there are 3 symbols for PDCCH but this can vary between 1 and 3. With the smallest bandwidth of 1.4 MHz the number of symbols varies between 2 and 4 to enable sufficient signaling capacity and enough bits to allow for good enough channel coding in range critical cases.

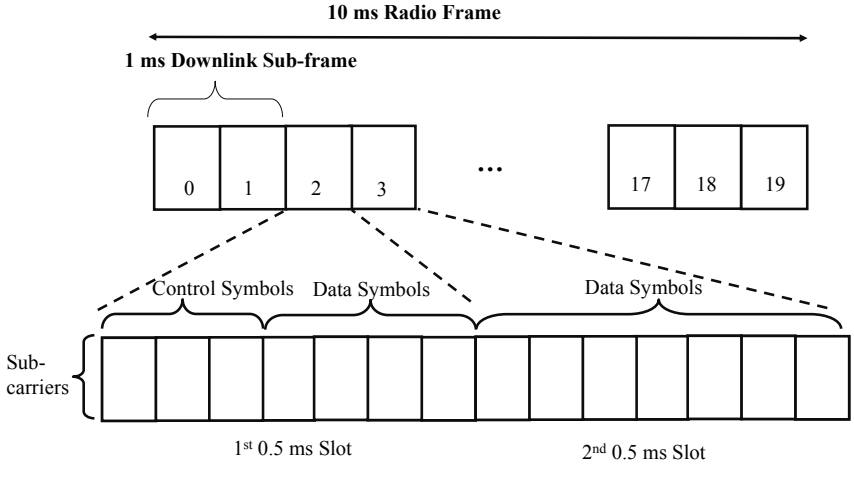


Figure 5.11 Downlink slot structure for bandwidths above 1.4 MHz

In addition to the control symbols for PDCCH, space from the user data is reduced due to the reference signals, synchronization signals and broadcast data. As discussed in Chapter 4, due to channel estimation it is beneficial when the reference symbols are distributed evenly in the time and frequency domains. This reduces the overhead needed, but requires some rules to be defined so that both receiver and transmitter understand the resource mapping in a similar manner. From the total resource allocation space over the whole carrier one needs to account for common channels, such as PBCH, that consume their own resource space. In Figure 5.12 an example of PDCCH and PDSCH resource allocation is presented. Note that the reference symbol placement in Figure 5.12 is purely illustrative and does not represent an actual reference symbol pattern.

The channel coding for user data in the downlink direction was also 1/3-rate turbo coding, as in the uplink direction. The maximum block size for turbo coding is limited to 6144 bits to reduce the processing burden, higher allocations are then segmented to multiple encoding blocks. Higher block sizes would not add anything to performance as the turbo encoder performance improvement effect for big block sizes has been saturated much earlier. Besides the turbo coding, downlink also has the physical layer HARQ with the same combining methods as in the uplink direction. The device categories also reflect the amount of soft memory available for retransmission combining. The downlink encoding chain is illustrated in Figure 5.13. There is no multiplexing to the same physical layer resources with PDCCH as they have their own separate resources during the 1 ms subframe.

Once the data have been encoded, the code words are provided onwards for scrambling and modulation functionality. The scrambling in the physical layer should not be confused with ciphering functionality but is just intended to avoid the wrong device successfully decoding the data should the resource allocations happen to be identical between cells. The modulation mapper applies the desired modulation (QPSK, 16QAM or 64QAM) and then symbols are fed for layer mapping and pre-coding. For multiple transmit antennas (2 or 4) the data are then divided into as many different streams and then mapped to correct resource elements available for PDSCH and then the actual OFDMA signal is generated, as shown in Figure 5.14 with an example of 2 antenna transmission. Should there be only a single transmit antenna available,

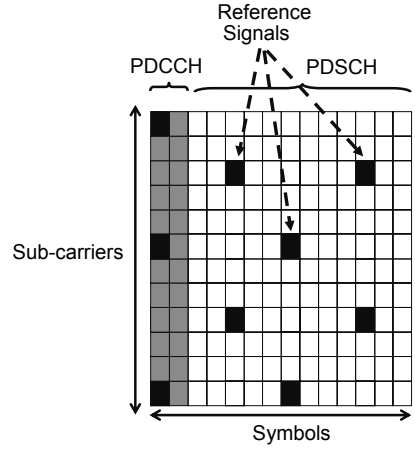


Figure 5.12 Example of downlink resource sharing between PDCCH and PDSCH

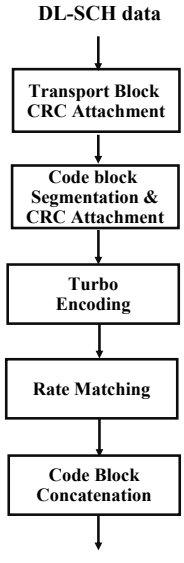


Figure 5.13 DL-SCH Channel Encoding Chain

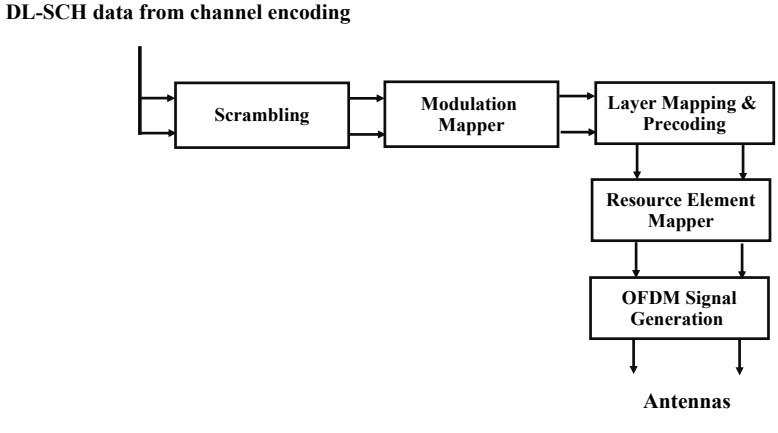


Figure 5.14 Downlink signal generation

then obviously the layer mapping and pre-coding functionalities do not have a role in signal transmission.

The resulting instantaneous data rate for downlink depends on:

- Modulation, with the same methods possible as in the uplink direction.
- Allocated amount of sub-carriers. Note that in the downlink the resource blocks are not necessary having continuous allocation in the frequency domain. The range of allocation is the same as in the uplink direction from 12 sub-carriers (180 kHz) up to the system bandwidth with 1200 sub-carriers.
- Channel encoding rate.
- Number of transmit antennas (independent streams) with MIMO operation.

The instantaneous peak data rate for downlink (assuming all resources to a single user and counting only the physical layer resources available) ranges between 0.7 Mbps and 170 Mbps. Even 300 Mbps or higher could be expected if using 4-by-4 antenna MIMO operation. There is no limit on the smallest data rate, and should the smallest allocation unit (1 resource block) be too high, then padding could be applied. Section 5.10 presents the maximum data rates taking the UE categories into account. The possible data rates for different bandwidth/coding/modulation combinations are presented in Chapter 9.

5.6 Uplink Physical Layer Signaling Transmission

Uplink Layer 1/Layer 2 (L1/L2) control signaling is divided into two classes in the LTE system:

- control signaling in the absence of uplink data, which takes place on PUCCH (Physical Uplink Control Channel);
- control signaling in the presence of uplink data, which takes place on PUSCH (Physical Uplink Shared Channel).

Due to single carrier limitations, simultaneous transmission of PUCCH and PUSCH is not allowed. This means that separate control resources are defined for the cases with and without uplink data. Alternatives considered were parallel transmission in the frequency domain (bad for the transmitter envelope) or pure time division (bad for control channel coverage). The selected approach maximizes the link budget for PUCCH and always maintains the single carrier properties on the transmitted signal.

PUCCH is a shared frequency/time resource reserved exclusively for User Equipment (UE) transmitting only L1/L2 control signals. PUCCH has been optimized for a large number of simultaneous UEs with a relatively small number of control signaling bits per UE.

PUSCH carries the uplink L1/L2 control signals when the UE has been scheduled for data transmission. PUSCH is capable of transmitting control signals with a large range of supported signaling sizes. Data and different control fields such as ACK/NACK and CQI are separated by means of Time Division Multiplexing (TDM) by mapping them into separate modulation symbols prior to the Discrete Fourier Transform (DFT). Different coding rates for control are achieved by occupying a different number of symbols for each control field.

There are two types of uplink L1 and L2 control-signaling information, as discussed in [6]:

- data-associated signaling (e.g. transport format and HARQ information), which is associated with uplink data transmission;
- data-non-associated signaling (ACK/NACK due to downlink transmissions, downlink CQI, and scheduling requests for uplink transmission).

It has been decided that there is no data-associated control signaling in the LTE uplink. Furthermore, it is assumed that eNodeB is not required to perform blind transport format detection. Basically this means that UE just obeys the uplink scheduling grant with no freedom in transport format selection. Furthermore, there is a new data indicator (1 bit) together with implicit information about the redundancy version included in the uplink grant [7]. This guarantees that the eNodeB always has exact knowledge about the uplink transport format.

5.6.1 Physical Uplink Control Channel (PUCCH)

From the single UE perspective, PUCCH consists of a frequency resource of one resource block (12 sub-carriers) and a time resource of one subframe. To handle coverage-limited situations, transmission of ACK/NACK spans the full 1 ms subframe. Furthermore, to support situations where coverage is extremely limited it has been agreed that ACK/NACK repetition is supported in the LTE uplink. Slot-based frequency hopping on the band edges symmetrically over the center frequency is always used on PUCCH, as shown in Figure 5.15. Frequency hopping provides the necessary frequency diversity needed for delay critical control signaling.

Different UEs are separated on PUCCH by means of Frequency Division Multiplexing (FDM) and Code Division Multiplexing (CDM). FDM is used only between the resource blocks whereas CDM is used inside the PUCCH resource block.

Two ways to realize CDM inside the PUCCH resource block are:

- CDM by means of cyclic shifts of a Constant Amplitude Zero Autocorrelation Codes (CAZAC)[1] sequence;
- CDM by means of block-wise spreading with the orthogonal cover sequences.

The main issue with CDM is the well-known near–far problem. Orthogonality properties of the considered CDM techniques were carefully studied during the Work Item phase of LTE standardization. We note that:

- channel delay spread limits the orthogonality between cyclically shifted CAZAC sequences;
- channel Doppler spread limits the orthogonality between block-wise spread sequences.

Orthogonality properties are optimized by means of a staggered and configurable channelization arrangement (see more details in section 5.6.2), proper configuration of block spreading, and a versatile randomization arrangement including optimized hopping patterns used for the control channel resources and the applied CAZAC sequences.

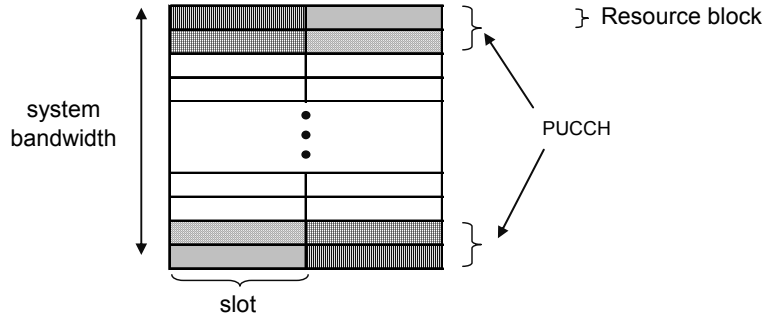


Figure 5.15 PUCCH resource

[1] The applied sequences are not true CAZAC but computer searched Zero-Autocorrelation (ZAC) sequences. The same sequences are applied as reference signals with a bandwidth allocation of one resource block.

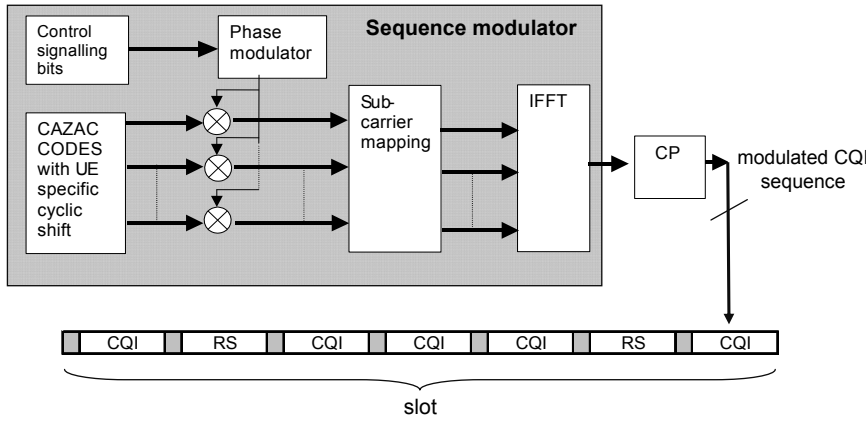


Figure 5.16 Block diagram of CAZAC sequence modulation applied for CQI

5.6.1.1 Sequence Modulation

Control signaling on PUCCH is based on sequence modulation. Cyclically shifted CAZAC sequences take care of both CDM and conveying the control information. A block diagram of the sequence modulator configured to transmit periodic CQI on PUCCH is shown in Figure 5.16. On the PUCCH application CAZAC sequences with a length of 12 symbols (1 resource block) are BPSK or QPSK modulated, thus carrying one or two information bits per sequence. Different UEs can be multiplexed into the given frequency/time resource by allocating different cyclic shifts of the CAZAC sequence for them. There are six parallel channels available per resource block, assuming that every second cyclic shift is in use.

5.6.1.2 Block-wise Spreading

Block-wise spreading increases the multiplexing capacity of PUCCH by a factor of the spreading factor (SF) used. The principle of block-wise spreading is shown in Figure 5.17, which illustrates

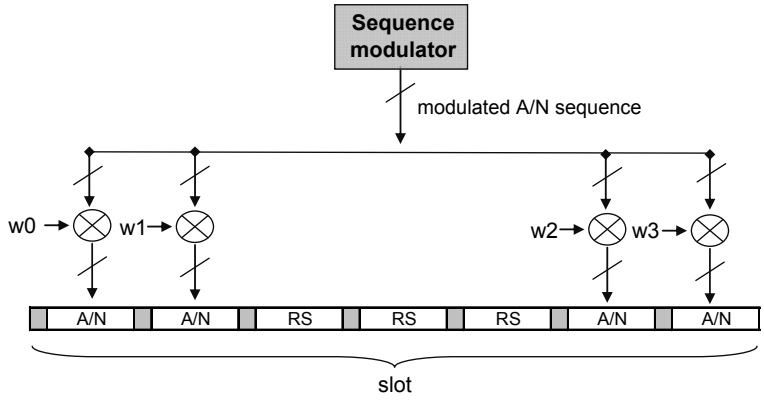


Figure 5.17 Principle of block spreading applied for ACK/NACK, spreading SF = 4

the block spreading operation made for an ACK/NACK data sequence transmitted on PUCCH. A separate block spreading operation is made for the reference signal and the ACK/NACK data parts but for simplicity, block processing related to the Reference Symbol (RS) part is neglected in Figure 5.17. In the example in Figure 5.17, the spreading factor applied for the ACK/NACK data and RS parts is equal to 4 and 3 respectively. Walsh–Hadamard codes are used as block spreading codes with SF=4 and SF=2, whereas DFT codes are used when SF=3.

5.6.1.3 PUCCH Formats

The available PUCCH formats are summarized in Table 5.1. PUCCH Format 1/1a/1b is based on the combination of CAZAC sequence modulation and block-wise spreading whereas PUCCH Format 2/2a/2b uses only CAZAC sequence modulation. As a result, Format 1/1a/1b can only carry one information symbol (1 to 2 bits) per slot while Format 2/2a/2b is capable of conveying 5 symbols per slot (20 coded bits + ACK/NACK per subframe). With Format 2/2a/2b, the CQI data are encoded using a punctured (20, *N*) Reed–Muller block code.

The supported control signaling formats were selected based on a careful evaluation process. The main issues of the evaluation phase were the link performance and multiplexing capacity as well as compatibility with other formats. It is also noted that the number of reference signal blocks were optimized separately for different formats.

Two different approaches were selected for signaling the ACK/NACK and CQI on PUCCH (Format 2a/2b):

- Normal cyclic prefix: ACK/NACK information is modulated in the second CQI reference signals of the slot. The RS modulation follows the CAZAC sequence modulation principle.
- Extended cyclic prefix: ACK/NACK bits and the CQI bits are jointly coded. No information is embedded in any of the CQI reference signals.

The main reason for having different solutions for normal and extended cyclic prefix lengths was that with an extended cyclic prefix there is only one reference signal per slot and hence the method used with the normal cyclic prefix cannot be used.

Support of Format 2a/2b is made configurable in the LTE uplink system. In order to guarantee ACK/NACK coverage, the eNodeB can configure a UE to drop the CQI when ACK/NACK and CQI would appear in the same subframe on PUCCH. In this configuration, Format 1a/1b is used instead of Format 2a/2b.

Table 5.1 PUCCH formats

PUCCH Formats	Control type	Modulation (data part)	Bit rate (raw bits/ subframe)	Multiplexing capacity (UE/RB)
1	Scheduling request	Unmodulated	– (on/off keying)	36, 18*, 12
1a	1-bit ACK/NACK	BPSK	1	36, 18*, 12
1b	2-bit ACK/NACK	QPSK	2	36, 18*, 12
2	CQI	QPSK	20	12, 6*, 4
2a	CQI + 1-bit ACK/NACK	QPSK	21	12, 6*, 4
2b	CQI + 2-bit ACK/NACK	QPSK	22	12, 6*, 4

*Typical value

5.6.1.4 Scheduling Request

One of the new features of the LTE uplink system is the support of fast uplink scheduling request mechanism for the active mode UE (RRC_CONNECTED) being synchronized by the eNodeB but having no valid uplink grant on PUSCH available. The supported scheduling request procedure is presented in Figure 5.18 [8].

The UE indicates the need for an uplink resource by a Scheduling Request Indicator (SRI). During the Release 8 LTE standardization process, the contention based synchronized RACH and non-contention based SRI mechanisms were compared. It was pointed out that a non-contention based approach is better suited to LTE uplink usage because it provides better coverage, a smaller system overhead and better delay performance than a non-contention based approach [9].

The SRI is transmitted using PUCCH Format 1. On-off keying based signaling is applied with SRI, i.e. only the positive SRI is transmitted. The positive SRI is transmitted using the ACK/NACK structure [1], the only difference between the SRI and the ACK/NACK formats is that with SRI the data part is not modulated. The benefit of this arrangement is that SRI and ACK/NACK can share the same physical resources.

5.6.2 PUCCH Configuration

Figure 5.19 shows the logical split between different PUCCH formats and the way in which the PUCCH is configured in the LTE specifications [1]. The number of resource blocks in a slot reserved for PUCCH transmission is configured by the N_{RB}^{HO}-parameter. This broadcasted system parameter can be seen as the maximum number of resource blocks reserved for PUCCH while actual PUCCH size changes dynamically based on Physical Control Format Indicator Channel (PCFICH) transmitted on the downlink control channel. The parameter is used to define the frequency hopping PUSCH region. The number of resource blocks reserved for periodic CQI (i.e. PUCCH Format 2/2a/2b) is configured by another system parameter, $N_{RB}^{(2)}$.

In general it makes sense to allocate separate PUCCH resource blocks for PUCCH Format 1/1a/1b and Format 2/2a/2b. With narrow system bandwidth options such as 1.4 MHz, however, this would lead to unacceptably high PUCCH overhead [10]. Therefore, sharing the PUCCH resources block between Format 1/1a/1b and Format 2/2a/2b users is supported in the LTE specifications. The mixed resource block is configured by the broadcasted system parameter $N_{CS}^{(1)}$, which is the number of cyclic shifts reserved for PUCCH Format 1/1a/1b on the mixed PUCCH resource block.

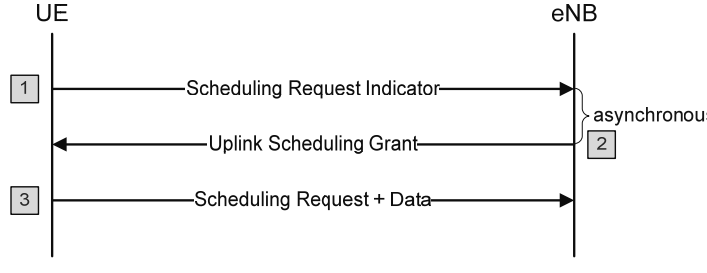


Figure 5.18 Scheduling request procedure

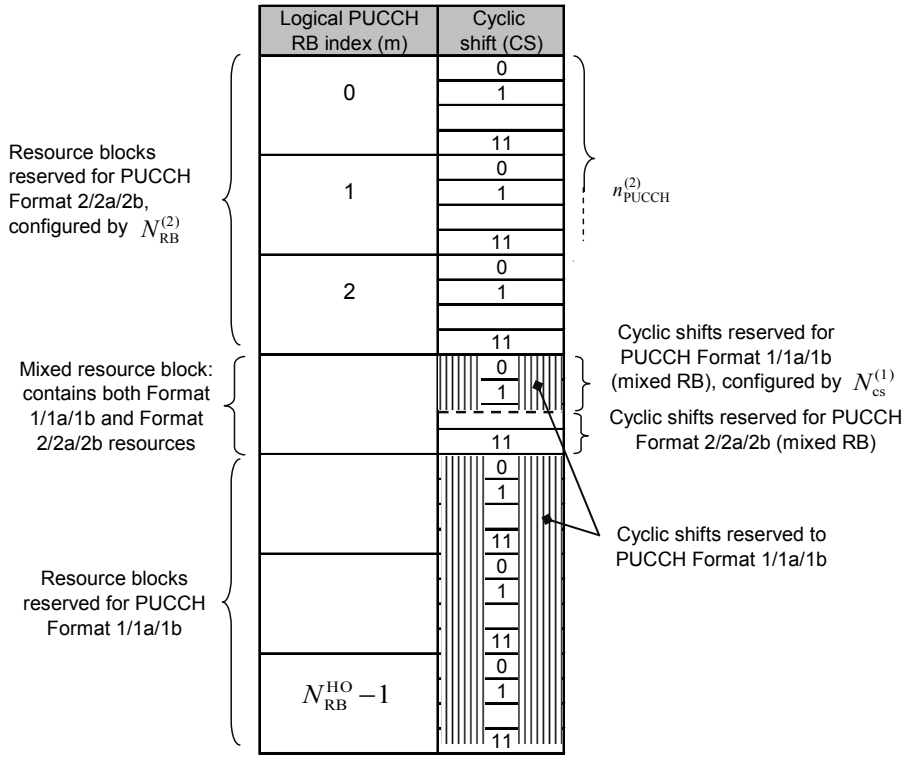


Figure 5.19 PUCCH configuration

Resources used for transmission of PUCCH Format 2/2a/2b are identified by a resource index $n_{PUCCH}^{(2)}$, which is mapped directly into a single CS resource. This parameter is explicitly signaled via UE-specific higher layer signaling.

5.6.2.1 Channelization and Resource Allocation for PUCCH Format 1/1a/1b

PUCCH Format 1/1a/1b resources are identified by a resource index $n_{PUCCH}^{(1)}$. Direct mapping between the PUCCH cyclic shifts and the resource indexes cannot be used with PUCCH Format 1/1a/1b due to the block spreading operation. Instead, PUCCH channelization is used to configure the resource blocks for PUCCH Format 1/1a/1b. The purpose of the channelization is to provide a number of parallel channels per resource block with optimized and adjustable orthogonality properties. Format 1/1a/1b channelization structure is configured by means of broadcasted system parameter, *Delta_shift*.

The number of PUCCH format 1/1a/1b resources per resource block, denoted as $N_{PUCCH\ Format\ 1}^{RB}$, can be calculated as follows:

$$N_{PUCCH\ Format\ 1}^{RB} = \frac{N_{RS}^{PUCCH} * 12}{Delta\_shift}, \tag{5.1}$$

where the *Delta_shift* parameter is the cyclic shift difference between two adjacent ACK/NACK resources using the same orthogonal cover sequence [11], and parameter $N_{\mathrm{RS}}^{\mathrm{PUCCH}}$ is the number of reference signals on PUCCH Format 1/1a/1b ($N_{\mathrm{RS}}^{\mathrm{PUCCH}}=3$ with normal CP and 2 with extended CP). Three values are allowed for the *Delta_shift* parameter, namely 1, 2 or 3. This means that the number of PUCCH Format 1/1a/1b resources per resource block equals 36, 18 or 12, with normal CP length.

An example of the PUCCH channelization within the resource block following the staggered resource structure is shown in Figure 5.20. In this example, *Delta_shift* is set to 2 and normal CP length is assumed.

The configuration of the PUCCH Format 1/1a/1b resource is shown in Figure 5.21. PUCCH Format 1/1a/1b resources are divided into available PUCCH resource blocks and are subject to

Cyclic shift	Orthogonal cover code		
	0	1	2
0	0		12
1		6	
2	1		13
3		7	
4	2		14
5		8	
6	3		15
7		9	
8	4		16
9		10	
10	5		17
11		11	

Figure 5.20 Principle of Format 1/1a/1b channelization within one resource block, *Delta_shift* = 2, normal CP

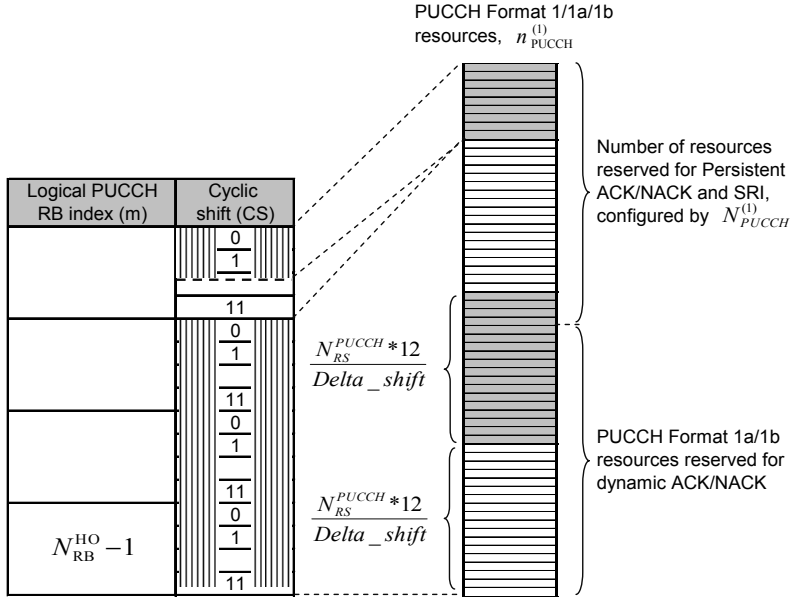


Figure 5.21 Configuration of PUCCH Format 1/1a/1b resource

channelization within the resource block as described earlier. Before that, the Format 1/1a/1b resource is split into persistent and dynamic parts. This is achieved using the broadcasted system parameter $N^{(1)}_{PUCCH}$, which is the number of resources reserved for persistent Format 1/1a/1b resources. These resources are used by the SRI and ACK/NACK related to persistently scheduled PDSCH. Both resources are allocated explicitly by resource index $n^{(1)}_{\mathrm{PUCCH}}$. Dynamic Format 1/1a/1b resources are placed at the end of logical PUCCH resources. Allocation of these ACK/NACK resources, which relate to dynamically scheduled PDSCH, is made implicitly based on the PDCCH allocation.

The idea of implicit allocation in dynamic ACK/NACK resources is to have one-to-one mapping to the lowest PDCCH Control Channel Element (CCE) index. The total number of CCEs depends on the system bandwidth and the number of OFDM symbols allocated for control signaling in a downlink subframe, which is signaled in each subframe using PCFICH (1, 2 or 3 OFDM symbols/subframe for bandwidths above 1.4 MHz, with 2, 3 or 4 OFDM symbols occupied for 1.4 MHz). There has to be a dedicated ACK/NACK resource for each CCE. This means that, for example, with a 20 MHz system bandwidth the number of CCEs can be up to 80 if three OFDM symbols are allocated for control signaling in a subframe.

5.6.2.2 Mapping of Logical PUCCH Resource Blocks into Physical PUCCH Resource Blocks

Mapping of logical resource blocks, denoted as *m,* into physical PUCCH resource blocks is shown in Figure 5.22. Taking into account the logical split between different PUCCH Formats, we note that PUCCH Format 2/2a/2b is located at the outermost resource blocks of the system bandwidth. ACK/NACK reserved for persistently scheduled PDSCH and SRI are located on the PUCCH resource blocks next to periodic CQI while the ACK/NACK resources reserved to dynamically scheduled PDSCH are located at the innermost resource blocks reserved for PUCCH.

An interesting issue from the system perspective is the fact that PUCCH defines the uplink system bandwidth. This is because PUCCH always exists and is located at both edges of the frequency spectrum. We note that proper PUCCH configuration allows narrowing down the active uplink system bandwidth by the resolution of two resource blocks. This can be made so that the PUCCH Format 2/2a/2b resource is over-dimensioned and at the same time the pre-defined PUCCH Format 2/2a/2b resources placed at the outermost resource blocks are left unused. Figure 5.23 shows the principle of this configuration.

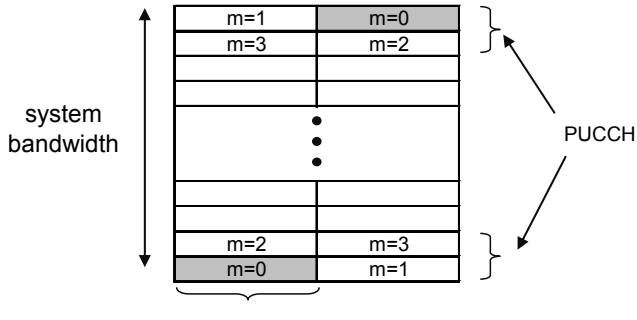


Figure 5.22 Mapping of logical PUCCH RBs into physical RBs [1]

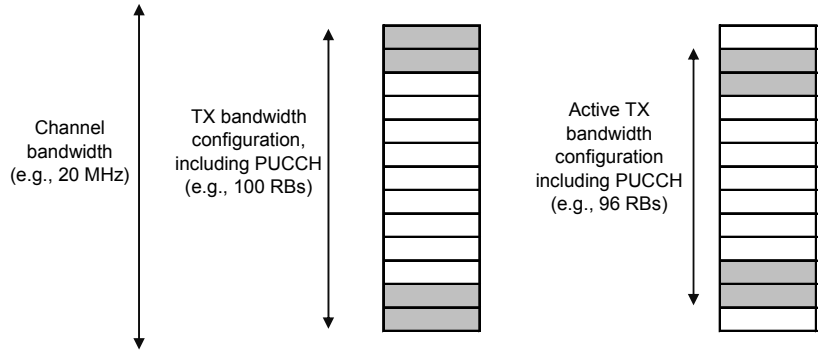


Figure 5.23 Changing the uplink system bandwidth via PUCCH configuration

5.6.3 Control Signaling on PUSCH

PUSCH carries the uplink L1/L2 control signals in the presence of uplink data. Control signaling is realized by a dedicated control resource, which is valid only during the uplink subframe when UE has been scheduled for data transmission on PUSCH. The main issues related to control signal design on PUSCH are:

- how to arrange multiplexing between uplink data and different control fields;
- how to adjust the quality of L1/L2 signals transmitted on PUSCH.

Figure 5.24 shows the principle of control and data multiplexing within the SC-FDMA symbol (block). To maintain the single carrier properties, transmitted signal data and different control symbols are multiplexed prior to the DFT. Data and different control fields (ACK/NAK, CQI/Pre-coding Matrix Indicator [PMI], Rank Indicator [RI]) are coded and modulated separately before multiplexing them into the same SC-FDMA symbol block. Block level multiplexing was also considered, but would have resulted in too large a control overhead [12]. Using the selected symbol level multiplexing scheme the ratio between the data symbols and control symbols can be accurately adjusted within each SC-FDMA block.

Figure 5.25 shows the principle of how uplink data and different control fields are multiplexed on the PUSCH. The actual mix of different L1/L2 control signals and their size vary from subframe to subframe. Both the UE and the eNodeB have the knowledge about the number of symbols reserved by the control part. The data part of PUSCH is punctured by the number of control symbols allocated in the given subframe.

Control and data multiplexing is performed so that control is present at both slots of the subframe. This guarantees that control channels can benefit from frequency hopping when it is

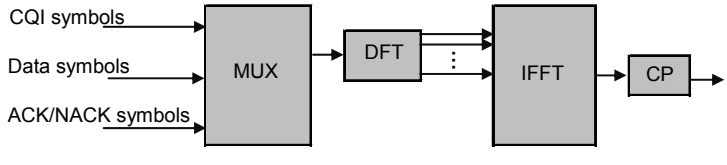


Figure 5.24 Principle of data and control modulation

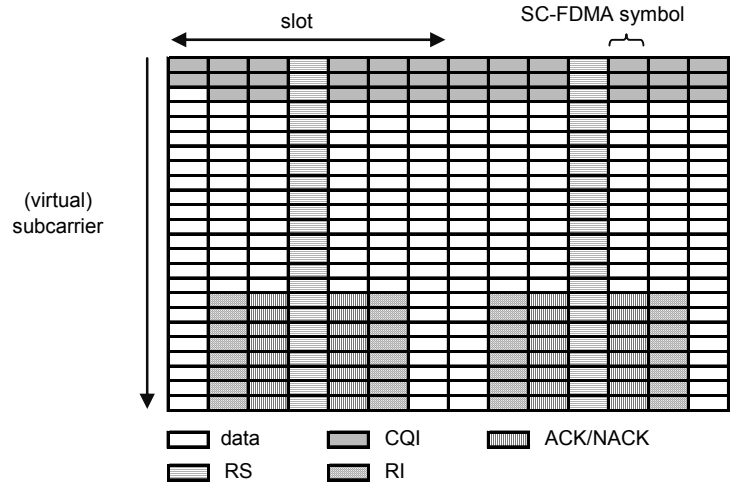


Figure 5.25 Allocation data and different control fields on PUSCH

applied. ACK/NACK is placed at the end of SC-FDMA symbols next to the reference signals. There is a maximum of 2 SC-FDMA symbols per slot allocated to ACK/NACK signaling. The same applies to RI, which is placed on the SC-FDMA symbols next to ACK/NACK. CQI/PMI symbols are placed at the beginning of the SC-FDMA symbols and they are spread over all the available SC-FDMA symbols.

CQI/PMI transmitted on PUSCH uses the same modulation scheme as the data part. ACK/NACK and RI are transmitted so that the coding, scrambling and modulation maximize the Euclidean distance at the symbol level. This means that a modulation symbol used for a ACK/NACK carrier is at most 2 bits of coded control information regardless of the PUSCH modulation scheme. The outermost constellation points having the highest transmission power are used to signal the ACK/NACK and RI for 16QAM and 64QAM. This selection provides a small power gain for ACK/NACK and RI symbols, compared to PUSCH data using higher order modulation.

Four different channel coding approaches are applied with control signals transmitted on PUSCH:

- repetition coding only: 1-bit ACK/NACK;
- simplex coding: 2-bit ACK/NACK/RI;
- (32, N) Reed–Muller block codes: CQI/PMI <11 bits;
- tail-biting convolutional coding (1/3): CQI/PMI ≥11 bits.

An important issue related to control signaling on PUSCH is how to keep the performance of control signaling at the target level. It is noted that power control will set the SINR target of PUSCH according to the data channel. Therefore, the control channel has to adapt to the SINR operation point set for data.

One way to adjust the available resources would be to apply different power offset values for data and different control parts. The problem of the power offset scheme is that single carrier properties are partially destroyed [13]. Therefore this scheme is not used in the LTE uplink

system. Instead, a scheme based on a variable coding rate for the control information is used. This is achieved by varying the number of coded symbols for control channel transmission. To minimize the overall overhead from the control signaling, the size of physical resources allocated to control transmission is scaled according to PUSCH quality. This is realized so that the coding rate to be used for control signaling is given implicitly by the Modulation and Coding Scheme (MCS) of PUSCH data. The linkage between data MCS and the size of the control field in made according to Equation 5.2 [14]:

$$O' = \left\lceil \frac{O \cdot offset \cdot M_{SC}^{PUSCH} \cdot N_{symb}^{PUSCH}}{K_{bits}^{PUSCH}} \right\rceil, \tag{5.2}$$

where O' is the number of coded control symbols for the given control type, O is the number of control signaling bits, K_{bits}^{PUSCH} is the number of transmitted bits after code block segmentation and $M_{SC}^{PUSCH} \cdot N_{symb}^{PUSCH}$ is the total number of sub-carriers per subframe carrying PUSCH. Offset is a semi-statically configured parameter related to the coding rate adjustment of control channel and is used to achieve a desired B(L)ER operation point for a given control signaling type. $\lceil \cdot \rceil$ is the operation rounding the control channel size to the nearest supported integer value, towards plus infinity.

As mentioned, the offset-parameter is used to adjust the quality of control signals for the PUSCH data channel. It is a UE-specific parameter configured by higher layer signaling. Different control channels need their own offset-parameter setting. There are some issues that need to be taken into account when configuring the offset-parameter:

- BLER operation point for the PUSCH data channel;
- B(L)ER operation point for the L1/L2 control channel;
- difference in coding gain between control and data parts, due to different coding schemes and different coding block sizes (no coding gain with 1-bit ACK/NACK);
- DTX performance.

Different BLER operation points for data and control parts are because HARQ is used for the data channels whereas the control channels do not benefit from HARQ. The higher the difference in the BLER operation point between data and control channels, the larger is the offset parameter (and vice versa). Similar behavior relates also to the packet size. The highest offset values are needed with ACK/NACK signals due to the lack of coding gain.

5.6.4 Uplink Reference Signals

In addition to the control and data signaling, there are reference signals as discussed in Section 5.4. The eNodeB needs to have some source of known data symbols to facilitate coherent detection, like in the WCDMA where uplink Physical Dedicated Control Channel (PDCCH) was carrying pilot symbols in the uplink direction. In LTE uplink, reference signals (RS) are used as demodulation reference signals (DM RS) on PUCCH and PUSCH. The new purpose of the reference signals, not part of WCDMA operation, is to use them as sounding reference signals (SRS). Additionally, reference signals are used for sequence modulation on PUCCH as discussed in section 5.6.1.1. The sequences used as reference signals are discussed in section

5.6.4.1, while demodulation reference signals and sounding reference signals are considered in sections 5.6.4.2 and 5.6.4.3, respectively.

5.6.4.1 Reference Signal Sequences

Considering first the sequences itself, the most important properties for the RS sequences in LTE uplink are:

- favorable auto- and cross-correlation properties;
- sufficient number of sequences;
- flat frequency domain representation facilitating efficient channel estimation;
- low cubic metric values comparable to the cubic metric of QPSK modulation.

The sequences also need to be suitable for supporting the numerous bandwidth options in uplink allocations. This means that sequences of various lengths, multiples of 12, are needed.

Constant Amplitude Zero Autocorrelation Codes (CAZAC) such as Zadoff–Chu [15] and Generalized Chirp-Like [16] polyphase sequences have most of the required properties. There exist also a reasonable number of Zadoff–Chu sequences when the sequence length is a prime number. However, the sequence lengths needed in LTE uplink are multiples of 12, for which only a modest number of Zadoff–Chu sequences exist. To obtain a sufficient number of RS sequences, computer generated sequences are used for sequence lengths of 12 and 24. They are constructed from QPSK alphabet in frequency domain. Longer sequences are derived from Zadoff–Chu sequences with length of prime number. They are circularly extended in frequency domain to the desired length. These sequences are frequently referred to as extended Zadoff–Chu sequences.

As a result, there are 30 reference signal sequences available for sequence lengths of 12, 24, and 36 and a larger number for longer sequence lengths. The RS sequences do not have constant amplitude in time and, thus, they are not actually CAZAC sequences. However, they have acceptable cubic metric values and zero autocorrelation and, thus, may be referred to as Zero Autocorrelation (ZAC) sequences.

As said, the RS sequences have a periodic autocorrelation function that is zero except for the zero shift value. In other words, the cyclic, or circular, shifts of a sequence, illustrated in Figure 5.26, are orthogonal to each other. This provides a convenient means to derive multiple orthogonal sequences from a single RS sequence, which is used in LTE to multiplex UE. To maintain the orthogonality, however, the time difference between the signals arriving at the base station should not exceed the time interval corresponding to the cyclic shift separation. To accommodate multi-path delay spread, the minimum time separation between the cyclic shifts available in LTE is 5.56 μs for Demodulation RS (DMRS) and 4.17 μs for Sounding RS (SRS). Correspondingly, there are 12 and 8 cyclic shifts specified for DMRS and SRS, respectively, with constant time separation between cyclic shifts irrespective of the reference signal bandwidth.

5.6.4.2 Demodulation Reference Signals

DMRS are primarily used for channel estimation needed for coherent detection and demodulation and it has the same bandwidth as the uplink data transmission. There is one DMRS in every

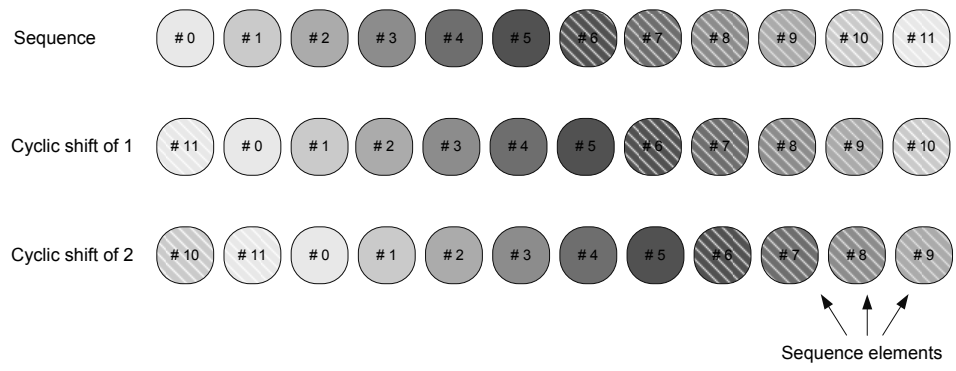


Figure 5.26 Cyclic shifts of a sequence

0.5 ms slot on PUSCH, whereas on PUCCH, there are 2–3 reference signals per slot depending on the used PUCCH format. For PUSCH, DMRS occupies the 4th SC-FDMA symbol in the slot, and the RS sequence length equals the number of allocated sub-carriers. The reference signal locations for PUCCH are discussed in section 5.6.1

As said, there are RS sequences of numerous lengths with 30 or more sequences for each sequence length. For simplicity, the sequences of various lengths are grouped into 30 sequence groups. Each sequence group contains RS sequences for all supported sequence lengths, with one RS sequence for allocations from 1 to 5 PRBs and two RS sequences for larger allocations. As a result, the used demodulation reference signal is defined with four parameters:

- sequence group, with 30 options – this is a cell specific parameter;
- sequence, with 2 options for sequence lengths of 6 PRBs or longer – this is a cell specific parameter;
- cyclic shift, with 12 options – this has both terminal and cell specific components, and 8 different values can be configured with the uplink allocation;
- sequence length, given by the uplink allocation.

Cyclic shifts are used to multiplex reference signals from different terminals within a cell, whereas different sequence groups are typically used in neighboring cells. While cyclic shifts provide good separation within a cell, a more complicated interference scenario is faced between cells. Simultaneous uplink allocations on neighboring cells can have different bandwidths and can be only partially overlapping in frequency. This alone prevents effective optimization of RS cross-correlations between cells. Thus, multiple hopping methods are included to LTE to randomize inter-cell interference for reference signals. The pseudo-random hopping patterns are cell specific and derived from the physical layer cell identity. For PUSCH and PUCCH, LTE supports:

- Cyclic shift hopping, which is always used. A cell specific cyclic shift is added on top of UE specific cyclic shifts. Cyclic shift hops for every slot on PUSCH. Inter-cell interference is expected to be more significant on PUCCH than on PUSCH due to CDM applied on PUCCH. To enhance inter-cell interference randomization, cyclic shift hops for every SC-FDMA symbol on PUCCH. Cyclic shift hopping is applied also for SC-FDMA symbols carrying control data due to the sequence modulation used on PUCCH.

- Sequence group hopping. Sequence group hopping pattern is composed of a group hopping pattern and a sequence-shift. The same group hopping pattern is applied to a cluster of 30 cells. To differentiate cells within a cluster, a cell specific sequence shift is added on top of the group hopping pattern. With this arrangement, the occasional use of the same sequence group simultaneously on neighboring cells is avoided within the cell cluster. Sequence group hopping can be also disabled, thus facilitating sequence planning. Sequence group hops for every slot.
- Sequence hopping, which means hopping between the two sequences within a sequence group. Sequence hopping can be applied for resource allocations larger than 5 RBs if sequence group hopping is disabled and sequence hopping is enabled.

On PUSCH, it is possible to configure cyclic shift hopping and sequence group hopping patterns so that the same patterns are used on neighboring cells. This means that the same sequence group is used on neighboring cells. This is not a feasible solution for PUCCH, however, due to a more intensive use of cyclic shifts. Thus, the hopping patterns are cell-specific and are derived from the cell identity on PUCCH. Therefore, a sequence group is configured separately for PUCCH and PUSCH with an additional configuration parameter for PUSCH.

5.6.4.3 Sounding Reference Signals

SRS is used to provide information on uplink channel quality on a wider bandwidth than the current PUSCH transmission or when a terminal has no transmissions on PUSCH. Channel is estimated on eNodeB and the obtained channel information can be used in the optimization of uplink scheduling as part of the uplink frequency domain scheduling operation. Thus, SRS is in a sense an uplink counterpart for the CQI reporting of downlink channel. SRS can also be used for other purposes, e.g. to facilitate uplink timing estimation for terminals with narrow or infrequent uplink transmissions. SRS is transmitted on the last SC-FDMA symbol of the subframe, as shown in Figure 5.27. Note that the SRS transmission does not need to be in the frequency area used by the PUSCH for actual data transmission but it may locate elsewhere as well.

On SRS, distributed SC-FDMA transmission is used. In other words, UE uses every second sub-carrier for transmitting the reference signal, as illustrated in Figure 5.28. Related sub-carrier offset defines a transmission comb for the distributed transmission. The transmission comb provides another means to multiplex UE reference signals in addition to the cyclic shifts. SRS

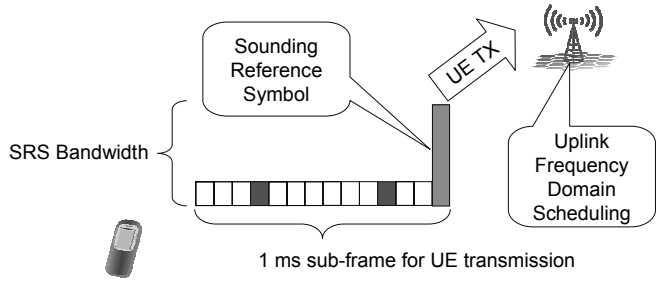


Figure 5.27 Sounding reference signal transmission in the frame

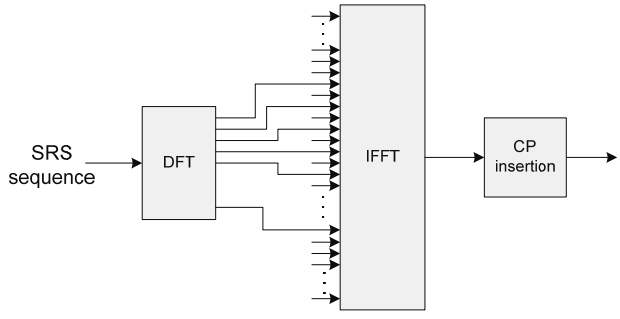


Figure 5.28 Sub-carrier mapping for sounding reference signal

uses the same sequences as DMRS. SRS sequence lengths are multiples of 24, or, correspondingly, SRS bandwidths are multiples of 4 RBs. This follows from the available RS sequence lengths combined with the definition of 8 cyclic shifts for SRS.

SRS transmissions can be flexibly configured. SRS transmission can be a single transmission or periodic with period ranging from 2 ms to 320 ms. There can be up to four different SRS bandwidth options available, depending on the system bandwidth and cell configuration. SRS transmission can also hop in frequency. This is particularly beneficial for the terminals on the cell edge, which cannot support wideband SRS transmissions. Frequency hopping can also be limited to a certain portion of system bandwidth which is beneficial for inter-cell interference coordination [17]. SRS configuration is explicitly signaled via terminal specific higher layer signaling.

Sounding reference signal transmissions from different terminals can be multiplexed in multiple dimensions:

- Time: Periodic SRS transmissions can be interleaved into different subframes with subframe offsets.
- Frequency: To facilitate frequency division multiplexing, the available SRS bandwidths follow a tree structure. This is illustrated in Figure 5.29, where a set of available SRS bandwidths is shown for a certain cell configuration. The SRS frequency hopping pattern also follows the tree structure, as shown in Figure 5.30 with an illustrative example based on the SRS bandwidth set of Figure 5.29.
- With cyclic shifts: Up to 8 cyclic shifts can be configured. The cyclic shift multiplexed signals, however, need to have the same bandwidth to maintain orthogonality. Due to the intensive use of cyclic shifts, the sequence group configured for PUCCH is used also for SRS.
- Transmission comb in the distributed transmission: Two combs are available. Contrary to cyclic shifts, transmission comb does not require that the multiplexed signals occupy the same bandwidth.

In addition to the terminal specific SRS configuration, cell specific SRS configuration defines the subframes that can contain SRS transmissions as well as the set of SRS bandwidths available in the cell. Typically SRS transmissions should not extend into the frequency band reserved for PUCCH. Therefore, multiple SRS bandwidth sets are needed for supporting flexible cell specific PUCCH configuration.

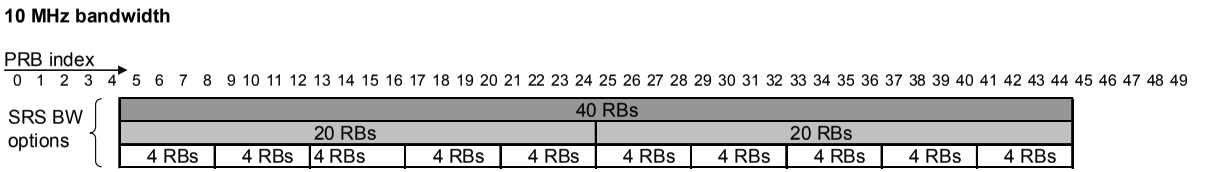


Figure 5.29 A SRS bandwidth set with a tree structure

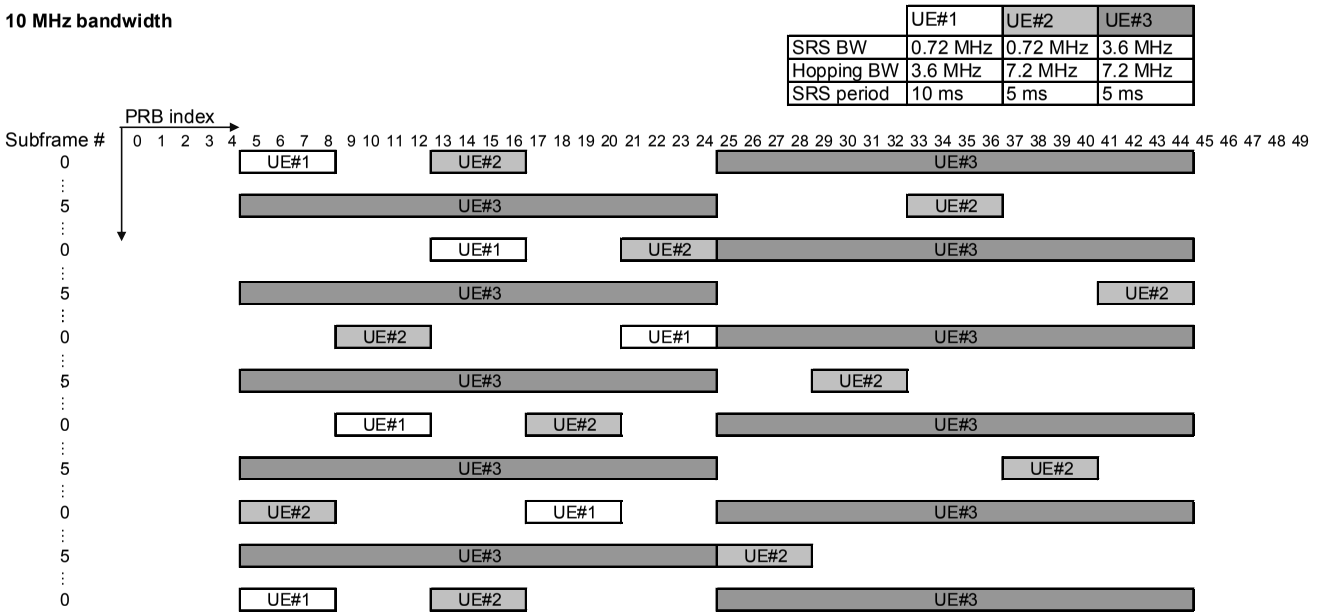


Figure 5.30 An example of SRS frequency hopping patterns

5.7 PRACH Structure

5.7.1 Physical Random Access Channel

Random access transmission is the only non-synchronized transmission in the LTE uplink. Although the terminal synchronizes to the received downlink signal before transmitting on RACH, it cannot determine its distance from the base station. Thus, timing uncertainty caused by two-way propagation delay remains on RACH transmissions.

Appropriately designed Physical Random Access Channel (PRACH) occurs reasonably frequently, provides a sufficient number of random access opportunities, supports the desired cell ranges in terms of path loss and uplink timing uncertainty, and allows for sufficiently accurate timing estimation. Additionally, PRACH should be configurable to a wide range of scenarios, both for RACH load and physical environment. For example, LTE is required to support cell ranges up to 100 km, which translates to 667 μs two-way propagation delay, as facilitated in the timing advance signaling range in the MAC layer.

In a LTE frame structure type 1 (FDD), only one PRACH resource can be configured into a subframe. The periodicity of PRACH resources can be scaled according to the expected RACH load, and PRACH resources can occur from every subframe to once in 20 ms. PRACH transmission is composed of a preamble sequence and a preceding cyclic prefix with four different formats as shown in Figure 5.31. Multiple preamble formats are needed due to the wide range of environments. For example, the long CP in preamble formats 1 and 3 assists with large cell ranges in terms of increased timing uncertainty tolerance whereas repeated preamble sequences in formats 2 and 3 compensate for increased path loss. The guard period that is necessary after an unsynchronized preamble is not explicitly specified, but PRACH location in the subframe structure provides a sufficient guard period. Particular considerations are needed only in very special cases. For each cell, 64 preamble sequences are configured and, thus, there are 64 random access opportunities per PRACH resource. PRACH occupies 1.08 MHz bandwidth, which provides reasonable resolution for timing estimation.

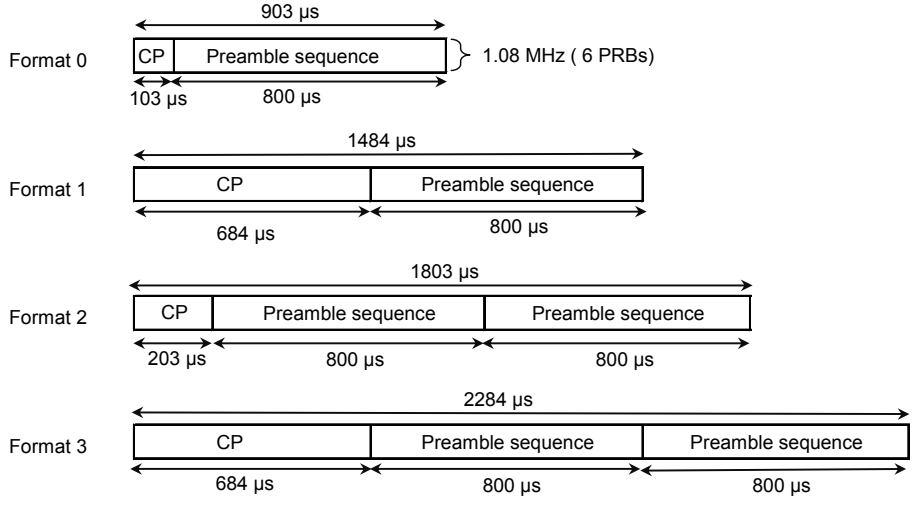


Figure 5.31 LTE RACH preamble formats for FDD

5.7.2 Preamble Sequence

Zadoff–Chu sequences [15] belonging to CAZAC are used as RACH preamble sequences due to several desirable sequence properties:

- They have a periodic autocorrelation function that is zero except for the zero shift value. This is desirable for preamble detection and timing estimation. Additionally, the cyclic, or circular, shifts of a sequence are orthogonal to each other. As a result, multiple preamble sequences are obtained from a single Zadoff–Chu sequence with cyclic shifts.
- As the preamble sequence length is set to a prime number of 839, there are 838 sequences with optimal cross-correlation properties.
- Sequences also have reasonable cubic metric properties.

The cyclic shifts used as different preambles need to have a sufficient separation. The cyclic shift separation needs to be sufficiently wide to accommodate the uplink timing uncertainty, as illustrated in Figure 5.32. The propagation delay and, thus, the cyclic separation are directly related to the cell range. To accommodate the wide range of cell ranges supported by LTE, 16 different cyclic shift separations can be configured for a cell, providing 1 to 64 preambles from a single Zadoff–Chu sequence.

A particular high speed mode, or a restricted set, is defined for RACH due to the peculiar properties of Zadoff–Chu sequences for Doppler spread. Essentially, high speed mode is a set of additional restrictions on the cyclic shifts that can be used as preambles.

Discrete Fourier Transform of Zadoff–Chu sequence is also a Zadoff–Chu sequence defined in frequency. Due to the Doppler, the transmitted preamble sequence spreads on cyclic shifts adjacent in frequency to the transmitted cyclic shift and particularly on the cyclic shift neighboring the transmitted cyclic shift in frequency. Particularly, a Doppler frequency corresponding to the inverse of sequence length, i.e. 1.25 kHz for LTE, transforms the transmitted cyclic shift (of sequence) completely to the cyclic shift neighboring the transmitted one in frequency. As a result, the received sequence is orthogonal with the transmitted sequence for a Doppler frequency of 1.25 kHz.

There is a tractable one-to-one relation between the cyclic shifts neighboring each other in frequency and the cyclic shifts in time. Thus, three uplink timing uncertainty windows can be defined for a preamble, with the windows separated by a shift distance as shown in Figure 5.33. The shift distance is a function of the Zadoff–Chu sequence index. The shifting of the timing uncertainty windows is circular, as illustrated in Figure 5.33.

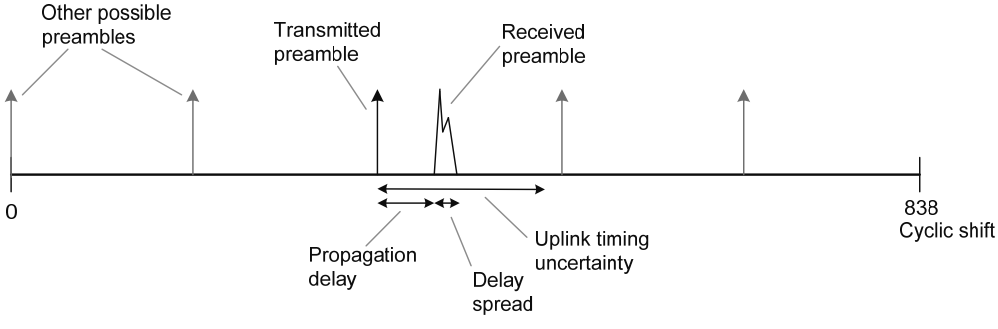


Figure 5.32 Illustration of cyclic shift separation N_{CS} between preamble sequences

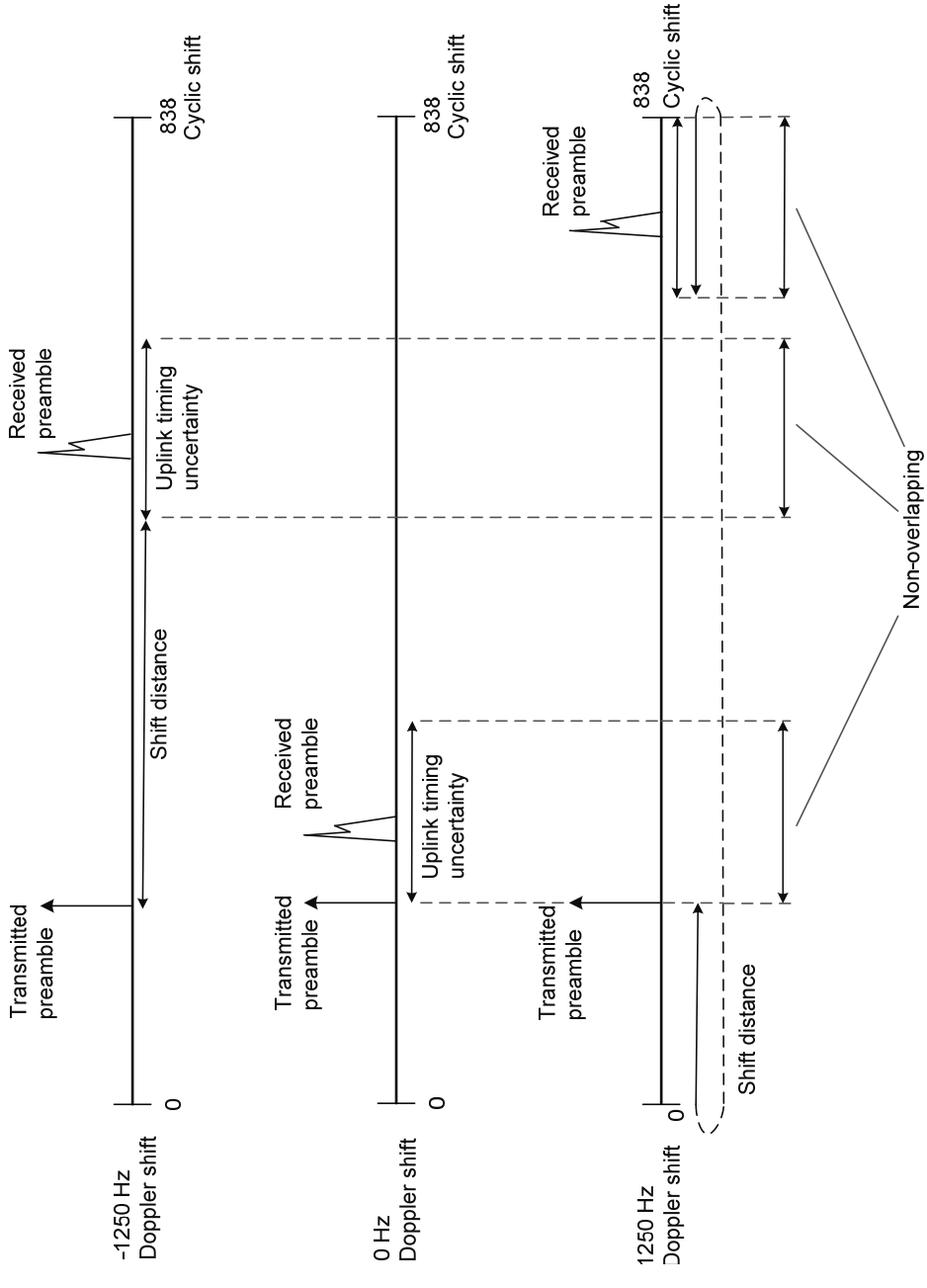


Figure 5.33 Illustration of uplink timing uncertainty windows for RACH preamble with considerable Doppler

There are several consequences of the aforementioned when detecting preambles with considerable Doppler:

- Signal energy needs to be collected from all three windows for reliable preamble detection.
- The windows of a preamble should not overlap each other to allow for initial uplink timing estimation.
- The windows of the different preambles should not overlap each other to prevent unnecessary false alarms.

As a result, the preamble cyclic shifts for high speed mode need to be selected so that the timing uncertainty windows do not overlap each other for each preamble as well as between preambles. Although these requirements, as well as the dependency of the shift distance on the sequence index, complicate the calculation of cyclic shift for preambles, tractable equations have been found and standardized [18].

5.8 Downlink Physical Layer Signaling Transmission

The control information in the downlink direction is carried using three different types of control messages:

- Control Format Indicator (CFI), which indicates the amount of resources devoted to control channel use. CFI is mapped to the Physical Control Format Indicator Channel (PCFICH).
- HARQ Indication (HI), which informs of the success of the uplink packets received. The HI is mapped on the Physical HARQ Indicator Channel (PHICH).
- Downlink Control Information (DCI), which controls with different formats basically all the physical layer resource allocation in both uplink and downlink direction and has multiple formats for different needs. The DCI is mapped on the Physical Downlink Control Channel (PDCCH)

5.8.1 Physical Control Format Indicator Channel (PCFICH)

The sole purpose of PCFICH is to dynamically indicate how many OFDMA symbols are reserved for control information. This can vary between 1 and 3 for each 1 ms subframe. From the PCFICH, UE knows which symbols to treat as control information. Location and modulation of PCFICH is fixed. The use of dynamic signaling capability allows the system to support both a large number of low data rate users (e.g. VoIP) as well as to provide sufficiently low signaling overhead when higher data rates are used by fewer simultaneously active users. The extreme situations are illustrated in Figure 5.34, where the PDCCH allocation is changed from 1 symbol to 3 symbols. When calculating the resulting overhead, note that PDCCH is only allocated to the first 0.5 ms slot in the 1 ms subframe, thus the overhead is from 1/14 to 3/14 of the total physical layer resource space.

With the 1.4 MHz operation, the PDCCH resource is 2, 3 or 4 symbols to ensure enough payload size and a sufficient range for all signaling scenarios. In big cells it is important to have enough room for channel coding together with signaling, especially for the operation with RACH.

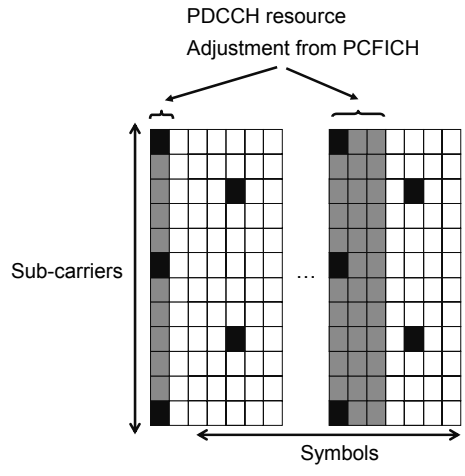


Figure 5.34 PDCCH resource allocation from PCFICH for bandwidths above 1.4 MHz

5.8.2 Physical Downlink Control Channel (PDCCH)

The UE will obtain from the PDCCH information for both uplink and downlink resource allocations the UE may use. The DCI mapped on the PDCCH has different formats and depending on the size DCI is transmitted using one or more Control Channel Elements (CCEs). A CCE is equal to 9 resource element groups. Each group in turn consists of 4 resource elements. The different PDCCH formats are shown in Table 5.2, where it can be seen that as PDCCH is using QPSK modulation, then a single resource element carries 2 bits and there are 8 bits in a resource element group.

The UE will listen to the set of PDCCHs and tries to decode them (checking all formats) in all subframes except during the ones where DRX is configured. The set of PDCCHs to monitor is up to 6 channels. Depending on the network parameterization, some of the PDCCHs are so-called common PDCCHs and may also contain power control information.

The DCI mapped to PDCCH has four different formats and further different variations for each format. It may provide the control information for the following cases:

- PUSCH allocation information (DCI Format 0);
- PDSCH information with one codeword (DCI Format 1 and its variants);
- PDSCH information with two codewords (DCI Format 2 and its variants);.
- Uplink power control information (DCI Format 3 and its variants)

Table 5.2 PDCCH format and their size

PDCCH format	Number of CCEs	Number of resource-element groups	Number of PDCCH bits
0	1	9	72
1	2	18	144
2	4	36	288
3	8	72	576

The PDCCH containing PDSCH related information is often referred to as the downlink assignment. The following information is carried on the downlink assignment when providing downlink resource allocation information related to PDSCH:

- Resource block allocation information. This indicates the position of the resources allocated for the user in question in the resource block domain. The allocation can be based on, for example, a bitmap pointing the given PRBs or an index of the first PRB and the number of contiguously allocated PRBs.
- The modulation and coding scheme used for downlink user data. The 5 bit signaling indicates the modulation order and the transport block size (TBS). Based on these parameters and the number of allocated resource blocks the coding rate can be derived.
- The HARQ process number needs to be signaled as the HARQ retransmission from the eNodeB point of view is asynchronous and the exact transmission instant is up to the eNodeB scheduler functionality. Without the process number the UE could confuse the different processes and combine the wrong data. This also prevents error propagation from this part if control signaling is lost for a single TTI. The number of HARQ processes was fixed to 8 in both uplink and downlink.
- A new data indicator to tell whether the transmission for the particular process is a retransmission or not. This again follows similar principles to those applied with HSDPA.
- Redundancy version is a HARQ parameter that can be used with incremental redundancy to tell which retransmission version is used.
- The power control commands for the PUCCH are also included on the PDCCH. The power control command has two bits and it can thus use 2 steps up and downwards to adjust the power.

Additionally, when MIMO operation is involved, there are MIMO specific signaling elements involved, as discussed with the MIMO in section 5.9.7.

The PDCCH containing PUSCH related information is also known as the uplink grant. The following information is carried on the uplink grant.

- Hopping flag and resource block assignment and hopping resource allocation. The number of bits for this depends on the bandwidth to be used. Uplink resource allocation is always contiguous and it is signaled by indicating the starting resource block and the size of the allocation in terms of resource blocks.
- Modulation and coding scheme and the redundancy version.
- A new data indicator, which is intended to be used for synchronizing the scheduling commands with the HARQ ACK/NACK message status.
- TPC command for scheduled PUSCH which can represent four different values.
- Cyclic shift for the Demodulation Reference Symbols – 3 bits.
- Aperiodic CQI report request.

Besides these purposes, the PDCCH can also carry power control information for several users. The options supported are both the 1 bit and 2 bit formats. The maximum number of users listening to a single PDCCH for power control purposes is N. Higher layers indicate to the device which bit or bits it needs to receive for the power control purposes.

As mentioned earlier, the PCFICH indicates the amount of resources reserved for the PDCCH with 1 ms basis. The UEs in the cell check the PCFICH and then blindly decode the PDCCH to see

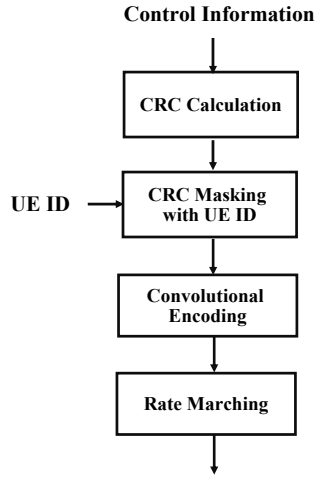


Figure 5. 35 PDCCH channel coding chain

which of the control channels (if any) is intended for them. The user identification is based on using UE specific CRC, which is generated after normal CRC generation by masking the CRC with the UE ID. The PDCCH channel coding is based on the same 1/3-rate convolutional coding as other convolutionally encoding channels. The PDCCH encoding chain is illustrated in Figure 5. 35.

5.8.3 Physical HARQ Indicator Channel (PHICH)

The task for the Physical HARQ Indicator Channel (PHICH) is simply to indicate in the downlink direction whether an uplink packet was correctly received or not. The device will decode the PHICH based on the uplink allocation information received on the PDCCH.

5.8.4 Downlink Transmission Modes

For robust and efficient system operation, it is important that the UE knows beforehand which type of transmission to expect. If the transmission mode could change dynamically from one subframe to another the UE would need to monitor all the possible DCI formats simultaneously, leading to a considerable increase in the number of blind decodings and receiver complexity (and possibly an increased number of signaling errors). Furthermore, the UE would not be able to provide meaningful channel feedback since, for example, the CQI value depends on the transmission mode assumed.

Therefore each UE is configured semi-statically via RRC signaling to one transmission mode. The transmission mode defines what kind of downlink transmissions the UE should expect, e.g. transmit diversity or closed loop spatial multiplexing, and restricts the channel feedback to modes corresponding to the desired operation. In LTE Release 8, seven transmission modes have been defined:

1 Single-antenna port; port 0
This is the simplest mode of operation with no pre-coding.

2 Transmit Diversity
Transmit diversity with two or four antenna ports using SFBC.
3 Open-loop spatial multiplexing
This is an open loop mode with the possibility to do rank adaptation based on the RI feedback. In the case of rank = 1 transmit diversity is applied similarly to transmission mode 2. With higher rank spatial multiplexing with up to four layers with large delay, CDD is used.
4 Closed-loop spatial multiplexing
This is a spatial multiplexing mode with pre-coding feedback supporting dynamic rank adaptation.
5 Multi-user MIMO
Transmission mode for downlink MU-MIMO operation.
6 Closed-loop Rank = 1 pre-coding
Closed loop pre-coding similar to transmission mode 5 without the possibility of spatial multiplexing, i.e. the rank is fixed to one.
7 Single-antenna port; port 5
This mode can be used in a beam forming operation when UE specific reference signals are in use.

5.8.5 Physical Broadcast Channel (PBCH)

The physical broadcast Channel (PBCH) carries the system information needed to access the system, such as RACH parameters, and as covered in more detail in Chapter 6. The channel is always provided with 1.08 MHz bandwidth, as shown in Figure 5.36, so the PBCH structure is independent of the actual system bandwidth being used, similar to other channels/signals needed for initial system access. The PBCH is convolutionally encoded as the data rate is not that high. As discussed in Chapter 6, the broadcast information is partly carried on the PBCH,

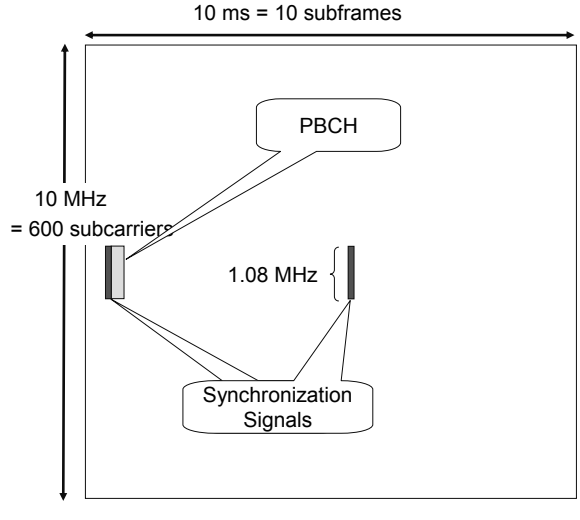


Figure 5.36 PBCH location at the center frequency

where the Master Information Block (MIB) is transmitted while the actual System Information Blocks (SIBs) are then on the PDSCH. The 600 sub-carriers in Figure 5.36 need only 9 MHz (50 resource blocks) in the resource domain but the system bandwidth needed for sufficient attenuation for the adjacent operator increases the total bandwidth needed to 10 MHz, as discussed in Chapter 11. With a 1.4 MHz system bandwidth there are no resource blocks on either side of the PBCH in the frequency domain in use, so effectively only 6 resource blocks may be in use to meet the spectrum mask requirements.

5.8.6 Synchronization Signal

There are 504 Physical Cell Identities (PCIs) values in the LTE system, compared with the 512 primary scrambling codes in WCDMA. The Primary Synchronization Signal (PSS) and the Secondary Synchronization Signals (SSS) are transmitted, similar to PBCH, always with the 1.08 MHz bandwidth, located in the end of 1st and 11th slots (slots 0 and 10) of the 10 ms frame, as shown in Figure 5.37.

The PSS and SSS jointly point the space of 504 unique Physical-layer Cell Identities (PCIs). The PCIs form 168 PCI groups, each of them having 3 PCIs (thus a total of 504 PCIs). The location and structure of the PCIs mean that taking a sample from the center frequency (with a bandwidth of 1.08 MHz) for a maximum of 5 ms contains the necessary information needed for cell identification.

5.9 Physical Layer Procedures

The key physical layer procedures in LTE are power control, HARQ, timing advance and random access. Timing advance is based on the signaling in the Medium Access Control (MAC) layer (as shown in the MAC section in Chapter 6), but as it is directly related to the physical layer, the timing advance details are covered in this chapter. The big contrast to WCDMA is that there are no physical layer issues related to macro-diversity since the UE is only connected to one base station at a time and hand handover is applied. Also, a specific means for dealing with inter-system and inter-frequency measurements such as compressed

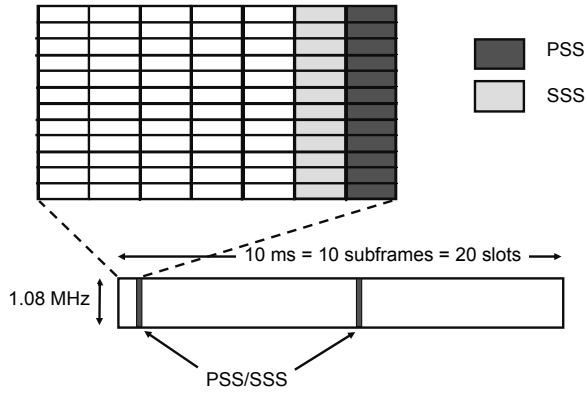


Figure 5.37 Synchronization signals in the frame

mode is not needed in LTE, as LTE by nature has a discontinuous operation that will facilitate the measurements by scheduling.

5.9.1 HARQ Procedure

The HARQ in LTE is based on the use of a stop-and-wait HARQ procedure. Once the packet is transmitted from the eNodeB, the UE will decode it and provide feedback in the PUCCH, as described in section 5.6. For negative acknowledgement (NACK) the eNodeB will send a retransmission. The UE will combine the retransmission with the original transmission and will run the turbo decoding again. Upon successful decoding (based on CRC check) the UE will send positive acknowledgement (ACK) for the eNode. After that eNodeB will send a new packet for that HARQ process. Due to the stop-and-wait way of operating, one needs to have multiple HARQ processes to enable a continuous data flow. In LTE the number of processes is fixed to 8 processes in both the uplink and downlink direction. An example of a single user continuous transmission is illustrated in Figure 5.38. For multiple users, it is dependent on the eNodeB scheduler when a retransmission is sent in the uplink or downlink direction, as a retransmission also requires that resources are allocated.

The HARQ operation in LTE supports both soft combining and the use of incremental redundancy. The use of soft combining means that retransmission has exactly the same rate matching parameters as the original transmission and thus exactly the same symbols are transmitted. For incremental redundancy, the retransmission may have different rate matching parameters like the original transmission. The minimum delay between the end of a packet and the start of a retransmission is 7 ms. The UE will send the ACK/NACK for a packet in frame *n*, in the uplink frame *n+4*. This leaves around 3 ms processing time for the UE, depending on the uplink/downlink timing offset controlled by the timing advance procedure. The downlink timing for a single transmitted downlink packet is shown in Figure 5.39. The retransmission instant in the downlink is subject to the scheduler in eNodeB and thus the timing shown in Figure 5.39 is the earliest moment for a retransmission to occur.

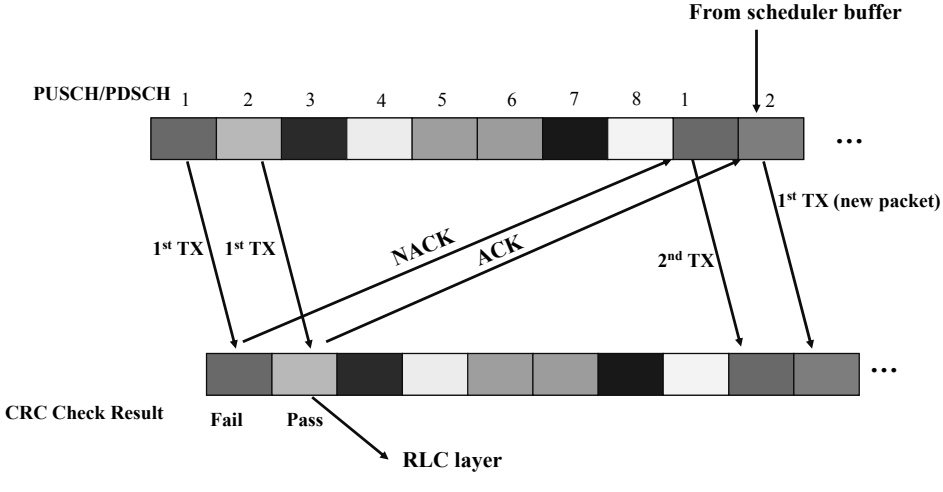


Figure 5.38 LTE HARQ operation with 8 processes

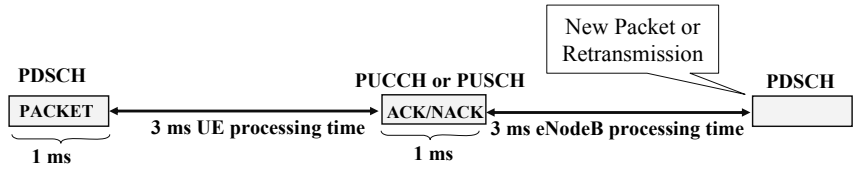


Figure 5.39 LTE HARQ timing for a single downlink packet

5.9.2 Timing Advance

The timing control procedure is needed so that the uplink transmissions from different users arrive at the eNodeB essentially within the cyclic prefix. Such uplink synchronization is needed to avoid interference between the users with uplink transmissions scheduled on the same subframe. The eNodeB continuously measures the timing of the UE uplink signal and adjusts the uplink transmission timing as shown in Figure 5.40. Timing advance commands are sent only when a timing adjustment is actually needed. The resolution of a timing advance command is 0.52 µs, and timing advance is defined relative to the timing of the downlink radio frame received on UE.

The timing advance value is measured from RACH transmission when the UE does not have a valid timing advance, i.e. the uplink for the UE is not synchronized. Such situations are system access, when the UE is in RRC_IDLE state or when the UE has had an inactivity period exceeding related timer, non-synchronized handover, and after radio link failure. Additionally, eNodeB can assign to UE a dedicated (contention-free) preamble on RACH for uplink timing measurement when eNodeB wants to establish uplink synchronization. Such situations are faced with handover or when downlink data arrive for a non-synchronized UE. From the range defined for timing advance, cell sizes up to 100 km would be facilitated, and even beyond by leaving some resources unused.

5.9.3 Power Control

For LTE, power control is slow for the uplink direction. In the downlink direction there is no power control. As the bandwidth varies due to data rate changes, the absolute transmission power of the UE will also change. The power control does not now actually control absolute power but rather the Power Spectral Density (PSD), power per Hz, for a particular device. What facilitates the use of a slower rate for power control is the use of orthogonal resources in the LTE uplink, which avoids the near–far problem that required fast power control in WCDMA. The key motivation for the power control is to reduce terminal power consumption and also to avoid an overly large dynamic range in the eNodeB receiver, rather than to mitigate interference. In the

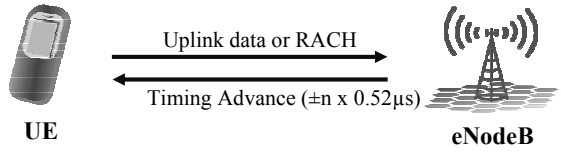


Figure 5.40 Uplink timing control

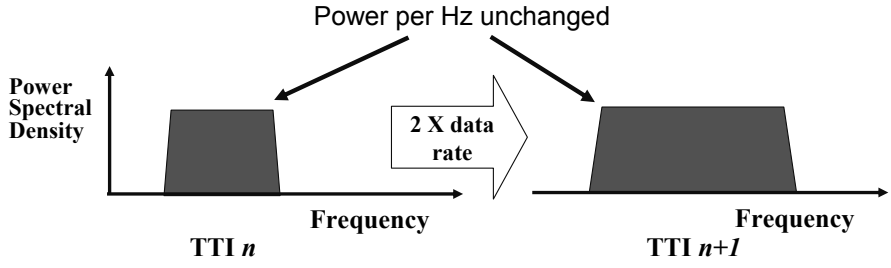


Figure 5.41 LTE uplink power with data rate change

receiver the PSDs of different users have to be reasonably close to each other so the receiver A/D converter has reasonable requirements and also the interference resulting from the non-ideal spectrum shape of the UE transmitter is kept under control. The LTE uplink power control principle is illustrated in Figure 5.41 where at the change of data rate the PSD stays constant but the resulting total transmission power is adjusted relative to the data rate change.

The actual power control is based on estimating the path loss, taking into account cell specific parameters and then applying the (accumulated) value of the correction factor received from the eNodeB. Depending on the higher layer parameter settings, the power control command is either 1 dB up or down or then the set of [−1dB, 0, +1dB, +3 dB] is used. The specifications also include power control that is absolute value based but, based on the text case prioritization, it is not foreseen that this will be used in the first phase networks. The total dynamic range of power control is slightly smaller than in WCDMA and the devices now have a minimum power level of −41 dBm compared to −50 dBm with WCDMA.

5.9.4 Paging

To enable paging, the UE is allocated a paging interval and a specific subframe within that interval where the paging message could be sent. The paging is provided in the PDSCH (with allocation information on the PDCCH). The key design criterion in paging is to ensure a sufficient DRX cycle for devices to save power and also to ensure a fast enough response time for the incoming call. The E-UTRAN may parameterize the duration of the paging cycle to ensure sufficient paging capacity (covered in more detail in Chapter 6).

5.9.5 Random Access Procedure

The LTE Random Access (RACH) operation resembles that of WCDMA because both use preambles and similar ramping of preamble power. The initial power is based on the measured path loss in DL, and power ramping is necessary because of the relatively coarse accuracy of the UE in path loss measurement and absolute power setting, and to compensate for uplink fading. Although LTE PRACH resources are separate from PUSCH and PUCCH, power ramping is useful for simultaneous detection of different preamble sequences and for minimizing the interference due to asynchronous PRACH transmission at the adjacent PUCCH and PUSCH resources. The steps of the physical layer procedure are as follows:

- Transmit a preamble using the PRACH resource, preamble sequence and power selected by MAC.
- Wait for the RACH response with matching preamble information (PRACH resource and preamble sequence). In addition to the preamble information, the response also contains the information on the uplink resource to use for further information exchange as well as the timing advance to be used. In the WCDMA RACH procedure, after acknowledging a preamble, the UE continues with a 10 or 20 ms frame duration of data, or even longer, as described for Release 8 HSPA operation in Chapter 13. The fundamental difference in LTE is that the device will move instead directly to the use of UL-SCH on reception of the random access response, which has the necessary information.
- If no matching random access response is received, transmit the preamble in the next available PRACH resource according to instructions of MAC, as shown in Figure 5.42.

Although the LTE specification models that the physical layer just transmits preambles and detects responses under the control of MAC, we describe below the complete procedure without focusing on the modeling of the specification.

Two fundamentally different random access procedures have been specified for LTE. The contention based procedure is what we normally understand with random access: UEs transmit randomly selected preambles on a common resource to establish a network connection or request resources for uplink transmission. The non-contention based random access is initiated by the network for synchronizing UE's uplink transmission, and the network can identify the UE from the very first uplink transmission. This procedure is nevertheless included under the LTE random access because it uses PRACH resources. Both procedures are common for TDD and FDD systems.

5.9.5.1 Contention and Non-contention Based Random Access

The contention based procedure follows the signaling diagram in Figure 5.43 (left half).

In the first step, the UE transmits a preamble sequence on PRACH. The details of PRACH and preamble sequences are explained in section 5.7. For each cell a total of 64 sequences are reserved, and these are grouped for the non-contention based and contention based procedures. The group reserved for the contention based procedures is divided further into two: by selecting the proper group, the UE sends one bit of information about the transport block size that the UE desires to send on PUSCH in Step 3.

In the second step, the UE receives a preamble response on DL-SCH resource that is assigned on PDCCH. The identity RA-RNTI that is used for this assignment is associated with the

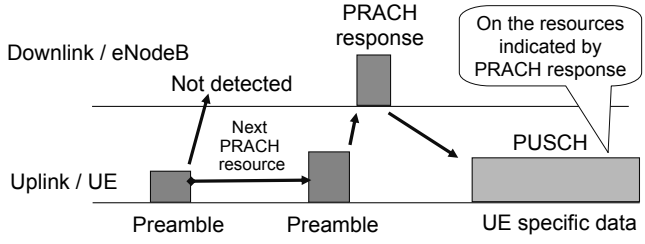


Figure 5.42 Power ramping in random access procedure

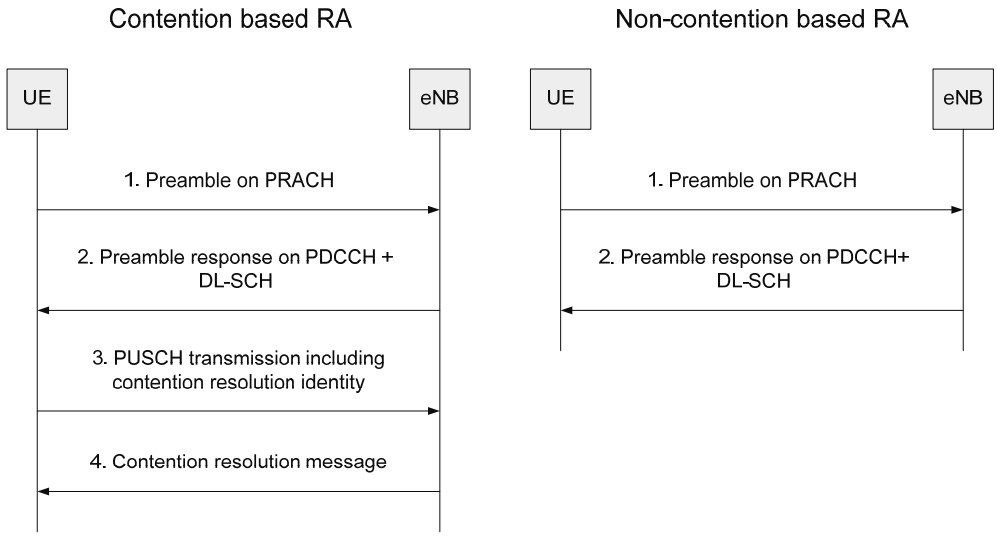


Figure 5.43 The contention and non-contention based random access procedures

frequency and time resource of the preamble. This permits bundling of the responses that are meant for preambles transmitted in the same PRACH frequency and time resource, which is important for saving PDCCH resources. eNodeB transmits the response in a time window that can be configured up to 10 ms in duration. The flexible window allows freedom for dimensioning of the RACH receiver and for scheduling of the responses.

The response (signaling part of the MAC layer as covered in Chapter 6) lists the sequence numbers of the observed preambles, and in addition the following information is given for each acknowledged preamble:

- A grant for the first transmission on PUSCH, including also information on the need for frequency hopping, power control command for uplink transmission and information on the need for CQI transmission and whether the PUSCH transmission needs to be delayed by one subframe from the nominal value.
- A timing alignment command.
- A temporary allocation for identity called temporary CRNTI, which is used for addressing PUSCH grants and DL-SCH assignments in Steps 3 and 4 of the procedure.

The typical probability of preamble collisions, meaning that two or more UEs are transmitting the same preamble sequence in the same frequency and time resource, is expected to be around 1%. These collisions are resolved in Steps 3 and 4: the UE includes its identity in the first message that it transmits on PUSCH in Step 3 and expects in Step 4 an acknowledgement that eNodeB has received the identity. There are two forms of acknowledgement: it can be either (1) a PUSCH grant or DL-SCH assignment addressed with CRNTI if the UE had included CRNTI to the message of Step 3, or (2) the UE's identity can be acknowledged with a message that is sent on a DL-SCH resource assigned with the temporary CRNTI. The first form of acknowledgement is for RRC connected UEs while the second form is used when a UE tries to establish or re-establish RRC connection. HARQ is used both in Step 3 and 4. In

Step 3 there is no difference compared with normal HARQ, but in Step 4 the UE never sends NAK, and ACK is sent only by the UE that wins the contention resolution. No special actions are taken after a lost contention resolution but the UE simply retransmits a preamble just like after failing to receive the preamble response.

The LTE network can control the RACH load rapidly. If the UE does not receive acknowledgement for its preamble in Step 2 or for its identity in Step 4, the UE retransmits the preamble with increased power if power ramp-up has been configured. Normally the retransmission can be done as soon as the UE is ready but the network can also configure a back-off parameter that forces the UE to add a random delay before the retransmission. When needed, the back-off parameter is included in the preamble response message, and the setting is obeyed by all the UE decoding the message. This allows much faster load control than in WCDMA where a similar load control parameter is in the broadcasted System Information.

The non-contention based procedure, shown in Figure 5.43 (right half), is used for time alignment during handover and when an RRC connected UE needs to be synchronized for downlink data arrival. The UE receives in the handover command or through PDCCH signaling an index of its dedicated preamble sequence, which it is allowed to transmit on PRACH. Besides the sequence index, some restrictions for the frequency and time resource can be signaled so that the same sequence can be simultaneously allocated for UEs that transmit on different PRACH subframes or, for TDD, at different PRACH frequencies. The preamble responses in the contention and non-contention based procedures are identical and they can thus be bundled to the same response message. As eNodeB knows the identity of the UE that has sent the dedicated preamble, the contention resolution with Steps 3 and 4 is not needed.

The non-contention based procedure provides delay and capacity enhancements compared with the contention based procedure. As the preamble collisions are absent and the contention resolution is not needed, a shorter delay can be guaranteed, which is especially important for handover. The sequence resource is in effective use because it is assigned to the UE only when needed and can be released as soon as eNodeB detects that the UE has received the preamble response.

An unsuccessful random access procedure ends based on preamble count or RRC timers. The preamble count is decisive only with two causes of random access: (a) an RRC connected UE, lacking scheduling request resources, asks resources because of uplink data arrival or (b) an RRC connected UE needs to be synchronized because of DL data arrival. If random access has been started because of RRC connection establishment or re-establishment or because of handover, the procedure continues until success or MAC reset in the expiry of the RRC timer corresponding to the cause of random access.

5.9.6 Channel Feedback Reporting Procedure

The purpose of the channel state feedback reporting is to provide the eNodeB with information about the downlink channel state in order to help optimize the packet scheduling decision. The principle of the channel state feedback reporting procedure is presented in Figure 5.44. The channel state is estimated by the UE based on the downlink transmissions (reference symbols, etc.) and reported to the eNodeB by using PUCCH or PUSCH. The channel state feedback reports contain information about the scheduling and link adaptation (MCS/TBS and MIMO) related parameters the UE can support in the data reception. The eNodeB can then take advantage of the feedback information in the scheduling decision in order to optimize the usage of the frequency resources.

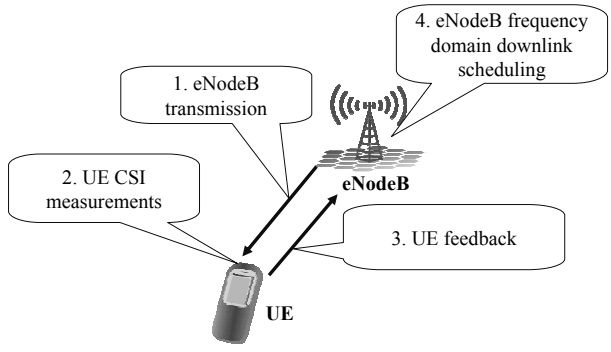


Figure 5.44 Channel State Information (CSI) reporting procedure

In general the channel feedback reported by the UE is just a recommendation and the eNodeB does not need to follow it in the downlink scheduling. In LTE the channel feedback reporting is always fully controlled by the eNodeB and the UE cannot send any channel state feedback reports without eNodeB knowing it beforehand. The corresponding procedure for providing information about the uplink channel state is called channel sounding and it is done using the Sounding Reference Signals (SRS) as presented in Section 5.6.4.3.

The main difference of the LTE channel state information feedback compared to WCDMA/HSDPA is the frequency selectivity of the reports, i.e. the information regarding the distribution of channel state over the frequency domain can also be provided. This is an enabler for Frequency Domain Packet Scheduling (FDPS), a method that aims to divide the radio resources in the frequency domain for different users so that system performance is optimized. In Figure 5.45 the gain from the FDPS is illustrated. As the UE speed increases, the CSI reports become more inaccurate and get outdated faster leading to reduced gains in high mobility.

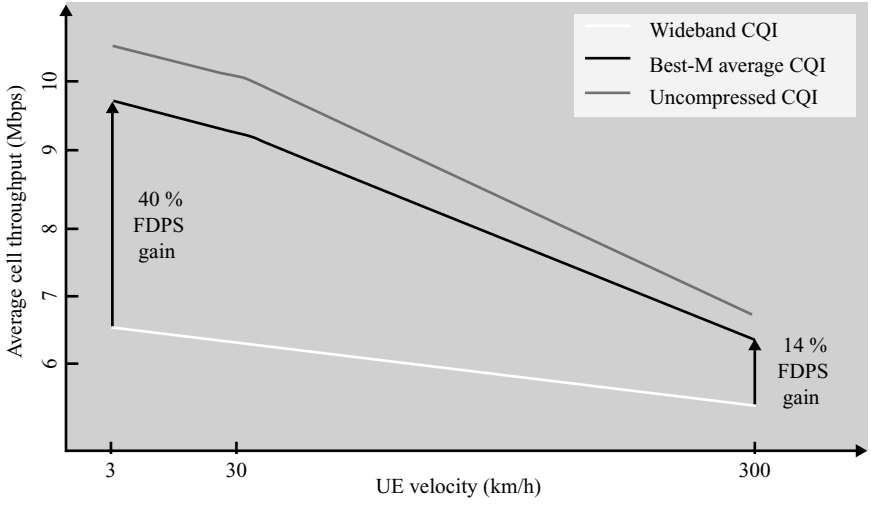


Figure 5.45 Comparison of the average cell throughputs for different CQI schemes and velocities

5.9.6.1 Channel Feedback Report Types in LTE

In LTE the UE can send three types of channel feedback information:

- CQI – Channel Quality Indicator
- RI – Rank Indicator
- PMI – Pre-coding Matrix Indicator

The most important part of channel information feedback is the Channel Quality Indicator (CQI). The CQI provides the eNodeB information about the link adaptation parameters the UE can support at the time (taking into account the transmission mode, the UE receiver type, number of antennas and interference situation experienced at the given time). The CQI is defined as a table containing 16 entries (Table 5.3) with Modulation and Coding Schemes (MCSs). The UE reports back to the eNodeB the highest CQI index corresponding to the MCS and TBS for which the estimated received DL transport block BLER shall not exceed 10%. The CQI operation has a high degree of similarity with HSDPA CQI use, as covered in [5]. Note that there are many more possibilities for MCS and TBS size values than only those 15 indicated by the CQI feedback.

Rank Indicator (RI) is the UE's recommendation for the number of layers, i.e. streams to be used in spatial multiplexing. RI is only reported when the UE is operating in MIMO modes with spatial multiplexing (transmission modes 3 and 4). In single antenna operation or TX diversity it is not reported. The RI can have values 1 or 2 with 2-by-2 antenna configuration and from 1 up to 4 with 4-by-4 antenna configuration. The RI is always associated with one or more CQI reports, meaning that the reported CQI is calculated assuming that particular RI value. Since the rank varies typically more slowly than the CQI it is normally reported less

Table 5.3 CQI table

CQI index	Modulation	Coding rate × 1024	Bits per resource element
0	out of range		
1	QPSK	78	0.1523
2	QPSK	120	0.2344
3	QPSK	193	0.3770
4	QPSK	308	0.6016
5	QPSK	449	0.8770
6	QPSK	602	1.1758
7	16QAM	378	1.4766
8	16QAM	490	1.9141
9	16QAM	616	2.4063
10	64QAM	466	2.7305
11	64QAM	567	3.3223
12	64QAM	666	3.9023
13	64QAM	772	4.5234
14	64QAM	873	5.1152
15	64QAM	948	5.5547

often. RI always describes the rank on the whole system band, i.e. frequency selective RI reports are not possible.

The PMI provides information about the preferred pre-coding matrix in codebook based pre-coding. Like RI, PMI is also relevant to MIMO operation only. MIMO operation with PMI feedback is called Closed Loop MIMO. The PMI feedback is limited to transmission modes 4, 5, and 6. The number of pre-coding matrices in the codebook depends on the number of eNodeB antenna ports: in the case of two antenna ports there are altogether six matrices to choose from, while with four antenna ports the total number is up to 64 depending on the RI and the UE capability. PMI reporting can be either wideband or frequency selective depending on the CSI feedback mode.

5.9.6.2 Periodic and Aperiodic Channel State Feedback Reporting

Although in principle the UE has up-to-date information about the changes in channel state, a channel state feedback report initiated by the UE would raise several issues. First, to detect the reports blind decoding would need to be performed at the eNodeB, which is not desirable from the receiver implementation point of view. Secondly, as the eNodeB is anyway fully in charge of the scheduling decisions, UE initiated reports would often be unnecessary. Furthermore, the reports initiated by UE would complicate the uplink resource allocation considerably, leading to increased signaling overhead. Hence it was agreed that in the LTE standardization channel state feedback reporting is always fully controlled by the eNodeB, i.e. the UE cannot send any channel state feedback reports without eNodeB knowing beforehand.

To fully exploit the gains from frequency selective packet scheduling, detailed CSI reporting is required. As the number of UEs reporting channel state feedback increases, however, the uplink signaling overhead becomes significant. Furthermore the PUCCH, which is supposed to carry primarily the control information, is rather limited in capacity: payload sizes of only up to 11 bits/subframe can be supported. On the PUSCH there are no similar restrictions on the payload size, but since PUSCH is a dedicated resource only one user can be scheduled on a single part of the spectrum.

To optimize the usage of the uplink resources while also allowing for detailed frequency selective CSI reports, a two-way channel state feedback reporting scheme has been adopted in LTE. Two main types of reports are supported: Periodic and Aperiodic. A comparison of the main features of the two reporting options is presented in Table 5.4.

Periodic reporting using PUCCH is the baseline mode for channel information feedback reporting. The eNodeB configures the periodicity parameters and the PUCCH resources via higher layer signaling. The size of a single report is limited up to about 11 bits depending on the reporting mode, and the reports contain little or no information about the frequency domain behavior of the propagation channel. Periodic reports are normally transmitted on PUCCH. If the UE is scheduled in the uplink, however, the Periodic report moves to PUSCH. The reporting period of RI is a multiple of CQI/PMI reporting periodicity. RI reports use the same PUCCH resource (PRB, Cyclic shift) as the CQI/PMI reports – PUCCH format 2/2a/2b.

When the eNodeB needs more precise channel state feedback information it can at any time request the UE to send an Aperiodic channel state feedback report on PUSCH. Aperiodic reports can be either piggybacked with data or sent alone on PUSCH. Using the PUSCH makes it possible to transmit large and detailed reports. When the transmission of Periodic and Aperiodic reports from the same UE might collide, only the Aperiodic report is sent.

Table 5.4 Comparison of Periodic and Aperiodic channel information feedback reporting

	Periodic reporting	Aperiodic reporting
When to send	Periodically every 2–160 ms	When requested by the eNodeB
Where to send	Normally on PUCCH, PUSCH used when multiplexed with data	Always on PUSCH
Payload size of the reports	4–11 bits	Up to 64 bits
Channel Coding	Linear block codes	Tail biting convolutional codes
CRC protection	No	Yes, 8 bit CRC
Rank Indicator	Sent in separate subframes at lower periodicity	Sent separately encoded in the same subframe
Frequency selectivity of the CQI	Only very limited amount of frequency information	Detailed frequency selective reports are possible
Frequency selectivity of the PMI	Only wideband PMI	Frequency selective PMI reports are possible

The two modes can also be used to complement each other. The UE can be, for example, configured to send Aperiodic reports only when it is scheduled, while Periodic reports can provide coarse channel information on a regular basis.

5.9.6.3 CQI Compression Schemes

Compared to the WCDMA/HSPA, the main new feature in the channel feedback is the frequency selectivity of the report. This is an enabler for the Frequency Domain Packet Scheduling (FDPS). Since providing a full 4-bit CQI for all the PRBs would mean excessive uplink signaling overhead of hundreds of bits per subframe, some feedback compression schemes are used.

To reduce feedback, the CQI is reporter per *sub-band* basis. The size of the sub-bands varies depending on the reporting mode and system bandwidth from two consecutive PRBs up to whole system bandwidth.

The main CQI compression methods are:

- wideband feedback
- Best-M average (UE-selected sub-band feedback)
- higher layer-configured sub-band feedback.

Additionally, delta compression can be used in combination with the above options, e.g. when a closed loop MIMO CQI for the 2nd codeword can be signaled as a 3-bit delta relative to the CQI of the 1st codeword. When the number of sub-bands is large this leads to considerable savings in signaling overhead.

5.9.6.4 Wideband Feedback

The simplest way to reduce the number of CQI bits is to use only wideband feedback. In wideband feedback only a single CQI value is fed back for the whole system band. Since

no information about the frequency domain behavior of the channel is included, wideband feedback cannot be used in FDPS. Often, however, this is still sufficient, e.g. PDCCH link adaptation or TDM-like scheduling in low loaded cells do not benefit from frequency selective CQI reporting. Also, when the number of scheduled UEs gets high, the total uplink signaling overhead due to detailed CQI reports may become excessive and the wideband CQI reports are the only alternative.

5.9.6.5 Best-M Average

Best-M average is an effective compromise between the system performance and the uplink feedback signaling overhead. The principle of the Best-M average compression is shown in Figure 5.46. In Best-M average reporting the UE first estimates the channel quality for each sub-band. Then it selects the M best ones and reports back to the eNodeB a single average CQI corresponding to the MCS/TBS the UE could receive correctly assuming that the eNodeB schedules the UE on those M sub-bands. The parameter M depends on the system bandwidth and corresponds to roughly 20% of the whole system bandwidth.

5.9.6.6 Higher Layer-configured Sub-band Feedback

In higher layer-configured sub-band feedback a separate CQI is reported for each sub-band using delta compression. This will result in the best performance at the cost of feedback overhead: the payload size of the reports can be as large 64 bits. To keep the signaling on a manageable level, the sub-band sizes with Full Feedback reporting are twice as large as with Best-M average. This will limit the accuracy and performance in very frequency selective channels, where Best-M average may be a better alternative.

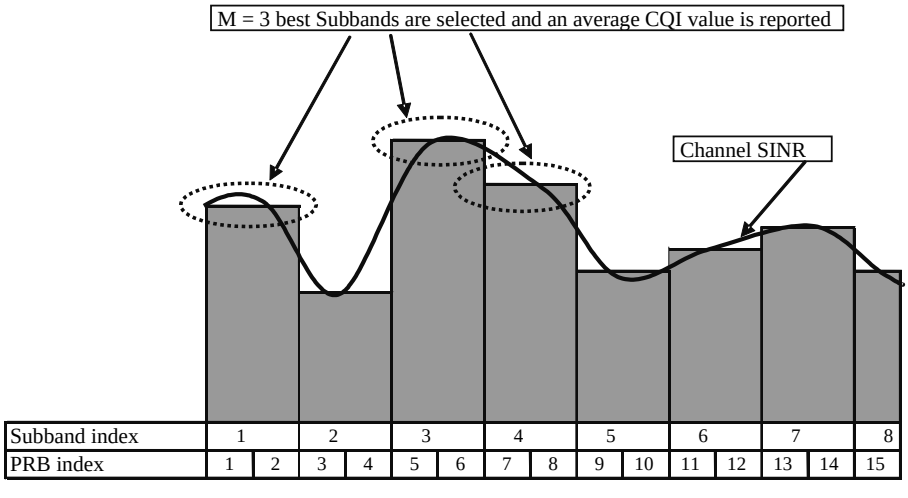


Figure 5.46 The principle of the Best-M average compression. An average CQI value is reported for the M best sub-bands

5.9.7 Multiple Input Multiple Output (MIMO) Antenna Technology

In the Release 8 LTE specifications there is support for multi-antenna operation both in terms of transmit diversity as well as spatial multiplexing with up to four layers. The use of MIMO with OFDMA has some favorable properties compared to WCDMA because of its ability to cope effectively with multi-path interference. In WCDMA, MIMO was introduced in Release 7 but has not yet been deployed. In LTE, MIMO is part of the first release and also part of the device categories with the exception of the simplest device type, as shown in section 5.9. The use of more than a single antenna in an eNodeB will not result in a problem with non-MIMO LTE devices because all devices can cope with the transmit diversity up to four antennas. Thus full transmit power can always be used regardless of the number of transmit antennas included. As shown in Chapter 4, the reference symbols are transmitted so that they have separate time and frequency domain resources for each antenna port to allow for good separation of different antennas and robust channel estimation, while the actual user data then overlap in the time and frequency domains.

Perhaps the most attractive mode of MIMO operation is spatial multiplexing, which allows an increase in the peak rates compared to the non-MIMO case by a factor of 2 or 4, depending on the eNodeB and the UE antenna configuration. In LTE the MIMO modes supporting spatial multiplexing are:

- Transmission mode 3 – Open-loop spatial multiplexing. The UE reports the RI but no pre-coding feedback. Based on the RI the eNodeB scheduler can select the number of layers used in spatial multiplexing. In the case of rank = 1, TX diversity is used. With a higher rank, large delay CDD is applied with deterministic pre-coding.
- Transmission mode 4 – Closed-loop spatial multiplexing. Here the UE reports both the RI and index of the preferred pre-coding matrix. Dynamic rank adaptation is supported on the eNodeB, which signals the applied pre-coding vector to the UE in the Downlink grant.

The UE provides feedback for the MIMO operation (as discussed in section 5.6.9.1) and the eNodeB selects the pre-coding vectors and the number of layers to use accordingly, as illustrated in Figure 5.47. This has some similarity with the MIMO solution in WCDMA, which is based on the closed loop feedback from the UE. In LTE some of the parameters in the downlink assignments are intended for MIMO control, either for informing the device of

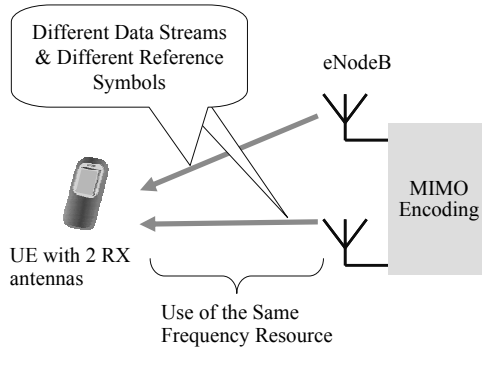


Figure 5.47 Single user MIMO principle

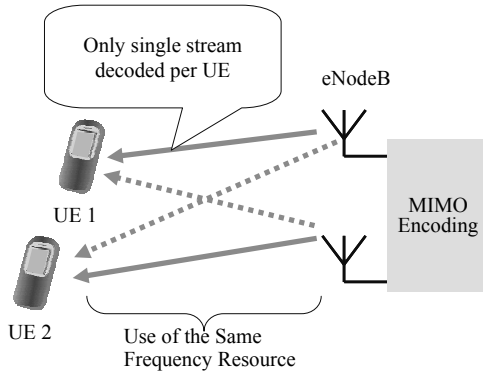


Figure 5.48 Multi-user MIMO transmission principle

the pre-coding applied or whether or not to swap the transport blocks between the code words in the retransmissions.

One specific mode of MIMO operation is the Multi-User MIMO (MU-MIMO) where the eNodeB is sending from two (also four possible from the specifications) antenna ports different data streams intended for different devices. The principle of MU-MIMO is illustrated in Figure 5.48.

5.9.8 Cell Search Procedure

The first action for a LTE device to be performed upon power on is the cell search. In LTE the cell search procedure is based on the use of synchronization signals. Similar to WCDMA, there are primary and secondary synchronization signals (called synchronization channels in the WCDMA specification). The cell search steps are as follows:

- The device will look for the primary synchronization signal at the center frequencies possible at the frequency band in question. There exist three different possibilities for the primary synchronization signal as described in section 5.8.6.
- Once the primary synchronization signal has been detected, the UE will look for the secondary synchronization signal.
- Once one alternative of the 168 possible secondary synchronization signals is detected, the UE has figured out the Physical Cell ID (PCI) value from the address space of 504 IDs.

From the PCI the UE has information about the parameters used for downlink reference signals and thus the UE can decode the PBCH. All this is independent of the actual system bandwidth in question. The LTE network specifications also support automated PCI planning as part of the Self Organization Networks (SON) functionality, as described in section 5.12.

5.9.9 Half Duplex Operation

The LTE specifications also enable half duplex operation, which is basically FDD mode operation (i.e. separate transmission and reception frequency) but transmission and reception do not occur simultaneously, as happens in TDD mode. The intention in 3GPP has been to have an

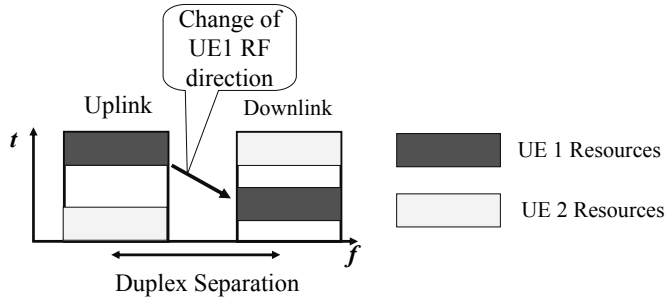


Figure 5.49 LTE FDD half duplex operation

option for the cases where due to the frequency arrangements the resulting requirements for the duplex filter would be unreasonable, resulting in high cost and high power consumption. The duplex operation principle is illustrated in Figure 5.49. It remains to be seen whether the performance requirements for some bands will be specified with the half duplex operation in mind. The eNodeB would obviously need to be aware if some devices are based on half duplex operation. The impact for the data rates would be that uplink and downlink would no longer be independent but the available data rate in one transmission direction would depend on the resources allocated for the other direction. Similar to TDD operation, one would need to schedule the data on a device basis (not on a system basis like in TDD) so that there is no conflict between uplink and downlink allocations for one UE. Also time would be needed for the UE to change between transmit and reception.

5.10 UE Capability Classes and Supported Features

In LTE there are five device capability classes defined. The supported data ranges run from 5 to 75 Mbps in the uplink direction and from 10 to 300 Mbps in the downlink direction. All devices support the 20 MHz bandwidth for transmission and reception, assuming that for the given frequency band this has been specified. It is foreseen that for most cases with frequency bands below 1 GHz the interest is with the smallest bandwidths and support for up to 20 MHz will not be specified. For bands above 1 GHz, bandwidths below 5 MHz will often not be needed. Only a category 5 device will do 64QAM in the uplink, others use QPSK and 16QAM. The receiver diversity and MIMO are in all categories, except in category 1, which does not support MIMO. The UE categories are shown in Table 5.5. The step in the data rates up to 300 Mbps

Table 5.5 LTE device categories

	Category 1	Category 2	Category 3	Category 4	Category 5
Peak rate DL/UL	10/5 Mbps	50/25 Mbps	100/50 Mbps	150/50 Mbps	300/75 Mbps
Modulation DL	QPSK/16QAM/64QAM				
Modulation UL	QPSK/16QAM	QPSK/16QAM	QPSK/16QAM	QPSK/16QAM	QPSK/16QAM + 64QAM
MIMO DL	Optional	2 × 2	2 × 2	2 × 2	4 × 4

with category 5 is achieved with the four antenna MIMO transmission, which is not supported by the other categories.

The actual device capabilities are behind other signaling not just these categories. While these categories mainly define the boundaries for the data rates, the test case prioritization in 3GPP also defines what the first phase devices will and will not support. 3GPP uses three different levels for prioritization of the test cases based on the input created by the operators planning to take the system into use. The latest version during writing of this can be found from [19]. The levels are:

- High priority, which refers to the features that are planned to be in the first phase deployments. This is the highest test case category for number of features (as some of the undesired issues were removed from the specifications in the early stages).
- Medium priority, which refers to features planned to be used a few years after initial deployments. Originally this priority level contained features such as large delay CDD, but then it proved difficult (impossible) to deploy later if there were devices not able to reach, for example, PBCH based on this. Similarly the extended cyclic prefix was originally in this class but deploying it later was found to be problematic if there is an existing population of devices unable to support it (due to lack of test cases).
- Low priority, which refers to features where no immediate plans exist for deployments from the 3GPP operator perspective. An example is the UE specific reference symbols for FDD mode operation, which reflects the lack of interest for adaptive antenna arrays in the industry for Release 8 FDD. This is similar to WCDMA where support for dedicated pilots was removed from a later version of the specifications due to lack of practical interest for introducing adaptive antennas. (The WCDMA system contains other means for dealing with adaptive antenna arrays as covered in [5].) In the TDD mode of operation, this feature is a high priority test case to show that many sites using the 1.28 Mcps TDD (TD-SCDMA) in China are based on adaptive antenna arrays. To enable smooth evolution toward LTE, then also LTE devices operating in TDD mode should allow the use of the same antenna structures along with other aspects of LTE TDD mode and TD-SCDMA co-existence as covered in Chapter 12.

5.11 Physical Layer Measurements

5.11.1 eNodeB Measurements

As all the radio functionalities are located in eNodeB, there are few eNodeB measurements that would need to be reported over any interface as there is no separate centralized RRM functionality like the radio network controller in WCDMA. The eNodeB measurements specified in the physical layer specifications in Release 8 in the downlink are as follows [4]:

- the power used (power contribution) for the resource elements that are used to transmit cell-specific reference signals from the eNodeB (in the system bandwidth);
- received interference power per physical resource block;
- thermal noise power over the system bandwidth.

The motivation for these measurements is to enable their consideration in the handover decisions of the relative base station strengths as well as to facilitate inter-cell interference

coordination, as illustrated by part of the X2-interface signaling in Chapter 5. eNodeB may use these internally for different purposes in addition to the other information available for eNodeB.

In 3GPP there are additional indicators as part of the Operation and Maintenance (O&M) specifications to support monitoring the system performance.

5.11.2 UE Measurements and Measurement Procedure

For the UE the following measurements are to be performed inside the LTE system:

- Reference Signal Received Power (RSRP), which for a particular cell is the average of the power measured (and the average between receiver branches) of the resource elements that contain cell-specific reference signals.
- Reference Signal Received Quality (RSRQ) is the ratio of the RSRP and the E-UTRA Carrier Received Signal Strength Indicator (RSSI), for the reference signals.
- E-UTRA RSSI, which is the total received wideband power on a given frequency. Thus it includes the noise 'from the whole universe' on the particular frequency, whether that is from interfering cells or any other noise source. E-UTRA RSSI is not reported by the UE as an individual measurement (as indicated in the early versions of [4] until June 2008), but it is only used in calculating the RSRQ value inside the UE.

The scheduling will create sufficient DTX/DRX gaps for the device to perform the measurement. In WCDMA there were specific frames in use with compressed mode when one needed to turn off the transmitter and/or receiver to measure other frequencies, as covered in [5], and most of the measurements were related to the Common Pilot Channel (CPICH).

The following measurements exist for the inter-system:

- UTRA FDD CPICH (CPICH) Received Signal Code Power (RSCP), which represents the power measured on the code channel used to spread the primary pilot channel on WCDMA.
- UTRA FDD (and TDD) carrier RSSI is the corresponding wideband power measurement as also defined for LTE.
- UTRA FDD CPICH Ec/No is the quality measurement, like RSRQ in LTE, and provides the received energy per chip energy over the noise.
- GSM carrier RSSI represents the wideband power level measured on a particular GSM carrier.
- UTRA TDD Primary Common Control Physical Channel (P-CCPCH) RSCP is the code power of the UTRA TDD broadcast channel.
- CDMA2000 1×RTT and HRPD Pilot Strengths are the power on the respective pilot channel (code) when considering the handover to cdma2000®.

5.12 Physical Layer Parameter Configuration

The physical layer parameters to configure for connection in a particular cell are the responsibility of the particular eNodeB. Several issues will come from the O&M settings, such as the cyclic prefix lengths to be used. For some parameters, 3GPP has been developing Self Organizing

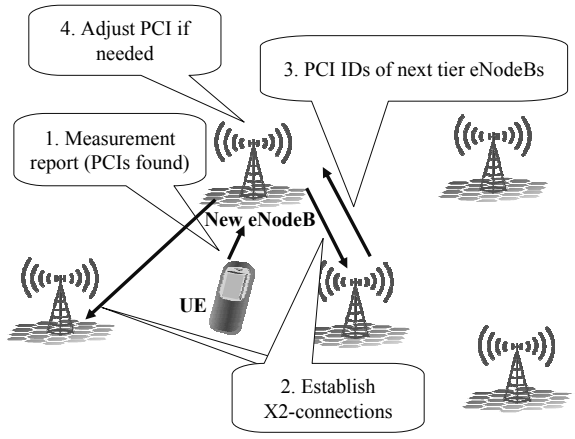


Figure 5.50 Self configuration for PCI

Network (SON) solutions. In the physical layer this covers the Physical Cell ID (PCI), as shown in Figure 5.50. When installing a new cell, the principle is that the cell could select the PCI randomly and once the first measurement report has been obtained from any UE, it learns the PCIs that are in use near by. After that eNodeB knows the neighbors and it can establish the X2 connections (UE then needs to be instructed also to decode the BCH to get the global cell ID after which the O&M system can provide the connection information for X2 creation). Once the X2 connections provide information about the PCI values used in nearby cells, the cell can confirm whether the selected PCI needs to be adjusted or not. Alternatively, the PCI could be obtained directly from O&M, thus avoiding initial conflicts for PCIs between nearby cells.

5.13 Summary

The LTE physical layer has been built on top of OFDMA (for downlink) and SC-FDMA (for uplink) technologies. The resource allocation is dynamic with 1 ms resolution in the time domain and 180 kHz in the frequency domain, making the radio resource utilization well matched to the packet nature of the communication. The physical layer facilitates advanced features known from HSDPA/HSUPA, such as physical layer retransmission, link adaptation, multi-antenna transmission and physical layer eNodeB based scheduling. Additionally there are new properties which take into account the possibilities of new radio access technology, such as frequency domain scheduling, that is facilitated in terms of both physical feedback methods and signaling to make the uplink and downlink resource allocations to the advantage of the frequency domain element. The maximum data rates within the Release 8 UE categories run up to 300 Mbps in downlink and 75 Mbps in the uplink direction. The physical layer design adapts well to the support of different system bandwidths as common channels are not dependent on the actual system bandwidth operated but are always based on the use of resources within the 1.08 MHz block at the center frequency.

References

[1] 3GPP Technical Specification, TS 36.211, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation', 3GPP, v 8.4.0 September 2008.

[2] 3GPP Technical Specification, TS 36.212, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding', 3GPP, v 8.4.0, September 2008.

[3] 3GPP Technical Specification, TS 36.213, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures', 3GPP, v 8.4.0, September 2008.

[4] 3GPP Technical Specification, TS 36.214, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer measurements', 3GPP, v 8.4.0, September 2008.

