



mobile communications series

Ramjee Prasad
Albena Mihovska
editors



NEW HORIZONS IN MOBILE AND WIRELESS COMMUNICATIONS

VOLUME 1 RADIO INTERFACES



New Horizons in Mobile and Wireless Communications

**Volume 1
Radio Interfaces**

For a listing of recent titles in the *Artech House Universal Personal Communications Series*, turn to the back of this book.

New Horizons in Mobile and Wireless Communications

Volume 1
Radio Interfaces

Ramjee Prasad
Albena Mihovska

Editors



ARTECH
HOUSE

BOSTON | LONDON
artechhouse.com

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the U.S. Library of Congress.

British Library Cataloguing in Publication Data

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

ISBN-13: 978-1-60783-967-5

ISBN-10: 1-60783-967-9

Cover design by Igor Valdman

© 2009 European Union

All rights reserved.

Artech House
685 Canton Street
Norwood, MA 02062

Printed and bound in the United States of America. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Artech House cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

10 9 8 7 6 5 4 3 2 1

Contents

Preface	<i>ix</i>
Acknowledgments	<i>xi</i>
CHAPTER 1	
Introduction	1
1.1 Mobile and Wireless Systems for Next Generation	1
1.1.1 Overview of Mobile and Wireless Communication Systems	2
1.1.2 Evolution and Migration Toward Next Generation Mobile Communication Systems	5
1.1.3 Radio Systems Planning and Optimization	6
1.2 Challenges and Requirements for Radio Interface Design	8
1.2.1 Technical Requirements	9
1.2.2 Enabling Technologies	13
1.2.3 Market Requirements and Services	13
1.3 Research and Standardization Research Activities Toward New Radio Interfaces	14
1.3.1 European-Funded Research Activities	15
1.3.2 Other Activities	17
1.4 Preview of the Book	18
References	20
CHAPTER 2	
Spectrum-Efficient Radio Interface Technologies	23
2.1 Introduction	23
2.1.1 Radio Interfaces for Ubiquitous Communications	25
2.1.2 OFDM-Based Radio Interfaces in the Scope of Next Generation Systems	27
2.1.3 Coexistence and Spectrum Sharing	32
2.1.4 Opportunities for Secondary Spectrum Use	34
2.1.5 Multiband Transmissions	36
2.2 Radio Interfaces Optimized for PANs	56
2.2.1 Scenarios and Radio Propagation Models for PANs	58
2.2.2 Increased Capacity and Throughput for a MIMO UWB System	64
2.3 Accurate Channel Modeling for Adaptive and Scalable Air Interfaces	65
2.3.1 Existing MIMO Channel Models	67

2.3.2	Channel Modeling	77
2.3.3	Simulation of Radio Systems	82
2.3.4	Classifying Channel Characteristics	86
2.3.5	General Descriptions of Basic Channel Models	90
2.4	Conclusions	92
	References	93

CHAPTER 3

	Coding and Modulation	101
3.1	Introduction	102
3.1.1	Analysis and Design of Coding Algorithms for Next Generation Systems	102
3.1.2	Analysis and Design of Modulation Schemes for Next Generation Systems	116
3.2	Design of Multiuser Space-Time Codes	143
3.2.1	Characterization of the Error Event Regions	144
3.2.2	Code Design Criteria	146
3.2.3	Multiuser MIMO Downlink Space-Time Block Coded Transmission Scheme	148
3.3	Advanced Modulation and Coding Techniques	153
3.3.1	Modulation and Coding Schemes with DTBCs	154
3.3.2	Quasicyclic Block LDPC Codes	163
3.3.3	Low-Rate Convolutional Codes for Broadcast Information	166
3.4	Conclusions	171
	References	171

CHAPTER 4

	Multiple-Access Schemes	177
4.1	Introduction	177
4.1.1	Selection Criteria for a Multiple-Access Scheme	179
4.1.2	Single-Carrier Versus Multicarrier Schemes	180
4.2	Analysis of TDMA/OFDMA for an IMT-A Candidate System	193
4.2.1	Key Design Aspects	193
4.2.2	Performance Modeling and Evaluation	197
4.3	Conclusions	214
	References	214

CHAPTER 5

	Smart Antennas and Related Technologies	217
5.1	Introduction	218
5.1.1	Multiple-Antenna Methods and Techniques	218
5.1.2	Benefits of Multiple Transmission Schemes	219
5.1.3	Summary of Multiple-Antenna Methods	221
5.1.4	Spatial Modes	229
5.2	Downlink Capacity Enhancement of IEEE 802.11a/g	229
5.2.1	Space and Time Diversity in the OBAN Scenario	232

5.2.2	Downlink Beamforming Under EIRP Constraints in WLAN OFDM Systems	234
5.2.3	Downlink Capacity Enhancement of IEEE 802.11a/g Using SDMA with a Multiple-Antenna Access Point	247
5.3	Modeling of Two- and Three-Dimensional Antenna Arrays for Channel Model Measurements	251
5.3.1	Overview of Antenna Models	252
5.3.2	Modeling of Antenna Characteristics	254
5.3.3	Broadband RF Module Implementation	265
5.4	Antenna Architectures for UWB Systems	267
5.4.1	Main Transceiver Principles	268
5.4.2	RF Front End and Antenna Implementation and Experiments	271
5.5	Conclusions	275
	References	276

CHAPTER 6

	Coverage Enhancement Technologies	281
6.1	Introduction	281
6.1.1	Standards Supporting Multihop Communications	283
6.1.2	Relay Technologies in Support of Ubiquitous Radio Access	288
6.2	RRM Schemes for Relay-Based Deployments	299
6.2.1	Spectrum Partitioning	300
6.2.2	Adaptive Soft Resource Partitioning	304
6.2.3	Multihop Relay Scenarios	310
6.3	Advanced Antenna Techniques for Relaying	312
6.3.1	Cooperative Relaying	312
6.3.2	Spatial Filtering and SDMA	326
6.4	Conclusions	330
	References	331

CHAPTER 7

	Interference Management	333
7.1	Introduction	334
7.1.1	Methods to Assess the Severity of Interference	335
7.1.2	Interference Management	340
7.1.3	Interference Mitigation	343
7.2	ICI Modeling for Ubiquitous Radio System Concept	347
7.2.1	Multicell Simulation Setup	349
7.2.2	Evaluation of ICI	351
7.3	Interference Scenarios for PAN-Optimized Radio Interfaces	361
7.3.1	Body-to-Body Channel Gain Modeling	362
7.3.2	Temporal Properties of Interferers	369
7.3.3	Unified Interference Modeling Derived from Link Budget Analysis	377
7.3.4	Cooperative Transmission for WPAN Scenarios	379
7.3.5	Interference Mitigation Through Dynamic Interleaving Codes	382

7.4	Conclusions	384
	References	386

CHAPTER 8

	Efficient Management of Radio Resources	389
8.1	Introduction	390
	8.1.1 User and Control Planes	390
	8.1.2 MAC Design Challenges	393
8.2	Resource Management Algorithms	398
	8.2.1 Scheduling Algorithms	398
	8.2.2 Link Adaptation	406
8.3	Cross-Layer Optimization and Design	407
	8.3.1 Cross-Layer Information and Signaling	409
	8.3.2 Scheduling	411
8.4	Cross-Optimization of FM-UWB and MC-SS Radio Interface Solutions	414
	8.4.1 AWA Algorithm	417
	8.4.2 Improved AWA	419
	8.4.3 Performance Analysis	419
8.5	Conclusions	425
	References	425

CHAPTER 9

	Conclusions	429
	About the Editors	431
	Index	435

Preface

ज्ञानं ज्ञेयं परिज्ञाता त्रिविधा कर्मचोदना ।
करणं कर्म कर्तेति त्रिविधः कर्मसंग्रहः ॥१८॥

jñānam jñeyam parijñātā
tri-vidhā karma-codanā
karaṇam karma karteti
tri-vidhaḥ karma-saṅgrahaḥ

jñānam—knowledge; *jñeyam*—the objective of knowledge;
parijñātā—the knower; *tri-vidhā*—of three kinds; *karma*—of

Knowledge, the object of knowledge, and the knower are the three factors that motivate action; the senses, the work, and the doer are the three constituents of action.

—*The Bhagavad Gita (18.18)*

European Research Framework programs are public policy instruments designed to strengthen European competitiveness through cooperation. Although they have a fixed time frame, determined research themes, and a specific expected impact, the achievements in research and development (R&D) made by these funded projects pave the way for a research continuum.

The Information Society Technologies (IST) research program was launched in 1999 as a successor to the Advanced Communications Technologies and Services (ACTS) research framework. Within this program, two consecutive frameworks were focused on advancements in the state of the art in the area of mobile and personal communications and systems: FP5, Satellite-Based Systems and Services, and FP6, Mobile and Wireless Systems Beyond 3G and Broadband for All. Under FP6, the European Union has been funding collaborative R&D activities in the field of telecommunications with a financial allocation of more than €370 million and the objective to make significant progress toward advanced communication technologies, systems, and services.

The FP6 IST research and development effort was a primary initiative that launched large Integrated Projects (IPs) alongside the smaller Specific Targeted Research Projects (STREPs), Specific Support Actions (SSAs), and Networks of Excellence (NoEs).

The enormous research effort concentrated in the various R&D project activities required a special supporting initiative that would span the entire domain of

projects, promote structure, and disseminate information about the research effort and results. This was the main idea behind the European FP6 IST project titled “SIDEMIRROR.” This book is part of a series of books that has resulted from the project effort supported by the European Union under FP6 in the areas of Mobile and Wireless Systems Beyond 3G and Broadband for All. The final research results of the numerous projects in the above-mentioned R&D European initiatives were collected and integrated with the objective of creating a permanent record of their achievements in four books.

In particular, this book is about advancements in the state of the art of radio interface technologies. At the end of FP5, the foundations had been laid that placed technology in a position to define and develop the network and service infrastructures of the future. The objective for the follow-up program, namely, the FP6, was to create an opportunity for generating new economic and market growth with new classes of radio communication systems, while reducing operational expenditures. At the end of FP6, standardization activities in the area of radio interface technologies had gained momentum because some of the designs under the EU FP6 umbrella radio system concepts are strong competitors in the standardization race toward next generation communication systems. The achievements described in this book are key to solving the challenges for future communication systems in terms of spectrum availability and management, multimode terminals, ubiquitous coverage, and successful coexistence of technologies.

The FP6 project work encompassed the design and development of radically new radio concepts (e.g., the WINNER radio system, optimized radio interfaces in support of a novel personal area network concept). Additionally, work was performed for the design, modeling, and measurements of novel channel and propagation models, in combination with advanced modulation and coding techniques and the use of smart antennas. In addition, a number of FP6 projects contributed to the development and adoption of relay-based models as part of the future communication system concept.

Particular emphasis was placed on integration of the developed concepts in the overall picture of future generation wireless broadband systems and networks, from cellular to broadband fixed radio access and broadband wireless local-area networks for both interactive and distributive services. A key aspect of the work undertaken within the FP6 R&D initiative was that of the validation and demonstration of the proposed concepts and solutions.

The material collected in this book was edited to provide useful reading material to senior and junior engineers, undergraduate and postgraduate students, and anyone else interested in the development of current and future mobile communications. It was impossible to include all project achievements; however, the book provides a useful tool in terms of R&D methodology that can be applied to the development of new concepts.

We hope all readers will experience the benefits and power of this knowledge.

Acknowledgments

First of all, the editors would like to acknowledge Ms. Dua Idris from CTIF at Aalborg University for her efforts toward the completion of this book. Dua was involved in the FP6 IST project SIDEMIRROR as a research engineer. Her consistent hard work and passionate efforts toward improving the quality of the content are very commendable.

