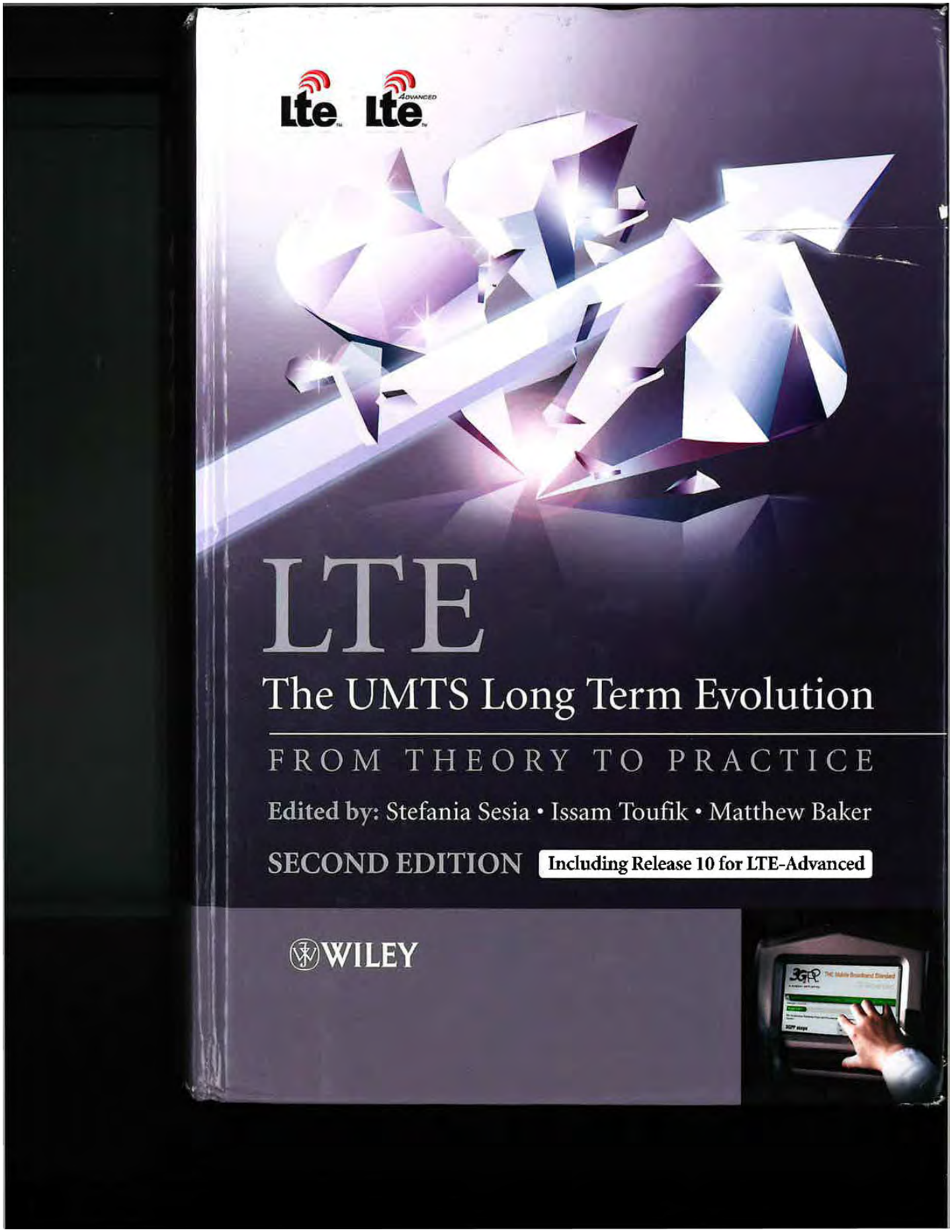

LTE
The UMTS Long Term Evolution
FROM THEORY TO PRACTICE
Edited by: Stefania Sesia • Issam Toufik • Matthew Baker
SECOND EDITION
Including Release 10 for LTE-Advanced
WILEY

APPLE 1013

LTE – The UMTS Long Term Evolution

From Theory to Practice

Second Edition

Stefania Sesia
ST-Ericsson, France

Issam Toufik
ETSI, France

Matthew Baker
Alcatel-Lucent, UK

WILEY
A John Wiley & Sons, Ltd., Publication

This edition first published 2011
© 2011 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 rights of the authors to be identified as the authors of this work have 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.

Photograph on cover courtesy of Alcatel-Lucent, from the *ngConnect* LTE-equipped car.
3GPP website reproduced by permission of © 3GPP™.

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.

Library of Congress Cataloging-in-Publication Data

Sesia, Stefania.
LTE–the UMTS long term evolution : from theory to practice / Stefania Sesia, Issam Toufik, Matthew Baker. – 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-66025-6 (hardback)
1. Universal Mobile Telecommunications System. 2. Long-Term Evolution (Telecommunications) I. Toufik, Issam. II. Baker, Matthew (Matthew P.J.) III. Title.
TK5103.4883.S47 2011
621.3845'6–dc22
2010039466

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

Print ISBN: 9780470660256 (H/B)
ePDF ISBN: 9780470978511
oBook ISBN: 9780470978504
epub ISBN: 9780470978641

Printed and bound in Great Britain by CPI Group (UK) Ltd, Croydon, CR0 4YY

5 2013

Dedication

To my family.
Stefania Sesia

To my parents for their sacrifices and unconditional love. To my brother and sisters for their love and continual support. To my friends for being what they are.
Issam Toufik

To the glory of God, who 'so loved the world that He gave His only Son, that whoever believes in Him shall not perish but have eternal life'. — The Bible.
Matthew Baker

Contents

Editors' Biographies xxi

List of Contributors xxiii

Foreword xxvii

Preface xxix

Acknowledgements xxxi

List of Acronyms xxxiii

1 Introduction and Background 1
Thomas Sälzer and Matthew Baker

1.1 The Context for the Long Term Evolution of UMTS 1
1.1.1 Historical Context 1
1.1.2 LTE in the Mobile Radio Landscape 2
1.1.3 The Standardization Process in 3GPP 5
1.2 Requirements and Targets for the Long Term Evolution 7
1.2.1 System Performance Requirements 7
1.2.2 Deployment Cost and Interoperability 12
1.3 Technologies for the Long Term Evolution 14
1.3.1 Multicarrier Technology 14
1.3.2 Multiple Antenna Technology 15
1.3.3 Packet-Switched Radio Interface 16
1.3.4 User Equipment Categories 17
1.3.5 From the First LTE Release to LTE-Advanced 19
1.4 From Theory to Practice 20
References 21

Part I Network Architecture and Protocols **23**

2 Network Architecture **25**
Sudeep Palat and Philippe Godin

2.1 Introduction . . . 25
2.2 Overall Architectural Overview . . . 26
2.2.1 The Core Network . . . 27
2.2.2 The Access Network . . . 30
2.2.3 Roaming Architecture . . . 31
2.3 Protocol Architecture . . . 32
2.3.1 User Plane . . . 32
2.3.2 Control Plane . . . 33
2.4 Quality of Service and EPS Bearers . . . 34
2.4.1 Bearer Establishment Procedure . . . 37
2.4.2 Inter-Working with other RATs . . . 38
2.5 The E-UTRAN Network Interfaces: S1 Interface . . . 40
2.5.1 Protocol Structure over S1 . . . 41
2.5.2 Initiation over S1 . . . 43
2.5.3 Context Management over S1 . . . 43
2.5.4 Bearer Management over S1 . . . 44
2.5.5 Paging over S1 . . . 44
2.5.6 Mobility over S1 . . . 45
2.5.7 Load Management over S1 . . . 47
2.5.8 Trace Function . . . 48
2.5.9 Delivery of Warning Messages . . . 48
2.6 The E-UTRAN Network Interfaces: X2 Interface . . . 49
2.6.1 Protocol Structure over X2 . . . 49
2.6.2 Initiation over X2 . . . 49
2.6.3 Mobility over X2 . . . 51
2.6.4 Load and Interference Management Over X2 . . . 54
2.6.5 UE Historical Information Over X2 . . . 54
2.7 Summary . . . 55
References . . . 55

3 Control Plane Protocols **57**
Himke van der Velde

3.1 Introduction . . . 57
3.2 Radio Resource Control (RRC) . . . 58
3.2.1 Introduction . . . 58
3.2.2 System Information . . . 59
3.2.3 Connection Control within LTE . . . 63
3.2.4 Connected Mode Inter-RAT Mobility . . . 73
3.2.5 Measurements . . . 75
3.2.6 Other RRC Signalling Aspects . . . 78
3.3 PLMN and Cell Selection . . . 78

3.3.1 Introduction . . . 78
3.3.2 PLMN Selection . . . 79
3.3.3 Cell Selection . . . 79
3.3.4 Cell Reselection . . . 80
3.4 Paging . . . 84
3.5 Summary . . . 86
References . . . 86

4 User Plane Protocols **87**
Patrick Fischer, SeungJune Yi, SungDuck Chun and YoungDae Lee

4.1 Introduction to the User Plane Protocol Stack . . . 87
4.2 Packet Data Convergence Protocol (PDCP) . . . 89
4.2.1 Functions and Architecture . . . 89
4.2.2 Header Compression . . . 90
4.2.3 Security . . . 92
4.2.4 Handover . . . 93
4.2.5 Discard of Data Packets . . . 95
4.2.6 PDCP PDU Formats . . . 97
4.3 Radio Link Control (RLC) . . . 98
4.3.1 RLC Entities . . . 99
4.3.2 RLC PDU Formats . . . 105
4.4 Medium Access Control (MAC) . . . 108
4.4.1 MAC Architecture . . . 108
4.4.2 MAC Functions . . . 111
4.5 Summary of the User Plane Protocols . . . 120
References . . . 120

Part II Physical Layer for Downlink **121**

5 Orthogonal Frequency Division Multiple Access (OFDMA) **123**
Andrea Ancora, Issam Toufik, Andreas Bury and Dirk Slock

5.1 Introduction . . . 123
5.1.1 History of OFDM Development . . . 124
5.2 OFDM . . . 125
5.2.1 Orthogonal Multiplexing Principle . . . 125
5.2.2 Peak-to-Average Power Ratio and Sensitivity to Non-Linearity . . . 131
5.2.3 Sensitivity to Carrier Frequency Offset and Time-Varying Channels . 133
5.2.4 Timing Offset and Cyclic Prefix Dimensioning . . . 135
5.3 OFDMA . . . 137
5.4 Parameter Dimensioning . . . 139
5.4.1 Physical Layer Parameters for LTE . . . 140
5.5 Summary . . . 142
References . . . 142

6 Introduction to Downlink Physical Layer Design **145**
Matthew Baker

6.1 Introduction . . . 145
6.2 Transmission Resource Structure . . . 145
6.3 Signal Structure . . . 148
6.4 Introduction to Downlink Operation . . . 149
References . . . 150

7 Synchronization and Cell Search **151**
Fabrizio Tomatis and Stefania Sesia

7.1 Introduction . . . 151
7.2 Synchronization Sequences and Cell Search in LTE . . . 151
7.2.1 Zadoff–Chu Sequences . . . 155
7.2.2 Primary Synchronization Signal (PSS) Sequences . . . 157
7.2.3 Secondary Synchronization Signal (SSS) Sequences . . . 158
7.3 Coherent Versus Non-Coherent Detection . . . 161
References . . . 163

8 Reference Signals and Channel Estimation **165**
Andrea Ancora, Stefania Sesia and Alex Gorokhov

8.1 Introduction . . . 165
8.2 Design of Reference Signals in the LTE Downlink . . . 167
8.2.1 Cell-Specific Reference Signals . . . 168
8.2.2 UE-Specific Reference Signals in Release 8 . . . 171
8.2.3 UE-Specific Reference Signals in Release 9 . . . 171
8.3 RS-Aided Channel Modelling and Estimation . . . 174
8.3.1 Time-Frequency-Domain Correlation: The WSSUS Channel Model . 175
8.3.2 Spatial-Domain Correlation: The Kronecker Model . . . 176
8.4 Frequency-Domain Channel Estimation . . . 178
8.4.1 Channel Estimate Interpolation . . . 178
8.4.2 General Approach to Linear Channel Estimation . . . 179
8.4.3 Performance Comparison . . . 180
8.5 Time-Domain Channel Estimation . . . 181
8.5.1 Finite and Infinite Length MMSE . . . 182
8.5.2 Normalized Least-Mean-Square . . . 184
8.6 Spatial-Domain Channel Estimation . . . 184
8.7 Advanced Techniques . . . 185
References . . . 186

9 Downlink Physical Data and Control Channels **189**
Matthew Baker and Tim Moulsley

9.1 Introduction . . . 189
9.2 Downlink Data-Transporting Channels . . . 189
9.2.1 Physical Broadcast Channel (PBCH) . . . 189
9.2.2 Physical Downlink Shared CHannel (PDSCH) . . . 192

9.2.3 Physical Multicast Channel (PMCH) 196
9.3 Downlink Control Channels 196
9.3.1 Requirements for Control Channel Design 196
9.3.2 Control Channel Structure 198
9.3.3 Physical Control Format Indicator CHannel (PCFICH) 198
9.3.4 Physical Hybrid ARQ Indicator Channel (PHICH) 200
9.3.5 Physical Downlink Control CHannel (PDCCH) 202
9.3.6 PDCCH Scheduling Process 212
References 214

10 Link Adaptation and Channel Coding **215**
Brian Classon, Ajit Nimbalker, Stefania Sesia and Issam Toufik

10.1 Introduction 215
10.2 Link Adaptation and CQI Feedback 217
10.2.1 CQI Feedback in LTE 218
10.3 Channel Coding 223
10.3.1 Theoretical Aspects of Channel Coding 223
10.3.2 Channel Coding for Data Channels in LTE 232
10.3.3 Channel Coding for Control Channels in LTE 244
10.4 Conclusions 245
References 246

11 Multiple Antenna Techniques **249**
Thomas Sälzer, David Gesbert, Cornelius van Rensburg, Filippo Tosato, Florian Kaltenberger and Tetsushi Abe

11.1 Fundamentals of Multiple Antenna Theory 249
11.1.1 Overview 249
11.1.2 MIMO Signal Model 252
11.1.3 Single-User MIMO Techniques 253
11.1.4 Multi-User MIMO Techniques 258
11.2 MIMO Schemes in LTE 262
11.2.1 Practical Considerations 263
11.2.2 Single-User Schemes 264
11.2.3 Multi-User MIMO 274
11.2.4 MIMO Performance 276
11.3 Summary 276
References 277

12 Multi-User Scheduling and Interference Coordination **279**
Issam Toufik and Raymond Knopp

12.1 Introduction 279
12.2 General Considerations for Resource Allocation Strategies 280
12.3 Scheduling Algorithms 283
12.3.1 Ergodic Capacity 283
12.3.2 Delay-Limited Capacity 285

12.4 Considerations for Resource Scheduling in LTE . . . 286
12.5 Interference Coordination and Frequency Reuse . . . 287
12.5.1 Inter-eNodeB Signalling to Support Downlink Frequency-Domain ICIC in LTE . . . 290
12.5.2 Inter-eNodeB Signalling to Support Uplink Frequency-Domain ICIC in LTE . . . 290
12.5.3 Static versus Semi-Static ICIC . . . 291
12.6 Summary . . . 291
References . . . 292

13 Broadcast Operation **293**
Himke van der Velde, Olivier Hus and Matthew Baker

13.1 Introduction . . . 293
13.2 Broadcast Modes . . . 293
13.3 Overall MBMS Architecture . . . 295
13.3.1 Reference Architecture . . . 295
13.3.2 Content Provision . . . 295
13.3.3 Core Network . . . 296
13.3.4 Radio Access Network – E-UTRAN/UTRAN/GERAN and UE . . . 296
13.3.5 MBMS Interfaces . . . 297
13.4 MBMS Single Frequency Network Transmission . . . 297
13.4.1 Physical Layer Aspects . . . 297
13.4.2 MBSFN Areas . . . 301
13.5 MBMS Characteristics . . . 303
13.5.1 Mobility Support . . . 303
13.5.2 UE Capabilities and Service Prioritization . . . 303
13.6 Radio Access Protocol Architecture and Signalling . . . 304
13.6.1 Protocol Architecture . . . 304
13.6.2 Session Start Signalling . . . 305
13.6.3 Radio Resource Control (RRC) Signalling Aspects . . . 306
13.6.4 Content Synchronization . . . 308
13.6.5 Counting Procedure . . . 310
13.7 Public Warning Systems . . . 312
13.8 Comparison of Mobile Broadcast Modes . . . 312
13.8.1 Delivery by Cellular Networks . . . 312
13.8.2 Delivery by Broadcast Networks . . . 313
13.8.3 Services and Applications . . . 313
References . . . 314

Part III Physical Layer for Uplink **315**

14 Uplink Physical Layer Design **317**
Robert Love and Vijay Nangia

14.1 Introduction . . . 317
14.2 SC-FDMA Principles . . . 318

14.2.1 SC-FDMA Transmission Structure . . . 318
14.2.2 Time-Domain Signal Generation . . . 318
14.2.3 Frequency-Domain Signal Generation (DFT-S-OFDM) . . . 320
14.3 SC-FDMA Design in LTE . . . 321
14.3.1 Transmit Processing for LTE . . . 321
14.3.2 SC-FDMA Parameters for LTE . . . 322
14.3.3 d.c. Subcarrier in SC-FDMA . . . 324
14.3.4 Pulse Shaping . . . 324
14.4 Summary . . . 325
References . . . 326

15 Uplink Reference Signals **327**
Robert Love and Vijay Nangia

15.1 Introduction . . . 327
15.2 RS Signal Sequence Generation . . . 328
15.2.1 Base RS Sequences and Sequence Grouping . . . 330
15.2.2 Orthogonal RS via Cyclic Time-Shifts of a Base Sequence . . . 330
15.3 Sequence-Group Hopping and Planning . . . 332
15.3.1 Sequence-Group Hopping . . . 332
15.3.2 Sequence-Group Planning . . . 333
15.4 Cyclic Shift Hopping . . . 333
15.5 Demodulation Reference Signals (DM-RS) . . . 335
15.6 Uplink Sounding Reference Signals (SRS) . . . 337
15.6.1 SRS Subframe Configuration and Position . . . 337
15.6.2 Duration and Periodicity of SRS Transmissions . . . 337
15.6.3 SRS Symbol Structure . . . 338
15.7 Summary . . . 340
References . . . 341

16 Uplink Physical Channel Structure **343**
Robert Love and Vijay Nangia

16.1 Introduction . . . 343
16.2 Physical Uplink Shared Data Channel Structure . . . 344
16.2.1 Scheduling on PUSCH . . . 345
16.2.2 PUSCH Transport Block Sizes . . . 347
16.3 Uplink Control Channel Design . . . 348
16.3.1 Physical Uplink Control Channel (PUCCH) Structure . . . 348
16.3.2 Types of Control Signalling Information and PUCCH Formats . . . 352
16.3.3 Channel State Information Transmission on PUCCH (Format 2) . . . 353
16.3.4 Multiplexing of CSI and HARQ ACK/NACK from a UE on PUCCH 355
16.3.5 HARQ ACK/NACK Transmission on PUCCH (Format 1a/1b) . . . 356
16.3.6 Multiplexing of CSI and HARQ ACK/NACK in the Same (Mixed) PUCCH RB . . . 363
16.3.7 Scheduling Request (SR) Transmission on PUCCH (Format 1) . . . 363
16.4 Multiplexing of Control Signalling and UL-SCH Data on PUSCH . . . 365
16.5 ACK/NACK Repetition . . . 367

16.6 Multiple-Antenna Techniques 367
16.6.1 Closed-Loop Switched Antenna Diversity 367
16.6.2 Multi-User 'Virtual' MIMO or SDMA 368
16.7 Summary 369
References 369

17 Random Access **371**
Pierre Bertrand and Jing Jiang

17.1 Introduction 371
17.2 Random Access Usage and Requirements in LTE 371
17.3 Random Access Procedure 372
17.3.1 Contention-Based Random Access Procedure 373
17.3.2 Contention-Free Random Access Procedure 376
17.4 Physical Random Access Channel Design 376
17.4.1 Multiplexing of PRACH with PUSCH and PUCCH 376
17.4.2 The PRACH Structure 377
17.4.3 Preamble Sequence Theory and Design 385
17.5 PRACH Implementation 396
17.5.1 UE Transmitter 397
17.5.2 eNodeB PRACH Receiver 398
17.6 Time Division Duplex (TDD) PRACH 404
17.6.1 Preamble Format 4 404
17.7 Concluding Remarks 405
References 406

18 Uplink Transmission Procedures **407**
Matthew Baker

18.1 Introduction 407
18.2 Uplink Timing Control 407
18.2.1 Overview 407
18.2.2 Timing Advance Procedure 408
18.3 Power Control 411
18.3.1 Overview 411
18.3.2 Detailed Power Control Behaviour 412
18.3.3 UE Power Headroom Reporting 419
18.3.4 Summary of Uplink Power Control Strategies 420
References 420

Part IV Practical Deployment Aspects **421**

19 User Equipment Positioning **423**
Karri Ranta-aho and Zukang Shen

19.1 Introduction 423
19.2 Assisted Global Navigation Satellite System (A-GNSS) Positioning 425
19.3 Observed Time Difference Of Arrival (OTDOA) Positioning 426

19.3.1 Positioning Reference Signals (PRS) . . . 427
19.3.2 OTDOA Performance and Practical Considerations . . . 430
19.4 Cell-ID-based Positioning . . . 431
19.4.1 Basic CID Positioning . . . 431
19.4.2 Enhanced CID Positioning using Round Trip Time and UE Receive Level Measurements . . . 431
19.4.3 Enhanced CID Positioning using Round Trip Time and Angle of Arrival . . . 432
19.5 LTE Positioning Protocols . . . 433
19.6 Summary and Future Techniques . . . 435
References . . . 436

20 The Radio Propagation Environment **437**
Juha Ylitalo and Tommi Jämsä

20.1 Introduction . . . 437
20.2 SISO and SIMO Channel Models . . . 438
20.2.1 ITU Channel Model . . . 439
20.2.2 3GPP Channel Model . . . 440
20.2.3 Extended ITU Models . . . 440
20.3 MIMO Channel Models . . . 441
20.3.1 SCM Channel Model . . . 442
20.3.2 SCM-Extension Channel Model . . . 444
20.3.3 WINNER Model . . . 445
20.3.4 LTE Evaluation Model . . . 446
20.3.5 Extended ITU Models with Spatial Correlation . . . 448
20.3.6 ITU Channel Models for IMT-Advanced . . . 449
20.3.7 Comparison of MIMO Channel Models . . . 453
20.4 Radio Channel Implementation for Conformance Testing . . . 454
20.4.1 Performance and Conformance Testing . . . 454
20.4.2 Future Testing Challenges . . . 454
20.5 Concluding Remarks . . . 455
References . . . 455

21 Radio Frequency Aspects **457**
Moray Rumney, Takaharu Nakamura, Stefania Sesia, Tony Sayers and Adrian Payne

21.1 Introduction . . . 457
21.2 Frequency Bands and Arrangements . . . 459
21.3 Transmitter RF Requirements . . . 462
21.3.1 Requirements for the Intended Transmissions . . . 462
21.3.2 Requirements for Unwanted Emissions . . . 467
21.3.3 Power Amplifier Considerations . . . 471
21.4 Receiver RF Requirements . . . 474
21.4.1 Receiver General Requirements . . . 474
21.4.2 Transmit Signal Leakage . . . 475
21.4.3 Maximum Input Level . . . 477
21.4.4 Small Signal Requirements . . . 478

21.4.5 Selectivity and Blocking Specifications . . . 482
21.4.6 Spurious Emissions . . . 488
21.4.7 Intermodulation Requirements . . . 489
21.4.8 Dynamic Range . . . 491
21.5 RF Impairments . . . 492
21.5.1 Transmitter RF Impairments . . . 492
21.5.2 Model of the Main RF Impairments . . . 495
21.6 Summary . . . 500
References . . . 501

22 Radio Resource Management **503**
Muhammad Kazmi

22.1 Introduction . . . 503
22.2 Cell Search Performance . . . 505
22.2.1 Cell Search within E-UTRAN . . . 505
22.2.2 E-UTRAN to E-UTRAN Cell Global Identifier Reporting Requirements . . . 509
22.2.3 E-UTRAN to UTRAN Cell Search . . . 510
22.2.4 E-UTRAN to GSM Cell Search . . . 511
22.2.5 Enhanced Inter-RAT Measurement Requirements . . . 512
22.3 Mobility Measurements . . . 513
22.3.1 E-UTRAN Measurements . . . 513
22.3.2 UTRAN Measurements . . . 514
22.3.3 GSM Measurements: GSM Carrier RSSI . . . 516
22.3.4 CDMA2000 Measurements . . . 516
22.4 UE Measurement Reporting Mechanisms and Requirements . . . 516
22.4.1 E-UTRAN Event Triggered Reporting Requirements . . . 517
22.4.2 Inter-RAT Event-Triggered Reporting . . . 517
22.5 Mobility Performance . . . 518
22.5.1 Mobility Performance in RRC_IDLE State . . . 518
22.5.2 Mobility Performance in RRC_CONNECTED State . . . 522
22.6 RRC Connection Mobility Control Performance . . . 525
22.6.1 RRC Connection Re-establishment . . . 525
22.6.2 Random Access . . . 525
22.7 Radio Link Monitoring Performance . . . 526
22.7.1 In-sync and Out-of-sync Thresholds . . . 526
22.7.2 Requirements without DRX . . . 527
22.7.3 Requirements with DRX . . . 527
22.7.4 Requirements during Transitions . . . 527
22.8 Concluding Remarks . . . 528
References . . . 529

23 Paired and Unpaired Spectrum **531**
Nicholas Anderson

23.1 Introduction . . . 531
23.2 Duplex Modes . . . 532

23.3 Interference Issues in Unpaired Spectrum 533
23.3.1 Adjacent Carrier Interference Scenarios 535
23.3.2 Summary of Interference Scenarios 543
23.4 Half-Duplex System Design Aspects 544
23.4.1 Accommodation of Transmit–Receive Switching 544
23.4.2 Coexistence between Dissimilar Systems 547
23.4.3 HARQ and Control Signalling for TDD Operation 548
23.4.4 Half-Duplex FDD (HD-FDD) Physical Layer Operation 551
23.5 Reciprocity 552
23.5.1 Conditions for Reciprocity 554
23.5.2 Applications of Reciprocity 558
23.5.3 Summary of Reciprocity Considerations 561
References 562

24 Picocells, Femtocells and Home eNodeBs **563**
Philippe Godin and Nick Whinnett

24.1 Introduction 563
24.2 Home eNodeB Architecture 564
24.2.1 Architecture Overview 564
24.2.2 Functionalities 565
24.2.3 Mobility 566
24.2.4 Local IP Access Support 568
24.3 Interference Management for Femtocell Deployment 569
24.3.1 Interference Scenarios 570
24.3.2 Network Listen Mode 574
24.4 RF Requirements for Small Cells 574
24.4.1 Transmitter Specifications 575
24.4.2 Receiver Specifications 576
24.4.3 Demodulation Performance Requirements 578
24.4.4 Time Synchronization for TDD Operation 579
24.5 Summary 580
References 580

25 Self-Optimizing Networks **581**
Philippe Godin

25.1 Introduction 581
25.2 Automatic Neighbour Relation Function (ANRF) 582
25.2.1 Intra-LTE ANRF 582
25.2.2 Automatic Neighbour Relation Table 583
25.2.3 Inter-RAT or Inter-Frequency ANRF 583
25.3 Self-Configuration of eNodeB and MME 584
25.3.1 Self-Configuration of eNodeB/MME over S1 585
25.3.2 Self-Configuration of IP address and X2 interface 585
25.4 Automatic Configuration of Physical Cell Identity 587
25.5 Mobility Load Balancing Optimization 587

25.5.1 Intra-LTE Load Exchange . . . 588
25.5.2 Intra-LTE Handover Parameter Optimization . . . 589
25.5.3 Inter-RAT Load Exchange . . . 590
25.5.4 Enhanced Inter-RAT Load Exchange . . . 590
25.6 Mobility Robustness Optimization . . . 591
25.6.1 Too-Late Handover . . . 591
25.6.2 Coverage Hole Detection . . . 591
25.6.3 Too-Early Handover . . . 592
25.6.4 Handover to an Inappropriate Cell . . . 592
25.6.5 MRO Verdict Improvement . . . 593
25.6.6 Handover to an Unprepared Cell . . . 594
25.6.7 Unnecessary Inter-RAT Handovers . . . 594
25.6.8 Potential Remedies for Identified Mobility Problems . . . 595
25.7 Random Access CHannel (RACH) Self-Optimization . . . 595
25.8 Energy Saving . . . 596
25.9 Emerging New SON Use Cases . . . 597
References . . . 598

26 LTE System Performance **599**
Tetsushi Abe

26.1 Introduction . . . 599
26.2 Factors Contributing to LTE System Capacity . . . 599
26.2.1 Multiple Access Techniques . . . 600
26.2.2 Frequency Reuse and Interference Management . . . 600
26.2.3 Multiple Antenna Techniques . . . 601
26.2.4 Semi-Persistent Scheduling . . . 601
26.2.5 Short Subframe Duration and Low HARQ Round Trip Time . . . 602
26.2.6 Advanced Receivers . . . 602
26.2.7 Layer 1 and Layer 2 Overhead . . . 602
26.3 LTE Capacity Evaluation . . . 603
26.3.1 Downlink and Uplink Spectral Efficiency . . . 605
26.3.2 VoIP Capacity . . . 608
26.4 LTE Coverage and Link Budget . . . 608
26.5 Summary . . . 610
References . . . 611

Part V LTE-Advanced **613**

27 Introduction to LTE-Advanced **615**
Dirk Gerstenberger

27.1 Introduction and Requirements . . . 615
27.2 Overview of the Main Features of LTE-Advanced . . . 618
27.3 Backward Compatibility . . . 619
27.4 Deployment Aspects . . . 620
27.5 UE Categories for LTE-Advanced . . . 621

References 622

28 Carrier Aggregation **623**
Juan Montojo and Jelena Damnjanovic

28.1 Introduction 623
28.2 Protocols for Carrier Aggregation 624
28.2.1 Initial Acquisition, Connection Establishment and CC Management . 624
28.2.2 Measurements and Mobility 625
28.2.3 User Plane Protocols 628
28.3 Physical Layer Aspects 631
28.3.1 Downlink Control Signalling 631
28.3.2 Uplink Control Signalling 636
28.3.3 Sounding Reference Signals 642
28.3.4 Uplink Timing Advance 642
28.3.5 Uplink Power Control 642
28.3.6 Uplink Multiple Access Scheme Enhancements 644
28.4 UE Transmitter and Receiver Aspects 648
28.4.1 UE Transmitter Aspects of Carrier Aggregation 648
28.4.2 UE Receiver Aspects of Carrier Aggregation 648
28.4.3 Prioritized Carrier Aggregation Scenarios 649
28.5 Summary 650
References 650

29 Multiple Antenna Techniques for LTE-Advanced **651**
Alex Gorokhov, Amir Farajidana, Kapil Bhattad, Xiliang Luo and Stefan Geirhofer

29.1 Downlink Reference Signals 651
29.1.1 Downlink Reference Signals for Demodulation 652
29.1.2 Downlink Reference Signals for Estimation of Channel State Information (CSI-RS) 654
29.2 Uplink Reference Signals 657
29.2.1 Uplink DeModulation Reference Signals (DM-RS) 657
29.2.2 Sounding Reference Signals (SRSs) 658
29.3 Downlink MIMO Enhancements 659
29.3.1 Downlink 8-Antenna Transmission 659
29.3.2 Enhanced Downlink Multi-User MIMO 661
29.3.3 Enhanced CSI Feedback 662
29.4 Uplink Multiple Antenna Transmission 666
29.4.1 Uplink SU-MIMO for PUSCH 666
29.4.2 Uplink Transmit Diversity for PUCCH 668
29.5 Coordinated MultiPoint (CoMP) Transmission and Reception 669
29.5.1 Cooperative MIMO Schemes and Scenarios 669
29.6 Summary 671
References 671

30 Relaying **673**
Eric Hardouin, J. Nicholas Laneman, Alexander Golitschek, Hidetoshi Suzuki, Osvaldo Gonsa
30.1 Introduction . . . 673
30.1.1 What is Relaying? . . . 673
30.1.2 Characteristics of Relay Nodes . . . 675
30.1.3 Protocol Functionality of Relay Nodes . . . 676
30.1.4 Relevant Deployment Scenarios . . . 677
30.2 Theoretical Analysis of Relaying . . . 679
30.2.1 Relaying Strategies and Benefits . . . 679
30.2.2 Duplex Constraints and Resource Allocation . . . 683
30.3 Relay Nodes in LTE-Advanced . . . 684
30.3.1 Types of RN . . . 684
30.3.2 Backhaul and Access Resource Sharing . . . 685
30.3.3 Relay Architecture . . . 687
30.3.4 RN Initialization and Configuration . . . 689
30.3.5 Random Access on the Backhaul Link . . . 690
30.3.6 Radio Link Failure on the Backhaul Link . . . 690
30.3.7 RN Security . . . 690
30.3.8 Backhaul Physical Channels . . . 691
30.3.9 Backhaul Scheduling . . . 696
30.3.10 Backhaul HARQ . . . 698
30.4 Summary . . . 699
References . . . 699

31 Additional Features of LTE Release 10 **701**
Teck Hu, Philippe Godin and Sudeep Palat
31.1 Introduction . . . 701
31.2 Enhanced Inter-Cell Interference Coordination . . . 701
31.2.1 LTE Interference Management . . . 703
31.2.2 Almost Blank Subframes . . . 703
31.2.3 X2 Interface Enhancements for Time-Domain ICIC . . . 705
31.2.4 UE Measurements in Time-Domain ICIC Scenarios . . . 706
31.2.5 RRC Signalling for Restricted Measurements . . . 708
31.2.6 ABS Deployment Considerations . . . 709
31.3 Minimization of Drive Tests . . . 710
31.3.1 Logged MDT . . . 711
31.3.2 Immediate MDT . . . 712
31.4 Machine-Type Communications . . . 712
References . . . 714

32 LTE-Advanced Performance and Future Developments **715**
Takehiro Nakamura and Tetsushi Abe
32.1 LTE-Advanced System Performance . . . 715
32.2 Future Developments . . . 718
References . . . 720

Index **721**

Editors' Biographies

Matthew Baker holds degrees in Engineering and Electrical and Information Sciences from the University of Cambridge. From 1996 to 2009 he worked at Philips Research where he conducted leading-edge research into a variety of wireless communication systems and techniques, including propagation modelling, DECT, Hiperlan and UMTS, as well as leading the Philips RAN standardization team. He has been actively participating in the standardization of both UMTS WCDMA and LTE in 3GPP since 1999, where he has been active in 3GPP TSG RAN Working Groups 1, 2, 4 and 5, contributing several hundred proposals. He now works for Alcatel-Lucent, which he joined in 2009, and he has been Chairman of 3GPP TSG RAN Working Group 1 since being elected to the post in August of that year. He is the author of several international conference papers and inventor of numerous patents. He is a Chartered Engineer, a Member of the Institution of Engineering and Technology and a Visiting Lecturer at the University of Reading, UK.

Stefania Sesia received her Ph.D. degree in Communication Systems and Coding Theory from both Eurecom (Sophia Antipolis, France) and ENST-Paris (Paris, France) in 2005. From 2002 to 2005 she worked at Motorola Research Labs, Paris, towards her Ph.D. thesis. In June 2005 she joined Philips/NXP Semiconductors (now ST-Ericsson) Research and Development Centre in Sophia Antipolis, France where she was technical leader and responsible for the High Speed Downlink Packet Access algorithm development. She has been participating in 3GPP TSG RAN Working Groups 1 and 4 standardization meetings. From 2007 to 2009 she was on secondment from NXP Semiconductors to the European Telecommunications Standard Institute (ETSI) acting as 3GPP TSG RAN and 3GPP TSG RAN Working Group 4 Technical Officer. She is currently back in ST-Ericsson as senior research and development engineer, actively participating in 3GPP TSG RAN Working Group 4 as a delegate. She is the author of several international IEEE conference and journal papers and many contributions to 3GPP, and inventor of numerous US and European patents.

Issam Toufik graduated in Telecommunications Engineering (majoring in Mobile Communication Systems) in 2002 from both ENST-Bretagne (Brest, France) and Eurecom (Sophia Antipolis, France). In 2006, he received his Ph.D. degree in Communication Systems from Eurecom/ENST-Paris, France. From June to August 2005 he worked for Samsung Advanced Institute of Technology (SAIT), South Korea, as a Research Engineer on LTE. In January 2007, he joined NXP Semiconductors/ST-Ericsson, Sophia Antipolis, France, as a Research and Development Engineer for UMTS and LTE algorithm development. In November 2009, he joined the European Telecommunications Standard Institute (ETSI) acting as 3GPP TSG

RAN and 3GPP TSG RAN Working Group 4 Technical Officer. He is the author of several international IEEE conference and journal papers and contributions to 3GPP, and inventor of numerous patents.

List of Contributors

Abe, Tetsushi, NTT DOCOMO
e-mail: abetet@nttdocomo.com

Ancora, Andrea, ST-Ericsson
e-mail: andrea.ancora@stericsson.com

Anderson, Nicholas, Research In Motion
e-mail: nianderson@rim.com

Baker, Matthew, Alcatel-Lucent
e-mail: matthew.baker@alcatel-lucent.com, m.p.j.baker.92@cantab.net

Bertrand, Pierre, Texas Instruments
e-mail: p-bertrand@ti.com

Bhattad, Kapil, Qualcomm
e-mail: kbhattad@qualcomm.com

Bury, Andreas, Blue Wonder Communications
e-mail: andreas.bury@bluwo.com

Chun, SungDuck,LG Electronics
e-mail: duckychun@lge.com

Classon, Brian, Huawei
e-mail: brian.classon@huawei.com

Damnjanovic, Jelena, Qualcomm
e-mail: jelenad@qualcomm.com

Farajidana, Amir, Qualcomm
e-mail: amirf@qualcomm.com

Fischer, Patrick, Bouygues Telecom
e-mail: pfischer@bouyguestelecom.fr

Geirhofer, Stefan, Qualcomm
e-mail: sgeirhofer@qualcomm.com

Gerstenberger, Dirk, Ericsson
e-mail: dirk.gerstenberger@ericsson.com

Gesbert, David, Eurecom
e-mail: david.gesbert@eurecom.fr

Godin, Philippe, Alcatel-Lucent
e-mail: philippe.godin@alcatel-lucent.com

Golitschek, Alexander, Panasonic
e-mail: alexander.golitschek@eu.panasonic.com

Gonsa, Osvaldo, Panasonic
e-mail: osvaldo.gonsa@eu.panasonic.com

Gorokhov, Alex, Qualcomm
e-mail: gorokhov@qualcomm.com

Hardouin, Eric, Orange Labs
e-mail: eric.hardouin@orange-ftgroup.com

Hu, Teck, Alcatel-Lucent
e-mail: teck.hu@alcatel-lucent.com

Hus, Olivier,
e-mail: olivierjhus@gmail.com

Jämsä, Tommi, Elektrobit
e-mail: tommi.jamsa@elektrobit.com

Jiang, Jing, Texas Instruments
e-mail: jing.jiang@ti.com

Kaltenberger, Florian, Eurecom
e-mail: florian.kaltenberger@eurecom.fr

Kazmi, Muhammad, Ericsson
e-mail: muhammad.kazmi@ericsson.com

Knopp, Raymond, Eurecom
e-mail: raymond.knopp@eurecom.fr

Laneman, J. Nicholas, University of Notre Dame
e-mail: jnl@nd.edu

Lee, YoungDae, LG Electronics
e-mail: leego@lge.com

Love, Robert, Motorola Mobility
e-mail: robert.love@motorola.com

Luo, Xiliang, Qualcomm
e-mail: xluo@qualcomm.com

Montojo, Juan, Qualcomm
e-mail: juanm@qualcomm.com

Moulsley, Tim, Fujitsu
e-mail: t.moulsley@btopenworld.com

Nakamura, Takaharu, Fujitsu
e-mail: n.takaharu@jp.fujitsu.com

Nakamura, Takehiro, NTT DOCOMO
e-mail: nakamurata@nttdocomo.co.jp

Nangia, Vijay, Motorola Mobility
e-mail: vijay.nangia@motorola.com

Nimbalker, Ajit, Motorola Mobility
e-mail: ajit.nimbalker@motorola.com

Palat, K. Sudeep, Alcatel-Lucent
e-mail: spalat@alcatel-lucent.com

Payne, Adrian, ERA Technology
e-mail: adrian.w.payne@gmail.com

Ranta-aho, Karri, Nokia Siemens Networks
e-mail: karri.ranta-aho@nsn.com

Rumney, Moray, Agilent
e-mail: moray_rumney@agilent.com

Sälzer, Thomas, Huawei
e-mail: thomas.salzer@huawei.com, thomas.salzer@gmx.de

Sayers, Tony, Ultra Electronics
e-mail: tony.sayers@talktalk.net

Sesia, Stefania, ST-Ericsson
e-mail: stefania.sesia@stericsson.com

Shen, Zukang, CATT
e-mail: shenzukang@catt.cn

Slock, Dirk, Eurecom
e-mail: dirk.slock@eurecom.fr

Suzuki, Hidetoshi, Panasonic
e-mail: Suzuki.Hidetoshi@jp.panasonic.com

Tomatis, Fabrizio, ST-Ericsson
e-mail: fabrizio.tomatis@stericsson.com

Tosato, Filippo, Toshiba
e-mail: filippo.tosato@toshiba-trel.com

Toufik, Issam, ETSI
e-mail: issam.toufik@etsi.org, issam.toufik@eurecom.fr

van der Velde, Himke, Samsung
e-mail: himke.vandervelde@samsung.com

van Rensburg, Cornelius, Huawei
e-mail: cdvanren@ieee.org

Whinnett, Nick, Picochip
e-mail: nickw@picochip.com

Yi, SeungJune, LG Electronics
e-mail: seungjune@lge.com

Ylitalo, Juha, Elektrobit
e-mail: juha.ylitalo@elektrobit.com

Foreword

GSM, and its evolution through GPRS, EDGE, WCDMA and HSPA, is the technology stream of choice for the vast majority of the world's mobile operators. Users have experienced increasing data rates, together with a dramatic reduction in telecommunications charges; they now expect to pay less but receive more. Therefore, in deciding the next steps, there must be a dual approach: seeking considerable performance improvement but at reduced cost. Improved performance must be delivered through systems which are cheaper to install and maintain. LTE and LTE-Advanced represent these next steps and will be the basis on which future mobile telecommunications systems will be built.

Many articles have already been published on the subject of LTE, varying from doctoral theses to network operator analyses and manufacturers' product literature. By their very nature, those publications have viewed the subject from one particular perspective, be it academic, operational or promotional. A very different approach is taken with this book. The authors come from a number of different spheres within the mobile telecommunications ecosystem and collectively bring a refreshing variety of perspectives. What binds the authors together is a thorough knowledge of the subject material which they have derived from their long experience within the standards-setting environment, the 3rd Generation Partnership Project (3GPP). LTE discussions started within 3GPP in 2004, so it is not a particularly new subject. In order to fully appreciate the thinking that conceived this technology, however, it is necessary to have followed the subject from the very beginning and to have witnessed the discussions that took place from the outset. Moreover, it is important to understand the thread that links academia, through research to standardization since it is widely acknowledged that by this route impossible dreams become market realities. Considerable research work has taken place to prove the viability of the technical basis on which LTE is founded and it is essential to draw on that research if any attempt is made to explain LTE to a wider audience. The authors of this book have not only followed the LTE story from the beginning but many have also been active players in WCDMA and its predecessors, in which LTE has its roots.

This book provides a thorough, authoritative and complete tutorial of the LTE system, now fully updated and extended to include LTE-Advanced. It gives a detailed explanation of the advances made in our theoretical understanding and the practical techniques that will ensure the success of this ground-breaking new radio access technology. Where this book is exceptional is that the reader will learn not just how LTE works but why it works.

I am confident that this book will earn its rightful place on the desk of anyone who needs a thorough understanding of the LTE and LTE-Advanced technology, the basis of the world's mobile telecommunications systems for the next decade.

Adrian Scrase, ETSI Vice-President,
International Partnership Projects

Preface to the Second Edition

> Research workers and engineers toil unceasingly on the development of wireless telegraphy. Where this development can lead, we know not. However, with the results already achieved, telegraphy over wires has been extended by this invention in the most fortunate way. Independent of fixed conductor routes and independent of space, we can produce connections between far-distant places, over far-reaching waters and deserts. This is the magnificent practical invention which has flowered upon one of the most brilliant scientific discoveries of our time!

These words accompanied the presentation of the Nobel Prize for Physics to Guglielmo Marconi in December 1909.

Marconi's success was the practical and commercial realization of wireless telegraphy – the art of sending messages without wires – thus exploiting for the first time the amazing capability for wireless communication built into our universe. While others worked on wireless telephony – the transmission of audio signals for voice communication – Marconi interestingly saw no need for this. He believed that the transmission of short text messages was entirely sufficient for keeping in touch.

One could be forgiven for thinking that the explosion of wireless voice communication in the intervening years has proved Marconi wrong; but the resurgence of wireless data transmission at the close of the twentieth century, beginning with the mobile text messaging phenomenon, or 'SMS', reveals in part the depth of insight Marconi possessed.

Nearly 100 years after Marconi received his Nobel prize, the involvement of thousands of engineers around the world in major standardization initiatives such as the 3rd Generation Partnership Project (3GPP) is evidence that the same unceasing toil of research workers and engineers continues apace.

While the first mobile communications standards focused primarily on voice communication, the emphasis now has returned to the provision of systems optimized for data. This trend began with the 3rd Generation Wideband Code Division Multiple Access (WCDMA) system designed in the 3GPP, and is now reaching fulfilment in its successor, the Long-Term Evolution (LTE). LTE was the first cellular communication system optimized from the outset to support packet-switched data services, within which packetized voice communications are just one part. Thus LTE can truly be said to be the heir to Marconi's heritage – the system, unknown indeed to the luminaries of his day, to which his developments have led.

LTE is an enabler. It is not technology for technology's sake, but technology with a purpose, connecting people and information to enable greater things to be achieved. It is already providing higher data rates than ever previously achieved in mobile communications,

combined with wide-area coverage and seamless support for mobility without regard for the type of data being transmitted. To provide this level of functionality and flexibility, it is inevitable that the complexities of the LTE system have far surpassed anything Marconi could have imagined.

One aim of this book, therefore, is to chart an explanatory course through the LTE specifications, to support those who design LTE equipment.

The LTE specification documents themselves do not tell the whole story. Essentially they are a record of decisions taken – decisions which are often compromises between performance and cost, theoretical possibility and practical constraints. We aim therefore to give the reader a detailed insight into the evaluations and trade-offs which lie behind the technology choices inherent in LTE. The specifications also continue to develop, as new releases are produced, and this Second Edition is therefore fully updated to cover Release 9 and the first release of LTE-Advanced, Release 10.

Since the first version of LTE was developed, the theoretical understanding which gave rise to LTE has continued to advance, as the 'unceasing toil' of thousands of engineers continues with the aim of keeping pace with the explosive growth of mobile data traffic. Where the first version of LTE exploited Multiple-Input Multiple-Output (MIMO) antenna techniques to deliver high data rates, the evolution of LTE towards LTE-Advanced extends such techniques further for both downlink and uplink communication, together with support for yet wider bandwidths; meanwhile, heterogeneous (or hierarchical) networks, relaying and Coordinated MultiPoint (CoMP) transmission and reception start to become relevant in LTE-Advanced.

It is particularly these advances in underlying scientific understanding which this book seeks to highlight.

In selecting the technologies to include in LTE and LTE-Advanced, an important consideration is the trade-off between practical benefit and cost of implementation. Fundamental to this assessment is ongoing enhancement in understanding of the radio propagation environment and scenarios of relevance to deployments of LTE and LTE-Advanced. This has been built on significant advances in radio-channel modelling.

The advances in techniques and theoretical understanding continue to be supported by developments in integrated circuit technology and signal processing power which render them feasible where they would have been unthinkable only a few years ago.

Changes in spectrum availability and regulation also influence the development path of LTE towards LTE-Advanced, reinforcing the need for the new technology to be adaptable, capable of being scaled and enhanced to meet new global requirements and deployed in a wide range of different configurations.

With this breadth and depth in mind, the authorship of the chapters of the second edition of this book is even wider than that of the first edition, and again is drawn from all fields of the ecosystem of research and development that has underpinned the design of LTE. They work in the 3GPP standardization itself, in the R&D departments of companies active in LTE, for network operators as well as equipment manufacturers, in universities and in other collaborative research projects. They are uniquely placed to share their insights from the full range of perspectives.

To borrow Marconi's words, where LTE and LTE-Advanced will lead, we know not; but we can be sure that these will not be the last developments in wireless telegraphy.

Matthew Baker, *Stefania Sesia* and *Issam Toufik*

Acknowledgements

Like the first edition, the fully updated and expanded second edition of this book is first and foremost the fruit of a significant team effort, which would not have been successful without the expertise and professionalism displayed by all the contributors, as well as the support of their companies. The dedication of all the co-authors to their task, their patience and flexibility in allowing us to modify and move certain parts of their material for harmonization purposes, are hereby gratefully acknowledged. Particular thanks are due to ST-Ericsson, Alcatel-Lucent and ETSI for giving us the encouragement and working environment to facilitate such a time-consuming project. The help provided by ETSI, 3GPP and others in authorizing us to reproduce certain copyrighted material is also gratefully acknowledged. We would like to express our gratitude to the many experts who kindly provided advice, feedback, reviews and other valuable assistance. We believe their input in all its forms has made this book a more accurate, valuable and even enjoyable resource. These experts include Jacques Achard, Kevin Baum, Martin Beale, Keith Blankenship, Yufei Blankenship, Federico Boccardi, Kevin Boyle, Sarah Boumendil, Alec Brusilovsky, Paul Bucknell, Richard Burbidge, Aaron Byman, Emilio Calvanese Strinati, Choo Chiap Chiau, Anand Dabak, Peter Darwood, Merouane Debbah, Vip Desai, Marko Falck, Antonella Faniuolo, Jeremy Gosteau, Lajos Hanzo, Lassi Hentilä, Shin Horng Wong, Paul Howard, Howard Huang, Alan Jones, Yoshihisa Kishiyama, Achilles Kogiantis, Pekka Kyösti, Daniel Larsson, Jung-Ah Lee, Thierry Lestable, Gert-Jan van Lieshout, Andrew Lillie, Matti Limingoja, Huiheng Mai, Caroline Mathieson, Darren McNamara, Juha Meinilä, Tarik Muharemovic, Gunnar Nitsche, Jukka-Pekka Nuutinen, SungJun Park, Roope Parviainen, Paul Piggin, Claudio Rey, Safouane Sfar, Ken Stewart, Miloš Tesanovic, Paolo Toccacelli, Ludo Tolhuizen, Li Wang, Tim Wilkinson and Steve Zhang.

We would also like to acknowledge the efforts of all participants in 3GPP who, through innumerable contributions and intense discussions often late into the night, facilitated the completion of the LTE specifications for Releases 8, 9 and 10 in such a short space of time.

We would especially like to thank the publishing team at John Wiley & Sons, especially Tiina Ruonamaa, Susan Barclay, Jasmine Chang, Mariam Cheok, Sheena Deuchars, Caitlin Flint, Sarah Hinton, Anna Smart and Sarah Tilley for their professionalism and extensive support and encouragement throughout the preparation of both the first and second editions of this book.

Finally, it should be noted that this book is intended only as a guide to LTE and LTE-Advanced, and the reader should refer to the specifications published by 3GPP for definitive information. Any views expressed in this book are those of the authors and do not necessarily reflect the views of their companies. The editors welcome any suggestions to improve future editions of this book.

The Editors

List of Acronyms

* An asterisk indicates that the acronym can have different meanings depending on the context. The meaning is clearly indicated in the text when used.

3GPP 3rd Generation Partnership Project

3GPP2 3rd Generation Partnership Project 2

ABS Almost Blank Subframe

AC Access Class

ACI Adjacent Channel Interference

ACIR Adjacent Channel Interference Ratio

ACK Acknowledgement

ACLR Adjacent Channel Leakage Ratio

ACS Adjacent Channel Selectivity

ADC Analogue to Digital Converter

ADSL Asymmetric Digital Subscriber Line

AGI Antenna Gain Imbalance

A-GNSS Assisted Global Navigation Satellite System

AM Acknowledged Mode

AMC Adaptive Modulation and Coding

AMPS Analogue Mobile Phone System

AMR Adaptive MultiRate

ANR Automatic Neighbour Relation

ANRF Automatic Neighbour Relation Function

AoA Angle-of-Arrival

AoD Angle-of-Departure

APN Access Point Name

APP A-Posteriori Probability

ARFCN Absolute Radio Frequency Channel Number

ARIB Association of Radio Industries and Businesses

ARP Almost Regular Permutation*

ARP Allocation and Retention Priority*

ARQ Automatic Repeat reQuest

AS Access Stratum*

AS Angular Spread*

A-SEM Additional SEM

ATDMA Advanced TDMA

ATIS Alliance for Telecommunications Industry Solutions

AuC Authentication Centre

AWGN Additive White Gaussian Noise

BCC Base station Colour Code

BCH Broadcast CHannel

BCCH Broadcast Control CHannel

BCJR Algorithm named after its inventors, Bahl, Cocke, Jelinek and Raviv

BER Bit Error Rate

BLER BLock Error Rate

BM-SC Broadcast-Multicast Service Centre

BP Belief Propagation

BPRE Bits Per Resource Element

bps bits per second

BPSK Binary Phase Shift Keying

BSIC Base Station Identification Code

BSR Buffer Status Reports

CAPEX CAPital EXpenditure

CAZAC Constant Amplitude Zero AutoCorrelation

CB Circular Buffer
CBF Coordinated Beamforming
CC Component Carrier
CCCH Common Control CHannel
CCE Control Channel Element
CCI Co-Channel Interference
CCO Cell Change Order
CCSA China Communications Standards Association
CDD Cyclic Delay Diversity
CDF Cumulative Distribution Function
CDL Clustered Delay Line
CDM Code Division Multiplex(ed/ing)
CDMA Code Division Multiple Access
C/I Carrier-to-Interference ratio
CID Cell ID
CIF Carrier Indicator Field
CF Contention-Free
CFI Control Format Indicator
CFO Carrier Frequency Offset
CINR Carrier-to-Interference-and-Noise Ratio
CIR Channel Impulse Response
CM Cubic Metric
CMAS Commercial Mobile Alert Service
CMHH Constant Modulus HouseHolder
CN Core Network
CoMP Coordinated MultiPoint
CODIT UMTS Code DIvision Testbed
COFDM Coded OFDM
CP Cyclic Prefix
CPICH Common PIlot CHannel
CPR Common Phase Rotation
CPT Control PDU Type
CQI Channel Quality Indicator
CRC Cyclic Redundancy Check
CRE Cell Range Expansion
C-RNTI Cell Radio Network Temporary Identifier
CRS Common Reference Signal
CS Circuit-Switched*
CS Cyclic Shift*
CSA Common Subframe Allocation
CSG Closed Subscriber Group
CSI Channel State Information
CSI-RS Channel State Information RS
CSIT Channel State Information at the Transmitter
CTF Channel Transfer Function
CVA Circular Viterbi Algorithm
CVQ Channel Vector Quantization
CW Continuous-Wave
DAB Digital Audio Broadcasting
DAC Digital to Analogue Converter
DAI Downlink Assignment Index
dB deci-Bel
d.c. direct current
DCCH Dedicated Control CHannel
DCFB Direct Channel FeedBack
DCI Downlink Control Information
DFT Discrete Fourier Transform
DFT-S-OFDM DFT-Spread OFDM
Diffserv Differentiated Services
DL DownLink
DL-SCH DownLink Shared CHannel
DMB Digital Mobile Broadcasting
DM-RS DeModulation-RS
DOA Direction Of Arrival
DPC Dirty-Paper Coding
DRB Data Radio Bearer
DRX Discontinuous Reception
DS-CDMA Direct-Sequence Code Division Multiple Access
DSP Digital Signal Processor
DTCH Dedicated Traffic CHannel
DTX Discontinuous Transmission
DVB-H Digital Video Broadcasting – Handheld
DVB-T Digital Video Broadcasting – Terrestrial
DwPTS Downlink Pilot TimeSlot
ECGI E-UTRAN Cell Global Identifier

ECM EPS Connection Management

EDGE Enhanced Data rates for GSM Evolution

EESM Exponential Effective SINR Mapping

eICIC enhanced Inter-Cell Interference Coordination

EMEA Europe, Middle East and Africa

EMM EPS Mobility Management

eNodeB evolved NodeB

EPA Extended Pedestrian A

EPC Evolved Packet Core

EPG Electronic Programme Guide

ePHR extended Power Headroom Report

EPS Evolved Packet System

E-RAB E-UTRAN Radio Access Bearer

E-SMLC Evolved Serving Mobile Location Centre

ESP Encapsulating Security Payload

ETSI European Telecommunications Standards Institute

ETU Extended Typical Urban

ETWS Earthquake and Tsunami Warning System

E-UTRA Evolved-UTRA

E-UTRAN Evolved-UTRAN

EVA Extended Vehicular A

EVM Error Vector Magnitude

FACH Forward Access CHannel

FB Frequency Burst

FCC Federal Communications Commission

FCCH Frequency Control CHannel

FDD Frequency Division Duplex

FDE Frequency-Domain Equalizer

FDM Frequency Division Multiplexing

FDMA Frequency Division Multiple Access

FDSS Frequency-Domain Spectral Shaping

FFT Fast Fourier Transform

FI Framing Info

FIR Finite Impulse Response

FMS First Missing SDU

FSTD Frequency Switched Transmit Diversity

FTP File Transfer Protocol

FTTH Fibre-To-The-Home

GBR Guaranteed Bit Rate

GCL Generalized Chirp-Like

GERAN GSM EDGE Radio Access Network

GGSN Gateway GPRS Support Node

GMSK Gaussian Minimum-Shift Keying

GNSS Global Navigation Satellite System

GPRS General Packet Radio Service

GPS Global Positioning System

GSM Global System for Mobile communications

GT Guard Time

GTP GPRS Tunnelling Protocol

GTP-U GTP-User plane

HARQ Hybrid Automatic Repeat reQuest

HD-FDD Half-Duplex FDD

HeNB Home eNodeB

HFN Hyper Frame Number

HII High Interference Indicator

HLR Home Location Register

HRPD High Rate Packet Data

HSDPA High Speed Downlink Packet Access

HSPA High Speed Packet Access

HSPA+ High Speed Packet Access Evolution

HSS Home Subscriber Server

HSUPA High Speed Uplink Packet Access

HTTP HyperText Transfer Protocol

ICI Inter-Carrier Interference

ICIC Inter-Cell Interference Coordination

IDFT Inverse Discrete Fourier Transform

IETF Internet Engineering Task Force

IFDMA Interleaved Frequency Division Multiple Access

IFFT Inverse Fast Fourier Transform

i.i.d. Independent identically distributed

IM Implementation Margin

IMD Inter-Modulation Distortion

IMS IP Multimedia Subsystem

IMSI International Mobile Subscriber Identity
IMT International Mobile Telecommunications
InH Indoor Hotspot
IP Internet Protocol
IR Incremental Redundancy
IRC Interference Rejection Combining
ISD Inter-Site Distance
ISI Inter-Symbol Interference
IST-WINNER Information Society Technologies - Wireless world INitiative NEw Radio
ITU International Telecommunication Union
ITU-R ITU Radiocommunication sector
J-TACS Japanese Total Access Communication System
JT Joint Transmission
LA Local Area
LAC Local Area Code
LB Long Block
LBP Layered Belief Propagation
LBRM Limited Buffer Rate Matching
LCID Logical Channel ID
LDPC Low-Density Parity Check
L-GW LIPA GateWay
LI Length Indicator
LIPA Local IP Access
LLR Log-Likelihood Ratio
LMMSE Linear MMSE
LNA Low Noise Amplifier
LO Local Oscillator
LOS Line-Of-Sight
LPP LTE Positioning Protocol
LS Least Squares
LSF Last Segment Flag
LTE Long-Term Evolution
MA Metropolitan Area
MAC Medium Access Control
MAC-I Message Authentication Code for Integrity
MAN Metropolitan Area Network
MAP Maximum A posteriori Probability
MBL Mobility Load Balancing
MBMS Multimedia Broadcast/Multicast Service
MBMS GW MBMS GateWay
MBR Maximum Bit Rate
MBSFN Multimedia Broadcast Single Frequency Network
MCCH Multicast Control CHannel
MCE Multicell Coordination Entity
MCH Multicast CHannel
MCL Minimum Coupling Loss
MCS Modulation and Coding Scheme
Mcps Megachips per second
MDS Minimum Discernible Signal
MDT Minimization of Drive Tests
MeNB Macro eNodeB
MIB Master Information Block
MIMO Multiple-Input Multiple-Output
MIP Mobile Internet Protocol
MISO Multiple-Input Single-Output
ML Maximum Likelihood
MLD Maximum Likelihood Detector
MME Mobility Management Entity
MMSE Minimum MSE
MO Mobile Originated
MOP Maximum Output Power
MPS Multimedia Priority Service
M-PSK M-ary Phase-Shift Keying
MQE Minimum Quantization Error
MRB Multicast Radio Bearer
MRC Maximum Ratio Combining
M-RNTI MBMS Radio Network Temporary Identifier
MRO Mobility Robustness Optimization
MSA MCH Subframe Allocation
MSAP MCH Subframe Allocation Pattern
MSB Most Significant Bit
MSD Maximum Sensitivity Degradation
MSE Mean Squared Error

MSI MCH Scheduling Information

MSISDN Mobile Station International Subscriber Directory Number

MSP MCH Scheduling Period

MSR Maximum Sensitivity Reduction

MTC Machine-Type Communications

MTCH Multicast Traffic CHannel

MU-MIMO Multi-User MIMO

MUE Macro User Equipment

NACC Network Assisted Cell Change

NACK Negative ACKnowledgement

NACS NonAdjacent Channel Selectivity

NAS Non Access Stratum

NCC Network Colour Code

NCL Neighbour Cell List

NDI New Data Indicator

NF Noise Figure

NGMN Next Generation Mobile Networks

NLM Network Listen Mode

NLMS Normalized Least-Mean-Square

NLOS Non-Line-Of-Sight

NMT Nordic Mobile Telephone

NNSF NAS Node Selection Function

NodeB The base station in WCDMA systems

NR Neighbour cell Relation

NRT Neighbour Relation Table

O&M Operation and Maintenance

OBPD Occupied Bandwidth Power De-rating

OBW Occupied BandWidth

OCC Orthogonal Cover Code

OFDM Orthogonal Frequency Division Multiplexing

OFDMA Orthogonal Frequency Division Multiple Access

OPEX OPerational Expenditure

OSG Open Subscriber Group

OTDOA Observed Time Difference Of Arrival

OI Overload Indicator

OMA Open Mobile Alliance

OOB Out-Of-Band

P/S Parallel-to-Serial

PA Power Amplifier

PAN Personal Area Network

PAPR Peak-to-Average Power Ratio

PBCH Physical Broadcast CHannel

PBR Prioritized Bit Rate

PCC Policy Control and Charging*

PCC Primary Component Carrier*

PCCH Paging Control CHannel

P-CCPCH Primary Common Control Physical CHannel

PCEF Policy Control Enforcement Function

PCell Primary serving Cell

PCFICH Physical Control Format Indicator CHannel

PCG Project Coordination Group

PCH Paging CHannel

PCI Physical Cell Identity

P-CPICH Primary Common PIlot CHannel

PCRF Policy Control and charging Rules Function

PDCCH Physical Downlink Control CHannel

PDCP Packet Data Convergence Protocol

PDN Packet Data Network

PDP Power Delay Profile

PDSCH Physical Downlink Shared CHannel

PDU Protocol Data Unit

PF Paging Frame

PFS Proportional Fair Scheduling

P-GW PDN GateWay

PHICH Physical Hybrid ARQ Indicator CHannel

PHR Power Headroom Report

PLL Phase-Locked Loop

PLMN Public Land Mobile Network

P-MCCH Primary MCCH

PMCH Physical Multicast CHannel

PMI Precoding Matrix Indicators

PMIP Proxy MIP

PN Pseudo-Noise
PO Paging Occasion
PRACH Physical Random Access CHannel
PRB Physical Resource Block
P-RNTI Paging RNTI
PRG Precoder Resource block Group
PRS Positioning Reference Signal
PS Packet-Switched
P-SCH Primary Synchronization CHannel
PSD Power Spectral Density
PSS Primary Synchronization Signal
PTI Precoder Type Indication
PUCCH Physical Uplink Control CHannel
PUSCH Physical Uplink Shared CHannel
PVI Precoding Vector Indicator
PWS Public Warning System
QAM Quadrature Amplitude Modulation
QCI QoS Class Identifier
QoS Quality-of-Service
QPP Quadratic Permutation Polynomial
QPSK Quadrature Phase Shift Keying
RA Random Access
RAC Routing Area Code
RACH Random Access CHannel
RAN Radio Access Network
RAR Random Access Response
RA-RNTI Random Access Radio Network Temporary Identifier
RAT Radio Access Technology
RB Resource Block
RE Resource Element
REG Resource Element Group
RF Radio Frequency*
RF Resegmentation Flag*
RFC Request For Comments
RI Rank Indicator
RIM RAN Information Management
RIT Radio Interface Technology
RLC Radio Link Control
RLF Radio Link Failure
RLS Recursive Least Squares
RM Rate Matching*
RM Reed-Muller*
RMa Rural Macrocell
RN Relay Node
RNC Radio Network Controller
RNTI Radio Network Temporary Identifier
RNTP Relative Narrowband Transmit Power
ROHC RObust Header Compression
RoT Rise over Thermal
R-PDCCH Relay Physical Downlink Control Channel
RPRE Received Power per Resource Element
RPF RePetition Factor
R-PLMN Registered PLMN
RRC Radio Resource Control*
RRC Root-Raised-Cosine*
RRH Remote Radio Head
RRM Radio Resource Management
RS Reference Signal
RSCP Received Signal Code Power
RSRP Reference Signal Received Power
RSRQ Reference Signal Received Quality
RSSI Received Signal Strength Indicator
RSTD Reference Signal Time Difference
RTCP Real-time Transport Control Protocol
RTD Round-Trip Delay
RTP Real-time Transport Protocol
RTT Round-Trip Time
RV Redundancy Version
S/P Serial-to-Parallel
S1AP S1 Application Protocol
SAE System Architecture Evolution
SAP Service Access Point
SAW Stop-And-Wait
SB Short Block*
SB Synchronization Burst*
SBP Systematic Bit Puncturing

SCC Secondary Component Carrier

SC-FDMA Single-Carrier Frequency Division Multiple Access

SCH Synchronization CHannel

SCM Spatial Channel Model

SCME Spatial Channel Model Extension

SCTP Stream Control Transmission Protocol

SDMA Spatial Division Multiple Access

SDO Standards Development Organization

SDU Service Data Unit

SeGW Security GateWay

SEM Spectrum Emission Mask

SFBC Space-Frequency Block Code

SFDR Spurious-Free Dynamic Range

SFN System Frame Number

SGSN Serving GPRS Support Node

S-GW Serving GateWay

SI System Information

SIB System Information Block

SIC Successive Interference Cancellation

SIMO Single-Input Multiple-Output

SINR Signal-to-Interference plus Noise Ratio

SIP Session Initiation Protocol

SIR Signal-to-Interference Ratio

SI-RNTI System Information Radio Network Temporary Identifier

SISO Single-Input Single-Output*

SISO Soft-Input Soft-Output*

SLP SUPL Location Platform

S-MCCH Secondary MCCH

SMS Short Message Service

SN Sequence Number

SNR Signal-to-Noise Ratio

SO Segmentation Offset

SON Self-Optimizing Networks

SORTD Space Orthogonal-Resource Transmit Diversity

SPA Sum-Product Algorithm

SPS Semi-Persistent Scheduling

SPS-C-RNTI Semi-Persistent Scheduling C-RNTI

SR Scheduling Request

SRB Signalling Radio Bearer

SRIT Set of Radio Interface Technology

SRNS Serving Radio Network Subsystem

SRS Sounding Reference Signal

S-SCH Secondary Syncronization CHannel

SSS Secondary Synchronization Signal

STBC Space-Time Block Code

S-TMSI SAE-Temporary Mobile Subscriber Identity

STTD Space-Time Transmit Diversity

SU-MIMO Single-User MIMO

SUPL Secure User Plane Location

SVD Singular-Value Decomposition

TA Tracking Area

TAC Tracking Area Code

TACS Total Access Communication System

TAI Tracking Area Identity

TB Transport Block

TCP Transmission Control Protocol

TDC Time-Domain Coordination

TDD Time Division Duplex

TDL Tapped Delay Line

TDMA Time Division Multiple Access

TD-SCDMA Time Division Synchronous Code Division Multiple Access

TEID Tunnelling End ID

TF Transport Format

TFT Traffic Flow Template

TM Transparent Mode

TMD Transparent Mode Data

TNL Transport Network Layer

TNMSE Truncated Normalized Mean-Squared Error

TPC Transmitter Power Control

TPD Total Power De-rating

TPMI Transmitted Precoding Matrix Indicator

TR Tone Reservation

TSC Training Sequence Code

TSG Technical Specification Group

TTA Telecommunications Technology Association

TTC Telecommunications Technology Committee

TTFF Time To First Fix

TTI Transmission Time Interval

TU Typical Urban

UCI Uplink Control Information

UDP User Datagram Protocol

UE User Equipment

UL UpLink

ULA Uniform Linear Array

UL-SCH UpLink Shared CHannel

UM Unacknowledged Mode

UMa Urban Macrocell

UMB Ultra-Mobile Broadband

UMi Urban Microcell

UMTS Universal Mobile Telecommunications System

UP Unitary Precoding

UpPTS Uplink Pilot TimeSlot

US Uncorrelated-Scattered

USIM Universal Subscriber Identity Module

UTRA Universal Terrestrial Radio Access

UTRAN Universal Terrestrial Radio Access Network

VA Viterbi Algorithm

VCB Virtual Circular Buffer

VCO Voltage-Controlled Oscillator

VoIP Voice-over-IP

VRB Virtual Resource Block

WA Wide Area

WAN Wide Area Network

WCDMA Wideband Code Division Multiple Access

WFT Winograd Fourier Transform

WG Working Group

WiMAX Worldwide interoperability for Microwave Access

WINNER Wireless world INitiative NEw Radio

WLAN Wireless Local Area Network

WPD Waveform Power De-rating

WRC World Radiocommunication Conference

WSS Wide-Sense Stationary

WSSUS Wide-Sense Stationary Uncorrelated Scattering

ZC Zadoff–Chu

ZCZ Zero Correlation Zone

ZF Zero-Forcing

ZFEP Zero-Forcing Equal Power

3

Control Plane Protocols

Himke van der Velde

3.1 Introduction

As introduced in Section 2.2.2, the Control Plane of the Access Stratum (AS) handles radio-specific functionalities. The AS interacts with the Non-Access Stratum (NAS), also referred to as the 'upper layers'. Among other functions, the NAS control protocols handle Public Land Mobile Network[1] (PLMN) selection, tracking area update, paging, authentication and Evolved Packet System (EPS) bearer establishment, modification and release.

The applicable AS-related procedures largely depend on the Radio Resource Control (RRC) state of the User Equipment (UE), which can be either RRC_IDLE or RRC_CONNECTED.

A UE in RRC_IDLE performs cell selection and reselection – in other words, it decides on which cell to camp. The cell (re)selection process takes into account the priority of each applicable frequency of each applicable Radio Access Technology (RAT), the radio link quality and the cell status (i.e. whether a cell is barred or reserved). An RRC_IDLE UE monitors a paging channel to detect incoming calls, and also acquires system information. The System Information (SI) mainly consists of parameters by which the network (E-UTRAN) can control the cell (re)selection process.

In RRC_CONNECTED, the E-UTRAN allocates radio resources to the UE to facilitate the transfer of (unicast) data via shared data channels.[2] To support this operation, the UE monitors an associated control channel[3] used to indicate the dynamic allocation of the shared transmission resources in time and frequency. The UE provides the network with reports of its

[1]The network of one operator in one country.

[2]The Physical Downlink Shared CHannel (PDSCH) and Physical Uplink Shared CHannel (PUSCH)– see Sections 9.2.2 and 16.2 respectively.

[3]The Physical Downlink Control CHannel (PDCCH) – see Section 9.3.5.

LTE – The UMTS Long Term Evolution: From Theory to Practice, Second Edition.
Stefania Sesia, Issam Toufik and Matthew Baker.
© 2011 John Wiley & Sons, Ltd. Published 2011 by John Wiley & Sons, Ltd.

buffer status and of the downlink channel quality, as well as neighbouring cell measurement information to enable E-UTRAN to select the most appropriate cell for the UE. These measurement reports include cells using other frequencies or RATs. The UE also receives SI, consisting mainly of information required to use the transmission channels. To extend its battery lifetime, a UE in RRC_CONNECTED may be configured with a Discontinuous Reception (DRX) cycle.

RRC, as specified in [1], is the protocol by which the E-UTRAN controls the UE behaviour in RRC_CONNECTED. RRC also includes the control signalling applicable for a UE in RRC_IDLE, namely paging and SI. The UE behaviour in RRC_IDLE is specified in [2].

Chapter 22 gives some further details of the UE measurements which support the mobility procedures.

Functionality related to Multimedia Broadcast/Multicast Services (MBMSs) is covered separately in Chapter 13.

3.2 Radio Resource Control (RRC)

3.2.1 Introduction

The RRC protocol supports the transfer of *common* NAS information (i.e. NAS information which is applicable to all UEs) as well as *dedicated* NAS information (which is applicable only to a specific UE). In addition, for UEs in RRC_IDLE, RRC supports notification of incoming calls (via paging).