[5] H. Holma, A. Toskala 'WCDMA for UMTS', 4th edition, Wiley, 2008.

[6] 3GPP Technical Report, TR 25.814, 'Physical layer aspect for evolved Universal Terrestrial Radio Access (UTRA)', 3GPP.

[7] 3GPP Tdoc R1–081126, 'Way forward on RV and NDI signaling', Panasonic et al.

[8] 3GPP Tdoc R1–070041, 'Dynamic Contention Free Scheduling Request', Motorola.

[9] 3GPP Tdoc R1–070378, 'Uplink Resource Request for LTE System', Nokia.

[10] 3GPP Tdoc R1–080931, 'ACK/NACK channelization for PRBs containing both ACK/NACK and CQI', Nokia Siemens Networks, Nokia, Texas Instruments.

[11] 3GPP Tdoc R1–080035, 'Joint proposal on uplink ACK/NACK channelization', Samsung, Nokia, Nokia Siemens Networks, Panasonic, TI.

[12] 3GPP Tdoc R1–062840, 'TDM based Multiplexing Schemes between L1/L2 Control and uplink Data', Nokia.

[13] 3GPP Tdoc R1–072224, 'Uplink control signal transmission in presence of data', Samsung.

[14] 3GPP Tdoc R1–081852, 'Linkage between PUSCH MCS and amount of control resources on PUSCH', Nokia Siemens Networks, Nokia.

[15] Chu, D.C., 'Polyphase codes with good periodic correlation properties,' IEEE Transactions Information Theory, vol. 18, pp. 531–532, July 1972.

[16] Popovic, B.M., 'Generalized chirp-like polyphase sequences with optimum correlation properties,' IEEE Transactions Information Theory, vol. 38, pp. 1406–1490, July 1992.

[17] 3GPP Tdoc R1–082570, 'Signalling for SRS Hopping Bandwidth', NTT DoCoMo, Panasonic.

[18] 3GPP Tdoc R1–074977, 'Specification of Formula for Restricted Cyclic Shift Set', LG Electronics, Nokia, Nokia Siemens Networks, Panasonic, Huawei, Texas Instruments.

[19] 3GPP Tdoc R5–083674, 'LTE Features and Test Prioritisation' NTT DoCoMo.

6

LTE Radio Protocols

Antti Toskala and Woonhee Hwang

6.1 Introduction

The role of the LTE radio interface protocols is to set up, reconfigure and release the Radio Bearer that provides the means for transferring the EPS bearer (as presented in Chapter 3). The LTE radio interface protocol layers above the physical layer include Layer 2 protocols; Medium Access Control (MAC), Radio Link Control (RLC) and Packet Data Convergence Protocol (PDCP). Layer 3 consists of the Radio Resource Control (RRC) protocol, which is part of the control plane. The protocol layer above (for the control plane) is the Non-Access Stratum (NAS) protocol that terminates in the core network side and was addressed in Chapter 3. This chapter describes the general radio interface protocol architecture and then the main functions of the MAC, RLC, PDCP and RRC layers are introduced. The radio protocol aspects of the control plane of the X2 interface – the interface between eNodeBs – is also covered. This chapter concludes with the introduction of the LTE Inter Operability Testing (IOT) bits intended for early UE handling in LTE.

6.2 Protocol Architecture

The overall LTE radio interface protocol architecture is shown in Figure 6.1, covering only the protocol part of the radio access in LTE. Additionally there are protocols in the core network that are between the UE and the core network but these are transparent to the radio layers and are generally referred to as Non-Access Stratum (NAS) signaling.

The physical layer carries the transport channels provided by the MAC layer. The transport channels, listed in Chapter 5, describe how and with what characteristics data are carried on the radio interface on the physical channels. The MAC layer offers the logical channels to the RLC layer. The logical channels characterize the type of data to be transmitted and are covered in detail in section 6.4. Above the RLC layer there is the PDCP layer, now used for both the control and the user plane, in contrast to WCDMA where it was only used for user plane data. Layer 2 provides upwards radio bearers. Signaling Radio Bearers (SRBs) carry the RRC

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6

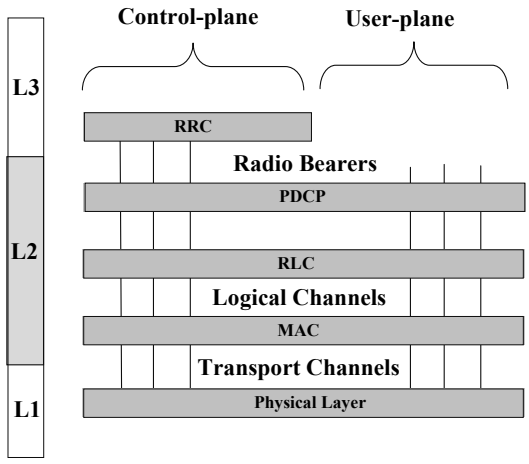


Figure 6.1 LTE Radio Protocol Stacks

signaling messages. Correspondingly the user plane Radio Bearers (RBs) carry the user data. As described in Figure 6.2 on radio protocol architecture for the control plane, MAC, RLC, PDCP and RRC are all located in the eNodeB. With WCDMA RRC and PDCP were both in the RNC, and the MAC layer was either in a NodeB (especially for HSPA) or in a Radio Network Controller (RNC) in case of Release'99 WCDMA.

The control plane signaling between different network elements is carried on the X2 interface for inter-eNodeB communications. For the traffic between the Mobility Management Entity (MME) and eNodeB the S1_MME is used for mobility management related signaling in the control plane. The control plane interfaces inside E-UTRAN and between E-UTRAN and MME are shown in Figure 6.3, as introduced in Chapter 3. All the radio mobility related decisions are still carried out in eNodeB but the MME acts as the mobility anchor when the UE moves between different eNodeBs. The X2 protocol stack itself is not shown, but rather the architectural aspects. The full protocol stack for X2, including the underlying transport layers, is described in Chapter 3, while later in this chapter more details are discussed for radio related signaling

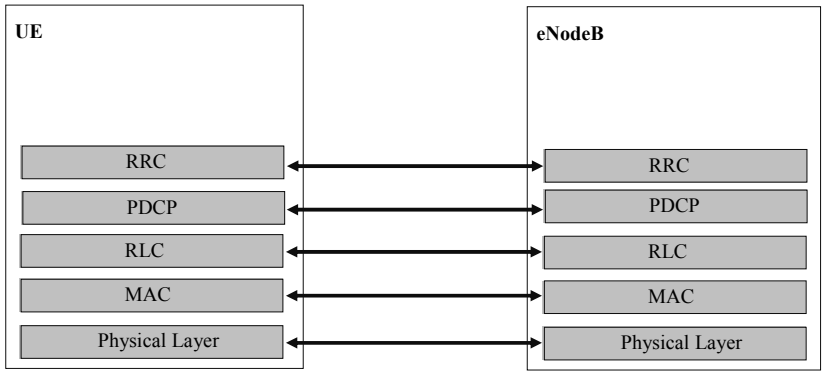


Figure 6.2 LTE control plane radio protocols in LTE architecture

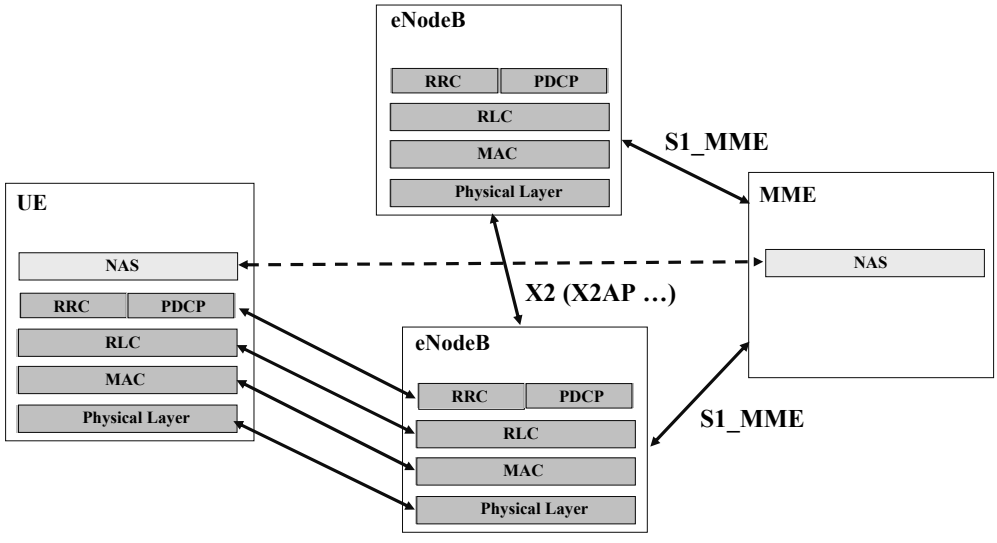


Figure 6.3 LTE control plane radio protocols in LTE architecture

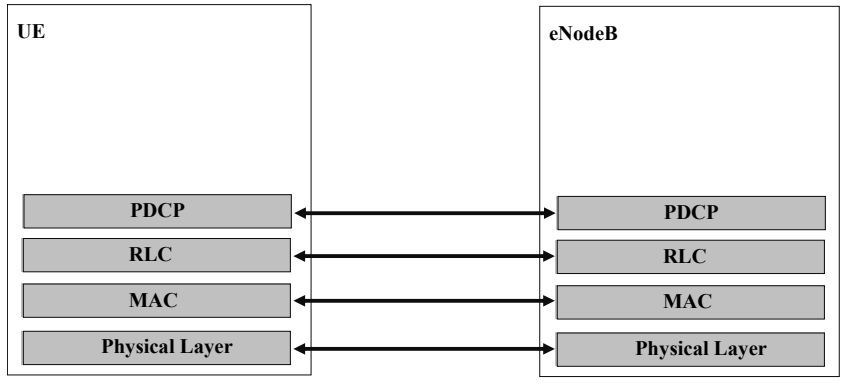


Figure 6.4 LTE user plane radio protocols in LTE architecture

in the X2 Application Protocol (X2AP). Note also that most of the RRC signaling traffic goes through the PDCP layer, thus PDCP is part of the control plane as well.

For the user plane, similarly all the network side user plane radio protocols are located in the eNodeB as shown in Figure 6.4. This was enabled from the system design point of view also because the radio technology chosen does not need macro-diversity, thus flat architecture was easily facilitated without the need for providing the user data continuously over the X2 interface.

6.3 Medium Access Control

The MAC layer maps the logical channels to transport channels, as shown in Figure 6.5. The other tasks of the MAC layer in LTE are:

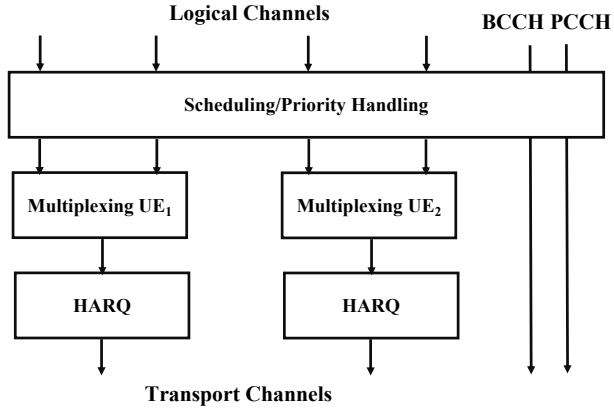


Figure 6.5 MAC layer (downlink)

- MAC layer multiplexing/demultiplexing of RLC Payload Data Units (PDUs) belonging to one or different radio bearers into/from Transport Blocks (TB) delivered to/from the physical layer on transport channels; also the Padding if a PDU is not fully filled with data.
- Traffic volume measurement reporting, to provide RRC layer information about the traffic volume experienced.
- Error correction through HARQ, to control the uplink and downlink physical layer retransmission handling in the eNodeB together with the scheduling functionality.
- Priority handling between logical channels of one UE and between UEs by means of dynamic scheduling, thus the scheduling in the eNodeB is considered as MAC layer functionality similar to HSPA.
- Transport format selection (as part of the link adaptation functionality in the eNodeB scheduler).

Compared to WCDMA there is no ciphering functionality in the MAC layer, neither is there transport channel type switching as the user data are only transmitted over a single type of transport channel (Uplink Shared Channel [UL-SCH] or Downlink Shared Channel [DL-SCH]). As the user identification is based on the physical layer signaling, there is no need to use the MAC layer for UE identification. The downlink MAC layer functionality in Figure 6.5 is otherwise identical in the uplink direction, but obviously the Broadcast Control Channel (BCCH) and Paging Control Channel (PCCH) are not present there and only one user is considered in the UE uplink MAC structure. The MAC layer details in 3GPP are covered in [1]. The MAC layer specification also covers the random access procedure description, including the power ramping parameterization as covered in connection with the physical layer description in Chapter 5.

6.3.1 Logical Channels

The MAC layer provides the service to the RLC layer by logical channels. Different logical channels are defined for different data transfer services in the uplink and downlink directions.

The uplink logical channels mapping to transport channels are shown in Figure 6.6. The following logical channels are defined in LTE uplink:

- Common Control Channel (CCCH) transfers control information between the UE and the network; used when no RRC connection exists between the UE and the network.
- Dedicated Control Channel (DCCH) is a point-to-point channel for dedicated control information between the UE and the network.
- Dedicated Traffic Channel (DTCH) carries all the user data for point-to-point connection.

In the uplink direction all the logical channels are mapped to the UL-SCH, there is no logical channel mapped on the Random Access Channel (RACH) as it is not carrying any information above the MAC layer.

The following logical channels are defined in the LTE downlink, with the mapping to the downlink transport channels as illustrated in Figure 6.7:

- CCCH, DCCH and DTCH have the same functionality as their uplink counterparts, now just delivering the control or user data information in the downlink direction. They are mapped to the DL-SCH in the transport channels.
- Multicast Control Channel and Multicast Traffic Channel are not included in the Release 8 version of the specifications but are expected to be part of Release 9 LTE (or a later release) version, due at the end of 2009. Their intention is to carry multicast data (and related control information) in a point-to-multipoint fashion, similar to the Multimedia Broadcast Multicast Service (MBMS) part of WCDMA in Release 6.
- BCCH carries the broadcast information, such as the parameters necessary for system access. It uses the Broadcast Channel (BCH) as the transport channel for the Master Information Block (MIB) while the actual System Information Blocks (SIBs) are mapped on DL-SCH.
- Paging Control Channel (PCCH) carries the paging information, to enable the network to page a device not in connected mode and is mapped to the Paging Channel (PCH).

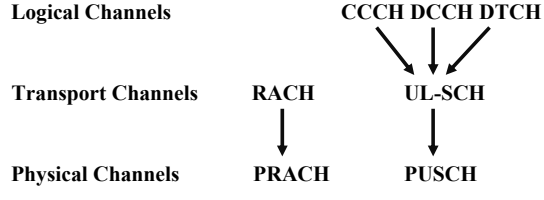


Figure 6.6 PDCP layer operation

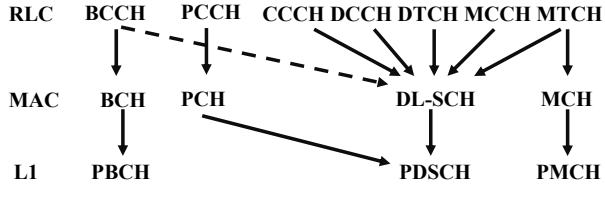


Figure 6.7 Mapping of the downlink logical and transport channels

6.3.2 Data Flow in MAC Layer

The MAC layer receives data from the RLC layer, as MAC Service Data Units (SDUs). The MAC PDU then consists of the MAC header, MAC SDUs and MAC control elements. The MAC control elements carry important control information that is used for several control functionalities, as shown in Figure 6.8. In the uplink direction (in connection with UL-SCH), in addition to data (in MAC SDU) the MAC payload can also contain several control information elements, such as:

- buffer status report, to tell how many data there are in the UE waiting for transmission and information about the priority of the data in the buffer;
- power headroom report to indicate the available uplink transmission power resources;
- contention resolution procedure information (CCCH and C-RNTI).

In the downlink direction (in connection with DL-SCH), the following control information can be carried in the MAC control elements:

- control of the Discontinuous Reception (DRX) operation (when to start, when to stop, etc.);
- timing advance commands to adjust uplink timing;
- contention resolution information.

The MAC header and payload for RACH are different, as is the key to enable the access procedure to be completed and deal with potential collisions or overload situations in the cell

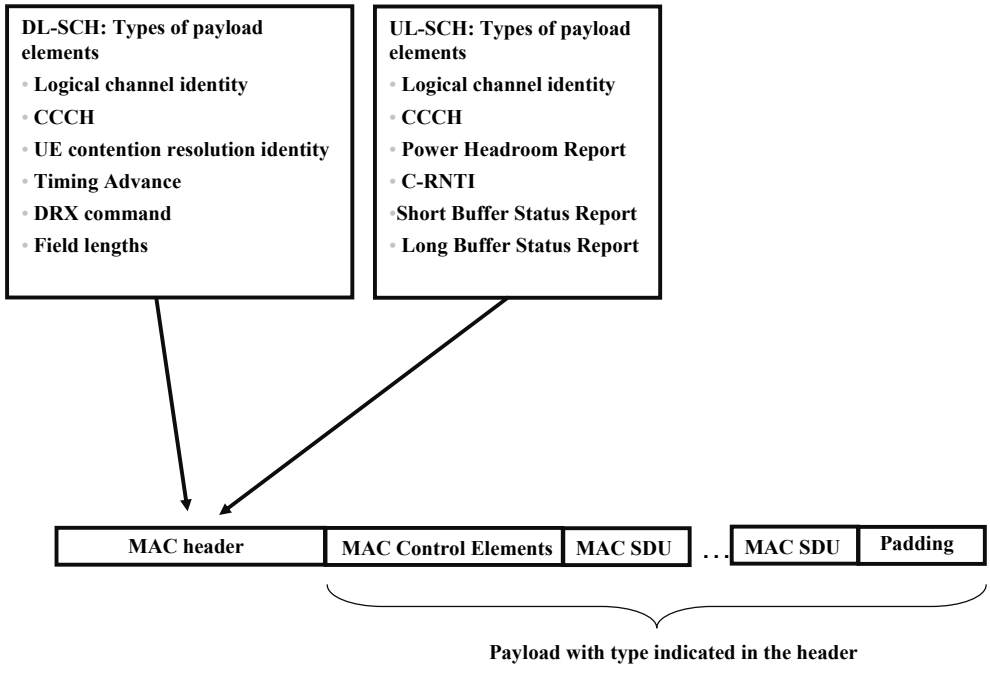


Figure 6.8 MAC PDU structure and payload types for DL-SCH and UL-SCH

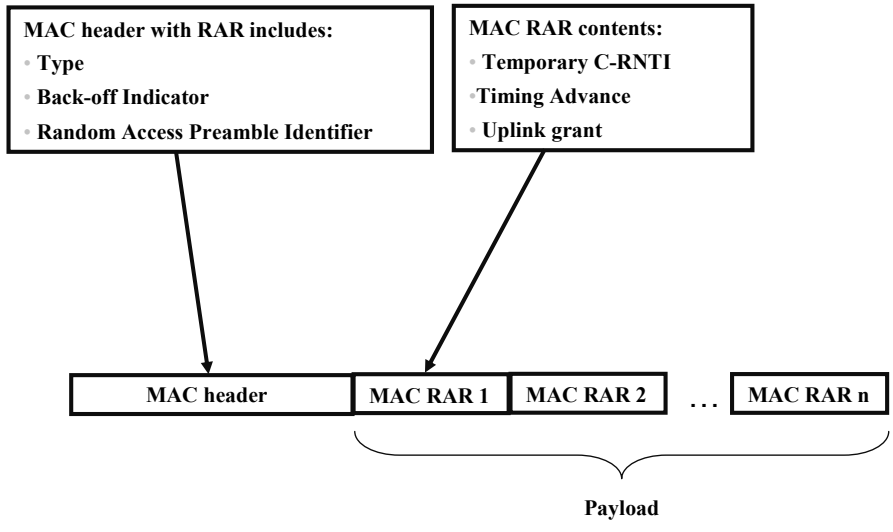


Figure 6.9 MAC PDU structure with random access response

that might have occurred. The MAC layer random access signaling is indicated in Figure 6.9, and it consists of the necessary control information needed to enable data transmission after a successful RACH procedure (detection of preamble after the ramping as covered in Chapter 5). The information transmitted in the MAC Random Access Response (RAR) covers:

- timing advance indicating the necessary adjustment needed for the uplink transmission to ensure different users fit without overlap or extra guard periods for the 1 ms resource allocations;
- uplink grant indicates the detailed resources (time and frequency domain) to be used in the uplink direction;
- temporary C-RNTI provides the UE with a temporary identity to be used in completing the random access operation;
- the header indicates possible back-off and the identification of the preamble received, which the UE had chosen (randomly) in connection with the physical layer random access procedure as described in Chapter 5.

6.4 Radio Link Control Layer

The RLC layer has the following basic functionalities:

- transferring the PDUs received from higher layers, i.e. from RRC (Common Control Channel) or PDCP (other cases, including user plane);
- then (depending on the RLC mode used), error correction with ARQ, concatenation/segmentation, in-sequency delivery and duplicate detection may be applied;
- protocol error handling to detect and recover from the protocol error states caused by, for example, signaling errors.

Fundamentally the key difference of the LTE RLC specifications [2] when compared to WCDMA is the lack of ciphering functionality in RLC. Also the re-segmentation before RLC retransmission is enabled. Thus compared to the order of sequences in [3], the buffer is before segmentation/concatenation functionality.

6.4.1 RLC Modes of Operation

The RLC can be operated in three different modes:

- Transparent Mode (TM). In the TM mode the RLC only delivers and receives the PDUs on a logical channel but does not add any headers to it and thus no track of received PDUs is kept between the receiving and transmitting entity. The TM mode of operation is only suited for services that do not use physical layer retransmissions or that are not sensitive to delivery order. Thus from the logical channel only BCCH, CCCH and PCCH can be operated in TM mode. In WCDMA the AMR speech call (also CS video) was used in transparent mode as well because there were no retransmissions, but now because all the user data have physical layer retransmissions and as the minimum in-sequence delivery is expected, other modes are used instead.
- Unacknowledged Mode (UM) of operation does provide more functionality, including in-sequence delivery of data which might be received out of sequence due to HARQ operation in lower layers. The UM Data (UMD) are segmented or concatenated to suitable size RLC SDUs and the UMD header is then added. The RLC UM header includes the sequence number for facilitating in-sequence delivery (as well as duplicate detection). As shown in Figure 6.10, the receiving side will then be based on the header information due to the re-ordering and then the necessary reassembly in response to the segmentation or concatenation is applied. In addition to the DCCH and DTCH, the UM RLC has been planned for use with multicast channels (MCCH/MTCH) expected to be completed beyond Release 8.
- Acknowledged Mode (AM) of RLC operation provides, in addition to the UM mode functionalities, also retransmission if PDUs are lost as a result of operations in the lower layers.

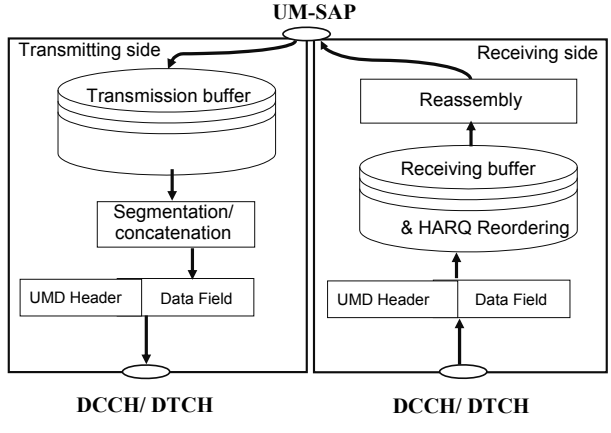


Figure 6.10 RLC UM operation

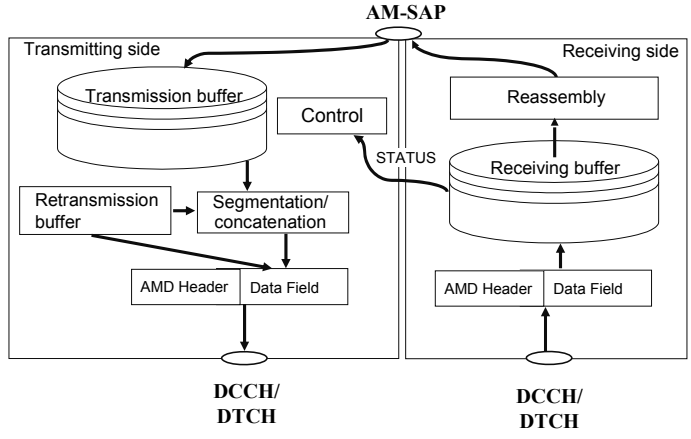


Figure 6.11 RLC AM mode operation

The AM Data (AMD) can also be re-segmented to fit the physical layer resources available for retransmissions, as indicated in Figure 6.11. The Header now contains information on the last correctly received packet on the receiving side along with the sequence numbers, as for the UM operation.

6.4.2 Data Flow in RLC Layer

For the AM RLC operation, the RLC layer receives the data from the PDCP layer, the data are stored in the transmission buffer and then, based on the resources available, segmentation or concatenation is used. In the downlink direction (DTCH/DCCH) the following control means are available for AM RLC purposes, either in the AMD PDU or in the status PDU:

- 10 bit Sequence Number (SN), to enable a long enough in-sequence delivery window;
- last correctly received SN and incorrectly received SN(s) (detected to be lost).

When the receiving entity gets a RLC PDU, it checks for possible duplicated parts and will then forward it further for re-assembly (assuming no SNs are missing in between) and provides data to the next protocol layer (in this case for PDCP) for further processing. If there was a SN missing, the receiving entity will request a retransmission from the transmitting side.

6.5 Packet Data Convergence Protocol

The Packet Data Convergence Protocol (PDCP) is located above the RLC layer of the user plane and PDCP is also used for most of the RRC messages. The key difference to WCDMA is that now all user data go via the PDCP layer, because ciphering is now in the PDCP, which is located in the eNodeB. In some early plans of the LTE architecture PDCP was on the other side of the S1 interface (on the core network) but was later placed in eNodeB along with all other radio protocols. The key functionalities of the PDCP are:

- Header compression and corresponding decompression of the IP packets. This is based on the Robust Header Compression (ROHC) protocol, specified in the Internet Engineering Task Force (IETF) [4–9] and also part of the WCDMA PDCP layer. Header compression is more important for smaller IP packets in question, especially in connection with the VoIP service, as the large IP header could be a significant source of overhead for small data rates.
- Ciphering and deciphering both the user plane and most of the control plane data, a functionality that in WCDMA was located in the MAC and RLC layers.
- Integrity protection and verification, to ensure that control information is coming from the correct source.

The PDCP layer receives PDCP SDUs from the NAS and RRC and after ciphering and other actions, as shown in Figure 6.12 for operation of packets that are associated to a PDCP SDU, the data are forwarded to the RLC layer. Correspondingly, in the receiving side the data are received from the RLC layer. Besides the functionalities listed above, the PDCP layer has specific functions in connection with the handover events (intra-LTE). The PDCP does the in-order delivery function in the downlink direction and detects duplicates. In the uplink direction, PDCP retransmits all the packets which have not been indicated by lower layers to be completed, as the lower layers will flush all the HARQ buffers with handover. In the downlink direction, the PDCP layer will forward the non-delivered packets to the new eNodeB as described in the Chapter 7. This is to ensure that no data are lost in connection with a handover event between LTE eNodeBs. The LTE PDCP specification in 3GPP is covered in [10].

6.6 Radio Resource Control (RRC)

Radio Resource Control messages are a major part of the control information exchanged between the UE and E-UTRAN. The RRC in E-UTRAN has been simplified significantly compared to that in UTRAN by reducing the number of messages and the redundancies in the messages. RRC uses the same protocol language as WCDMA – Abstract Syntax Notation One (ASN.1)

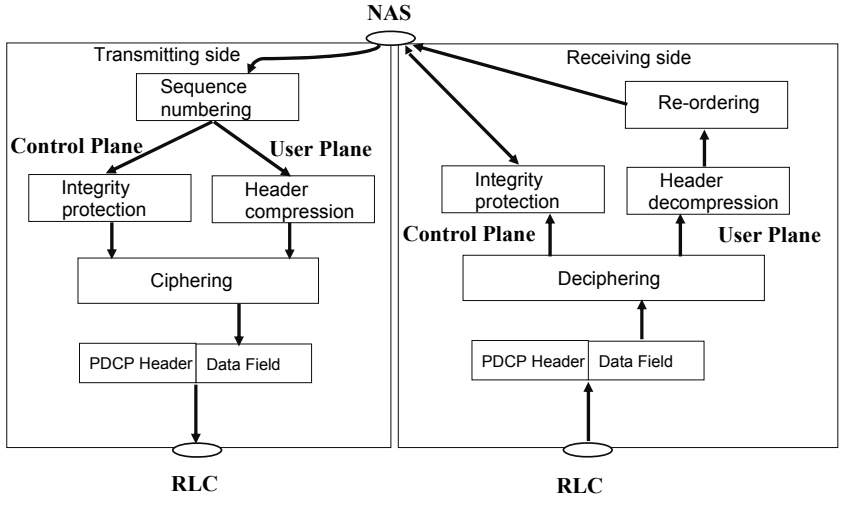


Figure 6.12 PDCP layer operation for the packets which are associated to a PDCP SDU

– as it has been efficient in facilitating the evolution between different release versions with an extension capability within ASN.1. The encoding on the X2 and S1 interface messages also uses ASN.1. The LTE RRC specification in [11] has the ASN.1 message descriptions at the end.

6.6.1 UE States and State Transitions Including Inter-RAT

Contrary to the UTRAN (WCDMA), UE states in E-UTRAN are also simplified significantly and there are only two states, i.e., RRC_CONNECTED and RRC_IDLE, depending on whether the RRC connection has been established or not.

In the RRC_IDLE state, the UE monitors a paging channel to detect incoming calls, acquires system information and performs neighboring cell measurement and cell (re)selection. In this state, a UE specific DRX may be configured by the upper layer and the mobility is controlled by the UE.

In the RRC_CONNECTED state, the UE transfers/receives data to/from the network. For this, the UE monitors control channels that are associated with the shared data channel to determine if data are scheduled for it and provides channel quality and feedback information to eNodeB. Also in this state, the UE performs neighboring cell measurement and measurement reporting based on the configuration provided by eNodeB. In contrast with the UTRAN system, the UE can acquire system information from BCCH during the RRC_CONNECTED state. At lower layers the UE may be configured with a UE specific DRX and the mobility is controlled by the network, i.e. handover.

6.6.1.1 E-UTRAN States and Inter-RAT State Mobility

Figure 6.13 shows the mobility support between E-UTRAN, UTRAN and GSM EDGE Radio Access Network (GERAN). As CELL_FACH in UTRAN is considered a very short period, a direct transition from UTRAN CELL_FACH to E-UTRAN RRC state is not supported.

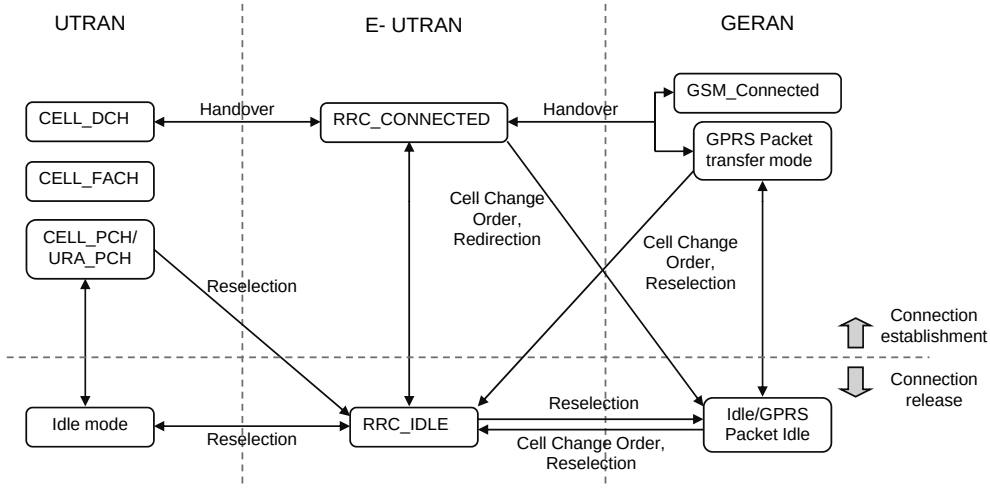


Figure 6.13 E-UTRAN RRC States and State transitions among 3GPP systems

6.6.1.2 Signaling Radio Bearers (SRB)

SRBs are special radio bearers which convey only RRC messages and NAS messages. There are three SRBs defined. SRB0 is used for RRC messages using CCCH, like during RRC connection set-up or during radio link failure. Thus, for instance, the following messages are transferred via SRB0: RRC Connection Request message, RRC Connection Setup message, RRC Connection Reject message, RRC Connection Reestablishment Request message, RRC Connection Reestablishment message, RRC Connection Reestablishment Reject message. Once a RRC connection is established, SRB1 is used to transfer both RRC messages using DCCH and NAS messages until the security is activated. Once the security is successfully activated, SRB2 is set up and NAS messages are transferred via SRB2 while RRC messages are still transferred via SRB1. SRB2 has a lower priority than SRB1.

6.6.2 RRC Functions and Signaling Procedures

The following functions are provided by RRC protocol layer:

- broadcast of system information
- paging
- establishment, maintenance and release of an RRC connection between the UE and e-UTRAN
- security functions including key management
- establishment, configuration, maintenance and release of point-to-point Radio Bearers
- UE measurement reporting and control of the reporting
- handover
- UE cell selection and reselection and control of cell selection and reselection
- context transfer between eNodeBs
- NAS direct message transfer between network and UE
- UE capability transfer
- generic protocol error handling
- support of self-configuration and self-optimization.

6.6.2.1 Broadcast of System Information

System information contains both non-access stratum (NAS) and access stratum (AS) related information. Based on the characteristics and usages of the information, the system information elements are grouped together into Master Information Block (MIB) and different System Information Blocks (SIBs).

As MIB is the most important information block, MIB is transferred on the BCH every 40 ms and is repeated within 40 ms. The first transmission of the MIB is scheduled at SFN mod $4 = 0$ in the subframe #0. UE will acquire MIB to decode SCH. The MIB contains a DL system bandwidth, PHICH configuration and a system frame number (SFN).

SIB1 is scheduled in a fixed manner with a periodicity of 80 ms and is repeated within 80 ms. The first transmission of the SIB1 is scheduled at SFN mod $8 = 0$ in the subframe #5. SIB 1 contains cell access related information (e.g. a PLMN identity list, tracking area code, cell identity, etc.), information for cell selection (e.g. minimum required Rx level in the cell

and offset), p-Max, a frequency band indicator, scheduling information, TDD configuration, SI-window length and system information value tag.

All other SIBs except SIB1 are contained in SI message(s). Each SI message is transmitted periodically in time domain windows (i.e. SI-Window) and SI-Windows for different SI messages do not overlap. The length of a SI-window is defined in SIB1 and is common for all SI messages.

Figure 6.14 shows how the UE can find each SI message to read the SIBs in it.

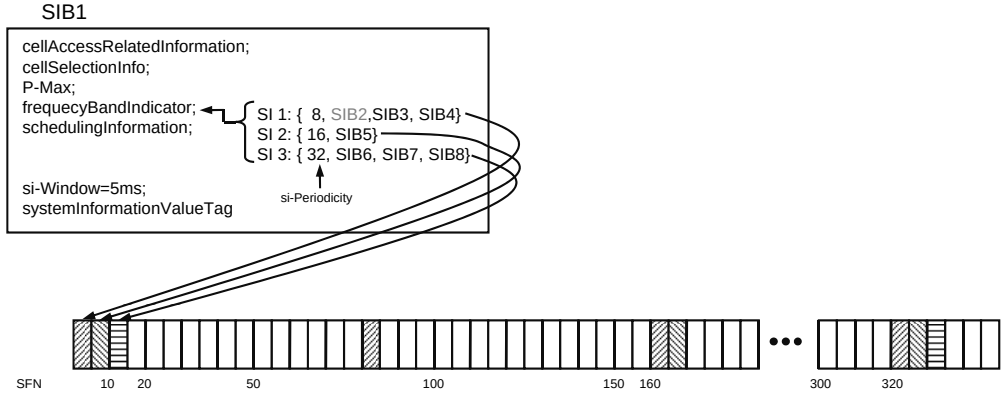
SIB2 is not listed in the scheduling information in SIB1 but the first SI message contains SIB2 as the first entry.

SIB2 contains radio resource configuration information which is common to all UEs. More specifically, SIB2 includes access barring information, radio resource configuration of common channels (RACH configuration, BCCH configuration, PCCH configuration, PRACH configuration, PDSCH configuration, PUSCH configuration and PUCCH configuration, sounding reference signal configuration, uplink power control information), timers and constants which are used by UEs, MBSFN subframe configuration, time alignment timer and frequency information (UL EARFCN, uplink bandwidth and additional spectrum emission).

SIB3 contains cell-reselection information that is common for intra-frequency, inter-frequency and/or inter-RAT cell re-selection. Speed dependent scaling parameters are also included in SIB3 to provide a different set of re-selection parameters depending on UE speed.

SIB4 contains neighbor cell related information only for intra-frequency cell re-selection. Thus in addition to intra-frequency neighboring cell list with Q_{offset}, an intra-frequency blacklisted cell list is also included in SIB4.

SIB5 contains information relevant only for inter-frequency cell re-selection like E-UTRAN inter-frequency neighboring cell related information and E-UTRAN inter-frequency blacklisted cell list.



determine integer value $x = (n - 1)*w$, where n is the order of entry in the list of SI message and w is the *si-WindowLength*

si-Window starts at subframe #a, where a = x mod 10, in the next radio frame for which SFN mod T = FLOOR (x/10) where T is the si-Periodicity of the concerned SI message

Figure 6.14 Acquisition of SI message

SIB6 contains information relevant only for cell re-selection to the UTRAN. Thus UTRA FDD and TDD frequency information is included in SIB6.

SIB7 contains information relevant only for cell re-selection to the GERAN such as GERAN neighboring frequency list.

SIB8 contains information relevant only for cell re-selection to the cdma2000® system. Thus SIB8 contains cdma2000® system time information, HRPD related parameters and 1xRTT related parameters.

SIB9 contains a home eNodeB identifier, which is a maximum of 48 octets so that the UE can display the home eNodeB identifier to help manual home eNodeB selection.

SIB10 contains Earthquake and Tsunami Warning System (ETWS) primary notification and SIB 11 contains ETWS secondary notification.

In case the system information is modified, the SI messages may be repeated during the *modification period*. The *modification period* starts at SFN mod *modification period* = 0. During the first modification period, system information modification indicator is sent to the UE by paging message and in the next modification period, the network transmits the updated system information.

The *modification period* is calculated as follows:

modification period (in number of radio frames) = modificationPeriodCoeff * defaultPaging-Cycle DIV 10 ms

where modificationPeriodCoeff is: signaled in BCCH-Configuration in SIB2, value = 1, 2, 4, 8 and defaultPagingCycle is: 320 ms, 640 ms, 1280 ms, 2560 ms (signaled in PCCH-configuration in SIB2).

The value tag in SIB1 is to indicate a change in SI messages. The SI message is considered to be valid for a maximum of 3 h from the moment it was acquired.

6.6.2.2 Paging

The main purpose of a paging message is to page UEs in RRC_IDLE mode for a mobile terminated call. Also the paging message can be used to inform UEs, in RRC_IDLE as well as in RRC_CONNECTED modes, that system information will be changed or that the ETWS notification is posted in SIB10 or SIB11.

As UE may use DRX in idle mode to reduce power consumption, UE should be able to calculate when to wake up and to check the correct subframe. For this, the UE stores the default paging cycle and the number for the paging group when it receives the necessary information from the SIB2 and applies it to the calculation as described in section 7 in TS36.304.

UE can be paged by either S-TMSI (temporary UE ID allocated by MME) or IMSI. In Release 8, IMEI paging is not considered.

For the CS fallback, if UE had registered to a CS core network and received a temporary ID from the CS core network, CS paging can be delivered to UE with a corresponding core network indicator (i.e. CS or PS). In this case, the UE should deduce the correct CS UE identity based on the PS UE identity included in the paging message and include the CS UE identity in the paging response once it is moved to another RAT. The details for the CS fallback procedure can be found in section 6.6.2.8.

6.6.2.3 UE Cell Selection and Reselection and Control of Cell Selection and Reselection

Based on idle mode measurements and cell selection parameters provided in SIBs, the UE selects the suitable cell and camps on it. In the LTE system, the priority parameter has been newly introduced. eNodeB can provide a priority per LTE frequency and per RAT in SIBs or RRC connection Release message. In case the UE is camping on a cell on the highest priority frequency, the UE doesn't need to measure cells in other frequencies or RAT as long as the signaling strength of the serving cell is above a certain level. On the other hand, if the UE is camped on a cell in the low priority layer, it should measure other cells under higher priority frequencies or RAT regularly.

This priority concept is also adapted to Release 8 UTRAN and GERAN (or GSM) system for inter-RAT cell reselection.

The LTE idle mode mobility is explained in Chapter 7 in more detail.

6.6.2.4 Establishment, Maintenance and Release of an RRC Connection Between the UE and E-UTRAN

The RRC Connection Setup procedure, as shown in Figure 6.15, is triggered by a request from the UE NAS layer for various reasons such as a mobile originated call, NAS signaling transfer or a paging response. Consequently the RRC connection is established between UE and eNodeB and SRB1 is set up. If the network is overloaded, eNodeB can set access class barring parameters in SIB2 appropriately and/or reject the RRC Connection Request with a wait timer.

If the UE has a valid S-TMSI, the UE includes it in the RRC connection request message. Otherwise the UE includes a 40 bit random value. Due to the limited size of the message, only five establishment causes are defined in the RRC Connection Request message, i.e. emergency, high priority access, mobility terminated access, mobile originated signaling and

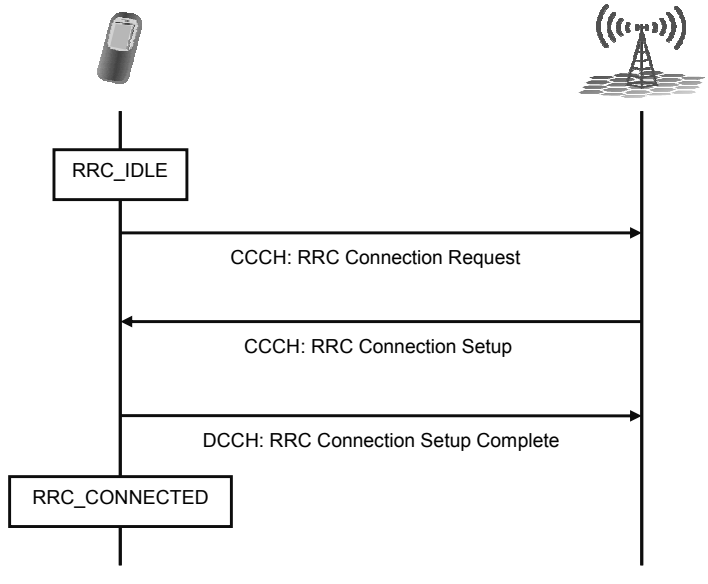


Figure 6.15 RRC Connection Setup procedure

mobile originated data. An NAS service request message is conveyed in the RRC Connection Setup Complete message.

SIB1 broadcasts at most 6 PLMN identities (which can also facilitate network sharing) and the UE selects one and reports it in the RRC Connection Setup Complete message. When the UE is registered to an MME, the UE also includes the identity of the registered MME in the RRC Connection Setup Complete message. Then eNodeB finds/selects the correct MME (which stores the UE idle mode context) in case that S1-flex is deployed and starts S1 connection setup.

After a successful RRC Connection setup procedure, the UE moves to the RRC_ CONNECTED state.

6.6.2.5 Security Functions Including Key Management

The security keys for the Access Stratum (AS), covering user data and RRC control signaling, are different from those used on the Evolved Packet Core (EPC) side. From the eNodeB perspective the following keys are necessary (ignoring the keys used for NAS signaling):

- K_{eNB} is derived by UE and MME from the 'master key' (K_{ASME}) and then provided by MME to eNodeB. K_{eNB} is used to derive the necessary keys for AS traffic and also for the derivation of K_{eNB}* during handover.
- K_{eNB}* is derived by UE and source eNodeB from either K_{eNB} or from a valid Next Hop (NH) in case Next Hop is available. At handover, UE and target eNodeB will derive a new K_{eNB} for AS traffic from K_{eNB}*.
- K_{UPenc} is used for the user plane traffic ciphering with the key and is derived from $K_{eNB.}$
- K_{RRCint} is derived from K_{eNB} to be used for RRC message integrity handling.
- K_{RRCenc} is derived from K_{eNB} to be used for RRC message ciphering.
- Next Hop (NH) is an intermediate key that is used for deriving K_{eNB}* for the provision of security for intra-LTE handover purposes. There is a related counter, Next Hop Chaining Counter (NCC), which determines whether the next K_{eNB}* needs to be based on the current K_{eNB} or if a fresh NH is needed. If no fresh NH is available, the K_{eNB}* is derived by UE and the source eNodeB from the target PCI and K_{eNB} (horizontal key derivation) or if a fresh NH is available, then derivation is based on the target PCI and a new NH is provided (vertical key derivation).

The EPC and UE will share the same master key after the security procedure is completed. This master key (K_{ASME}) is not provided outside the EPC but keys derived from it and necessary for the traffic between eNodeB and UE are delivered to the eNodeB. The eNodeB retains a given key as long as the UE is connected to it but will delete the keys when the UE is moving to idle mode (or to another eNodeB). The integrity and ciphering algorithm can be changed only upon handover. The AS keys are changed upon every handover and connection re-establishment. In the LTE system, a handover can be performed only after security is activated.

RRC integrity and ciphering are always activated together and never de-activated, but it is possible to use a NULL ciphering algorithm. In addition to the keys, security algorithms also use a sequence number as inputs, which consists of a PDCP Sequence Number and a Hyper Frame Number (HFN).

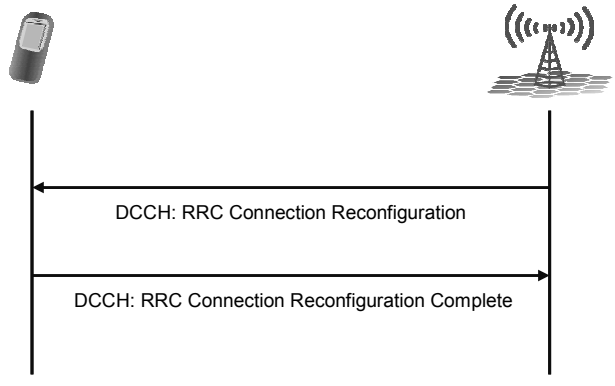


Figure 6.16 RRC Connection Reconfiguration procedure

6.6.2.6 Establishment, Maintenance and Release of Point-to-point Radio Bearers

A RRC connection reconfiguration procedure, as shown in Figure 6.16, is used to maintain and modify radio bearers and to release data radio bearers, i.e. SRBs cannot be released by the RRC connection reconfiguration procedure.

The parameters defined in the radio resource configuration information element group in the RRC Connection Reconfiguration message mainly deal with the configuration of the radio bearer. When the RRC Connection Reconfiguration message is used to set up a new Data Radio Bearer (DRB), a corresponding NAS message is also included within this message.

6.6.2.7 UE Measurement Reporting and Control of the Reporting

A measurement configuration parameter in the RRC Connection reconfiguration message is used to configure measurements in UEs (for measurements inside LTE or for measurements from other Radio Access Technologies [RATs] and a measurement report message is used for reporting.

The measurement configuration consists of the following parameters.

1 Measurement objects: The objects on which the UE shall perform the measurements. E-UTRAN configures only a single measurement object for a given frequency.
   (a) For intra-frequency (i.e. measurements at the same downlink carrier frequency as the serving cell) and inter-frequency measurement (i.e. measurements at frequencies which differ from the downlink carrier frequency of the serving cell), a measurement object is a single E-UTRA carrier frequency. For this carrier frequency, E-UTRAN can configure a list of cell specific offsets and a list of blacklisted cells.
   (b) For inter-RAT UTRA measurements, a measurement object is a set of cells on a single UTRA carrier frequency.
   (c) For inter-RAT GERAN measurements, a measurement object is a set of GERAN carrier frequencies.
   (d) For inter-RAT cdma2000® measurements, a measurement object is a set of cells on a single cdma2000® carrier frequency.

2 Reporting configuration: The reporting configuration contains the reporting criteria and reporting format. The reporting criteria are for the UE to trigger a measurement report and can be either event triggered or periodic. Periodic reporting is especially used for measuring an automatic neighbor cell search. Reporting format includes the quantities which should be included in the measurement report (e.g. the number of cells to report).
   (a) Reporting criteria: For an event triggered report, the criteria for E-UTRA measurements are A1, A2, A3, A4 and A5. For inter-RAT measurement, B1 and B2 are used.
   (b) Event A1: Serving becomes better than absolute threshold.
   (c) Event A2: Serving becomes worse than absolute threshold.
   (d) Event A3: Neighbor becomes amount of offset better than serving.
   (e) Event A4: Neighbor becomes better than absolute threshold.
   (f) Event A5: Serving becomes worse than absolute threshold 1 and Neighbor becomes better than another absolute threshold 2.
   (g) Event B1: Neighbor becomes better than absolute threshold.
   (h) Event B2: Serving becomes worse than absolute threshold 1 and Neighbor becomes better than another absolute threshold 2.
3 Measurement identities: Measurement identity binds one measurement object with one reporting configuration. The measurement identity is used as a reference number in the measurement report.
4 Quantity configurations: One quantity configuration can be configured per RAT (i.e. E-UTRA, UTRAN, GERAN and cdma2000® each) and contains the filter coefficient for the corresponding measurement type.
5 Measurement gaps: Periods that the UE may use to perform measurement. Measurement gaps are used for inter-frequency and inter-RAT measurements.

6.6.2.8 Handover

Usually eNodeB triggers handover based on the measurement result received from the UE. A handover can be performed inside E-UTRAN or to E-UTRAN from other RAT or from E-UTRAN to another RAT and is classified as intra-frequency intra-LTE handover, inter-frequency intra-LTE handover, inter-RAT towards LTE, inter-RAT towards UTRAN handover, inter-RAT towards GERAN handover, and inter-RAT towards cdma2000® system handover.

Intra-LTE handover

For intra-LTE handover, the source and target are both in the LTE system, and the RRC connection reconfiguration message with the mobility control information (including the parameters necessary for handover) is used as a handover command. When an X2 interface exists between the source eNodeB and the target eNodeB, the source eNodeB sends X2: Handover Request message to target eNodeB to prepare the target cell. Target eNodeB builds RRC connection reconfiguration message and transfers to source eNodeB in the X2: Handover Request Acknowledge message, as shown in Figure 6.17.

Inter-radio access technology handover to other radio access technology

Based on UE measurement reporting, the source eNodeB can decide the target RAT and in Release 8 the target RAT can be UTRAN, GERAN or cdma2000® system. To perform the inter-RAT handover to another RAT, the actual handover command message is built by the target RAT and is sent

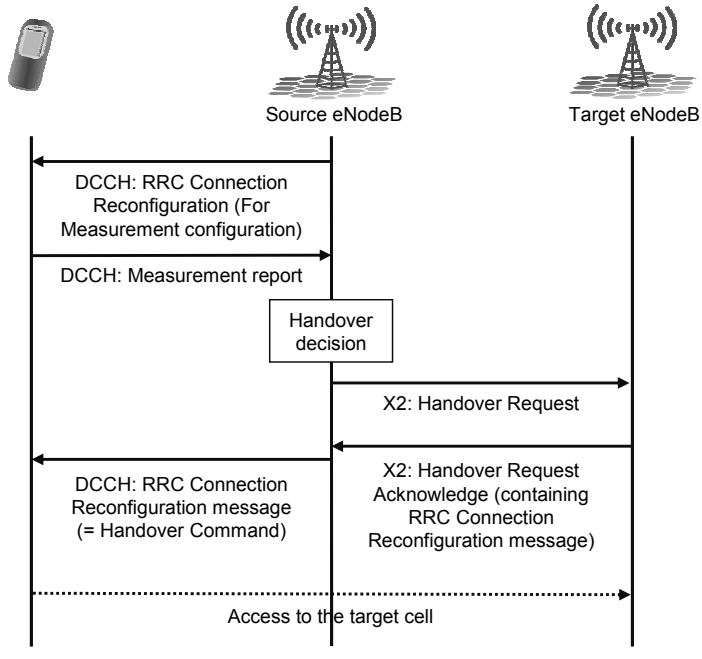


Figure 6.17 Inter-eNodeB Handover procedure

to the source eNodeB transparently. The source eNodeB includes this actual handover command in the Mobility From E-UTRA Command message as a bit string and sends it to the UE.

As the LTE system supports packet service only, if a multi-mode UE wants to start a CS service, CS fallback can be performed. A handover procedure is mainly used for CS fallback to the UTRAN system. Either a handover procedure or a cell change order procedure is used for CS fallback to the GERAN system. A RRC Connection release procedure is used for CS fallback to cdma2000® system. To align the behavior of cell change order with inter-RAT handover, the cell change order procedure uses the Mobility From EUTRA Command message. Also in case, no VoIP service is supported in the target RAT and if the network nodes support, SR-VCC (Single Radio Voice Call Continuity) handover can be performed.

Inter-radio access technology handover from other radio access technology

The handover is triggered based on the criteria defined in the source RAT system. When the target eNodeB receives a S1: Handover Request message from MME, it allocates the necessary resources, builds the RRC Connection Reconfiguration message (=Handover Command) and sends it in the S1: Handover Request Acknowledge message to MME. Thus this RRC Connection Reconfiguration message will be transferred to the UE via source RAT.

CS fallback

As initial LTE deployment may not support voice services as such, 3GPP agreed to provide a means to fallback a CS voice service to the GERAN or UTRAN system. For CS fallback, PS handover, Cell change order and RRC connection release can be used and eNodeB decides the actual method. In Release 8, 3GPP decided that for CS fallback to UTRAN system, only PS handover, and for CS fallback to GERAN system, PS handover and cell change order with the

network assistance (NACC) (i.e. providing some system information about the target cell to speed up the cell selection procedure) should be considered. The two methods are illustrated in Figure 6.18 and Figure 6.19.

For an incoming CS call, the UE is registered to the CS core network in addition to MME. In case the UE receives CS paging or wants to initiate a mobile originated CS call, the UE indicates

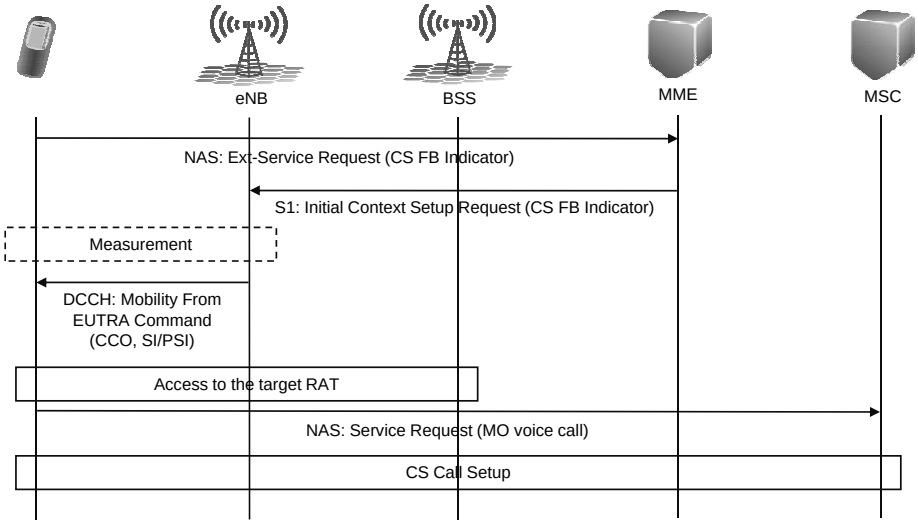


Figure 6.18 CS fallback with Cell Change Order

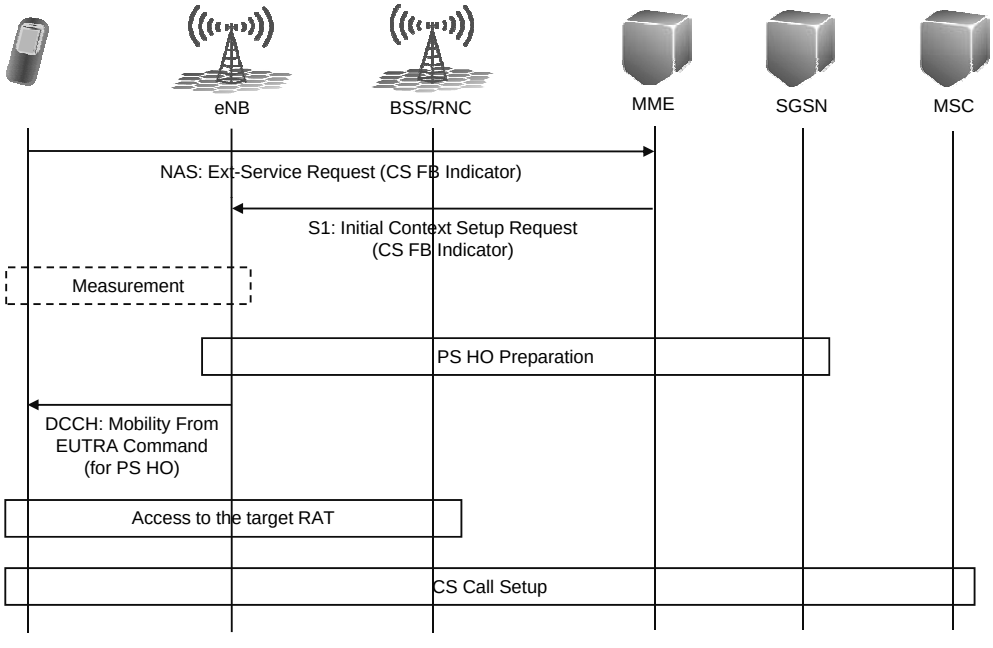


Figure 6.19 CS fallback with PS Handover

that the NAS: Ext-Service Request message is for the CS fallback to MME. Then MME indicates that S1: Initial Context Setup Request message is for the CS Fallback (CS FB). Depending on the target system and the capability of the target cell, eNodeB performs PS handover or cell change order procedure and the UE starts the CS call setup once it moves in to the target cell.

6.6.2.9 Context Transfer Between eNodeBs

During HO preparation, the source eNodeB should transfer the context of the UE to the target eNodeB so that the target eNodeB can take into account the configuration in the source cell and the handover command can contain only differences from the source configuration. This contains AS configuration, RRM configuration and AS Context. AS configuration contains the UE AS configuration configured via the RRC message exchange such as measurement configuration, radio resource configuration, security configuration, system information of source cell and UE ID. RRM configuration contains UE specific RRM information. AS Context contains UE Radio Access Capability and information for re-establishment.

6.6.2.10 NAS Direct Message Transfer Between Network and UE

NAS messages are transferred mainly via dedicated RRC messages, i.e. DL Information Transfer and UL Information transfer. During a bearer setup, however, necessary NAS messages should be transferred via the RRC Connection Complete message and RRC Connection Reconfiguration message.

6.6.2.11 UE capability transfer

UE transfers NAS UE capability (e.g. security UE capability) directly to MME via a NAS message and this is forwarded to eNodeB during the initial UE context setup. As UE capability may contain UTRAN and GERAN capability in addition to LTE capability, the information can be rather large. To reduce the overhead in the air interface during RRC_IDLE mode to RRC_CONNECTED mode transition, MME stores the UE AS capability as well and provides it to eNodeB during initial UE context setup. In case UE AS capability has been changed during RRC_IDLE or MME doesn't have the valid UE AS capability, however, MME does not provide UE AS capability during the initial UE context setup and eNodeB acquires UE AS capability directly from UE as shown in Figure 6.20.

6.6.2.12 Generic Error Handling

In UTRAN RRC, various error handling cases in the DCCH message had been specified thoroughly. For LTE, however, there was no desire to add complexity by specifying the handling of network error cases, such as ASN.1 errors. Thus only error handling for very important procedures are specified and the rest is unspecified. For example, when a RRC Connection Reconfiguration message has an ASN.1 error or integrity protection has failed, UE starts the RRC Connection Re-establishment procedure as shown in Figure 6.21. But for other error cases where the failure response message is not defined, the UE may just ignore the message. Also RRC Connection Re-establishment procedure is used to recover from radio link failure.

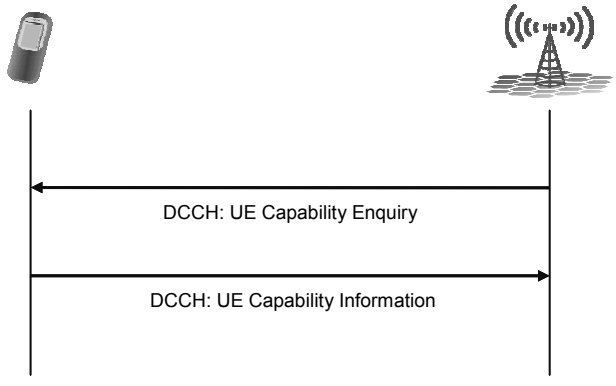


Figure 6.20 UE Capability Transfer procedure

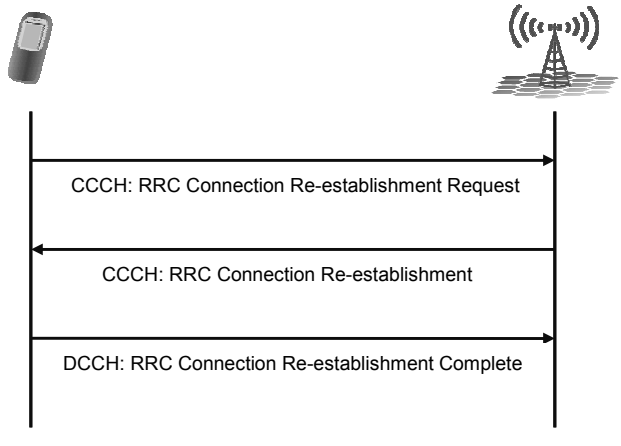


Figure 6.21 RRC Connection Re-establishment procedure

6.6.2.13 Support of Self-configuration and Self-optimization

To reduce the CAPEX and OPEX, there is natural interest in self-configuration and self-optimization solutions for the LTE system. For some key issues there are solutions agreed in Release 8 specifications to support the O&M system functionality in system setup and in fine tuning. X2 and S1 automatic setup has been specified in the X2 and S1 application protocol as addressed in Chapter 3, and the measurements to support the automatic neighbor cell relation are defined in the RRC specification as covered in Chapter 7.

6.7 X2 Interface Protocols

As introduced in Chapter 3, the LTE has the X2 interface between the eNodeBs. The X2 is a logical interface and even though normally drawn as a direct connection between eNodeBs, it is usually routed via the same transport connection as the S1 interface to the site. The X2 interface control and user plane protocol stacks are shown in Figure 6.22, as presented in Chapter 3. X2 is an open interface, similar to the Iur interface in WCDMA. Normally the X2 interface is only

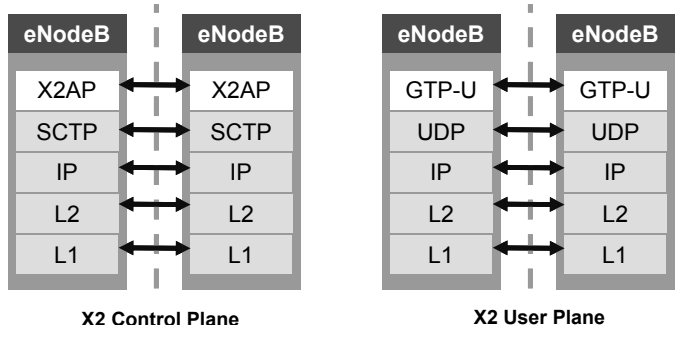


Figure 6.22 X2 interface User and Control Plane protocol stacks

used for control plane information but in connection with the handover, it can be temporarily used for user data forwarding. The key difference between the user plane and control plane X2 interface protocol stacks is the use of Stream Control Transmission Protocol (SCTP) for control plane transmission between eNodeBs. Use of SCTP enables reliable delivery of control plane information between eNodeBs, while for data forwarding, the User Datagram Protocol (UDP) was considered sufficient. The X2 Application Protocol (X2AP) covers the radio related signaling while the GPRS Tunneling Protocol, User Plane (GTP-U) is the protocol also used in the S1 interface for user data handling, as covered in Chapter 3.

The X2AP [12] functionalities are:

- Mobility management for Intra LTE mobility, as covered in more detail in Chapter 7. The handover message between eNodeBs is transmitted on the X2 interface.
- Load management to enable inter-cell interference coordination by providing information of the resource status, overload and traffic situation between different eNodeBs.
- Setting up and resetting of the X2 interface.
- Error handling for covering specific or general error cases.

6.7.1 Handover on X2 Interface

The X2 interface has a key role in the intra-LTE handover operation. The source eNodeB will use the X2 interface to send the Handover Request message to the target eNodeB. If the X2 interface does not exist between the two eNodeBs in question, then procedures need to be initiated to set one up before handover can be achieved, as explained in Chapter 6. The Handover Request message initiates the target eNodeB to reserve resources and it will send the Handover Request Acknowledgement message assuming resources are found, as shown in Figure 6.23.

There are different information elements provided (some optional) on the handover Request message, such as:

- Requested SAE bearers to be handed over.
- Handover restrictions list, which may restrict following handovers for the UE.
- Last visited cells the UE has been connected to if the UE historical information collection functionality is enabled. This has been considered to be useful in avoiding the ping-pong effects between different cells when the target eNodeB is given information on how the serving eNodeB has been changing in the past. Thus actions can be taken to limit frequent

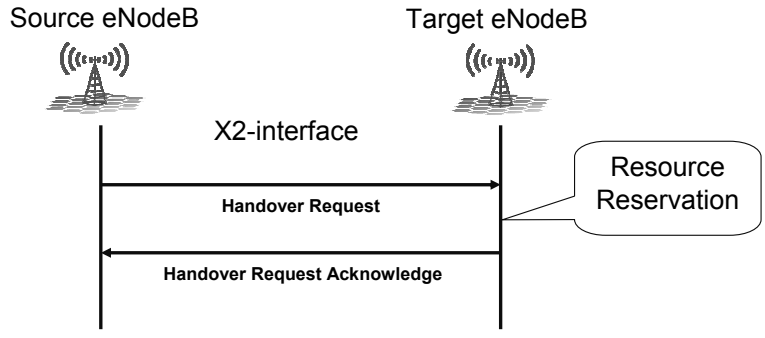


Figure 6.23 X2 Handover preparation over X2 interface

changes back and forth. Similar information is included in WCDMA Release 8 for consideration of similar cases with flat architecture.

Upon sending the handover Request message, the eNodeB will start the timer and if no response is received the handover preparation is cancelled (and an indication that the timer has expired is given to the target eNodeB). There is also a Handover Cancel message to cancel ongoing handover.

The Handover Request Acknowledge message contains more than just the information that the target eNodeB is able to accommodate (at least part) the requested SAE bearers to be handed over. The other information included is:

- GTP tunnel information for each SAE bearer to enable delivery of uplink and downlink PDUs;
- possible SAE bearers not admitted;
- a transparent container having handover information (actual RRC Connection Reconfiguration message) which the source eNodeB will then send to the UE as explained in section 6.6.2.8 and indicated as the handover commands in Chapter 7.

The SN Status Transfer procedure is intended to provide information on the status of each SAE bearer that will be transferred to the eNodeB. The PDCP-SN and HFN are provided for both the uplink and downlink direction for those SAE bearers where status presentation is applied. This will take place at the moment the source eNodeB stops assigning new PDCP SNs in downlink and stops forwarding data over the S1 interface to EPC.

The UE Context Release signals to the source eNodeB that 'the control point' for the UE has now moved to target eNodeB and basically that the source can now release all resources.

6.7.2 Load Management

There are three different measurements specified in Release 8 to support standardized operation for the load control and interference management over the X. The indications are:

- On the transmitter side there is the indication of the transmitted power level in the form of the Relative Narrowband Tx Power (RNTP) I.E.
- On the receiver side there are the received interference level and interference sensitivity.

All these are measured/estimated per Physical Resource Block (PRB) of 180 kHz. The received interference level measurement (or rather indication) on the X2 interface is illustrated in Figure 6.24.

In the downlink direction the intention of the indication of the transmission power per PRB is to facilitate interference coordination between different eNodeBs. The eNodeB receiving the UL High Interference Indication information element should avoid scheduling the cell edge users on those PRBs where interference level is high in the neighboring eNodeB. The measurement of the TX power per PRB is illustrated in Figure 6.25.

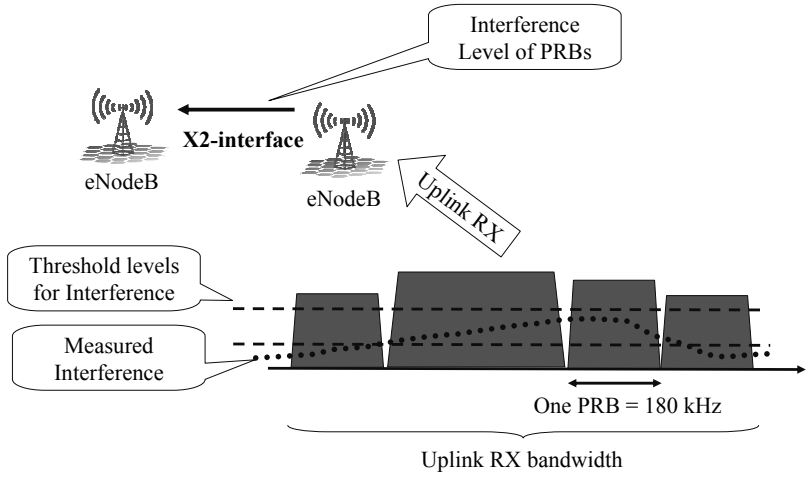


Figure 6.24 Interference level reporting over X2 interface

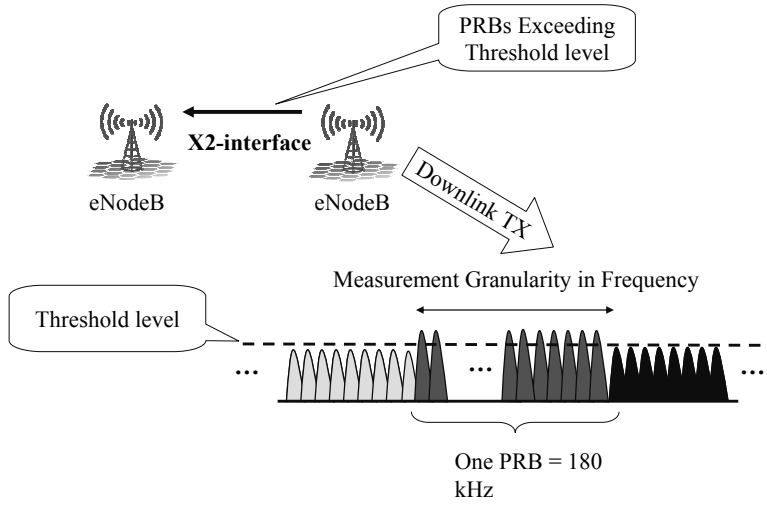


Figure 6.25 Transmission power vs threshold reporting over X2 interface

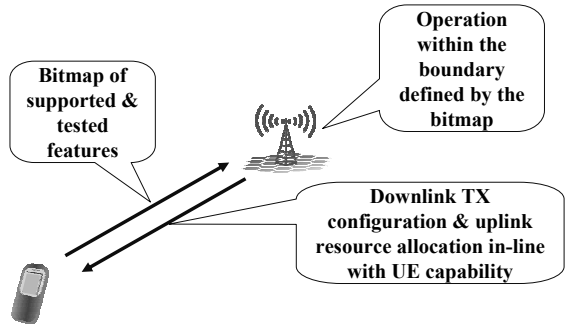


Figure 6.26 Signaling of the device capability at connection setup

6.8 Early UE Handling in LTE

In LTE, the 3GPP has requested feedback from the operators about the planned features for the first phase(s) of LTE deployment and then the first phase test cases will be based on the prioritization received. For features which have a lower priority in the testing area, UE capability signaling will be in place [11,13] to indicate to the network whether the feature is fully implemented and tested or not. This should avoid the problems faced with Release 99 deployment when many features mandatory for the UE were not supported in the networks (and thus not available for testing either) when the first networks were commercially opened. The latest plan for test case prioritization can be found in [14], with the expectation that work in 3GPP TSG RAN WG5 will be completed during 2009 for high priority cases (with some cases potentially still to be dropped from the list because of the high number of high priority test cases). The LTE test cases for signaling can be found from the recently created test specifications [15], and the completion of these test signaling test cases will naturally take longer than the actual radio protocol work in other groups as a stable signaling specification is needed before test cases can be finalized.

6.9 Summary

The LTE radio protocols follow a basic structure similar to WCDMA, but there are obvious differences due to the differences in radio technology. From the functional split the distribution is also partly different with the most notable difference being the handling of ciphering in the PDCP layer. The lack of an Iub-like interface reduces the need for radio protocol signaling on the internal interfaces as the eNodeB deals directly with RRC signaling between the UE and the network. The X2 interface is thus rather simple from the radio protocol perspective, with the key functionalities being to take care of the mobility related signaling for intra-LTE mobility as well as handling information exchange at the network level for cell level interference management. As with the WCDMA specifications, ASN.1 is used as the LTE radio protocol language to ensure straightforward extendibility with releases beyond LTE Release 8. For the UE capability signaling the input from the operator community has shaped what the first phased commercial UE will look like in terms of radio capability, and respectively bitmap is signaled in connection with the connection setup.

References

[1] 3GPP Technical Specification, TS 36.321, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) protocol specification'.
[2] 3GPP Technical Specification, TS 36.322, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Link Control (RLC) protocol specification'.
[3] H. Holma, A. Toskala 'WCDMA for UMTS', 4th edition, Wiley, 2007.
[4] IETF RFC 4995, 'The RObust Header Compression (ROHC) Framework'.
[5] IETF RFC 4996, 'RObust Header Compression (ROHC): A profile for TCP/IP (ROHC-TCP)'.
[6] IETF RFC 3095, 'RObust Header Compression (ROHC): Framework and four profiles: RTP, UDP, ESP and uncompressed'.
[7] IETF RFC 3843, 'RObust Header Compression (ROHC): A compression profile for IP'.
[8] IETF RFC 4815, 'RObust Header Compression (ROHC): Corrections and clarifications to RFC 3095'.
[9] IETF RFC 5225, 'RObust Header Compression (ROHC) Version 2: Profiles for RTP, UDP, IP, ESP and UDP Lite'.
[10] 3GPP Technical Specification, TS 36.323, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Packet Data Convergence Protocol (PDCP) specification'.
[11] 3GPP Technical Specification, TS 36.331, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC) protocol specification'.
[12] 3GPP Technical Specification, TS 36.423, 'X2 application protocol (XSAP)', Version 8.3.0, September 2008.
[13] 3GPP Technical Specification, TS 36.306, 'Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio access capabilities'.
[14] 3GPP Tdoc R5–083674, 'LTE Features and Test Prioritisation' NTT DoCoMo.
[15] 3GPP Technical Specification, TS 36.523–1, 'Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Packet Core User Equipment (UE) conformance specification; Part 1: Protocol conformance specification (Release 8) v1.0.0, September 2008.

7

Mobility

Chris Callender, Harri Holma, Jarkko Koskela and Jussi Reunanen

7.1 Introduction

Mobility offers clear benefits to the end users: low delay services such as voice or real time video connections can be maintained while moving even in high speed trains. Mobility is beneficial also for nomadic services, such as laptop connectivity, since it allows a reliable connection to be maintained in the areas between two cells where the best serving cell is changing. This also implies a simple setup of the broadband connection without any location related configurations. Mobility typically has its price in the network complexity: the network algorithms and the network management get complex. The target of the LTE radio network is to provide seamless mobility while simultaneously keeping network management simple.

The mobility procedures can be divided into idle mode and connected mode for the attached UE (see Figure 7.1). Idle mode mobility is based on UE autonomous cell reselections according to the parameters provided by the network, thus this is quite similar to the one in WCDMA/HSPA today. The connected mode mobility in LTE, on the other hand, is quite different in LTE than in WCDMA/HSPA radio networks. The UE transition between idle and RRC connected mode is controlled by the network according to the UE activity and mobility. The algorithms for RRC connection management are described in Chapter 8.

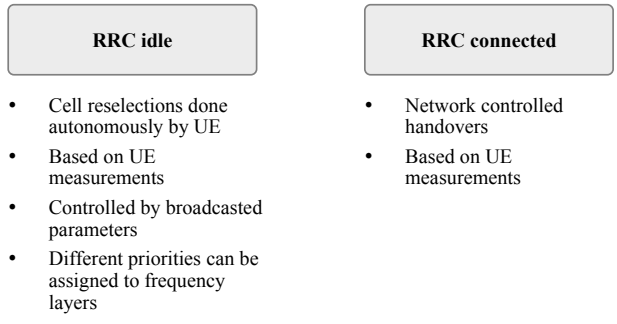


Figure 7.1 Idle mode and RRC connected mode mobility

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6

This chapter explains the mobility principles and procedures in the idle and active state. Idle state mobility management is described in section 7.2. The intra-frequency handover is presented in section 7.3. The inter-RAT handovers are covered in section 7.4. The differences in mobility between UTRAN and E-UTRAN are summarized in section 7.8. The voice handover from E-UTRAN Voice over IP to 2G/3G circuit switched voice is discussed in Chapter 10. The performance requirements related to mobility are described in Chapter 11.

The general description of LTE mobility is presented in [1], idle mode mobility is specified in [2], the performance requirements for radio resource management are defined in [3] and the relevant Radio Resource Control specifications in [4]. The rest of this chapter uses the term E-UTRAN for LTE and UTRAN for WCDMA.

7.2 Mobility Management in Idle State

7.2.1 Overview of Idle Mode Mobility

The UE chooses a suitable cell of the selected Public Land Mobile Network (PLMN) based on radio measurements. This procedure is known as cell selection. The UE starts receiving the broadcast channels of that cell and finds out if the cell is suitable for camping, which requires that the cell is not barred and that radio quality is good enough. After cell selection, the UE must register itself to the network thus promoting the selected PLMN to the registered PLMN. If the UE is able to find a cell that is deemed a better candidate for reselection according to reselection criteria (section 7.2.2), it reselects onto that cell and camps on it and again checks if the cell is suitable for camping. If the cell where the UE camps does not belong to at least one tracking area to which the UE is registered, location registration needs to be performed. An overview is shown in Figure 7.2.

A priority value can be assigned to PLMNs. The UE searches for higher priority PLMNs at regular time intervals and searches for a suitable cell if another PLMN has been selected. For example, the operator may have configured preferred roaming operators to the Universal Subscriber Identity Module (USIM) card. When the UE is roaming and not camped to the preferred operators, it tries periodically to find the preferred operator.

If the UE is unable to find a suitable cell to camp on or if the location registration fails, it attempts to camp on a cell irrespective of the PLMN identity, and enters a 'limited service' state which only allows emergency calls to be made.

The USIM must be inserted in the UE to perform the registration. UTRAN UE can use either the older SIM cards or the new USIM cards, while E-UTRAN uses only USIM. USIM can provide stronger protection against identity theft compared to SIM.

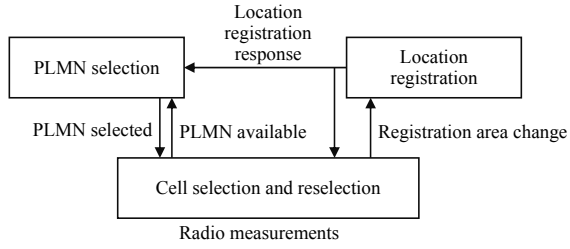


Figure 7.2 Idle mode overview

7.2.2 Cell Selection and Reselection Process

When UE is powered-on for the first time, it will start the Initial Cell Selection procedure. The UE scans all radio frequency (RF) channels in the E-UTRA bands according to its capabilities to find a suitable cell. On each carrier frequency, the UE needs only to search for the strongest cell. Once a suitable cell is found this cell is selected. Initial cell selection is used to ensure that the UE gets into service (or back to the service area) as fast as possible.

The UE may also have stored information about the available carrier frequencies and cells in the neighborhood. This information may be based on the system information or any other information the UE has acquired in the past – 3GPP specifications do not exactly define what kind of information the UE is required or allowed to use for Stored Information cell selection. If the UE does not find a suitable cell based on the stored information, the Initial Cell Selection procedure is started to ensure that a suitable cell is found.

For a cell to be suitable it has to fulfill S-criterion:

$$S_{rxlevel} > 0 \tag{7.1}$$

where

$$S_{rxlevel} > Q_{\text{rxlevelmeas}} - (Q_{rxlev\text{min}} - Q_{rxlevel\text{min}offset}) \tag{7.2}$$

$Q_{rxlevelmeas}$ is the measured cell received level (RSRP), $Q_{rxlevelmin}$ is the minimum required received level [dBm] and $Q_{rxlevelminoffset}$ is used when searching for a higher priority PLMN.

Whenever UE has camped to a cell, it will continue to find a better cell as a candidate for reselection according to the reselection criteria. Intra-frequency cell reselection is based on the cell ranking criterion. To do this the UE needs to measure neighbor cells, which are indicated in the neighbor cell list in the serving cell. The network may also ban the UE from considering some cells for reselection, also known as blacklisting of cells. To limit the need to carry out reselection measurements it has been defined that if $S_{ServingCell}$ is high enough, the UE does not need to make any intra-frequency, inter-frequency or inter-system measurements. The intra-frequency measurements must be started when $S_{ServingCell} \leq S_{intrasearch}$. The inter-frequency measurements must be started when $S_{ServingCell} \leq S_{nonintrasearch.}$.

In case the UE moves fast it is possible for the network to adjust the cell reselection parameters. The high (or medium) mobility state is based on the number of cell reselections, N_{CR}, within a predefined time, T_{RCmax}. High mobility is typically characterized by different parameter values for the hysteresis and for the reselection timer. To avoid adjusting reselection parameters in case the UE is ping-ponging between two cells, these cell reselections are not counted in the mobility state calculations. As the 'speed' estimation is based on the count of reselections it does not give an exact speed, just a rough estimate of UE movement, but nevertheless provides a means for the network to control UE reselection behavior dependent on the UE movement. The speed dependent scaling method is also applied in the RRC_CONNECTED state for connected mode mobility parameters.

7.2.2.1 Intra-frequency and Equal Priority Reselections

Cell ranking is used to find the best cell for UE camping for intra-frequency reselection or on reselection to equal priority E-UTRAN frequency. The ranking is based on the criterion R_s for the serving cell and R_n for neighboring cells:

$$R_s = Q_{meas,s} + Q_{hyst}$$
$$R_n = Q_{meas,n} + Q_{offset} \quad (7.3)$$

where Q_{meas} is the RSRP measurement quantity, Q_{hyst} is the power domain hysteresis to avoid ping-pong and Q_{offset} is an offset value to control different frequency specific characteristics (e.g. propagation properties of different carrier frequencies) or cell specific characteristics. In the time domain, $T_{reselection}$ is used to limit overly frequent reselections. The reselection occurs to the best ranked neighbor cell if it is better ranked than the serving cell for a longer time than $T_{reselection}$. The Q_{hyst} provides hysteresis by requiring any neighbor cell to be better than the serving cell by an RRC configurable amount before reselection can occur. The $Q_{offsets,n}$ and $Q_{offsetfrequency}$ make it possible to bias the reselection toward particular cells and/or frequencies. The cell reselection parameters are illustrated in Figure 7.3.

7.2.2.2 Inter-frequency/RAT reselections

In the UTRAN, inter-frequency and inter-RAT reselections were based on the same ranking as intra-frequency reselections. It was seen that this is very difficult for the network to control as the measurement quantities of different RATs are different and the network needs to be able to control reselection between multiple 3GPP RATs (or even non-3GPP technologies). And as it was seen that operators would like to control how UE prioritizes camping on different RATs or frequencies of E-UTRAN, a new method for reselection handling between different RATs/frequencies (called layers from now on) was chosen. The method is known as absolute priority based reselection. Each layer is given a priority and based on this information the UE tries to camp on the highest priority frequency/RAT if it can provide decent service. In order for the UE to decide if decent service can be provided the network allocates each frequency/RAT a threshold ($\text{Thresh}_{x,\,high}$) which has to be fulfilled before reselection to such a layer is performed. A similar $T_{reselection}$ as in intra-frequency reselections is utilized, i.e. the new layer needs to fulfill the threshold for consecutive time of $T_{reselection}$ before reselection is performed. This is used to eliminate reselections if just temporary fading for the evaluated frequency occurs. To make reselection to a lower priority layer the UE will not reselect to that if the higher priority layer is still above the threshold or if the lower priority frequency is not above another threshold ($\text{Thresh}_{x,\,low}$).

The main parameters for controlling idle mode mobility are shown in Table 7.1.

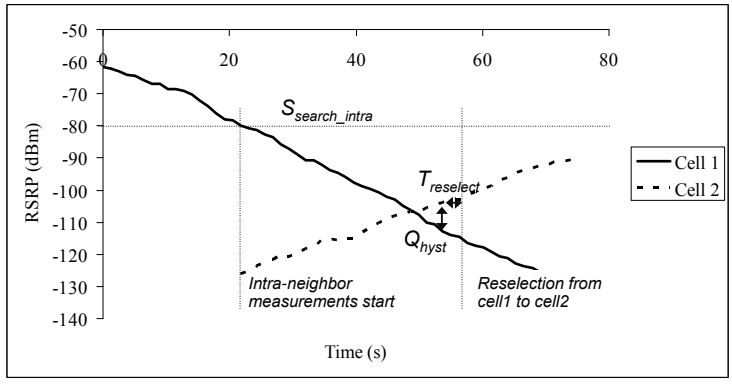


Figure 7.3 Idle mode intra-frequency cell reselection algorithm

Table 7.1 Main parameters for idle mode mobility

Parameters	Description
Q_{hyst}	Specifies the hysteresis value for ranking criteria. The hysteresis is required to avoid ping-ponging between two cells.
$T_{reselection}$	Gives the cell reselection timer value. The cell reselection timer together with the hysteresis is applied to control the unnecessary cell reselections.
$Q_{rxlevmin}$	Specifies the minimum required received level in the cell in dBm. Used in the S-criterion, i.e. in the cell selection process.
$S_{intrasearch}$, $S_{nonintrasearch}$	Specifies the threshold (in dB) for intra-frequency, inter-frequency and inter-RAN measurements. UE is not required to make measurements if $S_{servingcell}$ is above the threshold.
N_{CR}, T_{RCmax}	Specifies when to enter the medium or high mobility state: if the number of cell reselections within time T_{RCmax} is higher than N_{CR}, the UE enters the medium or higher mobility state.
$\text{Thresh}_{x,\,high}$	This specifies the threshold used by the UE when reselecting towards the higher priority frequency X than currently serving frequency.
$\text{Thresh}_{x,\,low}$	This specifies the threshold used in reselection towards frequency X from a higher priority frequency.

7.2.3 Tracking Area Optimization

The UE location in RRC idle is known by the Mobility Management Entity (MME) with the accuracy of the tracking area. The size of the tracking area can be optimized in network planning. A larger tracking area is beneficial to avoid tracking area update signaling. On the other hand, a smaller tracking area is beneficial to reduce the paging signaling load for the incoming packet calls. The corresponding concept in UTRAN is called a Routing area, which typically covers a few hundred base stations. The concept of the tracking area is illustrated in Figure 7.4.

UE can be assigned multiple tracking areas to avoid unnecessary tracking area updates at the tracking area borders, e.g. when the UE is ping-ponging between cells of two different tracking areas. UE can also be assigned both a LTE tracking area and a UTRAN routing area to avoid signaling when changing between the two systems.

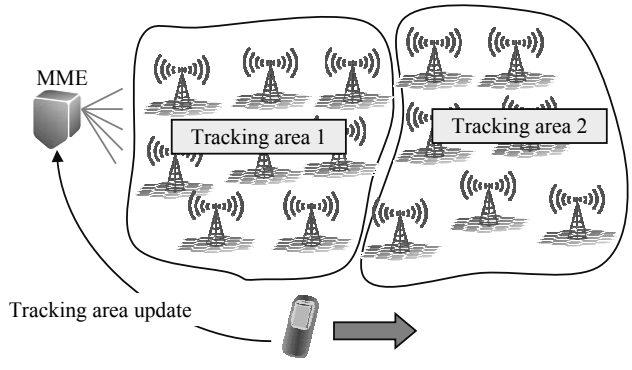


Figure 7.4 Tracking area concept

7.3 Intra-LTE Handovers

UE mobility is solely controlled by the handovers when the RRC connection exists. Thus there is no UTRAN type of state as CELL_PCH where UE based mobility is possible while UE is in RRC_CONNECTED state. The handovers in E-UTRAN are based on the following principles:

1 The handovers are network controlled. E-UTRAN decides when to make the handover and what the target cell is.
2 The handovers are based on the UE measurements. The UE measurements and measurement reporting is controlled by parameters given by E-UTRAN.
3 The handovers in E-UTRAN are targeted to be lossless by using packet forwarding between the source and the target eNodeB.
4 The core network S1 connection is updated only when the radio handover has been completed. This approach is called Late path switch. The core network has no control on the handovers.

The E-UTRAN handover procedure, measurements, signaling and neighborlist control are described below.

7.3.1 Procedure

An overview of the intra-frequency handover procedure is shown in Figure 7.5. The UE is moving from left to right. In the initial phase the UE has user plane connection to the source eNodeB and further to the System Architecture Evolution Gateway (SAE GW). The S1 signaling connection exists between eNodeB and MME. When the target cell fulfills the measurement reporting

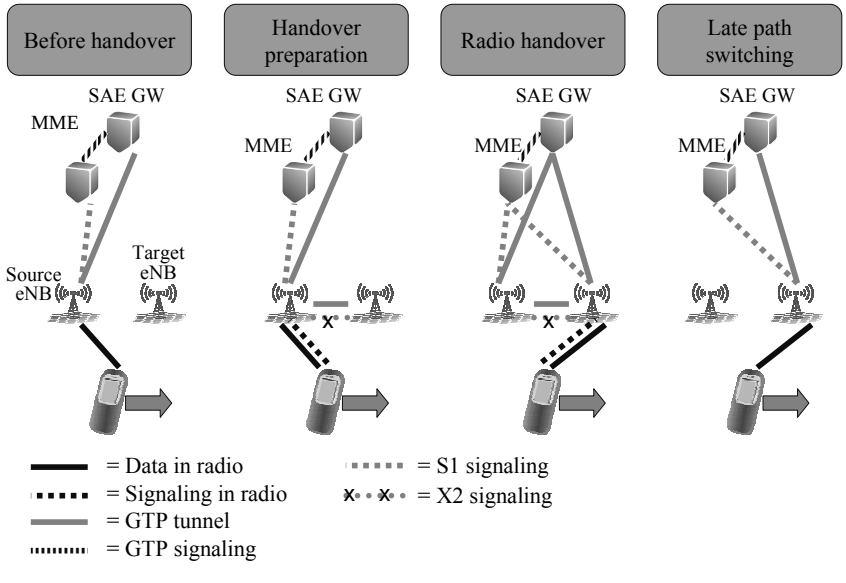


Figure 7.5 Intra-frequency handover procedure

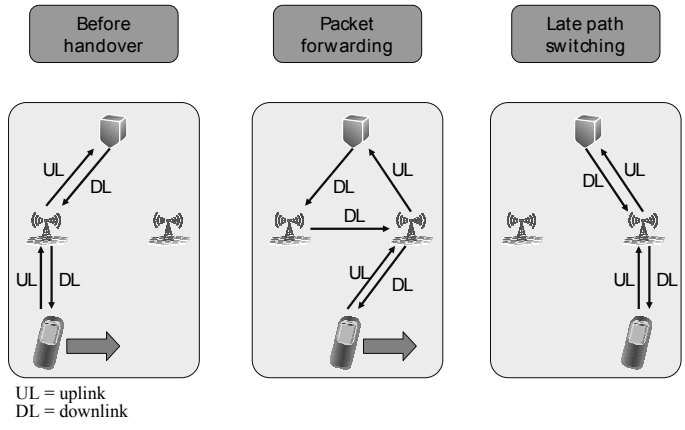


Figure 7.6 User plane switching in handover

threshold, the UE sends the measurement report to the eNodeB. The eNodeB establishes the signaling connection and GTP (GPRS Tunneling Protocol) tunnel to the target cell. Once the target eNodeB has the resource available, the source eNodeB sends the handover command to the UE. The UE can switch the radio connection from the source to the target eNodeB. The core network is not aware of the handover at this point in time. The core network connection is finally updated and this procedure is called Late path switching.

The user plane switching is illustrated in Figure 7.6. The user plane packets in downlink are forwarded from the source eNodeB to the target eNodeB over the X2 interface before Late path switching. The X2 interface enables lossless handover. The source eNodeB forwards all downlink RLC SDUs (Service Data Units) that have not been acknowledged by the UE. The target eNodeB re-transmits and prioritizes all downlink RLC SDUs forwarded by the source eNodeB as soon as it obtains them. It may happen that some packets were already received by the UE from the source eNodeB but the source eNodeB did not receive the acknowledgement. The target eNodeB will then retransmit the packets unnecessarily and the UE receives the same packet twice. Therefore, the UE must be able to identify and remove duplicate packets.

In the uplink direction, the source eNodeB forwards all successfully received uplink RLC SDUs to the packet core. The UE re-transmits the uplink RLC SDUs that have not been acknowledged by the source eNodeB. It may happen also in uplink than some packets are sent twice. The reordering and the duplication avoidance in uplink are done in the packet core.

The packet forwarding takes place only for a short time until the late path switching has been completed. Therefore, the packet forwarding does not consume excessive transport resources and X2 interface requirements are not significant from the network dimensioning point of view. The packets in the uplink direction can be sent directly from the target eNodeB to SAE GW.

7.3.2 Signaling

The detailed signaling messages during the handover procedure are described in this section. The procedure is divided into three parts: handover preparation (Figure 7.7), handover execution (Figure 7.8) and handover completion (Figure 7.9).

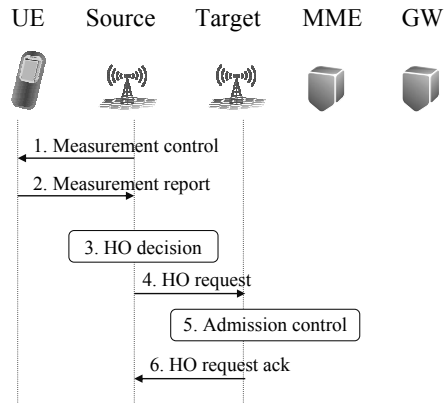


Figure 7.7 Handover preparation

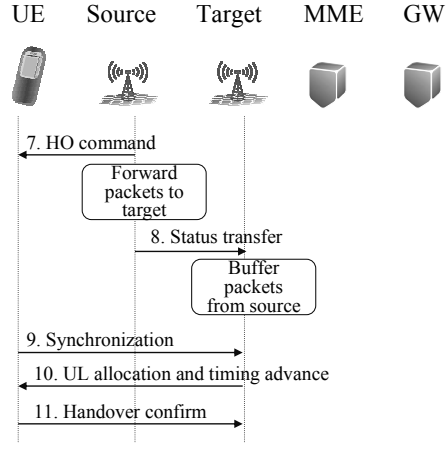


Figure 7.8 Handover execution

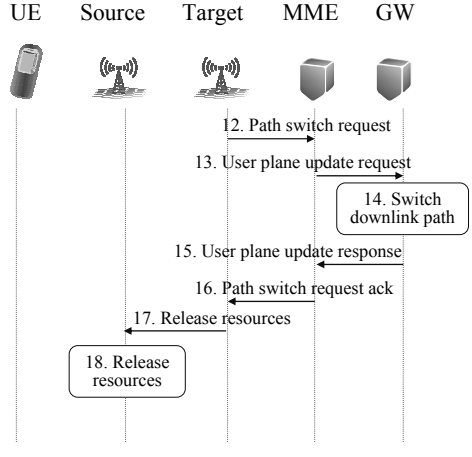


Figure 7.9 Handover completion

1 The source eNodeB configures the UE measurement procedures with a MEASUREMENT CONTROL message. The message defines the measurement reporting thresholds.
2 When the target cell fulfills the reporting threshold, the UE sends a MEASUREMENT REPORT to the eNodeB. The typical reporting is event triggered where the UE sends measurements only when the reporting threshold has been fulfilled. It is also possible to configure periodic reporting.
3 The source eNodeB makes the handover decision based on the UE report. For intra-frequency handovers, the UE needs to be connected to the cell with the lowest path loss and the network has practically no freedom in deciding the handover target. For inter-frequency and inter-RAT handovers, the eNodeB can also take the load and service information into account. The operator may want to balance the loading between frequencies and may want to push certain services to certain frequency layers or systems.
4 The source eNodeB sends a HANDOVER REQUEST to the target eNodeB.
5 The target eNodeB performs the admission control. For intra-frequency handovers the network has little freedom in blocking the new connection since the UE transmission will anyway cause the uplink interference to the target cell even if the UE does not have the connection to the target cell. Actually, the uplink interference can be minimized by allowing the UE to connect to the cell with lowest path loss. If there are simply no resources in the target cell, the network may need to release the connection to avoid excessive interference.
6 The target eNodeB sends the HANDOVER REQUEST ACKNOWLEDGE to the source eNodeB. The target eNodeB is now ready to receive the incoming handover.
7 The source eNodeB sends the HANDOVER COMMAND to the UE. The source eNodeB starts forwarding the downlink packets to the target eNodeB.
8 The source eNodeB sends the status information to the target eNodeB indicating the packets that were acknowledged by the UE. The target eNodeB starts buffering the forwarded packets.
9 The UE makes the final synchronization to target eNodeB and accesses the cell via a RACH procedure. The pre-synchronization is already obtained during the cell identification process.
10 The target eNodeB gives the uplink allocation and timing advance information to the UE.
11 The UE sends HANDOVER CONFIRM to the target eNodeB. The target eNodeB can now begin to send data to the UE.
12 The target eNodeB sends a PATH SWITCH message to the MME to inform it that the UE has changed cell.
13 The MME sends a USER PLANE UPDATE REQUEST message to the Serving Gateway.
14 The Serving Gateway switches the downlink data path to the target eNodeB.
15 The Serving Gateway sends a USER PLANE UPDATE RESPONSE message to the MME.
16 The MME confirms the PATH SWITCH message with the PATH SWITCH ACK message.
17 The target eNodeB sends RELEASE RESOURCE to the source eNodeB, which allows the source eNodeB to release the resources.
18 The source eNodeB can release radio and control plane related resources associated with the UE context.

7.3.3 Handover Measurements

Before the UE is able to send the measurement report, it must identify the target cell. The UE identifies the cell using the synchronization signals (see Chapter 5). The UE measures the signal level using the reference symbols. There is no need for the E-UTRAN UE to read the Broadcast channel during the handover measurements. In UTRAN the UE needs to decode the Broadcast channel to find the system frame number, which is required to time-align the soft handover transmissions in downlink. There is no such requirement in E-UTRAN as there is no soft handover.

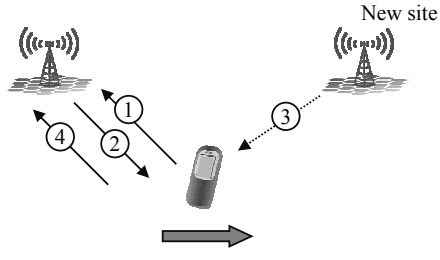
When the reporting threshold condition is fulfilled, the UE sends Handover measurements to the eNodeB

7.3.4 Automatic Neighbor Relations

The neighborlist generation and maintenance has turned out to be heavy work in the existing mobile networks especially when the networks are expanded and new sites are being added. A missing neighbor is a common reason for the call drops. The UE in E-UTRAN can detect the intra-frequency neighbors without neighborlists, leading to simpler network management and better network quality. The automatic neighbor relation functionality is illustrated in Figure 7.10. The UE is moving towards a new cell and identifies the Physical Cell Identity (PCI) based on the synchronization signals. The UE sends a measurement report to eNodeB when the handover reporting threshold has been fulfilled. The eNodeB, however, does not have X2 connection to that cell. The physical cell identity (ID) is not enough to uniquely identify the cell since the number of physical cell IDs is just 504 while the large networks can have tens of thousands of cells. Therefore, the serving eNodeB requests the UE to decode the global cell ID from the broadcast channel of the target cell. The global cell ID identifies the cell uniquely. Based on the global cell ID the serving eNodeB can find the transport layer address of the target cell using the information from the MME and set up a new X2 connection. The serving eNodeB can then proceed with the handover.

New X2 connections need to be created and some old unused connections can be removed when new cells are added to the network or when changes are done to the antenna systems.

The intra-frequency neighborlist generation is simple since the UE can easily identify all the cells within the same frequency. In order to create inter-RAT or inter-frequency neighbors the



(1) = UE reports neighborcell signal including Physical Cell ID
(2) = Request for Global Cell ID reporting
(3) = UE reads Global Cell ID from BCH
(4) = UE reports Global Cell ID

Figure 7.10 Automatic intra-frequency neighbor identification. BCH, broadcast channel

eNodeB must ask the UE to make inter-RAT or inter-frequency measurements. The eNodeB must also schedule gaps in the signal to allow the UE to perform the measurements.

7.3.5 Handover Frequency

The handover frequency in the network depends on a number of factors:

- The actual mobility of the users compared to the cell size. The statistics from HSPA networks can be used to estimate the impact of the mobility.
- The user activity. The typical voice activity is 30–50 mErl, which is 3–5% activity. Laptop users and smart phone users may have a RRC connection active all the time since the always-on applications keep sending packets at frequent intervals. User activity may therefore increase considerably in the future.
- The mobility solution of the system. The total number of handovers will likely be higher in LTE than in HSPA since part of the mobility in HSPA is carried out by the UE itself making cell reselections in Cell_PCH and Cell_FACH states, while LTE always uses handovers.

The impact of user mobility is illustrated by comparing the number of handovers for HSPA users, which are mostly laptop users today, and for WCDMA users, which are mostly voice users. An example of the WCDMA handover frequency is illustrated in Figure 7.11. The handover frequency is collected per cell per hour. The median active set update frequency

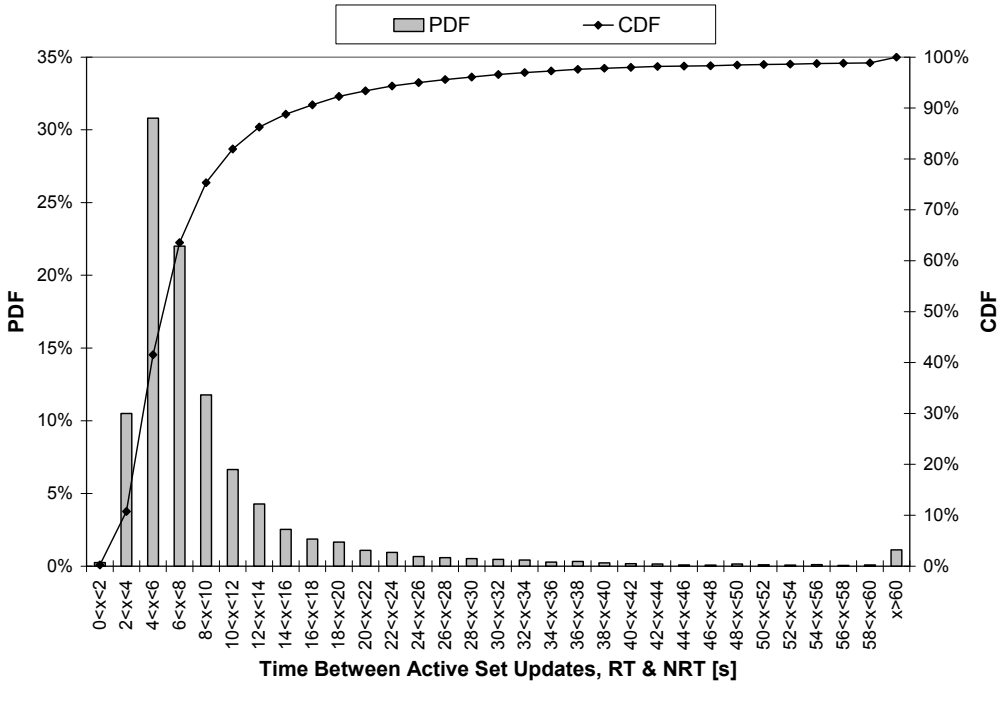


Figure 7.11 Example of WCDMA active set update frequency in 3G networks. PDF, probability density function; CDF, cumulative density function; RT, real time; NRT, non-real time

is 6 s, corresponding to approximately 12 s handover frequency. Each HSPA cell change or LTE handover corresponds typically to two active set updates in WCDMA: add and drop. An example of the HSPA cell change frequency is shown in Figure 7.12. Both Figure 7.11 and Figure 7.12 were collected from the same network at the same time. The median number of HSPA serving cell changes is one serving cell change for every 120–140 s for each active connection. That indicates that current laptop HSPA users have clearly lower mobility than the WCDMA voice users. The HSPA cell change used 2 dB handover hysteresis where the target HSPA serving cell has to have 2 dB higher Ec/No than the current HSPA serving cell.

It is expected that the LTE mobility will initially be similar to HSPA mobility since LTE devices are likely to be USB modems. The introduction of handheld LTE devices will increase the handover rate per connection.

Figure 7.11 and Figure 7.12 show that the handover frequency varies considerably depending the environment, cell size and user mobility. For 20% of the cells the HSPA service cell change rate is below 40 s. There are, however, 15% of cells where the cell change rate is more than 500 s.

The most typical application today is voice with a normal activity of 3–5%. The application activity will increase with always on applications. Laptops keep sending small packets in the background implying that the RRC connection needs to be maintained 100% of the time. The same situation applies to the handheld devices that run always on applications such as push-email, push-to-talk or presence. Even if the end user does not use the application actively, the RRC connection is still active from the system perspective due to the transmission of the

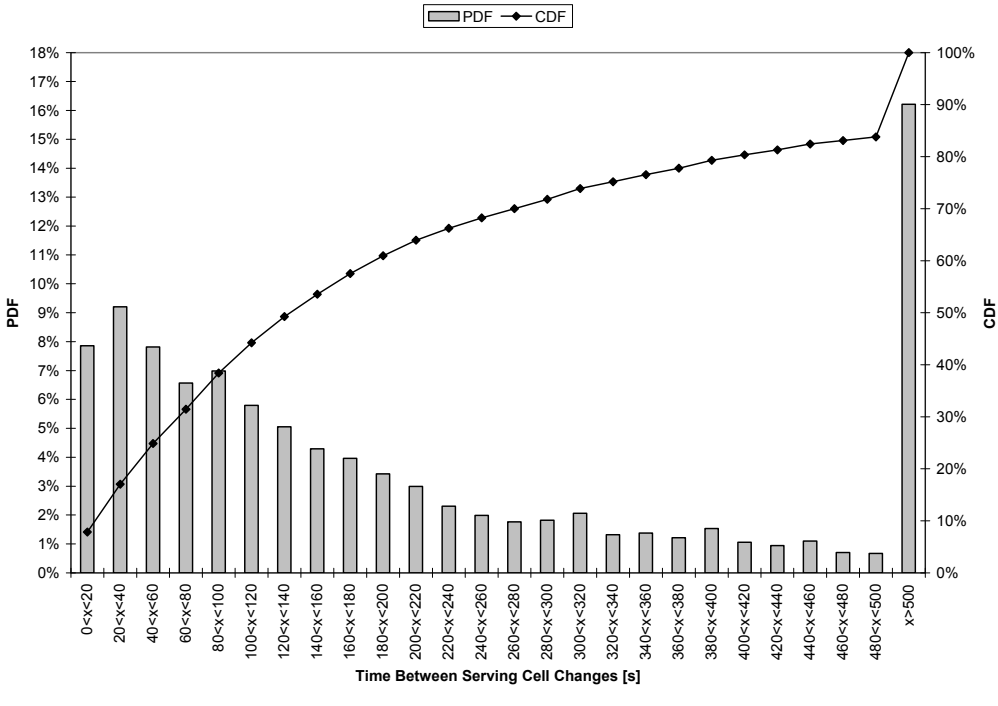


Figure 7.12 Example of HSPA Serving Cell Change frequency in 3G networks

keep alive messages, reception of emails and presence updates. The user activity may increase from 5% to 100% due to the new applications – that is a 20 times higher activity than today. Therefore, the number of handovers may be considerably higher with new applications in the LTE network where all the mobility is carried by handovers for RRC connected users.

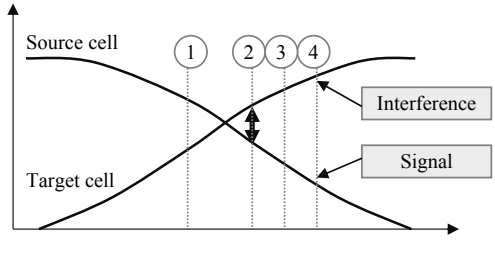
7.3.6 Handover Delay

The fast handover process is essential when the signal level from the source cell fades fast while the signal from the target cell increases. This may happen when the UE drives around the corner. The signal levels are illustrated in Figure 7.13. The critical point is that the UE must be able to receive the handover command from the source eNodeB before the signal-to-interference ratio gets too low. The handover reliability in this scenario can be improved by suitable window and averaging parameters as well as by minimizing the network delay in responding to the UE measurement report.

7.4 Inter-system Handovers

The inter-RAT handovers refer to the inter-system handovers between E-UTRAN and GERAN, UTRAN or cdma2000® both for real time and for non-real time services. The inter-RAT handover is controlled by the source access system for starting the measurements and deciding the handover execution. The inter-RAT handover is a backwards handover where the radio resources are reserved in the target system before the handover command is issued to the UE. As the GERAN system does not support Packet Switched Handover (PS HO), the resources are not reserved before the handover. The signaling is carried via the core network as there are no direct interfaces between the different radio access systems. The inter-RAT handover is similar to intra-LTE handover in the case where the packet core node is changed.

All the information from the target system is transported to the UE via the source system transparently. The user data can be forwarded from the source to the target system to avoid loss of user data. To speed up the handover procedure there is no need for the UE to have any



1 = UE identifies the target cell
2 = Reporting range fulfilled
3 = After UE has averaged the measurement, it sends measurement report to source eNodeB
4 = Source eNodeB sends handover command to the UE

Figure 7.13 Handover timings

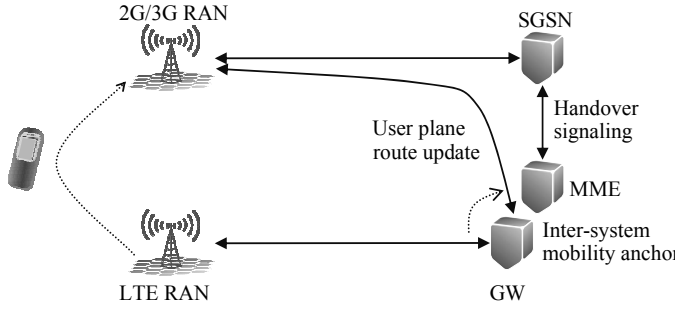


Figure 7.14 Overview of inter-RAT handover from E-UTRAN to UTRAN/GREAN

signaling to the core network. The security and QoS context is transferred from the source to the target system.

The Serving GW can be used as the mobility anchor for the inter-system handovers.

The overview of the inter-system handover is shown in Figure 7.14. The interfaces are described in Chapter 3.

The interruption time in inter-system handovers can be very small. The interruption time from the UE point of view is defined as the time between the last TTI containing a transport block on the E-UTRAN side and the time the UE starts transmission of the new uplink DPCCH on the UTRAN side. This interruption time can be as low as 50 ms plus the frame alignments. When the UE enters UTRAN, it will initially be connected to one cell and the soft handover links can be added separately. The UE cannot go directly into soft handover in UTRAN since the UE is not required to read the system frame number (SFN) from UTRAN before the inter-system handover.

E-UTRAN also supports mobility to and from non-3GPP radio systems, such as cdma2000®.

7.5 Differences in E-UTRAN and UTRAN Mobility

The main differences between UTRAN and E-UTRAN mobility are summarized below and listed in Table 7.2. The idle mode mobility is similar in UTRAN and E-UTRAN for intra-frequency reselections. But a new priority based reselection method was chosen in E-UTRAN to ease network planning for the presence of multiple different RATs. The flat architecture in E-UTRAN brings some differences to the active mode mobility.

In UTRAN the UE must update the location both to the circuit switched core network (Location areas) and to the packet core network (Routing areas) while E-UTRAN only uses Tracking areas (packet core). For so-called Circuit Switched Fall Back (CSFB) handover, the CS core network can send a paging message to the E-UTRAN UE. In that case the MME maps the Tracking area to the Location area. This procedure is explained in Chapter 10.

E-UTRAN uses handovers for all UEs that have RRC connection. The UE in UTRAN makes cell reselections without network control even if a RRC connection exists when the UE is in Cell_FACH, Cell_PCH or in URA_PCH states. Such states are not used in E-UTRAN. The UE location is known with UTRAN Registration area accuracy in URA_PCH state while the UE location is known with the accuracy of one cell in E-UTRAN whenever the RRC connection exists.

Table 7.2 UTRAN and E-UTRAN differences in mobility

UTRAN	E-UTRAN	Notes
Location area (CS core)	Not relevant since no CS connections	For CS fallback handover, MME maps Tracking area to Location area
Routing area	Tracking area	
Soft handover is used for WCDMA uplink and downlink and for HSUPA uplink	No soft handovers	
Cell_FACH, Cell_PCH, URA_PCH	No similar RRC states	E-UTRAN always uses handovers for RRC connected users
RNC hides most of mobility	Core network sees every handover	Flat architecture HSPA is similar to E-UTRAN
Neighbor cell lists required	No need to provide cell specific information, i.e. only carrier frequency is required, but the network can provide cell specific reselection parameters (e.g. Q_{offset}) if desired	Also UTRAN UE can use detected cell reporting to identify a cell outside the neighborlist

The WCDMA system uses soft handovers both in uplink and in downlink. The HSPA system uses soft handovers in uplink, but not in downlink. In the LTE system a need for soft handovers was no longer seen, thus simplifying the system significantly.

The Radio Network Controller (RNC) typically connects hundreds of base stations and hides the UE mobility from the core network. The packet core can only see Serving RNC (S-RNC) relocations when the SRNC is changed for the UE. For E-UTRAN, the packet core network can see every handover. 3GPP has specified the flat architecture option for UTRAN in Release 7. That architecture is known as Internet-HSPA. The RNC functionalities are embedded to the Node-B. Internet-HSPA mobility is similar to E-UTRAN mobility: serving RNC relocation over the Iur interface is used in the same way as E-UTRAN handover over the X2 interface.

7.6 Summary

3GPP E-UTRAN as well as the SAE standards are designed to support smooth mobility within E-UTRAN (LTE) and between E-UTRAN and GERAN/UTRAN (2G/3G RAN) and to cdma2000®. The idle mode mobility follows the same principles as in 2G/3G networks where the UE makes cell reselection autonomously. E-UTRAN adds the flexibility that the UE can be allocated multiple tracking areas. The network knows the UE location with the accuracy of the tracking area.

The handover is used in E-UTRAN RRC connected mode and is controlled by the network. The intra-frequency handover is based on fast hard handover in the radio followed by the core network path switching. The E-UTRAN flat architecture implies that the core network sees every handover while in UTRAN system most of the handovers are hidden by RNC. Additionally, the number of handovers in E-UTRAN will likely increase compared to UTRAN since the mobility is controlled by handovers for all RRC connected users.

The E-UTRAN system can be operated without pre-defined neighbor cell lists, which makes the network operation simpler than in GERAN or UTRAN networks.

References

[1] 3GPP Technical Specification 36.300 'E-UTRAN Overall Description; Stage 2', v. 8.6.0.
[2] 3GPP Technical Specification 36.304 'User Equipment (UE) procedures in idle mode', v. 8.3.0.
[3] 3GPP Technical Specification 36.133 'Requirements for support of radio resource management', v. 8.3.0.
[4] 3GPP Technical Specification 36.331 'Radio Resource Control (RRC); Protocol specification', v. 8.3.0.

8

Radio Resource Management

Harri Holma, Troels Kolding, Daniela Laselva, Klaus Pedersen, Claudio Rosa and Ingo Viering

8.1 Introduction

The role of Radio Resource Management (RRM) is to ensure that the radio resources are efficiently used, taking advantage of the available adaptation techniques, and to serve the users according to their configured Quality of Service (QoS) parameters. An overview of primary RRM functionalities is given in section 8.2, and the chapter then moves on to discuss QoS parameters and admission control issues in section 8.3. Next, details related to both downlink and uplink adaptation and scheduling methodologies are discussed in sections 8.4 and 8.5. Interference coordination is discussed in section 8.6, and a discussion of discontinuous reception is included in section 8.7. The Radio Resource Control (RRC) connection management is introduced in section 8.8, and the chapter is summarized in section 8.9.

8.2 Overview of RRM Algorithms

Figure 8.1 shows an overview of the user-plane and control-plane protocol stack at the eNodeB, as well as the corresponding mapping of the primary RRM related algorithms to the different layers. The family of RRM algorithms at the eNodeB exploits various functionalities from Layer 1 to Layer 3 as illustrated in Figure 8.1. The RRM functions at Layer 3, such as QoS management, admission control, and semi-persistent scheduling, are characterized as semi-dynamic mechanisms, since they are mainly executed during setup of new data flows. The RRM algorithms at Layer 1 and Layer 2, such as Hybrid Adaptive Repeat and Request (HARQ) management, dynamic packet scheduling, and link adaptation, are highly dynamic functions with new actions conducted every Transmission Time Interval (TTI) of 1 ms. The latter RRM functions are therefore characterized as *fast dynamic*.

The Channel Quality Indicator (CQI) manager at Layer 1 processes the received CQI reports (downlink) and Sounding Reference Signals (SRSs) (uplink) from active users in the cell. Each received CQI report and SRS is used by the eNodeB for scheduling decisions and

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6

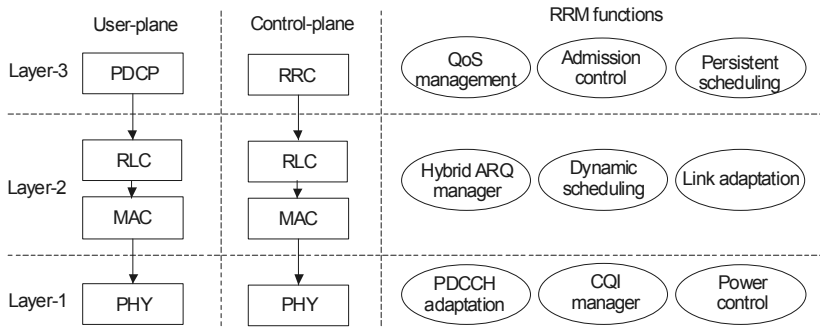


Figure 8.1 Overview of the eNodeB user plane and control plane protocol architecture, and the mapping of the primary RRM functionalities to the different layers. PHY = Physical layer; MAC = Medium access control; RLC = Radio link control; PDCP = Packet data convergence protocol; PDCCH = Physical downlink control channel

for link adaptation purposes in downlink and uplink, respectively. This chapter presents the Layer 3 and Layer 2 RRM functions except for the semi-persistent scheduling, which is part of the voice description in Chapter 10 as semi-persistent scheduling is typically used for voice service. The Layer 1 functions are covered in Chapter 5.

3GPP specifies the RRM related signaling but the actual RRM algorithms in the network are not defined in 3GPP – those algorithms can be vendor and operator dependent.

8.3 Admission Control and QoS Parameters

The eNodeB admission control algorithm decides whether the requests for new Evolved Packet System (EPS) bearers in the cell are granted or rejected. Admission control (AC) takes into account the resource situation in the cell, the QoS requirements for the new EPS bearer, as well as the priority levels, and the currently provided QoS to the active sessions in the cell. A new request is only granted if it is estimated that QoS for the new EPS bearer can be fulfilled, while still being able to provide acceptable service to the existing in-progress sessions in the cell having the same or higher priority. Thus, the admission control algorithm aims at only admitting new EPS bearers up to the point where the packet scheduler in the cell can converge to a feasible solution where the promised QoS requirements are fulfilled for at least all the bearers with high priority. The exact decision rules and algorithms for admission control are eNodeB vendor specific and are not specified by 3GPP. As an example, possible vendor specific admission control algorithms for OFDMA based systems are addressed in [1]. Similarly, the QoS-aware admission control algorithms in [2] and [3] can be extended to LTE.

Each LTE EPS bearer has a set of associated QoS parameters in the same way as GERAN and UTRAN radios. All the packets within the bearer have the same QoS treatment. It is possible to modify QoS parameters of the existing bearers dynamically. It is also possible to activate another parallel bearer to allow different QoS profiles for different services simultaneously. The new bearer can be initiated by the UE or by the packet core network.

The QoS profile of the EPS bearer consists of the following related parameters [4]:

- allocation retention priority (ARP)

- uplink and downlink guaranteed bit rate (GBR)
- QoS class identifier (QCI).

The GBR parameter is only specified for EPS GBR bearers. For non-GBR bearers, an aggregate MBR (AMBR) is specified. The ARP parameter, being an integer in the range 1–16, is primarily for prioritization when conducting admission control decisions. The QCI is a pointer to a more detailed set of QoS attributes. The QCI includes parameters like the Layer 2 packet delay budget, packet loss rate and scheduling priority. These parameters can be used by the eNodeB to configure the outer ARQ operation point for the RLC protocol, and the Layer 2 packet delay budget is used by the eNodeB packet scheduler, i.e. to prioritize certain queues in order to fulfill certain head-of-line packet delay targets.

3GPP specifications define a mapping table for nine different QCIs and their typical services, see Table 8.1. Further QCI values may be defined later depending on the need and on the new emerging services.

An additional QoS parameter called the prioritized bit rate (PBR) is specified for the uplink per bearer. The PBR is introduced to avoid the so-called uplink scheduling starvation problem that may occur for UE with multiple bearers. A simple rate control functionality per bearer is therefore introduced for sharing of uplink resources between radio bearers. RRC controls the uplink rate control function by giving each bearer a priority and a PBR. The PBR is not necessarily related to the GBR parameter signaled via S1 to the eNodeB, i.e. a PBR can also be defined for non-GBR bearers. The uplink rate control function ensures that the UE serves its radio bearer(s) in the following sequence [6]:

- all the radio bearer(s) in decreasing priority order up to their PBR;
- all the radio bearer(s) in decreasing priority order for the remaining resources.