Further, this book would not have been possible without the strong support of the European Union project officers from the Directorate General Information Society (DG INFSO) of the European Commission in Brussels, namely, Dr. Jorge M. Pereira, Dr. Manuel Monteiro, and Dr. Francisco Guirao Moya, who guided the project work and ensured strong cooperation with the rest of the FP6 IST projects in the area. Their efforts were an essential prerequisite for the successful editing of the available research results. The material in this book was collected and structured with the support of the European Commission within the framework of the IST supporting project SIDEMIRROR, in an effort that spanned a period of 5 years. It originated from the technical reports and documentation of the projects involved in the FP6 IST R&D initiatives Mobile and Wireless Systems Beyond 3G and Broadband for All.

The project SIDEMIRROR consisted of two member organizations, Aalborg University and Artech House Publishers. The editors would like to acknowledge the support of their colleagues from Artech House in the completion of this manuscript.

Finally, the editors would like to thank the administrative and IT staff from Aalborg University for providing the required administrative and IT project support.

*Ramjee Prasad
Albena Mihovska
Editors
Aalborg, Denmark
May 2009*

Introduction

In mobile communication systems, the air interface is the radio-based link between the mobile device and an active base station (BS). Designing and developing an air interface for future communication systems requires careful consideration of the envisioned scenarios for operation of these communication systems and identification of the expected challenges.

Research activities designed to improve transmission capabilities, including novel designs for radio air interfaces, were principally sustained by collaborative research programs in Europe such as the Sixth Framework Program (FP6), a large-scale research and development initiative in the area of IST, partially funded by major industries worldwide.

Huge state-funded initiatives have also been launched in different Asian countries, and activities have taken place in major North American companies along with developments in the wireless IT sector and its associated markets. In parallel, international organizations such as the International Telecommunication Union (ITU-R) [1] and the Wireless World Research Forum (WWRF) are actively structuring this rapidly growing sector through preparation work and standards.

This book describes the results and contributions of the European research effort performed within the framework of FP6 and supported by major industries, academia, and end-user organizations worldwide for the design and standardization of radio air interfaces for next generation communication systems.

This chapter introduces the thematic topic and puts the achieved results in perspective in terms of the worldwide effort and vision for future communication systems. The chapter is organized as follows. Section 1.1 describes the current state of the art and development trends in mobile and wireless communication systems in Europe and worldwide. Based on these trends, a vision for future communication systems is proposed. Section 1.2 defines the requirements imposed by users, technology, and businesses and identifies the challenges that need to be overcome including enabling technologies. The European research effort, as part of the IST R&D for communication and networking technologies, is outlined in terms of general major contributions to worldwide standardization efforts in Section 1.3. Specific outstanding contributions supporting the overall air interface development are also outlined. Section 1.4 provides a preview of the contents of the rest of this book.

1.1 Mobile and Wireless Systems for Next Generation

Mobile and wireless communication is an important driver of economic growth. Significantly improved transmission capabilities are increasingly required to support

content-rich, data-oriented services in order to connect people as well as machines to the information society. Mobile and wireless systems beyond third generation (3G) emerged as the means to realize the concept of “optimally connected anywhere, anyhow, and anytime” supported by all system levels, from access methods and networks to service platforms and architectures. These systems were modeled as a horizontal communication model in which different terrestrial access levels and technologies are combined to complement each other in an optimum way for different service requirements and radio environments [2]. On the other side, R&D efforts directed at optimizing access technologies, as a function of the operating environment and at affordable prices, were identified as critically important for a generalized introduction of broadband services in Europe and in less developed regions worldwide.

Broadband air interfaces are needed to provide for high-bit-rate wireless communications. To achieve this goal, consolidated research is required among areas such as communication theory, information theory, signal processing, propagation, antennas, and real-time hardware.

1.1.1 Overview of Mobile and Wireless Communication Systems

Current infrastructures are unable to support ubiquity because of problems such as interference, network congestion, and spectrum limitations. Another difficulty that is coming from the evolution to future wireless communications is the heterogeneity of radio access technologies (RATs). The International Mobile Telecommunications–Advanced (IMT-A) concept is from the ITU and covers mobile communication systems with capabilities that go further than that of IMT-2000 [1]. IMT-A was previously known as “Systems Beyond IMT-2000.” The support of broadband services for mobile and wireless applications in IMT-A systems [1], with excellent user experiences, are key trends for future RATs, providing deployment scenarios with reduced operator capital expenditure (CAPEX) and operational expenditure (OPEX) indices.

1.1.1.1 Enhancements of 3G and WLAN Systems

The development of IMT-2000 is expected to reach a limit of around 30 Mbps [3]. Many of the ongoing initiatives to define and implement systems beyond IMT-2000 are aimed at overcoming the above limitation. Research and standardization efforts are focused on the evolution of existing air interfaces, for example, the Long Term Evolution (LTE) [4] initiative of the Third Generation Partnership Project (3GPP) [5] for enhanced UTRAN; on the work with WiMAX [6]; on the development of the IEEE 802.16e standard for the adoption of mobile WiMAX [7] and other IEEE 802.x standards (e.g., IEEE 802.22); and on the development of new air interfaces.

Within the Universal Mobile Telecommunications Services (UMTS), high-speed downlink packet access (HSDPA) [5] provides a downlink packet data channel with peak data rates of up to 10 Mbps, which are shared between multiple users. This is enabled by the use of adaptive modulation and coding (AMC), hybrid ARQ (HARQ), fast scheduling, and a modified frame structure. Furthermore, multiple

input/multiple output (MIMO) antenna techniques and a complementary enhanced uplink channel are being specified in 3GPP and COST 278 as well as orthogonal frequency division multiplexing (OFDM), which has been under study as an additional UMTS modulation scheme for improved system capacity [8].

3G system enhancements based on new network topologies and broadcast services can be provided by the Wideband Distribution System (WDS) and the Multimedia Broadcast Multicast Services (MBMS) approach. The use of existing wireless local-area network (WLAN) standards as a mechanism to provide higher data capacities for hot spot coverage in UMTS systems has been investigated with respect to seamless interworking in terms of first identifying a common authentication, authorization, and accounting (AAA) mechanism for both systems and finally achieving seamless handover with guaranteed quality of service (QoS) [9, 10].

Further developments of these systems are ongoing in attempts to achieve higher data rates, better mobility, more flexibility and even more efficient support of packet transmission. As an improved version of HSDPA, followed by HSUPA, 3GPP introduced HSPA [5], a combination of the two technologies that is considered a spectrally efficient wireless solution.

The enhancements of HSDPA [5] for the wideband code division multiple-access (WCDMA) [11] air interface of UMTS, however, are limited by system-inherent restrictions, such as the allocated carrier bandwidth. The same is valid for WLANs [12]. These are expected to provide high throughput for low mobility but ubiquitous coverage cannot be achieved due to the small cell leading to high deployment costs.

The ITU's vision is that new wireless access technology may need to be developed around 2010 that is capable of supporting even higher data rates with high mobility, which could be widely deployed around 2015 in some countries [1]. The new capabilities of these IMT-A systems are envisaged to handle a wide range of supported data rates according to economic and service demands in multiuser environments with target peak data rates of up to approximately 100 Mbps for high mobility such as mobile access and up to approximately 1 Gbps for low mobility such as nomadic/local wireless access [3].

Recommendation ITU-R M.1645 [3] states that IMT-A will support the general cellular system requirements and features related to improved performance, in comparison to enhanced IMT-2000 systems, with respect to parameters, including spectrum efficiency and peak data rate, latency in order to enable new delay-sensitive applications, and cell size and cell-edge performance, which are required for adoption of mobile communication services.

On the other hand, IMT-A will support connectivity, with increased system performance for a variety of low-mobility environments, such as these:

- Stationary (fixed or nomadic terminals);
- Pedestrian (pedestrian speeds up to 3 km/h);
- Typical vehicular (vehicular speeds up to 120 km/h);
- High-speed vehicular (high-speed trains up to 350 km/h).

The inclusion of mobile WiMAX [7] under the IMT-2000 umbrella can offer significant benefits to the mobile community by ensuring the global delivery of

wireless broadband Internet services at the lowest cost [6]. Mobile WiMAX is based on orthogonal frequency-division multiple access (OFDMA), an access technology standardized in the IEEE 802.16e-2005 amendment to IEEE 802.16. OFDMA technology has become commonly accepted as the basis for the evolution of mobile technology to the next generation, because it can provide high data rate capability and excellent support for new features such as advanced antenna technologies to maximize coverage and the number of users supported by the network. OFDMA (specifically, the air interface designated “WirelessMAN-OFDMA” within IEEE 802.16) provides multipath and interference tolerance in non-line-of-sight (non-LOS) conditions to achieve ubiquitous broadband coverage in a wide range of operating environments and usage models, including full mobility [6].

ITU-R M.1645 [3] also recommends that the IMT-A systems provide the mobile user with an “always-on” experience while also taking into account and providing features needed to preserve battery life. The connectivity from the user terminal (UT) to the BS should be automatic and transparent to the user as the user moves between mobile networks.

Next generation communication systems are seen as organized in a layered structure, comprising a distribution layer, a cellular layer, a hot spot layer, a personal network layer, and a fixed (wired) layer [1]. These systems should also support the use of coverage enhancing technologies [3], such as mobile and fixed relays.

Extremely low power wireless connectivity among inexpensive fixed, wearable, and moving devices as part of the personal network layer can be provided by air interfaces based on short-range technologies, such as ultrawide band (UWB) and multicarrier-spread spectrum (MC-SS) [13]. When considering the criteria identified for such air interfaces, the question has been raised [13] of whether the existing WLANs standards (e.g., IEEE 802.11) can fulfill the personal network layer requirements and challenges. In that regard, it can be observed [14] that the primary design criteria of the WLANs are different from those of a personal area network (PAN). When the WLAN IEEE 802.11 standard is analyzed, the following observations can be made [14]:

- IEEE 802.11 does not provide balanced costs, which is the case for the PAN devices.
- IEEE 802.11 does not address the power consumption of the PAN class of devices.
- WLAN IEEE 802.11 does not address the reduced complexity requirements of the PAN class of devices.
- IEEE 802.11 is optimized for throughput, distance, and roaming, whereas PANs can be optimized for low cost and low power consumption in a small form factor.

To clarify the preceding differences, we provide the ITU’s definition of a device that uses UWB technology. According to the ITU [1], this refers to a type of equipment that incorporates, as an integral part or as an accessory, technology for short-range radio communication, involving the intentional generation and transmission of radio-frequency energy that spreads over a frequency range wider

than 50 MHz, which may overlap several frequency bands allocated to radio communication services.

The heterogeneity in technology and also in ownership leads to complex systems and interworking problems, which can be seen in the ability to establish and maintain connections with required quality, in fault detection and location, in resource allocation, and in charging of the usage of the network's resources. This is also the reason why researchers look for systems that are adaptive and flexible and provide coverage in various deployment modes (wide area, metropolitan area, and local area) [5, 13]. Cooperation and coexistence in this case extend to interworking between different physical layer modes of the same radio system [13].

The IST project WINNER [13] proposed an air interface based on OFDMA technology whose system comprises various functions that are intended to avoid data loss and minimize delay during handover, as well as ensure the coverage recommended by the ITU deployment scenarios. This is made possible by the WINNER simplified radio access network (RAN) architecture and supporting cooperation mechanisms [14, 15].

A simplified network architecture means that the functionalities as we know them from legacy systems (e.g., GSM, UMTS) will be distributed in a new way. In the case of evolved systems, such as LTE [4] and the WINNER RAN [13], the RAN's architecture moves the radio functionalities closer to the radio interface.

For successful QoS delivery as well as security and protection of data, any radio resource management (RRM) mechanism (i.e., mobility management, congestion and admission control, load control) must take these architectural changes into account, because the traditional ones will not be optimized to support such a distributed system architecture.

1.1.2 Evolution and Migration Toward Next Generation Mobile Communication Systems

The rising demand for fast, scalable, efficient, and robust data transfer over the air has produced a variety of RATs. Future wireless systems are supposed to provide high-bit-rate services in IP-based, real-time, person-to-person as well as machine-to-machine multimedia communications. These systems will include a number of coexisting subnetworks with different RATs. The interworking between radio subnetworks and, especially, the tight cooperation between them is very interesting in operating RATs for system capacity optimization. The reconfiguration technology provides adaptation of the radio interface to varying RATs, provision of possible applications and services, updating of software, and enabling full exploitation of the flexible resources and services of heterogeneous networks. Reconfigurable terminals, with embedded radio link layer functionalities according to network architecture, will allow for cooperation among multiple RATs.