The RRC protocol covers a number of functional areas:

- **System information** handles the broadcasting of SI, which includes NAS common information. Some of the system information is applicable only for UEs in RRC_IDLE while other SI is also applicable for UEs in RRC_CONNECTED.
- **RRC connection control** covers all procedures related to the establishment, modification and release of an RRC connection, including paging, initial security activation, establishment of Signalling Radio Bearers (SRBs) and of radio bearers carrying user data (Data Radio Bearers, DRBs), handover within LTE (including transfer of UE RRC context information[4]), configuration of the lower protocol layers,[5] access class barring and radio link failure.
- **Network controlled inter-RAT mobility** includes handover, cell change orders and redirection upon connection release, security activation and transfer of UE RRC context information.
- **Measurement configuration and reporting** for intra-frequency, inter-frequency and inter-RAT mobility, includes configuration and activation of measurement gaps.
- **Miscellaneous functions** including, for example, transfer of dedicated NAS information and transfer of UE radio access capability information.

[4]This UE context information includes the radio resource configuration including local settings not configured across the radio interface, UE capabilities and radio resource management information.

[5]Packet Data Convergence Protocol (PDCP), Radio Link Control (RLC), Medium Access Control (MAC), all of which are explained in detail in Chapter 4, and the physical layer which is explained in Chapters 5–11 and 14–18.

Dedicated RRC messages are transferred across SRBs, which are mapped via the PDCP and RLC layers onto logical channels – either the Common Control CHannel (CCCH) during connection establishment or a Dedicated Control CHannel (DCCH) in RRC_CONNECTED. System Information and Paging messages are mapped directly to logical channels – the Broadcast Control CHannel (BCCH) and Paging Control CHannel (PCCH) respectively. The various logical channels are described in more detail in Section 4.4.1.2.

SRB0 is used for RRC messages which use the CCCH, SRB1 is for RRC messages using DCCH, and SRB2 is for the (lower-priority) RRC messages using DCCH which only include NAS dedicated information.[6] All RRC messages using DCCH are integrity-protected and ciphered by the PDCP layer (after security activation) and use Automatic Repeat reQuest (ARQ) protocols for reliable delivery through the RLC layer. The RRC messages using CCCH are not integrity-protected and do not use ARQ in the RLC layer.

It should also be noted that the NAS independently applies integrity protection and ciphering.

Figure 3.1 illustrates the overall radio protocol architecture as well as the use of radio bearers, logical channels, transport channels and physical channels.

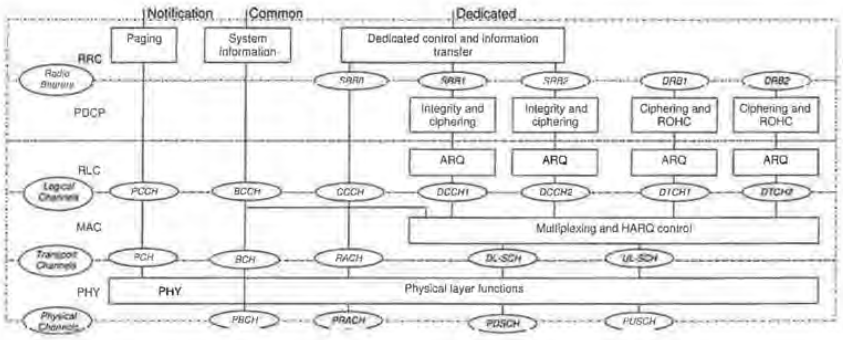


Figure 3.1: Radio architecture.

For control information for which low transfer delay is more important than reliable transfer (i.e. for which the use of ARQ is inappropriate due to the additional delay it incurs), MAC signalling is used provided that there are no security concerns (integrity protection and ciphering are not applicable for MAC signalling).

3.2.2 System Information

System information is structured by means of System Information Blocks (SIBs), each of which contains a set of functionally-related parameters. The SIB types that have been defined include:

[6]Prior to SRB2 establishment, SRB1 is also used for RRC messages which only include NAS dedicated information. In addition, SRB1 is used for higher priority RRC messages which only include NAS dedicated information.

- **The Master Information Block** (MIB), which includes a limited number of the most frequently transmitted parameters which are essential for a UE's initial access to the network.
- **System Information Block Type 1** (SIB1), which contains parameters needed to determine if a cell is suitable for cell selection, as well as information about the time-domain scheduling of the other SIBs.
- **System Information Block Type 2** (SIB2), which includes common and shared channel information.
- **SIB3–SIB8**, which include parameters used to control intra-frequency, inter-frequency and inter-RAT cell reselection.
- **SIB9**, which is used to signal the name of a Home eNodeB (HeNBs).
- **SIB10–SIB12**, which include the Earthquake and Tsunami Warning Service (ETWS) notifications and Commercial Mobile Alert System (CMAS) warning messages (See Section 13.7).
- **SIB13**, which includes MBMS related control information (See Section 13.6.3.2.

Three types of RRC message are used to transfer system information: the MIB message, the SIB1 message and SI messages. An SI message, of which there may be several, includes one or more SIBs which have the same scheduling requirements (i.e. the same transmission periodicity). Table 3.1 provides an example of a possible system information scheduling configuration, also showing which SIBs the UE has to acquire in the idle and connected states. The physical channels used for carrying the SI are explained in Section 9.2.1.

Table 3.1: Example of SI scheduling configuration.

Message	Content	Period (ms)	Applicability
MIB	Most essential parameters	40	Idle and connected
SIB1	Cell access related parameters, scheduling information	80	Idle and connected
1st SI	SIB2: Common and shared channel configuration	160	Idle and connected
2nd SI	SIB3: Common cell reselection information and intra-frequency cell reselection parameters other than the neighbouring cell information SIB4: Intra-frequency neighbouring cell information	320	Idle only
3rd SI	SIB5: Inter-frequency cell reselection information	640	Idle only
4th SI	SIB6: UTRA cell reselection information SIB7: GERAN cell reselection information	640	Idle only, depending on UE support of UMTS or GERAN

3.2.2.1 Time-Domain Scheduling of System Information

The time-domain scheduling of the MIB and SIB1 messages is fixed with a periodicities of 40 ms and 80 ms respectively, as explained in Sections 9.2.1 and 9.2.2.2.

The time-domain scheduling of the SI messages is dynamically flexible: each SI message is transmitted in a defined periodically-occurring time-domain window, while physical layer control signalling[7] indicates in which subframes[8] within this window the SI is actually scheduled. The scheduling windows of the different SI messages (referred to as SI-windows) are consecutive (i.e. there are neither overlaps nor gaps between them) and have a common length that is configurable. SI-windows can include subframes in which it is not possible to transmit SI messages, such as subframes used for SIB1, and subframes used for the uplink in TDD.

Figure 3.2 illustrates an example of the time-domain scheduling of SI, showing the subframes used to transfer the MIB, SIB1 and four SI messages. The example uses an SI window of length 10 subframes, and shows a higher number of 'blind' Hybrid ARQ (HARQ) transmissions[9] being used for the larger SI messages.

SI messages may have different periodicities. Consequently, in some clusters of SI-windows all the SI messages are scheduled, while in other clusters only the SIs with shorter repetition periods are transmitted. For the example of Table 3.1, the cluster of SI-windows beginning at System Frame Number (SFN) 0 contains all the SI messages, the cluster starting at SFN160 contains only the first SI message, that beginning at SFN320 contains the first and second SI messages, and the one starting at SFN480 contains only the first SI message.

Note that Figure 3.2 shows a cluster of SI-windows where all the SI messages are transmitted. At occasions where a given SI is not transmitted (due to a longer repetition period), its corresponding SI-window is not used.

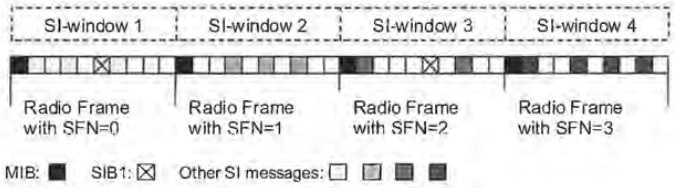


Figure 3.2: SI scheduling example.

3.2.2.2 Validity and Change Notification

SI normally changes only at specific radio frames whose System Frame Number is given by SFN mod $N = 0$, where N is configurable and defines the period between two radio frames at which a change may occur, known as the *modification period*. Prior to performing a change

[7] The Physical Downlink Control Channel – PDCCH; see Section 9.3.5.
[8] A subframe in LTE has a duration of 1 ms; see Section 6.2
[9] With blind HARQ retransmissions, there is no feedback to indicate whether the reception has been successful.

of the system information, the E-UTRAN notifies the UEs by means of a *Paging* message including a *SystemInfoModification* flag. Figure 3.3 illustrates the change of SI, with different shading indicating different content.

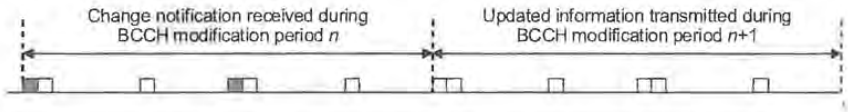


Figure 3.3: SI modification periods. Reproduced by permission of © 3GPP.

LTE provides two mechanisms for indicating that SI has changed:

1. A paging message including a flag indicating whether or not SI has changed.
2. A value tag in SIB1 which is incremented every time one or more SI message changes.

UEs in RRC_IDLE use the first mechanism, while UEs in RRC_CONNECTED can use either mechanism; the second being useful, for example, in cases when a UE was unable to receive the paging messages.

UEs in RRC_IDLE are only required to receive the paging message at their normal paging occasions – i.e. no additional wake-ups are expected to detect changes of SI. In order to ensure reliability of reception, the change notification paging message is normally repeated a number of times during the BCCH modification period preceding that in which the new system information is first transmitted. Correspondingly, the modification period is expressed as a multiple of the cell-specific default paging cycle.

UEs in RRC_CONNECTED are expected to try receiving a paging message the same number of times per modification period as UEs in RRC_IDLE using the default paging cycle. The exact times at which UEs in RRC_CONNECTED which are using this method have to try to receive a paging message are not specified; the UE may perform these tries at convenient times, such as upon wake-up from DRX, using any of the subframes which are configured for paging during the modification period. Since the eNodeB anyway has to notify all the UEs in RRC_IDLE, it has to send a paging message in all subframes which are configured for paging (up to a maximum of four subframes per radio frame) during an entire modification period. Connected mode UEs can utilize any of these subframes. The overhead of transmitting paging messages to notify UEs of a change of SI is considered marginal, since such changes are expected to be infrequent – at most once every few hours.

If the UE receives a notification of a change of SI, it starts acquiring SI from the start of the next modification period. Until the UE has successfully acquired the updated SI, it continues to use the existing parameters. If a critical parameter changes, the communication may be seriously affected, but any service interruption that may result is considered acceptable since it is short and infrequent.

If the UE returns to a cell, it is allowed to assume that the SI previously acquired from the cell remains valid if it was received less than 3 hours previously and the value tag matches.

3.2.3 Connection Control within LTE

Connection control involves:

- Security activation;
- Connection establishment, modification and release;
- DRB establishment, modification and release;
- Mobility within LTE.

3.2.3.1 Security Key Management

Security is a very important feature of all 3GPP RATs. LTE provides security in a similar way to its predecessors UMTS and GSM.

Two functions are provided for the maintenance of security: *ciphering* of both control plane (RRC) data (i.e. SRBs 1 and 2) and user plane data (i.e. all DRBs), and *integrity protection* which is used for control plane (RRC) data only. Ciphering is used in order to protect the data streams from being received by a third party, while integrity protection allows the receiver to detect packet insertion or replacement. RRC always activates both functions together, either following connection establishment or as part of the handover to LTE.

The hierarchy of keys by which the AS security keys are generated is illustrated in Figure 3.4. The process is based on a common secret key K_{ASME} (Access Security Management Entity) which is available only in the Authentication Centre in the Home Subscriber Server (HSS) (see Section 2.2.1) and in a secure part of the Universal Subscriber Identity Module (USIM) in the UE. A set of keys and checksums are generated at the Authentication Centre using this secret key and a random number. The generated keys, checksums and random number are transferred to the Mobility Management Entity (MME) (see Section 2.2.1), which passes one of the generated checksums and the random number to the UE. The USIM in the UE then computes the same set of keys using the random number and the secret key. Mutual authentication is performed by verifying the computed checksums in the UE and network using NAS protocols.

Upon connection establishment, the AS derives an *AS base-key* K_{eNB} (eNodeB-specific) and Next Hop (NH), from K_{ASME}.

K_{eNB} is used to generate three further security keys known as the *AS derived-keys*: one, called $K_{RRC\ int}$, is used for integrity protection of the RRC signalling (SRBs), one for ciphering of the RRC signalling known as $K_{RRC\ enc}$ and $K_{UP\ enc}$ used for ciphering of user data (i.e. DRBs).

NH is an intermediate key used to implement 'forward security'[10] [3]. It is derived by the UE and MME using K_{ASME} and K_{eNB} when the security context is established or using K_{ASME} and the previous NH otherwise. NH is associated with a counter called Next hop Chaining Counter (NCC) which is initially set to 0 at connection establishment.

In case of handover within E-UTRAN, a new AS base-key and new AS Derived-keys are computed from the AS base-key used in the source cell. An intermediate key, K_{eNB^*} is derived by the UE and the source eNodeB based on the Physical Cell Identity (PCI) of the target cell,

[10] Forward security refers to the property that, for an eNodeB sharing a K_{eNB} with a UE, it shall be computationally infeasible to predict any future K_{eNB}, that will be used between the same UE and another eNodeB

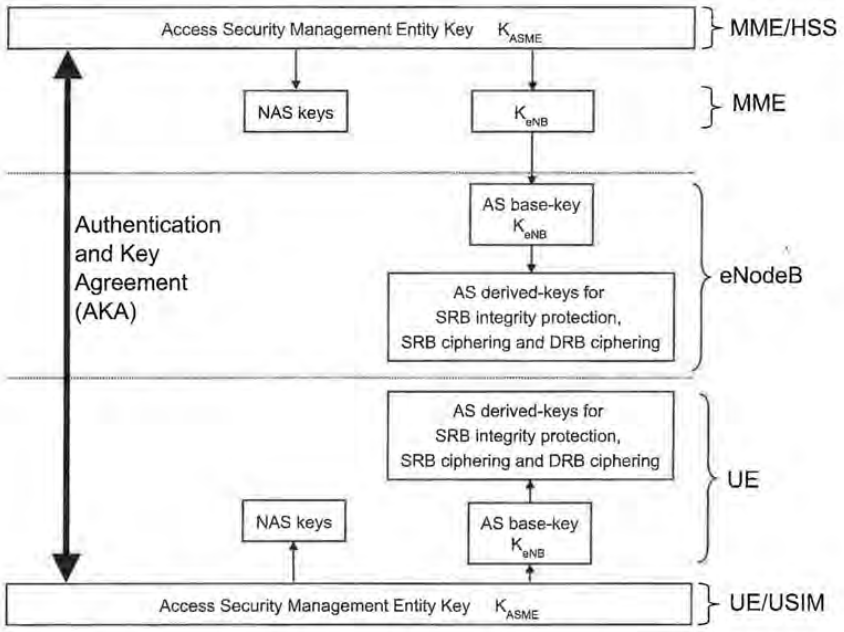


Figure 3.4: Security key derivation.

the target frequency and NH or K_{eNB}. If a fresh NH is available[11], the derivation of K_{eNB}* is based on NH (referred to as vertical derivation). If no fresh NH is available then the K_{eNB}* derivation is referred to as horizontal derivation and is based on K_{eNB}. K_{eNB}* is then used at the target cell as the new K_{eNB} for RRC and data traffic.

For handover to E-UTRAN from UTRAN or GERAN, the AS base-key is derived from integrity and ciphering keys used in the UTRAN or GERAN. Handover within LTE may be used to take a new K_{ASME} into account, i.e. following a re-authentication by NAS.

The use of the security keys for the integrity protection and ciphering functions is handled by the PDCP layer, as described in Section 4.2.3.

The security functions are never deactivated, although it is possible to apply a 'NULL' ciphering algorithm. The 'NULL' algorithm may also be used in certain special cases, such as for making an emergency call without a USIM.

[11] In this case the NCC is incremented and is then larger than that of the currently active K_{eNB}*

3.2.3.2 Connection Establishment and Release

Two levels of NAS states reflect the state of a UE in respect of connection establishment: the EPS Mobility Management (EMM) state (EMM-DEREGISTERED or EMM-REGISTERED) reflects whether the UE is registered in the MME, and the EPS Connection Management (ECM) state (ECM-IDLE or ECM-CONNECTED) reflects the connectivity of the UE with the Evolved Packet Core (EPC – see Chapter 2).

The NAS states, and their relationship to the AS RRC states, are illustrated in Figure 3.5.

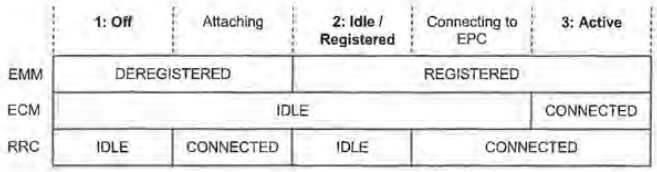


Figure 3.5: Possible combinations of NAS and AS states.

The transition from ECM-IDLE to ECM-CONNECTED not only involves establishment of the RRC connection but also includes establishment of the S1-connection (see Section 2.5). RRC connection establishment is initiated by the NAS and is completed prior to S1-connection establishment, which means that connectivity in RRC_CONNECTED is initially limited to the exchange of control information between UE and E-UTRAN.

UEs are typically moved to ECM-CONNECTED when becoming active. It should be noted, however, that in LTE the transition from ECM-IDLE to ECM-CONNECTED is performed within 100 ms. Hence, UEs engaged in intermittent data transfer need not be kept in ECM-CONNECTED if the ongoing services can tolerate such transfer delays. In any case, an aim in the design of LTE was to support similar battery power consumption levels for UEs in RRC_CONNECTED as for UEs in RRC_IDLE.

RRC connection release is initiated by the eNodeB following release of the S1-connection between the eNodeB and the Core Network (CN).

Connection establishment message sequence. RRC connection establishment involves the establishment of SRB1 and the transfer of the initial uplink NAS message. This NAS message triggers the establishment of the S1-connection, which normally initiates a subsequent step during which E-UTRAN activates AS-security and establishes SRB2 and one or more DRBs (corresponding to the default and optionally dedicated EPS bearers).

Figure 3.6 illustrates the RRC connection establishment procedure, including the subsequent step of initial security activation and radio bearer establishment.

Step 1: Connection establishment

- Upper layers in the UE trigger connection establishment, which may be in response to paging. The UE checks if access is barred (see Section 3.3.4.6). If this is not the case, the lower layers in the UE perform a contention-based random access procedure

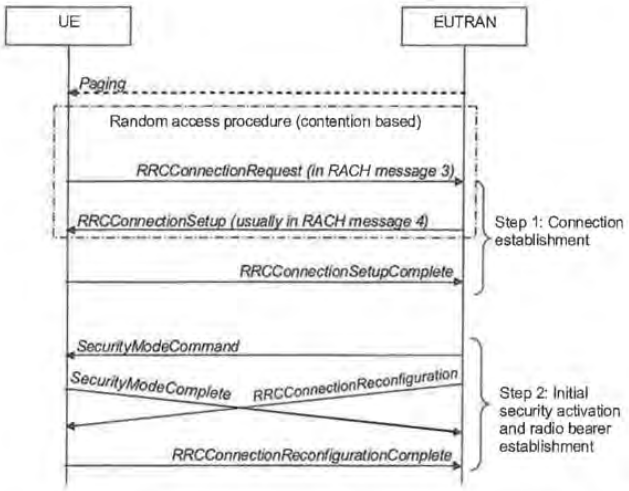


Figure 3.6: Connection establishment (see Section 17.3.1 for details of the contention-based RACH procedure).

as described in Section 17.3, and the UE starts a timer (known as *T300*) and sends the RRCConnectionRequest message. This message includes an initial identity (S-TMSI[12] or a random number) and an establishment cause.

- If E-UTRAN accepts the connection, it returns the RRCConnectionSetup message that includes the initial radio resource configuration including SRB1. Instead of signalling each individual parameter, E-UTRAN may order the UE to apply a default configuration – i.e. a configuration for which the parameter values are specified in the RRC specification [1].
- The UE returns the RRCConnectionSetupComplete message and includes the NAS message, an identifier of the selected PLMN (used to support network sharing) and, if provided by upper layers, an identifier of the registered MME. Based on the last two parameters, the eNodeB decides on the CN node to which it should establish the S1-connection.

Step 2: Initial security activation and radio bearer establishment

- E-UTRAN sends the SecurityModeCommand message to activate integrity protection and ciphering. This message, which is integrity-protected but not ciphered, indicates which algorithms shall be used.
- The UE verifies the integrity protection of the SecurityModeControl message, and, if this succeeds, it configures lower layers to apply integrity protection and ciphering to all subsequent messages (with the exception that ciphering is not applied to the

[12] S-Temporary Mobile Subscriber Identity.

response message, i.e. the SecurityModeComplete (or SecurityModeFailure) message).

- E-UTRAN sends the RRCConnectionReconfiguration message including a radio resource configuration used to establish SRB2 and one or more DRBs. This message may also include other information such as a piggybacked NAS message or a measurement configuration. E-UTRAN may send the RRCConnectionReconfiguration message prior to receiving the SecurityModeComplete message. In this case, E-UTRAN should release the connection when one (or both) procedures fail (because the two procedures result from a single S1-procedure, which does not support partial success).
- The UE finally returns the RRCConnectionReconfigurationComplete message.

A connection establishment may fail for a number of reasons, such as the following:

- Access may be barred (see Section 3.3.4.6).
- In case cell reselection occurs during connection establishment, the UE aborts the procedure and informs upper layers of the failure to establish the connection.
- E-UTRAN may temporarily reject the connection establishment by including a wait timer, in which case the UE rejects any connection establishment request until the wait time has elapsed.
- The NAS may abort an ongoing RRC connection establishment, for example upon NAS timer expiry.

3.2.3.3 DRB Establishment

To establish, modify or release DRBs, E-UTRAN applies the RRC connection reconfiguration procedure as described in Section 3.2.3.2.

When establishing a DRB, E-UTRAN decides how to transfer the packets of an EPS bearer across the radio interface. An EPS bearer is mapped (1-to-1) to a DRB, a DRB is mapped (1-to-1) to a DTCH (Dedicated Traffic CHannel – see Section 4.4.1.2) logical channel, all logical channels are mapped (n-to-1) to the Downlink or Uplink Shared Transport CHannel (DL-SCH or UL-SCH), which are mapped (1-to-1) to the corresponding Physical Downlink or Uplink Shared CHannel (PDSCH or PUSCH). This radio bearer mapping is illustrated in Figure 3.1.

The radio resource configuration covers the configuration of the PDCP, RLC, MAC and physical layers. The main configuration parameters / options include the following:

- For services using small packet sizes (e.g. VoIP), PDCP may be configured to apply indexheader!compressionheader compression to significantly reduce the signalling overhead.
- The RLC Mode is selected from those listed in Section 4.3.1. RLC Acknowledged Mode (AM) is applicable, except for services which require a very low transfer delay and for which reliable transfer is less important.
- E-UTRAN assigns priorities and Prioritized Bit-Rates (PBRs) to control how the UE divides the granted uplink resources between the different radio bearers (see Section 4.4.2.6).

- Unless the transfer delay requirements for any of the ongoing services are very strict, the UE may be configured with a DRX cycle (see Section 4.4.2.5).
- For services involving a semi-static packet rate (e.g. VoIP), semi-persistent scheduling may be configured to reduce the control signalling overhead (see Section 4.4.2.1). Specific resources may also be configured for reporting buffer status and radio link quality.
- Services tolerating higher transfer delays may be configured with a HARQ profile involving a higher average number of HARQ transmissions.

3.2.3.4 Mobility Control in RRC_IDLE and RRC_CONNECTED

Mobility control in RRC_IDLE is UE-controlled (cell-reselection), while in RRC_CONNECTED it is controlled by the E-UTRAN (handover). However, the mechanisms used in the two states need to be consistent so as to avoid ping-pong (i.e. rapid handing back and forth) between cells upon state transitions. The mobility mechanisms are designed to support a wide variety of scenarios including network sharing, country borders, home deployment and varying cell ranges and subscriber densities; an operator may, for example, deploy its own radio access network in populated areas and make use of another operator's network in rural areas.

If a UE were to access a cell which does not have the best radio link quality of the available cells on a given frequency, it may create significant interference to the other cells. Hence, as for most technologies, radio link quality is the primary criterion for selecting a cell on an LTE frequency. When choosing between cells on different frequencies or RATs the interference concern does not apply. Hence, for inter-frequency and inter-RAT cell reselection other criteria may be considered such as UE capability, subscriber type and call type. As an example, UEs with no (or limited) capability for data transmission may be preferably handled on GSM, while home customers or 'premium subscribers' might be given preferential access to the frequency or RAT supporting the highest data rates. Furthermore, in some LTE deployment scenarios, voice services may initially be provided by a legacy RAT only (as a Circuit Switching (CS) application), in which case the UE needs to be moved to the legacy RAT upon establishing a voice call (also referred to as *CS FallBack* (CSFB)).

E-UTRAN provides a list of neighbouring frequencies and cells which the UE should consider for cell reselection and for reporting of measurements. In general, such a list is referred to as a *white-list* if the UE is to consider only the listed frequencies or cells – i.e. other frequencies or cells are not available; conversely, in the case of a *black-list* being provided, a UE may consider any *un*listed frequencies or cells. In LTE, white-listing is used to indicate all the neighbouring frequencies of each RAT that the UE is to consider. On the other hand, E-UTRAN is not required to indicate all the neighbouring cells that the UE shall consider. Which cells the UE is required to detect by itself depends on the UE state as well as on the RAT, as explained below.

Note that for GERAN, typically no information is provided about individual cells. Only in specific cases, such as at country borders, is signalling[13] provided to indicate the group of cells that the UE is to consider – i.e. a white cell list.

[13]The 'NCC-permitted' parameter – see GERAN specifications.

Mobility in idle mode. In RRC_IDLE, cell reselection between frequencies is based on absolute priorities, where each frequency has an associated priority. Cell-specific default values of the priorities are provided via SI. In addition, E-UTRAN may assign UE-specific values upon connection release, taking into account factors such as UE capability or subscriber type. In case equal priorities are assigned to multiple cells, the cells are ranked based on radio link quality. Equal priorities are not applicable between frequencies of different RATs. The UE does not consider frequencies for which it does not have an associated priority; this is useful in situations such as when a neighbouring frequency is applicable only for UEs of one of the sharing networks.

Table 3.2 provides an overview of the SI parameters which E-UTRAN may use to control cell reselection. Other than the cell reselection priority of a frequency, no idle mode mobility-related parameters may be assigned via dedicated signalling. Further details of the parameters listed are provided in Section 3.3.

Mobility in connected mode. In RRC_CONNECTED, the E-UTRAN decides to which cell a UE should hand over in order to maintain the radio link. As with RRC_IDLE, E-UTRAN may take into account not only the radio link quality but also factors such as UE capability, subscriber type and access restrictions. Although E-UTRAN may trigger handover without measurement information (blind handover), normally it configures the UE to report measurements of the candidate target cells – see Section 22.3. Table 3.3 provides an overview of the frequency- and cell-specific parameters which E-UTRAN can configure for mobility-related measurement reporting.

In LTE the UE always connects to a single cell only – in other words, the switching of a UE's connection from a source cell to a target cell is a *hard* handover. The hard handover process is normally a 'backward' one, whereby the eNodeB which controls the source cell requests the target eNodeB to prepare for the handover. The target eNodeB subsequently generates the RRC message to order the UE to perform the handover, and the message is transparently forwarded by the source eNodeB to the UE. LTE also supports a kind of 'forward' handover, in which the UE by itself decides to connect to the target cell, where it then requests that the connection be continued. The UE applies this connection re-establishment procedure only after loss of the connection to the source cell; the procedure only succeeds if the target cell has been prepared in advance for the handover.

Besides the handover procedure, LTE also provides for a UE to be redirected to another frequency or RAT upon connection release. Redirection during connection establishment is not supported, since at that time the E-UTRAN may not yet be in possession of all the relevant information such as the capabilities of the UE and the type of subscriber (as may be reflected, for example, by the SPID, the Subscriber Profile ID for RAT/Frequency Priority). However, the redirection may be performed while AS-security has not (yet) been activated. When redirecting the UE to UTRAN or GERAN, E-UTRAN may provide SI for one or more cells on the relevant frequency. If the UE selects one of the cells for which SI is provided, it does not need to acquire it.

Message sequence for handover within LTE. In RRC_CONNECTED, the E-UTRAN controls mobility by ordering the UE to perform handover to another cell, which may be

Table 3.2: List of SI parameters which may be used to control cell reselection.

Parameter	Intra-Freq.	Inter-Freq.	UTRA	GERAN	CDMA2000
Common	(SIB3)	(SIB5)	(SIB6)	(SIB7)	(SIB8)
Reselection info	Q-Hyst MobilityStatePars Q-HystSF S-Search[a]		T-Reselect T-ReselectSF	T-Reselect T-ReselectSF	T-Reselect T-ReselectSF
Frequency list	(SIB3)	(SIB5)	(SIB6)	(SIB7)	(SIB8)
White frequency list	n/a	+	+	+	+
Frequency specific reselection info[b]	Priority $Thresh_{Serving\text{-}Low}$ $Thresh_{Serving\text{-}LowQ}$ T-Reselect T-ReselectSF $Thresh_{X\text{-}LowQ}$ T-Reselect T-ReselectSF	Priority Qoffset, $Thresh_{X\text{-}High}$, $Thresh_{X\text{-}Low}$ $Thresh_{X\text{-}HighQ}$,	Priority $Thresh_{X\text{-}High}$, $Thresh_{X\text{-}Low}$ $Thresh_{X\text{-}HighQ}$, $Thresh_{X\text{-}LowQ}$	Priority $Thresh_{X\text{-}High}$, $Thresh_{X\text{-}Low}$	Priority $Thresh_{X\text{-}High}$, $Thresh_{X\text{-}Low}$
Frequency specific suitability info[c]	Q-RxLevMin MaxTxPower Q-QualMin	Q-RxLevMin MaxTxPower Q-QualMin	Q-RxLevMin, MaxTxPower, Q-QualMin	Q-RxLevMin MaxTxPower	
Cell list	(SIB4)	(SIB5)	(SIB6)	(SIB7)	(SIB8)
White cell list	–	–	–	NCC permitted[d]	–
Black cell list	+	+	–	–	–
List of cells with specific info[e]	Qoffset	Qoffset	–	–	–

[a] Separate parameters for intra/ inter-frequency, both for RSRP and RSRQ.
[b] See Section 3.3.4.2.
[c] See Section 3.3.3.
[d] See GERAN specifications.
[e] See Section 3.3.4.3.

on the same frequency ('intra-frequency') or a different frequency ('inter-frequency'). Inter-frequency measurements may require the configuration of measurement gaps, depending on the capabilities of the UE (e.g. whether it has a dual receiver) – see Section 22.3.

The E-UTRAN may also use the handover procedures for completely different purposes, such as to change the security keys to a new set (see Section 3.2.3.1), or to perform a 'synchronized reconfiguration' in which the E-UTRAN and the UE apply the new configuration simultaneously.

The message sequence for the procedure for handover within LTE is shown in Figure 3.7. The sequence is as follows:

1. The UE may send a MeasurementReport message (see Section 3.2.5).

Table 3.3: Frequency- and cell-specific information which can be configured in connected mode.

Parameter	Intra-Freq.	Inter-Freq.	UTRA	GERAN	CDMA2000
Frequency list					
White frequency list	n/a	+	+	+	+
Frequency specific info	Qoffset	Qoffset	Qoffset	Qoffset	Qoffset
Cell list					
White cell list	–	–	+	NCC permitted	+
Black cell list	+	+	–	–	–
List of cells with specific info.	Qoffset	Qoffset	–	–	–

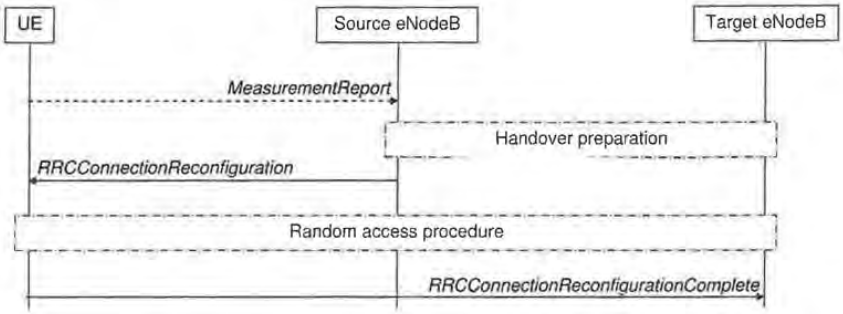


Figure 3.7: Handover within LTE.

2. Before sending the handover command to the UE, the source eNodeB requests one or more target cells to prepare for the handover. As part of this 'handover preparation request', the source eNodeB provides UE RRC context information[14] about the UE capabilities, the current AS-configuration and UE-specific Radio Resource Management (RRM) information. In response, the eNodeB controlling the target cell generates the 'handover command'. The source eNodeB will forward this command to the UE in the RRCConnectionReconfiguration message. This is done transparently (apart from performing integrity protection and ciphering) – i.e. the source eNodeB does not add or modify the protocol information contained in the message.

3. The source eNodeB sends the RRCConnectionReconfiguration message which to the UE orders it to perform handover. It includes mobility control information (namely the identity, and optionally the frequency, of the target cell) and the radio resource

[14]This UE context information includes the radio resource configuration including local settings not configured across the radio interface, UE capabilities and radio resource management information.

configuration information which is common to all UEs in the target cell (e.g. information required to perform random access. The message also includes the dedicated radio resource configuration, the security configuration and the C-RNTI[15] to be used in the target cell. Although the message may optionally include the measurement configuration, the E-UTRAN is likely to use another reconfiguration procedure for re-activating measurements, in order to avoid the RRCConnectionReconfiguration message becoming excessively large. If no measurement configuration information is included in the message used to perform inter-frequency handover, the UE stops any inter-frequency and inter-RAT measurements and deactivates the measurement gap configuration.

4. If the UE is able to comply with the configuration included in the received RRCConnectionReconfiguration message, the UE starts a timer, known as *T304*, and initiates a random access procedure (see Section 17.3), using the received Random Access CHannel (RACH) configuration, to the target cell at the first available occasion.[16] It is important to note that the UE does not need to acquire system information from the target cell prior to initiating random access and resuming data communication. However, the UE may be unable to use some parts of the physical layer configuration from the very start (e.g. semi-persistent scheduling (see Section 4.4.2.1), the PUCCH (see Section 16.3) and the Sounding Reference Signal (SRS) (see Section 15.6)). The UE derives new security keys and applies the received configuration in the target cell.

5. Upon successful completion of the random access procedure, the UE stops the timer T304. The AS informs the upper layers in the UE about any uplink NAS messages for which transmission may not have completed successfully, so that the NAS can take appropriate action.

For handover to cells broadcasting a Closed Subscriber Group (CSG) identity, normal measurement and mobility procedures are used to support handover. In addition, E-UTRAN may configure the UE to report that it is entering or leaving the proximity of cell(s) included in its CSG whitelist. Furthermore, E-UTRAN may request the UE to provide additional information broadcast by the handover candidate cell, for example the cell global identity, CSG identity or CSG membership status. E-UTRAN may use a indexproximity report'proximity report' to configure measurements and to decide whether or not to request the UE to provide additional information broadcast by the handover candidate cell. The additional information is used to verify whether or not the UE is authorized to access the target cell and may also be needed to identify handover candidate cells.[17] Further details of the mobility procedures for HeNBs can be found in Section 24.2.3.

[15]The Cell Radio Network Temporary Identifier is the RNTI to be used by a given UE while it is in a particular cell.

[16]The target cell does not specify when the UE is to initiate random access in that cell. Hence, the handover process is sometimes described as *asynchronous*.

[17]This may be the case if PCI confusion occurs, i.e. when the PCI that is included in the measurement report does not uniquely identify the cell.

3.2.3.5 Connection Re-Establishment Procedure

In a number of failure cases (e.g. radio link failure, handover failure, RLC unrecoverable error, reconfiguration compliance failure), the UE initiates the RRC connection re-establishment procedure, provided that security is active. If security is not active when one of the indicated failures occurs, the UE moves to RRC_IDLE instead.

To attempt RRC connection re-establishment, the UE starts a timer known as *T311* and performs cell selection. The UE should prioritize searching on LTE frequencies. However, no requirements are specified regarding for how long the UE shall refrain from searching for other RATs. Upon finding a suitable cell on an LTE frequency, the UE stops the timer T311, starts the timer T301 and initiates a contention based random access procedure to enable the RRCConnectionReestablishmentRequest message to be sent. In the RRCConnectionReestablishmentRequest message, the UE includes the identity used in the cell in which the failure occurred, the identity of that cell, a short Message Authentication Code and a cause.

The E-UTRAN uses the re-establishment procedure to continue SRB1 and to re-activate security without changing algorithms. A subsequent RRC connection reconfiguration procedure is used to resume operation on radio bearers other than SRB1 and to re-activate measurements. If the cell in which the UE initiates the re-establishment is not prepared (i.e. does not have a context for that UE), the E-UTRAN will reject the procedure, causing the UE to move to RRC_IDLE.

3.2.4 Connected Mode Inter-RAT Mobility

The overall procedure for the control of mobility is explained in this section; some further details can be found in Chapter 22.

3.2.4.1 Handover to LTE

The procedure for handover to LTE is largely the same as the procedure for handover within LTE, so it is not necessary to repeat the details here. The main difference is that upon handover to LTE the entire AS-configuration needs to be signalled, whereas within LTE it is possible to use 'delta signalling', whereby only the changes to the configuration are signalled.

If ciphering had not yet been activated in the previous RAT, the E-UTRAN activates ciphering, possibly using the NULL algorithm, as part of the handover procedure. The E-UTRAN also establishes SRB1, SRB2 and one or more DRBs (i.e. at least the DRB associated with the default EPS bearer).

3.2.4.2 Mobility from LTE

Generally, the procedure for mobility from LTE to another RAT supports both handover and Cell Change Order (CCO), possibly with Network Assistance (NACC – Network Assisted Cell Change). The CCO/NACC procedure is applicable only for mobility to GERAN. Mobility from LTE is performed only after security has been activated. When used for enhanced CSFB[18] to CDMA2000, the procedure includes support for parallel handover (i.e.

[18] See Section 2.4.2.1.

to both 1XRTT and HRPD), for handover to 1XRTT in combination with redirection to HRPD, and for redirection to HRPD only.

The procedure is illustrated in Figure 3.8.

1. The UE may send a MeasurementReport message (see Section 3.2.5 for further details).
2. In case of handover (as opposed to CCO), the source eNodeB requests the target Radio Access Network (RAN) node to prepare for the handover. As part of the 'handover preparation request' the source eNodeB provides information about the applicable inter-RAT UE capabilities as well as information about the currently-established bearers. In response, the target RAN generates the 'handover command' and returns this to the source eNodeB.
3. The source eNodeB sends a MobilityFromEUTRACommand message to the UE, which includes either the inter-RAT message received from the target (in case of handover), or the target cell/frequency and a few inter-RAT parameters (in case of CCO).
4. Upon receiving the MobilityFromEUTRACommand message, the UE starts the timer T304 and connects to the target node, either by using the received radio configuration (handover) or by initiating connection establishment (CCO) in accordance with the applicable specifications of the target RAT.

Upper layers in the UE are informed, by the AS of the target RAT, which bearers are established. From this, the UE can derive if some of the established bearers were not admitted by the target RAN node.

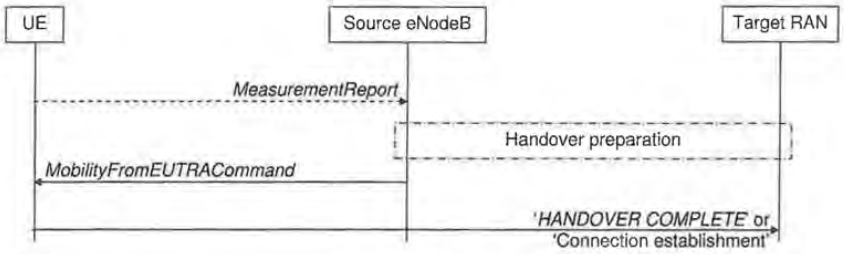


Figure 3.8: Mobility from LTE.

3.2.4.3 CDMA2000

For CDMA2000, additional procedures have been defined to support the transfer of dedicated information from the CDMA2000 upper layers, which are used to register the UE's presence in the target core network prior to performing the handover (referred to as preregistration). These procedures use SRB1.

3.2.5 Measurements

3.2.5.1 Measurement Configuration

The E-UTRAN can configure the UE to report measurement information to support the control of UE mobility. The following measurement configuration elements can be signalled via the RRCConnectionReconfiguration message.

1. **Measurement objects.** A measurement object defines on what the UE should perform the measurements – such as a carrier frequency. The measurement object may include a list of cells to be considered (white-list or black-list) as well as associated parameters, e.g. frequency- or cell-specific offsets.

2. **Reporting configurations.** A reporting configuration consists of the (periodic or event-triggered) criteria which cause the UE to send a measurement report, as well as the details of what information the UE is expected to report (e.g. the quantities, such as Received Signal Code Power (RSCP) (see Section 22.3.2.1) for UMTS or Reference Signal Received Power (RSRP) (see Section 22.3.1.1) for LTE, and the number of cells).

3. **Measurement identities.** These identify a measurement and define the applicable measurement object and reporting configuration.

4. **Quantity configurations**. The quantity configuration defines the filtering to be used on each measurement.

5. **Measurement gaps.** Measurement gaps define time periods when no uplink or downlink transmissions will be scheduled, so that the UE may perform the measurements. The measurement gaps are common for all gap-assisted measurements. Further details of the measurement gaps are discussed in Section 22.2.1.2.

The details of the above parameters depend on whether the measurement relates to an LTE, UMTS, GERAN or CDMA2000 frequency. Further details of the measurements performed by the UE are explained in Section 22.3. The E-UTRAN configures only a single measurement object for a given frequency, but more than one measurement identity may use the same measurement object. The identifiers used for the measurement object and reporting configuration are unique across all measurement types. An example of a set of measurement objects and their corresponding reporting configurations is shown in Figure 3.9.

In LTE it is possible to configure the quantity which triggers the report (RSCP or RSRP) for each reporting configuration. The UE may be configured to report either the trigger quantity or both quantities.

The RRC measurement reporting procedures include some extensions specifically to support Self-Optimizing Network (SON) functions such as the determination of Automatic Neighbour Relations (ANR) – see Section 25.2. The RRC measurement procedures also support UE positioning[19] by means of the enhanced cell identity method – see Section 19.4.

[19] See Chapter 19.

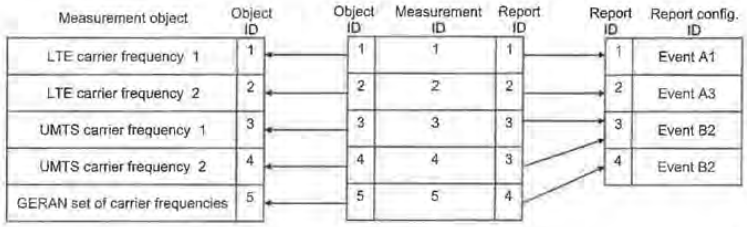


Figure 3.9: Example measurement configuration.

3.2.5.2 Measurement Report Triggering

Depending on the measurement type, the UE may measure and report any of the following:

- The serving cell;
- Listed cells (i.e. cells indicated as part of the measurement object);
- Detected cells on a listed frequency (i.e. cells which are not listed cells but are detected by the UE).

For some RATs, the UE measures and reports listed cells only (i.e. the list is a white-list), while for other RATs the UE also reports detected cells. For further details, see Table 3.3. Additionally, E-UTRAN can configure UTRAN PCI ranges for which the UE is allowed to send a measurement reports (mainly for the support of handover to UTRAN cells broadcasting a CSG identity).

For LTE, the following event-triggered reporting criteria are specified:

- **Event A1.** Serving cell becomes better than absolute threshold.
- **Event A2.** Serving cell becomes worse than absolute threshold.
- **Event A3.** Neighbour cell becomes better than an offset relative to the serving cell.
- **Event A4.** Neighbour cell becomes better than absolute threshold.
- **Event A5.** Serving cell becomes worse than one absolute threshold and neighbour cell becomes better than another absolute threshold.

For inter-RAT mobility, the following event-triggered reporting criteria are specified:

- **Event B1.** Neighbour cell becomes better than absolute threshold.
- **Event B2.** Serving cell becomes worse than one absolute threshold and neighbour cell becomes better than another absolute threshold.

The UE triggers an event when one or more cells meets a specified 'entry condition'. The E-UTRAN can influence the entry condition by setting the value of some configurable parameters used in these conditions – for example, one or more thresholds, an offset, and/or a hysteresis. The entry condition must be met for at least a duration corresponding to a 'timeToTrigger' parameter configured by the E-UTRAN in order for the event to be triggered.

The UE scales the timeToTrigger parameter depending on its speed (see Section 3.3 for further detail).

Figure 3.10 illustrates the triggering of event A3 when a timeToTrigger and an offset are configured.

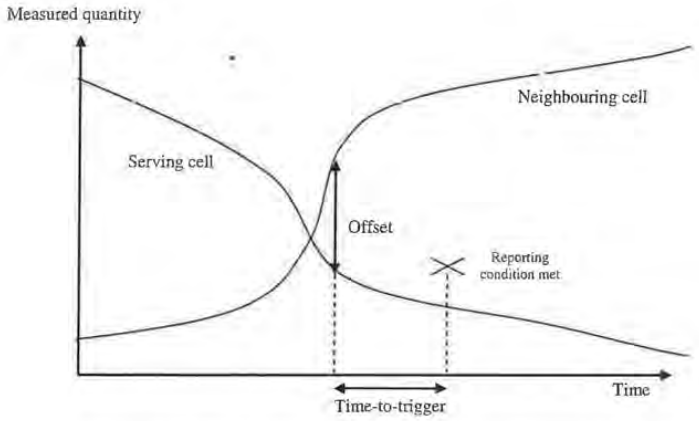


Figure 3.10: Event triggered report condition (Event A3).

The UE may be configured to provide a number of periodic reports after having triggered an event. This 'event-triggered periodic reporting' is configured by means of parameters 'reportAmount' and 'reportInterval', which specify respectively the number of periodic reports and the time period between them. If event-triggered periodic reporting is configured, the UE's count of the number of reports sent is reset to zero whenever a new cell meets the entry condition. The same cell cannot then trigger a new set of periodic reports unless it first meets a specified 'leaving condition'.

In addition to event-triggered reporting, the UE may be configured to perform periodic measurement reporting. In this case, the same parameters may be configured as for event-triggered reporting, except that the UE starts reporting immediately rather than only after the occurrence of an event.

3.2.5.3 Measurement Reporting

In a MeasurementReport message, the UE only includes measurement results related to a single measurement – in other words, measurements are not combined for reporting purposes. If multiple cells triggered the report, the UE includes the cells in order of decreasing value of the reporting quantity – i.e. the best cell is reported first. The number of cells the UE includes in a MeasurementReport may be limited by a parameter 'indexmaxReportCellsmaxReportCells'.

3.2.6 Other RRC Signalling Aspects

3.2.6.1 UE Capability Transfer

In order to avoid signalling of the UE radio access capabilities across the radio interface upon each transition from RRC_IDLE to RRC_CONNECTED, the core network stores the AS capabilities (both the E-UTRA and GERAN capabilities) while the UE is in RRC_IDLE/EMM-REGISTERED. Upon S1 connection establishment, the core network provides the capabilities to the E-UTRAN. If the E-UTRAN does not receive the (required) capabilities from the core network (e.g. due to the UE being in EMM-DEREGISTERED), it requests the UE to provide its capabilities using the UE capability transfer procedure. The E-UTRAN can indicate for each RAT (LTE, UMTS, GERAN) whether it wants to receive the associated capabilities. The UE provides the requested capabilities using a separate container for each RAT. Dynamic change of UE capabilities is not supported, except for change of the GERAN capabilities in RRC_IDLE which is supported by the tracking area update procedure.

3.2.6.2 Uplink/Downlink Information Transfer

The uplink/downlink information transfer procedures are used to transfer only upper layer information (i.e. no RRC control information is included). The procedure supports the transfer of 3GPP NAS dedicated information as well as CDMA2000 dedicated information.

In order to reduce latency, NAS information may also be included in the RRCConnectionSetupComplete and RRCConnectionReconfiguration messages. For the latter message, NAS information is only included if the AS and NAS procedures are dependent (i.e. they jointly succeed or fail). This applies for EPS bearer establishment, modification and release.

As noted earlier, some additional NAS information transfer procedures have also been defined for CDMA2000 for preregistration.

3.2.6.3 UE Information Transfer

The UE information transfer procedure was introduced in Release 9 to support SON (see Chapter 25). The procedure supports network optimization for mobility robustness by the reporting, at a later point in time, of measurement information available when a radio link failure occurs (see Section 25.6). E-UTRAN may also use the UE information transfer procedure to retrieve information regarding the last successful random access, which it may use for RACH optimization – see Section 25.7.

3.3 PLMN and Cell Selection

3.3.1 Introduction

After a UE has selected a PLMN, it performs *cell selection* – in other words, it searches for a suitable cell on which to camp (see Chapter 7). While camping on the chosen cell, the UE acquires the SI that is broadcast (see Section 9.2.1). Subsequently, the UE registers its presence in the tracking area, after which it can receive paging information which is used

to notify UEs of incoming calls. The UE may establish an RRC connection, for example to establish a call or to perform a tracking area update.

When camped on a cell, the UE regularly verifies if there is a better cell; this is known as performing *cell reselection*.

LTE cells are classified according to the service level the UE obtains on them: a *suitable cell* is a cell on which the UE obtains normal service. If the UE is unable to find a suitable cell, but manages to camp on a cell belonging to another PLMN, the cell is said to be an *acceptable cell*, and the UE enters a 'limited service' state in which it can only perform emergency calls (and receive public warning messages) – as is also the case when no USIM is present in the UE. Finally, some cells may indicate via their SI that they are barred or reserved; a UE can obtain no service on such a cell.

A category called 'operator service' is also supported in LTE, which provides normal service but is applicable only for UEs with special access rights.

Figure 3.11 provides a high-level overview of the states and the cell (re)selection procedures.

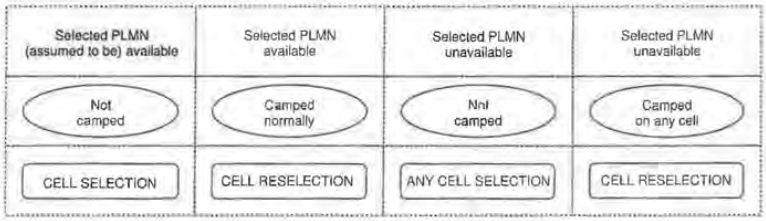


Figure 3.11: Idle mode states and procedures.

3.3.2 PLMN Selection

The NAS handles PLMN selection based on a list of available PLMNs provided by the AS. The NAS indicates the selected PLMN together with a list of equivalent PLMNs, if available. After successful registration, the selected PLMN becomes the *Registered* PLMN (R-PLMN).

The AS may autonomously indicate available PLMNs. In addition, NAS may request the AS to perform a full search for available PLMNs. In the latter case, the UE searches for the strongest cell on each carrier frequency. For these cells, the UE retrieves the PLMN identities from SI. If the quality of a cell satisfies a defined radio criterion, the corresponding PLMNs are marked as *high quality*; otherwise, the PLMNs are reported together with their quality.

3.3.3 Cell Selection

Cell selection consists of the UE searching for the strongest cell on all supported carrier frequencies of each supported RAT until it finds a suitable cell. The main requirement for cell selection is that it should not take too long, which becomes more challenging with the ever increasing number of frequencies and RATs to be searched. The NAS can speed up the

search process by indicating the RATs associated with the selected PLMN. In addition, the UE may use information stored from a previous access.

The cell selection criterion is known as the *S-criterion* and is fulfilled when the cell-selection receive level and the quality level are above a given value: Srxlev > 0 and Squal > 0, where

$$\text{Srxlev} = Q_{\text{rxlevmeas}} - (Q_{\text{rxlevmin}} - Q_{\text{rxlevminoffset}})$$

$$\text{Squal} = Q_{\text{qualmeas}}(Q_{\text{qualmin}} + Q_{\text{qualminoffset}})$$

in which $Q_{\text{rxlevmeas}}$ is the measured cell receive level value, also known as the RSRP (see Section 22.3.1.1), and Q_{rxlevmin} is the minimum required receive level in the cell. Q_{qualmeas} and Q_{qualmin} are the corresponding parameters for the quality level, also known as the RSRQ.

$Q_{\text{rxlevminoffset}}$ and $Q_{\text{qualminoffset}}$ are offsets which may be configured to prevent ping-pong between PLMNs, which may otherwise occur due to fluctuating radio conditions. The offsets are taken into account only when performing a periodic search for a higher priority PLMN while camped on a suitable cell in a visited PLMN.

The cell selection related parameters are broadcast within the SIB1 message.

For some specific cases, additional requirements are defined:

- Upon leaving connected mode, the UE should normally attempt to select the cell to which it was connected. However, the connection release message may include information directing the UE to search for a cell on a particular frequency.
- When performing 'any cell selection', the UE tries to find an acceptable cell of any PLMN by searching all supported frequencies on all supported RATs. The UE may stop searching upon finding a cell that meets the 'high quality' criterion applicable for that RAT.

Note that the UE only verifies the suitability of the strongest cell on a given frequency. In order to avoid the UE needing to acquire SI from a candidate cell that does not meet the S-criterion, suitability information is provided for inter-RAT neighbouring cells.

3.3.4 Cell Reselection

Once the UE camps on a suitable cell, it starts cell reselection. This process aims to move the UE to the 'best' cell of the selected PLMN and of its equivalent PLMNs, if any. As described in Section 3.2.3.4, cell reselection between frequencies and RATs is primarily based on absolute priorities. Hence, the UE first evaluates the frequencies of all RATs based on their priorities. Secondly, the UE compares the cells on the relevant frequencies based on radio link quality, using a ranking criterion. Finally, upon reselecting to the target cell the UE verifies the cell's accessibility. Further rules have also been defined to allow the UE to limit the frequencies to be measured, to speed up the process and save battery power, as discussed in Section 3.3.4.1. Figure 3.12 provides a high-level overview of the cell reselection procedure.

It should be noted that the UE performs cell reselection only after having camped for at least one second on the current serving cell.

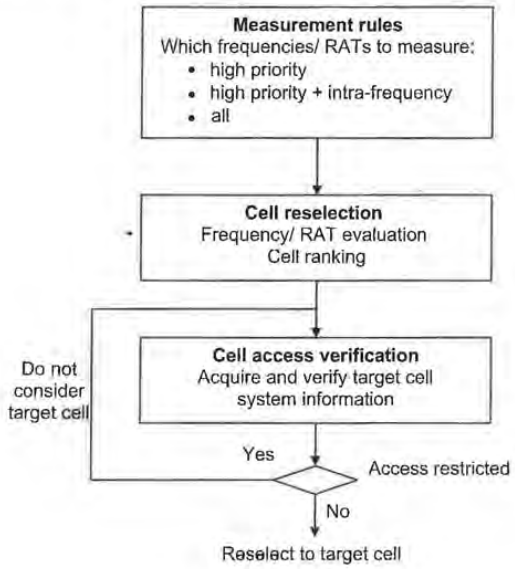


Figure 3.12: Cell reselection.

3.3.4.1 Measurement Rules

To enable the UE to save battery power, rules have been defined which limit the measurements the UE is required to perform. Firstly, the UE is required to perform intra-frequency measurements only when the quality of the serving cell is below or equal to a threshold ('SintraSearch'). Furthermore, the UE is required to measure other frequencies/RATs of lower or equal priority only when the quality of the serving cell is below or equal to another threshold ('SnonintraSearch'). The UE is always required to measure frequencies and RATs of higher priority. For both cases (i.e. intra-frequency and inter-frequency) the UE may refrain from measuring when a receive level and a quality criterion is fulfilled. The required performance (i.e. how often the UE is expected to make the measurements, and to what extent this depends on, for example, the serving cell quality) is specified in [4].

3.3.4.2 Frequency/RAT Evaluation

E-UTRAN configures an absolute priority for all applicable frequencies of each RAT. In addition to the cell-specific priorities which are optionally provided via SI, E-UTRAN can assign UE-specific priorities via dedicated signalling. Of the frequencies that are indicated in the system information, the UE is expected to consider for cell reselection only those for which it has priorities. Equal priorities are not applicable for inter-RAT cell reselection.

The UE reselects to a cell on a higher priority frequency if the S-criterion (see Section 3.3.3) of the concerned target cell exceeds a high threshold ($Thresh_{X\text{-}High}$) for longer

than a certain duration $T_{reselection}$. The UE reselects to a cell on a lower-priority frequency if the S-criterion of the serving cell is below a low threshold ($Thresh_{Serving\text{-}Low}$) while the S-criterion of the target cell on a lower-priority frequency (possibly on another RAT) exceeds a low threshold ($Thresh_{X\text{-}Low}$) during the time interval $T_{reselection}$, and in the same time no cell on a higher-priority frequency is available. The UE evaluates the thresholds either based on receive level or on quality level, depending on which parameters E-UTRAN configures. Figure 3.13 illustrates the condition(s) to be met for reselecting to a cell on a higher-priority frequency (light grey bar) and to a cell on a lower priority frequency (dark grey bars).

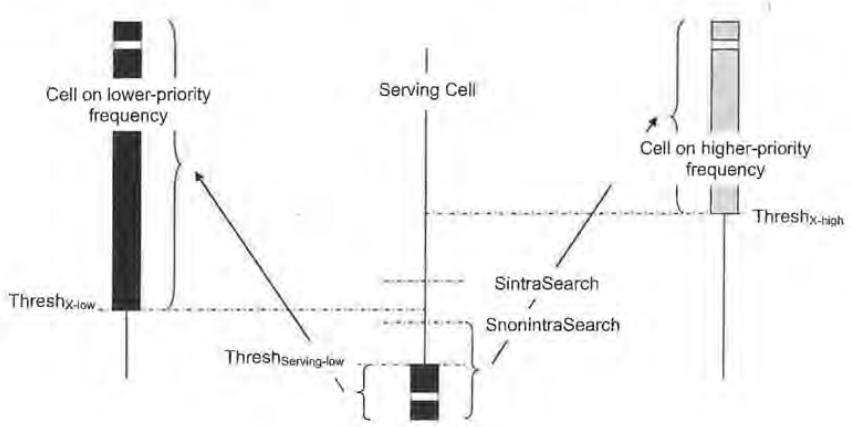


Figure 3.13: Frequency/RAT evaluation.

When reselecting to a frequency, possibly on another RAT, which has a different priority, the UE reselects to the highest-ranked cell on the concerned frequency (see Section 3.3.4.3).

Note that, as indicated in Section 3.2.3.4, thresholds and priorities are configured per frequency, while $T_{reselection}$ is configured per RAT.

From Release 8 onwards, UMTS and GERAN support the same priority-based cell reselection as provided in LTE, with a priority per frequency. Release 8 RANs will continue to handle legacy UEs by means of offset-based ranking. Likewise, Release 8 UEs should apply the ranking based on radio link quality (with offsets) unless UMTS or GERAN indicate support for priority-based reselection.

3.3.4.3 Cell Ranking

The UE ranks the intra-frequency cells and the cells on other frequencies having equal priority which fulfil the S-criterion using a criterion known as the *R-criterion*. The R-criterion generates rankings R_s and R_n for the serving cell and neighbour cells respectively:

For the serving cell: $R_s = Q_{meas,s} + Q_{hyst,s}$

For neighbour cells: $R_n = Q_{meas,n} + Q_{off\,s,n}$

where Q_{meas} is the measured cell received quality (RSRP) (see Section 22.3.1.1), $Q_{hyst,s}$ is a parameter controlling the degree of hysteresis for the ranking, and $Q_{off\,s,n}$ is an offset applicable between serving and neighbouring cells on frequencies of equal priority (the sum of the cell-specific and frequency-specific offsets).

The UE reselects to the highest-ranked candidate cell provided that it is better ranked than the serving cell for at least the duration of $T_{reselection}$. The UE scales the parameters $T_{reselection}$ and Q_{hyst}, depending on the UE speed (see Section 3.3.4.5 below).

3.3.4.4 Accessibility Verification

If the best cell on an LTE frequency is barred or reserved, the UE is required to exclude this cell from the list of cell reselection candidates. In this case, the UE may consider other cells on the same frequency unless the barred cell indicates (by means of field 'intraFreqReselection' within SIB1) that intra-frequency reselection is not allowed for a certain duration, unless the barred cell is an inaccessible Closed Subscriber Group (CSG) cell. If, however, the best cell is unsuitable for some other specific reason (e.g. because it belongs to a forbidden tracking area or to another non-equivalent PLMN), the UE is not permitted to consider any cell on the concerned frequency as a cell reselection candidate for a maximum of 300 s.

3.3.4.5 Speed Dependent Scaling

The UE scales the cell reselection parameters depending on its speed. This applies both in idle mode ($T_{reselection}$ and Q_{hyst}) and in connected mode (timeToTrigger). The UE speed is categorized by a mobility state (high, normal or low), which the UE determines based on the number of cell reselections/handovers which occur within a defined period, excluding consecutive reselections/handovers between the same two cells. The state is determined by comparing the count with thresholds for medium and high state, while applying some hysteresis. For idle and connected modes, separate sets of control parameters are used, signalled in SIB3 and within the measurement configuration respectively.

3.3.4.6 Cell Access Restrictions

The UE performs an access barring check during connection establishment (see Section 3.2.3.2). This function provides a means to control the load introduced by UE-originated traffic. There are separate means for controlling Mobile Originated (MO) calls and MO signalling. On top of the regular access class barring, Service Specific Access Control (SSAC) may be applied. SSAC facilitates separate control for MultiMedia TELephony (MMTEL) voice and video calls. Most of the SSAC functionality is handled by upper layers. In addition, separate access control exists to protect against E-UTRAN overload due to UEs accessing E-UTRAN merely to perform CSFB to CDMA2000.

Each UE belongs to an Access Class (AC) in the range 0–9. In addition, some UEs may belong to one or more high-priority ACs in the range 11–15, which are reserved for specific uses (e.g. security services, public utilities, emergency services, PLMN staff). AC10 is used for emergency access. Further details, for example regarding in which PLMN the high priority ACs apply, are provided in [5]. The UE considers access to be barred if access is barred for all its applicable ACs.

SIB2 may include a set of AC barring parameters for MO calls and/or MO signalling. This set of parameters comprises a probability factor and a barring timer for AC0–9 and a list of barring bits for AC11–15. For AC0–9, if the UE initiates a MO call and the relevant AC barring parameters are included, the UE draws a random number. If this number exceeds the probability factor, access is not barred. Otherwise access is barred for a duration which is randomly selected centred on the broadcast barring timer value. For AC11–15, if the UE initiates a MO call and the relevant AC barring parameters are included, access is barred whenever the bit corresponding to all of the UE's ACs is set. The behaviour is similar in the case of UE-initiated MO signalling.

For cell (re)selection, the UE is expected to consider cells which are neither barred nor reserved for operator or future use. In addition, a UE with an access class in the range 11–15 shall consider a cell that is (only) reserved for operator use and part of its home PLMN (or an equivalent) as a candidate for cell reselection. The UE is never allowed to (re)select a cell that is not a reselection candidate even for emergency access.

3.3.4.7 Any Cell Selection

When the UE is unable to find a suitable cell of the selected PLMN, it performs 'any cell selection'. In this case, the UE performs normal idle mode operation: monitoring paging, acquiring SI, performing cell reselection. In addition, the UE regularly attempts to find a suitable cell on other frequencies or RATs (i.e. not listed in SI). If a UE supporting voice services is unable to find a suitable cell, it should attempt to find an acceptable cell on any supported RAT regardless of the cell reselection priorities that are broadcast. The UE is not allowed to receive MBMS in this state.

3.3.4.8 Closed Subscriber Group

LTE supports the existence of cells which are accessible only for a limited set of UEs – a Closed Subscriber Group (CSG). In order to prevent UEs from attempting to register on a CSG cell on which they do not have access, the UE maintains a CSG white list, i.e. a list of CSG identities for which access has been granted to the UE. The CSG white list can be transferred to the UE by upper layers, or updated upon successful access to a CSG cell. To facilitate the latter, UEs support 'manual selection' of CSG cells which are not in the CSG white list. The manual selection may be requested by the upper layers, based on a text string broadcast by the cell. LTE also supports hybrid cells. Like CSG cells, hybrid cells broadcast a CSG identity; they are accessible as CSG cells by UEs whose CSG white lists include the CSG identity, and as normal cells by all other UEs (see Section 24.2.2).

3.4 Paging

To receive paging messages from E-UTRAN, UEs in idle mode monitor the PDCCH channel for an RNTI value used to indicate paging: the P-RNTI (see Section 9.2.2.2). The UE only needs to monitor the PDCCH channel at certain UE-specific occasions (i.e. at specific subframes within specific radio frames – see Section 6.2 for an introduction to the LTE radio frame structure.). At other times, the UE may apply DRX, meaning that it can switch off its receiver to preserve battery power.

The E-UTRAN configures which of the radio frames and subframes are used for paging. Each cell broadcasts a default paging cycle. In addition, upper layers may use dedicated signalling to configure a UE-specific paging cycle. If both are configured, the UE applies the lowest value. The UE calculates the radio frame (the Paging Frame (PF)) and the subframe within that PF (the Paging Occasion (PO)), which E-UTRAN applies to page the UE as follows:

$$\begin{aligned} \text{SFN mod } T &= (T/N) \times (\text{UE\_ID mod } N) \\ \text{i\_s} &= \lfloor \text{UE\_ID}/N \rfloor \text{ mod Ns} \\ T &= \text{UE DRX cycle (i.e. paging cycle)} - \min(T_c, T_{ue}) \\ N &= \min(T, \text{nB}) \\ \text{Ns} &= \max(1, \text{nB}/T) \end{aligned} \tag{3.1}$$

where:

T_c is the cell-specific default paging cycle {32, 64, 128, 256} radio frames,

T_{ue} is the UE-specific paging cycle {32, 64, 128, 256} radio frames,

N is the number of paging frames within the paging cycle of the UE,

UE_ID is the IMSI[20] mod 1024, with IMSI being the decimal rather than the binary number,

i_s is an index pointing to a pre-defined table defining the corresponding subframe,

nB is the number of 'paging subframes' per paging cycle (across all UEs in the cell),

Ns is the number of 'paging subframes' in a radio frame that is used for paging.

Table 3.4 includes a number of examples to illustrate the calculation of the paging radio frames (PF) and subframes (PO).

Table 3.4: Examples for calculation of paging frames and subframes.

Case	UE_ID	T_c	T_{ue}	T	nB	N	Ns	PF	i_s	PO
A	147	256	256	256	64	64	1	76	0	9
B	147	256	128	128	32	32	1	76	0	9
C	147	256	128	128	256	128	2	19	1	4

In cases A and B in Table 3.4, one out of every four radio frames is used for paging, using one subframe in each of those radio frames. For case B, there are 32 paging frames within the UE's paging cycle, across which the UEs are distributed based on the UE-identity. In case C, two subframes in each radio frame are used for paging, i.e. Ns = 2. In this case, there are 128 paging frames within the UE's paging cycle and the UEs are also distributed across the two subframes within the paging frame. The LTE specifications include a table that indicates the subframe applicable for each combination of Ns and i_s, which is the index that follows from Equation (3.1). Figure 3.14 illustrates cases B and C. All the shaded subframes can be used for paging; the darker ones are applicable for the UE with the indicated identity.

[20] International Mobile Subscriber Identity.

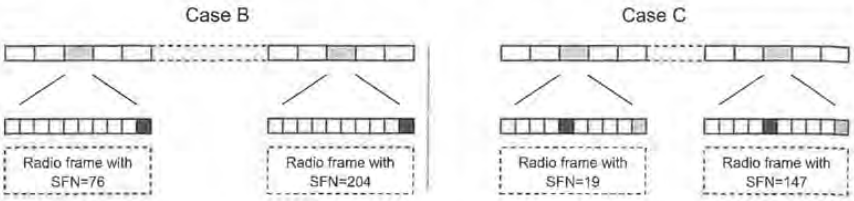


Figure 3.14: Paging frame and paging occasion examples.

3.5 Summary

The main aspects of the Control Plane protocols in LTE can be broken down into the Cell Selection and Reselection Procedures when the UE is in Idle Mode, and the RRC protocol when the UE is in Connected Mode.

The roles of these protocols include supporting security, mobility both between different LTE cells and between LTE and other radio systems, and establishment and reconfiguration of the radio bearers which carry control information and user data.

References[21]

[1] 3GPP Technical Specification 36.331, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol specification', www.3gpp.org.

[2] 3GPP Technical Specification 36.304, 'Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) procedures in idle mode (Release 9)', www.3gpp.org.

[3] 3GPP Technical Specification 33.401, '3GPP System Architecture Evolution; Security Architecture', www.3gpp.org.

[4] 3GPP Technical Specification 36.133, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Requirements for support of radio resource management', www.3gpp.org.

[5] 3GPP Technical Specification 22.011, 'Service accessibility', www.3gpp.org.

[21] All web sites confirmed 1st March 2011.

6

Introduction to Downlink Physical Layer Design

Matthew Baker

6.1 Introduction

The LTE downlink transmissions from the eNodeB consist of user-plane and control-plane data from the higher layers in the protocol stack (as described in Chapters 3 and 4) multiplexed with physical layer signalling to support the data transmission. The multiplexing of these various parts of the downlink signal is facilitated by the Orthogonal Frequency Division Multiple Access (OFDMA) structure described in Chapter 5, which enables the downlink signal to be subdivided into small units of time and frequency.

This subdivided structure is introduced below, together with an outline of the general steps in forming the transmitted downlink signal in the physical layer.

6.2 Transmission Resource Structure

The downlink transmission resources in LTE possess dimensions of time, frequency and space. The spatial dimension, measured in 'layers', is accessed by means of multiple 'antenna ports' at the eNodeB; for each antenna port a Reference Signal (RS) is provided to enable the User Equipment (UE) to estimate the radio channel (see Section 8.2); the techniques for using multiple antenna ports to exploit multiple spatial layers are explained in Section 11.2.

The time-frequency resources for each transmit antenna port are subdivided according to the following structure: the largest unit of time is the 10 ms radio frame, which is subdivided into ten 1 ms subframes, each of which is split into two 0.5 ms slots. Each slot comprises seven OFDM symbols in the case of the normal Cyclic Prefix (CP) length, or six if the

LTE – The UMTS Long Term Evolution: From Theory to Practice, Second Edition.
Stefania Sesia, Issam Toufik and Matthew Baker.
© 2011 John Wiley & Sons, Ltd. Published 2011 by John Wiley & Sons, Ltd.

extended CP is configured in the cell as explained in Section 5.4.1. In the frequency domain, resources are grouped in units of 12 subcarriers (thus occupying a total of 180 kHz with a subcarrier spacing of 15 kHz), such that one unit of 12 subcarriers for a duration of one slot is termed a Resource Block (RB).[1]

The smallest unit of resource is the Resource Element (RE), which consists of one subcarrier for a duration of one OFDM symbol. An RB thus comprises 84 REs in the case of the normal cyclic prefix length, and 72 REs in the case of the extended cyclic prefix.

The detailed resource structure is shown in Figure 6.1 for the normal cyclic prefix length.

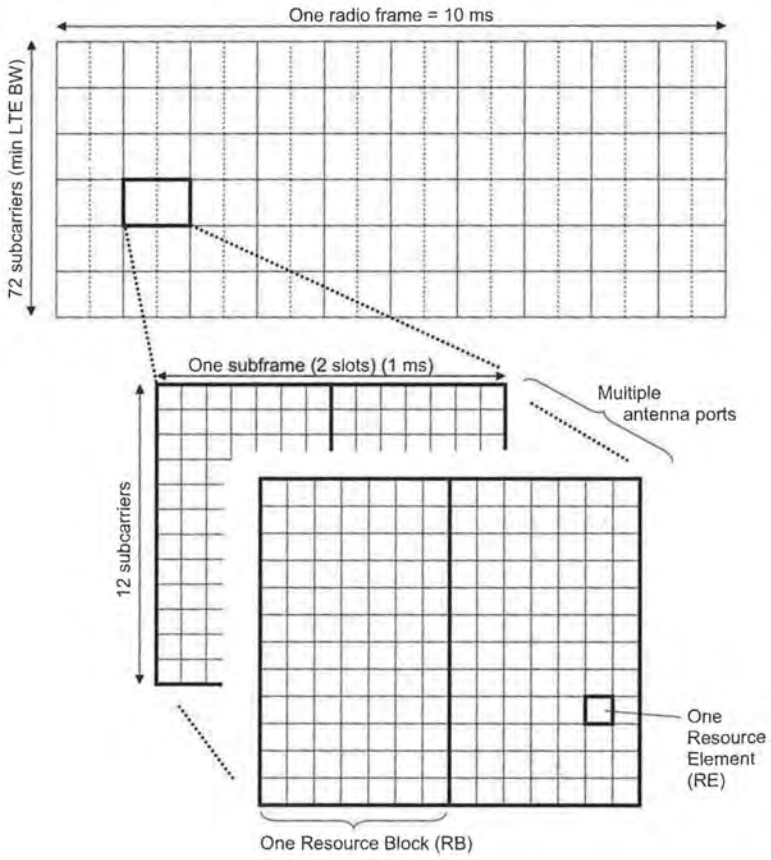


Figure 6.1: Basic time-frequency resource structure of LTE (normal cyclic prefix case).

[1] In the case of the 7.5 kHz subcarrier spacing which may be available for Multimedia Broadcast Multicast Service (MBMS) transmission in later releases of LTE (see Section 5.4.1 and Chapter 13), one RB consists of 24 subcarriers for a duration of one slot.

Within certain RBs, some REs are reserved for special purposes: synchronization signals (Chapter 7), Reference Signals (RSs – see Chapter 8), control signalling and critical broadcast system information (Chapter 9). The remaining REs are used for data transmission, and are usually allocated in pairs of RBs (the pairing being in the time domain).