When the PBR is set to zero for all RBs, the first step is skipped and the radio bearer(s) are served in strict priority order.

The QoS concept in LTE has been simplified compared to WCDMA/HSPA in which more than ten different QoS parameters are signaled over the Iu interface. LTE also enables a network activated GBR bearer without the need for the terminal application to initiate the request. There

Table 8.1 QCI characteristics for the EPS bearer QoS profile [4] [5]

QCI #	Priority	L2 packet delay budget	L2 packet loss rate	Example services
1 (GBR)	2	100 ms	10^{-2}	Conversational voice
2 (GBR)	4	150 ms	10^{-3}	Conversational video
3 (GBR)	5	300 ms	10^{-6}	Buffered streaming
4 (GBR)	3	50 ms	10^{-3}	Real-time gaming
5 (non-GBR)	1	100 ms	10^{-6}	IMS signaling
6 (non-GBR)	7	100 ms	10^{-3}	Live streaming
7 (non-GBR)	6	300 ms	10^{-6}	Buffered streaming, email, browsing, file download, file sharing, etc.
8 (non-GBR)	8	300 ms	10^{-6}	
9 (non-GBR)	9	300 ms	10^{-6}	

is, however, also more need for well performing QoS in LTE since circuit switched connections are not possible and voice is carried as Voice over IP (VoIP).

8.4 Downlink Dynamic Scheduling and Link Adaptation

Dynamic packet scheduling and link adaptation are key features to ensure high spectral efficiency while providing the required QoS in the cell. In this section, the general framework as well as specific algorithms and applications are presented.

8.4.1 Layer 2 Scheduling and Link Adaptation Framework

The controlling RRM entity at Layer 2 is the dynamic packet scheduler (PS), which performs scheduling decisions every TTI by allocating Physical Resource Blocks (PRBs) to the users, as well as transmission parameters including modulation and coding scheme. The latter is referred to as link adaptation. The allocated PRBs and selected modulation and coding scheme are signaled to the scheduled users on the PDCCH. The overall packet scheduling goal is to maximize the cell capacity, while making sure that the minimum QoS requirements for the EPS bearers are fulfilled and there are adequate resources also for best-effort bearers with no strict QoS requirements. The scheduling decisions are carried out on a per user basis even though a user has several data flows. Actually, each active user with an EPS bearer has at least two Layer 2 data flows; namely a control plane data flow for the RRC protocol and one or multiple user plane data flows for EPS bearers. Each of the data flows are uniquely identified with a 5-bit Logical Channel Identification (LCID) field. Given the scheduled Transport Block Size (TBS) for a particular user, the MAC protocol decides how many data are sent from each LCID.

As illustrated in Figure 8.2, the packet scheduler interacts closely with the HARQ manager as it is responsible for scheduling retransmissions. For downlink, asynchronous adaptive HARQ is supported with 8 stop-and-wait channels for each code-word, meaning that the scheduler has full flexibility to dynamically schedule pending HARQ retransmissions in the time and the frequency

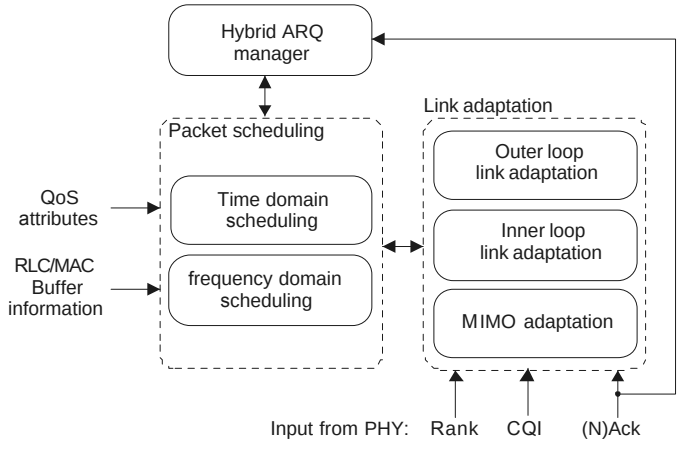


Figure 8.2 Layer 2 functionalities for dynamic packet scheduling, link adaptation, and HARQ management

domain. For each TTI, the packet scheduler must decide between sending a new transmission or a pending HARQ transmission to each scheduled user, i.e. a combination is not allowed. The link adaptation provides information to the packet scheduler of the supported modulation and coding scheme for a user depending on the selected set of PRBs. The link adaptation unit primarily bases its decisions on CQI feedback from the users in the cell. An outer loop link adaptation unit may also be applied to control the block error rate of first transmissions, based on, for example, HARQ acknowledgements (positive or negative) from past transmissions [7]. A similar outer loop link adaptation algorithm is widely used for HSDPA [8].

8.4.2 Frequency Domain Packet Scheduling

Frequency Domain Packet Scheduling (FDPS) is a powerful technique for improving the LTE system capacity. The basic principle of FDPS is illustrated in Figure 8.3. The FDPS principle exploits frequency selective power variations on either the desired signal (frequency selective fading) or the interference (fading or due to fractional other cell load) by only scheduling users on the PRBs with high channel quality, while avoiding the PRBs where a user experiences deep fades. A condition for achieving high FDPS gains is therefore that the radio channel's effective coherence bandwidth is less than the system bandwidth, which is typically the case for cellular macro and micro cell deployments with system bandwidths equal to or larger than 5 MHz.

As an example, Figure 8.4 shows the performance results from extensive macro cellular system level simulations versus the UE velocity. These results are obtained for a 10 MHz system

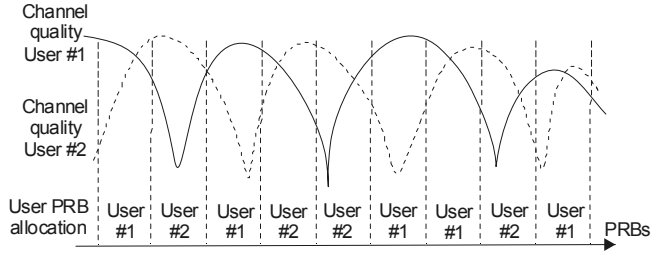


Figure 8.3 Frequency domain scheduling principle

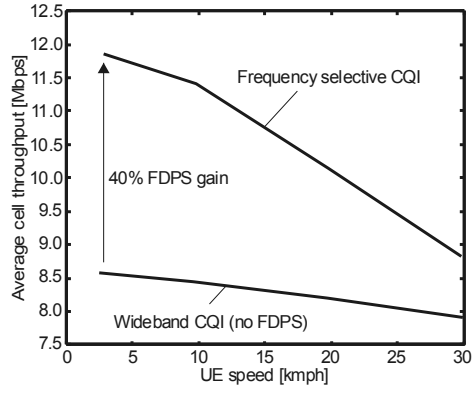


Figure 8.4 Capacity gain from Frequency Domain Packet Scheduling (FDPS)

bandwidth, a configuration with one transmit antenna and dual antenna UEs with interference rejection combining, Poisson arrival best effort traffic, and other assumptions in line with the 3GPP agreed simulation assumptions. The average cell throughput is reported. The results for simple wideband CQI correspond to cases with no FDPS, while the results with frequency selective CQI reporting are obtained with proportional FDPS. It is observed that a significant FDPS gain of approximately 40% is achievable for low to moderate UE speeds, while the FDPS gain decreases at higher UE speeds. The latter is observed because the radio channel cannot be tracked accurately due to the delay on uplink CQI reporting.

The time domain scheduling also can provide multi-user diversity gains when there is frequency flat fast fading. The gains depend on the amount of fading and on the speed of the fading. When there are deep fades, it gives more freedom for the scheduler to select optimally the user for the transmission. When the mobile speed is low enough, the scheduling is able to follow the fast fading. Therefore, the time domain scheduling gains generally get lower with:

- mobile antenna diversity since it reduces fading;
- base station transmit antenna diversity since it reduces fading;
- large bandwidth since it reduces fading due to frequency diversity;
- high mobile speed;
- multi-path propagation since it reduces fading due to multi-path diversity.

The time domain scheduling gains in LTE are relatively low since antenna diversity is a standard feature in the LTE terminal and also LTE uses typically large bandwidths.

Even though the spectral efficiency is maximized by operating with full frequency reuse [9], the system may start to operate in fractional load mode if there are only small data amounts in the eNodeB for the users in the cell. If such situations occur, the buffered data in the eNodeB are only transmitted on the required number of PRBs, while the rest of the PRBs are muted (i.e. no transmission). Under such fractional load conditions, the packet scheduler still aims at selecting the PRBs with the highest channel quality based on the CQI feedback. As illustrated in Figure 8.5 with a simple two-cell example, this allows convergence to a PRB allocation strategy between the two cells, where they aim at allocating complementary sets of PRBs to minimize the interference between the cells. As an example, it was shown in [10] that the average experienced Signal-to-Interference Noise Ratio (SINR) was improved by 10 dB compared to the full load case by operating a 3-sector network at 25% fractional load by using CQI assisted scheduling. The 10 dB improvement should be compared to the expected 6 dB improvement from a 25% fractional load (i.e. factor of 4 reduced load corresponds to 6 dB) with blind scheduling. The

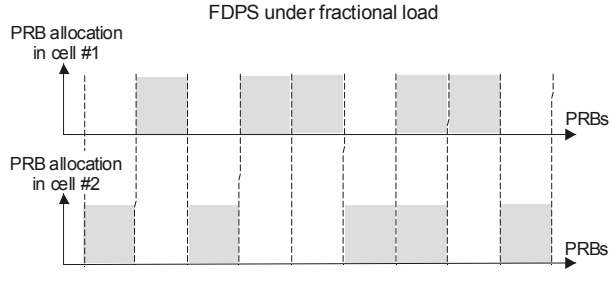


Figure 8.5 Frequency domain scheduling principle under fractional load conditions

CQI assisted scheduling scheme under fractional load resembles the well-known carrier sense allocation scheme from the Ethernet protocol. No explicit load information was shared over the X2 interface in this case.

8.4.3 Combined Time and Frequency Domain Scheduling Algorithms

Figure 8.6 illustrates the three-step packet scheduling implementation as proposed in [7]. The first step consists of the time-domain (TD) scheduling algorithm, which selects up to *N* users for potential scheduling during the next TTI. The time domain scheduler only selects users with (i) pending data for transmission and (ii) users configured for scheduling in the following TTI, i.e. excluding users in Discontinuous Reception (DRX) mode. The time domain scheduler assigns a priority metric for each of the *N* selected users. Secondly, a control channel schedule check is performed to evaluate if there is sufficient control channel capacity to schedule all the users selected by the time domain scheduler. This implies evaluating if there are enough transmission resources within the first 3 OFDM symbols to have reliable transmission of the PDCCH for each of the selected users. If the latter is not the case, then only a sub-set of the *N* users are passed to the frequency domain (FD) scheduler. In cases with control channel congestion, the users with the lowest priority metric from the time domain scheduler are blocked. Users with pending HARQ retransmissions can be given priority. Finally, the frequency domain scheduler decides how to schedule the remaining users across the available PRBs. Compared to a fully joint time/frequency domain packet scheduling method, this method provides similar performance at a much lower complexity as the frequency domain scheduler only has to consider a sub-set of the users for multiplexing on the PRBs.

Assuming that the number of users scheduled per TTI is smaller than the number of schedulable users per cell, the time domain scheduler becomes the primary instrument for QoS control, while the frequency domain scheduler task is to benefit from radio channel aware multi-user frequency diversity scheduling. This implies that the overall scheduling QoS concept from HSDPA (see e.g. [11] and [12]) can be generalized to also apply for the LTE time domain scheduler part for most cases. Thus, the barrier function based scheduler in [13], the QoS-aware schedulers with required activity detection in [14] and [15], and the generalized delay aware schedulers in [16] and [17] are all relevant for the LTE time domain scheduler. Similarly, the well-known proportional fair scheduler (as previously studied in [18] and [19], among others) can be extended for OFDMA based systems as discussed in [20], [21] and [22].

The frequency domain scheduler allocates PRBs to the users selected by the time domain scheduler. The overall objective is to allocate the available PRBs among the selected users to maximize the benefit from multi-user FDPS, while still guaranteeing a certain fairness. The frequency domain scheduler is responsible for allocating PRBs to users with new transmissions and also to users with pending HARQ retransmissions. Note, however, that it is not possible

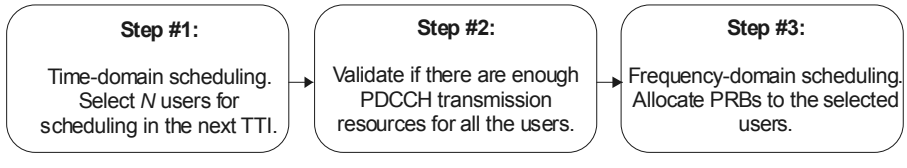


Figure 8.6 Illustration of the three-step packet scheduling algorithm framework

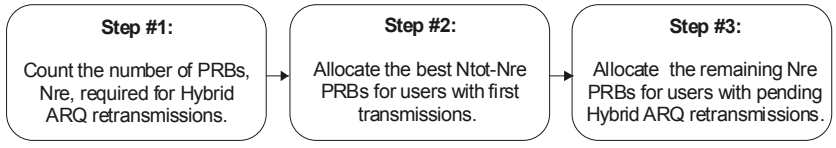


Figure 8.7 HARQ aware frequency domain packet scheduling

to simultaneously schedule new data and pending HARQ retransmissions to the same user in the same TTI. Pending HARQ retransmissions will typically be scheduled on the same number of PRBs as the original transmission. As scheduling of HARQ retransmissions creates a combining gain it is typically not critical to have those retransmissions allocated on the best PRBs. The latter argument has led to the proposed frequency domain scheduling framework summarized in Figure 8.7 [7]. In Step #1 the frequency domain scheduler computes the number of required PRBs (denoted by Nre) for scheduling of pending HARQ retransmissions for the selected users by the time domain scheduler. Assuming Ntot available PRBs for PDSCH transmission, users with new data are allocated the best Ntot-Nre PRBs in Step #2. Finally, the remaining Nre PRBs are allocated for the users with pending HARQ retransmissions. This framework has the advantage that HARQ retransmissions are given priority (i.e. reducing HARQ delays) while maximizing the channel capacity by allocating the best PRBs for users with new data transmissions. The latter is an attractive solution for scenarios where the experienced channel quality is less critical for HARQ retransmissions due to the benefits from the HARQ combining gain. For highly coverage limited scenarios where it is important to maximize the channel quality for second transmissions, however, Steps #2 and #3 in Figure 8.7 can be exchanged. The exact algorithms for allocating PRBs to the considered users in Step #2 and Step #3 can be a simple diversity scheduler or a modified version of proportional fair or maximum throughput schedulers as discussed in [20], [7] and [22], among others. The full benefit from FPDS requires that the considered users are configured to report frequency selective CQI reports such as full sub-band reporting or average best-M (see more details in Chapter 5 for the CQI schemes).

8.4.4 Packet Scheduling with MIMO

Multi-antenna transmission, Multiple Input Multiple Output (MIMO) with two transmit and two receive antennas can use either one transport block or two independently coded transport blocks to the user on virtual streams on the allocated PRBs. The coded transport blocks are also called code-words. The transmission of two independently coded transport blocks requires that the Rank of the MIMO radio channel is at least two. In the latter case, the CQI feedback from the user will include information that allows the eNodeB link adaptation unit to potentially select different modulation and coding schemes for the two independent transport blocks. Separate acknowledgements are transmitted for each code-word such that link adaptation and HARQ for the two streams are decoupled. The eNodeB can dynamically switch between transmission of one and two code-words depending on the CQI and Rank Indicator feedback from the scheduled user. As a simple rule of thumb, the average SINR has to be in excess of approximately 12 dB before transmission of two independent transport blocks becomes attractive on a 2×2 MIMO link with Rank 2. From a packet scheduling point of view, the single-user MIMO

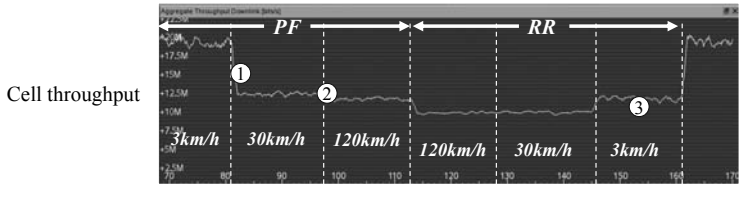
functionality is partially hidden as part of the link adaptation unit although supported TBS for the two streams is known by the packet scheduler. The different QoS-aware packet scheduling algorithms are equally applicable independent of whether advanced MIMO is being used or not. General discussions of MIMO schemes for OFDMA are available in [23].

LTE also supports multi-user MIMO transmission, where two users are scheduled on the exact same PRB(s), sending one independently coded transport block to each of the users by applying an appropriate set of antenna transmit weights. The two antenna transmission weights have to form orthogonal vectors while still being matched to achieve good received SINR for both users. To make multi-user MIMO attractive, i.e. maximizing the sum throughput served to the two users, the choice of the antenna transmission weights has to be accurate in both time and frequency domains. The latter also calls for coordinated packet scheduling actions, where users paired for transmission on the same PRB(s) are dynamically selected dependent on their joint short term radio channel characteristics. Hence, the usage of multi-user MIMO schemes also adds complexity to the packet scheduler and requires more control signaling (PDCCH).

8.4.5 Downlink Packet Scheduling Illustrations

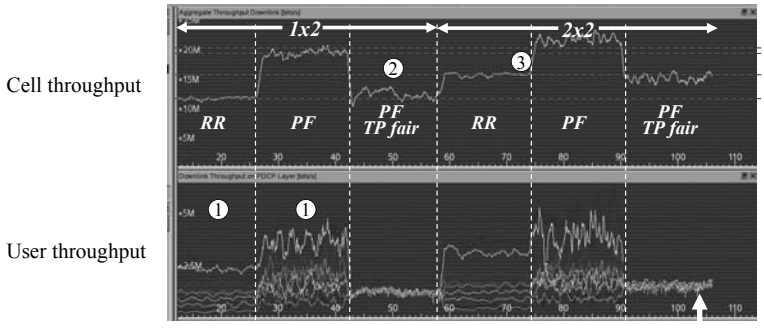
This section presents performance results of LTE downlink packet scheduling. We first illustrate the impact of basic packet scheduling, mobile speed and antenna configuration on the cell and user throughput using a network emulation tool. Later in the section, we evaluate the QoS impact on the cell throughput and cell capacity by conducting dynamic system level simulations with multiple cells and multiple QoS UE classes.

Figure 8.8 and Figure 8.9 illustrate the impact of the packet scheduling, mobile speed and antenna configuration to the cell and user throughput. The plots are taken from an LTE network emulation tool. Figure 8.8 shows the cell throughput with proportional fair scheduler and with round robin scheduler with 3, 30 and 120 km/h. The highest cell throughput is obtained with proportional fair scheduler at 3 km/h. The throughput drops considerably when the mobile speed increases to 30 km/h because the feedback is not able to follow the fast fading, and the frequency domain scheduling gain is lost. Increasing the mobile speed further to 120 km/h has only a marginal impact on the capacity. The round robin scheduler has a lower cell throughput and less sensitivity to the mobile speed. There is some capacity increase with the round robin scheduler at 3 km/h compared to 30 and 120 km/h because the link adaptation works more



① = Significant impact to cell throughput when mobile speed increases from 3 to 30 km/h because the scheduling gain is reduced
② = Negligible impact when mobile speed increases from 30 to 120 km/h
③ = Gain from more accurate link adaptation at 3 km/h

Figure 8.8 Example of downlink cell throughput with different schedulers and mobile speeds



① = Resource fair scheduling
② = Throughput fair reduces cell throughput
③ = 2×2 MIMO provides more gain over 1×2 with Round robin than with proportional fair scheduling

Figure 8.9 Example of downlink cell throughputs and user throughputs with different antenna configuration and different schedulers

accurately at 3 km/h. The capacity difference between proportional fair and round robin is small at 120 km/h, but increases at low mobile speeds.

Figure 8.9 shows the cell and user throughput with 1×2 antenna configuration (1 transmission branch at eNodeB and 2 receiver branches at UE) and with 2×2 MIMO at 3 km/h. Three different types of schedulers are presented: round robin, proportional fair with equal resource scheduling, and proportional fair with equal throughput. The proportional fair with equal resource scheduling increases both cell and user throughput with both antenna configurations. The relative gain from the proportional fair is higher with 1×2 since that antenna configuration has less diversity and experiences more fading than 2×2 MIMO. More fading gives more freedom for the scheduler to optimize the transmission. The throughput fair scheduling gives the same throughput for all users, but the cell throughput is clearly reduced compared to the resource fair scheduling. The cell edge users get higher throughput with throughput fair scheduling while those users in good locations get higher throughput with resource fair scheduling.

In the following we illustrate the design and the performance of a QoS-aware packet scheduler and the QoS impact on the cell throughput and cell capacity under full load. We consider the case of traffic mixes of best effort traffic (modeled as simple file download) with no QoS constraints and constant bit rate (CBR) applications with strict GBR QoS constraints. The results presented in the following are obtained with 1×2 interference rejection combining (IRC) antennas, for a system bandwidth of 10 MHz and for localized transmission.

The time domain (TD) packet scheduler is the primary entity to facilitate QoS differentiation, in terms of both service and user differentiation. An example of a QoS-aware packet scheduler is introduced in Table 8.2. The proposed scheduler is limited here to GBR-awareness for simplicity. It uses a GBR-aware time domain metric based on a modified version of the proportional fair scheduler with adjustment according to the estimated required activity level for a certain user. Such a Required Activity Detection (RAD) scheduler dynamically estimates the capacity used by serving the GBR users and it is able to share the remaining part, often called the excess capacity, according to any desired policy, e.g. by setting QoS class-specific shares. The frequency domain scheduler is based on a modification of the proportional fair metric, called proportional fair scheduler (PFsch). Similarly to the time domain proportional fair scheduler,

Table 8.2 Example of QoS-aware packet scheduler in LTE

Scheduler	Time domain scheduling metric, M_n	Frequency domain scheduling metric, $M_{n,k}$
(TD) QoS PS \	Required Activity Detection (RAD) [14], [15] $\frac{D_n}{\overline{R}_n} \cdot \left(\frac{GBR_n}{\overline{R}_{sch,n}} + Share_n \cdot ExcessCap \right)$	Proportional Fair scheduled (PFsch) [21] $\frac{d_{n,k}}{\overline{R}_{sch,n}}$

D_n = Instantaneous wideband achievable throughput for user n.
$\overline{R}_n$ = Past average throughput of user n.
$d_{n,k}$ = Instantaneous achievable throughput for user n on PRB k.
$\overline{R}_{sch,n}$ = Past average throughput over the TTIs where user n is selected by the TD scheduler.
GBR_n = Guaranteed bit rate of user n.
$ExcessCap$ = Excess capacity (i.e. capacity left after minimum QoS is fulfilled).
$Share_n$ = Share of the excess capacity of user n.

the PFsch scheduler is designed to opportunistically benefit from radio channel aware multi-user frequency diversity and it aims at maximizing the spectral efficiency while providing resource fairness among the users. This scheduler decouples the behavior of the frequency domain scheduler from the time domain scheduler decisions, and vice versa [21]. In fact, the time–frequency domain coupling occurs whenever the time and the frequency domain metrics perform the allocation of the activity and the frequency resources based on the same feedback, e.g. the past average throughput. The coupling may lead to conflicting decisions when users are assigned a high activity rate but a low share of frequency resources, or vice versa.

Figure 8.10 shows the performance of the QoS packet scheduler as described in Table 8.2. The environment is 3GPP macro cell case #1. The results are obtained assuming different traf-

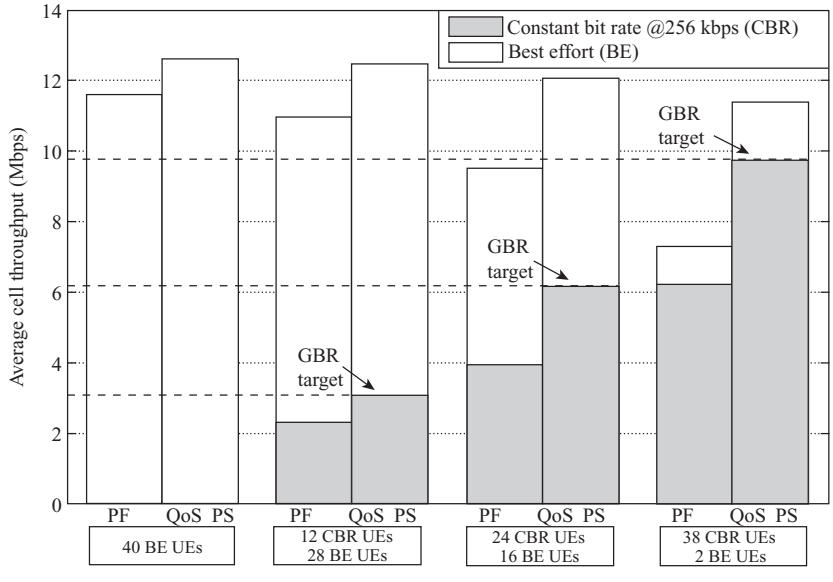


Figure 8.10 Average cell throughput for different traffic mixes of best effort and CBR traffic, for proportional fair scheduler and QoS-aware packet scheduler (see Table 8.2)

fic mixes of best effort traffic and constant bit rate (CBR) applications with strict GBR QoS constraints of 256 kbps. As reference, results with the time/frequency domain proportional fair scheduler are also depicted. As expected, the proportional fair scheduler is not able to meet the GBR demands of a large percentage of CBR users as some users do not have sufficient performance with the proportional fair even-resource allocation principle. The figure illustrates that the QoS-aware scheduler strictly provides QoS to 100% of the CBR users and it has the ability to distribute the excess capacity among the best effort users. It is also observed that the average cell throughput with only best effort traffic decreases when more CBR users are present in the cell as shown with 12, 24 and 38 CBR users per cell, respectively, over the total of 40 users. The latter is observed because the CBR users are strictly served with their GBR requirement regardless of their location in the cell, hence reducing the freedom for the packet scheduler to optimize the cell throughput.

To tackle the need for supporting high GBRs compared to the cell throughput, the QoS aware time domain scheduler alone is not sufficient. To fully understand this limitation one should consider that, with a certain FDPS policy, a CBR user may not be capable of meeting its GBR requirement from a link budget perspective by only being assigned the maximum activity (i.e. the time domain scheduler selects the user every TTI). Note that the highest supportable GBR depends on the system bandwidth, the environment, the frequency domain packet scheduler, and the number of UEs multiplexed per TTI. Therefore, to support high GBR, QoS awareness is required to be further included in the frequency domain scheduler. An example of an adjustment to be applied to the frequency domain metric of GBR users is the following weight factor:

$$w_{FD_n} = \frac{GBR_n}{\overline{R}_{schPerPRB,n}} \tag{8.1}$$

where $\overline{R}_{schPerPRB,n}$ is the past average throughput per PRB over the TTIs where user n is allocated PRB resources. The adjustment estimates the share of frequency resources required by a GBR user to meet the GBR requirement, and it aims at allocating a higher share, when needed. For the sake of simplification, details are omitted on the weight lower bound and on the normalization of the non-GBR frequency domain metrics needed to account for the constraint of a limited number of frequency resources.

Figure 8.11 illustrates the performance of the QoS PS scheduler (see Table 8.2) with a mix of 12 best effort users with no QoS and 8 CBR users with strict GBR QoS constraints increasing from 512 kbps, to 768 kbps and up to 1024 kbps. It can be observed that the QoS differentiation uniquely enforced by QoS aware time domain is not sufficient for supporting 100% of the CBR traffic when GBR equals 768 and 1024 kbps. When applying the frequency domain weight, w_{FD}, results show that a strict provision of the GBR requirements for all the CBR users is ensured with a limited cost in decreased FDPS gain.

8.5 Uplink Dynamic Scheduling and Link Adaptation

There are a few differences between uplink and downlink packet schedulers. The main differences are listed below.

1 eNodeB does not have full knowledge of the amount of buffered data at the UE. See section 8.5.1.2 for details on buffer status reports.

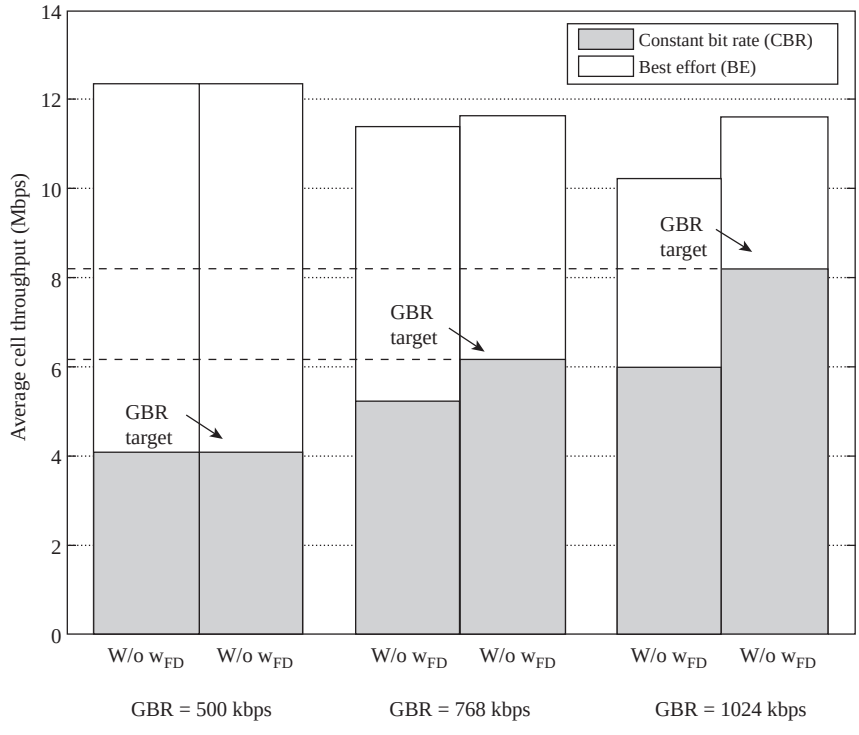


Figure 8.11 Average cell throughput for a traffic mix of 12 best effort and 8 CBR users per cell, when GBR equals 512, 768 and 1024 kbps. Results are obtained with the QoS aware packet scheduler as in Table 8.2, with and without the QoS-aware frequency domain weight

2 The uplink direction is more likely going to be power limited than the downlink due to low UE power compared to eNodeB. This means that especially in macro cell deployments users cannot always be allocated a high transmission bandwidth to compensate for poor SINR conditions.

3 Only adjacent physical resource blocks can be allocated to one user in uplink due to single carrier FDMA transmission. The fundamental difference between PRB allocation in LTE downlink and uplink stems from the uplink channel design and is illustrated with an example in Figure 8.12. The single-carrier constraint in LTE uplink limits both frequency and multi-user diversity.

4 The uplink is typically characterized by high interference variability [24] [25]. Interference variations from TTI to TTI in the order of 15–20 dB make it a hard task to estimate accurately the instantaneous uplink interference. Therefore in LTE uplink fast adaptive modulation and coding and channel-aware FDPS is typically based on channel state information about the own desired signal path rather than on the instantaneous interference level.

5 An uplink grant transmitted on PDCCH in TTI n refers to uplink transmission in TTI $n+4$. This 4 ms delay is due to PDCCH decoding and processing time at the UE and represents a further limitation to link adaptation and channel-aware packet scheduling in uplink. For more details, please refer to the PDCCH physical layer description in Chapter 5.

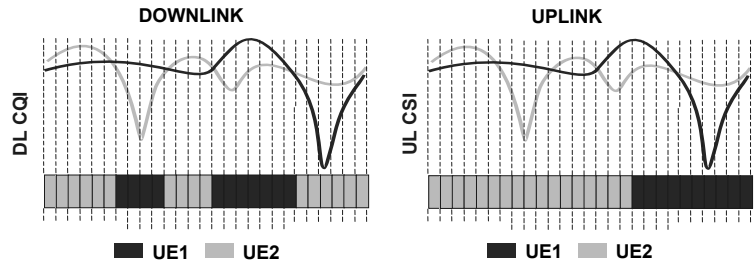


Figure 8.12 Example illustrating the single-carrier constraint to frequency domain packet scheduling in uplink

An overview of the uplink RRM functionalities and their interaction is given in Figure 8.13. As for the downlink, both uplink fast adaptive modulation and coding and frequency domain packet scheduler rely on frequency-selective channel state information (CSI). Uplink channel state information is estimated based on the SRS transmitted by the UE. The core of the uplink RRM functionality is the interaction between the packet scheduler, including fast Adaptive Transmission Bandwidth (ATB), and the so-called Link Adaptation (LA) unit, including Power Control (PC), adaptive modulation and coding and Outer-Loop Link Adaptation (OLLA). Buffer status reports and scheduling requests are also part of the core input to the uplink scheduler.

Interaction between the uplink dynamic packet scheduler and adaptive modulation and coding – The adaptive modulation and coding function is responsible for providing information to the packet scheduler on the channel state of a certain user in correspondence of a specific

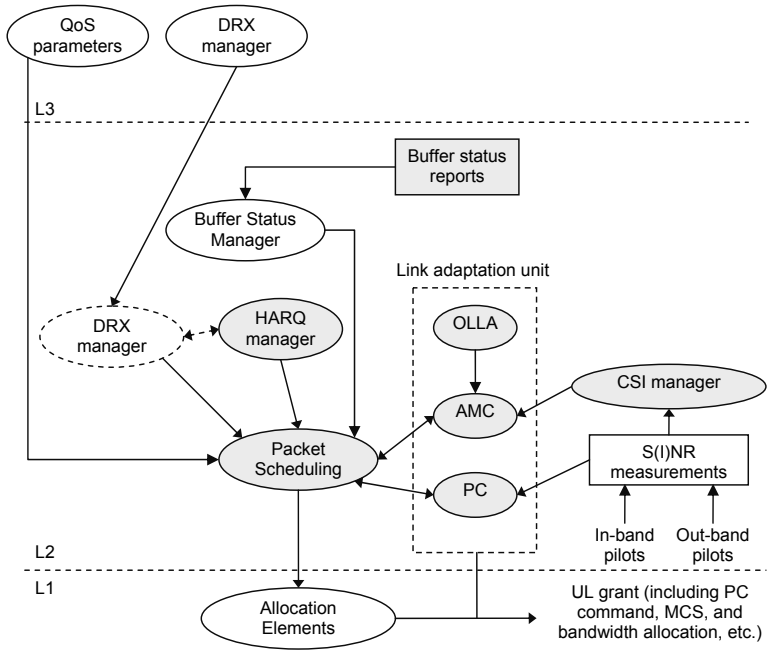


Figure 8.13 Inter-working between packet scheduling, link adaptation unit, and other uplink RRM functionalities

transmission bandwidth. In this sense the adaptive modulation and coding function is the link between the uplink packet scheduler and the uplink channel state information manager, which collects uplink channel state information based on the transmission of SRS in the uplink. Fast adaptive modulation and coding is also responsible for selecting the most appropriate MCS once the uplink packet scheduler has allocated a specific uplink band to the corresponding UE.

Interaction between packet scheduler and power control – The main scope of power control in LTE uplink is to limit inter-cell interference while respecting minimum SINR requirements based on QoS constraints, cell load and UE power capabilities. The uplink transmission power is set in the UE based on the following standardized formula [26]:

$$P_{\text{PUSCH}} = \min\{P_{\text{MAX}}, 10\log_{10}(M_{\text{PUSCH}}) + P_{\text{O\_PUSCH}} + \alpha \cdot PL + \Delta_{\text{MCS}} + f(\Delta_i)\} \text{ [dBm]} \qquad (8.2)$$

where P_{MAX} is the maximum allowed power and depends on the UE power class, M_{PUSCH} is the number of allocated PRBs on the Physical Uplink Shared Channel (PUSCH), PL is the downlink path loss measured by the UE (including distance-dependent path loss, shadowing and antenna gain), $P_{\text{0\_PUSCH}}$, α and Δ_{MCS} are power control parameters, and $f(\Delta_i)$ is the closed loop power control correction transmitted by the eNodeB [26].

Figure 8.14 [27] is obtained assuming $\Delta_{\text{MCS}} = 0$ and $f(\Delta_i) = 0$ in equation (8.1), and clearly illustrates how the distribution of uplink SINR – and hence the system performance – strongly depends on the power control parameters, as well as on the propagation scenario. In general power control in uplink determines the average SINR region a user is operated at, which is similar to the G-factor in downlink. Since the UE performs power scaling depending on the allocated transmission bandwidth, the packet scheduler needs to have information on the UE transmit power spectral density so that the allocated transmission bandwidth does not cause the UE power capabilities to be exceeded. For such use, power headroom reports have been standardized for LTE uplink; see section 8.5.3.1 for more details.

Interaction between outer loop link adaptation and adaptive modulation and coding – As already pointed out, the adaptive modulation and coding functionality is responsible for two main tasks: (i) providing the link between the packet scheduler and the channel state informa-

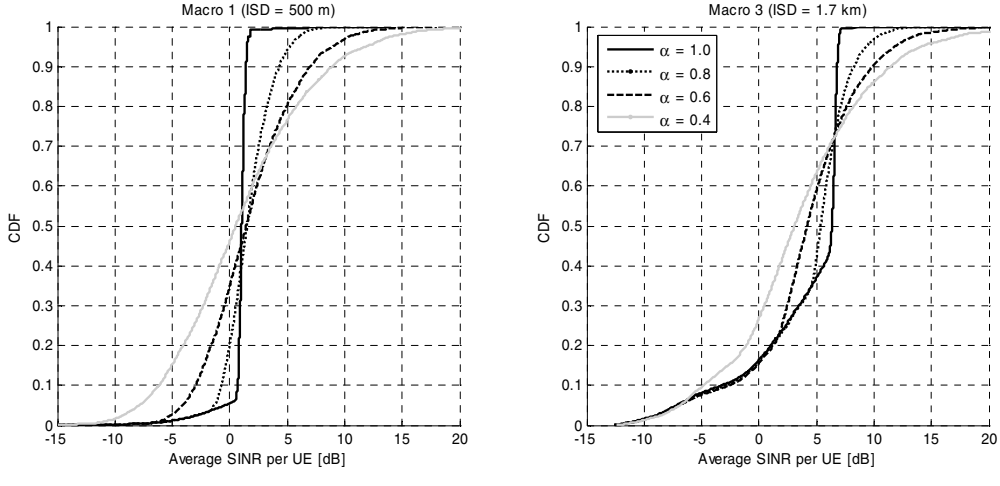


Figure 8.14 CDF of the average SINR per user for different propagation scenarios and values [27]

tion manager, and (ii) selecting the most appropriate MCS to be used for transmission on the selected uplink band by the uplink packet scheduler. In order to fulfill these tasks the adaptive modulation and coding functionality needs to map channel state information from the channel state information manager into SINR and/or MCS. This function is similar to downlink; the main goal of OLLA is to compensate for systematic errors in the channel state information at the output of the channel state information manager so that packet scheduler and adaptive modulation and coding can perform correctly.

8.5.1 Signaling to Support Uplink Link Adaptation and Packet Scheduling

Fast link adaptation and channel-aware scheduling in LTE uplink strongly relies on the availability of frequency-selective channel state information based on SRS measurements. Moreover, since the uplink transmission buffers are located in the UE, the information on the buffer status needs to be transmitted in the uplink as well. Also, the user needs to report power headroom measurements in order to signal to the uplink packet scheduler (and to the fast ATB functionality in particular) how close it is operating to its maximum power capabilities. While the SRS physical layer details are covered in Chapter 5, in the following section RRM aspects of the LTE uplink channel sounding concept are briefly discussed. Next, the transmission of the so-called uplink scheduling information (i.e. buffer status and power headroom reports) is also addressed.

8.5.1.1 Sounding Reference Signals (SRS)

The uplink sounding concept basically determines 'when and how' an uplink SRS is transmitted. As a consequence, the sounding concept and parameters have a significant impact on the accuracy of uplink SRS measurements, and hence on the performance of uplink link adaptation and packet scheduling. From an uplink RRM perspective, some of the most important sounding parameters are [28]:

- SRS bandwidth: Indicates the transmission bandwidth of the uplink SRS. The SRS bandwidth is semi-statically signaled via RRC.
- SRS period and time offset: Indicates System Frame Number (SFN) modulo number and periodicity of SRS from one UE. The SRS period and time offset are semi-statically signaled via RRC. 3GPP has also standardized the possibility for the eNodeB to disable SRS transmission on a per-UE basis.
- SRS duration: Indicates for how long a UE must keep on transmitting the uplink SRS. SRS duration is also semi-statically signaled via RRC.
- Transmission combination, Constant Amplitude Zero Autocorrelation (CAZAC) sequence index and cyclic shift: Necessary to guarantee orthogonality among users transmitting uplink SRS using the same transmission bandwidth (see more details in Chapter 5). In LTE uplink up to 6 UEs/cell can share the same SRS bandwidth without interfering with each other.
- SRS sub-band hopping sequence: Determines the hopping sequence in case, for example, the SRS bandwidth is much narrower than the available bandwidth for scheduling.

It is an uplink RRM task to distribute the limited SRS resources among active users so that accurate and up-to-date channel state information is available. For instance, typically there is a

tradeoff between measurement accuracy and SRS bandwidth (especially for power limited users): The narrower the SRS bandwidth, the higher the measurement accuracy [29]. On the other hand, several narrowband SRS transmissions are required to sound the entire schedulable bandwidth. Therefore, channel state information can become relatively outdated if the SRS periodicity and sub-band hopping are not configured properly. Another task of the uplink RRM functionality is to decide which users should be transmitting SRSs using the same time and frequency resources, since their orthogonality strongly depends on the received power level from the different UEs.

8.5.1.2 Buffer Status Reports (BSR)

The buffer status information is reported in uplink to inform the uplink packet scheduler about the amount of buffered data at the UE. LTE introduces a buffer status reporting mechanism that allows distinguishing between data with different scheduling priorities. The LTE buffer status reporting mechanism basically consists of two phases: (i) Triggering and (ii) Reporting.

Triggering
A Buffer Status Report (BSR) is triggered if any of the following events occur [30]:

- Uplink data arrive in the UE transmission buffer and the data belong to a radio bearer (logical channel) group with higher priority than those for which data already existed in the UE transmission buffer. This also covers the case of new data arriving in an empty buffer, The Buffer Status Report is referred to as 'regular BSR'.
- Uplink resources are allocated and number of padding bits is larger than the size of the BSR MAC control element, in which case the Buffer Status Report is referred to as 'padding BSR' (see below).
- A serving cell change occurs, in which case the Buffer Status Report is referred to as 'regular BSR' (see below).
- The periodic BSR timer expires, in which case the BSR is referred to as 'periodic BSR' (see below).

Reporting
The main uplink buffer status reporting mechanisms in LTE are the Scheduling Request (SR) and the Buffer Status Report (BSR) [30].

Scheduling request (SR): The SR is typically used to request PUSCH resources and is transmitted when a reporting event has been triggered and the UE is not scheduled on PUSCH in the current TTI. The SR can be conveyed to the eNodeB in two ways:

- Using a dedicated 'one-bit' BSR on the Physical Uplink Control Channel (PUCCH), when available. The occurrence of SR resources on PUCCH is configured via RRC on a per UE basis. It might be possible that no resources for SR are allocated on PUCCH.
- Using the Random Access procedure. Random Access is used when neither PUSCH allocation nor SR resources are available on PUCCH.

According to the 3GPP specifications [30], a SR is only transmitted as a consequence of the triggering of a 'regular BSR'. The triggering of 'periodic BSR' and 'padding BSR' does not cause the transmission of a SR.

Buffer status report (BSR): BSRs are transmitted using a Medium Access Control (MAC) control element when the UE is allocated resources on PUSCH in the current TTI and a reporting event has been triggered. Basically the buffer status report is transmitted as a MAC-C PDU with only header, where the field length is omitted and replaced with buffer status information [30].

In summary, when a reporting event is triggered:

- If the UE has resources allocated on PUSCH, then a buffer status report is transmitted.
- If a 'regular BSR' is triggered and the UE has no PUSCH allocation in the current TTI but has SR resources allocated on PUCCH, then a SR is transmitted on PUCCH at the first opportunity.
- If a 'regular BSR' is triggered and the UE has neither SR resources allocated on PUCCH nor PUSCH allocation, then a SR is issued using the Random Access Procedure.

The design of buffer status reporting schemes and formats for UTRAN LTE uplink has been driven by two important factors [6]:

- Separate buffer status reports for data flows with different QoS characteristics/requirements are needed to support QoS-aware radio resource allocation.
- The overhead from the BSR needs to be minimized since it directly affects the uplink capacity.

Therefore within 3GPP it has been agreed that buffer status is reported on a per Radio Bearer Group (RBG) basis. A RBG is defined as a group of radio bearers with similar QoS requirements. The maximum number of radio bearer groups, however, has been fixed and is four. An example showing the mapping from radio bearers to RBG for buffer status reporting is shown in Figure 8.15.

Two BSR formats are used in LTE uplink [30]: Short BSR (only one radio bearer group is reported) and long BSR (all four radio bearer groups are reported). In the first case 2 bits are required for radio bearer group identification, while in the latter case the 4 buffer size fields can be concatenated. In any case, the buffer size fields of the BSR are 6 bits long.

The basic idea is to transmit a short BSR if there are only data from one radio bearer group in the UE buffer, otherwise always transmit a long BSR. Some exceptions have been included in the standard in order to take into account the cases where, for example, a short BSR fits in the allocated transport block size, but a long BSR does not (see [30] for more details).

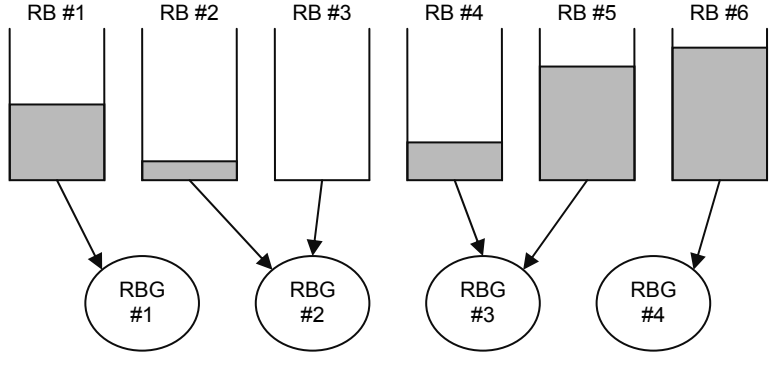


Figure 8.15 Example of mapping from RB to radio bearer group for buffer status reporting

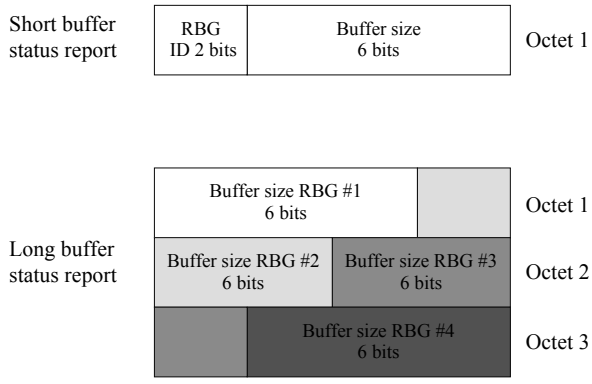


Figure 8.16 Short and long buffer status report types in LTE uplink [30]

8.5.1.3 Power Headroom Reports

Due to the open loop component in the standardized power control formula (see Chapter 5) the eNodeB cannot always know the power spectral density level used at the UE. Information on the power spectral density is important for performing correct RRM decisions at the eNodeB, especially when allocating the transmission format including bandwidth and modulation and coding scheme. Not knowing the power spectral density used by a certain terminal could, for example, cause the allocation of a too high transmission bandwidth compared to the maximum UE power capabilities, thus resulting in a lower SINR than expected. Information on the power spectral density used at the UE can be obtained from power headroom reports, provided that the eNodeB knows the corresponding transmission bandwidth. Information on the power spectral density is primarily critical for the PUSCH as the transmission format for this channel is adaptively adjusted. For all these reasons power headroom reports have been standardized in 3GPP. The power headroom is calculated at the UE as the difference between the maximum UE transmit power and the 'nominal' PUSCH transmit power in the corresponding subframe set according to the PUSCH power control formula in [26]. A power headroom report is triggered if any of the following criteria are met [30]:

- An appositely defined prohibit timer expires or has expired and the path loss has changed by more than a pre-defined threshold since the last power headroom report. The prohibit timer is introduced to limit unnecessarily frequent transmissions of power headroom reports.
- An appositely defined periodic timer expires.

More details on the triggering and transmission of power headroom reports can be found in [30].

8.5.2 Uplink Link Adaptation

Uplink link adaptation strongly relies on uplink SRS measurements. The uplink sounding concept and parameters basically determine how often and on which uplink physical resources the SRS measurements are available for a given UE, the measurement accuracy, etc. Then, similarly

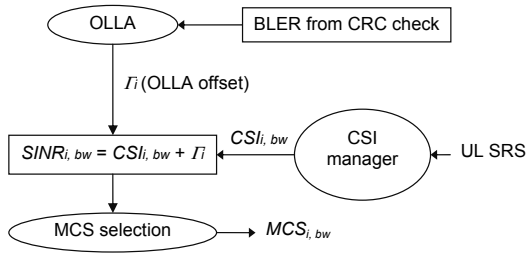


Figure 8.17 Schematic of the uplink fast adaptive modulation and coding functionality

to downlink CQI, uplink SRS measurements are used by the channel state information manager to determine the most appropriate MCS for transmission on a specific uplink bandwidth. Also, uplink SRS measurements can be used by the uplink frequency domain scheduler when, for example, computing the scheduling priority of a specific user in correspondence of a specific uplink band. The uplink link adaptation functionality based on fast adaptive modulation and coding is schematically illustrated in Figure 8.17.

Uplink SRS measurements are eNodeB vendor specific and as a consequence are not strictly specified by 3GPP. However, it has been shown that uplink fast adaptive modulation and coding based on instantaneous SINR measurements can suffer from uplink interference variability [24]. On the other hand, selecting the MCS based on instantaneous channel conditions can still give a significant gain compared to slow AMC schemes [25]. Therefore one feasible implementation is to use uplink SRS to perform frequency-selective signal strength measurements with the aim of tracking fast fading on the signal component, while interference can be treated as a constant in order to minimize the effect of interference unpredictability.

8.5.3 Uplink Packet Scheduling

The main responsibility of the uplink packet scheduler is to share the available radio resources between users taking into account limitations and/or requirements imposed by other RRM functionalities.

- Packet scheduler and DRX/DTX. Users cannot be scheduled for transmission on PUSCH unless they are listening to the L1/L2 control channel.
- Packet scheduler and power control. UE transmission power capabilities must be considered when, for example, packet scheduler (ATB) allocates the uplink transmission bandwidth to a specific UE.
- Packet scheduler and QoS. The packet scheduler is responsible for fulfilling the user QoS requirements.
- Packet scheduler and BSRs. Users should be scheduled for data transmission only if they have data to transmit. Also, the prioritization between users can be carried out based on the information conveyed by BSRs (e.g. users with high priority data are prioritized over users with low priority data).
- Packet scheduler and HARQ. Synchronous HARQ is used for LTE uplink. Hence, the UE must be scheduled if an earlier transmission has failed.
- Packet scheduler and MIMO. Uplink multi-user (MU-) MIMO can be used in Release 8 of the 3GPP specifications. With MU-MIMO the uplink packet scheduler can simultane-

ously allocate the same frequency resources to two users. The orthogonality between users is achieved by exploiting frequency-domain channel state information available via SRS measurements. No specific signaling (neither downlink nor uplink) is needed to support MU-MIMO in uplink.

As for the downlink (see Figure 8.6), the uplink packet scheduler is split in two sub-units: one that works in the time domain (TD) and one that works in the frequency domain (FD). The time domain metric is typically computed based on the experienced QoS vs QoS requirements and the estimated buffer occupancy. The channel-aware time domain scheduling metrics only take into account wideband channel state information. Basically, the time domain scheduler for uplink and downlink follows the same principle. The only difference is that the time domain scheduler in uplink must always prioritize users with pending retransmissions independently of other users' priority, QoS and channel conditions due to the uplink synchronous HARQ.

The allocation in the frequency domain is performed based on specific frequency domain scheduling metrics, which are typically computed based on frequency selective channel state information available from SRS measurements. QoS requirements and BSRs can also be considered when deriving frequency domain scheduling metrics. The main difference between uplink and downlink frequency domain PS is that in uplink the FDPS algorithm has to deal with the restrictions imposed by single-carrier transmission and limited UE power capabilities.

8.5.3.1 Example of Frequency Domain Packet Scheduler with Fast Adaptive Transmission Bandwidth

LTE supports fast Adaptive Transmission Bandwidth (ATB) in the uplink, i.e. the user transmission bandwidth can be modified on a TTI basis depending on channel conditions, traffic, fulfillment of QoS requirements, etc. With fast ATB the allocation of the user transmission bandwidth can be integrated in the FDPS algorithm. In this way, for instance, QoS-aware schedulers can be implemented in combination with channel-aware scheduling by simply introducing a QoS-based metric per UE depending on the level of fulfillment of its QoS requirements. In the time domain, users are ordered based on such a QoS-based metric, while in the frequency domain the frequency selective scheduling metric can potentially be weighted by the QoS-based metric. In this way fast ATB can guarantee automatic adaptation to variations in the cell load by, for example, differentiating resource allocation between different QoS users in the time or frequency domain depending on the temporary cell load.

Assuming that frequency-selective uplink channel state information is available at the uplink packet scheduler based on uplink SRS measurements (see section 8.5.1), a scheduling metric per user per PRB can then be defined based on uplink channel state information as well as on fulfillment of QoS requirements, buffers status information, etc. An example of a FDPS algorithm, which is illustrated in Figure 8.18, starts allocating the user with the highest scheduling metric on the corresponding PRB. The user is then 'expanded' respecting the single-carrier transmission constraint of the LTE uplink until another user with a higher metric is met, or the UE transmission power constraints are exceeded (known via power headroom information reported by the UE), or the UE has no more data to transmit (known via buffer status information reported by the UE). Then a similar procedure is repeated until the entire bandwidth is used or there are no more users to schedule.

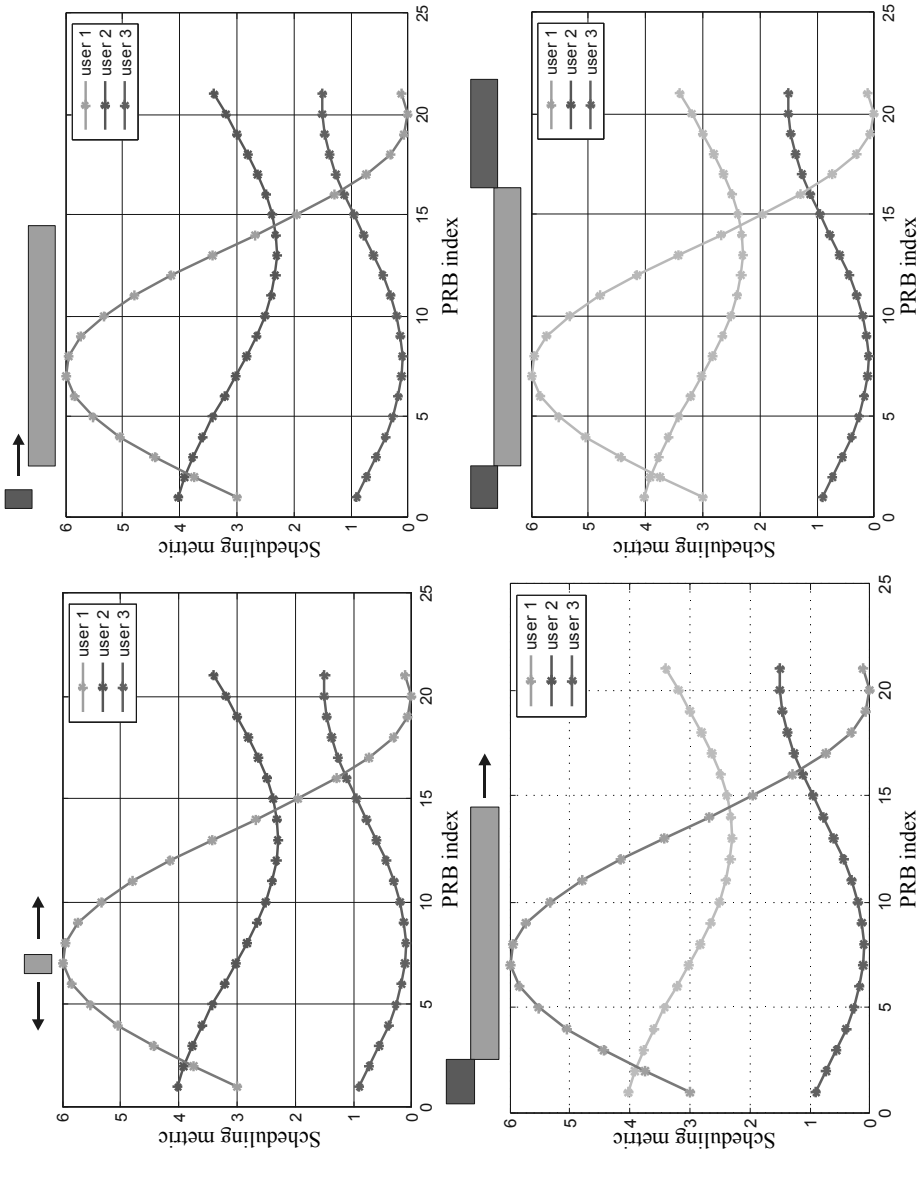


Figure 8.18 Example of combined fast adaptive transmission bandwidth and frequency domain packet scheduling algorithm in LTE uplink

The fast ATB framework described in this example is able to achieve most of the gain compared to more complex FDPS algorithms, which try to derive near-optimum FDPS solutions using more exhaustive approaches [31] [32]. Due to its iterative nature the complexity of the algorithm remains very high, however, unless the time domain framework effectively limits the scheduling order of the system. Because of this, many simulations that have been conducted during the 3GPP work item phase have been based on the assumption of fixed bandwidth uplink packet scheduling [33].

Figure 8.19 and Figure 8.20 show some examples of gain numbers of uplink FDPS (proportional fair) compared to blind scheduling in the frequency domain.

In general higher gain numbers can be observed when the capacity is measured with user throughput outage criteria. The reason is that the SINR gain from proportional fair scheduling typically maps into a higher throughput gain in the lower SINR regions. Also, note that for a given user transmission bandwidth, the cell throughput gain from frequency domain scheduling tends to saturate 15 simultaneous users.

By reducing the user transmission bandwidth the gain from frequency domain scheduling increases (see Figure 8.20). Due to the single-carrier constraint, however, a high number of users needs to be scheduled each TTI to fully exploit the frequency selectivity of the channel. In practice due to PDCCH limitations approximately 8–10 users can be scheduled in uplink each TTI, which means that the user transmission bandwidth is about 5–6 PRBs. Therefore while the potential gain from uplink frequency scheduling is about 40% in average cell throughput and 85% in outage user throughput, single-carrier constraint and practical PDCCH implementation limit the actually achievable gain to 25% and 50%, respectively.

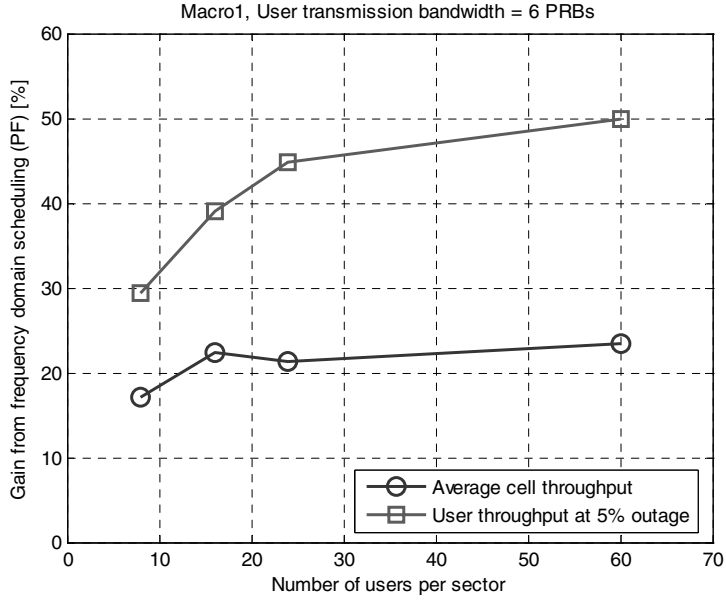


Figure 8.19 Gain from uplink frequency domain scheduling as a function of number of users for Macro 1 scenario and a fixed user transmission bandwidth of 6 PRBs

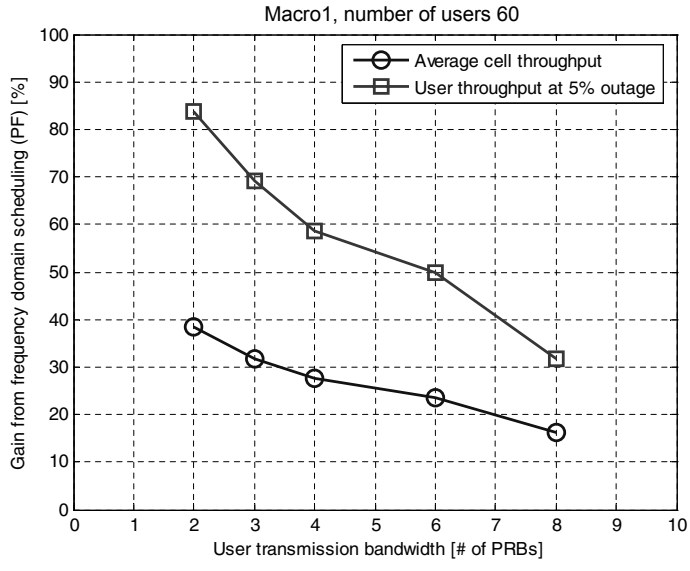


Figure 8.20 Gain from uplink frequency domain scheduling as a function of the user transmission bandwidth for Macro 1 propagation scenario and 60 simultaneous users

8.6 Interference Management and Power Settings

LTE includes a variety of mechanisms for controlling the interference between neighboring cells, also known as inter-cell interference control (ICIC). The standardized ICIC schemes for Release 8 primarily rely on frequency domain sharing between cells and adjustment of transmit powers. The X2 interface between eNodeBs includes standardized signaling for carrying interference information and scheduling information. The standardized ICIC methods are categorized as follows:

- Reactive schemes: Methods based on measurements of the past. Measurements are used to monitor the performance, and if the interference detected is too high, then appropriate actions are taken to reduce the interference to an acceptable level. Actions could include transmit power adjustments or packet scheduling actions to reduce the interference coupling between cells.
- Proactive schemes: Here an eNodeB informs its neighboring eNodeBs how it plans to schedule its users in the future (i.e. sending announcements), so that the neighboring eNodeB can take this information into account. Proactive schemes are supported via standardized signaling between eNodeBs over the X2 interface, e.g. by having eNodeBs sending frequency domain scheduling announcements to their neighbors.

3GPP Release 8 ICIC schemes are primarily designed for improving the performance of the uplink and downlink shared data channel (PDSCH and PUSCH). Hence, no explicit ICIC techniques are standardized for common channels like the BCCH, and control channels such as PDCCH and PUCCH (except for the use of power control). Note that for Release 8 time-synchronized networks (frame synchronized), the PDCCH and PUCCH always send at the same time and frequency.

8.6.1 Downlink Transmit Power Settings

Proactive downlink ICIC schemes are facilitated via the standardized Relative Narrowband Transmit Power (RNTP) indicator. The RNTP is an indicator (not a measurement) per PRB signaled to neighboring eNodeBs over the X2 interface. It indicates the maximum anticipated downlink transmit power level per PRB. Hence, from this information the neighboring eNodeBs will know at which PRBs a cell plans to use the most power, and the idea is thus that different power patterns can be used in those cells to improve the overall SINR conditions for the UEs. It is implementation specific how often a new RNTP indicator is sent from an eNodeB, and also the exact actions by an eNodeB receiving such messages are left unspecified by 3GPP.

By setting different downlink transmit powers per PRB, it is possible to dynamically configure different re-use patterns ranging from full frequency re-use (re-use one) with equal transmit power on all PRBs, to hard frequency re-use as shown in Figure 8.21, fractional frequency re-use as pictured in Figure 8.22, and soft frequency re-use as illustrated in Figure 8.23. The basic idea for fractional frequency re-use is that users close to their eNodeB are scheduled in the frequency band with frequency re-use one, while the cell-edge users are scheduled in the complementary frequency band with, for example, hard frequency re-use three in Figure 8.21. When using the soft frequency re-use configuration in Figure 8.23, users are scheduled over the entire bandwidth with different power levels. Cell-edge users are primarily scheduled in the bandwidth with the highest power level, while users closer to their serving eNodeB are scheduled on the PRBs with lower transmit power. Note that PRBs allocated to the same user should be transmitted with the same power level. Further information about performance results for different downlink transmit power profiles is presented in [9].

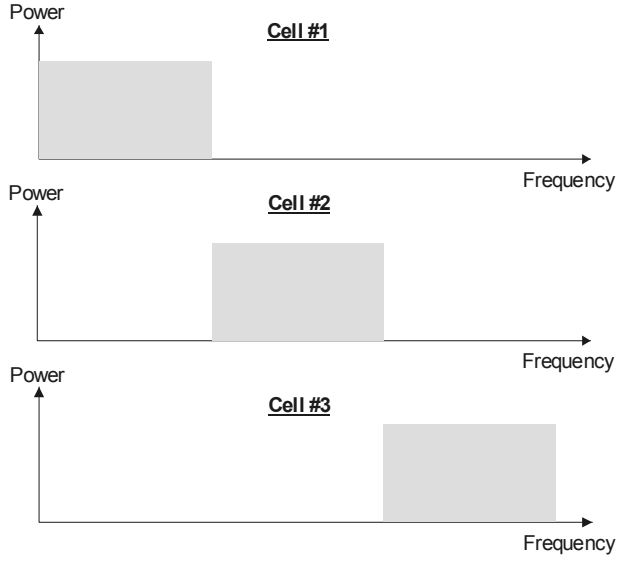


Figure 8.21 Example of downlink transmit power settings for hard frequency re-use three

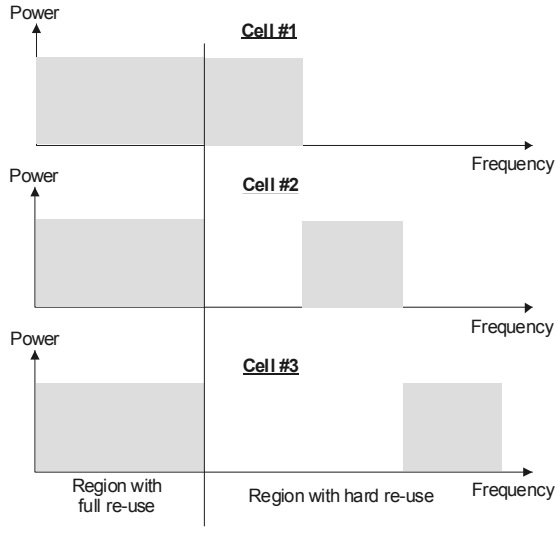


Figure 8.22 Example of downlink transmit power settings for a fractional frequency re-use

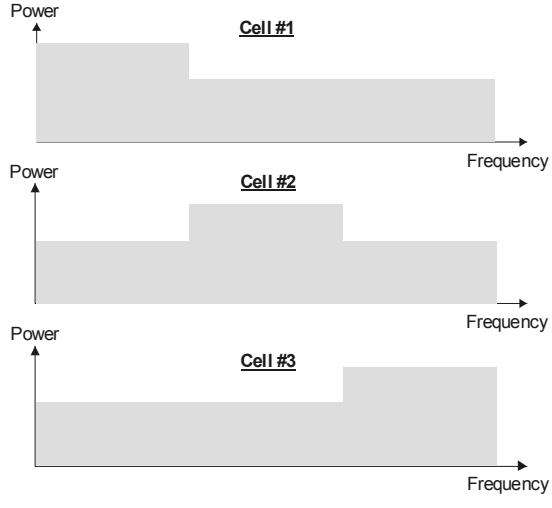


Figure 8.23 Example of downlink transmit power settings for a soft frequency re-use three

8.6.2 Uplink Interference Coordination

One proactive ICIC mechanism is standardized for the uplink, based on the High Interference Indicator (HII). The HII consists of sending a message over the X2 interface to the neighboring eNodeBs with a single bit per PRB, indicating whether the serving cell intends to schedule cell-edge UEs causing high inter-cell interference on those PRBs. Different HII messages can be sent from the serving cell to different eNodeBs. There are no standardized handshake procedures between eNodeBs. The latter means that it is open for eNodeB vendors (i.e. not

standardized by 3GPP) to decide when a new HII message is sent from an eNodeB, and the specific actions by the eNodeB receiving the message are also vendor specific. The basic idea behind the HII is that the serving cell will inform its neighboring eNodeBs at which PRBs it intends to schedule high interference users in the future. The neighboring eNodeBs should then subsequently aim at scheduling low interference users at those particular PRBs to avoid scheduling of cell-edge users at the same PRBs between two neighboring cells. Use of the HII mainly provides gain for fractional load cases, where there is only transmission on a sub-set of the PRBs per cell as in principle it can be used to dynamically achieve a hard/soft frequency re-use pattern in accordance with the offered traffic.

LTE also includes a reactive uplink ICIC scheme based on the Overload Indicator (OI). Here the basic idea is that the eNodeB measures the uplink interference+noise power, and creates OI reports based on this measurement. Low, medium, and high OI reports can be signaled over the X2 to neighboring cells. Hence, it is worth noting that the OI is primarily a function of the interference from other cells, and therefore does not include any information on the carried traffic or interference generated by the users in the serving cell. There is support for frequency selective OI, so that the aforementioned measurement and subsequent reporting to neighboring cells is per group of PRBs. One potential use is to dynamically adjust the uplink open loop power control parameters (e.g. Po) to maintain a certain maximum desirable uplink interference+noise level (or IoT level) based on the OI information exchanged between the eNodeBs.

8.7 Discontinuous Transmission and Reception (DTX/DRX)

LTE provides methods for the UE to micro-sleep even in the active state to reduce power consumption while providing high QoS and connectivity. DRX in the LTE sense means that the UE is not monitoring the PDCCH in the given subframe and is allowed to go to power saving mode. As uplink is scheduled in downlink PDCCH, the DRX parameter impacts both uplink and downlink performance for a UE.

The DRX concept contains different user-specific parameters that are configured via higher layer signaling. These parameters are described in Table 8.3 and illustrated in simplified form in Figure 8.24.

Basically, upon knowledge of the activity requirements in uplink and downlink for a certain UE, the regular DRX period including a certain planned on-time can be set. These two parameters are illustrated in Figure 8.24a. For highly predictable traffic (e.g. VoIP), the On Duration can be set to 1 subframe and the DRX Cycle to 20 ms or 40 ms if packet staggering is used. For traffic that is more dynamic and in bursts with tight delay requirements, it is possible to configure the user with a DRX Inactivity Timer where the packet scheduler can keep the UE awake by scheduling it within a certain time window. HARQ retransmissions are planned outside of the predefined DRX cycle to allow for a tighter DRX optimization without having to plan for worst-case retransmissions. There is a DRX retransmission timer defined so that a UE does not have to wait for a full DRX cycle for an expected retransmission that has been lost due to either ACK/NACK misinterpretation or PDCCH failed detection. Signaling of SRS and CQI is tied to the DRX parameters to reduce the use of PUCCH resources when signaling is not specifically needed for link adaptation and scheduling purposes and thus increases the power saving for the UE in long DRX periods. For very long DRX periods and a long time between SRS transmissions, the network and the UE may lose the time alignment and so the UE has to restart transmission to the network using RACH. Hence, this is also a consideration

Table 8.3 DRX related parameters and examples of their use/setting [34]

DRX Parameter	Description	Example settings and purpose
DRX Cycle	Specifies the periodic repetition of the On Duration followed by a possible period of inactivity.	The overall DRX cycle is set from knowledge of the service requirements of a user (e.g. responsiveness/latency requirements) as well as requirements for other mobility/update related requirements in the cell.
On Duration timer	This parameter specifies the number of consecutive subframes the UE follows the short DRX cycle after the DRX Inactivity Timer has expired.	Depending on the user's relative scheduling priority, this parameter can be set by the network to achieve a tradeoff among multi-user scheduling performance and single-user power saving.
DRX Inactivity Timer	Specifies the number of consecutive PDCCH-subframe(s) after successfully decoding a PDCCH indicating an initial uplink or downlink user data transmission for this UE.	The parameter provides a means for the network to keep a user 'alive' when there is suddenly a large need for scheduling a UE. Hence, the mechanism allows for large power saving while facilitating high QoS for bursty traffic types. Setting is a tradeoff among scheduler freedom and UE power saving as UE is forced to continue monitoring the PDCCH whenever scheduled.
DRX Retransmission Timer	Specifies the maximum number of consecutive PDCCH-subframe(s) where a downlink retransmission is expected by the UE after the first available retransmission time.	This parameter can be set to allow for asynchronous HARQ in the downlink also for HARQ users. If disabled, the network will need to send its retransmission at the first available time. Hence, the parameter is set as a tradeoff among scheduling freedom for retransmissions and UE power saving.
DRX Short Cycle	Specifies the periodic repetition of the On Duration followed by a possible period of inactivity for the short DRX cycle.	Short DRX is basically like an inactivity timer with gaps that is more suitable for services where the incoming data become available in 'gaps'.
DRX Short Cycle Timer	This parameter specifies the number of consecutive subframe(s) the UE follows the short DRX cycle after the DRX Inactivity Timer has expired.	This parameter is set from knowledge of the service pattern. A larger setting allows for more distributed data transmission but increases UE power consumption.

for the RRM management function to properly set available DRX/SRS parameters according to the planned load on the RACH channel.

Additionally, a short DRX cycle can be triggered that allows for periodic activity within the regular DRX cycle if additional and time-distributed scheduling resources are needed to facilitate the best power saving and QoS tradeoff for the given UE. The Short DRX concept is shown in Figure 8.24b and if the UE is scheduled in the current DRX window (e.g. regular On Duration window), new scheduling opportunities are created in distributed form within the current DRX cycle. The availability of the scheduling resources is controlled by the Short DRX Inactivity Timer and if it expires, the UE returns to the normal DRX pattern immediately.

The impact of the parameter settings on web browsing experience was studied in [35]. It was shown that with a DRX period set to 100 ms and the On Duration set to 1ms, the DRX

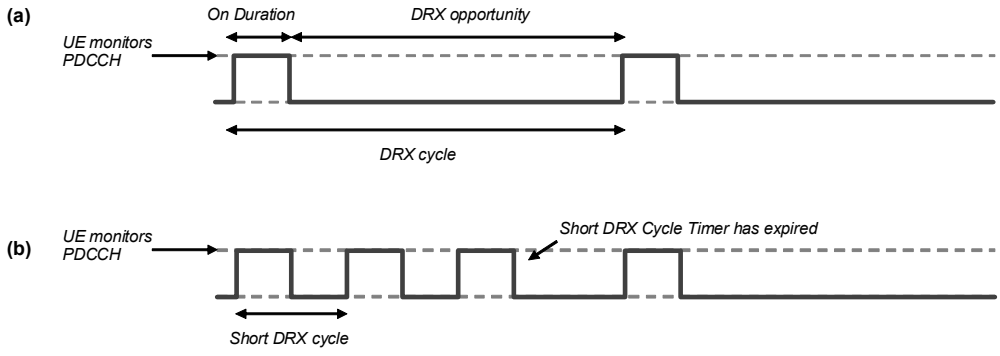


Figure 8.24 Simple illustration of DRX parameters

Inactivity timer could be set quite loosely from 25 to 50 ms while still achieving a large UE power saving, large scheduling freedom for the network, and allow for multiple web objects to be received within the active window thus improving the web browsing performance. With such settings the UE saw a 95% reduction in the radio-related power consumption while achieving 90% of the maximum achievable throughput.

Since DRX benefits UE power saving and thus is an important factor for ensuring the success of LTE, it is important to integrate its use tightly with the RRM concept in general. As the use of semi-static DRX takes the user out of the scheduling candidate set, the gain from multi-user packet scheduling may be significantly reduced. Hence, it is important that DRX parameters for a certain user are configured considering the multi-user situation in the cell as well as the service requirements for that particular user; e.g. activity needed to get the GBR requirement fulfilled or DRX period adjusted according to latency requirements. By using the span of the different DRX parameters, it is possible to effectively tune the tradeoff among scheduling flexibility for best cell-level performance and the UE's performance/power tradeoff.

8.8 RRC Connection Maintenance

The RRC connection maintenance is controlled by the eNodeB based on the RRM algorithms. When the UE has an RRC connection, the UE mobility is controlled by handovers. The handovers create some signaling traffic if the UE is moving. Therefore, if the UE is not transferring any data and is moving, it may be beneficial to release the RRC connection, while still maintaining the EPS bearer and the IP address. It may also be beneficial to release the RRC connection if the connection has been inactive for a long time or if the maximum number of RRC connections per base station is achieved. Example triggers for RRC connection release are illustrated in Figure 8.25. When the UE next transmits or receives data, the RRC connection is again re-established via RACH procedure.

8.9 Summary

The radio resource management algorithms are important to optimize the system capacity and end user performance. The network algorithms are not standardized but the network vendors

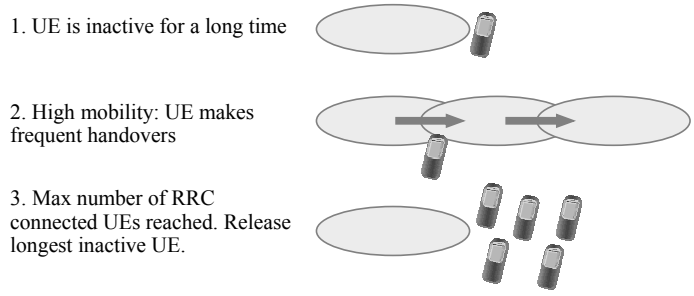


Figure 8.25 Triggers for RRC connection release

and operators can design and tune the algorithms according to the needs. The main algorithms in LTE are packet scheduling, admission control, power control and interference control. The LTE radio gives a lot of freedom in the packet scheduling area since the scheduling can be done in both the time and the frequency domain. It is shown that frequency domain scheduling can provide a clear capacity benefit at low mobile speeds compared to random scheduling. The uplink scheduling has less freedom compared to the downlink because the signaling from the network to the terminal takes time and because SC-FDMA transmission in uplink must use adjacent resource blocks.

The signaling in 3GPP has been defined to support efficient scheduling including the downlink Channel Quality Information (CQI), uplink Sound Reference Signals (SRS), uplink Buffer Status Reports (BSR) and uplink Power Headroom Reports (PHR).

The QoS differentiation allows separate treatment for different services, applications, subscribers or depending on the used amounts of data. The QoS is relevant in LTE since all the services are packet based including voice. The QoS concept in LTE is fully network controlled and is simplified compared to the existing 2G/3G networks. The QoS priority is typically achieved in time domain scheduling.

The LTE interface specifications support the inter-cell interference coordination to achieve dynamic frequency re-use configurations. The co-ordination can be used to optimize the resource allocation in the adjacent cells to maximize the cell edge throughput.

Discontinuous transmission and reception (DTX/DRX) is important to minimize the UE power consumption and to maximize the operating times for the handheld devices. The tuning of the DTX/DRX parameters also needs to consider the scheduling performance and the end to end performance.

References

[1] D. Niyato, E. Hossain, 'Connection Admission Control Algorithms for OFDM Wireless Networks', IEEE Proc. Globecomm, pp. 2455–2459, September 2005.

[2] P. Hosein, 'A Class-Based Admission Control Algorithm for Shared Wireless Channels Supporting QoS Services', in Proceedings of the Fifth IFIP TC6 International Conference on Mobile and Wireless Communications Networks, Singapore, October 2003.

[3] K.I. Pedersen, 'Quality Based HSDPA Access Algorithms', in IEEE Proc. Vehicular Technology Conference Fall, September 2005.

[4] 3GPP 23.401, 'Technical Specification Group Services And System Aspects; GPRS enhancements for E-UTRAN access (Release 8)', August 2007.

[5] 3GPP 23.203, 'Technical Specification Group Services and System Aspects; Policy and charging control architecture', v. 8.3.1.
[6] 3GPP 36.300 'Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 (Release 8)', October 2007.
[7] A. Pokhariyal, *et al*., 'HARQ Aware Frequency Domain Packet Scheduler with Different Degrees of Fairness for the UTRAN Long Term Evolution', IEEE Proc. Vehicular Technology Conference, pp. 2761–2765, April 2007.
[8] K.I. Pedersen, F. Frederiksen, T.E. Kolding, T.F. Lootsma, P.E. Mogensen, 'Performance of High Speed Downlink Packet Access in Co-existence with Dedicated Channels', IEEE Trans. Vehicular Technology, Vol. 56, No. 3, pp. 1261–1271, May 2007.
[9] A. Simonsson, 'Frequency Reuse and Intercell Interference Co-ordination in E-UTRA', IEEE Proc. Vehicular Technology Conference, pp. 3091–3095, April 2007.
[10] A. Pokhariyal, *et al*., 'Frequency Domain Packet Scheduling Under Fractional Load for the UTRAN LTE Downlink', IEEE Proc. Vehicular Technology Conference, pp. 699–703, April 2007.
[11] H. Holma, A. Toskala, 'WCDMA for UMTS – HSPA Evolution and LTE', 4th edition, Wiley, 2007.
[12] K.I. Pedersen, P.E. Mogensen, Troels E. Kolding, 'Overview of QoS Options for HSDPA', IEEE Communications Magazine, Vol. 44, No. 7, pp. 100–105, July 2006.
[13] P.A. Hosein, 'QoS Control for WCDMA High Speed Packet Data', IEEE Proc. Vehicular Technology Conference 2002.
[14] D. Laselva, *et al.*, 'Optimization of QoS-aware Packet Schedulers in Multi-Service Scenarios over HSDPA', 4th International Symposium on Wireless Communications Systems, pp. 123 – 127, ISWCS 2007, 17–19 October 2007.
[15] T.E. Kolding, 'QoS-Aware Proportional Fair Packet Scheduling with Required Activity Detection', IEEE Proc. Vehicular Technology Conference, September 2006.
[16] G. Barriac, J. Holtzman, ''Introducing Delay Sensitivity into the Proportional Fair algorithm for CDMA Downlink Scheduling'', IEEE Proc. ISSSTA, pp. 652–656, September 2002.
[17] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, P. Whiting, 'Providing Quality of Service over a Shared Wireless Link,' IEEE Communications Magazine, Vol. 39, No. 2, pp. 150–154, February 2001.
[18] F. Kelly, 'Charging and rate control for elastic traffic', European Trans. On Telecommunications, No. 8, pp. 33–37, 1997.
[19] J.M. Holtzman, ''Asymptotic Analysis of Proportional Fair Algorithm'', IEEE Proc. PIMRC, Personal Indoor and Mobile Radio Communication Conference, pp. F33–F37, September 2000.
[20] G. Song, Y. Li, 'Utility-Based Resource Allocation and Scheduling in OFDM-Based Wireless Broadband Networks', IEEE Communications Magazine, pp. 127–134, December 2005.
[21] G. Monghal, *et al*., 'QoS Oriented Time and Frequency Domain Packet Schedulers for The UTRAN Long Term Evolution', in IEEE Proc. VTC-2008 Spring, May 2008.
[22] C. Wengerter, J. Ohlhorst, A.G.E. v. ElbWart, 'Fairness and throughput analysis for Generalized Proportional Fair Frequency Scheduling in OFDMA', IEEE Proc. of Vehicular Technology Conference , pp. 1903–1907, Sweden, May 2005.
[23] H. Yang, 'A Road to Future Broadband Wireless Access: MIMO-OFDM-Based Air Interface', IEEE Communication Magazine, pp. 553–560, January 2005.
[24] I. Viering, A. Klein, M. Ivrlac, M. Castaneda, J.A. Nossek, 'On Uplink Intercell Interference in a Cellular System', IEEE International Conference on Communications (ICC), Vol. 5, pp. 2095–2100, Istanbul, Turkey, June 2006.
[25] C. Rosa, D. López Villa, C. Úbeda Castellanos, F.D. Calabrese, P. Michaelsen, K.I. Pedersen, Peter Skov, 'Performance of Fast AMC in E-UTRAN Uplink', IEEE International Conference on Communications (ICC 2008), May 2008.
[26] 3GPP TS 36.213, 'Technical Specification Group Radio Access Network; Physical layer procedures (Release 8)', November 2007.
[27] C. Úbeda Castellanos, D. López Villa, C. Rosa, K.I. Pedersen, F.D. Calabrese, P. Michaelsen, J. Michel, 'Performance of Uplink Fractional Power Control in UTRAN LTE', IEEE Trans. on Vehicular Technology, May 2008.
[28] R1-080900**, '**Physical-layer parameters to be configured by RRC', TSG RAN WG1 #52, Sorrento, Italy, 11–15 February 2008.

[29] R1-060048, 'Channel-Dependent Packet Scheduling for Single-Carrier FDMA in E-UTRA Uplink', 3GPP TSG RAN WG1 LTE Ad Hoc Meeting, Helsinki, Finland, 23–25 January 2006.
[30] 3GPP TS 36.321, 'Technical Specification Group Radio Access Network; Medium Access Control (MAC) protocol specification (Release 8)', March 2008.
[31] J. Lim, H.G. Myung, D.J. Goodman, 'Proportional Fair Scheduling of Uplink Single Carrier FDMA Systems', IEEE Personal Indoor and Mobile Radio Communication Conference (PIMRC), Helsinki, Finland, September 2006.
[32] F.D. Calabrese, P.H. Michaelsen, C. Rosa, M. Anas, C. Úbeda Castellanos, D. López Villa, K.I. Pedersen, P.E. Mogensen, 'Search-Tree Based Uplink Channel Aware Packet Scheduling for UTRAN LTE', in IEEE Proc. Vehicular Technology Conference, 11–14 May 2008.
[33] R1-060188, 'Frequency-domain user multiplexing for the E-UTRAN downlink', 3GPP TSG RAN WG1 LTE Ad Hoc Meeting, Helsinki, Finland, 23–25 January 2006.
[34] 3GPP 36.321 Technical Specifications 'Medium Access Control (MAC) protocol specification', v. 8.3.0.
[35] T. Kolding, J. Wigard, L. Dalsgaard, 'Balancing Power Saving and Single-User Experience with Discontinuous Reception in LTE', Proc. of IEEE International Symposium on Wireless Communication Systems (ISWCS), Reykjavik, Iceland, October 2008.

9

Performance

Harri Holma, Pasi Kinnunen, István Z. Kovács, Kari Pajukoski, Klaus Pedersen and Jussi Reunanen

9.1 Introduction

This chapter illustrates LTE capabilities from the end user's and from the operator's point of view. Radio performance has a direct impact on the cost of deploying the network in terms of the required number of base station sites and in terms of the transceivers required. The operator is interested in the network efficiency: how many customers can be served, how much data can be provided and how many base station sites are required. The efficiency is considered in the link budget calculations and in the capacity simulations. The end user application performance depends on the available bit rate, latency and seamless mobility. The radio performance defines what applications can be used and how these applications perform.

The link level studies in this chapter illustrate the main factors affecting LTE performance including mobile speed and transmission bandwidth. The LTE link performance is benchmarked with the theoretical Shannon limit. The capacity studies present the impact of the environment and show the bit rate distributions in typical network deployments. The relative capacity gains compared to HSPA networks are shown analytically and with simulations. The chapter presents also general dimensioning guidelines and the specific aspects of refarming LTE to the GSM spectrum. The network capacity management is illustrated with examples from High Speed Packet Access (HSPA) networks.

Practical performance is also impacted by the performance of the commercial UE and eNodeBs. To guarantee consistent performance, 3GPP has defined a set of radio performance requirements. 3GPP performance requirements are explained in detail in Chapter 11.

9.2 Layer 1 Peak Bit Rates

LTE provides high peak bit rates by using a large bandwidth up to 20 MHz, high order 64QAM modulation and multistream MIMO transmission. Quadrature Phase Shift Keying (QPSK) modulation carries 2 bits per symbol, 16QAM 4 bits and 64QAM 6 bits. 2 × 2 MIMO

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6

further doubles the peak bit rate up to 12 bits per symbol. Therefore, QPSK ½ rate coding carries 1 bps/Hz while 64QAM without any coding and with 2×2 Multiple Input Multiple Output (MIMO) carries 12 bps/Hz. The bandwidth is included in the calculation by taking the corresponding number of resource blocks for each bandwidth option: 6 with 1.4 MHz and 15 with 3 MHz bandwidth. The number of resource blocks for the bandwidths 5, 10, 15 and 20 MHz are 25, 50, 75 and 100 respectively. We assume the following control and reference signal overheads:

- Physical Downlink Control Channel (PDCCH) takes one symbol out of 14 symbols. That is the minimum possible PDCCH allocation. It is enough when considering single user peak bit rate. The resulting control overhead is 7.1% (=1/14).
- Downlink Reference Signals (RS) depend on the antenna configuration. Single stream transmission uses 2 RS out of 14 in every 3rd sub-carrier, 2×2 MIMO 4 symbols and 4×4 MIMO 6 symbols. The overhead varies between 4.8% and 14.3%. The RS partly overlap with PDCCH and the overlapping is taken into account.
- Other downlink symbols are subtracted: synchronization signal, Physical Broadcast Channel (PBCH), Physical Control Format Indicator Channel (PCFICH) and one group of Physical Hybrid Automatic Repeat Request Indicator Channel (PHICH). The overhead depends on the bandwidth ranging from below 1% at 20 MHz to approximately 9% at 1.4 MHz.
- No Physical Uplink Control Channel (PUCCH) is included in the calculation. PUCCH would slightly reduce the uplink data rate.
- Uplink reference signals take 1 symbol out of 7 symbols resulting in an overhead of 14.3% (=1/7).

The achievable peak bit rates are shown in Table 9.1 The highest theoretical data rate is approximately 172 Mbps. If 4×4 MIMO option is applied, the theoretical peak data rate increases to 325 Mbps. The bit rate scales down according to the bandwidth. The 5-MHz peak bit rate is 42.5 Mbps and 1.4-MHz 8.8 Mbps with 2×2 MIMO.

The uplink peak data rates are shown in Table 9.2: up to 86 Mbps with 64QAM and up to 57 Mbps with 16QAM with 20 MHz. The peak rates are lower in uplink than in downlink since single user MIMO is not specified in uplink in 3GPP Release 8. The single user MIMO

Table 9.1 Downlink peak bit rates (Mbps)

Resource blocks								
Modulation and coding	Bits/ symbol	MIMO usage	1.4 MHz 6	3.0 MHz 15	5.0 MHz 25	10 MHz 50	15 MHz 75	20 MHz 100
QPSK 1/2	1.0	Single stream	0.8	2.2	3.7	7.4	11.2	14.9
16QAM 1/2	2.0	Single stream	1.5	4.4	7.4	14.9	22.4	29.9
16QAM 3/4	3.0	Single stream	2.3	6.6	11.1	22.3	33.6	44.8
64QAM 3/4	4.5	Single stream	3.5	9.9	16.6	33.5	50.4	67.2
64QAM 1/1	6.0	Single stream	4.6	13.2	22.2	44.7	67.2	89.7
64QAM 3/4	9.0	2×2 MIMO	6.6	18.9	31.9	64.3	96.7	129.1
64QAM 1/1	12.0	2×2 MIMO	8.8	25.3	42.5	85.7	128.9	172.1
64QAM 1/1	24.0	4×4 MIMO	16.6	47.7	80.3	161.9	243.5	325.1

Table 9.2 Uplink peak bit rates (Mbps)

Resource blocks								
Modulation and coding	Bits/ symbol	MIMO usage	1.4 MHz 6	3.0 MHz 15	5.0 MHz 25	10 MHz 50	15 MHz 75	20 MHz 100
QPSK 1/2	1.0	Single stream	0.9	2.2	3.6	7.2	10.8	14.4
16QAM 1/2	2.0	Single stream	1.7	4.3	7.2	14.4	21.6	28.8
16QAM 3/4	3.0	Single stream	2.6	6.5	10.8	21.6	32.4	43.2
16QAM 1/1	4.0	Single stream	3.5	8.6	14.4	28.8	43.2	57.6
64QAM 3/4	4.5	Single stream	3.9	9.7	16.2	32.4	48.6	64.8
64QAM 1/1	6.0	Single stream	5.2	13.0	21.6	43.2	64.8	86.4

in uplink would require two power amplifiers in the terminal. MIMO can be used in Release 8 uplink as well to increase the aggregate cell capacity, but not single user peak data rates. The cell level uplink MIMO is called Virtual MIMO (V-MIMO) where the transmission from two terminals, each with single antenna, is organized so that the cell level peak throughput can be doubled.

The transport block sizes in [1] have been defined so that uncoded transmission is not possible. The maximum achievable bit rates taking the transport block sizes into account are shown in Table 9.3 for downlink and in Table 9.4 for uplink for the different modulation schemes. The downlink peak rate with 2×2 MIMO goes up to 150 Mbps and the uplink rate up to 75 Mbps. The calculations assume that uplink 16QAM uses Transport Block Size

Table 9.3 Downlink peak bit rates with transport block size considered (Mbps)

Resource blocks							
Modulation and coding	MIMO usage	1.4 MHz 6	3.0 MHz 15	5.0 MHz 25	10 MHz 50	15 MHz 75	20 MHz 100
QPSK	Single stream	0.9	2.3	4.0	8.0	11.8	15.8
16QAM	Single stream	1.8	4.3	7.7	15.3	22.9	30.6
64QAM	Single stream	4.4	6.5	18.3	36.7	55.1	75.4
64QAM	2×2 MIMO	8.8	8.6	36.7	73.7	110.1	149.8

Table 9.4 Uplink peak bit rates with transport block size considered (Mbps)

Resource blocks							
Modulation and coding	MIMO usage	1.4 MHz 6	3.0 MHz 15	5.0 MHz 25	10 MHz 50	15 MHz 75	20 MHz 100
QPSK	Single stream	1.0	2.7	4.4	8.8	13.0	17.6
16QAM	Single stream	3.0	7.5	12.6	25.5	37.9	51.0
64QAM	Single stream	4.4	11.1	18.3	36.7	55.1	75.4

(TBS) index 21, uplink QPSK uses TBS index 10, downlink 16QAM uses TBS index 15 and downlink QPSK uses TBS index 9.

The original LTE target was peak data rates of 100 Mbps in downlink and 50 Mbps in uplink, which are clearly met with the 3GPP Release 8 physical layer.

9.3 Terminal Categories

3GPP Release 8 has defined five terminal categories having different bit rate capabilities. Category 1 is the lowest capability with maximum bit rates of 10 Mbps downlink and 5 Mbps uplink while Category 5 is the highest capability with data rates of 300 Mbps in downlink and 75 Mbps uplink. The bit rate capability in practice is defined as the maximum transport block size that the terminal is able to process in 1 ms.

All categories must support all RF bandwidth options from 1.4 to 20 MHz, 64QAM modulation in downlink and 1–4 transmission branches at eNodeB. The receive antenna diversity is mandated via performance requirements. The support of MIMO transmission depends on the category. Category 1 does not need to support any MIMO while categories 2–4 support 2×2 MIMO and Category 5 support 4×4 MIMO. The uplink modulation is up to 16QAM in categories 1–4 while 64QAM is required in category 5.

The terminal categories are shown in Table 9.5. Further terminal categories may be defined in later 3GPP releases. The initial LTE deployment phase is expected to have terminal categories 2, 3 and 4 available providing downlink data rates up to 150 Mbps and supporting 2×2 MIMO.

Table 9.5 Terminal categories [2]

	Category 1	Category 2	Category 3	Category 4	Category 5
Peak rate downlink (approximately)	10 Mbps	50 Mbps	100 Mbps	150 Mbps	300 Mbps
Peak rate uplink (approximately)	5 Mbps	25 Mbps	50 Mbps	50 Mbps	75 Mbps
Max bits received within TTI	10 296	51 024	102 048	149 776	299 552
Max bits transmitted within TTI	5 160	25 456	51 024	51 024	75 376
RF bandwidth	20 MHz	20 MHz	20 MHz	20 MHz	20 MHz
Modulation downlink	64QAM	64QAM	64QAM	64QAM	64QAM
Modulation uplink	16QAM	16QAM	16QAM	16QAM	64QAM
Receiver diversity	Yes	Yes	Yes	Yes	Yes
eNodeB diversity	1–4 tx	1–4 tx	1–4 tx	1–4 tx	1–4 tx
MIMO downlink	Optional	2×2	2×2	2×2	4×4

9.4 Link Level Performance

9.4.1 Downlink Link Performance

The peak data rates are available only in extremely good channel conditions. The practical data rate is limited by the amount of interference and noise in the network. The maximum theoretical data rate with single antenna transmission in static channel can be derived using the Shannon formula. The formula gives the data rate as a function of two parameters: bandwidth and the received Signal-to-Noise Ratio (SNR).

$$\text{Bit rate [Mbps]} = \text{Bandwidth [MHz]} \cdot \log_2(1 + SNR) \tag{9.1}$$

The Shannon capacity bound in equation (9.1) cannot be reached in practice due to several implementation issues. To represent these loss mechanisms accurately, we use a modified Shannon capacity formula [3]:

$$\text{Bit rate [Mbps]} = BW\_eff \cdot \text{Bandwidth [MHz]} \cdot \log_2(1 + SNR/SNR\_eff) \tag{9.2}$$

where *BW_eff* accounts for the system bandwidth efficiency of LTE and *SNR_eff* accounts for the SNR implementation efficiency of LTE.

The bandwidth efficiency of LTE is reduced by several issues, as listed in Table 9.6. Due to the requirements of the Adjacent Channel Leakage Ratio (ACLR) and practical filter implementation, the bandwidth occupancy is reduced to 0.9. The overhead of the cyclic prefix is approximately 7% and the overhead of pilot assisted channel estimation is approximately 6% for single antenna transmission. For dual antenna transmission the overhead is approximately doubled to 11%. Note that here ideal channel estimation is used, which is the reason why the pilot overhead is not included in the link performance bandwidth efficiency but must be included in the system bandwidth level efficiency [3]. This issue also impacts the *SNR_eff*. The overall link-level bandwidth efficiency is therefore approximately 83%.

When fitting Equation 9.2 to the Shannon performance curve in Additive White Gaussian Noise (AWGN) channel conditions, we extract the best value for *SNR_eff* using the setting for *BW_eff* of 0.83 from Table 9.6 and the fitting parameters are indicated in parentheses as (*BW_eff*, *SNR_eff*). The results are presented in Figure 9.1. We can observe that LTE is performing less than 1.6~2 dB off from the Shannon capacity bound. There is nevertheless a minor discrepancy in both ends of the G-factor dynamic range. This is because the *SNR_eff*

Table 9.6 Link bandwidth efficiency for LTE downlink with a 10 MHz system

Impairment	*Link BW_eff*	*System BW_eff*
BW efficiency	0.90	0.90
Cyclic prefix	0.93	0.93
Pilot overhead	–	0.94
Dedicated and common control channels	–	0.715
Total	0.83	0.57

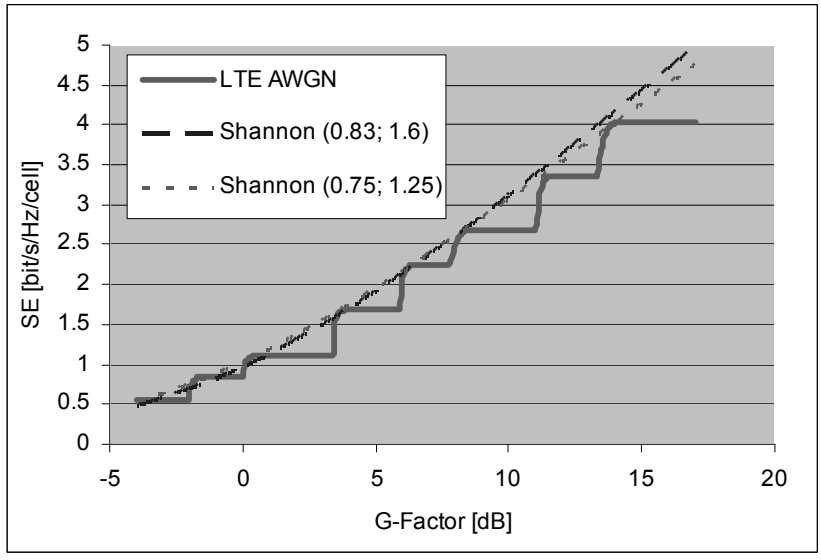


Figure 9.1 LTE spectral efficiency (SE) as a function of G-factor (in dB) including curves for best Shannon fit. The steps are due to the limited number of modulation and coding schemes in the simulation [3]. © 2007 IEEE

is not constant but changes with the G-factor. It is shown in [3] that this dependency can be accounted for using the fudge factor, η, multiplying the *BW_eff* parameter. For AWGN, $\eta = 0.9$ (*BW_eff*$*\eta = 0.75$) and *SNR_eff* = 1.0. The best fit to the link adaptation curve [3] is 1.25 dB.

In Figure 9.2 we show the spectral efficiency results versus G-factor from LTE link level studies and the best Shannon fit for a 3 km/h typical urban fading channel. Different antenna configurations are presented: single antenna transmission and reception Single Input Single Output (SISO) (1 × 1), single antenna transmission and antenna receive diversity Single Input Multiple Output (SIMO) (1 × 2) and Closed Loop MIMO (2 × 2). Using the G-factor dependent *SNR_eff* we achieve a visibly almost perfect fit to the link simulation results. For SISO it can be observed that the best Shannon fit parameters are significantly worsened compared to AWGN: the equivalent *BW_eff* has reduced from 0.83 to 0.56 (corresponding to fudge factor $\eta = 0.6$) and the *SNR_eff* parameter is increased from 1.6~2 dB to 2~3 dB. The match between the fitted Shannon curves and the actual link-level results is not perfect but sufficiently close for practical purposes.

These link-level results show that LTE downlink with ideal channel estimation performs approximately 2 dB from Shannon Capacity in AWGN, whereas the deviation between Shannon and LTE becomes much larger for a fading channel. This fading loss can, however, be compensated with the multi-user diversity gain when using frequency domain packet scheduling and the fitted Shannon curves can also include the system level scheduling gain [3]. With these adjustments, cell capacity results can be accurately estimated from the suggested modified Shannon formula and a G-Factor distribution according to a certain cellular scenario.

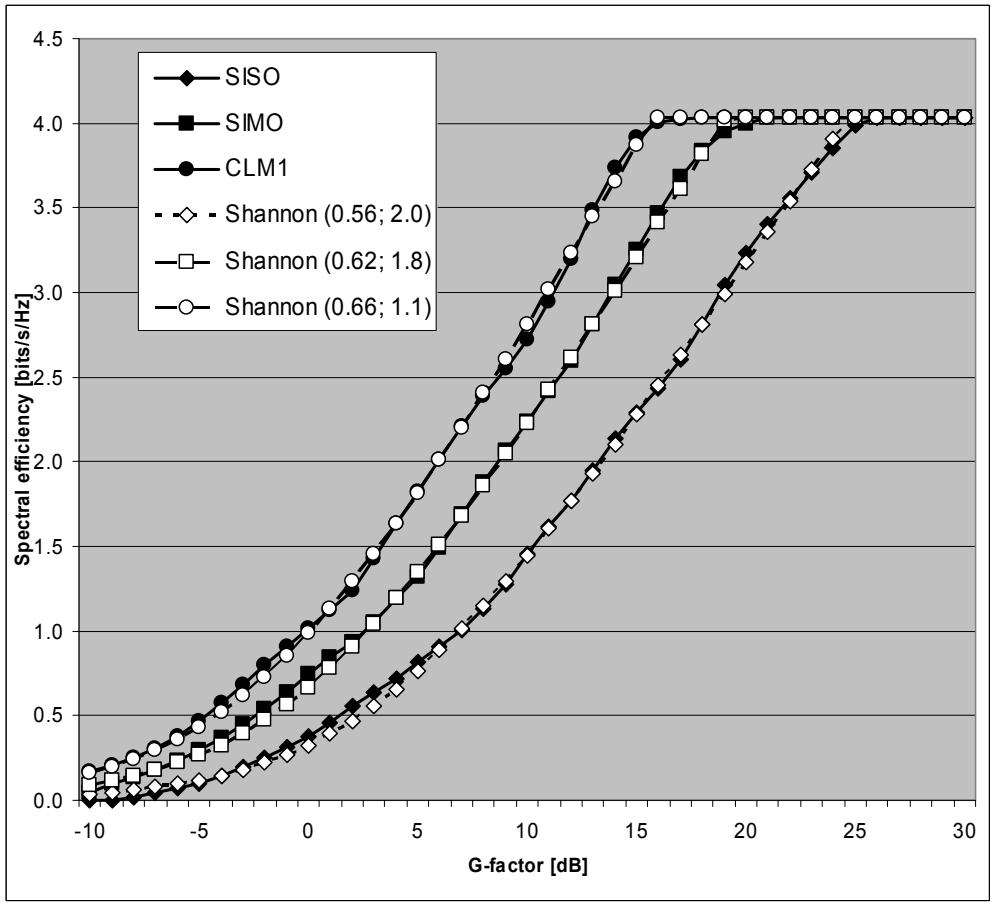


Figure 9.2 Spectral efficiency for SISO (1 × 1), SIMO (1 × 2) and Closed loop MIMO (2 × 2), as a function of G-factor. The best Shannon fit curves are plotted with parameters (*BW_eff*, *SNR_eff*) using Equation 9.2 [3]

9.4.2 Uplink Link Performance

9.4.2.1 Impact of Transmission Bandwidth

The uplink coverage can be optimized by bandwidth selection in LTE. The bandwidth adaptation allows the UE power to be emitted in an optimum bandwidth, which extends the coverage compared to having a fixed transmission bandwidth. A larger number of channel parameters need to be estimated for the wideband transmissions, which reduces the accuracy of the channel estimation due to noise. The optimized transmission bandwidth in LTE enables more accurate channel estimation for the lower data rates compared to WCDMA/HSUPA.

In Figure 9.3, the throughput as a function of SNR with different bandwidth allocation is shown. The smallest bandwidth allocation, 360 kHz, optimizes the coverage for bit rates below 100 kbps. The moderate bandwidth allocation, 1.08 MHz, gives the best coverage for bit rates

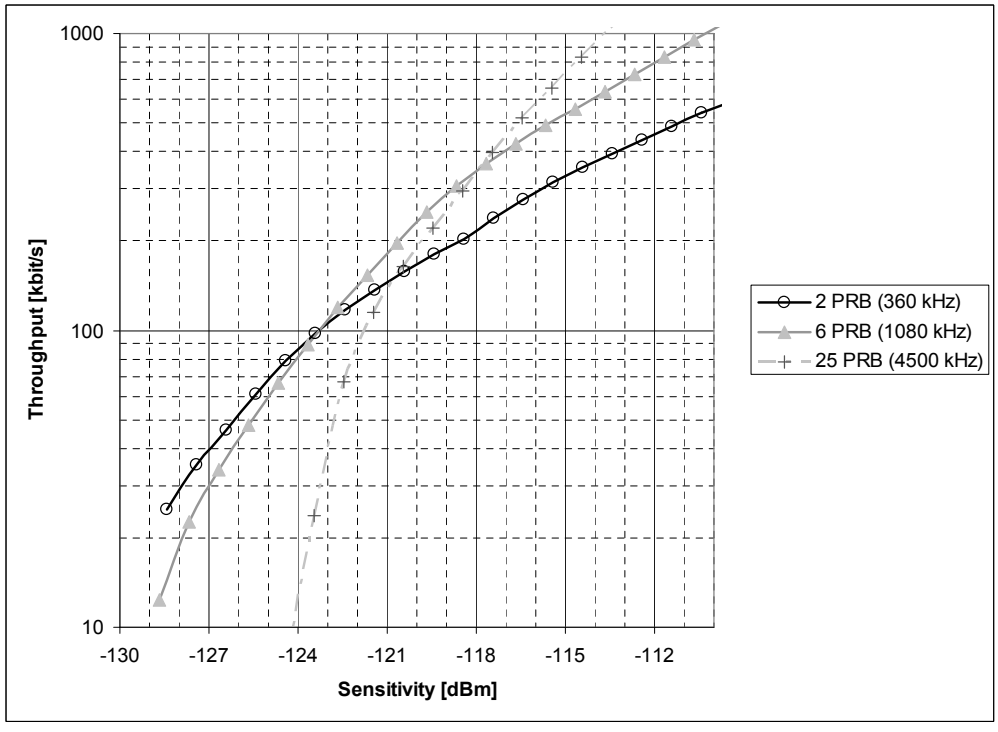


Figure 9.3 LTE eNodeB sensitivity as a function of received power with allocation bandwidths of 360 kHz, 1.08 MHz and 4.5 MHz

from 100 kbps to 360 kbps. The assumed base station noise figure is 2 dB and no interference margin is included.

9.4.2.2 Impact of Mobile Speed

Figure 9.4 shows the link adaptation curve of LTE uplink for the UE speeds of 3, 50, 120 and 250 km/h. The received pilot signals are utilized for channel estimation. The channel estimation is based on Wiener filter in both time and frequency domain. The impact of the mobile speed is small at low data rates while at high data rates the very high mobile speed of 250 km/h shows 30% lower throughput than at 3 km/h.

The required SNR values for various link efficiencies are given in Figure 9.5. In the efficiency range of QPSK (0.1–1 bits/Hz), the LTE uplink performance is affected only by 1–1.5 dB, when increasing UE speed from 3 to 250 km/h. For higher order modulations, 16QAM and 64QAM, the impact of the UE speed is higher – up to 3 dB for the efficiency of 2 bps/Hz/cell. In general, LTE uplink is robust against Doppler frequency shifts. The UE speed affects the throughput and SNR requirements mainly due to channel estimation performance: when the channel is changing fast, it is not possible to use a long averaging time and the accuracy of the channel estimation is impacted.

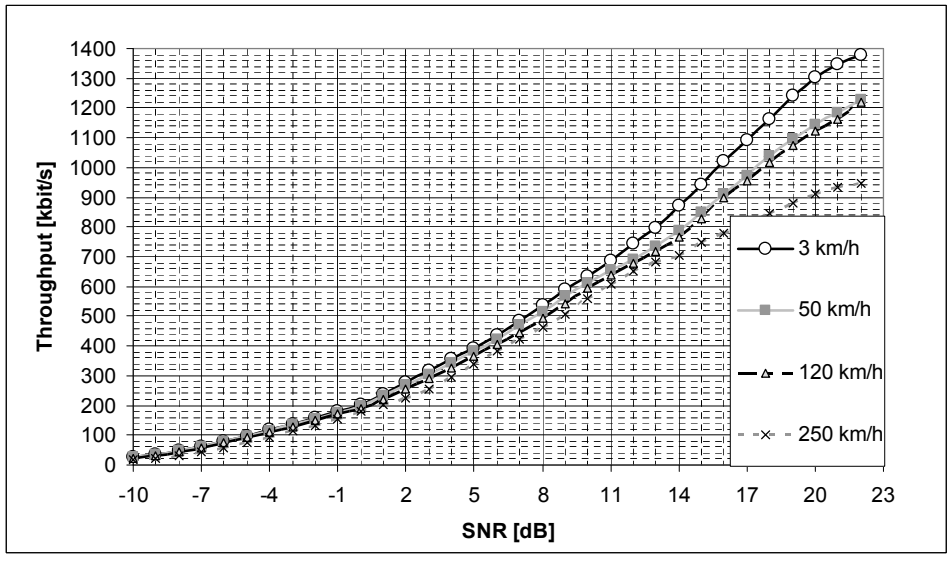


Figure 9.4 LTE eNodeB throughput as a function of SNR with UE speeds of 3 km/h, 50 km/h, 120 km/h and 250 km/h

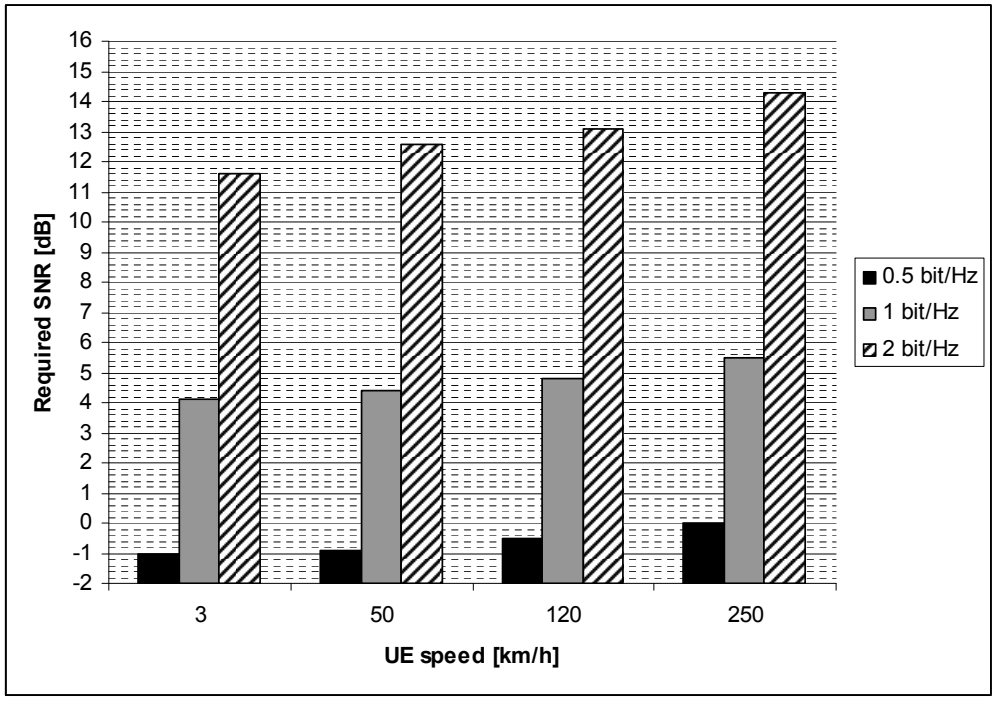


Figure 9.5 Required SNR values for various spectral efficiencies with UE speeds of 3 km/h, 50 km/h, 120 km/h and 250 km/h

9.5 Link Budgets

The link budget calculations estimate the maximum allowed signal attenuation, called path loss, between the mobile and the base station antenna. The maximum path loss allows the maximum cell range to be estimated with a suitable propagation model, such as Okumura–Hata. The cell range gives the number of base station sites required to cover the target geographical area. The link budget calculation can also be used to compare the relative coverage of the different systems. The relative link budget indicates how well the new LTE radio system will perform when it is deployed on the existing base station sites that are designed for GSM and WCDMA.

The parameters for the LTE link budget for uplink are introduced in Table 9.7, and the link budget is presented in Table 9.8. The corresponding LTE downlink link budget parameters are presented in Table 9.9, and the link budget in Table 9.10. The link budgets are calculated for 64 kbps uplink with 2-antenna base station receive diversity and 1 Mbps downlink with 2-antenna mobile receive diversity. For reference, the link budgets for GSM voice and for HSPA data are illustrated as well.

The LTE link budget in downlink has several similarities with HSPA and the maximum path loss is similar. The uplink part has some differences: smaller interference margin in LTE, no macro diversity gain in LTE and no fast fading margin in LTE. The maximum path loss values are summarized in Figure 9.6. The link budgets show that LTE can be deployed using existing GSM and HSPA sites assuming that the same frequency is used for LTE as for GSM and HSPA. LTE itself does not provide any major boost in the coverage. That is because the transmission power levels and the RF noise figures are also similar in GSM and HSPA technologies, and the link performance at low data rates is not much different in LTE than in HSPA.

The link budget was calculated for 64 kbps uplink, which is likely not a high enough data rate for true broadband service. If we want to guarantee higher data rates in LTE, we may need low frequency deployment, additional sites, active antenna solutions or local area solutions.

The coverage can be boosted by using a lower frequency since a low frequency propagates better than a higher frequency. The benefit of the lower frequency depends on the environment and on its use. The difference between 900 MHz and 2600 MHz is illustrated in Table 9.11. Part of the benefit of the lower frequency is lost since the antenna gains tend to get smaller at a lower frequency band. To maintain the antenna gain at lower frequency would require a physically larger antenna which is not always feasible at base station sites and in small terminals. The greatest benefits from low frequencies can be obtained when the base station site can use large antennas 2.5 m high and where the external antenna can be used in the terminal. This is fixed wireless deployment.

Example cell ranges are shown in Figure 9.7. Note that the *y*-axis has a logarithmic scale. The cell range is shown for 900 MHz, 1800 MHz, 2100 MHz and 2600 MHz frequency variants. These frequencies do not cover all possible LTE frequency variants. Typical US frequencies would be 700 MHz and 1700/2100 MHz, but the end result would be very similar to the values in Figure 9.7. The cell ranges are calculated using the Okumura–Hata propagation model with the parameters shown in Table 9.12. The urban cell range varies from 0.6 km to 1.4 km and suburban from 1.5 km to 3.4 km. Such cell ranges are also typically found in existing GSM and UMTS networks. The rural case shows clearly higher cell ranges: 26 km for the outdoor mobile coverage and even up to 50 km for the rural fixed installation at 900 MHz.

Note that the earth's curvature limits the maximum cell range to approximately 40 km with an 80 m high base station antenna assuming that the terminal is at ground level. The maximum

Table 9.7 Uplink link budget parameters for LTE

Field	Description	Typical value
a	UE maximum transmission power for power class 3. Different power classes would have different power levels. The power can be reduced depending on the modulation, see Chapter 11 for details.	23 dBm
b	UE antenna gain depends on the type of device and on the frequency band. Small handheld terminal at a low frequency band (like Band VIII) can have an antenna gain of −5 dBi while a fixed wireless terminal with directive antenna can have a gain of up to 10 dBi.	−5 to 10 dBi
c	Body loss is typically included for voice link budget where the terminal is held close to the user's head.	3 to 5 dB for voice
d	Calculated as a + b − c	
e	Base station RF noise figure. Depends on the implementation design. The minimum performance requirement is approximately 5 dB but the practical performance can be better.	2 dB
f	Terminal noise can be calculated as k (Boltzmann constant) × T (290K) × bandwidth. The bandwidth depends on bit rate, which defines the number of resource blocks. We assume two resource blocks for 64 kbps uplink.	−118.4 dBm for two resource blocks (360 kHz)
g	Calculated as e + f	
h	Signal-to-noise ratio from link simulations or measurements. The value depends on the modulation and coding schemes, which again depend on the data rate and on the number of resource blocks allocated.	−7 dB for 64 kbps and two resource blocks
i	Calculated as g + h	
j	Interference margin accounts for the increase in the terminal noise level caused by the interference from other users. Since LTE uplink is orthogonal, there is no intra-cell interference but we still need a margin for the other cell interference. The interference margin in practice depends heavily on the planned capacity – there is a tradeoff between capacity and coverage. The LTE interference margin can be smaller than in WCDMA/HSUPA where the intra-cell users are not orthogonal. In other words, the cell breathing will be smaller in LTE than in CDMA based systems.	1 to 10 dB
k	Cable loss between the base station antenna and the low noise amplifier. The cable loss value depends on the cable length, cable type and frequency band. Many installations today use RF heads where the RF parts are close to the antenna making the cable loss very small. The cable loss can also be compensated by using mast head amplifiers.	1 to 6 dB
l	Base station antenna gain depends on the antenna size and the number of sectors. Typical 3-sector antenna 1.3 m high at 2 GHz band gives 18 dBi gain. The same size antenna at 900 MHz gives smaller gain.	15 to 21 dBi for sectorized base station
m	Fast fading margin is typically used with WCDMA due to fast power control to allow headroom for the power control operation. LTE does not use fast power control and the fast fading margin is not necessary in LTE.	0 dB
n	Soft handover is not used in LTE	0 dB

Table 9.8 Uplink link budgets

	Uplink	GSM voice	HSPA	LTE
	Data rate (kbps)	12.2	64	64
	Transmitter – UE			
a	Max tx power (dBm)	33.0	23.0	23.0
b	Tx antenna gain (dBi)	0.0	0.0	0.0
c	Body loss (dB)	3.0	0.0	0.0
d	EIRP (dBm)	30.0	23.0	23.0
	Receiver – Node B			
e	Node B noise figure (dB)	–	2.0	2.0
f	Thermal noise (dB)	−119.7	−108.2	−118.4
g	Receiver noise (dBm)	–	−106.2	−116.4
h	SINR (dB)	–	−17.3	−7.0
i	Receiver sensitivity	−114.0	−123.4	−123.4
j	Interference margin (dB)	0.0	3.0	1.0
k	Cable loss (dB)	0.0	0.0	0.0
l	Rx antenna gain (dBi)	18.0	18.0	18.0
m	Fast fade margin (dB)	0.0	1.8	0.0
n	Soft handover gain (dB)	0.0	2.0	0.0
	Maximum path loss	162.0	161.6	163.4

cell range can be calculated with Equation 9.3, where R is the effective earth radius of 8650 km, h is the base station antenna height and d is the distance from the terminal to the base station. The calculation is illustrated in Figure 9.8. To achieve 100 km cell range, the required antenna height is 580 m, which in practice is feasible if the base station antenna is located on a mountain pointing towards the sea or other flat terrain.

$$\begin{aligned} R^2 + d^2 &= (R+h)^2 \\ \mathrm{h} &= \sqrt{R^2 + d^2} - R \end{aligned} \tag{9.3}$$

9.6 Spectral Efficiency

This section presents the typical spectral efficiency of LTE downlink and uplink in terms of bps/Hz/cell. The impact of the radio resource management algorithms on performance is discussed in Chapter 8.

9.6.1 System Deployment Scenarios

The specified LTA evaluation and test scenarios correspond to various radio propagation conditions and environments, similarly to HSPA [4]. The assumed macro and micro scenarios are shown in Table 9.13 and Table 9.14, and both assume a hexagonal-grid cellular layout.