The RATs are often divided into different "generations" of mobile radio networks. 2G systems had their focus on conversational services, but it was possible to use circuit-switched data services as well, though their performance was not satisfying [11]. As an extension to the Global System for Mobile Communications (GSM) the General Packet Radio Service (GPRS) was invented [16]. With enhanced

data rates up to 60 Kbps and its packet-switched nature, this service solved the weaknesses of circuit-switched systems.

Various RATs also refer to different transport technologies. Not too long ago (around 2001), heated debate arose regarding whether WLAN [12] systems like IEEE 802.11x [17] and HIPERLAN/2 [18] or the UMTS [19] would be more efficient in terms of costs for planning, building, and operating. It was suggested that these systems could be considered an extension of UMTS [5]. Hence, two or more of these systems may peacefully coexist in such a way that the user would always be served with the best possible QoS at the lowest possible price. UMTS is currently gaining momentum with increased deployment worldwide. At areas with high volumes of traffic, the so-called “hot spots” or “hot areas,” systems with high data rates are being deployed. In addition, new RATs covering a wide range of services and operating in different deployment concepts are the focus of research and development [1]. These heterogeneous systems must be capable of forwarding data streams and session context among each other, and a (vertical) handover should happen seamlessly. For example, the WINNER air interface is a new type of air interface that provides a seamless level of service across a coverage area similar to that achieved by the UTRAN and GERAN systems deployed today [20]. This covers rural areas and also provides a contiguous coverage layer in towns and cities, where it will overlap with metropolitan-area and local-area network deployments. Another important target is to support the full range of mobility scenarios, including use on high-speed trains. One requirement is that an ubiquitous radio system has to be self-contained, allowing it to target the chosen requirements without the need for interworking with other systems. Another requirement is that cooperation, whenever required, will be successfully ensured between any new or legacy systems.

To ensure successful adoption of any new technology, an evolutionary approach to the design and its smooth migration are very important. It is also expected that legacy systems will continue to serve users [5]. More complete integration of the various communication systems into an ambient “umbrella” network including highly reconfigurable core networks, radio access infrastructures, and user devices provides improved services to users [21]. Such an approach is beneficial to defining the path toward the evolution of next generation communication systems. One of the key challenges in the evolution of the telecommunications world will be to efficiently manage and control the optimization of radio resource and spectrum usage in the context of heterogeneous systems, encompassing the space, time, multiaccess, and multiowner dimensions.

1.1.3 Radio Systems Planning and Optimization

The deployment scenarios shown in Table 1.1 have been proposed for next generation systems.

A successful system deployment should enable economic solutions with respect to spectrum need, deployment, and equipment costs. Higher data rates and possibly higher carrier frequencies reduce the covered range significantly compared to existing systems. Ubiquitous coverage, therefore, requires a higher number of BSs and access points, which would increase the deployment costs [21]. For this reason,

Table 1.1 Proposed Deployments for Next Generation Systems

<i>Cell Range</i>	<i>Performance Target</i>
Up to 100m	Nomadic performance, up to 1 Gbps
Up to 5 km	Performance targets for at least 100 Mbps
5–30 km	Graceful degradation in system/edge spectrum efficiency
30–100 km	System should be functional (thermal noise-limited scenario)

new radio systems must ensure sufficient range and scalability with respect to system capacity and deployment costs. New frequency bands might also be required to provide for new high-data-rate services. Optimum spectrum use is one of the research challenges related to radio interface system design and optimization. Figure 1.1 shows an ubiquitous radio system concept, which was proposed by the IST project WINNER [21].

One key to ubiquitous deployment is the ability to operate in multiple scenarios. The system shown in Figure 1.1 is able to provide coverage in wide- as well as short-range areas. Relays offer benefits in terms of coverage extension, and technologies such as cooperative relaying are being considered in the context of next generation networks. Relaying requires specific channel models, relay to relay and mobile to relay. Such were proposed by the FP6 IST projects WINNER [21] and the FP6 IST project FIREWORKS [22].

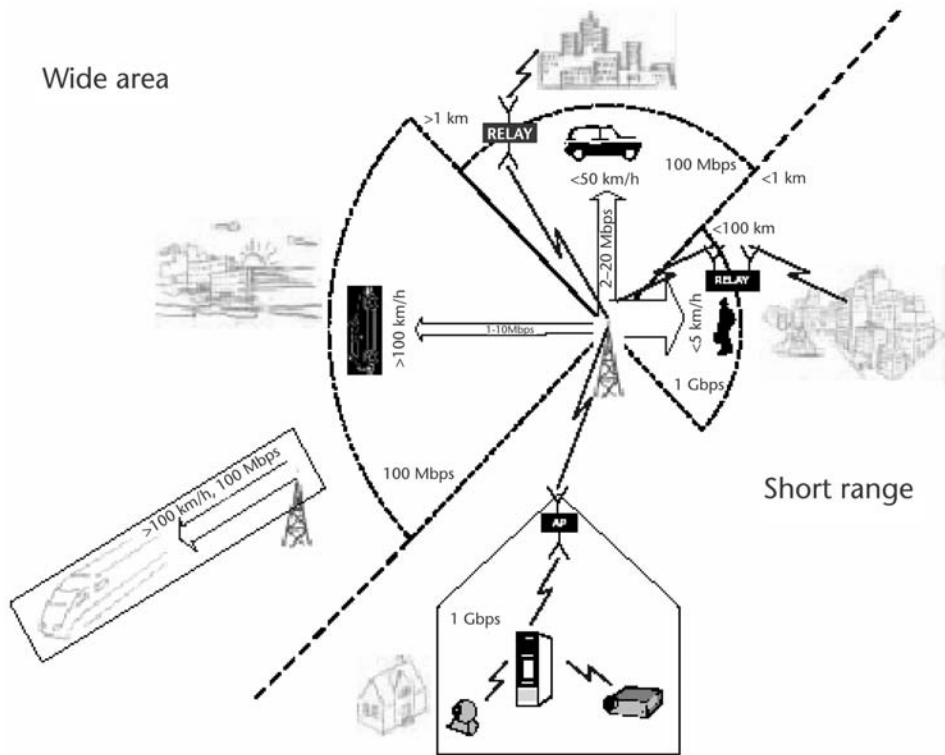


Figure 1.1 A concept of a radio system with ubiquitous coverage.

1.2 Challenges and Requirements for Radio Interface Design

The definition of a new air interface concept is an inherently iterative process determined not only by pure technical considerations but also by regulatory and other decisions or restrictions. These decisions and restrictions are in turn typically influenced by many, more or less technically oriented considerations and trade-offs that may often be linked to and/or originate from the definition process itself. Hence, in general, the definition and regulatory processes can be characterized as being in some sense mutually dependent. From a design perspective this is clearly a problem because the regulatory framework is needed to provide an optimized solution. On the other hand this is also a problem for the regulatory process that needs technical arguments and indications to define the system framework.

The design of an optimum air interface is a multidimensional problem. The appropriate configuration, which might be the optimum for each case, depends on several input factors such as the physical scenario and radio channel, application requirements, terminal capabilities, actual traffic load, and user subscription profile. The following dimensions can be mentioned [18]:

- The physical conditions, defined by the channel state information (CSI), determine the criteria for best exploitation of the radio resources in terms of bits per second per hertz given a target bit error rate (BER).
- Different UTs may be accommodated, and UT capabilities will also determine the optimum mode of operation. What is optimum regarding one aspect, such as channel quality, may not be optimum for another one, such as UT capability.
- Coverage area, capacity, and provided service QoS are the three most important aspects of a wireless communication network. The size of the coverage area is chosen based on predicted traffic and user distributions in a given geographical area: urban, suburban, rural. The average capacity, which can be provided in the coverage area, ultimately determines the number of users connected to the network and, thus, also the optimal resource allocation strategies to be used by the network operator.
- Different applications have different requirements in terms of real-time demands and bit rate, which imply diverse solutions for the air interface configuration. Specific QoS requirements are needed for the optimal transmit-receive implementation and to accommodate requirements imposed by the envisioned packet services. To focus the service requirements on quality, services and their associated QoS are usually grouped in classes, for example, conversational, streaming, interactive, and background [23].
- In overall system design, the clear trend is that cutting the cost of the core network will lead to an increased degree of network control being located in the BS. Ultimately, this means that the BSs will be direct access points to the IP backbone network. This trend is also driven by the growing demand of fast Internet multimedia services.
- There is no one, single evolutionary path from today's diverse radio systems to future mobile communication systems. The WLAN type of radio access schemes is evolving toward mobility support and at least to a low level of

QoS guarantee, which might imply that the development of future cellular networks may be to loosen the tight QoS requirements of current 2G and 3G systems [23]. On the other hand, new applications such as interactive teleconferencing, videoconferencing, and gaming will set much tighter requirements on some parameters, such as delay, with respect to what can be offered in current communication systems. Moreover, multiple access and multicast applications may require different QoS parameters. In addition to delay requirements, the system's overall performance and capacity are dependent on the QoS specifications.

One major factor involved in the final decision about which technologies will be the best candidates for next generation systems is how well any candidate technology reflects and deals with the international, national, and regional differences that exist in the use of the radio spectrum. With the identification of additional spectrum by the World Radiocommunication Conference (WRC) [24], an international forum for world agreement on spectrum allocation, it is likely that various subsets of the identified frequency bands will continue to be used in order to meet the needs of the various users worldwide. Therefore, any new air interface technology must be able to deal with the range of frequency bands in use as well as with the various user demands.

1.2.1 Technical Requirements

An air interface can be modeled by a protocol stack that consists of four basic protocol layers: the radio resource control (RRC) layer, the radio link control (RLC) layer, the medium access control (MAC) layer, and the physical (PHY) layer. The basics and specifics of each interface technical system requirement are usually determined based on information, such as the traffic intensity for which the system is to be designed, the admission and congestion strategies, the maximum allowed output power from different types of network nodes in different environments, spectrum availability, and so forth, all of which are required to calculate the absolute numbers of many key parameters. To obtain a design basis for the air interface, initial estimates of these parameters and their implications must be provided.

1.2.1.1 Physical Channel Requirements

An air interface must carry at least two types of information: system control data for network management purposes and user-shared or user-dedicated payload data. Therefore, the spectrum resources available for the downlinks (DLs) and uplinks (ULs) of the air interface may be divided into physical channels in order to provide means for efficient network management and control signaling as well as user data traffic scheduling.

In the downlink, system control data is transmitted to support the user terminals in performing the following:

- Cell detection;
- Synchronization;

- Acquisition of access point configuration/access information;
- Acquisition of UL and/or DL scheduling information.

Similar to user data, system control data can be divided into two different classes depending on its intended use. Shared system control data are typically used for cell detection, synchronization, and acquisition of configuration/access information purposes serving any users within the service area of the access point, whereas dedicated system control data is typically used to convey UL and DL scheduling information to one or a small group of known users.

The fundamental differences in conveying information using broadcast to a large number of unknown users and multicast to a small number of known users suggest a need for using different physical channel structures for different kinds of control data in order to optimize the efficiency of the air interface and to meet the requirements of different system functions. For example, in the case of broadcasting, preprocessing methods at the transmitter based on a priori information can typically not be exploited in order to improve the detection performance of the different receivers. However, utilizing diversity techniques offers an alternative solution—although one achieved at the price of an overall reduced air interface spectral efficiency in general. For multicast purposes to known users, on the other hand, scheduling or other preprocessing methods at the transmitter based on, for example, CSI knowledge, may be considered to achieve sufficient transmission quality without facing a spectrum efficiency loss. Hence, the physical channel structure must provide sufficient flexibility to support efficient transmission of different types of control data.