The structure shown in Figure 6.1 assumes that all subframes are available for downlink transmission. This is known as 'Frame Structure Type 1' and is applicable for Frequency Division Duplexing (FDD) in paired radio spectrum, or for a standalone downlink carrier. For Time Division Duplexing (TDD) in unpaired radio spectrum, the basic structure of RBs and REs remains the same, but only a subset of the subframes are available for downlink transmission; the remaining subframes are used for uplink transmission, or for special subframes which contain a guard period to allow for switching between downlink and uplink transmission. The guard period allows the uplink transmission timing to be advanced as described in Section 18.2. This TDD structure is known as 'Frame Structure Type 2', of which seven different configurations are defined, as shown in Figure 6.2; these allow a variety of downlink-uplink ratios and switching periodicities. Further details of TDD operation using this frame structure are described in Chapter 23.

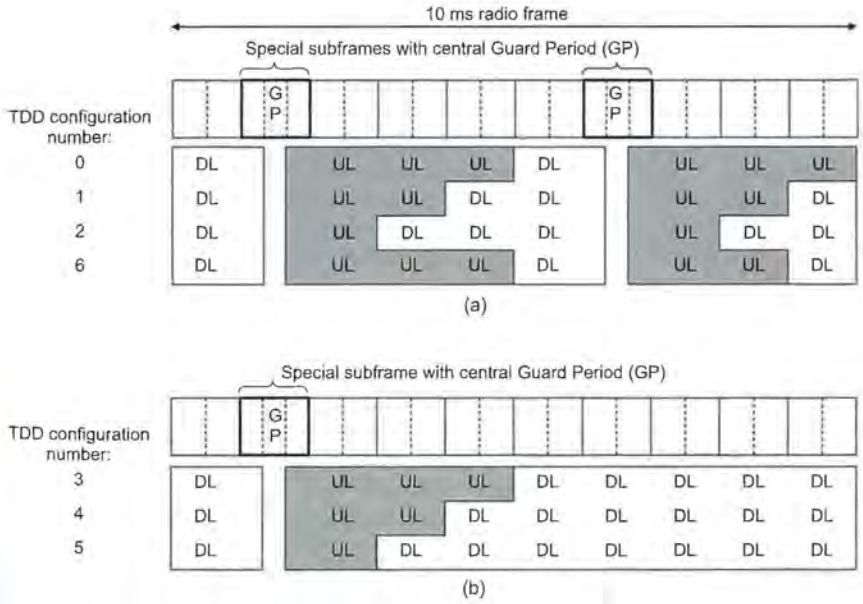


Figure 6.2: LTE subframe structure for TDD operation:
(a) configurations with 5 ms periodicity of switching from downlink (DL) to uplink (UL);
(b) configurations with 10 ms periodicity of switching from downlink (DL) to uplink (UL).

TERM EVOLUTION

arry data originating
nlink Shared Channel
onwards, the Physical
k control signalling is
ransmission resources

ers are summarized in
s.
hese channels to make
ation are explained in
ered in Chapter 11, and
ple users are described

m: North-Holland, 1998.
, MA: Kluwer Academic

7

Synchronization and Cell Search

Fabrizio Tomatis and Stefania Sesia

7.1 Introduction

A User Equipment (UE) wishing to access an LTE cell must first undertake a *cell search procedure*. This chapter focuses on the aspects of the physical layer that are designed to facilitate cell search. The way this relates to the overall mobility functionality and protocol aspects for cell reselection and handover is explained in Chapters 3 and 22. The performance requirements related to cell search and synchronization are explained in Section 22.2.

At the physical layer, the cell search procedure consists of a series of synchronization stages by which the UE determines time and frequency parameters that are necessary to demodulate the downlink and to transmit uplink signals with the correct timing. The UE also acquires some critical system parameters.

Three major synchronization requirements can be identified in the LTE system:

1. Symbol and frame timing acquisition, by which the correct symbol start position is determined, for example to set the Discrete Fourier Transform (DFT) window position;
2. Carrier frequency synchronization, which is required to reduce or eliminate the effect of frequency errors[1] arising from a mismatch of the local oscillators between the transmitter and the receiver, as well as the Doppler shift caused by any UE motion;
3. Sampling clock synchronization.

7.2 Synchronization Sequences and Cell Search in LTE

The cell search procedure in LTE begins with a synchronization procedure which makes use of two specially designed physical signals that are broadcast in each cell: the Primary Synchronization Signal (PSS) and the Secondary Synchronization Signal (SSS). The detection of

[1] Frequency offsets may arise from factors such as temperature drift, ageing and imperfect calibration.

LTE – The UMTS Long Term Evolution: From Theory to Practice, Second Edition.
Stefania Sesia, Issam Toufik and Matthew Baker.
© 2011 John Wiley & Sons, Ltd. Published 2011 by John Wiley & Sons, Ltd.

these two signals not only enables time and frequency synchronization, but also provides the UE with the physical layer identity of the cell and the cyclic prefix length, and informs the UE whether the cell uses Frequency Division Duplex (FDD) or Time Division Duplex (TDD).

In the case of initial synchronization (when the UE is not already camping on or connected to an LTE cell) after detecting the synchronization signals, the UE decodes the Physical Broadcast CHannel (PBCH), from which critical system information is obtained (see Section 9.2.1). In the case of neighbour cell identification, the UE does not need to decode the PBCH; it simply makes quality-level measurements based on the reference signals (see Chapter 8) transmitted from the newly detected cell and uses them for cell reselection (in RRC_IDLE state) or handover (in RRC_CONNECTED state); in the latter case, the UE reports these measurements to its serving cell.

The cell search and synchronization procedure is summarized in Figure 7.1, showing the information ascertained by the UE at each stage. The PSS and SSS structure is specifically designed to facilitate this acquisition of information.

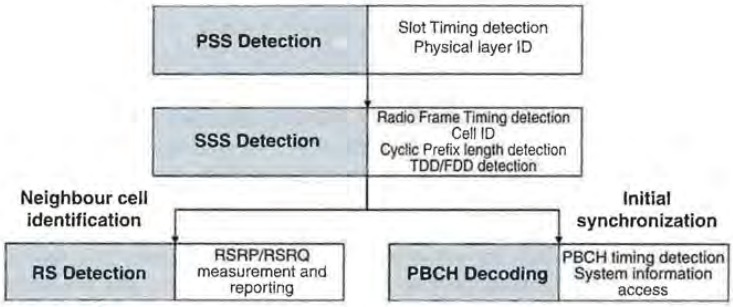


Figure 7.1: Information acquired at each step of the cell search procedure.

The PSS and SSS structure in time is shown in Figure 7.2 for the FDD case and in Figure 7.3 for TDD: the synchronization signals are transmitted periodically, twice per 10 ms radio frame. In an FDD cell, the PSS is always located in the last OFDM (Orthogonal Frequency Division Multiplexing) symbol of the first and 11^{th} slots of each radio frame (see Chapter 6), thus enabling the UE to acquire the slot boundary timing independently of the Cyclic Prefix (CP) length. The SSS is located in the symbol immediately preceding the PSS, a design choice enabling coherent detection[2] of the SSS relative to the PSS, based on the assumption that the channel coherence duration is significantly longer than one OFDM symbol. In a TDD cell, the PSS is located in the third symbol of the 3^{rd} and 13^{th} slots, while the SSS is located three symbols earlier; coherent detection can be used under the assumption that the channel coherence time is significantly longer than four OFDM symbols.

The precise position of the SSS changes depending on the length of the CP configured for the cell. At this stage of the cell detection process, the CP length is unknown a priori

[2]See Section 7.3.

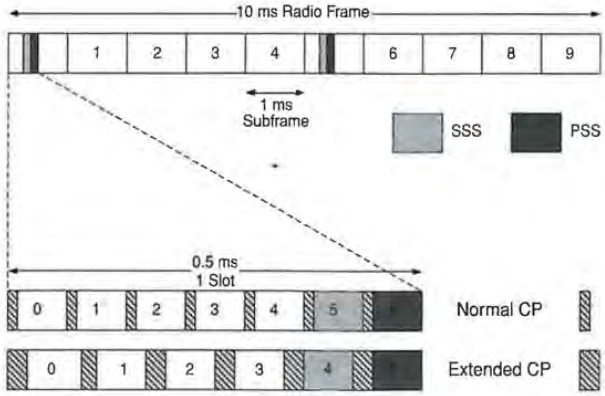


Figure 7.2: PSS and SSS frame and slot structure in time domain in the FDD case.

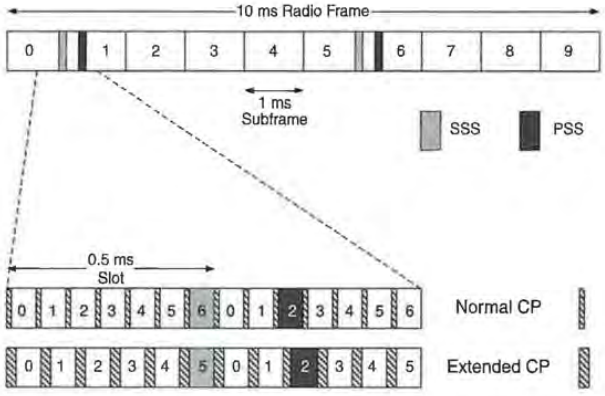


Figure 7.3: PSS and SSS frame and slot structure in time domain in the TDD case.

to the UE, and it is, therefore, blindly detected by checking for the SSS at the two possible positions.[3]

[3]Hence a total of four possible SSS positions must be checked if the UE is searching for both FDD and TDD cells.

While the PSS in a given cell is the same in every subframe in which it is transmitted, the two SSS transmissions in each radio frame change in a specific manner as described in Section 7.2.3, thus enabling the UE to establish the position of the 10 ms radio frame boundary.

In the frequency domain, the mapping of the PSS and SSS to subcarriers is shown in Figure 7.4. The PSS and SSS are transmitted in the central six Resource Blocks[4] (RBs), enabling the frequency mapping of the synchronization signals to be invariant with respect to the system bandwidth (which can in principle vary from 6 to 110 RBs to suit channel bandwidths between around 1.4 MHz and 20 MHz); this allows the UE to synchronize to the network without any a priori knowledge of the allocated bandwidth. The PSS and SSS are each comprised of a sequence of length 62 symbols, mapped to the central 62 subcarriers around the d.c. subcarrier, which is left unused. This means that the five resource elements at each extremity of each synchronization sequence are not used. This structure enables the UE to detect the PSS and SSS using a size-64 Fast Fourier Transform (FFT) and a lower sampling rate than would have been necessary if all 72 subcarriers were used in the central six resource blocks. The shorter length for the synchronization sequences also avoids the possibility in a TDD system of a high correlation with the uplink demodulation reference signals which use the same kind of sequence as the PSS (see Chapter 15).

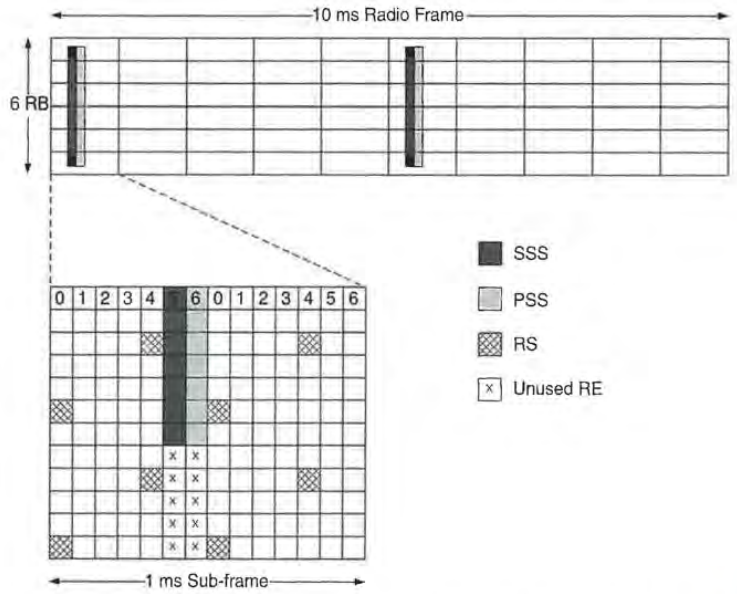


Figure 7.4: PSS and SSS frame structure in frequency and time domain for an FDD cell.

[4]Six Resource Blocks correspond to 72 subcarriers.

In the case of multiple transmit antennas being used at the eNodeB, the PSS and SSS are always transmitted from the same antenna port[5] in any given subframe, while between different subframes they may be transmitted from different antenna ports in order to benefit from time-switched antenna diversity.

The particular sequences which are transmitted for the PSS and SSS in a given cell are used to indicate the physical layer cell identity to the UE. There are 504 unique Physical Cell Identities (PCIs) in LTE, grouped into 168 groups of three identities. The three identities in a group would usually be assigned to cells under the control of the same eNodeB. Three PSS sequences are used to indicate the cell identity within the group, and 168 SSS sequences are used to indicate the identity of the group.[6]

The PSS uses sequences known as *Zadoff–Chu*. This kind of sequence is widely used in LTE, including for the uplink reference signals and random access preambles (see Chapters 15 and 17) in addition to the PSS. Therefore, Section 7.2.1 is devoted to an explanation of the fundamental principles behind Zadoff–Chu sequences, before discussing the specific constructions of the PSS and SSS sequences in the subsequent sections.

7.2.1 Zadoff–Chu Sequences

Zadoff–Chu (ZC) sequences (also known as *Generalized Chirp-Like* (GCL) sequences) are named after the authors of [1] and [2]. ZC sequences are non-binary unit-amplitude sequences [3], which satisfy a Constant Amplitude Zero Autocorrelation (CAZAC) property. CAZAC sequences are complex signals of the form e^{ja_k}. The ZC sequence of odd-length N_{ZC} is given by

$$a_q(n) = \exp\left[-j2\pi q \frac{n(n+1)/2 + ln}{N_{ZC}}\right] \tag{7.1}$$

where $q \in \{1, \ldots, N_{ZC} - 1\}$ is the ZC sequence root index, $n = 0, 1, \ldots, N_{ZC} - 1$, $l \in \mathbb{N}$ is any integer. In LTE, $l = 0$ is used for simplicity.

ZC sequences have the following three important properties:

Property 1. A ZC sequence has constant amplitude, and its N_{ZC}-point DFT also has constant amplitude. The constant amplitude property limits the Peak-to-Average Power Ratio (PAPR) and generates bounded and time-flat interference to other users. It also simplifies the implementation as only phases need to be computed and stored, not amplitudes.

Property 2. ZC sequences of any length have 'ideal' cyclic autocorrelation (i.e. the correlation with its circularly shifted version is a delta function). The zero autocorrelation property may be formulated as:

$$r_{kk}(\sigma) = \sum_{n=0}^{N_{ZC}-1} a_k(n) a_k^*[(n+\sigma)] = \delta(\sigma) \tag{7.2}$$

where $r_{kk}(\cdot)$ is the discrete periodic autocorrelation function of a_k at lag σ. This property is of major interest when the received signal is correlated with a reference sequence and the received reference sequences are misaligned. As an example, Figure 7.5 shows the difference

[5] The concept of an antenna port in LTE is explained in Section 8.2.

[6] A subset of the available PCIs can be reserved for Closed Subscriber Group (CSG) cells (e.g. Home eNodeBs). If this is the case, the reserved PCIs are indicated in System Information Block 4 (SIB4), enabling the UE to eliminate those PCIs from its search if CSG cells are not applicable for it.

between the periodic autocorrelation of a truncated Pseudo-Noise (PN) sequence (as used in WCDMA [4]) and a ZC sequence. Both are 839 symbols long in this example. The ZC periodic autocorrelation is exactly zero for $\sigma \neq 0$ and it is non-zero for $\sigma = 0$, whereas the PN periodic autocorrelation shows significant peaks, some above 0.1, at non-zero lags.

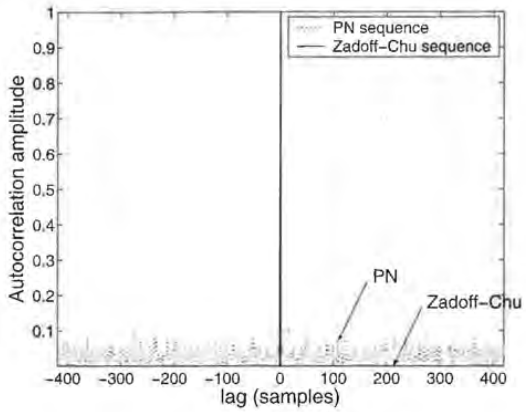


Figure 7.5: Zadoff–Chu versus PN sequence: periodic autocorrelation.

The main benefit of the CAZAC property is that it allows multiple orthogonal sequences to be generated from the same ZC sequence. Indeed, if the periodic autocorrelation of a ZC sequence provides a single peak at the zero lag, the periodic correlation of the same sequence against its cyclic shifted replica provides a peak at lag N_{CS}, where N_{CS} is the number of samples of the cyclic shift. This creates a Zero-Correlation Zone (ZCZ) between the two sequences. As a result, as long as the ZCZ is dimensioned to cope with the largest possible expected time misalignment between them, the two sequences are orthogonal for all transmissions within this time misalignment.

Property 3. The absolute value of the cyclic cross-correlation function between any two ZC sequences is constant and equal to $1/\sqrt{N_{ZC}}$,[7] if $|q_1 - q_2|$ (where q_1 and q_2 are the sequence indices) is relatively prime with respect to N_{ZC} (a condition that can be easily guaranteed if N_{ZC} is a prime number). The cross-correlation of $\sqrt{N_{ZC}}$ at all lags achieves the theoretical minimum cross-correlation value for any two sequences that have ideal autocorrelation.

Selecting N_{ZC} as a prime number results in $N_{ZC} - 1$ ZC sequences which have the optimal cyclic cross-correlation between any pair. However, it is not always convenient to use sequences of prime length. In general, a sequence of non-prime length may be generated by either cyclic extension or truncation of a prime-length ZC sequence. A further useful property of ZC sequences is that the DFT of a ZC sequence $x_u(n)$ (in Equation (7.1)) is a weighted cyclicly shifted ZC sequence $X_w(k)$ such that $w = -1/u \bmod N_{ZC}$. This means that a ZC sequence can be generated directly in the frequency domain without the need for a DFT operation.

[7]Note that this value corresponds to the normalized cross-correlation function.

8

Reference Signals and Channel Estimation

Andrea Ancora, Stefania Sesia and Alex Gorokhov

8.1 Introduction

A simple communication system can be generally modelled as in Figure 8.1, where the transmitted signal **x** passes through a radio channel **H** and suffers additive noise before being received. Mobile radio channels usually exhibit multipath fading, which causes Inter-Symbol Interference (ISI) in the received signal. In order to remove ISI, various kinds of equalization and detection algorithms can be utilized, which may or may not exploit knowledge of the Channel Impulse Response (CIR). Orthogonal Frequency Division Multiple Access (OFDMA) is particularly robust against ISI, thanks to its structure and the use of the Cyclic Prefix (CP) which allows the receiver to perform a low-complexity single-tap scalar equalization in the frequency domain, as described in Section 5.2.1.

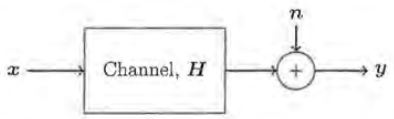


Figure 8.1: A simple transmission model.

As explained in Section 7.3, when the detection method exploits channel knowledge, it is generally said to be 'coherent'; otherwise it is called 'non-coherent'. Coherent detection can make use of both amplitude and phase information carried by the complex signals, and not of only amplitude information as with non-coherent detection. Optimal reception by coherent detection therefore typically requires accurate estimation of the propagation channel.

LTE – The UMTS Long Term Evolution: From Theory to Practice, Second Edition.
Stefania Sesia, Issam Toufik and Matthew Baker.
© 2011 John Wiley & Sons, Ltd. Published 2011 by John Wiley & Sons, Ltd.

The main advantage of coherent detection is the simplicity of implementation compared to the more complex algorithms required by non-coherent detection for equivalent performance. However, this simplicity comes at a price, namely the overhead needed in order to be able to estimate the channel. A common and simple way to estimate the channel is to exploit known signals which do not carry any data, but which therefore cause a loss in spectral efficiency. In general, it is not an easy task to find the optimal trade-off between minimizing the spectral efficiency loss due to the overhead and providing adequate ability to track variations in the channel.

Other possible techniques for channel estimation include exploiting the correlation properties of the channel or using blind estimation.

Once synchronization between an eNodeB and a UE has been achieved, LTE (in common with earlier systems such as GSM and UMTS) is a coherent communication system, for which purpose known Reference Signals (RSs) are inserted into the transmitted signal structure.

In general, a variety of methods can be used to embed RSs into a transmitted signal. The RSs can be multiplexed with the data symbols (which are unknown at the receiver) in either the frequency, time or code domains (the latter being used in the case of the common pilot channel in the UMTS downlink). A special case of time multiplexing, known as preamble-based training, involves transmitting the RSs at the beginning of each data burst. Multiplexing-based techniques have the advantage of low receiver complexity, as the data symbol detection is decoupled from the channel estimation problem. Alternatively, RSs may be superimposed on top of the unknown data, without the two necessarily being orthogonal. Note that multiplexing RSs in the code domain is a particular type of superposition with a constraint on orthogonality between known RSs and the unknown data. A comprehensive analysis of the optimization of RS design can be found in [1, 2].

Orthogonal RS multiplexing is by far the most common technique. For example, to facilitate channel estimation in the UMTS downlink, two types of orthogonal RS are provided. The first is code-multiplexed, available to all users in a cell, and uses a specific spreading code which is orthogonal to the codes used to spread the users' data. The second type is time-multiplexed dedicated RSs, which may in some situations be inserted into the users' data streams [3].

In the LTE downlink, the OFDM transmission can be described by a two-dimensional lattice in time and frequency, as shown in Figure 6.1 and described in Chapter 6. This structure facilitates the multiplexing of the RSs, which are mapped to specific Resource Elements (REs) of the two-dimensional lattice in Figure 6.1 according to patterns explained in Section 8.2.

In order to estimate the channel as accurately as possible, all correlations between channel coefficients in time, frequency and space should be taken into account. Since RSs are sent only on particular OFDM REs (i.e. on particular OFDM symbols on particular subcarriers), channel estimates for the REs which do not bear RSs have to be computed via interpolation. The optimal interpolating channel estimator in terms of mean-squared error is based on a two-dimensional Wiener filter interpolation [4]. Due to the high complexity of such a filter, a trade-off between complexity and accuracy is achieved by using one-dimensional filters. In Sections 8.4, 8.5 and 8.6, the problem of channel estimation is approached from a theoretical point of view, and some possible solutions are described.

The work done in the field of channel estimation, and the corresponding literature available, is vast. Nevertheless many challenges still remain, and we refer the interested reader to [2, 5] for general surveys of open issues in this area.

8.2 Design of Reference Signals in the LTE Downlink

In the LTE downlink, five different types of RS are provided [6, Section 6.10]:

- Cell-specific RSs (often referred to as 'common' RSs, as they are available to all UEs in a cell and no UE-specific processing is applied to them);
- UE-specific RSs (introduced in Release 8, and extended in Releases 9 and 10), which may be embedded in the data for specific UEs (also known as DeModulation Reference Signals (DM-RSs);
- MBSFN-specific RSs, which are used only for Multimedia Broadcast Single Frequency Network (MBSFN) operation and are discussed further in Section 13.4.1;
- Positioning RSs, which from Release 9 onwards may be embedded in certain 'positioning subframes' for the purpose of UE location measurements; these are discussed in Section 19.3.1;
- Channel State Information (CSI) RSs, which are introduced in Release 10 specifically for the purpose of estimating the downlink channel state and not for data demodulation (see Section 29.1.2).

Each RS pattern is transmitted from an *antenna port* at the eNodeB. An antenna port may in practice be implemented either as a single physical transmit antenna, or as a combination of multiple physical antenna elements. In either case, the signal transmitted from each antenna port is not designed to be further deconstructed by the UE receiver: the transmitted RS corresponding to a given antenna port defines the antenna port from the point of view of the UE, and enables the UE to derive a channel estimate for all data transmitted on that antenna port – regardless of whether it represents a single radio channel from one physical antenna or a composite channel from a multiplicity of physical antenna elements together comprising the antenna port. The designations of the antenna ports available in LTE are summarized below:

- Antenna Ports 0–3: cell-specific RSs – see Section 8.2.1;
- Antenna Port 4: MBSFN – see Section 13.4.1;
- Antenna Port 5: UE-specific RSs for single-layer beamforming – see Section 8.2.2;
- Antenna Port 6: positioning RSs (introduced in Release 9) – see Section 19.3.1;
- Antenna Ports 7–8: UE-specific RSs for dual-layer beamforming (introduced in Release 9) – see Section 8.2.3;
- Antenna Ports 9–14: UE-specific RSs for multi-layer beamforming (introduced in Release 10) – see Section 29.1.1;
- Antenna Ports 15–22: CSI RSs (introduced in Release 10) – see Section 29.1.2.

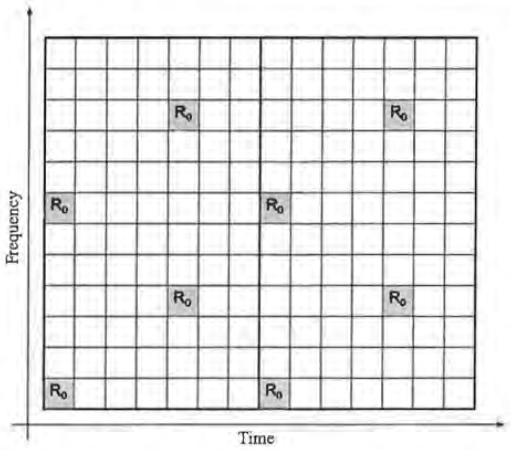


Figure 8.2: Cell-specific reference symbol arrangement in the case of normal CP length for one antenna port. Reproduced by permission of © 3GPP.

Details of the techniques provided for multi-antenna transmission in LTE can be found in Chapter 11 for Releases 8 and 9, and in Chapter 29 for the new techniques introduced in Release 10 for LTE-Advanced.

8.2.1 Cell-Specific Reference Signals

The cell-specific RSs enable the UE to determine the phase reference for demodulating the downlink control channels (see Section 9.3) and the downlink data in most transmission modes[1] of the Physical Downlink Shared Channel (PDSCH – see Section 9.2.2). If UE-specific precoding is applied to the PDSCH data symbols before transmission,[2] downlink control signalling is provided to inform the UE of the corresponding phase adjustment it should apply relative to the phase reference provided by the cell-specific RSs. The cell-specific RSs are also used by the UEs to generate Channel State Information (CSI) feedback (see Sections 10.2.1 and 11.2.2.4).

References [7, 8] show that in an OFDM-based system an equidistant arrangement of reference symbols in the lattice structure achieves the Minimum Mean-Squared Error (MMSE) estimate of the channel. Moreover, in the case of a uniform reference symbol grid, a 'diamond shape' in the time-frequency plane can be shown to be optimal.

In LTE, the arrangement of the REs on which the cell-specific RSs are transmitted follows these principles. Figure 8.2 illustrates the RS arrangement for the normal CP length.[3]

[1] Transmission modes 1 to 6.

[2] In PDSCH transmission modes 3 to 6.

[3] In the case of the extended CP, the arrangement of the reference symbols changes slightly, but the explanations in the rest of this chapter are no less valid. The detailed arrangement of reference symbols for the extended CP can be found in [6].

The LTE system is designed to work under high-mobility assumptions, in contrast to WLAN systems which are generally optimized for pedestrian-level mobility. WLAN systems typically use a preamble-based training sequence, and the degree of mobility such systems can support depends on how often the preamble is transmitted.

The required spacing in time between the reference symbols can be determined by considering the maximum Doppler spread (highest speed) to be supported, which for LTE corresponds to 500 km/h [9]. The Doppler shift is $f_d = (f_c\, v/c)$ where f_c is the carrier frequency, v is the UE speed in metres per second, and c is the speed of light ($3 \cdot 10^8$ m/s). Considering $f_c = 2$ GHz and $v = 500$ km/h, then the Doppler shift is $f_d \simeq 950$ Hz. According to Nyquist's sampling theorem, the minimum sampling frequency needed in order to reconstruct the channel is therefore given by $T_c = 1/(2f_d) \simeq 0.5$ ms under the above assumptions. This implies that two reference symbols per slot are needed in the time domain in order to estimate the channel correctly.

In the frequency direction, there is one reference symbol every six subcarriers on each OFDM symbol that includes reference symbols, but the reference symbols are staggered so that within each Resource Block (RB) there is one reference symbol every threesubcarriers, as shown in Figure 8.2. This spacing is related to the expected coherence bandwidth of the channel, which is in turn related to the channel delay spread. In particular the 90% and 50% coherence bandwidths[4] are given by $B_{c,90\%} = 1/50\sigma_\tau$ and $B_{c,50\%} = 1/5\sigma_\tau$ respectively, where σ_τ is the r.m.s delay spread. In [10], the maximum r.m.s channel delay spread considered is 991 ns, corresponding to $B_{c,90\%} = 20$ kHz and $B_{c,50\%} = 200$ kHz. In LTE, the spacing between two reference symbols in frequency, in one RB, is 45 kHz, thus allowing the expected frequency-domain variations of the channel to be resolved.

Up to four cell-specific antenna ports, numbered 0 to 3, may be used by an LTE eNodeB, thus requiring the UE to derive up to four separate channel estimates.[5] For each antenna port, a different RS pattern has been designed, with particular attention having been given to the minimization of the intra-cell interference between the multiple transmit antenna ports. In Figure 8.3, R_p indicates that the RE is used for the transmission of an RS on antenna port p. When an RE is used to transmit an RS on one antenna port, the corresponding RE on the other antenna ports is set to zero to limit the interference.

From Figure 8.3, it can be noticed that the density of reference symbols for the third and fourth antenna ports is half that of the first two; this is to reduce the overhead in the system. Frequent reference symbols are useful for high-speed conditions as explained above. In cells with a high prevalence of high-speed users, the use of four antenna ports is unlikely, hence for these conditions reference symbols with lower density can provide sufficient channel estimation accuracy.

All the RSs (cell-specific, UE-specific or MBSFN-specific) are QPSK modulated – a constant modulus modulation. This property ensures that the Peak-to-Average Power Ratio (PAPR) of the transmitted waveform is kept low. The signal can be written as

$$r_{l,n_s}(m) = \frac{1}{\sqrt{2}}[1 - 2c(2m)] + j\frac{1}{\sqrt{2}}[1 - 2c(2m+1)] \tag{8.1}$$

[4] $B_{c,x\%}$ is the bandwidth where the autocorrelation of the channel in the frequency domain is equal to $x\%$ of the peak.

[5] Any MBSFN and UE-specific RSs, if transmitted, constitute additional independent antenna ports in the LTE specifications.

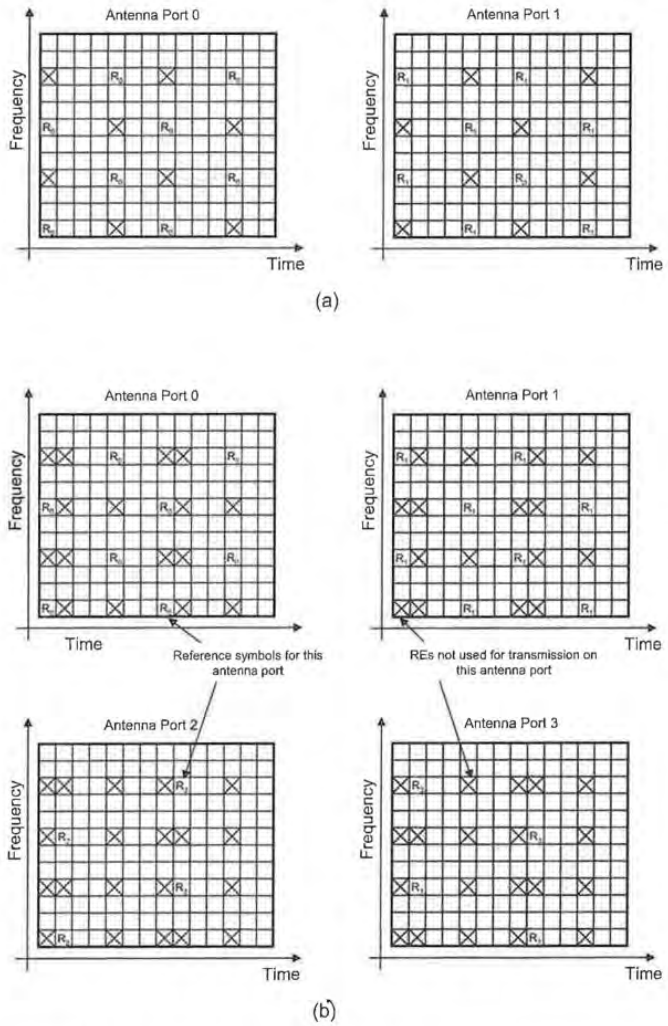


Figure 8.3: Cell-specific RS arrangement in the case of normal CP length for (a) two antenna ports, (b) four antenna ports. Reproduced by permission of © 3GPP.

where m is the index of the RS, n_s is the slot number within the radio frame and 'l' is the symbol number within the time slot. The pseudo-random sequence $c(i)$ is comprised of a length-31 Gold sequence, already introduced in Chapter 6, with different initialization values depending on the type of RSs. For the cell-specific RSs, the sequence is reinitialized at the

start of each OFDM symbol, with a value that depends on the cell identity, N_{ID}^{cell}. The cell-specific RS sequence therefore carries unambiguously one of the 504 different cell identities.

A cell-specific frequency shift is applied to the patterns of reference symbols shown in Figures 8.2 and 8.3, given by $N_{ID}^{cell} \bmod 6$.[6] This shift helps to avoid time-frequency collisions between cell-specific RSs from up to six adjacent cells. Avoidance of collisions is particularly relevant in cases when the transmission power of the RS is boosted, as is possible in LTE up to a maximum of 6 dB relative to the surrounding data symbols. RS power-boosting is designed to improve channel estimation in the cell, but if adjacent cells transmit high-power RSs on the same REs, the resulting inter-cell interference will prevent the benefit from being realized.

8.2.2 UE-Specific Reference Signals in Release 8

In Release 8 of LTE, UE-specific RSs may be transmitted in addition to the cell-specific RSs described above if the UE is configured (by higher-layer RRC signalling) to receive its downlink PDSCH data in transmission mode 7 (see Section 9.2.2.1). The UE-specific RSs are embedded only in the RBs to which the PDSCH is mapped for those UEs. If UE-specific RSs are transmitted, the UE is expected to use them to derive the channel estimate for demodulating the data in the corresponding PDSCH RBs. The same precoding is applied to the UE-specific RSs as to the PDSCH data symbols, and therefore there is no need for signalling to inform the UE of the precoding applied. Thus the UE-specific RSs are treated as being transmitted using a distinct antenna port (number 5), with its own channel response from the eNodeB to the UE.

A typical usage of the UE-specific RSs is to enable beamforming of the data transmissions to specific UEs. For example, rather than using the physical antennas used for transmission of the other (cell-specific) antenna ports, the eNodeB may use a correlated array of physical antenna elements to generate a narrow beam in the direction of a particular UE. Such a beam will experience a different channel response between the eNodeB and UE, thus requiring the use of UE-specific RSs to enable the UE to demodulate the beamformed data coherently. The use of UE-specific beamforming is discussed in more detail in Section 11.2.2.3.

As identified in [11], the structure shown in Figure 8.4 (for the normal CP) has been chosen because there is no collision with the cell specific RSs, and hence the presence of UE-specific RSs does not affect features related to the cell-specific RSs. The UE-specific RSs have a similar pattern to that of the cell-specific RSs, which allows a UE to re-use similar channel estimation algorithms. The density is half that of the cell-specific RS, hence minimizing the overhead. Unlike the cell-specific RSs, the sequence for the UE-specific RSs is only reinitialized at the start of each subframe, as the number of REs to which the sequence is mapped in one OFDM symbol may be very small (in the event of a small number of RBs being transmitted to a UE). The initialization depends on the UE's identity.

The corresponding pattern for use in case of the extended CP being configured in a cell can be found in [6, Section 6.10.3.2].

8.2.3 UE-Specific Reference Signals in Release 9

A new design for UE-specific RSs is defined in Release 9 of the LTE specifications in order to extend UE-specific RS support to dual layer transmission. This includes transmission of

[6] The mod6 operation is used because reference symbols are spaced apart by six subcarriers in the lattice grid.

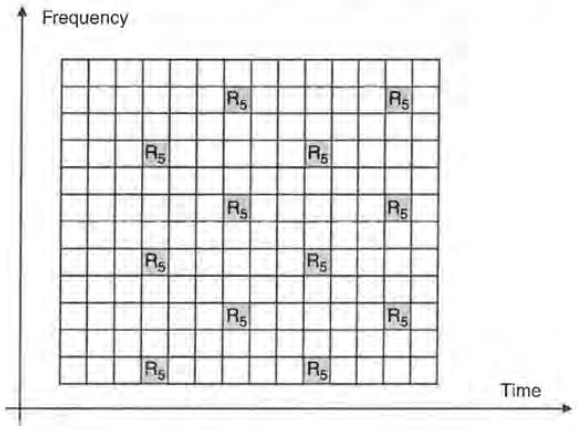


Figure 8.4: Release 8 UE-specific RS arrangement for PDSCH transmission mode 7 with normal CP. Reproduced by permission of © 3GPP.

two spatial layers to one UE, or single-layer transmission to each of two UEs as a Multi-User Multiple Input Multiple Output (MU-MIMO) transmission (see Chapter 11).

Dual layer beamforming in Release 9 is a precursor to higher-rank single- and multi-user transmissions in Release 10, and therefore scalability towards LTE-Advanced (see Section 29.1) was a key factor in the choice of design for the new RSs. This includes enabling efficient inter-cell coordination, which can be facilitated by choosing the set of REs to be non-cell-specific (i.e. without a frequency shift that depends on the cell identity).

A further consideration was backward compatibility with Release 8; this requires coexistence with the 'legacy' physical channels, and therefore the REs carrying the new UE-specific RSs have to avoid the cell-specific RSs and downlink control channels.

A final design requirement was to keep equal power spectral density of the downlink transmission in one RB. This favours designs where the UE-specific RSs for both layers are present in the same OFDM symbol.

Since UE-specific RSs are designed for time-frequency channel estimation within a given RB, the eigen-structure of the time and frequency channel covariance matrix provides insights into the optimal pattern of REs for the RSs for MMSE channel estimation. As an example, Figure 8.5 shows the eigenvalues and three dominant eigenvectors of the time-domain channel covariance matrix for a pair of RBs in a subframe for various mobile speeds assuming spatially uniform (Jakes) scattering. The same behaviour is observed in the frequency domain.

It can be seen that the three principal eigenvectors can be accurately approximated by constant, linear and quadratic functions respectively. Such a channel eigen-structure stems from the time-frequency variations across a pair of RBs being limited. Most of the energy is captured by the constant and linear components, and hence these should be the focus for optimization. While the reference symbol locations do not affect estimation accuracy for the

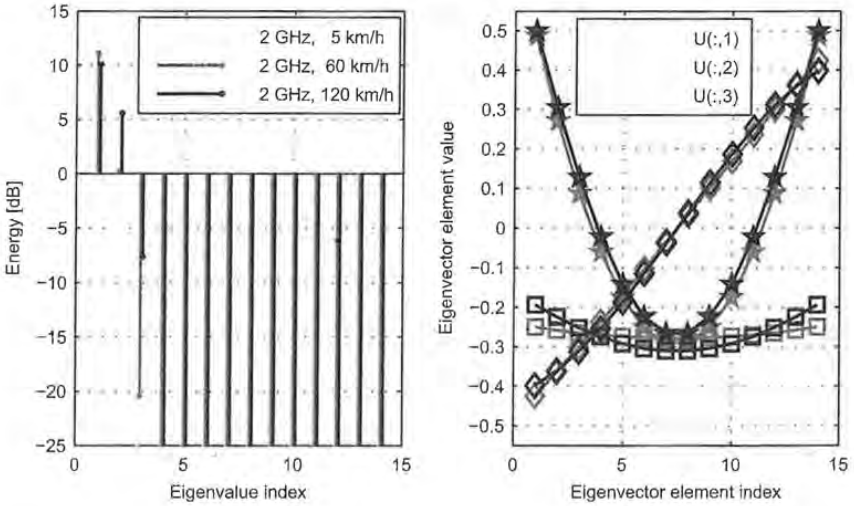


Figure 8.5: Time-domain channel covariance over a pair of RBs: eigenvalues (left) and eigenvectors (right), assuming spatially uniform (Jakes) scattering. The eigenvalue energies are normalized relative to the per-RE energy.

constant component, the best strategy for estimation of the linear component is to allocate RS energy at the edges of the measured region. Finally, the quadratic component is small but non-negligible, warranting three disjoint RS positions across the bandwidth of an RB.

Based on these observations, Figure 8.6 shows the pattern adopted for the new UE-specific RS pattern in Release 9.[7] The reference symbols are positioned in the earliest and latest available pairs of OFDM symbols that avoid collisions with the cell-specific RSs. Pairs of REs are used so that the UE-specific RSs for the two layers can be code-multiplexed. The UE-specific RSs for the two layers using this pattern are termed antenna ports 7 and 8. A UE configured to use the Release 9 dual-layer UE-specific RSs is configured in PDSCH transmission mode 8 (see Section 9.2.2.1).

Length-2 orthogonal Walsh codes are used for the code-multiplexing of the two RS ports.[8] Compared to the use of frequency-multiplexed RS ports, code multiplexing can improve the accuracy of interference estimation and potentially simplifies the implementation by maintaining the same set of RS REs regardless of the number of layers transmitted (thus facilitating dynamic switching between one and two layers).

Finally, the Release 9 UE-specific RS sequence is initialized using only the cell identity (without the UE identity used for the Release 8 UE-specific RSs), in order to enable the two orthogonally code-multiplexed UE-specific RS ports to be used for MU-MIMO (with the two RS ports assigned to different UEs). In addition, two different RS sequence initializations are

[7]Note that distributed RB mapping (see Section 9.2.2.1) is not supported in conjunction with UE-specific RSs.

[8]The mapping of the second Walsh code to each pair of REs reverses from one subcarrier to the next; this is to minimize the Peak-to-Average Power Ratio (PAPR) of the RS transmissions [12].

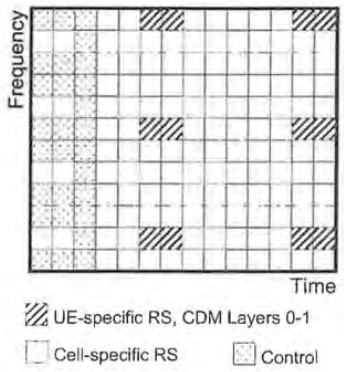


Figure 8.6: Dual-layer UE-specific RS arrangement in Release 9.

provided, enabling non-orthogonal UE-specific RS multiplexing to be used, for example for MU-MIMO with up to dual layer transmission to each UE. An example of a possible usage is shown in Figure 8.7.

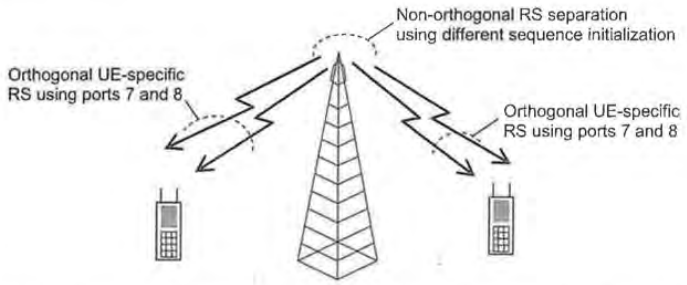


Figure 8.7: Example of usage of orthogonal and non-orthogonal UE-specific RSs in Release 9, for dual-layer transmission to each of two UEs with MU-MIMO.

8.3 RS-Aided Channel Modelling and Estimation

The channel estimation problem is related to the channel model, itself determined by the physical propagation characteristics, including the number of transmit and receive antennas, transmission bandwidth, carrier frequency, cell configuration and relative speed between eNodeB and UE receivers. In general,

- The carrier frequencies and system bandwidth mainly determine the scattering nature of the channel.

- The cell deployment governs its multipath, delay spread and spatial correlation characteristics.
- The relative speed sets the time-varying properties of the channel.

The propagation conditions characterize the channel correlation function in a three-dimensional space comprising frequency, time and spatial domains. In the general case, each MIMO multipath channel component can experience different but related spatial scattering conditions leading to a full three-dimensional correlation function across the three domains. Nevertheless, for the sake of simplicity, assuming that the multipath components of each spatial channel experience the same scattering conditions, the spatial correlation can be assumed to be independent from the other two domains and can be handled separately.

This framework might be suboptimal in general, but is nevertheless useful in mitigating the complexity of channel estimation as it reduces the general three-dimensional joint estimation problem into independent estimation problems.

For a comprehensive survey of MIMO channel estimation, the interested reader is referred to [13]. The following two subsections define the channel model and the corresponding correlation properties which are then used as the basis for an overview of channel estimation techniques.

8.3.1 Time-Frequency-Domain Correlation: The WSSUS Channel Model

The Wide-Sense Stationary Uncorrelated Scattering (WSSUS) channel model is commonly employed for the multipath channels experienced in mobile communications.

Neglecting the spatial dimension for the sake of simplicity, let $h(\tau; t)$ denote the time-varying complex baseband impulse response of a multipath channel realization at time instant t and delay τ.

Considering the channel as a random process in the time direction t, the channel is said to be delay Uncorrelated-Scattered (US) if

$$\mathbb{E}[h(\tau_a; t_1)^* h(\tau_b; t_2)] = \phi_h(\tau_a; t_1, t_2)\delta(\tau_b - \tau_a) \tag{8.2}$$

where $\mathbb{E}[\cdot]$ is the expectation operator. According to the US assumption, two CIR components a and b at relative delays τ_a and τ_b are uncorrelated if $\tau_a \neq \tau_b$.

The channel is Wide-Sense Stationary (WSS) uncorrelated if

$$\phi_h(\tau; t_1, t_2) = \phi_h(\tau; t_2 - t_1) \tag{8.3}$$

which means that the correlation of each delay component of the CIR is only a function of the *difference* in time between each realization.

Hence, the second-order statistics of this model are completely described by its delay cross-power density $\phi_h(\tau; \Delta t)$ or by its Fourier transform, the scattering function defined as

$$S_h(\tau; f) = \int \phi_h(\tau; \Delta t) e^{-j2\pi f \Delta t} \, \mathrm{d}\Delta t \tag{8.4}$$

with f being the Doppler frequency. Other related functions of interest include the *Power Delay Profile* (PDP)

$$\psi_h(\tau) = \phi_h(\tau; 0) = \int S_h(\tau; f) \, \mathrm{d}f$$

the *time-correlation function*

$$\overline{\phi}_h(\Delta t) = \int \phi_h(\tau; \Delta t)\, \mathrm{d}\tau$$

and the *Doppler power spectrum*

$$\overline{S}_h(f) = \int S_h(\tau; f)\, \mathrm{d}\tau.$$

A more general exposition of WSSUS models is given in [14]. Classical results were derived by Clarke [15] and Jakes [16] for the case of a mobile terminal communicating with a stationary base station in a two-dimensional propagation geometry.

These well-known results state that

$$\overline{S}_h(f) = \frac{1}{\sqrt{f_\mathrm{d}^2 - f^2}} \tag{8.5}$$

for $|f| \leq f_\mathrm{d}$ with f_d being the maximum Doppler shift and

$$\bar{\phi}_h(\Delta t) = J_0(2\pi f_\mathrm{d} \Delta t) \tag{8.6}$$

where $J_0(\cdot)$ is the zero$^{\text{th}}$-order Bessel function. Figure 8.8 shows the PSD of the classical Doppler spectrum described by Clarke and Jakes [15, 16]. The Clarke and Jakes derivations are based on the assumption that the physical scattering environment is chaotic and therefore the angle of arrival of the electromagnetic wave at the receiver is a uniformly distributed random variable in the angular domain. As a consequence, the time-correlation function is strictly real-valued, the Doppler spectrum is symmetric and interestingly there is a *delay-temporal separability* property in the general bi-dimensional scattering function $S_h(\tau, \Delta t)$. In other words,

$$S_h(\tau; f) = \psi_h(\tau)\overline{S}_h(f) \tag{8.7}$$

or equivalently

$$\phi_h(\tau; \Delta t) = \psi_h(\tau)\overline{\phi}_h(\Delta t) \tag{8.8}$$

If the time and frequency aspects can be assumed to be separable, the complexity of channel estimation can be significantly reduced, as the problem is reduced to two one-dimensional operations.

The continuous-time correlation properties of the channel $h(\tau, t)$ discussed above apply equivalently to the corresponding low-pass sampled discrete-time CIR, i.e. $h[l, k]$ [17], for every l^{th} delay sampled at the kT^{th} instant where T is the sampling period. Moreover, these properties are maintained if we assume $h[l, k]$ to be well-approximated by a Finite-Impulse Response (FIR) vector $\mathbf{h}[k] = [h[0, k], \cdots, h[L-1, k]]^{\mathrm{T}}$ with a maximum delay spread of L samples. For the sake of notational simplicity and without loss of generality, the index k will be dropped in the following sections.

8.3.2 Spatial-Domain Correlation: The Kronecker Model

While the frequency and time correlations addressed in the previous section are induced by the channel delays and Doppler spread, the spatial correlation arises from the spatial scattering conditions.

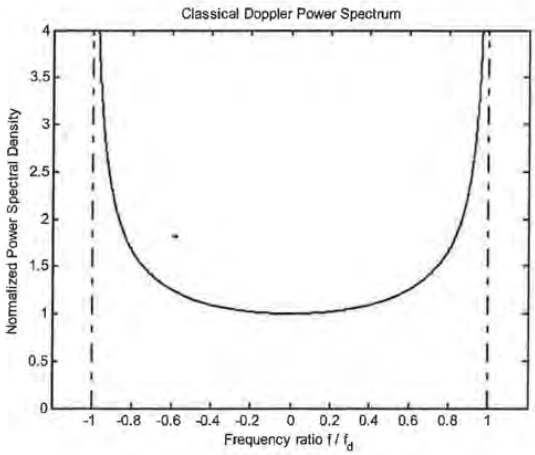


Figure 8.8: Normalized PSD for Clarke's model.

Among the possible spatial correlation models, for performance evaluation of LTE the Kronecker model is generally used [10]. Despite its simplicity, this correlation-based analytical model is widely used for the theoretical analysis of MIMO systems and yields experimentally verifiable results when limited to two or three antennas at each end of the radio link.

Let us assume the narrowband MIMO channel $N_{\mathrm{Rx}} \times N_{\mathrm{Tx}}$ matrix for a system with N_{Tx} transmitting antennas and N_{Rx} receiving antennas to be

$$\mathbf{H} = \begin{bmatrix} h_{0,0} & \cdots & h_{0,N_{\mathrm{Tx}}-1} \\ \vdots & & \vdots \\ h_{N_{\mathrm{Rx}}-1,N_{\mathrm{Tx}}-1} & \cdots & h_{N_{\mathrm{Rx}}-1,N_{\mathrm{Tx}}-1} \end{bmatrix} \tag{8.9}$$

The narrowband assumption particularly suits OFDM systems (see Section 5.2.1), where each MIMO channel component $h_{n,m}$ can be seen as the complex channel coefficient of each spatial link at a given subcarrier index.

The matrix $\mathbf{H}$ is rearranged into a vector by means of the operator $\mathrm{vec}(\mathbf{H}) = [\mathbf{h}_0^{\mathrm{T}}, \mathbf{h}_1^{\mathrm{T}}, \ldots, \mathbf{h}_{N_{\mathrm{Tx}}-1}^{\mathrm{T}}]^{\mathrm{T}}$ where $\mathbf{h}_i$ is the i^{th} column of $\mathbf{H}$ and $\{\cdot\}^{\mathrm{T}}$ is the transpose operation. Hence, the correlation matrix can be defined as

$$\mathbf{C}_{\mathrm{S}} = \mathbb{E}[\mathrm{vec}(\mathbf{H})\mathrm{vec}(\mathbf{H})^{\mathrm{H}}] \tag{8.10}$$

where $\{\cdot\}^{\mathrm{H}}$ is the Hermitian operation. The matrix in (8.10) is the full correlation matrix of the MIMO channel.

The Kronecker model assumes that the full correlation matrix results from separate spatial correlations at the transmitter and receiver, which is equivalent to writing the full correlation matrix as a Kronecker product ($\otimes$) of the transmitter and receiver correlation matrices:

$$\mathbf{C}_{\mathrm{S}} = \mathbf{C}_{\mathrm{Tx}} \otimes \mathbf{C}_{\mathrm{Rx}} \tag{8.11}$$

with $\mathbf{C}_{\mathrm{Tx}} = \mathbb{E}[\mathbf{H}^{\mathrm{H}}\mathbf{H}]$ and $\mathbf{C}_{\mathrm{Rx}} = \mathbb{E}[\mathbf{H}\mathbf{H}^{\mathrm{H}}]$.

Typical values assumed for these correlations according to the Kronecker model are discussed in Section 20.3.5.

8.4 Frequency-Domain Channel Estimation

In this section, we address the channel estimation problem over one OFDMA symbol (specifically a symbol containing reference symbols) to exploit the frequency-domain characteristics.

In the LTE context, as for any OFDM system with uniformly distributed reference symbols [18], the Channel Transfer Function (CTF) can be estimated using a maximum likelihood approach in the frequency domain at the REs containing the RSs by de-correlating the constant modulus RS. Using a matrix notation, the CTF estimate $\widehat{\mathbf{z}}_p$ on reference symbol p can be written as

$$\widehat{\mathbf{z}}_p = \mathbf{z}_p + \widetilde{\mathbf{z}}_p = \mathbf{F}_p\mathbf{h} + \widetilde{\mathbf{z}}_p \tag{8.12}$$

for $p \in (0, \ldots, P)$ where P is the number of available reference symbols and $\mathbf{h}$ is the $L \times 1$ CIR vector. $\mathbf{F}_p$ is the $P \times L$ matrix obtained by selecting the rows corresponding to the reference symbol positions and the first L columns of the $N \times N$ Discrete Fourier Transform (DFT) matrix where N is the FFT order. $\widetilde{\mathbf{z}}_p$ is a $P \times 1$ zero-mean complex circular white noise vector whose $P \times P$ covariance matrix is given by $\mathbf{C}_{\widetilde{\mathbf{z}}_p}$. The effective channel length $L \leq L_{\mathrm{CP}}$ is assumed to be known.

8.4.1 Channel Estimate Interpolation

8.4.1.1 Linear Interpolation Estimator

The natural approach to estimate the whole CTF is to interpolate its estimate between the reference symbol positions. In the general case, let $\mathbf{A}$ be a generic interpolation filter; then the interpolated CTF estimate at subcarrier index i can be written as

$$\widehat{\mathbf{z}}_i = \mathbf{A}\widehat{\mathbf{z}}_p \tag{8.13}$$

Substituting Equation (8.12) into (8.13), the error of the interpolated CTF estimate is

$$\widetilde{\mathbf{z}}_i = \mathbf{z} - \widehat{\mathbf{z}}_i = (\mathbf{F}_L - \mathbf{A}\mathbf{F}_p)\mathbf{h} - \mathbf{A}\widetilde{\mathbf{z}}_p \tag{8.14}$$

where $\mathbf{F}_L$ is the $N \times L$ matrix obtained by taking the first L columns of the DFT matrix and $\mathbf{z} = \mathbf{F}_L\mathbf{h}$. In Equation (8.14), it can be seen that the channel estimation error is constituted of a bias term (itself dependent on the channel) and an error term.

The error covariance matrix is

$$\mathbf{C}_{\widetilde{\mathbf{z}}_i} = (\mathbf{F}_L - \mathbf{A}\mathbf{F}_p)\mathbf{C}_{\mathbf{h}}(\mathbf{F}_L - \mathbf{A}\mathbf{F}_p)^{\mathrm{H}} + \sigma^2_{\widetilde{\mathbf{z}}_p}\mathbf{A}\mathbf{A}^{\mathrm{H}} \tag{8.15}$$

where $\mathbf{C}_h = \mathbb{E}[\mathbf{h}\mathbf{h}^{\mathrm{H}}]$ is the channel covariance matrix.

Recalling Equation (8.13), linear interpolation would be the intuitive choice. Such an estimator is deterministically biased, but unbiased from the Bayesian viewpoint regardless of the structure of $\mathbf{A}$.

8.4.1.2 IFFT Estimator

As a second straightforward approach, the CTF estimate over all subcarriers can be obtained by IFFT interpolation. In this case, the matrix $\mathbf{A}$ from Equation (8.13) becomes:

$$\mathbf{A}_{\text{IFFT}} = \frac{1}{P}\mathbf{F}_L\mathbf{F}_p^{\text{H}} \tag{8.16}$$

The error of the IFFT-interpolated CTF estimate and its covariance matrix can be obtained by substituting Equation (8.16) into Equations (8.14) and (8.15).

With the approximation of $\mathbf{I}_L \approx (1/P)\mathbf{F}_p^{\text{H}}\mathbf{F}_p$, where $\mathbf{I}_L$ is the $L \times L$ identity matrix, it can immediately be seen that the bias term in Equation (8.14) would disappear, providing for better performance.

Given the LTE system parameters and the patterns of REs used for RSs, in practice $(1/P)\mathbf{F}_p^{\text{H}}\mathbf{F}_p$ is far from being a multiple of an identity matrix. The approximation becomes an equality when $K = N$, $N/W > L$ and N/W is an integer,[9] i.e. the system would have to be dimensioned without guard-bands and the RS would have to be positioned with a spacing which is an exact factor of the FFT order N, namely a power of two.

In view of the other factors affecting the design of the RS RE patterns outlined above, such constraints are impractical.

8.4.2 General Approach to Linear Channel Estimation

Compared to the simplistic approaches presented in the previous section, more elaborate linear estimators derived from both deterministic and statistical viewpoints are proposed in [19–21]. Such approaches include Least Squares (LS), Regularized LS, Minimum Mean-Squared Error (MMSE) and Mismatched MMSE. These can all be expressed under the following general formulation:

$$\mathbf{A}_{\text{gen}} = \mathbf{B}(\mathbf{G}^{\text{H}}\mathbf{G} + \mathbf{R})^{-1}\mathbf{G}^{\text{H}} \tag{8.17}$$

where $\mathbf{B}$, $\mathbf{G}$ and $\mathbf{R}$ are matrices that vary according to each estimator as expressed in Table 8.1, where $\mathbf{0}_L$ is the all-zeros $L \times L$ matrix.

The LS estimator discussed in [19] is theoretically unbiased. However, as shown in [21], it is not possible to apply the LS estimator to LTE directly, because the expression $(\mathbf{F}_p\mathbf{F}_p^{\text{H}})^{-1}$ is ill-conditioned due to the unused portion of the spectrum corresponding to the unmodulated subcarriers.

To counter this problem, the classical robust regularized LS estimator can be used instead, so as to yield a better conditioning of the matrix to be inverted. A regularization matrix $\alpha\mathbf{I}_L$ is introduced [22] where α is a constant (computed off-line) chosen to optimize the performance of the estimator in a given Signal-to-Noise Ratio (SNR) working range.

The MMSE estimator belongs to the class of statistical estimators. Unlike deterministic LS and its derivations, statistical estimators need knowledge of the second-order statistics (PDP and noise variance) of the channel in order to perform the estimation process, normally with much better performance compared to deterministic estimators. However, second-order statistics vary as the propagation conditions change and therefore need appropriate re-estimation regularly. For this reason statistical estimators are, in general, more complex due

[9] W is the spacing (in terms of number of subcarriers) between reference symbols.

Table 8.1: Linear estimators.

	Components of interpolation filter (see Equation (8.17))		
Interpolation method	**B**	**G**	**R**
Simple interpolator	$\mathbf{A}$	$\mathbf{I}_P$	$\mathbf{0}_L$
FFT	$\frac{1}{P}\mathbf{F}_L\mathbf{F}_P^{\mathrm{H}}$	$\mathbf{I}_P$	$\mathbf{0}_L$
LS	$\mathbf{F}_L$	$\mathbf{F}_p$	$\mathbf{0}_L$
Regularized LS	$\mathbf{F}_L$	$\mathbf{F}_p$	$\alpha\mathbf{I}_L$
MMSE	$\mathbf{F}_L$	$\mathbf{F}_p$	$\sigma^2_{\tilde{\mathbf{z}}_p}\mathbf{C}_{\mathbf{h}}{}^{-1}$
Mismatched MMSE	$\mathbf{F}_L$	$\mathbf{F}_p$	$\sigma^2_{\tilde{\mathbf{z}}_p}/\sigma^2_{\mathbf{h}}\cdot\mathbf{I}_L$

to the additional burden of estimating the second-order statistics and computing the filter coefficients.

A mismatched MMSE estimator avoids the estimation of the second-order channel statistics and the consequent on-line inversion of an $L \times L$ matrix (as required in the straightforward application of MMSE) by assuming that the channel PDP is uniform[10] [20]. This estimator is, in practice, equivalent to the regularized LS estimator where the only difference lies in the fact that $\alpha = \sigma^2_{\tilde{\mathbf{z}}_p}/\sigma^2_{\mathbf{h}}$ is estimated and therefore adapted. The mismatched MMSE formulation offers the advantage that the filter coefficients can be computed to be real numbers because the uniform PDP is symmetric. Moreover, since the length of the CIR is small compared to the FFT size, the matrix A_{gen} can be considered to be 'low-density', so storing only the significant coefficients can reduce the complexity.

However, LTE does not use an exactly uniform pattern of reference symbols; for the cell-specific RSs this is the case around the d.c. subcarrier which is not transmitted. This implies that a larger number of coefficients needs to be stored.

8.4.3 Performance Comparison

For the sake of comparison between the performance of the different classes of estimator, we introduce the Truncated Normalized Mean Squared Error (TNMSE).

For each estimator, the TNMSE is computed from its covariance matrix $\mathbf{C}_{\tilde{\mathbf{z}}}$ and the true CTF $\mathbf{z}$ as follows:

$$\text{TNMSE}_{\tilde{\mathbf{z}}} = \frac{\text{Ttr}(\mathbf{C}_{\tilde{\mathbf{z}}})}{\text{Ttr}(\mathbf{F}_L\mathbf{C}_{\mathbf{h}}\mathbf{F}_L^{\mathrm{H}})} \tag{8.18}$$

where $\text{Ttr}\{\cdot\}$ denotes the truncated trace operator, where the truncation is such that only the K used subcarriers are included.

Figure 8.9 shows the performance of an LTE FDD downlink with 10 MHz transmission bandwidth ($N = 1024$) and Spatial Channel Model-A (SCM-A – see Section 20.3.4).

It can be seen that the IFFT and linear interpolation methods yield the lowest performance. The regularized LS and the mismatched MMSE perform equally and the curve of the latter is therefore omitted. As expected, the optimal MMSE estimator outperforms any other estimator.

[10] This results in the second-order statistics of the channel having the structure of an identity matrix.

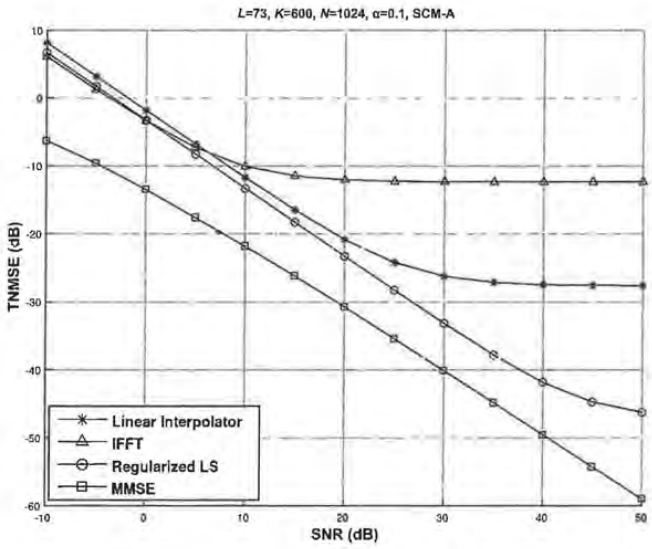


Figure 8.9: Frequency-domain channel estimation performance.

The TNMSE computed over all subcarriers actually hides the behaviour of each estimator in relation to a well-known problem of frequency-domain channel estimation techniques: the band-edge effect. This can be represented by the Gibbs [23] phenomenon in a finite-length Fourier series approximation; following this approach, Figure 8.10 shows that MMSE-based channel estimation suffers the least band-edge degradation, while all the other methods presented are highly adversely affected.

8.5 Time-Domain Channel Estimation

The main benefit of time-domain channel estimation is the possibility to enhance the channel estimate of one OFDM symbol containing reference symbols by exploiting its time correlation with the channel at previous OFDM symbols containing reference symbols. This requires sufficient memory for buffering soft values of data over several OFDM symbols while the channel estimation is carried out.

However, the correlation between consecutive symbols decreases as the UE speed increases, as expressed by Equation (8.6), and this sets a limit on the possibilities for time-domain filtering in high-mobility conditions.

Time-domain filtering is applied to the CIR estimate, rather than to the CTF estimate in the frequency domain. The use of a number of parallel scalar filters equal to the channel length L does not imply a loss of optimality, because of the WSSUS assumption.

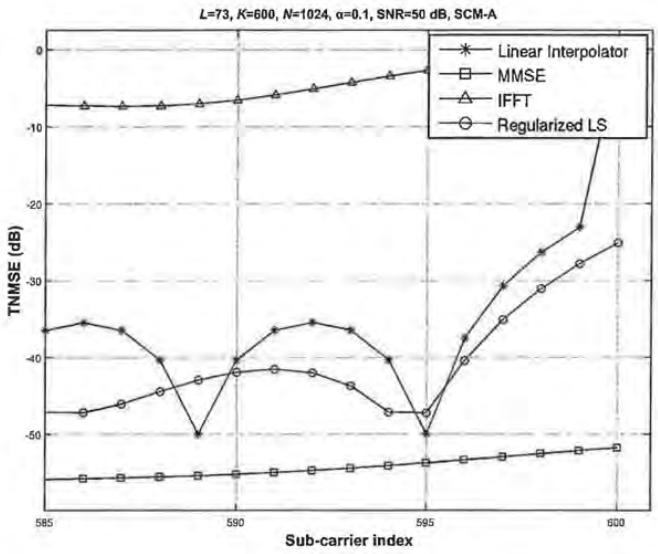


Figure 8.10: Band-edge behaviour of frequency-domain channel estimation.

8.5.1 Finite and Infinite Length MMSE

The statistical time-domain filter which is optimal in terms of Mean Squared Error (MSE) can be approximated in the form of a finite impulse response filter [24]. The channel at the l^{th} tap position and at time instant n is estimated as

$$\hat{\hat{h}}_{l,n} = \mathbf{w}_l^H \hat{\mathbf{h}}_{l,n}^M \tag{8.19}$$

where $\hat{\hat{h}}_l$ is the smoothed CIR l^{th} tap estimate which exploits the vector $\hat{\mathbf{h}}_{l,n}^M = [\hat{h}_{l,n}, \ldots, \hat{h}_{l,n-M+1}]^{\text{T}}$ of length M of the channel tap h_l across estimates at M time instants.[11] This is obtained by inverse Fourier transformation of, for example, any frequency-domain technique illustrated in Section 8.4 or even a raw estimate obtained by RS decorrelation.

The $M \times 1$ vector of Finite Impulse Response (FIR) filter coefficients $\mathbf{w}_l$ is given by

$$\mathbf{w}_l = (\mathbf{R}_h + \sigma_n^2 \mathbf{I})^{-1} \mathbf{r}_h \tag{8.20}$$

where $\mathbf{R}_h = \mathbb{E}[\mathbf{h}_l^M (\mathbf{h}_l^M)^{\text{H}}]$ is the l^{th} channel tap $M \times M$ correlation matrix, σ_n^2 the additive noise variance and $\mathbf{r}_h = \mathbb{E}[\mathbf{h}_l^M h_{l,n}^*]$ (the $M \times 1$ correlation vector between the l^{th} tap of the current channel realization and M previous realizations including the current one).

[11] $h_{l,k}$ is the l^{th} component of the channel vector $\mathbf{h}_k$ at time instant k. $\hat{h}_{l,k}$ is its estimate. Note that for the frequency-domain treatment, the time index was dropped.

In practical cases, the FIR filter length M is dimensioned according to a performance-complexity trade-off as a function of UE speed.

By setting M infinite, the upper bound on performance is obtained.

The MSE performance of the finite-length estimator of a channel of length L can be analytically computed as

$$\epsilon^{(M)} = 1 - \frac{1}{\sigma_h^2} \sum_{l=0}^{L-1} \mathbf{w}_l^{\mathrm{H}} \mathbf{r}_h \tag{8.21}$$

The upper bound given by an infinite-length estimator is therefore given by

$$\epsilon^{(\infty)} = \lim_{M \to \infty} \epsilon^{(M)} \tag{8.22}$$

From Equation (8.20) it can be observed that, unlike frequency-domain MMSE filtering, the size of the matrix to be inverted for a finite-length time-domain MMSE estimator is independent of the channel length L but dependent on the chosen FIR order M. Similarly to the frequency-domain counterpart, the time-domain MMSE estimator requires knowledge of the PDP, the UE speed and the noise variance.

Figure 8.11 shows the performance of time-domain MMSE channel estimation as a function of filter length M (Equation (8.21)) for a single-tap channel with a classical Doppler spectrum for low UE speed. The performance bounds derived for an infinite-length filter in each case are also indicated.

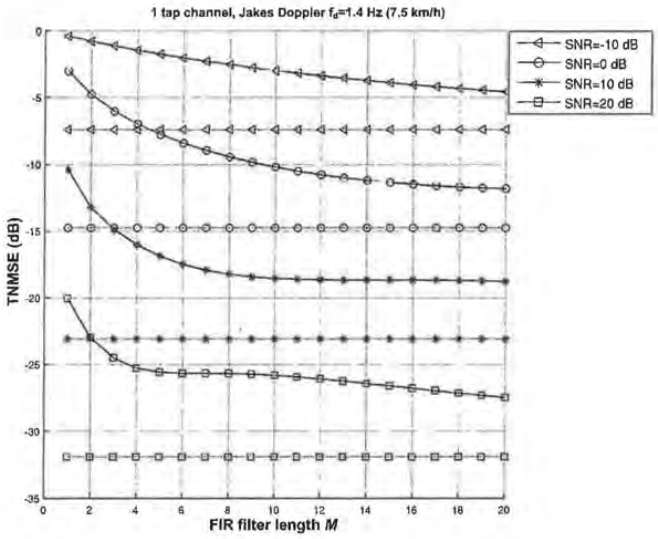


Figure 8.11: Time-domain channel estimation performance.

8.5.2 Normalized Least-Mean-Square

As an alternative to time-domain MMSE channel estimation, an adaptive estimation approach can be considered which does not require knowledge of second-order statistics of both channel and noise. A feasible solution is the Normalized Least-Mean-Square (NLMS) estimator.

It can be expressed exactly as in Equation (8.19) but with the $M \times 1$ vector of filter coefficients $\mathbf{w}$ updated according to

$$\mathbf{w}_{l,n} = \mathbf{w}_{l,n-1} + \mathbf{k}_{l,n-1} e_{l,n} \tag{8.23}$$

where M here denotes the NLMS filter order. The $M \times 1$ updated gain vector is computed according to the well-known NLMS adaptation:

$$\mathbf{k}_{l,n} = \frac{\mu}{\|\hat{\mathbf{h}}_{l,n}\|^2} \hat{\mathbf{h}}_{l,n}^{M} \tag{8.24}$$

where μ is an appropriately chosen step adaptation, $\hat{\mathbf{h}}_{l,n}^{M}$ is defined as for Equation (8.19) and

$$e_{l,n} = \hat{h}_{l,n} - \hat{h}_{l,n-1} \tag{8.25}$$

It can be observed that the time-domain NLMS estimator requires much lower complexity compared to time-domain MMSE because no matrix inversion or a priori statistical knowledge is required.

Other adaptive approaches could also be considered, such as Recursive Least Squares (RLS) and Kalman-based filtering. Although more complex than NLMS, the Kalman filter is a valuable candidate and is reviewed in detail in [25].

8.6 Spatial-Domain Channel Estimation

It is assumed that an LTE UE has multiple receiving antennas. Consequently, whenever the channel is correlated in the spatial domain, the correlation can be exploited to provide a further means for enhancing the channel estimate.

If it is desired to exploit spatial correlation, a natural approach is again offered by spatial domain MMSE filtering [26].

We consider here the case of a MIMO OFDM communication system with N subcarriers, N_{Tx} transmitting antennas and N_{Rx} receiving antennas. The $N_{\mathrm{Rx}} \times 1$ received signal vector at subcarrier k containing the RS sequence can be written as

$$\mathbf{r}(k) = \frac{1}{\sqrt{N_{\mathrm{Tx}}}} \mathbf{W}(k)\mathbf{s}(k) + \mathbf{n}(k) \tag{8.26}$$

where $\mathbf{W}(k)$ is the $N_{\mathrm{Rx}} \times N_{\mathrm{Tx}}$ channel frequency response matrix at RS subcarrier k, $\mathbf{s}(k)$ is the $N_{\mathrm{Tx}} \times 1$ known zero-mean and unit-variance transmitted RS sequence at subcarrier k, and $\mathbf{n}(k)$ is the $N_{\mathrm{Rx}} \times 1$ complex additive white Gaussian noise vector with zero mean and variance σ_n^2.