Table 9.9 Downlink link budget parameters for LTE

Field	Description	Typical value
a	Base station maximum transmission power. A typical value for macro cell base station is 20–60 W at the antenna connector.	43–48 dBm
b	Base station antenna gain. See uplink link budget.	
c	Cable loss between the base station antenna connector and the antenna. The cable loss value depends on the cable length, cable thickness and frequency band. Many installations today use RF heads where the power amplifiers are close to the antenna making the cable loss very small.	1–6 dB
d	Calculated as A + B – C	
e	UE RF noise figure. Depends on the frequency band, Duplex separation and on the allocated bandwidth. For details, see Chapter 12.	6–11 dB
f	Terminal noise can be calculated as k (Boltzmann constant) × T (290K) × bandwidth. The bandwidth depends on bit rate, which defines the number of resource blocks. We assume 50 resource blocks, equal to 9 MHz, transmission for 1 Mbps downlink.	–104.5 dBm for 50 resource blocks (9 MHz)
g	Calculated as E + F	
h	Signal-to-noise ratio from link simulations or measurements. The value depends on the modulation and coding schemes, which again depend on the data rate and on the number of resource blocks allocated.	–9 dB for 1 Mbps and 50 resource blocks
i	Calculated as G + H	
j	Interference margin accounts for the increase in the terminal noise level caused by the other cell. If we assume a minimum G-factor of –4 dB, that corresponds to 10*log10(1+10^(4/10)) = 5.5 dB interference margin.	3–8 dB
k	Control channel overhead includes the overhead from reference signals, PBCH, PDCCH and PHICH.	10–25% = 0.4–1.0 dB
l	UE antenna gain. See uplink link budget.	-5–10 dBi
m	Body loss. See uplink link budget.	3.5 dB for voice

The typical evaluation scenarios used for LTE are macro case 1 and macro case 3 with 10 MHz bandwidth and low UE mobility. The propagation models for macro cell scenario are based on the Okamura–Hata model presented in section 9.5.

The micro cell scenarios assume an outdoor base station location and an outdoor and/or indoor mobile station location. The propagation model for these micro cell scenarios is similar to the macro cell models, but with different parameter settings, specific to the dense urban deployment scenario. The assumed micro cell base station antenna configuration is omni-directional (one-sector) as opposed to the directional (three sectors) deployments for macro cell scenarios. For the micro cell outdoor-to-outdoor scenario a dual-slope distance dependent path loss model is used in order to account for the Line-of-Sight (LOS) and Non-Line-of-Sight (NLOS) propagation conditions. For the micro cell outdoor-to-indoor scenario a single-slope distance dependent path loss model is used with a higher path loss exponent compared to the macro cell Okamura–Hata and micro cell outdoor-to-outdoor models. Typically, for system evaluation both outdoor and indoor mobile stations are assumed simultaneously with a 50:50 distribution ratio.

Table 9.10 Downlink link budgets

	Downlink	GSM voice	HSPA	LTE
	Data rate (kbps)	12.2	1024	1024
	Transmitter – Node B			
a	Tx power (dBm)	44.5	46.0	46.0
b	Tx antenna gain (dBi)	18.0	18.0	18.0
c	Cable loss (dB)	2.0	2.0	2.0
d	EIRP (dBm)	60.5	62.0	62.0
	Receiver – UE			
e	UE noise figure (dB)	–	7.0	7.0
f	Thermal noise (dB)	−119.7	−108.2	−104.5
g	Receiver noise floor (dBm)	–	−101.2	−97.5
h	SINR (dB)	–	−5.2	−9.0
i	Receiver sensitivity (dBm)	−104.0	−106.4	−106.5
j	Interference margin (dB)	0.0	4.0	4.0
k	Control channel overhead (%)	0.0	20.0	20.0
l	Rx antenna gain (dBi)	0.0	0.0	0.0
m	Body loss (dB)	3.0	0.0	0.0
	Maximum path loss	161.5	163.4	163.5

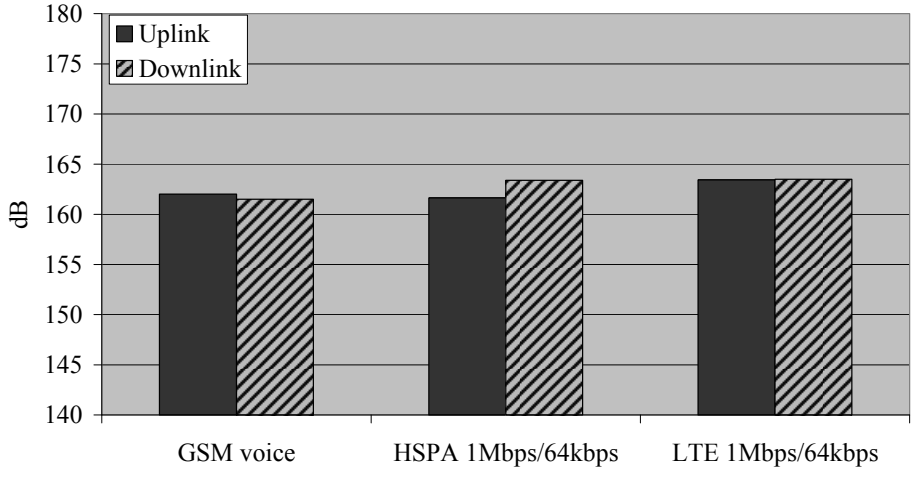


Figure 9.6 Maximum path loss values for GSM voice and HSPA and LTE data

The differences between the reference system deployment cases, as expected, yield different system performances. Generally, the macro cell scenarios provide the baseline LTE system performance numbers, while the special deployment cases, such as the micro cell

Table 9.11 Benefit of 900 MHz compared to 2600 MHz

	Urban	Rural	Rural fixed wireless
Propagation loss[a]	+14 dB	+14 dB	+14 dB
BTS antenna gain	−3 dB[b]	0 dB[c]	0 dB[c]
BTS cable loss[d]	+1 dB	+3 dB	+3 dB
UE antenna gain	−5 dB[e]	−5 dB[e]	0 dB[f]
UE sensitivity[g]	−1 dB	−1 dB	−1 dB
Total	+6 dB	+11 dB	+16 dB

[a]According to Okumura–Hata.
[b]Shared 1.3 m antenna for 900/2600 giving 15 and 18 dBi gain at 900 vs 2600 MHz.
[c]2.5 m antenna giving 18 dBi gain at 900 MHz.
[d]Cable 1/2'. Urban 30 m and rural 100 m.
[e]Based on 3GPP RAN4 contributions.
[f]External fixed antenna assumed.
[g]UE sensitivity can be up to 3 dB worse at 900 MHz but the difference in practice is less.

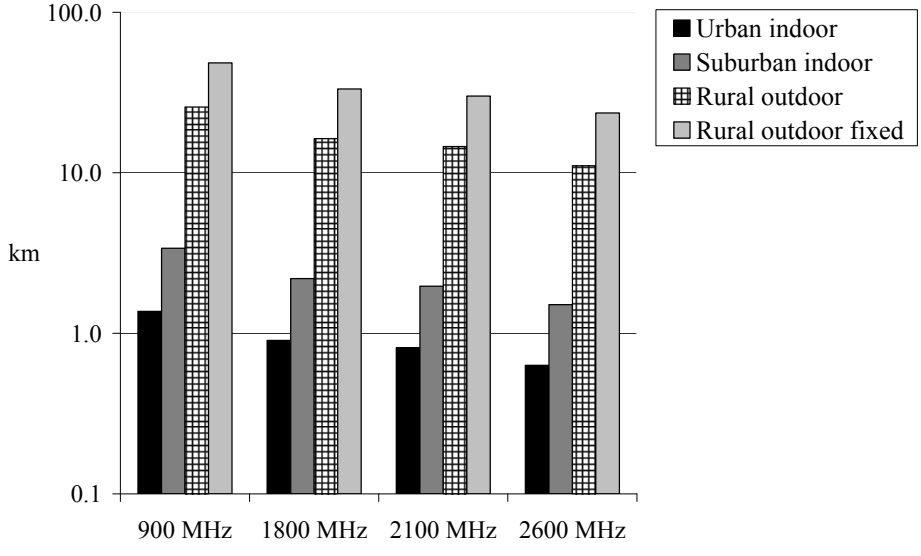


Figure 9.7 Cell ranges with Okumura–Hata propagation model

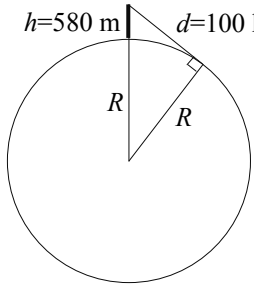
scenario, are used for the evaluation of adaptive spatial multiplexing MIMO enhancement techniques, which are able to better exploit the high Signal-to-Interference-and-Noise (SINR) conditions.

The results with Spatial Channel Model in Urban macro (SCM-C) radio channels and with a mixed velocity are also shown, as defined in reference [5].

The main simulation assumptions are shown in Table 9.15.

Table 9.12 Parameters for Okumura–Hata propagation model

	Urban indoor	Suburban indoor	Rural outdoor	Rural outdoor fixed
Base station antenna height (m)	30	50	80	80
Mobile antenna height (m)	1.5	1.5	1.5	5
Mobile antenna gain (dBi)	0.0	0.0	0.0	0.0
Slow fading standard deviation (dB)	8.0	8.0	8.0	8.0
Location probability	95%	95%	95%	95%
Correction factor (dB)	0	−5	−15	−15
Indoor loss (dB)	20	15	0	0
Slow fading margin (dB)	8.8	8.8	8.8	8.8
Max path loss at 1800/2100/2600 (dB)	163	163	163	163
Max path loss at 900 (dB)	160	160	160	160



R = 8650 km

(Effective earth radius for radio propagation is 4/3 larger than real radius)

Figure 9.8 Maximum cell range limitation caused by earth curvature

Table 9.13 LTE reference system deployment cases – macro cell scenarios

Simulation cases (macro)	Frequency (GHz)	Inter-site distance (m)	Bandwidth (MHz)	UE Speed (kmph)	Penetration loss (dB)
1	2.0	500	10	3	20
3	2.0	1732	10	3	20

Table 9.14 LTE reference system deployment cases – micro cell scenarios. The indoor:outdoor UE location ratio used is 50:50

Simulation cases (micro)	Frequency (GHz)	Inter-site distance (m)	Bandwidth (MHz)	UE Speed (kmph)	Penetration loss (dB)
Outdoor-to-outdoor	2.0	130	10	3	0
Outdoor-to-indoor	2.0	130	10	3	Included in the path loss model

9.6.2 Downlink System Performance

As a typical comparison between the different deployment environments, Figure 9.9 shows the distribution of the average wideband channel SINR also known as Geometry factor or G-factor.

Table 9.15 Simulation assumptions for simulation scenarios

Description	Settings
Number of active UEs per cell	10 UEs, uniformly distributed
Bandwidth	10 MHz
Channel	3GPP TU at 3 km/h (6 tap) SCM-C at 3 km/h and with correlations for spatial channels ITU modified Pedestrian B and Vehicular A with correlation Mixed velocity case: – 60% of UEs with ITU modified Pedestrian B at 3 km/h, – 30% of UEs with ITU modified Vehicular A at 30 km/h and – 10% of UEs with ITU modified Vehicular A at 120 km/h
Cell selection	Best cell selected with 0 dB margin
Transmission power	Uplink: Max 24 dBm (latest 3GPP specifications use 23 dBm)
Antenna configuration	Downlink: 1×2, 2×2 Uplink: 1×2, 1×4
Receiver algorithms	Downlink: LMMSE (Linear Minimum Mean Square Error) receiver Uplink: LMMSE receiver with Maximal Ratio Combining

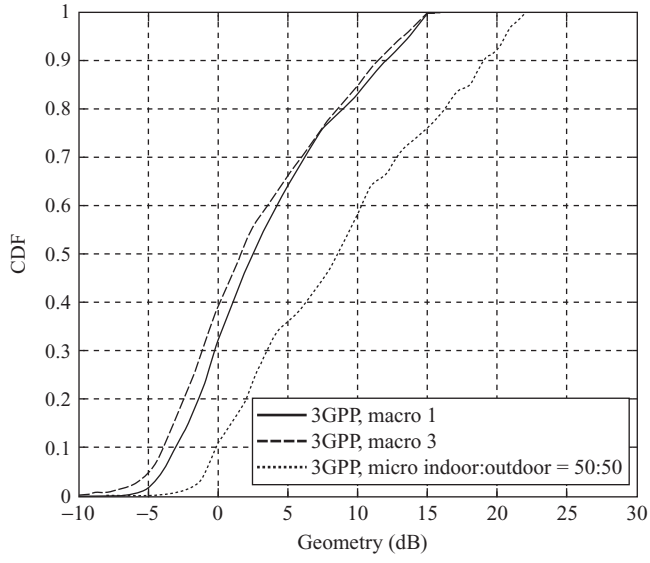


Figure 9.9 Distribution of the average wide-band channel SINR (geometry factor) macro case 1, case 3 and micro

The corresponding downlink system performance numbers for LTE SIMO and 2×2 MIMO transmission schemes are presented in Figure 9.10 and Figure 9.11, respectively.

Macro scenarios have median G-factors of 2–3 dB while the micro scenario has a G-factor above 8 dB. The macro cell has practically no values above 15 dB while the micro cell has 25% of the samples above 15 dB. The higher G-factor in micro cells turns into a higher throughput compared to the macro cells.

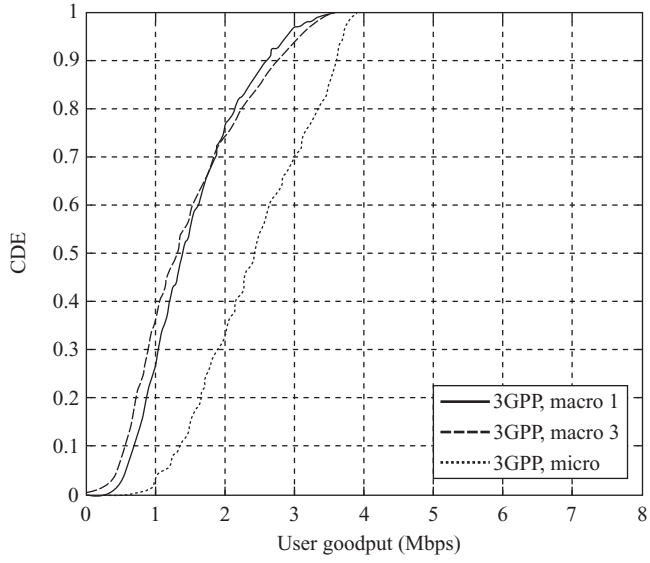


Figure 9.10 Downlink user throughput distribution simulation results in macro case 1, case 3 and micro test deployment scenarios for LTE 1 × 2 SIMO transmission scheme

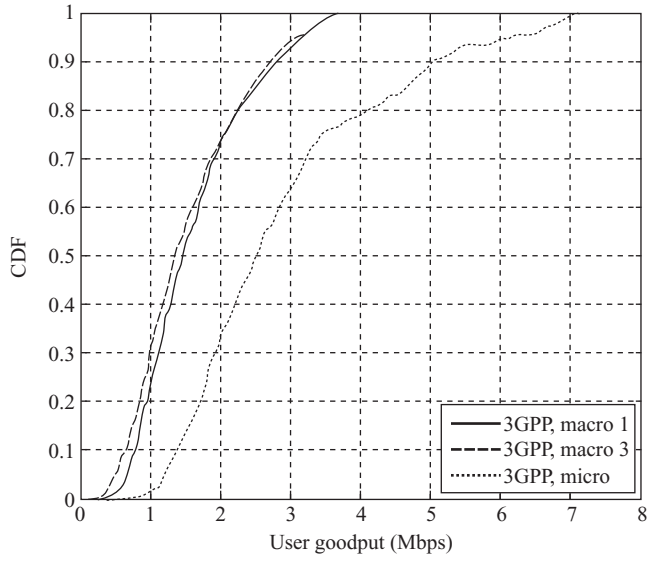


Figure 9.11 Downlink user throughput distribution simulation results in the macro case 1, case 3 and micro test deployment scenarios for LTE 2 × 2 MIMO transmission scheme

Figure 9.10 shows the user throughput with 10 users without MIMO. The median macro cell throughput is 1.4 Mbps while the micro cell provides 2.5 Mbps. The maximum values are 3–4 Mbps. Figure 9.11 presents the same case but with 2 × 2 MIMO. The median data rates are similar to the non-MIMO case but the highest data rates in micro cells are increased from 4 Mbps to 7 Mbps.

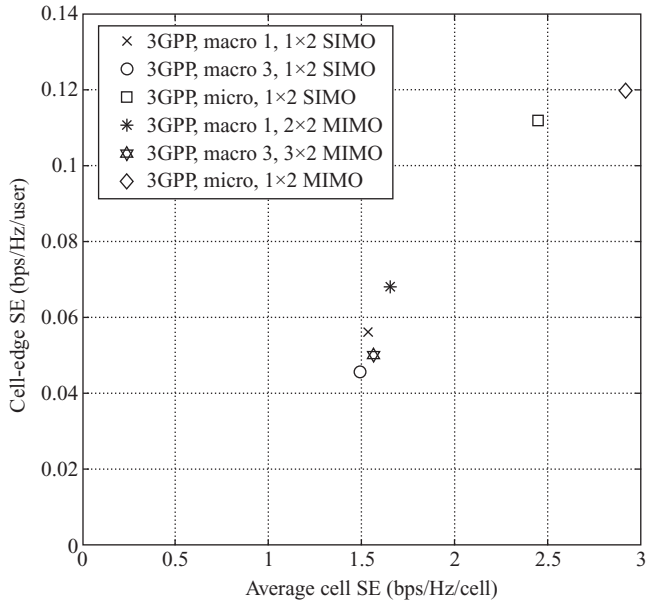


Figure 9.12 Downlink average cell and cell-edge spectral efficiency simulation results in the macro case 1, case 3 and micro test deployment scenarios for LTE 1×2 SIMO and 2×2 MIMO transmission scheme

Figure 9.12 shows the summary of the average cell spectral efficiency (SE) and the cell edge efficiency. Micro cell scenarios show considerably higher efficiency than macro scenarios. MIMO increases the average efficiency in micro cells but provides only marginal gains in the macro cell environment.

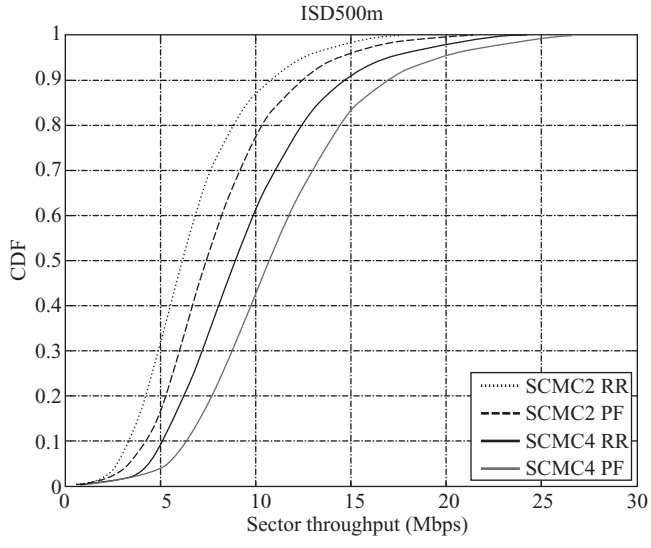
9.6.3 Uplink System Performance

Uplink baseline results with Best Effort data are shown for a two and four antenna receiver with Round Robin (RR) and Proportional Fair (PF) schedulers. In these results Interference over Thermal (IoT) probability at 20 dB was limited to 5%.

Figure 9.13 shows cell throughput for two types of scenarios for Inter-Site Distances (ISD) of 500 m and 1500 m. The cell throughput is presented with full Sounding Reference Signal (SRS) usage, thus giving a lower bound of system performance. From these results the PF scheduler gives about a 15–20% gain for mean cell throughput over the RR scheduler. Most of this gain comes from using SRS for better frequency domain scheduling. With higher UE speeds (e.g. 120 km/h) this gain will decrease towards the corresponding RR scheduler results. Furthermore, the cell edge throughput with the same cell mean throughput is about 125% higher in ISD 500 m than in a larger cell size. UE power limitation is seen at ISD 1500 m with lower cell edge performance. Note that the impact of the cell size is larger in uplink than in downlink due to the lower UE power compared to the eNodeB transmission power.

The 95% limit for IoT distribution is set to 20 dB. Smaller cell sizes like ISD 500 m are affected by this constraint because higher IoT values would be possible from the UE power

(a)



(b)

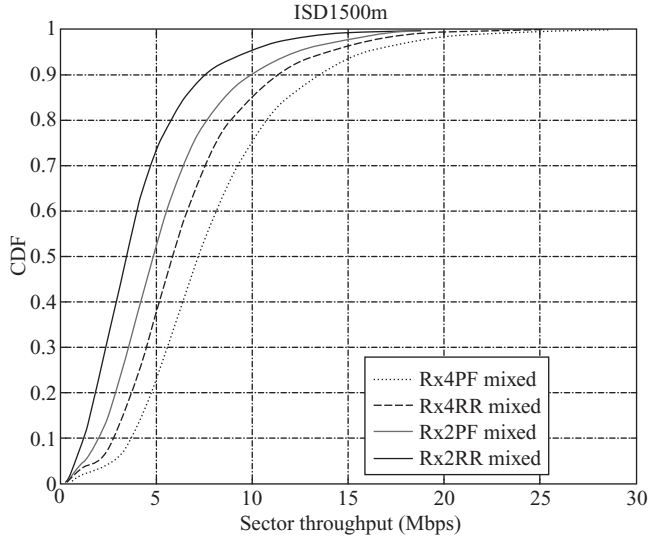


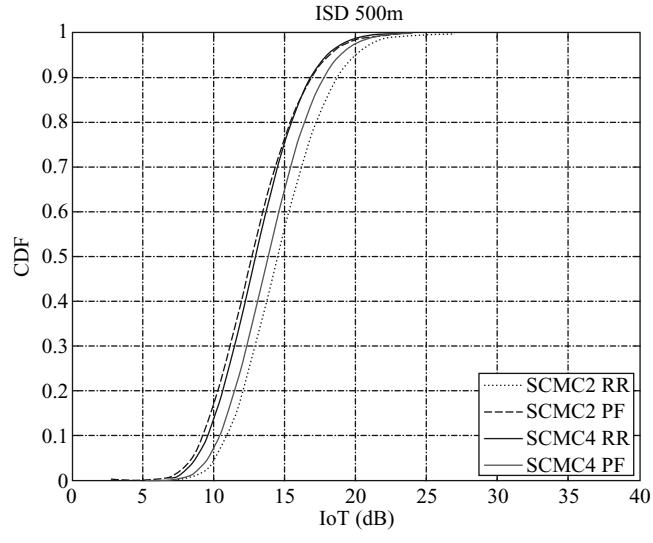
Figure 9.13 Uplink cell throughput for (a) ISD 500 m; (b) ISD 1500 m

point of view. The mean IoT is 13 dB in Figure 9.14. For a larger cell size (like ISD 1500m) IoT is not a limiting factor and the mean IoT is 2 dB.

The uplink power control range in ISD 500 m is seen in Figure 9.15. The power control range can be used well at a smaller cell size. The median UE transmission power is 15 dBm. This power is relatively high compared to voice users, but it is expected for the data connection. In a larger cell a higher proportion of users are limited by maximum power, thus moving the remaining distribution towards the right side.

Some spectral efficiency figures are collected into Table 9.16 for comparison with HSUPA. From these figures we observe that uplink performance is doubled from the cell mean spectral

(a)



(b)

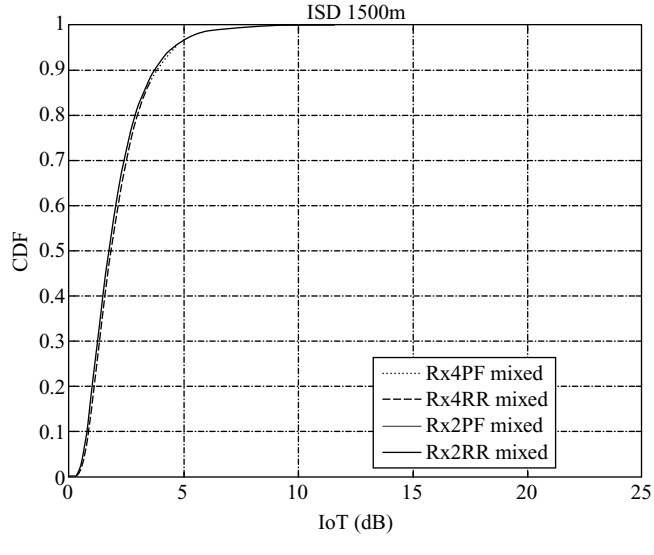


Figure 9.14 Uplink instantaneous Noise Raise for (a) ISD 500 m; (b) ISD 1500 m

efficiency point of view with RR scheduler and 2-branch reception. Furthermore, at the same time cell edge performance is tripled. With PF and 4-branch reception, the gain over HSUPA can be further improved.

The average spectral efficiency in different channel models and cell sizes is summarized in Figure 9.16 and the cell edge efficiency in Figure 9.17. The efficiency is higher in small cells than in large cells since the UE runs out of power in large cells. The impact of the cell size is most clearly visible in cell edge efficiency. We can also see that 4-branch reception and PF scheduling provide higher capacity compared to 2-branch reception and RR scheduling.

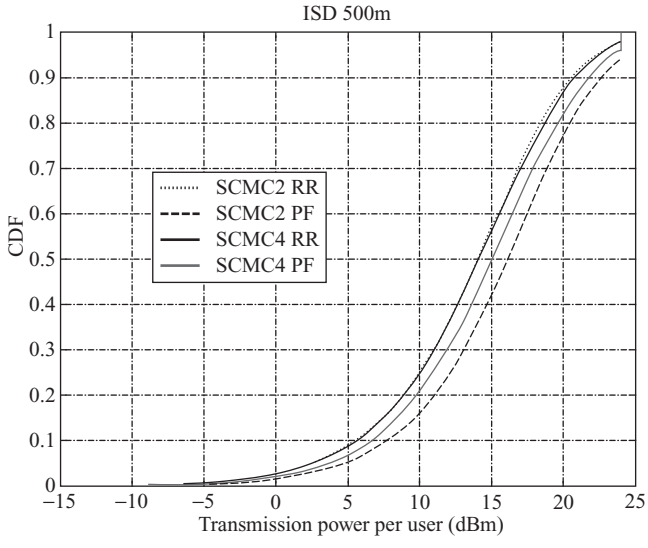


Figure 9.15 Uplink transmission power CDF per user in case of ISD 500 m

Table 9.16 Uplink spectral efficiency at the mean and at the cell edge for various multi-antenna configurations, FDE/MRC receivers and RR/PF schedulers

Best effort traffic		
Features	Spectral efficiency cell mean (b/s/Hz/cell)	Spectral efficiency cell edge (5%-ile) (b/s/Hz/cell)
HSUPA	0.33	0.009
1×2 RR	0.66	0.027
1×2 PF	0.79	0.033
1×4 RR	0.95	0.044
1×4 PF	1.12	0.053

9.6.4 Multi-antenna MIMO Evolution Beyond 2 × 2

The downlink LTE MIMO transmission schemes are specified to support up to a four transmit antenna port configuration. A brief summary of these schemes is given in Table 9.17, along with the employed terminology.

The MIMO rank identifies the number of spatial layers (streams) while the MIMO codeword is used to jointly encode up to two spatial layers. In this context the closed-loop MIMO operation is implemented using codebook based pre-coding, i.e. the antenna weights (vectors or matrices) to be applied at the base station side are selected from a fixed number of possibilities, which form a quantized codebook set. These codebook sizes are summarized in Table 9.17. Furthermore, for all the single-user spatial multiplexing MIMO schemes it is assumed that the multiplexed spatial layers are transmitted to the same mobile station.

In a practical system, to ensure proper performance of these MIMO schemes, the overall radio Layer 1 (L1) and Layer 2 (L2) resource management has to include a minimum set of

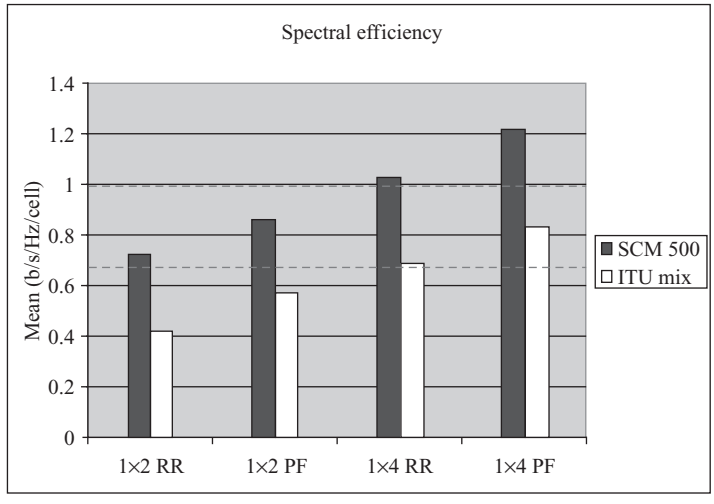


Figure 9.16 Uplink mean spectral efficiency

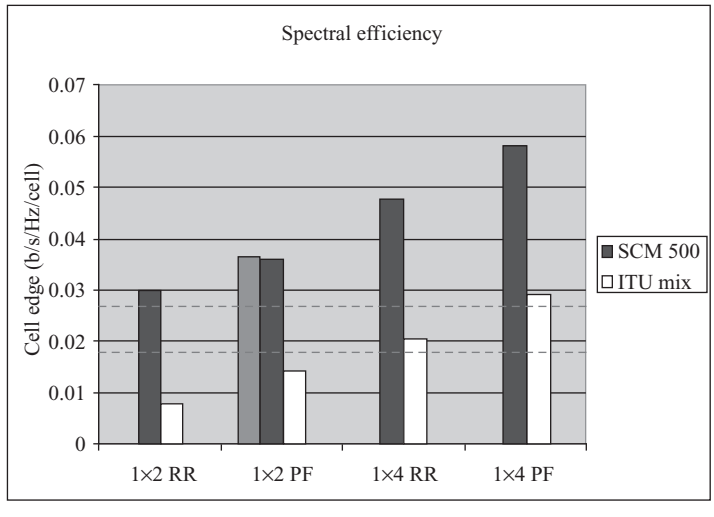


Figure 9.17 Uplink cell edge spectral efficiency

Table 9.17 LTE specification for downlink single-user MIMO transmission schemes. MIMO precoding codebook size and codeword transmission modes

MIMO scheme	Rank 1 (1 layer)	Rank 2 (2 layers)	Rank 3 (3 layers)	Rank 4 (4 layers)	Codeword transmission mode
2×1	4 vectors	No	No	No	Single codeword
2×2	4 vectors	3 matrices	No	No	Rank 2 with dual codeword
4×2	16 vectors	16 matrices	No	No	Rank 2 with dual codeword
4×4	16 vectors	16 matrices	16 matrices	16 matrices	Rank 2–4 with dual codeword

MIMO-aware mechanisms at both eNodeB and UE side. The minimum set of required feedback information from the UE to support downlink adaptive spatial multiplexing MIMO transmission comprises:

- channel state/quality information (CQI): direct (e.g. SINR) or indirect (e.g. MCS) information on the average channel conditions estimated on the physical resources and MIMO codewords;
- MIMO rank information: indicates the optimum number of spatial layers (streams) to be used and the corresponding pre-coding vector/matrix index (PMI);
- HARQ information: indicates the reception status (ACKnowledged/Not ACKnowledged) of the transmitted data packets on the scheduled MIMO spatial layers.

The first two items above, CQI and PMI, are estimated by the mobile station based on the downlink pilot measurements and are transmitted back to the base station in a quantized form. The HARQ mechanism has an impact on the freedom of the radio resource management blocks, when link adaptation and optimized packet scheduling has to be performed for single/dual codeword MIMO transmission.

The mobile station feedback information is heavily dependent on the channel/signal estimation performance, which has an upper bound due to the practical system design limitations (reference signals, measurement/estimation time, propagation delays, etc.). Furthermore, all thise feedback information requires uplink control channel capacity; thus it is desirable to minimize the actual number of information bits used to encode them and converge to a practical tradeoff between the achievable MIMO performance and the required uplink capacity.

The system level performance degradation due to limited and noisy feedback from the terminals has been studied and presented in the literature under practical LTE-like assumptions, e.g. [6, 7]. In the following the main concepts related to these performance tradeoffs in simple reference MIMO configuration cases are summarized. The radio resource management aspects are discussed in more detail in Chapter 8.

In theory, the optimal link adaptation operation with spatial-multiplexing MIMO schemes requires separate CQI measures estimated for each separately encoded spatial layer. Without significant loss of performance the CQI feedback has been reduced by grouping the spatial layers into two groups with a maximum of two layers each, so that two CQI measures are sufficient for the considered MIMO transmission schemes. Each group of spatial layers is separately encoded and corresponds to one codeword (see Table 9.17).

The optimized operation of the adaptive MIMO transmission schemes requires, in addition to the classical link adaptation, a MIMO rank adaptation mechanism. A straightforward and simple scheme is with the optimum MIMO rank selected at the terminal side such that it maximizes the instantaneous downlink user throughput. The optimum pre-coding vector/matrix can be determined based on various criteria: maximum beam-forming gain, maximum SINR or maximum throughput. For simplicity, here we use the first criterion because this implies the computation of signal power levels only. The latter options are more complex and require information about the number of allocated physical resources blocks, the interference levels and the modulation and coding used.

Based on the instantaneous channel conditions the terminal estimates the optimum MIMO rank, and the corresponding optimum pre-coding vector/matrix. Ideally, both the MIMO rank and the PMI should be estimated for each physical resource block, i.e. assuming frequency-selective MIMO adaptation. For large system bandwidths of 5 MHz and above, in practice this approach leads to a

very large signaling overhead introduced in the uplink control channel. A good compromise can be achieved by using MIMO feedback per group of several consecutive physical resource blocks. This approach is further motivated by the choice made for the CQI feedback in the LTE systems. The CQI feedback is described in Chapter 5. The MIMO feedback information is assumed to be transmitted to the serving base station together with, or embedded in, the CQI reports.

In the context of the LTE system, because the same modulation and coding scheme is used on all the allocated physical resource blocks, it is also reasonable to assume that the same MIMO rank can be used for all the allocated physical resource blocks. Furthermore, given that the actual resource allocation is determined by the packet scheduler at the base station, and not at the terminal side, a practical solution is to use one single MIMO rank determined for the entire monitored system bandwidth, i.e. frequency non-selective MIMO rank information is fed back to the base station.

The minimal system-level restrictions described above, still allow for quite a large degree of freedom in the MIMO adaptation mechanisms, in both the time and frequency domains. Some relevant examples are presented in the following.

The frequency granularity of the PMI feedback can be different from the granularity of the MIMO rank. Two reference PMI selection schemes can be assumed combined with the above described full-bandwidth MIMO rank selection:

- Frequency Non-Selective (FNS) PMI: one PMI is determined for the entire effective system bandwidth;
- Frequency Selective (FS) PMI: one PMI is determined for each group of two consecutive physical resource blocks, i.e. using the same $2 \times$PRB granularity as the CQI measures.

Using the MIMO and CQI feedback information from all the active terminals in the serving cell, the RRM algorithm in the serving base station performs the actual resource allocation and scheduling for each terminal. Previous investigations with low mobility scenarios have disclosed the influence on the overall system performance of the rate at which the MIMO adaptation is performed. The results show that a semi-adaptive scheme with slow update rate (~100 ms) based on the average channel conditions yields only small cell throughput losses of the order of 5% compared to the case with a fast-adaptive scheme with MIMO rank selected in each scheduling period. Thus, there are two reference schemes:

- G-factor based (GF): with rank update based on the average wideband channel SINR conditions;
- quasi-dynamic (QD): with rank selected only when a new (1st HARQ) transmission is scheduled.

The quasi-dynamic MIMO scheme is a tradeoff solution between the fast-adaptive (per TTI) and the G-factor based (slow, per 5 ms to 10 ms) adaptation schemes, thus it can be successfully used in both medium and low mobility scenarios.

In combination with these two L1–L2 mechanisms – MIMO rank adaptation, and MIMO pre-coding information feedback frequency domain granularity – the downlink adaptive 2×2 and 4×4 MIMO transmission schemes have been simulated. These evaluations assume limited, imperfect and delayed channel feedback from the terminals. The PMI and rank information feedback is assumed to be error free, while the CQI information is modeled with errors introduced by the channel estimation.

Figure 9.18 and Figure 9.19 show representative simulation results for the spectral efficiency in the macro cell case 1 and micro cell deployment scenarios (see section 9.6.1), respectively. The number of mobile stations is 20 simulated per cell/sector. Time–frequency proportional fair packet scheduling is employed.

For both the 2×2 and 4×4 MIMO transmission schemes, the macro cell investigations show gains in the average cell spectral efficiency and the cell-edge spectral efficiency of the order of 10% and 18%, respectively, when using quasi-dynamic MIMO rank adaptation and frequency selective closed-loop information feedback compared to the G-factor based and frequency non-selective cases. The performance in micro cell scenarios is less impacted, with gains of up to 4% and 15%, respectively. The latter results confirm the suitability and high potential of the downlink MIMO schemes in microcellular environments without introducing significant overhead in the uplink signaling and control channels payload.

9.6.5 Higher Order Sectorization (Six Sectors)

One of the methods for increasing the performance by using more antennas at the base station is to use higher order sectorization. The use of higher order sectorization is especially considered to be an option for macro cell installations, where antennas are mounted above rooftops. Typically, three-sector sites are assumed in most of the LTE macro cell performance evaluations, by using three separate panel antennas per site, each with a 3 dB beamwidth of 65 or 70 degrees. A first step could therefore be to increase the sectorization to six sectors per site by simply using six panel antennas with a narrower beamwidth of, for example, 35 degrees. This provides a simple method for increasing the capacity of the network at the sites with high offered traffic. The six sector antennas could also be implemented with three panel antennas and digital fixed beamforming.

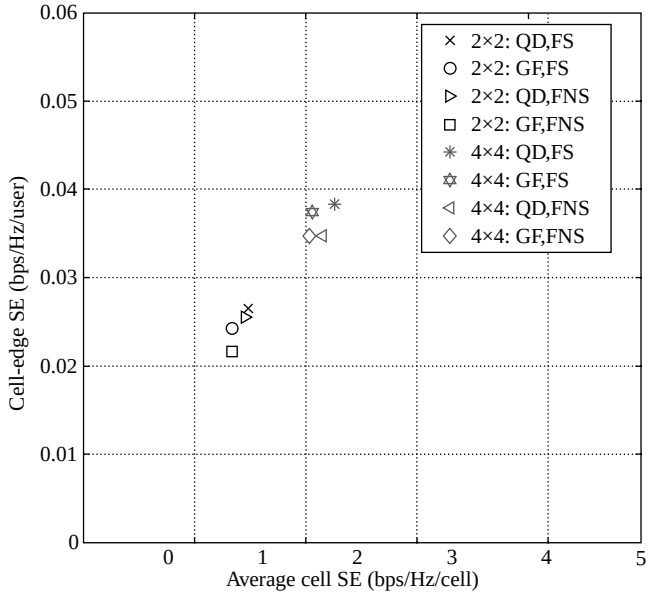


Figure 9.18 Average cell vs cell-edge user (coverage) spectral efficiency simulation results for MIMO schemes operating in the macro cell case 1 scenario

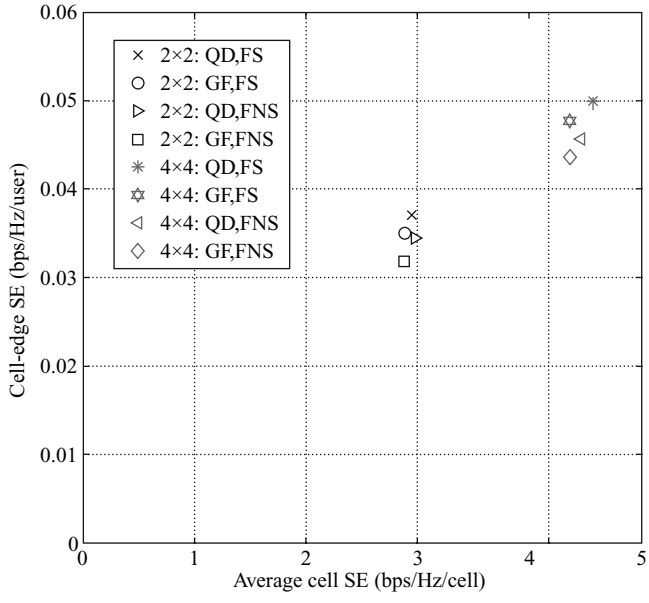


Figure 9.19 Average cell vs cell-edge user (coverage) spectral efficiency simulation results for MIMO schemes operating in the micro cell scenario

The performance improvements of six sectors vs three sectors have been evaluated by means of system simulations in a typical macro cell environment with 500 m site-to-site distance, i.e. similar to 3GPP macro cell case 1 with LTE 10 MHz bandwidth. The following four cases are considered:

- all sites have three sectors with antenna beamwidth of 70 degrees;.
- all sites have six sectors with antenna beamwidth of 35 degrees;
- only one centre site has six sectors, while other sites have three sectors (denoted Mode-1);
- a cluster of seven sites has six sectors, while all other sites have three sectors (denoted Mode-2).

Performance results for these four cases are presented in Figure 9.20, where the average throughput per site is reported. These results were obtained with a simple full buffer traffic model, assuming two antennas at the terminals, and standard proportional fair time-frequency packet scheduling. A significant performance gain is achieved by increasing the sectorization from three sectors to six sectors: the site capacity is increased by 88%. If only one site is upgraded to have six sectors (Mode-1), then the capacity of that particular site is increased by 98% compared to using three sectors. The latter result demonstrates that upgrading sites to using higher order sectorization is a possible solution for solving potential capacity bottlenecks that may arise in certain hotspot areas. Assuming that the number of users per site is the same, independent of whether three sectors or six sectors are used, then similar improvements are found in terms of user experienced throughput. However, the use of higher order sectorization does not increase the peak user throughput as is the case for the use of MIMO schemes. Similar

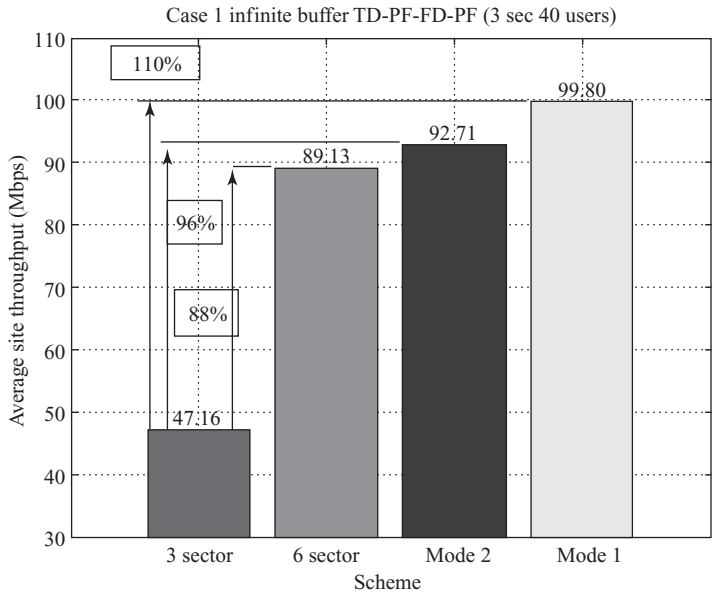


Figure 9.20 Average throughput per macro cell site for different sectorization configurations [8]. © 2008 IEEE

capacity gains of six sectors vs three sectors have also been reported in [9] for WCDMA networks. The simulated capacity has considered the downlink direction. It is expected that the gains are also similar in the uplink direction.

The impact of the beamwidth with a six-sector antenna was also studied and the results are shown in Figure 9.21. The motivation for presenting these results is that even though the raw antenna beamwidth is 35 degrees, the effective antenna beamwidth is larger due to the radio channels' azimuthal dispersion [10]. The azimuthal dispersion is found to be of the order of 5–10 degrees for typical urban macro cells. The simulation results show that the average site throughput with antenna beamwidth of 40 degrees is 1% lower compared to 35 degrees, and with 45 degrees it is 4% lower.

The six-sector deployment also improves the network coverage since the typical antenna gain is increased from 18 dBi in three-sector 70-degree antennas to 21 dBi in six-sector 35-degree antennas.

9.6.6 Spectral Efficiency as a Function of LTE Bandwidth

Most LTE system simulations assume 10 MHz bandwidth. The system bandwidth has some impact on the efficiency mainly due to the following factors:

- The frequency domain scheduling gain is higher with larger bandwidth since there is more freedom for the scheduler to optimize the transmission. For 1.4 MHz bandwidth the frequency domain scheduling gain is very low since the whole channel bandwidth likely has flat fading.
- The relative overhead from common channels and control channel – PBCH, Synchronization signal and PUCCH – is lower with larger bandwidths.

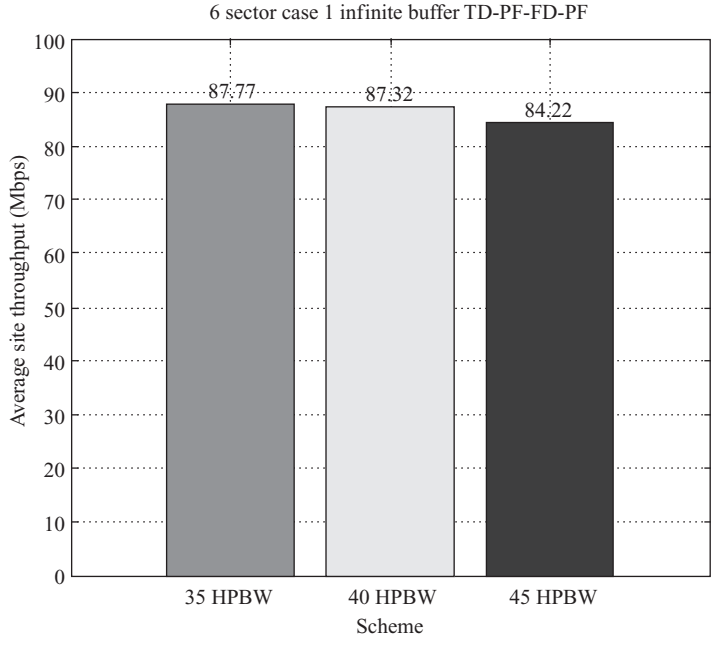


Figure 9.21 Comparison of site throughput for different degrees of azimuth spread for six-sector-site deployment. © 2008 IEEE

- The relative guard band is higher with a 1.4 MHz bandwidth: six resource blocks at 180 kHz equal to 1.08 MHz corresponding to 77% utilization compared to the channel spacing of 1.4 MHz. The utilization is 90% for other bandwidths.

The calculations for the relative efficiency of the bandwidths are shown in Table 9.18 for downlink and in Table 9.19 for uplink. The frequency domain scheduling gain depends on several factors including environment, mobile speed, services and antenna structure. We assume that the scheduling gain in uplink is half of the downlink gain for all bandwidths.

The results are summarized in Figure 9.22. The efficiency is quite similar for the bandwidths of 5 to 20 MHz while there is a 15% difference for 3 MHz. The efficiency with 1.4 MHz bandwidth is approximately 35–40% lower than for 10 MHz. The typical downlink cell throughput

Table 9.18 Relative efficiency of different LTE bandwidths in downlink

	LTE 1.4 MHz	LTE 3 MHz	LTE 5 MHz	LTE 10 MHz	LTE 20 MHz
Resource blocks	6	15	25	50	100
Guard band overhead	23%	10%	10%	10%	10%
BCH overhead	2.9%	1.1%	0.7%	0.3%	0.2%
SCH overhead	2.5%	1.0%	0.6%	0.3%	0.2%
Frequency domain scheduling gain	5%	20%	35%	40%	45%

Table 9.19 Relative efficiency of different LTE bandwidths in uplink

	LTE 1.4 MHz	LTE 3 MHz	LTE 5 MHz	LTE 10 MHz	LTE 20 MHz
Guard band overhead	23%	10%	10%	10%	10%
PUCCH overhead	16.7%	13.3%	8.0%	8.0%	8.0%
Frequency domain scheduling gain	2.5%	10%	17.5%	20%	22.5%

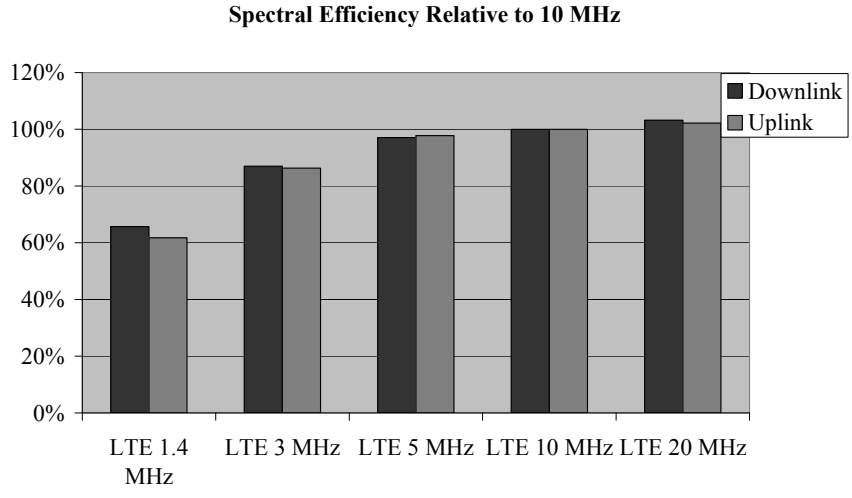


Figure 9.22 Relative spectral efficiency compared to 10 MHz bandwidth in macro cells

for 20 MHz is 1.74 bps/Hz/cell × 20 MHz = 35 Mbps, while for 1.4 MHz the cell throughput is 1.74 bps/Hz/cell × 60% x 1.4 MHz = 1.5 Mbps. The conclusion is that LTE should be deployed using as large a bandwidth as possible. The main motivation is to maximize the LTE data rates, but the second motivation is to optimize the spectral efficiency as well. The narrowband options are still useful for refarming purposes if it is not possible to allocate larger bandwidths initially. Even if LTE 1.4 MHz cell throughput of 1.5 Mbps seems relatively low, it is still more than what can be offered by narrowband 2G GPRS/EDGE networks.

9.6.7 Spectral Efficiency Evaluation in 3GPP

The LTE target was to provide spectral efficiency at least three times higher than HSDPA Release 6 in downlink and two times higher than HSUPA Release 6 in uplink. The relative efficiency was studied by system simulations in 3GPP during 2007. The results of the different company contributions are illustrated in Figure 9.23 for downlink and in Figure 9.24 for uplink. These results are for Case 1, which represents interference limited small urban macro cells. The average downlink efficiency increases from HSDPA 0.55 bps/Hz/cell to LTE 1.75 bps/Hz/cell, corresponding to three times higher efficiency. The uplink efficiencies are 0.33 for HSUPA and 0.75 for LTE, corresponding to more than two times higher efficiency. The performance

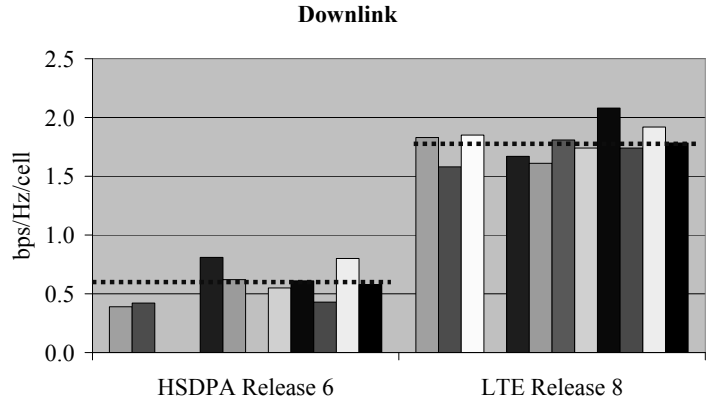


Figure 9.23 Downlink spectral efficiency results from different company contributions [11]

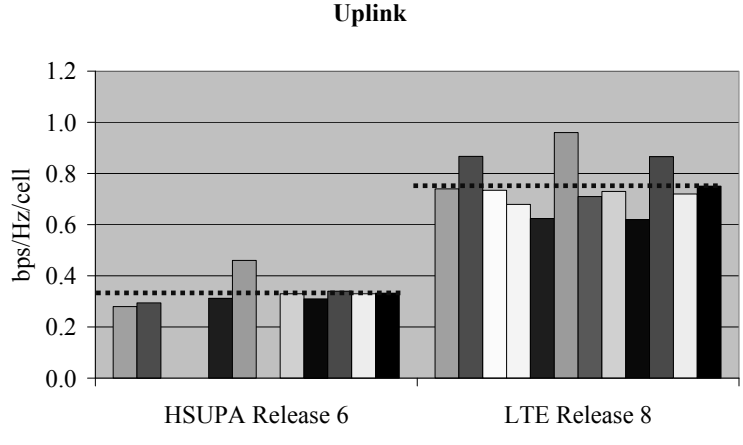


Figure 9.24 Uplink spectral efficiency results from different company contributions [12]

evaluation showed that LTE can fulfill those targets that have been defined for the system in the early phase.

HSPA Release 6 with two-antenna Rake receiver terminals was used as the reference case since it was the latest 3GPP release available at the time when LTE work started. HSPA Release 6 is based on 1×2 maximum ratio combining. LTE Release 8 uses 2×2 MIMO transmission.

9.6.8 Benchmarking LTE to HSPA

The spectral efficiency gain of LTE over HSDPA can be explained by a few characteristics in the LTE system design. Those factors are shown in Table 9.20 for the downlink. LTE uses orthogonal modulation to avoid intra-cell interference providing a major boost in capacity compared to HSDPA with Rake receiver. The CDMA transmission in HSDPA suffers from intra-cell interference caused by multi-path propagation. The CDMA codes are orthogonal but only in the single path channel. The LTE benefit over HSDPA depends on the amount of multi-path propagation.

Table 9.20 LTE downlink efficiency benefit over HSPA Release 6 in macro cells

LTE benefit	Gain	Explanation
OFDM with frequency domain equalization	Up to +70% depending on the multi-path profile	HSDPA suffers from intra-cell interference for the Rake receiver. Rake receiver is assumed in Release 6. However, most HSDPA terminals have an equalizer that removes most intra-cell interference.
Frequency domain packet scheduling	+40%	Frequency domain scheduling is possible in OFDM system, but not in single carrier HSDPA. The dual carrier HSDPA can get part of the frequency domain scheduling gain.
MIMO	+15%	No MIMO defined in HSDPA Release 6. The gain is relative to single antenna base station transmission. HSDPA Release 7 includes MIMO.
Inter-cell interference rejection combining	+10%	The interference rejection combining works better in OFDM system with long symbols.
Total difference	=3.0×	$1.7 \times 1.4 \times 1.15 \times 1.1$

Another major benefit in LTE is the frequency domain scheduling, which is not possible in a CDMA based system. The CDMA systems transmit the signal using the full bandwidth. MIMO provides some efficiency benefit in LTE since MIMO was not included in HSDPA Release 6. The inter-cell interference rejection combining is used in LTE simulation results and it gives some further capacity boost for LTE. Adding these factors together gives the result that LTE is indeed expected to provide approximately three times higher efficiency than HSDPA Release 6. The LTE performance gain over HSDPA Release 6 will be larger in small micro or indoor cells where the multistream MIMO transmission can be used more often for taking advantage of high LTE data rates.

The LTE uplink efficiency gain over HSUPA mainly comes from the orthogonal uplink while HSUPA suffers from intra-cell interference.

HSPA evolution brings some performance enhancements to HSPA capacity; for details see Chapter 13. MIMO is part of HSPA Release 7; terminal equalizers can remove intra-cell interference in downlink and base station interference cancellation can remove intra-cell interference in uplink. Dual carrier HSDPA transmission is included in Release 8, which enables some frequency domain scheduling gain. Even if all the latest enhancements from HSPA evolution are included, LTE still shows an efficiency benefit over HSPA.

The efficiencies from system simulations are illustrated in Figure 9.25. HSPA Release 7 includes here a UE equalizer and 2×2 MIMO in downlink and base station interference cancellation in uplink. HSPA Release 8 includes dual carrier HSDPA with frequency domain packet scheduling. HSPA Release 9 is expected to allow the combination of MIMO and DC-HSDPA.

9.7 Latency

9.7.1 User Plane Latency

User plane latency is relevant for the performance of many applications. There are several applications that do not require a very high data rate, but they do require very low latency. Such

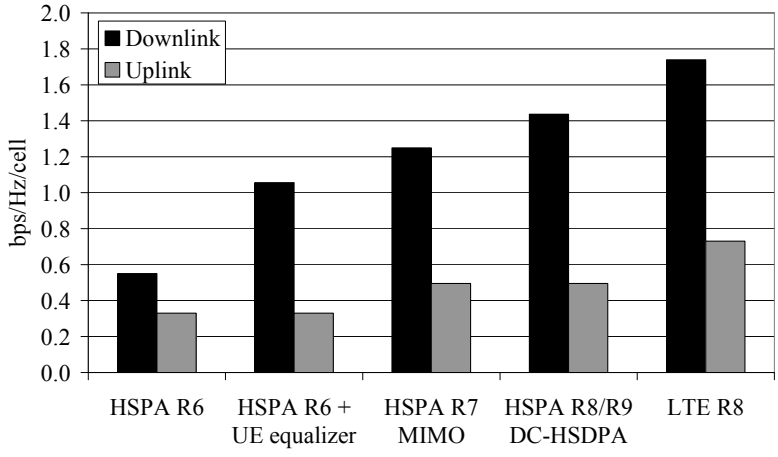


Figure 9.25 Spectral efficiency of HSPA and LTE

applications are, for example, voice, real time gaming and other interactive applications. The latency can be measured by the time it takes for a small IP packet to travel from the terminal through the network to the internet server, and back. That measure is called round trip time and is illustrated in Figure 9.26.

The end-to-end delay budget is calculated in Table 9.21 and illustrated in Figure 9.27. The 1-ms frame size allows a very low transmission time. On average, the packet needs to wait for 0.5 ms for the start of the next frame. The retransmissions take 8 ms at best and the assumed retransmission probability is 10%. The average delay for sending the scheduling request is

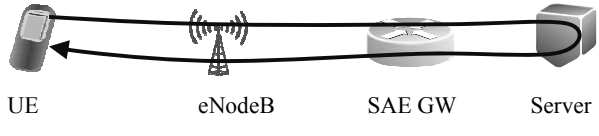


Figure 9.26 Round trip time measurement

Table 9.21 Latency components

Delay component	Delay value
Transmission time uplink + downlink	2 ms
Buffering time (0.5 × transmission time)	2 × 0.5 × 1 ms = 1 ms
Retransmissions 10%	2 × 0.1 × 8 ms = 1.6 ms
Uplink scheduling request	0.5 × 5 ms = 2.5 ms
Uplink scheduling grant	4 ms
UE delay estimated	4 ms
eNodeB delay estimated	4 ms
Core network	1 ms
Total delay with pre-allocated resources	13.6 ms
Total delay with scheduling	20.1 ms

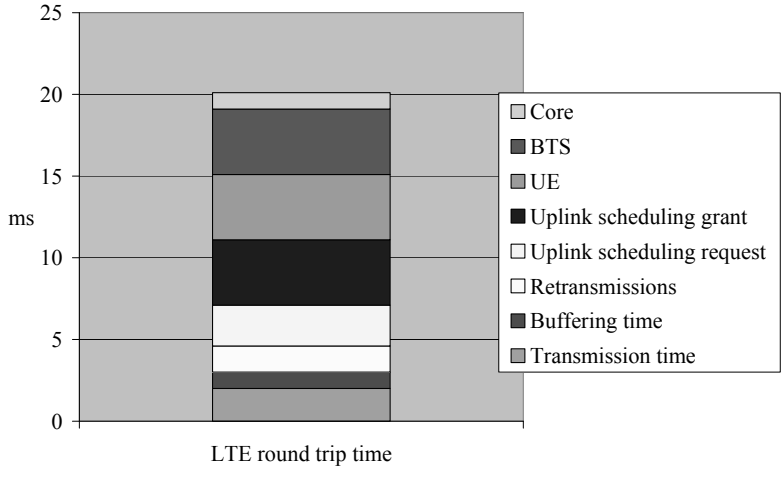


Figure 9.27 End-to-end round trip time including scheduling latency

2.5 ms and the scheduling grant 4 ms. We further assume a UE processing delay of 4 ms, an eNodeB processing delay of 4 ms and a core network delay of 1 ms.

The average round trip including retransmission can be clearly below 15 ms if there are pre-allocated resources. If the scheduling delay is included, the delay round trip time will be approximately 20 ms. These round trip time values are low enough even for applications with very tough delay requirements. The practical round trip time in the field may be higher if the transport delay is longer, or if the server is far away from the core network. Often the end-to-end round trip time can be dominated by non-radio delays, e.g. by the distance and by the other elements in the internet. The propagation time of 5000 km is more than 20 ms.

9.8 LTE Refarming to GSM Spectrum

LTE will be deployed in the existing GSM spectrum like 900 MHz or 1800 MHz. The flexible LTE bandwidth makes refarming easier than with WCDMA because LTE can start with 1.4 MHz or 3.0 MHz bandwidths and then grow later when the GSM traffic has decreased. The required separation of the LTE carrier to the closest GSM carrier is shown in Table 9.22. The required total spectrum for LTE can be calculated based on the carrier spacing. The coordinated

Table 9.22 Spectrum requirements for LTE refarming

	LTE–GSM carrier spacing		LTE total spectrum requirement	
	Coordinated	Uncoordinated	Coordinated	Uncoordinated
5 MHz LTE (25 RBs)	2.5 MHz	2.7 MHz	4.8 MHz	5.2 MHz
3 MHz LTE (125 RBs)	1.6 MHz	1.7 MHz	3.0 MHz	3.2 MHz
1.4 MHz LTE (6 RBs)	0.8 MHz	0.9 MHz	1.4 MHz	1.6 MHz

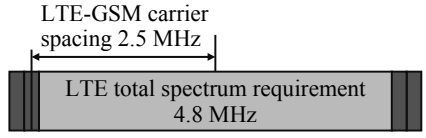


Figure 9.28 LTE 5-MHz refarming example

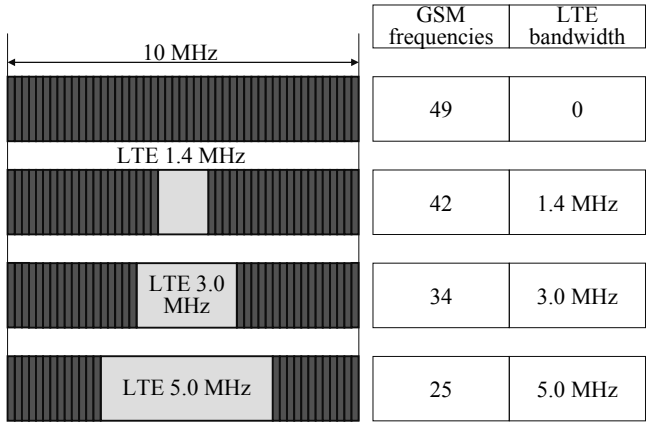


Figure 9.29 LTE refarming to GSM spectrum

case assumes that LTE and GSM use the same sites while the uncoordinated case assumes that different sites are used for LTE and GSM. The uncoordinated case causes larger power differences between the systems and leads to a larger guard band requirement. The coordinated case values are based on the GSM UE emissions and the uncoordinated values on LTE UE blocking requirements. It may be possible to push the LTE spectrum requirements down further for coordinated deployment depending on the GSM UE power levels and the allowed LTE uplink interference levels. The limiting factor is the maximum allowed interference to the PUCCH RBs that are located at the edge of the carrier.

The carrier spacing definition is illustrated in Figure 9.28. Figure 9.29 shows the expansion of the LTE carrier bandwidth when the GSM traffic decreases. Only seven GSM carriers need to be removed to make room for LTE 1.4 MHz and 15 GSM carriers for LTE 3.0 MHz.

9.9 Dimensioning

This section presents examples on how to convert the cell throughput values to the maximum number of broadband subscribers. Figure 9.30 shows two methods: a traffic volume based approach and a data rate based approach. The traffic volume based approach estimates the maximum traffic volume in gigabytes that can be carried by LTE 20 MHz 1 + 1 + 1 configuration. The spectral efficiency is assumed to be 1.74 bps/Hz/cell using 2×2 MIMO. The busy

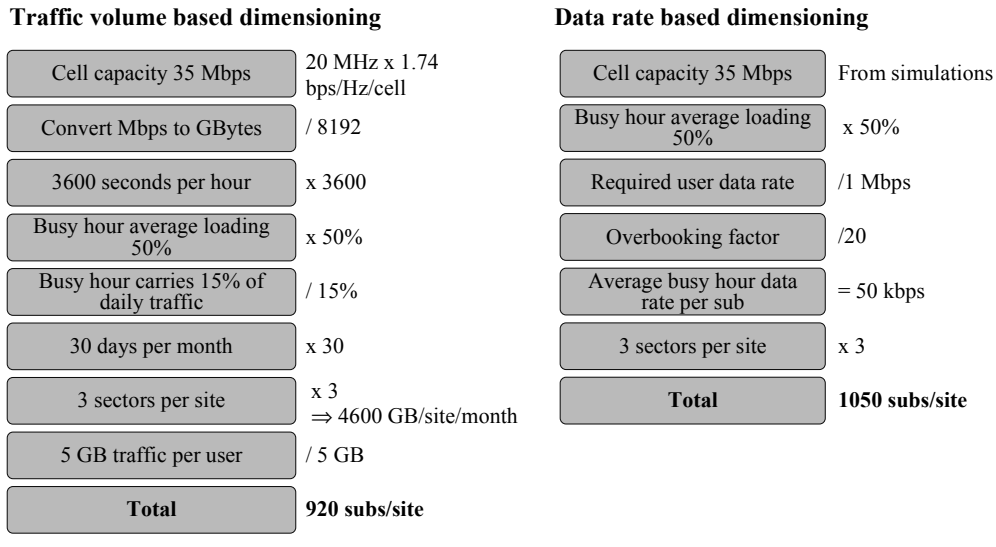


Figure 9.30 LTE dimensioning example for 1 + 1 + 1 at 20 MHz

hour is assumed to carry 15% of the daily traffic according to Figure 9.30 and the busy hour average loading is 50%. The loading depends on the targeted data rates during the busy hour: the higher the loading, the lower are the data rates. The maximum loading also depends on the applied QoS differentiation strategy: QoS differentiation pushes the loading closer to 100% while still maintaining the data rates for more important connections.

The calculation shows that the total site throughput per month is 4600 GB. To offer 5 GB data for every subscriber per month, the number of subscribers per site will be 920.

Another approach assumes a target of 1 Mbps per subscriber. Since only some of the subscribers are downloading data simultaneously, we can apply an overbooking factor, for example 20. This essentially means that the average busy hour data rate is 50 kbps per subscriber. The number of subscribers per site using this approach is 1050.

The calculation illustrates that LTE has the capability to support a large number of broadband data subscribers.

Figure 9.31 illustrates the technology and spectrum limits for the traffic growth assuming that HSPA and LTE use the existing GSM sites. The starting point is voice only traffic in a GSM network with 12 + 12 + 12 configuration, which corresponds to a high capacity GSM site found in busy urban areas. This corresponds to $12 \times 8 \times 0.016 = 1.5$ Mbps sector throughput assuming that each time slot carries 16 kbps voice rate. The voice traffic was the dominant part of the network traffic before flat rate HSDPA was launched. The data traffic has already exceeded the voice traffic in many networks in data volume. The second scenario assumes that the total traffic has increased 10 times compared to the voice only case. The sector throughput would then be 15 Mbps, which can be carried with three HSPA carriers using a 15 MHz spectrum.

The last scenario assumes 50 times more traffic compared to voice only, which leads to 75 Mbps and can be carried with two LTE carriers each of 20 MHz with a total 40 MHz of spectrum. The site throughput will be beyond 200 Mbps, setting corresponding requirements for the transport network capacity also.

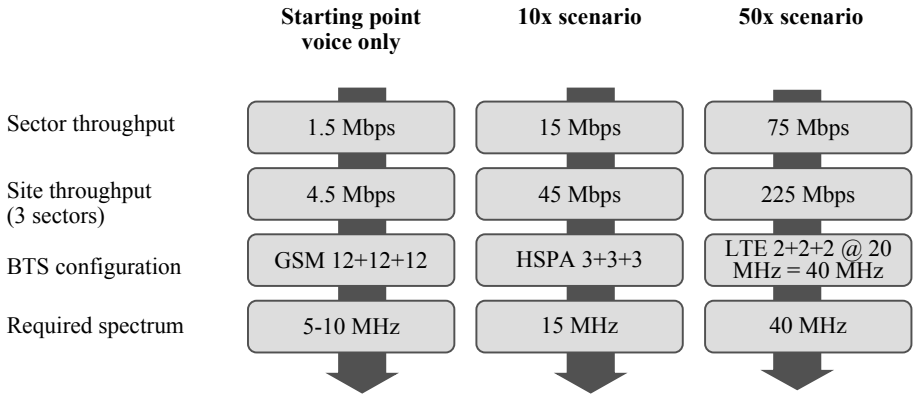


Figure 9.31 Traffic growth scenarios with 10 times and 50 times more traffic

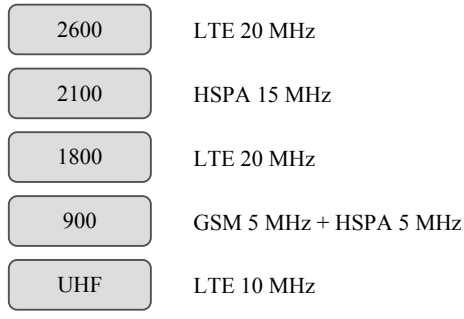


Figure 9.32 Example of a European operator with good spectrum resources

An example of the availability of the spectrum resources by a European operator within a few years is illustrated in Figure 9.32: GSM with 5 MHz, HSPA with 20 MHz and LTE with 50 MHz. Such an amount of spectrum would allow the traffic to increase more than 50 times compared to the voice only scenario. There are many operators in Asia and Latin America with less spectrum resources, which makes it more difficult to provide the high broadband wireless capacities.

9.10 Capacity Management Examples from HSPA Networks

In this section some of the HSDPA traffic analysis in RNC and at the cell level is shown and the implications discussed. It is expected that the analysis from broadband HSDPA networks will also be useful for the dimensioning of broadband LTE networks. All the analysis is based on the statistics of a single RNC and up to 200 NodeBs. The NodeBs were equipped with a 5-code round robin type shared HSDPA scheduler where five codes of HSDPA are shared among three cells and the transport solution was 2*E1 per NodeB. The maximum power for HSDPA was limited to 6 W. This area was selected to be analyzed as in this RNC the RF capability and transport capability are in line with each other, i.e. the transport solution can deliver the same throughput as the shared 5-code HSDPA scheduler. First it is necessary to evaluate the cell level data volume

fluctuations and contributions to RNC level total data volume so that the RNC throughput capacity can be checked. Then the cell level data volume and throughput limits are evaluated for when the new throughput improving features (proportional fair scheduler, cell dedicated scheduler, more codes, code multiplexing and so on) for the radio interface are introduced.

9.10.1 Data Volume Analysis

Figure 9.33 shows the data volume distribution over 24 h for the RNC on a typical working day. It can be seen that the single hour data volume share is a maximum of 6% from the total daily traffic volume and the fluctuation is just 3% to 6%. Also the hourly data volume share from the busy hour data volume share is 50% during the early morning hours and steadily increases towards the busiest hours from 8 pm to 1 am. The 3 hours from 9 pm to midnight are the busiest hours during the day. The usage increases heavily after about 6 pm, which indicates that as the working day ends then is the time for internet usage.

Looking into the individual cell contribution to the total RNC level data volume in Figure 9.34, it can be seen that during the night when the data volume is low and mobility is low, the traffic is also heavily concentrated on certain areas (cells) and during the day the share of cells contributing to the total data volume also increases.

As can be seen from Figure 9.34 during the early morning hours, 10% of the cells contribute 70–85% of the total RNC level data volume whereas during the busiest hours the same 70–85% data volume is contributed by 19–25% of the cells. During the early morning hours the data volume is very concentrated on just a couple of cells, which means that cells covering the residential areas should have very high data volume during the night time and early morning and due to low mobility the channel reservation times should be extremely long. This is shown in Figure 9.35, which indicates that during the early morning hours the

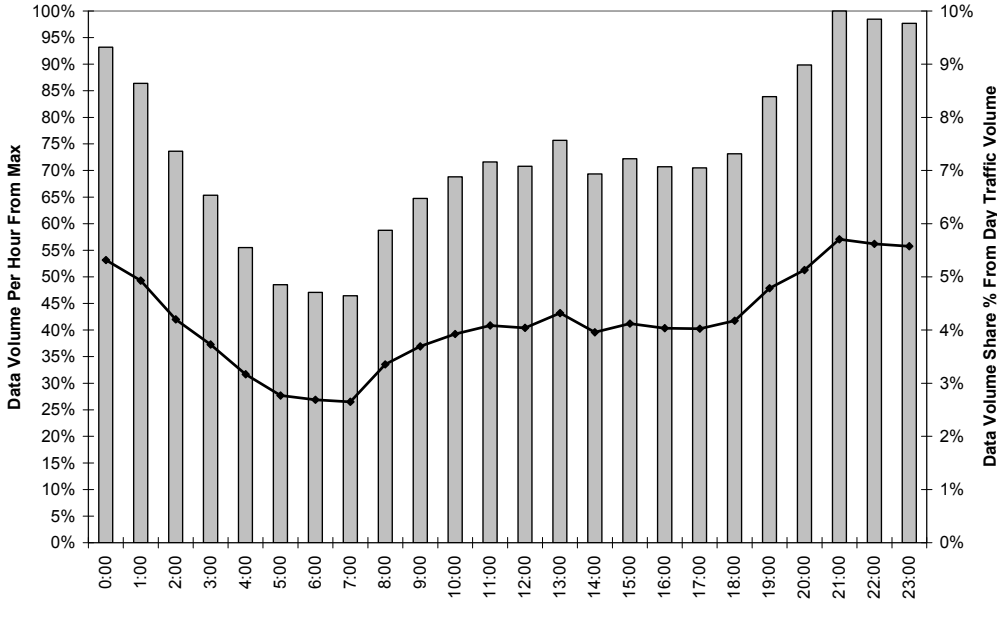


Figure 9.33 Daily RNC level hourly data volume deviation

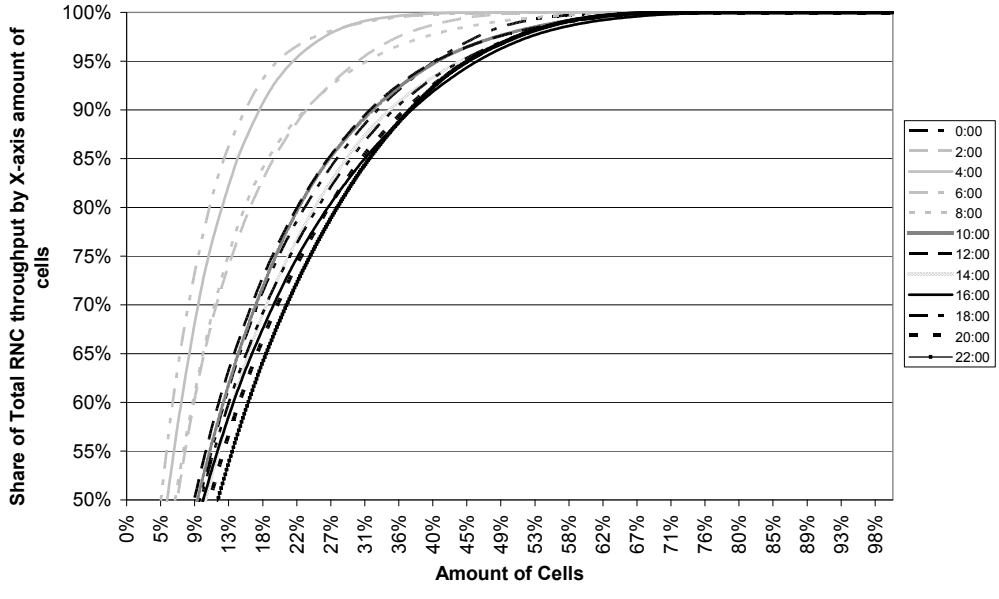


Figure 9.34 Cells' data volume contribution to total RNC data volume

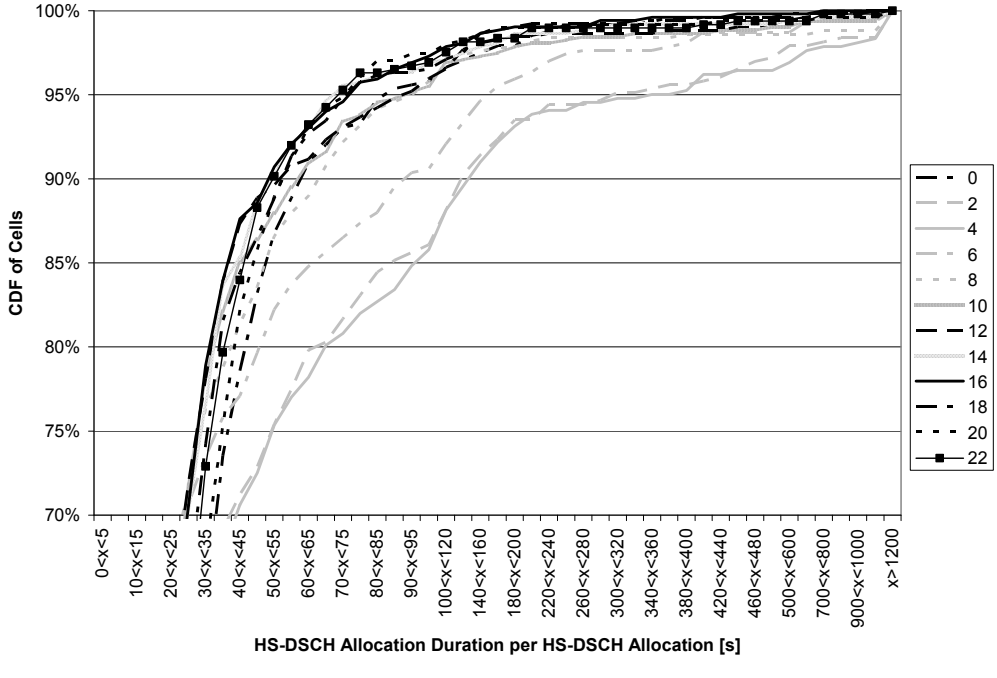


Figure 9.35 Average HS-DSCH allocation duration CDF of cells

average HS-DSCH allocation duration under the cell is a lot longer than during the highest usage. The lack of mobility and potentially some file sharing dominating usage during night time means that certain cells have extremely high data volumes at certain hours only and some cells have a fairly high usage and low data volume variation throughout the whole day. This makes efficient cell dimensioning a challenging task as if capacity is given to a cell according to the busy hour needs, then some cells are totally empty during most times of the day.

In Figure 9.36, the data volume fluctuation per hour is shown as the cumulative distribution of cells having a certain percentage of data volume from the highest data volume during the specific hour. From the graphs it can be seen that during the early morning and low data volume times there is also the highest number of cells having the lowest data volume. This indicates a very high fluctuation of data volume between all the cells.

Figure 9.37 shows the busy hour share of the total daily traffic on the cell level. The cells are sorted according to the data volume: the cells on the left have the highest traffic volume. The busy hour carries 10–20% of the daily traffic in busy cells. Figure 9.38 shows the distribution of the cells depending on the percentage of the data carried by the busy hour. The most typical values are 10–15% and 15–20%. We need to note that the busy hour share on the RNC level was only 6%. The difference between the cell and RNC level traffic distribution is explained by the trunking gain provided by RNC since the traffic is averaged over hundreds of cell.

9.10.2 Cell Performance Analysis

The cell performance analysis is carried out using the key indicators below:

- Active HS-DSCH Throughput – typically given in this analysis as kbps, it is the throughput under a cell when data have been sent in TTIs. Put simply it is the amount of data (kbit)/ number of active TTIs (s) averaged over a 1 h measurement period.
- HSDPA Data Volume – typically given in this analysis as Mbyte, kbit or Mbit, it is the amount of data sent per cell during the 1 h measurement period.
- Average number of simultaneous HSDPA users, during HSDPA usage – the amount of simultaneous users during the active TTIs, i.e. how many users are being scheduled during active TTIs per cell (the maximum amount depends on operator purchased feature set). When taking the used application into account, the average number of simultaneous users during HSDPA usage needs to be replaced by the number of active users who have data in the NodeB buffers. Averaged over the 1 h measurement period.
- Allocated TTI share – the average number of active TTIs during the measurement period (i.e. when there are data to send, the TTI is active) over all the possible TTIs during the 1 h measurement period.
- Average Throughput per User – typically given as kbps, it is the Active HS-DSCH Throughput/ Average number of simultaneous HSDPA users adjusted by the allocated TTI share; it is the average throughput that one single user experiences under a certain cell. When taking into account the used application, the Active HS-DSCH throughput needs to be divided by the number of active users who have data in the NodeB buffer. Averaged over the 1 h measurement period.
- Average reported Channel Quality Indicator (CQI) – every UE with HS-DSCH allocated measures and reports the CQI value back to the BTS. The average reported CQI is the

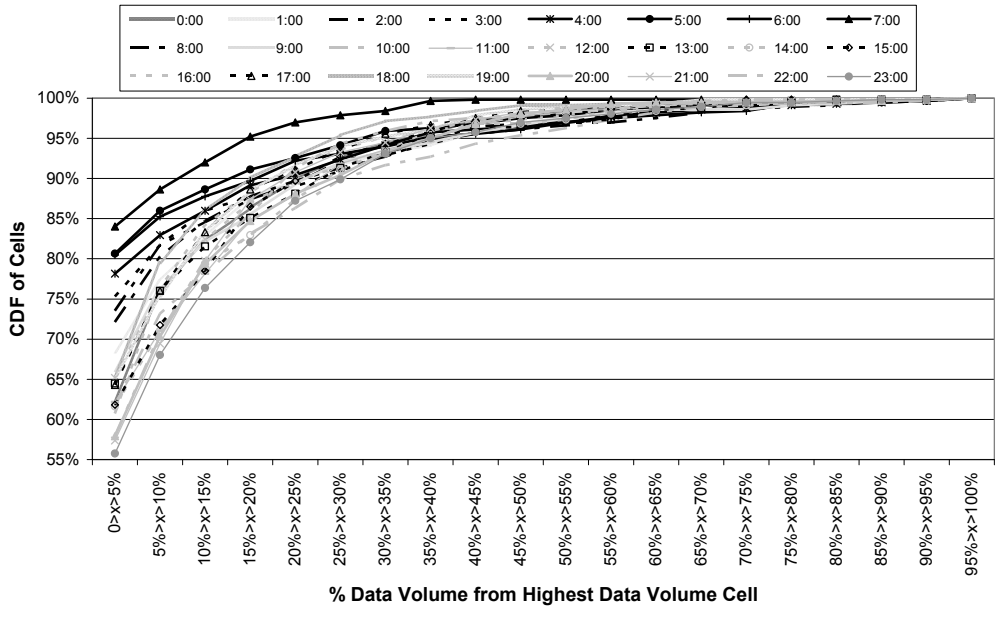


Figure 9.36 Percentage data volume per cell from highest data volume cell per hour

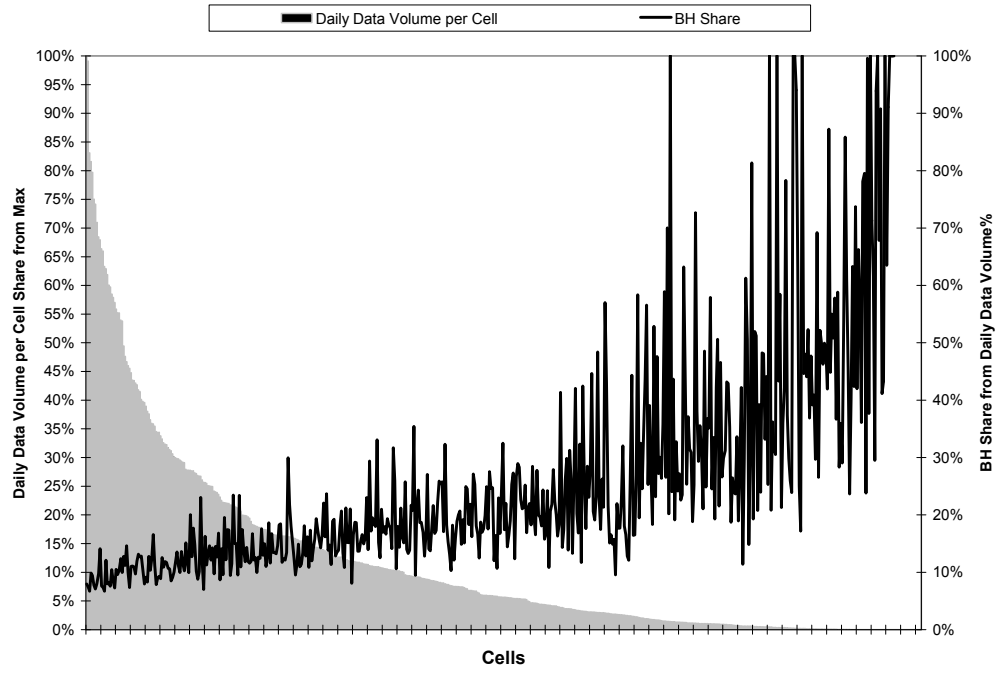


Figure 9.37 Busy hour share of cell level data volume

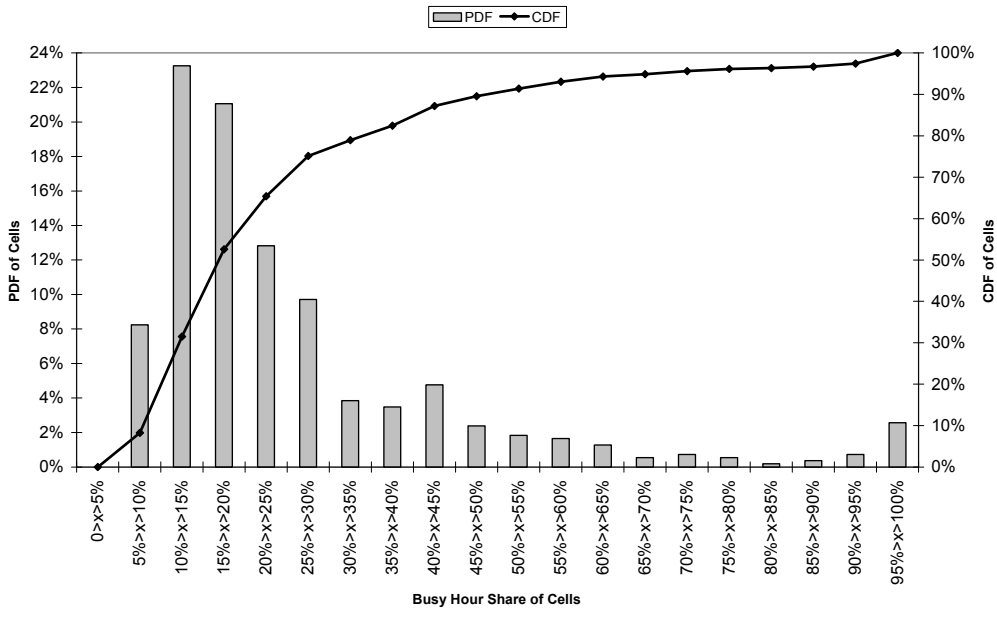


Figure 9.38 CDF and PDF of data volume busy hour share

CQI value reported by all the UEs under the cell averaged over the 1 h measurement period.

The end user throughput depends then on the average active HS-DSCH throughput and the average number of simultaneous HSDPA users. Figure 9.39 shows the average throughput as a function of the average number of simultaneous users. The average active HS-DSCH throughput is approximately 1.2 Mbps. When the allocated TTI share is low, the end user throughput approaches the HS-DSCH throughput. When the allocated TTI share increases, the end user throughput is reduced. This calculation is, however, pessimistic since it assumes that all users are downloading all the time.

When the used application activity is taken into account, i.e. the active HS-DSCH throughput is divided by the number of users who have data in the NodeB buffer, the average throughput per user is quite different from the formula used above. Figure 9.40 shows the comparison of the two different average throughput per user formulas and the Active HS-DSCH throughput. It should be noted that Figure 9.39 and Figure 9.40 are not from exactly the same time. The end user throughput is 400–800 kbps when only those users are considered that have data in the buffer. This can be explained when analyzing the difference between the number of simultaneous users and the number of simultaneous users who have data in the NodeB buffer as shown in Figure 9.41.

Figure 9.41 shows that the used application plays a significant role in the end user throughput evaluation and it should not be ignored. Therefore, the average user throughput may be low because the application simply does not need a high average data rate. The network performance analysis needs to separate the data rate limitations caused by the radio network and the actual used application data rate.

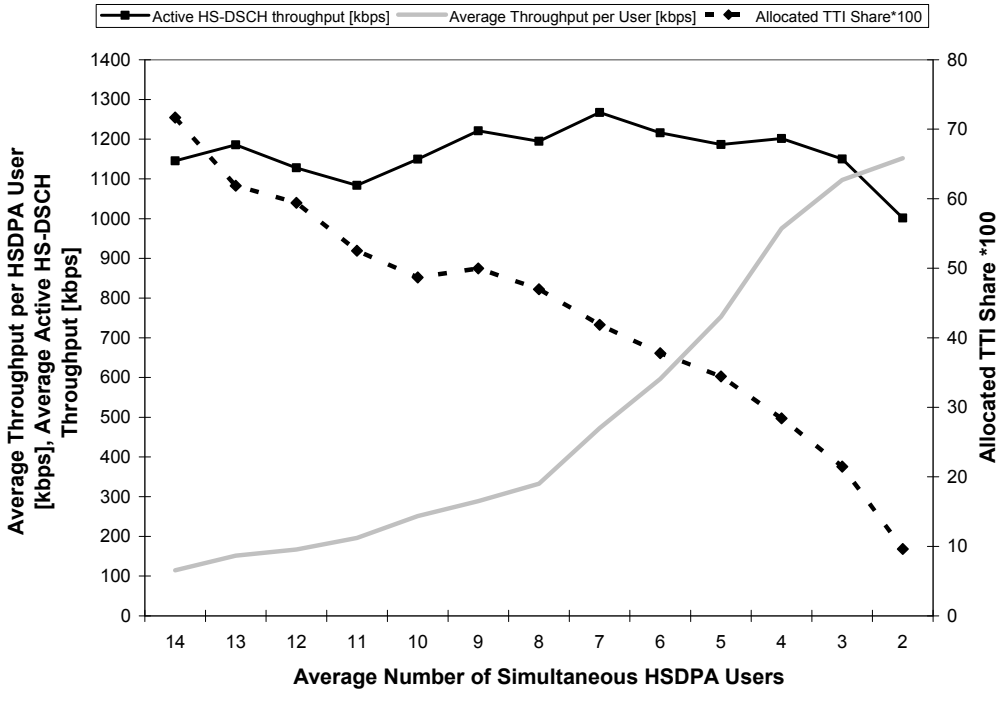


Figure 9.39 Average throughput per user

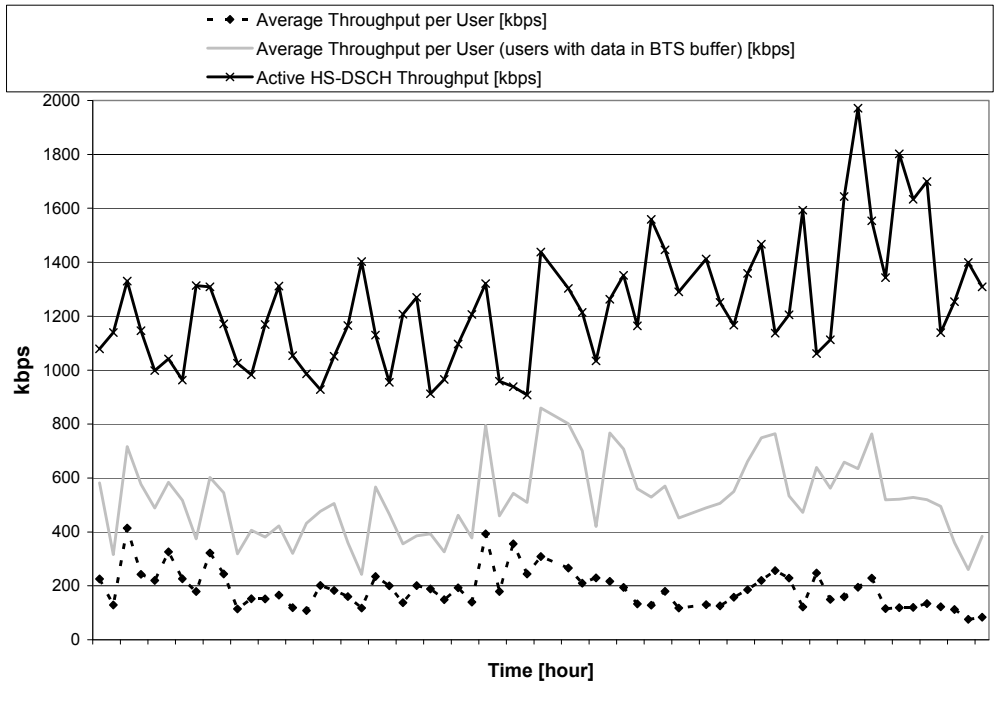


Figure 9.40 Average throughput per user, comparisons of different formulas

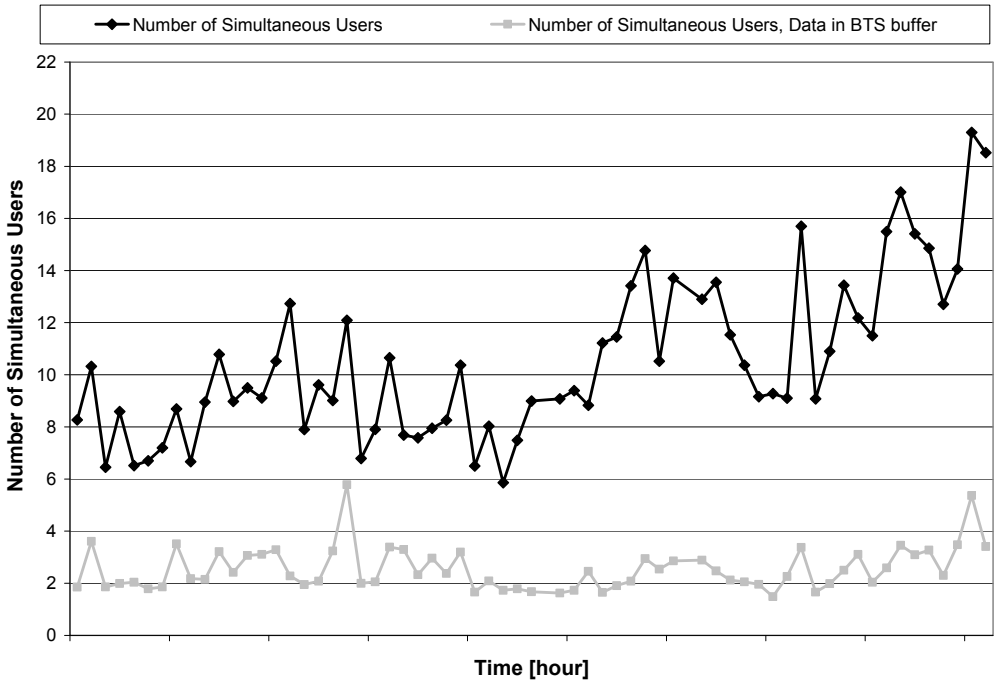


Figure 9.41 Number of simultaneous users with and without taking the used application into account

9.11 Summary

3GPP LTE Release 8 enables a peak bit rate of 150 Mbps in downlink and 50 Mbps in uplink with 2×2 MIMO antenna configuration in downlink and 16QAM modulation in uplink. The bit rates can be pushed to 300 Mbps in downlink with 4×4 MIMO and 75 Mbps in uplink with 64QAM. It is expected that the first LTE deployments will provide bit rates up to 150 Mbps.

The LTE link level performance can be modeled with the theoretical Shannon limit when suitable correction factors are included. LTE link level performance was shown to be robust with high mobile speeds, and the uplink LTE performance can be optimized by using adaptive transmission bandwidth. The link level simulations are used in the link budget calculations, which indicate that the LTE link budget is similar to the HSPA link budget with the same data rate and same spectrum.

The LTE system performance is optimized by orthogonal transmission schemes, by MIMO transmission and by frequency domain scheduling. The spectral efficiency can be further enhanced with multi-antenna transmission and higher order sectorization. The high efficiency can be maintained for different LTE bandwidths between 5 and 20 MHz, while the spectral efficiency is slightly lower with the narrowband LTE carriers 1.4 and 3.0 MHz. It was shown that LTE provides higher spectral efficiency compared to HSPA and HSPA evolution especially due to the frequency domain packet scheduling.

The user plane latency in LTE can be as low as 10–20 ms. The low latency is relevant for improving the end user performance since many applications and protocols benefit from low latency. The low latency is enabled by the short sub-frame size of 1 ms.

The flexible refarming of LTE to the GSM spectrum is enabled by the narrowband options: the refarming can be started with 1.4 or 3.0 MHz and later expanded when the GSM traffic has decreased. All UEs need to support all bandwidths between 1.4 and 20 MHz.

The dimensioning of the broadband wireless networks is different from voice networks. The examples from HSPA networks illustrate that the traffic distribution over the day, over the geographical area and the user mobility need to be considered.

References

[1] 3GPP Technical Specification 36.213 'Physical layer procedures', v. 8.3.0.
[2] 3GPP Technical Specification 36.306 v8.2.0: 'User Equipment (UE) radio access capabilities', August 2008.
[3] P. Mogensen *et al*. 'LTE Capacity compared to the Shannon Bound', *IEEE Proc. Vehicular Technology Conference*, pp. 699–703, April 2007.
[4] 3GPP Technical Specification 25.942 'Radio Frequency (RF) system scenarios', v. 7.0.0.
[5] 3GPP Technical Report 25.996 'Spatial channel model for Multiple Input Multiple Output (MIMO) simulations', V.7.0.0.
[6] I.Z. Kovács *et al*., 'Effects of Non-Ideal Channel Feedback on Dual-Stream MIMO OFDMA System Performance', *IEEE Proc. Veh. Technol. Conf.*, Oct. 2007
[7] I.Z. Kovács *et al.*, 'Performance of MIMO Aware RRM in Downlink OFDMA', *IEEE Proc. Veh. Technol. Conf.*, pp. 1171–1175, May 2008.
[8] S. Kumar, *et al*., 'Performance Evaluation of 6-Sector-Site Deployment for Downlink UTRAN Long Term Evolution', *IEEE Proc. Vehicular Technology Conference*, September 2008.
[9] B. Hagerman, D. Imbeni, J. Barta, A. Pollard, R. Wohlmuth, P. Cosimini, 'WCDMA 6 Sector Deployment-Case study of a real installed UMTS-FDD Network', *Vehicular Technology Conference*, Vol. 2, pp. 703–707, Spring 2006.
[10] K.I. Pedersen, P.E. Mogensen, B. Fleury, 'Spatial Channel Characteristics in Outdoor Environments and Their Impact on BS Antenna System Performance', *IEEE Proc. Vehicular Technology Conference*, pp. 719–723, May 1998.
[11] 3GPP TSG RAN R1-072444 'Summary of Downlink Performance Evaluation', May 2007.
[12] 3GPP TSG RAN R1-072261 'LTE Performance Evaluation – Uplink Summary', May 2007.

10

Voice over IP (VoIP)

Harri Holma, Juha Kallio, Markku Kuusela, Petteri Lundén, Esa Malkamäki, Jussi Ojala and Haiming Wang

10.1 Introduction

While the data traffic and the data revenues are increasing, the voice service still makes the majority of operators' revenue. Therefore, LTE is designed to support not only data services efficiently, but also good quality voice service with high efficiency. As LTE radio only supports packet services, the voice service will also be Voice over IP (VoIP), not Circuit Switched (CS) voice. This chapter presents the LTE voice solution including voice delay, system performance, coverage and inter-working with the CS networks. First, the general requirements for voice and the typical voice codecs are introduced. The LTE voice delay budget calculation is presented. The packet scheduling options are presented and the resulting system capacities are discussed. The voice uplink coverage challenges and solutions are also presented. Finally, the LTE VoIP inter-working with the existing CS networks is presented.

10.2 VoIP Codecs

GSM networks started with Full rate (FR) speech codec and evolved to Enhanced Full Rate (EFR). The Adaptive Multi-Rate (AMR) codec was added to 3GPP Release 98 for GSM to enable codec rate adaptation to the radio conditions. AMR data rates range from 4.75 kbps to 12.2 kbps. The highest AMR rate is equal to the EFR. AMR uses a sampling rate of 8 kHz, which provides 300–3400 Hz audio bandwidth. The same AMR codec was included also for WCDMA in Release 99 and is also used for running the voice service on top of HSPA. The AMR codec can also be used in LTE.

The AMR-Wideband (AMR-WB) codec was added to 3GPP Release 5. AMR-WB uses a sampling rate of 16 kHz, which provides 50–7000 Hz audio bandwidth and substantially better voice quality and mean opinion score (MOS). As the sampling rate of AMR-WB is double the sampling rate of AMR, AMR is often referred to as AMR-NB (narrowband). AMR-WB data rates range from 6.6 kbps to 23.85 kbps. The typical rate is 12.65 kbps, which is similar to the

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6

normal AMR of 12.2 kbps. AMR-WB offers clearly better voice quality than AMR-NB with the same data rate and can be called wideband audio with narrowband radio transmission. The radio bandwidth is illustrated in Figure 10.1 and the audio bandwidth in Figure 10.2. The smallest bit rates, 1.8 and 1.75 kbps, are used for the transmission of Silence Indicator Frames (SID).

This chapter considers AMR codec rates of 12.2, 7.95 and 5.9 kbps. The resulting capacity of 12.2 kbps would also be approximately valid for AMR-WB 12.65 kbps.

When calling outside mobile networks, voice transcoding is typically required to 64 kbps Pulse Code Modulation (PCM) in links using ITU G.711 coding. For UE-to-UE calls, the transcoding can be omitted with transcoder free or tandem free operation [1]. Transcoding generally degrades the voice quality and is not desirable within the network.

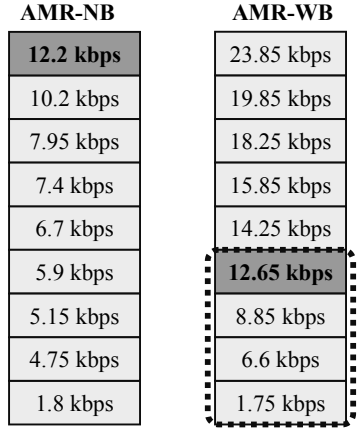


Figure 10.1 Adaptive Multirate (AMR) Voice Codec radio bandwidth

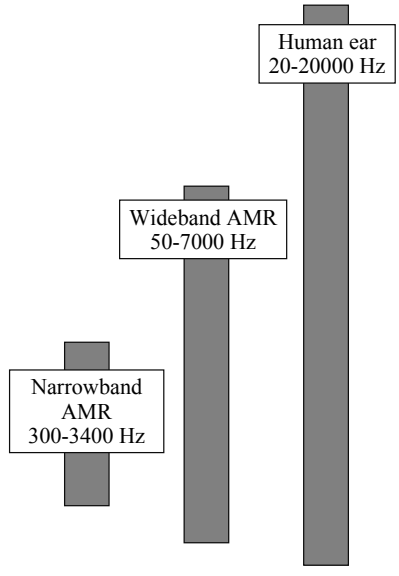


Figure 10.2 Adaptive Multirate (AMR) Voice Codec audio bandwidth

There are also a number of other voice codecs that are generally used for VoIP. A few examples are G.729, using an 8 kbps coding rate, and Internet Low Bit Rate Codec (iLBC), using 13 kbps which is used, for example, in Skype and in Googletalk.

10.3 VoIP Requirements

There is a high requirement set for the radio network to provide a reliable and good quality voice service. Some of the main requirements are considered below.

The impact of the mouth-to-ear latency on user satisfaction is illustrated in Figure 10.3. The delay preferably should be below 200 ms, which is similar to the delay in GSM or WCDMA voice calls. The maximum allowed delay for a satisfactory voice service is 280 ms.

IP Multimedia Subsystem (IMS) can be deployed to control VoIP. IMS is described in Chapter 3. IMS provides the information about the required Quality of Service (QoS) to the radio network by using 3GPP standardized Policy and Charging Control (PCC) [3]. The radio network must be able to have the algorithms to offer the required QoS better than Best Effort. QoS includes mainly delay, error rate and bandwidth requirements. QoS in LTE is described in Chapter 8.

The voice call drop rates are very low in the optimized GSM/WCDMA networks today – in the best case below 0.3%. VoIP in LTE must offer similar retainability including smooth interworking between GSM/WCDMA circuit switched (CS) voice calls. The handover functionality from VoIP in LTE to GSM/WCDMA CS voice is called Single radio Voice Call Continuity (SR-VCC) and described in detail in section 10.10.

The AMR 12.2 kbps packet size is 31 bytes while the IP header is 40-60 bytes. IP header compression is a mandatory requirement for an efficient VoIP solution. IP header compres-

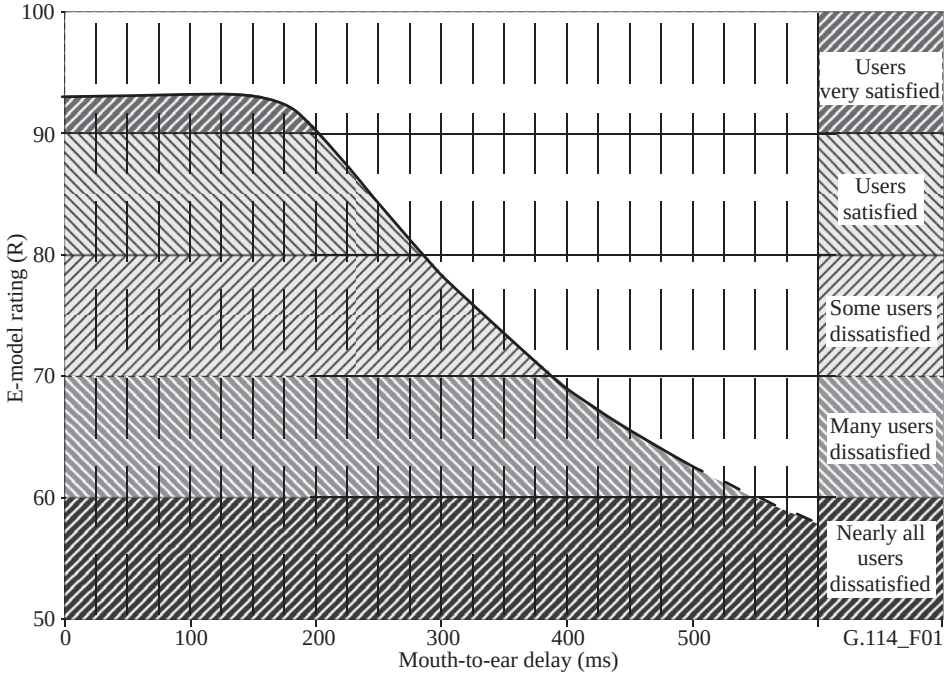


Figure 10.3 Voice mouth-to-ear delay requirements [2]

sion is required both in UE and in eNodeB. All the VoIP simulations in this chapter assume IP header compression.

The IP connectivity requires keep alive messages when the UE does not have a phone call running. The frequency of the keep alive messages depends on the VoIP solution: operator IMS VoIP can use fairly infrequent keep alive messages since IMS is within the operator's own network and no firewalls or Network Address Tables (NAT) are required in between. The internet VoIP requires very frequent keep alive message to keep the connectivity open through firewalls and NATs. The frequent keep alive message can affect UE power consumption and network efficiency.

VoIP roaming cases need further attention especially if there are some LTE networks designed for VoIP and data, while some networks are designed for data only transmission without the required voice features. VoIP roaming also requires IMS and GPRS roaming agreements and the use of visited GGSN/MME model. One option is to use circuit switched (CS) calls whenever roaming with CS Fallback for LTE procedures. Similarly CS calls can also be used for emergency calls since 3GPP Release 8 LTE specifications do not provide the priority information from the radio to the core network nor a specific emergency bearer.

10.4 Delay Budget

The end-to-end delay budget for LTE VoIP is considered here. The delay should preferably be below 200 ms, which is the value typically achieved in the CS network today. We use the following assumptions in the delay budget calculations. The voice encoding delay is assumed to be 30 ms including a 20 ms frame size, 5 ms look-ahead and 5 ms processing time. The receiving end assumes a 5 ms processing time for the decoding. The capacity simulations assume a maximum 50 ms air–interface delay in both uplink and downlink including scheduling delay and the time required for the initial and the HARQ retransmissions of a packet. We also assume a 5 ms processing delay in the UE, 5 ms in eNodeB and 1 ms in the SAE gateway. The transport delay is assumed to be 10 ms and it will depend heavily on the network configuration. The delay budget is presented in Figure 10.4. The mouth-to-ear delay is approximately 160 ms with these

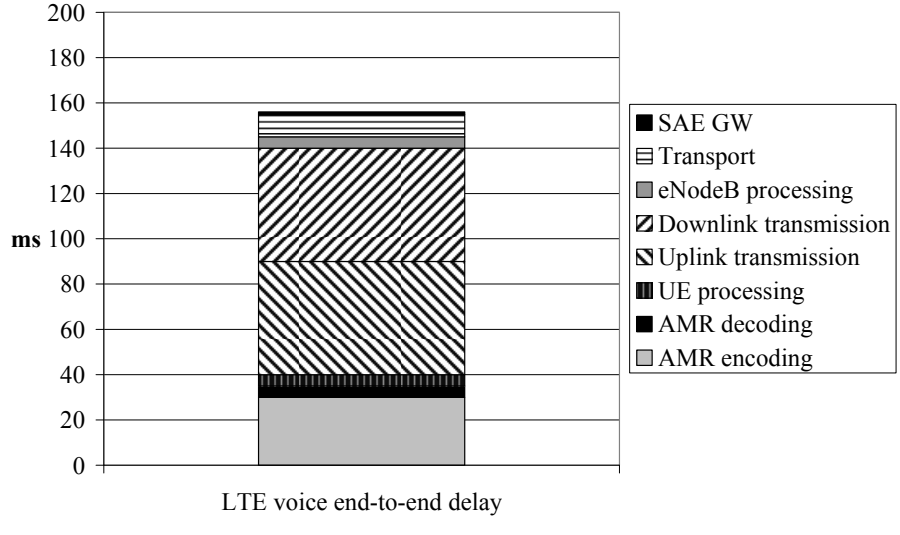


Figure 10.4 LTE voice end-to-end delay budget

assumptions, illustrating that LTE VoIP can provide lower end-to-end latency than CS voice calls today while still providing high efficiency.

10.5 Scheduling and Control Channels

By default the LTE voice service uses dynamic scheduling, where control channels are used to allocate the resources for each voice packet transmission and for the possible retransmissions. Dynamic scheduling gives full flexibility for optimizing resource allocation, but it requires control channel capacity. Multi-user channel sharing plays an important role when optimizing the air interface for VoIP traffic. Because each user is scheduled with control signaling with the dynamic scheduler, the control overhead might become a limiting factor for VoIP capacity. One solution in downlink for reducing control channel capacity consumption with a dynamic scheduler is to bundle multiple VoIP packets together at Layer 1 into one transmission of a user. The packet bundling is CQI based and is applied only for users whose channel conditions favor the use of bundling. The main benefit from the bundling is that more users could be fitted to the network with the same control channel overhead as users in good channel conditions are scheduled less often. The drawback from the bundling is that bundled packets experience a tighter delay budget, but the negative impact of this to VoIP capacity can be kept to a minimum by applying bundling for good enough users that are not relying on HARQ retransmissions. From the voice quality perspective it is also important to minimize the probability of losing consecutive packets. This can be achieved by making the link adaptation for the TTIs carrying bundled packets in a more conservative way leading to a reduced packet error rate for the first transmission. Due to the low UE transmission power and non-CQI based scheduling, packet bundling is not seen as an attractive technique in LTE uplink.

The 3GPP solution to avoid control channel limitation for VoIP capacity is the Semi-Persistent Scheduling (SPS) method [4, 5], where the initial transmissions of VoIP packets are sent without associated scheduling control information by using persistently allocated transmission resources instead. The semi-persistent scheduling is configured by higher layers (Radio Resource Control, RRC), and the periodicity of the semi-persistent scheduling is signaled by RRC. At the beginning of a talk spurt, the semi-persistent scheduling is activated by sending the allocated transmission resources by Physical Downlink Control Channel (PDCCH) and the UE stores the allocation and uses it periodically according to the periodicity. With semi-persistent scheduling only retransmissions and SID frames are scheduled dynamically, implying that the control channel capacity is not a problem for the semi-persistent scheduler. On the other hand, only limited time and frequency domain scheduling gains are available for semi-persistent scheduling. The semi-persistent allocation is released during the silence periods. Semi-persistent scheduling is adopted for 3GPP Release 8. The downlink operation of the semi-persistent scheduler is illustrated in Figure 10.5.

In the following the impact of control channel capacity to VoIP maximum capacity with dynamic scheduler is illustrated with theoretical calculations.

It is assumed that the number of Control Channel Elements (CCE) is approximately 20 per 5 MHz bandwidth. Two, four or eight CCEs can be aggregated per user depending on the channel conditions in low signal-to-noise ratios. We further apply the following assumptions for the calculations: voice activity 50%, voice packet arrival period 20 ms, downlink share of the traffic 50% and number of retransmissions 20%. The results are calculated without packet

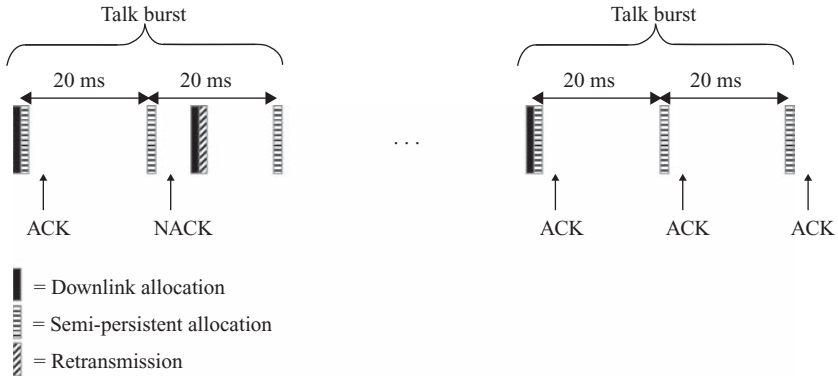


Figure 10.5 Semi-persistent scheduling in downlink

bundling and with two packet bundling. To simplify the calculations it is further assumed that SID frames are not taken into account. VoIP maximum capacity can be calculated by using Equation 10.1.

$$Max\_users = \frac{\#CCEs}{\#CCEs\_per\_user} \cdot \frac{Packet\_period[ms]}{Voice\_activity} \cdot Packet\_bundling \cdot Downlink\_share \cdot \frac{1}{1 + BLER} \tag{10.1}$$

The calculation results are shown in Figure 10.6 for a 5 MHz channel. As an example, the maximum capacity would be 330 users without CCE aggregation and without packet bundling. Assuming an average CCE aggregation of three brings the maximum capacity to 110 without packet bundling and 220 with packet bundling. These theoretical calculations illustrate the importance that multi-user channel sharing has for VoIP system performance, and furthermore, the maximum gain in capacity that packet bundling may provide with the dynamic

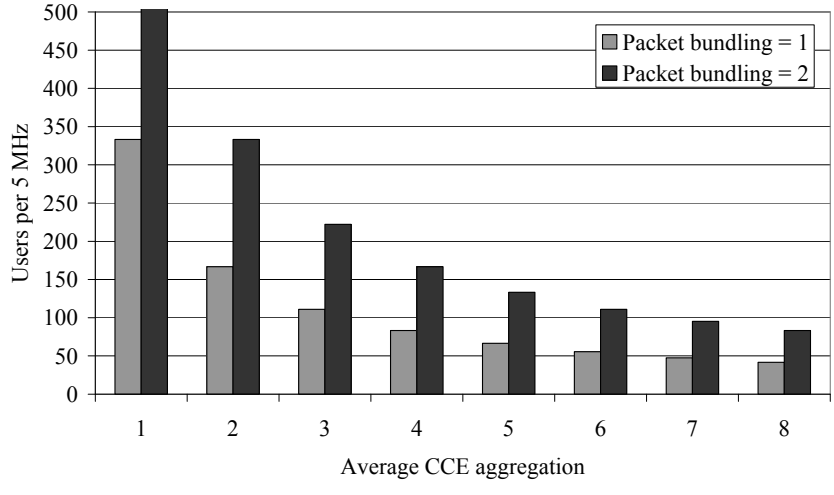


Figure 10.6 Maximum control channel limited capacity with fully dynamic scheduling

scheduler. According to the system level simulations, the average size of CCE aggregation is approximately 1.5 for a scheduled user. Comparison of the simulated downlink capacities[1] for a dynamic scheduler (presented in section 10.6) with the theoretical upper limits shows that the simulated capacities without packet bundling match rather nicely the theoretical calculations, being approximately 5% lower than the theoretical maximum. With packet bundling the simulated capacities are notably below the theoretical upper limits, as in practice the probability for using bundling is clearly limited below 100%.

10.6 LTE Voice Capacity

This section presents the system level performance of VoIP traffic in LTE at a 5 MHz system bandwidth. VoIP capacity is given as the maximum number of users that could be supported within a cell without exceeding 5% outage. A user is defined to be in an outage if at least 2% of its VoIP packets are lost (i.e. either erroneous or discarded) during the call. VoIP capacity numbers are obtained from system level simulations in a macro cell scenario 1 [6], the main system simulation parameters being aligned with [7]. An illustration of a VoIP capacity curve is presented in Figure 10.7, which shows how the outage goes up steeply when the maximum capacity is approached. This enables running the system with a relatively high load (compared to the maximum capacity) while still maintaining low outage.

The results of VoIP capacity simulations are summarized in Table 10.1 for three different AMR codecs (AMR 5.9, AMR 7.95 and AMR 12.2) and for both dynamic and semi-persistent schedulers. The capacity is 210–470 users in downlink and 210–410 users in uplink, which

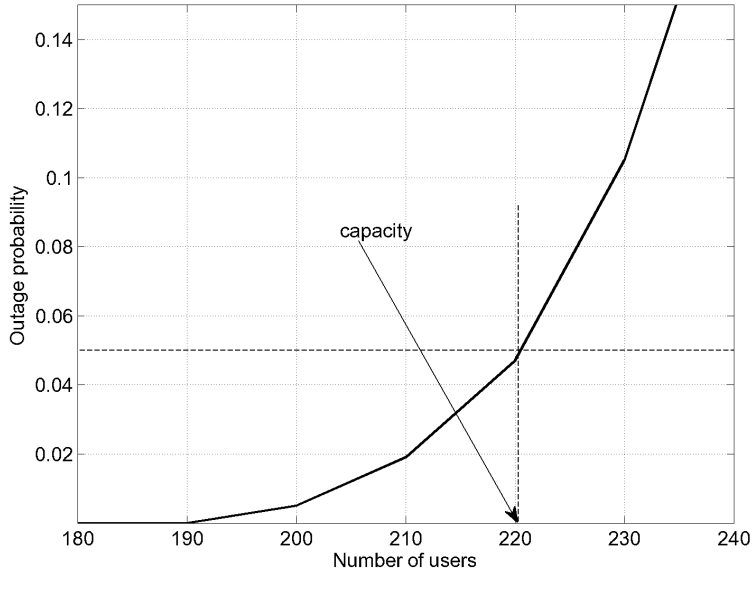


Figure 10.7 An example of VoIP capacity curve

[1] Simulated capacities should be up-scaled by 12.5% to remove the impact of SID transmissions on the capacity.

Table 10.1 VoIP capacity in LTE at 5 MHz

VoIP codec	AMR 5.9	AMR 7.95	AMR 12.2
Downlink capacity			
Dynamic scheduler, without bundling	210	210	210
Dynamic scheduler, with bundling	410	400	370
Semi-persistent scheduler	470	430	320
Uplink capacity			
Dynamic scheduler	230	230	210
Semi-persistent scheduler	410	320	240

corresponds to 42–94 users per MHz per cell in downlink and 42–82 users per MHz per cell in uplink. The lower AMR rates provide higher capacity than the higher AMR rates. The AMR codec data rate can be defined by the operator allowing a tradeoff between the capacity and the voice quality. The lower AMR rates do not increase the capacity with dynamic scheduling due to the control channel limitation.

The results also show that voice capacity is uplink limited, which can be beneficial when there is asymmetric data transmission on the same carrier taking more downlink capacity. The downlink offers higher capacity than uplink because the downlink scheduling uses a point-to-multipoint approach and can be optimized compared to the uplink.

In the following, the simulated capacities are analyzed in more detail. All the supporting statistics presented in the sequel are assuming load as close as possible to the 5% outage.

As described in section 10.5, VoIP system performance with the dynamic scheduler is limited by the available PDCCH resources, and therefore the dynamic scheduler cannot fully exploit the Physical Downlink Shared Channel (PDSCH) air interface capacity as there are not enough CCEs to schedule the unused Physical Resource Blocks (PRBs). This is illustrated in Figure 10.8 for downlink, which contains a cumulative distribution function for the scheduled PRBs per Transmission Time Interval (TTI) with AMR 12.2. The total number of PRBs on 5 MHz bandwidth is 25. As can be seen from Figure 10.8, the average utilization rate for the dynamic scheduler is only 40% if packet bundling is not allowed.

Due to the control channel limitations, the savings in VoIP packet payload size with lower AMR rates are not mapped to capacity gains with the dynamic scheduler, if packet bundling is not used. The different codecs provide almost identical performance in downlink whereas there are small gains in capacity in uplink with lower AMR rates. This uplink gain originates from the improved coverage as more robust MCS could be used when transmitting a VoIP packet in uplink – this gain does not exist in the downlink direction as it is not coverage limited due to higher eNodeB transmission power.

When packet bundling is enabled, the average size of an allocation per scheduled user is increased and hence the air interface capacity of PDSCH can be more efficiently exploited, which leads to an increased average PRB utilization rate of 70%. Gains of 75–90% are achieved for capacity when packet bundling is enabled. The probability of using packet bundling in a macro cell scenario is approximately 70% for AMR 12.2 codec, and it slightly increases when the VoIP packet payload decreases. This explains the small (< 10%) gains in capacity achieved by reducing VoIP packet payload.