Cellular system aspects may also place requirements on the physical channel structure. In contrast to the situation in isolated cell scenarios, in which the entire or suitable parts of the available bandwidth can be utilized by an access point to provide reliable transmissions to all parts of its service area, in cellular systems, frequency reuse larger than one is often desirable or even required to guarantee sufficient transmission quality. Thus, to support cellular system operation, the physical channel structure must provide a canalization that allows for reuse flexibility.

In addition to transmission efficiency and cellular system operation-oriented considerations, requirements may also be raised on the physical channel structure to provide means for maximizing the standby and active mode times of battery-driven devices. This implies that, besides selecting a physical channel structure that supports the use of power-saving protocols and scheduling strategies, the physical channel structure should—as much as possible—support the use of low-complexity implementations in the transmitters and receivers. The concern over power consumption, size, and realizable product cost related to wireless PAN (WPAN) communication systems is much greater and comes from the fact that WPANs are targeted to wearable and mobile devices [25].

In the uplink, system control data is transmitted to convey the following:

- Random access requests;
- Scheduling requests;
- Measurement reports.

Compared to the downlink where certain system control data are common to all users, in the uplink, the system control data will be dedicated only, and only transmitted when the terminals are in an active mode. In the same way as for the downlink, this implies that the physical channel structure should provide flexibility that allows different degrees of a priori information at the terminals and the access point to be utilized in order to support efficient transmission of the control data. Flexibility is also required to handle the diverse amounts of control data that can be expected from each terminal depending on what services it is supporting.

Another fundamental problem affecting the choice of the physical channel structure is the output power limitations of the different UTs. Typically these limitations are due to the power amplifier of the UTs and give rise to constraints on both the maximum peak power as well as the maximum output power density of the terminal as a function of the transmission bandwidth. This in turn has implications for the link budget and thus the uplink coverage. Hence, to provide both good coverage in cellular system scenarios and good capacity in isolated cell scenarios, the physical channel structure should take the power limitations of the user terminals into account. In addition, the channel structure must provide means for low processing delays in the access points to keep link setup delays, round-trip time delays, and so forth low.

The demand for ubiquitous coverage and a variety of data services requires a certain amount of adaptivity on the part of the air interface. The air interface must provide means for adaptation in terms of both different user data rates as well as quality levels in order to support different types of services and coverage scenarios. It must also utilize the available spectrum resource in an efficient manner to provide cost-efficient solutions to both users and operators of the air interface. These demands translate into requirements on the physical channels needed for carrying user payload data. The following requirement can be identified: The spectrum resource available for user payload data should be capable of being divided into portions suitable for scheduling in all applicable domains (i.e., frequency, time, and/or code domains) as required.

The physical channels of the air interface carry different types and combinations of logical information channels. Beside logical channels for control and/or payload data, one or more logical channels for channel estimation or other purposes such as synchronization, cell specific broadcast, and so forth, will in most cases be multiplexed onto each physical channel. To support the mapping of logical channels onto physical channels, the basic physical channel protocol structure of the air interface (i.e., the time structure and the extension of the physical channel definitions into resource blocks) must be defined.

1.2.1.2 Protocol Requirements

The performance of packet-oriented connections, both TCP and non-TCP based, is to a large extent determined by the end-to-end available bandwidth and the end-to-end round-trip time delay of the communication link between a packet source and its associated sink. Optimizing these two entities alone to maximize the end-user service experience, however, is not enough. The product of the available bandwidth and the RTT can be used to minimize the pipe capacity as well as to

minimize the variations in the available bandwidth and the RTT since such variations typically have a negative impact on the overall performance due to their effects on the congestion control algorithm of the used traffic protocol.

Beside the requirement of providing wireless connectivity between an access point and multiple users possibly at the same time, one of the key requirements of the future air interface is that it should support packet data traffic. Seen as an integral part of a larger network providing end-to-end packet transmissions, taking the performance limiting factors into account, a fundamental requirement of the future air interface is, therefore, that its contribution to the overall RTT must be small. For example, the RTT time for IMT-A candidate systems has been proposed to be 2 ms [1, 3, 21]. Furthermore, because the transmission and processing time of user data packets can be expected to be much larger than the time acceptable for transmission and processing of the associated protocol signaling (primarily ACK/NACK signaling), a natural requirement on the UL and DL physical channel protocols is that they must be able to support substantially different one-way delay requirements for different types of data.

1.2.1.3 Resource Blocks

The concept of resource blocks is fundamental to describing how data of different logical channels is to be mapped onto physical channels. A resource block is an extension of the physical channel definition, taking the physical channel time structure and the modulation format of the used transmission scheme into account. Thus, it defines a container of physical channel symbols and provides the basic structure on which logical channel data—after being mapped into channel symbols—can be multiplexed in order to be transmitted over the physical channel. This is shown in Figure 1.2 for the downlink of an OFDM-based air interface as proposed by the IST project WINNER.

1.2.1.4 Estimation Techniques

The choice of appropriate estimation techniques should be based on performance and complexity. For example, if the implementation of a flexible air interface that

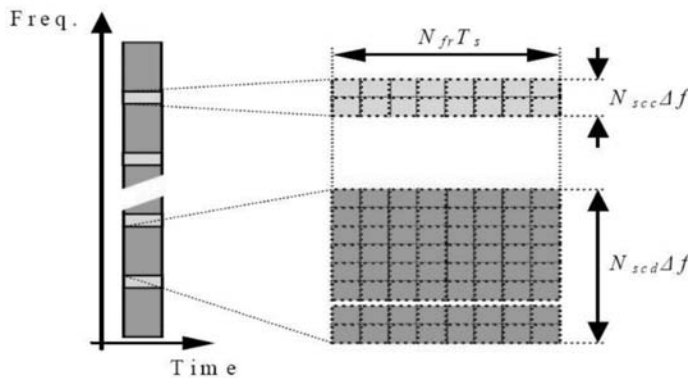


Figure 1.2 Concept of resource blocks as proposed in [26] for the downlink.

works in arbitrary channel conditions should be supported, then reliable estimation must be maintained for highly frequency-selective channels as well as for high velocities of the mobile users. Furthermore, high modulation cardinalities, different MIMO concepts, and various multiple-access schemes should be supported. Also, scalability is an important criterion, since there will mostly likely be different solutions for low- and high-end terminals. The requirements and algorithms for estimation schemes strongly depend on the system concept used, that is, whether single or multiple carriers, uplink or downlink, single or multiple antennas are deployed.

1.2.1.5 Link Adaptation

Power and bandwidth are crucial resources in wireless communication systems. In addition, transmission delay is an important aspect. In an adaptive link, on the other hand, the receiver has to be informed as to which demodulator/decoder parameters to employ for the received signal. This information can be conveyed, for example, together with the transmitted signal (for an OFDM-based system) at the cost of effective data throughput. A receiver can also estimate these parameters by the appropriate blind detection algorithms [27].

1.2.2 Enabling Technologies

There is a wide agreement that for high link data rate systems on the order of many megabits per second, multiple-carrier radio interfaces, and more specifically those based on OFDM, offer advantages, although single-carrier interfaces in combination with spectrum domain equalization techniques receive support from some parties due to the potential for less transmitter complexity [21]. In addition, there is a widespread support for combining OFDM with CDMA, to optimize the RAN capacity [21]. It has been proven theoretically that multiple-antenna systems are capable of providing capacity gains over single-antenna ones, at least on the link level. The additional capacity gain is still an open issue for system level investigations. Due to its inherent advantages, the technology combination of OFDM with MIMO antennas is receiving the most support. Promising results using MIMO antennas in the handheld terminals are available even under the condition of correlated receive signals at closely packed antennas. However, the implementation complexity is an issue that needs to be resolved.

Wideband communications also involve appropriate channel models, with both spatial and temporal characteristics.

1.2.3 Market Requirements and Services

An important requirement for communication systems is interoperability between different vendors' equipment. In that way, service providers can build up their access networks by using devices from multiple vendors, subscribers can exploit their user terminals in different operators' networks, and manufacturers can sell equipment from the same production lines to all customers. The main advantage of interoperability is that resources are exploited in an efficient way and that

economies of scale can prevail. This contributes to the benefits of subscribers, service providers, manufacturers, and to the entire economy and society as a whole.

The description and modeling of the radio interface is normally based on the Open Systems Interconnection (OSI) reference model that is explained in Recommendation ITU-T X.200. The interface structure is also consistent with ITU-R M-series Recommendations ITU-R M.1035, ITU-R M.1457, and ITU-R M.1645 and also the ITU conceptual frameworks for IMT-2000 and IMT-A.

Further, next generation radio systems should support applications that conform to open standards and protocols, including, but not limited to, video, full graphical Web browsing, e-mail, file uploading and downloading without size limitations (e.g., FTP), streaming video and streaming audio, IP multicast, location-based services, virtual private network (VPN) connections, VoIP, instant messaging, and online multiplayer gaming.

Therefore, one important consequence is an ability to cope with heterogeneity. In such a scenario, connections would span over several networks that deploy different transport technologies, and the networks would be owned and operated by separate organizations. When, for example, a customer subscribes to an Internet connection or a VPN service with specified quality, it is almost impossible for the customer to make sure that the promised quality is fulfilled. On the other hand, it is also difficult for the Internet service provider (ISP) to guarantee the promised quality in a multitechnology and multioperator environment. If the service does not provide the promised quality, it is equally difficult to identify the cause for that.

1.3 Research and Standardization Activities Toward New Radio Interfaces

Research and standardization activities in the area of radio interfaces are focused on developing, optimizing, and evaluating an IMT-A candidate proposal by integrating innovative and cost-effective additional concepts and functions and providing an evolution path aimed at further improving the performance of IMT-A systems. Figure 1.3 shows a summary of the evolution process of the various technologies that are being considered as IMT-A candidates [28].

The activities shown in Figure 1.3 mainly target already identified and/or allocated frequency bands for mobile and wireless communication, which results in constraints on the feasible carrier bandwidth and thereby on the potentially achievable system throughput. This has focused research [29] on the investigation of potential IMT-A candidate technologies that have even more challenging requirements in additional frequency bands still to be identified by the WRC [1].

This development is ongoing in a globally competitive environment. Technology candidate proposals are expected from Europe, the United States, China, Japan, and Korea. There are similar evolution paths for CDMA2000 and IEEE 802.16 systems toward LTE-like systems. To reduce fragmentation and ensure a competitive position in the global context, manufacturers and operators in Europe and worldwide are mobilized for a collaborative research effort.

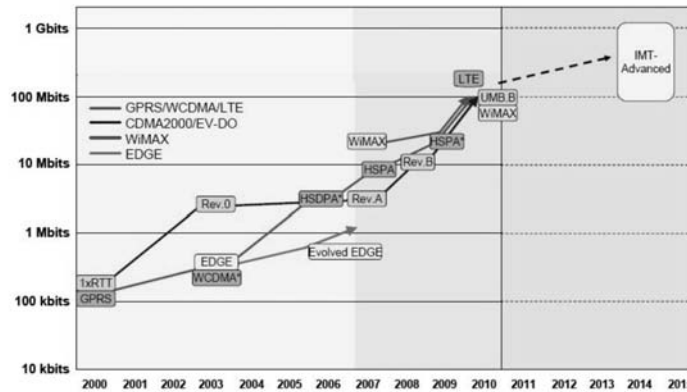


Figure 1.3 Evolution of mobile technologies toward the next generation.

As one step toward a fruitful collaboration, the European Community identified the FP6 as a continuation of the R&D achievements at the end of FP5 and in support of overall worldwide standardization efforts.

1.3.1 European-Funded Research Activities

At the end of FP5, the issue of a new air interface for Beyond 3G systems was still open, with targeted bandwidths up to 100 MHz and spectral efficiency up to 10 bits per second per hertz. The modulation scheme had first to be chosen. The trend toward multicarrier modulation was present at the end of FP5, but no choice had been made yet. Research on various spectrum efficiency enhancement tools, such as multiple antennas and relays, was ongoing. Single-user MIMO schemes had already been widely studied, though receiver complexity seemed still an issue. Multiuser MIMO scheme studies had just started. Relay benefits in terms of coverage extension had been acknowledged, but cooperative relaying, as well as ad hoc and routing, only started to be considered in the context of cellular or fixed wireless access networks. On the other hand, topics such as PANs were still in their infancy.