The CTF at subcarrier k, $\mathbf{W}(k)$, is obtained by DFT from the CIR matrix at the l^{th} tap $\mathbf{H}_l$ as

$$\mathbf{W}(k) = \sum_{l=0}^{L-1} \mathbf{H}_l \exp\left[-j2\pi \frac{lk}{N}\right] \tag{8.27}$$

The vector $\mathbf{h}$ is obtained by rearranging the elements of all $\mathbf{H}_l$ channel tap matrices as follows:

$$\mathbf{h} = [\text{vec}(\mathbf{H}_0)^{\text{T}}, \text{vec}(\mathbf{H}_1)^{\text{T}}, \ldots, \text{vec}(\mathbf{H}_{L-1})^{\text{T}}]^{\text{T}} \tag{8.28}$$

The correlation matrix of $\mathbf{h}$ is given by

$$\mathbf{C}_h = \mathbb{E}[\mathbf{h}\mathbf{h}^{\text{H}}] \tag{8.29}$$

Using Equations (8.27) and (8.28), Equation (8.26) can now be rewritten as

$$\mathbf{r}(k) = \frac{1}{\sqrt{N_{\text{Tx}}}}\mathbf{G}(k)\mathbf{h} + \mathbf{n}(k) \tag{8.30}$$

where $\mathbf{G}(k) = [\mathbf{D}_0(k), \mathbf{D}_1(k), \ldots, \mathbf{D}_{L-1}(k)]$ and $\mathbf{D}_l(k) = \exp[-j2\pi(lk/N)]\mathbf{s}^{\text{T}}(k) \otimes \mathbf{I}_{N_{\text{Rx}}}$.

Rearranging the received signal matrix $\mathbf{G}$ and the noise matrix at all N subcarriers into a vector as follows,

$$\mathbf{r} = [\mathbf{r}^{\text{T}}(0), \mathbf{r}^{\text{T}}(1), \ldots, \mathbf{r}^{\text{T}}(N-1)]^{\text{T}} \tag{8.31}$$

Equation (8.30) can be rewritten more compactly as

$$\mathbf{r} = \frac{1}{\sqrt{N_{\text{Tx}}}}\mathbf{G}\mathbf{h} + \mathbf{n} \tag{8.32}$$

Hence, the spatial domain MMSE estimation of the rearranged channel impulse vector $\mathbf{h}$ can be simply obtained by

$$\hat{\mathbf{h}} = \mathbf{Q}\mathbf{r} \tag{8.33}$$

with

$$\mathbf{Q} = \frac{1}{\sqrt{N_{\text{Tx}}}}\mathbf{C}_h\mathbf{G}^{\text{H}}\left(\frac{1}{N_{\text{Tx}}}\mathbf{G}\mathbf{C}_h\mathbf{G}^{\text{H}} + \sigma_n^2\mathbf{I}_{N\cdot N_{\text{Rx}}}\right)^{-1} \tag{8.34}$$

Therefore the NMSE can be computed as

$$\text{NMSE}_{\text{SD-MMSE}} = \frac{\text{tr}(\mathbf{C}_{\text{SD-MMSE}})}{\text{tr}(\mathbf{C}_h)} \tag{8.35}$$

where $\mathbf{C}_{\text{SD-MMSE}}$ is the error covariance matrix.

Figure 8.12 shows the spatial domain MMSE performance given by Equation (8.35) compared to the performance obtained by the ML channel estimation on the subcarriers which carry RS.

8.7 Advanced Techniques

The LTE specifications do not mandate any specific channel estimation technique, and there is therefore complete freedom in implementation provided that the performance requirements are met and the complexity is affordable.

Particular aspects, such as channel estimation based on the UE-specific RSs, band-edge effect reduction or bursty reception, might require further improvements which go beyond the techniques described here.

Blind and semi-blind techniques are promising for some such aspects, as they try to exploit not only the a priori knowledge of the RSs but also the unknown data structure. A comprehensive analysis is available in [27] and references therein.

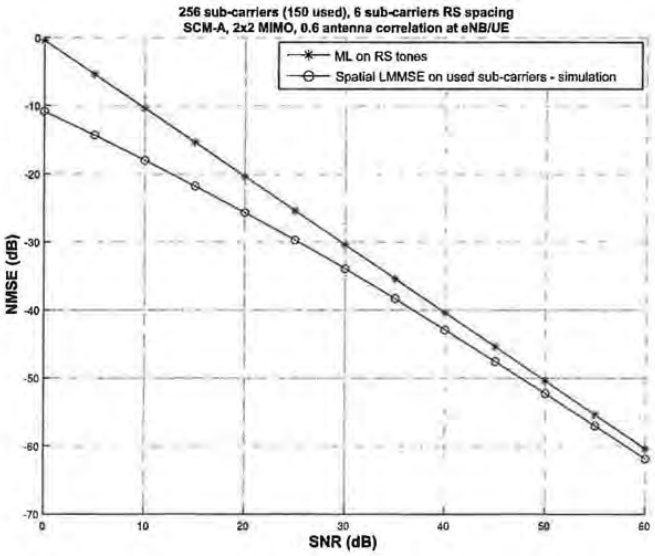


Figure 8.12: Spatial-domain channel estimation performance: CIR NMSE versus SNR.

References[12]

[1] A. Vosoughi and A. Scaglione, 'The Best Training Depends on the Receiver Architecture' in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004)*, (Montreal, Canada), May 2004.

[2] A. Vosoughi and A. Scaglione, 'Everything You Always Wanted to Know about Training: Guidelines Derived using the Affine Precoding Framework and the CRB'. *IEEE Trans. on Signal Processing*, Vol. 54, pp. 940–954, March 2006.

[3] 3GPP Technical Specification 25.211, 'Physical Channels and Mapping of Transport Channels onto Physical Channels (FDD)', www.3gpp.org.

[4] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. New York: John Wiley & Sons, Ltd, 1949.

[5] D. T. M. Slock, 'Signal Processing Challenges for Wireless Communications' in *Proc. International Symposium on Control, Communications and Signal Processing (ISCCSP' 2004)*, (Hammamet, Tunisia), March 2004.

[6] 3GPP Technical Specification 36.211, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation', www.3gpp.org.

[7] R. Negi and J. Cioffi, 'Pilot Tone Selection for Channel Estimation in a Mobile OFDM System'. *IEEE Trans. on Consumer Electronics*, Vol. 44, pp. 1122–1128, August 1998.

[12] All web sites confirmed 1st March 2011.

[8] I. Barhumi, G. Leus and M. Moonen, 'Optimal Training Design for MIMO OFDM Systems in Mobile Wireless Channels'. *IEEE Trans. on Signal Processing*, Vol. 51, pp. 1615–1624, June 2003.

[9] 3GPP Technical Report 25.913, 'Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN) (Release 7)', www.3gpp.org.

[10] 3GPP Technical Specification 36.101, 'User Equipment (UE) Radio Transmission and Reception', www.3gpp.org.

[11] Motorola, Nortel, Broadcomm, Nokia, NSN, NTT DoCoMo, NEC, Mitsubishi, Alcatel-Lucent, CATT, Huawei, Sharp, Texas Instrument, ZTE, Panasonic, Philips and Toshiba, 'R1-081108: Way Forward on Dedicated Reference Signal Design for LTE Downlink with Normal CP', www.3gpp.org, 3GPP TSG RAN WG1, meeting 52, Sorrento, Italy, February 2008.

[12] NTT DoCoMo and NEC, 'R1-094337: Remaining Issues for Rel. 9 Downlink DM-RS Design', www.3gpp.org, 3GPP TSG RAN WG1, meeting 58bis, Miyazaki, Japan, Octboer 2009.

[13] P. Almers, E. Bonek, A. Burr, N. Czink, M. Debbah, V. Degli-Esposti, H. Hofstetter, P Kyösty, D. Laurenson, G. Matz, F. Molisch, C. Oestges, and H. Özcelik, 'Survey of Channel and Radio Propagation Models for Wireless MIMO Systems'. *EURASIP Journal on Wireless Communications and Networking*, pp. 957–1000, July 2007.

[14] J. G. Proakis, *Digital Communications*. New York: McGraw Hill, 1995.

[15] R. H. Clarke, 'A Statistical Theory of Mobile-radio Reception'. *Bell Syst. Tech. J.*, pp. 957–1000, July 1968.

[16] W. C. Jakes, *Microwave Mobile Communications*. New York: John Wiley & Sons, Ltd/Inc., 1974.

[17] P. Hoeher,'A statistical discrete-time model for the WSSUS multipath channel' in *IEEE Transactions on Vehicular Technology*, Vol 41, No. 4, pp. 461–468, 1992.

[18] J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Börjesson, 'On Channel Estimation in OFDM Systems' in *Proc. VTC'1995, Vehicular Technology Conference (VTC 1995)*, Chicago, USA, July 1995.

[19] M. Morelli and U. Mengali, 'A Comparison of Pilot-aided Channel Estimation Methods for OFDM Systems'. *IEEE Trans. on Signal Processing*, Vol. 49, pp. 3065–3073, December 2001.

[20] P. Hoher, S. Kaiser and P. Robertson, 'Pilot-Symbol-aided Channel Estimation in Time and Frequency' in *Proc. Communication Theory Mini-Conf. (CTMC) within IEEE Global Telecommun. Conf. (Globecom 1997)* (Phoenix, USA), July 1997.

[21] A. Ancora, C. Bona and D. T. M. Slock, 'Down-Sampled Impulse Response Least-Squares Channel Estimation for LTE OFDMA' in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2007)*, Honolulu, Hawaii, April 2007.

[22] P. C. Hansen, 'Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion', Society for Industrial Mathematics, 1998.

[23] J. W. Gibbs, 'Fourier Series'. *Nature*, Vol. 59, p. 200, 1898.

[24] D. Schafhuber, G. Matz and F. Hlawatsch, 'Adaptive Wiener Filters for Time-varying Channel Estimation in Wireless OFDM Systems' in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2003)*, Hong Kong, April 2003.

[25] J. Cai, X. Shen and J. W. Mark, 'Robust Channel Estimation for OFDM Wireless Communication Systems. An H∞ Approach'. *IEEE Trans. Wireless Communications*, Vol. 3, November 2004.

[26] H. Zhang, Y. Li, A. Reid and J. Terry, 'Channel Estimation for MIMO OFDM in Correlated Fading Channels' in *Proc. 2005 IEEE International Conference on Communications (ICC 2005)*, South Korea, May 2005.

[27] H. Bölcskei, D. Gesbert and C. Papadias, *Space-time Wireless Systems: From Array Processing to MIMO Communications*. Cambridge: Cambridge University Press, 2006.

ERM EVOLUTION

io Access (E-UTRA);

).org, 3GPP TSG RAN

10

Link Adaptation and Channel Coding

Brian Classson, Ajit Nimbalker, Stefania Sesia and Issam Toufik

10.1 Introduction

The principle of link adaptation is fundamental to the design of a radio interface which is efficient for packet-switched data traffic. Unlike the early versions of UMTS (Universal Mobile Telecommunication System), which used fast closed-loop power control to support circuit-switched services with a roughly constant data rate, link adaptation in HSPA (High Speed Packet Access) and LTE adjusts the transmitted information data rate (modulation scheme and channel coding rate) dynamically to match the prevailing radio channel capacity for each user. Link adaptation is therefore very closely related to the design of the channel coding scheme used for forward error correction.

For the downlink data transmissions in LTE, the eNodeB typically selects the modulation scheme and code rate depending on a prediction of the downlink channel conditions. An important input to this selection process is the Channel Quality Indicator (CQI) feedback transmitted by the User Equipment (UE) in the uplink. CQI feedback is an indication of the data rate which can be supported by the channel, taking into account the Signal-to-Interference-plus-Noise Ratio (SINR) and the characteristics of the UE's receiver. Section 10.2 explains the principles of link adaptation as applied in LTE; it also shows how the eNodeB can select different CQI feedback modes to trade off the improved downlink link adaptation enabled by CQI against the uplink overhead caused by the CQI itself.

The LTE specifications are designed to provide the signalling necessary for interoperability between the eNodeB and the UEs so that the eNodeB can optimize the link adaptation,

LTE – The UMTS Long Term Evolution: From Theory to Practice, Second Edition.
Stefania Sesia, Issam Toufik and Matthew Baker.
© 2011 John Wiley & Sons, Ltd. Published 2011 by John Wiley & Sons, Ltd.

but the exact methods used by the eNodeB to exploit the information that is available are left to the manufacturer's choice of implementation.

In general, in response to the CQI feedback the eNodeB can select between QPSK, 16QAM and 64QAM[1] schemes and a wide range of code rates. As discussed further in Section 10.2.1, the optimal switching points between the different combinations of modulation order and code rate depend on a number of factors, including the required quality of service and cell throughput.

The channel coding scheme for forward error correction, on which the code rate adaptation is based, was the subject of extensive study during the standardization of LTE. The chapter therefore continues with a review of the key theoretical aspects of the types of channel coding studied for LTE: convolutional codes, turbo codes with iterative decoding, and a brief introduction of Low-Density Parity Check (LDPC) codes. The theory of channel coding has seen intense activity in recent decades, especially since the discovery of turbo codes offering near-Shannon limit performance, and the development of iterative processing techniques in general. 3GPP was an early adopter of these advanced channel coding techniques, with the turbo code being standardized in the first version of the UMTS as early as 1999. Later releases of UMTS (in HSPA) added more advanced channel coding features with the introduction of link adaptation, including Hybrid Automatic Repeat reQuest (HARQ), a combination of ARQ and channel coding which provides more robustness against fading; these schemes include incremental redundancy, whereby the code rate is progressively reduced by transmitting additional parity information with each retransmission. However, the underlying turbo code from the first version of UMTS remained untouched. Meanwhile, the academic and research communities were generating new insights into code design, iterative decoding and the implementation of decoders. Section 10.3.2 explains how these developments impacted the design of the channel coding for LTE, and in particular the decision to enhance the turbo code from UMTS by means of a new contention-free interleaver, rather than to adopt a new LDPC code.

For the LTE uplink transmissions, the link adaptation process is similar to that for the downlink, with the selection of modulation and coding schemes also being under the control of the eNodeB. An identical channel coding structure is used for the uplink, while the modulation scheme may be selected between QPSK, 16QAM and, for the highest category of UE, also 64QAM. The main difference from the downlink is that instead of basing the link adaptation on CQI feedback, the eNodeB can directly make its own estimate of the supportable uplink data rate by channel sounding, for example using the Sounding Reference Signals (SRSs) which are described separately in Section 15.6.

A final important aspect of link adaptation is its use in conjunction with multi-user scheduling in time and frequency, which enables the radio transmission resources to be shared efficiently between users as the channel capacity to individual users varies. The CQI can therefore be used not only to adapt the modulation and coding rate to the channel conditions, but also for the optimization of time/frequency selective scheduling and for inter-cell interference management as discussed in Chapter 12.

[1] Quadrature Phase Shift Keying (QPSK) and Quadrature Amplitude Modulation (QAM).

11

Multiple Antenna Techniques

Thomas Sälzer, David Gesbert, Cornelius van Rensburg, Filippo Tosato, Florian Kaltenberger and Tetsushi Abe

11.1 Fundamentals of Multiple Antenna Theory

11.1.1 Overview

The value of multiple antenna systems as a means to improve communications was recognized in the very early ages of wireless transmission. However, most of the scientific progress in understanding their fundamental capabilities has occurred only in the last 20 years, driven by efforts in signal and information theory, with a key milestone being achieved with the invention of so-called Multiple-Input Multiple-Output (MIMO) systems in the mid-1990s.

Although early applications of beamforming concepts can be traced back as far as 60 years in military applications, serious attention has been paid to the utilization of multiple antenna techniques in mass-market commercial wireless networks only since around 2000. Today, the key role which MIMO technology plays in the latest wireless communication standards for Personal, Wide and Metropolitan Area Networks (PANs, WANs and MANs) testifies to its importance. Aided by rapid progress in the areas of computation and circuit integration, this trend culminated in the adoption of MIMO for the first time in a cellular mobile network standard in the Release 7 version of HSDPA (High Speed Downlink Packet Access); soon after, the development of LTE broke new ground in being the first global mobile cellular system to be designed with MIMO as a key component from the start.

This chapter first provides the reader with the theoretical background necessary for a good understanding of the role and advantages promised by multiple antenna techniques in

LTE – The UMTS Long Term Evolution: From Theory to Practice, Second Edition.
Stefania Sesia, Issam Toufik and Matthew Baker.
© 2011 John Wiley & Sons, Ltd. Published 2011 by John Wiley & Sons, Ltd.

wireless communications in general.[1] The chapter focuses on the intuition behind the main technical results and show how key progress in information theory yields practical lessons in algorithm and system design for cellular communications. As can be expected, there is still a gap between the theoretical predictions and the performance achieved by schemes that must meet the complexity constraints imposed by commercial considerations.

We distinguish between single-user MIMO and multi-user MIMO, although a common set of concepts captures the essential MIMO benefits in both cases. Single-user MIMO techniques dominated the algorithms selected for the first version of LTE, with multi-user MIMO becoming more established in Releases 9 and 10.

The second part of the chapter describes the actual methods adopted for LTE, paying particular attention to the combinations of theoretical principles and system design constraints that led to these choices.

While traditional wireless communications (Single-Input Single-Output (SISO)) exploit time- or frequency-domain pre-processing and decoding of the transmitted and received data respectively, the use of additional antenna elements at either the base station (eNodeB) or User Equipment (UE) side (on the downlink or uplink) opens an extra spatial dimension to signal precoding and detection. Space-time processing methods exploit this dimension with the aim of improving the link's performance in terms of one or more possible metrics, such as the error rate, communication data rate, coverage area and spectral efficiency (expressed in bps/Hz/cell).

Depending on the availability of multiple antennas at the transmitter and/or the receiver, such techniques are classified as Single-Input Multiple-Output (SIMO), Multiple-Input Single-Output (MISO) or MIMO. Thus in the scenario of a multi-antenna-enabled base station communicating with a single antenna UE, the uplink and downlink are referred to as SIMO and MISO respectively. When a multi-antenna terminal is involved, a full MIMO link may be obtained, although the term MIMO is sometimes also used in its widest sense, thus including SIMO and MISO as special cases. While a point-to-point multiple-antenna link between a base station and one UE is referred to as Single-User MIMO (SU-MIMO), Multi-User MIMO (MU-MIMO) features several UEs communicating simultaneously with a common base station using the same frequency- and time-domain resources.[2] By extension, considering a multicell context, neighbouring base stations sharing their antennas in virtual MIMO fashion [4] to communicate with the same set of UEs in different cells comes under the term Coordinated MultiPoint (CoMP) transmission/reception. This latter scenario is not supported in the first versions of LTE but it is being studied for possible inclusion in later releases of LTE-Advanced as described in Section 29.5. The overall evolution of MIMO concepts, from the simplest diversity setup to the advanced joint-processing CoMP technique, is illustrated in Figure 11.1.

Despite their variety and sometimes perceived complexity, single-user and multi-user MIMO techniques tend to revolve around just a few fundamental principles, which aim at leveraging some key properties of multi-antenna radio propagation channels. As introduced

[1] For more exhaustive tutorial information on MIMO systems, the reader is referred, for example, to [1–3].

[2] Note that, in LTE, a single eNodeB may in practice control multiple cells; in such a case, we consider each cell as an independent base station for the purpose of explaining the MIMO techniques; the simultaneous transmissions in the different cells address different UEs and are typically achieved using different fixed directional physical antennas; they are therefore not classified as multi-user MIMO.

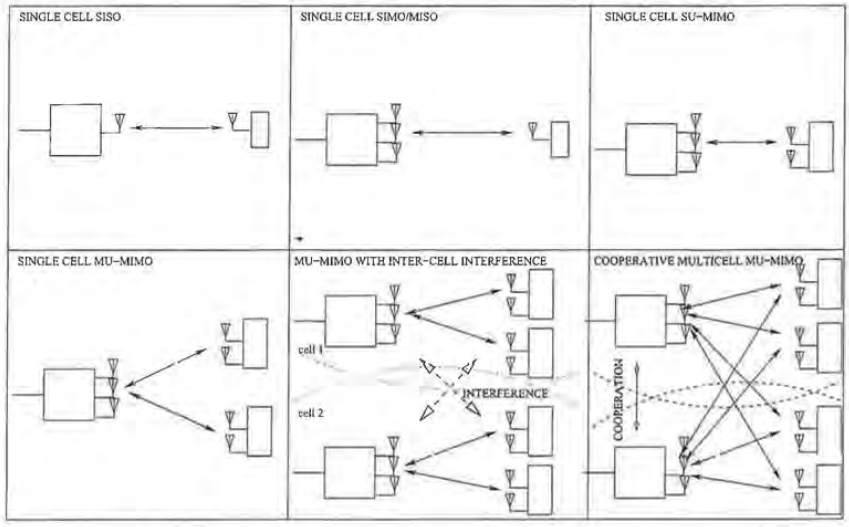


Figure 11.1: The evolution of MIMO technology, from traditional single antenna communication, to multi-user MIMO scenarios, to the possible multicell MIMO networks of the future.

in Section 1.3, there are basically three advantages associated with such channels (over their SISO counterparts):

- Diversity gain.
- Array gain.
- Spatial multiplexing gain.

Diversity gain corresponds to the mitigation of the effect of multipath fading, by means of transmitting or receiving over multiple antennas at which the fading is sufficiently decorrelated. It is typically expressed in terms of an *order*, referring either to the number of effective independent diversity branches or to the slope of the Bit Error Rate (BER) curve as a function of the Signal-to-Noise Ratio (SNR) (or possibly in terms of an SNR gain in the system's link budget).

While diversity gain is fundamentally related to improvement of the *statistics* of instantaneous SNR in a fading channel, array gain and multiplexing gain are of a different nature, rather being related to the geometry and the theory of vector spaces. Array gain corresponds to a spatial version of the well-known matched-filter gain in time-domain receivers, while multiplexing gain refers to the ability to send multiple data streams in parallel and to separate them on the basis of their spatial signature. The latter is much akin to the multiplexing of users separated by orthogonal spreading codes, timeslots or frequency assignments, with the great advantage that, unlike Code, Time or Frequency Division Multiple Access (CDMA, TDMA or FDMA), MIMO multiplexing does not come at the cost of bandwidth expansion; it does,

however, suffer the expense of added antennas and signal processing complexity and its gains strongly depend on the spatial characteristics of the propagation channel (see Chapter 20).

These aspects are analysed further in the following sections, considering first theoretically optimal transmission schemes and then popular MIMO approaches. First a base station to single-user communication model is considered (to cover single-user MIMO); then the techniques are generalized to multi-user MIMO.

The schemes adopted in LTE Releases 8 and 9, specifically for the downlink, are addressed subsequently. We focus on the Frequency Division Duplex (FDD) case. Discussion of some aspects of MIMO which are specific to Time Division Duplex (TDD) operation can be found in Section 23.5.

11.1.2 MIMO Signal Model

Let $\mathbf{Y}$ be a matrix of size $N \times T$ denoting the set of (possibly precoded) signals being transmitted from N distinct antennas over T symbol durations (or, in the case of some frequency-domain systems, T subcarriers), where T is a parameter of the MIMO algorithm (defined below). Thus the n^{th} row of $\mathbf{Y}$ corresponds to the signal emitted from the n^{th} transmit antenna. Let $\mathbf{H}$ denote the $M \times N$ channel matrix modelling the propagation effects from each of the N transmit antennas to any one of M receive antennas, over an arbitrary subcarrier whose index is omitted here for simplicity. We assume $\mathbf{H}$ to be invariant over T symbol durations. The matrix channel is represented by way of example in Figure 11.2. Then the $M \times T$ signal $\mathbf{R}$ received over T symbol durations over this subcarrier can be conveniently written as

$$\mathbf{R} = \mathbf{HY} + \mathbf{N} \tag{11.1}$$

where $\mathbf{N}$ is the additive noise matrix of dimension $M \times T$ over all M receiving antennas. We will use $\mathbf{h}_i$ to denote the i^{th} column of $\mathbf{H}$, which will be referred to as the *receive spatial signature* of (i.e. corresponding to) the i^{th} transmitting antenna. Likewise, the j^{th} row of $\mathbf{H}$ can be termed the *transmit spatial signature* of the j^{th} receiving antenna.

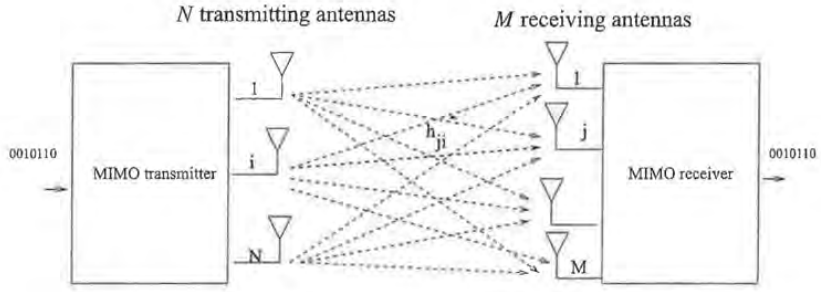


Figure 11.2: Simplified transmission model for a MIMO system with N transmit antennas and M receive antennas, giving rise to an $M \times N$ channel matrix, with MN links.

Let $\mathbf{X} = (x_1, x_2, \ldots, x_P)$ be a group of P QAM[3] symbols to be sent to the receiver over the T symbol durations; these symbols must be *mapped* to the transmitted signal $\mathbf{Y}$ before launching into the air. The choice of this mapping function $\mathbf{X} \rightarrow \mathbf{Y}(\mathbf{X})$ determines which one out of several possible MIMO transmission methods results, each yielding a different combination of the diversity, array and multiplexing gains. Meanwhile, the so-called *spatial* rate of the chosen MIMO transmission method is given by the ratio P/T.

Note that, in the most general case, the considered transmit (or receive) antennas may be attached to a single transmitting (or receiving) device (base station or UE), or distributed over different devices. The symbols in $(x_1, x_2, \ldots, x_P)$ may also correspond to the data of one or possibly multiple users, giving rise to the single-user MIMO or multi-user MIMO models.

11.1.3 Single-User MIMO Techniques

Several classes of SU-MIMO transmission methods are discussed below, both optimal and suboptimal.

11.1.3.1 Optimal Transmission over MIMO Systems

The optimal way of communicating over the MIMO channel involves a channel-dependent precoder, which fulfils the roles of both transmit beamforming and power allocation across the transmitted streams, and a matching receive beamforming structure. Full channel knowledge is therefore required at the transmit side for this method to be applicable. Consider a set of $P = NT$ symbols to be sent over the channel. The symbols are separated into N streams (or layers) of T symbols each. Stream i consists of symbols $[x_{i,1}, x_{i,2}, \ldots, x_{i,T}]$. Note that in an ideal setting, each stream may adopt a distinct code rate and modulation. This is discussed in more detail below. The transmitted signal can now be written as

$$\mathbf{Y}(\mathbf{X}) = \mathbf{V}\mathbf{P}\tilde{\mathbf{X}} \tag{11.2}$$

where

$$\tilde{\mathbf{X}} = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,T} \\ \vdots & \vdots & & \vdots \\ x_{N,1} & x_{N,2} & \cdots & x_{N,T} \end{pmatrix} \tag{11.3}$$

$\mathbf{V}$ is an $N \times N$ transmit beamforming matrix and $\mathbf{P}$ is a $N \times N$ diagonal power-allocation matrix with $\sqrt{p_i}$ as its i^{th} diagonal element, where p_i is the power allocated to the i^{th} stream. Of course, the power levels must be chosen so as not to exceed the available transmit power, which can often be conveniently expressed as a constraint on the total normalized transmit power P_t.[4] Under this model, the information-theoretic capacity of the MIMO channel in bps/Hz can be obtained as [3]

$$C_{\text{MIMO}} = \log_2 \det(I + \rho \mathbf{H}\mathbf{V}\mathbf{P}^2\mathbf{V}^{\text{H}}\mathbf{H}^{\text{H}}) \tag{11.4}$$

where $\{\cdot\}^{\text{H}}$ denotes the Hermitian operator for a matrix or vector and ρ is the so-called transmit SNR, given by the ratio of the transmit power over the noise power.

[3] Quadrature Amplitude Modulation.

[4] In practice, there may be a limit on the maximum transmission power from each antenna.

The optimal (capacity-maximizing) precoder (**VP**) in Equation (11.4) is obtained by the concatenation of *singular vector beamforming* and the so-called *waterfilling power allocation*.

Singular vector beamforming means that $\mathbf{V}$ should be a unitary matrix (i.e. $\mathbf{V}^H\mathbf{V}$ is the identity matrix of size N) chosen such that $\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^H$ is the Singular-Value Decomposition[5] (SVD) of the channel matrix $\mathbf{H}$. Thus the i^{th} right singular vector of $\mathbf{H}$, given by the i^{th} column of $\mathbf{V}$, is used as a transmit beamforming vector for the i^{th} stream. At the receiver side, the optimal beamformer for the i^{th} stream is the i^{th} left singular vector of $\mathbf{H}$, obtained as the i^{th} row of $\mathbf{U}^H$:

$$\mathbf{u}_i^H\mathbf{R} = \lambda_i \sqrt{p_i}[x_{i,1}, x_{i,2}, \ldots, x_{i,T}] + \mathbf{u}_i^H\mathbf{N} \tag{11.5}$$

where λ_i is the i^{th} singular value of $\mathbf{H}$.

Waterfilling power allocation is the optimal power allocation and is given by

$$p_i = [\mu - 1/(\rho\lambda_i^2)]_+ \tag{11.6}$$

where $[x]_+$ is equal to x if x is positive and zero otherwise. μ is the so-called 'water level', a positive real variable which is set such that the total transmit power constraint is satisfied.

Thus the optimal SU-MIMO multiplexing scheme uses SVD-based transmit and receive beamforming to decompose the MIMO channel into a number of parallel non-interfering subchannels, dubbed 'eigen-channels', each one with an SNR being a function of the corresponding singular value λ_i and chosen power level p_i.

Contrary to what would perhaps be expected, the philosophy of optimal power allocation across the eigen-channels is *not* to equalize the SNRs, but rather to render them more unequal, by 'pouring' more power into the better eigen-channels, while allocating little power (or even none at all) to the weaker ones because they are seen as not contributing enough to the total capacity. This waterfilling principle is illustrated in Figure 11.3.

The underlying information-theoretic assumption here is that the information rate on each stream can be adjusted finely to match the eigen-channel's SNR. In practice, this is done by selecting a suitable Modulation and Coding Scheme (MCS) for each stream.

11.1.3.2 Beamforming with Single Antenna Transmitter or Receiver

In the case where either the receiver or the transmitter is equipped with only a single antenna, the MIMO channel exhibits only one active eigen-channel, and hence multiplexing of more than one data stream is not possible.

In *receive* beamforming, $N = 1$ and $M > 1$ (assuming a single stream). In this case, one symbol is transmitted at a time, such that the symbol-to-transmit-signal mapping function is characterized by $P = T = 1$, and $\mathbf{Y}(\mathbf{X}) = \mathbf{X} = x$, where x is the one QAM symbol to be sent. The received signal vector is given by

$$\mathbf{R} = \mathbf{H}x + \mathbf{N} \tag{11.7}$$

The receiver combines the signals from its M antennas through the use of weights $\mathbf{w} = [w_1, \ldots, w_M]$. Thus the received signal after antenna combining can be expressed as

$$z = \mathbf{w}\mathbf{R} = \mathbf{w}\mathbf{H}s + \mathbf{w}\mathbf{N} \tag{11.8}$$

[5]The reader is referred to [5] for an explanation of generic matrix operations and terminology.

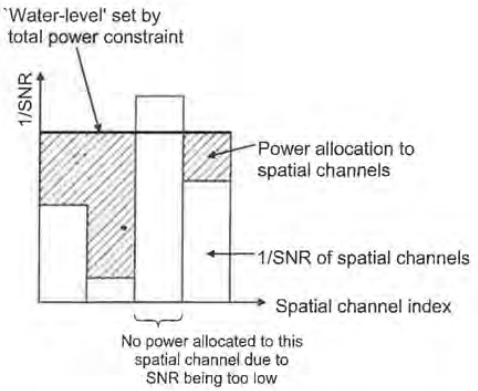


Figure 11.3: The waterfilling principle for optimal power allocation.

After the receiver has acquired a channel estimate (as discussed in Chapter 8), it can set the beamforming vector $\mathbf{w}$ to its optimal value to maximize the received SNR. This is done by aligning the beamforming vector with the UE's channel, via the so-called Maximum Ratio Combining (MRC) $\mathbf{w} = \mathbf{H}^{\mathrm{H}}$, which can be viewed as a spatial version of the well-known matched filter. Note that cancellation of an interfering signal can also be achieved, by selecting the beamforming vector to be orthogonal to the channel from the interference source. These simple concepts are illustrated in Figure 11.4.

The maximum ratio combiner provides a factor of M improvement in the received SNR compared to the $M = N = 1$ case – i.e. an array gain of $10\log_{10}(M)$ dB in the link budget.

In *transmit* beamforming, $M = 1$ and $N > 1$. The symbol-to-transmit-signal mapping function is characterized by $P = T = 1$, and $\mathbf{Y}(\mathbf{X}) = \mathbf{w}x$, where x is the one QAM symbol to be sent and $\mathbf{w}$ is the transmit beamforming vector of size $N \times 1$, computed based on channel knowledge, which is itself often obtained via a receiver-to-transmitter feedback link.[6] Assuming perfect channel knowledge at the transmitter side, the SNR-maximizing solution is given by the transmit MRC, which can be seen as a matched prefilter,

$$\mathbf{w} = \frac{\mathbf{H}^{\mathrm{H}}}{\|\mathbf{H}\|} \tag{11.9}$$

where the normalization by $\|\mathbf{H}\|$ enforces a total power constraint across the transmit antennas. The transmit MRC pre-filter provides a similar gain as its receive counterpart, namely $10\log_{10}(N)$ dB in average SNR improvement.

[6] In some situations, other techniques such as receive/transmit channel reciprocity may be used, as discussed in Section 23.5.2.

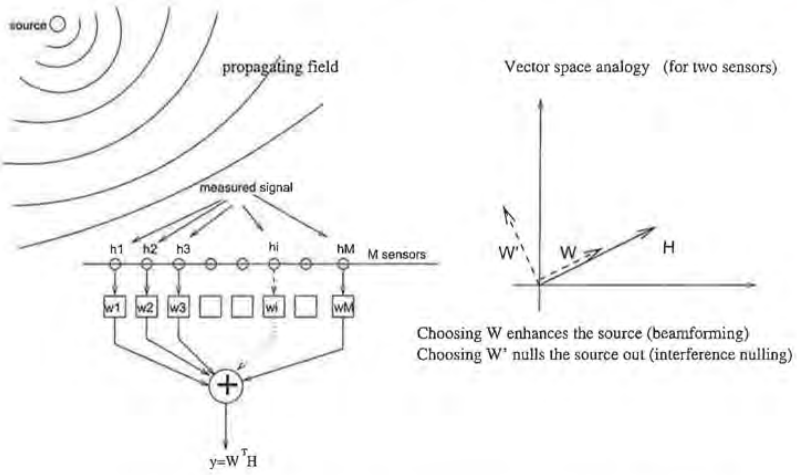


Figure 11.4: The beamforming and interference cancelling concepts.

11.1.3.3 Spatial Multiplexing without Channel Knowledge at the Transmitter

When $N > 1$ and $M > 1$, multiplexing of up to $\min(M, N)$ streams is theoretically possible even without channel knowledge at the transmitter. Assume for instance that $M \geq N$. In this case, one considers N streams, each transmitted using a different transmit antenna. As the transmitter does not have knowledge of the matrix $\mathbf{H}$, the precoder is simply the identity matrix. In this case, the symbol-to-transmit-signal mapping function is characterized by $P = NT$ and by

$$\mathbf{Y}(\mathbf{X}) = \tilde{\mathbf{X}} \tag{11.10}$$

At the receiver, a variety of linear and non-linear detection techniques may be implemented to recover the symbol matrix $\tilde{\mathbf{X}}$. A low-complexity solution is offered by the linear case, whereby the receiver superposes N beamformers $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_N$ so that the i^{th} stream $[x_{i,1}, x_{i,2}, \ldots, x_{i,T}]$ is detected as follows,

$$\mathbf{w}_i\mathbf{R} = \mathbf{w}_i\mathbf{H}\hat{\tilde{\mathbf{X}}} + \mathbf{w}_i\mathbf{N} \tag{11.11}$$

The design criterion for the beamformer $\mathbf{w}_i$ can be interpreted as a compromise between single-stream beamforming and cancelling of interference (created by the other $N-1$ streams). Inter-stream interference is fully cancelled by selecting the Zero-Forcing (ZF) receiver given by

$$\mathbf{W} = \begin{pmatrix} \mathbf{w}_1 \\ \mathbf{w}_2 \\ \vdots \\ \mathbf{w}_N \end{pmatrix} (\mathbf{H}^{\mathrm{H}}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{H}}. \tag{11.12}$$

However, for optimal performance, $\mathbf{w}_i$ should strike a balance between alignment with respect to $\mathbf{h}_i$ and orthogonality with respect to all other signatures $\mathbf{h}_k$, $k \neq i$. Such a balance is achieved by, for example, a Minimum Mean-Squared Error (MMSE) receiver.

Beyond classical linear detection structures such as the ZF or MMSE receivers, more advanced but non-linear detectors can be exploited which provide a better error rate performance at the chosen SNR operating point, at the cost of extra complexity. Examples of such detectors include the Successive Interference Cancellation (SIC) detector and the Maximum Likelihood Detector (MLD) [3]. The principle of the SIC detector is to treat individual streams, which are channel-encoded, as layers which are 'peeled off' one by one by a processing sequence consisting of linear detection, decoding, remodulating, re-encoding and subtraction from the total received signal $\mathbf{R}$. On the other hand, the MLD attempts to select the most likely set of all streams, simultaneously, from $\mathbf{R}$, by an exhaustive search procedure or a lower-complexity equivalent such as the sphere-decoding technique [3].

Multiplexing gain

The multiplexing gain corresponds to the multiplicative factor by which the spectral efficiency is increased by a given scheme. Perhaps the single most important requirement for MIMO multiplexing gain to be achieved is for the various transmit and receive antennas to experience a sufficiently different channel response. This translates into the condition that the spatial signatures of the various transmitters (the $\mathbf{h}_i$'s) (or receivers) be sufficiently decorrelated and linearly independent to make the channel matrix $\mathbf{H}$ invertible (or, more generally, well-conditioned). An immediate consequence of this condition is the limitation to $\min(M, N)$ of the number of independent streams which may be multiplexed into the MIMO channel, or, more generally, to $\text{rank}(\mathbf{H})$ streams. As an example, single-user MIMO communication between a four-antenna base station and a dual antenna UE can, at best, support multiplexing of two data streams, and thus a doubling of the UE's data rate compared with a single stream.

11.1.3.4 Diversity Techniques

Unlike the basic multiplexing scenario in Equation (11.10), where the design of the transmitted signal matrix $\mathbf{Y}$ exhibits no redundancy between its entries, a diversity-oriented design will feature some level of repetition between the entries of $\mathbf{Y}$. For 'full diversity', each transmitted symbol $x_1, x_2, \ldots, x_P$ must be assigned to each of the transmit antennas at least once in the course of the T symbol durations. The resulting symbol-to-transmit-signal mapping function is called a Space-Time Block Code (STBC). Although many designs of STBC exist, additional properties such as the orthogonality of matrix $\mathbf{Y}$ allow improved performance and easy decoding at the receiver. Such properties are realized by the Alamouti space-time code [6], explained in Section 11.2.2.1. The total diversity order which can be realized in the N to M MIMO channel is MN when entries of the MIMO channel matrix are statistically uncorrelated. The intuition behind this is that $MN - 1$ represents the number of SISO links simultaneously in a state of severe fading which the system can sustain while still being able to convey the information to the receiver. The diversity order is equal to this number plus one. As in the previous simple multiplexing scheme, an advantage of diversity-oriented transmission is that the transmitter does not need knowledge of the channel $\mathbf{H}$, and therefore no feedback of this parameter is necessary.

Diversity versus multiplexing trade-off

A fundamental aspect of the benefits of MIMO lies in the fact that any given multiple antenna configuration has a limited number of degrees of freedom. Thus there exists a compromise between reaching full beamforming gain in the detection of a desired stream of data and the perfect cancelling of undesired, interfering streams. Similarly, there exists a trade-off between the number of streams that may be multiplexed across the MIMO channel and the amount of diversity that each one of them will enjoy. Such a trade-off can be formulated from an information-theoretic point of view [7]. In the particular case of spatial multiplexing of N streams over an N to M antenna channel, with $M \geq N$, and using a linear detector, it can be shown that each stream can enjoy a maximum diversity order of $M - N + 1$.

To some extent, increasing the spatial load of MIMO systems (i.e. the number of spatially multiplexed streams) is akin to increasing the user load in CDMA systems. This correspondence extends to the fact that an optimal load level exists for a given target error rate in both systems.

11.1.4 Multi-User MIMO Techniques

11.1.4.1 Comparing Single-User and Multi-User MIMO

The set of MIMO techniques featuring data streams being communicated to (or from) antennas located on distinct UEs is referred to as Multi-User MIMO (MU-MIMO), which is one of the more recent developments of MIMO technology.

Although the MU-MIMO situation is just as well described by our model in Equation (11.1), the MU-MIMO scenario differs in a number of crucial ways from its single-user counterpart. We first explain these differences qualitatively, and then present a brief survey of the most important MU-MIMO transmission techniques.

In MU-MIMO, K UEs are selected for simultaneous communication over the same time-frequency resource, from a set of U active UEs in the cell. Typically, K is much smaller than U. Each UE is assumed to be equipped with J antennas, so the selected UEs together form a set of $M = KJ$ UE-side antennas. Since the number of streams that may be communicated over an N to M MIMO channel is limited to $\min(M, N)$ (if complete interference suppression is intended using linear combining of the antennas), the upper bound on the number of streams in MU-MIMO is typically dictated by the number of base station antennas N. The number of streams which may be allocated to each UE is limited by the number of antennas J at that UE. For instance, with single-antenna UEs, up to N streams can be multiplexed, thus achieving the maximum multiplexing gain, with a distinct stream being allocated to each UE. This is in contrast to SU-MIMO, where the transmission of N streams necessitates that the UE be equipped with at least N antennas. Therefore a great advantage of MU-MIMO over SU-MIMO is that the MIMO multiplexing benefits are preserved even in the case of low-cost UEs with a small number of antennas. As a result, it is generally assumed that, in MU-MIMO, it is the base station which bears the burden of spatially separating the UEs, be it on the uplink or the downlink. Thus the base station performs receive beamforming from several UEs on the uplink and transmit beamforming towards several UEs on the downlink.

Another fundamental contrast between SU-MIMO and MU-MIMO comes from the difference in the underlying channel model. While in SU-MIMO the decorrelation between the spatial signatures of the antennas requires rich multipath propagation or the use of

orthogonal polarizations, in MU-MIMO the decorrelation between the signatures of the different UEs occurs naturally due to fact that the separation between such UEs is typically large relative to the wavelength.

However, all these benefits of MU-MIMO depend on the level of Channel State Information at the Transmitter (CSIT) that the eNodeB receives from each UE. In the case of SU-MIMO, it has been shown that even a small number of feedback bits per antenna can be very beneficial in steering the transmitted energy accurately towards the UE's antenna(s) [8–11]. More precisely, in SU-MIMO channels, the accuracy of CSIT only causes an SNR offset, but does not affect the slope of the cell throughput-versus-SNR curve (i.e. the multiplexing gain). Yet for the MU-MIMO downlink, the level of Channel State Information (CSI) available at the transmitter does affect the multiplexing gain, because a MU-MIMO system with finite-rate feedback is essentially interference-limited after the crucial interference rejection processing has been carried out by the transmitter. Hence, providing accurate channel feedback is considerably more important for MU-MIMO than for SU-MIMO. On the other hand, in a system with a large number of UEs, the uplink resources required for accurate CSI feedback can become large and must be carefully controlled in the system design.

11.1.4.2 Techniques for Single-Antenna UEs

In considering the case of MU-MIMO for single-antenna UEs, it is worth noting that the number of antennas available to a UE for transmission is typically less than the number available for reception. We therefore examine first the uplink scenario, followed by the downlink.

With a single antenna at each UE, the MU-MIMO uplink scenario is very similar to the one described by Equation (11.10): because the UEs in mobile communication systems such as LTE typically cannot cooperate and do not have knowledge of the uplink channel coefficients, no precoding can be applied and each UE simply transmits an independent message. Thus, if K UEs are selected for transmission in the same time-frequency resource, with each UE k transmitting symbol s_k, the received signal at the base station, over a single $T = 1$ symbol period, is written as

$$\mathbf{R} = \mathbf{H}\bar{\mathbf{X}} + \mathbf{N} \tag{11.13}$$

where $\bar{\mathbf{X}} = (x_1, \ldots, x_K)^T$. In this case, the columns of $\mathbf{H}$ correspond to the receive spatial signatures of the different UEs. The base station can recover the transmitted symbol information by applying beamforming filters, for example using MMSE or ZF solutions. Note that no more than N UEs can be served (i.e. $K \leq N$) if inter-user interference is to be suppressed fully.

MU-MIMO in the uplink is sometimes referred to as 'Virtual MIMO' as, from the point of view of a given UE, there is no knowledge of the simultaneous transmissions of the other UEs. This transmission mode and its implications for LTE are discussed in Section 16.6.2.

On the downlink, which is illustrated in Figure 11.5, the base station must resort to transmit beamforming in order to separate the data streams intended for the various UEs. Over a single $T = 1$ symbol period, the signal received by UEs 1 to K can be written compactly as

$$\mathbf{R} = (r_1, \ldots, r_K)^T = \mathbf{HVP}\bar{\mathbf{X}} + \mathbf{N} \tag{11.14}$$

This time, the *rows* of $\mathbf{H}$ correspond to the transmit spatial signatures of the various UEs. $\mathbf{V}$ is the transmit beamforming matrix and $\mathbf{P}$ is the (diagonal) power allocation matrix selected

such that it fulfils the total normalized transmit power constraint P_t. To cancel out fully the inter-user interference when $K \leq N$, a transmit ZF beamforming solution may be employed (although this is not optimal due to the fact that it may require a high transmit power if the channel is ill-conditioned). Such a solution would be given by $\mathbf{V} = \mathbf{H}^{\mathrm{H}}(\mathbf{H}\mathbf{H}^{\mathrm{H}})^{-1}$.

Note that regardless of the channel realization, the power allocation must be chosen to satisfy any power constraints at the base station, for example such that $\mathrm{trace}(\mathbf{V}\mathbf{P}\mathbf{P}^{\mathrm{H}}\mathbf{V}^{\mathrm{H}}) = P_t$. It is important to note that ZF precoding requires accurate channel knowledge at the transmitter, which in most cases necessitates terminal feedback in the uplink.

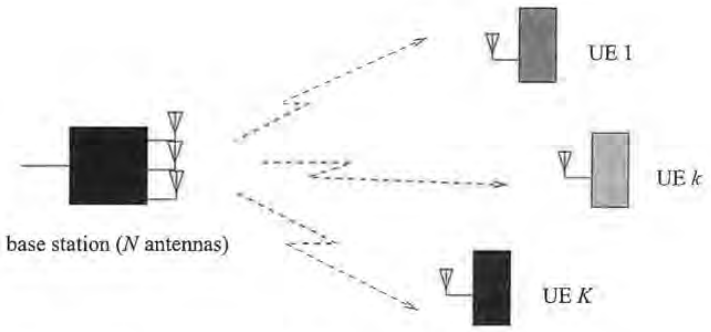


Figure 11.5: An MU-MIMO scenario in the downlink with single-antenna UEs: the base station transmits to K selected UEs simultaneously. Their signals are separated by multiple-antenna precoding at the base station side, based on channel knowledge.

11.1.4.3 Techniques for Multiple-Antenna UEs

The principles presented above for single-antenna UEs can be generalized to the case of multiple-antenna UEs. There could, in theory, be essentially two ways of exploiting the additional antennas at the UE side. In the first approach, the multiple antennas are simply treated as multiple *virtual* UEs, allowing high-capability terminals to receive or transmit more than one stream, while at the same time spatially sharing the channel with other UEs. For instance, a four-antenna base station could theoretically communicate in a MU-MIMO fashion with two UEs equipped with two antennas each, allowing two streams per UE, resulting in a total multiplexing gain of four. Another example would be that of two single-antenna UEs, receiving one stream each, and sharing access with another two-antenna UE, the latter receiving two streams. Again, the overall multiplexing factor remains limited to the number of base-station antennas.

The second approach for making use of additional UE antennas is to treat them as extra degrees of freedom for the purpose of strengthening the link between the UE and the base

station. Multiple antennas at the UE may then be combined in MRC fashion in the case of the downlink, or, in the case of the uplink, space-time coding could be used. Antenna selection is another way of extracting more diversity out of the uplink channel, as discussed in Section 16.6.

11.1.4.4 Comparing Single-User and Multi-User Capacity

To illustrate the gains of multi-user multiplexing over single-user transmission, we compare the sum-rate achieved by both types of system from an information-theoretic standpoint, for single antenna UEs. We compare the Shannon capacity in single-user and multi-user scenarios both for an idealized synthetic channel and for a channel obtained from real measurement data.

The idealized channel model assumes that the entries of the channel matrix $\mathbf{H}$ in Equation (11.13) are independently and identically distributed (i.i.d.) Rayleigh fading. For the measured channel case, a channel sounder[7] was used to perform real-time wideband channel measurements synchronously for two UEs moving at vehicular speed in an outdoor semi-urban hilly environment with Line-Of-Sight (LOS) propagation predominantly present. The most important parameters of the platform are summarized in Table 11.1.

Table 11.1: Parameters of the measured channel for SU-MIMO/MU-MIMO comparison. More details can be found in [12, 13].

Parameter	Value
Centre frequency	1917.6 MHz
Bandwidth	4.8 MHz
Base station transmit power	30 dBm
Number of antennas at base station	4 (2 cross polarized)
Number of UEs	2
Number of antennas at UE	1
Number of subcarriers	160

The sum-rate capacity of a two-UE MU-MIMO system (calculated assuming a ZF precoder as described in Section 11.1.4) is compared with the capacity of an equivalent MISO system serving a single UE at a time (i.e. in TDMA), employing beamforming (see Section 11.1.3.2). The base station has four antennas and the UE has a single antenna. Full CSIT is assumed in both cases.

Figure 11.6 shows the ergodic (i.e. average) sum-rate of both schemes in both channels. The mean is taken over all frames and all subcarriers and subsequently normalized to bps/Hz. It can be seen that in both the ideal and the measured channels, MU-MIMO yields a higher sum-rate than SU-MISO in general. In fact, at high SNR, the multiplexing gain of the MU-MIMO system is two while it is limited to one for the SU-MISO case.

However, for low SNR, the SU-MISO TDMA and MU-MIMO schemes perform very similarly. This is because a sufficiently high SNR is required to excite more than one MIMO

[7] The Eurecom MIMO OpenAir Sounder (EMOS) [12].

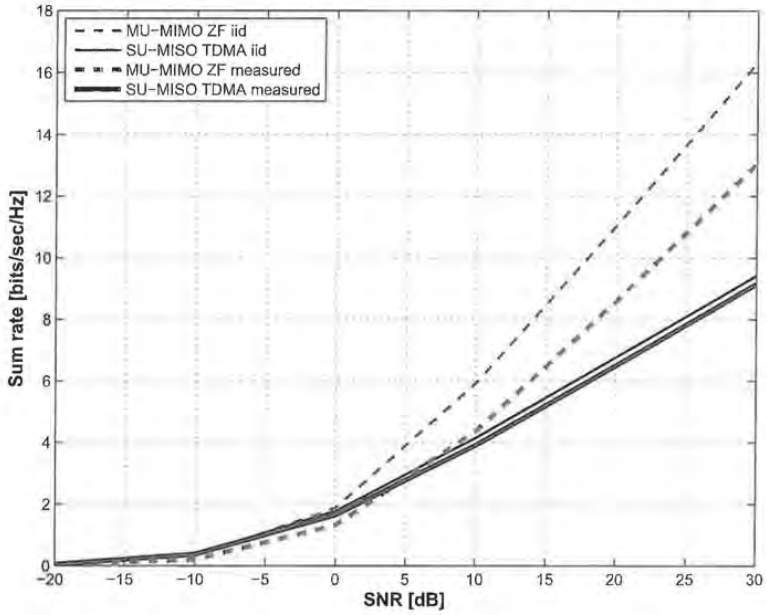


Figure 11.6: Ergodic sum-rate capacity of SU-MISO TDMA and MU-MIMO with two UEs, for an i.i.d. Rayleigh fading channel and for a measured channel.

transmission mode. Interestingly, the performance of both SU-MISO TDMA and MU-MIMO is slightly worse in the measured channels than in the idealized i.i.d. channels. This can be attributed to the correlation of the measured channel in time (due to the relatively slow movement of the users), in frequency (due to the LOS propagation), and in space (due to the transmit antenna correlation). In the MU-MIMO case, the difference between the i.i.d. and the measured channel is much higher than in the single-user TDMA case, since these correlation effects result in a rank-deficient channel matrix.

11.2 MIMO Schemes in LTE

Building on the theoretical background of the previous section, the MIMO schemes adopted for LTE Releases 8 and 9 are now reviewed and explained. These schemes relate to the downlink unless otherwise stated. The MIMO enhancements introduced for LTE-Advanced in Release 10 are explained in Chapter 29.

11.2.1 Practical Considerations

First, a few important practical constraints are briefly reviewed which affect the real-life performance of the theoretical MIMO systems considered above. Such constraints include practical deployment and antenna configurations, propagation conditions, channel knowledge, and implementation complexity. These aspects are often decisive for assessing the performance of a particular transmission strategy in a given propagation and system setting and were fundamental to the selection of MIMO schemes for LTE.

It was argued above that the full MIMO benefits (array gain, diversity gain and multiplexing gain) assume ideally decorrelated antennas and full-rank MIMO channel matrices. In this regard, the propagation environment and the antenna design (e.g. the spacing and polarization) play a significant role. If the antenna separation at the base station and the UE is small relative to the wavelength, strong correlation will be observed between the spatial signatures (especially in a LOS situation), limiting the usefulness of spatial multiplexing. However, this can to some extent be circumvented by means of dual-polarized antennas, whose design itself provides orthogonality even in LOS situations.[8] Beyond such antennas, the condition of spatial signature independence with SU-MIMO can only be satisfied with the help of rich random multipath propagation. In diversity-oriented schemes, the invertibility of the channel matrix is not required, yet the entries of the channel matrix should be statistically decorrelated. Although a greater LOS to non-LOS energy ratio will tend to correlate the fading coefficients on the various antennas, this effect will be counteracted by the reduction in fading delivered by the LOS component.

Another source of discrepancy between theoretical MIMO gains and practically achieved performance lies in the (in-)ability of the receiver and, whenever needed, the transmitter, to acquire reliable knowledge of the propagation channel. At the receiver, channel estimation is typically performed over a finite sample of Reference Signals (RSs), as discussed in Chapter 8. In the case of transmit beamforming and MIMO precoding, the transmitter then has to acquire this channel knowledge from the receiver usually through a limited feedback link, which causes further degradation to the available CSIT.

The potential advantages of MU-MIMO over SU-MIMO include robustness with respect to the propagation environment and the preservation of spatial multiplexing gain even in the case of UEs with small numbers of antennas. However, these advantages rely on the ability of the base station to compute the required transmit beamformer, which requires accurate CSIT; if no CSIT is available and the fading statistics are identical for all the UEs, the MU-MIMO gains disappear and the SU-MIMO strategy becomes optimal [14]. As a consequence, one of the most difficult challenges in making MU-MIMO practical for cellular applications, and particularly for an FDD system, is devising feedback mechanisms that allow for accurate CSI to be delivered efficiently by the UEs to the base station. This requires the use of appropriate *codebooks* for quantization (see Section 11.2.2.3). A recent account of the literature on this subject may be found in [15].

Another issue which arises for practical implementations of MIMO schemes is the interaction between the physical layer and the scheduling protocol. As noted in Section 11.1.4.1, in both uplink and downlink cases the number of UEs which can be served in a MU-MIMO fashion is typically limited to $K = N$, assuming linear combining structures. Often

[8] Horizontal and vertical polarizations, or, better, +45° and −45° polarizations, give a twofold multiplexing capability even in LOS.

one may even decide to limit K to a value strictly less than N to preserve some degrees of freedom for per-user diversity. As the number of active users U will typically exceed K, a selection algorithm must be implemented to identify which set of users will be scheduled for simultaneous transmission over a particular time-frequency Resource Block (RB). This algorithm is not specified in LTE and various approaches are possible; as discussed in Chapter 12, a combination of rate maximization and QoS constraints will typically be considered. It is important to note that the choice of UEs that will maximize the sum-rate (the sum over the K individual rates for a given subframe) is one that favours UEs exhibiting not only good instantaneous SNR but also spatial separability among their signatures.

11.2.2 Single-User Schemes

In this section, we examine the downlink multi-antenna schemes adopted in LTE for individual UEs. We consider first the diversity schemes used on the transmit side, then the single-user spatial multiplexing transmission modes, and finally beamforming using precoded UE-specific RSs (including the combination of beamforming and spatial multiplexing as introduced in Release 9).

11.2.2.1 Transmit Diversity Schemes

The theoretical aspects of transmit diversity were discussed in Section 11.1. Here we discuss the two main transmit diversity techniques defined in LTE. In LTE, transmit diversity is only defined for 2 and 4 transmit antennas and one data stream, referred to in LTE as one *codeword* since one transport block CRC[9] is used per data stream. To maximize diversity gain, the antennas typically need to be uncorrelated, so they need to be well separated relative to the wavelength or have different polarization. Transmit diversity is valuable in a number of scenarios, including low SNR, or for applications with low delay tolerance. Diversity schemes are also desirable for channels for which no uplink feedback signalling is available (e.g. Multimedia Broadcast/Multicast Services (MBMS) described in Chapter 13, Physical Broadcast CHannel (PBCH) in Chapter 9 and Synchronization Signals in Chapter 7), or when the feedback is not sufficiently accurate (e.g. at high speeds); transmit diversity is therefore also used as a fallback scheme in LTE when transmission with other schemes fails.

In LTE the multiple-antenna scheme is independently configured for the control channels and the data channels, and in the case of the data channels (Physical Downlink Shared CHannel – PDSCH) it is assigned independently per UE.

In this section we discuss in detail the transmit diversity techniques of Space-Frequency Block Codes (SFBCs) and Frequency-Switched Transmit Diversity (FSTD), as well as the combination of these schemes as used in LTE. These transmit diversity schemes may be used in LTE for the PBCH and Physical Downlink Control CHannel (PDCCH), and also for the PDSCH if it is configured in transmit diversity mode[10] for a UE.

Another transmit diversity technique which is commonly associated with OFDM is Cyclic Delay Diversity (CDD). CDD is not used in LTE as a diversity scheme in its own right but

[9]Cyclic Redundancy Check.

[10]PDSCH transmission mode 2 – see Section 9.2.2.1. As noted above, transmit diversity may also be used as a fallback in other transmission modes.

rather as a precoding scheme for spatial multiplexing on the PDSCH; we therefore introduce it in Section 11.2.2.2 in the context of spatial multiplexing.

Space-Frequency Block Codes (SFBCs)

If a physical channel in LTE is configured for transmit diversity operation using two eNodeB antennas, pure SFBC is used. SFBC is a frequency-domain version of the well-known Space-Time Block Codes (STBCs), also known as Alamouti codes [6]. This family of codes is designed so that the transmitted diversity streams are orthogonal and achieve the optimal SNR with a linear receiver. Such orthogonal codes only exist for the case of two transmit antennas.

STBC is used in UMTS, but in LTE the number of available OFDM symbols in a subframe is often odd while STBC operates on pairs of adjacent symbols in the time domain. The application of STBC is therefore not straightforward for LTE, while the multiple subcarriers of OFDM lend themselves well to the application of SFBC.

For SFBC transmission in LTE, the symbols transmitted from the two eNodeB antenna ports on each pair of adjacent subcarriers are defined as follows:

$$\begin{bmatrix} y^{(0)}(1) & y^{(0)}(2) \\ y^{(1)}(1) & y^{(1)}(2) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \\ -x_2^* & x_1^* \end{bmatrix} \tag{11.15}$$

where $y^{(p)}(k)$ denotes the symbols transmitted from antenna port p on the k^{th} subcarrier.

Since no orthogonal codes exist for antenna configurations beyond 2×2, SFBC has to be modified in order to apply it to the case of 4 transmit antennas. In LTE, this is achieved by combining SFBC with FSTD.

Frequency Switched Transmit Diversity (FSTD) and its Combination with SFBC

General FSTD schemes transmit symbols from each antenna on a different set of subcarriers. In LTE, FSTD is only used in combination with SFBC for the case of 4 transmit antennas, in order to provide a suitable transmit diversity scheme where no orthogonal rate-1 block code exists. The LTE scheme is in fact a combination of two 2×2 SFBC schemes mapped to independent subcarriers as follows:

$$\begin{bmatrix} y^{(0)}(1) & y^{(0)}(2) & y^{(0)}(3) & y^{(0)}(4) \\ y^{(1)}(1) & y^{(1)}(2) & y^{(1)}(3) & y^{(1)}(4) \\ y^{(2)}(1) & y^{(2)}(2) & y^{(2)}(3) & y^{(2)}(4) \\ y^{(3)}(1) & y^{(3)}(2) & y^{(3)}(3) & y^{(3)}(4) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & 0 & 0 \\ 0 & 0 & x_3 & x_4 \\ -x_2^* & x_1^* & 0 & 0 \\ 0 & 0 & -x_4^* & x_3^* \end{bmatrix} \tag{11.16}$$

where, as previously, $y^{(p)}(k)$ denotes the symbols transmitted from antenna port p on the k^{th} subcarrier. Note that the mapping of symbols to antenna ports is different in the 4 transmit-antenna case compared to the 2 transmit-antenna SFBC scheme. This is because the density of cell-specific RSs on the third and fourth antenna ports is half that of the first and second antenna ports (see Section 8.2.1), and hence the channel estimation accuracy may be lower on the third and fourth antenna ports. Thus, this design of the transmit diversity scheme avoids concentrating the channel estimation losses in just one of the SFBC codes, resulting in a slight coding gain.

11.2.2.2 Spatial Multiplexing Schemes

We begin by introducing some terminology used to describe spatial multiplexing in LTE:

- A spatial *layer* is the term used in LTE for one of the different streams generated by spatial multiplexing as described in Section 11.1. A layer can be described as a mapping of symbols onto the transmit antenna ports.[11] Each layer is identified by a precoding vector of size equal to the number of transmit antenna ports and can be associated with a radiation pattern.
- The *rank* of the transmission is the number of layers transmitted.[12]
- A *codeword* is an independently encoded data block, corresponding to a single Transport Block (TB) delivered from the Medium Access Control (MAC) layer in the transmitter to the physical layer, and protected with a CRC.

For ranks greater than 1, two codewords can be transmitted. Note that the number of codewords is always less than or equal to the number of layers, which in turn is always less than or equal to the number of antenna ports.

In principle, a SU-MIMO spatial multiplexing scheme can either use a single codeword mapped to all the available layers, or multiple codewords each mapped to one or more different layers. The main benefit of using only one codeword is a reduction in the amount of control signalling required, both for Channel Quality Indicator (CQI) reporting, where only a single value would be needed for all layers, and for HARQ ACK/NACK[13] feedback, where only one ACK/NACK would have to be signalled per subframe per UE. In such a case, the MLD receiver is optimal in terms of minimizing the BER. At the opposite extreme, a separate codeword could be mapped to each of the layers. The advantage of such a mapping is that significant gains are possible by using SIC, albeit at the expense of more signalling being required. An MMSE-SIC receiver can be shown to approach the Shannon capacity [7]. Note that an MMSE receiver is viable for both transmitter structures. For LTE, a middle-way was adopted whereby at most two codewords are used, even if four layers are transmitted. The codeword-to-layer mapping is static, since only minimal gains were shown for a dynamic mapping method. The mappings are shown in Table 11.2. Note that in LTE all RBs belonging to the same codeword use the same MCS, even if a codeword is mapped to multiple layers.

Precoding

The PDSCH transmission modes for open-loop spatial multiplexing[14] and closed-loop spatial multiplexing[15] use precoding from a defined 'codebook' to form the transmitted layers. Each codebook consists of a set of predefined precoding matrices, with the size of the set being a trade-off between the number of signalling bits required to indicate a particular matrix in the codebook and the suitability of the resulting transmitted beam direction.

[11] See Section 8.2 for an explanation of the concept of an antenna port.

[12] Note that the rank of a matrix is the number of linearly independent rows or columns, or equivalently the dimension of the subspace generated by the rows or columns of the matrix. Hence the rank of the channel matrix may be higher than the number of transmitted layers.

[13] Hybrid Automatic Repeat reQuest ACKnowledgement/Negative ACKnowledgement.

[14] PDSCH transmission mode 3 – see Section 9.2.2.1.

[15] PDSCH transmission modes 4 and 6 – see Section 9.2.2.1.

Table 11.2: Codeword-to-layer mapping in LTE.

Transmission rank	Codeword 1	Codeword 2
Rank 1	Layer 1	
Rank 2	Layer 1	Layer 2
Rank 3	Layer 1	Layer 2 and Layer 3
Rank 4	Layer 1 and Layer 2	Layer 3 and Layer 4

In the case of closed-loop spatial multiplexing, a UE feeds back to the eNodeB the index of the most desirable entry from a predefined codebook. The preferred precoder is the matrix which would maximize the capacity based on the receiver capabilities. In a single-cell, interference-free environment the UE will typically indicate the precoder that would result in a transmission with an effective SNR following most closely the largest singular values of its estimated channel matrix.

Some important properties of the LTE codebooks are as follows:

- **Constant modulus property.** LTE uses precoders which mostly comprise pure phase corrections – that is, with no amplitude changes. This is to ensure that the Power Amplifier (PA) connected to each antenna is loaded equally. The one exception (which still maintains the constant modulus property) is using an identity matrix as the precoder. However, although the identity precoder may completely switch off one antenna on one layer, since each layer is still connected to one antenna at constant power the net effect across the layers is still constant modulus to the PA.
- **Nested property.** The nested property is a method of arranging the codebooks of different ranks so that the lower rank codebook is comprised of a subset of the higher rank codebook vectors. This property simplifies the CQI calculation across different ranks. It ensures that the precoded transmission for a lower rank is a subset of the precoded transmission for a higher rank, thereby reducing the number of calculations required for the UE to generate the feedback. For example, if a specific index in the codebook corresponds to columns 1, 2 and 3 from the precoder $\mathbf{W}$ in the case of a rank 3 transmission, then the same index in the case of rank 2 transmission must consist of either columns 1 and 2 or columns 1 and 3 from $\mathbf{W}$.
- **Minimal 'complex' multiplications.** The 2-antenna codebook consists entirely of a QPSK[16] alphabet, which eliminates the need for any complex multiplications since all codebook multiplications use only ± 1 and $\pm j$. The 4-antenna codebook does contain some QPSK entries which require a $\sqrt{2}$ magnitude scaling as well; it was considered that the performance gain of including these precoders justified the added complexity.

The 2-antenna codebook in LTE comprises one 2×2 identity matrix and two DFT (Discrete Fourier Transform) matrices:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 1 \\ j & -j \end{bmatrix} \tag{11.17}$$

[16] Quadrature Phase Shift Keying.

Here the columns of the matrices correspond to the layers.

The 4 transmit antenna codebook in LTE uses a Householder generating function:

$$\mathbf{W}_H = \mathbf{I} - 2\mathbf{u}\mathbf{u}^H/\mathbf{u}^H\mathbf{u}, \tag{11.18}$$

which generates unitary matrices from input vectors $\mathbf{u}$, which are defined in [16, Table 6.3.4.2.3-2]. The advantage of using precoders generated with this equation is that it simplifies the CQI calculation (which has to be carried out for each individual precoder in order to determine the preferred precoder) by reducing the number of matrix inversions. This structure also reduces the amount of control signalling required, since the optimum rank-1 version of $\mathbf{W}_H$ is always the first column of $\mathbf{W}_H$; therefore the UE only needs to indicate the preferred precoding matrix, and not also the individual vector within it which would be optimal for the case of rank-1 transmission.

Note that both DFT and Householder matrices are unitary.

Cyclic Delay Diversity (CDD)

In the case of *open*-loop spatial multiplexing,[17] the feedback from the UE indicates only the *rank* of the channel, and not a preferred precoding matrix. In this mode, if the rank used for PDSCH transmission is greater than 1 (i.e. more than one layer is transmitted), LTE uses Cyclic Delay Diversity (CDD) [17]. CDD involves transmitting the same set of OFDM symbols on the same set of OFDM subcarriers from multiple transmit antennas, with a different delay on each antenna. The delay is applied before the Cyclic Prefix (CP) is added, thereby guaranteeing that the delay is cyclic over the Fast Fourier Transform (FFT) size. This gives CDD its name.

Adding a time delay is identical to applying a phase shift in the frequency domain. As the same time delay is applied to all subcarriers, the phase shift will increase linearly across the subcarriers with increasing subcarrier frequency. Each subcarrier will therefore experience a different beamforming pattern as the non-delayed subcarrier from one antenna interferes constructively or destructively with the delayed version from another antenna. The diversity effect of CDD therefore arises from the fact that different subcarriers will pick out different spatial paths in the propagation channel, thus increasing the frequency-selectivity of the channel. The channel coding, which is applied to a whole transport block across the subcarriers, ensures that the whole transport block benefits from the diversity of spatial paths.

Although this approach does not optimally exploit the channel in the way that ideal precoding would (by matching the precoding to the eigenvectors of the channel), it does help to ensure that any destructive fading is constrained to individual subcarriers rather than affecting a whole transport block. This can be particularly beneficial if the channel information at the transmitter is unreliable, for example due to the feedback being limited or the UE velocity being high.

The general principle of the CDD technique is illustrated in Figure 11.7. The fact that the delay is added before the CP means that any delay value can be used without increasing the overall delay spread of the channel. By contrast, if the delay had been added after the addition of the CP, then the usable delays would have had to be kept small in order to ensure that the delay spread of the delayed symbol is no more than the maximum channel delay spread, in order to obviate any need to increase the CP length.

[17] PDSCH transmission mode 3 – see Section 9.2.2.1.

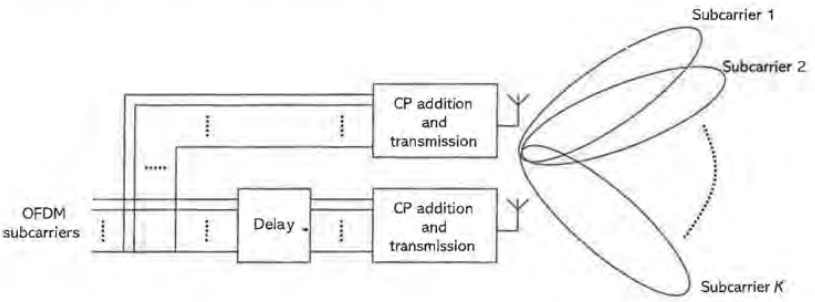


Figure 11.7: Principle of Cyclic Delay Diversity.

The time-delay/phase-shift equivalence means that the CDD operation can be implemented as a frequency-domain precoder for the affected antenna(s), where the precoder phase changes on a per-subcarrier basis according to a fixed linear function. In general, the implementation designer can choose whether to implement CDD in the time domain or the frequency domain. However, one advantage of a frequency-domain implementation is that it is not limited to delays corresponding to an integer number of samples.

As an example of CDD for a case with 2 transmit antenna ports, we can express mathematically the received symbol r_k on the k^{th} subcarrier as

$$r_k = h_{1k}x_k + h_{2k}e^{j\phi k}x_k \tag{11.19}$$

where h_{pk} is the channel from the p^{th} transmit antenna and $e^{j\phi k}$ is the phase shift on the k^{th} subcarrier due to the delay operation. We can see clearly that on some subcarriers the symbols from the second transmit antenna will add constructively, while on other subcarriers they will add destructively. Here, $\phi = 2\pi d_{\text{cdd}}/N$, where N is the FFT size and d_{cdd} is the delay in samples.

The number of peaks and troughs created by CDD in the received signal spectrum across the subcarriers therefore depends on the length of the delay d_{cdd}: as d_{cdd} is increased, the number of peaks and troughs in the spectrum also increases.

In LTE, the frequency-domain phase shift values ϕ are π, $2\pi/3$ and $\pi/2$ for 2-, 3- and 4-layer transmission respectively. For a size-2048 FFT, for example, these values correspond to $d_{\text{cdd}} = 1024,\ 682.7,\ 512$ sample delays respectively.

For the application of CDD in LTE, the eNodeB transmitter combines CDD delay-based phase shifts with additional precoding using fixed unitary DFT-based precoding matrices.

The application of precoding in this way is useful when the channel coefficients of the antenna ports are correlated, since then *virtual* antennas (formed by fixed, non-channel-dependent precoding) will typically be uncorrelated. On its own, the benefit of CDD is reduced by antenna correlation. The use of uncorrelated virtual antennas created by fixed precoding can avoid this problem in correlated channels, while not degrading the performance if the individual antenna ports are uncorrelated.

For ease of explanation, the above discussion of CDD has been in terms of a rank-1 transmission – i.e. with a single layer. However, in practice, CDD is only applied in LTE when the rank used for PDSCH transmission is greater than 1. In such a case, each layer benefits independently from CDD in the same way as for a single layer. For example, for a rank-2 transmission, the transmission on the second antenna port is delayed relative to the first antenna port for each layer. This means that symbols transmitted on both layers will experience the delay and hence the increased frequency selectivity.

For multilayer CDD operation, the mapping of the layers to antenna ports is carried out using precoding matrices selected from the spatial multiplexing codebooks described earlier. As the UE does not indicate a preferred precoding matrix in the open loop spatial multiplexing transmission mode in which CDD is used, the particular spatial multiplexing matrices selected from the spatial multiplexing codebooks in this case are predetermined.

In the case of 2 transmit antenna ports, the predetermined spatial multiplexing precoding matrix **W** is always the same (the first entry in the 2 transmit antenna port codebook, which is the identity matrix). Thus, the transmitted signal can be expressed as follows:

$$\begin{bmatrix} y^{(0)}(k) \\ y^{(1)}(k) \end{bmatrix} = \mathbf{W}\mathbf{D}_2\mathbf{U}_2\mathbf{x} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & e^{j\phi_1 k} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x^{(0)}(i) \\ x^{(1)}(i) \end{bmatrix} \tag{11.20}$$

In the case of 4 transmit antenna ports, ν different precoding matrices are used from the 4 transmit antenna port codebook where ν is the transmission rank. These ν precoding matrices are applied in turn across groups of ν subcarriers in order to provide additional decorrelation between the spatial streams.

11.2.2.3 Beamforming Schemes

The theoretical aspects of beamforming were described in Section 11.1. This section explains how it is implemented in LTE Release 8 and Release 9.