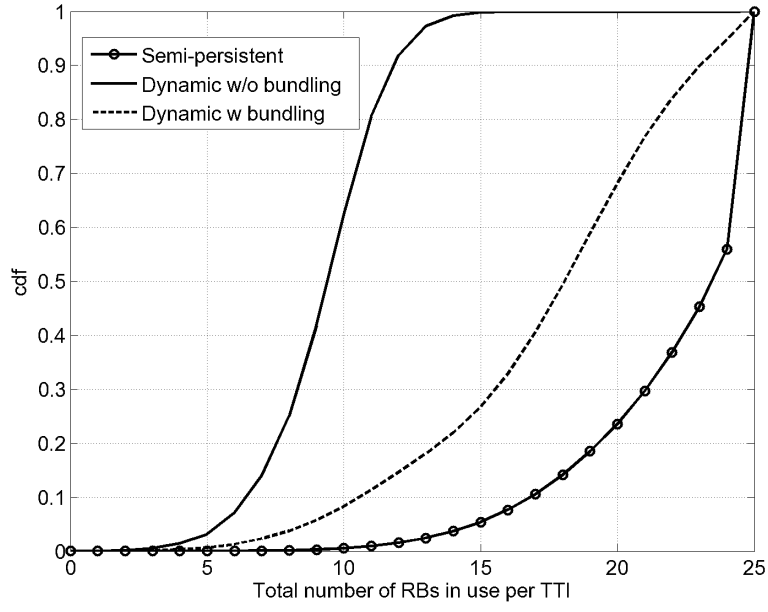


Figure 10.8 Cumulative distribution function for scheduled Physical Resource Blocks (PRB) per TTI in downlink (out of a total of 25 PRBs on 5 MHz bandwidth)

The performance of the semi-persistent scheduler does not suffer from control channel limitations, as initial transmissions are scheduled without associated control information by using persistently allocated time and frequency resources instead. The difference in control channel consumption for simulated packet scheduling methods is presented for the downlink direction in Figure 10.9, which contains the cumulative distribution function for the total number of consumed CCEs per TTI. With the dynamic scheduler all control channel elements are in use 70% of the time, if bundling is not allowed. With packet bundling, load is higher at 5% outage and hence control channel capacity could be more efficiently exploited implying that all CCEs are used almost 100% of the time. With the semi-persistent scheduler, the control channel overhead is clearly lower, the probability of using all CCEs being only slightly higher than 10%. Similar observations can be made from the distribution of consumed CCEs per TTI in the uplink direction, which is presented in Figure 10.10. As expected, the CCE consumption for semi-persistent scheduling is much lower than that for dynamic scheduling. Here we assume that the total number of CCEs reserved for downlink/uplink scheduling grants is 10.

The performance of the semi-persistent scheduler is not limited by the control channel resources, but it is limited by the available PDSCH bandwidth, which is illustrated in Figure 10.8, showing that the average PDSCH utilization rate for a semi-persistent scheduler is more than 90%. As the performance of the semi-persistent scheduler is user plane limited, the size of a VoIP packet has a significant impact on the capacity: an approximately 35% higher capacity is obtained if AMR 7.95 is used instead of AMR 12.2, whereas the use of AMR 5.9 instead of AMR 7.95 provides capacity gains of only 10%. The reason for the reduced gain for AMR 5.9 is presented in Figure 10.11, which shows the distribution of the size of the persistent resource allocation in terms of PRBs for all simulated codecs in the downlink direction. According to

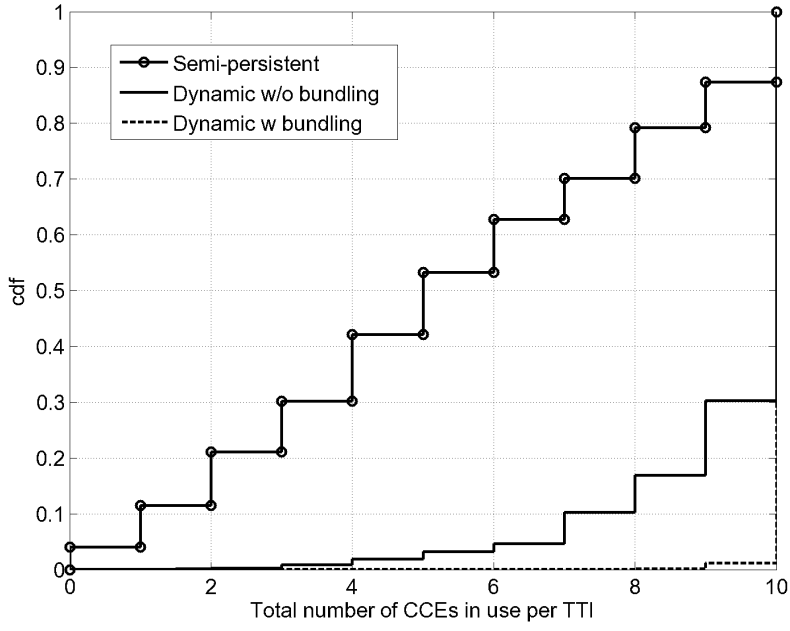


Figure 10.9 Cumulative distribution function for the total number of CCEs per TTI in downlink

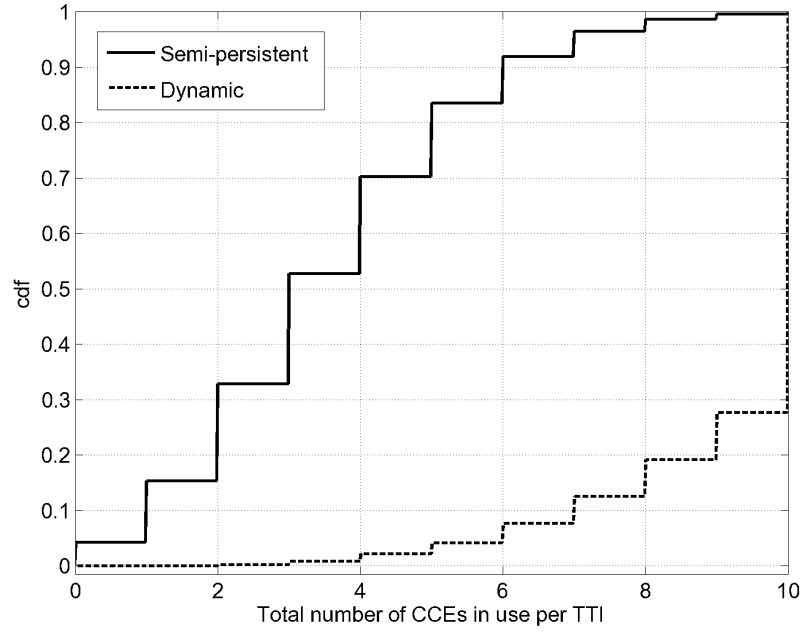


Figure 10.10 Cumulative distribution function for the total number of CCEs per TTI in uplink

the distribution, the size of the persistent allocation is one (1) PRB for almost 50% of the users with AMR 7.95, and hence no savings in allocated bandwidth can be achieved for those users when using AMR 5.9 instead of AMR 7.95. In the downlink the size of the persistent allocation

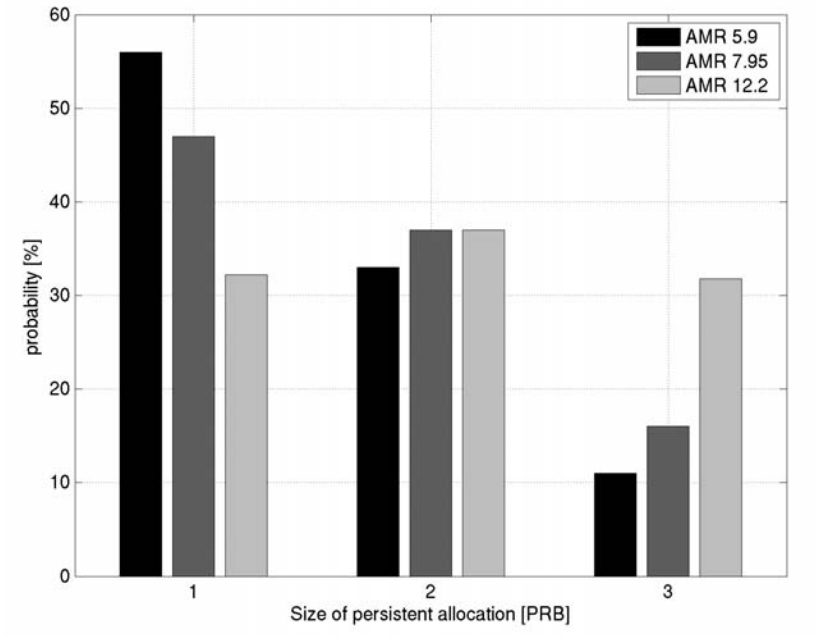


Figure 10.11 Probability distribution function for the size of the persistent resource allocation for different codecs in downlink

is calculated dynamically from wideband CQI for each talk spurt separately, whereas in the uplink, due to practical constraints, the size of the persistent resource allocation is decided from the path loss measurements of a user. The reason for the uplink solution is that the sounding is not accurate enough with a large number of users. The size of the persistent resource allocation has the following distribution during the uplink simulation: with AMR12.2 all the users have a persistent allocation of size two (2) PRBs, whereas with AMR 7.95 half of the users have a persistent allocation of size one (1) PRB, and half of the users have a persistent allocation of size two (2) PRBs. With AMR 5.9, 90% of the users have a persistent allocation of size one (1) PRB, and others have a persistent allocation of size two (2) PRB. An approximately 33% higher capacity is obtained if AMR 7.95 is used instead of AMR 12.2. Furthermore, AMR 5.9 provides approximately 28% gains in capacity over AMR 7.95. The reduced gain is mainly due to a slightly increased number of retransmissions for AMR 5.9.

When comparing the schedulers in the downlink direction for experienced air interface delay, it is observed that due to the control channel limitations the relative amount of packets experiencing delay close to the used delay bound (50 ms) is clearly higher for the dynamic scheduler. With packet bundling, control channel limitations can be relaxed, and therefore delay critical packets can be scheduled earlier. Bundling itself adds delays in steps of 20 ms. For the semi-persistent scheduler, the time the packet has to wait in the packet scheduler buffer before initial scheduling takes place depends on the location of the persistent resource allocation in the time domain, which has a rather even distribution with a full load. Therefore the packet delay distribution for the semi-persistent scheduler is smooth. In all these schemes, retransmissions are scheduled at the earliest convenience, after an 8 ms HARQ cycle. Due to the first HARQ retransmission the distributions plateau at 10 ms. A cumulative distribution function for the packet delay is presented in Figure 10.12.

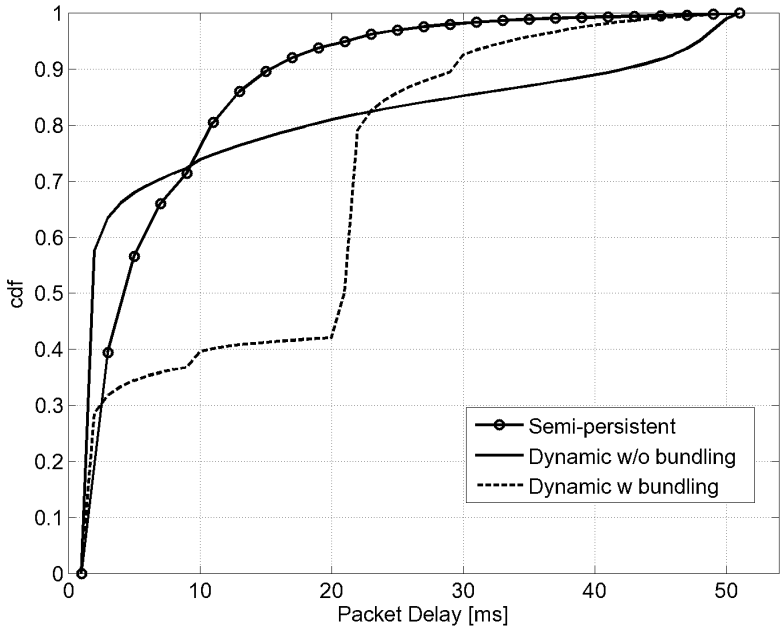


Figure 10.12 Cumulative distribution function for the experienced air interface delay in downlink with AMR 12.2

When summarizing VoIP system performance in downlink, it is concluded that the performance of the dynamic scheduler is control channel limited without packet bundling and hence the semi-persistent scheduler is able to have 50–125% higher capacities than the dynamic scheduler. The performance of the semi-persistent scheduler is user plane limited and the gains in capacity over the dynamic scheduler are smallest for AMR 12.2, which has the highest data rate amongst the simulated codecs and hence the biggest VoIP packet size. When the packet bundling is used with AMR 12.2, the control channel limitation for the performance of the dynamic scheduler is less significant compared with lower rate codecs as the number of supported users at 5% outage is lower. Therefore, the dynamic scheduler can have a 15% gain in capacity over the semi-persistent scheduler if AMR 12.2 is used with the packet bundling. When bundling is used with lower rate codecs, the control channel capacity starts to limit the performance of the dynamic scheduler due to the high number of supported users at 5% outage, and hence the semi-persistent scheduler achieves 8–15% higher capacities than the dynamic scheduler.

When comparing the performances of different scheduling methods in the uplink, it is observed that (similar to downlink) the performance of the dynamic scheduler is purely control channel limited whereas the semi-persistent scheduler suffers much less from control channel limitation due to a looser control channel requirement. Therefore, the semi-persistent scheduler is able to have 40–80% capacity gains over the dynamic scheduler with AMR 7.95 and AMR 5.9. Moreover, even with AMR 12.2 the semi-persistent scheduler achieves a 14% higher capacity than the dynamic scheduler.

The presented downlink results are obtained by using frequency dependent CQI reporting. In practice the large number of supported VoIP users in LTE may necessitate the use of wideband CQI to keep the overhead from CQI feedback at a reasonable level. This would mean a lower

Table 10.2 Relative loss in capacity (%) due to increased CQI granularity

	CQI reporting sub-band size	
	4 PRBs	Wideband CQI
Semi-persistent scheduling	2%	7%
Dynamic without bundling	0%	3%
Dynamic with bundling	7%	15%

uplink signaling overhead from CQI feedback at the cost of reduced capacity, as the frequency domain scheduling gains will be lost. To keep the capacity reduction as small as possible, the impact of lost frequency domain scheduling gains to performance should be compensated with more efficient use of frequency diversity. Therefore the use of the semi-persistent scheduler becomes even more attractive for VoIP traffic in LTE, as the performance of the semi-persistent scheduler is less dependent on the frequency domain scheduling gains than the dynamic scheduler. This is illustrated in Table 10.2, which presents the relative losses in capacity with AMR 12.2 when increasing the size of CQI reporting sub-band in the frequency domain. Relative losses are calculated against the presented capacity numbers, which were obtained assuming narrowband CQI reporting [7].

According to the simulations, the capacity of the dynamic scheduler is reduced by 15% due to the use of wideband CQI, whereas for the semi-persistent scheduler the corresponding loss is only 7%. Hence with wideband CQI the semi-persistent scheduler provides a similar performance to the dynamic scheduler for AMR 12.2, and for lower rate codecs the gains in capacity over the dynamic scheduler are increased further compared to the gains achieved with narrowband CQI.

Finally, in the light of the above results analysis, it seems that semi-persistent scheduling is the most attractive scheduling method for VoIP traffic in LTE. Nevertheless, as the used persistent allocation is indicated to the user via PDCCH signaling, some sort of combination of dynamic and semi-persistent scheduling methods seems to be the best option for VoIP in LTE. Furthermore, when comparing the performances of downlink and uplink, it is concluded that the VoIP performance in LTE is uplink limited. Hence downlink capacity can accommodate additional asymmetric data traffic, e.g. web browsing.

10.7 Voice Capacity Evolution

This section presents the evolution of the voice spectral efficiency from GSM to WCDMA/HSPA and to LTE. Most of the global mobile voice traffic is carried with GSM EFR or AMR coding. The GSM spectral efficiency can typically be measured with Effective Frequency Load (EFL), which represents the percentage of the time slots of full rate users that can be occupied in case of frequency reuse one. For example, EFL = 8% corresponds to 8% × 8 slots/200 kHz = 3.2 users/MHz. The simulations and the network measurements show that GSM EFR can achieve EFL 8% and GSM AMR can achieve 20% assuming all terminals are AMR capable and the network is optimized. The GSM spectral efficiency can be further pushed by the network feature Dynamic Frequency and Channel Allocation and by using Single Antenna Interference Cancellation (SAIC), known also as Downlink Advanced Receiver Performance (DARP). We assume up to EFL 35% with those enhancements. For further information see [8] and [9]. The

overhead from the Broadcast Control Channel (BCCH) is not included in the calculations. BCCH requires higher reuse than the hopping carriers.

WCDMA voice capacity is assumed to be 64 users with AMR 12.2 kbps, and 100 users with AMR 5.9 kbps on a 5 MHz carrier. HSPA voice capacity is assumed to be 123 users with AMR 12.2 kbps, and 184 users with AMR 5.9 kbps. For HSPA capacity evolution, see Chapter 13.

Capacity evolution is illustrated in Figure 10.13. LTE VoIP 12.2 kbps can provide efficiency which is 15× more than GSM EFR. The high efficiency can squeeze the voice traffic into a smaller spectrum. An example calculation is shown in Figure 10.14 assuming 1500 subscrib-

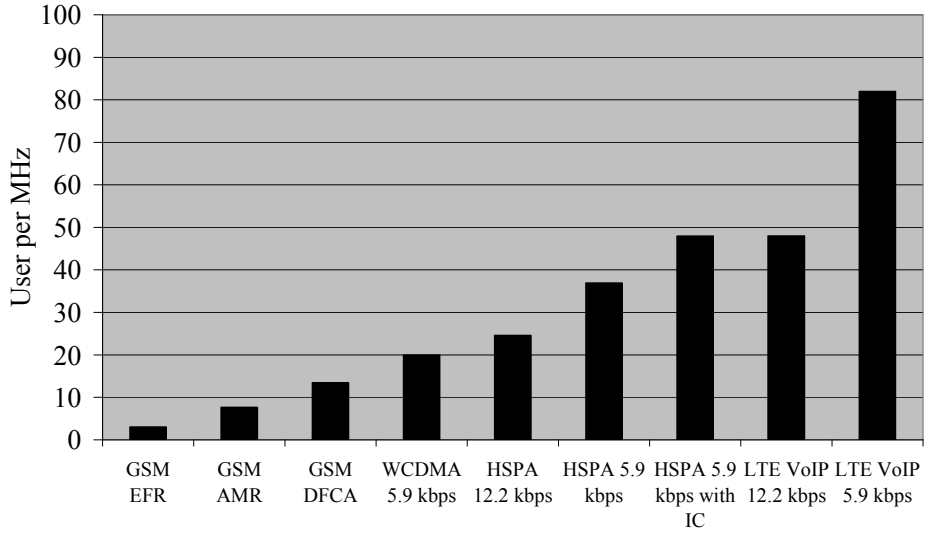


Figure 10.13 Voice spectral efficiency evolution (IC = Interference Cancellation)

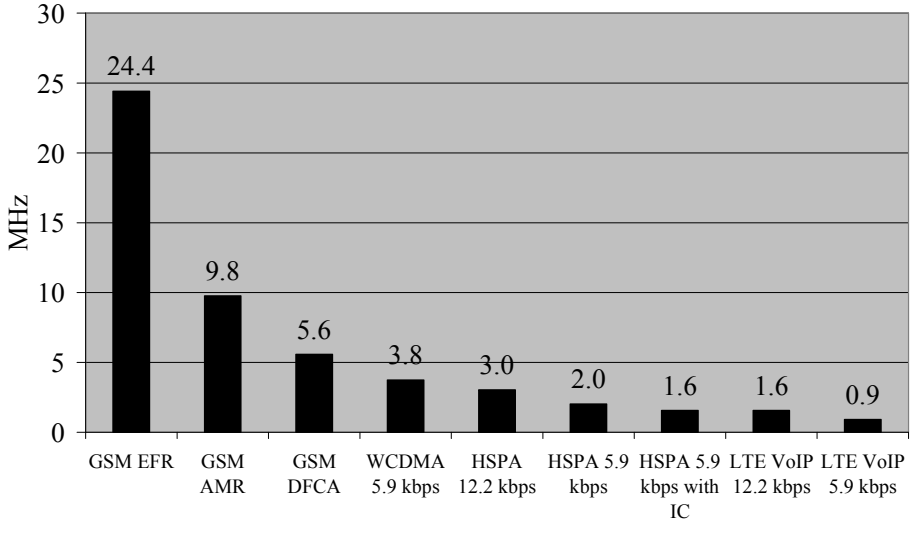


Figure 10.14 Spectrum required for supporting 1500 subscribers per sector at 40 mErl

ers per sector and 40 mErl traffic per subscriber. GSM EFR would take 25 MHz of spectrum while LTE will be able to carry that voice traffic in less than 2 MHz. LTE can then free up more spectrum for data usage.

10.8 Uplink Coverage

Uplink coverage can be maximized when UE transmits continuously with maximum power. Since VoIP packets are small (< 40 bytes), they can easily fit into 1 ms TTI. The AMR packets arrive every 20 ms leading to only 5% activity in the uplink. The uplink coverage could be improved by increasing the UE transmission time by using retransmissions and TTI bundling. These solutions are illustrated in Figure 10.15. The number of retransmissions must be limited for VoIP to remain within the maximum delay budget. The maximum number of retransmissions is six assuming a maximum 50 ms total delay, since the retransmission delay is 8 ms in LTE.

TTI bundling can repeat the same data in multiple (up to four) TTIs [4,10]. TTI bundling effectively increases the TTI length allowing the UE to transmit for a longer time. A single transport block is coded and transmitted in a set of consecutive TTIs. The same hybrid ARQ process number is used in each of the bundled TTIs. The bundled TTIs are treated as a single resource where only a single grant and a single acknowledgement are required. TTI bundling can be activated with a higher layer signaling per UE. The trigger could be, for example, UE reporting its transmit power is getting close to the maximum value.

The resulting eNodeB sensitivity values with retransmission and TTI bundling are shown in Table 10.3. The eNodeB sensitivity can be calculated as follows:

$$eNodeB\_sensitivity\,[dBm] = -174 + 10 \cdot \log_{10}(Bandwidth) + Noise\_figure + SNR \quad (10.2)$$

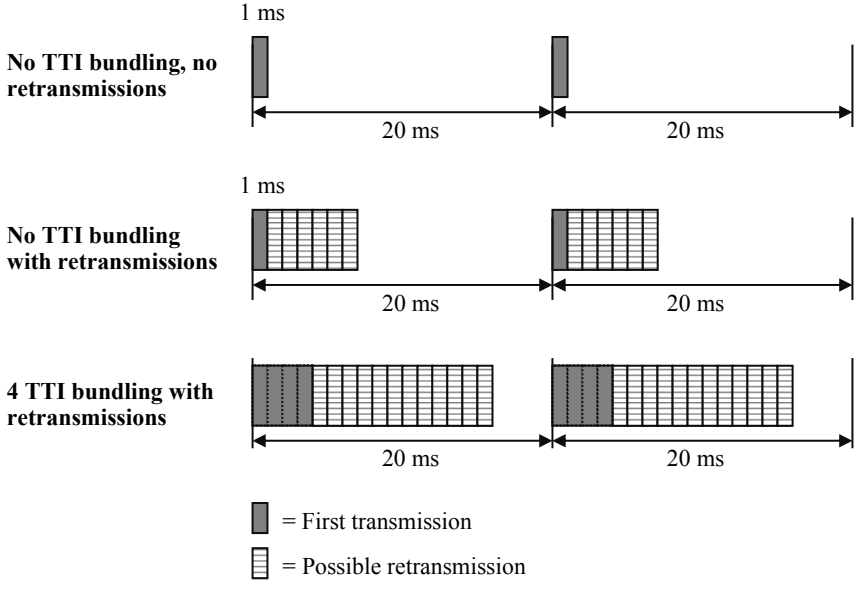


Figure 10.15 TTI bundling for enhancing VoIP coverage in uplink

Table 10.3 Uplink VoIP sensitivity with TTI bundling [11]

Number of TTIs bundled	1	4
Transmission bandwidth	360 kHz (2 resource blocks)	360 kHz (2 resource blocks)
Number of retransmissions	6	3
Required SNR	−4.2 dB	−8.0 dB
Receiver sensitivity	−120.6 dBm	−124.4 dBm

The eNodeB receiver noise figure is assumed to be 2 dB, two resource blocks are used for voice and no interference margin is included. The bundling can improve the uplink voice coverage by 4 dB.

The WCDMA NodeB sensitivity can be estimated by assuming Eb/N0 = 4.5 dB [12], which gives −126.6 dBm. To get similar voice coverage in LTE as in WCDMA we need to apply TTI bundling with a sufficient number of retransmissions. In terms of Hybrid-ARQ (HARQ) acknowledgement and retransmission timing, the TTI bundling method illustrated in Figure 10.16 is adopted for the LTE specifications [4,10]. According to this method four (4) subframes are in one bundle for the Frequency Division Duplex (FDD) system. Within a bundle, the operation is like autonomous retransmission by the UE in consecutive subframes without waiting for ACK/NACK feedback. The Redundancy Version (RV) on each autonomous retransmission in consecutive subframes changes in a pre-determined manner.

HARQ acknowledgement is generated after receiving the last subframe in the bundle. The timing relation between the last subframe in the bundle and the transmission instant of the HARQ acknowledgement is identical to the case of no bundling. If the last subframe in a bundle is subframe N then the acknowledgement is transmitted in subframe $N+4$. If the first subframe in a bundle is subframe k then any HARQ retransmissions begin in subframe k + 2 × HARQ RTT.

In the following, the impact of four (4) TTI bundling on capacity is studied via system level simulations in a bad coverage limited macro cell scenario 3 [2]. Main system simulation parameters are aligned with [7]. Table 10.4 presents both the energy accumulation and obtained

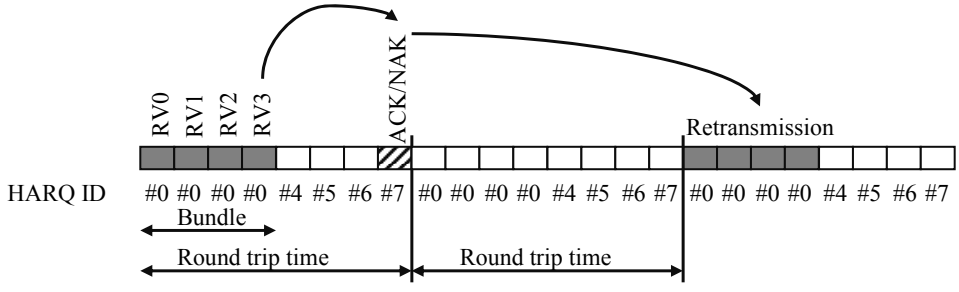


Figure 10.16 TTI bundling combining with HARQ

Table 10.4 Impact of TTI bundling on performance

	No bundling	TTI bundling
Maximal collected TTI per packet	7	12
VoIP capacity with TTI bundling at 5 MHz	<120	144

Table 10.5 Performance comparison for different delay budgets

Delay budget (ms)	50	60	70
Maximal collected TTI per packet	12	16	20
VoIP capacity with TTI bundling at 5 MHz	144	155	163

capacity for AMR 12.2 when dynamic scheduling with 50 ms packet delay budget is used. By using TTI bundling, 2.34 dB energy accumulation gain can be achieved, so VoIP capacity can be improved by 20% at least by using TTI bundling compared to no bundling.

Moreover, as the results in Table 10.5 show, the air interface delay has a clear impact on the achieved VoIP capacity and also on the gains that can be obtained from TTI bundling, i.e., with a looser delay budget capacity gains from TTI bundling are increased. This is evident as the longer air interface delay allows more energy to be accumulated from bundled TTIs, e.g. for 50 ms air interface delay, the packet will cumulate the energy from at most 12 TTIs assuming four (4) TTI bundling. Furthermore, for 70 ms air interface delay the corresponding number is increased to 20 TTIs, which would mean that 66% more energy could be accumulated. Hence the longer air interface delay implies bigger combining gains for the received signal from the bundled TTIs, and therefore the coverage is increased as the air interface delay is increased. This is again mapped to capacity gains. Besides, a longer delay budget means that the realized time domain scheduling gains are increased as each packet has more opportunities to be scheduled, and this further improves the coverage and hence improves the capacity. All in all, due to this reasoning the capacity with a 70 ms delay bound is approximately 13% higher than the capacity with a 50 ms delay bound.

10.9 Circuit Switched Fallback for LTE

Voice service with LTE terminals can also be offered before high quality VoIP support is included into LTE radio and before IP Multimedia Subsystem (IMS) is deployed. The first phase can use so-called Circuit Switched (CS) fallback for LTE where the LTE terminal will be moved to GSM, WCDMA or CDMA network to provide the same services that exists in CS networks today, for example voice calls, video call or Location Services [13]. The CS fallback procedures require that the Mobility Management Entity (MME) as well as Mobile Switching Center (MSC)/Visiting Location Register (VLR) network elements are upgraded to support procedures described in this section. The CS fallback is described in [13].

When LTE UE executes the attach procedure towards the LTE core network, the LTE core network will also execute a location update towards the serving CS core network to announce the presence of the terminal to the CS core network via the LTE network in addition to executing the normal attach procedures. The UE sends the attach request together with specific 'CS Fallback Indicator' to the MME, which starts the location update procedure towards MSC/VLR via an IP based SGs interface. The new Location Area Identity (LAI) is determined in the MME based on mapping from the Tracking area. A mapped LAI could be either GERAN or UTRAN based on the operator configuration. Additionally GERAN and UTRAN can be served by different MSC/VLR in the network architecture, which causes execution of a roaming retry for the CS fallback procedure as described later. The VLR creates an association with

the MME. This procedure is similar to the combined LA/RA update supported in GERAN by using Network Mode of Operation 1 (NMO1) with Gs interface.

If MME supports the connection to the multiple core network nodes similarly as SGSN does in GERAN/UTRAN, then a single MME can be connected to multiple MSC/VLR network elements in the CS core network. The benefit is the increased overall network resiliency compared to the situation when the MME is connected to only a single MSC/VLR.

For a mobile terminated call, MSC/VLR sends a paging message via SGs interface to the correct MME based on the location update information. eNodeB triggers the packet handover or network assisted cell change from LTE to the target system. The UE sends the paging response in the target system and proceeds with the CS call setup.

An overview of the mobile terminated fallback handover is shown in Figure 10.17.

An overview of the mobile originated call is shown in Figure 10.18. The UE sends a CS call request to the eNodeB, which may ask the UE to measure the target cell. eNodeB triggers the handover to the target system. the UE starts the normal CS call establishment procedure in the target cell. Once the CS call ends, the UE again reselects LTE to get access to high data rate capabilities.

If the service areas covered by LTE Tracking Area (TA) and GERAN/UTRAN Location Area (LA) are not similar, the serving MSC/VLR of LTE UE is different from MSC/VLR that serves the target LA. The target MSC/VLR is not aware of the UE and is not expecting a response to the paging message. This case is handled with a procedure called as Roaming Retry for CS fallback. When LTE UE has received the paging request, it executes an additional

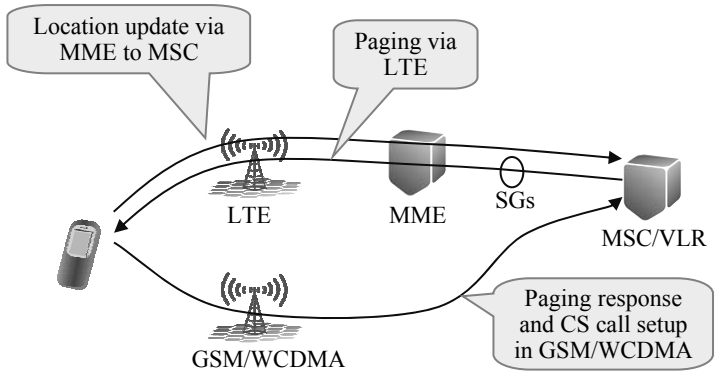


Figure 10.17 Circuit switched fallback handover – mobile terminated call

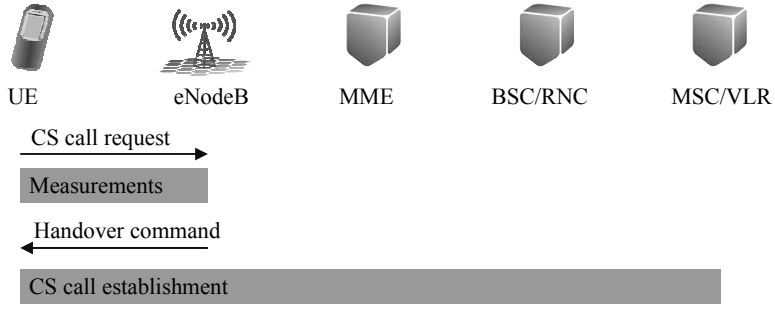


Figure 10.18 Circuit switched fallback for LTE – mobile originated call

location update procedure to the target MSC/VLR to make the target MSC/VLR aware of the existence of the UE. Based on this location update the HLR of the LTE UE will send a cancel message to the MSC/VLR that was serving the LTE UE prior to the CS fallback attempt. This cancel message will stop the paging process in that MSC/VLR and will trigger the resume call handling procedure. This procedure is similar to the Mobile Terminating Roaming Retry Call procedure used in GERAN/UTRAN to request a call to be re-routed to a new MSC/VLR from a gateway MSC. The gateway MSC will re-execute the mobile terminating call procedure by executing a HLR enquiry resulting in the call being routed to the correct MSC/VLR.

If there is an active packet session in LTE and if the target system supports simultaneous voice and data connection, the packet session is relocated from LTE to the target system. If the target system does not support simultaneous voice and data, the data connection is suspended during the CS voice call. This happens when LTE UE is moved to a GERAN that does not support simultaneous voice and data connection, called Dual Transfer Mode (DTM).

Voice is one of the most demanding services in terms of delay requirements and in terms of call reliability. Many optimized GERAN (GSM) and UTRAN (WCDMA) networks today can provide call drop rates below 0.5% or even below 0.3%. The advantage of the fallback handover is that there is no urgent need to implement QoS support in LTE or SAE. The network optimization may also be less stringent if the voice is not supported initially. Also, there is no need to deploy IMS just because of the voice service. From the end user point of view, the voice service in GSM or in WCDMA is as good as voice service in LTE, except that the highest data rates are not available during the voice call. On the other hand, the CS fallback for LTE adds some delay to the call setup process as the UE must search for the target cell. With high mobility, the CS fallback for LTE may also affect the call setup success rate, which naturally needs to be addressed in practical implementation.

Similar fallback procedures can also be used for other services familiar from today's network such as Unstructured Supplementary Service Data (USSD) and Location Services (LCS). Due to the popularity of SMS, the number of SMSs can be very high and the CS fallback handover may create a large number of reselections between LTE and GERAN/UTRAN. Therefore, it was seen as the preferred way to transmit SMSs over LTE even if the voice calls would use the CS fallback solution. If the core network has support for IMS, then SMS body can be transferred over SIP messages in LTE as defined by 3GPP [14].

10.10 Single Radio Voice Call Continuity (SR-VCC)

Once the VoIP call is supported in LTE, there may still be a need to make a handover from LTE VoIP to a GSM, WCDMA or CDMA CS voice network when running out of LTE coverage. The handover functionality from VoIP to the CS domain is referred to as Single Radio Voice Call Continuity (SR-VCC). The solution does not require UE capability to simultaneously signal on two different radio access technologies – therefore it is called a Single Radio Solution. SR-VCC is illustrated in Figure 10.19. SR-VCC is defined in [15]. The Dual radio Voice call continuity was already defined in 3GPP Release 7 for voice call continuation between WLAN and GSM/WCDMA networks [16].

The selection of the domain or radio access is under the network control in SR-VCC and SR-VCC enhanced MSC Server (called 'S-IWF' in this chapter) deployed into the CS core network. This architecture has been defined to enable re-use of already deployed CS core network assets to provide the necessary functionality to assist in SR-VCC. This architecture option is illustrated in Figure 10.20.

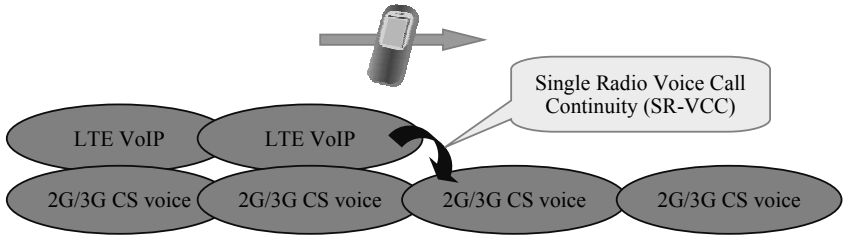


Figure 10.19 Single radio voice call continuity (SR-VCC)

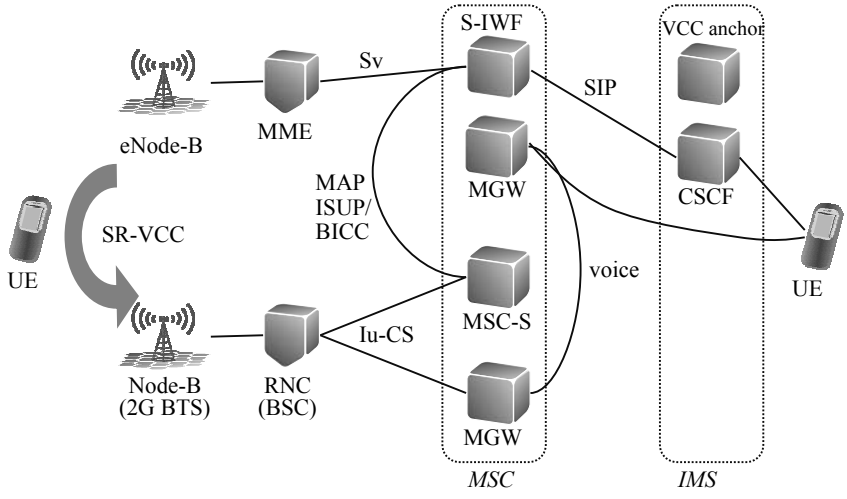


Figure 10.20 Architecture for SR-VCC. CSCF, call session control function; MGW, media gateway; MAP, mobile application port; ISUP, ISDN user part; BICC, bearer independent call control protocol

S-IWF, co-located in the MSC Server, uses a new GTP based Sv interface towards MME function in the LTE domain to trigger the SR-VCC procedure. Architecturally the S-IWF acts similarly to an anchor MSC Server towards target 2G/3G CS domain and is also responsible for preparing the needed target radio access network resources jointly with target GERAN or UTRAN. S-IWF also connects the speech path from the target CS domain towards the other subscriber in the voice call and together with the VCC anchor also hides mobility between LTE VoIP and 2G/3G CS domain from the other side of the call. The VCC anchor is located at the IMS application server and based on the same concept that was defined by 3GPP for Release 7 WLAN Voice Call Continuity.

S-IWF implementation has been specified so that it may be deployed with IMS Centralized Service (ICS) architecture so that the S-IWF then uses Mw interface (acting as the MSC Server enhanced for ICS) towards IMS instead of, for instance, the ISUP interface. In the former situation, handover causes registration prior to a new call towards IMS on behalf of the served subscriber whereas in the latter situation no registration is needed. This kind of support for SR-VCC with both ICS and non-ICS architectures is essential to achieve the required deployment flexibility.

During the LTE attach procedure, the MME will receive the required SR-VCC Domain Transfer Number from HSS that is then given to S-IWF via the Sv interface. S-IWF uses this

number to establish a connection to the VCC anchor during the SR-VCC procedure. When the VoIP session is established, it will be anchored within IMS in the VCC anchor to use SR-VCC in case it is needed later in the session. This anchoring occurs in both the originating and terminating sessions based on IMS subscription configuration.

The SR-VCC procedure is presented in Figure 10.21. LTE eNodeB first starts inter-system measurements of the target CS system. eNodeB sends the handover request to the MME, which then triggers the SR-VCC procedure via the Sv interface to the MSC-Server S-IWF functionality with Forward Relocation Request. S-IWF in the MSC-Server initiates the session transfer procedure towards IMS by establishing a new session towards the VCC anchor that originally anchored the session. The session is established by S-IWF using the SR-VCC number provided by MME. S-IWF also coordinates the resource reservation in the target cell together with the target radio access network. The MSC Server then sends a Forward Relocation Response to MME, which includes the necessary information for the UE to access the target GERAN/UTRAN cell.

After the SR-VCC procedure has been completed successfully the VoIP connection is present from Media Gateway that is controlled by S-IWF towards the other side of an ongoing session. The CS connection exists towards the radio access network to which the UE was moved during the procedure.

S-IWF functionality is not involved in a normal LTE VoIP session if the handover to CS domain is not required.

For a simultaneous voice and non-voice data connection, the handling of the non-voice bearer is carried out by the bearer splitting function in the MME. The MME may suppress

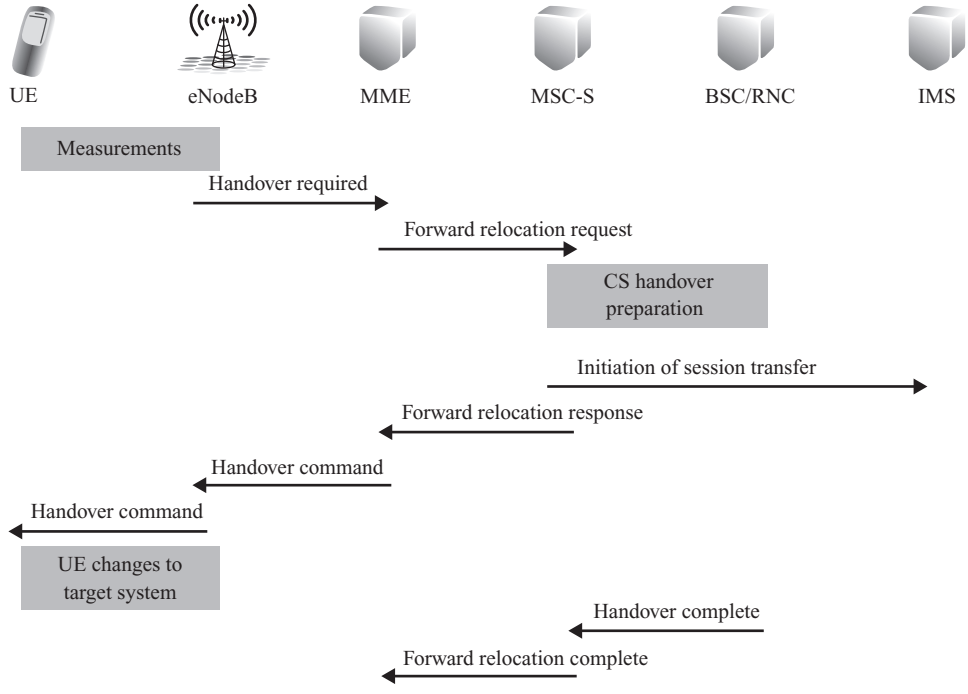


Figure 10.21 SR-VCC procedure

the handover of a non-voice PS bearer during the SR-VCC procedure. This may happen if the target GERAN system does not support simultaneous voice and data functionality (DTM). If the non-voice bearer is also handed over, the process is done in the same way as the normal inter-system handover for packet services. The MME is responsible for coordinating the Forward Relocation Response from the SR-VCC and the packet data handover procedure.

For roaming, the Visited Public Land Mobile Network (PLMN) controls the radio access and domain change while taking into account any related Home PLMN policies.

10.11 Summary

LTE radio is not only optimized for high data rates and high data capacities, but also for high quality voice with high capacity. The system simulations show that LTE can support 50–80 simultaneous voice users per MHz per cell in the macro cellular environment with AMR data rates between 12.2 and 5.9 kbps. The lower AMR rates offer the highest capacity. The highest capacity is achieved using semi-persistent packet scheduling where the first transmission has pre-allocated resources and only retransmissions are scheduled. The capacity with fully dynamic scheduling is limited by control channel capacity since each voice packet needs to be scheduled separately.

Voice quality depends on the voice codec and on the end-to-end delay. It is assumed that in the first phase, VoIP in LTE uses already standardized 3GPP speech codec AMR narrowband and AMR wideband due to backward compatibility with the legacy network. For end-to-end voice delay, LTE can provide values below 200 ms offering a similar or lower delay than the existing CS networks.

LTE uplink voice coverage can be optimized by using TTI bundling, which allows the combination of multiple 1-ms TTIs to increase the uplink transmission time and the average transmitted energy. The LTE voice coverage with TTI bundling is similar to the existing 2G/3G voice service.

As LTE supports only packet based services, then the voice service is also based on Voice over IP (VoIP). LTE specifications include definitions for inter-working with the existing Circuit Switched (CS) networks. One option is to use CS fallback for LTE where the UE moves to the 2G/3G network to make the voice call using a CS connection and returns to LTE when the voice call is over. Another option is to use VoIP in LTE and then allow inter-system handover to the 2G/3G CS network when LTE coverage is running out. This handover is called Single Radio Voice Call Continuity (SR-VCC). The CS fallback handover has been assumed to enable re-use of existing deployed CS core network investments as well as to provide a similar look and feel for the end users who are accustomed to current mobile network services. In some cases this kind of solution can be assumed to provide the first phase solution, but the IMS based core network architecture with support for VoIP and Multimedia Telephony is considered to be the long term solution.

Both CS fallback for LTE as well as the LTE based VoIP with SR-VCC solution can be deployed either in a phased manner or simultaneously depending on terminal capabilities. Additionally it is also possible to deploy IMS for non-VoIP services such as video sharing, presence or instant messaging, even primary voice service, by using CS fallback for LTE procedures. This kind of architecture ensures maximum flexibility for the operator to plan and deploy LTE as well as IMS solutions.

References

[1] 3GPP Technical Specifications 23.053 'Tandem Free Operation (TFO); Service description; Stage 2', V.8.3.0
[2] ITU-T Recommendation G.114 'One way transmission time', May 2003.
[3] 3GPP Technical Specification 23.203 'Policy and charging control architecture', V.8.3.1.
[4] 3GPP Technical Specifications 36.321 'Medium Access Control (MAC) protocol specification', V.8.3.0.
[5] 3GPP TS36.300, 'Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 (Release 8)'.
[6] 3GPP TR 25.814 v7.0.0, 'Physical layer aspect for evolved Universal Terrestrial Radio Access (UTRA)'.
[7] 'Next Generation Mobile Networks (NGMN) Radio Access Performance Evaluation Methodology', A White paper by the NGMN Alliance, January 2008.
[8] T. Halonen, J. Romero, J. Melero, 'GSM, GPRS and EDGE Performance', 2nd Edition, John Wiley & Sons, 2003.
[9] A. Barreto, L. Garcia, E. Souza, 'GERAN Evolution for Increased Speech Capacity', Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th, April 2007.
[10] 3GPP R2-082871, 'LS on TTI Bundling', RAN1#62, Kansas City, USA, May 2008.
[11] R1-081856, 'Coverage Comparison between PUSCH, PUCCH and RACH', Nokia/Nokia Siemens Networks, Catt, RAN1#54, Kansas City, USA, May 2008.
[12] H. Holma, A. Toskala, 'WCDMA for UMTS', 4th Edition, John Wiley & Sons, 2007.
[13] 3GPP Technical Specifications 23.272 'Circuit Switched (CS) fallback in Evolved Packet System (EPS)', V. 8.1.0.
[14] 3GPP Technical Specifications 23.204 'Support of Short Message Service (SMS) over generic 3GPP Internet Protocol (IP) access', V.8.3.0.
[15] 3GPP Technical Specifications 23.216 'Single Radio Voice Call Continuity (SRVCC)', V. 8.1.0.
[16] 3GPP Technical Specifications 23.206 'Voice Call Continuity (VCC) between Circuit Switched (CS) and IP Multimedia Subsystem (IMS)', V.7.5.0.

11

Performance Requirements

Andrea Ancora, Iwajlo Angelow, Dominique Brunel, Chris Callender, Harri Holma, Peter Muszynski, Earl McCune and Laurent Noël

11.1 Introduction

3GPP defines minimum Radio Frequency (RF) performance requirements for terminals (UE) and for base stations (eNodeB). These performance requirements are an essential part of the LTE standard as they facilitate a consistent and predictable system performance in a multi-vendor environment. The relevant specifications are given in [1], [2], [3] covering both duplex modes of LTE: Frequency Division Duplex (FDD) and Time Division Duplex (TDD).

Some of the RF performance requirements, e.g. limits for unwanted emissions, facilitate the mutual coexistence of LTE with adjacent LTE systems or adjacent 2G/3G systems run by different operators without coordination. The corresponding requirements may have been derived from either regulatory requirements or from coexistence studies within 3GPP; see [4].

Other performance requirements, e.g. modulation accuracy and baseband demodulation performance requirements, provide operators with assurance of sufficient performance within their deployed LTE carrier.

[2] is the basis for the eNodeB test specification [5], which defines the necessary tests for type approval. Correspondingly, [1] is the basis on the UE side for the test specifications [6] and [7].

This chapter presents the most important LTE minimum performance requirements, the rationale underlying these requirements and the implications for system performance and equipment design. Both RF and baseband requirements are considered. First the eNodeB requirements are considered and then the UE requirements.

11.2 Frequency Bands and Channel Arrangements

11.2.1 Frequency Bands

Table 11.1 lists the currently defined LTE frequency bands, together with the corresponding duplex mode (FDD or TDD). There are currently 17 bands defined for FDD and 8 bands for

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6
DigRF is a service mark of The MIPI Alliance, Inc. and this chapter is printed with the permission of The MIPI Alliance, Inc.

Table 11.1 LTE frequency bands

LTE Band	Uplink eNode B receive UE transmit			Downlink eNode B transmit UE receive			Duplex mode
1	1920 MHz	–	1980 MHz	2110 MHz	–	2170 MHz	FDD
2	1850 MHz	–	1910 MHz	1930 MHz	–	1990 MHz	FDD
3	1710 MHz	–	1785 MHz	1805 MHz	–	1880 MHz	FDD
4	1710 MHz	–	1755 MHz	2110 MHz	–	2155 MHz	FDD
5	824 MHz	–	849 MHz	869 MHz	–	894 MHz	FDD
6	830 MHz	–	840 MHz	875 MHz	–	885 MHz	FDD
7	2500 MHz	–	2570 MHz	2620 MHz	–	2690 MHz	FDD
8	880 MHz	–	915 MHz	925 MHz	–	960 MHz	FDD
9	1749.9 MHz	–	1784.9 MHz	1844.9 MHz	–	1879.9 MHz	FDD
10	1710 MHz	–	1770 MHz	2110 MHz	–	2170 MHz	FDD
11	1427.9 MHz	–	1452.9 MHz	1475.9 MHz	–	1500.9 MHz	FDD
12	698 MHz	–	716 MHz	728 MHz	–	746 MHz	FDD
13	777 MHz	–	787 MHz	746 MHz	–	756 MHz	FDD
14	788 MHz	–	798 MHz	758 MHz	–	768 MHz	FDD
17	704 MHz	–	716 MHz	734 MHz	–	746 MHz	FDD
18	815 MHz	–	830 MHz	860 MHz	–	875 MHz	FDD
19	830 MHz	–	845 MHz	875 MHz	–	890 MHz	FDD
…							
33	1900 MHz	–	1920 MHz	1900 MHz	–	1920 MHz	TDD
34	2010 MHz	–	2025 MHz	2010 MHz	–	2025 MHz	TDD
35	1850 MHz	–	1910 MHz	1850 MHz	–	1910 MHz	TDD
36	1930 MHz	–	1990 MHz	1930 MHz	–	1990 MHz	TDD
37	1910 MHz	–	1930 MHz	1910 MHz	–	1930 MHz	TDD
38	2570 MHz	–	2620 MHz	2570 MHz	–	2620 MHz	TDD
39	1880 MHz	–	1920 MHz	1880 MHz	–	1920 MHz	TDD
40	2300 MHz	–	2400 MHz	2300 MHz	–	2400 MHz	TDD

TDD. Whenever possible, the RF requirements for FDD and TDD have been kept identical to maximize the commonality between the duplex modes.

While the physical layer specification and many RF requirements are identical for these frequency bands, there are some exceptions to this rule for the UE RF specifications, as will be discussed further in UE related sections in this chapter. On the other hand, the eNodeB RF requirements are defined in a frequency band agnostic manner as there are less implementation constraints for base stations. If the need arises, further LTE frequency bands can be easily added affecting only isolated parts of the RF specifications. Furthermore, the specified LTE frequency variants are independent of the underlying LTE release feature content (Release 8, 9, etc.). Those frequency variants that will be added during the Release 9 time frame can still be implemented using just Release 8 features.

The band numbers 15 and 16 are skipped since those numbers were used in ETSI specifications.

Not all of these bands are available in each of the world's regions. Table 11.2 illustrates the typical deployment areas for the different FDD frequency variants.

The most relevant FDD bands in Europe are:

Table 11.2 Usage of the frequency variants within the world's regions

Band	Europe	Asia	Japan	Americas
1	X	X	X	
2				X
3	X	X		
4				X
5		X		X
6			X	
7	X	X		
8	X	X		
9			X	
10				X
11			X	
12				X
13			X	
14				X
15				X
16				X
17				X
18			X	

- Band 7 is the new 2.6 GHz band. The 2.6 GHz auctions have been running in a few countries during 2007 and 2008, and continue during 2009.
- Band 8 is currently used mostly by GSM. The band is attractive from a coverage point of view due to the lower propagation losses. The band can be reused for LTE or for HSPA. The first commercial HSPA900 network started in Finland in November 2007. LTE refarming is considered in Chapter 9.
- Band 3 is also used by GSM, but in many cases Band 3 is not as heavily used by GSM as Band 8. That makes refarming for LTE simpler. Band 3 is also important in Asia where Band 7 is not generally available.

Correspondingly, the new bands in the USA are Bands 4, 12, 13, 14 and 17. Bands 2 and 5 can be used for LTE refarming. The LTE deployment in Japan will use Bands 1, 9, 11 and 18. In summary, LTE deployments globally will use several different frequency bands from the start.

Further frequencies for International Mobile Telephony (IMT) have been identified at the ITU WRC 2007 conference (see Chapter 1).

11.2.2 Channel Bandwidth

The width of a LTE carrier is defined by the concepts of Channel bandwidth ($BW_{Channel}$) and Transmission bandwidth configuration (N_{RB}) (see Figure 11.1). Their mutual relationship is specified in Table 11.3.

The transmission bandwidth configuration, N_{RB}, is defined as the maximum number of Resource Blocks (RB) that can be allocated within a LTE RF channel. A RB comprises 12

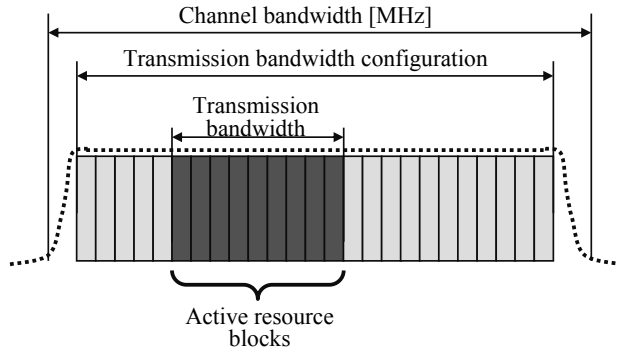


Figure 11.1 Definition of Channel Bandwidth and Transmission Bandwidth Configuration for one LTE carrier

Table 11.3 Transmission bandwidth configuration N_{RB} in LTE channel bandwidths

Channel bandwidth $BW_{Channel}$ [MHz]	1.4	3	5	10	15	20
Transmission bandwidth configuration N_{RB}	6	15	25	50	75	100

sub-carriers and can thus be understood to possess a nominal bandwidth of 180 kHz. While the physical layer specifications allow N_{RB} to assume any value within the range $6 \leq N_{RB} \leq 110$, all RF requirements (and thus a complete LTE specification) are only defined for the values in Table 11.3. This flexibility within the LTE specifications, however, supports the addition of further options for transmission bandwidth configurations (channel bandwidths), should the need arise.

The channel bandwidth is the relevant RF related parameter for defining Out-of-band (OOB) emission requirements (e.g. spectrum emission mask, Adjacent Channel Leakage Ratio [ACLR]). The RF channel edges are defined as the lowest and highest frequencies of the carrier separated by the channel bandwidth, i.e. at $F_C \pm BW_{Channel}/2$.

Note from Table 11.3 that the transmission bandwidth configuration measures only 90% of the channel bandwidth for 3, 5, 10, 15, and 20 MHz LTE and less so for 1.4 MHz LTE; e.g. for 5 MHz LTE a transmission bandwidth of $25 \times 180\,\text{kHz} = 4.5\,\text{MHz}$ is obtained. In fact, the basic OFDM spectrum comprises only slowly decaying sidelobes, with a rolloff as described by the *sinc()* function. Therefore, efficient usage of the spectrum requires the use of filtering to effectively confine the OOB emissions. Such filters, however, require a certain amount of transition bandwidth in order to be (a) practical and (b) to consume only a small amount of the cyclic prefix duration due to the inevitable time dispersion they will cause. The values in Table 11.3, in particular for the LTE channel bandwidth of 1.4 MHz, are therefore a compromise between in- and out-of-channel distortions and were extensively studied in 3GPP. The transmission bandwidth is 77% of the channel bandwidth for LTE 1.4 MHz.

The 1.4 and 3 MHz channel bandwidth options have been chosen to facilitate a multitude of cdma2000® and GSM migration scenarios within Bands 2, 3, 5 and 8.

Not all combinations of LTE frequency bands and channel bandwidth options are meaningful, and Table 11.4 provides the combinations supported by the standard. These combinations were based on inputs from the operators. Note from Table 11.4 that the options with a wider channel

Table 11.4 Supported transmission bandwidths with normal sensitivity ('X') and relaxed sensitivity ('O')

Channel bandwidth						
E-UTRA Band	1.4 MHz	3 MHz	5 MHz	10 MHz	15 MHz	20 MHz
1			X	X	X	X
2	X	X	X	X	O	O
3	X	X	X	X	O	O
4	X	X	X	X	X	X
5	X	X	X	O		
6			X	O		
7			X	X	X	O
8	X	X	X	O		
9			X	X	O	O
10			X	X	X	X
11			X	O	O	O
12						
13	X	X	O	O		
14	X	X	O	O		
17						
…						
33			X	X	X	X
34			X	X	X	
35	X	X	X	X	X	X
36	X	X	X	X	X	X
37			X	X	X	X
38			X	X		
39			X	X	X	X
40				X	X	X

bandwidth are typically supported for bands with a larger amount of available spectrum, e.g. 20 MHz LTE is supported in Bands 1, 2, 3 and 4 but not in Bands 5, 8, etc. Conversely, LTE channel bandwidths below 5 MHz are typically supported in frequency bands with either a smaller amount of available spectrum (e.g. Bands 5, 8, etc.) or bands exhibiting 2G migration scenarios (e.g. Bands 2, 5 and 8).

Furthermore [1] defines two levels of UE sensitivity requirements depending on the combination of bandwidths and operating band; these are also listed in Table 11.4. This is because there will be an increased UE self-interference in those frequency bands which possess a small duplex separation and/or gap. For large bandwidth and small duplex separation, certain relaxations of the UE performance are allowed or UE functionality is limited. For other cases, the UE needs to meet the baseline RF performance requirements.

11.2.3 Channel Arrangements

The channel raster is 100 kHz for all E-UTRA frequency bands, which means that the carrier centre frequency must be an integer multiple of 100 kHz. For comparison, UMTS uses a 200 kHz channel raster and for some frequency bands requires additional RF channels, offset by 100 kHz relative to the baseline raster, in order to effectively support the deployment within

a 5 MHz spectrum allocation. The channel raster of 100 kHz for LTE will support a wide range of migration cases.

The spacing between carriers will depend on the deployment scenario, the size of the frequency block available and the channel bandwidths. The nominal channel spacing between two adjacent LTE carriers is defined as follows:

$$\text{Nominal channel spacing} = (BW_{Channel(1)} + BW_{Channel(2)})/2 \tag{11.1}$$

where $BW_{Channel(1)}$ and $BW_{Channel(2)}$ are the channel bandwidths of the two respective LTE carriers. The LTE–LTE coexistence studies within 3GPP [4] assumed this channel spacing.

The nominal channel spacing can be adjusted, however, to optimize performance in a particular deployment scenario, e.g. for coordination between adjacent LTE carriers.

11.3 eNodeB RF Transmitter

LTE eNodeB transmitter RF requirements are defined in [2] and the corresponding test specification in [5]. In this section we discuss two of the most important transmitter requirements:

- Unwanted emissions, both inside and outside the operating band. The corresponding requirements will ensure the RF compatibility of the LTE downlink with systems operating in adjacent (or other) frequency bands.
- Transmitted signal (modulation) quality, or also known as Error Vector Magnitude (EVM) requirements. These requirements will determine the in-channel performance for the transmit portion of the downlink.

Coexistence of LTE with other systems, in particular between the LTE FDD and TDD modes, will also be discussed, both in terms of the 3GPP specifications as well as within the emerging regulative framework within Europe for Band 7.

11.3.1 Operating Band Unwanted Emissions

For WCDMA the spurious emissions requirements as recommended in ITU-R SM.329 are applicable for frequencies that are greater than 12.5 MHz away from the carrier centre frequency. The 12.5 MHz value is derived as 250% of the necessary bandwidth (5 MHz for WCDMA) as per ITU-R SM.329. The frequency range within 250% of the necessary bandwidth around the carrier centre may be referred to as the OOB domain. Transmitter intermodulation distortions manifest themselves predominantly within the OOB domain and therefore more relaxed emission requirements, such as ACLR, are typically applied within the OOB domain.

In LTE, the channel bandwidth can range from 1.4 to 20 MHz. A similar scaling by 250% of the channel bandwidth would result in a large OOB domain for LTE: for the 20 MHz LTE channel bandwidth option the OOB domain would extend to a large frequency range of up to ± 50 MHz around the carrier centre frequency. To protect services in adjacent bands in a more predictable manner and to align with WCDMA, the LTE spurious domain is defined to start from 10 MHz below the lowest frequency of the eNodeB transmitter operating band and from 10 MHz above the highest frequency of the eNodeB transmitter operating band, as shown in

Figure 11.2. The operating band plus 10 MHz on each side are covered by the LTE operating band unwanted emissions.

LTE spurious domain emission limits in [2] are defined as per ITU-R SM.329 and are divided into several categories, where Categories A and B are applied as regional requirements. Within Europe the Category B limit of −30 dBm/MHz is required in the frequency range 1 GHz ↔ 12.75 GHz while the corresponding Category A limit applied in the Americas and in Japan is −13 dBm/MHz.

In addition to the ITU-R SM.329 spurious emission limits, [2] defines more stringent limits across the operating bands of various wireless systems, including WCDMA, GSM and Personal Handyphone System (PHS).

The Operating band unwanted emission limits are defined as absolute limits by a mask that stretches from 10 MHz below the lowest frequency of the eNodeB transmitter operating band up to 10 MHz above the highest frequency of the eNodeB transmitter operating band, as shown in Figure 11.2.

This mask depends on the LTE channel bandwidth and is shown in Figure 11.3 for LTE bands >1 GHz and the limit of −25 dBm in 100 kHz as the lower bound for the unwanted emission limits. The limit of -25 dBm in 100 kHz is consistent with the level used for UTRA as a spurious emission limit inside the operating band (−15 dBm/MHz). The measurement bandwidth of 100 kHz was chosen to be of similar granularity as the bandwidth of any victim system's smallest radio resource allocation (LTE 1 RB at 180 kHz; GSM 200 kHz carrier).

The operating band unwanted emission limits must also be consistent with the limits according to ITU-R SM.329 for all LTE channel bandwidth options. This means that outside 250% of the necessary bandwidth from the carrier centre frequency, the corresponding Category B limit of −25 dBm/100 kHz must be reached (see Figure 11.3). Even though the frequency range in which transmitter intermodulation distortions occur scales with the channel bandwidth, it was found feasible to define a common mask for the 5, 10, 15 and 20 MHz LTE options which meets the SM.329 limits at 10 MHz offset from the channel edge. However, this −25 dBm/100 kHz limit must already be reached with offsets of 2.8 MHz and 6 Mhz respectively from the channel edge for the LTE 1.4 MHz and LTE 3 MHz, and this necessitated definition of separate masks. For

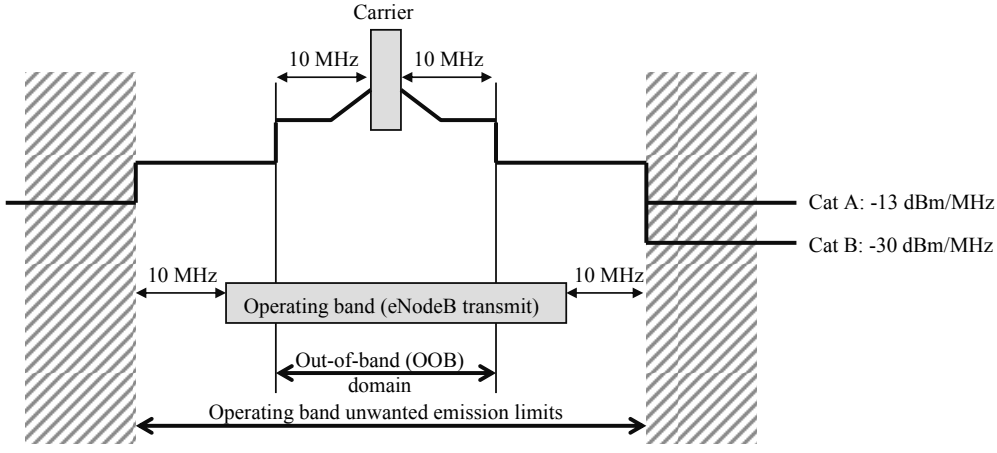


Figure 11.2 Defined frequency ranges for spurious emissions and operating band unwanted emissions

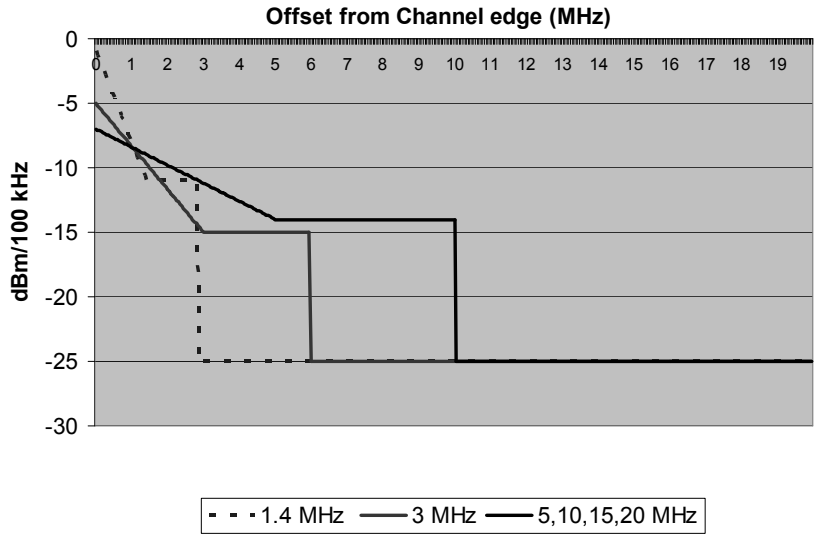


Figure 11.3 Operating band unwanted emission requirements levels relative to channel edge (E-UTRA bands >1 GHz)

the 1.4, 3 and 5 MHz LTE the same total eNodeB transmission power of 46 dBm was assumed resulting in a higher power spectral density for the smaller LTE bandwidth options, hence the mask permits higher emission levels at the channel edge.

For the North American LTE Bands (Bands 2, 4, 5, 10, 12, 13, 14, 17) an additional unwanted emission limit is derived in [2] from FCC Title 47 Parts 22, 24 and 27.These requirements are interpreted as −13 dBm in a measurement bandwidth defined as 1% of the '−26 dB modulation bandwidth' within the first MHz from the channel edge and −13 dBm/MHz elsewhere.

11.3.2 Coexistence with Other Systems on Adjacent Carriers Within the Same Operating Band

The RAN4 specifications also include ACLR requirements of 45 dBc for the 1st and 2nd adjacent channels of (a) the same LTE bandwidth and (b) 5 MHz UTRA (see Figure 11.4).

[4] contains simulation results for the downlink coexistence of LTE with adjacent LTE, UTRA or GSM carriers. The required Adjacent Channel Interference Ratio (ACIR) to ensure ≤5% cell edge throughput loss for the victim system was found to be about 30 dB for all of these cases. With an assumed UE Adjacent Channel Selectivity (ACS) of 33 dB for LTE and UTRA, the choice of 45 dB ACLR ensures minimal impact from the eNodeB transmit path and also aligns with the UTRA ACLR1 minimum performance requirements.

For LTE–LTE coexistence the ACLR1, ACLR2 are only specified for the same LTE bandwidth, i.e. mixed cases such as 5 MHz LTE ↔ 20 MHz LTE are not covered by explicit ACLR requirements according to all possible values for the victim carrier bandwidth. Analysis in [4], however, has shown that these cases are sufficiently covered by the 45 dBc ACLR requirements measured within the same bandwidth as the aggressing carrier. A large matrix of ACLR require-

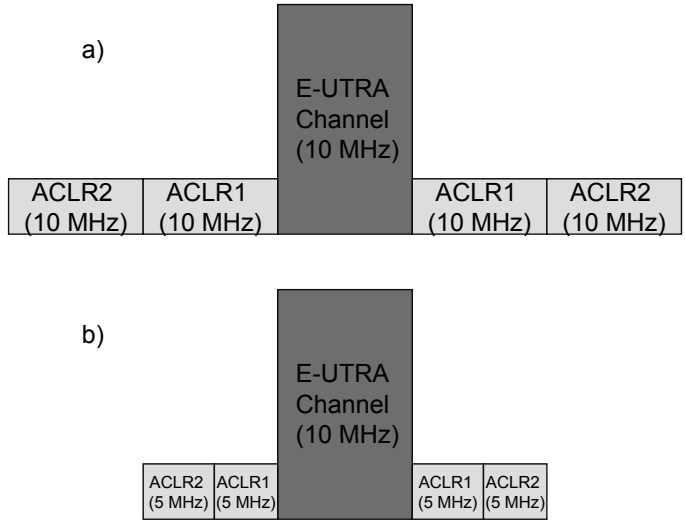


Figure 11.4 The two defined ACLR measures, one for 1st and 2nd adjacent E-UTRA carriers and one for 1st and 2nd adjacent UTRA carrier

ments for all LTE bandwidth options and, additionally, for multiple other radio technologies would have led to a large eNodeB conformance testing effort without materially adding more robust protection.

Furthermore, the unwanted emission mask specified for LTE provides a baseline protection for any victim system. For this reason no ACLR requirements were specified, for example, for GSM victim carriers – the corresponding ACLR values can be obtained by integrating the unwanted emission mask accordingly and are well above the required ~30 dB suggested by system simulations [4]. However, ACLR requirements for the 1st and 2nd WCDMA channels have been added to explicitly protect adjacent 'legacy' WCDMA carriers by the same level from aggressing LTE systems as specified for WCDMA as LTE migration within WCDMA bands was considered an important deployment scenario.

The ACLR2/3.84 MHz for a WCDMA victim is specified to be the same value as the ACLR1/3.84 MHz (45 dBc), not 50 dBc as for WCDMA. This is reasonable, as the second adjacent channel interference contributes only little to overall ACIR,

$$ACIR = \frac{1}{\frac{1}{ACLR1} + \frac{1}{ACS1} + \frac{1}{ACLR2} + \frac{1}{ACS2}} \tag{11.2}$$

because the WCDMA UE ACS2 in the second adjacent channel is significantly higher (~42 dBc) than the ACS1 (~33 dBc). Therefore, decreasing the ACLR2/3.84 MHz from 50 to 45 dBc has only negligible impact. On the other hand, an ACLR2/3.84 MHz requirement of 50 dBc for 10, 15 and 20 MHz LTE would be overly restrictive from an eNodeB implementation perspective

as the WCDMA ACLR2 frequency region (within 5 … 10 MHz offset from the LTE carrier edge) would still be within the ACLR1 region of the transmitting LTE carrier for which 45 dBc is a more appropriate requirement.

The ACLR values specified for LTE can also be compared to the corresponding values obtained by integrating the unwanted emission mask, together with an assumption on the eNodeB output power, e.g. 46 dBm. In fact, the integrated ACLR values obtained from the mask are ~2 … 5 dB more relaxed when compared with the specified ACLR values. This was specified on purpose, as downlink power control on RBs may lead to some 'ripples' of the leakage power spectrum within the OOB domain and the LTE unwanted emission mask was designed to be merely a 'roofing' requirement. These ripples would be averaged out within the ACLR measurement bandwidth and are therefore not detrimental from a coexistence perspective. Hence, in LTE, unlike WCDMA, the transmitter (PA) linearity is set by the ACLR and not the unwanted emission mask requirements, facilitating increased output power – for more information see [8].

Finally, unwanted emission mask and ACLR requirements apply whatever the type of transmitter considered (single carrier or multi-carrier). For a multi-carrier eNodeB, the ACLR requirement applies for the adjacent channel frequencies below the lowest carrier frequency used by the eNodeB and above the highest carrier frequency used by the eNodeB, i.e. not within the transmit bandwidth supported by the eNodeB, which is assumed to belong to the same operator.

11.3.3 Coexistence with Other Systems in Adjacent Operating Bands

[2] contains additional spurious emission requirements for the protection of UE and/or eNodeB of other wireless systems operating in other frequency bands within the same geographical area. The system operating in the other frequency band may be GSM 850/900/1800/1900, Personal Handyphone System (PHS), Public Safety systems within the US 700 MHz bands, WCDMA FDD/TDD and/or LTE FDD/TDD. Most of these requirements have some regulatory background and are therefore mandatory within their respective region (Europe, Japan, North America). The spurious emission limits for the protection of 3GPP wireless systems assume typically ~67 dB isolation (Minimum Coupling Loss, MCL) between the aggressor and victim antenna systems, including antenna gain and cable losses.

Moreover, to facilitate co-location (i.e. assuming only 30 dB MCL) with the base stations of other 3GPP systems, an additional set of optional spurious emission requirements has been defined in [2].

However, as shown by Figure 11.2, spurious emissions do not apply for the 10 MHz frequency range immediately outside the eNodeB transmit frequency range of an operating band. This is also the case when the transmit frequency range is adjacent to the victim band, hence the above coexistence requirements do not apply for FDD/TDD transition frequencies at, e.g., 1920, 2570, 2620 MHz. However, the deployment of FDD and TDD technologies in adjacent frequency blocks can cause interference between the systems as illustrated in Figure 11.5. There may be interference between FDD and TDD base stations and between the terminals as well.

From an eNodeB implementation perspective it is challenging to define stringent transmitter and receiver filter requirements to minimize the size of the needed guard band around the FDD/TDD transition frequencies. Furthermore, different regions of the world may impose local regulatory requirements to facilitate FDD/TDD deployment in adjacent bands. It was therefore

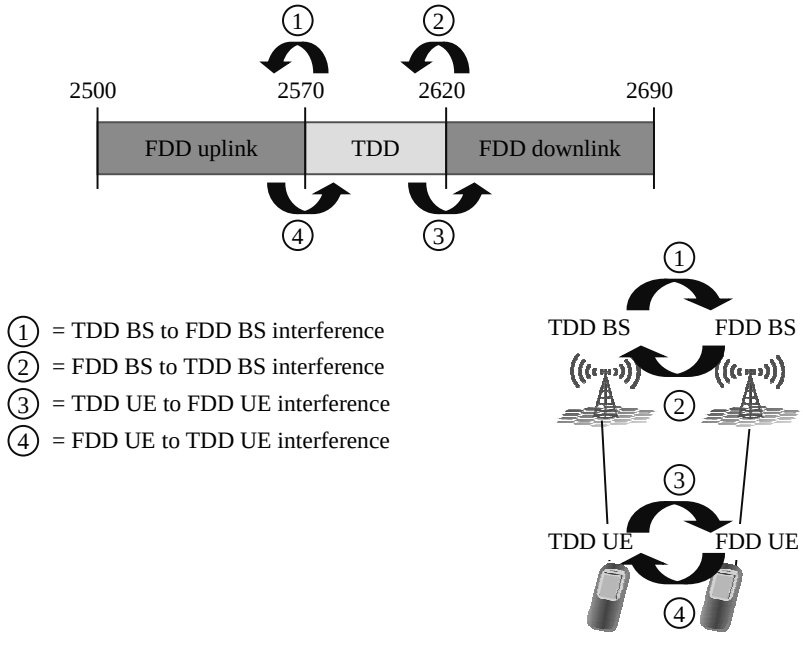


Figure 11.5 Interference scenarios between FDD and TDD (example for Band 7)

seen as impractical in 3GPP to make generic assumptions about the available guard bands and subsequently to define general FDD/TDD protection requirements, which would need to be in line with the various regulatory requirements throughout the world. Hence, additional FDD/TDD protection emission limits for the excluded frequency range of a 10 MHz frequency range immediately outside the eNodeB transmit frequency range of an operating band are left to local or regional requirements.

As an example of such local requirements, the remainder of this section will discuss the rules defined in [9] for using FDD and TDD within the European 2.6 GHz band allocation (3GPP Band 7). The cornerstones of the emission requirements in [9] are as follows:

- In-block and out-of-block Equivalent Isotropic Radiated Power (EIRP) emission masks are defined without any reference to a specific wireless system's BS standard, i.e. the limits can be considered as technology neutral. The in-block EIRP limits provide a safeguard against blocking of an adjacent system's BS receive path. The out-of-block masks, also called Block Edge Masks (BEM), impose radiated unwanted emissions limits across adjacent license block uplink/downlink receive band frequency ranges.
- The BEMs should facilitate FDD/TDD coexistence without any detailed coordination and cooperation arrangements between operators in neighboring spectrum blocks, for base station separations greater than 100 m. In this, the assumed MCL has been > 53 dB and the interference impact relative to BS thermal noise, *I/N*, has been assumed as −6 dB, with a noise figure of 5 dB. Antenna isolation measurements show that 50 dB MCL can be easily achieved on the same rooftop by using separate, vertically and/or horizontally displaced, antennas for FDD and TDD, respectively.

- Some of the BEMs have been derived from the WCDMA BS Spectrum Emission Mask (SEM), by converting the WCDMA SEM (i.e. a conducted emission requirement to be met at the BS antenna connector) into an EIRP mask by making assumptions on antenna gain and cable losses (17 dBi).
- EIRP limits are requirements for radiated (not conducted) unwanted emissions and they apply within or outside a license block. Hence, these are not testable BS equipment requirements as such, but they involve the whole RF site installation, including antennas and cable losses. Furthermore, a BEM requirement does not necessarily start at the RF channel (carrier) edge, but at the edge of the licensee's block, which may differ from the channel edge depending on the actual deployment scenario. Finally, the BEM must be met for the sum of all radiated out-of-block emissions, for example from multiple RF carriers or even multiple BSs present at a site, not only for those pertaining to a single RF channel as is the case for the UTRA SEM.

These requirements will be discussed in the following in more detail assuming the Band 7 frequency arrangements with 70 MHz FDD uplink, 50 MHz TDD and 70 MHz FDD downlink.

Figure 11.6 shows the maximum allowed in-block EIRP limits. The nominal EIRP limit of 61 dBm/5 MHz can be relaxed by local regulation up to 68 dBm/5 MHz for specific deployments, for example in areas of low population density where higher antenna gain and/or BS output power are desirable.

The lower 5 MHz part of the TDD allocation is a restricted block with a maximum allowed base station EIRP of 25 dBm/5 MHz where antennas are placed indoors or where the antenna height is below a certain height specified by the regulator. The lower transmission power in the restricted 5 MHz TDD block serves to relax the FDD BS selectivity and blocking requirements just below 2570 MHz.

Figure 11.7 shows the resulting out-of-block EIRP emission mask (BEM) for an unrestricted (high power) 20 MHz TDD block allocation starting 5 MHz above the FDD uplink band allocation. The following can be noted:

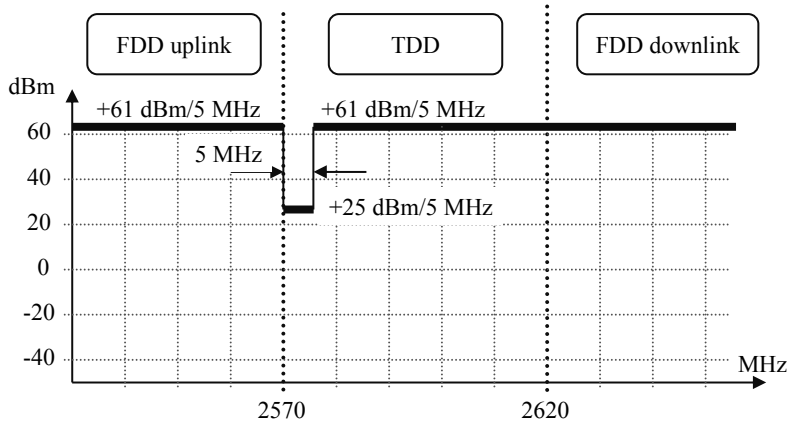


Figure 11.6 Maximum allowed in-block EIRP including restricted block

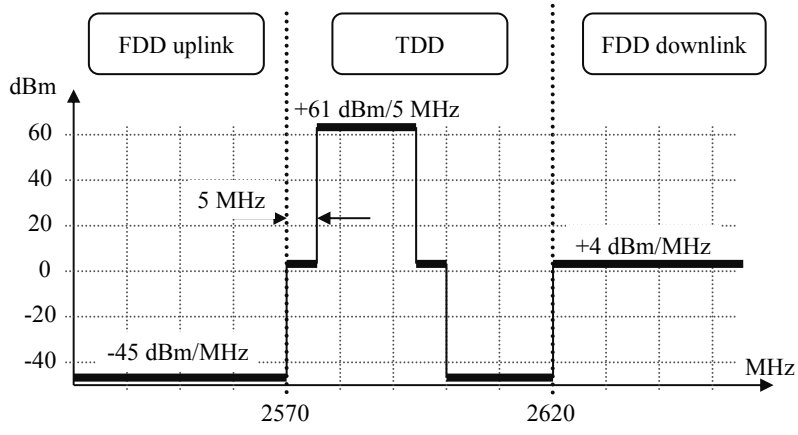


Figure 11.7 Out-of-block EIRP emission mask (BEM) for a 20 MHz TDD block above the FDD uplink band allocation

- The whole FDD uplink band allocation is protected by a −45 dBm/MHz EIRP limit, which is chosen as a baseline requirement for protection of the uplink in the BS transmitter → BS receiver interference scenario. Given a MCL > 53 dB, there will be minimal impact on the FDD BS uplink.
- Possible unsynchronized TDD operation is also protected by the −45 dBm/MHz EIRP limit from 5 MHz above the TDD block edge onwards.
- As the relevant interference mechanism across the FDD downlink band will be TDD BS → FDD UE, a more relaxed limit of +4 dBm is sufficient. This limit is also stipulated as out-of-block EIRP limit for a FDD BS across this part of the band (see Figure 11.7).
- Not shown in Figure 11.7 is the possible use of a TDD restricted block within 2570–2575 MHz, which has a tighter BEM requirement across the first 5 MHz just below the 2570 MHz transition frequency in order to protect the FDD uplink.
- Also not shown is the detailed slope of the BEM within the first MHz from either side of the TDD block edge, which follows essentially the shape of the UTRA SEM.

Figure 11.8 shows the resulting out-of-block EIRP emission mask (BEM) for 10 MHz FDD block allocation starting just above the TDD band allocation. The following can be noted:

- The first 5 MHz TDD block just below the FDD downlink block may receive a higher level of interference due to the more relaxed BEM EIRP limit of +4 dBm. This is to facilitate FDD BS transmit filters; no restricted power blocks are foreseen for the FDD blocks.
- The rest of the TDD band and FDD uplink band allocation is protected by the −45 dBm/MHz EIRP baseline requirement.

11.3.4 Transmitted Signal Quality

[2] contains the modulation-specific Error Vector Magnitude (EVM) requirements to ensure sufficiently high quality of the transmitted base station signal (see Table 11.5).

Typical impairments of the transmitted signal modulation accuracy are analog RF distortions (frequency offset, local oscillator phase noise, amplitude/phase ripple from analog fil-

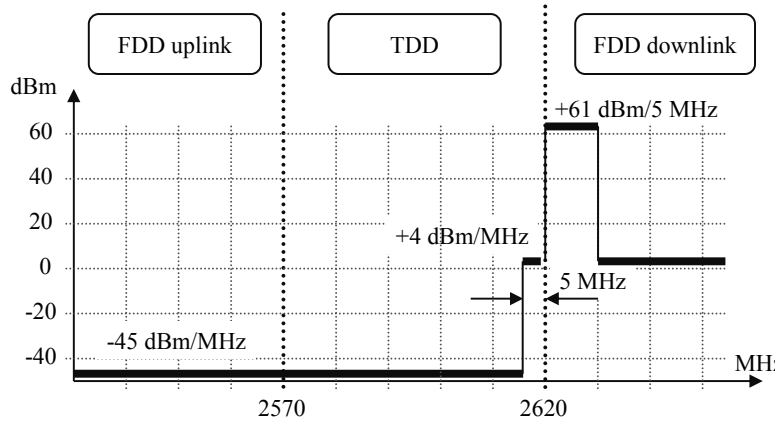


Figure 11.8 Out-of-block EIRP emission mask (BEM) for a 10 MHz FDD block above the TDD band allocation

Table 11.5 EVM requirements

Modulation scheme for PDSCH	Required EVM (%)
QPSK	17.5%
16QAM	12.5%
64QAM	8%

ters) as well as distortions created within the digital domain such as inter-symbol interference from digital filters used for spectrum shaping, finite wordlength effects and, most important for conditions near maximum transmit power, the clipping noise from peak-to-average radio reduction schemes.

The EVM requirement ensures that the downlink throughput due to the base station non-ideal waveform is only marginally reduced, typically by 5% assuming ideal reception in the UE. The required EVM must be fulfilled for all transmit configurations and across the whole dynamic range of power levels used by the base station.

11.3.4.1 Definition of the EVM

More precisely, the EVM is a measure of the difference between the ideal constellation points and the measured constellation points obtained after equalization by a defined 'reference receiver'. Unlike WCDMA, the EVM is not measured on the transmitted composite time-domain signal waveform, but within the frequency domain, after the FFT, by analyzing the modulation accuracy of the constellation points on each sub-carrier. Figure 11.9 shows the reference point for the EVM measurement.

The OFDM signal processing blocks prior to the EVM reference point, in particular the constrained Zero-Forcing (ZF) equalizer, are fundamentally not different from those of an actual UE. The rationale for this rather complex EVM definition has been that only those transmitter impairments should 'pollute' the EVM that would not be automatically removed by the UE recep-

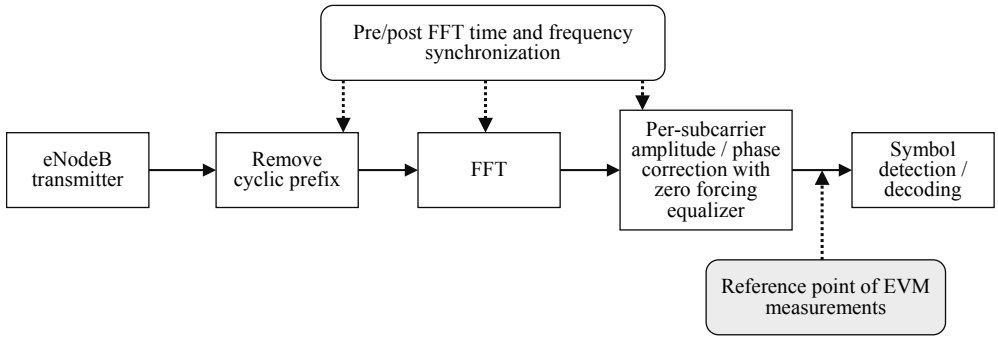


Figure 11.9 Reference point for EVM measurement

tion anyway. For example, clipping noise from peak-to-average ratio reduction has AWGN-like characteristics and cannot be removed by UE equalizers, whereas a non-linear phase response from analog filter distortion can be estimated from the reference signals and subsequently be removed by the ZF equalizer. The qualifier 'constrained' within the ZF-equalization process refers to a defined frequency averaging process of the raw channel estimates used to compute the ZF equalizer weights. Some form of frequency averaging of the raw channel estimates at the reference signal locations will be present within an actual UE receiver implementation and this will therefore limit the removal of some of the impairments such as filter amplitude (or phase) ripple with a spatial frequency in the order of the sub-carrier spacing. Therefore an equivalent of this frequency averaging process has been defined for the EVM measurement process.

Yet another feature of the EVM definition is the chosen time synchronization point for the FFT processing. The EVM is deliberately measured not at the ideal time synchronization instant, which would be approximately the centre of the cyclic prefix, but rather at two points shortly after the beginning and before the end of the cyclic prefix, respectively. This also ensures that in the presence of pre- and post-cursors from the transmit spectrum shaping filters, the UE will see sufficiently low inter-symbol interference even under non-ideal time tracking conditions.

The details of this EVM measurement procedure, including all the aspects above are defined within [5], Annex F.

11.3.4.2 Derivation of the EVM requirement

Next we provide a rationale for the EVM requirements based on an analytic approach presented in [10]. Let us calculate the required EVM for 5% throughput loss for each instantaneous C/I value, corresponding to a certain selected Modulation and Coding Scheme (MCS).

To start with, we consider the MCS throughput curves and approximating MCS envelope shown in Figure 11.10. The MCS curves were generated from link level simulations under AWGN channel conditions for a 1×1 antenna configuration (1 transmit and 1 receive antenna) using ideal channel estimation within the receiver.

We approximate the set of these throughput curves by the expression for Shannon's channel capacity with the fitting parameter $\alpha = 0.65$.

$$C = \alpha \log 2 (1 + \frac{S}{N}) \tag{11.3}$$

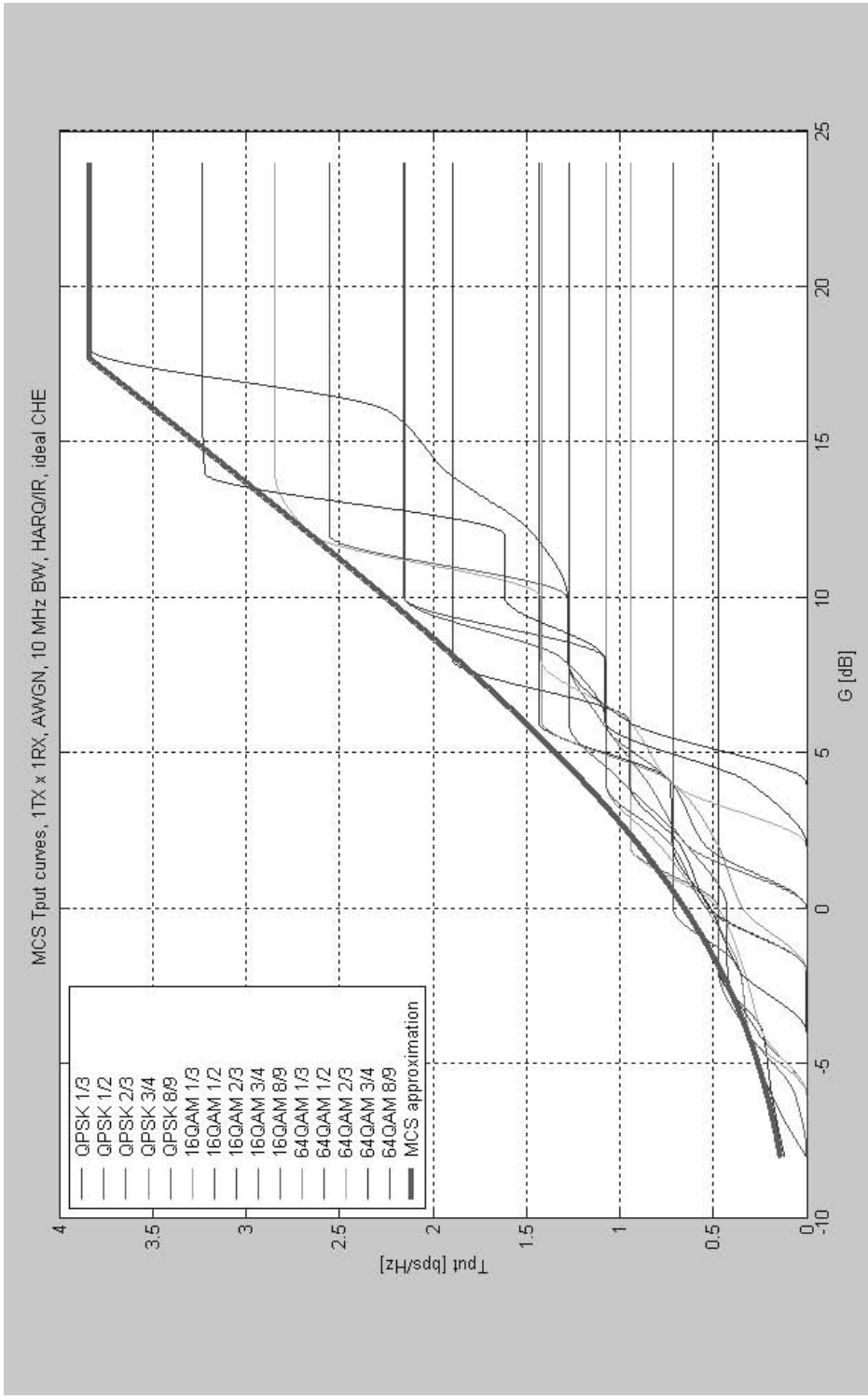


Figure 11.10 MCS throughput curves and approximating MCS envelope

where S is the signal power and N is the noise power. Assume next that the transmit impairments can be modeled as AWGN with power M and the receiver noise as AWGN with power N. Then the condition that the transmitter AWGN should lead to 5% throughput loss can be expressed as:

$$\alpha \log 2(1+\frac{S}{M+N}) = 0.95 \cdot \alpha \log 2(1+\frac{S}{N}) \tag{11.4}$$

Solving this condition for $EVM_{req} \equiv \frac{M}{S}$ we obtain

$$EVM_{req} \equiv \frac{M}{S} = \left[\left(1+\frac{S}{N}\right)^{0.95} - 1\right]^{-1} - \frac{N}{S} \tag{11.5}$$

The required EVM is plotted in Figure 11.11 for the C/I range of the approximating envelope MCS throughput curve of Figure 11.10. As can been seen ~6.3% EVM would be required for the highest throughput MCS (64QAM 8/9, operating S/N ~17.7 dB). However, when assuming a single EVM requirement for all 64QAM modulated MCSs, this would be too stringent as 64QAM may be chosen from a C/I range from 12 to 17.7 dB according to the chosen MCS set. This would indicate a required EVM in the range of 10–6.3%. Looking at the midpoint S/N of ~15 dB for 64QAM MCS selection one obtains a 7.9% EVM requirement. Similarly, one obtains for:

- 16QAM: C/I range from 6 to 12 dB, midpoint C/I of ~9 dB with an EVM requirement of 12.9%;
- QPSK: C/I range from −8 to 6 dB, midpoint C/I of ~ −1 dB with EVM requirements of 29.6% and 16.3% respectively (for 6 dB C/I).

The required EVM values for 64QAM, 16QAM and QPSK are 8%, 12.5% and 17.5%.

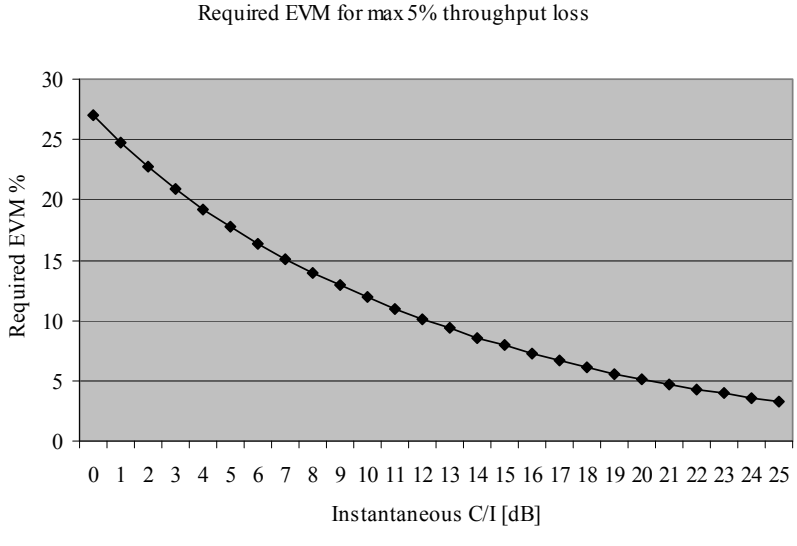


Figure 11.11 Required EVM for 5% throughput loss

However, the above derivation assumes a smooth approximating MCS envelope. In reality, the MCS envelope has a 'waterfall' shape in AWGN, as shown in Figure 11.10 and may exhibit either a larger impact from EVM in regions with steep slopes or a smaller impact in regions with flat response. In an actual system scenario the C/I distribution will average these unequal throughput losses across the 'waterfall' MCS envelope. The quasi-static system simulations in [10] do indeed verify that the resulting average throughput loss for 64QAM is in line with the above simplified derivation.

11.4 eNodeB RF Receiver

The purpose of the eNodeB RF receiver requirements is to verify different RF impairments that have an impact on the network performance. These impairments include noise figure, receiver EVM, selectivity on different frequencies, including adjacent channel, etc. The following base station RF receiver requirements are described: reference sensitivity level, dynamic range, in-channel selectivity, Adjacent Channel Selectivity (ACS), blocking, receiver spurious emissions, and receiver intermodulation.

11.4.1 Reference Sensitivity Level

The reference sensitivity level is the minimum mean power received at the antenna connector, at which a throughput requirement is fulfilled. The throughput will be equal to or higher than 95% of the maximum throughput for a specified reference measurement channel.

The purpose of this requirement is to verify the receiver noise figure. Other receiver impairments such as receiver EVM are included within the receiver demodulation performance requirements at high SNR points. Therefore, the maximum throughput is defined at low SNR points for the sensitivity case. The reference measurement channel is based on a QPSK modulation and 1/3 coding rate.

For channel bandwidths lower than or equal to 5 MHz, the reference measurement channels are defined on the basis of all resource blocks allocated to this channel bandwidth. For channel bandwidths higher than 5 MHz, the sensitivity is measured using consecutive blocks consisting of 25 RBs.

The reference sensitivity level calculation is given by Equation 11.6. Noise Figure (NF) is equal to 5 dB and implementation margin (IM) is equal to 2 dB.

$$P_{REFSENS}[\text{dBm}] = -174[\text{dBm/Hz}] + 10\log_{10}(N_{RB} \cdot 180k) + NF + SNR + IM \qquad (11.6)$$

For example, for a 10 MHz channel bandwidth (N_{RB}=25), the reference sensitivity level is equal to −101.5 dBm. The simulated SNR for 95% of the maximum throughput is equal to −1.0 dB.

The eNodeB noise figure is relevant for the coverage area. The uplink link budget calculation in Chapter 9 assumes a base station noise figure of 2 dB, so clearly better than the minimum performance requirement defined in 3GPP specifications. The typical eNodeB has better performance than the minimum requirement since the optimized performance is one of the selling arguments for the network vendors.

11.4.2 Dynamic Range

The dynamic range requirement is a measure of the capability of the receiver to receive a wanted signal in the presence of an interfering signal inside the received frequency channel, at which a throughput requirement is fulfilled. The throughput will be equal to or higher than 95% of the maximum throughput for a specified reference measurement channel.

The intention of this requirement is to ensure that the base station can receive a high throughput in the presence of increased interference and high wanted signal levels. Such a high interference may come from neighboring cells in the case of small cells and high system loading. This requirement measures the effects of receiver impairments such as receiver EVM and is performed at high SNR points. The mean power of the interfering signal (AWGN) is equal to the receiver noise floor increased by 20 dB, in order to mask the receiver's own noise floor. The maximum throughput is defined at high SNR points, thus the reference measurement channel is based on 16QAM modulation and 2/3 coding rate.

The mean power calculation of the wanted signal is given by Equation 11.7. NF is equal to 5 dB and IM is equal to 2.5 dB.

$$P_{\text{wanted}}[\text{dBm}] = -174[\text{dBm/Hz}] + 10\log(N_{\text{RB}} \cdot 180\text{k}) + 20 + \text{NF} + \text{SNR} + \text{IM} \qquad (11.7)$$

For example, for 10 MHz channel bandwidth ($N_{\text{RB}} = 25$), the wanted signal mean power is equal to −70.2 dBm. The simulated SNR for 95% of the maximum throughput is equal to 9.8 dB. The interfering signal mean power is set equal to −79.5 dBm. The dynamic range measurement for the 5 MHz case is illustrated in Figure 11.12.

11.4.3 In-channel Selectivity

The in-channel selectivity requirement is a measure of the receiver's ability to receive a wanted signal at its assigned resource block locations in the presence of an interfering signal received at a larger power spectral density, at which a throughput requirement is fulfilled. The throughput will equal or exceed 95% of the maximum throughput for a specified reference measurement channel.

The intention of this requirement is to address in-band adjacent resource block selectivity, i.e. the reception of simultaneous user signals at greatly different power spectral density levels due to used modulation format, power control inaccuracies, other-cell interference levels, etc. The uplink signal is created by two signals, where one is the wanted QPSK modulated signal

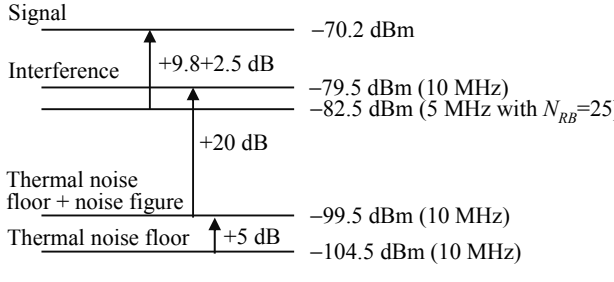


Figure 11.12 Dynamic range measurement for 10 MHz LTE

Table 11.6 Number of resource blocks allocated for wanted and interfering signal

Channel bandwidth (MHz)	Wanted signal	Interfering signal
1.4	3	3
3	9	6
5	15	10
10	25	25
15	25	25
20	25	25

and the other is the interfering 16QAM modulated signal at elevated power. Table 11.6 presents the allocation of resource blocks for wanted and interfering signals, for different channel bandwidths. The high power level difference may happen if the interfering user is close to the base station and can use a high signal-to-noise ratio while the wanted user is far from the base station and can only achieve a low signal-to-noise ratio.

For channel bandwidths equal to 10, 15 and 20 MHz, the 25 resource block allocations of the wanted and interfering signals are adjacently around Direct Current (DC), in order to be sensitive to the RF impairments of the receiver image leakage, EVM, 3rd order intermodulation (IMD3) and the local oscillator phase noise.

The mean power of the interfering signal is equal to the receiver noise floor increased by 9.5 dB (required SNR) and additionally by 16 dB (assumed interference over thermal noise). The desensitization of the wanted resource block allocation, in the presence of the interfering resource block allocation, is equal to 3 dB.

The wanted signal mean power calculation is given by Equation 11.8. NF is equal to 5 dB and IM is equal to 2 dB.

$$P_{wanted}[dBm] = -174[dBm/Hz] + 10\log(N_{RB} \cdot 180k) + 3 + NF + SNR + IM \qquad (11.8)$$

For example, for 10 MHz channel bandwidth ($N_{RB} = 25$), the wanted signal mean power is equal to –98.5 dBm. The simulated SNR for 95% of the maximum throughput is equal to −1.0 dB. The interfering signal mean power is equal to −77 dBm. The in-channel measurement case is presented in Figure 11.13.

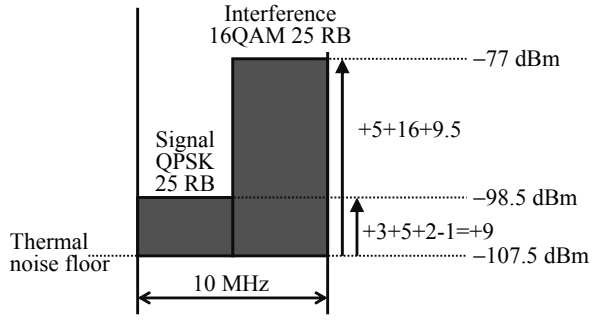


Figure 11.13 In-channel selectivity measurement for 5 MHz LTE

11.4.4 Adjacent Channel Selectivity (ACS) and Narrow-band Blocking

The ACS (narrow-band blocking) requirement is a measure of the receiver's ability to receive a wanted signal at its assigned channel frequency in the presence of an interfering adjacent channel signal, at which a throughput requirement is fulfilled. The throughput will equal or exceed 95% of the maximum throughput, for a specified reference measurement channel.

The intention of this requirement is to verify the selectivity on the adjacent channel. The selectivity and the narrowband blocking are important to avoid interference between operators. The adjacent operator's mobile may use a high transmission power level if its own base station is far away. If such a mobile happens to be close to our base station, it can cause high interference levels on the adjacent channel.

Table 11.7 presents the ACS requirement relationship between E-UTRA interfering signal channel bandwidth, wanted signal channel bandwidth, wanted signal mean power and interfering signal centre frequency offset to the channel edge of the wanted signal. The interfering signal mean power is equal to −52 dBm. Table 11.8 shows how the adjacent channel selectivity can be calculated from the performance requirements for 10 MHz channel bandwidth with 25 resource blocks. The adjacent channel selectivity test case is illustrated in Figure 11.14. Base station noise floor is given by Equation 11.9. NF is equal to 5 dB and $N_{RB} = 25$.

$$D[\text{dBm}] = -174[\text{dBm/Hz}] + 10\log(N_{RB} \cdot 180\text{k}) + \text{NF} \tag{11.9}$$

The narrow-band blocking measurement is illustrated in Figure 11.15. The wanted signal mean power is equal to $P_{REFSENS} + 6$ dB. The interfering 1RB E-UTRA signal mean power is equal to −49 dBm. The interfering signal is located in the first five worst case resource block allocations. Additionally, for 3 MHz channel bandwidth, the interfering signal is located in every third resource block allocation. For 5, 10, 15 and 20 MHz channel bandwidths, the interfering signal is located additionally in every fifth resource block allocation. Such a location of the interfering signal verifies different possible impairments of the receiver performance. For GSM

Table 11.7 Derivation of ACS requirement

Wanted signal channel bandwidth [MHz]	Wanted signal mean power [dBm]	Offset [MHz]	Interfering signal channel bandwidth [MHz]
1.4	$P_{REFSENS} + 11$	0.7	1.4
3	$P_{REFSENS} + 8$	1.5	3
5, 10, 15, 20	$P_{REFSENS} + 6$	2.5	5

Table 11.8 Calculation of the ACS requirement for 10 MHz channel bandwidth

Interfering signal mean power (A)	–52 dBm
Base station noise floor with 5 dB noise figure (B)	–102.5 dBm
Allowed desensitization (C)	6 dB
Total base station noise (D = B + C in dB)	–96.5 dBm
Allowed base station interference (E = D – B in absolute value)	–98 dBm
Base station adjacent channel selectivity (F = A – E)	46 dB

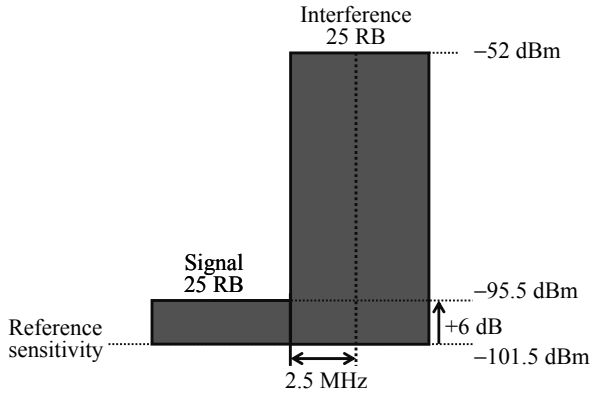


Figure 11.14 Adjacent channel selectivity measurement for 10 MHz LTE

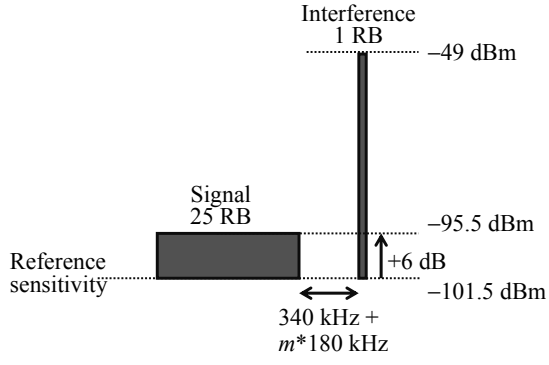


Figure 11.15 Narrowband blocking measurement for 10 MHz LTE

band refarming, the blocker can be a narrowband GSM signal. There are no specific requirements with a GSM signal used as the blocker. The GSM signal, however, is sufficiently similar to 1 RB LTE blocker for practical performance purposes.

11.4.5 Blocking

The blocking requirement is a measure of the receiver's ability to receive a wanted signal at its assigned channel frequency in the presence of an unwanted interferer, at which a throughput requirement is fulfilled. The throughput will equal or exceed 95% of the maximum throughput for a specified reference measurement channel.

The intention of this requirement is to verify the selectivity on different frequencies, excluding the adjacent channel. The in-band blocking can also be called adjacent channel selectivity for the second adjacent channel.

For in-band blocking the unwanted interferer is an LTE signal. For example, for Operating Band 1 (1920–1980 MHz), the in-band blocking refers to the centre frequencies of the interfering signal from 1900 MHz to 2000 MHz. For out-of-band blocking, the unwanted interferer

is a Continuous Wave (CW) signal. For Operating Band 1, the out-of-band blocking refers to the centre frequencies of the interfering signal from 1 MHz to 1900 MHz and from 2000 MHz to 12 750 MHz. The measurements are shown in Figure 11.16.

The in-band blocking requirement for 10 MHz LTE is illustrated in Figure 11.17**,** and out-of-band in Figure 11.18. The wanted signal mean power is equal to $P_{REFSENS}$ + 6 dB. The E-UTRA interfering signal mean power is equal to −43 dBm. The CW interfering signal mean power is equal to −15 dBm.

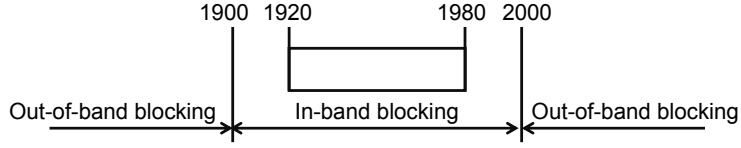


Figure 11.16 In-band and out-of-band blocking measurement for Band 1

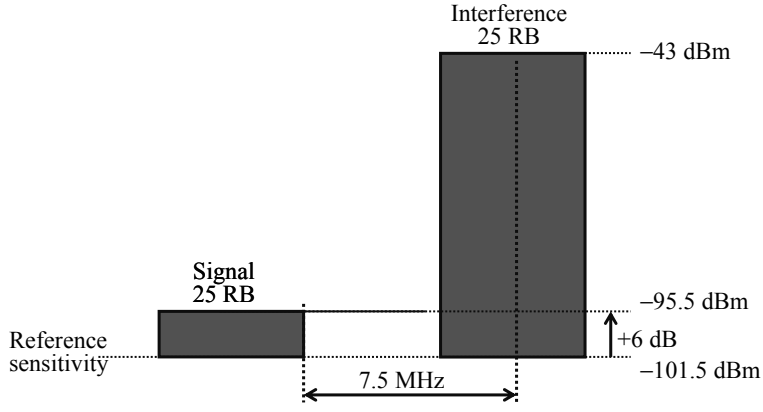


Figure 11.17 In-band blocking measurement for 10 MHz LTE

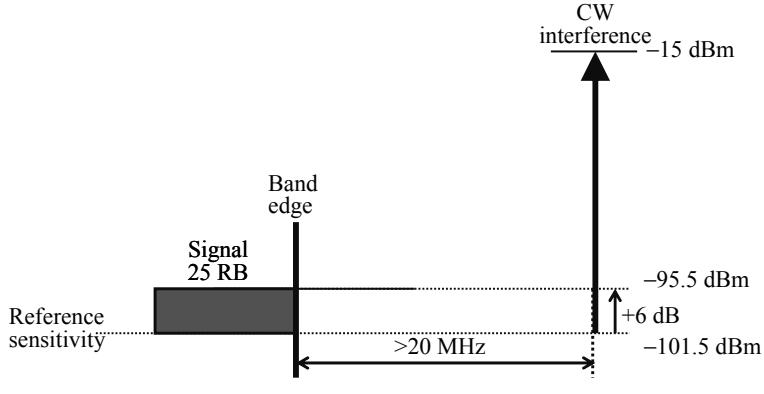


Figure 11.18 Out-of-band blocking measurement for 10 MHz LTE

11.4.6 Receiver Spurious Emissions

The spurious emissions power is the power of emissions generated or amplified in a receiver that appears at the base station receiver antenna connector.

The frequency range between 2.5 × channel bandwidth below the first carrier frequency and 2.5 × channel bandwidth above the last carrier frequency, transmitted by the base station, may be excluded from the requirement. Frequencies that are more than 10 MHz below the lowest frequency or more than 10 MHz above the highest frequency of the base station transmitter operating band will not be excluded from the requirement.

Additionally, the power of any spurious emission will not exceed the levels specified for coexistence with other systems in the same geographical area and for protection of the LTE FDD base station receiver of own or different base station.

11.4.7 Receiver Intermodulation

Intermodulation response rejection is a measure of the capability of the receiver to receive a wanted signal on its assigned channel frequency in the presence of two interfering signals which have a specific frequency relationship to the wanted signal, at which a throughput requirement is fulfilled. The throughput will equal or exceed 95% of the maximum throughput for a specified reference measurement channel.

The intermodulation requirement is relevant when there are two terminals from other operators transmitting at high power level close to our base station.

The mean power of the wanted signal is equal to $P_{REFSENS} + 6$ dB (Figure 11.19). The interfering signal mean power (both LTE and CW) is equal to −52 dBm. The offset between the CW interfering signal centre frequency and the channel edge of the wanted signal is equal to 1.5 LTE interfering signal channel bandwidth. The offset between the LTE interfering signal centre frequency and the channel edge of the wanted signal is specified on the basis of the worst case scenario – the intermodulation products fall on the edge resource blocks of an operating channel bandwidth.

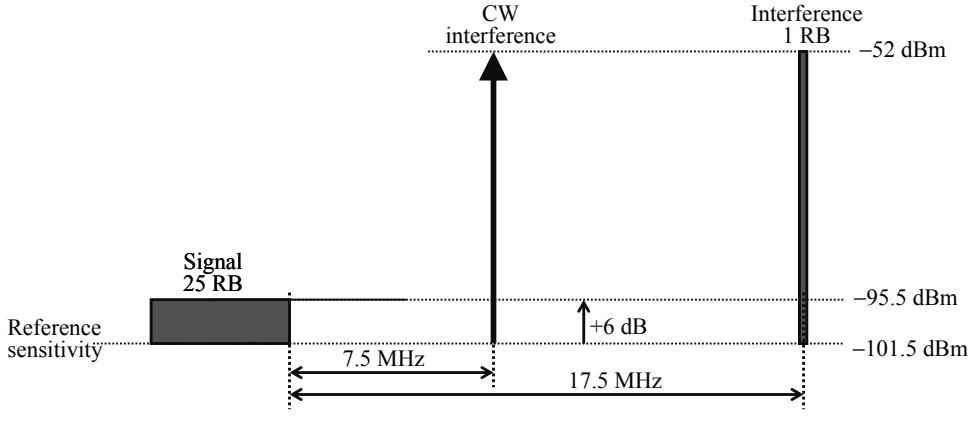


Figure 11.19 Intermodulation measurement for 5 MHz LTE

11.5 eNodeB Demodulation Performance

The purpose of the base station demodulation performance requirements is to estimate how the network is performing in practice and to verify different base station impairments which have an impact on the network performance. These impairments include RF and baseband impairments, receiver EVM, time and frequency tracking, frequency estimation, etc. The base station demodulation performance requirements are described for the following uplink channels: Physical Uplink Shared Channel (PUSCH), Physical Uplink Control Channel (PUCCH) and Physical Random Access Channel (PRACH).

11.5.1 PUSCH

The PUSCH demodulation performance requirements are determined by a minimum required throughput for a given SNR. The throughput will equal or exceed 30% or 70% of the maximum throughput for a specified reference measurement channel that contains data and reference signals only.

The PUSCH demodulation performance requirements are specified for all LTE channel bandwidths. Additionally, for each channel bandwidth, the following various network related parameters are selected to match different radio system configurations:

- number of receive antennas – 2 or 4;
- modulation and coding scheme – QPSK 1/3, 16QAM 3/4 or 64QAM 5/6;
- channel model – EPA5, EVA5, EVA70, ETU70 or ETU300, where the number indicates the Doppler shift. PA is Pedestrian A, VA is Vehicular A and TU is Typical Urban channel model;
- cyclic prefix type – normal or extended;
- number of resource blocks allocated for channel bandwidth – single or all possible.

Each channel model is described by the Doppler frequency which corresponds to various velocities depending on the frequency band. For example, EPA5 corresponds to velocities equal to 7.7 kmph, 2.7 kmph and 2.1 kmph for 0.7 GHz (Band 12), 2 GHz (Band 1) and 2.6 GHz (Band 7), respectively.

The incremental redundancy HARQ allowing up to three retransmissions is used. Table 11.9 presents a set of PUSCH base station tests for each channel bandwidth and for each configuration of receive antennas.

For a base station supporting multiple channel bandwidths, only tests for the largest and the smallest channel bandwidths are applicable.

The SNR requirements were specified on the basis of average link level simulation results with implementation margins presented by various companies during 3GPP meetings.

An example of the relationship between the net data rate and required SNR follows: 10 MHz channel bandwidth, normal cyclic prefix and QSPK 1/3 modulation and coding scheme were taken into account.

For a fixed reference channel A3–5 [2], the SNR requirements were specified for full resource block allocation (50 resource blocks), for 2 and 4 receive antennas, for EPA5 and EVA70 channel models and for 30% and 70% fractions of the maximum throughput. Table 11.10 presents the SNR requirements for a 30% fraction of the maximum throughput.

Table 11.9 Set of PUSCH base station tests

Cyclic prefix	Channel model, RB allocation	Modulation and coding scheme	Fraction of maximum throughput [%]
Normal	EPA5, all	QPSK 1/3 16QAM 3/4 64QAM 5/6	30 and 70 70 70
	EVA5, single	QPSK 1/3 16QAM 3/4 64QAM 5/6	30 and 70 30 and 70 70
	EVA70, all	QPSK 1/3 16QAM 3/4	30 and 70 30 and 70
	ETU70, single	QPSK 1/3	30 and 70
	ETU300, single	QPSK 1/3	30 and 70
Extended	ETU70, single	16QAM 3/4	30 and 70

Table 11.10 SNR requirements for 30% fraction of the maximum throughput (fixed reference channel A3–5)

Number of receive antennas	Channel model	SNR requirement [dB]
2	EPA5 EVA70	–4.2 –4.1
4	EPA5 EVA70	–6.8 –6.7

For the fixed reference channel A3–5 the payload size is equal to 5160 bits and corresponds to a 5.16 Mbps instantaneous net data rate. For one resource block it is 103.2 kbps, accordingly.

A 30% fraction of the maximum throughput corresponds to 30.9 kbps and 61.9 kbps for one and two resource blocks, respectively.

The link budget in Chapter 9 assumed an SNR equal to −7 dB with 64 kbps and two resource blocks corresponding to 32 kbps and one resource block. The link budget assumes better eNodeB receiver performance because of more assumed HARQ retransmissions and also because typical eNodeB performance is closer to the theoretical limits than the 3GPP minimum requirement.

Additionally, the uplink timing adjustment requirement for PUSCH was specified in [2]. The rationale for this requirement is to check if the base station sends timing advance commands with the correct frequency and if the base station estimates appropriate uplink transmission timing. The uplink timing adjustment requirements are determined by a minimum required throughput for a given SNR and are specified for moving propagation conditions as shown in Figure 11.20. The time difference between the reference timing and the first tap is described by Equation 11.10, where $A = 10\ \mu s$.

$$\Delta\tau = \frac{A}{2} \cdot \sin(\Delta\omega \cdot t) \tag{11.10}$$

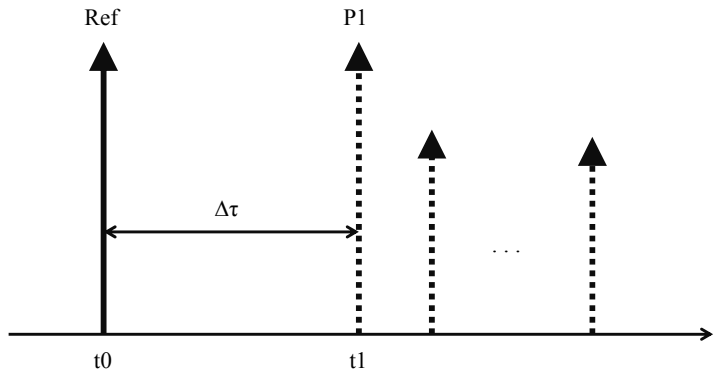


Figure 11.20 Moving propagation conditions

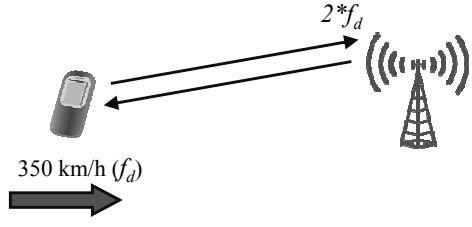


Figure 11.21 High speed train demodulation requirement

The uplink timing adjustment requirement is specified for normal and extreme conditions. For normal conditions the ETU channel model and UE speed of 120 km/h is considered ($\Delta\omega = 0.04$ s^{-1}). Uplink timing adjustment in extreme conditions is an optional requirement and corresponds to AWGN channel model and UE speed of 350 km/h ($\Delta\omega = 0.13$ s^{-1}).