The objective of the European consolidated approach under FP6 in terms of enabling technologies for future highly spectrally air interfaces spanned a large number of projects, namely, 4MORE, MASCOT, UNITE, WINDECT, WINNER I and II [21], FIREWORKS [22], MAGNET and MAGNET Beyond [13], PULSERS, NEWCOM, ACE, URANUS, WIDENS, OBAN, E2R II, MEMBRANE, and SURFACE [30].

MIMO technologies, virtual MIMO (relay), appropriate channel models, efficient coding schemes, and advanced system-on-chips (network-on-chips) were at the heart of studied areas.

The WINNER project developed a generic model, covering 12 mobile and 5 fixed propagation environments. The models are applicable over the frequency range of 2 to 6 GHz and their RF bandwidth is 100 MHz. In similar conditions the complexity of this modeling approach is comparable or even smaller than for the alternative modeling methods. WINNER II developed a single ubiquitous radio

access system adaptable to a comprehensive range of mobile communication scenarios from short range to wide area. This is based on a single radio access technology for the PHY and MAC layers. A toolbox of components was developed to enable such an adaptive system.

The PULSERS project members contributed to IEEE 802.15.4a channel models. The MAGNET project developed novel channel models for both wideband (at 2.4 and 5.2 GHz) and UWB (at 3–5, 17, and 60 GHz) systems, which incorporate the main PAN characteristics: user dynamics and device handling, body area and body proximity effects, and peer-to-peer communications scenarios. At 3 to 5 GHz, different configurations of multiple-antenna structures were tested in order to highlight its role in UWB MIMO system.

The OBAN project modified the classical beamforming algorithm taking into account equivalent isotropic radiated power (EIRP) constraints. A significant range (coverage) extension gain was obtained compared to the single-antenna access point, for example, an average two to four times gain for the four-antenna WLAN. OBAN also proposed an alternating time-offset spatial-division multiple access (SDMA) solution, which supports legacy IEEE 802.11a/g terminals and requires some modifications to the transmission protocol at the AP. It has been shown that the downlink capacity in a conference room environment can be almost doubled for low levels of channel reciprocity errors at the access point equipped with three or four antennas depending on the propagation conditions.

The 4MORE project developed a complete and consistent air interface definition, based on multicarrier CDMA, that is suitable for 4G transmission. It encompasses physical parameters, complete baseband functional chains for both 2×2 and 4×2 MIMO configurations, together with the definitions of temporal frames. Finally, elements of a MAC protocol solution have been developed. This solution is close enough to the current emerging standards (e.g., 3GPP LTE) to be considered as a reference for a 4G OFDMA solution.

The FIREWORKS project developed multiple transmit and receive antennas physical layer algorithms for relay and cooperative systems. Some of the major techniques investigated were space-time block codes (STBC), golden codes, and multiuser precoding and detection. Relaying and cooperative schemes at the physical layer were also considered with a particular focus on how to amplify and forward/decode and forward distributed STBC, compress and forward techniques, and selective relaying.

The URANUS project proposed a reconfigurable architecture for various standards (UMTS, WiMAX, 802.11x, GSM, and so on), based on the concept of generalized multicarrier (GMC) signals. A novel receiver architecture was proposed that is capable of processing both single- and multiple-carrier baseband signals under a common fast Fourier transform (FFT) module, using the concept of Gabor frames. It relies on the one-tap channel model assumption, which was validated for WCDMA downlink transmission. An appropriate channel estimation scheme was developed. Issues of peak-to-average power ratio reduction have been considered. A novel initial synchronization algorithm based on GMC was developed.

The MEMBRANE project investigated reconfigurable antenna techniques for wireless backhauling applicable to a wide range of system configurations. In MEMBRANE a data-splitting algorithm for MIMO relay systems was developed. The

algorithm exploits the propagation characteristics of the relay links to optimally split the transmission data stream into several substreams forwarded to the final destination. Algorithms for a two-path relaying protocol were developed and a selection criterion between these algorithms was deduced. An interference cancellation algorithm at the receiver side using multiple antennas was also proposed.

The UNITE project derived an optimal decoding and power allocation for SIC, MMSE, and matched filters, using asymptotic tools from random matrix theory. Power allocation is thus performed in a decentralized manner.

The WINDECT project used IEEE 802.11e with a minor modification to achieve seamless handover from access point to access point and, with voice processing, to achieve greatly improved bit error control. New approaches to minimizing power consumption were proposed, making WINDECT devices broadly equivalent to traditional DECT.

The SURFACE project addressed the optimization of linear MIMO transceivers under a minimum BER criterion. To tackle the limitations arising from the common practice of a priori fixing the number of symbols to be transmitted, a multimode procedure was developed that adaptively selects the optimum number of modes for a given target rate. The proposed multimode minimum BER design achieves the full diversity of the channel in both the perfect and partial CSI scenarios.

The ACE project developed multiple-antenna techniques for cooperative communications distributed schemes aiming to exploit spatial diversity, multiuser MIMO cross layer designs, including multidimensional packet scheduling, and a framework and methodology for context-aware optimization with smart antennas.

The MAGNET and MAGNET Beyond projects developed and optimized WPAN-specific air interfaces. Two PHY layers were developed: a high-data rate (HDR) MC-SS PHY layer and a low data rate (LDR) FM-UWB PHY layer. In addition, measurements proved significant power level gains for a simple switched diversity system in a handheld to handheld scenario. Also interference to the HDR mode from other systems was modeled and mitigation mechanisms were introduced, such as power loading, interleaving, and space diversity (MIMO) techniques.

Most of the projects developed demonstrators as proofs of concepts. These have been successfully demonstrated at a large number of worldwide events. For example, the MAGNET Beyond project developed and implemented a low-band IC and a high-band RF IC for the FM-UWB air interface [voltage control oscillator (VCO), prescaler, wideband demodulator (WBDM), LNA, subcarrier processor, direct digital synthesizers (DDS)]. Further, an RF board was designed and implemented for the low-band system that uses the MAGNET chipset and that is interfaced to the LDR digital board. The resulting implementation is shown in Figure 1.4.

1.3.2 Other Activities

In 3GPP2, Revision A of the evolution path for the IMT-2000 CDMA multiple carrier called Ultra Broadband Mobile (UMB) is already finished and the work on Revision B is ongoing. The goal of IEEE 802.16m is to develop an IMT-A system candidate, based on IEEE 802.16e (WiMAX), and the goal of IEEE 802.11n is to develop a high-performance ad hoc extension to be used in IMT-A systems. These

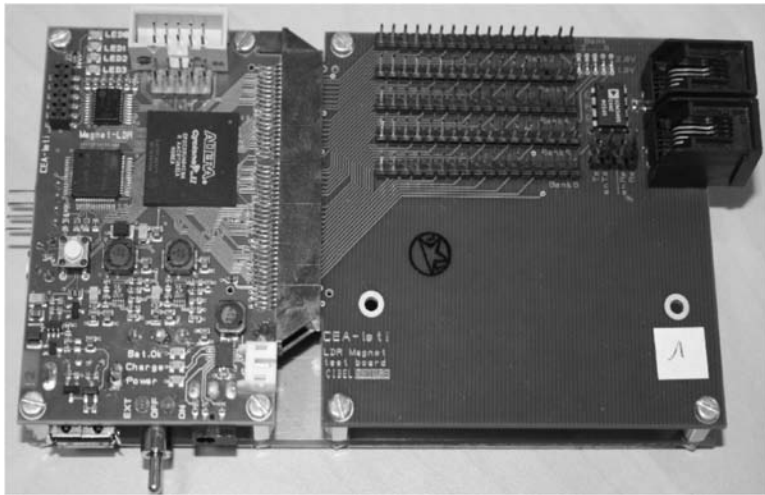


Figure 1.4 An implementation of a LDR WPAN air interface (PHY+MAC board).

systems do not yet fulfill the technical requirements of IMT-A. However, they provide an evolution path toward more enhanced systems. Finally, these different systems are expected to provide similar capabilities.

The global research targets for IMT-A have been agreed on in ITU-R and WWRF [31] and they have initiated huge research programs in different regions, which will be exploited during the forthcoming global standardization process.

The IEEE 802.11 standardization body is the leading organization worldwide driving the development of WLANs. The 5-GHz OFDM-based 802.11a WLAN provides variable data rates up to 54 Mbps, based on a link adaptation mechanism. The task group 802.11g is working on an 802.11b compatible high-speed extension that provides data rates up to 54 Mbps with a physical layer that employs OFDM or packet binary convolutional coding (PBCC) as an option in IEEE 802.22g. Other task groups are addressing issues such as dynamic frequency selection (DFS), transmit power control (TPC), and QoS support. In addition, a high-throughput study group (HTSG) was created whose goal is to provide significantly higher throughput by improving the PHY and MAC layers.

In terms of WPANs and WBANs, the IEEE 802.15 standardization body is one organization driving the standardization process for short-range wireless networks. Also, within the European Telecommunications Standards Institute (ETSI), such activities are supported especially for applications in the health sector. After the finalization of the Bluetooth standard, PANs that are especially designed for low- and high-data-rate applications are now being defined and improved further.

1.4 Preview of the Book

This book is organized into nine chapters. Chapter 2 focuses on spectrum-efficient radio interface technologies. It describes the efforts of the IST WINNER I and II

projects and the MAGNET and MAGNET Beyond projects to design and optimize air interfaces in compliance with the activities of bodies such as the ITU. Methods for efficient and flexible spectrum use and spectrum sharing are important contributions to the development of spectrum demand estimation methods in preparation for the spectrum-related decisions of regulatory bodies. Issues for this chapter include the topics of accurate channel modeling, use of MIMO techniques to enhance the performance of the PHY layer technologies, interference modeling, and other technologies for better spectrum use and for minimizing interference to the adjacent spectrum. MIMO or beamforming is one candidate technology for this requirement. Some of the projects discussed in this chapter are WINNER, MAGNET, FIREWORKS, and B-BONE.

Chapter 3 focuses on modulation and coding as techniques for enhancing the performance of air interface transmission technologies. Here, we outline the time versus the frequency domains, the single versus the multicarrier techniques, advanced coding techniques, and link layer techniques. Some contributing projects are WINNER, WINNER II, and MAGNET.

Chapter 4 describes the importance of multiple-access schemes for achieving high data rates. Multiple access is discussed for both the uplink and downlink. Further, some techniques for spectrum sharing are also discussed. The main contributing projects are WINNER, B-BONE/C-MOBILE, 4MORE, MAGNET, and MASCOT.

Chapter 5 focuses on smart antennas and related technologies. It discusses smart transmitter and receiver configurations. Further, it discusses the use of spatial techniques in spectrum shared scenarios. The main contributing projects are ACE and WINNER.

Chapter 6 discusses how coverage can be enhanced based on use of relay techniques and what is needed to consider when employing such techniques. Cell coverage is a very important issue because support for larger cell sizes should not compromise the performance of smaller cells. This chapter comprises topics such as spectrum partitioning and relay positioning for cellular enhanced systems (with two-hop), the role of RRM in relay-based deployments, resource allocation and scheduling in multiple-hop relaying networks, joint routing and resource partitioning in relay-enhanced cellular networks, cooperative relaying, and spatial temporal processing and relaying. Projects contributing to this area are WINNER, WINNER II, and FIREWORKS.

Chapter 7 describes research efforts related to the management of interference problems. Intercell interference cancellation and averaging are introduced followed by interference mitigation using smart antennas. An overview of an intercell interference avoidance scheme is provided. Resource management and partitioning, and resource allocation and scheduling are also covered. WINNER and ACE are the major projects contributing to this chapter's content.