In LTE, two beamforming techniques are supported for the PDSCH:

- **Closed-loop rank-1 precoding.**[18] This mode amounts to beamforming since only a single layer is transmitted, exploiting the gain of the antenna array. However, it can also be seen as a special case of the SU-MIMO spatial multiplexing using codebook based precoding discussed in Section 11.2.2.2. In this mode, the UE feeds channel state information back to the eNodeB to indicate suitable precoding to apply for the beamforming operation. This feedback information is known as Precoding Matrix Indicators (PMIs) and is described in detail in Section 11.2.2.4.
- **Beamforming with UE-specific RSs.**[19] In this case, the MIMO precoding is not restricted to a predefined codebook, so the UE cannot use the cell-specific RS for PDSCH demodulation and precoded UE-specific RS are therefore needed.

In this section we focus on the latter case which supports a single transmitted layer in LTE Release 8 and was extended to support dual-layer beamforming in Release 9. These modes are primarily mechanisms to extend cell coverage by concentrating the eNodeB power in the directions in which the radio channel offers the strongest path(s) to reach the UE. They are

[18] PDSCH transmission mode 6 – see Section 9.2.2.1.

[19] PDSCH transmission mode 7 (from Release 8) and 8 (introduced in Release 9) – see Section 9.2.2.1.

typically implemented with an array of closely spaced antenna elements (or pairs of cross-polarized elements) for creating directional transmissions. Due to the fact that no precoding codebook is used, an arbitrary number of transmit antennas is theoretically supported; in practical deployments up to 8 antennas are typically used (e.g. an array of 4 cross-polarized pairs). The signals from the correlated antenna elements are phased appropriately so that they add up constructively at the location where the UE is situated. The UE is not really aware that it is receiving a precoded directional transmission rather than a cell-wide transmission (other than by the fact that the UE is directed to use the UE-specific RS as the phase reference for demodulation). To the UE, the phased array of antennas 'appears' as just one antenna port.

In LTE Release 8, only a single UE-specific antenna port and associated RSs is defined (see Section 8.2.2). As a consequence the beamformed PDSCH can only be transmitted with a single spatial layer. The support of two UE-specific antenna ports was introduced in Release 9 (see Section 8.2.3), thus enabling beamforming with two spatial layers with a direct one-to-one mapping between the layers and antenna ports. The UE-specific RS patterns for the two Release 9 antenna ports were designed to be orthogonal in order to facilitate separation of the spatial layers at the receiver. The two associated PDSCH layers can then be transmitted to a single user in good propagation conditions to increase its data rate, or to multiple users (MU-MIMO) to increase the system capacity. LTE Release 10 further extends the support for beamforming with UE-specific RSs to higher numbers of spatial layers, as described in Chapter 29.

When two layers are transmitted over superposed beams they share the available transmit power of the eNodeB, so the increase in data rate comes at the expense of a reduction of coverage. This makes fast rank adaptation (which is also used in the codebook-based spatial multiplexing transmission mode) an important mechanism for the dual-layer beamforming mode. To support this, it is possible to configure the UE to feed back a Rank Indicator (RI) together with the PMI (see Section 11.2.2.4).

It should also be noted that beamforming in LTE can only be applied to the PDSCH and not to the downlink control channels, so although the range of a given data rate on the PDSCH can be extended by beamforming, the overall cell coverage may still be limited by the range of the control channels unless other measures are taken. The trade-off between throughput and coverage with beamforming in LTE is illustrated in Figure 11.8.

Like any precoding technique, these beamforming modes require reliable CSIT which can, in principle, be obtained either from feedback from the UE or from estimation of the uplink channel.

In the single-layer beamforming mode of Release 8, the UE does not feed back any precoding-related information, and the eNodeB deduces this information from the uplink, for example using Direction Of Arrival (DOA) estimations. It is worth noting that in this case calibration of the eNodeB RF paths may be necessary, as discussed in Section 23.5.2.

The dual-layer beamforming mode of Release 9 provides the possibility for the UE to help the eNodeB to derive the beamforming precoding weights by sending PMI feedback in the same way as for codebook-based closed-loop spatial multiplexing. For beamforming, this is particularly suitable for FDD deployments where channel reciprocity cannot be exploited as effectively as in the case of TDD. Nevertheless, this PMI feedback can only provide partial CSIT, since it is based on measurements of the common RSs which, in beamforming deployments, are usually transmitted from multiple antenna elements with broad beam patterns; furthermore, the PMI feedback uses the same codebook as for closed-loop spatial

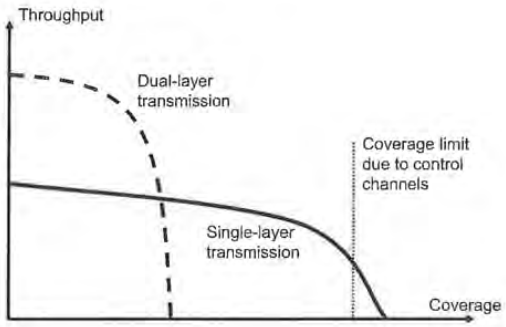


Figure 11.8: Trade-off between throughput and coverage with beamforming.

multiplexing and is therefore constrained by quantization while the beamforming precoding is not. As a consequence, additional uplink channel estimates are still required if the full benefits of beamforming are to be realized.

For the Channel Quality Indicator (CQI) feedback for link adaptation, the eNodeB must take into account the fact that in the beamforming modes the channel quality experienced by the UE on the beamformed PDSCH resources will typically be different (hopefully better) than on the cell-specific RSs from which the CQI is derived. The UE-specifc RSs cannot be used for CQI estimation because they are only present in the RBs in which beamforming transmission is used; UE-specific RSs that are precoded for a different UE are obviously not useful for channel estimation. The eNodeB has to adjust the MCS used for transmission relative to the CQI reports by estimating the expected beamforming gain depending on the number of transmitted layers; the HARQ acknowledgement rate can be useful in deriving such an adjustment.

It is worth mentioning that beamforming also affects the interference caused to neighbouring cells. If the beamforming is intermittent, it can result in a problem often known as the 'flash-light effect', where strong intermittent interference may disturb the accuracy of the UEs' CQI reporting in adjacent cells. This effect was shown to have the potential to cause throughput reductions in HSDPA [18]. In LTE the possibility for frequency-domain scheduling (as discussed in Section 12.5) provides an additional degree of freedom to avoid such issues. In addition, beamforming may even be coordinated to take advantage of the spatial interference structure created by beamforming in multiple cells, especially among the cells controlled by one eNodeB.

11.2.2.4 Feedback Computation and Signalling

In order to support MIMO operation. the UE can be configured to report Precoding Matrix Indicators (PMIs) and Rank Indicators(RIs), in addition to CQI reporting as discussed in Section 10.2.1.

In the case of codebook-based spatial multiplexing transmissions, the precoding at the eNodeB transmitter is applied relative to the transmitted phase of the cell-specific RSs for each antenna port. Thus if the UE knows the precoding matrices that could be applicable (as defined in the configured codebook) and it knows the transfer function of the channels from the different antenna ports (by making measurements on the RSs), it can determine which **W** is most suitable under the current radio conditions and signal this to the eNodeB. The preferred **W**, whose index constitutes the PMI report, is the precoder that maximizes the aggregate number of data bits which could be received across all layers. The number of RBs to which each PMI corresponds in the frequency domain depends on the feedback mode configured by the eNodeB:

- **Wideband PMI.** The UE reports a single PMI corresponding to the preferred precoder assuming transmission over the whole system bandwidth. This is applicable for PMI feedback that is configured to be sent periodically (on the PUCCH[20] unless the PUSCH[21] is transmitted), and also for PMI feedback that is sent aperiodically in conjunction with UE-selected sub-band CQI reports on the PUSCH (see Section 10.2.1).
- **Sub-band PMI.** The UE reports one PMI for each sub-band across the whole system bandwidth, where the sub-band size is between 4 and 6 RBs depending on the system bandwidth (as shown in Table 10.3). This is applicable for PMI feedback that is sent aperiodically on the PUSCH in conjunction with a wideband CQI report. It is well suited to a scenario where the eNodeB wishes to schedule a wideband data transmission to a UE while the channel direction varies across the bandwidth; the precoder used by the eNodeB can change from sub-band to sub-band within one transport block.
- **UE-selected sub-band PMI.** The UE selects a set of M preferred sub-bands, each of size k RBs (where M and k are the same as for UE-selected sub-band CQI feedback as given in Table 10.4) and reports a single PMI corresponding to the preferred precoder assuming transmission over all of the M selected sub-bands. This is applicable in conjunction with UE-selected sub-band CQI reports on the PUSCH.

The eNodeB is not bound to use the precoder requested by the UE, but clearly if the eNodeB chooses another precoder then the eNodeB will have to adjust the MCS accordingly since the reported CQI could not be assumed to be applicable.

The eNodeB may also restrict the set of precoders which the UE may evaluate and report. This is known as *codebook subset restriction*. It enables the eNodeB to prevent the UE from reporting precoders which are not useful, for example in some eNodeB antenna configurations or in correlated fading scenarios. In the case of open-loop spatial multiplexing, codebook subset restriction amounts simply to a restriction on the rank which the UE may report.

For each PDSCH transmission to a UE, the eNodeB indicates via a 'Transmitted Precoding Matrix Indicator' (TPMI) in the downlink assignment message[22] on the PDCCH whether it is applying the UE's preferred precoder, and if not, which precoder is used. This enables the UE to derive the correct phase reference relative to the cell-specific RSs in order to

[20] Physical Uplink Control CHannel.
[21] Physical Uplink Shared CHannel.
[22] The TPMI is sent in DCI Formats 2, 2A and 2B – see Section 9.3.5.1.

demodulate the PDSCH data. The ability to indicate that the eNodeB is applying the UE's preferred precoder is particularly useful in the case of sub-band PMI reporting, as it enables multiple precoders to be used across the bandwidth of a wideband transmission to a UE without incurring a high signalling overhead to notify the UE of the set of precoders used.

The UE can also be configured to report the channel rank via a RI, which is calculated to maximize the capacity over the entire bandwidth, jointly selecting the preferred precoder(s) to maximize the capacity on the assumption of the selected transmission rank. The CQI values reported by the UE correspond to the preferred transmission rank and precoders, to enable the eNodeB to perform link adaptation and multi-user scheduling as discussed in Sections 10.2 and 12.1. The number of CQI values reported normally corresponds to the number of codewords supported by the preferred transmission rank. Further, the CQI values themselves will depend on the assumed transmission rank: for example, the precoding matrix for layer 1 will usually be different depending on whether or not the UE is assuming the presence of a second layer.

Although the UE indicates the transmission rank that would maximize the downlink data rate, the eNodeB can also indicate to the UE that a different transmission rank is being used for a PDSCH transmission. This gives flexibility to the eNodeB, since the UE does not know the amount of data in the downlink buffer; if the amount of data in the buffer is small, the eNodeB may prefer to use a lower rank with higher reliability.

11.2.3 Multi-User MIMO

The main focus of MIMO in the first release of LTE was on achieving transmit antenna diversity or single-user multiplexing gain, neither of which requires such a sophisticated CSI feedback mechanism as MU-MIMO. As pointed out in Section 11.2.1, the availability of accurate CSIT is the main challenge in making MU-MIMO schemes attractive for cellular applications.

Nevertheless, basic support for MU-MIMO was provided in Release 8, and enhanced schemes were added in Releases 9 and 10.

The MU-MIMO scheme supported in LTE Release 8[23] is based on the same codebook-based scheme as is defined for SU-MIMO (see Section 11.2.2.2). In particular the same 'implicit' feedback calculation mechanism is used, whereby the UE makes hypotheses on the precoder the eNodeB would use for transmission on the PDSCH and reports a recommended precoding matrix (PMI) corresponding to the best hypothesis – i.e. an *implicit* representation of the CSI. The recommended precoder typically depends on the type of receiver implemented by the UE, e.g. MRC, linear MMSE or MMSE Decision Feedback Equalizer (DFE). However, in the same way as for SU-MIMO, the eNodeB is not bound to the UE's recommendation and can override the PMI report by choosing a different precoder from the codebook and signalling it back to the scheduled UEs using a TPMI on the PDCCH.

In the light of the similarities between implicit channel feedback by precoder recommendation and explicit channel vector quantization, it is instructive to consider the implications of the fact that the LTE 2-antenna codebook and the first eight entries in the 4-antenna codebook are derived from a DFT matrix (as noted in Section 11.2.2.2). The main reasons for this choice are as follows:

[23] PDSCH transmission mode 5 – see Section 9.2.2.1.

- Simplicity in the PMI/CQI calculation, which can be done in some cases without complex multiplications;
- A DFT matrix nicely captures the characteristics of a highly correlated MISO channel.[24]

In LTE Release 8, a UE configured in MU-MIMO mode assumes that an eNodeB transmission on the PDSCH would be performed on one layer using one of the rank-1 codebook entries defined for two or four antenna ports. Therefore, the MU-MIMO transmission is limited to a single layer and single codeword per scheduled UE, and the RI and PMI feedback are the same as for SU-MIMO rank 1.

The eNodeB can then co-schedule multiple UEs on the same RB, typically by pairing UEs that report orthogonal PMIs. The eNodeB can therefore signal to those UEs (using one bit in the DCI format 1D message on the PDCCH – see Section 9.3.5.1) that the PDSCH power is reduced by 3 dB relative to the cell-specific RSs. This assumes that the transmission power of the eNodeB is equally divided between the two codewords and implies that in practice a maximum of two UEs may be co-scheduled in Release 8 MU-MIMO mode.

Apart from this power offset signalling, MU-MIMO in Release 8 is fully transparent to the scheduled UEs, in that they are not explicitly aware of how many other UEs are co-scheduled or of which precoding vectors are used for any co-scheduled UEs.

CQI reporting for MU-MIMO in Release 8 does not take into account any interference from transmissions to co-scheduled users. The CQI values reported are therefore equivalent to those that would be reported for rank-1 SU-MIMO. It is left to the eNodeB to estimate a suitable adjustment to apply to the MCS if multiple UEs are co-scheduled.

In Release 9, the support for MU-MIMO is enhanced by the introduction of dual-layer beamforming in conjunction with precoded UE-specific RSs, as discussed in Sections 8.2.3 and 11.2.2.3. This removes the constraint on the eNodeB to use the feedback codebook for precoding, and gives flexibility to utilize other approaches to MU-MIMO user separation, for example according to a zero-forcing beamforming criterion. Note that in the limit of a large UE population, a zero-forcing beamforming solution converges to a unitary precoder because, with high probability, there will be N UEs with good channel conditions reporting orthogonal channel signatures.

The use of precoded UE-specific RSs removes the need to signal explicitly the details of the chosen precoder to the UEs. Instead, the UEs estimate the effective channel (i.e. the combination of the precoding matrix and physical channel), from which they can derive the optimal antenna combiner for their receivers.

The UE-specific RS structure described in Section 8.2.3 allows up to 4 UEs to be scheduled for MU-MIMO transmissions from Release 9.

Furthermore, the precoded UE-specific RSs mean that for Release 9 MU-MIMO it is not necessary to signal explicitly a power offset depending on the number of co-scheduled UEs. As the UE-specific RSs for two UEs are code-division multiplexed, the eNodeB transmission power is shared between the two antenna ports in the same way as the transmitted data, so the UE can use the UE-specific RSs as both the phase reference and the amplitude reference for demodulation. Thus the MU-MIMO scheduling is fully transparent in Release 9.

[24] It is not difficult to see that if the first N rows are taken from a DFT matrix of size N_q, each of the N_q column vectors comprises the spatial signature of the i^{th} element of a uniform linear antenna array with spacing l, such that $\frac{l}{\lambda}\sin\theta = \frac{i}{N_q}$ [7].

11.2.4 MIMO Performance

The multiple antenna schemes supported by LTE contribute much to its spectral efficiency improvements compared to UMTS.

Figure 11.9 shows a typical link throughput performance comparison between the open-loop SU-MIMO spatial multiplexing mode of LTE Release 8 with rank feedback but no PMI feedback, and closed-loop SU-MIMO spatial multiplexing with PMI feedback. Four transmit and two receive antennas are used. Both modes make use of spatial multiplexing gain to enhance the peak data rate. Closed-loop MIMO provides gain from the channel-dependent precoding, while open-loop MIMO offers additional robustness by SFBC for single-layer transmissions and CDD for more than one layer.

The gain of closed-loop precoding can be observed in low mobility scenarios, while the benefit of open-loop diversity can clearly be seen in the case of high mobility.

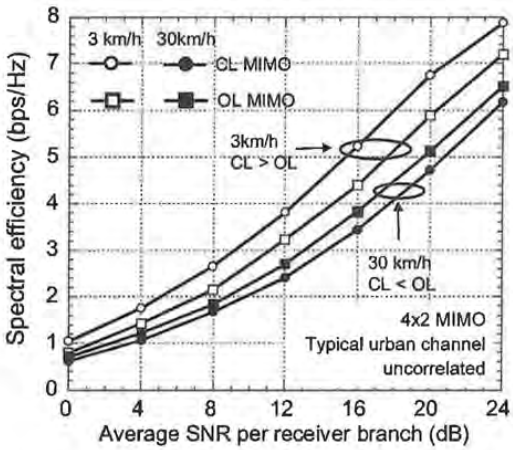


Figure 11.9: Example of closed-loop and open-loop MIMO performance.

11.3 Summary

In this chapter we have reviewed the predominant families of MIMO techniques (both single-user and multi-user) of relevance to LTE. LTE breaks new ground in drawing on such approaches to harness the power of MIMO systems not just for boosting the peak per-user data rate but also for improving overall system capacity and spectral efficiency. Single-User MIMO techniques are well-developed in Release 8, providing the possibility to benefit from precoding while avoiding a high control signalling overhead. Support for beamforming and Multi-User MIMO is significantly enhanced in Release 9. LTE-Advanced further develops and enhances these techniques, as outlined in Chapter 29.

References[25]

[1] D. Gesbert, M. Shafi, P. Smith, D. Shiu and A. Naguib, 'From Theory to Practice: An Overview of MIMO Space-time Coded Wireless Systems'. *IEEE Journal on Selected Areas in Communications*. Special Issue on MIMO systems, guest edited by the authors, April 2003.

[2] H. Bölcskei, D. Gesbert, C. Papadias and A. J. van der Veen (eds), *Space-Time Wireless Systems: From Array Processing to MIMO Communications*. Cambridge: Cambridge University Press, 2006.

[3] A. Goldsmith, *Wireless Communications*. Cambridge: Cambridge University Press, 2005.

[4] D. Gesbert, S. Hanly, H. Huang, S. Shamai, O. Simeone and W. Yu, 'Multi-Cell MIMO Cooperative Networks: A New Look at Interference', to appear in *IEEE Journal on Selected Areas in Communications*, December 2010.

[5] G. Golub and C. van Loan, *Matrix Computations*. Baltimore, MA: John Hopkins University Press, 1996.

[6] S. M. Alamouti, 'A Simple Transmitter Diversity Technique for Wireless Communications'. *IEEE Journal on Selected Areas in Communications*, Vol. 16, pp. 1451–1458, October 1998.

[7] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge: Cambridge University Press, 2004.

[8] A. Narula, M. J. Lopez, M. D. Trott and G. W. Wornell, 'Efficient Use of Side Information in Multiple Antenna Data Transmission over Fading Channels'. *IEEE Journal on Selected Areas in Communications*, Vol. 16, pp. 1223–1436, October 1998.

[9] D. Love, R. Heath Jr and T. Strohmer, 'Grassmannian Beamforming for Multiple-input Multiple output Wireless Systems'. *IEEE Trans. on Information Theory*, Vol. 49, pp. 2735–2747, October 2003.

[10] D. Love, R. Heath Jr, W. Santipach, and M. Honig, 'What is the Value of Limited Feedback for MIMO Channels?'. *IEEE Communication Magazine*, Vol. 42, pp. 54–59, October 2003.

[11] N. Jindal, 'MIMO Broadcast Channels with Finite Rate Feedback'. *IEEE Trans. Information Theory*, Vol. 52, pp. 5045–5059, November 2006.

[12] F. Kaltenberger, L. S. Cardoso, M. Kountouris, R. Knopp and D. Gesbert, 'Capacity of Linear Multi-user MIMO Precoding Schemes with Measured Channel Data', in *Proc. IEEE Workshop on Sign. Proc. Adv. in Wireless Comm. (SPAWC)*, Recife, Brazil, July 2008.

[13] F. Kaltenberger, M. Kountouris, D. Gesbert and R. Knopp, 'Correlation and Capacity of Measured Multi-user MIMO Channels', in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Cannes, France, September 2008.

[14] Caire, G. and Shamai, S. (Shitz), 'On the Achievable Throughput of a Multi-antenna Gaussian Broadcast Channel'. *IEEE Trans. on Information Theory*, Vol. 49, pp. 1691–1706, July 2003.

[15] R. Heath, V. Lau, D. Love, D. Gesbert, B. Rao and A. Andrews, (eds), 'Exploiting Limited Feedback in Tomorrow's Wireless Communications Networks'. Special issue of *IEEE Journal on Selected Areas in Communications*, October 2008.

[16] 3GPP Technical Specification 36.211, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation', www.3gpp.org.

[17] Y. Li, J. Chuang and N. Sollenberger, 'Transmitter Diversity for OFDM Systems and Its Impact on High-Rate Data Wireless Networks'. *IEEE Journal on Selected Areas in Communications*, Vol. 17, pp. 1233–1243, July 1999.

[18] A. Osseiran and A. Logothetis, 'Closed Loop Transmit Diversity in WCDMA HS-DSCH', in *Proc. IEEE Vehicular Technology Conference*, May 2005.

[25] All web sites confirmed 1st March 2011.

12

Multi-User Scheduling and Interference Coordination

Issam Toufik and Raymond Knopp

12.1 Introduction

The eNodeB in an LTE system is responsible, among other functions, for managing resource scheduling for both uplink and downlink channels. The ultimate aim of this function is typically to fulfil the expectations of as many users of the system as possible, taking into account the Quality of Service (QoS) requirements of their respective applications.

A typical single-cell cellular radio system is shown in Figure 12.1, comprising K User Equipments (UEs) communicating with one eNodeB over a fixed total bandwidth B. In the uplink, each UE has several data queues corresponding to different uplink logical channel groups, each with different delay and rate constraints. In the same way, in the downlink the eNodeB may maintain several buffers per UE containing dedicated data traffic, in addition to queues for broadcast services. The different traffic queues in the eNodeB would typically have different QoS constraints. We further assume for the purposes of this analysis that only one user is allocated a particular Resource Block (RB) in any subframe, although in practice multi-user MIMO schemes may result in more than one UE per RB, as discussed in Sections 11.2.3 and 16.6.2 for the downlink and uplink respectively.

The goal of a resource scheduling algorithm in the eNodeB is to allocate the RBs and transmission powers for each subframe in order to optimize a function of a set of performance metrics, for example maximum/minimum/average throughput, maximum/minimum/average delay, total/per-user spectral efficiency or outage probability. In the downlink the resource allocation strategy is constrained by the total transmission power of the eNodeB, while in the uplink the main constraints on transmission power in different RBs arises from a multicell view of inter-cell interference and the power headroom of the UEs.

In this chapter we provide an overview of some of the key families of scheduling algorithms which are relevant for an LTE system and highlight some of the factors that an eNodeB can advantageously take into account.

LTE – The UMTS Long Term Evolution: From Theory to Practice, Second Edition.
Stefania Sesia, Issam Toufik and Matthew Baker.
© 2011 John Wiley & Sons, Ltd. Published 2011 by John Wiley & Sons, Ltd.

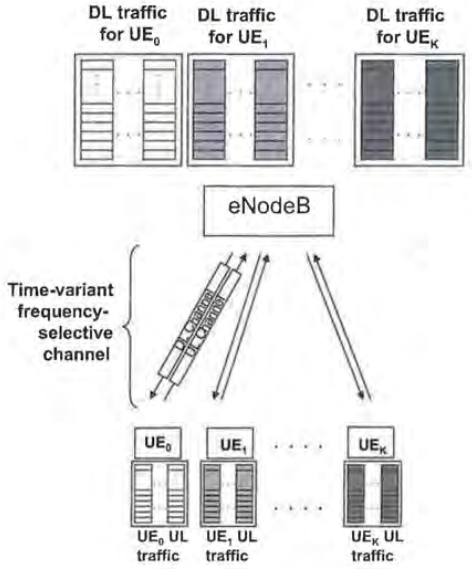


Figure 12.1: A typical single-cell cellular radio system.

12.2 General Considerations for Resource Allocation Strategies

The generic function of a resource scheduler, as shown for the downlink case in Figure 12.2, is to schedule data to a set of UEs on a shared set of physical resources.

It is worth noting that the algorithm used by the resource scheduler is also tightly coupled with the Adaptive Modulation and Coding (AMC) scheme and the retransmission protocol (Hybrid Automatic Repeat reQuest, HARQ, see Sections 4.4 and 10.3.2.5). This is due firstly to the fact that, in addition to dynamic physical resource allocation, the channel measurements are also used to adapt the Modulation and Coding Scheme (MCS) (i.e. transmission spectral-efficiency) as explained in Section 10.2. Secondly, the queue dynamics, which impact the throughput and delay characteristics of the link seen by the application, depend heavily on the HARQ protocol and transport block sizes. Moreover, the combination of channel coding and retransmissions provided by HARQ enables the spectral efficiency of an individual transmission in one subframe to be traded off against the number of subframes in which retransmissions take place. Well-designed practical scheduling algorithms will necessarily consider all these aspects.

In general, scheduling algorithms can make use of two types of measurement information to inform the scheduling decisions, namely Channel State Information (CSI) and traffic measurements (volume and priority). These are obtained either by direct measurements at

the eNodeB, or via feedback signalling channels, or a combination of both. The amount of feedback used is an important consideration, since the availability of accurate CSI and traffic information helps to maximize the data rate in one direction at the expense of more overhead in the other. This fundamental trade-off, which is common to all feedback-based resource scheduling schemes, is particularly important in Frequency Division Duplex (FDD) operation where uplink-downlink reciprocity of the radio channels cannot be assumed (see Chapter 20). For Time Division Duplex (TDD) systems, coherence between the uplink and downlink channels may be used to assist the scheduling algorithm, as discussed in Section 23.5.

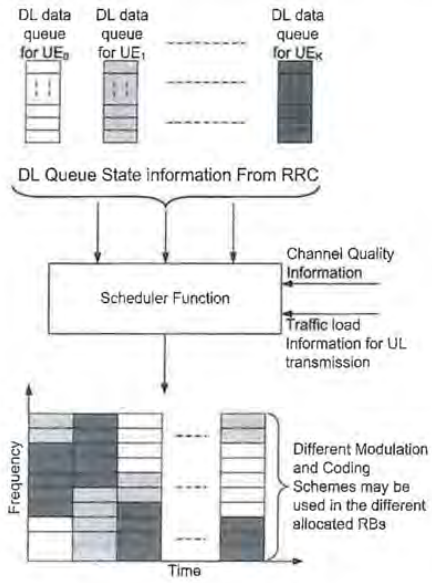


Figure 12.2: Generic view of a wideband resource scheduler.

Based on the available measurement information, the eNodeB resource scheduler must manage the differing requirements of all the UEs in the cells under its control to ensure that sufficient radio transmission resources are allocated to each UE within acceptable latencies to meet their QoS requirements in a spectrally efficient way. The details of this process are not standardized as it is largely internal to the eNodeB, allowing for vendor-specific algorithms to be developed which can be optimized for specific scenarios in conjunction with network operators. However, the key inputs available to the resource scheduling process are common.

In general, two extremes of scheduling algorithm may be identified: *opportunistic scheduling* and *fair scheduling*. The former is typically designed to maximize the sum of the transmitted data rates to all users by exploiting the fact that different users experience different channel gains and hence will experience good channel conditions at different times and frequencies. A fundamental characteristic of mobile radio channels is the fading effects

arising from the mobility of the UEs in a multipath propagation environment, and from variations in the surrounding environment itself (see Chapter 20). In [1–3] it is shown that, for a multi-user system, significantly more information can be transmitted across a fading channel than a non-fading channel for the same average signal power at the receiver. This principle is known as *multi-user diversity*. With proper dynamic scheduling, allocating the channel at each given time instant only to the user with the best channel condition in a particular part of the spectrum can yield a considerable increase in the total throughput as the number of active users becomes large.

The main issue arising from opportunistic resource allocation schemes is the difficulty of ensuring fairness and the required QoS. Users' data cannot always wait until the channel conditions are sufficiently favourable for transmission, especially in slowly varying channels. Furthermore, as explained in Chapter 1, it is important that network operators can provide reliable wide area coverage, including to stationary users near the cell edge – not just to the users which happen to experience good channel conditions by virtue of their proximity to the eNodeB.

The second extreme of scheduling algorithm, fair scheduling, therefore pays more attention to latency and achieving a minimum data rate for each user than to the total data rate achieved. This is particularly important for real-time applications such as Voice-over-IP (VoIP) or video-conferencing, where a certain minimum rate must be guaranteed independently of the channel state.

In practice, most scheduling algorithms fall between the two extremes outlined above, including elements of both to deliver the required mix of QoS. A variety of metrics can be used to quantify the degree of fairness provided by a scheduling algorithm, the general objective being to avoid heavily penalizing the cell-edge users in an attempt to give high throughputs to the users with good channel conditions. One example is based on the Cumulative Distribution Function (CDF) of the throughput of all users, whereby a system may be considered sufficiently fair if the CDF of the throughput lies to the right-hand side of a particular line, such as that shown in Figure 12.3. Another example is the *Jain index* [4], which gives an indication of the variation in throughput between users. It is calculated as:

$$J = \frac{\overline{T}^2}{\overline{T}^2 + \mathrm{var}(T)}, \quad 0 \le J \le 1, \tag{12.1}$$

where $\overline{T}$ and $\mathrm{var}(T)$ are the mean and variance respectively of the average user throughputs. The more similar the average user throughputs ($\mathrm{var}(T) \rightarrow 0$), the higher the value of J.

Other factors also need to be taken into account, especially the fact that, in a coordinated deployment, individual cells cannot be considered in isolation – nor even the individual set of cells controlled by a single eNodeB. The eNodeBs should take into account the interference generated by co-channel cells, which can be a severe limiting factor, especially for cell-edge users. Similarly, the performance of the system as a whole can be enhanced if each eNodeB also takes into account the impact of the transmissions of its own cells on the neighbouring cells. These inter-cell aspects, and the corresponding inter-eNodeB signalling mechanisms provided in LTE, are discussed in detail in Section 12.5.

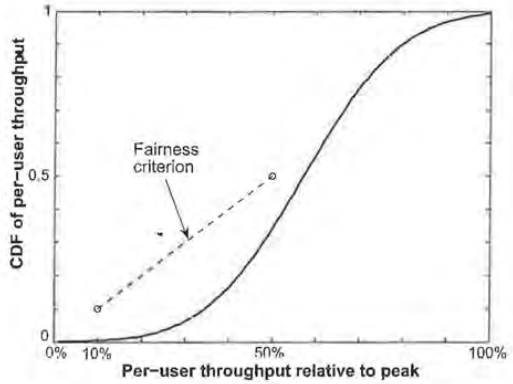


Figure 12.3: An example of a metric based on the throughput CDF over all users for scheduler fairness evaluation (10–50 metric).

12.3 Scheduling Algorithms

Multi-user scheduling finds its basis at the interface between information theory and queueing theory, in the theory of capacity-maximizing resource allocation. Before establishing an algorithm, a capacity-related metric is first formulated and then optimized across all possible resource allocation solutions satisfying a set of predetermined constraints. Such constraints may be physical (e.g. bandwidth and total power) or QoS-related.

Information theory offers a range of possible capacity metrics which are relevant in different system operation scenarios. Two prominent examples are explained here, namely the so-called *ergodic capacity* and *delay-limited capacity*, corresponding respectively to soft and hard forms of rate guarantee for the user.

12.3.1 Ergodic Capacity

The ergodic capacity (also known as the Shannon capacity) is defined as the maximum data rate which can be sent over the channel with asymptotically small error probability, averaged over the fading process. When CSI is available at the Transmitter (i.e. CSIT), the transmit power and mutual information[1] between the transmitter and the receiver can be varied depending on the fading state in order to maximize the average rates. The ergodic capacity metric considers the long-term average data rate which can be delivered to a user when the user does not have any latency constraints.

[1]The mutual information bewteen two random variables X and Y with probability density function $p(X)$ and $p(Y)$ respectively is defined as $I(X;Y) = \mathbb{E}_{X,Y}\,[\log p(X;Y)/(p(X)p(Y))]$, where $\mathbb{E}[.]$ is the expectation function [5].

12.3.1.1 Maximum Rate Scheduling

It has been shown in [3, 6] that the maximum total ergodic sum rate $\sum_{k=1}^{K} R_k$, where R_k is the total rate allocated to user k, is achieved with a transmission power given by

$$P_k(m, f) = \begin{cases} \left[\frac{1}{\lambda_k} - \frac{P_N}{|H_k(m, f)|^2}\right]^+ & \text{if } |H_k(m, f)|^2 \geq \frac{\lambda_k}{\lambda_{k'}}|H_{k'}(m, f)|^2 \\ 0 & \text{otherwise} \end{cases} \tag{12.2}$$

where $[x]^+ = \max(0, x)$, $H_k(m, f)$ is the channel gain of user k in RB m of subframe f and λ_k are constants which are chosen in order to satisfy an average per-user power constraint.

This approach is known in the literature as *maximum sum rate* scheduling. The result in Equation (12.2) shows that the maximum sum rate is achieved by orthogonal multiplexing where in each subchannel (i.e. each RB in LTE) only the user with the best channel gain is scheduled. This orthogonal scheduling property is in line with, and thus justifies, the philosophy of the LTE multiple-access schemes. The input power spectra given by Equation (12.2) are water-filling formulae in both frequency and time (i.e. allocating more power to a scheduled user when his channel gain is high and less power when it is low).[2]

A variant of this resource allocation strategy with no power control (relevant for cases where the power control dynamic range is limited or zero) is considered in [2]. This allocation strategy is called 'maximum-rate constant-power' scheduling, where only the user with the best channel gain is scheduled in each RB, but with no adaptation of the transmit power. It is shown in [2] that most of the performance gains offered by the maximum rate allocation in Equation (12.2) are due to multi-user diversity and not to power control, so an on-off power allocation can achieve comparable performance to maximum sum rate scheduling. This not only allows some simplification of the scheduling algorithm but also is well suited to a downlink scenario where the frequency-domain dynamic range of the transmitted power in a given subframe is limited by constraints such as the dynamic range of the UE receivers and the need to transmit wideband reference signals for channel estimation.

12.3.1.2 Proportional Fair Scheduling

The ergodic sum rate corresponds to the optimal rate for traffic which has no delay constraint. This results in an unfair sharing of the channel resources. When the QoS required by the application includes latency constraints, such a scheduling strategy is not suitable and other fairer approaches need to be considered. One such approach is the well-known Proportional Fair Scheduling (PFS) algorithm. PFS schedules a user when its instantaneous channel quality is high relative to its own average channel condition over time. It allocates user $\hat{k}_m$ in RB m in any given subframe f if [7]

$$\hat{k}_m = \underset{k=1,\ldots,K}{\operatorname{argmax}} \frac{R_k(m, f)}{T_k(f)} \tag{12.3}$$

where $T_k(f)$ denotes the long-term average throughput of user k computed in subframe f and $R_k(m, f) = \log(1 + \mathrm{SNR}_k(m, f))$ is the achievable rate by user k in RB m and subframe f.

[2] Water-filling strategies are discussed in more detail in the context of MIMO in Section 11.1.3.

The long-term average user throughputs are recursively computed by

$$T_k(f) = \left(1 - \frac{1}{t_c}\right) T_k(f-1) + \frac{1}{t_c} \sum_{m=1}^{M} R_k(m, f) I\{\hat{k}_m = k\} \tag{12.4}$$

where t_c is the time window over which fairness is imposed and $I\{\cdot\}$ is the indicator function equal to one if $\hat{k}_m = k$ and zero otherwise. A large time window t_c tends to maximize the total average throughput; in fact, in the limit of a very long time window, PFS and maximum-rate constant-power scheduling result in the same allocation of resources. For small t_c, the PFS tends towards a round-robin[3] scheduling of users [7].

Several studies of PFS have been conducted in the case of WCDMA/HSDPA[4], yielding insights which can be applied to LTE. In [8] the link level system performance of PFS is studied, taking into account issues such as link adaptation dynamic range, power and code resources, convergence settings, signalling overhead and code multiplexing. [9] analyses the performance of PFS in the case of VoIP, including a comparison with round-robin scheduling for different delay budgets and packet-scheduling settings. A study of PFS performance in the case of video streaming in HSDPA can be found in [10].

12.3.2 Delay-Limited Capacity

Even though PFS introduces some fairness into the system, this form of fairness may not be sufficient for applications which have a very tight latency constraint. For these cases, a different capacity metric is needed; one such example is referred to as the 'delay-limited capacity'. The delay-limited capacity (also known as zero-outage capacity) is defined as the transmission rate which can be guaranteed in all fading states under finite long-term power constraints. In contrast to the ergodic capacity, where mutual information between the transmitter and the receiver vary with the channel, the powers in delay-limited capacity are coordinated between users and RBs with the objective of maintaining constant mutual information independently of fading states. The delay-limited capacity is relevant to traffic classes where a given data rate must be guaranteed throughout the connection time, regardless of the fading dips.

The delay-limited capacity for a flat-fading multiple-access channel is characterized in [11]. The wideband case is analysed in [7, 12].

It is shown in [7] that guaranteeing a delay-limited rate incurs only a small throughput loss in high Signal to Noise Ratio (SNR) conditions, provided that the number of users is large. However, the solution to achieve the delay-limited capacity requires non-orthogonal scheduling of the users, with successive decoding in each RB. This makes this approach basically unsuitable for LTE.

It is, however, possible to combine orthogonal multiple access with hard QoS requirements. Examples of algorithms achieving such a compromise can be found in [12, 13]. It can be shown that even under hard fairness constraints it is possible to achieve performance which is very close to the optimal unfair policy; thus hard fairness constraints do not necessarily introduce a significant throughput degradation, even with orthogonal resource allocation, provided that the number of users and the system bandwidth are large. This may be a relevant

[3] A round-robin approach schedules each user in turn, irrespective of their prevailing channel conditions.
[4] Wideband Code Division Multiple Access / High Speed Downlink Packet Access.

scenario for a deployment targeting VoIP users, where densities of several hundred users per cell are foreseen (as mentioned in Section 1.2), with a latency constraint typically requiring each packet to be successfully delivered within 50 ms.

In general this discussion of scheduling strategies assumes that all users have an equal and infinite queue length – often referred to as a full-buffer traffic model. In practice this is not the case, especially for real-time services, and information on users' queue lengths is necessary to guarantee system stability. If a scheduling algorithm can be found which keeps the average queue length bounded, the system is said to be stabilized [14]. One approach to ensuring that the queue length remains bounded is to use the queue length to set the priority order in the allocation of RBs. This generally works for lightly loaded systems. In [15] and [16] it is shown that in wideband frequency-selective channels, low average packet delay can be achieved even if the fading is very slow.

12.4 Considerations for Resource Scheduling in LTE

In LTE, each logical channel has a corresponding QoS description which should influence the behaviour of the eNodeB resource scheduling algorithm. Based on the evolution of the radio and traffic conditions, this QoS description could potentially be updated for each service in a long-term fashion. The mapping between the QoS descriptions of different services and the resource scheduling algorithm in the eNodeB can be a key differentiating factor between radio network equipment manufacturers.

An important constraint for the eNodeB scheduling algorithm is the accuracy of the eNodeB's knowledge of the channel quality for the active UEs in the cell. The manner in which such information is provided to the scheduler in LTE differs between uplink and downlink transmissions. In practice, for the downlink this information is provided through the feedback of Channel Quality Indicators (CQIs)[5] by UEs as described in Chapter 10, while for the uplink the eNodeB may use Sounding Reference Signals (SRSs) or other signals transmitted by the UEs to estimate the channel quality, as discussed in Chapter 15. The frequency with which CQI reports and SRS are transmitted is configurable by the eNodeB, allowing for a trade-off between the signalling overhead and the availability of up-to-date channel information. If the most recent CQI report or SRS was received a significant time before the scheduling decision is taken, the performance of the scheduling algorithm can be significantly degraded.

In order to perform frequency-domain scheduling, the information about the radio channel needs to be frequency-specific. For this purpose, the eNodeB may configure the CQI reports to relate to specific sub-bands to assist the downlink scheduling, as explained in Section 10.2.1. Uplink frequency-domain scheduling can be facilitated by configuring the SRS to be transmitted over a large bandwidth. However, for cell-edge UEs the wider the transmitted bandwidth the lower the available power per RB; this means that accurate frequency-domain scheduling may be more difficult for UEs near the cell edge. Limiting the SRS to a subset of the system bandwidth will improve the channel quality estimation on these RBs but restrict the ability of the scheduler to find an optimal scheduling solution for all users. In general, provided that the bandwidth over which the channel can be estimated for

[5]More generally, in the context of user selection for MU-MIMO, full CSI is relevant, to support user pairing as well as rate scheduling.

scheduling purposes is greater than the intended scheduling bandwidth for data transmission by a sufficient factor, a useful element of multi-user diversity gain may still be achievable.

As noted above, in order to support QoS- and queue-aware scheduling, it is necessary for the scheduler to have not only information about the channel quality, but also information on the queue status. In the LTE downlink, knowledge of the amount of buffered data awaiting transmission to each UE is inherently available in the eNodeB; for the uplink, Section 4.4.2.2 explains the buffer status reporting mechanisms available to transfer such information to the eNodeB.

12.5 Interference Coordination and Frequency Reuse

One limiting aspect for system throughput performance in cellular networks is inter-cell interference, especially for cell-edge users. Careful management of inter-cell interference is particularly important in systems such as LTE which are designed to operate with a frequency reuse factor of one.

The scheduling strategy of the eNodeB may therefore include an element of Inter-Cell Interference Coordination (ICIC), whereby interference from and to the adjacent cells is taken into account in order to increase the data rates which can be provided for users at the cell edge. This implies for example imposing restrictions on what resources in time and/or frequency are available to the scheduler, or what transmit power may be used in certain time/frequency resources.

The impact of interference on the achievable data rate for a given user can be expressed analytically. If a user k is experiencing no interference, then its achievable rate in an RB m of subframe f can be expressed as

$$R_{k,\text{no-Int}}(m,f) = W\log\left(1+\frac{P^s(m,f)|H_k^s(m,f)|^2}{P_N}\right) \tag{12.5}$$

where $H_k^s(m,f)$ is the channel gain from the serving cell s to user k, $P^s(m,f)$ is the transmit power from cell s, P_N is the noise power and W is the bandwidth of one RB (i.e. 180 kHz). If neighbouring cells are transmitting in the same time-frequency resources, then the achievable rate for user k reduces to

$$R_{k,\text{Int}}(m,f) = W\log\left(1+\frac{P^s(m,f)|H_k^s(m,f)|^2}{P_N+\sum_{i\neq s}P^i(m,f)|H_k^i(m,f)|^2}\right) \tag{12.6}$$

where the indices i denote interfering cells.

The rate loss for user k can then be expressed as

$$\begin{aligned} R_{k,\text{loss}}(m,f) &= R_{k,\text{no-Int}}(m,f) - R_{k,\text{Int}}(m,f) \\ &= W\log\left(\frac{1+\text{SNR}}{1+\left[\frac{1}{\text{SNR}}+\frac{\sum_{i\neq s}P^i(m,f)|H_k^i(m,f)|^2}{P^s(m,f)|H_k^s(m,f)|^2}\right]^{-1}}\right) \end{aligned} \tag{12.7}$$

Figure 12.4 plots the rate loss for user k as a function of the total inter-cell interference to signal ratio $\alpha = \left(\sum_{i\neq s}P^i(m,f)|H_k^i(m,f)|^2\right)/\left(P^s(m,f)|H_k^s(m,f)|^2\right)$, with SNR = 0 dB. It can

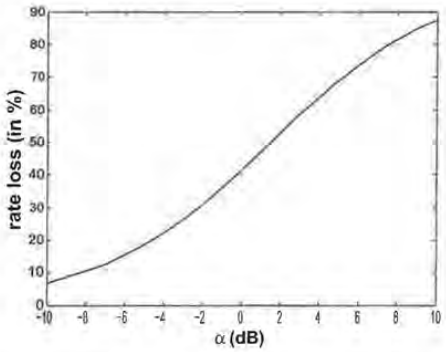


Figure 12.4: User's rate loss due to interference.

easily be seen that for a level of interference equal to the desired signal level (i.e. $\alpha \approx 0$ dB), user k experiences a rate loss of approximately 40%.

In order to demonstrate further the significance of interference and power allocation depending on the system configuration we consider two examples of a cellular system with two cells (s_1 and s_2) and one active user per cell (k_1 and k_2 respectively). Each user receives the wanted signal from its serving cell, while the inter-cell interference comes from the other cell.

In the first example, each user is located near its respective eNodeB (see Figure 12.5(a)). The channel gain from the interfering cell is small compared to the channel gain from the serving cell ($|H_{k_1}^{s_1}(m,f)| \gg |H_{k_1}^{s_2}(m,f)|$ and $|H_{k_2}^{s_2}(m,f)| \gg |H_{k_2}^{s_1}(m,f)|$). In the second example (see Figure 12.5(b)), we consider the same scenario but with the users now located close to the edge of their respective cells. In this case the channel gain from the serving cell and the interfering cell are comparable ($|H_{k_1}^{s_1}(m,f)| \approx |H_{k_1}^{s_2}(m,f)|$ and $|H_{k_2}^{s_2}(m,f)| \approx |H_{k_2}^{s_1}(m,f)|$).

The capacity of the system with two eNodeBs and two users can be written as

$$R_{\text{Tot}} = W\left(\log\left(1 + \frac{P^{s_1}|H_{k_1}^{s_1}(m,f)|^2}{P_N + P^{s_2}|H_{k_1}^{s_2}(m,f)|^2}\right) + \log\left(1 + \frac{P^{s_2}|H_{k_2}^{s_2}(m,f)|^2}{P_N + P^{s_1}|H_{k_2}^{s_1}(m,f)|^2}\right)\right) \tag{12.8}$$

From this equation, it can be noted that the optimal transmit power operating point in terms of maximum achievable throughput is different for the two considered cases. In the first scenario, the maximum throughput is achieved when both eNodeBs transmit at maximum power, while in the second the maximum capacity is reached by allowing only one eNodeB to transmit. It can in fact be shown that the optimal power allocation for maximum capacity for this situation with two base stations is binary in the general case; this means that either both base stations should be operating at maximum power in a given RB, or one of them should be turned off completely in that RB [17].

From a practical point of view, this result can be exploited in the eNodeB scheduler by treating users in different ways depending on whether they are cell-centre or cell-edge users. Each cell can then be divided into two parts – inner and outer. In the inner part, where users experience a low level of interference and also require less power to communicate with the

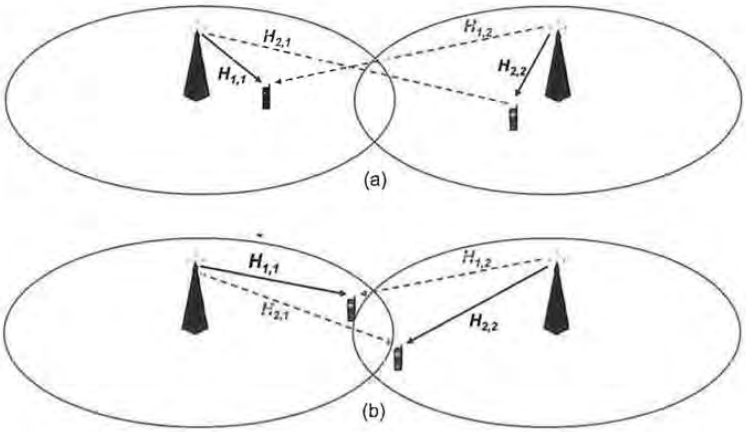


Figure 12.5: System configuration: (a) users close to eNodeBs; (b) users at the cell edge.

serving cell, a frequency reuse factor of 1 can be adopted. For the outer part, scheduling restrictions are applied: when the cell schedules a user in a given part of band, the system capacity is optimized if the neighbouring cells do not transmit at all; alternatively, they may transmit only at low power (probably to users in the inner parts of the neighbouring cells) to avoid creating strong interference to the scheduled user in the first cell. This effectively results in a higher frequency reuse factor at the cell-edge; it is often referred to as 'partial frequency reuse' or 'soft frequency reuse', and is illustrated in Figure 12.6.

In order to coordinate the scheduling in different cells in such a way, communication between neighbouring cells is required. If the neighbouring cells are managed by the

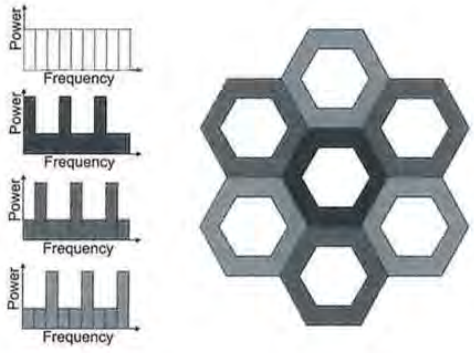


Figure 12.6: Partial frequency reuse.

same eNodeB, a coordinated scheduling strategy can be followed without the need for standardized signalling. However, where neighbouring cells are controlled by different eNodeBs, standardized signalling is important, especially in multivendor networks. The main mechanism for ICIC in LTE Releases 8 and 9 is normally assumed to be frequency-domain-based, at least for the data channels, and the Release 8/9 inter-eNodeB ICIC signalling explained in the following two sections is designed to support this. In Release 10, additional time-domain mechanisms are introduced, aiming particularly to support ICIC for the PDCCH and for heterogeneous networks comprising both macrocells and small cells; these mechanisms are explained in Section 31.2.

12.5.1 Inter-eNodeB Signalling to Support Downlink Frequency-Domain ICIC in LTE

In relation to the downlink transmissions, a bitmap termed the Relative Narrowband Transmit Power (RNTP[6]) indicator can be exchanged between eNodeBs over the X2 interface. Each bit of the RNTP indicator corresponds to one RB in the frequency domain and is used to inform the neighbouring eNodeBs if a cell is planning to keep the transmit power for the RB below a certain upper limit or not. The value of this upper limit, and the period for which the indicator is valid into the future, are configurable. This enables the neighbouring cells to take into account the expected level of interference in each RB when scheduling UEs in their own cells. The reaction of the eNodeB in case of receiving an indication of high transmit power in an RB in a neighbouring cell is not standardized (thus allowing some freedom of implementation for the scheduling algorithm); however, a typical response could be to avoid scheduling cell-edge UEs in such RBs. In the definition of the RNTP indicator, the transmit power per antenna port is normalized by the maximum output power of a base station or cell. The reason for this is that a cell with a smaller maximum output power, corresponding to smaller cell size, can create as much interference as a cell with a larger maximum output power corresponding to a larger cell size.

12.5.2 Inter-eNodeB Signalling to Support Uplink Frequency-Domain ICIC in LTE

For the uplink transmissions, two messages may be exchanged between eNodeBs to facilitate some coordination of their transmit powers and scheduling of users:

A reactive indicator, known as the 'Overload Indicator' (OI), can be exchanged over the X2 interface to indicate physical layer measurements of the average uplink interference plus thermal noise for each RB. The OI can take three values, expressing low, medium, and high levels of interference plus noise. In order to avoid excessive signalling load, it cannot be updated more often than every 20 ms.

A proactive indicator, known as the 'High Interference Indicator' (HII), can also be sent by an eNodeB to its neighbouring eNodeBs to inform them that it will, in the near future, schedule uplink transmissions by one or more cell-edge UEs in certain parts of the bandwidth, and therefore that high interference might occur in those frequency regions. Neighbouring cells may then take this information into consideration in scheduling their own users to limit the interference impact. This can be achieved either by deciding not to schedule their own

[6]RNTP is defined in [18, Section 5.2.1].

cell-edge UEs in that part of the bandwidth and only considering the allocation of those resources for cell-centre users requiring less transmission power, or by not scheduling any user at all in the relevant RBs. The HII is comprised of a bitmap with one bit per RB and, like the OI, is not sent more often than every 20 ms. The HII bitmap is addressed to specific neighbour eNodeBs.

In addition to frequency-domain scheduling in the uplink, the eNodeB also controls the degree to which each UE compensates for the path-loss when setting its uplink transmission power. This enables the eNodeB to trade off fairness for cell-edge UEs against inter-cell interference generated towards other cells, and can also be used to maximize system capacity. This is discussed in more detail in Section 18.3.2.

12.5.3 Static versus Semi-Static ICIC

In general, ICIC may be static or semi-static, with different levels of associated communication required between eNodeBs.

For *static interference coordination*, the coordination is associated with cell planning, and reconfigurations are rare. This largely avoids signalling on the X2 interface, but it may result in some performance limitation since it cannot adaptively take into account variations in cell loading and user distributions.

Semi-static interference coordination typically refers to reconfigurations carried out on a time-scale of the order of seconds or longer. The inter-eNodeB communication methods over the X2 interface can be used as discussed above. Other types of information such as traffic load information may also be used, as discussed in Section 2.6.4. Semi-static interference coordination may be more beneficial in cases of non-uniform load distributions in eNodeBs and varying cell sizes across the network.

12.6 Summary

In summary, it can be observed that a variety of resource scheduling algorithms may be applied by the eNodeB depending on the optimization criteria required. The prioritization of data will typically take into account the corresponding traffic classes, especially in regard to balancing throughput maximization for delay-tolerant applications against QoS for delay-limited applications in a fair way.

It can be seen that multi-user diversity is an important factor in all cases, and especially if the user density is high, in which case the multi-user diversity gain enables the scheduler to achieve a high capacity even with tight delay constraints. Finally, it is important to note that individual cells, and even individual eNodeBs, cannot be considered in isolation. System optimization requires some degree of coordination between cells and eNodeBs, in order to avoid inter-cell interference becoming the limiting factor. Considering the system as a whole, the best results are in many cases realized by simple 'on-off' allocation of resource blocks, whereby some eNodeBs avoid scheduling transmissions in certain resource blocks which are used by neighbouring eNodeBs for cell-edge users.

LTE Releases 8 and 9 support standardized signalling between eNodeBs to facilitate frequency-domain data-channel ICIC algorithms for both downlink and uplink. Additional time-domain ICIC mechanisms of particular relevance to the control signalling and to heterogeneous networks are introduced in Release 10 and explained in Section 31.2.

References[7]

[1] K. Knopp and P. A. Humblet, 'Information Capacity and Power Control in Single-cell Multiuser Communications', in *Proc. IEEE International Conference on Communications*, Seattle, WA, 1995.

[2] R. Knopp and P. A. Humblet, 'Multiple-Accessing over Frequency-selective Fading Channels', in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, September 1995.

[3] D. N. C. Tse and S. D. Hanly, 'Multiaccess Fading Channels. I: Polymatroid Structure, Optimal Resource Allocation and Throughput Capacities'. *IEEE Trans. on Information Theory*, Vol. 44, pp. 2796–2815, November 1998.

[4] R. Jain, W. Hawe and D. Chiu, 'A quantitative measure of fairness and discrimination for resource allocation in Shared Computer Systems', Digital Equipment Corporation Technical Report 301, September 1984.

[5] T. M. Cover and J. Thomas. *Elements of Information Theory*. Wiley, New York, 1991.

[6] R. Knopp and P. A. Hamblet, 'Multiuser Diversity', Technical Report, Eurecom Institute, Sophia Antipolis, France, 2002.

[7] G. Caire, R. Müller and R. Knopp, 'Hard Fairness Versus Proportional Fairness in Wireless Communications: The Single-cell Case'. *IEEE Trans. on Information Theory*, Vol. 53, pp. 1366–1385, April 2007.

[8] T. E. Kolding, 'Link and System Performance Aspects of Proportional Fair Scheduling in WCDMA/HSDPA', in *Proc. IEEE Vehicular Technology Conference*, Orlando, FL, October 2003.

[9] B. Wang, K. I. Pedersen, P. E. Kolding and T. E. Mogensen, 'Performance of VoIP on HSDPA', in *Proc. IEEE Vehicular Technology Conference*, Stockholm, May 2005.

[10] V. Vukadinovic and G. Karlsson, 'Video Streaming in 3.5G: On Throughput-Delay Performance of Proportional Fair Scheduling', in *IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems*, Monterey, CA, September 2006.

[11] S. D. Hanly and D. N. C. Tse, 'Multiaccess Fading Channels. II: Delay-limited Capacities'. *IEEE Trans. on Information Theory*, Vol. 44, pp. 2816–2831, November 1998.

[12] I. Toufik, 'Wideband Resource Allocation for Future Cellular Networks'. Ph.D. Thesis, Eurecom Institute, 2006.

[13] I. Toufik and R. Knopp, 'Channel Allocation Algorithms for Multi-carrier Systems', in *Proc. IEEE Vehicular Technology Conference*, Los Angeles, CA, September 2004.

[14] M. J. Neely, E. Modiano and C. E. Rohrs, 'Power Allocation and Routing in Multi-beam Satellites with Time Varying Channels'. *IEEE Trans. on Networking*, Vol. 11, pp. 138–152, February 2003.

[15] M. Realp, A. I. Pérez-Neira and R. Knopp, 'Delay Bounds for Resource Allocation in Wideband Wireless Systems', in *Proc. IEEE International Conference on Communications*, Istanbul, June 2006.

[16] M. Realp, R. Knopp and A. I. Pérez-Neira, 'Resource Allocation in Wideband Wireless Systems' in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Berlin, September 2005.

[17] A. Gjendemsjoe, D. Gesbert, G. Oien and S. Kiani, 'Binary Power Control for Sum Rate Maximization over Multiple Interfering Links'. *IEEE Trans. Wireless Communications*, August 2008.

[18] 3GPP Technical Specification 36.213, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures', www.3gpp.org.

[7] All web sites confirmed 1st March 2011.

case of $M = N$, where the DFT operation cancels the IFFT of the OFDM modulator resulting in the data symbols being transmitted serially in the time domain. However, this simplistic construction would not provide any oversampling or pulse-shape filtering.

As with the time-domain approach, DFT-S-OFDM is capable of generating both localized and distributed transmissions:

- **Localized transmission.** The subcarrier mapping allocates a group of M adjacent subcarriers to a user. $M < N$ results in zeros being appended to the output of the DFT spreader resulting in an upsampled/interpolated version of the original M QAM data symbols at the IFFT output of the OFDM modulator. The transmitted signal is thus similar to a narrowband single carrier with a CP (equivalent to time-domain generation with repetition factor $L = 1$) and 'sinc' pulse-shaping filtering.
- **Distributed transmission.** The subcarrier mapping allocates M equally spaced subcarriers (e.g. every L^{th} subcarrier). $(L - 1)$ zeros are inserted between the M DFT outputs, and additional zeros are appended to either side of the DFT output prior to the IFFT ($ML < N$). As with the localized case, the zeros appended on either side of the DFT output provide upsampling or 'sinc' interpolation, while the zeros inserted between the DFT outputs produce waveform repetition in the time domain. This results in a transmitted signal similar to time-domain IFDMA with repetition factor L and 'sinc' pulse-shaping filtering.

Like the time-domain SC-FDMA signal generation (in Section 14.2.2), orthogonality between different users with different data rate requirements can be achieved by assigning each user a unique set of subcarriers. The CP structure is the same as for the time-domain signal generation, and therefore the same efficient FDE techniques can be employed at the receiver [8, 9].

It is worth noting that, in principle, any unitary matrix can be used instead of the DFT for the spreading operation with similar performance [10]. However, the use of non-DFT spreading would result in increased CM/PAPR since the transmitted signal would no longer have the single carrier characteristic.

14.3 SC-FDMA Design in LTE

Having outlined the key principles of SC-FDMA transmission, we now explain the application of SC-FDMA to the LTE uplink.

14.3.1 Transmit Processing for LTE

Although the frequency-domain generation of SC-FDMA (DFT-S-OFDM) is functionally equivalent to the time-domain SC-FDMA signal generation, each technique requires a slightly different parameterization for efficient signal generation [3]. The pulse-shaping filter used in the time domain SC-FDMA generation approach in practice has a non-zero excess bandwidth, resulting in bandwidth efficiency which is smaller than that achievable with the frequency-domain method with its inherent 'sinc' (zero excess bandwidth) pulse-shaping filter which arises from the zero padding and IFFT operation. The non-zero excess bandwidth pulse-shaping filter in the time-domain generation also requires ramp-up and ramp-down times of 3–4 samples duration, while for DFT-S-OFDM there is no explicit pulse-shaping

filter, resulting in a much shorter ramp time similar to OFDM. However, the pulse-shaping filter in the time-domain generation does provide the benefit of reduced CM by approximately 0.25–0.5 dB compared to DFT-S-OFDM, as shown in Figure 14.4. Thus there is a trade-off between bandwidth efficiency and CM/PAPR reduction between the time- and frequency-domain SC-FDMA generation methods.

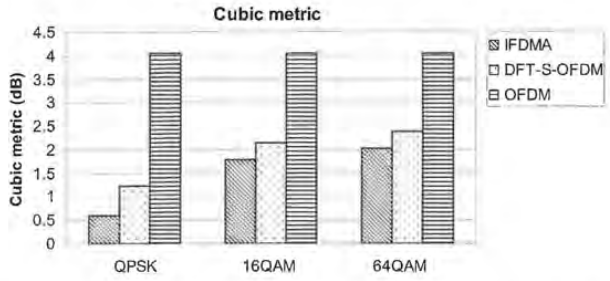


Figure 14.4: Cubic metric comparison of time-domain SC-FDMA generation (IFDMA), frequency-domain SC-FDMA generation (DFT-S-OFDM) and OFDM.

Frequency-domain signal generation for the LTE uplink has a further benefit in that it allows a very similar parameterization to be adopted as for the OFDM downlink, including the same subcarrier spacing, number of occupied subcarriers in a given bandwidth, and CP lengths. This provides maximal commonality between uplink and downlink, including for example the same clock frequency.

For these reasons, the SC-FDMA parameters chosen for the LTE uplink have been optimized under the assumption of frequency-domain DFT-S-OFDM signal generation.

An important feature of the LTE SC-FDMA parameterization is that the numbers of subcarriers which can be allocated to a UE for transmission are restricted such that the DFT size in LTE can be constructed from multiples of 2, 3 and/or 5. This enables efficient, low-complexity mixed-radix FFT implementations.

14.3.2 SC-FDMA Parameters for LTE

The same basic transmission resource structure is used for the uplink as for the downlink: a 10 ms radio frame is divided into ten 1 ms subframes each consisting of two 0.5 ms slots. As LTE SC-FDMA is based on the same fundamental processing as OFDM, it uses the same 15 kHz subcarrier spacing as the downlink. The uplink transmission resources are also defined in the frequency domain (i.e. before the IFFT), with the smallest unit of resource being a Resource Element (RE), consisting of one SC-FDMA data block length on one subcarrier. As in the downlink, a Resource Block (RB) comprises 12 REs in the frequency domain for a duration of 1 slot, as detailed in Section 6.2. The LTE uplink SC-FDMA physical layer parameters for Frequency Division Duplex (FDD) and Time Division Duplex (TDD) deployments are detailed in Table 14.1.

Table 14.1: LTE uplink SC-FDMA physical layer parameters.

Parameter	Value	Comments
Subframe duration	1 ms	
Slot duration	0.5 ms	
Subcarrier spacing	15 kHz	
SC-FDMA symbol duration	66.67 μs	
CP duration	Normal CP:	5.2 μs first symbol in each slot, 4.69 μs all other symbols
	Extended CP:	16.67 μs all symbols
Number of symbols per slot	Normal CP: 7 Extended CP: 6	
Number of subcarriers per RB	12	

Two CP durations are supported – a normal CP of duration 4.69 μs and an extended CP of 16.67 μs, as in the downlink (see Section 5.4.1). The extended CP is beneficial for deployments with large channel delay-spread characteristics, and for large cells.

The 1 ms subframe allows a 1 ms scheduling interval (or Transmission Time Interval (TTI)), as for the downlink, to enable low latency. However, one difference from the downlink is that the uplink coverage is more likely to be limited by the maximum transmission power of the UE. In some situations, this may mean that a single Voice-over-IP (VoIP) packet, for example, cannot be transmitted in a 1 ms subframe with an acceptable error rate. One solution to this is to segment the VoIP packet at higher layers to allow it to be transmitted over several subframes. However, such segmentation results in additional signalling overhead for each segment (including resource allocation signalling and Hybrid ARQ acknowledgement signalling). A more efficient technique for improving uplink VoIP coverage at the cell edge is to use the so-called *TTI bundling*, where a single transport block from the MAC layer is transmitted repeatedly in multiple consecutive subframes, with only one set of signalling messages for the whole transmission. The LTE uplink allows groups of 4 TTIs to be 'bundled' in this way, in addition to the normal 1 ms TTI.

In practice in LTE Releases 8 and 9, all the uplink data transmissions are localized, using contiguous blocks of subcarriers. This simplifies the transmission scheme. Frequency-diversity can still be exploited by means of frequency hopping, which can occur both within one subframe (at the boundary between the two slots) and between subframes. In the case of frequency hopping within a subframe, the channel coding spans the two transmission frequencies, and therefore the frequency-diversity gain is maximized through the channel decoding process. The only instance of distributed transmission in the LTE uplink (using an IFDMA-like structure) is for the Sounding Reference Signals (SRSs) which are transmitted to enable the eNodeB to perform uplink frequency-selective scheduling; these are discussed in Section 15.6.

Like the downlink, the LTE uplink supports scalable system bandwidths from approximately 1.4 MHz up to 20 MHz with the same subcarrier spacing and symbol duration for all bandwidths. The uplink scaling for the bandwidths supported in LTE is shown in

Table 14.2: LTE Uplink SC-FDMA parametrization for selected carrier bandwidths.

	Carrier bandwidth (MHz)					
	1.4	3	5	10	15	20
FFT size	128	256	512	1024	1536	2048
Sampling rate: M/N × 3.84 MHz	1/2	1/1	2/1	4/1	6/1	8/1
Number of subcarriers	72	180	300	600	900	1200
Number of RBs	6	15	25	50	75	100
Bandwidth efficiency (%)	77.1	90	90	90	90	90

Table 14.2. Note that the sampling rates resulting from the indicated FFT sizes are designed to be small rational multiples of the UMTS 3.84 MHz chip rate, for ease of implementation in a multimode UE.

Note that in the OFDM downlink parameter specification, the d.c. subcarrier is unused in order to support direct conversion (zero IF[3]) architectures. In contrast, no unused d.c. subcarrier is possible for SC-FDMA (as shown in Table 14.2) as it can affect the low CM/PAPR property of the transmit signal.

14.3.3 d.c. Subcarrier in SC-FDMA

Direct conversion transmitters and receivers can introduce distortion at the carrier frequency (zero frequency or d.c. in baseband), for example arising from local oscillator leakage [11]. In the LTE downlink, this is addressed by the inclusion of an unused d.c. subcarrier. However, for the uplink, the same solution would cause significant degradation to the low CM property of the transmitted signal. In order to minimize the impact of such distortion on the packet error rate and the CM/PAPR, the d.c. subcarrier of the SC-FDMA signal is modulated in the same way as all the other subcarriers but the subcarriers are all frequency-shifted by half a subcarrier spacing (±7.5 kHz), resulting in an offset of 7.5 kHz relative to d.c. Thus two subcarriers straddle the d.c. location, and hence the amount of distortion affecting any individual RB is halved. This is illustrated in Figure 14.5 for deployments with even and odd numbers of RBs across the system bandwidth.

14.3.4 Pulse Shaping

As explained in Sections 14.2.3 and 14.3.1, one of the benefits of frequency-domain processing for SC-FDMA is that, in principle, there is no need for explicit pulse-shaping thanks to the implicit 'sinc' pulse-shaping. Nevertheless, an additional explicit pulse-shaping filter can further reduce the CM/PAPR, but at the expense of spectral efficiency (due to the resulting non-zero excess filter bandwidth similar to that of the time domain SC-FDMA signal generation described in Section 14.2.2).

The use of pulse-shaping filters, such as RRC[4] [12] or Kaiser window [13], was considered for LTE in order to reduce CM/PAPR, especially for lower-order modulations such as QPSK

[3]Intermediate Frequency.
[4]Roor-Raised-Cosine.

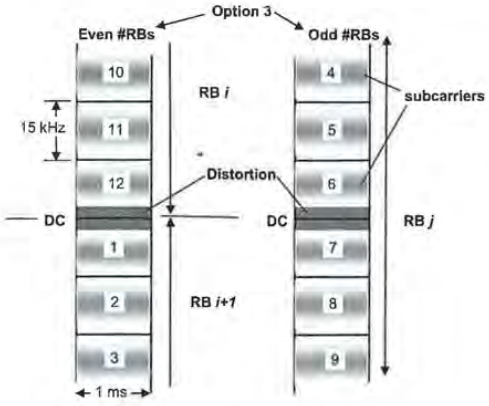


Figure 14.5: 7.5 kHz subcarrier shift for LTE system bandwidths with even or odd numbers of RBs.

to enhance uplink coverage. However, the bandwidth expansion caused by pulse-shaping filtering would require the code rate to be increased for a given data rate, and, in order to provide a net performance benefit, power boosting would be required with the size of the boost being approximately equal to the CM reduction achieved by the pulse-shaping.

Simulation results show that the CM benefits obtained from the use of pulse shaping (RRC or Kaiser window) are small. Moreover, pulse-shaping could not be applied to reference symbols, in order to prevent degradation of the channel estimation.

Finally, as discussed in Chapter 21, CM is not a problem for UEs operating at a maximum power level for resource allocations of up to 10 RBs. This means that there is no benefit possible from pulse-shaping given that UEs at the cell edge would often have less than 10 RBs allocated.

Therefore, no pulse-shaping is used in LTE.

14.4 Summary

The important properties of the SC-FDMA transmission scheme used for the LTE uplink are derived from its multicarrier OFDM-like structure with single-carrier characteristic. The multicarrier-based structure gives the LTE uplink the same robustness against ISI as the LTE downlink, with low-complexity frequency-domain equalization being facilitated by the CP. At the same time, the DFT-based precoding ensures that the LTE uplink possesses the low CM required for efficient UE design. Crucially, the LTE uplink is designed to be orthogonal in the frequency domain between different UEs, thus virtually eliminating the intra-cell interference associated with CDMA.

The parameters of the LTE uplink are designed to ensure maximum commonality with the downlink, and to facilitate frequency-domain DFT-S-OFDM signal generation.

The localized resource allocation scheme of the LTE uplink allows both frequency-selective scheduling and the exploitation of frequency diversity, the latter being achieved by means of frequency hopping.

The following chapters explain the application of SC-FDMA in LTE in more detail, showing how reference signals for channel estimation, data transmissions and control signalling are multiplexed together.

Enhancements to the uplink transmission scheme in Release 10 for LTE-Advanced are described in Section 28.3.6.

References[5]

[1] Motorola, 'R1-050397: EUTRA Uplink Numerology and Frame Structure', www.3gpp.org, 3GPP TSG RAN WG1, meeting 41, Athens, Greece, May 2005.

[2] Motorola, 'R1-050584: EUTRA Uplink Numerology and Design', www.3gpp.org, 3GPP TSG RAN WG1, meeting 41bis, Sophia Antipolis, France, June 2005.

[3] Motorola, 'R1-050971: Single Carrier Uplink Options for E-UTRA', www.3gpp.org, 3GPP TSG RAN WG1, meeting 42, London, UK, August 2005.

[4] U. Sorger, I. De Broeck and M. Schnell, 'Interleaved FDMA – A New Spread-Spectrum Multiple-Access Scheme', in *Proc. IEEE International Conference on Communications*, pp. 1013–1017, 1998.

[5] K. Bruninghaus and H. Rohling, 'Multi-carrier Spread Spectrum and its Relationship to Single-carrier Transmission', in *Proc. IEEE Vehicular Technology Conference*, May 1998.

[6] D. Galda and H. Rohling, 'A Low Complexity Transmitter Structure for OFDM-FDMA Uplink Systems', in *Proc. IEEE Vehicular Technology Conference*, May 2002.

[7] R. Dinis, D. Falconer, C. T. Lam and M. Sabbaghian, 'A Multiple Access Scheme for the Uplink of Broadband Wireless Systems', in *Proc. IEEE Global Telecommunications Conference*, November 2004.

[8] H. Sari, G. Karam and I. Jeanclaude, 'Frequency-Domain Equalization of Mobile Radio and Terrestrial Broadcast Channels', in *Proc. IEEE Global Telecommunications Conference*, November 1994.

[9] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar and B. Eidson, 'Frequency Domain Equalization for Single-carrier Broadband Wireless Systems'. *IEEE Communications Magazine*, Vol. 40, pp. 58–66, April 2002.

[10] V. Nangia and K. L. Baum, 'Experimental Broadband OFDM System – Field Results for OFDM and OFDM with Frequency Domain Spreading', in *Proc. IEEE International Conference on Vehicular Technology*, Vancouver, BC, Canada, September 2002.

[11] Motorola, 'R1-062061: Uplink DC Subcarrier Distortion Considerations in LTE', www.3gpp.org, 3GPP TSG RAN WG1, meeting 46, Tallinn, Estonia, August 2006.

[12] NTT DoCoMo, NEC, and SHARP, 'R1-050702: DFT-spread OFDM with Pulse Shaping Filter in Frequency Domain in Evolved UTRA Uplink', www.3gpp.org, 3GPP TSG RAN WG1, meeting 42, London, UK, August 2005.

[13] Huawei, 'R1-051434: Optimum family of spectrum-shaping functions for PAPR reduction in SC-FDMA', www.3gpp.org, 3GPP TSG RAN WG1, meeting 43, Seoul, Korea, November 2005.

[5] All web sites confirmed 1st March 2011.

consecutive subframes, together with typically three retransmissions at 16 ms intervals, then enables both Chase combining gain and Incremental Redundancy (IR) gain to be achieved.