[2] also includes eNodeB decoding requirements for the high speed train conditions up to 350 km/h. eNodeB can experience a two times higher Doppler shift in the worst case if the UE synchronizes to the downlink frequency including the Doppler shift (f_d) (see Figure 11.21). The maximum Doppler shift requirement is 1340 Hz which corresponds to 350 km/h at 2.1 GHz assuming that eNodeB experiences double Doppler shift.

11.5.2 PUCCH

The PUCCH performance requirements are specified for PUCCH format 1a and for PUCCH format 2. The PUCCH format 1a performance requirements are determined by a minimum required DTX to ACK probability and ACK missed detection probability for a given SNR. The DTX to ACK probability, i.e. the probability that ACK is detected when nothing is sent, will not exceed 1%. The ACK missed detection probability, i.e. the probability that ACK is not detected properly, will not exceed 1%.

The PUCCH format 1a performance requirements are specified for all E-UTRA channel bandwidths. Additionally, for each channel bandwidth, the following various network related parameters were selected to match different radio system configurations:

Table 11.11 Set of PUCCH base station tests

Cyclic prefix	Channel model
Normal	EPA5 EVA5 EVA70 ETU300
Extended	ETU70

- number of receive antennas – 2 or 4;
- channel model – EPA5, EVA5, EVA70, ETU70 or ETU300;
- cyclic prefix type – normal or extended.

Table 11.11 presents a set of PUCCH base station tests for each channel bandwidth and for each configuration of receive antennas.

The PUCCH format 2 performance requirements are determined by a CQI missed detection BLER probability for a given SNR. The CQI missed detection probability, i.e. the probability that CQI is not detected properly, will not exceed 1%. The PUCCH format 2 performance requirements are specified for all E-UTRA channel bandwidths, normal cyclic prefix, 2 receive antennas and ETU70 channel model only.

For base station supporting multiple channel bandwidths, only tests for the greatest and the smallest channel bandwidths are applicable.

11.5.3 PRACH

The PRACH performance requirements are specified for burst format 0, 1, 2, 3 and are determined by a minimum required total false alarm probability and missed detection probability for a given SNR. The total false alarm probability, i.e. the probability that preamble is detected (the sum of all errors from all detectors) when nothing is sent, will not exceed 0.1%. The missed detection probability will not exceed 1% and depends on the following errors:

- preamble is not detected properly;
- different preamble is detected than the one that is sent;
- correct preamble is detected but with wrong timing estimation.

The PRACH performance requirements are specified for 2 and 4 receive antennas and for AWGN and ETU70 channel models. For AWGN and ETU70, the timing estimation error occurs if the estimation error of the timing of the strongest path is larger than 1.04 µs and 2.08 µs, respectively. For ETU70, the strongest path for the timing estimation error refers to the strongest path in the power delay profile, i.e. the average of the delay of all paths having the same highest gain equal to 310 ns.

The PRACH performance requirements are specified for normal mode and for high speed mode. Different frequency offsets are tested for these modes. For high speed mode, when the receiver is in demanding conditions, additional frequency offsets are specified – 625 Hz

Table 11.12 Set of PRACH base station tests for normal mode

Channel model	f offset [Hz]
AWGN	0
ETU70	270

Table 11.13 Set of PRACH base station tests for high speed mode

Channel model	f offset [Hz]
AWGN	0
ETU70	270
ETU70	625
ETU70	1340

and 1340 Hz. For a frequency offset of 625 Hz, which corresponds to half of the preamble length (0.5·1/0.8 ms), the receiver is in difficult conditions because several correlation peaks are observed; a 1340 Hz frequency offset corresponds to a velocity of 350 km/h at 2.1 GHz frequency band.

Table 11.12 and Table 11.13 present a set of PRACH base station tests for each burst format and each configuration of receive antennas, for normal mode and for high speed mode, respectively.

11.6 UE Design Principles and Challenges

11.6.1 Introduction

The main requirements related to the LTE UE design are described in this section. As all LTE UEs will have to support legacy air interfaces, the LTE functionality will be constructed on top of the 2G GSM/EDGE and 3.5G WCDMA/HSPA architecture. This section presents the main differences and new challenges of an LTE terminal compared to a WCDMA terminal. Both data card and phone design are discussed.

11.6.2 RF Subsystem Design Challenges

11.6.2.1 Multi-mode and Multi-band Support

LTE equipment has to provide connection to legacy air interfaces to offer customers roaming capability in areas where LTE base stations are not yet deployed. It is essential for the acceptance of the new technology that there is continuity in the service to the user. The equipment must also support different operator, regional and roaming requirements, which results in the need to support many RF bands. For the successful introduction of a new technology the performance

of the UE must be competitive with existing technology in terms of key criteria such as cost, size and power consumption [11].

The definition of 3GPP bands can be found in Table 11.1. Although the Phase Locked Loop (PLL) and RF blocks of a Transceiver (TRX) can be designed to cover almost all the bands, the designer still has to decide how many bands will be supported simultaneously in a given phone to optimize the radio. This drives the number and frequency range of Low Noise Amplifiers (LNAs) and transmit buffers. The same considerations exist for the Front-End (FE) components in terms of the number and combination of Power Amplifiers (PAs), filters and thus the number of antenna switch ports. Similarly, the number of supported bands required in diversity path needs to be decided.

The support of legacy standards in combination with band support results in a complex number of use cases. These need to be studied to provide an optimum solution in terms of size and cost. Here are some of the anticipated multi-mode combinations:

- EGPRS + WCDMA + LTE FDD
- EGPRS + TD-SCDMA + LTE TDD
- EVDO + LTE FDD.

The first two combinations are natural migration paths through 3GPP cellular technologies and standardization has taken care of measurement mechanisms to allow handover back and forth between each technology without requiring operation (receive or transmit) in two modes simultaneously. This allows all three modes to be covered with a single receive or transmit path, or two receive paths when diversity is required. In the latter, handover support is more complex but again a transmit and receive path combination is feasible since two receivers are available from the LTE diversity requirement.

These multi-mode requirements can be supported by one highly reconfigurable TRX Integrated Circuit (IC), which is often better known under the well-used term of 'Software Defined Radios'. In reality, this should be understood as software reconfigurable hardware. This re-configurability for multi-mode operation takes place mainly in the analogue baseband section of the receiver and the transmitter. The multi-band operation takes place in the RF and Local Oscillator (LO) section of the TRX and in the RF-FE.

The combination of the high number of bands to be supported together with the need for multi-mode support is driving RF sub-system architectures that optimize hardware reuse, especially in the FE where the size and number of components becomes an issue. Improvement in this area builds on the optimizations already realized in EGPRS/WCDMA terminals, driving them further to meet LTE functionality. This means that not only does the LTE functionality need to comply with the architecture framework used for 2G and 3G but it also explores new opportunities of hardware reuse:

- LTE performance should be achieved without the use of external filters: neither between LNA and mixer nor between the transmitter and the PA as this is already realized in some WCDMA designs. This not only removes two filters per band but also allows simplification of the design of a multi-band TRX IC. This is especially critical for the FDD mode and where large channel bandwidth is used in bands where the duplex spacing is small. These new design tradeoffs are discussed in section 11.8.2. Similarly the interface to the Baseband (BB*)* needs to multiplex every mode to a minimum number of wires; this is best achieved by a digital interface as described in section 11.6.3.

- Reuse of the same RF FE path for any given band irrespective of its operation mode. This implies using:
  - Co-banding: reuse of the same receive filter for any mode, especially EGPRS (half duplex) reuses the duplexer required for FDD modes.
  - Multi-mode PA: reuse of same PA whatever mode or band.

The details associated with these two techniques are developed in further sections. However it is instrumental to illustrate their benefits in terms of reduced complexity and thus size and cost. Table 11.14 shows the difference in complexity between a design with none or all of the above optimizations. The example is that of a US phone with international roaming supporting: quad-band EGPRS (bands 2/3/5/8), quad-band WCDMA (bands 1/2/4/5), and triple band LTE (bands 4/7/13). The fully optimized solution block diagram is illustrated in Figure 11.22.

The two scenarios show a difference of almost a factor of two in the number of components. In addition the lower component count also significantly simplifies the TRX IC and the FE layout and size. A number of partially optimized solutions also exist, however, that fall somewhere between these scenarios.

11.6.2.2 Antenna Diversity Requirement

One of the main features introduced with LTE is MIMO operation to enhance the available data rate. MIMO operation requires the UE to be equipped with two receive antennas and paths. The conformance testing is done through RF connectors and assumes totally uncorrelated antennas. This scenario is far from representative of real operation in the field especially for small terminals operating in the lower 700 MHz frequency band. Small terminals such as smart phones have dimensions that only allow a few centimeters of separation between the two antennas. At low frequencies this distance results in a high correlation between the two signals received at each antenna, which results in degraded diversity gain. Furthermore at these frequencies and with small terminal sizes, the hand effect (modification of antenna's radiation patterns due to the hand or head proximity) further degrades the diversity gain. For devices demanding higher data rates, such as laptops or PC tablets, the device sizes allow proper antenna design. Also in higher frequency bands even a small terminal can provide sufficient antenna separation to grant good MIMO operation.

Table 11.14 Component count for different front end design

Blocks	Implementation	
	Un-optimized 'side-by-side'	Fully optimized
Low Noise Amplifiers	13	10
Duplex filters	6	5.5*
Band-pass filters	17	5
Power Amplifiers	8	2
Switches	2	4
Transceivers	3	1
Total number of components	49	27

*Reuse of RX side for band 1 and 4.

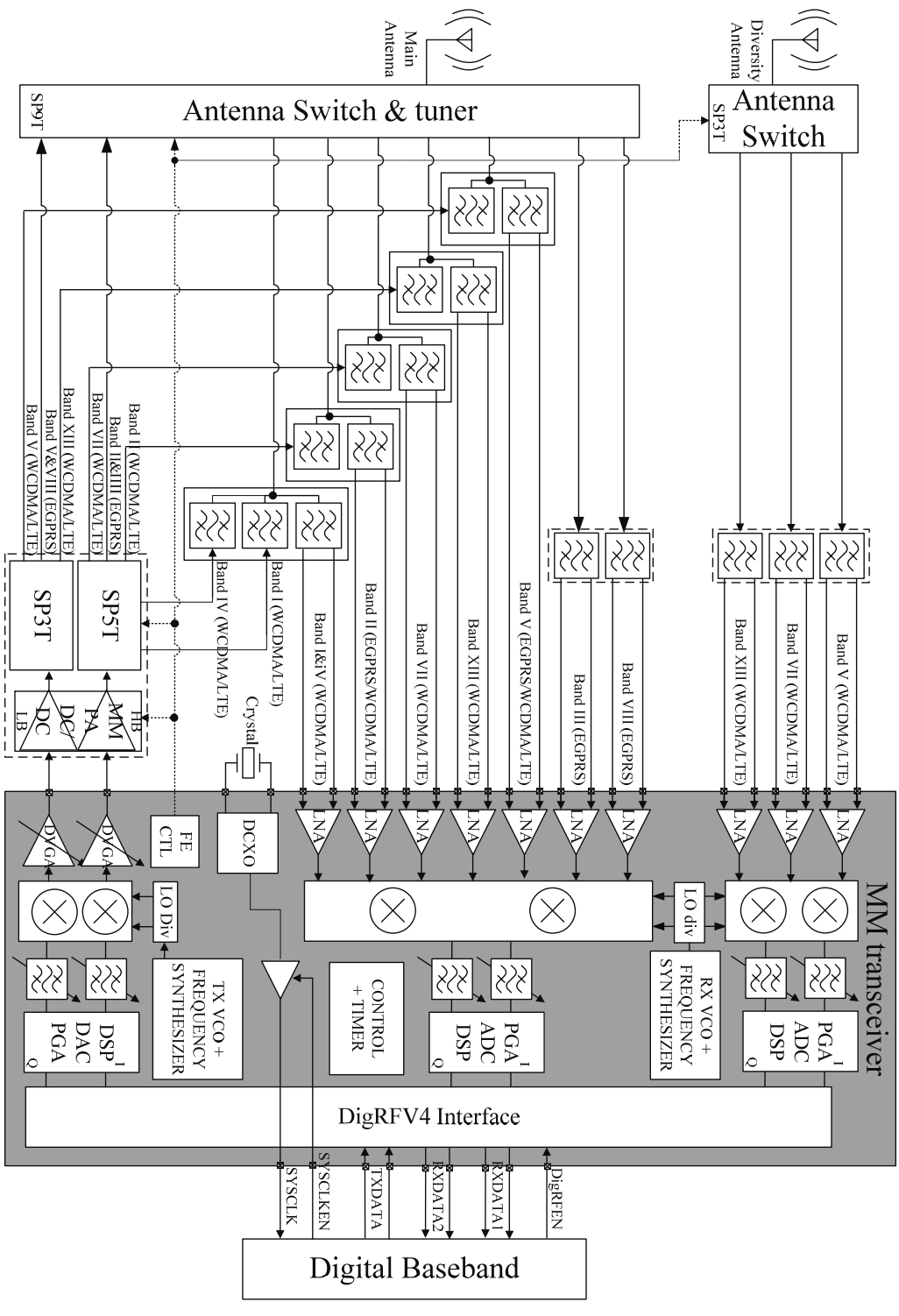


Figure 11.22 EGPRS/WCDMA/LTE optimized RF subsystem block diagram. HB, high band; LB, low band; DCXO, digital crystal oscillator; PGA, programmable gain amplifier; MM, multi-mode; FE, front-end

When the antenna design of an LTE UE is considered relative to current WCDMA phones the following design challenges are encountered:

1 The overall band support: the current WCDMA antenna frequency range is from 824 to 2170 MHz whereas future LTE devices will have to cover 698 to 2690 MHz. This stretches the current state of the art for antenna matching and also for maintaining antenna gain throughout the larger bandwidth. This will probably drive the introduction of new technology such as active antenna tuning modules.
2 Some LTE bands create new antenna coupling issues with other systems potentially present in the UE. The other antennas to be considered are the Global Positioning System (GPS) antenna, the Bluetooth (BT) and WLAN antenna, the analog Frequency Modulated (FM) radio antenna, and the Digital TV (Digital Video Broadcast – Handheld, DVB-H) antenna. The related critical coexistence use cases are discussed in section 11.6.2.3.
3 The support of antenna diversity: introducing an extra antenna in an already complex and relatively small UE presents a significant challenge if reasonable diversity gain is required.

The last two issues are easier to handle in data card designs where only the cellular modem is present. To some extent this is also true for the larger portable devices such as PC tablets and video viewers but smart phone mechanical design may prove particularly challenging for antenna placement.

11.6.2.3 New RF Coexistence Challenges

In the context of the multi-mode UE where multiple radio systems and multiple modems, such as BT, FM radio, GPS, WLAN and DVB-H must coexist, the larger bandwidth, the new modulation scheme and the new bands introduced in LTE create new coexistence challenges. In general coexistence issues are due to the transmit (TX) signal of one system (aggressor) negatively impacting another system's receiver ('RX'-victim) performance and notably its sensitivity. There are two aspects to consider: the direct rise of the victim's noise floor by the aggressor's transmitter out-of-band noise in the receiver band, and degradation of the receiver's performance due to blocking mechanisms.

Noise leakage from an aggressor TX in a victim's RX band is added to the RX noise floor further degrading its sensitivity, as shown in Figure 11.23. This noise leakage is only a function of the intrinsic TX noise, TX output filter attenuation in the victim's band and antenna isolation. This desensitization mechanism depends on the aggressor's TX design.

The desensitization of a victim due to an aggressor's transmitter out-of-band noise can be calculated as follows:

$$DES_{OOBN} = 10 * \mathrm{Log}(10^{(-174+NFvictim)} + 10^{(POUTaggressor-OOBN-ANTisol)}) + 174 - NFvictim \quad (11.11)$$

Where DES_{OOBN} is the resulting degradation of the victim's sensitivity in dB due to aggressor's TX noise, *NFvictim* is the victim's RX Noise Figure (NF) referred to the antenna in dB, *POUTaggressor* is the aggressor's TX maximum output power in dBm, *Out-of-band Noise (OOBN)* is the aggressor's TX noise in the victim's band in dBc/Hz, and *ANTisol* is the antenna isolation in dB.

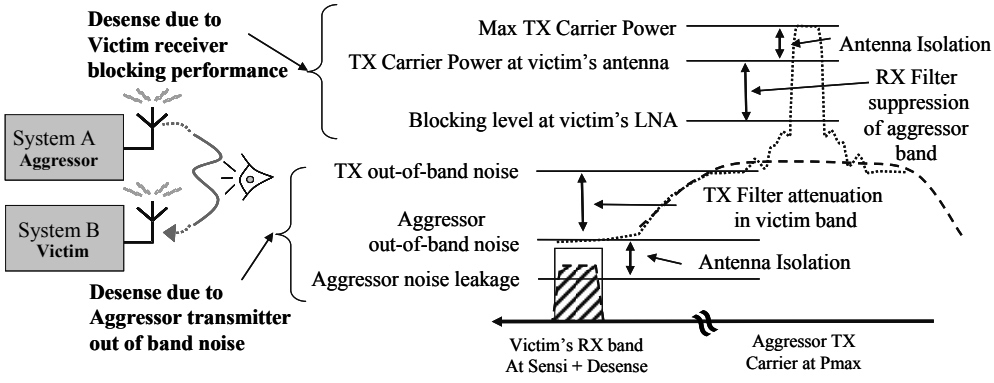


Figure 11.23 Victim/aggressor block diagram and aggressor transmitter leakage to victim receiver

The blocker power level present at the victim's LNA input depends on the aggressor maximum transmitted signal power, the antenna isolation and the victim's FE filter attenuation in the aggressor's TX band. Mitigation techniques can only be implemented in the victim's RX design. In both cases improved antenna isolation helps but the degrees of freedom are limited in small form factor UE especially when the aggressor and the victim operating frequencies are close to one another.

Victim desensitization due to the presence of a blocker may result from multiple mechanisms, as described in Figure 11.24.

A victim's RX LO phase noise is added to the wanted signal due to reciprocal mixing with the aggressor's transmit signal leakage:

1 Victim's RX reduced gain for wanted signal due to cross-compression on the TX signal leakage.
2 Second order intermodulation distortion (IMD_2) product of the TX signal leakage AM content falling at DC for Direct Conversion Receivers (DCR) which potentially overlap with the wanted signal.

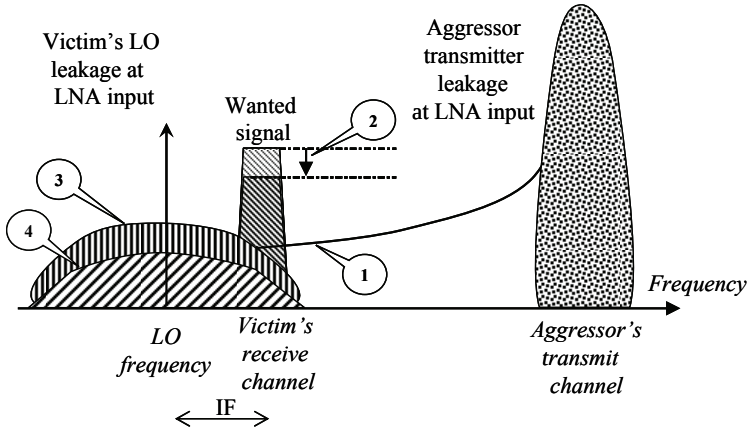


Figure 11.24 Desensitization mechanisms

3 Third order intermodulation product of victim's LO leakage with TX signal leakage AM content falling at DC which potentially overlap with wanted signal (known as cross-modulation).

For the first two cases, the actual aggressor signal characteristics and bandwidth do not play a role in the interference mechanism:

- Cross-compression is directly related to the signal peak power, which has been kept constant between WCDMA and LTE uplink thanks to the Maximum Power Reduction (MPR) introduced in 3G.
- Reciprocal mixing is directly related to interference leakage and the victim's LO Phase Noise at an offset equal to the difference between the aggressor's TX frequency and the victim's RX frequency. The larger distance improves both the selectivity on the aggressor leakage and the LO Phase Noise.

Similarly the aggressor's noise leakage is lower for higher frequency distance due to higher selectivity and lower Phase Noise at larger frequency offsets. From the above it can be seen that for mechanisms such as TX noise leakage, cross-compression and reciprocal mixing, the prime factor is the vicinity in the frequency domain of the aggressor and the victim. New band allocation for LTE has created new cases described in the band specific coexistence challenge section and in section 11.8.2.

The cross-modulation and the IMD_2 products, however, have a spectrum bandwidth (BW) directly related to the interferer's BW. Furthermore, the spectrum shape of these distortions depends on the signal statistics and modulation scheme. For these contributors it is difficult to propose a generic analysis of how the LTE interference compares to WCDMA. Analysis of which portion of the interfering spectrum is captured within the victim's channel bandwidth needs to be conducted for each LTE bandwidth and modulation combination. A comparison of IMD_2 products for WCDMA and LTE QPSK uplink modulated carriers can be found in section 11.8.2.1.

Band Specific Coexistence Challenges

As far as the LTE out-of-band TX noise emissions are concerned, the duplexer must provide enough attenuation in the following frequency bands:

- In band 11, the TX frequency is very close to GPS band. With the anticipated antenna isolation, more than 40 dB attenuation needs to be provided. GPS receivers may also have to improve their blocker handling capability to cope with LTE TX signal leakage.
- In bands 11/12/13/14, the TX frequencies are close to some of the TV bands, and consequently adequate attenuation of LTE TX noise is required.
- In bands 7/38/40, the TX frequencies are very close to the 2.4-GHz ISM band where BT and WLAN operate. With the anticipated antenna isolation, approximately 40 dB attenuation needs to be provided. Similarly the BT and WLAN TX noise needs to be kept low in the Band 7/38/40 RX frequencies. Given the blocker handling capability already required for the LTE FDD RX, the BT and WLAN blockers are not anticipated to be an issue. BT/WLAN receivers, however, may have to improve their blocker handling capability to cope with the LTE TX signal leakage.

As far as the LTE TX modulated carrier is concerned, with a 23 dBm output power, special attention must be paid to controlling the LTE TX second harmonic power level in the following aggressor/victim combinations:

- In Band 7/38, the TX second harmonic falls within the 5 GHz WLAN band, thus requiring the LTE transmitter to attenuate it to close to −90 dBm. In comparison to the 3GPP LTE which requires a harmonic level to be lower than −40 dBm, this represents a significant challenge. Also, the WLAN RX must be sufficiently linear to prevent regenerating this second harmonic in the presence of the LTE fundamental TX signal leakage.
- In Band 13, the TX second harmonic falls within the GPS band thus requiring the LTE transmitter to attenuate it to lower than −110 dBm. This level of attenuation might be hard to achieve practically. Similarly to the above coexistence test case, the LTE TX leakage presence at the GPS RX input also imposes linearity requirements to prevent regenerating this second harmonic.

There are also other cases where a harmonic of the LTE TX falls in the 5 GHz WLAN or UWB bands. These are usually higher harmonic orders, which should be less of an issue.

11.6.3 RF–Baseband Interface Design Challenges

In mobile handsets, the RF transceiver and the BB processor are often implemented on separate ICs. The TRX IC normally contains analog signal processing, while the BB IC is predominantly digital. Therefore analog-to-digital (A/D) and digital-to-analog (D/A) conversions are required in receive and transmit paths respectively. The location of these converters is a critical choice in wireless system design. If the converters are implemented in the RF transceiver, discrete time domain (digital) data are transferred across the interface between the BB and the TRX. On the other hand, if the converters are located in the BB IC, the interface comprises continuous time domain (analog) differential I/Q signals.

Figure 11.25 shows a typical example of an analog based interface with receive diversity. Here this mobile is designed to support all of these radio operating options: '3G' Wideband Code-Division Multiple Access (WCDMA) on one frequency band plus GSM on four frequency bands. It is readily seen that 24 separate interconnections are required.

Despite being in mass production from many vendors, analog interface solutions face the following challenges and criticisms. The large number of interconnecting pins increases package size and cost on both sides, and complicates printed circuit board routing. This is a particular problem because the sensitive analog interconnections also require special shielding to achieve the required performance. Furthermore, such interfaces are proprietary, forcing handset makers to use particular pairs of BB and RF devices. While some IC vendors prefer this forced restriction, it is a disservice to the industry by restricting competition and disabling creative approaches in either the BB or RF side alone from adoption in products. BB devices must include the necessary large number of analog blocks, increasing their die size.

This last point is particularly important, involving much more than the obvious economic issues. Clearly, larger die areas increase the cost of IC manufacture. Also these analog designs do not generally shrink as well as digital cells with progressively smaller CMOS processes. Thus, the fraction of the BB IC that these analog designs take up increases. Additionally, the analog circuitry yield in production may be lower than the digital circuitry, so a perfectly fine

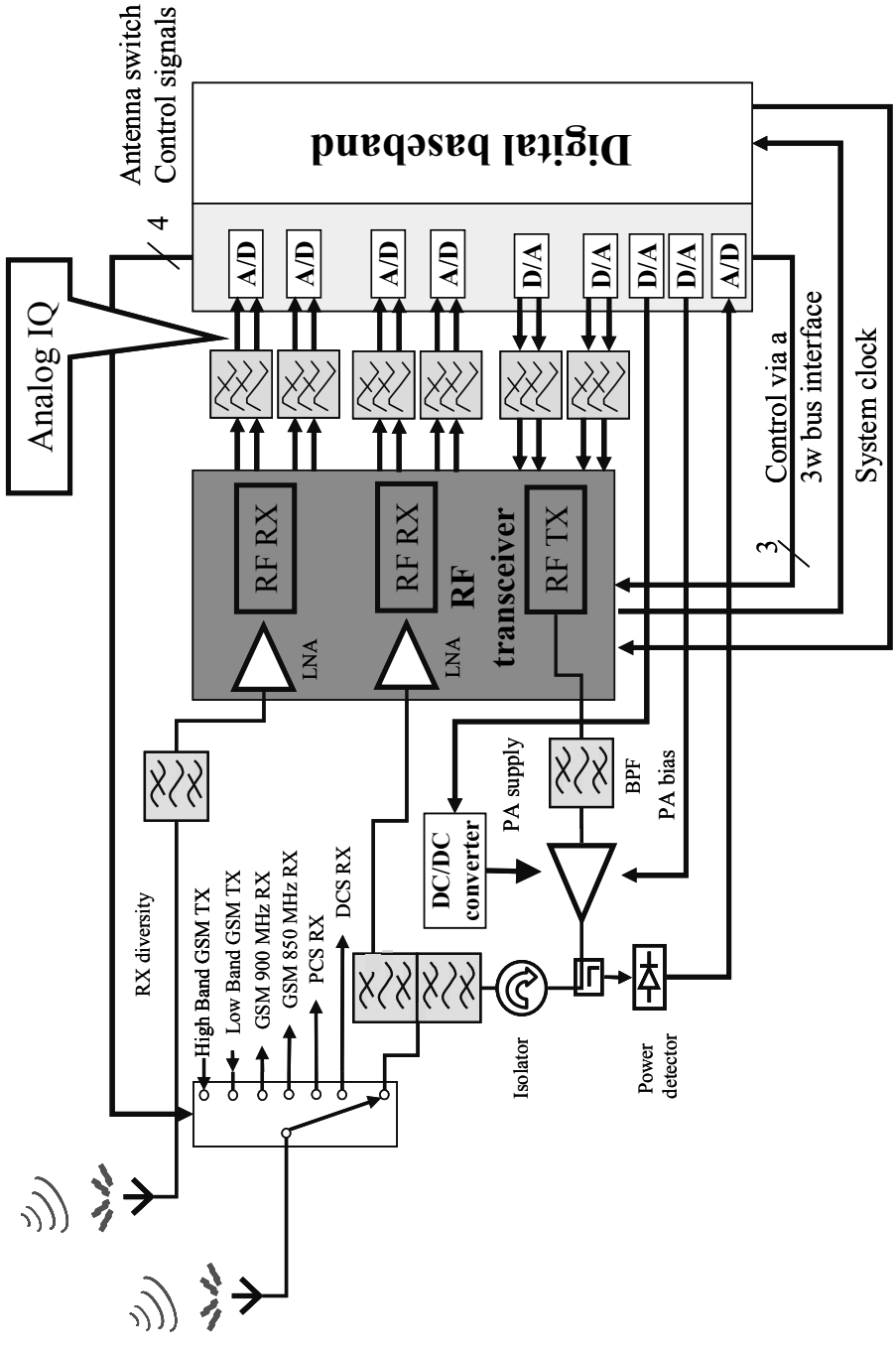


Figure 11.25 Typical HSPA monoband 3G, quad band 2G transceiver/BB block partitioning in analog I/Q interface

digital BB would be thrown away when an analog section fails. Even more of a problem is that analog design information on a new CMOS process is always provided later than the digital design information – sometimes much later. When new CMOS processes become available, this forces BB designs to wait until all the necessary design information is available and sufficiently qualified. In the newer CMOS processes, analog designs are actually getting more difficult, not easier.

Taking all of this together, it becomes clear that a purely digital interface is greatly desired, and would be a huge benefit to the mobile industry. The main organization leading the standardization effort is MIPI® – the Mobile Industry Processor Interface Alliance [12]. Combining the terms 'digital' and 'RF' together into the name 'DigRF^{SM}', this interface is already in its third evolutionary step, as listed in Table 11.15.

Adopting these digital interfaces changes the earlier block diagrams to that of Figure 11.26. The dual objective of eliminating analog designs from the BB device and reducing pin count down to only 7 pins is met. For DigRF^{SM} v4, the interface data rate necessary to support just a single antenna receive chain in a 20 MHz LTE application reaches 1248 Mbps.

One of the biggest challenges in DigRF^{SM} v4 is EMI control. EMI was not a major concern for DigRF v2 and v3, because these interface data rates are well below the mobile's radio operating frequencies. With all the bands now under consideration for LTE mobile device use, the internal radio Low Noise Amplifier (LNA) section is sensitive to energy across frequencies from 700 MHz to nearly 6 GHz. The DigRFv4^{SM} data rates and common mode spectrum emissions now exceed the radio frequencies of several mobile operating bands. This is a huge problem, clearly seen in Figure 11.29(a).

Before addressing EMI mitigation techniques, it is essential to understand how much signal isolation is available from the physical packaging of the RF IC. Adopting the vocabulary of EMI engineering, the LNA input is called the 'victim', while the interface is called the 'aggressor'. As we can see in Figure 11.27 (left), there are many possible paths for energy on the interface to couple into the 'victim' LNA. Figure 11.27 (right – plain dots) shows one example of practically available isolation.

The upper limit to the maximum amount of aggressor noise PSD allowed at the victim's input pins is established with the set of characteristic curves shown in Figure 11.28. Considering worst cases, a cellular LNA with an intrinsic NF of 3 dB may be degraded to a 3.5 dB NF. According to these charts, the interfering noise must be at least 6 dB below the kT floor, i.e. at or below −180 dBm/Hz. Similarly, a more sensitive GPS LNA with an intrinsic NF of 2 dB may be degraded by at most 0.25 dB. Evaluating these two cases, we conclude that the interference must be below −180 dBm/Hz and −184 dBm/Hz for a cellular and a GPS RX respectively.

The limit curve shown in Figure 11.29 (a–c) is based on combining the isolation model with the cellular LNA noise tolerance and shows violation in Figure 11.29(a). Mitigation techniques are therefore required to ensure product success. First on the mitigation techniques list is to

Table 11.15 DigRF version evolutions

DigRF version	Standard	Interface bitrate (Mbps)
v2: 2G	GSM/GPRS/EDGE	26
v3: 3G	2G + HSPA	312
v4: 4G	3G + LTE	1248, 1456, 2498, 2912

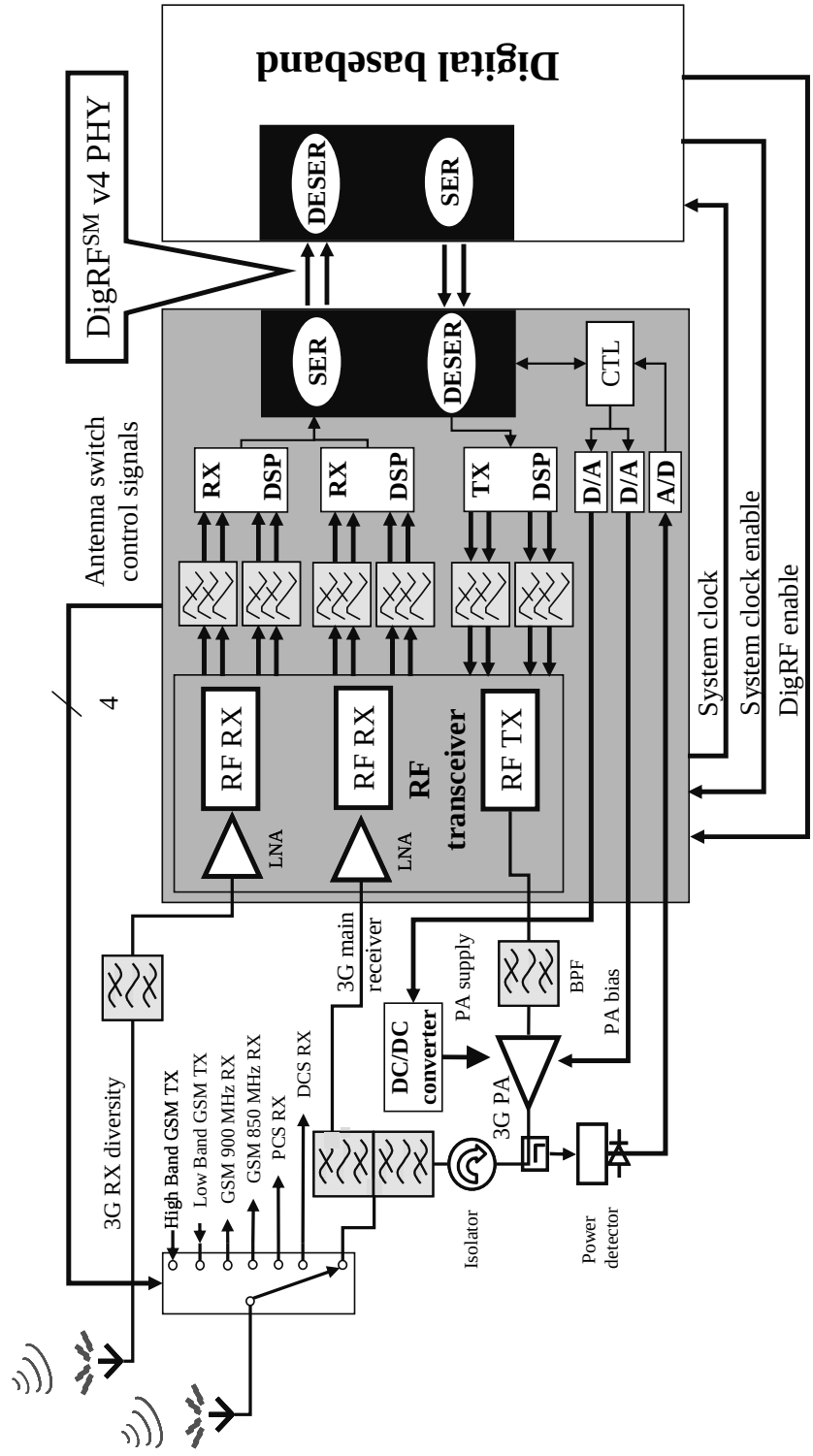


Figure 11.26 Example of digital RF–BB interface application using DigRF^SM v4 with RX diversity. A/D = analog to digital converter, D/A = digital to analog converter, SER = serialize, DESER = deserialize. DSP functions include A/D, D/A, digital filtering and decimation (specific to RX)

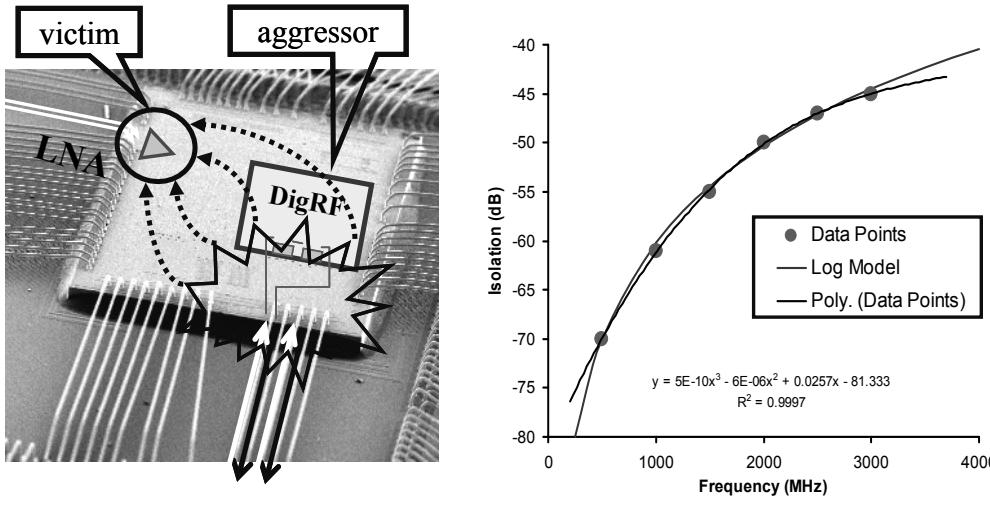


Figure 11.27 Left: examples of EMI coupling paths within an RF IC. Right: in-package isolation example at several frequencies

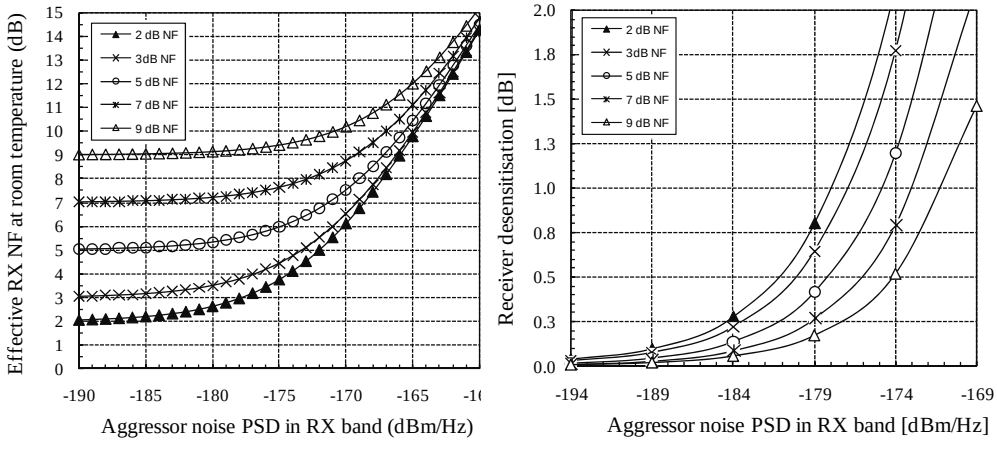


Figure 11.28 Response of a 'victim' LNA to additional noise provided at its input. Left: effective noise rise of a victim NF referred to its LNA input (room temperature). Right: corresponding victim's desense vs aggressor noise PSD

use control of the bit edge slew rate, and the situation is greatly improved as shown in Figure 11.29(b). Secondly, an alternate interface frequency is provided, which can be switched to if a mobile product is experiencing spurious interference from the interface. This alternate interface data rate, 1456 Mbps, is also particularly useful in protecting a local GPS RX input by ensuring the LNA is now operating close to a frequency null of the interface spectrum, as shown in Figure 11.29 (c).

As Figure 11.30 shows at 1248 Mbps, however, a single antenna 20 MHz LTE application requires 70% duty cycle of the interface. Adding the LTE diversity RX, the requirement exceeds 100%! One solution consists in doubling the number of 1248 Mbps lanes. Alternatively, for

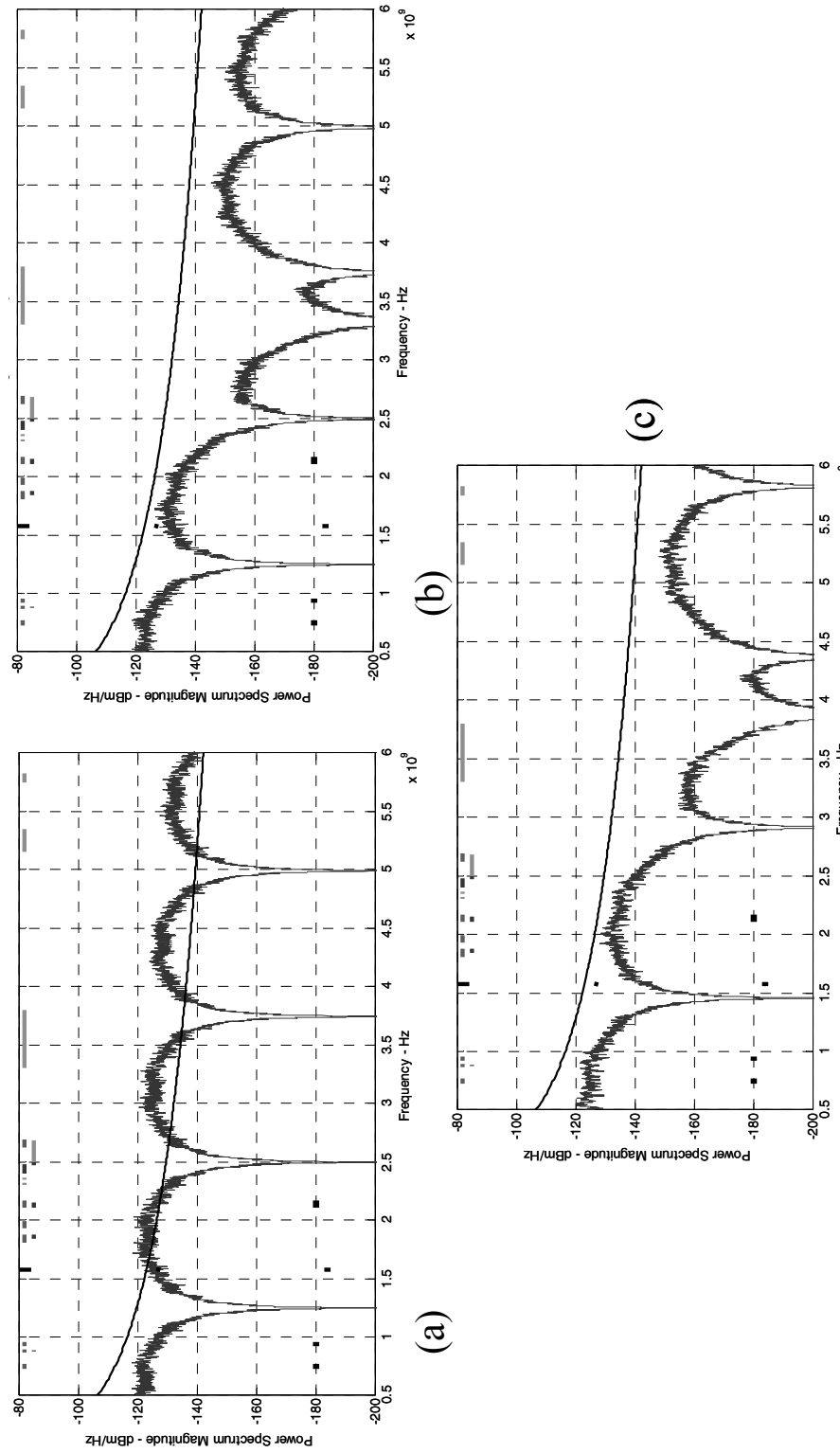


Figure 11.29 Effects from adding EMI mitigation features to the DigRF/M–PHY interface: (a) imperfect differential signaling, (b) slew rate control, and (c) alternate frequency selection of 1456 Mbit/s. Horizontal bars at top of each diagram indicate the location of telecommunication standards victim's frequency bands from 700 MHz to 6 GHz

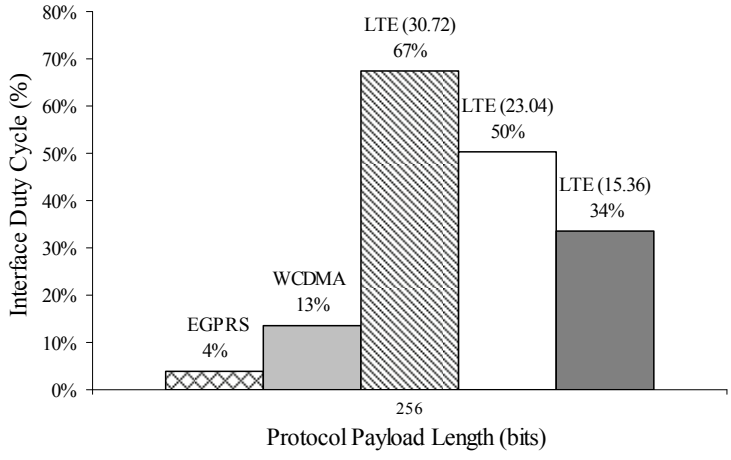


Figure 11.30 Interface duty-cycles for DigRF^{SM} v4 with 256 bit payload field size at 1.248 Gbps

applications which do not have significant EMI sensitivity, doubling the interface bit-rate to 2496 Mbps solves the capacity issue, but now generates a main lobe of the interface spectrum that spans over all bands, including the super-sensitive GPS band.

11.6.4 LTE vs HSDPA Baseband Design Complexity

11.6.4.1 Equalization

LTE aims at reducing the cost per bit, among other challenging goals such as increasing spectral efficiency and allowing flexible bandwidth deployment. From the receiver point of view, one key measure is the required complexity and in particular the comparison with previous releases such as WCDMA and HSDPA. Figure 11.31 shows the estimated complexity based on the baseline receiver for all transmission modes, as introduced in section 11.9, excluding channel decoding operation complexity. Note that the complexity of the LTE receiver grows linearly with respect to system bandwidth and the corresponding maximum nominal throughput. Interestingly, MIMO mode requires less than double the SIMO mode complexity. Comparing the estimated complexity of a Release 6 optimistic HSDPA receiver assuming a low complexity chip-level equalizer category 10 device with 13 Mbps throughput, LTE is shown to be an attractive technology at least from an inner receiver complexity point of view because of OFDM choice. Assuming 5 MHz bandwidth and 16QAM modulation, LTE offers the same throughput with nearly half of the required complexity of HSDPA.

Nevertheless, despite this advantage in the relative comparison for the 5 MHz bandwidth case, the LTE receiver requires a considerable complexity increase compared to HSDPA since the UE must support a system bandwidth of 20 MHz as a minimum requirement and therefore still constitutes a challenge for mobile and chip-set manufacturers. The LTE class 4 device with 150 Mbps capability has a complexity approximately four times higher than a HSDPA category 10 device with 13 Mbps capability.

Further details of the repartition of complexity among inner receiver stages are also given in Figure 11.32. Along the supported modes, the complexity associated with FFT operations – always

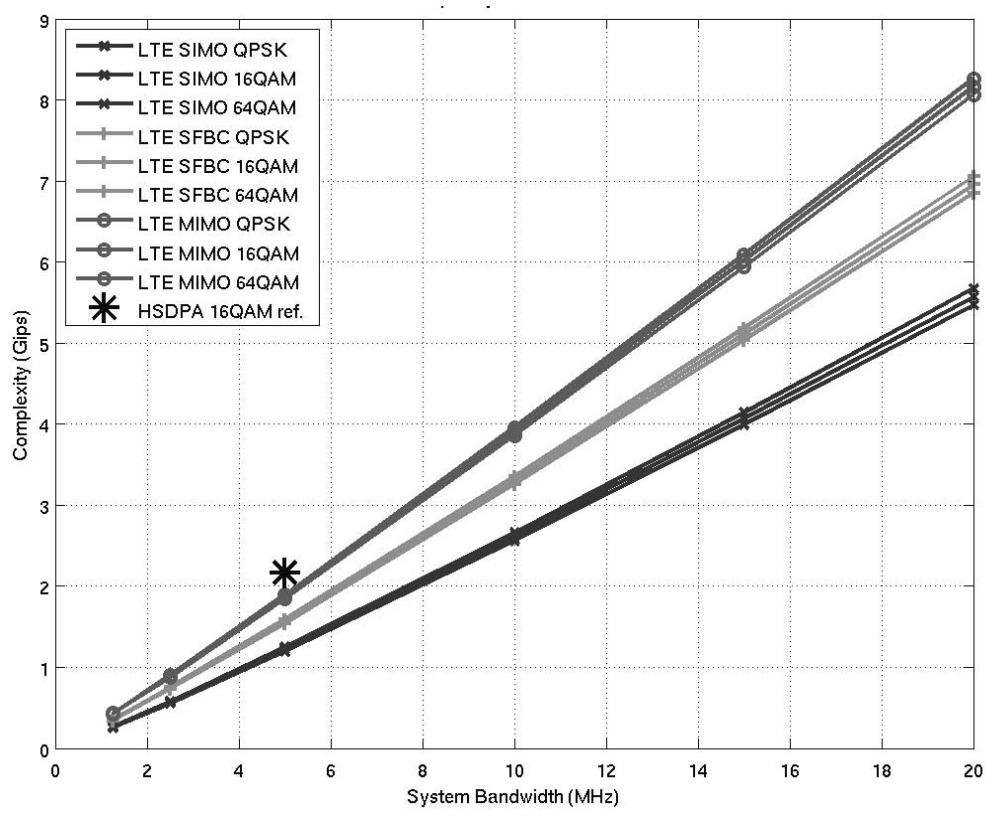


Figure 11.31 Complexity of LTE receiver

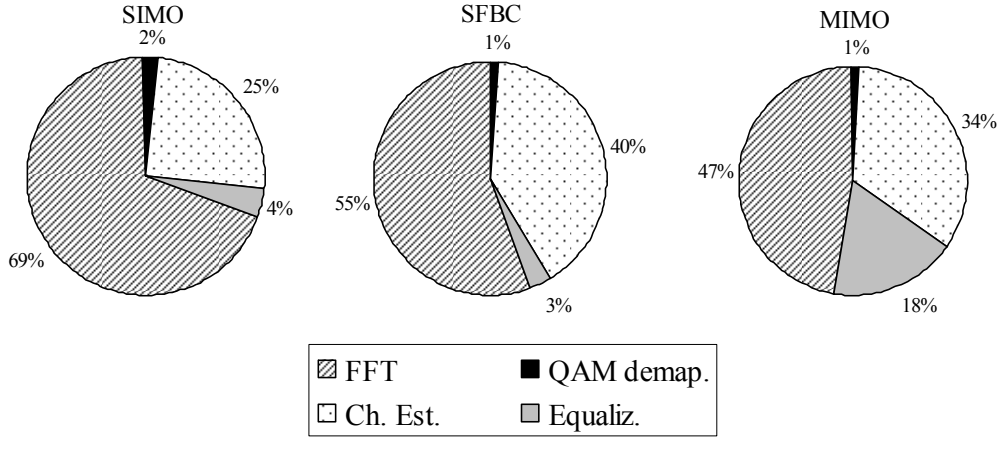


Figure 11.32 Complexity repartition in LTE receiver

equal to two as the number of receiving antennas – becomes less important for overall complexity because of the increased needs of channel estimation and equalization. With MIMO the operation FFT takes 47% of the calculations while channel estimation takes 34% and equalization 18%.

11.6.4.2 Turbo Decoder

In LTE, the efficiency and the complexity of the channel decoding operation grows considerably as the maximum nominal data rate increases compared to previous releases. The Turbo decoder must support rates up to 150 Mbps in a category 4 device. The Turbo code algorithm in LTE is similar to the Turbo code in HSDPA. The Turbo decoder processing delay, and hence its throughput, is roughly linearly proportional to the number of turbo decoding iterations and the size of the transport block. As a consequence, we can describe the Turbo decoder efficiency as

$$\eta = \frac{N_{it} \cdot r_{max}}{f_{clock}} \tag{11.12}$$

where N_{it} is the number of iterations, r_{max} is the maximum data rate and f_{clock} is the working clock frequency of the turbo decoder. Table 11.16 and Table 11.17 show the efficiency/clock-frequency tradeoff for the HSDPA and LTE cases assuming a number of iterations equal to 8 and a maximum data rate of 13 Mbps and 150 Mbps respectively.

Efficiency can be seen as a measure of the parallelization required within the Turbo decoding operation. Comparing the two tables, it seems evident that to support a LTE data rate within reasonable clock frequencies – strictly dictated by power consumption and hardware technology constraints – a high level of parallelization is imposed.

Table 11.16 Turbo decoder efficiency/clock frequency tradeoff for 13 Mbps

Efficiency η (b/s)	Clock frequency (MHz)
5.5	20
2.7	40
1.8	60
1.4	80
1.1	100
0.9	120
0.8	140
0.7	160
0.6	180

Table 11.17 Turbo decoder efficiency/clock frequency tradeoff for 150 Mbps

Efficiency η (b/s)	Clock frequency (MHz)
12.6	100
10.5	120
9.0	140
7.9	160
7.0	180
6.3	200
5.7	220
5.2	240
4.8	260

It is worth mentioning that Turbo decoding parallelization was not attainable given the construction of the interleaver used for the encoding and decoding procedure of previous releases of the 3GPP standard. This problem was solved for the LTE specification and a contention-free interleaver has been built to allow any level of parallelization expressed as a power of two (i.e. 2, 4, 8, 16).

Nevertheless, parallelization itself does not come for free and also constitutes a challenge for manufacturers: it demands an in-depth change in the Turbo decoding algorithm, increases the surface and gate count of an amount proportional to the parallelization factor and, in principle, an operational higher clock frequency.

As a final consideration, it is worth noting that Turbo Decoder complexity is in principle much higher than the complexity of the rest of the mobile receiver signal processing: complexity grows linearly with the target throughput and number of iterations. It can be approximated by [48]:

$$C_{TD} = 200 \cdot N_{it} \cdot r_{max} (\text{MIPS}) \tag{11.13}$$

For LTE, with a maximum data rate of 150 Mbps, the complexity approaches 240 Gips while the FFT, equalization and channel estimation complexity together are in the order of 9 Gips.

The Turbo Decoder then seems to require nearly 96% of the overall complexity in the receiver but this is a rather misleading conclusion. Signal processing involved in the Turbo Decoding is mainly addition operations in the max-log-MAP case. Equalization and associated functions require multiplication operations instead and in principle larger fixed-point sizes. For the implementation of choices making a difference, signal processing complexity is a valuable measure of the challenge required among standard evolutions for the same functionality but cannot allow for more general comparisons.

11.7 UE RF Transmitter

11.7.1 LTE UE Transmitter Requirement

11.7.1.1 Transmitter Output Power

The LTE specified maximum output power window is the same as in WCDMA: 23 dBm with a tolerance of ±2 dB. The earlier WCDMA specifications used 24 dBm with a tolerance of +1/−3 dB. The SC-OFDMA [13] has a higher Peak-to-Average Ratio (PAR) than the HPSK modulation of WCDMA. Figure 11.33 shows the ACLR performance of a WCDMA PA under QPSK modulated SC-OFDMA signal. The WCDMA operation assumes 24 dBm and LTE operation at 23 dBm output power. The main difference is related to the spectral shape and the fact that the occupied BW is slightly higher in LTE (4.5 MHz) than that of WCDMA (99% energy in 4.2 MHz), and consequently the ACLR is slightly degraded.

In a similar manner to HSDPA and HSUPA, a Maximum Power Reduction (MPR) has been introduced in LTE to take into account the higher PAR of 16QAM modulation and some resource block allocation. Again this ensures a proper ACLR under a complex set of TX modulations. Compared to WCDMA, where the only direct interference falling in the RX band is related to spurs and the OOB noise, the LTE TX linearity should also be considered. It is particularly important for 10 MHz BW in the 700 MHz bands where the duplex distance is only 30 MHz

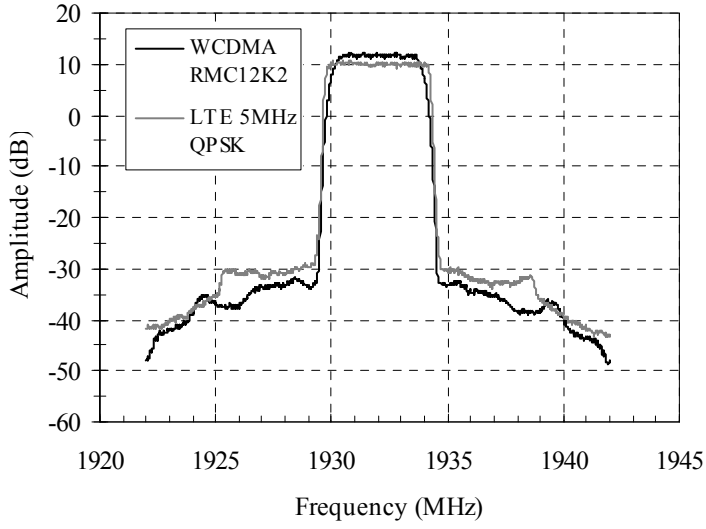


Figure 11.33 24 dBm WCDMA and 23 dBm LTE 5 MHz QPSK spectrums

and where 5th and 7th order intermodulation products of the TX overlap with the RX channel. This phenomenon is discussed in section 11.8.2.

LTE power control ranges from −40 dBm to +23 dBm. WCDMA transmitters offer −50 dBm to +23 dBm of Transmit Power Control (TPC) with 1 dB resolution. The same TPC techniques can be applied in LTE taking into account all MPR cases.

11.7.2 LTE Transmit Modulation Accuracy, EVM

The LTE 16QAM modulation case places stringent requirements on the TX imperfections to meet the Error Vector Magnitude (EVM) budget. Overall the errors should be further minimized for LTE. Each contributor can be analyzed separately to meet the 17.5% EVM budget in QPSK and 12.5% in 16QAM. EVM measurements are done after Zero-Forcing (ZF) equalization in the test equipment. For comparison, the WCDMA HPSK EVM budget is equal to 17.5%, but does not assume any equalization in the test instrument. In practice the design target is around 8%. The TX imperfections contributing to the EVM budget are considered in the next section.

- Carrier rejection: As the measurement is made after ZF-equalization the test equipment partially removes this contribution. LTE standardization has set a separate specification for carrier rejection to ensure that the carrier level stays within reasonable limits for the ZF algorithm. At low output power small DC offsets in the TX chain generate a high carrier leakage to a point where TPC accuracy may not be met anymore. This problem is already severe in WCDMA where the TPC range is 10 dB larger at the bottom end. For this reason carrier leakage compensation is implemented in RF TRXs and this technique generally achieves 40 dB of carrier rejection throughout the TPC range making this contribution negligible in the EVM budget.
- Even order distortion: Even order non-linearity contributes mainly to ACLR as the main effect is to enlarge the transmitted spectrum. As ACLR requirements are 33 dB and 43 dB in adjacent and alternate channels, these contributions are usually negligible in terms of EVM. AM/AM in the PA can be considered similarly.

- LO phase noise: The induced jitter generates phase error in the modulation constellation thus contributing to EVM. As the SC-OFDMA used in transmit is single carrier the phase noise has a contribution similar to the WCDMA case.
- PA distortion: AM/PM distortion (together with AM/AM) has a contribution to EVM and also generates asymmetrical spectral energy in the adjacent channels. AM/PM and AM/AM distortion increases as the transmit power is increased. Hence, the LTE MPR scheme ensures that PAs dimensioned for WCDMA will not dominate the overall EVM budget.
- Image: The signal image generated by the quadrature imperfections in the up-mixing process can be considered as band noise contributing to the TX signal SNR. The use of 16QAM requires better control of such imperfections relative to WCDMA [14].
- Group delay distortion: Compared to WCDMA where the I/Q BB BW is 2 MHz for a minimum duplex distance of 45 MHz, LTE has an I/Q BW of 5 MHz for 30 MHz duplex distance or 10 MHz for 80 MHz duplex distance. Significant BB filtering is required to ensure that BB I/Q noise at the duplex frequency offset is very low (to allow SAW-less transmit architecture). This means that stop-band attenuation for LTE BB filter requires a more stringent filter specification than for WCDMA, potentially introducing more in-band distortion contributing to EVM.

Overall, the LTE EVM specification requires a similar effort to WCDMA for RF imperfection but special attention needs to be paid to the higher bandwidth and smaller duplex distance. For the same reasons, the TX out-of-band noise is more difficult to achieve in LTE.

11.7.3 Desensitization for Band and Bandwidth Combinations (Desense)

Although TX out-of-band noise in RX band is not an explicit requirement in the 3GPP WCDMA specifications, reference sensitivity measurements are made in the presence of the full TX power. To meet the reference sensitivity the TX noise leakage levels must stay below the thermal noise floor. Recent efforts in TX architecture have allowed the removal of the filter between the TRX IC and the PA. Interstage filters were used to clean the TRX noise before further amplification. Removal of the filter is made feasible by careful design of every noise sources in the RF TRX IC.

As discussed in section 11.6, it is essential that the addition of the LTE functionality and the support of new bands do not require reintroduction of these filters. Two new challenges must be solved for this to be feasible: the smaller duplex separation of certain new band configurations (12/13/14) or a wider channel bandwidth. In some specific cases these two challenges are combined. This issue has been recognized by the standardization body, as shown in Table 11.4 where the relaxed sensitivity requirements have been defined for certain combinations of band and bandwidth. The severity of the UE self-desense is further described in section 11.8.2.

11.7.4 Transmitter Architecture

11.7.4.1 Transmit RF Modulator

Direct up conversion is the obvious choice for a 2G/3G/LTE multi-mode TX. It is the de facto standard for WCDMA. The large BW requirements of LTE would pose further challenges to alternative architectures such as polar modulation or other non-Cartesian approaches [15], [16].

This is especially true if the external filter requirement is to be relaxed for FDD. As briefly discussed in the previous sections, thanks to a pragmatic approach in the LTE standardization there is a minimum number of modifications needed to provide LTE transmit capability from a Cartesian transmitter already covering 2G and WCDMA. The main modifications lie in the higher bandwidth requirement on the BB DAC and filter to cover all the different BWs with low out-of-band noise. In addition, extra RF bands need to be supported, which requires extension of the PLLs' tuning range and RF buffer bandwidths.

11.7.4.2 Multi-mode Power Amplifier

As discussed in section 11.6, one essential simplification of the worldwide RF FE is the use of a single PA line-up covering multiple bands and multiple modes. The band coverage can be clustered in the following way:

- one Low Band (LB) PA line-up covering all bands between 698 MHz and 915 MHz;
- one High band (HB) PA line-up covering all bands between 1710 MHz and 2025 MHz;
- one Higher band PA line-up covering all bands between 2300 MHz and 2620 MHz.

The only band that is not covered is the Japanese Band 11, which can be added for this specific phone configuration or even replace one of the other wideband PAs depending on the overall band support.

Each of these line-ups has to support different modulation schemes and maximum output power depending on band–mode combinations. These combinations can be found in Table 11.18, where 2 dB and 3 dB PA to antenna losses are considered for TDD and FDD modes respectively.

Taking into account the different PAR inherent to each modulation scheme and the required back-off to meet ACLR requirements, a given line-up has to meet a range of saturated output power (P_{outsat}) capabilities to achieve best efficiency/linearity tradeoffs. For example the LB PA has to achieve 35 dBm P_{outsat} capability for GSM, GMSK having only phase modulation the PA can be driven into saturation and achieve best efficiency. In WCDMA mode, it needs close to 31 dBm P_{outsat} capability to allow sufficient back-off. A GMSK capable PA would have a very low efficiency in the WCDMA mode if nothing was done to decrease its output power capability. This would prove the multi-mode PA architecture to be uncompetitive in terms of performance especially in 3G modes, which are already challenging for talk time. The only

Table 11.18 Modulation scheme and maximum output power per band configurations for a multi-mode PA

Modes	Sub-bands		
	Low band 698–915 MHz	High band 1710–2025 MHz	Higher band 2300–2620 MHz
GSM (GMSK)	35 dBm	32 dBm	n-a
EDGE (8PSK)	29 dBm	28 dBm	n-a
WCDMA (HPSK)	27 dBm	27 dBm	27 dBm
LTE (QPSK)	26 dBm	26 dBm	26 dBm

way to reach the best PA efficiency in all modes is to tune the output stage load line. This can be achieved in two ways:

- Tuning the output matching to transform the load impedance (usually 50 Ω) into the desired load line for every mode. This technique can be achieved for a small set of impedances and usually results in a lower Q output matching.
- Tuning the PA supply: In this case the saturated output power capability is proportional to the square of the supply voltage. If the supply is varied using a DC/DC converter then efficiency can be optimized for every mode. This technique is becoming more and more popular [17] and has the benefit of allowing optimum efficiency for every mode.

11.7.4.3 Conclusion

Although this section does not provide a detailed analysis of the LTE transmitter requirement it discusses how these requirements can be achieved by simple extrapolation from GSM/EDGE/WCDMA architecture and performance. It is shown that all three modes can be supported with a single transmitter architecture with the minimum of extra hardware when techniques such as co-banding and DC/DC controlled multi-mode PA are introduced. This ensures easy migration towards LTE for mobile equipment manufacturers.

11.8 UE RF Receiver Requirements

The purpose of this section is to highlight the main differences between Frequency Division Duplex (FDD) WCDMA and Full Duplex (FD) – FDD LTE UE RF receiver (RX) system requirements. Throughout this section, WCDMA and LTE are used to refer to UTRA and E-UTRA respectively. The objective of the LTE UE RF test requirements listed in section 7 of [1] is to quantify RF impairments that have an impact on the network performance. These impairments include Noise Figure (NF), receiver Error Vector Magnitude (EVM), selectivity at different frequencies, including adjacent channel, etc. The test requirements have been derived to ensure the industry can make the best possible re-use of IPs developed for WCDMA UEs. This is highlighted in the description of the UE reference sensitivity level and the Adjacent Channel Selectivity (ACS) system requirements. For this reason, the chapter focuses on some of the novel design challenges that are specific to the LTE downlink modulation schemes and its associated new frequency bands. In this respect, the following challenges are discussed: RX self-desensitization, ADC design challenges, and the impact of RX EVM contributors in OFDM vs single carrier.

11.8.1 Reference Sensitivity Level

The reference sensitivity power level is the minimum mean power applied to both UE antenna ports at which a minimum throughput requirement will be fulfilled. The throughput will equal or exceed 95% of the maximum throughput for a specified Fixed Reference Channel (FRC). FRCs are similar to the WCDMA reference measurement channel and in the sensitivity test case the downlink carrier uses QPSK modulation and 1/3 coding rate.

The sensitivity requirement verifies the UE RX NF, which for FDD operation may include noise contributions due to the presence of the UE uplink modulated carrier as described in the next section. Other receiver impairments such as EVM are included within the demodulation performance requirements where a higher Signal-to-Noise Ratio (SNR) is applied. Therefore, the selected FRC provides a reference maximum throughput defined for low SNR operation. Beyond this purpose, the UE reference sensitivity is of primary importance in 36.101 since it also serves as a baseline to set the downlink carrier power for ACS and blocker test requirements.

With LTE, the reference sensitivity requirements present several major differences compared to the WCDMA system requirements:

- LTE flexible bandwidth requires the RF receiver to implement reconfigurability of its channel select filters to support 1.4, 3, 5, 10, 15 and 20 MHz bandwidths.
- The reference sensitivity must be tested on two antenna ports: main and diversity receiver.
- New frequency bands with small Duplex Gap (DG), such as Bands 5, 6, 8 and 11 introduce novel UE self-desense considerations when UE uplink transmission bandwidths are greater than 5 MHz. Also Bands 12, 13, 14 and the more recent Band 17, all falling under the terminology UMTS 700, are impacted by these limitations. This is not a major concern in, for example, WCDMA Band I devices because the UE self-desense is primarily dominated by the UE transmitter chain noise floor at large frequency offsets, as can be seen in Figure 11.34 (left). Adding bands with small Duplex Distance (DD) and small DG for which large BW deployment is planned, now places the UE receiver directly inside the bandwidth of the transmitter chain adjacent channel leakage shoulders, as shown in Figure 11.34 (right). The test requirement therefore includes several relaxations, which can be found in section 11.8.2.

The reference sensitivity level calculation is given by Equation 11.14. The differences with the eNodeB equation (11.6) reside in the UE dimensioning assumptions: the NF is set to 9 dB, the Implementation Margin (IM) is equal to 2.5 dB, and a 3 dB correction factor is applied to account for the dual antenna reception gain. In addition, Equation 11.14 includes a frequency band specific relaxation factor, D_{FB}. Note that D_{FB} is not official 3GPP terminology and is equal to 1, 2 and 3 dB for bands in which the DD/DG ratio is greater than 1.5, 2 and 4 dB respectively, and 0 otherwise [18].

$$P_{REFSENS}[\text{dBm}] = -174[\text{dBm/Hz}] + 10\log_{10}(N_{RB} \cdot 180k) + NF + SNR + IM - 3[\text{dB}] + D_{FB} \quad (11.14)$$

D_{FB} is a metric which reflects design challenges for front-end components such as the duplexer. For a given technology resonator quality factor (Q factor), the higher the DD/DG ratio:

- The higher the Insertion Loss (IL) of each of the duplexer Band-Pass Filters (BPFs). In RX, every decibel (dB) lost directly adds to the UE NF, therefore justifying a sensitivity relaxation. In TX, every dB lost causes heat dissipation in the duplexer.
- The sharper the TX BPF roll-off requirements to prevent TX noise from polluting the RX band. Due to heat dissipation, and mass production process variations, TX BPF must be designed with a slightly increased 3 dB cut-off frequency 'F_c' to ensure carriers located at

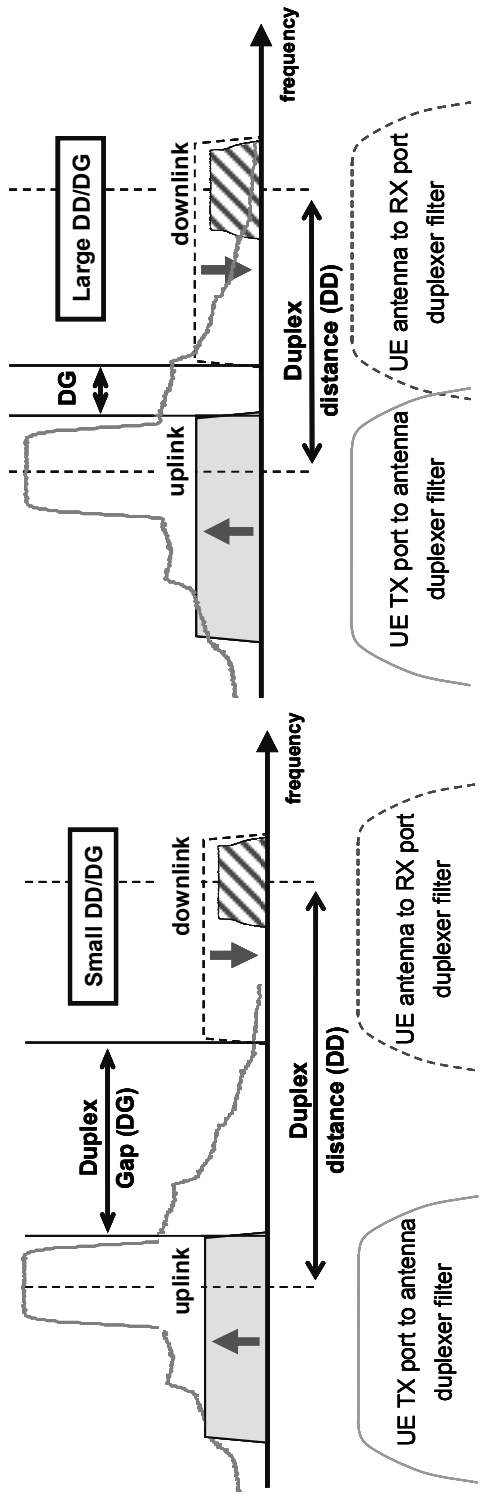


Figure 11.34 Illustration of large duplex gap (left) vs small duplex gap (right) frequency bands

the edge of the band do not suffer from insertion losses that are too high. In these bands the RX is more vulnerable to TX noise leakage, thereby justifying an extra relaxation.

With these assumptions listed, with the exception of the grey shaded values for which UE self-desense applies (cf. section 11.8.2), most of the FDD mode sensitivity levels listed in Table 11.19 can be computed.

For example, Figure 11.35 shows that for a 5 MHz channel bandwidth ($N_{RB}=25$), the reference sensitivity level is equal to −100 dBm. The simulated SNR for 95% of the maximum throughput is assumed to be equal to −1.0 dB. Recent link level simulations [19], [20], indicate slight variations around this nominal with SNRs of −1.4 and −0.9 dB respectively (FRC A1–3). The throughput is approximately 2.1 Mbps for 5 MHz channel in a sensitivity test.

Note that the 9 dB NF assumption used in LTE is very similar to the assumptions made for WCDMA UEs.

Depending on the balance of the link budget, the UE NF is a relevant parameter when planning cell coverage area. The fact that LTE NF requirements are similar to WCDMA commercial

Table 11.19 UE reference sensitivity levels applied to each antenna port for QPSK modulation. Values denoted with * in grey shaded cells are combinations for which a relaxation in N_{RB} and maximum output power is allowed to prevent UE self-desense.

	Channel bandwidth								
Band	1.4 MHz (dBm)	3 MHz (dBm)	5 MHz (dBm)	10 MHz (dBm)	15 MHz (dBm)	20 MHz (dBm)	Duplex Mode	DD/DG	D_{FB} (dB)
1	–	–	–100	–97	–95.2	–94	FDD	1.46	0
2	–104.2	–100.2	–98	–95	–93.2*	–92*	FDD	4	2
3	–103.2	–99.2	–97	–94	–92.2*	–91*	FDD	4.75	3
4	–106.2	–102.2	–100	–97	–95.2	–94	FDD	1.13	0
5	–104.2	–100.2	–98	–95*			FDD	2.25	2
6	–	–	–100	–97*			FDD	1.29	0
7	–	–	–98	–95	–93.2*	–92*	FDD	2.4	2
8	–103.2	–99.2	–97	–94*			FDD	4.5	3
9	–	–	–99	–96	–94*	–93*	FDD	1.58	1
10	–	–	–100	–97	–95.2	–94	FDD	1.18	0
11	–	–	–98	–95*	–93.2*	–92*	FDD	2.09	2
12	–103.2	–99.2	–97	–94*			FDD	2.5	3*
13	–103.2	–99.2	–97	–94*			FDD	1.48	3*
14									
…									
17	–104.2	–100.2	–98	–95*			FDD	1.67	1*
…									
33	–	–	–100	–97	–95.2	–94	TDD		
34	–	–	–100	–97	–95.2	–94	TDD		
35	–106.2	–102.2	–100	–97	–95.2	–94	TDD		
36	–106.2	–102.2	–100	–97	–95.2	–94	TDD		
37	–	–	–100	–97	–95.2	–94	TDD		
38	–	–	–100	–97	–95.2	–94	TDD		
39	–	–	–100	–97	–95.2	–94	TDD		
40	–	–	–100	–97	–95.2	–94	TDD		

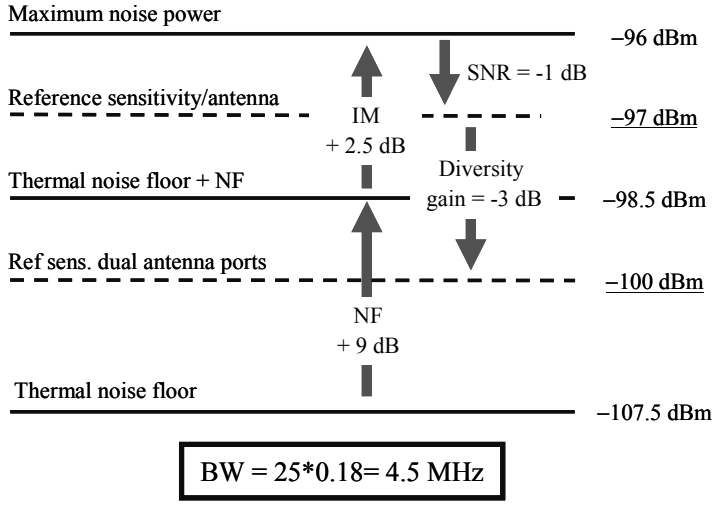


Figure 11.35 Reference sensitivity budget for LTE 5 MHz QPSK ($N_{RB} = 25$)

devices eases early delivery of UEs. In this way, the LTE standard provides a NF requirement small enough to guarantee good cell coverage, but not too small to allow system solution designers to deliver devices with better performance than the minimum requirement. This last point is important since sensitivity is most often a parameter used as a key selling argument. In this respect, LTE commercial devices from different vendors will deliver different sensitivity levels just like their WCDMA and GSM predecessors. An example of variability in WCDMA commercial reference sensitivity is shown in Figure 11.36 [21].

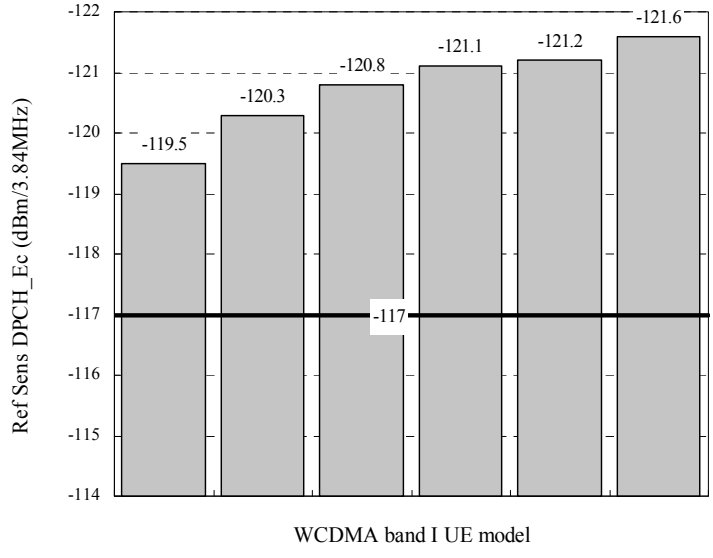


Figure 11.36 Example of class 3 WCDMA band I UE reference sensitivity performance

11.8.2 Introduction to UE Self-desensitization Contributors in FDD UEs

Many of the following items are common to all FDD systems, and in particular, the reader is encouraged to refer to [11] for a detailed discussion on the impact of both TX noise and TX carrier leakage in Direct Conversion Receivers (DCR). This section focuses on the key differences in LTE.

In Figure 11.37, it can be seen that the most sensitive victim is the main receiver LNA, since it is only protected by the duplexer TX to RX port isolation. The diversity LNA benefits from the TX to antenna port duplexer rejection plus the extra antenna to antenna isolation. During conformance tests, this latter is fairly high since the test is performed using coaxial shielded cables, therefore the only coupling mechanisms are those due to PCB cross-talks. Refer to section 11.6.2.2 for details on radiated antenna coupling mechanisms. Both victims operate in the presence of at least two aggressor noise sources: their own PA noise emissions, and the wide-band common mode noise of DigRF^{SM} v4 lines (cf. section 11.6.3).

11.8.2.1 Transmitter Noise Falling into the Receiver Band

Assumptions made in section 11.6.3 are adopted here to illustrate UE self-desense. The victim is a cellular band UE with a 3 dB intrinsic NF and a maximum desense of 0.5 dB is allowed. From section 11.6.3, the maximum aggressor noise PSD must be below −180 dBm/Hz.

11.8.2.2 Large Duplex Distance Frequency Bands

The situation is identical to that experienced in existing WCDMA Band I handsets. Assuming a worst case duplexer isolation in RX of 43 dB, the maximum noise PSD falling in the RX band measured at the PA output port must be less than −180 dBm/Hz + 43 dB = −137 dBm/Hz. Most PAs tested with an ideal signal generator, i.e. a generator which provides a noise floor close to thermal noise at the duplex distance, just barely manage to meet this level [22]. This is one of the reasons why it remains a challenge to design RF IC modulators which can deliver such low noise floors to the PA. The simplest solution is to use an inter-stage BPF, but with the ever-increasing number of bands to be supported, this solution has become unacceptable because of the associated increase of the Bill of Material (BOM). Designing filterless TX RF solutions is a subtle tradeoff exercise between the amount of desense tolerated, the RF modulator current consumption, the BOM cost, and delivering to the customer a competitive reference sensitivity level.

11.8.2.3 Large Transmission Bandwidth in Small Duplex Distance Frequency Bands

Although avoiding receiver desense is not trivial in large DG bands, solutions do exist and this leaves opportunities for innovation. For the small DG, the situation is unfortunately more critical because the aggressor is no longer out-of-band PA noise floor, but the PA ACLR shoulders, as shown in Figure 11.38(a). Therefore, a 3GPP relaxation is the only way to ensure adequate system operation.

An example of ACLR measurements performed using a commercial WCDMA band I PA with a 10 MHz (N_{RB} = 50) QPSK modulated carrier is illustrated in Figure 11.38(b). UMTS 700 MHz front-end IL are emulated with the insertion of a 3 dB resistive attenuator at the PA output port[10].

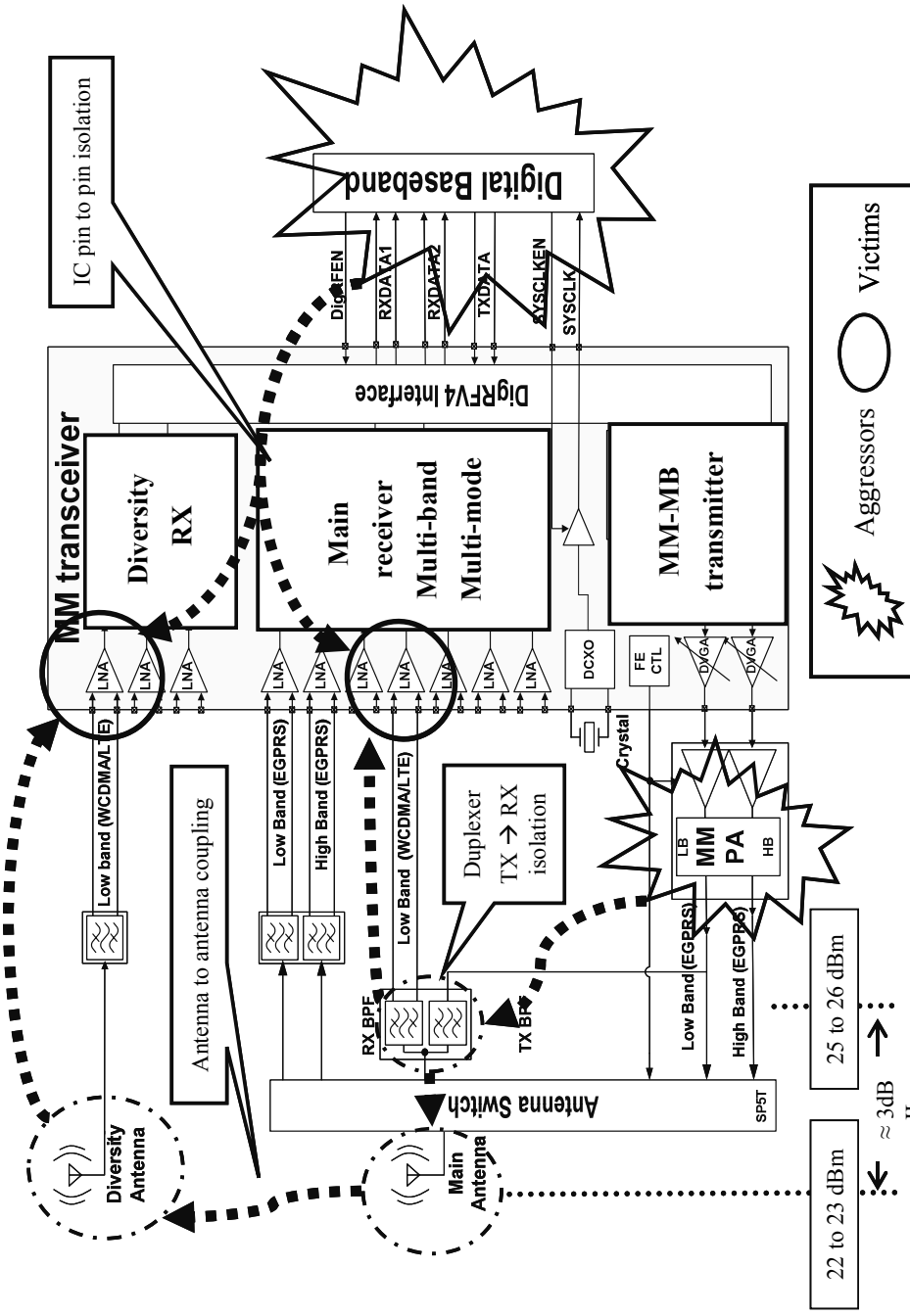


Figure 11.37 Example of aggressors and victims in an optimized quad band 2G – mono band 3G/LTE RF sub-system. Coupling mechanisms are shown in dashed lines. DCXO, digital crystal oscillator

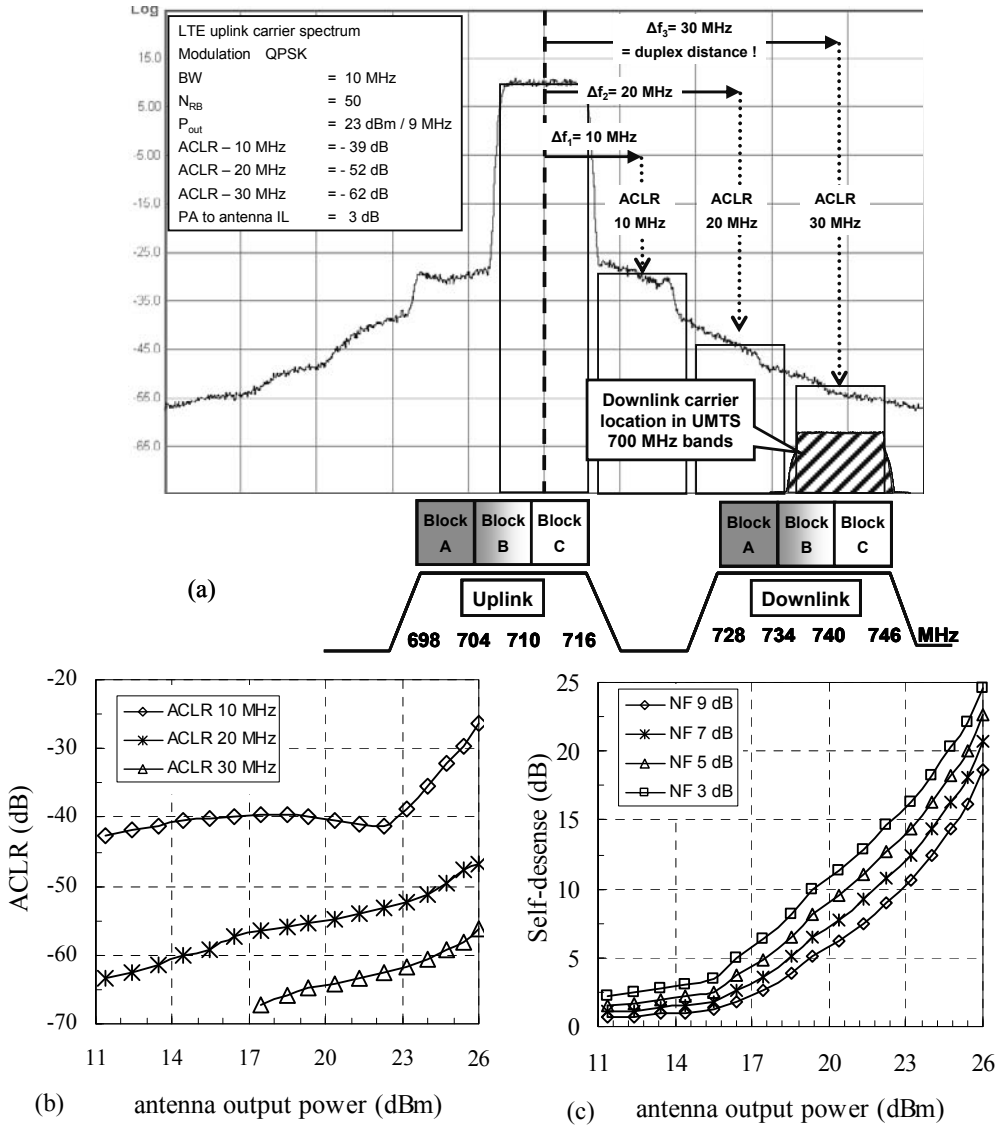


Figure 11.38 Example of LTE 10 MHz QPSK uplink ACLR overlapping the receiver band in the lower UMTS 700 MHz band. (a) ACLR spectral plot; (b) ACLR measured at 10, 20 and 30 MHz offset against antenna output power; (c) self desensitization vs. LNA input referred NF[1]

As can be seen from Figure 11.38(c), at 23 dBm output power, the desensitization reaches in a band 12 example, 16 and 10 dB for 3 and 9 dB NF respectively. To solve this issue, two mitigation techniques have been proposed in RAN4:

[1] Measurements in Figure 11.38 assume a Band XII application: minimum duplexer TX – RX isolation of 43dB, maximum PA to antenna port insertion loss of 3 dB (preliminary measurements on prototypes indicate 2.5 dB maximum in duplexer, 0.5 dB in antenna switch), 23dBm output power at antenna port.

- Maximum Sensitivity Degradation (MSD) [23]: 'the victim's relaxation' technique, which consists of relaxing the reference sensitivity level by an amount similar to those of Figure 11.38(c). The proposal maintains the UE at maximum output power (P_{outmax}) to pass conformance test.
- Point B approach: 'the aggressor relaxation' technique [24], in which the reference sensitivity level is kept intact. This technique maintains the UE at P_{outmax} for a number Resource Block (RB) limited by a point called 'B'. Then, for N_{RB} > point 'B', a progressive back-off of the UE output power is allowed to prevent UE self-desense. Thus, point 'B' corresponds to the maximum number of RBs at which P_{outmax} can be maintained, while point A corresponds to an output power back-off 'X' at which the maximum number of RBs can be supported, as shown in Figure 11.39.

At the time of writing, the MSD approach is adopted. Initial MSD values are proposed for certain bands [26]. Finally, it is worth noting that Half Duplex (HD)-FDD operation has been accepted in RAN 1 [27]. HD-FDD is an alternative solution to the self-interference problem since the transmitter and the receiver do not operate simultaneously. In HD-FDD, the duplexer is no longer needed. This mode of operation can significantly simplify the UE RF front-end architecture as shown in [28].

11.8.2.4 Impact of Transmitter Carrier Leakage Presence at the Receiver LNA Input

In DCRs, differential structure imbalances and self-mixing are known as some of the mechanisms generating second order intermodulation distortion (IMD_2) products [29]. Self-mixing is due to finite isolation between the RF and the LO port of the down-conversion mixer. Under these conditions, the mixer behavior may be approximated as that of a squarer and therefore generates IMD_2 products. The squaring of CW blockers is a simple DC term which can be easily rejected using High Pass Filters (HPF). Squaring an AM modulated blocker, however, generates a wideband noise like an IMD_2 product, which can degrade the wanted signal SNR. In a RX filter-less architecture, the mobile's own TX leakage places the most stringent requirements on the mixer IIp2, which must receive a weak input signal (≈ −85 dBm), in the presence of a

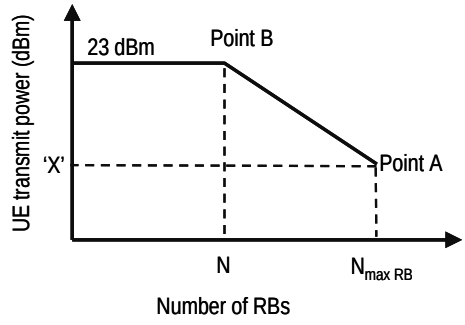


Figure 11.39 Point B approach to prevent UE self-desense

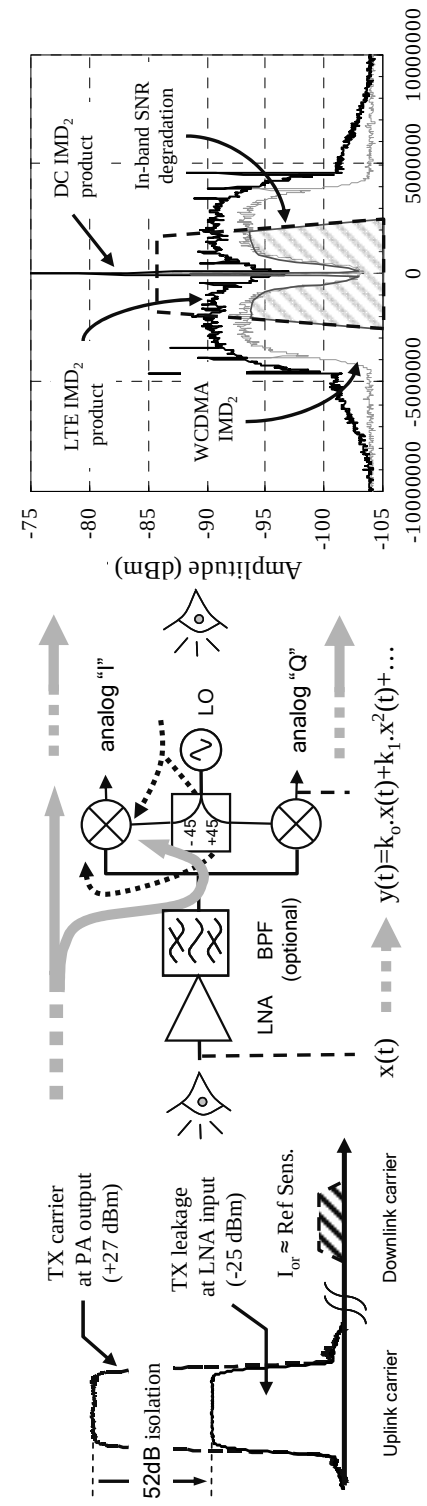


Figure 11.40 Self mixing in direct conversion receivers: Left: TX leakage at LNA input; right: I/Q spectrum observed at mixer output; dashed lines represent the wanted channel

TX leakage mean power of approximately[2] −10.5 dBm at mixer input. The simplest solution consists of using an inter-stage BPF, but this is not the preferred option for the reasons explained in section 11.6.2. A comparison between WCDMA and LTE QPSK uplink modulated carrier IMD_2 products is shown in Figure 11.40.

In a WCDMA RX filter-less application, the mixer IIp2 requirement to ensure a small SNR degradation is in the range of 70 dBm [11], a figure that is extremely challenging to achieve. From Figure 11.40, it can be seen that the higher LTE IMD_2 PSD sets slightly higher mixer IIp2 requirements than for WCDMA receivers.

11.8.3 ACS, Narrowband Blockers and ADC Design Challenges

Both ACS and Narrow-band (NB) blocking requirements are a measure of the receiver ability to receive a wanted signal at its assigned channel frequency in presence of an interfering Adjacent Channel Interferer (ACI), at which a throughput requirement will be fulfilled. The throughput will equal or exceed 95% of the maximum throughput, for a specified reference measurement channel.

The intention of this requirement is to verify the ACI Rejection (ACIR). Both tests are important to avoid UE dropped calls in cells where eNodeBs from adjacent operators are non co-located. In a similar fashion to the WCDMA specifications [43], the LTE requirements are based on a 33 dB ACS budget, which has been derived through extensive coexistence simulations [4]. In order to prevent stringent selectivity requirements for the 15 and 20 MHz BW, a 3 dB and 6 dB ACS relaxation is proposed respectively. The resulting test cases are summarized in Table 11.20[3].

Table 11.20 Relationship between interfering and wanted signals for ACS requirements

Rx Parameter	Units	Channel bandwidth					
		1.4 MHz	3 MHz	5 MHz	10 MHz	15 MHz	20 MHz
ACS test case I							
Wanted signal mean power	dBm	Refsens + 14 dB					
$P_{Interferer}$	dBm	Refsens + 45.5 dB				Refsens + 42.5 dB	Refsens + 39.5 dB
$BW_{Interferer}$	MHz	1.4	3	5	5	5	5
$F_{Interferer}$ (offset)	MHz	1.4	3	5	7.5	10	12.5
Assumed ACS	dB	33	33	33	33	30	27
ACS test case II							
Wanted signal mean power	dBm	–56.5	–56.5	–56.5	–56.5	–53.5	–50.5
$P_{Interferer}$	dBm	–25					

[2] Transmit leakage mean input power at mixer input ≈ PA output power (+27dBm) – isolator/coupler losses (0.5dB) – duplexer isolation (52dB) + LNA gain (15 dB) = −10.5 dBm.

[3] For ACS test case I, the transmitter will be set to 4 dB below the supported maximum output power. For ACS test case II, the transmitter will be set to 24 dB below the supported maximum output power. At the time of publication, there are ongoing discussions to add a small frequency offset to the $F_{Interferer}$ (offset) value to prevent the interfering signal from being orthogonal to the wanted signal [44].

- The ACS test case I is performed with a mean signal power set to 14 dB above the reference sensitivity level, and a variable interferer power for each wanted channel bandwidth as shown in Figure 11.41.
- The ACS test case II is the test which stresses the UE dynamic range by setting the interferer power constant to −25 dBm and variable interferer power so that the 33 dB ACS test condition is met. For example, in a 5 MHz BW (N_{RB}=25), the Carrier to Interferer power Ratio (CIR) is also equal to −56.5 dBm – (−25 dBm) = −31.5 dB.

One of the UE DCR blocks affected by the ACS II is the LNA, for which the gain must be sufficiently low to prevent overloading of the I/Q mixer inputs, and thereby relaxing the mixer linearity requirements, but also sufficiently high to prevent the UE NF from failing the minimum SNR requirements imposed by the highest MCS test requirements. Additionally, the presence of such a strong interferer, for which the PAPR in LTE is higher than that for WCDMA, sets a linearity requirement on the LNA. LNA non-linearities may generate Adjacent Channel Leakage (ACL), which would overlap the wanted signal and thereby directly degrade the wanted carrier SNR. Figure 11.42 illustrates this challenge.

The NB blocking test is illustrated in Figure 11.43 for 5 MHz BW (N_{RB}=25). This test ensures LTE UEs can be deployed in regions where other telecommunication standards, such as GSM/EDGE, are in service. To minimize the cellular capacity loss, it is important to minimize the guard bands. For this reason, the NB blocker test places the UE wanted channel in the presence of a blocker located at small frequency offsets. This test only differs slightly from its WCDMA equivalent: in LTE, the blocker is CW (vs GMSK in 3G), the frequency offset[4]

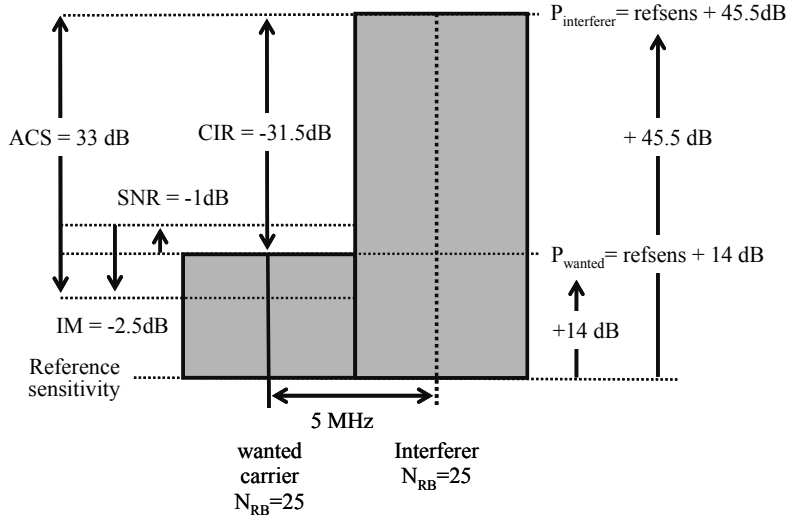


Figure 11.41 Adjacent channel selectivity test case I requirements for 5 MHz LTE

[4] In LTE, the introduction of a small frequency offset has been proposed to ensure that the interferer does not fall in the spectral nulls of the receiver's FFT operation. The offset is an odd multiple of half the tone spacing (2k+1).7.5 kHz, refer to Figure 11.42.

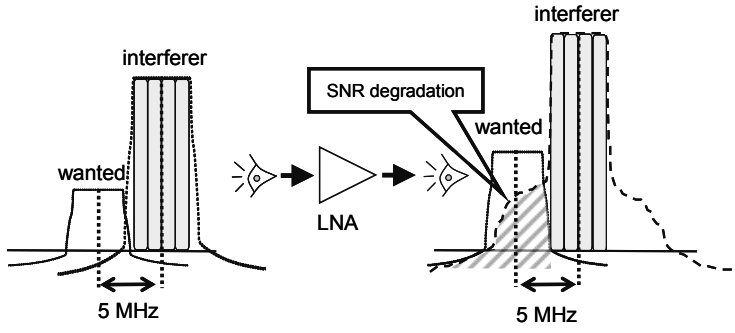


Figure 11.42 Example of LNA non-linearities in the presence of the -25 dBm ACS II interferer. Left: spectrum at LNA input; right: spectrum at the LNA output in the presence of LNA ACL

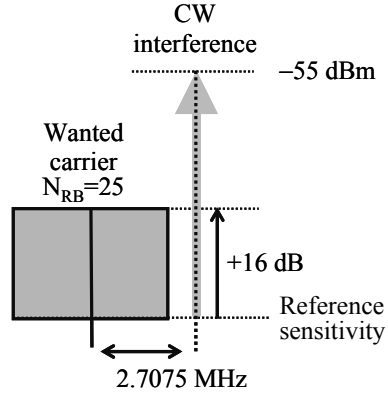


Figure 11.43 Narrowband blocking measurement for 5 MHz LTE

is set to 2.7075 MHz (vs 2.7 MHz in 3G) and the wanted channel benefits from 16 dB desense in 5 MHz BW (vs 10 dB in 3G).

It can be seen that in 5 MHz BW operation, LTE filtering requirements are similar to the existing WCDMA devices. In flexible bandwidth systems the filter design strategy is a compromise articulated around two extreme scenarios: at one end, a receiver with infinite ACIR minimizes the ADC resolution and power consumption at the expense of a sharp analog filter, which may distort the wanted signal. At the other extreme, a receiver which provides no or little ACIR imposes an ADC with a *Dynamic Range (DR)* large enough to cover the worst case 3GPP DR requirements. The following discussion highlights this tradeoff by firstly introducing the impact of bandwidth flexibility on *Analog Channel Filters (ACF)*, and secondly on the ADC.

11.8.3.1 Impact of Flexible Bandwidth Requirements on the Analog Channel Filter Design Strategy

OFDM systems overcome Inter-Symbol and Inter-Carrier Interference (ISI and ICI) due to propagation through time-dispersive channels by introducing a Cyclic-Prefix (CP). The

CP acts in a similar fashion to time guard bands between successive OFDM symbols. Therefore, the longer the CP, the better the resilience to large delay spreads at the expense of an energy loss. The CP length must also be selected so as to avoid the signal smearing caused by the Group Delay Distortion (GDD) of analog filters [30]. Yet, selecting a filter family for the worst case delay spread foreseen in the standard (such as the ETU model, for example) is perhaps not the best strategy. For example, [31] suggests that in most cases the delay spread experienced by the UE is less than the CP length, and therefore the estimation of delay spread by the BB channel estimator can be used to dynamically re-program the transfer function of a post ADC digital FIR filter. This elegant solution provides enhanced ACS performance for a given pre-ADC ACF. So, what is the limit to scaling the 3 dB cut-off frequency 'F_c' of the ACF?

Figure 11.44 illustrates the impact of scaling a baseline[5] filter's F_c optimized for WCDMA ACS and NB blocking onto the experienced GDD. The filter's F_c is either stretched or shrunk by a factor proportional to the LTE BW ratios.

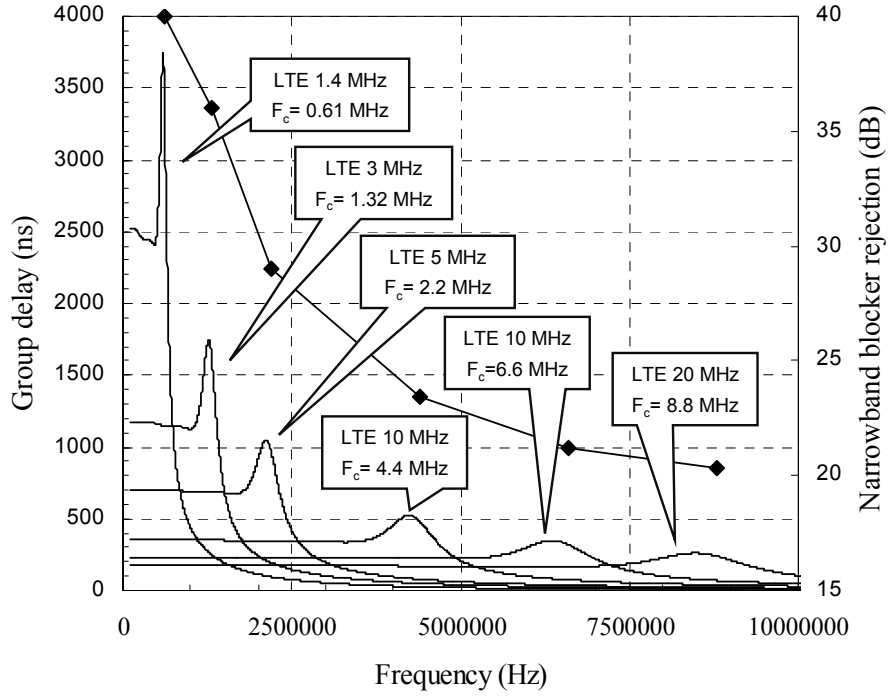


Figure 11.44 Impact of scaling 'Fc' of a 5 MHz baseline ACF optimized for WCDMA proportional to the LTE's operating BW. Left y-axis: filter's group delay; right y-axis: NB blocker filter's rejection performance (diamonds)

[5]The filter used is similar to that in [31]: it has 4 zeros, 8 poles, with a 3dB cut-off frequency "F_c" of 2.2 MHz in 5 MHz BW, and a notch located at 2.8 MHz offset. To reduce the group delay overshoot associated with the presence of a notch so close to the edge of the wanted signal, some of the poles and zeros are used to implement an analog group delay equalizer.

From Figure 11.44, it can be seen that the lower the filter's 'F_c', the higher the filter's delay, GDD, and NB blocker ACIR, so much so that in the 1.4 MHz BW of operation the GDD is slightly greater than 1 µs. This latter case 'consumes' a large amount of the normal CP length of 4.7 µs and could impact ISI in large delay spread channels. From this example, it can be concluded that proportional scaling of the ACF's F_c is probably not the best tradeoff. One alternative takes advantage of the 3GPP relaxations to tailor the ACF ACIR to just meet the ADC DR requirements at 15 and 20 MHz operation. The ADC DR enhancements in small operating BW can then be used to relax the filter's sharpness so as to take full benefit of the CP length.

11.8.3.2 Impact of Flexible Bandwidth Requirements on the ADC DR

For the sake of simplicity[6], the minimum required ADC resolution is estimated by assessing the ADC EVM (EVM_{ADC}) based upon the pseudo-quantization noise (PQN) model. Figure 11.45 (left) shows the optimum Signal to Distortion Quantization Noise Ratio (SDQNR) at the ADC output is met for ADC clipping ratios, also denoted 'ADC back-off' (ADC BO), ranging from 10 to 14 dB. The resulting EVM_{ADC} is plotted in Figure 11.45 (right).

The UE EVM budget is estimated from the required SNR to achieve the highest SIMO LTE throughput which corresponds to the 64QAM 8/9 MCS. From section 11.3.4, 5% throughput loss is met if the total composite EVM is less than 6.3%. Assuming that each EVM impairment

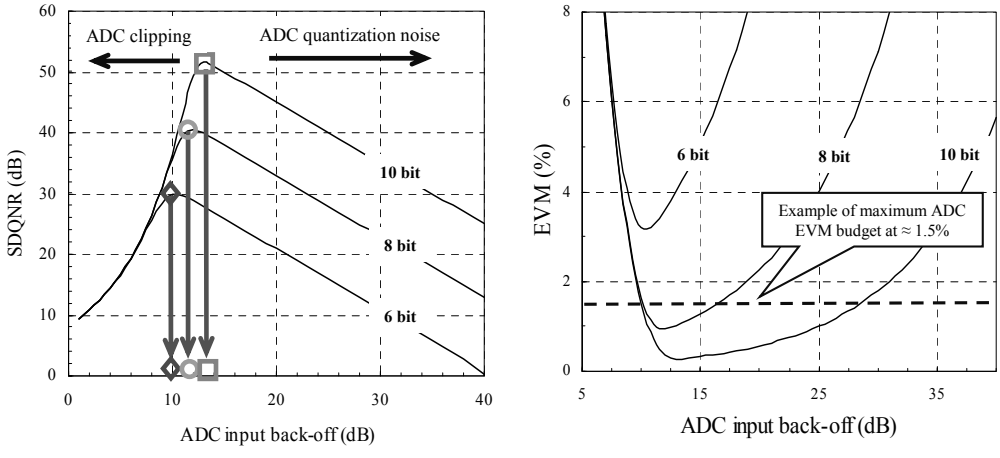


Figure 11.45 Left: SDQNR (in dB) at ADC output against ADC resolution and ADC input back-off (in dB) or ADC clipping ratio. The optimal ADC BO is respectively highlighted with a diamond, circle and square for 6, 8 and 10 bits ADC resolution. Right: corresponding ADC output EVM performance

[6] The PQN model considers the quantization error as an additive noise, which has a white and a uniformly distributed spectrum and which is uncorrelated with the input signal. The reality is more subtle: the recent work published in [32] shows that the PQN approach is not entirely sufficient to model the quantization errors due to the much larger peak to average power ratio of OFDM signals.

is AWGN like, taking an example where the eNodeB EVM is equal to 4.5%, and the UE RF RX EVM performance is equal to 4%, this leaves an EVM_{ADC} budget of 1.5%[7].

Let us first assume an ideal RF–BB chain, which provides an infinite ACIR, and an ideal Analog Gain Control (AGC) loop, so that the optimal BO is always exactly met. The situation is captured in Figure 11.46 (left) and indicates that the lowest acceptable ADC resolution is 8 bit. In a real AGC loop system, the ADC BO over the duration of a call is no longer that of a discrete point, but a spread of BO statistically distributed as shown by the histogram in Figure 11.46. Taking one particular example of AGC loop inaccuracies[8] into account, it can be seen that 10 bit is the minimum ADC resolution, which provides approximately 12 dB headroom (Δ_{RF}) for RF IC imperfections, such as imperfect DC offset cancellation and ACIR (Figure 11.46 right). This requirement is equivalent to a CW DR of 60 dB. One of the differences in LTE is that the UE AGC loop must also cope with time-varying amplitude of in-band signals due to dynamically scheduled users with a varying number of active RBs transmitted at different power levels in the downlink. With 10-bit resolution, Δ_{RF} is sufficiently large to

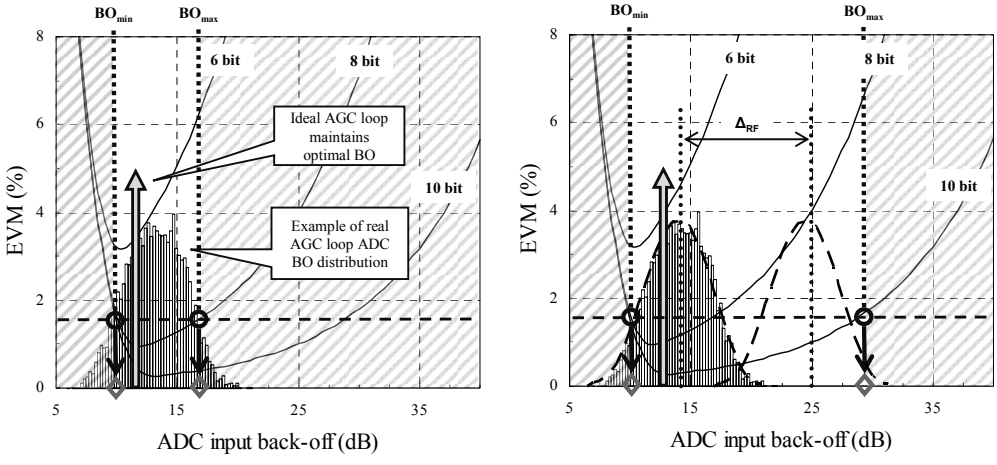


Figure 11.46 Impact of AGC loop imperfections. Left: 8 bit ADC; right: 10 bit ADC. Dashed areas indicate regions which do not fulfill the MCS EVM ADC requirements. The dashed histogram envelope illustrates the ADC BO margin due to RF imperfections (Δ_{RF}) present at ADC input

[7]Assuming AWGN like behavior of each EVM contributor, $EVM_{ADC} = \sqrt{6.3\%^2 - (EVM_{eNobeB}^2 + EVM_{RFRX}^2)}$. Note that in conformance tests, the eNodeB emulator EVM is extremely low (typically 1% or less). Also note that the recent WiMax (802.16e) RF transceiver design in [33] can achieve ≈1.5% RX EVM, thereby relaxing the overall EVM downlink budget.

[8]The histogram shows a recorded ADC BO distribution captured over a 10 minute long WCDMA BLER measurement performed in a fading test case-1, 3 km/h. The AGC loop updates the analog gain at 10 ms intervals. An AGC loop for LTE is likely to operate at a faster update rate to provide better control accuracy of the ADC BO. The resulting histogram would present a smaller spread of BO, thereby relaxing the margins for RF imperfections.

accommodate one example of RF imperfection margins[9]. Note that each of the assumptions[9] is listed as an example of minimum requirements since these are implementation specific.

In conclusion, the most challenging mode of operation for a multiple-standard ADC is the LTE 20 MHz operation, for which a minimum of 60 dB DR must be delivered. Sigma delta ADCs represent an attractive solution to these requirements [34]. These converters shape the Quantization Noise (QN) floor via a high pass transfer function, thereby digging a notch into the QN PSD within the BW of the wanted carrier, and rejecting the noise out-of-band. The BW of the QN notch can be optimized for each BW of operation by either changing the sampling frequency and/or by reconfiguring the noise shaping filter's transfer function. An example of sigma-delta ADC flexibility is shown in Figure 11.47. It can be seen that by ensuring the DR requirements are met for LTE 20 MHz, the ADC offers a 20 dB improvement for operation at LTE 1.4 MHz. Every decibel gained in the DR can be used as an extra relaxation of the ACF filter design. In particular, the large DR performance at low operating BW relaxes significantly the ACF rejection in GSM mode for which the design of a sharp filter can be expensive in terms of die area.

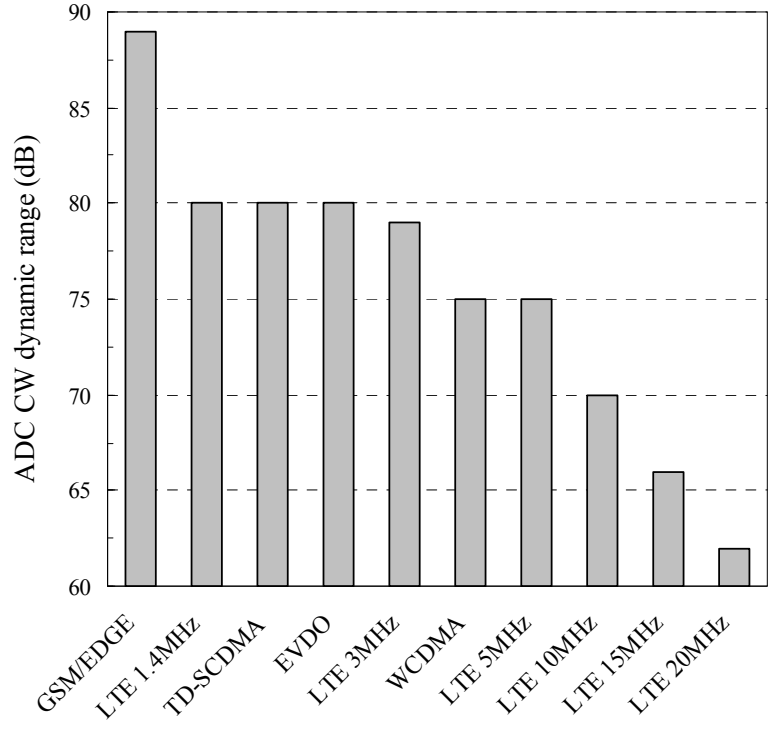


Figure 11.47 Achievable DR with sigma delta ADCs over a wide range of system's BW [35]

9 Following margins are assumed: 1) 2 dB margin for imperfect RF receiver DC offset removal at ADC input, 2) 4 dB to account for relative power difference between a 64QAM user and a QPSK user in the cell [2] - this is a minimum requirement, 3) Imperfect RF RX ACIR forces the AGC loop to lower the effective wanted signal target total ADC BO. With the previous assumptions 1 & 2, the DCR is allowed to present at the ADC input a CIR of ≈4 dB (eNodeB TPC DR) + 2 dB (leaked DC offset) – 12 dB (Δ_{RF}) = –6 dB.

11.8.4 EVM Contributors: A Comparison Between LTE and WCDMA Receivers

In section 11.3.4, the downlink LTE budget is set to approximately 8%. Compared to WCDMA, where 10–12% is acceptable even for HSDPA operation, this requirement appears tougher. However, the novelty in LTE is that EVM measurements make use of a zero-forcing equalizer (cf. section 11.3.4). Thus, a distinction must be made between AWGN like contributors, and contributors that can be equalized, and therefore become transparent to the UE receiver. This is an important difference with WCDMA where, for large enough SF, each EVM contributor behaves like AWGN [36]. This section aims at illustrating these differences with only a few selected impairments: I/Q gain and phase imbalance, distortions due to ACF, and DCR local oscillator phase noise.

11.8.4.1 Impact of Finite Image Rejection due to I/Q Amplitude and Phase Mismatches

In real world analog/RF designs, it is nearly impossible to design DCRs with perfectly identical gain and phase response between I and Q branches. Therefore DCRs come with two natural impairments: amplitude and phase mismatches, denoted ΔA and $\Delta\Phi$[10] respectively, leading to a finite Image Rejection (IR). Finite IR results in each sub-carrier being overlapped with the sub-carrier which is located at its frequency image mirror, as shown in Figure 11.48[11]. The power ratio between a sub-carrier and its image is by definition the IR. Assuming the symbols carried by each sub-carrier are uncorrelated, the impact of IR on LTE is no different to that of a single carrier system and can be considered an AWGN source [37].

Figure 11.49 shows EVM measurements performed with an Agilent™ 89600 Vector Signal Analyzer, which delivers the Error Vector Spectrum (EVS). EVS is a tool which plots the magnitude of the error vector against each sub-carrier. Figure 11.49(b) shows the composite LTE EVS, i.e. it is an overlay of each physical channel's EVS, where the darker line shows the

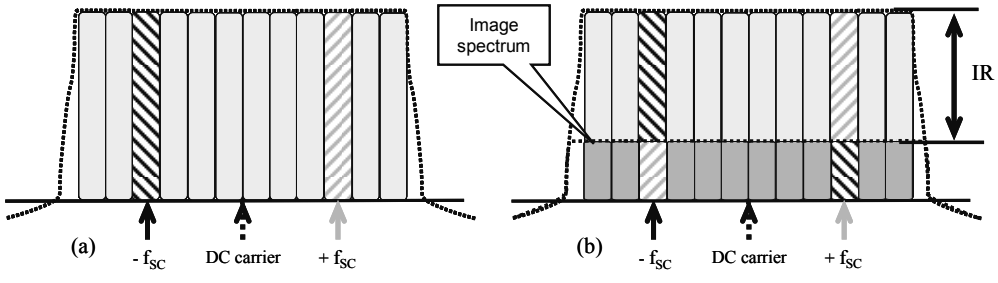


Figure 11.48 (a) Transmitted OFDM carrier; (b) equivalent i+jQ baseband complex spectrum at ADC input in presence of a finite image rejection equal across all sub-carriers

[10] $\Delta\Phi$ may originate from either a local oscillator which does not drive each I and Q mixer in exact quadrature, or from a tiny mismatch in the 3 dB cut-off frequency of each I/Q LPF. In the latter case, $\Delta\Phi$ is not identical for all sub-carriers. The net result is a frequency dependent image rejection (IR). Note that IR can be calibrated in modern DCR designs as shown in [33].

[11] In field trials, incoming sub-carriers are notched due to frequency selective fading and image rejection is not constant across the entire BW of the DCR I/Q bandwidth.

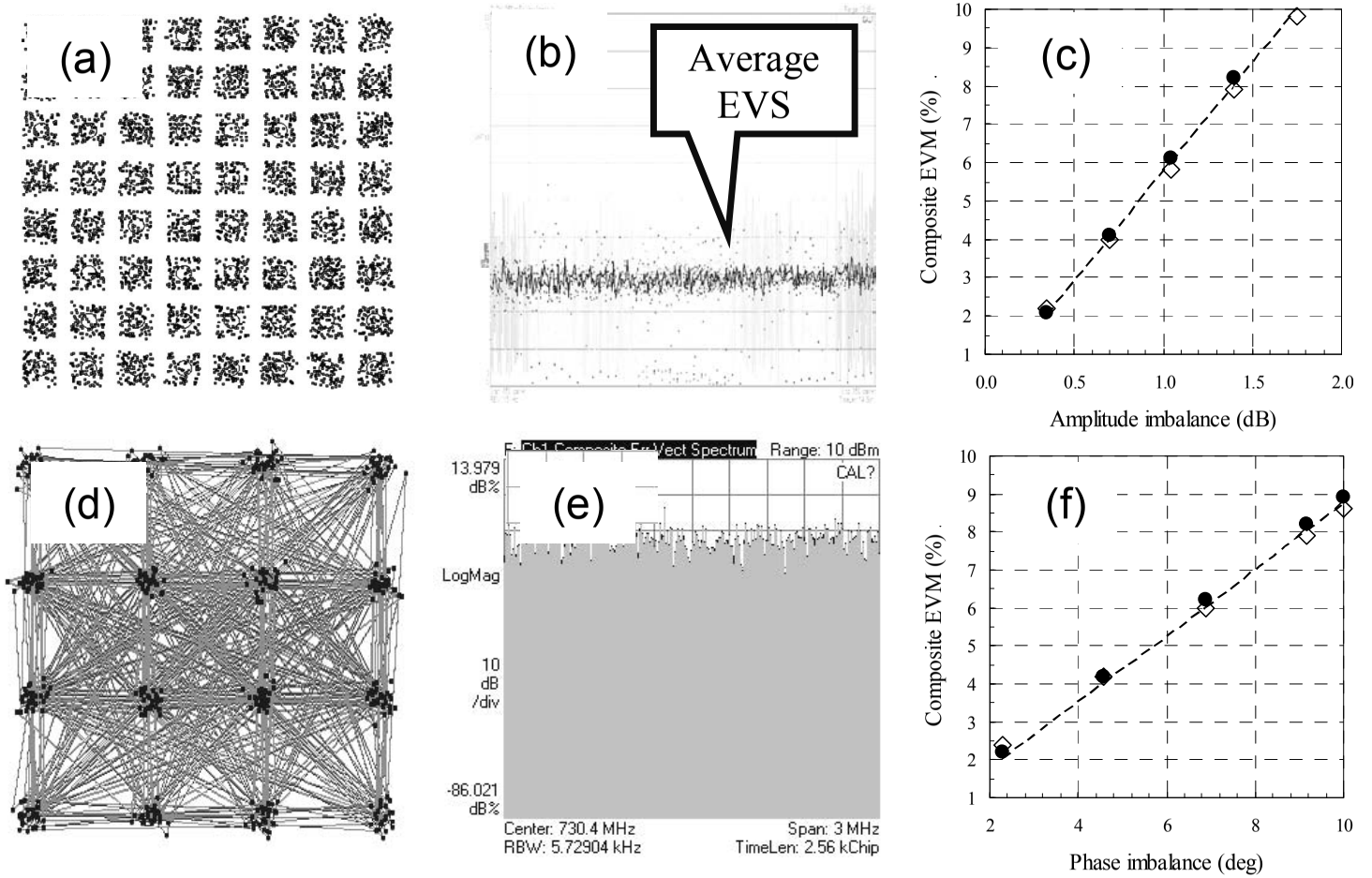


Figure 11.49 Impact of ΔΦ and ΔA on WCDMA and OFDM downlink carriers. Impact of ΔA = 1.75 dB on: (a) LTE 64QAM state spreading; (b) LTE carrier EVS; (d) 16QAM HS-DSCH WCDMA; (e) WCDMA EVM spectrum. (c) and (f): comparison of LTE (filled circles) and WCDMA (empty diamonds) EVM performance against ΔA and ΔΦ

average EVS. The EVS spectral flatness in Figure 11.49 confirms the AWGN-like behavior of EVM due to IR for both standards.