The techniques for managing radio resources efficiently are covered in Chapter 8. Concepts such as duplexing, scheduling, and resource allocation within a cell and intercell coordination are explained. IST projects such as AROMA, EVEREST, WINNER, DAIDALOS, and PHOENIX have carried out research work related to this chapter's topics. Chapter 9 concludes the book.

References

- [1] International Telecommunication Union, www.itu.int.
- [2] FP6 Overview, <http://cordis.europa.eu/ist/so/mobile-wireless/home.html>.
- [3] Recommendation ITU-R M.1645, "Framework and Overall Objectives of the Future Development of IMT 2000 and Systems Beyond IMT 2000," www.itu.int.
- [4] Long Term Evolution, <http://www.3gpp.org/Highlights/LTE/LTE.htm>.
- [5] Third Generation Partnership Project, www.3gpp.org.
- [6] WiMAX Forum, www.wimaxforum.org.
- [7] Standard IEEE 802.16e, www.ieee802.org/16/tge.
- [8] UTRAN Radio Interface Protocol Architecture, Release 5, TS 25.301, V5.2.0, www.3gpp.org, September 2002.
- [9] "Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; Requirements and Architectures for Interworking Between HIPERLAN/2 and 3rd Generation Cellular Systems," ETSI TR 101 957, V1.1.1 (2001-08).
- [10] "Feasibility Study on 3GPP System to Wireless Local Area Network (WLAN) Interworking Rel-6," 3GPP TR 22.934, V1.0.0.
- [11] Ojanpera, T., and R. Prasad, *Wideband CDMA for Third Generation Mobile Communications: Universal Personal Communications*, Norwood, MA: Artech House, 1998.
- [12] "WLAN," http://en.wikipedia.org/wiki/Wireless_LAN.
- [13] FP6 IST Project My Personal Adaptive Global Net (MAGNET) and MAGNET Beyond, www.ist-magnet.org.
- [14] Lott, M., et al., "Cooperation Mechanisms for Efficient Resource Management Between 4G and Legacy RANs," *13th Wireless World Research Forum*, Seoul, Korea, March 2005.
- [15] Mino, E., et al., "Scalable and Hybrid Radio Resource Management for Future Wireless Networks," *Proc. IST Mobile Summit 07*, Budapest, Hungary, July 2007.
- [16] GPRS Technology, www.gsmworld.com/technology/gprs/index.shtml.
- [17] IEEE 802.11, The Working Group Setting the Standards for Wireless LANs, www.ieee802.org/11.
- [18] HIPERLAN standards, www.etsi.org.
- [19] UMTS, www.umts-forum.org.
- [20] Sdralia, V., et al., "Cooperation of Radio Access Networks: The IST FP6 WINNER Project Approach," *11th Wireless World Research Forum*, Oslo, Norway, June 2004.
- [21] FP6 IST Project Wireless Interface New Radio, WINNER and WINNER II, www.ist-winner.org.
- [22] FP6 IST Project Flexible Relay Wireless OFDM-Based Networks (FIREWORKS), www.fireworks.intranet.gr.
- [23] "Services and Service Capabilities," 3GPP TS 22.105 V8.0.0 (2006-06).
- [24] www.itu.int/ITU-R/conferences/wrc.
- [25] FP6 IST Project MAGNET, "D3.2.2a Candidate Air Interfaces and Enhancements," www.ist-magnet.org, October 2004.
- [26] IST WINNER Project, "Identification of Radio Link Technologies," Deliverable 2.1, July 2004.
- [27] Hanzo, L., C. H. Wong, and M. S. Yee, *Adaptive Wireless Transceivers: Turbo-Coded, Turbo-Equalized and Space-Time Coded TDMA, CDMA and OFDM Systems*, New York: John Wiley, 2002.
- [28] Walko, J., "Mobile Operators Under Pressure in Barcelona—3GSM Report," *Picochip, EETimes Europe*, February 19–March 4, 2007.
- [29] CELTIC WINNER+ Project, www.celtic-winner-org.

-
- [30] IST SURFACE Project, www.ist-surface.org/deliverables.htm.
 - [31] World Wireless Research Forum, www.world-wireless.org.

Spectrum-Efficient Radio Interface Technologies

The need for next generation wireless communication systems to deliver high data rates places increased demand on the spectrum. Given the number of wireless systems already deployed today and the increased use of these, identification of exclusive spectrum for new radio systems is becoming increasingly difficult. The next generation radio access interface(s) should be spectrum efficient to meet this increasing demand for flexible use of the spectrum for emerging new services and applications [1].

Spectrum efficiency refers to the amount of information that can be transmitted over a given bandwidth in a specific communication system. It is a measure of how efficiently a limited frequency spectrum is utilized by the physical layer protocol and sometimes by the media access control layer.

This chapter addresses the challenge of designing spectrum-efficient air interfaces for next generation communication systems. Technologies for improving the spectral efficiency of next generation communication systems involve use of MIMO techniques, also in combination with adaptive modulation and coding, cognitive radio technologies, and so forth. This chapter also addresses the importance of accurate channel modeling as a means of achieving a spectrally efficient PHY layer. It describes the various approaches to channel modeling depending on the chosen underlying transmission technology and how these translate into improved spectral efficiency.

The chapter is organized as follows. Section 2.1 introduces the topic of spectrum-efficient radio interfaces. Section 2.2 describes various spectrum-efficient radio interface technologies proposed in the context of next generation communication systems and as part of the R&D efforts within the IST FP6 program. In particular, IST projects WINNER and WINNER II [2], MAGNET and MAGNET Beyond [3], SURFACE [4], and PULSERS invested a major effort in this area. Section 2.3 describes advancements in the area of advanced channel modeling. Section 2.4 concludes the chapter.

2.1 Introduction

An important aspect of the design of spectrum-efficient radio interface technologies requires an understanding of the spectrum needs of next generation communication systems and the market. In [5], the capabilities of next generation communication

systems (i.e., IMT-A) are defined for achieving high-quality multimedia applications within a wide range of services and platforms, providing a significant improvement in performance and QoS. Among these, of importance are the required high degree of commonality of functionality worldwide while retaining the flexibility to support a wide range of services and applications in a cost-efficient manner; the ability to interwork with other radio access systems; enhanced peak data rates to support advanced services and applications; and worldwide roaming.

Next generation wireless systems will be required to support very high aggregate bit rates (tens to hundreds of megabits per second) in nonideal radio propagation environments. Other requirements include flexibility and adaptability to user needs and application scenarios, moderate UT and BS hardware costs, high spectral efficiency, and scalability of the cost of terminals with respect to their maximum bit rate capabilities.

The services provided by next generation systems should be available where they are needed by users. The correlation between geographical location and type of service needed will be reduced (e.g., not only “home” services in the home, or “work” services in the office) [6]. The next generation system will need to support a variety of services, with different requirements, and the end-to-end QoS for these services should be negotiable and controllable. Some changes are expected concerning the roles of the service, content, and network providers. Users may play an active role in some content creation. Each user will use more than one service at a time, including background services that could eventually be used continuously (e.g., context discovery, interpretation, and management; agents and push services). End-to-end services are expected to be built primarily on IPv6. Location information from satellite systems, such as GPS and GALILEO, will gain momentum for the provision of ubiquitous coverage and QoS.

Several causes have been identified that lead to differences between service class requirements expressed at the user’s application level and the slightly different (tighter) requirements when they are translated to the air interface. First of all, note that in general, for the selected scenarios for next generation systems, a user can demand several different service classes at the same time.

Several inefficiencies are introduced when crossing the different layers of the stack, from the application layer to the air interface and vice versa. These inefficiencies modify the air interface requirements mainly at two points: bit rate and delay. A first inefficiency is introduced when performing the segmentation onto network layer packets. Next generation communication systems will be IP based, thus conveying the raw data on IP packets and possibly adding an IP overhead. IPv6 is the a priori enabler protocol for packet transfer in future core networks. In [7] models of IPv6 packet overhead are described in the context of new air interfaces.

A second inefficiency factor is the segmentation within and beneath Layer 2, MAC and PHY. Packets arriving at the air interface undergo a successive scheduling, thus breaking the relatively large IP packets into smaller transmission units. A problem arises when different data packet sizes are related to different service classes. For example, the data packet sizes of service class “file exchange” and “short control messages and signaling” are deemed to be of a rather different size. The problem must be overcome providing the required granularity when scheduling at the MAC and PHY level. This granularity, however, is to be properly traded

off for other figures of merit such as coding efficiency and retransmission delay. A two-stage scheduler joining two bypassable inner and outer encoders, coding, respectively, long and short transmission units, has been proposed as a solution to this problem [8]. Thus, transmission granularity can be chosen arbitrarily depending on the need, and inefficiency is minimized to a nonsignificant level.

Finally, service classes demanding very low BERs or highly interactive, delay-sensitive service classes will require an effective higher data rate through the air interface. In the first case, the higher data rate is required because deeper error correction coding is needed. In the second case, because the retransmission protocol use is bounded, the peak throughput must be properly tuned, that is, slightly increased.

Use of reconfigurable air interfaces is one way to comply with the preceding challenges and requirements, because those are based on frequency-domain transmission and reception methods, which are essential to meet the requirements, by adaptively selecting the uplink and downlink modulation and multiple-access scheme that is most appropriate for the channel, interference, traffic, and cost constraints. This approach also leads to a general unified framework for possible air interfaces.

In this context, technologies for achieving ubiquitous coverage and proper channel models are very important.

Wideband communications require appropriate channel models, with both spatial and temporal characteristics. MIMO techniques are seen as crucial for spectrally efficient next generation systems.

The overall spectral efficiency of a system can be improved with good coexistence properties, good spectrum sharing capabilities, and flexibility in the spectrum use. Next generation systems need to be able to coexist efficiently with each other and with different radio services in the same or adjacent frequency bands without a need for wide guard bands or restrictive technical or usage limitations. The ability to share spectrum with other systems will significantly increase the efficiency as well as the acceptability of the system. The overall spectral efficiency of the system can be also increased with a flexible use of spectrum that adapts to the spatial and temporal variations in the traffic and environment characteristics. The flexibility and scalability of the system are also important in order to simplify the network deployment under spectrum arrangements that may vary from region to region. Built-in capabilities for flexibility and sharing may significantly ease the task of spectrum identification for the system [9].

2.1.1 Radio Interfaces for Ubiquitous Communications

Spectral efficiency is required for the new air interfaces because they should provide high flexibility and coverage in a wide range of application scenarios, namely, at least the following:

- *Base coverage urban*: an urban macrocellular environment targeting continuous coverage for pedestrians up to fast vehicular users in built-up areas;

- *Microcellular*: an urban microcellular environment with higher user density focusing on pedestrian and slow vehicular users;
- *Indoor*: an indoor hot-spot environment targeting isolated cells at home or in small offices based on stationary and pedestrian users.
- *High speed*: macrocell environments with high-speed vehicular needs, such as use on trains.

Three of the above environments are rather similar to the ones that were used for IMT-2000 [5]: indoor-office, outdoor-indoor and pedestrian, and vehicular; modifications might not be needed related to fulfilling the air interface requirements for these three types of environments. The new environment is the high-speed environment, because subscribers nowadays also require connections in this type of environment. Figure 2.1 illustrates the relative positioning of three of the identified communication environments.

The initial focus for deployment and most of the challenges in IMT-A system design and performance will be encountered in populated areas. However, the provisions for ubiquitous coverage and the associated performance also in rural areas need to be addressed. The deployment of IMT-A is believed to be around 2015 on the mass market level, and at that point in time the majority of countries should have rather good coverage of pre-IMT-2000 systems as well as IMT-2000 systems and their enhancements. Also the interworking with other radio access technologies and spectrum sharing possibilities will be key parts for the choice of the IMT-A system.