16.3 Uplink Control Channel Design

In general, uplink control signalling in mobile communications systems can be divided into two categories:

- **Data-associated control signalling** is control signalling which is always transmitted together with uplink data and is used in the processing of that data. Examples include transport format indications, 'New Data' Indicators (NDIs) and MIMO parameters.
- **Control signalling not associated with data** is transmitted independently of any uplink data packet. Examples include HARQ Acknowledgements (ACK/NACK) for downlink data packets, Channel Quality Indicators (CQIs) to support link adaptation (see Section 10.2), and MIMO feedback such as Rank Indicators (RIs) and Precoding Matrix Indicators (PMIs) (see Section 11.2.2.4) for downlink transmissions. Scheduling Requests (SRs) for uplink transmissions also fall into this category (see Section 4.4.2.2).

In LTE, the low signalling latency afforded by the short subframe duration of 1 ms, together with the orthogonal nature of the uplink multiple access scheme which necessitates centralized resource allocation, make it appropriate for the eNodeB to be in full control of the uplink transmission parameters. Consequently uplink data-associated control signalling is not necessary in LTE, as the relevant information is already known to the eNodeB. Therefore only data-non-associated control signalling exists in the LTE uplink.

When simultaneous uplink PUSCH data and control signalling are scheduled, the control signalling is normally multiplexed together with the data prior to the DFT spreading[4], in order to preserve the single-carrier low Cubic Metric (CM) property of the uplink transmission. The uplink control channel, PUCCH, is used by a UE to transmit any necessary control signalling only in subframes in which the UE has not been allocated any RBs for PUSCH transmission. In the design of the PUCCH, special consideration was given to maintaining a low CM [8].

16.3.1 Physical Uplink Control Channel (PUCCH) Structure

The control signalling on the PUCCH is transmitted in a frequency region that is normally configured to be on the edges of the system bandwidth.

In order to minimize the resources needed for transmission of control signalling, the PUCCH in LTE is designed to exploit frequency diversity: each PUCCH transmission in one subframe comprises a single (0.5 ms) RB at or near one edge of the system bandwidth, followed (in the second slot of the subframe) by a second RB at or near the opposite edge of the system bandwidth, as shown in Figure 16.4; together, the two RBs are referred to as a *PUCCH region*. This design can achieve a frequency diversity benefit of approximately 2 dB compared to transmission in the same RB throughout the subframe.

[4]In Release 10, the possibility of simultaneous transmission of PUSCH and PUCCH is introduced; this is explained in Section 28.3.2.

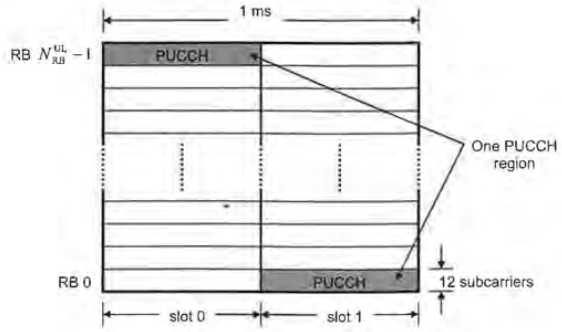


Figure 16.4: A PUCCH region.

At the same time, the narrow bandwidth of the PUCCH in each slot (only a single RB) maximizes the power per subcarrier for a given total transmission power (see Figure 16.5) and therefore helps to fulfil stringent coverage requirements.

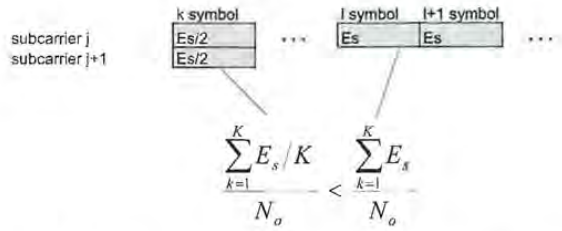


Figure 16.5: The link budget of a two-slot narrowband transmission exceeds that of a one-slot wider-band transmission, given equal coding gain.

Positioning the control regions at the edges of the system bandwidth has a number of advantages, including the following:

- The frequency diversity achieved through frequency hopping is maximized by allowing hopping from one edge of the band to the other.
- Out-Of-Band (OOB) emissions are smaller if a UE is only transmitting on a single RB per slot compared to multiple RBs. The PUCCH regions can serve as a kind of guard band between the wider-bandwidth PUSCH transmissions of adjacent carriers and can therefore improve coexistence [9].
- Using control regions on the band edges maximizes the achievable PUSCH data rate, as the entire central portion of the band can be allocated to a single UE. If the control regions were in the central portion of a carrier, a UE bandwidth allocation would be

limited to one side of the control region in order to maintain the single-carrier nature of the signal, thus limiting the maximum achievable data rate.

- Control regions on the band edges impose fewer constraints on the uplink data scheduling, both with and without inter-/intra-subframe frequency hopping.

The number of RBs (in each slot) that can be used for PUCCH transmission within the cell is $N_{\mathrm{RB}}^{\mathrm{HO}}$ (parameter 'pusch-HoppingOffset'). This is indicated to the UEs in the cell through broadcast signalling. Note that the number of PUCCH RBs per slot is the same as the number of PUCCH regions per subframe. Some typical expected numbers of PUCCH regions for different LTE bandwidths are shown in Table 16.1.

Table 16.1: Typical numbers of PUCCH regions.

Bandwidth (MHz)	Number of RBs per subframe	Number of PUCCH regions
1.4	2	1
3	4	2
5	8	4
10	16	8
20	32	16

Figures 16.6 and 16.7 respectively show examples of even and odd numbers of PUCCH regions being configured in a cell.

In the case of an even number of PUCCH regions (Figure 16.6), both RBs of each RB-pair (e.g. RB-pair 2 and RB-pair $N_{\mathrm{RB}}^{\mathrm{UL}} - 3$) are used for PUCCH transmission. However, for the case of an odd number of PUCCH regions (Figure 16.7), one RB of an RB-pair in each slot is not used for PUCCH (e.g. one RB of RB-pair 2 and RB-pair $N_{\mathrm{RB}}^{\mathrm{UL}} - 3$ is unused); the eNodeB

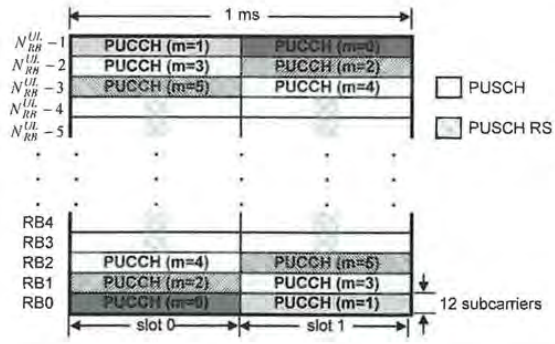


Figure 16.6: PUCCH uplink control structure with an even number of 'PUCCH Control Regions' ($N_{\mathrm{RB}}^{\mathrm{PUCCH}} = 6$).

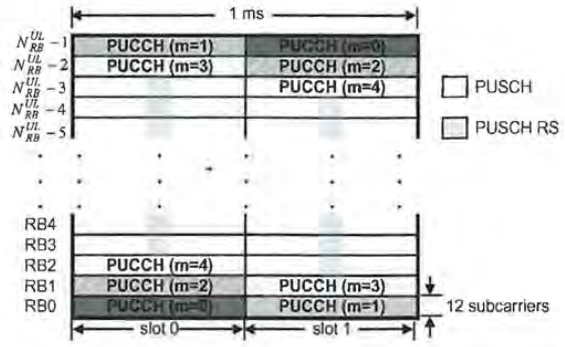


Figure 16.7: Example of an odd number of PUCCH regions ($N_{RB}^{HO} = 5$).

may schedule a UE with an intra-subframe frequency hopping PUSCH allocation to exploit these unused RBs.

Alternatively, a UE can be assigned a localized allocation which includes the unused RB-pair (e.g. RB-pair 2 or RB-pair $N_{RB}^{UL} - 3$). In this case, the UE will transmit PUSCH data on both RBs of the RB-pair, assuming that neither of the RBs are used for PUCCH by any UE in the subframe. Thus, the eNodeB scheduler can appropriately schedule PUSCH transmission on the PUCCH RBs when they are under-utilized.

The eNodeB may also choose to schedule low-power PUSCH transmission (e.g. from UEs close to the eNodeB) in the outer RBs of the configured PUCCH region, while the inner PUCCH region is used for PUCCH signalling. This can provide further reduction in OOB emissions which is necessary in some frequency bands, by moving higher-power PUCCH transmission (e.g. those from cell-edge UEs) slightly away from the edge of the band.

16.3.1.1 Multiplexing of UEs within a PUCCH Region

Control signalling from multiple UEs can be multiplexed into a single PUCCH region using orthogonal Code Division Multiplexing (CDM). In some scenarios this can have benefits over a pure Frequency Division Multiplexing (FDM) approach, as it reduces the need to limit the power differentials between the PUCCH transmissions of different UEs. One technique to provide orthogonality between UEs is by using cyclic time shifts of a sequence with suitable properties, as explained in Section 15.2.2. In a given SC-FDMA symbol, different cyclic time shifts of a waveform (e.g. a Zadoff–Chu (ZC) sequence as explained in Section 7.2.1) are modulated with a UE-specific QAM symbol carrying the necessary control signalling information, with the supported number of cyclic time shifts determining the number of UEs which can be multiplexed per SC-FDMA symbol. As the PUCCH RB spans 12 subcarriers, and assuming the channel is approximately constant over the RB (i.e. a single-tap channel), the LTE PUCCH supports up to 12 cyclic shifts per PUCCH RB.

For control information transmissions with a small number of control signalling bits, such as 1- or 2-bit positive/negative acknowledgements (ACK/NACK), orthogonality is achieved

between UEs by a combination of cyclic time shifts within an SC-FDMA symbol and SC-FDMA symbol time-domain spreading with orthogonal spreading codes, i.e. modulating the SC-FDMA symbols by elements of an orthogonal spreading code [10]. CDM of multiple UEs is used rather than Time Domain Multiplexing (TDM) because CDM enables the time duration of the transmission to be longer, which increases the total transmitted energy per signalling message in the case of a power-limited UE.

Thus, the LTE PUCCH control structure uses frequency-domain code multiplexing (different cyclic time shifts of a base sequence) and/or time-domain code multiplexing (different orthogonal block spreading codes), thereby providing an efficient, orthogonal control channel which supports small payloads (up to 22 coded bits) from multiple UEs simultaneously, together with good operational capability at low SNR.

16.3.2 Types of Control Signalling Information and PUCCH Formats

The Uplink Control Information (UCI) can consist of:

- Scheduling Requests (SRs) (see Section 4.4.2.2).
- HARQ ACK/NACK in response to downlink data packets on the Physical Downlink Shared CHannel(PDSCH); one ACK/NACK bit is transmitted in the case of single-codeword downlink transmission while two ACK/NACK bits are used in the case of two-codeword downlink transmission.
- Channel State Information (CSI), which includes Channel Quality Indicators (CQIs) as well as the MIMO-related feedback consisting of RIs and PMI. 20 bits per subframe are used for the CSI.

The amount of UCI a UE can transmit in a subframe depends on the number of SC-FDMA symbols available for transmission of control signalling data (i.e. excluding SC-FDMA symbols used for RS transmission for coherent detection of the PUCCH). The PUCCH supports seven different formats (with an eighth added in Release 10), depending on the information to be signalled. The mapping between PUCCH formats and UCI is shown in Table 16.2 (see [6, Section 10.1] and [3, Table 5.4-1]).

Table 16.2: Supported uplink control information formats on PUCCH.

PUCCH Format	Uplink Control Information (UCI)
Format 1	Scheduling request (SR) (unmodulated waveform)
Format 1a	1-bit HARQ ACK/NACK with/without SR
Format 1b	2-bit HARQ ACK/NACK with/without SR
Format 2	CSI (20 coded bits)
Format 2	CSI and 1- or 2-bit HARQ ACK/NACK for extended CP only
Format 2a	CSI and 1-bit HARQ ACK/NACK (20 + 1 coded bits)
Format 2b	CSI and 2-bit HARQ ACK/NACK (20 + 2 coded bits)
Format 3	Multiple ACK/NACKs for carrier aggregation: up to 20 ACK/NACK bits plus optional SR, in 48 coded bits; see Section 28.3.2.1 for details.

The physical mapping of the PUCCH formats to the PUCCH regions is shown in Figure 16.8.

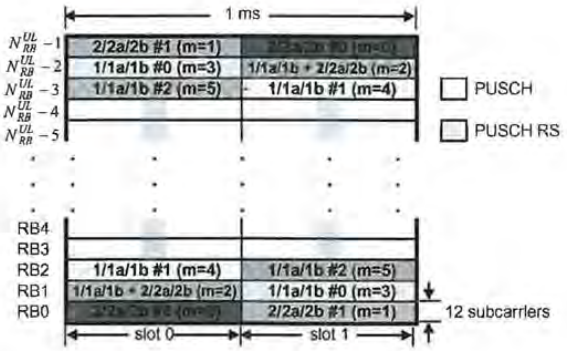


Figure 16.8: Physical mapping of PUCCH formats to PUCCH RBs or regions.

It can be seen that the PUCCH CSI formats 2/2a/2b are mapped and transmitted on the band-edge RBs (e.g. PUCCH region $m = 0, 1$) followed by a mixed PUCCH RB (if present, e.g. region $m = 2$) of CSI format 2/2a/2b and SR/HARQ ACK/NACK format 1/1a/1b, and then by PUCCH SR/HARQ ACK/NACK format 1/1a/1b (e.g. region $m = 4, 5$). The number of PUCCH RBs available for use by CSI format 2/2a/2b, N_{RB}^{2}, is indicated to the UEs in the cell by broadcast signalling.

16.3.3 Channel State Information Transmission on PUCCH (Format 2)

The PUCCH CSI channel structure (Format 2) for one slot with normal CP is shown in Figure 16.9. SC-FDMA symbols 1 and 5 are used for DM-RS transmissions in the case of the normal CP (while in the case of the extended CP only one RS is transmitted, on SC-FDMA symbol 3).

The number of RS symbols per slot results from a trade-off between channel estimation accuracy and the supportable code rate for the UCI bits. For a small number of UCI bits with a low SNR operating point (for a typical 1% target error rate), improving the channel estimation accuracy by using more RS symbols is more beneficial than being able to use a lower channel code rate. However, with larger numbers of UCI bits the required SNR operating point increases, and the higher code rate resulting from a larger overhead of RS symbols becomes more critical, thus favouring fewer RS symbols. In view of these factors, two RS symbols per slot (in case of normal CP) was considered to provide the best trade-off in terms of performance and RS overhead, given the payload sizes required.

10 CSI bits are channel coded with a rate 1/2 punctured (20, k) Reed-Muller code (see [2, Section 5.2.3.3]) to give 20 coded bits, which are then scrambled (in a similar

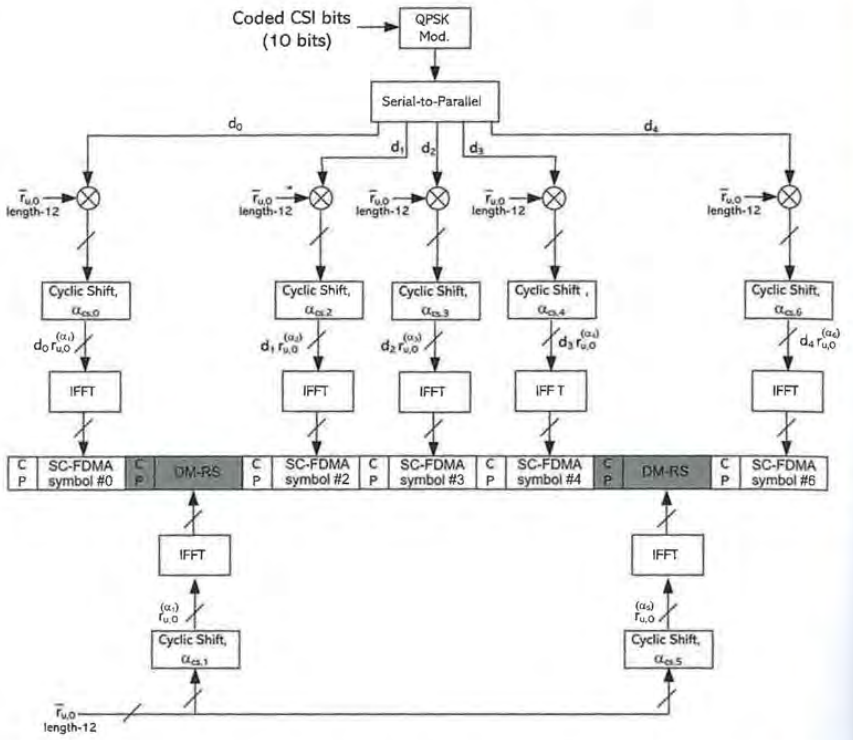
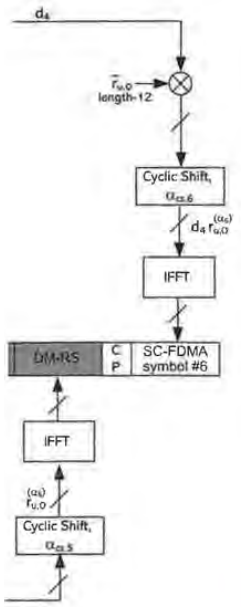


Figure 16.9: CSI channel structure for PUCCH format 2/2a/2b with normal CP for one slot.

way to PUSCH data with a length-31 Gold sequence) prior to QPSK constellation mapping. One QPSK modulated symbol is transmitted on each of the 10 SC-FDMA symbols in the subframe by modulating a cyclic time shift of the base RS sequence of length 12 prior to OFDM modulation. The 12 equally spaced cyclic time shifts allow 12 different UEs to be orthogonally multiplexed on the same PUCCH format 2 RB.

The DM-RS signal sequence (on the 2nd and 6th SC-FDMA symbols for the normal CP and the 4th symbol for the extended CP) is similar to the frequency-domain CSI signal sequence but without the CSI data modulation.

In order to provide inter-cell interference randomization, cell-specific symbol-level cyclic time-shift hopping is used, as described in Section 15.4. For example, the PUCCH cyclic time-shift index on SC-FDMA symbol l in even slots n_s is obtained by adding (modulo-12) a pseudo-random cell-specific PUCCH cyclic shift offset to the assigned cyclic time shift $n_{\mathrm{RS}}^{\mathrm{PUCCH}}$. Intra-cell interference randomization is achieved by cyclic time-shift remapping in the second slot as explained in Section 15.4.



ı normal CP for one slot.

K constellation mapping.
C-FDMA symbols in the
nce of length 12 prior to
w 12 different UEs to be

ols for the normal CP and
nain CSI signal sequence

ecific symbol-level cyclic
mple, the PUCCH cyclic
d by adding (modulo-12)
ssigned cyclic time shift
: time-shift remapping in

A UE is semi-statically configured by higher layer signalling to report periodically different CQI, PMI, and RI types (see Section 10.2.1) on the CSI PUCCH using a PUCCH *resource index* $n_{\text{PUCCH}}^{(2)}$, which indicates both the PUCCH region and the cyclic time shift to be used. The PUCCH region m used for the PUCCH format 2/2a/2b transmission (see Figure 16.8) is given by (see [3, Section 5.4.3])

$$m = \left\lfloor \frac{n_{\text{PUCCH}}^{(2)}}{12} \right\rfloor \tag{16.3}$$

and the assigned cyclic time shift, $n_{\text{RS}}^{\text{PUCCH}}$, is given by

$$n_{\text{RS}}^{\text{PUCCH}} = n_{\text{PUCCH}}^{(2)} \bmod 12. \tag{16.4}$$

16.3.4 Multiplexing of CSI and HARQ ACK/NACK from a UE on PUCCH

The simultaneous transmission of HARQ ACK/NACK and CSI on the PUCCH can be enabled by UE-specific higher layer signalling.[5] If simultaneous transmission is not enabled and the UE needs to transmit HARQ ACK/NACK on the PUCCH in the same subframe in which a CSI report has been configured, the CSI is dropped and only HARQ ACK/NACK is transmitted using the transmission structure detailed in Section 16.3.5.

In subframes where the eNodeB scheduler allows for simultaneous transmission of CSI and HARQ ACK/NACK on the PUCCH from a UE, the CSI and the 1- or 2-bit HARQ ACK/NACK information needs to be multiplexed in the same PUCCH RB, while maintaining the low CM single-carrier property of the signal. The method used to achieve this is different for the cases of normal CP and extended CP, as described in the following sections.

16.3.4.1 Normal CP (Format 2a/2b)

The transmission structure for CSI is the same as described in Section 16.3.3. In order to transmit a 1- or 2-bit HARQ ACK/NACK together with CSI, the HARQ ACK/NACK bits (which are not scrambled) are BPSK/QPSK modulated as shown in Figure 16.10, resulting in a single HARQ ACK/NACK modulation symbol, d_{HARQ}. A positive acknowledgement (ACK) is encoded as a binary '1' and a negative acknowledgement (NACK) is encoded as a binary '0' (see [2, Section 5.2.3.4]).

The single HARQ ACK/NACK modulation symbol, d_{HARQ}, is then used to modulate the second RS symbol (SC-FDMA symbol 5) in each slot – i.e. ACK/NACK is signalled using the RS. This results in PUCCH formats 2a/2b. It can be seen from Figure 16.10 that the modulation mapping is such that a NACK (or NACK, NACK in the case of two downlink MIMO codewords) is mapped to +1, resulting in a default NACK in case neither ACK nor NACK is transmitted (so-called Discontinuous Transmission (DTX)), as happens if the UE fails to detect the downlink grant on the PDCCH. In other words, a DTX (no RS modulation) is interpreted as a NACK by the eNodeB, triggering a downlink retransmission.

As one of the RSs in a slot is modulated by the HARQ ACK/NACK symbol, a variety of ACK/NACK and CSI detection schemes are possible. In low-Doppler environments with

[5] Note that Release 10 also provides the option of configuring simultaneous PUCCH and PUSCH transmission from a UE – see Section 28.3.2.

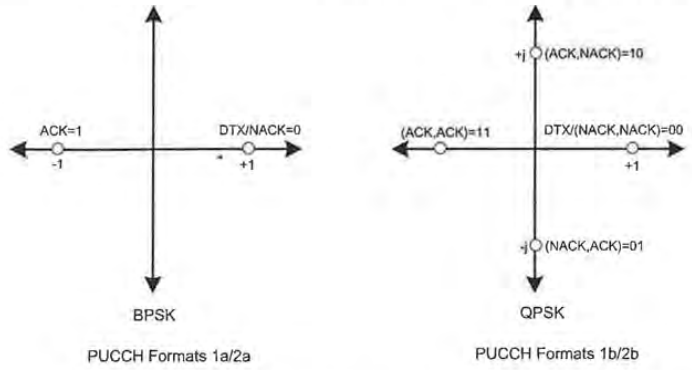


Figure 16.10: Constellation mapping for HARQ ACK/NACK.

little channel variation over the 0.5 ms slot, coherent detection of ACK/NACK and CSI can be achieved by using only the first RS symbol in the slot as the phase reference. Alternatively, to improve the channel estimation quality for CSI detection, an estimate of the HARQ ACK/NACK symbol can be used to undo the modulation on the second RS in the slot so that both RS symbols can be used for channel estimation and demodulation of CSI. In high-Doppler environments in which significant channel variations occur over a slot, relying on a single RS symbol for coherent detection degrades the performance of the ACK/NACK and CSI demodulation. In such cases, blind decoding or multiple hypothesis testing of the different ACK/NACK combinations can be used to decode the ACK/NACK and CSI, selecting the hypothesis that maximizes the correlation between the received signal and the estimated CSI [11] (i.e. a Maximum Likelihood detection).

16.3.4.2 Extended CP (Format 2)

In the case of the extended CP (with one RS symbol per slot), the 1- or 2-bit HARQ ACK/NACK is jointly encoded with the CSI resulting in a $(20, k_{CQI} + k_{ACK/NACK})$ Reed-Muller-based block code. A 20-bit codeword is transmitted on the PUCCH using the channel structure described in Section 16.3.3. The joint coding of the ACK/NACK and CSI is performed as shown in Figure 16.11. The largest number of information bits supported by the block code is 13, corresponding to $k_{CSI} = 11$ CSI bits and $k_{ACK/NACK} = 2$ bits (for two-codeword transmission in the downlink).

16.3.5 HARQ ACK/NACK Transmission on PUCCH (Format 1a/1b)

The PUCCH channel structure for HARQ ACK/NACK transmission with no CSI is shown in Figure 16.12 for one slot with normal CP. Three (two in case of extended CP) SC-FDMA symbols are used in the middle of the slot for RS transmission, with the remaining four SC-FDMA symbols being used for ACK/NACK transmission. Due to the small number of ACK/NACK bits, three RS symbols are used to improve the channel estimation accuracy for a lower SNR operating point than for the CSI structure described in Section 16.3.3.

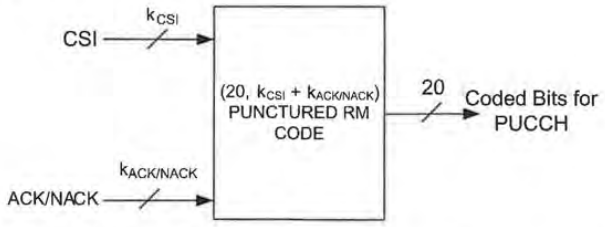


Figure 16.11: Joint coding of HARQ ACK/NACK and CSI for extended CP.

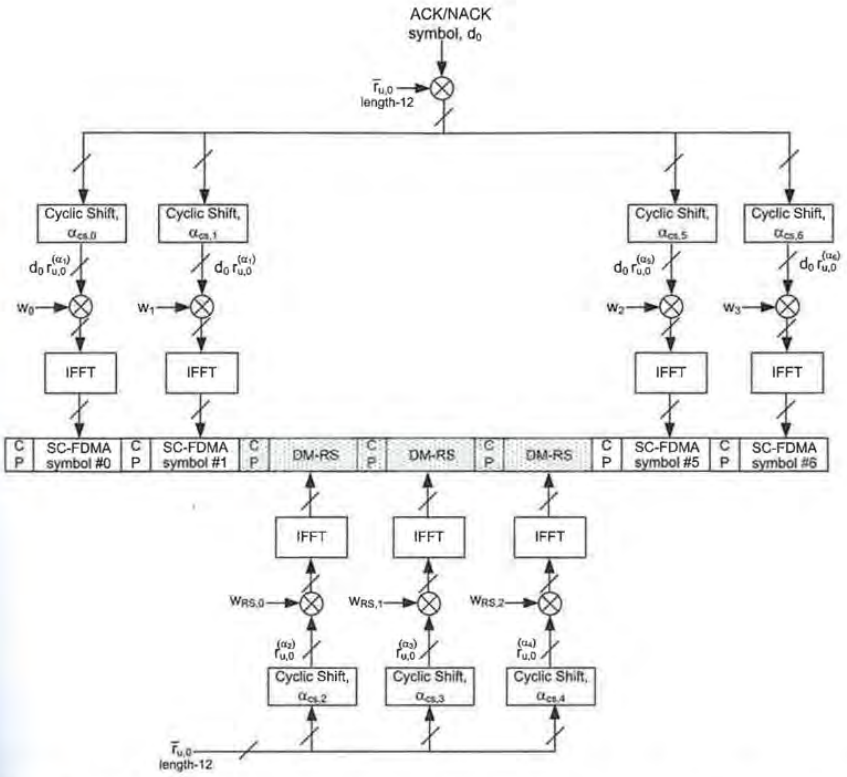


Figure 16.12: ACK/NACK structure – users are multiplexed using different cyclic shifts and time-domain spreading.

Both 1- and 2-bit acknowledgements are supported using BPSK and QPSK modulation respectively. The HARQ ACK/NACK bits are BPSK/QPSK modulated according to the modulation mapping shown in Figure 16.10 (see [3, Table 5.4.1-1]) resulting in a single HARQ ACK/NACK modulation symbol. An ACK is encoded as a binary '1' and a NACK as a binary '0' (see [2, Section 5.2.3.4]). The modulation mapping is the same as the mapping for 1- or 2-bit HARQ ACK/NACK when multiplexed with CSI for PUCCH formats 2a/2b. The modulation symbol is scrambled on a per-slot basis by either 1 or $e^{(-j\pi/2)}$ depending on the PUCCH resource index.

As in the case of CSI transmission, the one BPSK/QPSK modulated symbol (which is phase-rotated by 90 degrees in the second slot) is transmitted on each SC-FDMA data symbol by modulating a cyclic time shift of the base RS sequence of length 12 (i.e. frequency-domain CDM) prior to OFDM modulation. In addition, as mentioned in Section 16.3.1.1, time-domain spreading with orthogonal (Walsh–Hadamard or DFT) spreading codes is used to code-division-multiplex UEs. Thus, a large number of UEs (data and RSs) can be multiplexed on the same PUCCH RB using frequency-domain and time-domain code multiplexing. The RSs from the different UEs are multiplexed in the same way as the data SC-FDMA symbols.

For the cyclic time-shift multiplexing, the number of cyclic time shifts supported in an SC-FDMA symbol for PUCCH HARQ ACK/NACK RBs is configurable by a cell-specific higher-layer signalling parameter $\Delta_{\text{shift}}^{\text{PUCCH}} \in \{1,\ 2,\ 3\}$, indicating 12, 6, or 4 shifts respectively (see [3, Section 5.4.1]). The value selected by the eNodeB for $\Delta_{\text{shift}}^{\text{PUCCH}}$ can be based on the expected delay spread in the cell.

For the time-domain spreading CDM, the length-2 and length-4 orthogonal block spreading codes are based on Walsh–Hadamard codes, and the length-3 spreading codes are based on DFT codes as shown in Table 16.3. The number of supported spreading codes is limited by the number of RS symbols, as the multiplexing capacity of the RSs is smaller than that of the data symbols due to the smaller number of RS symbols. Therefore a subset of size s orthogonal spreading codes of a particular length L ($s \leq L$) is used depending on the number of RS SC-FDMA symbols. For the normal CP with four data SC-FDMA symbols and three supportable orthogonal time spreading codes (due to there being three RS symbols), the indices 0, 1, 2 of the length-4 orthogonal spreading codes are used for the data time-domain block spreading.

Table 16.3: Time-domain orthogonal spreading code sequences.
Reproduced by permission of © 3GPP.

Orthogonal code sequence index	Length 2 Walsh–Hadamard	Length 3 DFT	Length 4 Walsh–Hadamard
0	$[+1 \quad +1]$	$[+1 \quad +1 \quad +1]$	$[+1 \quad +1 \quad +1 \quad +1]$
1	$[+1 \quad -1]$	$[1 \quad e^{j2\pi/3} \quad e^{j4\pi/3}]$	$[+1 \quad -1 \quad +1 \quad -1]$
2	N/A	$[1 \quad e^{j4\pi/3} \quad e^{j2\pi/3}]$	$[+1 \quad -1 \quad -1 \quad +1]$
3	N/A	N/A	$[+1 \quad +1 \quad -1 \quad -1]$

Similarly, for the extended CP case with four data SC-FDMA symbols but only two RS symbols, orthogonal spreading code indices 0 and 2 of length 4 are used for the data block spreading codes. For the length-4 orthogonal codes, the code sequences used are such that subsets of the code sequences result in the minimum inter-code interference in high-Doppler conditions where generally the orthogonality between the code sequences breaks down [12]. Table 16.4 summarizes the time-domain orthogonal spreading code lengths (i.e. spreading factors) for data and RS. The number of supportable orthogonal spreading codes is equal to the number of RS SC-FDMA symbols, N_{RS}^{PUCCH}.

Table 16.4: Spreading factors for time-domain orthogonal spreading codes for data and RS for PUCCH formats 1/1a/1b for normal and extended CP.

	Normal CP		Extended CP	
	Data, N_{SF}^{PUCCH}	RS, N_{RS}^{PUCCH}	Data, N_{SF}^{PUCCH}	RS, N_{RS}^{PUCCH}
Spreading factor	4	3	4	2

It should be noted that it is possible for the transmission of HARQ ACK/NACK and SRS to be configured in the same subframe. If this occurs, the eNodeB can also configure (by cell-specific broadcast signalling) the way in which these transmissions are to be handled by the UE. One option is for the ACK/NACK to take precedence over the SRS, such that the SRS is not transmitted and only HARQ ACK/NACK is transmitted in the relevant subframe, according to the PUCCH ACK/NACK structure in Figure 16.12. The alternative is for the eNodeB to configure the UEs to use a shortened PUCCH transmission in such subframes, whereby the last SC-FDMA symbol of the ACK/NACK (i.e. the last SC-FDMA symbol in the second slot of the subframe is not transmitted; this is shown in Figure 16.13).

This maintains the low CM single-carrier property of the transmitted signal, by ensuring that a UE never needs to transmit both HARQ ACK/NACK and SRS symbols simultaneously, even if both signals are configured in the same subframe. If the last symbol of the ACK/NACK is not transmitted in the second slot of the subframe, this is known as a *shortened* PUCCH format, as shown in Figure 16.14.[6] For the shortened PUCCH, the length of the time-domain orthogonal block spreading code is reduced by one (compared to the first slot shown in Figure 16.12). Hence, it uses the length-3 DFT basis spreading codes in Table 16.3 in place of the length-4 Walsh–Hadamard codes.

The frequency-domain HARQ ACK/NACK signal sequence on SC-FDMA symbol n is defined in [3, Section 5.4.1].

The number of HARQ ACK/NACK resource indices $N_{PUCCH,RB}^{(1)}$ corresponding to cyclic-time-shift/orthogonal-code combinations that can be supported in a PUCCH RB is given by

$$N_{PUCCH,RB}^{(1)} = c \cdot P, \quad c = \begin{cases} 3 & \text{normal cyclic prefix} \\ 2 & \text{extended cyclic prefix} \end{cases} \tag{16.5}$$

[6] Note that configuration of SRS in the same subframe as channel quality information or SR is not valid. Therefore the shortened PUCCH formats are only applicable for PUCCH formats 1a and 1b.

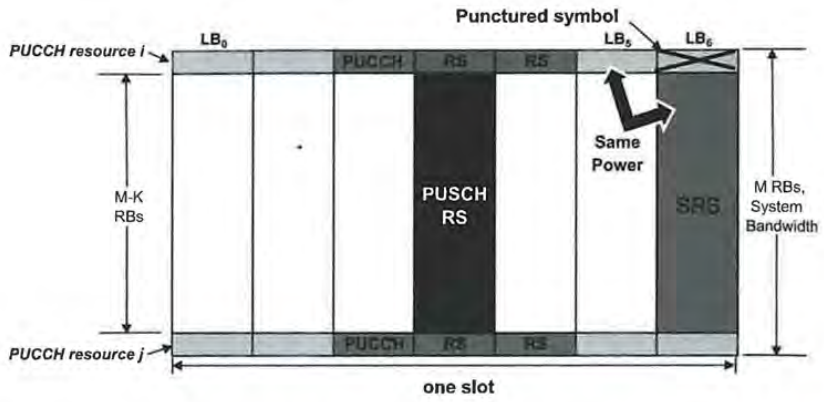


Figure 16.13: A UE may not simultaneously transmit on SRS and PUCCH or PUSCH, in order to avoid violating the single-carrier nature of the signal. Therefore, a PUCCH or PUSCH symbol may be punctured if SRS is transmitted.

where c is equal to the number of RS symbols, $P = 12/\Delta_{\text{shift}}^{\text{PUCCH}}$, and $\Delta_{\text{shift}}^{\text{PUCCH}} \in \{1, 2, 3\}$ is the number of equally spaced cyclic time shifts supported. For example, with the normal CP and $\Delta_{\text{shift}}^{\text{PUCCH}} = 2$, ACK/NACKs from 18 different UEs can be multiplexed in one RB.

As in the case of CSI (see Section 16.3.3), cyclic time shift hopping (described in Section 15.4) is used to provide inter-cell interference randomization.

In the case of semi-persistently scheduled downlink data transmissions on the PDSCH (see Section 4.4.2.1) without a corresponding downlink grant on the control channel PDCCH, the PUCCH ACK/NACK resource index $n_{\text{PUCCH}}^{(1)}$ to be used by a UE is semi-statically configured by higher layer signalling. This PUCCH ACK/NACK resource is used for ACK/NACK transmission corresponding to initial HARQ transmission. For dynamically-scheduled downlink data transmissions (including HARQ retransmissions for semi-persistent data) on the PDSCH (indicated by downlink assignment signalling on the PDCCH), the PUCCH HARQ ACK/NACK resource index $n_{\text{PUCCH}}^{(1)}$ is implicitly determined based on the index of the first Control Channel Element (CCE, see Section 9.3) of the PDCCH message.

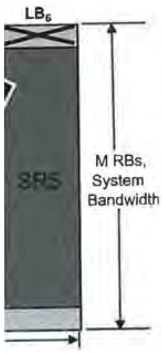
The PUCCH region m used for the HARQ ACK/NACK with format 1/1a/1b transmission for the case with no mixed PUCCH region (shown in Figure 16.8), is given by

$$m = \left\lfloor \frac{n_{\text{PUCCH}}^{(1)}}{N_{\text{PUCCH, RB}}^{(1)}} \right\rfloor + N_{\text{RB}}^{(2)} \tag{16.6}$$

where $N_{\text{RB}}^{(2)}$ is the number of RBs that are available for PUCCH formats 2/2a/2b and is a cell-specific broadcast parameter (see [3, Section 5.4.3]).

The PUCCH resource index $n^{(1)}(n_s)$, corresponding to a combination of a cyclic time shift and orthogonal code ($n_{\text{RS}}^{\text{PUCCH}}$ and n_{oc}), within the PUCCH region m in even slots is given by

$$n^{(1)}(n_s) = n_{\text{PUCCH}}^{(1)} \mod N_{\text{PUCCH, RB}}^{(1)} \quad \text{for } n_s \bmod 2 = 0 \tag{16.7}$$



CH or PUSCH, in
re, a PUCCH or

UCCH ∈ {1, 2, 3} is
hift
with the normal CP
in one RB.
ping (described in

on the PDSCH (see
l channel PDCCH,
3 is semi-statically
source is used for
. For dynamically-
for semi-persistent
the PDCCH), the
nined based on the
PDCCH message.
/1a/1b transmission
en by

(16.6)

its 2/2a/2b and is a

of a cyclic time shift
ven slots is given by

) (16.7)

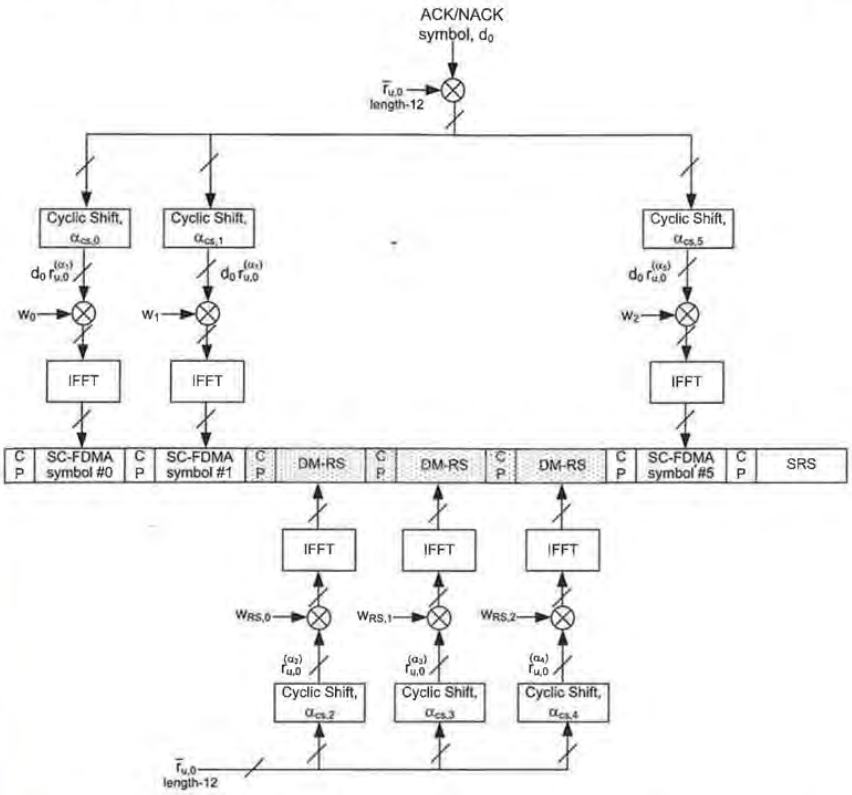


Figure 16.14: Shortened PUCCH ACK/NACK structure when simultaneous SRS and ACK/NACK is enabled in the cell.

The PUCCH resource index (n_{RS}^{PUCCH}, n_{oc}) allocation within a PUCCH RB format 1/1a/1b, is shown in Tables 16.5, 16.6 and 16.7, for $\Delta_{shift}^{PUCCH} \in \{1, 2, 3\}$ with 36, 18, and 12 resource indices respectively for the normal CP case [13]. For the extended CP, with two time-domain orthogonal spreading code sequences, only the first two columns of the orthogonal code sequence index, $n_{oc} = 1, 2$ are used, resulting in 24, 12 and 8 resource indices for $\Delta_{shift}^{PUCCH} \in \{1, 2, 3\}$ respectively.

The PUCCH resources are first indexed in the cyclic time-shift domain, followed by the orthogonal time spreading code domain.

The cyclic time shifts used on *adjacent* orthogonal codes can also be staggered, providing the opportunity to separate the channel estimates prior to de-spreading. As high-Doppler breaks down the orthogonality between the spread blocks, offsetting the cyclic time shift

Table 16.5: PUCCH RB format 1/1a/1b resource index allocation, $\Delta_{\text{shift}}^{\text{PUCCH}} = 1$, 36 resource indices, normal CP.

	Orthogonal code sequence index, n_{oc}		
Cyclic shift index, n_{RS}^{PUCCH}	$n_{oc} = 0$	$n_{oc} = 1$	$n_{oc} = 2$
0	0	12	24
1	1	13	25
2	2	14	26
3	3	15	27
4	4	16	28
5	5	17	29
6	6	18	30
7	7	19	31
8	8	20	32
9	9	21	33
10	10	22	34
11	11	23	35

Table 16.6: PUCCH RB format 1/1a/1b resource index allocation, $\Delta_{\text{shift}}^{\text{PUCCH}} = 2$, 18 resource indices, normal CP.

	Orthogonal code sequence index, n_{oc}		
Cyclic shift index, n_{RS}^{PUCCH}	$n_{oc} = 0$	$n_{oc} = 1$	$n_{oc} = 2$
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	

values within each SC-FDMA symbol can restore orthogonality at moderate delay spreads. This can enhance the tracking of high-Doppler channels [14].

In order to randomize intra-cell interference, PUCCH resource index remapping is used in the second slot [15]. Index remapping includes both cyclic shift remapping and orthogonal block spreading code remapping (similar to the case of CSI – see Section 16.3.3).

The PUCCH resource index remapping function in an odd slot is based on the PUCCH resource index in the even slot of the subframe, as defined in [3, Section 5.4.1].

Table 16.7: PUCCH RB format 1/1a/1b resource index allocation, $\Delta_{\text{shift}}^{\text{PUCCH}} = 3$, 12 resource indices, normal CP.

	Orthogonal code sequence index, n_{oc}		
Cyclic shift index, n_{RS}^{PUCCH}	$n_{oc} = 0$	$n_{oc} = 1$	$n_{oc} = 2$
0	0		
1		4	
2			7
3	1		
4		5	
5			8
6	2		
7			
8			
9	3		
10		6	
11			9

16.3.6 Multiplexing of CSI and HARQ ACK/NACK in the Same (Mixed) PUCCH RB

The multiplexing of CSI and HARQ ACK/NACK in different PUCCH RBs can in general simplify the system. However, in the case of small system bandwidths, such as 1.4 MHz, the control signalling overhead can become undesirably high with separate CSI and ACK/NACK RB allocations (two out of a total of 6 RBs for control signalling in 1.4 MHz system bandwidths). Therefore, multiplexing of CSI and ACK/NACK from different UEs in the same mixed PUCCH RB is supported in LTE to reduce the total control signalling overhead.

The ZC cyclic time-shift structure facilitates the orthogonal multiplexing of channel quality information and ACK/NACK signals with different numbers of RS symbols. This is achieved by assigning different sets of adjacent cyclic time shifts to CSI and ACK/NACK signals [16] as shown in Table 16.8. As can be seen from this table, $N_{cs}^{(1)} \in \{0, 1, \ldots, 7\}$ cyclic time shifts are used for PUCCH ACK/NACK formats 1/1a/1b in the mixed PUCCH RB case, where $N_{cs}^{(1)}$ is a cell-specific broadcast parameter (see [3, Section 5.4]) restricted to integer multiples of $\Delta_{\text{shift}}^{\text{PUCCH}}$. A guard cyclic time shift is used between the ACK/NACK and CSI cyclic shift resources to improve orthogonality and channel separation between UEs transmitting CSI and those transmitting ACK/NACK. To avoid mixing of the cyclic time shifts for ACK/NACK and CSI, the cyclic time shift (i.e. the PUCCH resource index remapping function) for the odd slot of the subframe is not used; the same cyclic time shift as in the first slot of the subframe is used.

16.3.7 Scheduling Request (SR) Transmission on PUCCH (Format 1)

The structure of the SR PUCCH format 1 is the same as that of the ACK/NACK PUCCH format 1a/1b explained in Section 16.3.5, where a cyclic time shift of the base RS sequence is modulated with time-domain orthogonal block spreading. The SR uses simple On–Off

Table 16.8: Multiplexing of ACK/NACK (format 1/1a/1b) and CSI (format 2/2a/2b) from different UEs in the same (mixed) PUCCH RB by using different sets of cyclic time shifts.

Cyclic shift index	Cyclic shift index allocation
0	Format 1/1a/1b (HARQ ACK/NACK, SR) cyclic shifts
1	
2	
⋮	
$N_{cs}^{(1)}$	
$N_{cs}^{(1)} + 1$	Guard cyclic shift
$N_{cs}^{(1)} + 2$	Format 2/2a/2b (CSI) cyclic shifts
⋮	
10	
11	Guard cyclic shift

keying, with the UE transmitting an SR using the modulation symbol $d(0) = +1$ (i.e. the same constellation point as is used for DTX/NACK or DTX/(NACK,NACK) for ACK/NACK transmission using PUCCH format 1a/1b – see Figure 16.10) to request a PUSCH resource (positive SR transmission), and transmitting nothing when it does not request to be scheduled (negative SR).

Since the HARQ ACK/NACK structure is reused for the SR, different PUCCH resource indices (i.e. different cyclic time shift/orthogonal code combinations) in the same PUCCH region can be assigned for SR (Format 1) or HARQ ACK/NACK (Format 1a/1b) from different UEs. This results in orthogonal multiplexing of SR and HARQ ACK/NACK in the same PUCCH region. The PUCCH resource index to be used by a UE for SR transmission, $n_{PUCCH,SRI}^{(1)}$, is configured by UE-specific higher-layer signalling.

If a UE needs to transmit a positive SR in the same subframe as a scheduled CSI transmission, the CSI is dropped and only the SR is transmitted, in order to maintain the low CM of the transmit signal. In the case of SRS coinciding with an ACK/NACK or positive SR in the same subframe, the UE drops its SRS transmission if the parameter 'ackNackSRS-SimultaneousTransmission' is set to 'FALSE' and transmits the SRS otherwise (see [6, Section 8.2]).

If an SR and ACK/NACK happen to coincide in the same subframe, the UE transmits the ACK/NACK on the assigned SR PUCCH resource for a positive SR and transmits ACK/NACK on its assigned ACK/NACK PUCCH resource in case of a negative SR (see [6, Section 7.3]).

The expected behaviour in some other cases of different types of UCI and SRS coinciding in the same subframe can be found in [17].

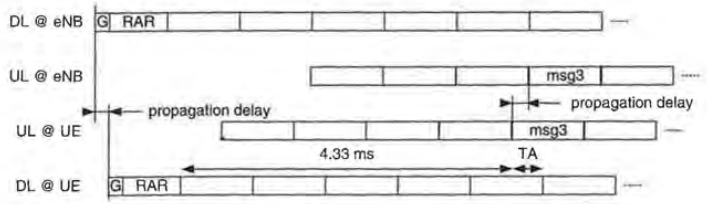


Figure 17.3: Timing of the message 3 transmission.

start another one. The UE's behaviour upon reception of the contention resolution message therefore has three possibilities:

- The UE correctly decodes the message and detects its own identity: it sends back a positive ACKnowledgement, 'ACK'.
- The UE correctly decodes the message and discovers that it contains another UE's identity (contention resolution): it sends nothing back (Discontinuous Transmission, 'DTX').
- The UE fails to decode the message or misses the DL grant: it sends nothing back ('DTX').

17.3.2 Contention-Free Random Access Procedure

The slightly unpredictable latency of the random access procedure can be circumvented for some use cases where low latency is required, such as handover and resumption of downlink traffic for a UE, by allocating a dedicated signature to the UE on a per-need basis. In this case the procedure is simplified as shown in Figure 17.4. The procedure terminates with the RAR.

17.4 Physical Random Access Channel Design

The random access preamble part of the random access procedure is mapped at the physical layer onto the PRACH. The design of the preamble is crucial to the success of the RACH procedure, and therefore the next section focuses on the details of the PRACH design.

17.4.1 Multiplexing of PRACH with PUSCH and PUCCH

The PRACH is time- and frequency-multiplexed with PUSCH and PUCCH as illustrated in Figure 17.5. PRACH time-frequency resources are semi-statically allocated within the PUSCH region, and repeat periodically. The possibility of scheduling PUSCH transmissions within PRACH slots is left to the eNodeB's discretion.

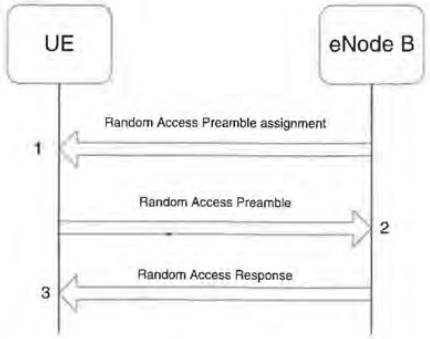


Figure 17.4: Contention-free random access procedure. Reproduced by permission of © 3GPP.

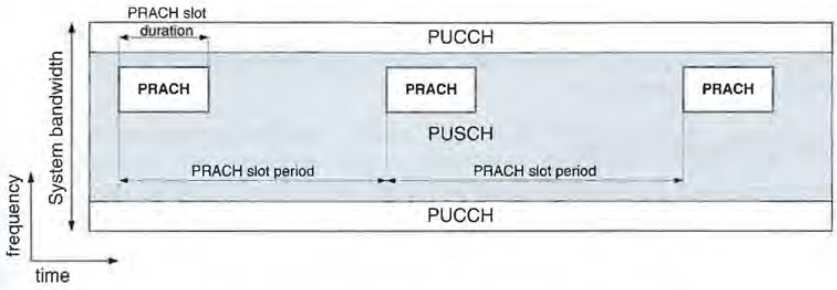


Figure 17.5: PRACH multiplexing with PUSCH and PUCCH.

17.4.2 The PRACH Structure

17.4.2.1 DFT-S-OFDM PRACH Preamble Symbol

Similarly to WCDMA, the LTE PRACH preamble consists of a complex sequence. However, it differs from the WCDMA preamble in that it is also an OFDM symbol following the DFT-S-OFDM structure of the LTE uplink, built with a CP, thus allowing for an efficient frequency-domain receiver at the eNodeB. As shown in Figure 17.6, the end of the sequence is appended at the start of the preamble, thus allowing a periodic correlation at the PRACH receiver (as opposed to a less efficient aperiodic correlation) [1].

The UE aligns the start of the random access preamble with the start of the corresponding uplink subframe at the UE assuming a timing advance of zero (see Section 18.2), and the preamble length is shorter than the PRACH slot in order to provide room for a Guard Time (GT) to absorb the propagation delay. Figure 17.6 shows two preambles at the eNodeB received with different timings depending on the propagation delay: as for a conventional

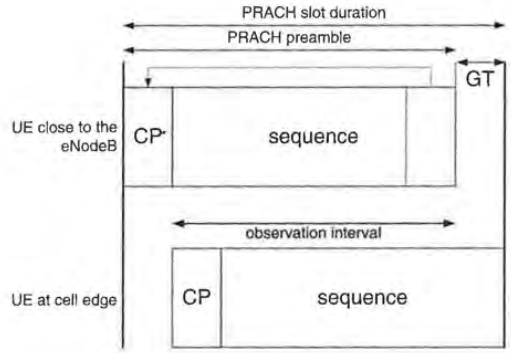


Figure 17.6: PRACH preamble received at the eNodeB.

OFDM symbol, a single observation interval can be used regardless of the UE's delay, within which periodic correlation is possible.

As further elaborated in Section 17.4.3, the LTE PRACH preamble sequence is optimized with respect to its periodic autocorrelation property. The dimensioning of the CP and GT is addressed in Section 17.4.2.4.

17.4.2.2 PRACH Formats

Four Random Access preamble formats are defined for Frequency Division Duplex (FDD) operation [2]. Each format is defined by the durations of the sequence and its CP, as listed in Table 17.1. The format configured in a cell is broadcast in the System Information.

Table 17.1: Random access preamble formats. Reproduced by permission of © 3GPP.

Preamble format	T_{CP} (μs)	T_{SEQ} (μs)	Typical usage
0	103.13	800	Normal 1 ms random access burst with 800 μs preamble sequence, for small to medium cells (up to ~14 km)
1	684.38	800	2 ms random access burst with 800 μs preamble sequence, for large cells (up to ~77 km) without a link budget problem
2	203.13	1600	2 ms random access burst with 1600 μs preamble sequence, for medium cells (up to ~29 km) supporting low data rates
3	684.38	1600	3 ms random access burst with 1600 μs preamble sequence, for very large cells (up to ~100 km)

The rationale behind these choices of sequence duration, CP and GT, are discussed below.

17.4.2.3 Sequence Duration

The sequence duration, T_{SEQ}, is driven by the following factors:

- Trade-off between sequence length and overhead: a single sequence must be as long as possible to maximize the number of orthogonal preambles (see Section 7.2.1), while still fitting within a single subframe in order to keep the PRACH overhead small in most deployments;
- Compatibility with the maximum expected round-trip delay;
- Compatibility between PRACH and PUSCH subcarrier spacings;
- Coverage performance.

Maximum round-trip time. The lower bound for T_{SEQ} must allow for unambiguous round-trip time estimation for a UE located at the edge of the largest expected cell (i.e. 100 km radius), including the maximum delay spread expected in such large cells, namely 16.67 μs. Hence

$$T_{SEQ} \geq \frac{200 \cdot 10^3}{3 \cdot 10^8} + 16.67 \cdot 10^{-6} = 683.33\ \mu s \tag{17.1}$$

Subcarrier spacing compatibility. Further constraints on T_{SEQ} are given by the Single-Carrier Frequency Division Multiple Access (SC-FDMA – see Chapter 14) signal generation principle (see Section 17.5), such that the size of the DFT and IDFT[3], N_{DFT}, must be an integer number:

$$N_{DFT} = f_s T_{SEQ} = k, \quad k \in \mathbb{N} \tag{17.2}$$

where f_s is the system sampling rate (e.g. 30.72 MHz). Additionally, it is desirable to minimize the orthogonality loss in the frequency domain between the preamble subcarriers and the subcarriers of the surrounding uplink data transmissions. This is achieved if the PUSCH data symbol subcarrier spacing Δf is an integer multiple of the PRACH subcarrier spacing Δf_{RA}:

$$\Delta f_{RA} = \frac{f_s}{N_{DFT}} = \frac{1}{T_{SEQ}} = \frac{1}{kT_{SYM}} = \frac{1}{k}\Delta_f, \quad k \in \mathbb{N} \tag{17.3}$$

where $T_{SYM} = 66.67$ μs is the uplink subframe symbol duration. In other words, the preamble duration must be an integer multiple of the uplink subframe symbol duration:

$$T_{SEQ} = kT_{SYM} = \frac{k}{\Delta f}, \quad k \in \mathbb{N} \tag{17.4}$$

An additional benefit of this property is the possibility to reuse the FFT/IFFT[4] components from the SC-FDMA signal processing for the PUSCH to implement the large DFT/IDFT blocks involved in the PRACH transmitter and receiver (see Section 17.5).

[3] Discrete Fourier Transform and Inverse Discrete Fourier Transform.
[4] Fast Fourier Transform / Inverse Fast Fourier Transform.

Coverage performance. In general a longer sequence gives better coverage, but better coverage requires a longer CP and GT in order to absorb the corresponding round-trip delay (Figure 17.6). The required CP and GT lengths for PRACH format 0, for example, can therefore be estimated from the maximum round-trip delay achievable by a preamble sequence which can fit into a 1 ms subframe.

Under a noise-limited scenario, as is typical of a low density, medium to large suburban or rural cell, coverage performance can be estimated from a link budget calculation. Under the assumption of the Okumura-Hata empirical model of distance-dependent path-loss $L(r)$ [3,4] (where r is the cell radius in km), the PRACH signal power P_{RA} received at the eNodeB baseband input can be computed as follows:

$$P_{\mathrm{RA}}(r) = P_{\max} + G_a - L(r) - \mathrm{LF} - P_L \ (\mathrm{dB}) \tag{17.5}$$

where the parameters are listed in Table 17.2 (mainly from [5]).

Table 17.2: Link budget parameters for analysis of PRACH preamble coverage.

Parameter	Value
Carrier frequency (f)	2000 MHz
eNodeB antenna height (h_b)	30 m / 60 m
UE antenna height (h_m)	1.5 m
UE transmitter EIRP[a] ($P_{\max}$)	24 dBm (250 mW)
eNodeB Receiver Antenna Gain (including cable loss) (G_a)	14 dBi
Receiver noise figure (N_f)	5.0 dB
Thermal noise density (N_0)	−174 dBm/Hz
Percentage of the area covered by buildings (α)	10%
Required E_p/N_0 (eNodeB with 2 Rx antenna diversity)	18 dB (six-path Typical Urban channel model)
Penetration loss (P_L)	0 dB
Log-normal fade margin (LF)	0 dB

[a]Equivalent Isotropic Radiated Power.

The required PRACH preamble sequence duration T_{SEQ} is then derived from the required preamble sequence energy to thermal noise ratio E_p/N_0 to meet a target missed detection and false alarm probability, as follows:

$$T_{\mathrm{SEQ}} = \frac{N_0 N_f}{P_{\mathrm{RA}}(r)} \frac{E_p}{N_0} \tag{17.6}$$

where N_0 is the thermal noise power density (in mW/Hz) and N_f is the receiver noise figure (in linear scale).

Assuming that $E_p/N_0 = 18$ dB is required to meet missed detection and false alarm probabilities of 10^{-2} and 10^{-3} respectively (see Section 17.4.3.3), Figure 17.7 plots the coverage performance of the PRACH sequence as a function of the sequence length T_{SEQ}.

It can be observed from Figure 17.7 that the potential coverage performance of a 1 ms PRACH preamble is in the region of 14 km. As a consequence, the required CP and GT

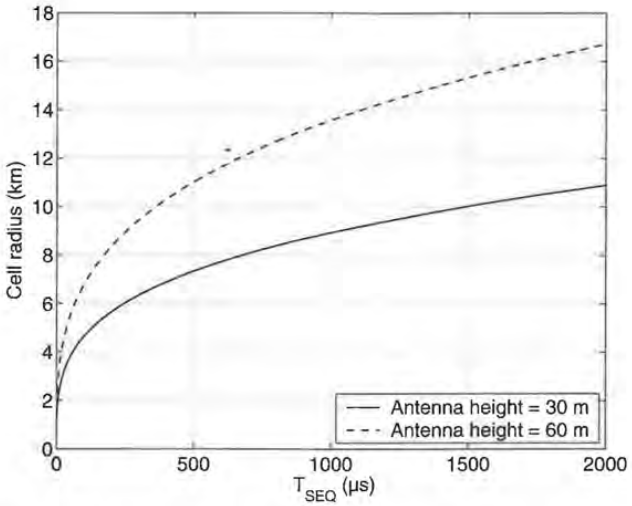


Figure 17.7: PRACH coverage performance versus sequence duration.

lengths are approximately $(2 \cdot 14000)/(3 \cdot 10^8) = 93.3$ μs, so that the upper bound for T_{SEQ} is given by

$$T_{SEQ} \leq 1000 - 2 \cdot 93.33 = 813 \text{ μs} \tag{17.7}$$

Therefore, the longest sequence simultaneously satisfying Equations (17.1), (17.4) and (17.7) is $T_{SEQ} = 800$ μs, as used for preamble formats 0 and 1. The resulting PRACH sub-carrier spacing is $\Delta f_{RA} = 1/T_{SEQ} = 1.25$ kHz.

The 1600 μs preamble sequence of formats 2 and 3 is implemented by repeating the baseline 800 μs preamble sequence. These formats can provide up to 3 dB link budget improvement, which is useful in large cells and/or to balance PUSCH/PUCCH and PRACH coverage at low data rates.

17.4.2.4 CP and GT Duration

Having chosen T_{SEQ}, the CP and GT dimensioning can be specified more precisely.

For formats 0 and 2, the CP is dimensioned to maximize the coverage, given a maximum delay spread d: $T_{CP} = (1000 - 800)/2 + d/2$ μs, with $d \approx 5.2$ μs (corresponding to the longest normal CP of a PUSCH SC-FDMA symbol). The maximum delay spread is used as a guard period at the end of CP, thus providing protection against multipath interference even for the cell-edge UEs.

In addition, for a cell-edge UE, the delay spread energy at the end of the preamble is replicated at the end of the CP (see Figure 17.8) and is therefore within the observation interval. Consequently, there is no need to include the maximum delay spread in the GT dimensioning. Hence, instead of locating the sequence in the centre of the PRACH slot, it is shifted later by half the maximum delay spread, allowing the maximum Round-Trip Delay

(RTD) to be increased by the same amount. Note that, as for a regular OFDM symbol, the residual delay spread at the end of the preamble from a cell-edge UE spills over into the next subframe, but this is taken care of by the CP at the start of the next subframe to avoid any Inter-Symbol Interference (ISI).

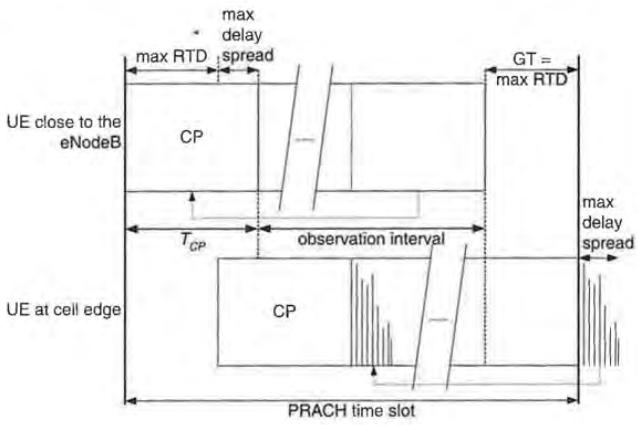


Figure 17.8: PRACH CP/GT dimensioning for formats 0 and 2.

For formats 1 and 3, the CP is dimensioned to address the maximum cell range in LTE, 100 km, with a maximum delay spread of $d \approx 16.67$ μs. In practice, format 1 is expected to be used with a 3-subframe PRACH slot; the available GT in 2 subframes can only address a 77 km cell range. It was chosen to use the same CP length for both format 1 and format 3 for implementation simplicity. Of course, handling larger cell sizes than 100 km with suboptimal CP dimensioning is still possible and is left to implementation.

Table 17.3 shows the resulting cell radius and delay spread ranges associated with the four PRACH preamble formats. The CP lengths are designed to be integer multiples of the assumed system sampling period for LTE, $T_S = 1/30.72$ μs.

Table 17.3: Field durations and achievable cell radius of the PRACH preamble formats.

Preamble format	Number of allocated subframes	CP duration		GT duration		Max. delay spread (μs)	Max. cell radius (km)
		in μs	as multiple of T_S	in μs	as multiple of T_S		
0	1	103.13	3168	96.88	2976	6.25	14.53
1	2	684.38	21024	515.63	15840	16.67	77.34
2	2	203.13	6240	196.88	6048	6.25	29.53
3	3	684.38	21024	715.63	21984	16.67	100.16

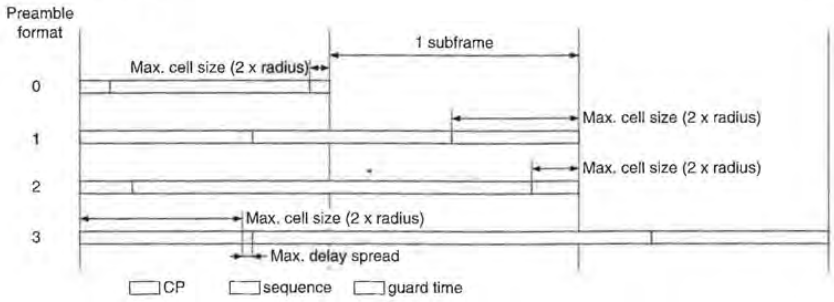


Figure 17.9: PRACH preamble formats and cell size dimensioning.

These are also illustrated in Figure 17.9.

17.4.2.5 PRACH Resource Configurations

The PRACH slots shown in Figure 17.5 can be configured to occur in up to 16 different layouts, or *resource configurations*. Depending on the RACH load, one or more PRACH resources may need to be allocated per PRACH slot period. The eNodeB has to process the PRACH very quickly so that message 2 of the RACH procedure can be sent within the required window, as shown in Figure 17.2. In case of more than one PRACH resource per PRACH period it is generally preferable to multiplex the PRACH resources in time (Figure 17.10 – right) rather than in frequency (Figure 17.10 – left). This helps to avoid processing peaks at the eNodeB.

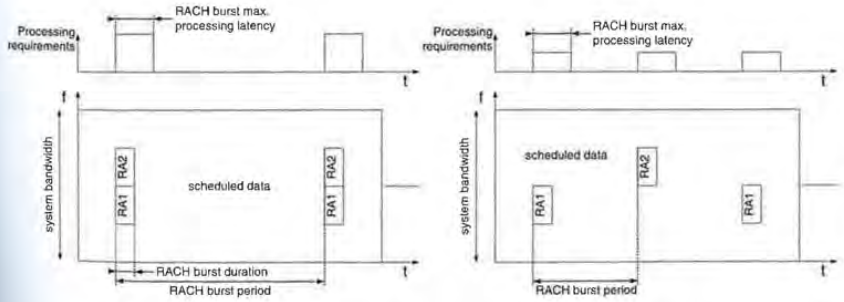


Figure 17.10: Processing peaks of the random access preamble receiver.

Extending this principle, the available slot configurations are designed to facilitate a PRACH receiver which may be used for multiple cells of an eNodeB, assuming a periodic pattern with period 10 ms or 20 ms.

Assuming an operating collision probability per UE, $p_{\text{coll}}^{\text{UE}} = 1\%$, one PRACH time-frequency resource (with 64 signatures) per 10 ms per 5 MHz can handle an offered load $G = -64 \ln(1 - p_{\text{coll}}^{\text{UE}}) = 0.6432$ average PRACH attempts, which translates into 128 attempts per second in 10 MHz. This is expected to be a typical PRACH load in LTE.

Assuming a typical PRACH load, example usages and cell allocations of the 16 available resource configurations are shown in Figure 17.11 for different system bandwidths and different numbers of cells per eNodeB. Resource configurations 0–2 and 15 use a 20 ms PRACH period, which can be desirable for small bandwidths (e.g. 1.4 MHz) in order to reduce the PRACH overhead at the price of higher waiting times.

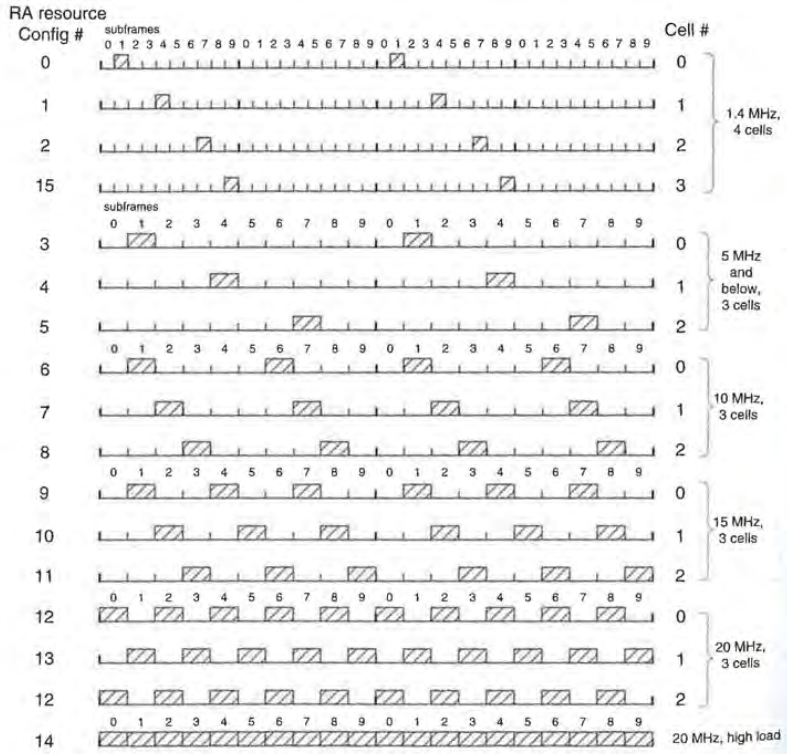
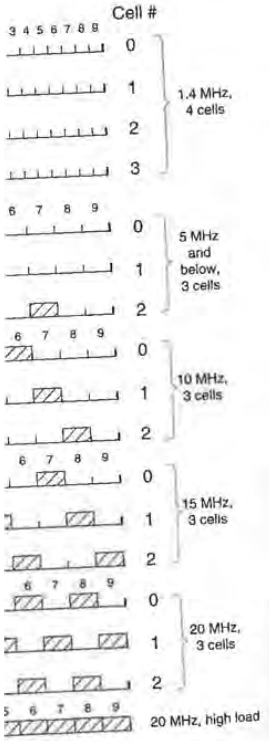


Figure 17.11: Random Access resource configurations.

As can be observed from Figure 17.11, in a three-cell scenario, time collision of PRACH resources can always be avoided except for the 20 MHz case, where collisions can be minimized to two PRACH resources occurring in the same subframe. It should also be noted that in a scenario with six cells per eNodeB, at most two PRACH resources will occur in the same subframe for bandwidths below 20 MHz.

= 1%, one PRACH time-
:an handle an offered load
hich translates into 128
'RACH load in LTE.
ocations of the 16 available
nt system bandwidths and
ıs 0–2 and 15 use a 20 ms
(e.g. 1.4 MHz) in order to



)nfigurations.

:nario, time collision of PRACH
: case, where collisions can be
subframe. It should also be noted
:ACH resources will occur in the

The variety of configurations provided therefore enables efficient dimensioning of a multicell PRACH receiver.

17.4.3 Preamble Sequence Theory and Design

As noted above, 64 PRACH signatures are available in LTE, compared to only 16 in WCDMA. This can not only reduce the collision probability, but also allow for 1 bit of information to be carried by the preamble and some signatures to be reserved for contention-free access (see Section 17.3.2). Therefore, the LTE PRACH preamble called for an improved sequence design with respect to WCDMA. While Pseudo-Noise (PN) based sequences were used in WCDMA, in LTE prime-length Zadoff–Chu (ZC) [6, 7] sequences have been chosen (see Chapter 7 for an overview of the properties of ZC sequences). These sequences enable improved PRACH preamble detection performance. In particular:

- The power delay profile is built from periodic instead of aperiodic correlation;
- The intra-cell interference between different preambles received in the same PRACH resource is reduced;
- Intra-cell interference is optimized with respect to cell size: the smaller the cell size, the larger the number of orthogonal signatures and the better the detection performance;
- The eNodeB complexity is reduced;
- The support for high-speed UEs is improved.

The 800 μs LTE PRACH sequence is built by cyclicly-shifting a ZC sequence of prime-length N_{ZC}, defined as

$$x_u(n) = \exp\left[-j\frac{\pi un(n+1)}{N_{ZC}}\right], \quad 0 \le n \le N_{ZC} - 1 \tag{17.8}$$

where u is the ZC sequence index and the sequence length $N_{ZC} = 839$ for FDD.

The reasons that led to this design choice are elaborated in the next sections.

17.4.3.1 Preamble Bandwidth

In order to ease the frequency multiplexing of the PRACH and the PUSCH resource allocations, a PRACH slot must be allocated a bandwidth BW_{PRACH} equal to an integer multiple of Resource Blocks (RBs), i.e. an integer multiple of 180 kHz.

For simplicity, BW_{PRACH} in LTE is constant for all system bandwidths; it is chosen to optimize both the detection performance and the timing estimation accuracy. The latter drives the lower bound of the PRACH bandwidth. Indeed, a minimum bandwidth of ~1 MHz is necessary to provide a one-shot accuracy of about ±0.5 μs, which is an acceptable timing accuracy for PUCCH/PUSCH transmissions.

Regarding the detection performance, one would intuitively expect that the higher the bandwidth, the better the detection performance, due to the diversity gain. However, it is important to make the comparison using a constant signal energy to noise ratio, E_p/N_0, resulting from accumulation (or despreading) over the same preamble duration, and the same false alarm probability, p_{fa}, for all bandwidths. The latter requires the detection threshold

to be adjusted with respect to the search window size, which increases with the bandwidth. Indeed, it will become clear from the discussion in Section 17.5.2.3 that the larger the search window, the higher p_{fa}. In other words, the larger the bandwidth the higher the threshold relative to the noise floor, given a false alarm target p_{fa_target} and cell size L; equivalently, the larger the cell size, the higher the threshold relative to the noise floor, given a target p_{fa_target} and bandwidth. As a result, under the above conditions, a smaller bandwidth will perform better than a large bandwidth in a single-path static AWGN channel, given that no diversity improvement is to be expected from such a channel.

Figure 17.12 shows simulation results for the TU-6 fading channel,[5] comparing detection performance of preamble bandwidths BW_{PRACH} of 6, 12, 25 and 50 RBs. For each bandwidth, the sequence length is set to the largest prime number smaller than $1/BW_{PRACH}$, the false alarm rate is set to $p_{fa\_target} = 0.1\%$, the cell radius is 0.7 km and the receiver searches for 64 signatures constructed from 64 cyclic shifts of one root ZC sequence.