11.8.4.2 Impact of In-band Analog Filter Amplitude and Group Delay Distortions

Zero-IF Low Pass Filter Contribution

Figure 11.50 shows the measured[12] impact of a prototype I/Q channel filter[13] similar to that presented in [31] on a 5 MHz 16QAM LTE and a WCDMA downlink carrier. The equal spreading of each constellation point in Figure 11.50(c) confirms the AWGN-like behavior of EVM_{LPF} for WCDMA, and in this example, results in 8% EVM performance. Figure 11.50(b) shows that without equalization the outermost sub-carriers are severely impacted, while sub-carriers located in the middle are less vulnerable. The use of the ZF-equalizer flattens EVS and brings the composite EVM down to ≈1.2%.

It can be concluded that LTE relaxes the LPF impairment budget compared to WCDMA modulation.

Zero-IF High Pass Filter Contribution

IMD_2 products described in section 11.8.1.2 generate a strong DC term which can lead to saturation through the several decibels of I/Q amplification required to meet the ADC target BO. In WCDMA the DC term can be cancelled with High Pass Filters (HPF), and [36] has shown that its impact on EVM is AWGN-like. HPF design is a compromise between EVM, capacitor size, die area, and DC settling time[14]. A passive 4.5 kHz first order HPF with group delay distortion in excess of the CP length has been deliberately chosen to illustrate the impact on LTE. In contrast to the LPF or BPF test case, the sub-carriers located close to the center of the carrier are the most vulnerable as can be seen in Figure 11.51. For example, the Primary Synchronization Signal (PSS), which occupies the center of the carrier over a 1.4 MHz BW, experiences a 7.5% EVM, while the QPSK user EVM improves as the allocated BW is increased (Figure 11.51(a) and (c)). The distortion impacts so few RS that the ZF-equalizer cannot flatten the impact of the HPF. LTE therefore calls for a careful design of the DC offset compensation scheme.

11.8.4.3 Impact of Phase Locked Loop Phase Noise

If the downlink OFDM signal was just a set of unmodulated closely spaced CW tones, the resulting I or Q output of the DCR mixer would be that of each CW tone multiplied with LO Phase Noise (PN) profile as shown in Figure 11.52(a). Clearly, any PN exceeding the sub-carrier spacing causes SNR degradation. In most PN studies [38, 39], a distinction is made

[12] All EVM measurements performed with an Agilent™ 89600 VSA.

[13] Both I and Q filters are nearly perfectly matched and deliver an IR across the entire receiver bandwidth better than –40 dBc. This guarantees that the EVM performance is dominated by the LPF non-linear phase and amplitude distortions.

[14] In BiCMOS designs, it is difficult to implement analog HPF with cut-off frequencies less than 3 kHz due to cost constraints in die area. The use of RFCMOS allows designing this loop partly in the digital domain with a much smaller cut-off frequency at nearly no impact on die area.

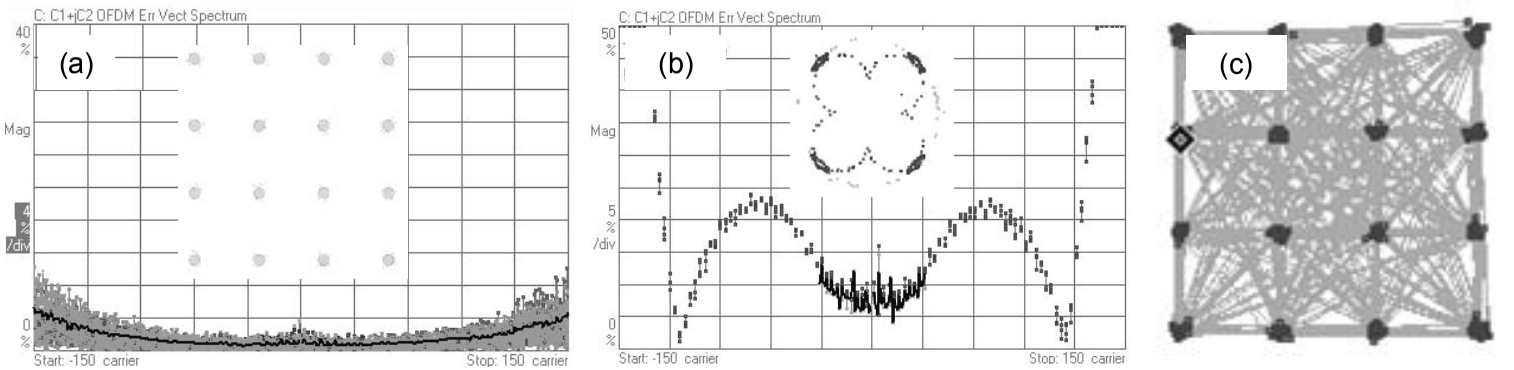


Figure 11.50 Measured impact of an I/Q ACF on the composite EVM of a 5 MHz LTE ($N_{RB} = 25$) and a WCDMA downlink carrier. (a) LTE composite EVS and 16QAM user data constellation with equalizer 'ON'; (b) LTE reference signals EVS and constellation with equalizer 'OFF'; (c) WCDMA HS-DPCH constellation, EVM ≈8%

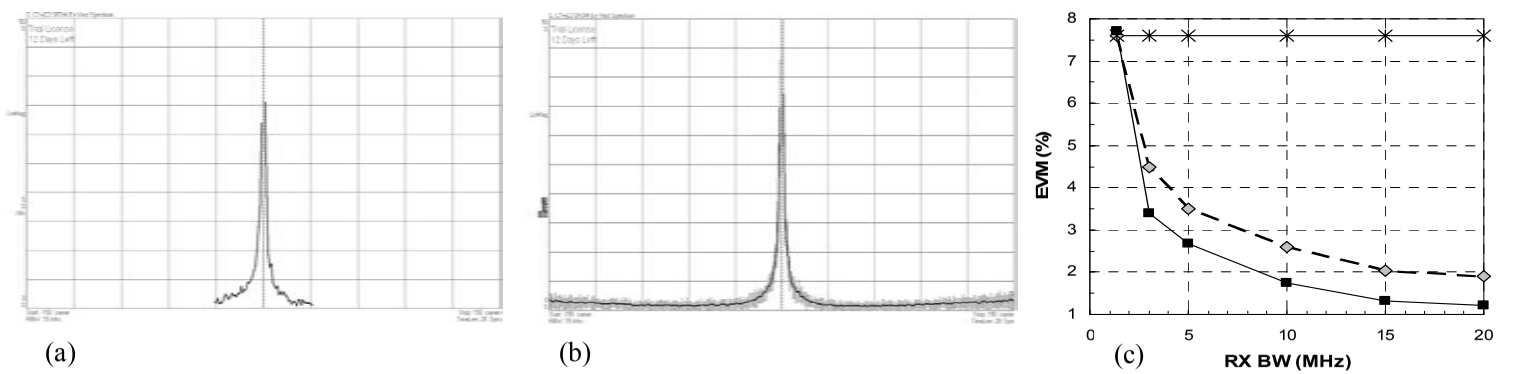


Figure 11.51 Measured[12] impact of a 4.5 kHz HPF prototype with ZF equalizer 'ON'. (a) EVS of PSS in 5 MHz (25 RB); (b) EVS of a 5 MHz (25 RB) QPSK user; (c) EVM vs carrier's BW. Stars = average PSS EVM; diamonds = average composite EVM; squares = QPSK user average EVM

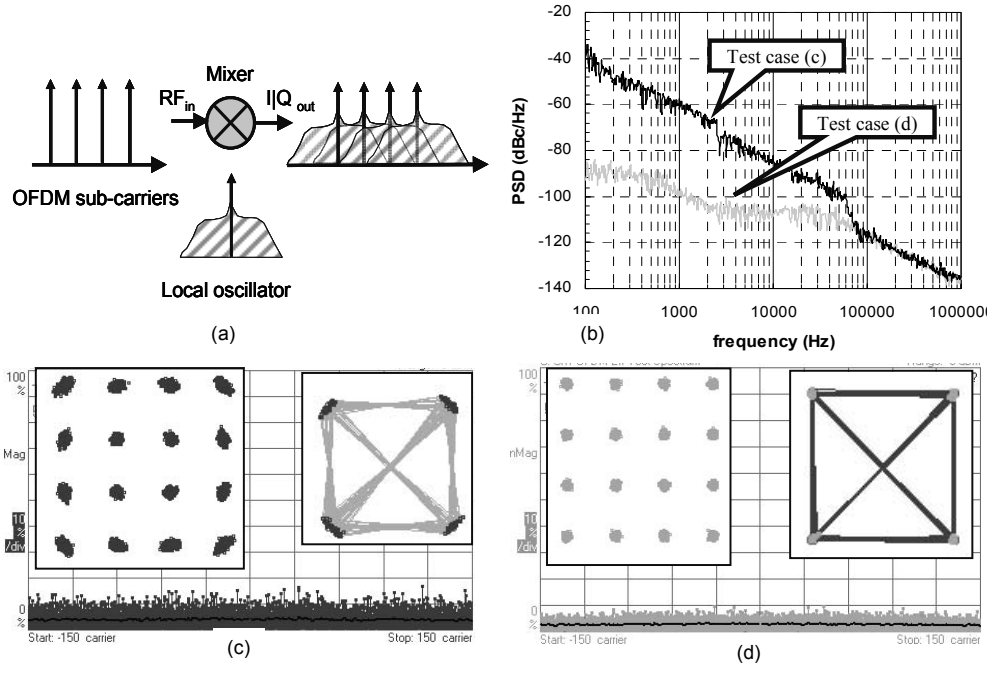


Figure 11.52 (a) Illustration of LO PN multiplication on OFDM sub-carriers; (b) LO PN profiles: degraded PN in black ('test case "c"), 'near ideal' in grey (test case 'd'); (c) and (d) LTE 16QAM and WCDMA QPSK constellations are overlaid on the LTE EVS; (c): easurements with degraded PN; (d) measurements in near ideal PN profile

between PN contributions at high and low frequency offsets from the carrier. The close-in PN produces an identical rotation of all sub-carriers and is also often referred to Common Phase Error (CPE). CPE can be estimated and therefore can be corrected for. The high frequency offset PN components generate ICI. This contribution can be modeled as AWGN and cannot be corrected. An illustration of both CPE and ICI due to a degraded PN profile is presented in Figure 11.52(c) where the LTE 16QAM user data suffer from both constellation rotation and state spreading.

In conclusion, the OFDM sub-carrier spacing used in LTE introduces a new LO phase noise requirement for frequency offsets ≥ 15 kHz offset. This differs from WCDMA where EVM for this impairment only depends on the integrated LO phase error.

11.9 UE Demodulation Performance

11.9.1 Transmission Modes

OFDM modulation in LTE downlink allows optimization of the transmission in time, frequency and antenna domains. The UE 3GPP standard compliancy must be ensured by satisfying the requirements covering a wide range of modes comprising Transmit/Receive diversity and spatial multiplexing. We examine in this section the principles of each mode and the corresponding receiver schemes for each case.

11.9.1.1 Single Input Multiple Output (SIMO) Mode

Plain OFDM Single Input Single Output (SISO) transmission is not supported by LTE as UEs are required to have at least two receiving antennas. The SIMO case constitutes then the simplest scheme and it is a straightforward extension of the SISO case. Let us assume a transmission bandwidth of $B = 20$ MHz and a sampling frequency of $f_s = 30.72$ MHz and normal cyclic prefix mode. At the receiver side, for each receiving antenna, base-band digital signal samples after RF and analog-to-digital conversion are buffered over an OFDM symbol of duration $Ts = 71.3\,\mu s$ (or $71.9\,\mu s$ for the first OFDM symbol of a slot since the cyclic prefix samples are then slightly more). The cyclic prefix samples are discarded, assuming timing offset recovery is safely done, and $N = 2048$ samples are FFT transformed into a frequency domain equivalent signal. Due to LTE standard assumptions, at the output of the FFT, only $K = 1200$ samples are kept and the others are discarded. Depending on the physical channel that needs to be modulated and its specific allocation over the sub-carriers, an operation of demultiplexing and re-ordering of the complex digital samples takes place.

Even for multi-path propagation, providing that the channel coherence time is much larger than the OFDM symbol duration, it is well known that in OFDM the signal at the output of the FFT at each sub-carrier position can be considered as affected by flat fading. In this case the optimal demodulation scheme is the simple Matched Filter and the effect of the channel is undone by multiplication by the complex conjugate of the channel coefficient at any given sub-carrier. The QAM symbols are obtained by combining the derotated symbols on the same sub-carriers across the two receiving antenna paths. This operation is also generally known as Maximum-Ratio-Combining and allows optimal benefit from the additional antenna diversity.

After MRC operation, the QAM symbols are demapped and the sequence of soft-bits for channel decoding operation is buffered until the total expected amount is available after demodulation of several OFDM symbols. The soft-bits are rate-matched for the parameters specific to the physical channel under consideration; eventually an operation of Hybrid-Retransmission-Request (HARQ) combining is performed and the channel decoding operation is invoked. At its output, the sequence of hard-decisions is verified against the transported Cyclic-Redundancy-Check (CRC) bits to decide whether the decoded bits are correct or not. The channel decoder to be used depends on the nature of the decoded physical channel: dedicated channels, e.g. PDSCH, are always turbo-encoded while channels carrying control information, e.g. PDCCH, are convolutionally encoded, and thus decoded by means of a Viterbi decoder.

For channels supporting the HARQ protocol, the result of the redundancy check operation is fed back to the BS. The receiver performance is computed upon the success rate of the CRC in terms of throughput, which is the measure of net successfully decoded bits after the HARQ process.

11.9.1.2 Transmit Diversity Mode

Transmit diversity is implemented in LTE by means of Spatial-Frequency-Block-Coding (SFBC). SFBC is the Alamouti encoding of two QAM symbols lying in neighboring sub-carriers. For transmit-diversity transmissions then, at the receiver side, symbols at the output of the FFT need to be re-ordered in pairs according to the SFBC encoding carried out at the transmitter and the Alamouti scheme is undone via a linear operation.

11.9.1.3 MIMO Mode

MIMO transmission mode is the key enabler of a high data rate of LTE (up to 150 Mbps for 20 MHz bandwidth) and allows the transmission of one or two independent data streams depending on the channel conditions experienced by the UE. In MIMO mode, the channel at each sub-carrier is represented by a matrix whose size is given by the number of transmitting (N_{Tx}) and receiving (N_{Rx}) antennas. If the propagation conditions result from a rich scattering environment, the rank of the channel matrix is full and in these conditions spatial multiplexing of two data streams can be supported. If instead the channel matrix rank is not full, i.e. a rank is equal to one, only one codeword is transmitted. As for HARQ acknowledgements, MIMO closed loop mode requires continuous feedback from the UE to the BS on a sub-frame basis. Together with the channel rank information, the UE also provides the BS with the indexes of the pre-coding codebook vectors to be used at the transmitter side. The closed loop MIMO pre-coding mechanism, at the expense of additional signaling overhead, is the method used in LTE to effectively exploit the MIMO channel diversity. This is because the pre-coding vector indexes, requested by the UE, are chosen such that the SNR is maximized, therefore also maximizing the throughput. The SNR is computed on the overall equivalent channel constituted by the cascade of the pre-coding matrix and the instantaneous propagation channel matrix.

The standard does not mandate for a particular detection scheme but instead assumes a linear Minimum-Mean-Squared-Error (MMSE) equalizer as a baseline detector to establish the minimum performance requirement. The MIMO transceiver scheme is shown in Figure 11.53. The equalizer coefficients are adjusted depending on the channel matrix coefficients, the pre-coding vector and the interference power. It can equivalently be regarded as a 2×2 matrix multiplication of the 2×1 vector constituted by the complex signal at each sub-carrier at the output of the FFT of the two receiving antennas as follows:

$$\underline{\hat{x}}_i = \underline{\underline{G}}_{i,MMSE}\underline{y}_i = \underline{\underline{G}}_{i,MMSE}(\underline{\underline{\tilde{H}}}_i\underline{x}_i + \underline{n}_i) = \underline{\underline{G}}_{i,MMSE}(\underline{\underline{H}}_i\underline{\underline{P}}_i\underline{\hat{x}}_i + \underline{n}_i) \tag{11.15}$$

Where:

- $\underline{\hat{x}}_i$ is the 2×1 detected symbol vector at sub-carrier i;
- $\underline{\underline{G}}_{i,MMSE}$ is the MMSE equalizer 2×2 matrix at sub-carrier i;
- $\underline{y}_i$ is the 2×1 received signal vector at sub-carrier i;
- $\underline{\underline{\tilde{H}}}_i$ is the 2×2 equivalent channel matrix resulting from the cascade of the 2×2 pre-coding matrix $\underline{\underline{P}}_i$ and the actual 2×2 channel matrix $\underline{\underline{H}}_i$ at sub-carrier i;
- $\underline{n}_i$ is the 2×1 interference vector received signal at sub-carrier i.

11.9.2 Channel Modeling and Estimation

11.9.2.1 3GPP Guidelines for Channel Estimation

The coherent detection schemes mentioned earlier require the availability of a reliable channel estimate for each sub-carrier, for each OFDM symbol and for each link between transmitting and receiving antennas. For this purpose, LTE systems provide a Reference Signal (RS) whose resource elements are disposed in the time–frequency plane in a diamond-shaped uniform grid.

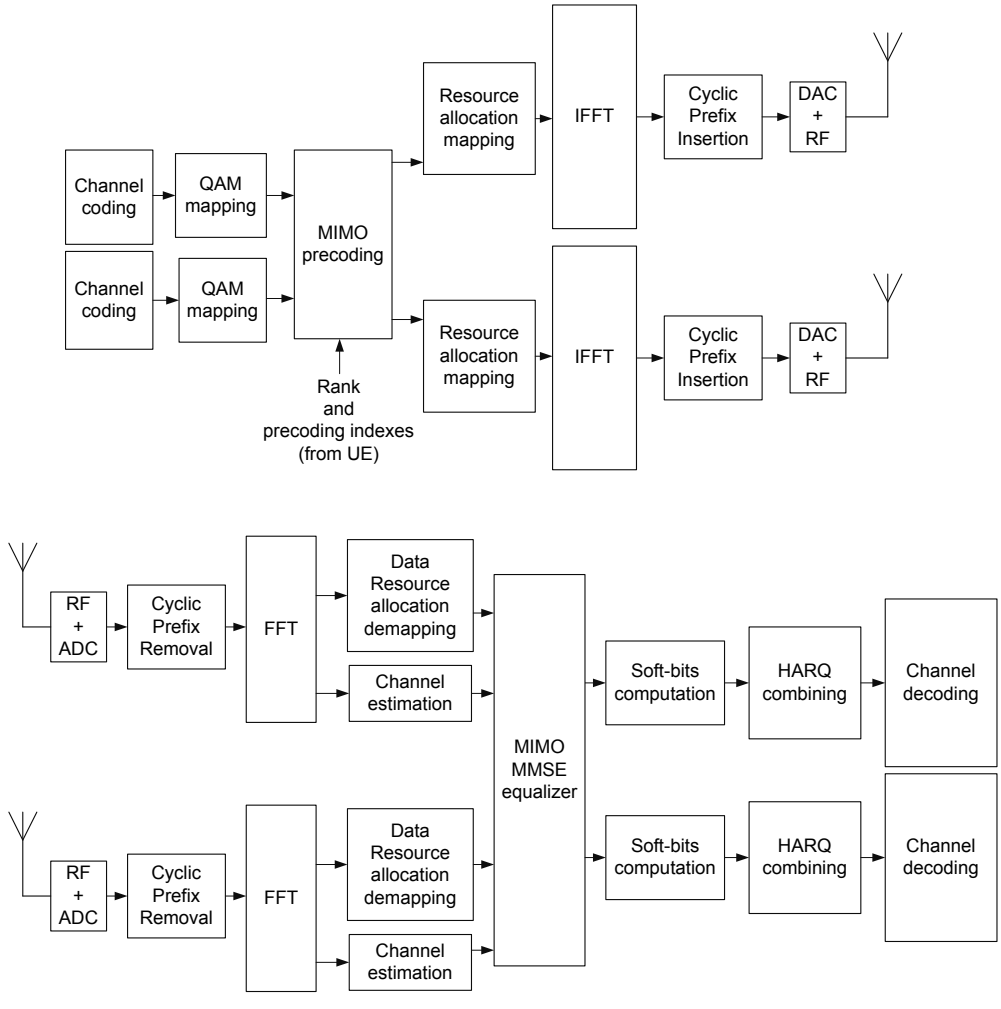


Figure 11.53 MIMO transceiver

For multiple antennas transmissions, RS are interleaved and zeroed out in correspondence to the RS of other antennas to minimize mutual interference. Hence, thanks to this particular structure introduced specifically in LTE, the channel estimation over antennas can be performed independently for each link between each transmitting antenna and each receiving antenna. This pilot signaling scheme considerably simplifies the channel estimation task in MIMO applications.

As for the receiver scheme, the standard gives freedom in the implementation of the frequency–time channel estimation, even if some companies provided practical guidelines along the standardization discussions for the definition of the performance requirements. These guidelines proposed the channel estimation to be implemented as the cascade of two 1D filtering processes: the first dimension consists of Wiener/MMSE filtering of the OFDM symbols containing RS in the frequency direction. The second dimension follows as a Wiener/MMSE

filter in the time domain to obtain the full channel transfer function for all sub-carrier indexes and OFDM symbols within a sub-frame.

Only the reference signals belonging to the current sub-frame are used for time domain interpolation. The coefficients used for the frequency domain filtering process are chosen from a pre-computed set as a function of signal-to-noise ratio only [45].

It is worth noting that since channel estimation is performed after the FFT operation, the actual channel being estimated is indeed the convolution of several filters' impulse responses, amongst which the time variant component is that of the physical air interface propagation channel, and the other filters are either the eNodeB channel filters, or those of the UE RF front end section. In that respect, in-band distortions introduced by RF filters are naturally compensated for as long as the total delay spread does not exceed the CP length.

11.9.3 Demodulation Performance

11.9.3.1 PDSCH Fixed Reference Channels

The performance requirements in [1] include a set of Fixed Reference Channels for each transmission mode – namely SIMO, transmit diversity and MIMO modes. The FRCs so far agreed, for the SIMO case and cell specific RS signals, fall into three categories involving a restricted set of Modulation and Coding Schemes: QPSK with coding rate 1/3, 16QAM with coding rate ½ and 64QAM with coding rate ¾.

The choice of such FRC categories was made so as to reduce to a minimum the number of tests while having representative MCS in the overall set of 29 MCS combinations.

Within each FRC category, one test is specified for a given system bandwidth and therefore characterized by a specific transport block length, code-block segmentation parameters assuming an allocation spanning the entire available bandwidth.

An additional FRC category is also defined for a single resource block allocation happening on the band-edge and making use of QPSK modulation.

11.9.3.2 PDSCH Demodulation Performance Requirements

The performance requirements are expressed as the required SNR $\hat{I}_{or}/N_{oc}$ in dB to obtain a given fraction of the nominal throughput of a given FRC and in given propagation conditions. The fraction of the nominal throughput is always chosen to be 30% or 70%. It is worth noting that for RF section performance tests, the metric has been set to a throughput greater or equal to 95% of the maximum throughput of the reference measurement channel. An example of the performance requirement for 64QAM as stated in the 3GPP standard is presented in Table 11.21.

The required SNR is agreed by consensus among the companies participating to the standard. The value is computed upon the decoding performances of a full-floating point receiver chain where an implementation margin is taken to account for the degradation induced by fixed point or other signal-processing non-idealities. The implementation margin is in the range of 1.5–2.0 dB. The principle just explained for the derivation of the performance requirement is graphically depicted in Figure 11.54.

Table 11.21 Minimum performance 64QAM (FRC)

Bandwidth	Propagation Condition	Correlation Matrix	Reference value	
			Fraction of Maximum Throughput (%)	SNR $\hat{I}_{or}/N_{oc}$ (dB)
10 MHz	EVA5	Low	70	17.7
	ETU70	Low	70	19.0
	EVA5	High	70	19.1

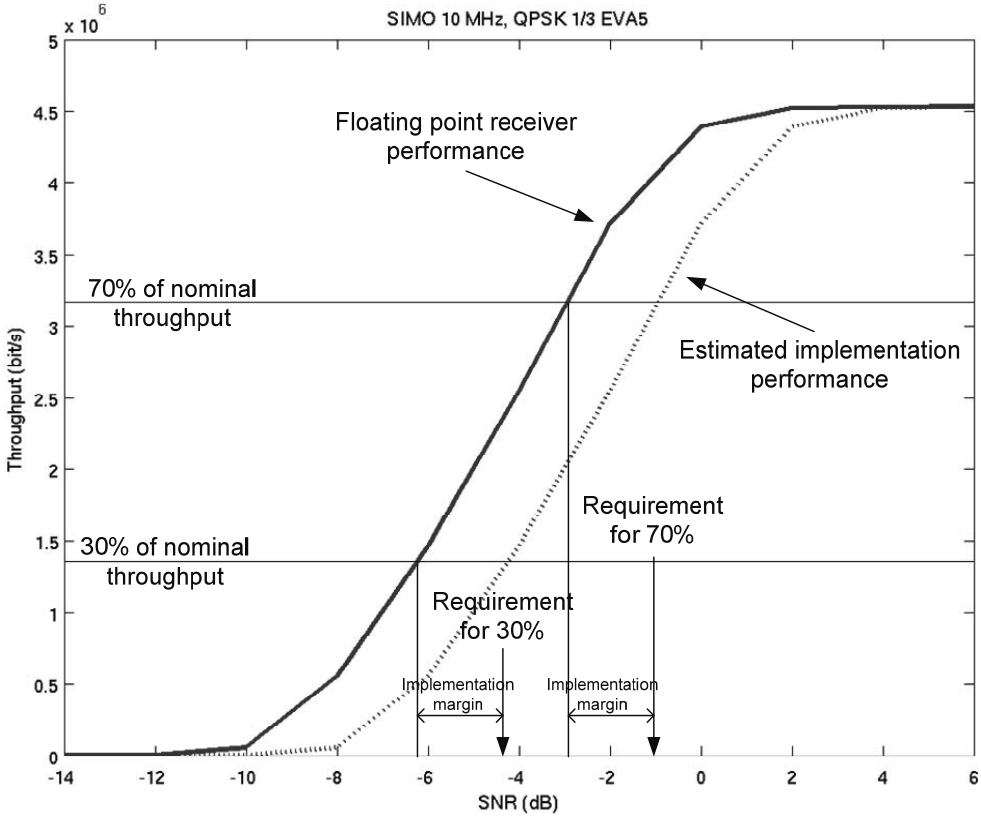


Figure 11.54 Principle of performance requirement derivation

11.9.3.3 EVM Modeling and Impact onto Performance

With the aim of deriving a realistic performance requirement, LTE assumes that the transmitter is modeled as a non-ideal signal source characterized by transmitter Error Vector Magnitude. The non-ideality in general results from non-linearities happening because of OFDM signal high Peak-to-Average Power Ratio and limited dynamic of the RF transmitter amplifier.

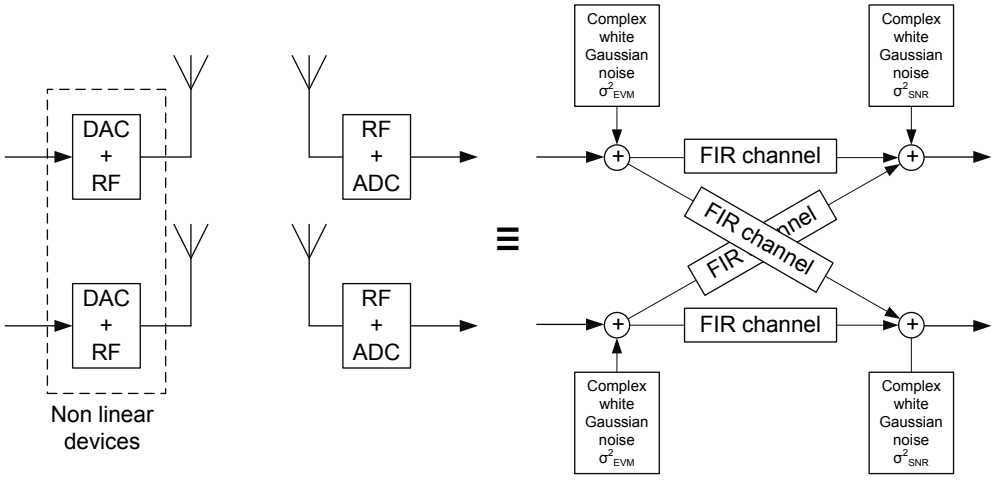


Figure 11.55 Equivalent EVM model used for performance simulations

Table 11.22 Error Vector Magnitude assumption for performance requirements

Modulation	Level	Equivalent per Tx antenna noise variance σ^2_{EVM} (dB)
QPSK	17.5%	–15.13924
16QAM	12.5%	–18.0618
64QAM	8%	–21.9382

Precise descriptions of the EVM modeling for LTE performance requirement derivation were provided, for example, in [47]. The model counts non-linearities as an additive distortion source at the transmitter thanks to the Bussgang theorem [48], as shown in Figure 11.55.

Table 11.22 presents the EVM levels to be assumed in performance simulation. The effect of the additive distortion source is to limit the attainable capacity as the receiver effective SNR cannot increase as other cell interference vanishes because of the irreducible distortion term, as Figure 11.56 shows. As an additional consequence, the MCS set gets limited and upper-bounded as the general system capacity.

11.10 Requirements for Radio Resource Management

The mobility procedures are described in general in Chapter 7. This section focuses on the performance requirements for mobility. The performance requirements for radio resource management (RRM) for E-UTRA are defined in [3] and follow a similar basic structure to those defined for UTRA in [40]. There are requirements supporting mobility in both E-UTRA RRC_Idle, and RRC_Connected states and mobility for E-UTRA intra-frequency, and E-UTRA inter-frequency as well as performance requirements for handover and reselection to other Radio Access Technologies (RATs) including UTRA FDD, UTRA TDD and GSM. Performance is also specified for mobility to non-3GPP RATs such as CDMA2000 1× and High Rate Packet Data (HRPD), also known as CDMA EV-DO.

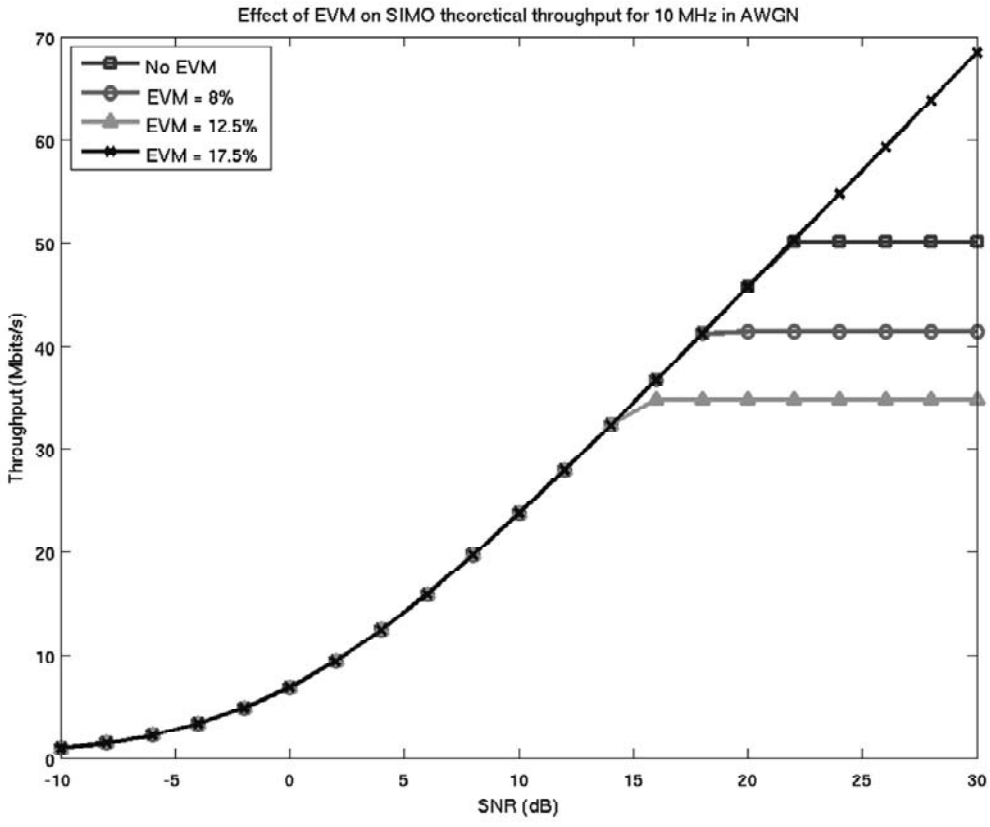


Figure 11.56 Effect of EVM on theoretical SIMO throughput for 10 MHz in AWGN

In RRC_Idle state, mobility is based on UE autonomous reselection procedures, whereas in RRC_Connected state measurement reports are made by the UE, to which the E-UTRA network may respond by issuing handover commands. Due to the autonomous nature of the idle state reselection procedures, and the relative importance of power saving in RRC_Idle state, typical idle performance requirements allow for more implementation freedom, whereas in RRC_Connected state, consistent UE measurement reporting is of prime importance, and so more details of measurement performance are specified to ensure that network based mobility algorithms can work with all UEs. One important difference in E-UTRA compared to UTRA is the specification of large DRX cycles in RRC_Connected state. The DRX cycle can be, for example, up to 2.56 s. The large DRX cycles are similar in length to idle state DRX cycles, but measurements are reported to the E-UTRA network for handover decision making.

As well as the performance requirements for supporting reselection and measurements for handover, [3] also includes performance requirements for radio related RRC procedures including RRC re-establishment and random access, and additionally contains requirements for UE transmit timing, and UE timer accuracy. This section gives an overview of the mobility related requirements.

11.10.1 Idle State Mobility

From a minimum performance perspective, [3] specifies how quickly the UE should be able to detect and evaluate the possibility of reselection to newly detectable target cells. In this context, an E-UTRA cell is considered detectable if it has signal to noise and interference ratio SCH_RP/$I_{ot} \geq -3$ dB. For cells that have previously been detected, [3] also specifies the minimum rate at which additional measurements should be made, and the minimum acceptable filtering and spacing of the samples used to support reselection. Maximum allowed evaluation times for cells that have been previously detected are also given. Largely, the requirements for measurement rate and maximum evaluation time are copied from those used previously for UTRA since basic needs for mobility should be similar.

Additionally, [3] specifies that reselection criteria will be re-evaluated at least every DRX cycle, since a new serving cell measurement, and probably also new neighbor cell measurements must be performed at least every DRX cycle. Also, once the reselection decision has been made, it is expected that the UE will make a short interruption in receiving the downlink paging channels when performing reselections. For reselections to E-UTRA and UTRA target cells, this is specified as the time needed to read system information of the target cell and an additional 50 ms implementation allowance.

11.10.2 Connected State Mobility when DRX is Not Active

In RRC_Connected state, the UE continuously searches for, and measures intra-frequency cells, and may additionally search for, and measure inter-frequency and inter-RAT cells if certain conditions are fulfilled, including the configuration of a measurement gap sequence if one is needed by the UE. Both periodic and event triggered reporting mechanisms are defined in RRC specifications to provide measurement results to the E-UTRA network. The relevant measurement quantities are shown in Table 11.23 for different radio technologies.

In general terms, [3] defines minimum performance requirements which indicate how quickly newly detectable cells should be identified and reported, measurement periods for cells that have been detected, and the applicable accuracy requirements and measurement report mapping for each of the measurement quantities in a measurement report. Furthermore, there are requirements in chapter 5 of [3] defining some aspects of how quickly the handover should be performed by the UE, once network mobility management algorithms have made the decision to perform a handover, and transmitted an RRC message to the UE to initiate the handover.

Table 11.23 Measurement quantities for different radio access technologies

Radio technology	Applicable measurement quantities
E-UTRA	Reference Signal Received Power (RSRP). Reference Signal Received Quality (RSRQ)
UTRA FDD	CPICH Received Symbol Code Power (RSCP). CPICH Ec/Io
UTRA TDD	PCCPCH received symbol code power (PCCPCH RSCP)
GSM	RSSI Note: BSIC confirmation may be requested for GSM measurements

11.10.2.1 Cell Identification

When in RRC_Connected state, the UE continuously attempts to search for, and identify new cells. Unlike UTRA, there is no explicit neighbor cell list containing the physical identities of E-UTRA neighbor cells. For both E-UTRA FDD and E-UTRA TDD intra-frequency cell identification, there are rather similar requirements and the UE is required to identify a newly detectable cell within no more than 800 ms. This should be considered a general requirement applicable in a wide range of propagation conditions when the SCH_RP/$I_{ot} \geq -6$ dB and the other side conditions given in [3] are met. The 800 ms requirement was agreed after a simulation campaign in 3GPP where cell identification was simulated by a number of companies at different SCH_RP/I_{ot} ratios.

It is also important to note that cell identification includes a 200 ms measurement period after the cell has been detected. Therefore, to comply with the requirement to identify cells within 800 ms, the UE needs to be able to detect cells internally in a shorter time to allow for the measurement period..

When less than 100% of the time is available for intra-frequency cell identification purposes – because, for example, the intra-frequency reception time is punctured by measurement gaps – the 800 ms time requirement is scaled to reflect the reduced time available.

For E-UTRA inter-frequency measurements, a similar approach is taken. Since inter-frequency cell identification is performed in measurement gaps, the basic identification time $T_{basic\_identify}$ is scaled according to the configured gap density so that

$$T_{Identify\_Inter} = T_{Basic\_Identify\_Inter} \cdot \frac{T_{measurement\_Period\_Inter\_FDD}}{T_{Inter}} \text{ [ms]} \tag{11.16}$$

The $T_{measurement\_Period\_Inter}$ is multiplied by the number of E-UTRA carriers which the UE is monitoring (denoted as N_{freq}), which in turn means that the identification time $T_{Identify\_Inter}$ is also multiplied by N_{freq} and so the requirement for inter-frequency cell identification becomes longer the more frequency layers are configured.

Similar requirements are also defined in [3] for UTRA cell identification when in E-UTRA RRC_Connected state.

11.10.2.2 Measurement of Identified Cells

Once cells are identified, the UE performs measurements on them, over a defined measurement period. 3GPP specifications do not define the sample rate at which the UE Layer 1 (L1) is required to make measurement (or even that uniform sampling is necessary) but the measurements specified are filtered over a standardized measurement period to ensure consistent UE behavior. The UE is also expected to meet the accuracy requirements (discussed further in section 11.10.2.3) over the measurement period, which places constraints on how many L1 samples of the measurement need to be taken and filtered during the measurement period.

For intra-frequency cells, minimum capabilities for measurement are defined in [3]. In summary, the UE is required to have the capability to measure 8 intra-frequency cells when 100% of the time is available for intra-frequency measurements. When less time is available, for example due to inter-frequency measurement gaps, then the requirement is scaled down accordingly. The period for intra-frequency measurements is specified as 200 ms, although this would again be scaled up if less time is available for intra-frequency measurements.

For inter-frequency cells, there are two measurement period requirements defined in [3], one of which is a mandatory requirement, and one of which may be optionally supported by the UE. The mandatory requirement is based on a measurement bandwidth of 6 resource blocks, in which case the measurement period is $480 \times N_{freq}$ ms. When it is indicated by signaling that a bandwidth of at least 50 resource blocks is in use throughout a particular frequency layer, the UE may optionally support a measurement period of $240 \times N_{freq}$ ms. In this case, the UE knows from the signaling that it is safe to make the measurement over the wider bandwidth, and may therefore be able to achieve the required accuracy while filtering fewer measurement samples in the time domain, making possible a reduced measurement period.

For inter-frequency measurements, the minimum requirement is that the UE is capable of supporting three E-UTRA carrier frequencies (in addition to the intra-frequency layer) and on each of these three carrier frequencies it should be capable of measuring at least four cells.

11.10.2.3 Accuracy Requirements and Report Mapping

For both RSRP and RSRQ, absolute and relative accuracy requirements are defined in [3] chapter 9. Absolute accuracy considers the difference between the actual and the ideal measurements for a single cell, whereas relative accuracy considers how much error can be expected when comparing the levels of two cells.

For RSRP, both intra-frequency and inter-frequency absolute and relative accuracy requirements are defined. For comparison of two cells measured on the same frequency, the main sources of inaccuracy can be studied in link level simulations of the measurements, and this was the approach used to define the accuracy requirements. For absolute measurements of RSRP, uncertainty in the RF gain setting is a significant additional source of error, and this is reflected in the accuracy requirements, especially in extreme temperature and voltage conditions. When considering inter-frequency relative RSRP accuracy, some of the uncertainties in RF gain setting cancel out, since they apply to both the cells which are being compared, so the performance requirement for inter-frequency RSRP relative accuracy is somewhat tighter than the inter-frequency absolute accuracy requirement, but not as tight as the intra-frequency relative accuracy requirement.

When considering the accuracy of RSRQ, it should be noted that comparison of two RSRQ measurements on the same frequency is not particularly useful. The reason is that the RSSI component of the RSRQ in both measurements will be similar because for intra-frequency both cells must be measured on the same frequency layer. For this reason, only absolute accuracy requirements are defined for intra-frequency RSRQ measurements. Since RF gain setting uncertainties to an extent affect both the measurement of the RSRP and RSSI components in RSRQ, uncertainties will somewhat cancel out, and as such, RSRQ absolute accuracy is required to be better than RSRP absolute accuracy.

[3] also defines RSRP and RSRQ report mapping, which defines how a UE measured quantity value should be mapped to a signaled information element. This defines the range and resolution of the values which may be reported by the UE.

11.10.3 Connected State Mobility when DRX is Active

One new feature of E-UTRA compared to UTRA which is important from a mobility aspect is that rather large DRX cycles (e.g. up to 2.56 s) may be used when the UE is in RRC_Connected

state. This means that inactive UEs may still be kept in the connected state, and will use connected state mobility procedures such as handover. To allow power saving opportunities for such devices, [3] specifies different measurement performance requirements which are applicable when large DRX cycles (80 ms and greater) are active. The basis of the large DRX measurement performance requirements is to ensure that the UE would be able to perform the mobility related measurements at or around the time when the receiver is active in the DRX cycle anyway. Power saving can be performed for the remainder of the DRX. To facilitate this, measurement periods and cell identification times are specified as a function of the configured DRX cycle length.

11.10.4 Handover Execution Performance Requirements

Similarly to UTRA specifications, handover performance is characterized by two different delays, which are illustrated in Figure 11.57. Total handover delay is denoted by $D_{handover}$ and is the total time between the end of transmission of the handover command on the source cell, and the start of UE transmissions to the target cell. In addition, maximum allowable interruption time is separately specified, and represents the time for which the UE is no longer receiving and transmitting to the source cell, and has not yet started transmission to the target cell.

RRC procedure delays are specified in [42] section 11.2 and the interruption time is given by

$$T_{interrupt} = T_{search} + T_{IU} + 20\,\text{ms} \tag{11.17}$$

where T_{search} is the time taken to search for the target cell. This will be 0 ms when the target cell has previously been measured by the UE within the last 5 s, and is only not zero for a blind handover to a cell not known to the UE. T_{IU} relates to the uncertainty in timing between the physical channel structure of the old cell. An additional 20 ms implementation margin is also included in the requirement.

As indicated in Figure 11.57, the total handover delay is either $D_{handover}$ = RRC procedure delay + $T_{interrupt}$ or activation time + $T_{interrupt}$ if an activation time is given which is greater than the RRC procedure delay in the future.

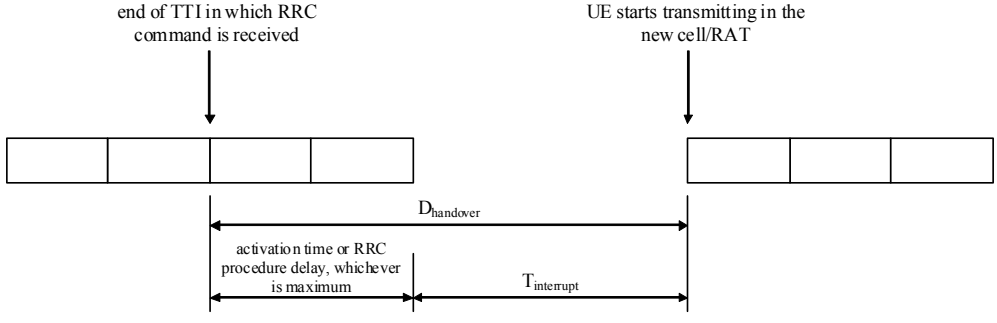


Figure 11.57 Handover delay and interruption time

11.11 Summary

Clearly defined performance requirements are an essential part of an open well-functioning standard. The requirements are needed to provide predictable performance within an operator's own band in terms of data rates, system capacity, coverage and mobility with different terminals and different radio network vendors. The requirements are also needed to facilitate the coexistence of LTE in the presence of other systems as well as the coexistence of adjacent LTE operators in the same area. The regulatory requirements are considered when defining the emission limits. 3GPP has defined minimum RF performance requirements for LTE terminals (UE) and for base stations (eNodeB) facilitating a consistent and predictable system performance in a multi-vendor environment.

3GPP LTE frequency variants cover all the relevant cellular bands. The first version of the specifications covers 17 different variants for FDD and 8 for TDD bands. More frequency variants will be added in the future. The frequency variants are independent of the other content of the releases.

3GPP LTE requirements are defined in such a way that it enables efficient multimode GSM/WCDMA/LTE device implementation from the RF requirements perspective. Yet, the need to support receive diversity required for MIMO operation, and the multiple frequency band combinations present a significant challenge for the front-end components optimization of such mobile devices. The SC-FDMA transmission in uplink has similar requirements as the current HSUPA transmission. The baseband processing requirements are naturally increased due to the high data rates in LTE but the simple front-end structure in OFDMA and the Turbo decoding parallelization simplifies the practical implementations.

References

[1] 3GPP Technical Specification 36.101 'User Equipment (UE) radio transmission and reception', v. 8.3.0.
[2] 3GPP Technical Specification 36.104 'Base Station (BS) radio transmission and reception', v. 8.3.0.
[3] 3GPP Technical Specification 36.133 'Requirements for support of radio resource management', v. 8.3.0.
[4] 3GPP Technical Report 36.942 'Radio Frequency (RF) system scenarios', v. 8.0.0.
[5] 3GPP Technical Specifications 36.141 'Base Station (BS) conformance testing', v. 8.0.0.
[6] 3GPP Technical Specifications 36.108 'Common test environments for User Equipment (UE); Conformance testing', v. 8.0.0.
[7] 3GPP Technical Specifications 36.114 'User Equipment (UE)/Mobile Station (MS) Over The Air (OTA) antenna performance; Conformance testing', v. 8.0.0.
[8] 3GPP R4–070342 'On E-UTRA spectrum emission mask and ACLR requirements', April 2007.
[9] European Communications Committee (ECC)/European Conference of Postal and Telecommunications Administrations (CEPT) Report #19, 'Report from CEPT to the European Commission in response to the Mandate to develop least restrictive technical conditions for frequency bands addressed in the context of WAPECS', December 2007.
[10] 3GPP R4–070124 'System simulation results for derivation of E-UTRA BS EVM requirements', February 2007.
[11] Holma, H., Toskala, A., 'WCDMA for UMTS: HSPA Evolution and LTE', 4th edition, Wiley, 2007.
[12] The Mobile Industry Processor Interface (MIPI^{SM}) Alliance, www.mipi.org
[13] Myung, H.G., Junsung Lim, Goodman, D.J., 'Peak-To-Average Power Ratio of Single Carrier FDMA Signals with Pulse Shaping', Personal, Indoor and Mobile Radio Communications, 2006 IEEE 17th International Symposium on September 2006: 1–5.
[14] Valkama, M., Anttila, L., Renfors, M., 'Some radio implementation challenges in 3G-LTE context', Signal Processing Advances in Wireless Communications, 2007. SPAWC 2007. IEEE 8th Workshop on 17–20 June 2007: 1–5.

[15] Priyanto, B.E., Sorensen, T.B., Jensen, O.K., Larsen, T., Kolding, T., Mogensen, P., 'Impact of polar transmitter imperfections on UTRA LTE uplink performance', Norchip, 2007, 19–20 November 2007: 1–4.
[16] Talonen, M., Lindfors, S., 'System requirements for OFDM polar transmitter', Circuit Theory and Design, 2005. Proceedings of the 2005 European Conference, Volume 3, 28 August–2 September 2005: III/69–III/72.
[17] Chow, Y.H., Yong, C.K., Lee, Joan, Lee, H.K., Rajendran, J., Khoo, S.H., Soo, M.L., Chan, C.F., 'A variable supply, (2.3–2.7) GHz linear power amplifier module for IEEE 802.16e and LTE applications using E-mode pHEMT technology'. Microwave Symposium Digest, 2008 IEEE MTT-S International 15–20 June 2008: 871–874.
[18] Ericsson, ' TP 36.101: REFSENS and associated requirements', TSG-RAN Working Group 4 (Radio) meeting #46bis, R4–080696, Shenzhen, China, 31 March–4 April, 2008.
[19] Ericsson, 'Simulation results for reference sensitivity and dynamic range with updated TBS sizes', TSG-RAN Working Group 4 (Radio) meeting #47bis, R4–081612, Munich, 16th–20th June 2008.
[20] Nokia Siemens Networks, 'Ideal simulation results for RF receiver requirements', 3GPP TSG-RAN WG4 Meeting #48, R4–081841, Jeju, Korea, 18th–22nd August, 2008.
[21] STN-wireless internal WCDMA handset benchmarking test report.
[22] Anadigics AWT6241 WCDMA Power Amplifier datasheet, http://www.anadigics.com/products/handsets_data-cards/wcdma_power_amplifiers/awt6241
[23] Ericsson, 'Introduction of MSD (Maximum Sensitivity Degradation)', TSG-RAN Working Group 4 (Radio) meeting #49, R4–082972, Prague, Czech Republic, 10–14 November 2008.
[24] NTT DoCoMo, Fujitsu, Panasonic, ' Performance requirements on Self interference due to transmitter noise', TSG-RAN Working Group 4 Meeting #47, R4–080873, Kansas, USA, 5–9 May 2008.
[25] Ericsson, ' Introduction of MSD (Maximum Sensitivity Degradation)', 36.101 Change Request # 90, R4–083164, 3GPP TSG-RAN WG4 Meeting #49, Prague, Czech Republic, 10–14 November 2008.
[26] NTT DOCOMO, ' Maximum Sensitivity Reduction values for Band 6/9/11', R4–083043, TSG-RAN Working Group 4 Meeting #49, Prague, Czech Republic, Nov. 10–14, 2008.
[27] 'Half Duplex FDD Operation in LTE', 3GPP TSG RAN WG4 Meeting #46, R4–080255, Sorrento (Italy), February 11th to 15th 2008.[28] Ericsson, 'HD-FDD from a UE perspective', TSG-RAN Working Group 4 (Radio) meeting #46bis,R4–080695 Shenzhen, China, 31 March – 4 April 2008.
[29] Manstretta, D., Brandolini, M., Svelto, F., 'Second-Order Intermodulation Mechanisms in CMOS Downconverters', IEEE Journal of Solid-State Circuits, Vol. 38, No. 3, March 2003, 394–406.
[30] Faulkner, M., 'The Effect of Filtering on the Performance of OFDM Systems', IEEE Transactions on Vehicular Technology, vol. 49, no. 5, September 2000: 1877–1883.
[31] Lindoff, B., Wilhelmsson, L., 'On selectivity filter design for 3G long-term evolution mobile terminals', Signal Processing Advances in Wireless Communications, 2007. SPAWC 2007. IEEE 8th Workshop on Volume, Issue, 17–20 June 2007: :1–5.
[32] Dardari, D., 'Joint Clip and Quantization Effects Characterization in OFDM Receivers', IEEE Transactions On Circuits And Systems – I: Regular Papers, vol. 53, no. 8, August 2006: 1741–1748.
[33] Locher, M. *et al*., 'A Versatile, Low Power, High Performance BiCMOS, MIMO/Diversity Direct Conversion Transceiver IC for Wi-Bro/WiMax (802.16e)', IEEE Journal of Solid State Circuits, vol. 43, no. 8, August 2008: 1–10.
[34] Norsworthy, S.R., Schreier, R., Temes, G.C., 'Delta-Sigma data converters, theory, design and simulation', Wiley Interscience, 1996.
[35] Internal STN-wireless data extracted with simulation parameter modifications, From: 'Ouzounov, S. van Veldhoven, R. Bastiaansen, C. Vongehr, K. van Wegberg, R. Geelen, G. Breems, L. van Roermund, A., "A 1.2V 121-Mode ΣΔModulator for Wireless Receivers in 90nm CMOS", Solid-State Circuits Conference, 2007. ISSCC 2007. Digest of Technical Papers. IEEE International Publication Date: 11–15 February 2007: 242–600.
[36] Martel, P., Lossois, G., Danchesi, C., Brunel, D., Noël, L., 'Experimental EVM budget investigations in Zero-IF WCDMA receivers', submitted to Electronics Letters.
[37] Windisch, M., Fettweis, G., 'Performance Degradation due to I/Q Imbalance in Multi-Carrier Direct Conversion Receivers: A Theoretical Analysis', Communications, 2006. ICC apos;06. IEEE International Conference on Volume 1, June 2006: 257–262.
[38] Robertson, P., Kaiser, S., 'Analysis of the effects of phase-noise in orthogonal frequency division multiplex (OFDM) systems', Communications, 1995. ICC 95 Seattle, Gateway to Globalization, 1995 IEEE International Conference on Volume 3, 18–22 June 1995: 1652–1657.
[39] Pollet, T., Moeneclaey, M., Jeanclaude, I., Sari, H., 'Effect of carrier phase jitter on single-carrier and multi-carrier QAM systems', Communications, 1995. ICC 95 Seattle, Gateway to Globalization, 1995 IEEE International Conference on Volume 2, 18–22 June 1995:1046–1050.

[40] 3GPP Technical Specification 25.133 'Requirements for support of radio resource management (UTRA)', v. 8.3.0.
[41] 3GPP Technical Specification 36.213 'Physical layer procedures', v. 8.3.0.
[42] 3GPP Technical Specification 36.331 'Radio Resource Control (RRC); Protocol specification', v. 8.3.0.
[43] 3GPP Technical Specification 25.101 'User Equipment (UE) radio transmission and reception (UTRA)', v. 8.3.0.
[44] Qualcomm Europe, Change Request 'R4–082669: UE ACS frequency offset', TSG-RAN Working Group 4 (Radio) meeting #49, Prague, CZ, 10–14 November, 2008.
[45] Motorola, 'Reference Channel and Noise Estimators', 3GPP E-mail reflector document for TSG-RAN Working Group 4 Meeting #46, Sorrento, Italy, 11–15 February, 2008.
[46] Almers, P. *et al*., 'Survey of channel and radio propagation models for wireless MIMO systems'. EURASIP Journal on Wireless Communications and Networking, 957–1000, July 2007.
[47] Ericsson, 'R4–071814: Clarification of TX EVM model', TSG-RAN Working Group 4 (Radio) meeting #44, Athens, Greece, 20–24 August 2007.
[48] Bannelli, P., Cacopardi, S., 'Theoretical Analysis and Performance of OFDM Signals in Nonlinear AWGN Channels,' IEEE trans. on Communications, vol. 48, no. 3, March 2000.
[49] Berkmann J. *et al*., 'On 3G LTE Terminal Implementation – Standard, Algorithms, Complexities and Challenges', Wireless Communications and Mobile Computing International Conference, 2008. IWCMC '08.

12

LTE TDD Mode

Che Xiangguang, Troels Kolding, Peter Skov, Wang Haiming and Antti Toskala

12.1 Introduction

With full coverage in the 3GPP Release 8 specifications of both Time Division Duplex (TDD) and Frequency Division Duplex (FDD) modes of operation, LTE can effectively be deployed in both the paired and unpaired spectrum. LTE TDD and FDD modes have been greatly harmonized in the sense that both modes share the same underlying framework, including radio access schemes OFDMA in downlink and SC-FDMA in uplink, basic subframe formats, configuration protocols, etc. As clear indication of the harmonization, the TDD mode is included together with the FDD mode in the same set of specifications, including the physical layer where there are just a few differences due to the uplink/downlink switching operation. In terms of architecture there are no differences between FDD and TDD and the very few differences in the MAC and higher layer protocols relate to TDD specific physical layer parameters. Procedures are kept the same. Thus there will be high implementation synergies between the two modes allowing for efficient support of both TDD and FDD in the same network or user device. Coexistence would of course still require careful analysis.

Another key feature of the LTE TDD mode (known also as TD-LTE) is the commonality with TD-SCDMA. This is an advantage as in, e.g. China, where the Release 4 based TD-SCDMA (including enhancements from later releases) has opened up a large-scale TDD system deployment, paving the way for further deployment of 3GPP based LTE TDD using the available unpaired spectrum. As presented in Chapter 11, there is a global trend to reserve significant unpaired spectrum allocations.

In this chapter, the detailed aspects of LTE TDD that differ from the FDD mode are introduced. Further, information related to both the link and system performance of the LTE TDD mode of operation is given.

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6

12.2 LTE TDD Fundamentals

The basic principle of TDD is to use the same frequency band for transmission and reception but to alternate the transmission direction in time. As shown in Figure 12.1 this is a fundamental difference compared to FDD, where different frequencies are used for continuous UE reception and transmission. Like FDD, LTE TDD supports bandwidths from 1.4 MHz up to 20 MHz but depending on the frequency band, the number of supported bandwidths may be less than the full range. For example, for the 2.5 GHz band, it is not likely that the smallest bandwidths will be supported. Since the bandwidth is shared between uplink and downlink and the maximum bandwidth is specified to be 20 MHz in Release 8, the maximum achievable data rates are lower than in LTE FDD. This way the same receiver and transmitter processing capability can be used with both TDD and FDD modes enabling faster deployment of LTE.

The TDD system can be implemented on an unpaired band (or in two paired bands separately) while the FDD system always requires a pair of bands with a reasonable separation between uplink and downlink directions, known as the duplex separation. In a FDD UE implementation this normally requires a duplex filter when simultaneous transmission and reception is facilitated. In a TDD system the UE does not need such a duplex filter. The complexity of the duplex filter increases when the uplink and downlink frequency bands are placed in closer proximity. In some of the future spectrum allocations it is foreseen that it will be easier to find new unpaired allocations than paired allocations with sensible duplex separation thereby increasing further the scope of applicability for TDD.

However, since uplink and downlink share the same frequency band, the signals in these two transmission directions can interfere with each other. This is illustrated in Figure 12.2

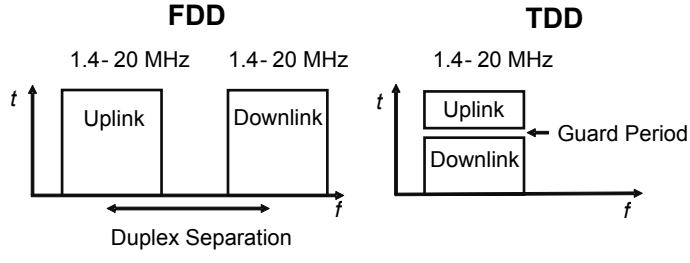


Figure 12.1 Principles of FDD and TDD modes of operation

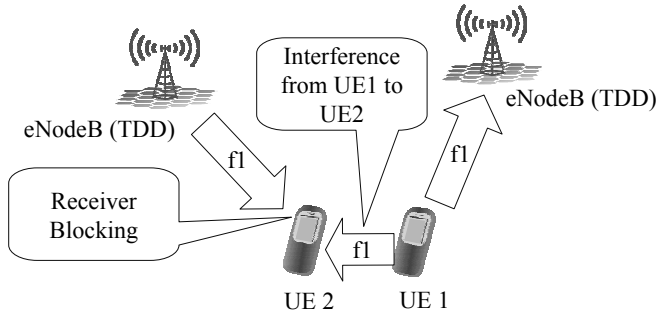


Figure 12.2 Interference from uplink to downlink in uncoordinated TDD operation

with the use of TDD on the same frequency without coordination and synchronization between sites in the same coverage area. For uncoordinated deployment (unsynchronized) on the same frequency band, the devices connected to the cells with different timing and/or different uplink/downlink allocation may cause blocking for other users. In LTE TDD the base stations need to be synchronized to each other at frame level in the same coverage area to avoid this interference. This can be typically done by using, for example, satellite based solutions like GPS or Galileo or by having another external timing reference shared by the LTE TDD base stations within the same coverage area. LTE FDD does not need the base station synchronization. There is no interference between uplink and downlink in FDD due to the duplex separation of the carriers.

Two adjacent LTE TDD operators (on adjacent carriers) should preferably synchronize the base stations and allocate the same asymmetry between uplink and downlink to avoid potentially detrimental interference to the system reliability. If the two operators do not coordinate the deployments, there is instead a need for guard bands and additional filtering. These requirements are discussed in Chapter 11.

12.2.1 LTE TDD Frame Structure

As the single frequency block is shared in the time domain between uplink and downlink (and also between users), the transmission in LTE TDD is not continuous. While often also being the case for data transmissions towards a certain user in LTE FDD mode, the level of discontinuity then depends entirely on the scheduling function (except for half-duplex FDD terminals). For control channels, e.g. the PDCCH and the PHICH, the transmission for FDD is continuous. For LTE TDD all uplink transmissions need to be on hold while any downlink resource is used and, conversely, the downlink needs to be totally silent when any of the UE is transmitting in the uplink direction. Switching between transmission directions has a small hardware delay (for both UE and eNodeB) and must be compensated. To control the resulting switching transients a Guard Period (GP) is allocated which compensates for the maximum propagation delay of interfering components (e.g. depends on cell size and level of available cell isolation). The impact of discontinuous uplink and downlink on the link budget in LTE TDD is covered specifically in section 12.5.

To explain the exact implementation of the mechanism for switching between downlink and uplink and vice versa, consider the example setup of the LTE TDD frame structure in Figure 12.3.

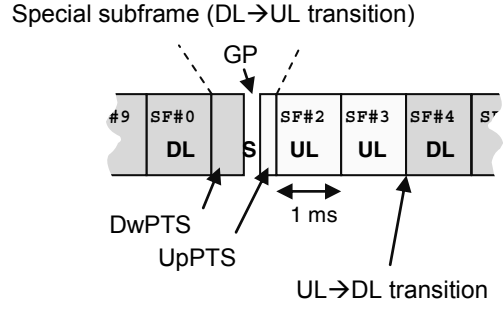


Figure 12.3 Simple illustration of the DL→UL and UL→DL transitions implemented in downlink (DL), uplink (UL), and special (S) subframes (SF)

The subframe denoted by either uplink (UL) or downlink (DL) has a design in common with LTE FDD with some minor but significant differences related to common control channels. In LTE TDD there is maximally one DL→UL and one UL→DL transition per 5 ms period (half-frame). The UL→DL transition is carried out for all intra-cell UE by the process of time alignment. The eNodeB instructs each UE to use a specific time offset so that all UE signals are aligned when they arrive at the eNodeB. Hence, uplink is synchronous as is the case for FDD. To ensure that the UE has sufficient time to shut down its transmission and switch to listening mode, the UE does not transmit a signal during the last 10–20 ms of subframe #3 in Figure 12.3. This procedure ensures that there is no UE transmission power from the own cell that spills over into the downlink transmission. Although eNodeBs in different cells are fully synchronized, this method does not prevent UE from other cells spilling their interference into the downlink transmission of the current sector. However, in practice this is less of a problem since any individual UE has limited transmission power.

While the UL→DL switching is merely an intra-cell method, the DL→UL switching method ensures that the high-power downlink transmissions from eNodeBs from other neighbor cells do not interfere when the eNodeB UL reception is ongoing in the current cell. Adopting the methodology of TD-SCDMA, LTE TDD introduces a special (S) subframe that is divided into three parts; the Downlink Pilot Time Slot (DwPTS), the GP, and the Uplink Pilot Time Slot (UpPTS). The special subframe replaces what would have been a normal subframe #1. The individual time duration in OFDM symbols of the special subframe parts are to some extent adjustable and the exact configuration of the special time slot will impact the performance. The GP implements the DL→UL transition and the GP has to be sufficiently long to cover the propagation delay of all critical downlink interferers on the same or adjacent carriers as well as the hardware switch-off time. Hence, the correct setting of the GP depends on network topology, antenna configurations, etc. To fit into the general LTE frame numerology, the total duration of DwPTS, GP, and UpPTS is always 1 ms.

DwPTS is considered as a 'normal' downlink subframe and carries control information as well as data for those cases when sufficient duration is configured. High commonality is achieved by rescaling the transport block size according to its length. In this way the effective coding rate for a certain selection of payload and transmission bandwidth will stay the same. UpPTS is primarily intended for sounding reference signal (SRS) transmissions from the UE, but LTE TDD also introduces the short RACH concept so that this space may be used for access purposes as well. The flexibility for the different components of the special subframe is summarized in Table 12.1 when using the normal cyclic prefix. The GP can be up to about 700 µs thus supporting a cell range up to 100 km. For more details related to the many possible configurations, including the extended cyclic prefix, the reader is referred to [1]. When discussing coexistence with TD-SCDMA in section 12.5 more details related to special subframe configuration are given.

Table 12.1 Possible durations configurable for special subframe components. Total duration of the three components is always 1 ms (normal cyclic prefix) [1]

Component	Unit	Range of duration
Downlink pilot time slot (DwPTS)	# OFDM symbols (71 µs)	3–12
Guard period (GP)	# OFDM symbols (71 µs)	1–10
Uplink pilot time slot (UpPTS)	# OFDM symbols (71 µs)	1–2

12.2.2 Asymmetric Uplink/Downlink Capacity Allocation

A key advantage to TDD is the ability to adjust the available system resources (time and frequency) to either downlink or uplink to match perfectly the uplink and downlink traffic characteristics of the cell. This is done by changing the duplex switching point and thus moving capacity from uplink to downlink, or vice versa. The LTE TDD frame structure can be adjusted to have either 5 ms or 10 ms uplink-to-downlink switch point periodicity. The resulting resource split is either balanced or between two extreme cases:

- single 1 ms subframe for uplink and 8 times 1 ms subframe for downlink per 10 ms frame;
- if we want to maximally boost the uplink capacity, then we can have three subframes for uplink and one for downlink per 5 ms. Thus the resulting uplink activity factor can be adjusted from 10% to 60% (if we do not consider the UpPTS).

For continuous coverage across several cells, the chosen asymmetry is normally aligned between cells to avoid the interference between transmission directions as described earlier. Hence, synchronization in a LTE TDD system deployed in a wide area is conducted both at frame and uplink–downlink configuration level. In practice it is expected that the uplink–downlink configuration is changed in the network only very seldom and the associated signaling of LTE TDD in Release 8 has been optimized according to this assumption, e.g. a LTE TDD Release 8 cellular system would be characterized as a static or semi-static TDD system. The UE is informed about the active uplink–downlink configuration via the System Information Block (SIB-1), which is broadcast via the Dynamic Broadcast Channel (D-BCH) to the cell with an update rate of 80 ms. The knowledge about which uplink–downlink configuration is active in the cell is essential for the UE to know the location of critical control channels and the timing of adaptation methods such as HARQ.

12.2.3 Co-existence with TD-SCDMA

As mentioned, there is a strong legacy between LTE TDD and TD-SCDMA. The coexistence of these systems on the same or adjacent carriers has therefore been a key discussion point during the standardization process. For sharing a site with TD-SCDMA and using the same frequency band, the systems need to have an aligned uplink/downlink split to avoid the interference between different base station transceivers. As TD-SCDMA slot duration does not match the LTE TDD subframe duration, the LTE subframe parameterization was designed to accommodate coexistence. From the knowledge of the relative uplink/downlink division, the relative timing of TD-SCDMA and LTE TDD can be adjusted to allow coexistence without BTS to BTS interference, as shown in Figure 12.4. Note the duration of fields in LTE TDD subframe with uplink/downlink that vary depending on the configuration, thus the timings shown can take different values.

Besides timing alignment, the exact configuration of the special subframe in LTE TDD also plays an important role when allowing TD-SCDMA/LTE TDD coexistence. Some of the configurations with a good match are listed in Table 12.2, including the required detailed settings for DwPTS, GP and UpPTS. Normal cyclic prefix is assumed. For more options for configuration, the reader is referred to [1].

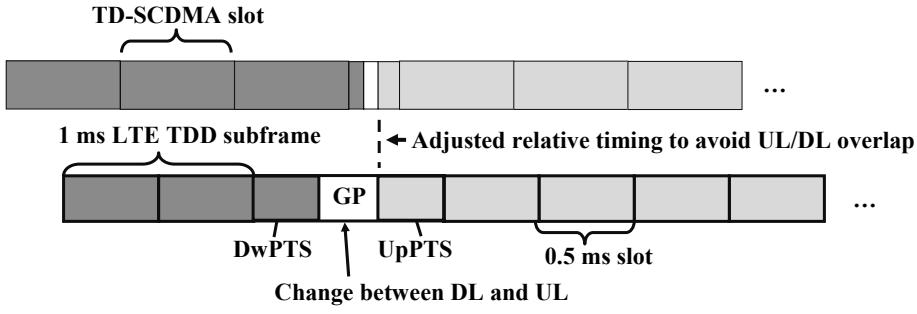


Figure 12.4 Ensuring TD-SCDMA and LTE TDD coexistence with compatible UL/DL timings

Table 12.2 Example modes for coexistence between TD-SCDMA and LTE TDD (normal cyclic prefix)

TD-SCDMA configuration	LTE TDD configuration	Special subframe configuration (OFDM symbols)			
		Configuration	DwPTS	GP	UpPTS
5DL-2UL	#2 (3DL-1UL)	#5	3	9	2
4DL-3UL	#1 (2DL-2UL)	#7	10	2	2
2DL-5UL	#0 (1DL-3UL)	#5	3	9	2

12.2.4 Channel Reciprocity

One interesting aspect for TDD systems is that the signals in uplink and downlink travel through the same frequency band. Assuming symmetric antenna configurations and adequate RF chain calibration, there is a high correlation of the fading on the own signal between uplink and downlink known as channel reciprocity. The key associated benefit is that the measurements in one direction may fully or partially be used to predict the other direction thereby enabling a reduction of the signaling overhead needed for link adaptation, packet scheduling and the use of advanced transmit antenna techniques. As already mentioned antenna configurations need to be reciprocal for full reciprocity to be applicable. The baseline configuration for LTE is two antennas in both eNodeB and UE, so in this case antenna configuration seems to be reciprocal. However, as the UE only has one power amplifier, it can only transmit on one antenna at a time so the only way to measure the full channel in UL is for the UE to switch the sounding transmission between the two antennas. If this is done we can say that both channel and antenna configuration are reciprocal.

An additional practical limitation towards the exploitation of reciprocity is that most adaptation schemes, including link adaptation methods and fast packet scheduling, rely on SINR or throughput assessments. As the interference level may vary significantly between uplink and downlink, the LTE TDD specifications therefore do not rely on the availability of channel reciprocity and allow for a full decoupling of uplink and downlink thus using the very same uplink and downlink reporting methods as are available for LTE FDD. However, within the specifications there is a large degree of freedom available to optimize the total signaling overhead in cases where reciprocity has significant gain and is applicable.

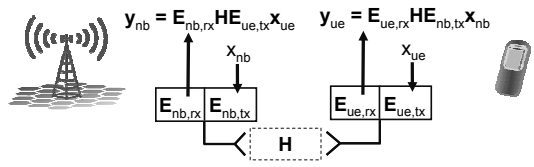


Figure 12.5 Model to illustrate the impact from RF units to channel reciprocity. Capital letters identify matrixes

A simple linear model for the TDD channel including impact from RF-chains is shown in Figure 12.5. From this it is observed that even though the channel represented by **H** is identical for UL and DL then the channel seen by physical layer is different. In the typical application of channel state information we would like to know the channel as seen by physical layer (in DL $\mathbf{E}_{nb,tx}\mathbf{HE}_{ue,rx}$) because pre-coding and decoding are all done after the signal has passed through the RF parts. This information cannot be estimated from UL sounding. It is possible to improve the channel reciprocity by adding calibration procedures within the Node B and UE but to calibrate the full RF chain so that $\mathbf{x}_{nb} = \mathbf{x}_{ue}$ would imply $\mathbf{y}_{nb} = \mathbf{y}_{ue}$ would require a standardized procedure.

Relying on channel reciprocity for pre-coding is very challenging as reciprocity of the full complex valued channel is assumed. For other purposes, such as reciprocity based fast power control, we only rely on reciprocity of the channel attenuation and this would be less sensitive towards calibration errors [2].

12.2.5 Multiple Access Schemes

In LTE TDD the applied multiple access schemes are the same as with LTE FDD. Hence, OFDMA is used for the downlink and SC-FDMA is used for the uplink direction as shown in Figure 12.6. The reasoning behind the selection was shared by TDD and FDD modes respectively. The selection of a power efficient uplink scheme was very important especially for a TDD device because of the more stringent uplink link budget resulting from the limited

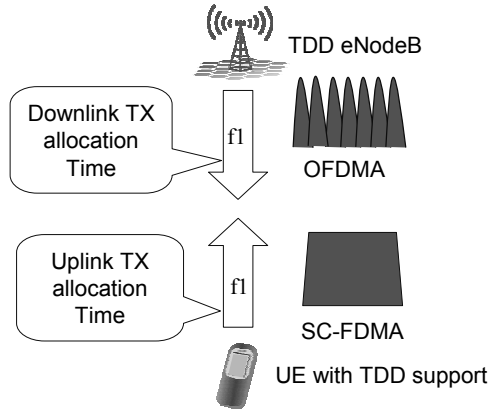


Figure 12.6 LTE TDD multiple access schemes in uplink and downlink

uplink transmission time allocation. From an implementation complexity viewpoint, this is an improvement over, for example, WCDMA, where different multiple access schemes between FDD (with WCDMA) and TDD (with TD/CDMA) modes required different receiver and transmitter solutions (even with the identical chip rate in Release 99) in the baseband implementation. Hence, complexity of any dual mode TDD/FDD implementation is reduced in LTE when compared to UTRAN.

12.3 TDD Control Design

Due to the special frame structure, the placement of critical control channels is different for LTE TDD than it is for LTE FDD. The exact placement of control channels depends on which uplink–downlink configuration is active in the cell and is thus key information for the UE to know to set up its connection with the network. In the following, the placement of the critical channels as well as their special meaning for LTE TDD is described. For general information about the meaning of the control channels, the reader is referred to Chapter 5, where the dynamic use of Physical Downlink Shared Control Channel (PDCCH) for uplink and downlink resource allocation is covered.

For FDD operation, where downlink and uplink are in separate frequencies with continuous transmission and reception on its dedicated frequency, the shared control channel design is straightforward due to one-to-one associated downlink and uplink subframes. However, for TDD operation where downlink and uplink share the same frequency for transmission and reception but alternate the transmission direction in time, physical control channel design is a bit more challenging. This was further complicated by the high available flexibility to adjust the downlink/uplink allocation.

12.3.1 Common Control Channels

To illustrate the placement of the different control channels, TDD uplink–downlink configuration #1 is used as reference. The critical system control channels are shown in Figure 12.7 and described in the following.

The Primary Synchronization Signal (PSS) is placed at the third symbol in subframes #1 and #6, which is different from LTE FDD where the PSS is placed at the last symbol of the first slot in subframes #0 and #5. The Secondary Synchronization Signal (SSS) is placed at the

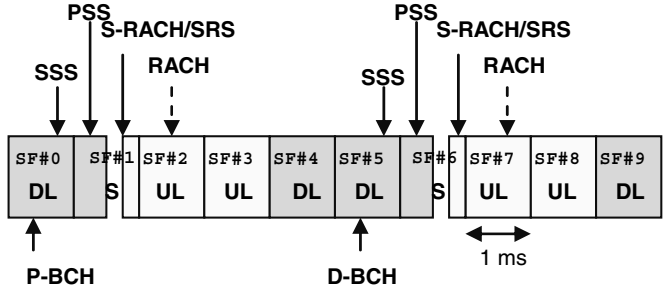


Figure 12.7 Mapping of critical control channels to TDD configuration #1

last symbol in subframes #0 and #5, which is also different from LTE FDD where the SSS is placed at the second last symbol of the first slot in subframes #0 and #5. Synchronization via the PSS is typically a robust process and UE can survive UE-to-UE interference of up to 100 dB for the initial cell search. This high measuring robustness is facilitated by using normalization of samples achieved over a long measuring window.

There are five RACH preamble formats defined for LTE TDD. Four of the RACH preamble formats are common with LTE FDD, while the LTE TDD specific RACH preamble format 4 is known as SHORT RACH (S-RACH) due to the short preamble sequence duration. As shown in Figure 12.7 the S-RACH is transmitted on the UpPTS within the special subframe.

RACH needs to be protected to make access reliable and to allow for the coexistence of multiple LTE TDD users. The RACH channel is fairly flexible regarding the density and placement in time and frequency. The RACH density, similarly to that for LTE FDD, can be 0.5, 1, 2, 3, 4, 5 or 6 in every 10 ms radio frame. The exact placement of RACH in LTE TDD within the available UL resources (i.e. UL subframes and UpPTS) in the configured UL/DL configuration is selected by proper network planning taking into account the requirement for coexistence with TD-SCDMA, traffic load, network topology, etc. All together, there are 58 configurations for RACH placement in time and frequency for each UL/DL configuration with the common principle that RACH channels are first distributed in the time domain among all available UL resources before multiplexed in the frequency domain. This principle is exemplified in Figure 12.8 for UL/DL configuration #0, and the exact RACH configurations can be found in [1]. The maximum available number of RACH channels per UL subframe or UpPTS (for S-RACH only) is six, which is different from FDD where at most one RACH channel is available per UL subframe. This ensures that even with limited UL resources available, sufficient RACH opportunities for high traffic load can be created. As illustrated in Figure 12.9, when there is more than one RACH channel in one UL subframe or UpPTS, the RACH channels are distributed towards both edges of the system bandwidth for RACH in UL subframe or placed from only one edge (either bottom or top) when in UpPTS. For the latter case, the frequency position is altered in the next RACH instance as shown in the figure.

The placement of the Primary Broadcast Channel (P-BCH) and the Dynamic Broadcast Channel (D-BCH) is the same as it is in LTE FDD. However, the detection of D-BCH is slightly different from that in LTE FDD. This is because to detect D-BCH, the size of the control channel in that DL subframe needs to be known to the UE. However, this raises a chicken-and-egg problem since the size of the control channel depends on the active UL/DL configuration and the UE does not yet know the active UL/DL configuration before it correctly detects D-BCH during the initial cell search. For all UL/DL configurations, there are three different sizes for the control channel in any DL subframe, thus the UE will need to perform three hypotheses on control channel size to detect the D-BCH. Once the UE correctly receives the D-BCH, it knows the UL/DL configuration in the cell, after which the hypothesis detection is no longer needed when detecting the D-BCH.

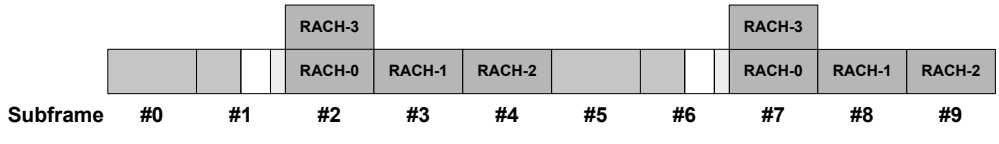


Figure 12.8 RACH distribution/placement in time and frequency in LTE TDD

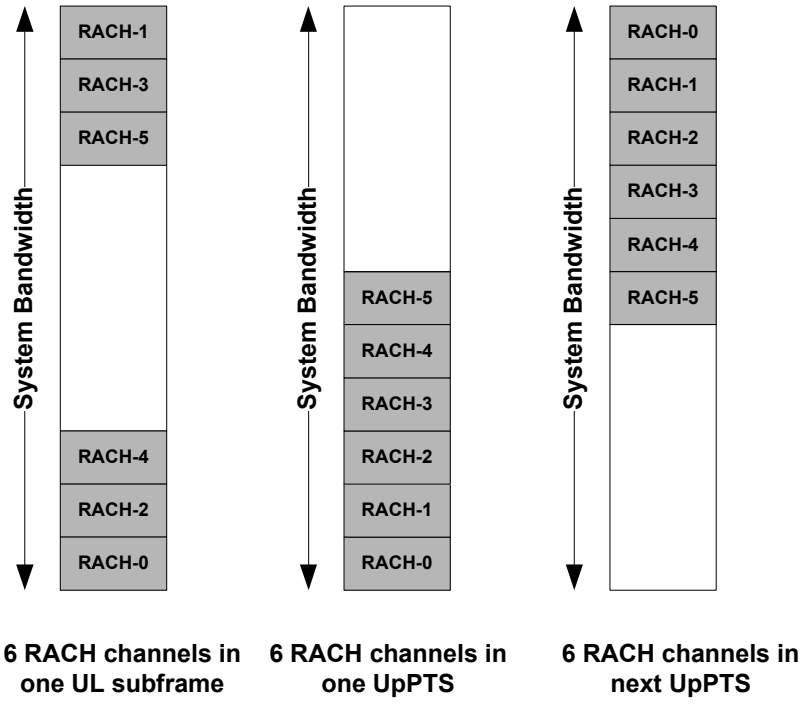


Figure 12.9 RACH placement in the frequency domain in LTE TDD

The paging procedure is the same for LTE TDD and FDD; however, the exact subframes being used for paging are slightly different in LTE TDD. For FDD, the subframes #0, #4, #5 and #9 can be used for paging. For TDD, the subframes #0, #1, #5 and #6 can be used for paging.

12.3.2 Sounding Reference Signal

In LTE TDD the Sounding Reference Signal (SRS) for any UE can be transmitted not only in the last symbol of one UL subframe as in LTE FDD, but also in one or both symbols of UpPTS. Since the UpPTS is anyway available and cannot carry uplink data traffic, it is expected to be the primary location for SRS in LTE TDD. The support of SRS is considered mandatory for all LTE TDD devices due to the installed base of TD-SCDMA base stations, which often have existing beamforming antenna configurations.

12.3.3 HARQ Process and Timing

The key to designing and configuring HARQ for TDD operation is to determine the required processing time of the eNodeB and the UE for Uplink (UL) and Downlink (DL) respectively. The relevant processing times are as follows:

- DL UE: Duration from the time when the last sample of the packet is received in downlink until a HARQ-ACK/NACK is transmitted in UL.
- DL eNodeB: Duration from the time when an HARQ-ACK/NACK is transmitted in UL until the eNodeB can (re)-transmit data on the same HARQ process.
- UL eNodeB: Duration from the time when the last sample of the packet is received in UL until a HARQ-ACK/NACK (or a new allocation on same HARQ process) is transmitted in DL.
- UL UE: Duration from the time when a UL grant (or HARQ-ACK/NACK) is given until the UE is able to transmit the associated packet in UL.

In FDD, the eNodeB and UE processing times for both DL and UL are fixed and assumed to be 3 ms due to invariant DL and UL subframe configuration and continuous DL and UL transmission and reception. In TDD, although the eNodeB and UL (minimum) processing times are the same as for FDD, the actual HARQ timing (i.e. DL UE/eNodeB, UL eNodeB/UE) varies depending on the active DL/UL configuration. This is exemplified in Figure 12.10, where it is seen that compared to the minimum processing time of 4 ms, there is sometimes an additional delay incurred due to unavailable DL or UL subframe after 3 ms.

In addition to the minimum HARQ processing time, a few other decisions were made related to the special subframe to fix the required number of HARQ processes and timing in the specifications:

- There shall be always a PDCCH in the DwPTS at least for UL grant or PHICH. If DwPTS spans more than three OFDM symbols it also contains a PDSCH.
- UpPTS does not contain HARQ control information nor a PUSCH.

With the given assumptions of processing time and special subframe behavior, the number of the DL HARQ process and the UL HARQ process varies with TDD between 4 and 15 for DL and 1 and 7 in UL (in FDD that was always constant). As for FDD, in TDD the DL HARQ is asynchronous and the UL HARQ is synchronous. The resulting delay thus varies depending on the subframe being used as well as on the uplink/downlink split applied, and can take values between 4 and 7 ms for the delay k_1 between the associated UL grant (or an intended PHICH) on the PDCCH and the UL data transmission over the PUSCH. Respectively the delay k_2 between the associated UL data transmission over PUSCH and PHICH (or a UL grant for re-/new transmission) over the PDCCH also varies between 4 and 7 ms.

The multi-TTI scheduling scheme in TDD allows for efficient use of the downlink shared control channel resources (PDCCH) in case less downlink resources are available and further decreases the UE complexity of decoding PDCCH (see Figure 12.11).

The DL HARQ timing delay k_3 between the associated DL data transmission over PDSCH and the UL HARQ-ACK transmission over PUCCH or PUSCH varies between 4 and 13 ms depending again on the uplink/downlink split applied. This is now simplified since TDD DL HARQ is operated in asynchronous mode like LTE FDD. The intended DL HARQ timing is derived by $n + k_3$, where n is the addressed subframe index number.

Due to the discontinuous UL/DL transmission, the TDD frame structure adds some delay to the user plane transmission when HARQ is involved. While the HARQ Round Trip Time (RTT) in LTE FDD is 8 ms (as described in Chapter 11) the corresponding LTE TDD RTT is at least 10 ms and up to 16 ms. The relatively small difference is because even for LTE FDD the core RTT is dominated by the UE and eNodeB HARQ processing times.

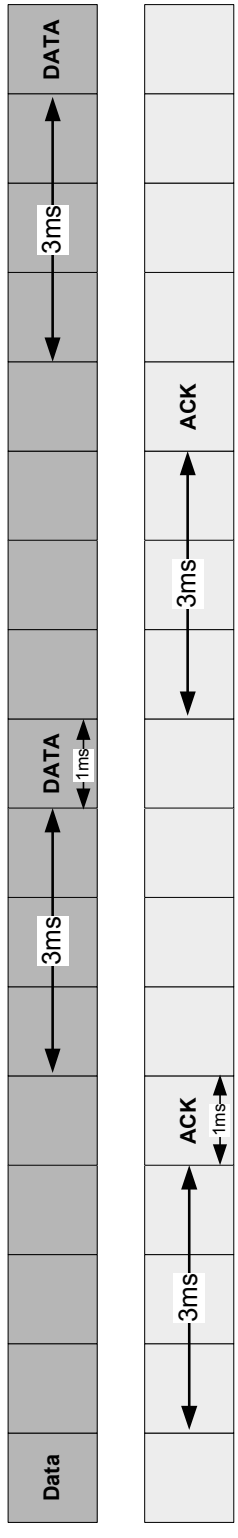


Figure 12.10 HARQ Timing: (a) LTE FDD; (b) LTE TDD

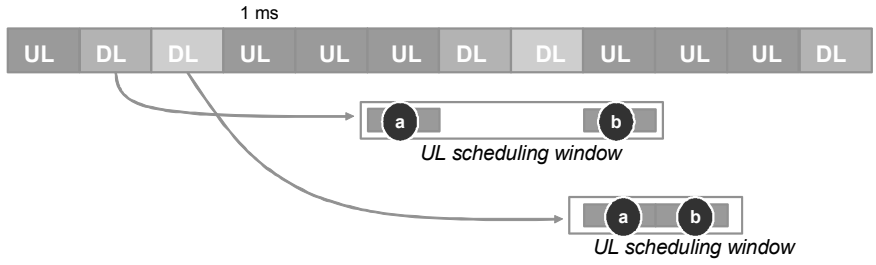


Figure 12.11 Possible uplink multi-TTI scheduling

12.3.4 HARQ Design for UL TTI Bundling

UL TTI bundling is used to improve the UL coverage by combining multiple UL TTIs for one HARQ process transmission. In a system where it is expected that many retransmissions are needed to successfully transmit a packet in uplink, UL TTI bundling provides fast automatic retransmission so that the delay penalty is minor. Although many details of UL TTI bundling are the same for LTE FDD and TDD, specifically the HARQ timing is different due to the inherited special UL HARQ timing in LTE TDD. In terms of the 3GPP discussion, only TDD UL/DL configurations #0**,** #1 and #6 have bundling fully defined in Release 8. To control the aspect of TTI bundling timing and HARQ process number after TTI bundling, the starting point of UL TTI bundling is limited and the number of bundled HARQ processes is fixed according to the bundle size. Thus the number of bundled HARQ processes is defined by:

$$Bundled\_HARQ\_No = \left\lceil \frac{Original\_HARQ\_No \times Bundled\_HARQ\_RTT}{Bundled\_size \times Original\_HARQ\_RTT} \right\rceil \quad (12.1)$$

In Equation 1, *Original_HARQ_No* and *Original_HARQ_RTT* are fixed for each TDD configuration in the LTE uplink system, so if assuming for instance TDD configuration #1 and a bundle size of 4, the number of bundled HARQ processes is 2, as shown in Figure 12.12.

As to the principle of ACK/NACK timing, it is always tied to the last subframe in a bundle, which is exactly the same as the rule in FDD. The uplink TTI bundling for LTE FDD is described in Chapter 10.

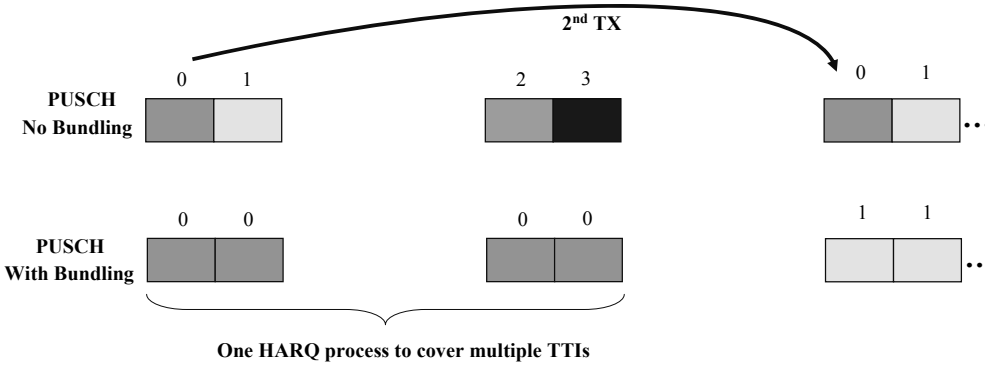


Figure 12.12 TTI bundling with TDD configuration #1

12.3.5 UL HARQ-ACK/NACK Transmission

In the same way as for FDD, the UL HARQ-ACK/NACK for LTE TDD is transmitted over the PHICH on PDCCH. The PHICH mapping and indexing is mostly the same for LTE FDD and TDD, i.e. the PDCCH in one DL subframe only contains the PHICH associated to a single UL subframe PUSCH. The only exception to this rule is the TDD UL/DL configuration #0 where the PHICH associated has specific exception specified.

12.3.6 DL HARQ-ACK/NACK Transmission

For both LTE FDD and TDD, the DL HARQ-ACK/NACK is transmitted on the PUCCH or the PUSCH depending on whether UL has simultaneous data transmission in the same UL subframe or not. In many cases the DL HARQ-ACK/NACK associated from more than one PDSCH, e.g. up to 9, will be mapped into a single UL subframe. This so-called multiple UL ACK/NACK transmission is, however, notably different from the FDD MIMO case in which the DL HARQ-ACK/NACK associated from a single PDSCH (e.g. two codewords) is mapped into a single UL subframe.

The fact that multiple downlink transmissions may need to be acknowledged within a single uplink subframe, makes the design for good UL coverage for control channels in TDD even more challenging. A very special design arrangement has been created to accomplish this task while simultaneously respecting the single carrier property of the UL multiple access scheme when UE has to transmit multiple DL HARQ-ACK/NACKs. There are two DL HARQ ACK/NACK feedback modes supported in TDD operation of LTE which are configured by a higher layer on a per-UE basis:

- ACK/NACK bundling feedback mode (the default mode), where a logical AND operation is performed per codeword's HARQ ACK/NACK across multiple DL subframes PDSCH whose associated HARQ ACK/NACK is mapped into the same UL subframe.
- ACK/NACK multiplexing feedback mode, where a logical AND operation is performed across spatial codewords within a DL HARQ ACK/NACK process. In Release 8 LTE TDD up to 4 bits DL HARQ ACK/NACK is supported per UL subframe, hence for UL/DL configuration #5 this feedback mode is not supported.

The ACK/NACK bundling feedback mode is the most aggressive mode to relieve the coverage problem of multiple UL ACK/NACK transmission in TDD. The allocated DL resources have been decoupled from the required UL feedback channel capability; i.e. only a single DL HARQ-ACK/NACK is transmitted in a single UL subframe regardless of the number of associated DL subframes carrying PDSCH for the user. The single ACK/NACK is then created by performing a logical AND operation over all associated HARQ ACK/NACK per UL subframe. This way, TDD has the same number of HARQ ACK/NACK feedback bits and thus transmission formats on PUCCH as FDD per UL subframe. The ACK/NACK encoding and transmission format in PUSCH is the same as it is in FDD.

Without proper compensation in the link adaptation and packet scheduling functions, the probability of DL HARQ NACK will increase causing more unnecessary DL retransmissions when using ACK/NACK bundling. Thus the control channel reliability is the key with ACK/NACK bundling. The second mode is more attractive for when the UE has sufficient UL cov-

erage to support multiple ACK/NACK bits on PUCCH. When in ACK/NACK multiplexing feedback mode, the status of each DL subframe HARQ-ACK/NACK, i.e. ACK, NACK or DTX (no data received in the DL subframe), one of the QPSK constellation points in certain derived PUCCH channels is selected for transmission on the UE side, and the eNodeB can decode the multi-bit HARQ ACK/NACK feedback by monitoring all constellation points from all associated PUCCH channels. The exact mapping table can be found in [3].

12.3.7 DL HARQ-ACK/NACK Transmission with SRI and/or CQI over PUCCH

When in ACK/NACK bundling mode, if both bundled HARQ-ACK/NACK and SRI are to be transmitted in the same UL subframe, the UE transmits the bundled ACK/NACK on its derived/assigned PUCCH ACK/NACK resource for a negative SRI transmission, or transmits the bundled HARQ-ACK/NACK on its assigned SRI PUCCH resource for a positive SRI transmission. This operation is exactly the same as for FDD.

When in ACK/NACK multiplexing mode, if both multiple HARQ-ACK/NACK and SRI are transmitted in the same UL subframe, the UE transmits the multiple ACK/NACK bits according to section 12.3.5 for a negative SRI transmission, and transmits 2-bit information mapped from multiple ACK/NACK input bits on its assigned SRI PUCCH resource for a positive SR transmission using PUCCH format 1b. The mapping between multiple ACK/NACK input bits and 2-bit information depends on the number of generated HARQ-ACK among the received DL subframe PDSCH within the associated DL subframes set K. The exact mapping table can be found in [3].

12.4 Semi-persistent Scheduling

Semi-persistent Scheduling (SPS) can be used with all the TDD UL/DL configurations. Many details of SPS are the same for LTE FDD and TDD, but in this section some TDD specific aspects related to SPS are detailed. The reader is referred to Chapter 10 for more information about SPS. To match the special frame structure of TDD, the SPS resource interval must be set to equal a multiple of the UL/DL allocation period (i.e. 10 ms) to avoid the conflict of non-matching UL/DL subframes because the UL subframe and DL subframe do not exist simultaneously. Furthermore, LTE UL uses synchronous HARQ and there are some problems for most UL/DL configurations in TDD because the HARQ RTT is 10 ms. When UL SPS is used for VoIP traffic (AMR codec periodicity is 20 ms), the second retransmission of a previous packet will collide with the next SPS allocation, since the period of SPS resource is two times the RTT. The collision case is shown in Figure 12.13. In the figure the numbers 1, 2 and 3 indicate different VoIP packets. For example, at the 20 ms point, the second retransmission of VoIP packet #1 collides with the initial VoIP packet #2. To solve this problem, two solutions are available to the network: dynamic scheduling and two-interval SPS patterns.

Although configured for SPS, the UE will anyway listen to dynamic allocations on the PDCCH. Such allocations will always override an existing persistent allocation. By using dynamic scheduling at the collision point it is possible to mitigate the problem as shown in Figure 12.14: if the UE is asked for a retransmission, the UE will perform a retransmission unless it has an empty buffer. With these definitions, if there are other following idle subframes available, the eNodeB will next schedule the retransmission in the current subframe, and

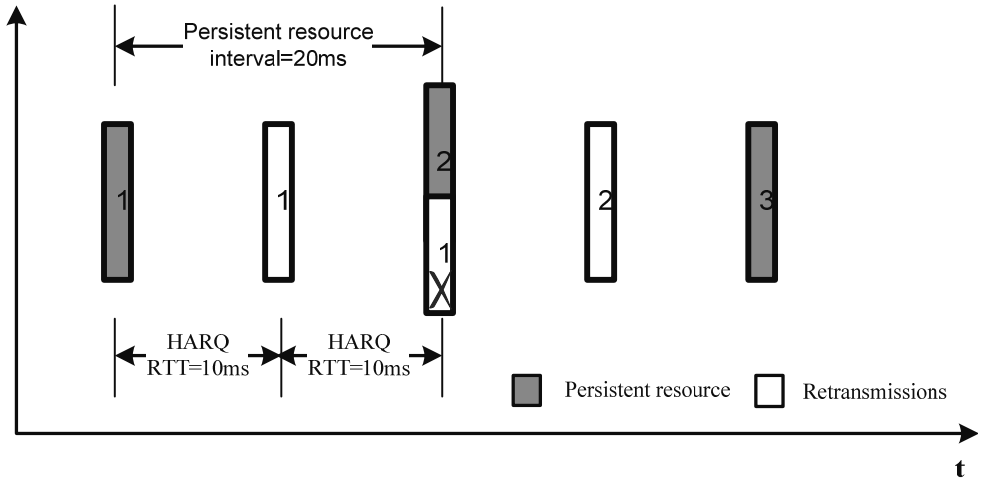


Figure 12.13 Collision between retransmissions and a new transmission

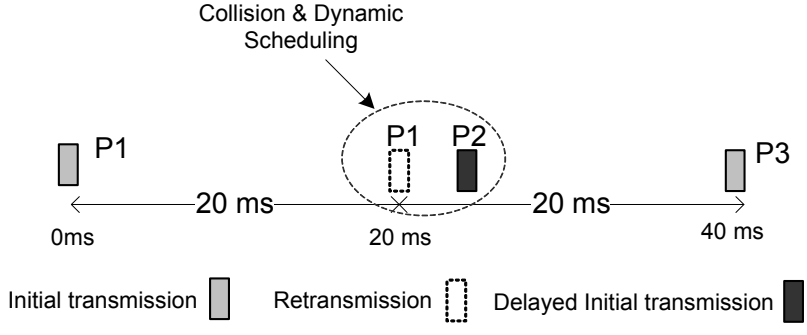


Figure 12.14 Dynamic scheduling at collision point

reschedule the initial transmission that was supposed to take place on the SPS resources in following subframes at the collision point.

The second solution is to use a two-interval SPS pattern. Here a two-interval SPS pattern means that two periods are used for semi-persistent scheduling while only one semi-persistent scheduling period is used and persistent allocation is carried out based on the pre-defined period in the conventional scheme. With a two-interval SPS pattern, a resource pattern with two different intervals (i.e. T1, T2, T1, T2 …) is used to avoid the main collision between the second retransmission of the previous packet and the SPS allocation. The procedure is given in Figure 12.15, in which T1 is not equal to T2 and the combined set of T1 and T2 is always a multiple of 10 ms. The offset between T1 and T2 is several subframes, and a variable subframe offset is used to indicate the offset. The following formulas are used to calculate T1 and T2:

$$\mathrm{T1} = \mathit{SPS\ periodicity} + \mathit{subframe\_offset} \tag{12.2}$$

$$\mathrm{T2} = \mathit{SPS\ periodicity} - \mathit{subframe\_offset} \tag{12.3}$$

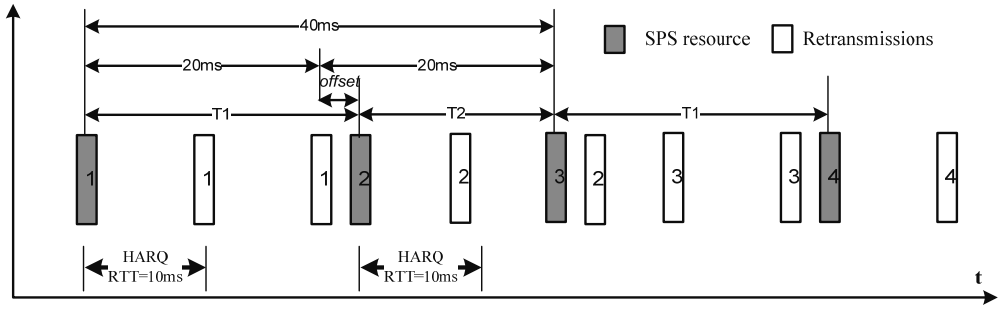


Figure 12.15 Two-interval SPS patterns

'*SPS periodicity*' will be signaled by RRC signaling; '*subframe_offset*' (positive value in Figure 12.15, but can also be negative) is implicitly defined according to the different TDD configurations and the starting point of a two-interval SPS pattern), and then T1 (the first time periodicity) and T2 (the second time periodicity) can be computed in terms of these equations. The allocation period always starts with the first time period T1. However, even configured with a two-interval SPS pattern, some residual collisions might still exist if the number of required retransmissions is large, i.e. 4. Any residual collisions can be avoided by means of dynamic scheduling as described earlier.

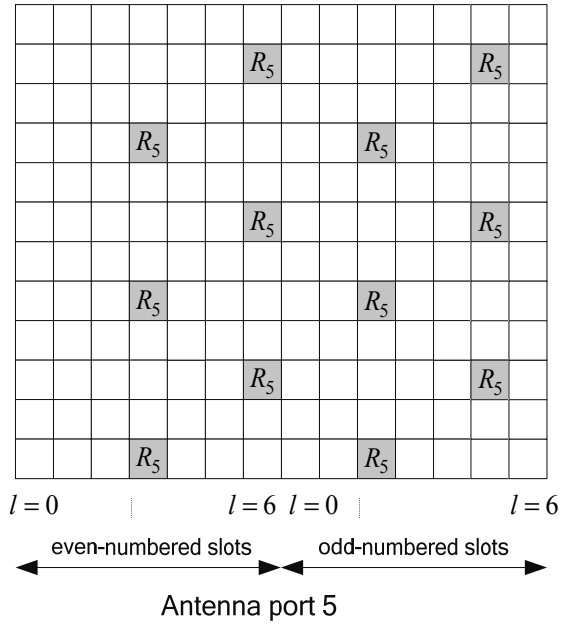
12.5 MIMO and Dedicated Reference Signals

LTE supports a number of different MIMO modes in DL, as already described in Chapter 5, covering both closed loop schemes with UE feedback information to the Node B. Together with the information of actually adapted downlink parameters from eNodeB this adds up to a significant amount of signaling to handle the DL closed loop MIMO. For TDD mode the earlier mentioned channel reciprocity can be exploited to mimic closed loop MIMO with a reduced signaling overhead. Under the assumption that the channel is identical for UL and DL we can estimate the channel in UL by using SRSs transmitted by the UE and then apply this channel knowledge for selecting the best DL pre-coding matrix. In this way UE feedback can be reduced or even avoided.

Going a step further, we can also eliminate the pre-coding matrix indication in the DL allocation message by using UE Specific Reference Signals (URS) to transfer this information. Moreover, the use of URS decouples the physical transmit antenna from the UE detection complexity and system overhead resulting from having a cell specific reference signal for each transmit antenna. URS are specified for LTE Release 8, and can be used for both FDD and TDD modes. However, this transmission mode is especially attractive when considered in a TDD setting where channel reciprocity is available.

URS are transmitted on antenna port 5 and they are generated with the same procedure as cell specific reference signals. The only difference is that the UE RNTI impacts the seed of the pseudo-random generator used to generate the code. The pattern for normal cyclic prefix and extended cyclic prefix can be seen in Figure 12.16(a) and Figure 12.16(b), respectively. There are 12 resource elements used for URS per PRB per 1 ms subframe so the additional overhead is rather large. With a normal cyclic prefix, cell specific reference signals on antenna port 0

(a)



(b)

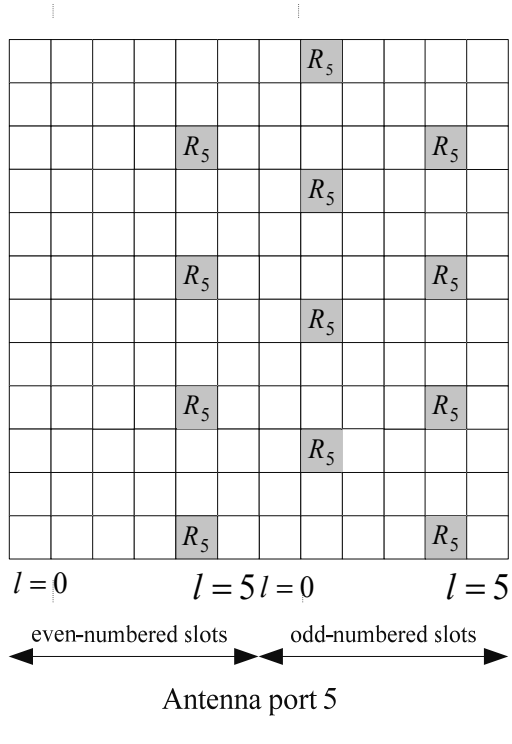


Figure 12.16 (a) URS pattern for normal CP; (b) URS pattern for extended CP

and 1 and a control channel region of three symbols antenna, enabling URS will reduce the number of resources for data transmission by 10%. On the other hand this gives a very robust performance at the cell edge or for high UE velocity.

One advantage of using URS is that pre-coding does not need to be quantified. As data and reference signals are using the same pre-coding matrix, the combined channel can be directly estimated from the UE specific reference signals and then used for demodulating the signal.

Due to the rather large overhead of URS, this mode is mainly expected to be used when Node B deploys more than four antennas. The Release 8 specifications do not allow for more than four different common reference signal ports, so in this case the only way to acheive UE specific pre-coding (beamforming) is to use URS.

An issue which is vendor specific and where eNodeB implementations may differ is how antennas are deployed and pre-coding is determined when URS are used. Two different scenarios will be discussed here. In the typical macro cell with antennas mounted above the roof-top, the azimuth spread of the channel would be low and the optimal solution would be to have antennas with narrow spacing and use traditional angular beamforming to create a narrow beam which directs the power towards the UE while reducing the interference to other UEs. In this case, the system only relies on estimating the Direction of Arrival (DoA) from the UL transmission and this can be done without relying on channel reciprocity. Moreover, the standard allows for reuse of time and frequency resources by other users in the same sector. Sufficient angular separation should be ensured to maintain performance.

In a scenario with many scatterers around the base station, the azimuth spread would be larger and angular beamforming might not work very well. Another solution could then be to rely on pre-coding determined from the eigen-vectors of the complex channel matrix determined from UL sounding. When the azimuth spread is increased, the rank of the channel could become larger than one and UE could potentially benefit from dual stream transmission. Although dual stream transmission is not supported in Release 8, it is a potential candidate for future releases. For more details on MIMO, beamforming and channel modeling see [4].

12.6 LTE TDD Performance

In this section the performance for LTE TDD is analyzed. For FDD mode, extensive analysis of performance has already been addressed in Chapter 9. As many of the observations provided there apply equally well to TDD mode, in the following we try to focus on the areas where TDD mode is different from FDD mode. First we look at the link performance, i.e. how well the receivers in eNodeB and UE can decode the physical channels. In general this is an area where there is little difference between FDD and TDD as reference signal patterns and channel coding are duplexing mode agnostic. After link performance, we discuss the link budget. Link budget for TDD is different from FDD because of the discontinuous transmission, so coverage for different bit rates in TDD will invariably be worse than for FDD. However, there are still a number of details to pay attention to in order to evaluate the TDD link budget properly.

This section ends with a discussion on system performance. First we look at the best effort type of service where a fairly small number of users per sector are assumed to download large amounts of data. Assuming that we are designing a network to deliver a certain bit rate for a certain number of active users, then the most important difference between networks based on FDD and TDD technologies is that UL and DL in TDD mode would need double the system bandwidth compared to FDD. Data transmission would also need to be carried out with a larger

bandwidth in the TDD system to achieve bit rates similar to FDD. Both system and transmission bandwidths affect the operation of RRM algorithms and in this the impact to the system performance is analyzed.

Performance assuming VoIP service is also evaluated and this gives a somewhat different perspective on the system performance as bit rates are low and users are many when we load the system with 100% VoIP traffic. As VoIP is a symmetric service, TDD systems can have a higher VoIP capacity than FDD because the split between UL and DL resources can be adjusted. The increased HARQ round trip time for the TDD system is also shown to have some effect on the VoIP performance where many UEs are coverage limited.

12.6.1 Link Performance

The two most important factors impacting the link performance for data transmission are channel estimation and channel coding. As reference signal design and channel coding are very similar for TDD and FDD, the link performance is also very similar. One source of difference is the discontinuous transmission in TDD. FDD receivers can use the reference signals from the previous subframe to estimate the channel. This is especially important for the DL link performance where the UE receiver should start decoding the control channel region as soon as it is received. In a TDD system, when a UE receiver decodes subframes transmitted right after UL→DL switching, it could not rely on reference signals from previous subframes and this could introduce some degradation to the channel estimation and thus coverage of control channels. The importance of such potential loss will depend on the UE implementation. If, for example, the UE could wait for the second column of reference signals then the performance degradation could be reduced.

Another potential difference in link performance between FDD and TDD is related to the special subframe. As explained earlier the guard period is created by adjusting the number of symbols in DwPTS. When the DwPTS length is reduced we also eliminate some reference signals as the rule is not to move them to new locations. The potential loss is minor as reference signals from a previous subframe could be taken into use to improve channel estimation. In the special case of UE specific reference signals we lose one column of reference signals even with full DwPTS length. In this case the UE cannot use reference signals from a previous subframe as these could have been transmitted with a different pre-coding.

Short RACH performance is clearly worse compared to 1 ms RACH preamble and thus should be used only in environments where link budget is not foreseen to be an issue.

12.6.2 Link Budget and Coverage for TDD System

The link budget calculation aims at estimating the range of different bit rates. A detailed description of how to calculate link budgets for LTE is already given in Chapter 9. Here we focus on the differences between link budgets for TDD and FDD modes. The differences relate mainly to the limited maximum UE transmit power and in the following we therefore focus our attention on UL link budgets.

The TDD UE cannot transmit continuously since the transmission must be switched off during the downlink reception, The UE will thus need to transmit with a larger bandwidth and a lower power density to achieve a similar bit rate to a FDD system. The lower power density is because the UE transmitter is limited on total maximum power, not on power per Hz.

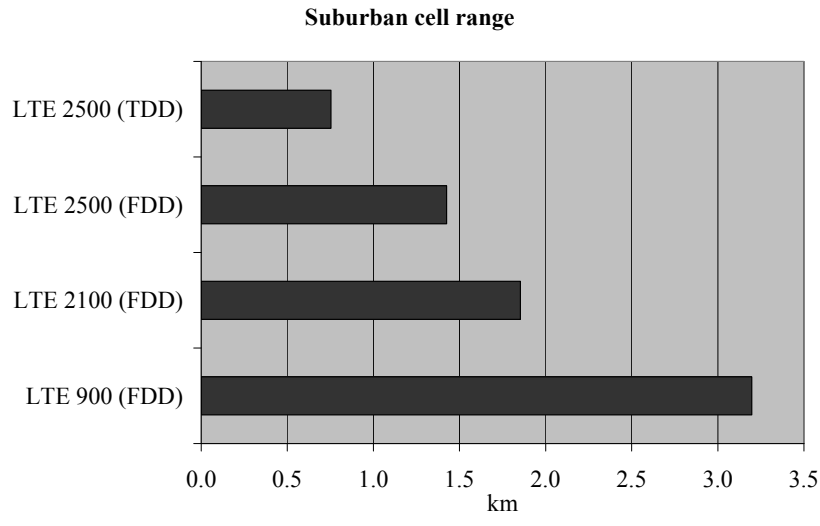


Figure 12.17 Uplink cell range for LTE FDD and TDD systems for 64 kbps

As a simple example, if the downlink : uplink share is 3 : 2, the UE transmission power density is reduced by $10 \times \log10(2/5) = -4$ dB as we need roughly 5/2 times the bandwidth for the TDD UL transmission. Another way of viewing this is that at a fixed distance from the base station the maximum achievable FDD bit rate will roughly be 2½ times larger than the bit rate achieved with maximum UE transmit power in a TDD system. Note that for DL, the power density can be assumed to be similar between FDD and TDD mode as the size of the power amplifier in eNodeB can be adapted to the system bandwidth.

So from a UL coverage perspective, FDD based systems do have an advantage over TDD systems due to the continuous transmission. Moreover, coverage with the TDD system can also be challenging because the TDD spectrum is typically situated at higher frequencies such as 2.3 GHz or 2.5 GHz. A cell range comparison for a suburban propagation environment is shown in Figure 12.17. The best coverage is obtained by using a FDD system at low frequency. The cell range for LTE900 FDD is four times larger (cell area 16 times larger) and LTE2500 FDD is 80% larger than LTE2500 TDD. The assumed data rate is 64 kbps and the cell range is calculated with the Okumura–Hata propagation model with 18 dB indoor penetration loss, 50 m base station antenna height and −5 dB correction factor. For maximum LTE coverage, LTE TDD deployment at high frequency could be combined with LTE FDD deployment at a lower frequency.

12.6.2.1 MCS Selection and UE Transmission Bandwidth for Coverage

For a certain target bit rate different combinations of MCS and transmission bandwidth have different coverage and different spectral efficiency. From Shannon's information theory [5] we know that if we want to maximize coverage under a fixed total transmission power constraint we should increase MCS when physical layer spectral efficiency (SE) is < 1 bit/s/Hz and increase bandwidth when SE is > 1 b/s/Hz. As adjusting BW does not impact SE, but increasing MCS does, the link should be operated with an MCS which achieves a SE of 1 bit/s//Hz (for LTE QPSK 2/3 would do the job). Then bandwidth can be adjusted to achieve the required bit rate with optimal coverage.

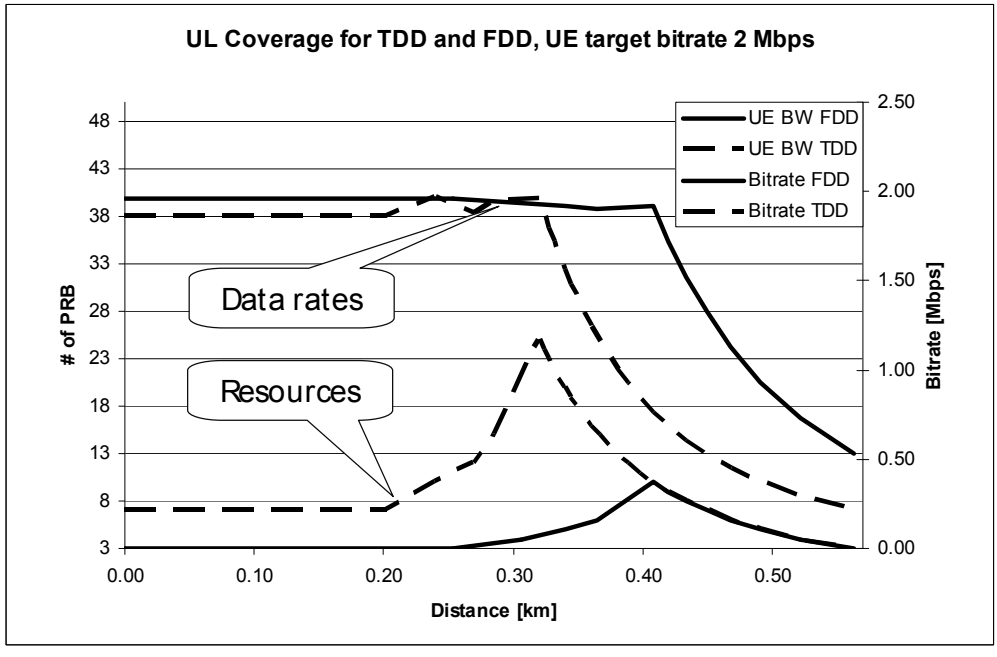


Figure 12.18 Coverage and required number of physical resource blocks for a 2 Mbps target bit rate for FDD and TDD UL. MCS in coverage limited region is QPSK ¾

In Figure 12.18 an example of this process is given. A UE at the cell edge selects MCS QPSK ¾ and sets the transmission bandwidth according to the available transmit power and required SINR. Curves for both TDD and FDD UE are shown and assumptions are as given in [6]. From the figure we can see that when the UE moves towards the Node B, path loss is reduced and the link gain can be used to increase the UE transmission bandwidth. As the transmission bandwidth increases, the UE bit rate also increases. When the UE target bit rate is reached, and if the UE path loss is reduced further, we can start to increase MCS and reduce bandwidth to improve the SE while keeping the bit rate on target. At some point the maximum MCS is reached and only at this point can we start to reduce the UE total transmit power. While the TDD system supports 2 Mbps in UL 300 m from the eNodeB, FDD system increases the coverage for 2 Mbps to 400 m. For more details on the interaction of UE transmit power, MCS and transmission bandwidth see Chapter 9.

12.6.2.2 Coverage for Low Bit Rates

When the data bit rate decreases, the relative overhead from header and CRC will increase. The effect of this is that even in FDD mode it does not make sense to schedule UEs in coverage problems with very narrow bandwidth and low MCS. An example for a data bit rate of 64 kbps is illustrated in Table 12.3. From this we can see that due to excessive overhead, the UE bandwidth increase for UL in a TDD 3DL/2UL configuration is only a factor 1.6, corresponding to a transmit power loss of 2 dB, not 4 dB as in the example given above where overhead was not taken into account.

Table 12.3 Required UE transmission bandwidth to support 64 kbps for TDD and FDD

System	TDD UL with 3DL/2UL	FDD UL
Service bit rate (kbps)	64	64
MCS	QPSK 1/10	QPSK 1/10
Data bits per TTI (bits)	160	64
Header (3 byte)	48	48
CRC (3 byte) (bits)		
Total bits per TTI (bits)	208	112
Payload per PRB in physical layer (2 symbols for DM RS 1 symbol for SRS)	27 bits	27 bits
Required number of PRBs	8	5

Table 12.4 Required SINR for VoIP in TDD and FDD with and without TTI bundling. UE transmission bandwidth assumed to be 360 kHz (2 PRB)

System	TDD (3DL/2UL or 2DL/3UL)		FDD	
Bundling enabled	No	Yes	No	Yes
Number of transmissions	5	8	7	12
Required SINR	−3.3 dB	−5.3 dB	−4.7 dB	−7.04 dB

One way of reducing the overhead for low bit rates is to use TTI bundling, as described in section 12.3.3. When TTI bundling is enabled the maximum number of retransmissions within a certain time limit is increased. In that way we can have more aggressive MCS selection and thus lower the relative overhead from protocol header and CRC. Due to the reduced number of UL TTIs in TDD, the potential link budget improvement from TTI bundling is not as important as for FDD. As shown in Table 12.4, for a VoIP service in 3DL/2UL configuration we can improve the number of transmissions of one VoIP packet within 50 ms from 5 to 8 TTIs, which is a 2 dB improvement to the link budget. We note that the link budget gain from TTI bundling in TDD mode is similar in both 2DL/3UL and 3DL/2ULconfiguration.

Finally, when operating at very low UL bit rates far away from the base station, coverage on UL control channels could also become a limiting factor. The control channel for TDD has been designed so that if ACK/NACK bundling mode is selected then the required receiver sensitivity is similar between TDD and FDD.

12.6.3 System Level Performance

TDD mode is in many aspects similar to FDD and this is also valid when we analyze the systems from a performance point of view. In general, when we compare time and frequency duplexing there are some differences related to the different ways of using the spectrum which make it difficult to make a 100% fair comparison of spectral efficiencies. Whereas a TDD mode system needs a guard period between UL and DL, a FDD system needs a large separation in frequency. Secondly if TDD systems have the same partition of UL and DL resources they can operate in adjacent bands; if not they need to be separated in a similar way to UL and DL for FDD.

Another spectrum related issue is that for the TDD system to provide a similar capacity to a FDD system the DL and UL system bandwidth needs to be double that of a FDD system. This impacts the RRM and users will typically need to operate with larger transmission bandwidths. That can be challenging for UL data transmission due to the limited transmission power of the UE.