The minimum performance capabilities for IMT-A candidate radio transmission technologies are shown in Table 2.1. Such stringent requirements set tight challenges both on the radio network solutions and radio interface techniques. Despite the technical challenges, the overall system development also has to result in acceptable

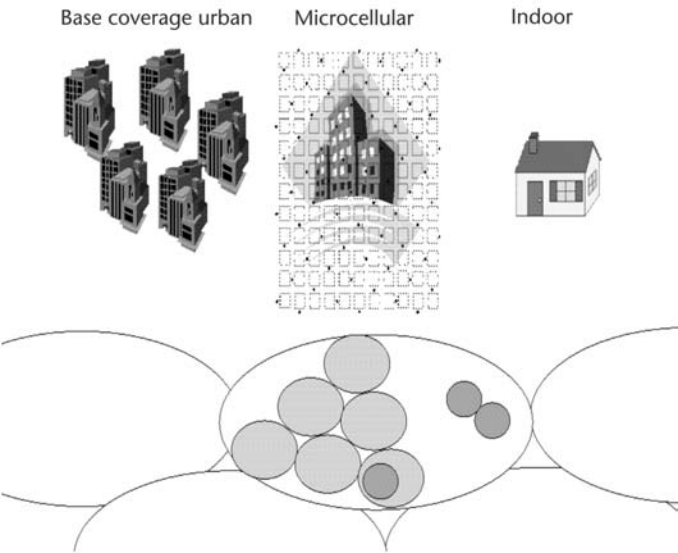


Figure 2.1 Illustrative representation of three of the deployment scenarios envisaged for IMT-A.

Table 2.1 Minimum Performance Capabilities for IMT-A Candidate Radio Transmission Technologies Derived from IMT-2000

<i>Types of Environments for IMT-A</i>	<i>Base Coverage Urban</i>	<i>Indoor</i>	<i>Microcellular</i>	<i>High Speed</i>
Mobility considerations	High	Low	Medium	Very high
Handover	Required	Required	Required	Required
Support of general service capabilities	Required/ not required	Required/ not required	Required/ not required	Required/ not required
Packet data	Required	Required	Required	Required
Asymmetric services	Required	Required	Required	Required
Multimedia	Required	Required	Required	Required
Variable bit rate	Required	Required	Required	Required

solutions from the viewpoints of economics as well as standardization and regulatory processes. Because spectrum is a very scarce resource, a next generation mobile communication system needs to be capable of efficient and flexible use of spectrum. Efficient spectrum use is one of the crucial design targets of a new air interface [10].

System scalability and adaptability need to be achieved in an economically sound manner. This means that the ease of implementation and reconfigurability, the cost-efficient deployment, and the stepwise increase in complexity and performance characteristics, allowing also for simple and cheap equipment, must be emphasized in the system design. Therefore, the next generation ubiquitous system must be based on a minimum number of variants for radio interface elements that can be combined in the most appropriate way to efficiently support any given scenario. The radio interface elements should have high commonality and may comprise one generic modulation scheme or multiple-access scheme [11].

2.1.2 OFDM-Based Radio Interfaces in the Scope of Next Generation Systems

OFDM is a digital modulation method that converts a high-data-rate signal into multiple low-speed narrowband signals and transmits them in parallel in the frequency domain to improve resistance to multipath interference and achieve high spectral efficiency. Due to its advantages in terms of its effect on multipath propagation and the orthogonality properties, the OFDM technique has been chosen for transmissions with high data rates over a frequency selective radio channel with a large maximum delay compared to the symbol duration. Some wireless communication systems based on OFDM have been proposed and implemented. An overview of these systems is shown in Table 2.2.

The OFDM technique has been used for broadcast applications; for example, in the European DAB and DVB-T systems for a data rate of 2 and 34 Mbps, respectively. For broadcasting in wide-area environments the concept of an SFN is introduced. In an SFN all transmitters operate in the same frequency band transmitting an identical signal.

In interactive communication systems two standards have been introduced so far, namely, HIPERLAN/2 [13] and IEEE 802.11a [14]. These two systems are

Table 2.2 Overview of OFDM-Based Systems According to Different Categories [12]

<i>OFDM-Based Systems</i>	<i>Single Cell</i>	<i>Cellular Networks</i>
Broadcasting	DAB DVB-T DRM	Single frequency network (SFN) Multifrequency network (MFN)
Interactive communication	HIPERLAN/2 IEEE 802.11a	MFN: HIPERLAN/2 SFN: WINNER

intended to supply mainly stationary terminals in hot spots, such as airports, train station, and hotels, with high data rates of up to 54 Mbps in a frequency range of 5 GHz. But these two standards are limited to single-cell networks or to multiple-cell networks that cover only a limited area. Therefore, an OFDM-based cellular network for the next generation of mobile communications was proposed by the research community [1] and in [2] to address the missing part.

To increase the flexibility and efficiency of a cellular communication system, an adapted SFN can be applied based on the OFDM transmission technique [12]. There are some important differences between such a network and a conventional cellular network, however. First, in a cellular SFN network, all BSs and mobile terminals (MTs) share all of the resources and can access them at any time. This means that resources used by BSs and MTs in different cells are not separated by any guard bands. But in this case each BS transmits different signals. Second, in the proposed network concept, there are no BS controllers. Instead, a radio resource management (RRM) system that uses self-organized dynamic channel allocation (SO-DCA) is operated in each BS. MTs in different cells communicate with their BSs by means of a multiple-access scheme, which in combination with the OFDM technique will support a very high degree of flexibility and adaptivity for an air interface.

2.1.2.1 Synchronization Algorithms

The well-known challenge of OFDM systems is their strong sensitivity to frequency offsets. The frequency, as well as the time, offsets are caused by the mismatch between the transmitting and the receiving oscillators and timing clocks, respectively. Additionally, channel delay and Doppler frequency also have an influence. A carrier frequency offset, for instance, results in a shift of the received signal spectrum in the frequency domain, so that the received data signals are erroneously demodulated. In addition, a time offset can cause the receiver observation window to be filled with samples belonging to consecutive OFDM symbols, resulting in intersymbol interference (ISI). Some possible forms of offset include the following [15]:

- Carrier frequency offset (caused by poor stability of oscillators and Doppler shift in a mobile environment);
- Phase noise in the RF oscillators;
- Sampling or clock frequency offset;
- Sampling or clock jitter;

- Symbol timing offset caused by the different timing clock at the transmitter and receiver sides.

Phase noise in the RF oscillators, sampling or clock frequency offset, and clock jitter in general can be considered frequency offsets.

Numerous techniques are available to achieve synchronization, each of which has its own advantages and disadvantages. Generally the synchronization techniques can be grouped in two categories, using either cyclic prefix (CP) or pilot signals.

The CP methods can be utilized to implement time and frequency synchronization. The main idea of these methods is that the redundancy introduced by the cyclic prefix allows synchronization to be performed without additional pilots. The most popular method is maximum likelihood estimation, which is based on the correlation of the cyclic prefix. This method operates well in fast fading environments and gives frequency offset estimates of better than 1%, which can be used in a tracking mode. Timing offset estimates are poor and are recommended for use in an acquisition mode [16, 17].

The CP methods are sometimes insufficient for proper synchronization. Some proposals have modified this scheme by using pilot signals. The basic idea of synchronization by pilot signals (preamble, postamble, pilot, null carrier, and so forth) is that known symbols are sent on predefined subchannels at predefined positions. A synchronization procedure is then applied, based on the information of the pilot signals. Furthermore, these can be used for channel estimation. An example is synchronization using a PN sequence, which is transmitted as part of an OFDM preamble. At the receiver side, a received PN sequence is then correlated with a locally generated PN sequence. This method yields good synchronization, however, its performance is poor in fast fading environments [16–20].

A summary of the types of synchronization is shown in Table 2.3.

Let's consider a cellular OFDM-based system in the scope of next generation systems, where every cell is allowed to transmit in the same frequency band at the same time. In this network each BS independently selects radio resources for communicating with its own MTs, which exhibit the smallest amount of interference from adjacent cells. An overview of such a network is shown in Figure 2.2.

In the system under consideration, a BS can observe MTs located in its own cell and in adjacent cells. Therefore, an MT should synchronize itself to a BS in order to receive the signal from that BS, but not from other BSs. Each MT decides to choose a BS that it receives with the highest signal power. The MT belongs to

Table 2.3 Types of Synchronization for Cellular Systems [12]

<i>Category</i>	<i>Method</i>	<i>Advantages</i>	<i>Disadvantages</i>
Cyclic prefix	Maximum likelihood estimation	Good performance in fast fading environment Frequency offset estimates of better than 1%	Poor performance in timing offset estimation
Pilot signals	PN sequence	Sharp synchronization peak	Unsuitable in fast fading environment

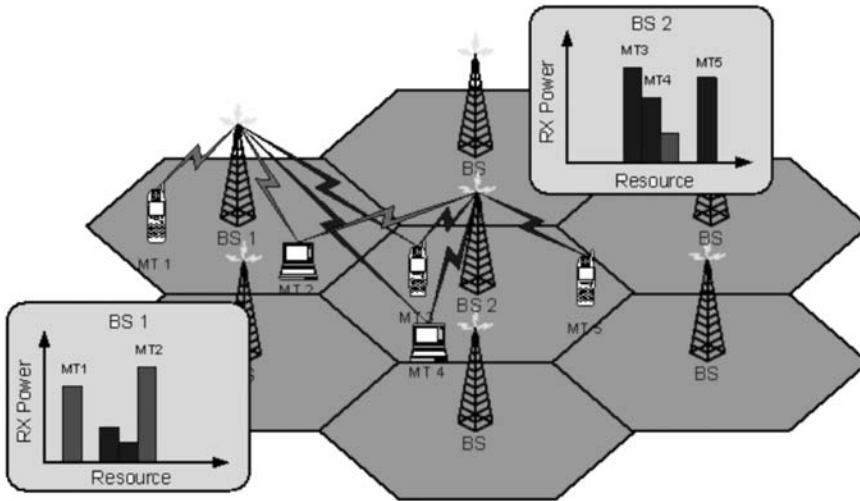


Figure 2.2 General structure of an OFDM-based cellular single-frequency network [12].

that BS's own cell or, in other words, it only synchronizes itself to the signal it receives from that BS. This procedure is applied in the DL period.

The main issue that should be discussed is how to synchronize all BSs in the network without the need for a BS controller. It is not possible for a BS to receive signals that are transmitted directly from other BSs, especially when a time-division duplex (TDD) frame structure is considered. In [12] a proposed concept for synchronization was described that allows all MTs and BSs within a cellular network to be synchronized in a self-organized way. The basic idea behind this concept is to use MTs as the means to synchronize the BSs. This procedure is carried out in the UL period and described later.

The synchronization process is separated into three phases: coarse synchronization, fine synchronization, and tracking.

The coarse synchronization is responsible for aligning the TDD frame structure of all cells and, therefore, is also referred to as frame synchronization. After having achieved a coarse synchronization between adjacent cells, there still remains a residual time and frequency offset. This offset is less than the duration of an OFDM symbol and a subcarrier spacing.

A coarse synchronization need only be performed rarely. If the BS and MT inside the network are synchronized after the initial system start-up, they will stay in that state as long as possible. The coarse synchronization will only be lost if no estimates are available from adjacent cells anymore. This situation might only occur if no terminals are present in the network. No data will be transmitted during this coarse synchronization phase in the uplink and downlink. The system then enters the fine synchronization phase.

The fine synchronization phase will remove the residual time and frequency offset, which remains after the coarse system synchronization. Because it is this offset that causes severe intercarrier interference (ICI), it has to be completely removed for the successful operation of an SFN. Even during this fine synchroniza-

tion, data transmission will not be possible and must, therefore, be suspended until the cell synchronization has reached a sufficient accuracy. Such accuracy will be continuously monitored by the BS within the network.

After the system has reached a state in which time and frequency offsets of adjacent cells observed at a BS do not exceed a certain threshold, the system will be considered synchronized. In this state the data transmission during the uplink and downlink will be allowed.

If a small cell size is considered [12], an assumption can be made that signal delays from its own cell and from adjacent cells are not larger than the guard period at the receivers. That leads to the important fact that no ISI will occur. As mentioned, MTs communicate with BSs by means of a certain multiple-access scheme. A combination of the OFDM transmission technique with a frequency-division multiple-access (FDMA) concept, for example, has some advantages in a short-range scenario. An overview of a resource allocation in such a system is shown in Figure 2.3.