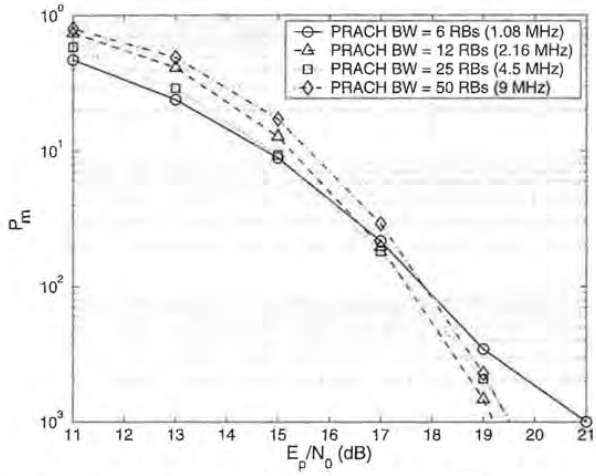


Figure 17.12: PRACH missed detection performance comparison for different BW_{PRACH} of 6, 12, 25 and 50 RBs.

We can observe that the best detection performance is achieved by preambles of 6 RBs and 12 RBs for low and high SNRs respectively. The 25-RB preamble has the overall best performance considering the whole SNR range. Thus the diversity gain of large bandwidths only compensates the increased detection threshold in the high SNR region corresponding to misdetection performances in the range of 10^{-3} and below. At a typical 10^{-2} detection probability target, the 6-RB allocation only has 0.5 dB degradation with respect to the best case.

Therefore, a PRACH allocation of 6 RBs provides a good trade-off between PRACH overhead, detection performance and timing estimation accuracy. Note that for the smallest

[5]The six-path Typical Urban (TU) channel model [8].

ncreases with the bandwidth.
.2.3 that the larger the search
idth the higher the threshold
l cell size L; equivalently, the
e floor, given a target $p_{\text{fa\_target}}$
aller bandwidth will perform
annel, given that no diversity

channel,[5] comparing detection
150 RBs. For each bandwidth,
er than $1/BW_{\text{PRACH}}$, the false
id the receiver searches for 64
quence.



rison for different BW_{PRACH} of

hieved by preambles of 6 RBs
preamble has the overall best
ersity gain of large bandwidths
igh SNR region corresponding
ow. At a typical 10^{-2} detection
adation with respect to the best

od trade-off between PRACH
racy. Note that for the smallest

system bandwidth (1.4 MHz, 6 RBs) the PRACH overlaps with the PUCCH; it is left to the eNodeB implementation whether to implement scheduling restrictions during PRACH slots to avoid collisions, or to let PRACH collide with PUCCH and handle the resulting interference.

Finally, the exact preamble transmission bandwidth is adjusted to isolate PRACH slots from surrounding PUSCH/PUCCH allocations through guard bands, as elaborated in the following section.

17.4.3.2 Sequence Length

The sequence length design should address the following requirements:

- Maximize the number of ZC sequences with optimal cross-correlation properties;
- Minimize the interference to/from the surrounding scheduled data on the PUSCH.

The former requirement is guaranteed by choosing a prime-length sequence. For the latter, since data and preamble OFDM symbols are neither aligned nor have the same durations, strict orthogonality cannot be achieved. At least, fixing the preamble duration to an integer multiple of the PUSCH symbol provides some compatibility between preamble and PUSCH subcarriers. However, with the 800 μs duration, the corresponding sequence length would be 864, which does not meet the prime number requirement. Therefore, shortening the preamble to a prime length slightly increases the interference between PUSCH and PRACH by slightly decreasing the preamble sampling rate.

The interference from PUSCH to PRACH is further amplified by the fact that the operating E_s/N_0 of PUSCH (where E_s is the PUSCH symbol energy) is much greater than that of the PRACH (typically as much as 24 dB greater if we assume 13 dB E_s/N_0 for 16QAM PUSCH, while the equivalent ratio for the PRACH would be −11 dB assuming $E_p/N_0 =$ 18 dB and adjusting by $-10\log_{10}(864)$ to account for the sequence length). This is illustrated in Figure 17.13 showing the missed detection rate (P_m) with and without data interference adjacent to the PRACH. The simulations assume a TU-6 channel, two receive antennas at the eNodeB and 15 km/h UE speed. The PRACH shows about 1 dB performance loss at $P_m = 1\%$.

The PRACH uses *guard bands* to avoid the data interference at preamble edges. A cautious design of preamble sequence length not only retains a high inherent processing gain, but also allows avoidance of strong data interference. In addition, the loss of spectral efficiency (by reservation of guard subcarriers) can also be well controlled at a fine granularity ($\Delta f_{\text{RA}} =$ 1.25 kHz). Figure 17.13 shows the missed detection rate for a cell radius of 0.68 km, for various preamble sequence lengths with and without 16QAM data interference.

In the absence of interference, there is no significant performance difference between sequences of similar prime length. In the presence of interference, it can be seen that reducing the sequence length below 839 gives no further improvement in detection rate. No effect is observed on the false alarm rate.

Therefore the sequence length of 839 is selected for LTE PRACH, corresponding to 69.91 PUSCH subcarriers in each SC-FDMA symbol, and offers 72 − 69.91 = 2.09 PUSCH subcarriers protection, which is very close to one PUSCH subcarrier protection on each side of the preamble. This is illustrated in Figure 17.14; note that the preamble is positioned

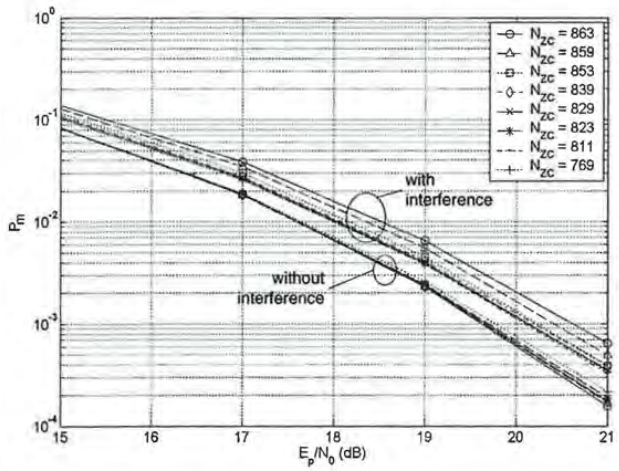


Figure 17.13: Missed detection rates of PRACH preamble with and without 16QAM interferer for different sequence lengths (cell radius of 0.68 km).

centrally in the block of 864 available PRACH subcarriers, with 12.5 null subcarriers on each side.

Finally, the PRACH preamble signal $s(t)$ can therefore be defined as follows [2]:

$$s(t) = \beta_{\text{PRACH}} \sum_{k=0}^{N_{ZC}-1} \sum_{n=0}^{N_{ZC}-1} x_{u,v}(n) \cdot \exp\left[-j\frac{2\pi nk}{N_{ZC}}\right] \times \exp[j2\pi[k + \varphi + K(k_0 + \tfrac{1}{2})]\Delta f_{\text{RA}}(t - T_{\text{CP}})] \quad (17.9)$$

where $0 \le t < T_{\text{SEQ}} + T_{\text{CP}}$, β_{PRACH} is an amplitude scaling factor and $k_0 = n_{\text{PRB}}^{\text{RA}} N_{\text{SC}}^{\text{RB}} - N_{\text{RB}}^{\text{UL}} N_{\text{SC}}^{\text{RB}}/2$. The location in the frequency domain is controlled by the parameter $n_{\text{PRB}}^{\text{RA}}$, expressed as an RB number configured by higher layers and fulfilling $0 \le n_{\text{PRB}}^{\text{RA}} \le N_{\text{RB}}^{\text{UL}} - 6$. The factor $K = \Delta f/\Delta f_{\text{RA}}$ accounts for the ratio of subcarrier spacings between the PUSCH and PRACH. The variable φ (equal to 7 for LTE FDD) defines a fixed offset determining the frequency-domain location of the random access preamble within the RBs. $N_{\text{RB}}^{\text{UL}}$ is the uplink system bandwidth (in RBs) and $N_{\text{SC}}^{\text{RB}}$ is the number of subcarriers per RB, i.e. 12.

17.4.3.3 Cyclic Shift Dimensioning (N_{CS}) for Normal Cells

Sequences obtained from cyclic shifts of *different* ZC sequences are not orthogonal (see Section 7.2.1). Therefore, orthogonal sequences obtained by cyclically shifting a single root sequence should be favoured over non-orthogonal sequences; additional ZC root sequences should be used only when the required number of sequences (64) cannot be generated by cyclic shifts of a single root sequence. The cyclic shift dimensioning is therefore very important in the RACH design.

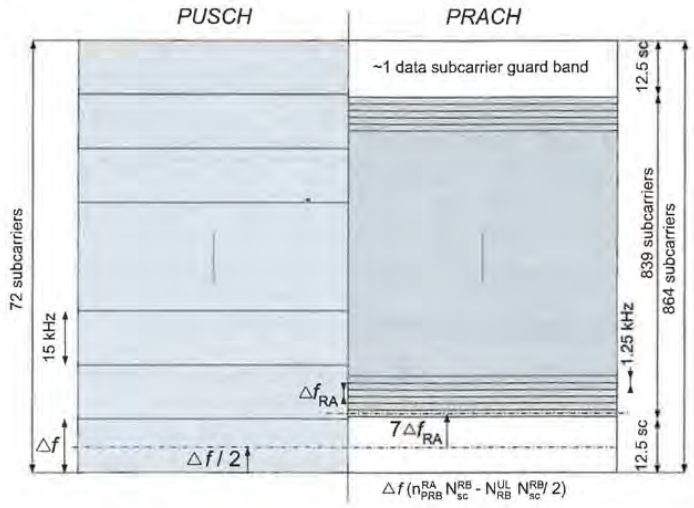


Figure 17.14: PRACH preamble mapping onto allocated subcarriers.

The cyclic shift offset N_{CS} is dimensioned so that the Zero Correlation Zone (ZCZ) (see Section 7.2.1) of the sequences guarantees the orthogonality of the PRACH sequences regardless of the delay spread and time uncertainty of the UEs. The minimum value of N_{CS} should therefore be the smallest integer number of sequence sample periods that is greater than the maximum delay spread and time uncertainty of an uplink non-synchronized UE, plus some additional guard samples provisioned for the spill-over of the pulse shaping filter envelope present in the PRACH receiver (Figure 17.15).

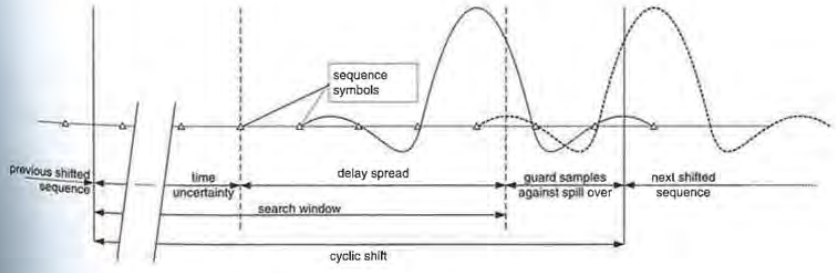


Figure 17.15: Cyclic shift dimensioning.

Table 17.4: Cell scenarios with different cyclic shift increments.

Cell scenario	Number of cyclic shifts per ZC sequence	Number of ZC root sequences	Cyclic shift size N_{CS} (samples)	Cell radius (km)
1	64	1	13	0.7
2	32	2	26	2.5
3	18	4	46	5
4	9	8	93	12

The resulting lower bound for cyclic shift N_{CS} can be written as

$$N_{CS} \geq \left\lceil \left(\frac{20}{3} r - \tau_{ds} \right) \frac{N_{ZC}}{T_{SEQ}} \right\rceil + n_g \tag{17.10}$$

where r is the cell size (km), τ_{ds} is the maximum delay spread, $N_{ZC} = 839$ and T_{SEQ} are the PRACH sequence length and duration (measured in μs) respectively, and n_g is the number of additional guard samples due to the receiver pulse shaping filter.

The delay spread can generally be assumed to be constant for a given environment. However, the larger the cell, the larger the cyclic shift required to generate orthogonal sequences, and consequently, the larger the number of ZC root sequences necessary to provide the 64 required preambles.

The relationship between cell size and the required number of ZC root sequences allows for some system optimization. In general, the eNodeB should configure N_{CS} independently in each cell, because the expected inter-cell interference and load (user density) increases as cell size decreases; therefore smaller cells need more protection from co-preamble interference than larger cells.

Some practical examples of this optimization are given in Table 17.4, showing four cell scenarios resulting from different N_{CS} values configured by the eNodeB. For each scenario, the total number of sequences is 64, but resulting from different combinations of the number of root sequences and cyclic shifts.

N_{CS} set design. Given the sequence length of 839, allowing full flexibility in signalling N_{CS} would lead to broadcasting a 10-bit parameter, which is over-dimensioning. As a result, in LTE the allowed values of N_{CS} are quantized to a predefined set of just 16 configurations. The 16 allowed values of N_{CS} were chosen so that the number of orthogonal preambles is as close as possible to what could be obtained if there were no restrictions on the value of N_{CS} [9]. This is illustrated in Figure 17.16 (left), where the cell radii are derived assuming a delay spread of 5.2 μs and 2 guard samples n_g for the pulse shaping filter.

The effect of the quantization is shown in Figure 17.16 (right), which plots the probability p_2 that two UEs randomly select two preambles on the same root sequence, as a function of the cell radius, for both the quantized N_{CS} set and an ideal unquantized set. The larger p_2, the better the detection performance. The figure also shows an ideal unquantized set. It can be seen that the performance loss due to the quantization is negligible.

Figure 17.17 illustrates the range of N_{CS} values and their usage with the various preamble formats. Note that this set of N_{CS} values is designed for use in low-speed cells. LTE also

ft increments.

ic shift N_{CS} nples)	Cell radius (km)
13	0.7
26	2.5
46	5
93	12

n as

(17.10)

, $N_{ZC} = 839$ and T_{SEQ} are the tively, and n_g is the number of r.

ant for a given environment. uired to generate orthogonal root sequences necessary to

r of ZC root sequences allows onfigure N_{CS} independently in (user density) increases as cell rom co-preamble interference

Table 17.4, showing four cell he eNodeB. For each scenario, nt combinations of the number

ng full flexibility in signalling over-dimensioning. As a result, ed set of just 16 configurations. ber of orthogonal preambles is no restrictions on the value of ell radii are derived assuming a haping filter.

ght), which plots the probability root sequence, as a function of quantized set. The larger p_2, the ideal unquantized set. It can be ligible.

usage with the various preamble se in low-speed cells. LTE also

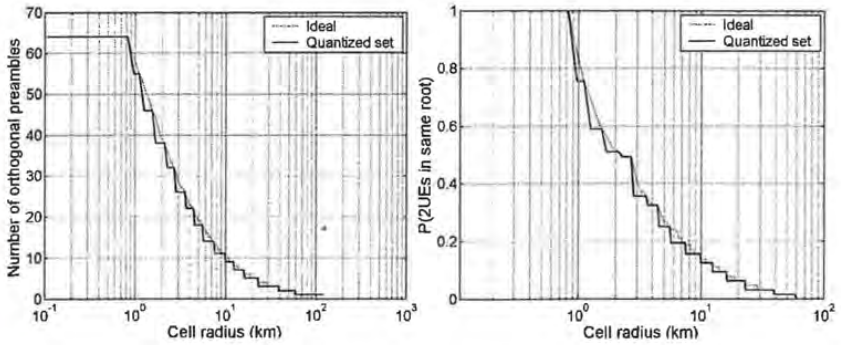


Figure 17.16: Number of orthogonal preambles (left) and probability that two UEs select two orthogonal preambles (right).

provides a second N_{CS} set specially designed for high-speed cells, as elaborated in the following sections.

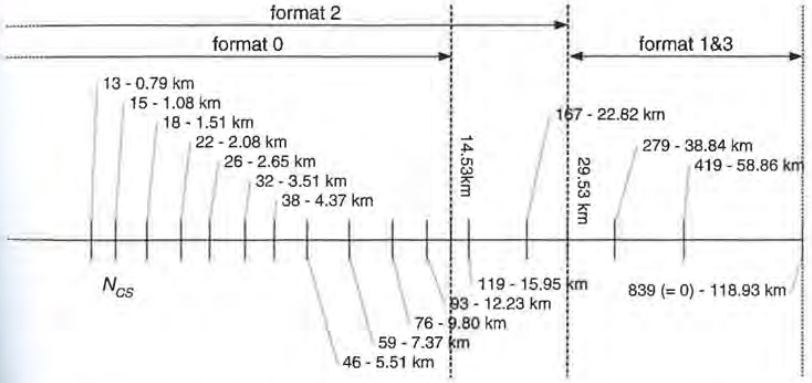


Figure 17.17: N_{CS} values and usage with the various preamble formats (low speed cells).

17.4.3.4 Cyclic Shift (N_{CS}) Restriction for High-Speed Cells

The support of 64 RACH preambles as described above assumes little or no frequency shifting due to Doppler spread, in the presence of which ZC sequences lose their zero auto-correlation property. In the presence of a frequency offset δf, it can be shown that the PRACH

ZC sequence in Equation (17.8) is distorted as follows:

$$x_u(n, \delta f) = \exp\left[-j\pi \frac{u(n-1/u)(n-1/u+1)}{N_{ZC}}\right] \times \exp\left[j2\pi \frac{n}{N_{ZC}}(\delta f\, T_{SEQ} - 1)\right] \exp\left[-\frac{j\pi}{N_{ZC}} \frac{u-1}{u}\right] = x_u(n-1/u) \exp\left[j2\pi \frac{n}{N_{ZC}}(\delta f\, T_{SEQ} - 1)\right] e^{j\Phi_u} \quad (17.11)$$

A similar expression can be written for the opposite frequency offset.

As can be observed, frequency offsets as large as one PRACH subcarrier ($\delta f = \pm\Delta f_{RA} = \pm 1/T_{SEQ} = \pm 1.25$ kHz) result in cyclic shifts $d_u = (\pm 1/u) \bmod N_{ZC}$ on the ZC sequence $x_u(n)$. (Note that $u \cdot d_u \bmod N_{ZC} = \pm 1$.) This frequency offset δf can be due to the accumulated frequency uncertainties at both UE transmitter and eNodeB receiver and the Doppler shift resulting from the UE motion in a Line of Sight (LOS) radio propagation condition.

Figure 17.18 illustrates the impact of the cyclic shift distortion on the received Power Delay Profile (PDP): it creates false alarm peaks whose relative amplitude to the correct peak depends on the frequency offset. The solution adopted in LTE to address this issue is referred to as 'cyclic shift restriction' and consists of 'masking' some cyclic shift positions in the ZC root sequence. This makes it possible to retain an acceptable false alarm rate, while also combining the PDPs of the three uncertainty windows, thus also maintaining a high detection performance even for very high-speed UEs.

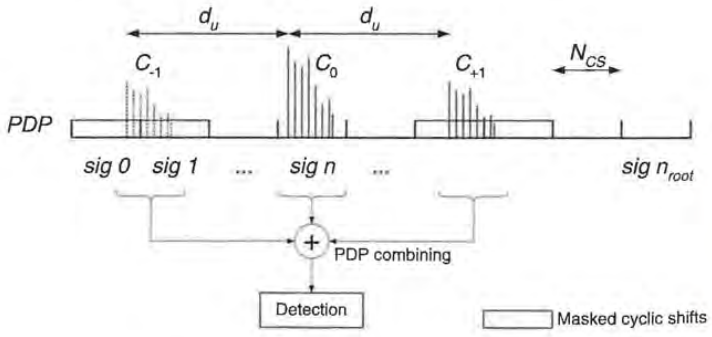


Figure 17.18: Side peaks in PDP due to frequency offset.

It should be noted that at $|\delta f| = \Delta f_{RA}$, the preamble peak completely disappears at the desired location (as per Equation (17.11)). However, the false image peak begins to appear even with $|\delta f| < \Delta f_{RA}$. Another impact of the side peaks is that they restrict the possible cyclic shift range so as to prevent from side peaks from falling into the cyclic shift region (see Figure 17.19). This restriction on N_{CS} is captured by Equations (17.12) and (17.13) and is important for the design of the high-speed N_{CS} set (explained in Section 17.4.3.5) and the

order in which the ZC sequences are used (explained in Section 17.4.3.6):

$$N_{CS} \le d < (N_{ZC} - N_{CS})/2 \tag{17.12}$$

where

$$d = \begin{cases} d_u, & 0 \le d_u < N_{ZC}/2 \\ N_{ZC} - d_u, & d_u \ge N_{ZC}/2 \end{cases} \tag{17.13}$$

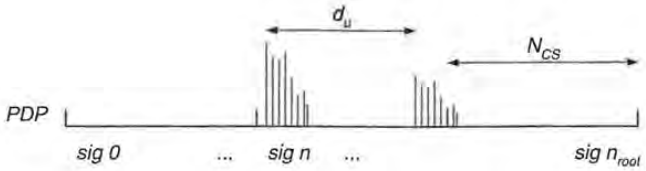


Figure 17.19: Side peaks within the signature search window.

We use C_{-1} and C_{+1} to denote the two wrong cyclic shift windows arising from the frequency offset, while C_0 denotes the correct cyclic shift window (Figure 17.18). The cyclic restriction rule must be such that the two wrong cyclic shift windows C_{-1} and C_{+1} of a cyclicly-shifted ZC sequence overlap none of C_0, C_{-1} or C_{+1} of other cyclicly-shifted ZC sequences, nor the correct cyclic shift window C_0 of the same cyclicly-shifted ZC sequence, nor each other [10]. Finally, the restricted set of cyclic shifts is obtained such that the minimum difference between two cyclic shifts is still N_{CS} but the cyclic shifts are not necessarily multiples of N_{CS}.

It is interesting to check the speed limit beyond which it is worth considering a cell to be a high-speed cell. This is done by assessing the performance degradation of the PRACH at the system-level as a function of the UE speed when no cyclic shift restriction is applied.

For this analysis, we model the RACH access attempts of multiple concurrent UEs with a Poisson arrival rate. A preamble detection is considered to be correct if the timing estimation is within 2 μs. A target E_p/N_0 of 18 dB is used for the first preamble transmission, with a power ramping step of 1 dB for subsequent retransmissions. The cell radius is random between 0.5 and 12 km, with either AWGN or a six-path TU channel, and a 2 GHz carrier frequency. eNodeB and UE frequency errors are modelled randomly within ±0.05 ppm. The access failure rate is the measure of the number of times a UE unsuccessfully re-tries access attempts (up to a maximum of three retransmissions), weighted by the total number of new access attempts.

Figure 17.20 shows the access failure rate performance for both channel types as a function of the UE speed, for various offered loads G. It can be observed that under fading conditions, the RACH failure rates experience some degradation with the UE speed (which translates into Doppler spread), but remains within acceptable performance even at 350 km/h. For the AWGN channel (where the UE speed translates into Doppler shift) the RACH failure rate stays below 10^{-2} up to UE speeds in the range 150 to 200 km/h. However, at 250 km/h and above, the throughput collapses. Without the cyclic shift restrictions the upper bound for useful performance is around 150–200 km/h.

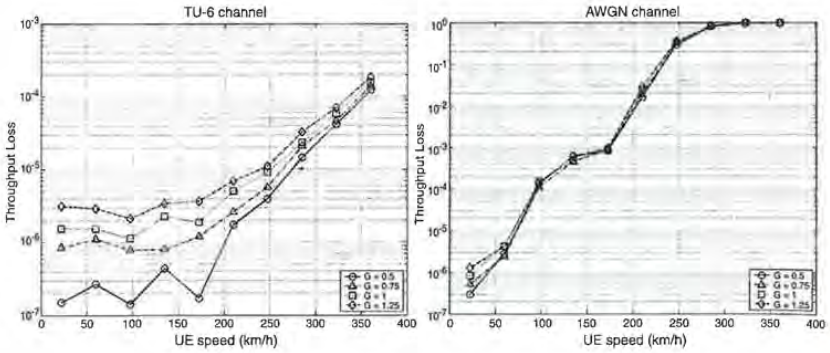


Figure 17.20: Random access failure rate as a function of UE speed.

17.4.3.5 Cyclic Shift Configuration for High-Speed Cells

The cyclic shift dimensioning for high-speed cells in general follows the same principle as for normal cells, namely maximizing the sequence reuse when group quantization is applied to cyclic shift values. However, for high-speed cells, the cyclic shift restriction needs to be considered when deriving the sequence reuse factor with a specific cyclic shift value. Note that there is no extra signalling cost to support an additional set of cyclic shift configurations for high-speed cells since the one signalling bit which indicates a 'high-speed cell configuration' serves this purpose.

The N_{CS} values for high-speed cells are shown in Figure 17.21 for the number of available and used preambles, with both consecutive and non-consecutive (quantized) cyclic shift values. The number of available preambles assumes no cyclic shift restriction at all, as in low-speed cells. It should be noted that with the cyclic shift restriction above, the largest usable high-speed cyclic shift value among all root sequences is 279 (from Equation (17.12)). As is further elaborated in the next section, only the preambles with Cubic Metric (CM) (see Section 21.3.3) below 1.2 dB are considered in Figure 17.21.

Since for small N_{CS} values the sequence usage is not so tight with a generally high sequence reuse factor, a way to simplify design, while still achieving a high reuse factor, is to reuse the small N_{CS} values for normal cells. In Figure 17.21, N_{CS} values up to 46 are from the normal cyclic shift values, corresponding to a cell radius up to 5.8 km. At the high end, the value of 237 rather than 242 is chosen to support a minimum of two high-speed cells when all the 838 sequences are used. The maximum supportable high-speed cell radius is approximately 33 km, providing sufficient coverage for preamble formats 0 and 2.

17.4.3.6 Sequence Ordering

A UE using the contention-based random access procedure described in Section 17.3.1 needs to know which sequences are available to select from. As explained in Section 17.4.3.3, the full set of 64 sequences may require the use of several ZC root sequences, the identity of each of which must be broadcast in the cell. Given the existence of 838 root sequences, in LTE the

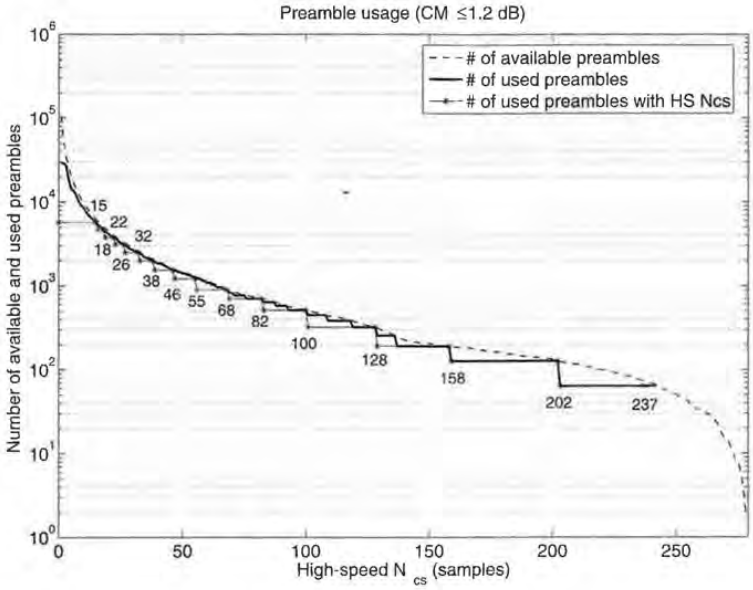


Figure 17.21: Number of available and used preambles in low CM group.

signalling is streamlined by broadcasting only the index of the *first* root sequence in a cell, and the UE derives the other preamble signatures from it given a predefined ordering of all the sequences.

Two factors are taken into account for the root sequence ordering, namely the CM [11,12] of the sequence, and the maximum supportable cell size for high-speed cells (or equivalently the maximum supported cyclic shift). Since CM has a direct impact on cell coverage, the first step in ordering the root sequences is to divide the 838 sequences into a low CM group and a high CM group, using the CM of Quadrature Phase Shift Keying (QPSK) (1.2 dB) as a threshold. The low CM group would be used first in sequence planning (and also for high-speed cells) since it is more favourable for coverage.

Then, within each CM group, the root sequences are classified into subgroups based on their maximum supportable cell radius, to facilitate sequence planning including high-speed cells. Specifically, a sequence subgroup g is the set of all root sequences with their maximum allowed cyclic shifts ($S_{\max}$) derived from Equation (17.12) lying between two consecutive high-speed N_{CS} values according to

$$N_{CS}(g) \le S_{\max} < N_{CS}(g+1), \text{ for } g = 0, 1, \ldots, G-2, \text{ and}$$
$$S_{\max} \ge N_{CS}(G-1) \tag{17.14}$$

for G cyclic shift values, with the set of N_{CS} values being those for high-speed cells. Sequences in each subgroup are ordered according to their CM values.

Figure 17.22 shows the CM and maximum allowed cyclic shift values at high speed for the resulting root sequence index ordering used in LTE. This ordering arrangement reflects a continuous CM transition across subgroups and groups, which ensures that consecutive sequences always have close CM values when allocated to a cell. Thus, consistent cell coverage and preamble detection can be achieved in one cell.

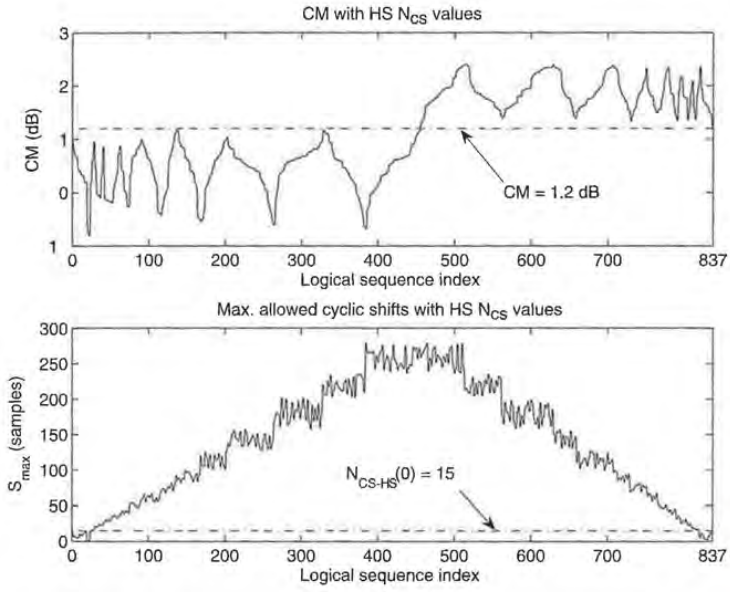


Figure 17.22: CM and maximum allowed cyclic shift of reordered group sequences.

The LTE specifications define the mapping from root sequence index u to a reordered index in a table [2], an enhanced extract of which is given in Figure 17.23. The first 16 subgroups are from the low CM group, and the last 16 from the high CM group. Figure 17.23 shows the corresponding high-speed N_{CS} value for each subgroup. Note that sequences with S_{max} less than 15 cannot be used by any high-speed cells, but they can be used by any normal cell which requires no more than 24 root sequences from this group for a total of 64 preambles. Note also that the ordering of the physical ZC root sequence indices is pairwise, since root sequence indices u and $N_{ZC} - u$ have the same CM and S_{max} values. This helps to simplify the PRACH receiver, as elaborated in Section 17.5.2.2.

17.5 PRACH Implementation

This section provides some general principles for practical implementation of the PRACH function.

22

Radio Resource Management

Muhammad Kazmi

22.1 Introduction

Radio Resource Management (RRM) encompasses a wide range of techniques and procedures, including power control, scheduling, cell search, cell reselection, handover, radio link or connection monitoring, and connection establishment and re-establishment. Advanced features like interference management, location services, Self-Optimizing Networks[1] (SON) and some network planning methods make use of RRM-related techniques based on radio related measurements made by the User Equipment (UE) or eNodeB. In this chapter, we address the RRM techniques and reporting mechanisms that support UE mobility in the LTE network (E-UTRAN), including cell search, radio measurements, cell reselection, handover and radio link monitoring. We focus here on the performance requirement aspects, while the procedures themselves are described in Chapters 3, 4, 7, 12, 17 and 18.

The RRM-related actions undertaken by the UE can be broadly divided into those relevant in the RRC_IDLE state and those relevant in the RRC_CONNECTED state, as illustrated in Figures 22.1 and 22.2 respectively.

RRM in E-UTRAN is designed to handle the challenges posed by the fundamental characteristics of the LTE system, including:

- The packet-oriented transmission of LTE, realized by fast time- and frequency-domain scheduling, may lead to large and swift interference fluctuations. This may affect the accuracy of the signal quality estimates required for mobility decisions in certain scenarios. Time-domain filtering is used to help smooth out the interference variations.
- The provision of a wide range of Discontinuous Reception (DRX) cycle lengths, including periods up to 2.5 seconds, yields substantial benefits in terms of UE power saving. However, such DRX cycles also mean that it is not feasible to enforce the same mobility performance in all cases.

[1]Details on user equipment positioning and self optimization of the network are described in Chapters 19 and 25 respectively.

LTE – The UMTS Long Term Evolution: From Theory to Practice, Second Edition.
Stefania Sesia, Issam Toufik and Matthew Baker.
© 2011 John Wiley & Sons, Ltd. Published 2011 by John Wiley & Sons, Ltd.

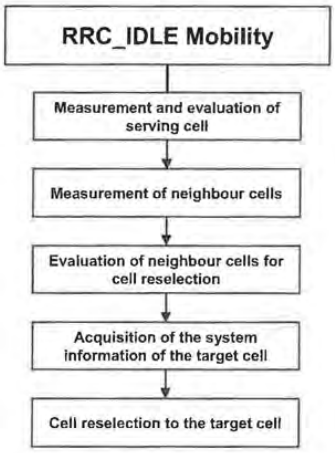


Figure 22.1: UE actions related to RRM Procedures in RRC_IDLE state.

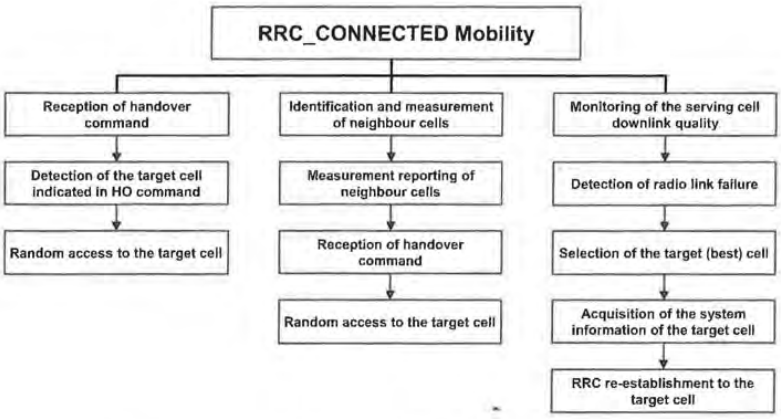


Figure 22.2: UE actions related to RRM Procedures in RRC_CONNECTED state.

- LTE is designed to support seamless mobility not only within itself (both intra- and inter-frequency) but also with the legacy 3GPP Radio Access Technologies (RATs) (e.g. GERAN[2] and UTRAN[3]) and with certain non-3GPP RATs (e.g. CDMA2000

[2]GSM EDGE Radio Access Network.
[3]Universal Terrestrial Radio Access Network.

1xRTT and HRPD[4]). Mobility between these different RATs requires the UE to detect and measure cells on the target technologies as requested by the network, each with its own different channel structure.

- LTE is designed to support a wide range of cell sizes (ranging from a few tens of metres to tens of kilometres) and deployment/propagation scenarios, yet the RRM techniques are required to be as generic as possible and to exhibit consistent mobility performance.
- Despite the differences between the Frequency Division Duplex (FDD) and Time Division Duplex (TDD) frame structures in LTE (see Section 6.2), and the fact that in FDD deployments the eNodeBs may be either synchronized or unsynchronized, mobility requirements are expected to be the same (or at least similar) in order to minimize UE implementation complexity and simplify network planning.
- The low overall latency requirements in LTE place constraints on the time taken to make and report measurements and to perform handover. Additionally, to meet the challenges of certain specific mobility procedures, enhanced requirements with substantially lower latency are also specified.
- The time-multiplexing of Multimedia Broadcast/Multicast Service (MBMS) and unicast data on a single LTE carrier is attractive in that it obviates the need for a separate carrier frequency to offer mobile broadcast services. However, it reduces the opportunities for the UEs to perform downlink cell measurements to support mobility; nevertheless, the RRM performance should remain consistent.

22.2 Cell Search Performance

Cell search is one of the most fundamental aspects of mobility. As explained in Chapter 7, it enables the UE to acquire the carrier frequency, timing and cell identity of cells. In LTE the cell search performance requirements relate to the case of neighbour cell search under the assumption that the UE has already acquired carrier frequency synchronization; the requirements are only applicable in RRC_CONNECTED.

22.2.1 Cell Search within E-UTRAN

The objective of cell search within E-UTRAN is to identify one of the 504 unique Physical Cell Identities (PCIs) (see Chapter 7 and [1, Section 6.11]). The cell search algorithm is not specified and is left for UE implementation; however, typically, the UE performs cell search in a hierarchical manner (see Section 7.2).

An important characteristic of the cell identification requirements is that the same requirements are applicable in a wide range of propagation conditions and for both FDD (with or without synchronization of the eNodeBs) and TDD (where synchronization of the eNodeBs can be assumed).

The requirements are specified in terms of the maximum permissible cell identification delay, which includes the time taken for Reference Signal Received Power (RSRP) or Reference Signal Received Quality (RSRQ) physical layer measurements (see Sections 22.3.1.1 and 22.3.1.2 respectively).

[4]High Rate Packet Data.

22.2.1.1 E-UTRAN Intra-frequency Cell Search

In case of intra-frequency cell search, the UE identifies E-UTRA cells on the same carrier frequency as that of the serving cell. The time required to detect a cell depends upon a number of factors, most notably the received quality of the synchronization signals, the received level of the Reference Signals (RSs) and the time available for performing the search. The latter factor stems from the fact that the available time for intra-frequency measurements may be reduced by measurement gaps for inter-frequency or inter-RAT measurements as explained in Section 22.2.1.2. The cell search delay also depends upon the configured DRX cycle period.

If no DRX is configured, and for DRX cycles up to 40 ms,[5] the UE is required to detect an E-UTRA FDD or TDD intra-frequency target cell within 800 ms if no inter-frequency measurement gaps are configured, provided that the target cell's received synchronization signal quality $\hat{E}_s/I_{ot}$ (defined as the energy per Resource Element (RE) of the synchronization signals divided by the total received energy of noise and interference on the same RE) is at least −6 dB. This is the 'minimum', or worst case, requirement.

The cell search delay can be shorter if the received signal quality is higher than the minimum cell detection threshold. The performance in some typical deployment conditions is illustrated in Figures 22.3 and 22.4. Here, scenarios covering both synchronized and unsynchronized eNodeBs are analysed, as summarized in Table 22.1. ETU5 (Extended Typical Urban with UE speed 5 km/h), ETU300 (UE speed 300 km/h) and EPA5 (Extended Pedestrian A with UE speed 5 km/h) propagation models are used,[6] and two receive antennas are assumed at the UE.[7] Further details of the modelled scenarios can be found in [2].

Table 22.1: Cell identification test parameters.

	Unit	Cell1	Cell2	Cell3 (target cell)
Relative delay of 1st path for synchronized case	ms	0	0	Half CP length
Relative delay of 1st path for unsynchronized case	ms	0	1.5	3
SNR	dB	5.18	0.29	−0.75 (worst case)
PSS for case of different PSS		PSS1	PSS2	PSS3
PSS for case of same PSS		PSS1	PSS2	PSS1

The cell search performance is measured in terms of the 90-percentile cell identification delay, i.e. the maximum time required to detect the target cell 90% of the time.

Various scenarios are analysed to examine the impact on the detection performance of different combinations of PSS[8] and SSS[9] sequences as indicated in Table 22.2. More detailed performance results can be found in [3].

[5] This is designed to ensure robust mobility performance for delay-sensitive services like Voice over IP (VoIP), which typically requires short DRX cycles.

[6] Further details of these propagation models are given in Chapter 20.

[7] No margin is included for non-ideal UE receiver implementation or reporting delay for the RSRP measurement to the network.

[8] Primary Synchronization Signal.

[9] Secondary Synchronization Signal.

Table 22.2: Cell identification test scenarios.

Test case (synch, asynch eNodeBs)	Cell3 (Target)		Cell1 (Interference)		Cell2 (Interference)	
1,5	PSS3	SSS3a, SSS3b	PSS1	SSS1a, SSS1b	PSS2	SSS2a, SSS2b
2,6	PSS1	SSS3a, SSS3b	PSS1	SSS1a, SSS1b	PSS2	SSS2a, SSS2b
3,7	PSS1	SSS1a, SSS3b	PSS1	SSS1a, SSS1b	PSS2	SSS2a, SSS2b
4,8	PSS3	SSS1a, SSS1b	PSS1	SSS1a, SSS1b	PSS2	SSS2a, SSS2b

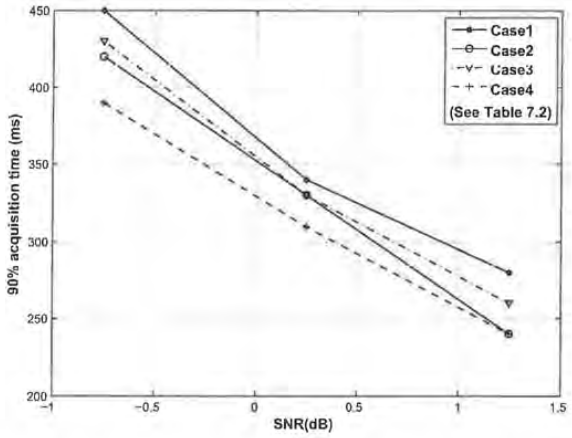


Figure 22.3: Cell search performance with synchronized eNodeBs.
Reproduced by permission of © NXP Semiconductors.

When measurement gaps are configured for inter-frequency or inter-RAT measurements the UE has less opportunity to detect a cell. Hence, in this case the cell identification delay may be larger than the baseline 800 ms delay, with the actual value depending upon the periodicity of the gaps. Furthermore, for DRX cycles larger than 40 ms, the cell identification delay increases in proportion to the length of the DRX cycle, allowing UE to save battery power.

22.2.1.2 E-UTRAN Inter-frequency Cell Search

In the case of inter-frequency cell search, the UE identifies E-UTRA cells operating on carrier frequencies other than that of the serving cell (and possibly also in different frequency bands and/or with different duplex modes). Inter-frequency measurements, including cell identification, are performed during periodic measurement gaps unless the UE has more than one receiver. Two possible gap patterns can be configured by the network, each with a gap length of 6 ms: in gap pattern #0, the gap occurs every 40 ms, while in gap pattern #1 the

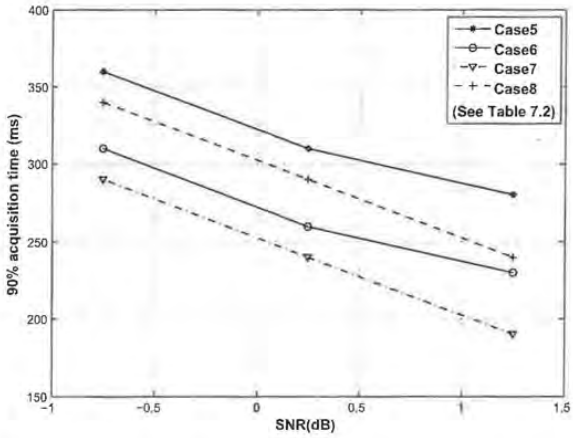


Figure 22.4: Cell search performance with unsynchronized eNodeBs. Reproduced by permission of © NXP Semiconductors.

gap occurs every 80 ms, as shown in Figure 22.5. There is an obvious trade-off between these two gap patterns: the former yields a shorter cell identification delay but a greater interruption in data transmission and reception. Only one gap pattern can be configured at a time for measuring all frequency layers (both inter-frequency and inter-RAT).

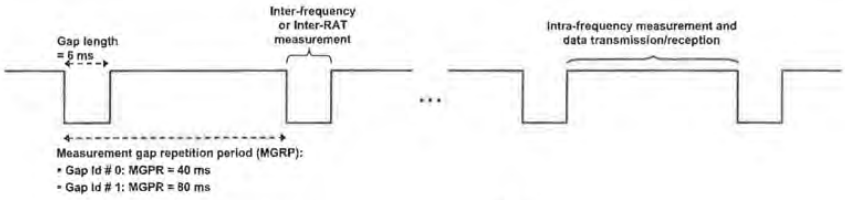


Figure 22.5: Measurement gap patterns for inter-frequency and inter-RAT cell search and measurements.

If no DRX is used, or if the DRX cycle length is less than or equal to 160 ms, the UE is required to identify an E-UTRA FDD or TDD inter-frequency cell within 3.84 s, provided the received synchronization signal quality is at least −4 dB (assuming gap pattern #0). As for intra-frequency cell search, for DRX cycles larger than 160 ms the cell identification delay increases proportionally.

22.2.2 E-UTRAN to E-UTRAN Cell Global Identifier Reporting Requirements

In addition to the basic PCI cell search requirements specified in LTE Release 8, E-UTRAN Cell Global Identifier[10] (ECGI) reporting requirements were introduced in Release 9 for both FDD and TDD. The requirements are the same for FDD and TDD for their respective frequency bands, and therefore the description in the subsequent sections is generic for both.

The UE is required to identify and report the ECGI of a target E-UTRA cell upon request from the serving eNodeB. The ECGI measurement report can be used by the eNodeB for various features. For instance, it can assist the eNodeB to establish the neighbour cell relations automatically (see Chapter 25.2) when a new neighbour cell is added or an existing one is removed [5]. The ECGI measurement report can also be used by the serving eNodeB for performing inbound mobility to a Home eNodeB[11] (HeNB) in RRC_CONNECTED state (see Chapter 24 for more details on HeNBs). Due to the small size of HeNB cells, it is not unlikely that a PCI may be used more than once among the HeNBs located within the coverage area of a macro-eNodeB. Hence an important objective of the ECGI report is to enable the serving eNodeB to identify unambiguously the target HeNB.

The acquisition of ECGI requires the UE first to read the Master Information Block (MIB), which is transmitted on Physical Broadcast CHannel (PBCH) with a periodicity of 40 ms (see Section 9.2.1). This is followed by the reading of the System Information Block type 1 (SIB1), which contains the ECGI and is transmitted with a periodicity of 80 ms on the DownLink Shared CHannel (DL-SCH) (see Section 3.2.2). The UE is not required to receive or transmit in the serving cell while acquiring the ECGI of a target cell; a UE is permitted to create gaps autonomously in downlink reception and uplink transmission to read the MIB and SIB1 of a target cell.

HeNBs can be deployed on either a shared carrier or a dedicated carrier. Similarly the SON Automatic Neighbouring Relation (ANR) is applicable to both intra- and inter-frequency cells, and therefore EGCI requirements are specified for both intra-frequency and inter-frequency ECGI reporting.

22.2.2.1 Intra-frequency E-UTRAN Cell Global Identifier Reporting

The intra-frequency ECGI reporting requirement is specified in terms of the maximum time allowed for the UE to identify the ECGI of a new intra-frequency target E-UTRA cell whose synchronization signal quality $\hat{E}_s/I_{ot}$ (as defined in Section 22.2.1.1) is at least -6 dB; the identification delay is not allowed to exceed 150 ms with 90% confidence, excluding the procedure delay of the ECGI request message. The same ECGI identification delay requirements are applicable with and without DRX.

The initial autonomous gap created by the UE for correcting the frequency and for decoding the first MIB block of the target cell comprises 9 subframes. Each subsequent autonomous gap for decoding the remaining MIB and SIB1 blocks to read the whole of the target cell's ECGI typically consists of 4 subframes. The maximum aggregated duration of autonomous gaps created by the UE is not allowed to exceed 90 subframes when identifying

[10] The E-UTRAN Cell Global Identifier is a globally unique identifier broadcast by the cell. It is composed of 3-bytes of Public Land Mobile Network (PLMN) Identity and 28 bits to identify the cell within that PLMN – see [4, Section 6.3.4], and Chapter 25.

[11] Inbound mobility to Home eNodeB refers to the case of a UE moving into a Home eNodeB cell.

an intra-frequency cell's ECGI; this is designed to ensure that certain minimum serving cell reception and transmission performance is maintained.

22.2.2.2 Inter-frequency E-UTRAN Cell Global Identifier Reporting

The inter-frequency ECGI requirements (i.e. ECGI identification delay and serving cell reception/transmission performance) are the same as those specified for the intra-frequency requirement, except that the synchronization signal quality under which the requirement should be satisfied is relaxed to -4 dB. An autonomous gap created by the UE for decoding the initial MIB block of the target inter-frequency cell comprises 9 subframes, and each of the subsequent gaps typically consists of 4 subframes. The maximum aggregated duration of autonomous gaps created by the UE is likewise not allowed to exceed 90 subframes when identifying an inter-frequency cell's ECGI. This ensures that certain minimum serving cell reception and transmission performance is maintained while reading the inter-frequency target cell ECGI.

22.2.3 E-UTRAN to UTRAN Cell Search

The E-UTRAN to UTRAN cell search procedure allows the UE to identify a target FDD or TDD UTRA cell. The E-UTRAN-UTRAN cell search requirements apply until an explicit neighbour cell list is received by the UE from the serving E-UTRA cell. This list indicates the primary scrambling codes of up to 32 neighbour cells per UTRA FDD or TDD carrier. The signals provided by UTRA FDD and TDD differ, and therefore separate cell identification requirements are specified for each technology. In each case, the cell search algorithm is not specified, but the search is typically carried out in hierarchical manner. In UTRA FDD, where Primary and Secondary Synchronization CHannels (P-SCH and S-SCH) are provided, the UE typically identifies the target UTRA cell via the following steps:

Slot timing acquisition: The UE acquires the slot timing by performing a matched filter correlation over the Primary Synchronization Code (PSC), which is transmitted in the first 256 chips in every slot (0.666 ms). The PSC is common to all UTRA cells.

Frame timing acquisition and code group identification: The UE acquires the frame timing and the identity of the group of 8 codes to which the cell's primary scrambling code belongs by performing correlations over the Secondary Synchronization Codes (SSCs) transmitted in the initial 256 chips of every slot (i.e. at the same time as the PSC). A sequence of 15 SSCs in one 10 ms radio frame uniquely identifies the code group. In good propagation conditions the information contained within three slots is sufficient to identify both the frame timing and the code group.

Primary scrambling code identification: A given UTRA cell uses one code from the group indicated by the SSC as the scrambling code for all downlink channels, including the Primary Common PIlot CHannel (P-CPICH). The UE therefore performs a simple correlation against the known CPICH sequence (which is the same in all UMTS cells) scrambled in turn by each of the eight possible scrambling sequences in the code group, in order to determine which primary scrambling code is being used in the cell, and hence the cell identity.

In order to report an identified cell, the UE needs to measure the P-CPICH E_c/N_0 (where E_c is the CPICH energy per chip and N_0 the noise power spectral density) or Received Signal Code Power (RSCP) over the 'physical layer measurement period'.

The same measurement gap patterns are used for E-UTRA inter-frequency and inter-RAT measurements. In both DRX and non-DRX scenarios, a target UTRA FDD cell is considered detectable if the CPICH E_c/I_0 is at least -20dB and the synchronization channel SCH_E_c/I_0 is at least -17dB, where I_0 is the power spectral density of the total input signal at the UE antenna connector and SCH_E_c is the transmit energy per chip of the SCH, including the downlink signal transmitted by the serving cell [6]. However, in terms of reporting delay, separate sets of UTRA cell identification requirements are specified for the cases with and without DRX.

If no DRX is used, or for DRX cycles up to 40 ms, the UE is required to identify a detectable UTRA FDD cell within $2.4 * N_{\text{freq}}$ seconds and $4.8 * N_{\text{freq}}$ seconds using gap patterns #0 and #1 respectively (where N_{freq} is the number of carrier frequencies to be monitored). For DRX cycles larger than 40 ms, the UTRA cell search delay increases proportionally and is expressed in terms of a number of DRX cycles. For example in case of DRX cycle length of 1.28 seconds, the maximum permissible cell search delay is $25.6 * N_{\text{freq}}$ seconds or $20 *$ DRX cycle length $* N_{\text{freq}}$; where the factor of 20 incorporates the required UTRA FDD cell search occasions and CPICH measurement samples.

22.2.4 E-UTRAN to GSM Cell Search

In GSM the Broadcast Control CHannel (BCCH) frequency reuse factor is greater than one. Therefore in principle the GSM BCCH Received Signal Strength Indicator (RSSI) measurement, which is measured on the BCCH carrier (see Section 22.3.3) can identify a GSM cell, since a different RSSI measurement can be reported for each carrier. However, if small cluster sizes are used with BCCH frequency reuse, the GSM carrier RSSI alone cannot always reliably identify the correct GSM cell: for example, the network may not be able to distinguish RSSI measurement reports for GSM cells whose BCCHs use the same carrier. This may lead to calls being dropped, especially in deployment scenarios with tight reuse. Therefore, the UE is also required to decode the Base Station Identity Code (BSIC)[12]. In E-UTRAN, GSM measurements are always reported together with the verified BSIC [4], in contrast to the legacy GSM system, where measurements can be reported with or without a verified BSIC. The BSIC verification comprises initial BSIC identification and BSIC re-confirmation. These are summarized below; A more extensive description of the GSM common channels is provided in [7].

Initial BSIC identification comprises the BSIC decoding for a GSM cell for the first time when the UE does not have any knowledge about the relative timing between the E-UTRA and GSM cells. The UE is required to decode the BSIC of the 8 strongest BCCH carriers

[12]The BSIC allows a UE to distinguish two different GSM cells which share the same beacon frequency. It is a 6-bit field composed of two 3- bit fields: the Base station Colour Code (BCC) and the Network Colour Code (NCC). The BCC is used to identify the Training Sequence Code (TSC) to be used when reading the BCCH. The NCC is used to differentiate between operators utilizing the same frequencies, e.g. on an international border when both operators have been allocated the same frequency or frequencies.

of the GSM cells that are provided in the GSM inter-RAT cell information list[13] [4]. The UE decodes the BSIC of the BCCH carriers in order of decreasing RSSI. Through the BSIC decoding the UE acquires the initial timing information of these GSM cells.

BSIC re-confirmation consists of the BSIC decoding of a GSM cell after the initial BSIC identification has been performed. The UE updates its stored values of the timing of up to 8 identified GSM cells every time the BSIC is decoded, in order to compensate for cell timing drift relative to the serving cell to which the UE is locked.

BSIC verification requirements are applicable down to the reference sensitivity level defined in [8]. Initial BSIC identification and BSIC re-confirmation requirements are expressed in terms of the time required to decode the BSIC for the first time, $T_{identity,GSM}$ and the time required to re-confirm the initially identified BSIC $T_{re\text{-}confirm,GSM}$ under the above constraints, which depend on the total number of carrier frequency layers of all RATs since the same gap pattern is shared by all RATs.

22.2.5 Enhanced Inter-RAT Measurement Requirements

As mentioned earlier, the measurement requirements are generally minimum requirements for worst-case scenarios. In order to support Circuit-Switched Fall Back (CSFB)[14] scenarios in E-UTRAN, which require shorter overall delay, enhanced measurement requirements for UTRAN FDD and GSM were introduced in LTE Release 9.

For both UTRA FDD and GSM, the enhanced measurement requirements are specified under the assumption that CSFB from a serving E-UTRA cell to a target UTRA FDD or GSM cell is used when the serving cell and the target cell have overlapping coverage. Under these conditions the received signal quality of the target cell is typically higher than the minimum levels specified for normal cell search.

In the case of UTRA FDD, the enhanced UTRA FDD cell identification requirements are therefore specified assuming that both received CPICH E_c/I_0 and synchronization channel SCH_E_c/I_0 levels from the UTRA FDD cell are at least -15 dB. At these levels when no DRX is used or when DRX ≤ 40 ms, the UTRA FDD cell search delay is about 1 s (assuming gap pattern #0). Hence, compared to the minimum requirements, the enhanced UTRA FDD measurement reporting delay of an unknown UTRA FDD cell is reduced by a factor of 2.

In the case of GSM, the enhanced BSIC verification requirements are specified assuming that the target cell is received at about 10 dB above the reference sensitivity level or reference interference levels (as specified in [8]). Under this condition and without DRX or with DRX ≤ 40 ms, the UE can complete initial BSIC decoding in about 1.3 s, assuming gap pattern #0 is used only for monitoring GSM measurements.

[13] When camped on an LTE cell, the UE is provided with a Neighbour Cell List (NCL) containing at least 32 GSM carrier numbers (i.e. Absolute Radio Frequency Channel Numbers (ARFCNs)) indicating the frequencies of neighbouring cells, and optionally an associated BSIC for each GSM carrier in the NCL.

[14] CSFB enables voice and other circuit-switched services to be provided to UEs served by E-UTRAN by reusing the existing circuit switched infrastructure such as GERAN or UTRAN [9] – see Section 2.4.2.1.

22.3 Mobility Measurements

In order to support mobility within E-UTRAN and between E-UTRAN and other RATs (UTRAN FDD and TDD, GERAN and CDMA2000), a number of radio-related UE measurements are specified. All the E-UTRAN and inter-RAT measurements described in the following sections are reported by the UE to the serving eNodeB only in RRC_CONNECTED state. In RRC_IDLE, the measurements are not reported but may be used autonomously by the UE for cell reselection. Therefore although a particular measurement may be used in both RRC_CONNECTED and RRC_IDLE states, the corresponding requirements only apply to RRC_CONNECTED.

22.3.1 E-UTRAN Measurements

For mobility within E-UTRAN and from other RATs to E-UTRAN, two UE measurements are defined: RSRP and RSRQ [10].

22.3.1.1 Reference Signal Received Power (RSRP)

RSRP is measured by the UE over the cell-specific Reference Signals (RSs) within the measurement bandwidth over a measurement period. RSRP is a type of signal strength measurement and is indicative of the cell coverage. It is defined as the linear average over the power contributions (in Watts) of the REs that carry cell-specific RSs within the considered measurement frequency bandwidth. Normally the RSs transmitted on the first antenna port are used for RSRP determination, but the RSs on the second antenna port can also be used if the UE can reliably determine that they are being transmitted. If receive diversity is in use by the UE, the combined RSRP must be at least as large as the RSRP of any of the individual diversity branches. It is applicable in both RRC_IDLE and RRC_CONNECTED states and is used for cell reselection and handover within E-UTRAN (intra-frequency and inter-frequency) and to E-UTRAN from any of UTRAN FDD or TDD, GSM, CDMA2000 1xRTT or HRPD.

RSRP is therefore considered to be the most important measurement quantity for E-UTRAN.

An accuracy requirement is defined for the RSRP measurement (both intra- and inter-frequency); this is applicable for received signal quality $\hat{E}_s/I_{ot}$ as low as -6 dB. A measurement bandwidth equivalent to the central 6 Resource Blocks (RBs) is assumed. In the time domain the physical layer measurement periods when no DRX is used (and for short DRX cycles) are 200 ms and 480 ms for intra-frequency and inter-frequency RSRP respectively. The inter-frequency measurement naturally takes longer as it can only be performed during the measurement gaps; furthermore, the inter-frequency RSRP physical layer measurement period increases linearly with the number of carrier frequencies that has to be measured during the gaps, N_{freq}. The physical layer measurement period also increases in proportion to the DRX cycle for DRX cycles larger than 40 ms for intra-frequency RSRP and larger than 80 ms for inter-frequency RSRP. The measurement sampling rate is not specified but is left to the UE implementation.

The UE is required to be able to measure RSRP from at least 8 identified intra-frequency cells over the physical layer measurement period.[15] Similarly, the UE is also required to measure RSRP from at least 4 identified inter-frequency cells per inter-frequency carrier for up to 3 carriers (i.e. a total of 12 inter-frequency cells).

22.3.1.2 Reference Signal Received Quality (RSRQ)

RSRQ is the ratio of RSRP to RSSI for an E-UTRA carrier. The RSSI part of RSRQ is the total received power including interference from all sources, including serving and non-serving cells, adjacent channel interference and thermal noise. Unlike RSRP, it is measured on all REs in the OFDM symbols containing RSs for antenna port 0 (not just the REs containing RSs themselves), within the measurement bandwidth.[16]

This interference component of RSRQ enables the UE to quantify the received signal quality considering both signal strength and interference, which may vary with the UE's location in the cell [11].

In the first release of LTE (Release 8), RSRQ was applicable only in RRC_CONNECTED state. It is therefore used for handover within E-UTRAN, and from other RATs to E-UTRAN.

However, in order to prevent outages caused by high interference situations, in Release 9 RSRQ was also introduced for RRC_IDLE state; this gives the network the option to configure the UE to use RSRQ as a metric for performing cell reselection, at least in the cases of cell reselection within E-UTRAN, from UTRAN FDD to E-UTRAN and from GSM to E-UTRAN.

Both intra- and inter-frequency RSRQ measurement requirements are specified in [12]. The RSRQ and RSRP together have been shown to be particularly beneficial for performing inter-frequency quality-based handover [13]. RSRQ is inherently a relative quantity which to some extent eliminates absolute measurement errors and leads to better accuracy than is possible for RSRP. Like RSRP, the RSRQ accuracy requirements are applicable for $\hat{E}_s/I_{ot}$ down to −6 dB (based on a measurement bandwidth equivalent to the central 6 RBs). The time domain physical layer measurement periods for RSRP are equivalent to those of RSRP for the corresponding DRX cases. The measurement sampling rate is also UE implementation dependent.

The UE is also required to measure RSRQ from the same number of intra- and inter-frequency cells as for RSRP.

22.3.2 UTRAN Measurements

The UTRAN measurements can be categorized into those specific to UTRAN FDD and those for E-UTRAN TDD. For mobility to UTRAN FDD, CPICH[17]RSCP, CPICH E_c/N_0[18] and RSSI measurements are defined [10]. The CPICH measurements are the most important for

[15] When measurement gaps are used, and depending upon the gap periodicity, the number of measured cells may be lower than 8 due to the fact that the measurement gaps reduce the time available for the UE to make intra-frequency measurements.

[16] Note that in Release 10 the RSRQ can be configured to be measured on all OFDM symbols if mechanisms for time-domain inter-cell interference coordination are configured – see Section 31.2.4.2.

[17] Common PIlot CHannel.

[18] Energy per chip divided by the noise power spectral density.

mobility between E-UTRAN and UTRAN FDD systems. For mobility to UTRAN TDD P-CCPCH[19] RSCP and RSSI measurements are used.

All these UTRAN measurements can be used in both RRC_IDLE and RRC_CONNECTED states. However, the network has the freedom to configure specific measurements in different mobility scenarios. Therefore, all the measurement requirements (such as accuracies) are applicable only in RRC_CONNECTED state.

22.3.2.1 UTRAN FDD CPICH RSCP

UTRAN FDD CPICH RSCP is measured on the Primary-CPICH of the target UTRAN cell and is used for both cell reselection and handover from E-UTRAN. For inter-RAT mobility, the CPICH RSCP absolute accuracy requirement should be fulfilled over the physical layer measurement period, whose length depends upon the gap pattern and DRX cycle (e.g. 480 ms for gap pattern #0 with DRX cycle not greater than 40 ms). The UE is required to measure the CPICHs from 6 identified UTRA FDD cells.

22.3.2.2 UTRAN FDD CPICH E_c/N_0

The UTRAN FDD CPICH E_c/N_0 is the ratio of the CPICH RSCP to the UTRA carrier RSSI. CPICH E_c/N_0 can be used for both cell reselection and handover.

The measurement periods for CPICH E_c/N_0 with and without DRX are the same as for CPICH RSCP.

22.3.2.3 UTRAN FDD Carrier RSSI

The UTRAN FDD carrier RSSI is the total received wideband power, including thermal noise and noise generated in the receiver, within the bandwidth defined by the receiver pulse shaping filter. In practice this measurement quantity is rarely used in RRC_IDLE, and its usage is mainly for performing RRC_CONNECTED handovers between E-UTRAN and UTRAN FDD. The measurement periods are the same as those described in Section 22.3.2.1.

22.3.2.4 UTRAN TDD P-CCPCH RSCP

The UTRAN TDD P-CCPCH RSCP is the received power of the P-CCPCH of a UTRA TDD neighbour cell. It is considered to be the most fundamental measurement quantity for all mobility scenarios between E-UTRAN and UTRAN TDD. It is therefore used for performing both cell reselection and handover to UTRA TDD. The P-CCPCH RSCP physical layer measurement period depends on the configured gap periodicity and the DRX cycle in the same way as for UTRAN FDD.

22.3.2.5 UTRAN TDD carrier RSSI

The UTRAN TDD carrier RSSI is the total received wideband power, including thermal noise and noise generated in the receiver, within the bandwidth defined by the receiver pulse shaping filter within a specified timeslot. In principle it can be used for both cell reselection

[19]Primary Common Control Physical CHannel.

and handover from E-UTRAN to UTRAN TDD. The measurement periods of UTRAN TDD carrier RSSI with and without DRX are the same as for the P-CCPCH RSCP measurement.

22.3.3 GSM Measurements: GSM Carrier RSSI

For mobility from E-UTRAN to GSM in RRC_IDLE mode (i.e. cell reselection) and RRC_CONNECTED mode (i.e. handover) the GSM carrier RSSI is used.

The GSM carrier RSSI is measured on a GSM BCCH carrier [10]. In accordance with the GSM core specification [14] the UE is required to take at least 3 GSM carrier RSSI samples per GSM carrier as evenly spaced as possible during the physical layer measurement period. The GSM carrier RSSI physical layer measurement period without DRX and for DRX cycles up to 80 ms is 480 ms regardless of the gap pattern, and increases linearly with the number of carrier frequencies, N_{freq}, monitored during the measurement gaps. For DRX cycles larger than 80 ms, the measurement period increases with the length of the DRX cycle.

22.3.4 CDMA2000 Measurements

For mobility from E-UTRAN to CDMA2000 systems the CDMA2000 1xRTT pilot strength and CDMA2000 High Rate Packet Data (HRPD) Pilot Strength measurement quantities are used in RRC_IDLE and RRC_CONNECTED states.

22.3.4.1 CDMA2000 1xRTT Pilot Strength

CDMA2000 1xRTT Pilot Strength is the strength of the received pilot signals as defined in [15, Section 2.6.6.2.2]. The required measurement accuracy, specified in [16, Section 3.2.4], is to be fulfilled over the measurement periods of about $1.1 * N_{\text{freq}}$ seconds and $2.1 * N_{\text{freq}}$ seconds for gap patterns #0 and #1 respectively. It is used for cell reselection and handover from E-UTRAN to CDMA2000 1xRTT. A CDMA20000 1xRTT-capable UE is required to monitor CDMA2000 1xRTT Pilot Strength of CDMA2000 1xRTT cells for up to five CDMA2000 1xRTT carriers.

22.3.4.2 CDMA2000 HRPD Pilot Strength

CDMA2000 HRPD pilot strength is the strength of the received pilot signals as defined in [17, Section 8.7.6.1.2.3]. It is used for cell reselection and handover from E-UTRAN to CDMA2000 HRPD. An HRPD-capable UE is required to monitor CDMA2000 HRPD Pilot Strength of HRPD cells for up to five HRPD carriers.

22.4 UE Measurement Reporting Mechanisms and Requirements

As explained in Section 22.3, in LTE the UE reports measurements to the serving eNodeB only in RRC_CONNECTED state using the Dedicated Control CHannel (DCCH). The measurement reporting mechanism can be periodic, event-triggered or event-triggered and periodic.

These reporting mechanisms are applicable both with and without DRX. The interval between periodic reports is configurable by the eNodeB and ranges from 120 ms to 3600 s [4]. The events which can trigger reporting by the UE are described in Section 3.2.5.1.

The total number of measurement reporting criteria which can be evaluated by the UE in parallel is limited to 21. This includes 9 E-UTRA intra-frequency reporting criteria, 7 E-UTRA inter-frequency reporting criteria and 5 inter-RAT reporting criteria. The actual number and types of reporting criteria are configured by the network and can be based on any of the three possible reporting mechanisms mentioned above.

The event triggered reporting delay requirements, which also apply to the first report in event-triggered periodic reporting, are explained in the following sections for different E-UTRA and inter-RAT cases.

22.4.1 E-UTRAN Event Triggered Reporting Requirements

The E-UTRAN event-triggered reporting consists of intra- and inter-frequency event reporting. Five such events (A1-A5) and their corresponding triggering thresholds are configured in the UE by the serving eNodeB as explained in section 3.2.5.2. When an event is triggered, the UE reports the event to the eNodeB, which may take an appropriate mobility decision.

For both intra-frequency and inter-frequency event-triggered measurements, the reporting delay[20] is less than the cell search delay (see Sections 22.2.1.1 and 22.2.1.2) for the corresponding cases. Due to signal variations and user mobility, the measured quality of a cell may vary between detectable and undetectable threshold levels.

To cater for this typical scenario, requirements are specified for both intra- and inter-frequency event reporting as follows: Regardless of the number of times the cell quality varies between the detectable and undetectable levels, whenever an event (i.e. any of the intra- or inter-frequency events A1-A5) is triggered the UE is required to send the event-triggered measurement report within a duration less than the relevant physical layer measurement period without layer 3 filtering (e.g. within 200 ms RSRP/RSRQ for the intra-frequency non-DRX case), provided certain conditions are met. When layer 3 filtering is used, a reporting delay longer than the physical layer measurement period is expected. For an intra-frequency event reporting, these conditions require that the cell which has been detectable over the intra-frequency cell search delay, does not become undetectable for more than 5 s. It is also required that the timing to the intra-frequency cell does not change by more than $\pm 50T_s$.[21] In case of inter-frequency event reporting the corresponding condition requires that the timing to the inter-frequency cell does not change by more than $\pm 50T_s$ during periods when the measurement gaps are not available (i.e. between successive measurement gaps). This is because in the absence of measurement gaps the UE cannot appropriately track the timing of an inter-frequency cell.

22.4.2 Inter-RAT Event-Triggered Reporting

Event-triggered reporting requirements are specified for E-UTRAN to UTRAN FDD, E-UTRAN to UTRAN TDD and E-UTRAN to GSM inter-RAT scenarios. Two inter-RAT

[20] Here we refer to the delay for the transmitted measurement reports without layer 3 filtering being applied. The behaviour of the Layer 3 filters is standardized and their configuration provided by RRC signalling. See [4, Section 5.5.3.2] for more details.

[21] T_s is the basic time unit and is equal to $1/(15000 \times 2048)$ s.

events (B1 and B2) and their corresponding triggering thresholds are also configured in the UE by the serving eNodeB as explained in section Section 3.2.5.2. When one of these events is triggered, the UE reports it to the eNodeB, allowing it to take an appropriate inter-RAT mobility decision.

E-UTRAN to UTRAN event-triggered reporting. The UTRA FDD and TDD event-triggered measurement reporting delays without Layer 3 filtering are less than the corresponding cell search delays (see Section 22.2.3). However, if a UTRA FDD cell, which was previously detectable over a period equal to the UTRA FDD cell search delay, enters or leaves the reporting range, then the event-triggered measurement reporting delay is less than the UTRA FDD physical layer measurement period without layer 3 filtering provided that the timing to that UTRA FDD cell has not changed more than ± 32 chips when the measurement gap has not been available. Note also that the inter-RAT physical layer measurement period is typically a few times shorter than the corresponding inter-RAT cell search delay.

E-UTRAN to GSM event-triggered reporting. The event-triggered measurement reporting delay for a GSM cell with BSIC verified without layer 3 filtering is less than twice the GSM measurement period (see Section 22.3.3).

22.5 Mobility Performance

The mobility procedures comprising of cell selection and cell reselection in RRC_IDLE state and handover in RRC_CONNECTED state are described in Chapter 3. Here we focus on the performance requirements for these mobility procedures.

22.5.1 Mobility Performance in RRC_IDLE State

The performance requirements for mobility in RRC_IDLE state aim to ensure that a UE camps on a cell which guarantees good paging reception, that substantial UE battery power saving is achieved and that the interruption in paging reception during cell reselection is minimized.

All intra-frequency, inter-frequency and inter-RAT cell reselection requirements have the following characteristics:

- Requirements are specified for a selected set of typical DRX cycles: 0.32 s, 0.64 s, 1.28 s and 2.56 s;
- Cell reselection involves detection of new neighbour cells (both E-UTRA and inter-RAT), and measurement of those cells and of previously detected neighbour cells;
- Measurement and evaluation of cells are carried out at specific rates, which depend upon the DRX cycle in use;
- Cell reselection decisions are autonomously taken by the UE but are governed by predefined standardized rules, network control parameters and performance requirements;
- No performance requirements are specified for E-UTRA or inter-RAT cell identification, RSRP/RSRQ or inter-RAT measurements in RRC_IDLE, since in RRC_IDLE the UE is not required to report any event or measurement to the network.

The key elements of cell reselection in E-UTRAN are described in the following subsections.

22.5.1.1 Measurement and Evaluation of Serving Cell

The UE is required to measure both RSRP and RSRQ of the serving cell in order to evaluate the serving cell selection criterion (the 'S-criterion' – see Section 3.3) at least once every DRX cycle.

Failure of the S-criterion may imply that the UE is on the verge of losing the serving cell and it is therefore important that the UE identifies a potential new serving cell. Therefore if the S-criterion for the serving cell is not met over a certain number of consecutive DRX cycles (depending on the DRX cycle length in use), the UE initiates measurements of all the neighbour cells (i.e. over all frequency layers indicated by the serving cell) regardless of the measurement rules. For example in case of a 1.28 s DRX cycle, the UE starts measuring all the neighbour cells if the S-criterion for the serving cell is not met for 2 consecutive DRX cycles (i.e. for 2.56 s).

22.5.1.2 Intra-frequency Cell Reselection

Cell reselection to a neighbour cell on the same frequency is governed by a set of requirements for measurement and evaluation of intra-frequency cells.

Measurement of intra-frequency cells. Unlike the measurement of the serving cell, the neighbour cell measurements may not be performed every DRX cycle. The UE initiates the measurement of intra-frequency neighbour cells when the serving cell's RSRP or RSRQ fall below their respective thresholds.