One advantage for the TDD RRM solution is the possibility of exploiting channel reciprocity. In the current RRM framework (see Chapter 8) two parallel mechanisms are available for obtaining channel state information. For DL, the UE can be configured to feedback CQI, PMI and RI reports based on measurements of DL reference signals. In UL the UE can transmit SRSs so that Node B can measure the radio channel. For TDD mode, when channel reciprocity is present we ideally need only one of these mechanisms as the DL channel state can be inferred from the UL channel state or inversely. As mentioned earlier, there are challenges before this could work in a practical RRM solution, such as differences in UL/DL interference levels, different UL/DL antenna configurations and lack of UL/DL radio chain calibration, but on the other hand gains could be important. UL sounding, for example, can take up more than 10% of the UL system capacity. See section 12.2.4 for a further discussion of channel reciprocity.

12.6.3.1 Round Trip Time for TDD Systems

Feedback control loops are used for quite a few purposes in the LTE system. As TDD systems do not have continuous transmission and reception we might expect that the round trip time for such control loops would be increased for TDD systems, potentially degrading the system performance. However, due to the need for processing time, i.e. time for the UE or the Node B to decode and encode the control information, the typical round trip times between TDD and FDD are quite similar. Since the resulting delays are usually about 10 ms compared to 8 ms for the FDD RTT, the impact of the TDD frame structure is rather low and unlikely to impact on the performance of TCP/IP regardless of being run over LTE TDD or FDD.

12.6.3.2 Scheduling

One of the key LTE RRM features is channel aware scheduling, which is available for both UL and DL. Under the right conditions this feature can bring gains in spectral efficiency of up to 50% with even more important improvements to the coverage. To maximize the scheduling gain it is important to have frequency selective channel knowledge and flexibility in the control signaling to allocate UEs to the optimal frequency resources. Obtaining detailed frequency selective channel state information and enabling flexible resource allocation in the frequency domain is very costly in terms of control signaling.

For DL a number of different frequency resource allocation schemes are specified in the standard. For channel aware scheduling, the most effective allocation scheme specified gives a bit mask where each bit corresponds to a sub-band consisting of a number of continuous PRBs. To obtain the channel state for different sub-bands the CQI report contains a similar bit mask indicating the best M sub-bands and one indication of supported MCS on these sub-bands. The number of PRBs in one sub-band and the value of M are fixed in the specification. As the size of the bit mask increases with the system bandwidth, it was decided to increase sub-band size as the system bandwidth increases. If we compare a 10 MHz FDD DL and a 20 MHz TDD DL the scheduling gain is most likely to have some loss for 20 MHz especially for channels with a

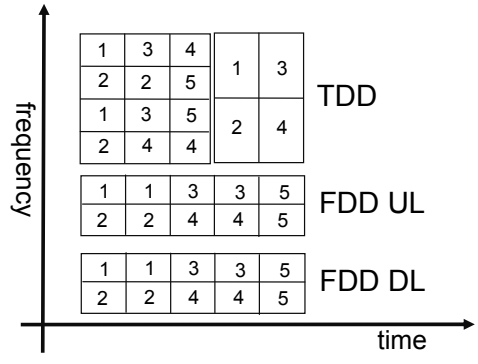


Figure 12.19 TDD and FDD resource allocations in LTE. Due to limitations on the number of users scheduled per TTI, TDD mode UEs are forced to transmit at a higher bandwidth

narrow frequency correlation bandwidth. The increase in system bandwidth could on the other hand give a TDD system an advantage due to the increased frequency diversity.

For UL due to the single carrier constraint, frequency resources have to be allocated contiguously; so here there is no difference for the resource allocation signaling between a lower and higher system bandwidth. For obtaining channel state information, we can enable SRS transmission from the UE. Sounding takes up resources and in contrast to the situation for DL, the larger the bandwidth that needs to be sounded, the more resources are needed. On the other hand it is not necessary to know the channel state information over the full bandwidth to obtain channel aware scheduling gain.

The resource allocation signaling consumes a lot of DL control channel resources and as these are limited the typical number of users that can be scheduled in one TTI is 10–20 (5–10 UL users and 5–10 DL users). As we assume that system bandwidth in TDD is double that of FDD, the TDD UE needs to operate at double bandwidth to have full bandwidth use in the system (see Figure 12.19). The increase in UE bandwidth can especially impact the UL channel aware scheduling gain. As the frequency resources are forced to be allocated contiguously, it is difficult to allocate resources according to channel quality in the frequency domain.

The total amount of DL control channel resources varies with the UL/DL ratio as up to three symbols can be reserved in each DL subframe (two for DwPTS). 2DL/3UL TDD potentially has fewer resources than in FDD while for other configurations there are more. The common case of 3DL/2UL has slightly more DL control channel resources than its FDD counterpart (8 symbols in 20 MHz for TDD compared to 15 symbols in 10 MHz for FDD) so there are no important differences between FDD and TDD. It should be mentioned that the TDD allocation messages (DCI) have a slightly higher payload, both due to higher system bandwidth but also due to additional bits for special features such as multi TTI and ACK/NACK bundling.

12.6.3.3 Uplink Power Control

The uplink power control is another very important RRM function as it has high impact on the distribution of bit rates in the sector. Moreover, fine tuning of power control parameters can improve both cell spectral efficiency and coverage. As we have discussed earlier the typical TDD UE will need to transmit at a higher bandwidth compared to an FDD UE. This puts extra

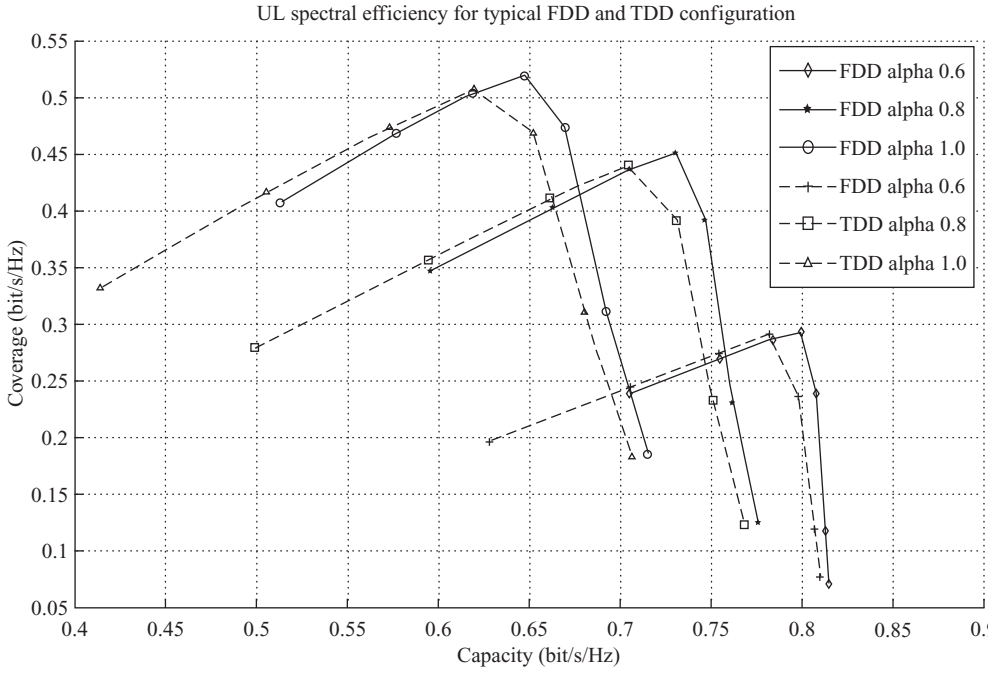


Figure 12.20 Uplink spectral efficiency for TDD and FDD systems with different power control settings. See [6] for assumptions

pressure on uplink transmit power resources, which are already limited, and can lead to poorer coverage for TDD systems. However, where interference limits the system performance, the performance difference is expected to be minor. In Figure 12.20 an uplink spectral efficiency comparison between a 20 MHz TDD system and a 2×10 MHz FDD system is shown where the TDD UEs transmit with a bandwidth of 12 PRB and FDD PRB with 6 PRB. From this figure we can see that both cell and coverage performance are very similar but slightly better for the FDD system. This difference is caused by reduced channel aware scheduling gain due to the higher bandwidth. Note that the optimal uplink interference level in the system depends on the bandwidth; for 12 PRB UE bandwidth it is 3 dB lower than for 6 PRB independent of the value of alpha. When site density decreases, the coverage performance of the TDD system will decrease faster than the FDD system and we will eventually see the effect from the differences in the link budget.

12.6.3.4 Downlink HARQ ACK/NACK Bundling

Finally we mention the effect on performance of the downlink HARQ ACK/NACK bundling feature described in section 12.3.5. As only one ACK/NACK is available for multiple subframes and these subframes are from different HARQ processes at different numbers of retransmissions then there is a risk that excessive retransmission will occur. If we consider particularly the very downlink heavy TDD configurations with aggressive MCS selection, then the resulting high retransmission probability will drag down the downlink system performance.

This effect can be offset by operating HARQ with relatively low BLER for the first transmission; however, this means that downlink HARQ will operate at a suboptimal operating point. Another way to reduce excessive retransmission from ACK/NACK bundling is to schedule users less often but with higher bandwidth. This on the other hand could reduce channel aware scheduling gain. In the most common TDD configurations with 3 downlink and 2 uplink subframes the impact from ACK/NACK bundling will be minor as a maximum of 2 ACK/NACKs will be bundled.

Simulations show that the loss in the case of 4 downlink/1 uplink is less than 10%. However, for users in coverage problems which rely on HARQ for the link budget, the excessive retransmission could give more important performance problems as one NACK typically will trigger retransmission of multiple TTIs.

12.6.3.5 Performance of VoIP

Voice over Internet protocol (VoIP) traffic should be supported as efficiently as possible in LTE TDD mode. However, supporting VoIP in LTE TDD faces very similar challenges to the FDD mode:

1 tight delay requirements with much longer HARQ RTT;
2 various control channel restrictions for different uplink/downlink configurations; and
3 even more serious coverage problems in uplink due to discontinuous transmission – as there are multiple uplink/downlink configurations possible, this can be used to achieve higher total VoIP capacity.

The following part presents the system level performance of VoIP traffic in LTE TDD mode at 5 MHz system bandwidth. The capacity evaluation criterion is the same as FDD: defined as the maximum number of per sector VoIP users that can be supported without exceeding a 5% outage level. VoIP capacity numbers are obtained from system level simulations in macro cell scenario 1 and macro cell scenario 3 [2], the main system simulation parameters being aligned with [7].

Similar to FDD in Chapter 10, the simulation results of VoIP capacity are summarized in Table 12.5 and Table 12.6 for two different AMR codecs, AMR 7.95 and AMR 12.2, and for

Table 12.5 VoIP capacity in LTE TDD at 5 MHz for macro case 1

	Configuration 0 (2DL/3UL)		Configuration 1 (3DL/2UL)	
VoIP codec	AMR 12.2	AMR 7.95	AMR 12.2	AMR 7.95
Downlink capacity				
Dynamic scheduler, without packet bundling	64	64	112	112
Dynamic scheduler, with packet bundling	114	122	194	206
Semi-persistent scheduler	102	140	168	220
Uplink capacity				
Dynamic scheduler	88	92	88	114
Semi-persistent scheduler	134	174	86	110

Table 12.6 VoIP capacity in LTE TDD at 5 MHz for macro case 3

	Configuration 0 (2DL/3UL)		Configuration 1 (3DL/2UL)	
Downlink capacity				
VoIP codec	AMR 12.2	AMR 7.95	AMR 12.2	AMR 7.95
Dynamic scheduler, without packet bundling	54	54	94	94
Dynamic scheduler, with packet bundling	90	94	158	168
Semi-persistent scheduler	84	114	144	194
Uplink capacity				
VoIP codec	AMR 12.2		AMR 12.2	
Dynamic scheduler, without TTI bundling	<40 (50 ms delay bound)		<40 (50 ms delay bound)	
Dynamic scheduler, with TTI bundling	<40 (50 ms delay budget)		<40 (50 ms delay budget)	
	50 (60 ms delay budget)		50 (60 ms delay budget)	
	72 (70 ms delay budget)		54 (70 ms delay budget)	

both dynamic and semi-persistent schedulers and for both macro case 1 and macro case 3 as defined in [6].

For both UL and DL simulations, the number of control symbols used in the special subframe is set to two, which contains six Control Channel Elements (CCEs). In a normal DL subframe, there are three control symbols available, which provide ten CCEs for UL and DL control.

For DL VoIP capacity for LTE TDD, performance of fully dynamic PS is seriously control channel limited if packet bundling is not allowed for both simulation scenarios. However, with packet bundling the users having good CQI are scheduled less often and hence the capacity can be boosted up to 70–90% because more control channel resources are released to be used by other users. There is no difference in the packet bundling concept between FDD and TDD. The reader is referred to Chapter 10 for more information about packet bundling. Semi-persistent PS performs very well in control channel limited situations, and always outperforms dynamic PS if packet bundling is not allowed. Fully dynamic PS outperforms semi-persistent PS only if packet bundling is allowed and the VoIP packet payload is high enough (e.g. AMR12.2 or higher) so that control channel limitations can be avoided. The performance difference between scenario 1 and scenario 3 is up to 15–20%, which is not so large because DL transmission power is not so limited even in scenario 3.

For UL VoIP capacity for LTE TDD, configuration 0 is regarded as seriously control channel limited, so the performance of dynamic scheduling is quite poor due to the limited number of control channels available. Performance loss reaches 40–80% compared to semi-persistent scheduling for different AMR codecs. In contrast, configuration 1 is a loose control channel case, so the performance of dynamic scheduling is slightly better than semi-persistent scheduling due to flexible retransmission and frequency domain packet scheduling (FDPS) gain. The performance difference between macro case 1 and macro case 3 at over 50% is a huge loss. The loss is very serious because macro case 3 is a very coverage-limited scenario. The UE transmit power is not large enough to even provide a proper link budget. For macro case 3, TTI bundling and longer delay bound can provide great gains on performance. For example, by using TTI bundling in a 70 ms delay budget for TDD configuration 0, up to 3 dB of energy accumulation

gain can be achieved. However, without TTI bundling, it is hard to achieve good VoIP capacity (i.e. <40 users/sector) because of the deficient energy accumulation.

To carry out a fair comparison between FDD and TDD, the FDD capacity results should be scaled for different TDD configurations and scheduling options. In DL, the scaling factor used for the dynamic scheduler should depend on the control symbol assumptions because it has control overhead limited performance for dynamic scheduling. With two control symbols, special subframe only has six CCEs available. Hence FDD results for the dynamic scheduler should be scaled with the scaling factor 26/50 = 0.52 for configuration 1 and 16/50 = 0.32 for configuration 0 in DL. However, in DL the scaling factor used for SPS should depend on the number of data symbol assumptions because it has data limited performance for SPS. With two control symbols, the special subframe only has seven data symbols available. Hence, FDD results for SPS should be scaled with the scaling factor 27/50 = 0.54 for configuration 1 and 17/50 = 0.34 for configuration 0.

In UL, there are no data symbols available in the special subframe, only UL control signaling, so we just use the scaling factors 0.6 and 0.4 for configuration 0 and configuration 1 simply in terms of the UL data symbol ratio.

Compared with the scaled FDD results in Chapter 10, we make the following observations. In downlink macro case 1, TDD results for both configurations are at most 5% lower than the scaled FDD results. For dynamic scheduling in the coverage limited macro case 3, the average amount of transmissions per packet is increased. As scheduling flexibility in the time domain is decreased for TDD and HARQ RTT is increased, the scaled TDD performances are about 10% lower than the scaled FDD results – losses are naturally higher for configuration 0 due to the small number of DL subframes.

In UL TDD configuration 0 of macro case 1, FDD has about a 7–10% gain over TDD for semi-persistent scheduling due to more diversity gain in FDD. Moreover, FDD has up to 40% gain over TDD for dynamic scheduling due to the limited control channel resources in TDD. For uplink TDD configuration 1 of macro case 1, TDD has 5–20% gain for AMR 12.2 and AMR 7.95 respectively over FDD for dynamic scheduling due to loose control channel limitation. FDD has ~10% gain for AMR 12.2 and AMR 7.95 respectively over TDD for semi-persistent scheduling due to more diversity.

For macro case 3 with TTI bundling and with a slightly longer packet delay budget (i.e. 70 ms), the difference between FDD and TDD reaches 15–25% for different TDD configurations because there is less energy accumulation in the TDD mode. However, with the shorter delay budget (i.e. 50 ms), TTI bundling is not able to provide visible capacity improvement although 2 dB energy accumulation gain is achieved compared to no bundling. Table 12.7 shows the energy accumulation with a 50 ms packet delay budget. With TTI bundling, a 1.76 dB energy accumulation gain can be achieved in FDD compared to the value in TDD. So the improvement for VoIP capacity is obvious.

Table 12.7 Impact from energy accumulation with TTI bundling

	Configuration 0 (2DL/3UL)	Configuration 1 (3DL/2UL)	FDD
Without bundling	5	5	7
With bundling	8	8 (for most of packets)	12

12.6.4 Evolution of LTE TDD

As mentioned earlier, TDD provides a very efficient way of enabling large peak data rates in unpaired spectrum conditions or when paired spectra with significant bandwidth or duplexing distance can be found. As LTE progresses towards LTE Advanced, introduced in Chapter 2, local area deployment and the outlook for reaching 1 Gbit/s data rates will become a key driver in further developing the TDD mode. In the path towards unleashing the transmission over a very wide bandwidth (e.g. multiple component carriers), TDD poses some interesting challenges for scalability while ensuring backwards compatibility for a smooth evolution.

Another important aspect related to TDD is that of coexistence. Based on Release 8 it seems that synchronization is the key to control interference and provide a reliable system operation in a network with full coverage for UE. However, for indoor or low-cost deployments, the existing techniques for synchronization may become too expensive and hence techniques for more lean synchronization, e.g. Over The Air (OTA) synchronization, could be of interest.

An aspect related to TDD is that of providing dynamic switching of the uplink/downlink configuration depending on the required capacity in the network. It is not yet known if the 3GPP LTE TDD standard will in the future provide a direct means for multiple operators to coexist within the same geographical area or if such coexistence will be guaranteed mainly by means of regulation. A further step in this direction is to look at protocols for slowly or semi-statically modifying the TDD configuration according to the varying load over time either within a coverage area covered by a single operator or in an area with multiple operators, in which case semi-static TDD adaptation would need to be considered together with schemes for flexible spectrum usage and sharing. Finally, it has yet to be seen if fast dynamic TDD for capturing the instantaneous load in a single cell will become possible in practice.

Finally, given the experience with the first deployments it will be interesting to follow to what extent reciprocity can be assumed to be available taking the practical antenna aspects into account. There are several ways in which the LTE TDD standard could be further optimized towards more efficient signaling (and thus performance, given fixed signaling overhead budget) in the case of reciprocity including topics coming as part of the LTE-Advanced work in upcoming 3GPP releases.

12.7 Summary

The LTE TDD operation is very similar to LTE FDD, including exactly the same multiple access solutions, channel encoding, etc. From the 3GPP specifications perspective the close similarity between FDD and TDD is also demonstrated via a single set of specifications for the physical layer. The resulting differences are due to the uplink/downlink split inherent for the TDD operation and cannot be really avoided as in FDD one does not need to change transmission direction but frames are continuous. Still as the resource allocation was 1 ms resolution both in FDD and TDD, the slots, excluding the ones when transmission direction changes, have the same parameters and timing. The sharing of the same spectrum for uplink and downlink gives flexibility for data asymmetry and makes finding new spectrum allocations easier for IMT-Advanced which needs to reach the 1 Gbit/s target. On the other hand the nature of the TDD operation sets the requirements for using a synchronized network to avoid uplink/downlink interference as well as resulting in a shorter range than with a FDD configuration. The LTE TDD mode (or TD-LTE) deployment has also been designed to accommodate coexistence with TD-SCDMA by ensuring compatible parameterization for uplink/downlink split when

sharing the same sites and antennas or for other close proximity installations. The spectrum availability for the TDD operation varies on a country basis, but in China the spectrum has been allocated and also in Europe an additional TDD spectrum (on top of the older allocations like 1900 to 1920 MHz) is being allocated from 2.5 GHz in several countries (with first licensing activities already completed). Also in the 3.5 GHz spectrum there have been recent allocations in certain countries with some unpaired parts, as in some countries there are existing deployments on that band for fixed mobile broadband use. Further development of LTE TDD mode proceeds together with the LTE FDD mode towards the LTE-Advanced to meet the ITU-R IMT-Advanced requirements.

References

[1] 3GPP Technical Specification, TS 36.211 V8.4.0, September 2008.
[2] 3GPP Technical Report, TR 25.814, 'Physical Layer Aspects for Evolved Universal Terrestrial Radio Access (UTRA)', 3GPP TSG RAN, September 2006.
[3] 3GPP Technical Specification, TS 36.213 V8.4.0 September 2008.
[4] C. Oestges, B. Clerckx, 'MIMO Wireless Communications', 1st edition, Academic Press, 2007.
[5] J.G. Proakis, 'Digital Communications', 3rd edition, McGraw-Hill Book Co., 1995.
[6] CATT, 'Enhanced Beamforming Technique for LTE-A', 3GPP T-doc R1-082972.
[7] 'Next Generation Mobile Networks (NGMN) Radio Access Performance Evaluation Methodology', A White paper by the NGMN Alliance, January 2008.

13

HSPA Evolution

Harri Holma, Karri Ranta-aho and Antti Toskala

13.1 Introduction

High Speed Packet Access (HSPA) was included in Third Generation Partnership Project (3GPP) Releases 5 and 6 for downlink and for uplink. The 3GPP Releases 7 and 8 have brought a number of HSPA enhancements providing major improvements to the end user performance and to the network efficiency. The work continues further in Release 9.

The HSPA evolution work has progressed in parallel to LTE work in 3GPP. HSPA evolution deployments in practice take place in parallel to LTE deployments. Many of the technical solutions in HSPA evolution and LTE are also similar. The overview of the HSPA evolution and LTE roles is illustrated in Figure 13.1. HSPA evolution is optimized for coexistence with WCDMA/HSPA supporting legacy Release 99 UEs on the same carrier and designed for a simple

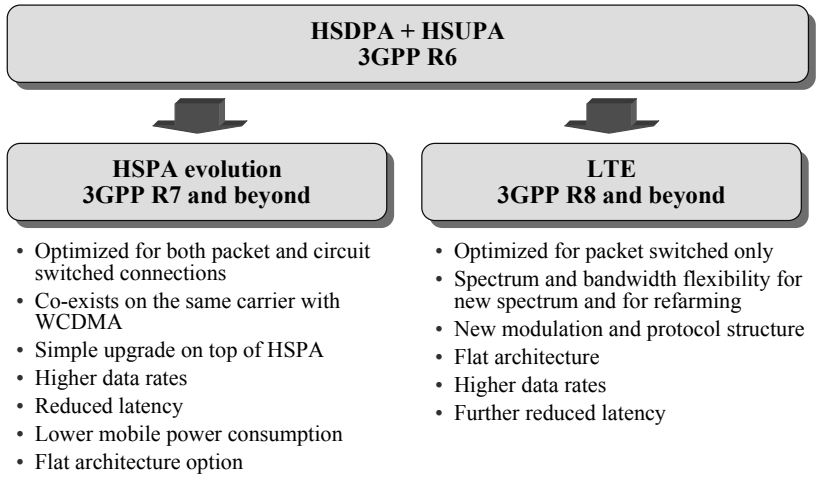


Figure 13.1 Overview of HSPA evolution and LTE roles

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala
© 2009 John Wiley & Sons, Ltd. ISBN: 978-0-470-99401-6

upgrade on top of HSPA. The HSPA evolution target is to improve end user performance by lower latency, lower power consumption and higher data rates. The HSPA evolution features are introduced in this chapter. The HSPA evolution is also known as HSPA+.

HSPA evolution also includes inter-working with LTE, which enables both packet handovers and voice handovers from LTE Voice over IP (VoIP) to HSPA circuit switched (CS) voice. The handovers are covered in Chapter 7 and the voice call continuity in Chapter 10.

13.2 Discontinuous Transmission and Reception (DTX/DRX)

Technology evolution in general helps to decrease the mobile terminal power consumption. Also, fast and accurate power control in WCDMA helps to minimize the transmitted power levels. The challenge in 3GPP from Release 99 to Release 6 is still the continuous reception and transmission when the mobile terminal is using HSDPA/HSUPA in Cell_DCH (Dedicated Channel) state. HSPA evolution introduces a few improvements to HSDPA/HSUPA that help to reduce the power consumption for CS voice calls and for all packet services.

With 3GPP Release 6 the UE keeps transmitting the physical control channel even if there is no data channel transmission. The control channel transmission and reception continues until the network commands the UE to Cell_FACH (Forward Access Channel) or Cell_PCH (Paging Channel) state. The Release 7 UE can cut off the control channel transmission as soon as there is no data channel transmission, allowing it to shut down the transmitter completely. This solution is called discontinuous uplink transmission and it brings clear savings in transmitter power consumption [1].

A similar concept is also introduced in the downlink where the UE needs to wake up only occasionally to check if the downlink data transmission is starting again. The UE can use power saving mode during other parts of the frame if there are no data to be received. This solution is called downlink discontinuous reception. The discontinuous transmission concept is illustrated in Figure 13.2 for web browsing. As soon as the web page is downloaded, the connection enters discontinuous transmission and reception.

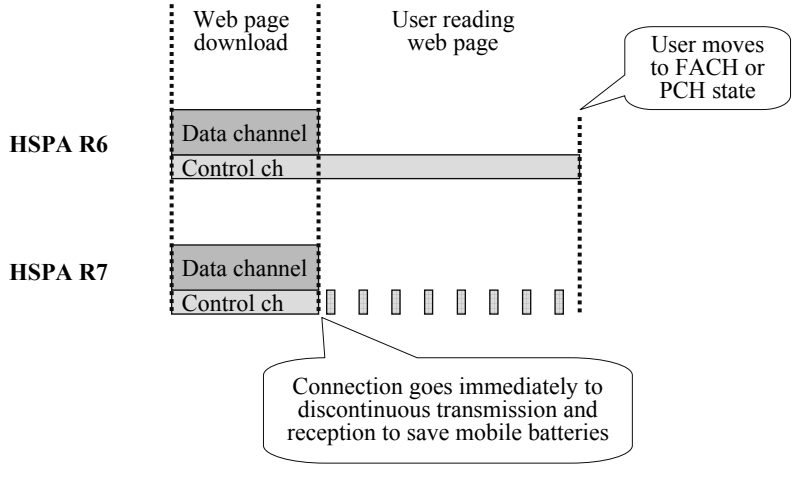


Figure 13.2 Discontinuous transmission and reception with continuous packet connectivity

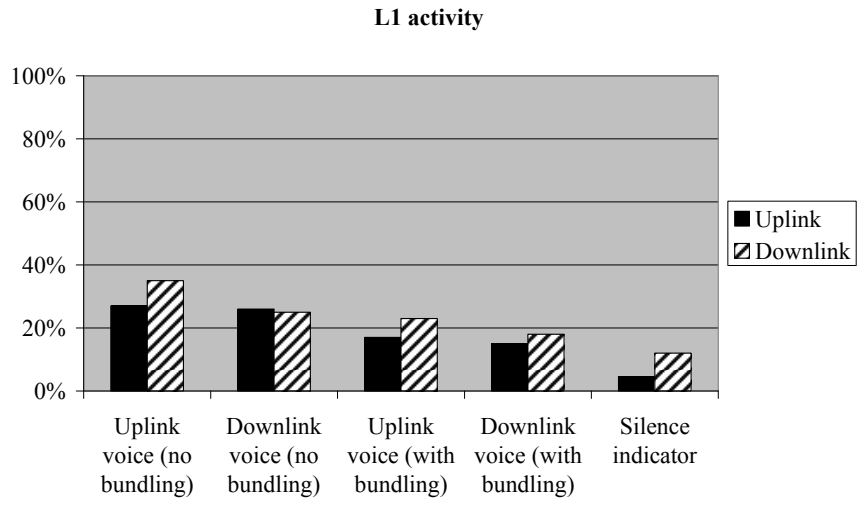


Figure 13.3 Layer 1 activity for voice call

Layer 1 (L1) activity with DTX/DRX for voice service is estimated in Figure 13.3. The packet bundling refers to the solution where two voice packets are transmitted together in the air interface halving the uplink active time as the UE needs to actually send data in the uplink only once every two voice packets arriving from the voice encoder. The activity is reduced from 100% in Release 6 to 20–40% in HSPA evolution. This activity reduction can translate into >50% longer voice talk time. The usage time improvement can be even more for bursty packet data connections.

The Release 99 FACH solution requires continuous reception by the UE, which is challenging from the power consumption point of view especially for always-on applications transmitting frequent keep alive messages. Each keep alive message forces the UE to move to the Cell_FACH state and stay there until the network inactivity timer expires. The discontinuous reception is also introduced for Cell_FACH state in HSPA evolution helping stand-by times when these types of applications are used.

Discontinuous transmission and reception is included in LTE from the beginning in Release 8 specifications. The power saving potential in LTE is even higher than in HSPA because the Transmission Time Interval (TTI) size is shorter than in HSPA (1 ms vs 2 ms) and because there is no need for fast power control related signaling.

13.3 Circuit Switched Voice on HSPA

Voice has remained an important service for the mobile operators. WCDMA Release 99 supports CS voice on Dedicated Channel (DCH) with quite high spectral efficiency. Efficient Voice over IP (VoIP) capability on top of HSPA was defined in Release 7, but VoIP mass market has not yet started. Therefore, CS voice over HSPA was also defined in HSPA evolution. CS voice over HSPA is part of Release 8 but because of the capability indication for the UE, support of the feature was introduced to Release 7, hence it is possible to implement this feature before other Release 8 features. There are two main benefits when running voice on HSPA: UE power

consumption is reduced because of DTX/DRX and the spectral efficiency is improved with HSPA features. Now these benefits can also be used for CS voice calls.

The different 3G voice options are illustrated in Figure 13.4. CS voice over DCH is used currently in commercial networks. Dedicated Release 99 channel is used in L1 and Transparent mode RLC in Layer 2 (L2). From the radio point of view, CS voice over HSPA and VoIP over HSPA use exactly the same L1 VoIP including unacknowledged mode RLC on L2. IP header compression is not needed for CS voice. From the core network point of view, there is again no difference between CS voice over DCH and CS voice over HSPA. In fact, CS core network is not aware if the radio maps CS voice on DCH or on HSPA. CS voice over HSPA could be described as CS voice from the core point of view and VoIP from the radio point of view.

CS voice over HSPA brought the following changes to 3GPP specifications:

- Iu interface = no changes
- Physical layer = no changes
- MAC layer = no changes
- RLC layer = forwarding RLC-UM sequence numbers to upper layers
- PDCP layer = modification of header to identify and timestamp the CS AMR frames + interfacing/inclusion of Jitter Buffer Management
- New de-jitter buffer = UE/RNC implementation dependent entity that absorbs the radio jitter created by HSPA operation so that the CS AMR frames can be delivered in a timely constant fashion to upper layers. The algorithm for the de-jitter buffer is not standardized in 3GPP.

The CS voice over HSPA concept is presented in Figure 13.5. The CS voice connection can be mapped on DCH or on HSPA depending on the UE capability and RNC algorithms. The AMR data rate adaptation can be controlled by RNC depending on the system loading. When CS voice over HSPA is used, there is a clear need for QoS differentiation in HSPA scheduling to guarantee low delays for voice packets also during the high packet data traffic load. Since the packet scheduling and the prioritization are similar for VoIP and for CS voice over HSPA, it will be simple from the radio perspective to add VoIP support later on top of CS voice over HSPA. Therefore, CS voice over HSPA is paving the way for future VoIP introduction. The similar radio solutions also make the handover simpler between VoIP and CS domains.

CS voice on HSPA can take advantage of IP protocol and packet transmissions in all interfaces: air interface carried by HSPA, Iub over IP, Iu-CS over IP and the backbone between

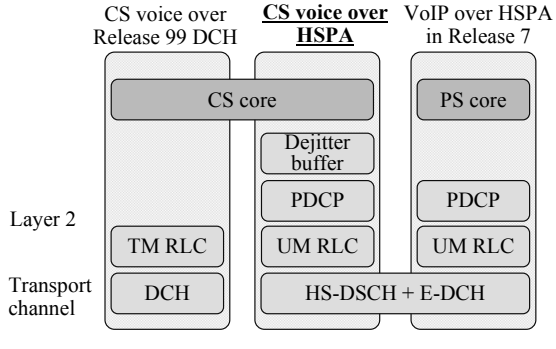


Figure 13.4 Voice options in WCDMA/HSPA

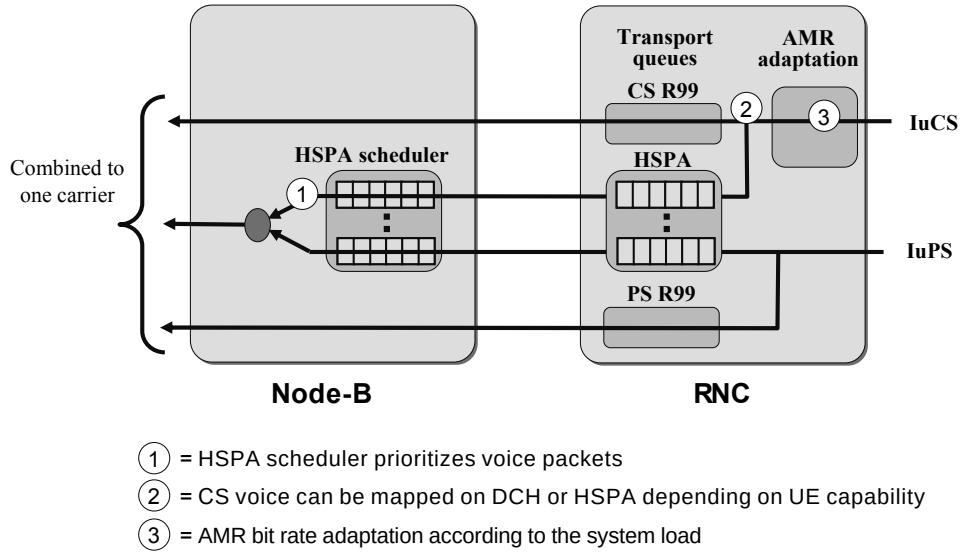


Figure 13.5 CS voice over HSPA overview

Media Gateways (MGW) using IP. The call control signaling is still based on CS protocols [2]. The use of the interfaces is shown in Figure 13.6. CS voice over HSPA can take advantage of the packet network performance and cost while maintaining the existing end-to-end protocols and ecosystem. No changes are required to charging, emergency calls or to roaming.

CS voice over HSPA also increases the spectral efficiency compared to voice over DCH because voice can also take advantage of the HSPA physical layer enhancements:

- UE equalizer increases downlink capacity. Equalizer is included in practice in all HSPA terminals [3].
- Optimized L1 control channel in HSPA reduces control channel overhead. The downlink solution is Fractional DPCH with discontinuous reception and the uplink solution is discontinuous transmission. Also transmission without HS-SCCH can be used on downlink.
- L1 retransmissions can also be used for voice on HSPA since the retransmission delay is only 14 ms.
- HSDPA optimized scheduling allows improvement of the capacity even if the tough delay requirements for voice limit the scheduling freedom compared to best effort data.

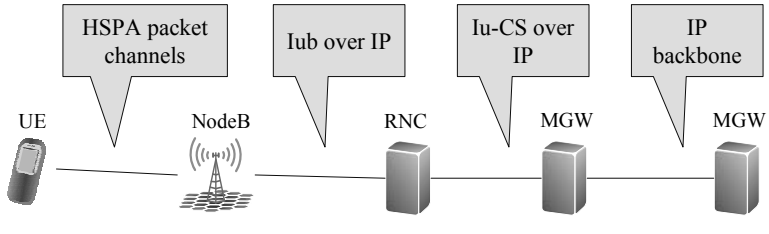


Figure 13.6 CS voice in different interfaces

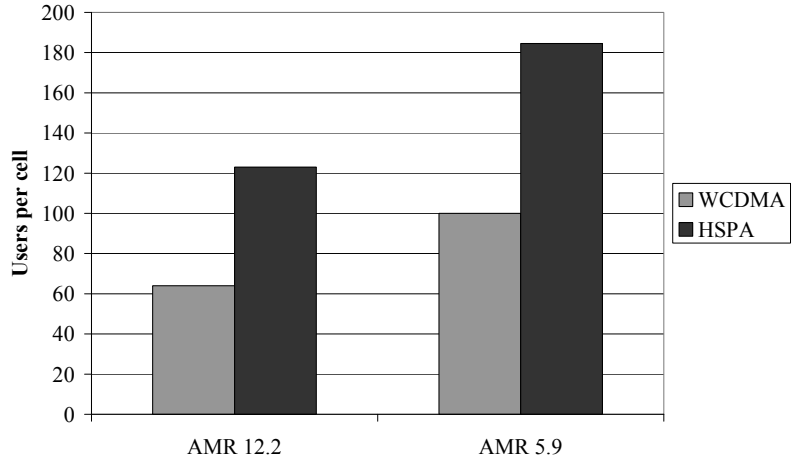


Figure 13.7 Circuit switched voice spectral efficiency with WCDMA and HSPA

The gain in spectral efficiency with CS voice over HSPA is estimated at 50–100% compared to CS voice over DCH. The voice capacity is illustrated in Figure 13.7. The voice capacity evolution including LTE is covered in more detail in Chapter 10.

The voice codec with CS over HSPA can be Narrowband Adaptive Multirate Codec (AMR) or Wideband AMR.

13.4 Enhanced FACH and RACH

WCDMA network data rate and latency are improved with the introduction of Release 5 HSDPA and Release 6 HSUPA. The end user performance can be further improved by minimizing the packet call setup time and the channel allocation time. The expected packet call setup time with Release 7 will be below 1 s. Once the packet call has been established, user data can flow on HSDPA/HSUPA in Cell_DCH state. When the data transmission is inactive for a few seconds, the UE is moved to the Cell_PCH state to minimize the mobile terminal power consumption. When there are more data to be sent or received, the mobile terminal is moved from Cell_PCH to Cell_FACH and to the Cell_DCH state. Release 99 Random Access Channel (RACH) and FACH can be used for signaling and for small amounts of user data. The RACH data rate is very low, typically below 10 kbps, limiting the use of the common channels. Release 5 or Release 6 do not provide any improvements in RACH or FACH performance. The idea in Release 7 Enhanced FACH and Release 8 Enhanced RACH is to use the Release 5 and Release 6 HSPA transport and physical channels also in the Cell_FACH state for improving the end user performance and system efficiency. The concept is illustrated in Figure 13.8. Enhanced FACH and RACH bring a few performance benefits:

- RACH and FACH data rates can be increased beyond 1 Mbps. The end user could get immediate access to relatively high data rates without the latency of channel allocation.
- The state transition from Cell_FACH to Cell_DCH would be practically seamless. Once the network resources for the channel allocation are available, a seamless transition can take place to Cell_DCH since the physical channel is not changed.

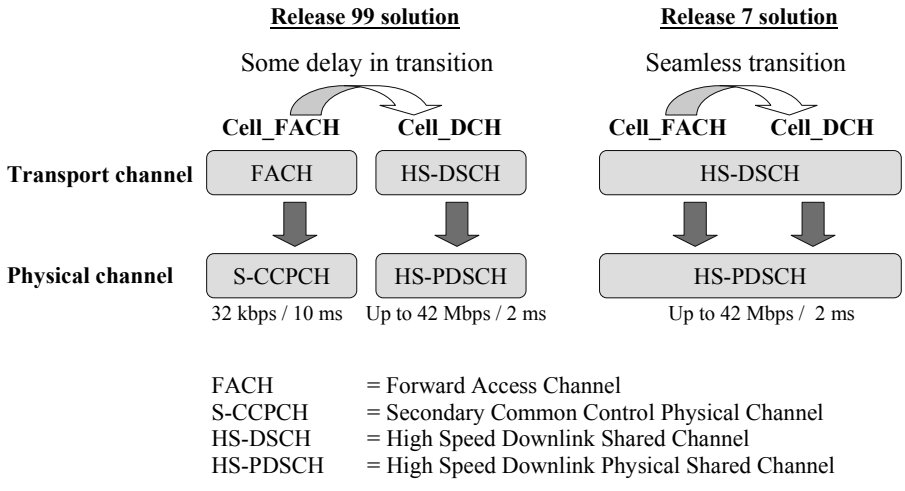


Figure 13.8 Enhanced FACH concept [4]

- Unnecessary state transitions to Cell_DCH can be avoided when more data can be transmitted in Cell_FACH state. Many applications create some background traffic that is today carried on Cell_DCH. Therefore, Enhanced RACH and FACH can reduce the channel element consumption in NodeB.
- Discontinuous reception could be used in Cell_FACH to reduce the power consumption. The discontinuous reception can be implemented since Enhanced FACH uses short 2 ms TTI instead of 10 ms as in Release 99. The discontinuous reception in Cell_FACH state is introduced in 3GPP Release 8.

Since the existing physical channels are used in Enhanced FACH, there are only minor changes in L1 specifications which allow fast implementation of the feature. Enhanced FACH can coexist with Release 99 and with HSDPA/HSUPA on the same carrier. No new power allocation is required for Enhanced FACH since the same HSDPA power allocation is used as for the existing HSDPA.

13.5 Downlink MIMO and 64QAM

The downlink peak data rate with Release 6 HSDPA is 10.8 Mbps with ¾ coding and 14.4 Mbps without any channel coding. There are a number of ways in theory to push the peak data rate higher: larger bandwidth, higher order modulation or multi-antenna transmission with Multiple Input Multiple Output (MIMO). All these solutions are part of HSPA evolution. MIMO and higher order modulation are included in HSPA evolution in Release 7 and Dual carrier (dual cell) HSDPA in Release 8. The 3GPP MIMO concept uses two transmit antennas in the base station and two receive antennas in the terminal and uses a closed loop feedback from the terminal for adjusting the transmit antenna weighting. The diagram of the MIMO transmission is shown in Figure 13.9.

Higher order modulation allows higher peak bit rate without increasing the transmission bandwidth. Release 6 supported Quadrature Phase Shift Keying (QPSK) and 16QAM (Quadrature Amplitude Modulation) transmission in the downlink and dual Binary Phase Shift

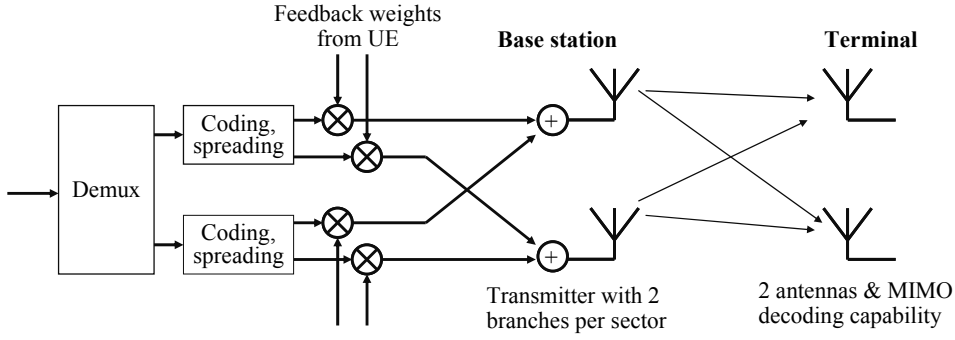


Figure 13.9 2×2 MIMO transmission concept

Keying (BPSK) in the uplink. Dual-channel BPSK modulation is similar to QPSK. Release 7 introduces 64QAM transmission for the downlink and 16QAM for the uplink. 16QAM can double the bit rate compared to QPSK by transmitting 4 bits instead of 2 bits per symbol. 64QAM can increase the peak bit rate by 50% compared to 16QAM since 64QAM transmits 6 bits with single symbol. On the other hand, the constellation points are closer to each other for the higher order modulation and the required signal-to-noise ratio for correct reception is higher. The difference in the required signal-to-noise ratio is approximately 6 dB between 16QAM and QPSK and also between 64QAM and 16QAM. Therefore, downlink 64QAM and uplink 16QAM can be used only when the channel conditions are favorable.

The system simulation results with 64QAM in macro cells are illustrated in Figure 13.10. The 64QAM modulation improves the user data rate with 10–25% probability depending on the scheduling (RR = round robin, PF = proportional fair). The rest of the time the channel conditions are not good enough to enable the reception of 64QAM modulation. The typical capacity gain from 64QAM is less than 10%.

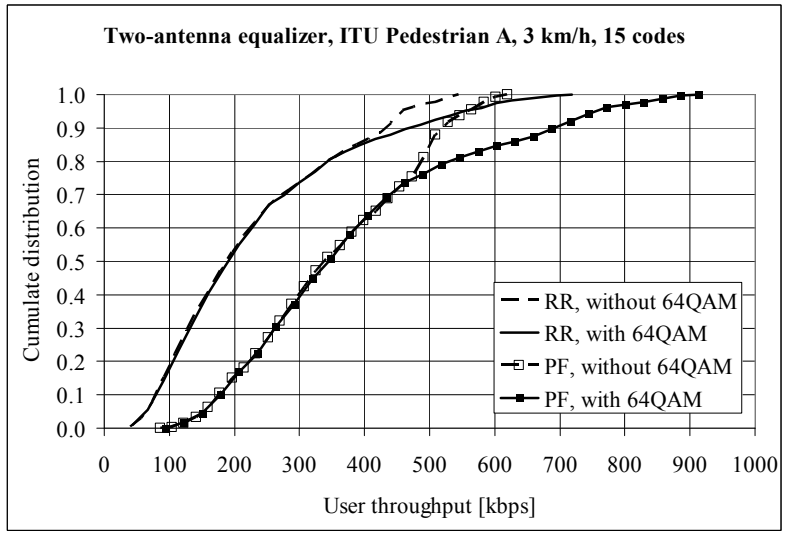


Figure 13.10 Macro cell data rates per user with 64QAM with 20 active users in a cell [5]

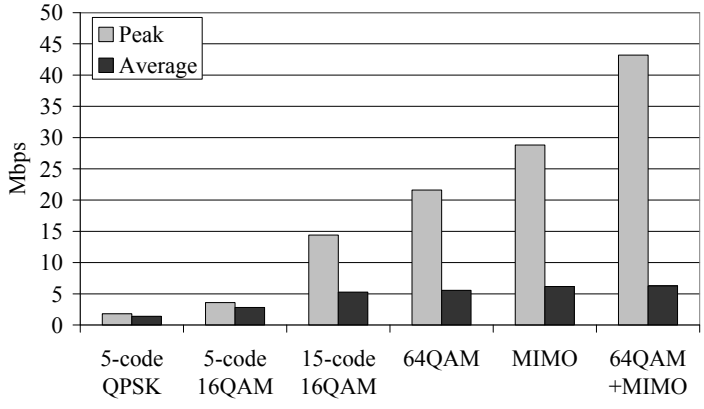


Figure 13.11 Downlink peak rate and average cell rate evolution with HSDPA features

64QAM and MIMO together improve the peak rate by 200% from 14 Mbps to 42 Mbps, but the average cell capacity is only improved by 20% since those high data rate features are useful only for part of the cell area. 64QAM activation can improve the cell capacity by approximately 5% and MIMO by 10%. The evolution of the peak and average rates is illustrated in Figure 13.11. The average rate corresponds to fully loaded macro cells. The practical networks are not always fully loaded, and therefore the gains of 64QAM and MIMO features can be higher in real networks than in fully loaded simulations.

13.6 Dual Carrier HSDPA

The LTE radio improves the data rates compared to HSPA because LTE can use a transmission bandwidth up to 20 MHz compared to 5 MHz in HSPA. The dual carrier (dual cell) HSDPA was specified as part of Release 8 enabling HSDPA to benefit from two adjacent HSDPA carriers in the transmission to a single terminal using a total 10 MHz downlink bandwidth. The uplink solution in Release 8 is still using single 5 MHz carrier. The concept is illustrated in Figure 13.12.

The benefit of the dual carrier for the user data rate is illustrated in Figure 13.13. The dual carrier can double the user data rate at low loading since the user can access the capacity of two carriers instead of just one. The relative benefit decreases when the loading increases. There is still some capacity benefit at high load due to frequency domain scheduling and due to dynamic balancing of the load if both carriers are not 100% loaded all the time. Node B scheduling can optimize the transmission between the two carriers based on the CQI reporting – obtaining partly similar gains as in LTE with frequency domain scheduling. The frequency domain scheduling gain in Dual carrier HSDPA is smaller than with LTE since the scheduling in Dual carrier HSDPA is done in 5 MHz blocks while LTE scheduling is done with 180 kHz resolution.

Both Dual carrier HSDPA and MIMO can boost HSDPA data rates. Those two solutions are compared in Table 13.1. Both solutions can provide the same peak rate of 42 Mbps with 64QAM modulation. MIMO can improve spectral efficiency due to two antenna transmission, while the dual carrier HSDPA brings some improvement to the high loaded case with frequency domain scheduling and a larger trunking gain. The dual carrier solution looks attractive because the data rate improvement is available over the whole cell area equally while MIMO improves the

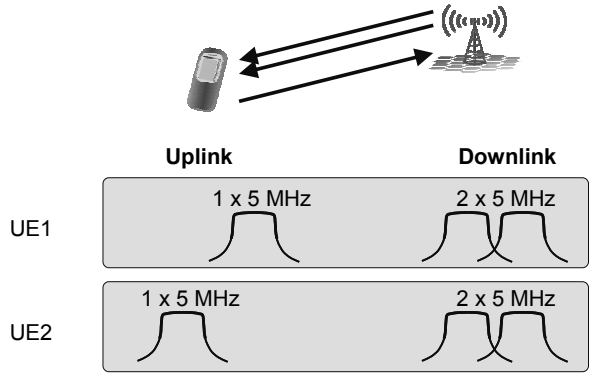


Figure 13.12 Downlink Dual carrier HSDPA concept

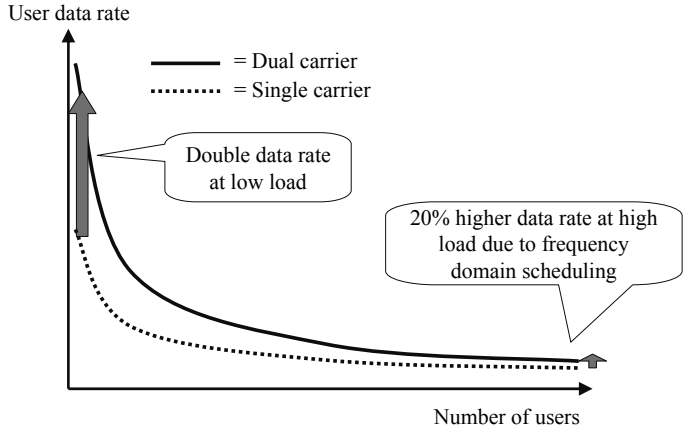


Figure 13.13 Data rate benefit of the dual carrier HSDPA

Table 13.1 Benchmarking of Dual carrier HSDPA and MIMO

	Dual carrier	MIMO
Peak bit rate	42 Mbps	42 Mbps
Spectral efficiency improvement	20% due to frequency domain scheduling and larger trunking gain	10% due to two antenna transmissions
Data rate gain	Similar gain over the whole cell area	Largest gain close to NodeB where dual stream transmission is feasible
NodeB RF requirements	Single power amplifier per sector	Two power amplifiers per sector
UE RF requirements	Possible with 1 antenna terminal	2 antennas required

data rates mostly close to the NodeB. Also, dual carrier HSDPA tends to be easier to upgrade the network since it can be implemented with single 10 MHz power amplifier per sector while MIMO requires two separate power amplifiers.

Table 13.2 HSDPA terminal categories

Category	Codes	Modulation	MIMO	DC-HSPA	Coding	Peak	3GPP
12	5	QPSK	–	–	3/4	1.8 Mbps	Release 5
6	5	16QAM	–	–	3/4	3.6 Mbps	Release 5
8	10	16QAM	–	–	3/4	7.2 Mbps	Release 5
9	15	16QAM	–	–	3/4	10.1 Mbps	Release 5
10	15	16QAM	–	–	1/1	14.0 Mbps	Release 5
13	15	64QAM	–	–	5/6	17.6 Mbps	Release 7
14	15	64QAM	–	–	1/1	21.1 Mbps	Release 7
15	15	16QAM	2×2	–	5/6	23.4 Mbps	Release 7
16	15	16QAM	2×2	–	1/1	28.0 Mbps	Release 7
17	15	64QAM or MIMO		–	5/6	23.4 Mbps	Release 7
18	15	64QAM or MIMO		–	1/1	28.0 Mbps	Release 7
19	15	64QAM	2×2	–	5/6	35.3 Mbps	Release 8
20	15	64QAM	2×2	–	1/1	42.2 Mbps	Release 8
21	15	16QAM	–	Yes	5/6	23.4 Mbps	Release 8
22	15	16QAM	–	Yes	1/1	28.0 Mbps	Release 8
23	15	64QAM	–	Yes	5/6	35.3 Mbps	Release 8
24	15	64QAM	–	Yes	1/1	42.2 Mbps	Release 8

3GPP Release 8 does not define the use of MIMO and dual carrier HSDPA at the same time. Therefore, the peak data rate in Release 8 is still 42 Mbps even with dual carrier solution. It would be possible to double the data rate to 84 Mbps by combining MIMO and dual carrier transmission in later 3GPP releases.

New HSDPA terminal categories with 64QAM, MIMO and DC-HSDPA were added in Releases 7 and 8. The HSDPA categories 13 and 14 include 64QAM and categories 15–18 MIMO offering peak rates of 21.1 Mbps and 28.0 Mbps. The combination of 2×2 MIMO and 64QAM is part of categories 19 and 20 pushing the peak rate to 42.2 Mbps. The DC-HSDPA categories are 21–24. Categories 21–22 support 16QAM and categories 23–24 support 64QAM. HSDPA terminal categories are listed in Table 13.2.

13.7 Uplink 16QAM

HSUPA Release 6 uplink uses QPSK modulation providing 5.76 Mbps. The uplink data rate within 5 MHz can be increased by using higher order modulation or MIMO transmission. The challenge with uplink MIMO is that the UE needs to have two power amplifiers. Therefore, uplink single user MIMO is not part of HSPA evolution nor LTE in Release 8. The higher order 16QAM modulation was adopted as part of Release 7 for HSUPA, doubling the peak rate to 11.5 Mbps. HSUPA terminal categories are listed in Table 13.3.

The higher order modulation improves the downlink spectral efficiency because the downlink has a limited number of orthogonal resources. The same is not true for uplink since the HSUPA uplink is not orthogonal and there is a practically unlimited number of codes available in uplink. The highest uplink spectral efficiency can be achieved by using QPSK modulation only. In other words, HSUPA 16QAM is a peak data rate feature, not a capacity feature.

The multi-path propagation affects high data rate performance. Therefore, the UE equalizer is used on HSDPA terminals. Also the NodeB receiver can improve the uplink high bit rate

Table 13.3 HSUPA terminal categories

Category	TTI	Modulation	MIMO	Coding	Peak	3GPP
3	10 ms	QPSK	–	3/4	1.4 Mbps	Release
5	10 ms	QPSK	–	3/4	2.0 Mbps	Release
6	2 ms	QPSK	–	1/1	5.7 Mbps	Release
7	2 ms	16QAM	–	1/1	11.5 Mbps	Release

HSUPA performance in multi-path channels by using an equalizer. Another solution is to use four antenna reception in uplink.

High uplink data rates require also high Ec/N0. Fixed reference channel 8 with 8 Mbps and 70% throughput requires Ec/N0 = 12–16 dB [6]. The corresponding uplink noise rise will also be similar to 12–16 dB impacting the coverage of other simultaneous users. It is therefore beneficial to use the uplink interference cancellation to subtract the high bit rate interference from other simultaneous users. The LTE radio uses an orthogonal uplink solution avoiding the intra-cell interference.

The uplink bit rate could be increased from 11.5 Mbps to 23 Mbps by using dual carrier transmission. The uplink dual carrier is not defined in Release 8 but could be considered for later 3GPP releases.

13.8 Layer 2 Optimization

WCDMA Release 99 specification was based on the packet retransmissions running from the Radio Network Controller (RNC) to the UE on L2. The L2 Radio Link Control (RLC) packets had to be relatively small to avoid the retransmission of very large packets in case of transmission errors. Another reason for the relatively small RLC packet size was the need to provide sufficiently small step sizes for adjusting the data rates for Release 99 channels. The RLC packet size in Release 99 is not only small, but it is also fixed for Acknowledged Mode Data and there are just a limited number of block sizes in Unacknowledged Mode Data. This limitation is due to transport channel data rate limitations in Release 99.

The RLC payload size is fixed to 40 bytes in Release 99 for Acknowledged Mode Data. The same RLC solution is applied to HSDPA Release 5 and HSUPA Release 6 as well: the 40-byte packets are transmitted from RNC to the base station for HSDPA. An additional configuration option to use an 80-byte RLC packet size was introduced in Release 5 to avoid extensive RLC protocol overhead, L2 processing and RLC transmission window stalling. With the 2 ms TTI used with HSDPA this leads to possible data rates being multiples of 160 kbps and 320 kbps respectively.

As the data rates are further increased in Release 7, increasing the RLC packet size even further would significantly impact on the granularity of the data rates available for HSDPA scheduling and the possible minimum data rates.

3GPP HSDPA and HSUPA allow the optimization of the L2 operation since L1 retransmissions are used and the probability of L2 retransmissions is very low. Also, the Release 99 transport channel limitation does not apply to HSDPA/HSUPA since the L2 block sizes are independent of the transport formats. Therefore, it is possible to use flexible and considerably larger RLC sizes and introduce segmentation to the Medium Access Control (MAC) layer in the base station.

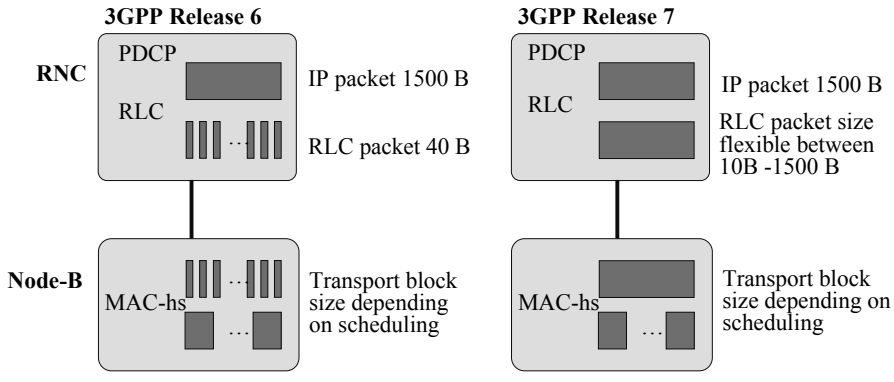


Figure 13.14 Flexible RLC concept

This optimization is included for downlink in Release 7 and for uplink in Release 8 and it is called flexible RLC and MAC segmentation solution. The RLC block size in flexible RLC solution can be as large as an Internet Protocol (IP) packet, which is typically 1500 bytes for download. There is no need for packet segmentation in RNC. By introducing the segmentation to the MAC, the MAC can perform the segmentation of the large RLC PDU based on physical layer requirements when needed. The flexible RLC concept in downlink is illustrated in Figure 13.14.

The flexible RLC and MAC segmentation brings a number of benefits for L2 efficiency and peak bit rates.

- The relative L2 overhead is reduced. With the RLC header of 2 bytes the RLC overhead is 5% for the 40-byte RLC packet. When the RLC packet size increases to 1500 bytes, the RLC header overhead is reduced to below 0.2%. This reduction of the overhead can improve the effective application data throughput.
- The RLC block size can be selected flexibly according to the packet size of each application. This flexibility helps to avoid unnecessary padding, which is no longer needed in the flexible RLC solution. This is relevant especially for small IP packet sizes, which are typical in VoIP or streaming applications.
- Less packet processing is required in the RNC and in UE with octet aligned protocol headers. The number of packets to be processed is reduced since the RLC packet size is increased and octet aligned protocol headers avoid bit shifting in high data rate connections. Both reduce L2 processing load and make the high bit rate implementation easier.
- Full flexibility and resolution of available data rates for the HSDPA scheduler.

13.9 Single Frequency Network (SFN) MBMS

Multimedia broadcast multicast service (MBMS) was added to 3GPP as part of Release 6. 3GPP Release 6 can use soft combining of the MBMS transmission from the adjacent cells. The soft combining considerably improves MBMS performance at the cell edge compared to receiving the signal from a single cell only. Release 6 MBMS therefore provides a very good starting point for broadcast services from the performance point of view.

Even if the soft combining can be used in Release 6, the other cell signals still cause interference to the MBMS reception since the adjacent cells are not orthogonal due to different scrambling codes. If the same scrambling code was used in all cells together with a terminal equalizer, the other cells transmitting the same signal in a synchronized network would be seen just as a single signal with time dispersion. This solution essentially provides a single frequency network with practically no neighboring cell interference. The MBMS over a Single Frequency Network (MBSFN) can enhance MBMS data rates and capacity. The single frequency can be realized with network synchronization and by using the same scrambling code for MBMS transmissions from multiple cells. The MBSFN is included in Release 7 and extended to unpaired bands in Release 8. The unpaired band solution is called Integrated Mobile Broadcast (IMB). The MBSFN solution requires a dedicated carrier for MBMS only transmission, which makes MBSFN a less flexible solution from the spectrum point of view compared to Release 6 MBMS. Release 6 MBMS can coexist with point-to-point traffic on the same carrier.

13.10 Architecture Evolution

3GPP networks will be increasingly used for IP based packet services. 3GPP Release 6 has four network elements in the user and control plane: base station (NodeB), RNC, Serving GPRS Support Node (SGSN) and Gateway GPRS Support Node (GGSN). The architecture in Release 8 LTE will have only two network elements: base station in the radio network and Access Gateway (a-GW) in the core network. The a-GW consists of control plane Mobility Management Entity (MME) and user plane System Architecture Evolution Gateway (SAE GW). The flat network architecture reduces the network latency and thus improves the overall performance of IP based services. The flat model also improves both user and control plane efficiency. The flat architecture is considered beneficial also for HSPA and it is specified in Release 7. The HSPA flat architecture in Release 7 and LTE flat architecture in Release 8 are exactly the same: NodeB is responsible for the mobility management, ciphering, all retransmissions and header compression both in HSPA and in LTE. The architecture evolution in HSPA is designed to be backwards compatible: existing terminals can operate with the new architecture and the radio and core network functional split is not changed. The architecture evolution is illustrated in Figure 13.15.

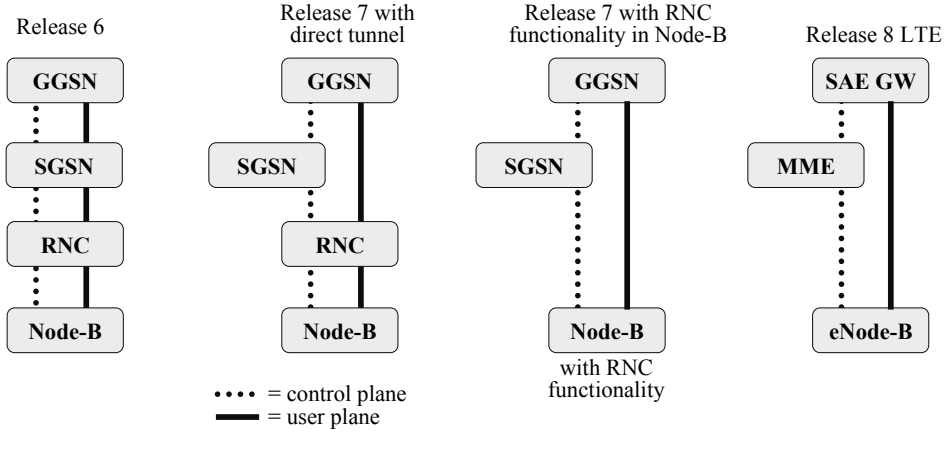


Figure 13.15 Evolution towards flat architecture

Also the packet core network has flat architecture in Release 7. It is called direct tunnel solution and allows the user plane to bypass SGSN. When having the flat architecture with all RNC functionality in the base station and using a direct tunnel solution, only two nodes are needed for user data operation. This achieves flexible scalability and allows the introduction of the higher data rates with HSPA evolution with minimum impact to the other nodes in the network. This is important for achieving a low cost per bit and enabling competitive flat rate data charging. As the gateway in LTE has similar functionality as the GGSN, it is foreseen to enable LTE and HSPA deployment where both connect directly to the same core network element for user plane data handling directly from the base station.

3GPP Release 8 includes a new architecture to support small home NodeBs as well. The home NodeBs are also called femto access points. The home NodeBs are installed by people at home in the same way as WLAN access points are installed now. These home NodeBs use the operator's licensed frequency and are connected to the operator's core network. The output power level is low, typically 100 mW or below. The transport connection uses fixed DSL connections in homes. The handovers and idle mode selections between home NodeBs and the macro network are controlled by the network algorithms and parameters. The flat architecture from Figure 13.15 is not optimized for when there are a very large number of small base stations because the number of core network connections would be very large. Therefore, a new network element – called Home NodeB gateway – is introduced to hide the large number of home NodeBs from the core network. The gateway is located in the operator's premises. The new interface between home NodeBs and the gateway is called Iuh. The interface between the gateway and the core network is the normal Iu interface. The home NodeB architecture is illustrated in Figure 13.16 [7].

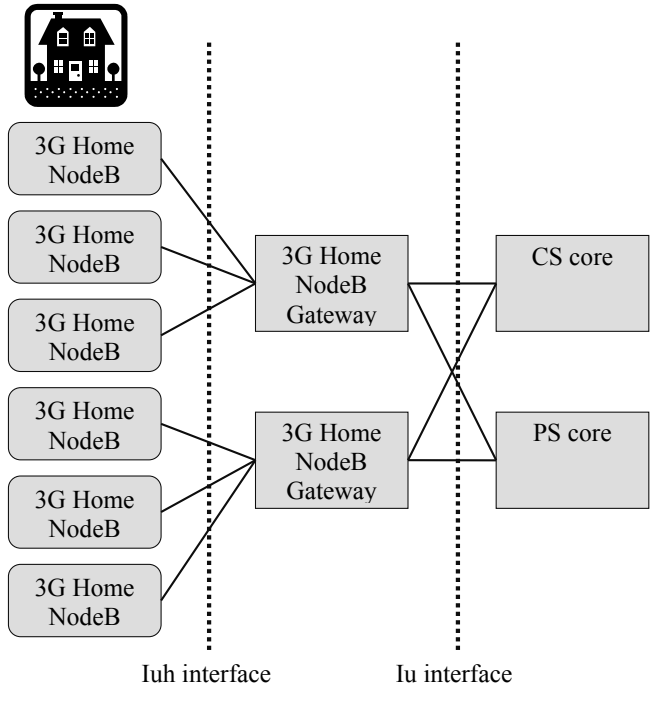


Figure 13.16 Home NodeB architecture [7]

13.11 Summary

While 3GPP is working on a new radio system – LTE (Long Term Evolution) – in Release 8, there are a number of important enhancements also brought to existing High Speed Packet Access (HSPA) standards. HSPA enhancements provide major improvements to the end user performance in increasing peak bit rates, reducing mobile terminal power consumption and reducing channel allocation latency. The peak bit rates can be tripled with 64QAM and 2×2 MIMO from 14 Mbps to 42 Mbps. These peak rate features also bring some improvements to the spectral efficiency and cell capacity. The downlink dual carrier HSDPA solution also enhances the data rates by using two adjacent HSDPA carriers for a single user transmission. The combination of Dual carrier and MIMO could push the peak rate to 84 Mbps in further 3GPP releases.

The UE power consumption is considerably reduced in HSPA evolution due to discontinuous transmission and reception. The expected voice call talk time will improve by 50% and the usage time of bursty data applications increases even more. The practical talk times will improve further when the digital and RF technologies evolve and general power consumption is reduced.

HSPA evolution enhances basic voice service by mapping Circuit Switched (CS) voice on top of HSPA channels – a concept called CS voice over HSPA. It is essentially a combination of Voice over IP (VoIP) in the radio and CS voice in the core network. CS voice over HSPA

		Peak rate	Average rate (capacity)	Cell edge rate	Latency gain	Talk time
Downlink	HSDPA 64QAM[1]	+50%	<10%	-	-	-
	HSDPA 2x2MIMO	+100%	<30%	<20%	-	-
	DC-HSDPA	+100%	+20-100%	+20-100%	-	-
Uplink	HSUPA 10 ms (2.0 Mbps)[2]	+600%	+20-100%	<100%	Gain 20 ms	-
	HSUPA 2 ms (5.8 Mbps)	+200%	<30%	-	Gain 15 ms	-
	HSUPA 16QAM	+100%	-	-	-	-
	Advanced NodeB receiver	-	>30%	-	-	-
	DTX/DRX	-	-	-	-	>+50%
	HS-FACH / HS-RACH	-	-	-	Setup time <0.1 s	-
	CS voice over HSPA	-	+80% (voice)	-	-	>+50%

[1]Baseline WCDMA Release 5 downlink 14.4 Mbps

[2]Baseline WCDMA Release 99 uplink 384 kbps

= clear gain >30%

= moderate gain <30%

Figure 13.17 Summary of HSPA evolution features and their benefits

reduces UE power consumption and increases spectral efficiency compared to WCDMA voice while having no impact on the core network and end-to-end ecosystem. It will be easy to upgrade CS voice over HSPA later to VoIP over HSPA since the radio solutions are very similar.

The setup times are reduced by also mapping the RACH/FACH common channels on top of HSDPA/HSUPA. Actually, all the services including signaling can be mapped on top of HSPA in HSPA evolution. There are only a few physical layer channels left from the Release 99 specifications in Release 8 – otherwise, everything is running on top of HSPA. This explains also why the end user performance and the network efficiency are considerably improved compared to Release 99.

The performance benefits of the HSPA evolution features are summarized in Figure 13.17. The exact gains depend on the implementation choices and on the deployment environment.

3GPP Release 7 allows simplification of the network architecture. The number of network elements for the user plane can be reduced from four in Release 6 to two in Release 7. Release 7 architecture is exactly the same as used in LTE in Release 8, which makes the network evolution from HSPA to LTE straightforward. Most of the 3GPP Release 7 and 8 enhancements are expected to be relatively simple upgrades to the HSPA networks as was the case for HSDPA and HSUPA in earlier releases.

References

[1] 3GPP, 'Continuous connectivity for packet data users', 3GPP TR25.903 V7.0.0, March 2007.
[2] 3GPP, Technical Specification 24.008 'Mobile radio interface Layer 3 specification; Core network protocols', V.8.3.0.
[3] Kurjenniemi, J., Nihtilä, T., Lampinen, M., Ristaniemi, T. 'Performance of WCDMA HSDPA Network with Different Advanced Receiver Penetrations', Wireless Personal Multimedia Communications (WPMC), Aalborg, Denmark, 17–22 September 2005.
[4] 3GPP, 'Further discussion on delay enhancements in Rel7', 3GPP R2-061189, August 2006.
[5] 3GPP, '64QAM for HSDPA', 3GPP R1-063335, November 2006.
[6] 3GPP, Technical Specifications 25.104 'Base Station (BS) radio transmission and reception (FDD)', V.8.3.0.
[7] 3GPP, Technical Report 25.820 '3G Home NodeB Study Item Technical Report', V.8.1.1.

Index

3GPP *see* 3rd Generation Partnership Project (3GPP) *under Th*

absolute priority based reselection 168
Abstract Syntax Notation One (ASN.1) 18, 146
Access Stratum (AS) 148
  security keys 152
ACKnowledged/Not ACKnowledged reception status (ACK/NACK) 93–7, 100–3, 114, 118, 119, 236, 274, 389, 391–3
  block-wise spreading 95
  bundling mode 377–81
  and CQI 93–4
  misinterpretation 207
Adaptive Multi-Rate (AMR) 144, 259–60, 262, 265–73, 275, 280, 381, 402–4
  AMR-Wideband (AMR-WB) codec 259–60, 394
  narrowband or wideband 404
  radio and audio bandwidth 260
Adaptive Transmission Bandwidth (ATB) 194, 196, 200–1, 203
Additive White Gaussian Noise (AWGN) 217, 301
  Shannon capacity 218
Adjacent Channel Interference Ratio (ACIR) 290
Adjacent Channel Leakage Ratio (ACLR) 217, 286, 288, 291
Adjacent Channel Selectivity (ACS) 290
  derivation and calculation of ACS requirement 303
  and narrow-band blocking 303
  test cases I and II 342
Admission control (AC) 182
Aggregate Maximum Bit Rate (AMBR) 64, 183
aggressor relaxation technique 339
aggressor transmitter leakage to victim receiver 315–16
aggressor/victim block diagram 315
Alamouti encoding 353
algorithms
  admission control, eNodeB 182–3
  combined fast adaptive transmission bandwidth and frequency domain packet scheduling, LTE uplink 202
  emergency calls, NULL ciphering 152
  Frequency Domain Packet Scheduling (FDPS) 201–2
  idle mode intra-frequency cell reselection 168
  RRM
    at eNodeB 181
    HARQ 181
  three-step packet scheduling 187
  time domain (TD) packet scheduling 187
  vendor-specific admission control 182
Allocation and Retention Priority (ARP) 64, 182, 271, 332, 343, 345, 347
  ARP parameter 183
Analog Channel Filters (ACF) 343
Analog-to-Digital (A/D) conversions 318
Application Function (AF) 60
Application Server (AS) 64
Authentication, Authorization and Accounting (AAA) 48–50, 52, 55, 207
  server interfaces and main functions 49
Authentication Center (AuC) 34, 50
Automatic Neighbour Relationship (ANR) functionality 35, 174
azimuth spread, sectorization configurations 241

backwards compatibility 18
Band-Pass Filter (BPF) 332

LTE for UMTS: OFDMA and SC-FDMA Based Radio Access Edited by Harri Holma and Antti Toskala

© 2009 John Wiley & Sons, Ltd.

ISBN: 978-0-470-99401-6

bandwidths
  Adaptive Transmission Bandwidth (ATB) 194, 196, 200–1, 203
    fast ATB 201
  channel and transmission bandwidths 77, 286–7
  combined fast adaptive transmission bandwidth and frequency domain packet scheduling, LTE uplink 202
  E-UTRA 287
  fixed bandwidth uplink packet scheduling 203
  LTE 219
    combined fast adaptive transmission bandwidth and frequency domain packet scheduling algorithm in LTE uplink 202
    efficiency for downlink with 10 MHz system 217
    relative efficiency, uplink/downlink 241–2
    spectral efficiency 240, 242
    system simulations 240
  selection, uplink link performance 219
  spectrum bandwidth 317
  supported transmission bandwidths with normal/relaxed sensitivity 287
  transmission bandwidth 77
  transmission bandwidth configuration 286
baseband (BB) 312, 318
Basic System Architecture Configuration
  CP protocol structure 35–6
  EPS Mobility Management (EMM) protocol 36
  Logical Elements 26–8
  roaming 39–40
  Serving Gateway (S-GW) 29–30
  summary of interfaces and protocols 39
Bearer Binding and Event Reporting Function (BBERF) 30, 32, 50, 61–2
Best-M average 128
Binary Phase Shift Keying (BPSK) 79, 86, 95, 406
Block Edge Masks (BEM) 293–5
block-wise spreading 95
Breakout Gateway Control Function (BGCF) 59
broadband data usage 2–3
Broadcast Channel (BCH) 84, 141, 148
  see also Physical Broadcast Channel (PBCH)
Broadcast Control Channel (BCCH) 140, 272
Buffer Status Reports (BSRs) 197–9, 210
  formats used in LTE uplink 198
  mapping from RB to radio bearer group for buffer status reporting 198
  short and long buffer status report types in LTE uplink 199
Bussgang theorem 358

Call State Control Function (CSCF)
  central element in SIP signaling 57–9
  Emergency CSCF (E-CSCF) 58–9
  and IMS architecture 57
  Interrogating CSCF (I-SCSF) 58
  Proxy CSCF (P-CSCF) 58
  Serving CSCF (S-CSCF) 58
capital expenditure (CAPEX) 158
carrier level 328
Carrier to Interferer power Ratio (CIR) 342
cdma2000* 1xRTT, protocols and interfaces 56
cdma2000* Access Networks 51–60
cdma2000* ANs 53
cdma2000* High Rate Packet Data (HRPD) 24, 45, 48, 133, 150, 358
  additional and updated logical elements 54–5
  HRPD Serving Gateway (HSGW) 52, 54
  inter-working 51–5
  protocols and interfaces 55–6
Cell Change Order 156
cell ranking criterion, intra-frequency cell reselection 167
cell search procedure 130
cell selection and reselection process 167–8
cell throughput
  traffic mix of 12 best effort and 8 CBR users per cell 193
  values, dimensioning 247–8
cells
  average HS-DSCH allocation duration CDF 251
  data volume contribution to total RNC data volume 251
  dual cell HSDPA 68
  performance analysis 252–6
  range limitation, caused by earth's curvature 228
chained case 51
channel arrangements 287
Channel Quality Indicator (CQI) 89, 125, 194, 200, 207, 210
  compression schemes 127–8
  feedback 85
  granularity 271
  higher layer-configured sub-band feedback 128

manager at Layer 1 181–8
wideband feedback 127–8
channel raster, E-UTRA frequency bands 287
channel reciprocity 372–3
channel state feedback reporting procedure 123–4
periodic and aperiodic 126–7
Channel State Information (CSI), reporting procedure 124
channel-aware time domain scheduling metrics 201
Circuit Switched (CS) voice 414
Circuit Switched Fall Back (CSFB) 45
handover, mobile terminated call 26, 276
for LTE 275–7
Closed Loop MIMO 126, 218–19
Code Division Multiplexing (CDM) 94
commercial reference sensitivity
Common Control Channel (CCCH) 141
Common Phase Error (CPE) 352
Common Pilot Channel (CPICH) 73
Constant Amplitude Zero Autocorrelation Codes (CAZAC) 94–5, 104
sequence index and cyclic shift 196
constant bit rate (CBR) 190, 192
Continuous Wave (CW) signal 305
Control Channel Elements (CCE)
index 100
number 263–4
number per TTI in downlink and uplink 268
control channel limitation
maximum capacity with fully dynamic scheduling 264
multi-user channel sharing 263
VoIP capacity, Semi-Persistent Scheduling (SPS) method 263
Control Format Indicator (CFI) 112
Coordinated Multiple Point (CoMP) transmission 20
CS fallback 155
with Cell Change Order 156
with PS Handover 156
Cubic Metric (CM) 75, 79, 85
with OFDMA and SC-FDMA 79
Cumulative Density Function (CDF)
average HS-DSCH allocation duration CDF 251
and PDF of data volume busy hour share 254
per user 234
Cyclic Prefix (CP) 343–4
Cyclic Redundancy Check (CRC) 353
cyclic shifts
hopping 105–6
separation NCS between preamble sequences 110
of a sequence 105

Data Radio Bearer (DRB) 153
data volume
busy hour share, CDF and PDF 254
distribution, 24 h over RNC, typical day 250
Dedicated Channel (DCH) 73
Dedicated Control Channel (DCCH) 141
Dedicated Reference Signals (DRSs) 383
Dedicated Traffic Channel (DTCH) 141
demodulation performance, PDSCH Fixed Reference Channels 356
demodulation reference signals (DM RS) 103–4, 105–7
desensitization for band and bandwidth combinations (desense) 329
mechanisms 316
design of User Equipment 311–27
Digital Video Broadcasting for Handhelds (DVB-H) 18
digital-to-analog (D/A) conversions 318, 321
DigRF version evolutions 320
Direct Conversion Receivers (DCR) 316
Direct Tunnel concept 24
Direction of Arrival (DoA) 385
Discontinuous Transmission and Reception (DTX/DRX) 207–9, 359, 400–2
with continuous packet connectivity 400
DRX related parameters and examples of use/setting 208
Layer 1 (L1) activity for voice service 401
power consumption 150, 207–10, 400–2
Discrete Fourier Transform (DFT) 70, 93
DL-SCH Channel Encoding Chain 92
Doppler shift requirement, eNodeB 309
Downlink Advanced Receiver Performance (DARP) 271
Downlink cell and user throughputs, different antenna configuration and schedulers 190
Downlink Control Information (DCI) 112
Downlink ICIC schemes 205
Downlink link performance 217–19
budgets 226
Downlink peak bit rates (Mbps) 214
with transport block size considered 215
Downlink Physical Layer Signaling Transmission (DPLST) 112–31

Downlink Pilot Time Slot (DwPTS) 370, 386
Downlink resource sharing, between PDCCH and PDSCH 91
Downlink Shared Channel (DLSCH) 84, 140
Downlink signal generation 92
Downlink slot structure 90
Downlink system performance 228–9
Downlink to Uplink transition 369–90
Downlink transmission modes 115
Downlink transmit power settings, fractional frequency and soft frequency re-use 206
Downlink User Data Transmission 89–90
Downlink user throughput distribution, with/without MIMO 230
Downlink vs Uplink packet schedulers 192–3
DRX concept 207–9
  see also Discontinuous Transmission and Reception (DTX/DRX)
DSMIPv6 47
Dual Transfer Mode (DTM) 277, 280
Duplex Gap (DG) 332
  large vs small duplex gap frequency bands 333
Dynamic Range (DR) 343
dynamic scheduling 140, 263–7, 269–71, 275, 280, 381–3, 393–5
  downlink 184–92
  downlink vs uplink 192–3
  uplink 192–3

E-HRPD
  connecting to EPC 55
  elements 54
  network 52
  tunnelled pre-registration to eHRPD and to E-UTRAN 53
E-HRPD to E-UTRAN
  handover 54
  idle mode mobility 54
E-UTRA 21
  dynamic range measurement for 10 MHz E-UTRA 301
  measurement quantities 360
  Radio Resource Management (RRM) 358–64
  signal, unwanted interferer 304
  specification, for downlink single-user MIMO transmission schemes 235
E-UTRA carriers
  definition of channel transmission bandwidth Configuration 286
  nominal channel spacing 288
E-UTRA FDD, measurement quantities 360
E-UTRA frequency bands 286–7
  channel raster 100 kHz 287
  unwanted emission limit, North America 290
E-UTRA reference system
  deployment cases
    macro cell scenarios 228
    micro cell scenarios 228
  evaluation and test scenarios 224–5
  RSSI 133
E-UTRA TDD, measurement quantities 360
E-UTRA victim, ACLR2/3.84 MHz 291
E-UTRAN 21, 24, 39–44, 43
  architecture and network elements 40
  bearer management procedures 36
  handover to UTRAN or GERAN 42
  handovers 170
  mobility 178–80
  mobility support 147
  and Non-3GPP Access Networks 45
  RRC connection between UE and E-UTRAN 151
E-UTRAN Access Network 25–7
  system architecture for E-UTRAN-only network 25
E-UTRAN Node B *see* eNodeB
E-UTRAN to E-HRPD
  handover 53
  idle mode mobility 54
earth's curvature, maximum cell range limitation 228
Effective Frequency Load (EFL) 271
EGPRS/WCDMA/LTE optimized RF subsystem block diagram 312–14
emergency calls, NULL ciphering algorithm 152
Emergency CSCF (E-CSCF) 58–9
EMI
  coupling paths within an RF IC 322
  mitigation features, imperfect differential signaling; slew rate control, ; alternate frequency selection 323
EMI control 320
Enhanced Data rates for GSM Evolution (EDGE) 7
  see also GSM Evolution (EDGE)
Enhanced Full Rate (EFR) speech codec 259
Enhanced Packet Core Network (EPC) 25, 49
  alternative protocols for the S5/S8 interface 37
  control plane protocol stack in EPS 36
Enhanced Packet Data Gateway (EPDG) 48, 49, 61

eNodeB 6, 27–8, 83, 132
admission control algorithm 182–3
connections to other logical nodes and main functions 28
demodulation performance 307–12
high speed train demodulation requirement 309
Doppler shift requirement 309
downlink resource allocation 90
GPRS Tunneling Protocol, tunnel to target cell 171
handover Request message 160
measurement 132
noise 300
power spectral density level used at UE 199
RF Receiver 300–6
RF Transmitter 288–99
security keys for the Access Stratum (AS) 152
self-configuration steps 35
sensitivity as function of received power with allocation bandwidths of 360 kHZ, 1.08 and 4.5 MHz 220
termination point 26
transfer context of UE to target eNodeB 157
transport layer address 173
triggering of handover 154–5
user plane and control plane protocol architecture, mapping of primary RRM functionalities 182
X2 interface for inter-eNodeB communications 138
eNodeB RF Receiver 300–6
adjacent channel selectivity and narrow-band blocking 303–4
blocking, in-band 304–5
dynamic range 301
in-channel selectivity 301–2
receiver intermodulation 306
receiver spurious emissions 306
reference sensitivity level 300
eNodeB RF Transmitter 288–96
coexistence with other systems on adjacent carriers in same operating band 290–2
coexistence with other systems in adjacent operating bands 292–5
operating band unwanted emissions 288–90
transmitted signal modulation accuracy 295–6
EPS Mobility Management (EMM) protocol 36
Equivalent Isotropic Radiated Power (EIRP) emission masks 293
error handling 157
Error Vector Magnitude (EVM) 288, 295, 296–7, 300, 357–8
budget 328
defined 296–7
derivation of EVM requirement 297–8
required EVM for 5throughput loss 299
ETU70 channel 307–8, 310–11, 357
Evolved HRPD *see* E-HRPD
Evolved Packet Core (EPC) 6, 43–4, 152
Evolved Packet System (EPS) 6, 26, 182
user plane protocol stack 38
Evolved Universal Terrestrial Radio Access *see* E-UTRA
Extensible Authentication Protocol Method for 3rd Generation Authentication and Key Agreement (EAP-AKA) 48

Fast Fourier Transform (FFT), block 70–2
femto access points 413
fixed bandwidth uplink packet scheduling 203
Forward Access Channel (FACH) 175, 178–9, 400–1, 404–6, 414–15
enhanced FACH and RACH 404
Frequency Division Duplex (FDD) mode 14, 331
FDD capacity results, various TDD configurations 395
Half Duplex (HD)-FDD operation 339
Frequency Division Multiple Access (FDMA) principle 67–8
Frequency Division Multiplexing (FDM) 94
Frequency Domain (FD) scheduler 187
Frequency Domain Packet Scheduling (FDPS) 124, 185–6, 394
algorithm 201–2
fast Adaptive Transmission Bandwidth (ATB) 201
interaction between outer loop link adaptation and adaptive modulation and coding 195
interaction between packet scheduler and power control 195
in LTE 244
single-carrier constraint 194
time-domain (TD) scheduling algorithm 187
under fractional load conditions 186
Frequency Non-Selective/Selective (FNS/FS) 237
Full Duplex (FD) 331
Full rate (FR) speech codec 259

Gateway GPRS Support Node (GGSN) 41, 412

General Packet Radio Service (GPRS)
bearer model 63
Serving GPRS Support Nodes (SGSNs) 43
VoIP roaming 262
see also GPRS Tunneling Protocol; GPRS/EDGE networks
Generalized Chirp-Like polyphase sequences 104
Generic Routing Encapsulation (GRE) 38
Geometry factor or G-factor 228–9
reference scheme 237
GERAN (GSM EDGE Radio AN) 43, 151
mobility support 147
voice call drop rates 277
GERAN/UTRAN 276, 277, 279
Globally Unique Temporary Identity (GUTI) 28
GMSK 330, 342
3GPP *see* 3rd Generation Partnership Project (3GPP)
GPRS Support Node (GGSN) 41, 412
GPRS Tunneling Protocol 24
Control Plane (GTP-C) 37
tunnel to target cell 171
User Plane (GTP-U) 37–8, 159
GPRS/EDGE networks 242
DigRF version evolution 320
Group Delay Distortion (GDD) 329, 344
GSM
Band 8 285–6
measurement quantities 360
GSM AN *see* GERAN (GSM EDGE Radio AN)
GSM Association, recommendations 39
GSM carrier 246
RSSI 133
GSM EDGE Radio Access Network (GERAN) 43, 147, 151
GSM Evolution (EDGE) 7, 19, 21, 311, 331, 342, 347
defined in 3GPP 7
enhanced data rates 7
LTE refarming to GSM spectrum 246–7
GSM signal 14–15, 248, 304, 318, 330, 347, 358, 360
existing base station sites 222
FR and EFR speech codec 259, 271–3
GSM spectral efficiency, measurement 271
GSM spectrum 246–7, 257
GSM voice 224, 226
HSPA and LTE data, maximum path loss values 226
GSM/WCDMA
voice call drop rates in optimized GSM/WCDMA 261
voice evolution 7, 277
GSM/WCDMA/LTE device 364
Guaranteed Bit Rate (GBR) 64, 183
constraints on 192
Guard Period (GP) 369

Half Duplex (HD)-FDD operation 339
handover
completion 172
delay 174
E-UTRAN 170
eNodeB triggering 154, 155
execution 172
inter-radio access technology handover to other radio access technology 154
inter-RAT (inter-system between LTE and GERAN) 177–8
interruption time 178
intra-frequency handover procedure 170
intra-LTE handovers 154, 170–3
preparation 172
WCDMA soft handovers in uplink and downlink 179
with/without Optimizations 45, 48
X2 interface enables lossless handover 171
handover frequency in network 175
handover performance 363–4
High Interference Indicator (HII) 206
High Pass Filters (HPF) 339, 350
High Rate Packet Data (HRPD) *see* cdma2000* High Rate Packet Data (HRPD)
High Speed Downlink Packet Access (HSDPA)
busy hour network level throughput 2
data growth 2
data volume vs voice volume 3
dual carrier 407–9
improvements 14
optimization of L2 410
scheduler operation 73–4
Serving Cell Change frequency in 3G networks 176
spectral efficiency gain of LTE over HSDPA 243
terminal categories 409
traffic analysis in RNC and cell level 249
High Speed Packet Access (HSPA) 213–14
capacity management examples from HSPA Networks 249–50

handovers 175
HSPA evolution 244, 399–415
HSPA monoband 3G, quad band 2G transceiver/BB block partitioning in analog I/Q interface 319
improvements 14
performance targets 4–5
Release 6 4
LTE downlink efficiency benefit in macro cells 244
with two-antenna Rake receiver terminals 243
standards 8
High Speed Packet Access (HSPA) evolution 244, 399–415
architecture evolution 412–13
circuit switched voice 401–4
dual carrier HSDPA 407–9
enhanced FACH and RACH 404–6
features and their benefits 414
HSPA evolution vs LTE roles 399
Layer 2 optimization 410–11
single frequency network (SFN) MBMS 411–12
uplink 16QAM 409–10
high speed train demodulation requirement 309
High Speed Uplink Packet Access (HSUPA)
RLC payload size 410
standards 8
terminal categories 410
higher order sectorization 238
Home Agent (HA) function 49
Home NodeB gateway 413
Home Routed model 51
Home Subscription Server (HSS) 28, 34, 50, 59
HSPA monoband 3G, quad band 2G transceiver/BB block partitioning in analog I/Q interface 319
Hybrid Adaptive Repeat and Request (HARQ) 88, 307
HARQ aware frequency domain packet scheduling 188
HARQ procedure 118
HARQ protocol 353
HARQ retransmission 114
information 236
see also ACKnowledged/Not ACKnowledged reception status (ACK/NACK)
management 181
packet scheduler interaction 184
Physical HARQ Indicator Channel (PHICH) 112, 115
Hyper Frame Number (HFN) 152

Idle Mode Mobility *see* mobility
IEEE 802.16 4
IEEE-ISTO 320
Image Rejection (IR) 348
Implementation Margin (IM) 332
IMS-Media Gateway (IMS-MGW) 60
in-channel selectivity 301–2
Insertion Loss (IL) 332
Integrated Circuit (IC) 312
Inter-Carrier Interference (ICI) 343
Inter-Carrier Interference (ICI) control (ICIC) 204
Inter-Cell Interference control (ICIC), reactive and proactive schemes 204
inter-frequency/RAT reselections 168–70
Inter-Site Distances (ISD) 231
Inter-Symbol Interference (ISI) 343
Interconnection Border Control Function (IBCF) 59
interference management and power settings 204–7
Interference over Thermal (IoT) probability 231
intermodulation distortion (IMD2) 316
products 339
intermodulation distortion (IMD3) 302
LO leakage 317
International Mobile Subscriber Identity (IMSI) 28–9
International Mobile Subscriber-Media Gateway (IMS-MGW) 60
International Mobile Telecommunications (IMT)
IMT-Advanced 11
new frequencies identified for IMT in WRC-07 10
new spectrum 10
internet connections 3
Internet Protocol (IP) 37
Internet Protocol (IP) Connectivity Layer 26
Interrogating CSCF (I-SCSF) 58
intra-frequency cell reselection 167–8
intra-frequency handover procedure 170
Inverse Fast Fourier Transform (IFFT) 68, 71
IP Multimedia Subsystem (IMS) 26, 34, 56–60, 261
architecture 57
IP header compression 261
QoS 261

ITU-R IMT-Advanced schedules 20
ITU-R World Radiocommunication Conference (WRC-07) 10

Japan, LTE deployments 9

late path switch 170
latency 244–6
  components 245
  end-to-end round trip time including scheduling latency 246
Layer 1 peak bit rates 213–14
Line-of-Sight (LOS) and Non-Line-of-Sight (NLOS) 225
link bandwidth efficiency for LTE downlink with a 10 MHz system 217
link budget calculations 222–4
link level performance 217–19
Local Mobility Anchor (LMA) 49
  power control 328
Local Oscillator (LO) 312
  Phase Noise 317, 329
Location Retrieval Function (LRF) 58
Logical Channel Identification (LCID) field 184
logical channels 140–1
Logical Elements, Basic System Architecture Configuration 26–8
Long Term Evolution (LTE)
  bandwidths 202, 217, 219, 240–2
    relative efficiency in downlink 241
    relative efficiency in uplink 242
    spectral efficiency (SE) 240
  benchmarking LTE to HSPA 243–4
  beyond Release 8 18
  cell search procedure 130
  combined fast adaptive transmission bandwidth and frequency domain packet scheduling algorithm in LTE uplink 202
  complexity of LTE receiver 325
  control plane radio protocols in LTE architecture 138, 139
  definition of targets 4
  device categories 131
  dimensioning example for 1+1+1 at 20 MHz 248
  downlink efficiency benefit over HSPA Release 6 in macro cells 244
  downlink vs uplink packet schedulers 192–5
  FDD frame structure 87
  first set of approved physical layer specifications 17–18
  Frequency Domain Packet Scheduling (FDPS) 185–6, 244
  half duplex operation 130–1
  HARQ operation with 8 processes 118
  intra-LTE handovers 170–1
  latency 15
  LTE-Advanced 11
    for IMT-Advanced 19–21
    resource sharing with LTE 20
    schedules 20
  MBMS 18
  modulation constellations 85
  multiple access background 67–70
  network architecture 6, 17
  overview 5–7
  packet switched domain optimized 14
  performance 213–57, 283–364
  performance targets 5
  Physical Layer (PHY) 37, 83–135
  power control 119
  Radio Protocols 137–63
    original network architecture 17
  Random Access (RACH) operation 120–1
  reactive uplink ICIC scheme based on OI 207
  refarming to GSM spectrum 246–7
  Release 8 160
    load control and interference management 160–1
  specifications 21–2
  standardization 13–22
    phases 16–18
  target setting for LTE feasibility 14
  TDD Mode 14, 367–97, 379–83, 385–97
    commonality with TD-SCDMA 367
    see also Time Division Duplex (TDD) operation
  terminal vs WCDMA terminal 311
  transmit modulation accuracy, EVM 328
  user plane radio protocols in LTE architecture 139
  voice over IP (VoIP), VoIP capacity 265–6
Low Noise Amplifier (LNA) 312, 320, 339–40, 342, 343

Master Information Block (MIB) 141, 148
Maximum Bit Rate (MBR) 64
  aggregate MBR (AMBR) 183
maximum path loss values for GSM voice and HSPA and LTE data 226
Maximum Power Reduction (MPR) 317
  LTE 327

Maximum Ratio Combining (MRC) 353
Maximum Sensitivity Degradation (MSD) 339
Media Gateway Control Function (MGCF) 59, 60
Medium Access Control (MAC) 37, 139–41, 198
  data flow 142–4
  layer 117, 137–41
  layer in base station 410
  PDU structure and payload types for DL-SCH and UL-SCH 142
  Random Access Response (RAR) 143
  segmentation 411
micro cell scenarios 225
Minimum Mean Squared Error (MMSE) 354
Mobile Industry Processor Interface (MIPI) consortium 320
mobile speeds 221
mobile subscribers
  data growth 2
  growth in penetration worldwide 1–2
Mobile Switching Center (MSC)/Visiting Location Register (VLR) network 275–80
mobility 165–80
  cell selection and reselection process 167–8
  differences between UTRAN and E-UTRAN 178–80
  idle mode intra-frequency cell reselection algorithm 168
  Idle Mode Mobility 166–70, 168
    idle vs connected mode 165
    parameters 169
Mobility Management Entity (MME) 6, 28–31, 169, 412
  connections to other logical nodes and main functions 30
  and new S101 interface towards eHRPD RAN 54
  self-configuration of S1-MME and X2 interfaces 34–5
  tracking area 169
Modulation and Coding Scheme (MCS) 297
  throughput curves and approximating MCS envelope 298
multi-antenna MIMO evolution beyond 2X2 234
multi-bandwidth and multi-radio access technology 19
multi-carrier principle 68
multi-user channel sharing 263–4
Multicast Channel (MCH) 84
Multicast Control Channel (MCCH) 141, 144
Multicast Traffic Channel (MCTH) 141, 144
Multimedia Broadcast Multicast Service (MBMS)
  part of Release 6 411
  Single Frequency Network (MBSFN) 412
  WCDMA in Release 6 141
Multimedia Resource Function (MRF) 59
  controller (MRFC) and processor (MRFP) 59
Multiple Input Multiple Output (MIMO) 80–7, 354–6
  2 X 2 MIMO transmission concept 406
  basics 80–7
  Closed Loop MIMO 126, 218
  Dedicated Reference Signals 383–4
  downlink EUTRA MIMO transmission schemes 234
  feedback information 237
  HSPA evolution 405
  improvement of spectral efficiency 407–9
  multi-antenna MIMO evolution beyond 2X2 234
  multi-antenna operation 129, 132
  multi-user MIMO 81
  packet scheduling 188–9
  quasi-dynamic (QD) 237
  single/dual codeword MIMO transmission 236
  spectral efficiency (SE), average cell vs cell-edge user (coverage) simulation results 238
  two-by-two antenna configuration 80
  v-MIMO 88, 215
multiplexing, uplink control and data 89

narrowband
  Adaptive Multi-Rate (AMR) 404
  blocking 303–4, 341–3
    ACS 303
    measurement for 5 MHz E-UTRA 343
    test 342
  Relative Narrowband Transmit Power (RNTP) 160, 205
neighborlist
  automatic neighbor relation functionality 35, 174
  generation and maintenance 174
  intra-frequency neighborlist generation 174
Network Mode of Operation 1 (NMO1), with Gs interface 276
NodeBs
  Home NodeB gateway 413
  small home NodeBs 413
  see also eNodeB

Noise Figure (NF) 315
noise leakage 315
non-3GPP Inter-working System Architecture 45–6
  interfaces and protocols 50–1
  roaming with non-3GPP accesses 51
  trusted non-3GPP AN networks 48
  untrusted non-3GPP AN networks 48–9
Non-Access Stratum (NAS) 35, 148
  message transfer 157
  protocol 137
Non-Guaranteed Bit-Rate (Non-GBR) 64
Non-Line-of-Sight (NLOS) 225
NULL ciphering algorithm, emergency calls 152

Offline Charging System (OFCS) 61
Okamura-Hata model 225
  cell ranges 227
  parameters for model 228
Online Charging System (OCS) 61
operating band
  frequency ranges for spurious emissions 289
  unwanted emission limits 288–90
operational expenditure (OPEX) 158
Orthogonal Frequency Division Multiple Access (OFDMA) 5–6, 16, 67–82
  basics 70–4
  challenges 70
  creation of guard interval for OFDM symbol 72
  reference signals 72, 81
  resource allocation in LTE 74
  transmitter and receiver 71
    windowing for shaping spectral mask 73
  vendor specific admission control algorithms 182
  see also Physical Layer (PHY)
Orthogonal Frequency Division Multiplex (OFDM) 352, 353
  sub-carrier spacing used in LTE 352
  symbol of duration 353
Out-of-Band (OOB) emission requirements 286, 288
Outer-Loop Link Adaptation (OLLA) 194–5
Overload Indicator (OI) 207

Packet Data Convergence Protocol (PDCP) 17, 37, 145–6
  PDCP Sequence Number 152
Packet Data Network Gateway (P-GW) 26, 31–2, 49
Packet Switched Handover (PSHO) 177
Paging Channel (PCH) 84
Paging Control Channel (PCCH) 140, 141
paging messages 150
Parallel Concanated Convolution Coding (PCCC) 88
Payload Data Units (PDUs) 140
Peak-to-Average Ratio (PAR) 85, 327
Pedestrian A (PA) 307
  distortion 329
performance (LTE) 213–57
  cell performance analysis 252–6
  comparison for different delay budgets 275
  requirements 283–364
  targets 5
Personal Handyphone System (PHS) 289
Phase Locked Loop (PLL) 312
Phase Noise (PN), multiplication on OFDM sub-carriers 352
Physical Broadcast Channel (PBCH) 84, 116–17, 214
Physical Cell ID (PCI) value 130, 134
Physical Control Format Indicator Channel (PCFICH) 97, 112–15
Physical Downlink Control Channel (PDCCH) 90, 184
  allocated transmission resources 263
  limitations 203, 214
Physical Downlink Shared Channel (PDSCH) 84, 89, 263
  air interface capacity 266
  available bandwidth 267
  average HS-DSCH allocation duration CDF 251
  utilization rate for semi-persistent scheduler 267
Physical HARQ Indicator Channel (PHICH) 112, 115
Physical Hybrid Automatic Repeat Request Indicator Channel (PHICH) 214
Physical Layer (PHY) 37, 83–135
  eNodeB measurements 132–4
  modulation 85–7
  parameter configuration 133–5
  procedures 117–31
Physical Random Access Channel (PRACH) 84, 109–12, 310–11
  ETU70 channel 310–11
  illustration of uplink timing uncertainty windows for RACH preamble with considerable Doppler 111

LTE RACH preamble formats for FDD 109
Physical Resource Blocks (PRBs) 89, 161, 184
Nre PRBs 188
Physical Uplink Control Channel (PUCCH) 93, 94–100, 309–10
configurations 97, 98
demodulation reference signals (DM RS) 103
format 1/1a/1b resources 96, 98–100
format 1a performance requirements 309
one-bit BSR 197
restrictions on payload size 126
uplink data rate reduction 214
Physical Uplink Shared Channel (PUSCH) 84, 93, 101–8, 307–9, 356–7
allocation data and different control fields 102
control signaling 101
number of allocated PRBs 195
PSD information 199
PUSCH base station tests 308
scheduling request (SR) 197
SNR requirements 308
Policy and Charging Control (PCC) 261
functions in roaming with PMIP, home routed vs breakout models 62
and QoS 60–6
summary of PCC interfaces 63
Policy and Charging Enforcement Function (PCEF) 61
Policy and Charging Resource Function (PCRF) 26, 32–3, 60
connections to other logical nodes and main functions 33
power amplifier, back-off requirements for different input waveforms 75
power consumption 4, 15, 42, 74, 75, 79, 82, 119, 131, 262, 312, 343, 346, 404–5
benefit of 900 MHz compared to 2600 MHz 227
DTX/DRX 150, 207–10, 399–402, 414–15
radio-related, 95reduction 209
Power Headroom Reports (PHR) 199–200, 210
Power Spectral Density (PSD) 119
Pre-coding Matrix Indicator (PMI) 101, 125–6, 236
Frequency Non-Selective/Selective (FNS/FS) 237
Primary Common Control Physical Channel (P-CCPCH) 133
Primary Synchronization Signals (PSS) 117, 374
prioritized bit rate (PBR) 183
probability distribution function, persistent resource allocation for different codecs in downlink 269
Proportional Fair (PF) schedulers 190, 231
protocol architecture 137–8
Proxy CSCF (P-CSCF) 58
Proxy Mobile IP (PMIP) 37, 61, 62
Public Land Mobile Network (PLMN) 148, 152, 280
cell selection 166–7
S2a or S2b interface 47
Pulse Code Modulation (PCM) 260

QoS Class Identifier (QCI) 64
characteristics for the EPS bearer QoS profile 183
Quadrature Amplitude Modulation (QAM) modulator 64, 85, 299–301, 307, 324, 327–9, 329, 345–58, 405
uplink 16QAM 409–10
Quadrature Phase Shift Keying (QPSK) 79, 85, 213–16, 220, 296, 299–302, 307–8, 317, 327–30, 334–6, 341, 347, 350–2, 356, 358, 381, 387–9, 405
Quality of Service (QoS) 27, 32–3, 37, 63–6
parameters for QCI 64, 183
and PCC 60–3
performance of QoS PS scheduler 192
QoS Class Identifier (QCI) 64, 183
QoS-aware packet schedulers in LTE 201
example 191
time domain (TD) packet scheduler 190
Quantization Noise (QN) 347
quasi-dynamic (QD) reference scheme 237

Radio Access Technology (RAT) 19
Radio Bearer Group (RBG), defined 198
Radio Bearers (RBs) 138
Data Radio Bearer (DRB) 153
Signaling Radio Bearers (SRBs) 137, 148
radio frequency (RF), channels in E-UTRA bands 167
Radio Link Control (RLC) 37
Acknowledged Mode (AM) 144
Payload Data Units (PDUs) 140
Transparent Mode (TM) 144
Unacknowledged Mode (UM) 144
Radio Link Control (RLC) layer 143–5
data flow 145
Service Data Units (SDUs) 171
Radio Network Controller (RNC) 43
3GPP Release 6 architecture 6

Radio Network Controller (*continued*)
connecting base stations 179
data volume distribution over 24 h, typical day 250
radio performance *see* performance
radio protocols 137
Radio Resource Control (RRC) 37, 146–7, 263
connection maintenance 209
connection re-establishment procedure 158
connection reconfiguration procedure 153
functions and signaling procedures 148
RRC-IDLE and RRC-CONNECTED state 147, 167
system information 148
triggers for RRC connection release 210
Radio Resource Management (RRM) 181–212
algorithms at eNodeB 181
HARQ 181
control of uplink rate control function 183
dynamic packet scheduling and link adaptation 184–92
Transmission Time Interval (TTI) 181
uplink RRM functionalities and their interaction 194
RAKE receiver 72, 243
RAN4, TSG RAN working groups 290, 338
Random Access, Contention and Non-contention Based Random Access 121–2
Random Access Channel (RACH) 84, 120–1, 370, 374–6, 386, 404–6, 414–15
enhanced FACH and RACH 404–6
placement in frequency domain in LTE TDD 376
see also Physical Random Access Channel (PRACH)
random access procedure, power ramping 121
Rank Indicator (RI) 101, 125, 127
feedback 188
Received Signal Code Power (RSCP) 133, 360
Received Signal Strength Indicator (RSSI) 133
Reference Signal Received Power (RSRP) 133, 360, 362
measurement quantity 167, 168
Reference Signal Received Quality (RSRQ) 133, 360, 362
Reference Signals (RS) 103, 104, 354
downlink 214
Relative Narrowband Transmit Power (RNTP) 160
indicator 205
relay nodes 19
Release 8 *see* Long Term Evolution (LTE); 3rd Generation Partnership Project (3GPP)
releases and process *see* 3rd Generation Partnership Project (3GPP)
Remote Access Technologies (RATs) 358
Required Activity Detection (RAD) scheduler 190
reselection handling 168
retransmissions, collision with new transmission 382
RF-BB interface application, using DigRF v4 with RX diversity 321
roaming
Basic System Architecture Configuration 39–40
home routed and local breakout 40
with non-3GPP accesses 51
visited Public Land Mobile Network (PLMN) 280
Roaming Agreement 39
Round Robin (RR) schedulers 231
round trip time measurement 245

S-criterion 167
S-IWF (SR-VCC enhanced MSC Server) 277–9
S1 Application Protocol (S1AP) 37
S1-MME and X2 interfaces, self-configuration 34–5
S9 interface 61
schedulers, Proportional Fair (PF) schedulers 190
Scheduling Request Indicator (SRI) 97
Scheduling request (SR) 197
Secondary Synchronization Signals (SSS) 117, 374
sectorization configurations 238–40
azimuth spread 241
higher order sectorization 238
impact of beamwidth with six-sector antenna 240
Self Optimized Networks (SON) enhancements 18
Self Organizing Network (SON) solutions 134
Semi-Persistent Scheduling (SPS) method 263, 271, 381–3, 395
control channel limitation for VoIP capacity 263
in downlink 264
with fully dynamic scheduling 264
user plane limited 270
sequence group hopping 106
Service Data Units (SDUs) 171

Service Request 36
Services Connectivity Layer 26
Services domain 34
Serving CSCF (S-CSCF) 58
Serving Gateway (S-GW) 26, 29–30, 44
  connections to other logical nodes and main functions 31
  and Packet Data Network Gateway (P-GW) combination 26
Serving GPRS Support Nodes (SGSNs) 43, 412
Session Initiation Protocol (SIP) 34
Shannon
  formula 217
  information theory 387
Shannon limit 213, 217–19, 256, 297, 387
  in AWGN 218
Signal to Distortion Quantization Noise Ratio (SDQNR) 345
Signal-to-Interference Noise Ratio (SINR) 186, 188, 189, 193, 195, 228
  distribution of average wide-band channel SINR macro case 1, case 3 and micro 229
  for VoIP in TDD and FDD 389
Signal-to-Noise Ratio (SNR) 217, 301, 332
  LTE eNodeB throughput as a function of SNR 221
  various spectral efficiencies 221
Signaling Radio Bearers (SRBs) 137, 148
Silence Indicator Frames (SIDs) 260, 263
simulation scenarios, assumptions 229
Single Antenna Interference Cancellation (SAIC) 271
Single Carrier Frequency Division Multiple Access (SC-FDMA) 5–6, 16, 67–82
  adjusting data rate 77
  basics 76–80
  transmitter and receiver with frequency domain signal generation 76
Single Frequency Network (SFN) MBMS 411–12
Single Input Multiple Output (SIMO) 218, 353
Single Input Single Output (SISO) 218, 353
Single Radio Voice Call Continuity (SR-VCC) 45, 56, 261, 277–80
Single User MIMO (SU-MIMO) 20
SN Status Transfer procedure 160
software reconfigurable hardware 312
Sounding Reference Signals (SRSs) 104, 106–7, 181, 210, 231
  CAZAC 196
  frequency hopping patterns 108
  important sounding parameters 196–7
  sub-carrier mapping 107
Spatial Channel Model in Urban macro (SCM-C), radio channels 227–8
Spatial Frequency Block Coding (SFBC) 353
spectral efficiency (SE) 224–8
  average cell vs cell-edge user (coverage) simulation results for MIMO 238
  compared to 10 MHz bandwidth in macro cells 242
  downlink and uplink 243
    HSPDA and LTE 224
  evaluation in 3GPP 242–3
  as function of G-factor 218
  HSPA and LTE 245
  LTE bandwidth 240
  relative to 10 MHz 242
  for SISO, SIMO and closed loop MIMO 219
  system deployment scenarios 224–5
  uplink mean and cell edge 235
    various multi-antenna configurations 234
spectrum bandwidth 317
Spectrum Emission Mask (SEM) 294
spectrum resources, example of European operator 249
spreading factor (SF), block-wise spreading 95
Stream Control Transmission Protocol (SCTP) and Internet Protocol (IP) 37
Subscription Locator Function (SLF) 59
subscription profile and service connectivity 29
Synchronization Signals, Primary (PSS) and Secondary (SSS) 117, 374
System Architecture Evolution Gateway (SAEGW) 6, 26–8, 170, 412
System Architecture Evolution (SAE) 23–66
  bearer model 63
  bearers to be handed over 159
System Frame Number (SFN) 196
System Information Blocks (SIBs) 85, 141, 148–9
  acquisition of SI message 149

Telephony Application Server (TAS) 59
terminal categories 216–18
Terminal Equipment (TE) 27
3rd Generation Partnership Project (3GPP)
  3GPP and non-3GPP access networks 46
  3GPP TSG RAN LTE specification development 21–2
  AAA server interfaces and main functions 49
  additional interfaces and protocols

3rd Generation Partnership Project (*continued*)
inter-working with legacy 3GPP CS 45
Inter-working System Architecture configuration 44
ANs, E-UTRAN, UTRAN and GERAN connected to EPC 40–1
architecture
evolution towards flat architecture 24
Inter-working System Architecture Configuration 42–4
frequency bands, paired bands vs unpaired bands 9
Guidelines for Channel Estimation 354–5
LTE frequency bands in 3GPP specifications 8–9
LTE milestones 16
market share development of 3GPP technologies 7
minimum R performance requirements for terminals (UE) and for base stations (eNodeB) 283
peak data rate evolution 8
Policy and Charging Control (PCC) 261
Release 5 HSDPA 409
Release 6 architecture 6
Release 6 HSPA 4
Release 6 HSUPA 404
Release 6 UE 400
Release 7 400–2, 404–5, 409
Release 7 allows simplification of network architecture 415
Release 7 HSDPA 409
Release 7 UE 400, 401
Release 7 WLAN Voice Call Continuity 278
Release 8 160–1, 263
Release 8 CS voice over HSPA 401
Release 8 HSDPA 409
Release 8 ICIC schemes 204
Release 8 new architecture to support small home NodeBs 413
Release 8 peak data rate 409
Release 8 specifications 385
Release 8 terminal categories 216–18
Release 9 closure 13
Release 10 20–1
releases and process 13–14
evolution towards flat architecture 412
schedule of 3GPP standards and their commercial deployments 7
semi-persistent scheduling 263
system architecture for 3GPP access networks 41–7
simplified showing only S2c 47
TSG RAN 16, 21–2
WG RAN4 16, 21–2, 290, 338
WG5 completion during 2009 162
see also Long Term Evolution (LTE)
3rd order intermodulation (IMD3) 302
Time Division Duplex (TDD) operation 14, 379–83, 385–97
basic principle 368–9
channel aware scheduling 390–1
channel reciprocity 372–3
co-existence with TD-SCDMA 371–2
control channels 374–81
data bit rate decreases 388–9
LTE TDD evolution 396
LTE TDD performance 385–96
link budget calculation 386–7
link performance 386–7
LTE TDD transmission 369–70
MCS selection and UE transmission bandwidth for coverage 387–9
MIMO and Dedicated Reference Signals 383–4
multiple access schemes 373–4
resource allocation signaling 391
round trip times between TDD and FDD 390
system level performance 389–90
uplink and downlink traffic 371
uplink power control 391
VoIP performance 393–4
Time Division Multiple Access (TDMA) systems 76
Time Division Multiplexing (TDM) 93
time domain (TD) packet scheduler 190
scheduling algorithm 187
timing advance value 119
tracking area concept/optimization 169–70
Tracking Area Updating (TAU) 36
traffic growth scenarios, 10 times and 100 times more traffic 249
transceiver (TRX), design 312
Transition Gateway (TrGW) 59
transmission bandwidth configuration 286
Transmission Control Protocol (TCP) 37
Transmission Time Interval (TTI) 181, 184, 185, 187
HARQ design for UL TTI bundling 379–80
TTI bundling 273–5, 280, 379, 395
TTI bundling vs no bundling 275, 389, 394
Transport Block Size (TBS) 123
Transport Blocks (TB) 140

transport channels 140–1
TSG RAN working groups 16, 21–2
  RAN4 290, 338
  working groups 162, 290, 338
Turbo decoder 326–7
Typical Urban (TU) channel 307

UMTS, spurious emissions requirements 288
Unit Data Protocol (UDP) and IP 37
Universal Integrated Circuit Card (UICC) 27
Universal Subscriber Identity Module (USIM) 26, 166
Universal Terrestrial Radio Access *see* UTRA
Universal Terrestrial Radio Access network (UTRAN) 21, 43
  mobility 178–80
  mobility support 147
  see also E-UTRA; E-UTRAN
uplink cell throughput for ISD 500m and ISD 1500m 232
uplink control and data, multiplexing 89
uplink coverage 273–5
  DL to UL transition 369–90
uplink and downlink shared data channel (PDSCH and PUSCH) 204
uplink fast adaptive modulation and coding functionality 200
uplink instantaneous noise raise for ISD 500m and ISD 1500m 233
uplink interference coordination 206–7
uplink link adaptation, uplink SRS measurements 199–200
uplink link budgets 224
  parameters 223
uplink link performance 219–21
  bandwidth selection in LTE 219
uplink packet scheduling 200–1
uplink peak bit rates (Mbps) 215
  with transport block size considered (Mbps) 215
Uplink Physical Layer Signaling Transmission 93–5
Uplink pilot time slot (UpPTS) 370
uplink power control 391–2
uplink resource allocation, controlled by eNodeB scheduler 86
Uplink Shared Channel (UL-SCH) 84, 140
uplink slot structure, with short and extended cyclix prefix 87
uplink spectral efficiency, mean and cell edge, various multi-antenna configurations 234
uplink system performance 231–2
uplink transmission power, CDF per user, ISD 500m 234
Uplink User Data Transmission 86
uplink vs downlink packet schedulers 192–3
USA, WCDMA networks 9
User Equipment (UE) 26–8
  capability classes and supported data 131–3
  capability transfer 157–8
  cell selection parameters provided in SIBs 151
  context release signals to source eNodeB 160
  demodulation performance 352–8
    3GPP guidelines for channel estimation 354–5
    EVM modeling and impact onto performance 357–8
    Multiple Input Multiple Output (MIMO) 80–7, 354–6
    Single Input Multiple Output (SIMO) Mode 353
    transmission in time and frequency 352–3
    transmit diversity 353–5
  design 311–27
    antennae and paths 313–15
    component count for different front end design 313
    EGPRS/WCDMA/LTE optimized RF subsystem block diagram 314
    equalization 324
    LTE vs HSDPA baseband design complexity 324–6
    multi-mode and multi-band support 311–12
    new coexistence challenges 315–16
    RF subsystem design challenges 311–12
    RF-baseband interface design challenges 318–24
    Turbo decoder 326–7
  eNodeB triggers handover 154
  inter-working 42
  LTE deployment 162
  measurement
    configuration parameters 153–4
    identities 154
    LTE system 133–4
    report 154, 173–4
    reporting thresholds 173
  periodic reporting 173
  Radio Resource Management (RRM) 358–64
    accuracy requirements 362
    cell identification 361–2

User Equipment, RRM (*continued*)
channel aware scheduling 390
handover performance 363–4
Idle vs Connected state 359–61
report mapping 362–3
reference sensitivity levels applied to each antenna port for QPSK modulation 334
reporting configuration 154
RF Receiver 331–52
ACS and narrowband (NB) blocking requirements 341
AGC loop imperfections 346
aggressors and victims, optimized quad band 2G - mono band 3G/LTE RF sub-system 337
duplexer isolation 336
EVM, LTE vs WCDMA 348
finite image rejection due to I/Q amplitude and phase mismatches 348–50
flexible bandwidth requirements
analog channel filter 343–5
onto the ADC DR 345–7
in-band analog filter amplitude and group delay distortions 350
interfering vs wanted signals for ACS requirements 341
large transmission bandwidth in small duplex distance frequency bands 336
LTE 10 MHz QPSK uplink ACLR overlapping receiver band 338
phase locked loop phase noise 350–1
reference sensitivity power level 331–2
self-desensitization contributors in FDD UEs 336
transmitter carrier leakage presence at receiver LNA input 339–40
RF Transmitter 327–31
desensitization for band and bandwidth combinations (desense) 329
modulation scheme and maximum output power per band configurations for a multi-mode PA 330
multi-mode power amplifier 330
output power window 327–8
transmitter architecture 329
RRC Connection Setup procedure 151
Specific Reference Signals (URS) 383–5
summary 364
user numbers
Macro 1 propagation scenario and 60 simultaneous users 204
Macro 1 scenario and a fixed user transmission bandwidth of 6 PRBs 203
maximum capacity calculation 264
scheduling per TTI 203
with/without taking used application into account 256
User Plane (UP) 24
GTP-U GPRS Tunneling Protocol 38, 159
latency 244–6
user satisfaction, mouth-to-ear latency 261

Vehicular A (VA) 307
vendor-specific admission control, algorithm 182
Very High Data Rate Digital Subscriber Line (VDSL2), bit rates 3
Victim/Aggressor block diagram and Aggressor transmitter leakage to victim receiver 315–16
virtual Multiple Input Multiple Output (V-MIMO) 88
Visiting Location Register (VLR) network 275–80
voice, mouth-to-ear delay requirements 261
voice call delay, end-to-end delay budget 262
voice call drop rates, optimized GSM/WCDMA networks 261
voice capacity evolution 271–3
voice communication, mobile network coverage 2
voice over IP (VoIP) 18–19, 45, 259–81, 393–5
capacity in LTE at 5 MHz 266
codes 259–60
control channel limitation, Semi-Persistent Scheduling (SPS) method 263, 381–3
end-to-end delay budget for LTE VoIP 262
example of VoIP capacity curve 265
maximum capacity calculation 264
packet bundling enabled 266
requirements 261–2
roaming, GPRS 262
scheduling and control channels 263–4
Semi-Persistent Scheduling (SPS) method 263, 271, 381–3, 395
system level performance 265–8
on top of HSPA, defined in Release 7 401
TTI bundling for enhancing VoIP coverage in uplink 273
uplink VoIP sensitivity with TTI bundling 274
VoIP capacity at 5 MHz 265–6
voice spectral efficiency, from GSM to WCDMA/HSPA 271–2

Wideband Code Division Multiple Access (WCDMA) 318
  Abstract Syntax Notation One (ASN.1) 146
  adaptive antenna arrays 132
  commercial reference sensitivity, variability 335
  dual cell HSDPA 68
  first deployments (2002) 8, 14
  and HSPA 7, 8
  impact of scaling a baseline filter's Fc optimized 344
  Release 99 supports CS voice on Dedicated Channel (DCH) 401
  soft handovers in uplink and downlink 179
  terminal, vs LTE terminal 311
  terminal power consumption 15
Wiener/MMSE filtering 355
wireless broadband access 3
Wireless Local Area Network (WLAN) 69
  access points 413
  Inter-Working (WLAN IW) 48
wireless spectrum 8–10
wireless vs wireline users, data rates 4
wireline networks, data rates 3
World Radiocommunication Conference (WRC-07) 10

X2 interfaces
  control and user plane protocol stacks 38
  inter-eNodeB communications 138
  interference level reporting 161
  intra-LTE handover operation 159
  lossless handover 171
  protocols 158–61
    structure 38
    User and Control Plane protocol stacks 159
  self-configuration 34–5
  transmission power vs threshold reporting 161

Zadoff–Chu sequences 85, 104, 110
Zero Autocorrelation (ZAC) sequences 104
Zero-Forcing (ZF) equalization 328