However, the cell ranking for cell reselection is only based on RSRP. The UE is required to detect and measure the RSRP of neighbour cells whose received quality is above the following thresholds, without an explicit intra-frequency neighbour cell list:

- Synchronization signal $\hat{E}_s/I_{ot} \geq -6$ dB
- RSRP and synchronization signal received power $\geq$ respective thresholds, which depend on the frequency band e.g. -124 dBm for Band 1.

The RSRP of the identified intra-frequency cells is measured once every $T_{measure,\ E\text{-}UTRA\_intra}$, as shown in Table 22.3.

Table 22.3: Measurement, detection and evaluation rates for intra-frequency cells.

DRX cycle length (s)	$T_{detect,EUTRAN\_intra}$ (s) (number of DRX cycles)	$T_{measure,EUTRAN\_intra}$ (s) (number of DRX cycles)	$T_{evaluate,EUTRAN\_intra}$ (s) (number of DRX cycles)
0.32	11.52 (36)	1.28 (4)	5.12 (16)
0.64	17.92 (28)	1.28 (2)	5.12 (8)
1.28	32(25)	1.28 (1)	6.4 (5)
2.56	58.88 (23)	2.56 (1)	7.68 (3)

Evaluation of intra-frequency cells. The measured intra-frequency cells are evaluated for possible cell reselection based on cell ranking (see Section 3.3.4.3). .

If an intra-frequency cell is detectable but not yet detected then the UE is required to evaluate whether it meets the reselection criteria based on ranking within $T_{detect,EUTRAN\_Intra}$ (see Table 22.3) when $T_{reselection} = 0$. On the other hand if the cell is already detected then the UE is required to evaluate whether this intra-frequency cell meets the reselection criteria based on ranking within $T_{evaluate,EUTRAN\_Intra}$ (see Table 22.3) when $T_{reselection} = 0$, provided that the target intra-frequency cell is ranked at least 3 dB above the serving cell. If the timer $T_{reselection}$ has a non-zero value and the target intra-frequency cell is found to be better ranked than the serving cell over the $T_{reselection}$ time, then that intra-frequency cell is selected.

22.5.1.3 Inter-frequency Cell Reselection

Similarly to the reselection of intra-frequency cells, requirements are also defined for measurement and evaluation of inter-frequency cells.

Measurement of inter-frequency cells. An inter-frequency layer may have a lower, equal or higher priority than that of the serving frequency layer. The UE is required to detect and measure the relevant measurement quantity under the same conditions as are applicable for intra-frequency cells.

If the serving cell's RSRP and RSRQ are above their respective thresholds the UE searches higher-priority inter-frequency layers at least once every $T_{higher\_priority\_search}$:

$$T_{higher\_priority\_search} = (60 * N_{layers}) \text{ seconds} \quad (22.1)$$

where N_{layers} is the total number of configured higher-priority inter-frequency and inter-RAT frequency layers. The relevant measurement quantities should be measured at least every $T_{measure,\ E\text{-}UTRAN\_Inter}$ as defined in Table 22.4.

If the serving cell's RSRP and RSRQ become equal to or fall below their respective thresholds, the UE searches and measures all inter-frequency cells regardless of their priority. The relevant measurement quantity of the identified inter-frequency cells is measured every $K_{carrier} * T_{measure,\ E\text{-}UTRA\_Inter}$ seconds, as shown in Table 22.3, where $K_{carrier}$ is the number of inter-frequency carriers configured by the serving cell.

Table 22.4: Measurement and Evaluation of Inter-frequency Cells.

DRX cycle length (s)	$T_{detect,EUTRAN\_Inter}$ (s) (number of DRX cycles)	$T_{measure,EUTRAN\_Inter}$ (s) (number of DRX cycles)	$T_{evaluate,EUTRAN\_Inter}$ (s) (number of DRX cycles)
0.32	11.52 (36)	1.28 (4)	5.12 (16)
0.64	17.92 (28)	1.28 (2)	5.12 (8)
1.28	32 (25)	1.28 (1)	6.4 (5)
2.56	58.88 (23)	2.56 (1)	7.68 (3)

Evaluation of inter-frequency cells. The evaluation of inter-frequency cells for possible cell reselection is based on cell ranking for layers of equal priority and on absolute priorities for layers of different priority. Only RSRP is allowed for cell ranking, while the network can configure either RSRP or RSRQ for the evaluation of cells with unequal priorities [18].

If an inter-frequency cell of lower or equal priority is detectable but not yet detected then the UE is required to evaluate whether it meets the reselection criteria (see Sections 3.3.4.2 and 3.3.4.3) within $K_{carrier} * T_{detect,EUTRAN\_Inter}$ seconds (see Table 22.4) provided that the cell reselection criteria can be met by a certain margin (5 dB for cell ranking and 6 dB for absolute priority based reselection) and the timer value $T_{reselection}$ is zero.

However, if an inter-frequency cell is already detected then the UE is required to evaluate whether it meets the reselection criteria within a shorter duration, $K_{carrier} * T_{evaluate,EUTRAN\_Inter}$ seconds (see Table 22.4) provided the same two conditions above are satisfied.

If the timer $T_{reselection}$ has a non-zero value and the target inter-frequency cell is found to be better ranked than the serving cell over the time $T_{reselection}$, then that inter-frequency cell is reselected.

22.5.1.4 Inter-RAT Cell Reselection

Inter-RAT cell reselection covers reselection to UTRAN FDD, UTRAN TDD, GSM, CDMA2000 1xRTT and CDMA2000 HRPD.

As for inter- and intra-frequency reselection, cell reselection to an inter RAT neighbour cell is also governed by the measurement and evaluation of the inter-RAT cells.

Measurement of inter-RAT cells. An inter-RAT frequency layer may have either lower or higher priority than that of the serving frequency layer. If the quality of the serving cell is above the threshold 'Snonintrasearch'[22] the UE searches higher priority inter-RAT frequency layers at least once every $T_{higher\_priority\_search}$ according to Equation (22.1). The detected higher-priority inter-RAT cells of RAT_j are to be measured at least every $T_{measure,RAT,j}$ as defined in Table 22.5.

If the serving cell quality becomes equal to or falls below the Snonintrasearch threshold the UE searches and measures all inter-RAT cells of RAT_j (i.e. cells of higher, lower or equal priority frequency layers of RAT_j). The relevant measurement quantity of the identified inter-RAT cells is measured every $N_{RAT,j} * T_{measure,\ RAT,j}$, as illustrated in Table 22.5, where $N_{RAT,j}$ is the number of frequency layers of RAT_j configured by the serving cell (for GSM, $N_{RAT} = 1$). The relevant inter-RAT measurement quantities in RRC_IDLE states are described in Section 22.3.

Evaluation of inter-RAT cells. Evaluation of inter-RAT cells for possible cell reselection is based only on priority (see Section 3.3.4.2).

If an inter-RAT cell is detectable but not yet detected then the UE is required to evaluate whether it meets the reselection criteria within $N_{RAT,j} * T_{detect,\ RAT,\ j}$ (see Table 22.5) provided that the cell reselection criteria can be met by a 6 dB margin (in the case of UTRA) and the timer value $T_{reselection} = 0$.

[22] See Section 3.3.4.1; 'Snonintrasearch' is defined in [18, Section 5.2.4.7].

Table 22.5: Measurement and Evaluation of Inter-RAT Cells.

DRX cycle length (s)	apply to UTRA FDD and TDD	$T_{detect, RAT,j}$ (s); $T_{measure,RAT,j}$ (s) (number of DRX cycles); apply to all RATs	$T_{evaluate, RAT,j}$ (s) (number of DRX cycles); apply to UTRA FDD, TDD and CDMA2000
0.32	30	5.12 (16)	15.36 (48)
0.64	30	5.12 (8)	15.36 (24)
1.28	30		6.4 (5) 19.2 (15)
2.56	60	7.68 (3)	23.04 (9)

For an inter-RAT cell that has already been detected, the UE is required to evaluate that it meets the reselection criteria within a shorter duration, $N_{RAT,j} * T_{evaluate, RAT, j}$ (see Table 22.4) provided that the cell reselection criteria of the respective RAT are satisfied.

22.5.1.5 Paging Interruption during Cell Reselection

During intra-frequency, inter-frequency or inter-RAT cell reselection procedures, the UE monitors the serving cell for paging messages until the UE is able to monitor the paging channels of the target cell. In order to complete the cell reselection successfully and camp on the new cell the UE has to acquire the relevant system information of the target cell. Therefore, during this period paging interruption might occur as the UE is not required simultaneously to receive paging and acquire the system information of the target cell.

The interruption in paging reception should not exceed $T_{Interrupt} = T_{SI\text{-}RAT} + 50$(ms), where $T_{SI\text{-}RAT}$ is the time required for acquiring all the relevant system information of the target RAT.

22.5.2 Mobility Performance in RRC_CONNECTED State

In E-UTRAN, only hard handovers are possible (see Section 3.2.3.4) resulting in a delay including a short interruption. In order to limit the length of the interruption, requirements are specified for various E-UTRAN handover scenarios as described in the following subsections. These requirements are expressed in terms of handover delay, which is the sum of the RRC procedure delay and the interruption time. This principle applies whether the target cell is known to the UE (referred to as *non-blind handover*), or unknown (*blind handover*). The interruption time is defined as the time from the end of the last subframe containing the handover command on the Physical Downlink Shared CHannel (PDSCH) from the serving cell and the moment the UE starts transmission on the relevant uplink physical channel in the target cell. Figure 22.6 illustrates the radio interface signalling of the handover procedure to a known target.

As explained in Section 3.2.3.4, the network does not require measurement reports from target cells for performing a blind handover. This is particularly useful in case of multiple frequency layers, which can only be monitored using the same gap pattern (i.e. either pattern #0 or pattern #1). In such a scenario the network may request the UE to perform measurements on only a subset of the frequency layers while relying on blind handovers for

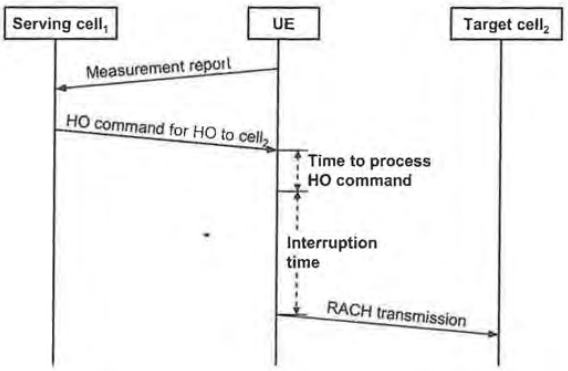


Figure 22.6: Handover to a known target cell.

the remaining ones. Handover to an unknown target cell results in a longer interruption time since the UE has to detect the target cell prior to accessing it. The process of handover to an unknown target cell is shown in Figure 22.7.

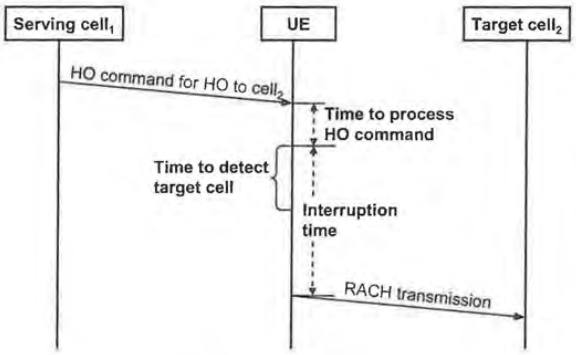


Figure 22.7: Handover to an unknown target cell.

22.5.2.1 E-UTRAN to E-UTRAN Handover

An E-UTRAN to E-UTRAN handover is completed when the UE starts transmission on the Physical Random Access CHannel (PRACH) in the target E-UTRA cell. The interruption time is expressed as $T_{interrupt} = T_{search} + T_{IU} + 20$ (ms), where:

- T_{search} is the cell search delay, which is 0 for known target cells and 80 ms for unknown target cells. The target cell is considered to be known if the cell search requirements were met in the last 5 s; otherwise the cell is considered to be unknown.
- T_{IU} is the timing uncertainty due to the timing of the PRACH occasions, and thus depends upon the PRACH configuration (see Section 17.4.2.5).

As an example, suppose the PRACH is allowed once in the middle of every frame (i.e. in every 4th subframe). If the target cell is known, the total intra-frequency or inter-frequency handover delay is 50 ms including 15 ms of RRC procedure delay (see [4, Section 11.2]). For the same PRACH configuration, in case of blind handover, the total handover delay is 130 ms including the 15 ms RRC procedure delay. In this case the cell search is much shorter than the usual cell search requirements (see Section 22.2.1) due to the fact that it is assumed that the target cell is sufficiently strong to be detected by the UE on the first correlation attempt.

22.5.2.2 Handover to Other 3GPP RATs

E-UTRAN to UTRAN handover

UTRA FDD or TDD cells can be known or unknown for the purposes of handover. The target cell is known if it has been measured by the UE during the last 5 s; otherwise it is considered unknown. An E-UTRAN-UTRAN FDD handover is completed when the UE starts transmission on the uplink Dedicated Physical Control CHannel (DPCCH) in the target cell. An E-UTRAN-UTRAN TDD handover is completed when the UE starts transmission of either the uplink DPCCH or an uplink synchronization code ('SYNC-UL') in the target cell.

Depending upon the choice of parameters, the interruption time when the target cell is known is typically in the order of 100–150 ms. When the target cell is unknown, the interruption increases by about 100 ms for UTRA FDD and 160 ms for UTRA TDD. The RRC procedure delay is 50 ms (see [12, Sections 5.3.1 and 5.3.2]) as this requirement is based on the UTRAN requirement [19]. Hence the overall handover delay can be of the order of 150–200 ms when the target UTRA (FDD or TDD) cell is known and 250–300 ms when the target cell is unknown. The exact handover delay can be derived from the relevant expressions in [12, Sections 5.3.1 and 5.3.2].

E-UTRAN to GSM handover

Handovers from the serving eNodeB to both known and unknown target GSM cells are supported in E-UTRAN. A handover to GSM is completed when the UE starts uplink transmission in the target GSM cell.

The total handover delays to known and unknown target GSM cells are 90 ms and 190 ms respectively, which includes 50 ms delay for processing the handover command (see [12, Section 5.3.3.1]).

22.5.2.3 Handover to CDMA2000 1xRTT or HRPD

Non-3GPP RAT handovers defined from E-UTRAN are for CDMA2000 1xRTT and HRPD. Handovers to both known and unknown target CDMA2000 1xRTT or HRPD cells are

possible. In both of these CDMA2000 technologies, a cell is considered known if it has been measured by the UE during the last 5 s. When performing handover to CDMA2000 1xRTT or HRPD the interruption time is influenced primarily by the following factors :

- Uncertainty due to changing the timing from the old E-UTRA serving cell to the new CDMA2000 1xRTT or HRPD cell (the delay can be up to one frame, i.e. 20 ms for CDMA2000 1xRTT and 26.7 ms for HRPD;
- The number of known and unknown target cells, which is signalled to the UE in the E-UTRAN System Information Block 8 (SIB8);
- The search window sizes for searching the known and unknown target cells; this is expressed in CDMA2000 1xRTT chips and signalled to the UE in SIB8.

The RRC procedure delay for processing the handover command is 130 ms for CDMA2000 1xRTT and 50 ms for HRPD. In order to derive the handover delay for different combinations of parameters, the reader is referred to [12, Section 5.4.2]. As a typical example the handover delay can be about 200 ms for CDMA2000 1xRTT and 130 ms for HRPD, assuming one target cell (regardless of whether it is known or unknown) and a typical window size of 60 chips.

22.6 RRC Connection Mobility Control Performance

In LTE the establishment and maintenance of the RRC connection is governed by the two main control plane functions: RRC Connection Re-establishment and Random Access.

22.6.1 RRC Connection Re-establishment

RRC connection re-establishment is initiated when a UE in RRC_CONNECTED state loses its RRC connection (e.g. due to radio link failure, handover failure or RRC connection reconfiguration failure), as specified in [4, Section 5.3.7.3]. The UE tries to re-establish the RRC connection with the strongest E-UTRA cell. Successful RRC re-establishment means that the UE is able to send the 'RRCConnectionReestablishmentRequest' message within $T_{re\text{-}establish\_delay}$, including the delay in acquiring the uplink grant for sending the message and $T_{UE\text{-}re\text{-}establish\_delay}$ which is defined as the delay from the moment when the UE detects the need for RRC re-establishment until it transmits a random access signal to the target cell:

$$T_{UE\text{-}re\text{-}establish\_delay} = 50\text{ms} + N_{freq} * T_{search} + T_{SI} + T_{PRACH} \quad (22.2)$$

where N_{freq} is the total number of E-UTRA carrier frequencies available for RRC re-establishment, T_{search} is the target cell search delay and depends upon whether the target cell is known or unknown to the UE, T_{SI} is the time required to read the target cell System Information (SI) and T_{PRACH} is the delay due to random access.

22.6.2 Random Access

In LTE the random access procedure serves several purposes as described in Sections 17.2 and 17.3. Requirements are specified for both contention-based and contention-free random

access. The primary objectives of the requirements are to ensure the correct UE behavior when performing random access, and that the UE transmit power accuracy and transmit timing error when sending random access are within suitable limits. These requirements include:

- If the UE does not receive a random access response matching its PRACH preamble identity within the random access response window, it should retransmit a preamble, up to and not exceeding the maximum number of preamble transmissions configured by the eNodeB.
- The transmit power accuracy of the UE's PRACH preamble transmissions should fulfil the requirements defined in [20, Sections 6.3.5.1 and 6.3.5.2].
- The required transmit timing accuracy of the UE's PRACH preamble transmissions depends upon the transmission bandwidth, since at lower bandwidth the UE typically uses lower sampling rate. The transmit timing error due to all PRACH transmissions (initial and subsequent preamble transmissions) should be within $\pm 24T_s$ and $\pm 12T_s$ for system bandwidths equal to 1.4 MHz and ≥ 3 MHz respectively, as specified in [12, Section 7.1.2].

22.7 Radio Link Monitoring Performance

The purpose of the radio link monitoring function in the UE is to monitor the downlink radio link quality of the serving cell in RRC_CONNECTED state and is based on the cell-specific RSs (see Section 8.2.1). This in turn enables the UE when in RRC_CONNECTED state to determine whether it is *in-sync* or *out-of-sync* with respect to its serving cell (see [21, Section 4.2.1]).

In case of a certain number of consecutive out-of-sync indications (called 'N310'), the UE starts a network-configured radio link failure timer 'T310'. The timer is stopped if a number 'N311' of consecutive in-sync indications are reported by the UE's physical layer. Both the out-of-sync and in-sync counters (N310 and N311) are configurable by the network. Upon expiry of the timer T310, Radio Link Failure (RLF) occurs. As a consequence the UE turns off its transmitter to avoid interference and is required to re-establish the RRC connection within $T_{\text{UE-re-establish\_delay}}$ as explained in Section 22.6.1. The various actions pertaining to radio link monitoring and the subsequent RRC re-establishment to the target cell are shown in Figure 22.8.

22.7.1 In-sync and Out-of-sync Thresholds

The UE's estimate of the downlink radio link quality is compared with out-of-sync and in-sync thresholds, Q_{out} and Q_{in}, for the purpose of radio link monitoring. These thresholds are expressed in terms of the BLock Error Rate (BLER) of a hypothetical Physical Downlink Control Channel (PDCCH) transmission from the serving cell (see [12, Section 7.6] for details). Specifically, Q_{out} corresponds to a 10% BLER while Q_{in} corresponds to a 2% BLER. The same threshold levels are applicable with and without DRX.

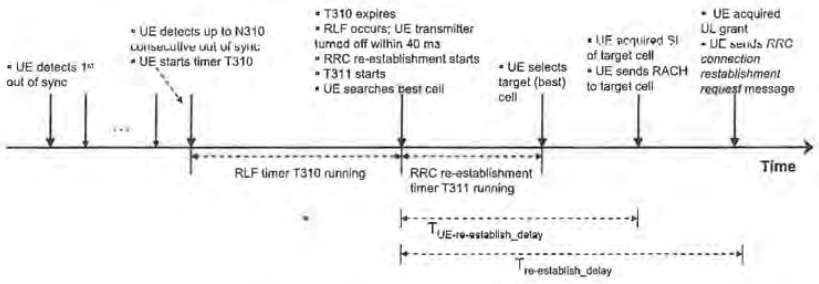


Figure 22.8: Radio link monitoring of the serving cell followed by RRC re-establishment to the target cell.

The mapping between the cell specific RS based downlink quality and the hypothetical PDCCH BLER is up to the UE implementation. However, the performance is verified by conformance tests defined for various environments [22].

22.7.2 Requirements without DRX

When no DRX is configured, out-of-sync occurs when the downlink radio link quality estimated over the last 200 ms period becomes worse than the threshold Q_{out}. Similarly without DRX the in-sync occurs when the downlink radio link quality estimated over the last 100 ms period becomes better than the threshold Q_{in}. Upon detection of out-of-sync, the UE initiates the evaluation of in-sync. The occurrences of out-of-sync and in-sync are reported internally by the UE's physical layer to its higher layers, which in turn may apply layer 3 (i.e. higher layer) filtering for the evaluation of RLF.

22.7.3 Requirements with DRX

When DRX is in use, in order to enable sufficient UE power saving the out-of-sync and in-sync evaluation periods are extended and depend upon the configured DRX cycle length. The UE starts in-sync evaluation whenever out-of-sync occurs. Therefore the same period ($T_{Evaluate\_Q_{out\_DRX}}$) is used for the evaluation of out-of-sync and in-sync. However, upon starting the RLF timer (T310) until its expiry, the in-sync evaluation period is shortened to 100 ms, which is the same as without DRX. If the timer T310 is stopped due to N311 consecutive in-sync indications, the UE performs in-sync evaluation according to the DRX-based period ($T_{Evaluate\_Q_{out\_DRX}}$).

22.7.4 Requirements during Transitions

In LTE a transition phase is caused by switching between DRX and non-DRX operation or switching between short and long DRX or vice versa. These scenarios can occur often, and

therefore the UE behaviour is specifically defined for the evaluation of the radio link quality during the transition period. There are two main aspects of the requirements:

- Length of the transition period, T_P;
- Evaluation period, $T_{Evaluate\_Transition}$, during the transition.

The transition period T_P is equal to the evaluation period of the mode after the transition. During this phase the evaluation period $T_{Evaluate\_Transition}$ is defined as follows:

$$T_{Evaluate\_Transition} \geq \min(T_{Evaluate\_mode1}, T_{Evaluate\_mode2}) \quad (22.3)$$

where $T_{Evaluate\_mode1}$ and $T_{Evaluate\_mode2}$ correspond to the evaluation periods of the first and the second mode respectively.

Equation (22.3) applies to both in-sync and out-of-sync evaluations. After the transition period, the UE uses an evaluation period corresponding to the second mode. The evaluation periods during the transition from short to long DRX and vice versa are illustrated in Figure 22.9.

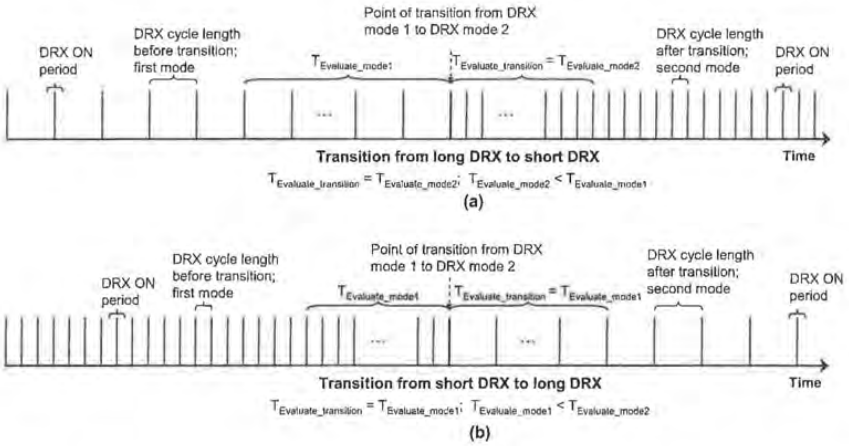


Figure 22.9: Radio link monitoring evaluation period (a) during transition from long DRX to short DRX, and (b) during transition from short DRX to long DRX.

22.8 Concluding Remarks

In this chapter we have explained the main aspects of the LTE RRM performance requirements. These requirements play a key role in ensuring that the UE meets the desired mobility performance in a wide range of practical scenarios envisaged for LTE. More specifically, by virtue of these requirements, robust mobility in both RRC_IDLE and RRC_CONNECTED

states is ensured, both within E-UTRAN and between E-UTRAN and other RATs including UTRAN FDD/TDD, GSM and CDMA2000 technologies.

References[23]

[1] 3GPP Technical Specification 36.211, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation', www.3gpp.org.

[2] Texas Instruments, NXP, Motorola, Ericsson, and Nokia, 'R4-072215: Simulation Assumptions for Intra-frequency Cell Identification', 3GPP TSG RAN WG4, meeting 45, Jeju, Korea, November 2007, www.3gpp.org.

[3] NXP, 'R4-080691: LTE Cell Identification Performance in Multi-cell Environment', www.3gpp.org 3GPP TSG RAN WG4, meeting 46bis, Shenzen, China, February 2008.

[4] 3GPP Technical Specification 36.331, ' Evolved Universal Terrestrial Radio Access (E-UTRA), Radio Resource Control (RRC); Protocol specification', www.3gpp.org.

[5] 3GPP Technical Specification 36.300, 'Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2', www.3gpp.org.

[6] 3GPP Technical Specification 25.101, 'User Equipment (UE) Radio Transmission and Reception (FDD)', www.3gpp.org.

[7] M. Mouly and M.-B. Pautet, *The GSM System for Mobile Communications*. Cell & Sys, 1992.

[8] 3GPP Technical Specification 45.005, ' Radio transmission and reception', www.3gpp.org.

[9] 3GPP Technical Specification 23.272, 'Circuit Switched (CS) fallback in Evolved Packet System (EPS); Stage 2', www.3gpp.org.

[10] 3GPP Technical Specification 36.214, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Measurements', www.3gpp.org.

[11] Ericsson, 'R1-073041: Reference Signal Received Quality, RSRQ, Measurement', www.3gpp.org, 3GPP TSG RAN WG1, meeting 49bis, Orlando, USA, June 2007.

[12] 3GPP Technical Specification 36.133, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Requirements for Support of Radio Resource Management', www.3gpp.org.

[13] M. Kazmi et al., 'Evaluation of Inter-Frequency Quality Handover Criteria in E-UTRAN', IEEE 69th Vehicular Technology Conference, Spring 2009.

[14] 3GPP Technical Specification 45.008, 'Radio subsystem link control', www.3gpp.org.

[15] 3GPP2 Technical Specification C.S0005-D, 'Upper Layer (Layer 3) Signaling Specification for cdma2000 Spread Spectrum Systems', www.3gpp2.org.

[16] 3GPP2 Technical Specification C.S0011-C, 'Recommended Minimum Performance Standards for cdma2000 Spread Spectrum Mobile Stations', www.3gpp2.org.

[17] 3GPP2 Technical Specification C.S0024-A, 'cdma2000 High Rate Packet Data Air Interface Specification', www.3gpp2.org.

[18] 3GPP Technical Specification 36.304, 'Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) Procedures in Idle Mode', www.3gpp.org.

[19] 3GPP Technical Specification 25.331, 'Radio Resource Control (RRC) Protocol Specification', www.3gpp.org.

[20] 3GPP Technical Specification 36.101, 'Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) Radio Transmission and Reception', www.3gpp.org.

[23] All web sites confirmed 1st March 2011.

[21] 3GPP Technical Specification 36.213, 'Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures', www.3gpp.org.

[22] 3GPP Technical Specification 36.521-3, 'Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) conformance specification; Radio Transmission and Reception; Part 3: Radio Resource Management Conformance Testing', www.3gpp.org.

29

Multiple Antenna Techniques for LTE-Advanced

Alex Gorokhov, Amir Farajidana, Kapil Bhattad, Xiliang Luo and Stefan Geirhofer

Multiple antenna techniques play a key role in LTE-Advanced. In the downlink, the goals are to support higher data rates than LTE Releases 8 and 9 through high-order Single-User Multiple-Input Multiple-Output (SU-MIMO), and higher spectral efficiency via enhanced Multi-User MIMO (MU-MIMO) techniques. To support these advances, new reference signals and enhanced UE feedback are introduced. In the uplink, SU-MIMO is introduced, and the control channel performance is enhanced using transmit diversity.

This chapter explains the techniques adopted for Release 10 and gives some insight into additional MIMO enhancement features that may be developed in future releases of LTE-Advanced.

29.1 Downlink Reference Signals

As discussed in detail in Section 8.2, LTE Release 8 provides cell-specific Reference Signals (RSs), also known as Common Reference Signals (CRSs) for up to 4 antenna ports. Cell-specific RSs are used by UEs both to perform channel estimation for demodulation of data and to derive feedback on the quality and spatial properties of the downlink radio channel. Together with signalling of the precoder used for data transmissions to the UE, the four cell-specific RS ports enable spatial multiplexing of up to four layers using codebook-based precoding, as explained in Section 11.2.2.2. LTE Release 9 additionally supports two-layer beamforming spatial multiplexing using precoded UE-specific RSs, which enable non-codebook-based precoding to be used, as explained in Section 11.2.2.3.

LTE – The UMTS Long Term Evolution: From Theory to Practice, Second Edition.
Stefania Sesia, Issam Toufik and Matthew Baker.
© 2011 John Wiley & Sons, Ltd. Published 2011 by John Wiley & Sons, Ltd.

For LTE-Advanced, downlink SU-MIMO transmission is extended to support up to eight spatial layers, and for this purpose the precoded UE-specific RS approach is further developed for the data demodulation. This has the advantage that transmission of the new Release 10 UE-specific RSs can be limited to only those Resource Blocks (RBs) where they are needed for demodulation, thus avoiding incurring a large overhead across the whole system bandwidth (as would have been the case with an extension to the cell-specific RSs, which would also have adversely impacted 'legacy' pre-Release 10 UEs that are unaware of the new RS structure).

In addition, in order to enable the UE to estimate and feed back the Channel State Information (CSI) corresponding to up to eight antenna ports across a wide bandwidth, new RSs referred to as CSI-RSs are provided. Since CSI-RSs are used only for feedback purposes they can be sparse and incur only a small overhead.

These two new types of RS are explained in detail in the following subsections.

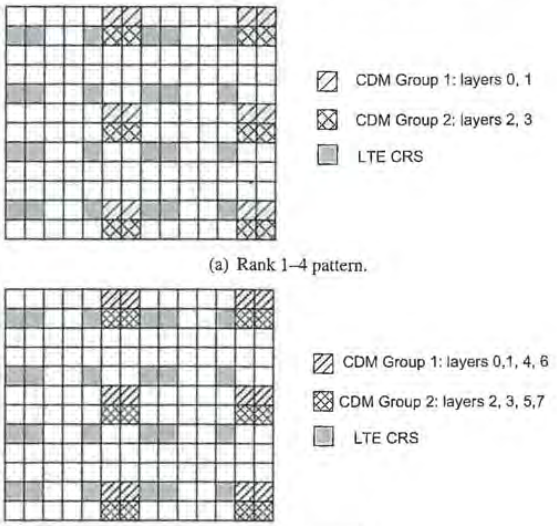
29.1.1 Downlink Reference Signals for Demodulation

In the same way as the UE-specific RSs of earlier releases (see Sections 8.2.3 and 8.2.3), the extended UE-specific RSs of LTE-Advanced are embedded in the RBs used for the Physical Downlink Shared CHannel (PDSCH) for a specific UE. The UE-specific RSs for each layer undergo the same precoding as the data symbols, and therefore there is no need for explicit signalling of precoding information to the UE. A variety of multi-antenna beamforming techniques can therefore be supported efficiently and transparently to the UE.

The pattern of Resource Elements (REs) designed for the new UE-specific RSs had to satisfy certain criteria. Firstly, it had to avoid overlapping with the cell-specific RSs and the control channels in order to ensure backward compatibility. Secondly, the UE-specific RSs of different layers should be orthogonally multiplexed to avoid inter-layer RS interference degrading the channel estimation accuracy; this could in principle be achieved by assigning different sets of REs to the RSs of different layers (by Frequency-Division and/or Time-Division Multiplexing FDM/TDM), and/or by Code Division Multiplexing (CDM). CDM has the advantages of more flexible power balancing across different layers and potentially better interference estimation accuracy in MU-MIMO scenarios.

For Release 10, the UE-specific RS pattern for up to 2 layers (referred to as 'rank-2' transmission) is identical to that defined in Release 9, in order to ensure backward compatibility. For up to 4 layers, the pattern is obtained by extending the Release 9 rank-2 UE-specific RS pattern in a hybrid CDM/FDM fashion, as shown in Figure 29.1(a) (for the case of the normal Cyclic Prefix (CP) length)[1]. The four precoded layers (antenna ports 7–10) are grouped into two groups of 2 REs, with the same length-2 Walsh–Hadamard Orthogonal Cover Codes (OCC) as in Release 9 (i.e. [1, 1] and [1, −1]) being used to multiplex the layers within each group. The UE-specific RSs in different groups are frequency-multiplexed on adjacent subcarriers. This pattern has been shown to provide reasonable performance for a variety of channel conditions and UE speeds [2, 3].

[1] By following the same principles, similar patterns are defined in Release 10 for the case of the extended CP and for the Downlink Pilot TimeSlot (DwPTS) field of the special mixed downlink-uplink subframe in TDD operation – see [1, Section 6.10.3]. Use of the extended CP with more than two UE-specific antenna ports is not supported in Release 10.



(a) Rank 1–4 pattern.

(b) Rank 5–8 pattern.

Figure 29.1: Release 10 UE-specific RS patterns.

For 8-layer transmission (antenna ports 7–14), the UE-specific RS structure of Figure 29.1(a) is further extended using hybrid CDM/FDM with two CDM groups each using a Walsh–Hadamard OCC of length 4 as shown in Figure 29.1(b). Exactly the same set of REs are used as for the rank-4 case, with the UE-specific RSs for each layer now being spread across four REs, all having the same frequency location but different time locations within the subframe. This approach maintains the power-balancing property of the rank-4 design and is optimized for pedestrian mobility – the most likely scenario for 8-layer SU-MIMO.

The length-2 and length-4 OCCs have a nested structure (the OCC of length two is nested into the OCC of length four) which ensures backward compatibility with the Release 9 design.

The mapping of the OCCs to REs is shown in Figure 29.2 for the case of normal CP. The mapping for odd and even RBs (in the frequency domain) is different in order to provide orthogonality in both time and frequency. Such orthogonality improves the performance of channel estimation in higher Doppler scenarios and reduces the inter-subcarrier interference across the two CDM groups in the presence of frequency offsets. This OCC mapping also enables power balancing across OFDM[2] symbols by peak power randomization [4].

The RS sequence for up to 8 layers is the same as in Release 9 (as described in Section 8.2.3), except that for antenna ports 9–14 only one sequence initialization is possible.

[2]Orthogonal Frequency Division Multiplexing.

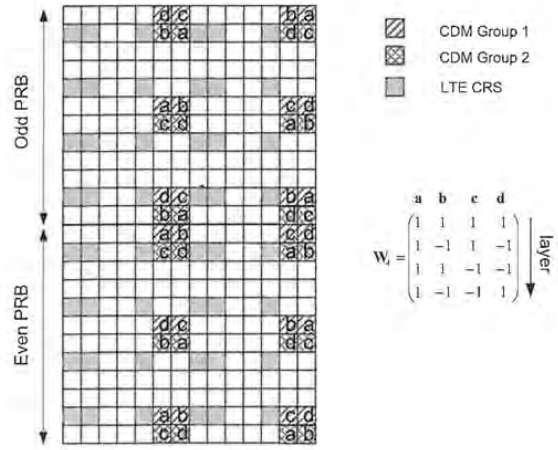


Figure 29.2: Length-4 OCC mapping to REs.

29.1.2 Downlink Reference Signals for Estimation of Channel State Information (CSI-RS)

The main goal of CSI-RSs is to obtain channel state feedback for up to eight transmit antenna ports to assist the eNodeB in its precoding operations. Release 10 supports transmission of CSI-RS for 1, 2, 4 and 8 transmit antenna ports. CSI-RSs also enable the UE to estimate the CSI for multiple cells rather than just its serving cell, to support future multicell cooperative transmission schemes (see Section 29.5.1).

The following general design principles can be identified for the CSI-RS:

- In the frequency domain, uniform spacing of CSI-RS locations is highly desirable, as explained in Chapter 8.
- In the time domain, it is desirable to minimize the number of subframes containing CSI-RS, so that a UE can estimate the CSI for different antenna ports and even different cells with a minimal wake-up duty cycle when the UE is in Discontinuous Reception (DRX) mode, to preserve battery life.
- The overall CSI-RS overhead involves a trade-off between accurate CSI estimation for efficient operation and minimizing the impact on legacy pre-Release 10 UEs which are unaware of the presence of CSI-RS and whose data are punctured by the CSI-RS transmissions. Figure 29.3 shows that a CSI-RS density of one RE per RB per antenna port is a good choice, as the throughput degradation compared to ideal CSI estimation is negligible.
- CSI-RSs of different antenna ports within a cell, and, as far as possible, from different cells, should be orthogonally multiplexed to enable accurate CSI estimation.

- To ensure backward compatibility, CSI-RSs should avoid REs used for cell-specific RSs and control channels, as well as avoiding REs used for the Release 10 UE-specific RSs.

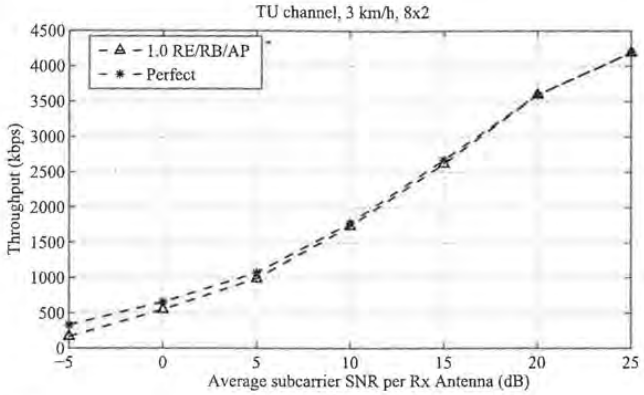


Figure 29.3: Throughput performance with a CSI-RS density of 1 RE per RB per antenna port: uncorrelated 8×2 SU-MIMO, 5 MHz, Typical Urban channel model, 3 km/h.

Taking these considerations into account, the CSI-RS patterns selected for Release 10 are shown in Figure 29.4. CDM codes of length 2 are used, so that CSI-RSs on two antenna ports share two REs on a given subcarrier.

The pattern shown in Figure 29.4(a) can be used in both frame structure 1 (FDD)[3] and frame structure 2 (TDD).[4] In Figure 29.4, the REs used for CSI-RSs are labelled using two letters, the first indicating the cell index and the second referring to the antenna ports of the CSI-RS transmitted on that RE. These patterns follow a 'nested' structure, meaning that the REs used in the case of two CSI-RS antenna ports are a subset of those used for four and eight antenna ports; this helps to simplify the implementation. The total number of supported antenna ports is 40, which can be used to give a frequency-reuse factor of 5 between cells with 8 antenna ports per cell, or a factor of 20 in the case of 2 antenna ports. It can be seen that collisions may occur with REs used for the UE-specific RS antenna ports defined in Release 8 for PDSCH transmission mode 7 (see Section 8.2.2); it may therefore be desirable to avoid scheduling UEs in transmission mode 7 in subframes containing CSI-RS.

The pattern shown in Figure 29.4(b) can only be used for frame structure 2 for TDD operation. This pattern is designed to avoid collisions with the Release 8 UE-specific RS antenna port, as this port is more suited to TDD operation where channel reciprocity can be more effectively exploited to support the beamforming. However, this pattern only offers frequency reuse factors of between 3 and 12 depending on the number of antenna ports per cell.

[3]Frequency Division Duplex.
[4]Time Division Duplex.

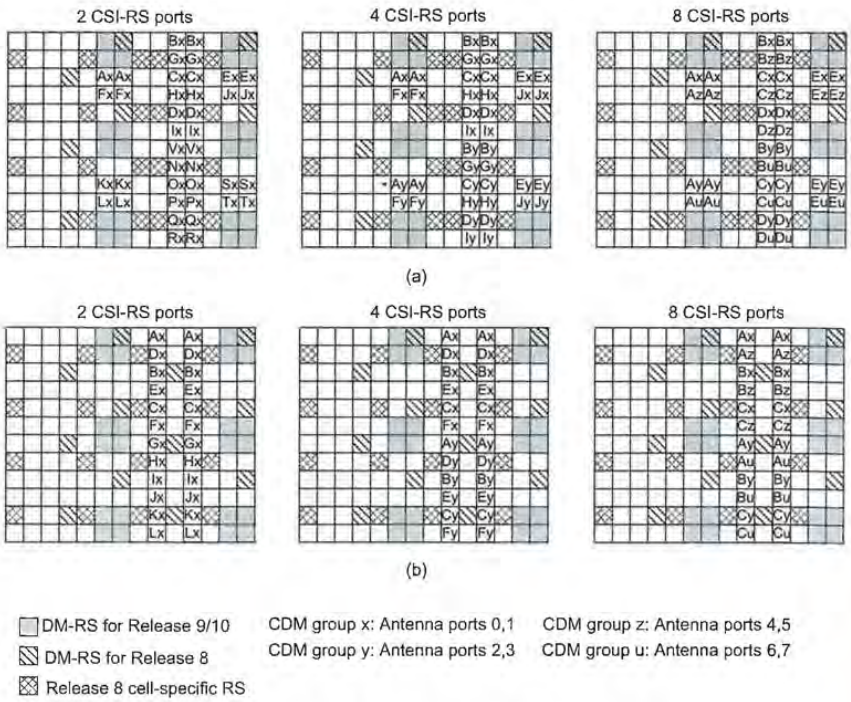


Figure 29.4: CSI-RS patterns for LTE-Advanced.

Similar CSI-RS patterns are provided for the case of the extended CP and the DwPTS field of the special mixed downlink-uplink subframe in frame structure 2 (see [1, Section 6.10.5]).

The CSI-RS configuration is UE-specific. When configured, CSI-RSs are present only in some specific subframes following a given duty cycle and subframe offset. The duty cycle and offset of the subframes containing CSI-RSs and the CSI-RS pattern used in those subframes are provided to a Release 10 UE through RRC signalling. The duty cycle and subframe offset are jointly coded, while the CSI-RS pattern is configured independently of these two parameters.

It should be noted that, in subframes containing CSI-RSs, rate-matching (see Section 10.3.2.4) for PDSCH transmissions to Release 10 UEs assumes that the CSI-RS REs are not available for PDSCH data, and the coded PDSCH data is only mapped to the surrounding REs. However, PDSCH transmissions to Release 8 and 9 UEs are punctured by CSI-RS transmissions. The density and periodicity of the CSI-RSs should therefore be chosen such that the impact of puncturing on the performance of these users is acceptable.

In the context of cooperative MIMO, it may be possible to improve the performance of channel estimation, and especially interference estimation, by coordinating CSI-RS

transmissions across multiple cells. In Release 10 it is therefore possible to 'mute' a specific set of REs in data transmissions from a cell. The locations of these REs, known as the 'muting pattern', can be chosen to avoid colliding with CSI-RS transmissions from other cells and hence improve the inter-cell measurement quality. The muting pattern is indicated to the UEs by a 16-bit bitmap, where each bit corresponds to a 4-port CSI-RS configuration for frame structure 1 or 2 as shown in Figure 29.4(a). All the REs in an indicated 4-port CSI-RS configuration are muted, and the UE can assume zero transmission power on those REs unless they are used to transmit CSI-RSs.

In subframes in which muting is configured, data transmissions for Release 10 UEs are rate-matched around the muted REs in the same way as for CSI-RSs. Data transmitted to Release 8 and 9 UEs is punctured in the muted REs.[5]

29.2 Uplink Reference Signals

In Chapter 15, the LTE Release 8 uplink RS design is explained, including DeModulation RSs (DM-RSs) and Sounding RSs (SRSs). In Release 10, the uplink DM-RSs are extended to support uplink SU-MIMO transmission with up to four spatial layers, and the SRSs are enhanced to provide improved support for channel sounding.

29.2.1 Uplink DeModulation Reference Signals (DM-RS)

In order to support uplink SU-MIMO transmission with up to 4 spatial layers, DM-RSs for all spatial layers need to be multiplexed together to enable channel estimation of each layer at the eNodeB. Like the downlink UE-specific RSs, the uplink DM-RSs are precoded using the same precoding as the Physical Uplink Shared CHannel (PUSCH) data transmissions in the same RBs when multiple antenna transmission is used.

Release 10 uses different Cyclic Shifts (CSs) of the DM-RS base sequence to multiplex the DM-RSs for the different layers. For example, in a rank-4 SU-MIMO transmission, CS offsets 0, 6, 3 and 9 are used relative to the base sequence, for layers 1, 2, 3 and 4 respectively. In addition to separation of the antenna ports by selection of different CSs, two length-2 OCCs can be applied to the two DM-RS symbols in the two slots of one subframe to further separate the multiplexed DM-RSs. The codes used are [+1, +1] and [+1 −1]; they enable better orthogonality to be achieved between DM-RSs that use close CSs. Figure 29.5 shows an example of the link-level throughput performance improvement that can be achieved thanks to the better orthogonality provided by the use of the OCCs in addition to the CS separation.

Different OCCs can also be assigned to different UEs for uplink MU-MIMO transmission. In this way, the DM-RSs from different UEs can be kept orthogonal to each other even if the PUSCH transmission bandwidths differ between the UEs (provided that there is no sequence group hopping (see Section 15.3.1) between the two slots within one subframe).

In order to reduce the control signalling overhead and make the benefits of the OCCs available to both SU-MIMO and MU-MIMO UEs, the identities of the OCCs applied to each spatial layer are implicitly derived from the 3-bit cyclic shift index signalled in the corresponding PDCCH uplink grants (see Section 9.3.5.1). The cyclic shift indices are divided into two groups: one group of 4 cyclic shift indices is targeted for SU-MIMO, and

[5] See [5, 6] for more background.

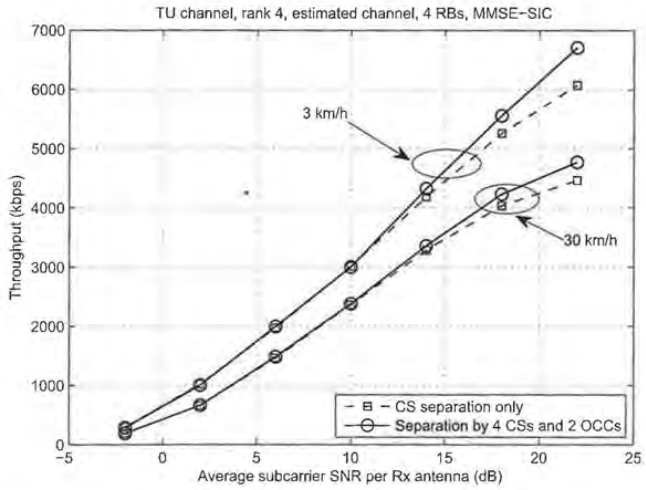


Figure 29.5: The benefit of the uplink DM-RS OCCs in terms of link throughput.

the layers with close CSs are assigned different OCCs; the other group is optimized for MU-MIMO, and the same OCC is applied to the different layers.

29.2.2 Sounding Reference Signals (SRSs)

The SRSs are important for uplink channel sounding to support dynamic uplink resource allocation, as well as for reciprocity-aided beamforming in the downlink, especially in TDD operation.

As explained in Section 15.6, the Release 8 SRSs allow for uplink sounding from a single transmit antenna only. With the introduction of uplink multiple-antenna transmission in Release 10, SRS transmission from all the uplink antenna ports is needed in order to enable all the spatial dimensions of the channel to be sounded. The same mechanisms as are available in Release 8 are used to separate the SRS transmissions from different antenna ports – i.e. different CSs and different transmission comb indices of the IFDMA[6] structure.

Nevertheless, the need for SRSs from up to four antenna ports increases the demand on SRS resources if all transmit antennas are to be sounded at a reasonable rate, and therefore semi-static configuration of the resources via higher layer (RRC) signalling is less appropriate than it was for Release 8. Therefore, Release 10 introduces the possibility of dynamically triggering aperiodic SRS transmissions via the Physical Downlink Control CHannel (PDCCH); these dynamic aperiodic SRS transmissions are known as 'type 1' SRSs, while the Release 8 RRC-configured SRSs described in Section 15.6 are known as 'type 0' in Release 10.

[6]Interleaved Frequency Division Multiple Access.

An indicator in an uplink resource grant on the PDCCH can be used to trigger a single type 1 SRS transmission. This facilitates rapid channel sounding to respond to changes in traffic or channel conditions, without tying up SRS resources for a long period. In DCI Format 0, one new bit can indicate activation of a type 1 SRS[7] according to a set of parameters that is configured beforehand by RRC signalling. In DCI Format 4, which is used for scheduling uplink SU-MIMO transmissions, two new bits allow one of three sets of RRC-configured type 1 SRS transmission parameters to be triggered (the remaining state indicates no type 1 SRS activation).

In the case of uplink multiple-antenna transmission (see Section 29.4), SRSs may be triggered on multiple antenna ports simultaneously.

Other enhancements to the SRSs, including coordination of SRS transmission resource configurations and frequency-hopping patterns, may also be considered for cooperative MIMO schemes and inter-cell interference coordination in the future (see, for example, [7] and references therein).

29.3 Downlink MIMO Enhancements

The main enhancements to the downlink multiple antenna transmission schemes in Release 10 are the extension of SU-MIMO to support 8-layer transmission, and improved support for MU-MIMO. These are introduced in a new PDSCH transmission mode, known as Transmission Mode 9.

29.3.1 Downlink 8-Antenna Transmission

The new RSs to support transmissions from 8 antenna ports in Release 10 are explained in Section 29.1.

For the data transmissions from 8 antenna ports, the Release 8/9 transmit diversity and spatial multiplexing modes are extended. For transmit diversity, no new schemes are specified in Release 10. Instead, if an eNodeB has more than four antennas and wishes to use transmit diversity, the Release 8 SFBC/FSTD[8] schemes are re-used via 'antenna virtualization'; this means that multiple antennas are treated as a single port by the application of suitable precoding, with a single Release 8 cell-specific RS being transmitted from all the antennas comprising the antenna port. The antenna virtualization is thus transparent to the UE.

For 8-layer spatial multiplexing, two aspects of the codeword-to-layer mapping needed to be considered, namely the number of codewords and the mapping of the codewords to layers. An increase in the number of codewords from the two supported in Release 8 could in theory have provided a throughput gain for some receivers such as those based on Minimum Mean Squared Error Serial Interference Cancellation (MMSE-SIC), but on the other hand a larger signalling overhead would have been incurred in both downlink and uplink, and the complexity of link adaptation would have been increased. Figure 29.6 shows the throughput performance of an 8×8 SU-MIMO transmission scheme with different numbers of codewords for a SIC receiver and a Linear MMSE (LMMSE) receiver. It can be seen that the performance gain from increasing the number of codewords beyond two is small even for the MMSE-SIC receiver.

[7] Only applicable if the DCI message is sent in the UE-specific search space – see Section 9.3.5.5.

[8] Space-Frequency Block Code / Frequency Switched Transmit Diversity (SFBC/FSTD).

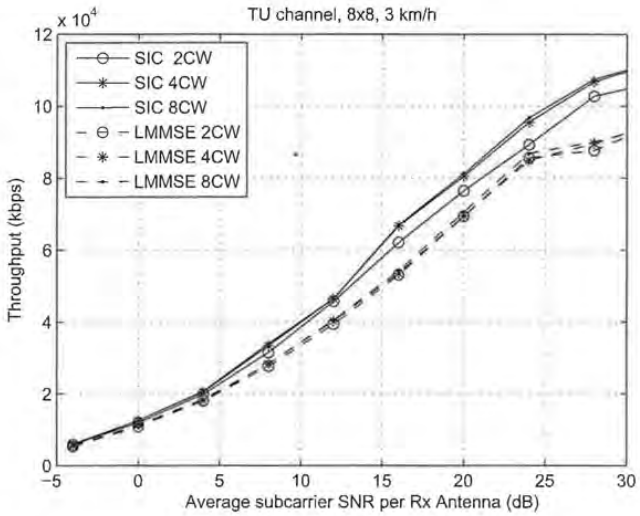


Figure 29.6: Throughput performance of LMMSE and MMSE-SIC receivers with different numbers of codewords.

Therefore Release 10 continues to use a maximum of two codewords even in the case of eight antenna ports. The codeword-to-layer mapping for 5 to 8 layers is illustrated in Figure 29.7. Note that, for up to 4 antenna ports, the mapping is the same as in Release 8.

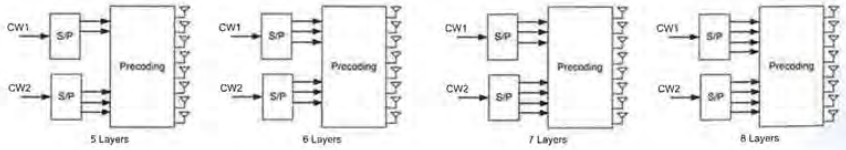


Figure 29.7: Codeword-to-layer mapping for up to 8 antenna ports.

As noted in Section 29.1, the Release 10 high-rank PDSCH transmissions use UE-specific RSs. This means that they are not constrained to being transmitted in subframes where the Release 8 cell-specific RSs are transmitted, and PDSCH transmission mode 9 in Release 10 therefore allows the PDSCH to be transmitted in the data region of 'MBSFN'[9] subframes, provided that those subframes are not actually used for MBMS[10] transmissions.

The use of non-codebook-based precoding for the PDSCH means that the eNodeB has a substantial degree of freedom in determining the actual precoding to apply for each UE

[9]Multimedia Broadcast Single Frequency Network.
[10]Multimedia Broadcast/Multicast Service.

and for each RB. However, in some cases the eNodeB may use the same precoding matrix across several RBs. For example, when a UE is configured to feed back sub-band Precoding Matrix Indicators (PMIs), the eNodeB is quite likely to use the same precoding matrix across the RBs in a sub-band, in the absence of other relevant information. In this case, the UE can improve its channel estimation quality if it has information about the RBs on which the same precoding matrix is applied. Therefore, in Release 10 it is specified that when the UE is configured to feed back PMI, it can assume that the precoder is the same across all the RBs within a *Precoding Resource block Group* (PRG). In resource allocation type 0 (see Section 9.3.5.4), the data allocation to the UE is performed in groups of contiguous RBs called Resource Block Groups (RBG); in this case, the PRG size is chosen such that it is a factor of the RBG size, in order to be consistent with the resource allocations to individual UEs. The sizes of the PMI sub-bands, RBGs and PRGs for different bandwidths are shown in Table 29.1.

Table 29.1: PRG sizes.

System bandwidth	Sub-band size	RBG size	PRG size
<10	4	1	1
11–26	4	2	2
27–63	6	3	3
64–110	8	4	2

29.3.2 Enhanced Downlink Multi-User MIMO

MU-MIMO was introduced in LTE Releases 8 and 9 as explained in Section 11.1.4. Further enhancements to MU-MIMO operation are included in PDSCH transmission mode 9 in Release 10.

Like PDSCH transmission mode 8 in Release 9, transmission mode 9 in Release 10 supports dynamic switching between SU-MIMO and MU-MIMO. The provision of UE-specific RS patterns with rank greater than one enables the scheduling of more than one spatial stream to UEs participating in MU-MIMO. Cell-specific scrambling of the UE-specific RS sequences, as well as orthogonal multiplexing of UE-specific RS ports assigned to co-scheduled UEs, enables accurate channel estimation and improved interference nulling at the UE.

Although Release 10 supports orthogonal multiplexing of up to 8 layers, the maximum intended multi-user multiplexing order is 4 in Release 10. The gains, if any, of supporting more than 4 layers with MU-MIMO are expected to be small since the channel estimation quality is expected to drastically decrease, and feedback inaccuracies make efficient selection of UEs for co-scheduling difficult.

Another salient aspect of MU-MIMO that was introduced in LTE Release 9 and is carried over into LTE-Advanced is the transparency of MU-MIMO operation from the point of view of the UE. Specifically, no signalling is provided to inform UEs of the presence of

co-scheduled transmissions, nor of any parameters of such transmissions (such as rank, antenna port, RB allocation or power offset); in particular, the use of UE-specific RSs whose amplitude can be varied along with the data symbols means that the Release 8 signalling of a power offset to indicate that the eNodeB's transmit power in an RB is divided between multiple UEs is no longer necessary. Control signalling is limited to indicating the set of UE-specific RS ports allocated to the UE itself. As in Release 9, the 4-layer multiplexing is achieved by using two scrambling sequences and the rank-2 UE-specific RS pattern.

A new DCI format (Format 2C) is introduced in PDSCH transmission mode 9 to support dynamic switching between SU-MIMO transmission (up to rank 8) and MU-MIMO transmission (up to rank 4). DCI Format 2C is based on Format 2B with the addition of 2 new bits; also, the SCrambling IDentity (SCID) bit of Format 2B is reused to encode jointly the number of layers, UE-specific RS SCID and antennna port mapping. The encoding is shown in Table 29.2.[11]

Table 29.2: Encoding of antenna ports, layers and SCID in DCI Format 2C.

One codeword: Codeword 0 enabled, Codeword 1 disabled		Two codewords: Both codewords enabled	
State	Message	State	Message
0	1 layer, port 7, SCID=0	0	2 layers, ports 7–8, SCID=0
1	1 layer, port 7, SCID=1	1	2 layers, ports 7–8, SCID=1
2	1 layer, port 8, SCID=0	2	3 layers, ports 7–9
3	1 layer, port 8, SCID=1	3	4 layers, ports 7–10
4	2 layers, ports 7–8	4	5 layers, ports 7–11
5	3 layers, ports 7–9	5	6 layers, ports 7–12
6	4 layers, ports 7–10	6	7 layers, ports 7–13
7	Reserved	7	8 layers, ports 7–14

It is also well known that the system capacity of MU-MIMO is highly dependent on the accuracy of the CSI feedback from the UEs. Therefore, improved feedback is also a key factor for improved MU-MIMO performance; this is discussed in the next section.

29.3.3 Enhanced CSI Feedback

In Release 10, the same CQI/PMI/RI[12] feedback types are used as in Release 8. As explained in Section 11.2.3, this is often described as *implicit* feedback, as it provides an implicit representation of the channel consisting of an indication of the data rate that could be achieved if the eNodeB used a certain precoder.

[11] Note that transmissions with more than one layer but a single codeword only occur in the case of retransmissions of SU-MIMO transmissions. Hence, for transmission with one codeword and 2 layers, only one scrambling sequence is used.

[12] Channel Quality Indicator / Precoding Matrix Indicator / Rank Indicator.

The concept of implicit (i.e. recommended precoder) feedback can be compared to what is sometimes called *explicit feedback* (not supported in LTE or LTE-Advanced), whereby a UE would instead explicitly report a quantized representation of the physical CSI without making assumptions about the nature of the eNodeB precoder. Since the capacity achieved by MU-MIMO is very dependent on the accuracy of the CSI at the transmitter, explicit feedback mechanisms are often favoured in theoretical studies. However, the use of implicit feedback for MU-MIMO in LTE does have some practical advantages, including minimizing the implementation effort required to support both SU- and MU-MIMO. Furthermore, the accuracy of a UE's implicit feedback can very easily be tested, by transmitting data to the UE using the UE's recommended precoder and the Modulation and Coding Scheme (MCS) corresponding to the UE's CQI report.

Moreover, it can be argued that, under certain assumptions, implicit and explicit feedback schemes have similar results. As an example, in the case of rank-1 feedback, an explicit short-term CSI report would signal the direction of the highest channel gain, i.e. the strongest eigenvector of the channel matrix, after some form of quantization to reduce the overhead to a reasonable level. On the other hand, with implicit feedback, a UE would select the vector from the precoding codebook that yields the largest Signal-to-Interference-plus-Noise Ratio (SINR), which is necessarily an approximation of the strongest eigenvector, since the latter identifies the direction of maximum beamforming gain of the channel.

29.3.3.1 PMI Feedback Codebook Enhancements

Enhanced spatial feedback in PDSCH transmission mode 9 is available for the case when 8 CSI-RS ports are configured.

The principle of this enhanced feedback is based on an observation that large antenna arrays at the eNodeB typically comprise two cross-polarized subarrays, in which each subarray is a Uniform Linear Array (ULA). A close spacing of the antenna elements (of the order of half a wavelength) gives rise to fairly strong long-term correlation within a subarray, which is seen in particular in many practical scenarios in macrocellular networks where the angular spread at the eNodeB is relatively small. A high long-term correlation within each subarray can be exploited to reduce the rate of feedback required for each subarray, in time and/or frequency.

In general, the relatively persistent nature of the spatial signatures of the subarrays facilitates robust MU-MIMO scheduling with predictable multi-user interference. Conversely, the low correlation between the cross-polarized subarrays yields improved diversity for SU-MIMO scheduling. Following these principles, the PMI feedback codebooks for 8 CSI-RS antenna ports have been designed such that precoders within each 4-element ULA subarray are represented by a grid of beams, with the number of available beams from which to choose being 32 when 1- or 2-layer transmission is preferred, 16 for 3 or 4 layers, 4 for 5 to 7 layers and only one for 8 layers. Cross-polarized dimensions (corresponding to low correlation) are combined with *co-phasing factors* represented by quaternary symbols, following the same codebook design principle as in Release 8 for uncorrelated antennas.

An example of the resulting Release 10 codebook design for the case of 2-layer transmission being preferred is shown in Table 29.3.3.1. Full details of the PMI feedback codebooks for 8 CSI-RS antenna ports when other transmission ranks are preferred can be found in [1, Section 6.3.4.2.3].

Table 29.3: PMI feedback codebook design for the case of 2-layer feedback with 8 CSI-RS antenna ports.

i_2	0	1	2	3
Codebook	$W_{2i_1,2i_1,0}$	$W_{2i_1,2i_1,1}$	$W_{2i_1+1,2i_1+1,0}$	$W_{2i_1+1,2i_1+1,1}$
i_2	4	5	6	7
Codebook	$W_{2i_1+2,2i_1+2,0}$	$W_{2i_1+2,2i_1+2,1}$	$W_{2i_1+3,2i_1+3,0}$	$W_{2i_1+3,2i_1+3,1}$
i_2	8	9	10	11
Codebook	$W_{2i_1,2i_1+1,0}$	$W_{2i_1,2i_1+1,1}$	$W_{2i_1+1,2i_1+2,0}$	$W_{2i_1+1,2i_1+2,1}$
i_2	12	13	14	15
Codebook	$W_{2i_1,2i_1+3,0}$	$W_{2i_1,2i_1+3,1}$	$W_{2i_1+1,2i_1+3,0}$	$W_{2i_1+1,2i_1+3,1}$

$$W_{m,k,n} = \frac{1}{4}\begin{bmatrix} \mathbf{v}_m & \mathbf{v}_k \\ e^{jn\pi/2}\mathbf{v}_m & -e^{jn\pi/2}\mathbf{v}_k \end{bmatrix}, \quad \mathbf{v}_m = \begin{bmatrix} 1 & e^{j2\pi m/32} & e^{j4\pi m/32} & e^{j6\pi m/32} \end{bmatrix}^T.$$

The index $0 \leq i_1 \leq 15$ selects a cluster of beams from a set of partially overlapping clusters in a co-polarized subarray. The LSB of the index $0 \leq i_2 \leq 15$ selects one of the two co-phasing factors while the 3 MSBs of i_2 select one of 8 pairs of beams within a cluster.

This codebook structure can be viewed as a two-stage codebook, in which the precoding matrix $\mathbf{W}$ corresponding to the UE's CSI feedback can be decomposed into a product of two matrices, such that $\mathbf{W} = \mathbf{W}_1 \cdot \mathbf{W}_2$, where $\mathbf{W}_1$ corresponds to the ULA grid-of-beams component of the PMI and $\mathbf{W}_2$ corresponds to the frequency-selective co-phasing component of the PMI. The codebooks for PMI feedback when less than 8 CSI-RS antenna ports are configured, and indeed the Release 8/9 codebooks, can be considered to have the same structure with $\mathbf{W}_1$ being set to the identity matrix.

29.3.3.2 Periodic and Aperiodic CSI Reporting Enhancements

Aperiodic CSI feedback on the Physical Uplink Shared Channel (PUSCH) as described in Section 10.2.1.1 follows the same design in Release 10 as in Release 8, except for the use of the new PMI codebooks described above if 8 CSI-RS ports are configured. However, the periodic feedback on the PUCCH (Physical Uplink Control CHannel) described in Section 10.2.1.2 is extended in Release 10 to take advantage of the new two-stage PMI feedback codebook structure explained above.

As spatial signatures within a co-polarized subarray are not expected to change much in the frequency domain, sub-band PMI feedback for this component (referred to as $\mathbf{W}_1$ above) was not considered to be necessary. On the other hand, sub-band reporting can be useful for the spatial component that captures co-phasing across polarizations (referred to as $\mathbf{W}_2$).

These, together with the aim of keeping the uplink overhead as low as possible, led to the following PUCCH feedback modes being supported for the case of 8 CSI-RS ports:

Wideband periodic feedback. The RI/PMI/CQI information is carried in two separate types of PUCCH feedback report: the first contains the RI, while the second contains associated PMI and CQI information. Two alternative modes of operation exist in this case, known as submode 1 and submode 2, which are selected by means of an RRC signalling parameter 'PUCCH_format1-1_CSI_reporting_mode'. In submode 1, the wideband grid-of-beams component of the PMI is included in the first feedback report together with the RI. In submode 2, all of the PMI information (i.e. including both the grid-of-beams component and the sub-array co-phasing component) is contained in the second PUCCH report; in order to achieve this, the PMI feedback codebook has to be subsampled (i.e. only certain specified entries are available from which the UE may select) so that the second report can fit within the available number of bits on the PUCCH. Submode 2 minimizes the code rate for the RI, whereas submode 1 avoids the need for subsampling of the PMI codebook.

Sub-band periodic feedback. This mode allows the fact that the grid-of-beams component of the feedback may be relatively time-invariant [8] to be exploited, by reducing the feedback rate for that component and thereby making some feedback bits available for sub-band reporting of the co-phasing component. The RI/PMI/CQI information is carried in three separate types of PUCCH feedback report. A one-bit Precoder Type Indication (PTI) report is added to the RI reports so that the UE can indicate the contents of the second and third reports. If the PTI is set to '0', the second report contains the grid-of-beams PMI component and the third report contains the wideband co-phasing component of the PMI together with the associated CQI; on the other hand, if the UE estimates that the grid-of-beams component has not changed since the previous time it was reported, the PTI is set to '1', the second report contains wideband co-phasing PMI feedback together with associated CQI, and the third report contains sub-band co-phasing PMI and associated CQI; multiple such third reports are transmitted, one for each sub-band. An example of this sub-band periodic feedback using the PTI is illustrated in Figure 29.8, showing how the sub-band resolution of the co-phasing feedback is increased if the grid-of-beams state is static.

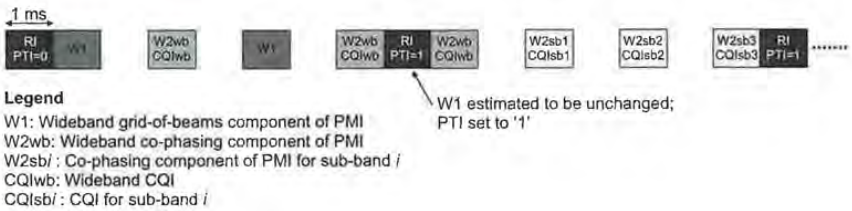
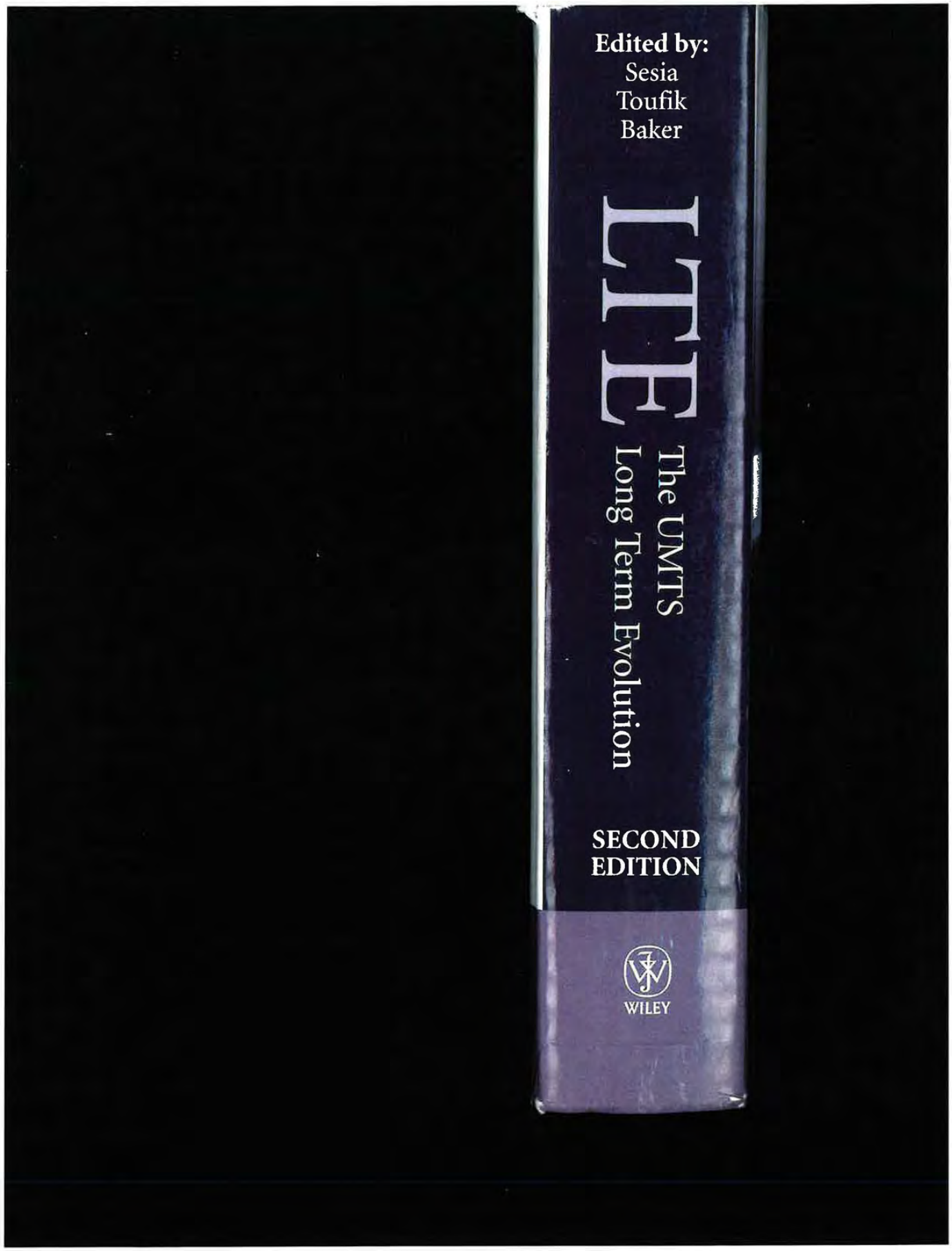


Figure 29.8: Example of a possible sequence of PUCCH CSI feedback reports when 8 CSI-RS ports and sub-band PMI feedback are configured.


Edited by:
Sesia
Toufik
Baker
LTE
The UMTS
Long Term Evolution
SECOND
EDITION
WILEY

SECOND EDITION

Including Release 10 for LTE-Advanced

LTE The UMTS Long Term Evolution

FROM THEORY TO PRACTICE



Edited by:
Stefania Sesia, *ST-Ericsson, France*
Issam Toufik, *ETSI, France*
Matthew Baker, *Alcatel-Lucent, UK*

> *"Where this book is exceptional is that the reader will learn not just how LTE works but why it works."*
> Adrian Scrase, ETSI Vice-President, International Partnership Projects

The first edition of this book is globally recognised as the ultimate reference book on LTE. Following the success of its predecessor, this second edition is fully updated, covering the latest additions to LTE and a full description of LTE-Advanced.

LTE - The UMTS Long Term Evolution, second edition offers a comprehensive system-level understanding built on explanations of the underlying theory. It includes complete coverage of Release 9 and the first release of LTE-Advanced, Release 10. The book is a collaborative effort of more than 40 key experts representing over 20 companies actively participating in the development of LTE, as well as academia. The book highlights practical implications, illustrates the expected performance, and draws comparisons with the well-known WCDMA/HSPA standards. The authors not only pay special attention to the physical layer, giving an insight into the fundamental concepts of OFDMA, SC-FDMA and MIMO, but also cover the higher protocol layers and system architecture to enable the reader to gain an overall understanding of the system.

Key Features:

- **Comprehensively updated with the latest changes of the LTE Release 8 and 9 specifications, including improved coverage of Radio Resource Management, RF aspects and performance requirements**
- **Provides detailed coverage of the new LTE Release 9 features, including: eMBMS, dual-layer beamforming, user equipment positioning, home eNodeBs/femtocells, picocells and Self-Optimizing Networks (SON)**
- **Evaluates the LTE system performance**
- **Explains LTE-Advanced, beginning with its context and motivation, and covering the key new features of carrier aggregation and high-order MIMO, as well as relaying, enhanced Inter-Cell Interference Coordination (eICIC), network energy saving and minimization of drive tests**
- **Discusses future developments such as Cooperative Multi-Point (CoMP) transmission and Machine-Type Communications (MTC)**
- **Includes an accompanying website containing a complete list of acronyms related to LTE and LTE-Advanced, with a brief description of each (http://www.wiley.com/go/sesia_theumts)**

This book is an invaluable reference for all research and development engineers involved in implementation of LTE or LTE-Advanced, as well as graduate and PhD students in wireless communications. Network operators, service providers and R&D managers will also find this book insightful.

Companion Website

Cover design by Dan Jubb

WILEY
wiley.com

ISBN 978-0-470-66025-6