To achieve the fine synchronization, a set of dedicated test signals is introduced for the synchronization of MTs during the downlink. The DL test signals are embedded in the regular frame structure as a preamble. This is shown on the left in Figure 2.4.

During every frame a new test signal is randomly selected from the set of available test signals. By choosing a new test signal in the DL period for every frame, the influence of collisions between the test signals transmitted by adjacent cells can be reduced. This preamble is used by the MTs to synchronize themselves in time and frequency to the BS.

The accuracy of the employed synchronization procedure determines the resulting amount of interference that will be observed by a BS during the UL

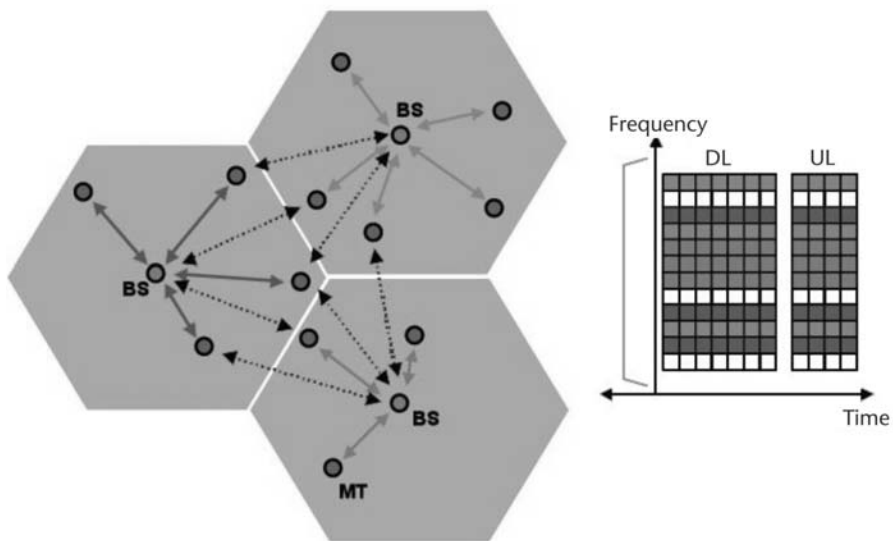


Figure 2.3 Bandwidth allocation in a TDD OFDM-FDMA SFN [12].

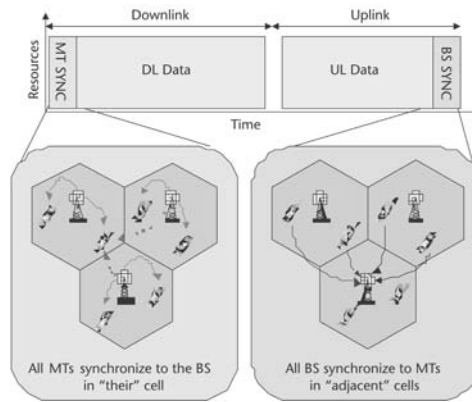


Figure 2.4 MAC frame structure with preamble and postamble test signals [12].

transmission. Thus, the performance that can be achieved in the data transmission during the uplink only depends on the synchronization accuracy that can be achieved by the MTs during the DL synchronization phase. No additional synchronization is carried out at the receiver during the uplink.

2.1.3 Coexistence and Spectrum Sharing

Next generation wireless systems will offer higher capacity, greater user bandwidth, and more reliable operation at the same or lower cost than second and third generation systems. At the same time they must operate with much higher spectral efficiency than current systems, since there is not likely to be an abundance of available frequency bands. A significant ingredient in the solution of the spectrum availability problem is the ability of different wireless systems, users, and service providers to share spectrum with one another, on an as-needed basis. The result will be a much more efficient use of common spectrum through sharing, as opposed to today's situation in which most of the electromagnetic spectrum is dedicated to exclusive "owners," some of whom use the spectrum only sporadically.

Coexistence is the concurrent operation of different services or RANs in the same or adjacent frequency bands without causing degradation to any service, with emphasis on the indicated limitations in terms of, for example, frequency separation, physical separation, and transmission powers.

Sharing is the use of the same frequency band by different RANs or services, either with coordination or possibly without any coordination between the systems, with emphasis on the spectrum access schemes and methods.

Efficient sharing capabilities are required because most frequency bands that are well suited for next generation emerging systems are already allocated and used to some extent by existing services. Therefore, future systems may not get sufficient dedicated frequency bands of their own, but will have to be able to use the same bands with other services. Of course, it must be guaranteed that existing services will not be significantly negatively affected. Sharing of the same frequency band by different services or technologies is only possible through well-defined limitations on the spectrum use, output power level, equipment density, transmission masks,

and so forth. Technical requirements that facilitate sharing capabilities are the transmitter power control, low-level protocols, and dynamic frequency selection [21].

Spectrum sharing includes the common use of bands between new technologies with existing ones, as well as between several other new technologies that are deployed in parallel in the same bands. An example of a new technology that shares bands with many existing ones is UWB, whereas WLAN and Bluetooth are a typical example for technologies, which are deployed at the same time and make use of overlapping frequency bands.

This spectrum sharing will be made possible by the concept of “cognitive radio”; entities wishing to use part of a frequency band for transmission first scan to determine which portions of the band are currently unused, quickly establish temporary occupancy of suitable subbands and transmit on them, and finally vacate the subbands when the transmission is finished, or hop to another set of subbands if a higher priority user requires the original set. Cognitive radio [22] becomes possible with recent advances in software-defined radios (SDRs), with intelligent spectrum-sensing techniques, and with transmitter and receiver digital signal processing (DSP) and RF technologies that accommodate signals whose spectra are split into several disjoint subbands and that can quickly change their transmitted spectra to adapt to the current multiuser environment without causing interference.

Figure 2.5 shows an example of a user signal with total bandwidth $B_1 + B_2$, which must be transmitted in available vacant subbands in a spectrum crowded with other user signals. It is split into two subbands to fill two available vacancies.

If the width of the common sharable spectrum block lies within the capabilities of the analog-to-digital and digital-to-analog converters and DSP implemented in the transceivers of the users, then most of the multiband transceiver signal processing can be done digitally. In this case the transmitter and receiver signal processing architectures can advantageously be based on efficient FFT operations [23].

If the total extent of the available band or bands exceeds state-of-the-art digital transceiver capabilities, analog processing at radio frequencies is necessary for generating and extracting user signals. This would require a study of the impact of the multiband scenarios on the radio front ends. Results of such studies are reported in [24, 25]. Another important topic related to feasibility of the spectrum use is the effects of the nonlinear power amplifier on the transmitted power spectrum. Coding, interleaving and predistortion considerations are also important in this context.

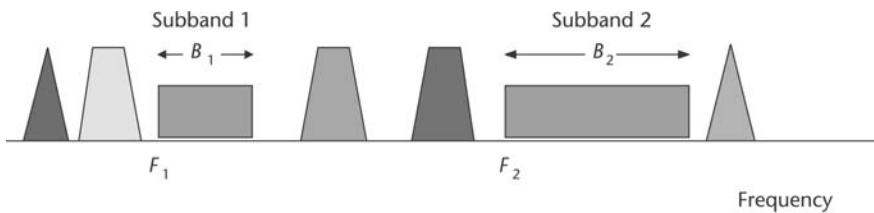


Figure 2.5 Example of a user spectrum, which must be split into two subbands to fit into a frequency band crowded with other users' spectra [23].

Future networks will very likely have to be able to coexist with some services and systems using the same or adjacent frequency bands. The deployment of these networks should be easy from the spectrum coordination point of view and, thus, the required coordination for transmitters and systems should be kept to a minimum. Coexistence of different services on adjacent bands is facilitated by grouping the services according to their power levels and interference sensitivity.

The issue of coexisting systems is closely related to the concept of dynamic spectrum allocation (DSA), because it is concerned with different radio systems sharing the same block of spectrum. However, in coexistence the concern is how to find the effect on the performance of the systems when they are simply both operating in the same spectrum band. This can either be investigated for systems operating on exactly the same frequencies (i.e., the cochannel case), or when the systems are operating in adjacent frequencies (i.e., the adjacent channel case). An interesting study of this concept deployed in 3G systems (UTRAN) and digital video broadcasting (DVB-T) is presented in [26].

2.1.4 Opportunities for Secondary Spectrum Use

Recent measurements of the spectrum occupancy as a function of time and frequency indicate that great parts of the spectrum, although dedicated to some service, are actually not used for significant periods of time. In these measurements, the received power at some location is measured and plotted over time and frequency. Spectrum is regarded as unoccupied when no power is detected at a certain time and frequency. Note that detecting whether a frequency band is in use is not a trivial task. Spread-spectrum systems in particular can operate at signal powers well below the noise level and may thus not be detected by simply measuring the power level. Figure 2.6 shows measurements of the spectrum occupancy that suggest how to make use of the so-called “spectrum holes” [27].

Although the measurement of the power spectral density over time in a limited frequency range seems feasible, it is more complicated to identify a spectrum hole that can be used by a nonprimary system without causing interference to the primary system. A basic problem of these measurements in the time-frequency plane is that the spatial dimension is missing. Although a frequency band is unused at the location where the measurement is taken, at another location within the

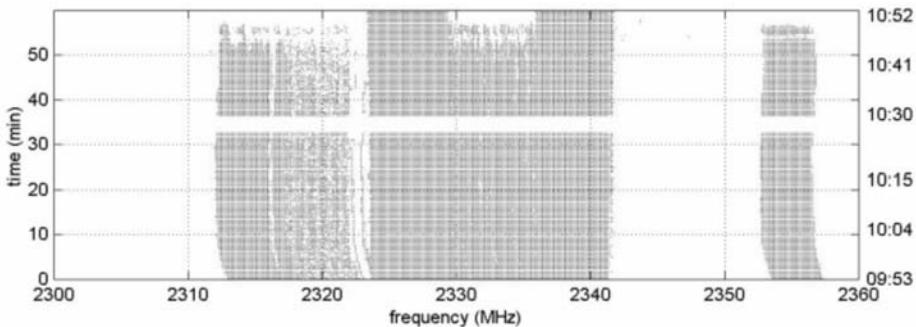


Figure 2.6 Measurement about the spectrum occupancy.

range of the secondary transmitter, it might be heavily used. On the other hand, there might be areas where a certain frequency band is not used over a very long time. These are referred to as temporal and spatial spectrum holes. There is no clear distinction, and most spectrum holes will be a combination of both, but often either the spatial or the temporal component will dominate. Possible examples for both cases can be summarized as follows:

- *Temporal holes* imply sporadic use of the spectrum by certain services (e.g., military use). Services that use the spectrum in a sporadic manner, that is, only a few times per day, are possible candidates for spectrum sharing.
- *Spatial holes* imply that some services use spectrum only in reduced geographical areas (e.g., WLAN in central urban areas or at airports).

If no coordination between the primary and the secondary systems is possible, the extent to which spectrum sharing is possible depends strongly on the characteristics of the primary system. Properties of the secondary system that favor uncoordinated spectrum sharing are as follows:

- *Directivity of radio emissions*, for example, in satellite links, radio links, or fixed wireless access with directive antennas. The more directive, the better.
- *Bidirectional links with fixed transmit powers* make it much easier for the secondary system to locate primary transceivers and estimate the path loss between primary (victim) receiver and secondary transmitters.
- Very narrowband or very broadband systems. If the primary system is narrowband, the influence of a broadband, low-power system might be negligible and vice versa.

Properties of the primary system that make uncoordinated spectrum sharing difficult are as follows:

- Broadcast systems or unidirectional links (in broadcast systems, generally no temporal holes are present and, what is more important, it is impossible for a secondary system to determine the locations of the receivers);
- Varying transmitting powers, for example, power control in CDMA;
- Adaptive modulation with power control;
- Spread spectrum with low SNR;
- Mobility of terminals.

This list suggests that IMT-2000 (i.e., 3G) systems are less adequate for uncoordinated secondary spectrum use. This can be explained by the fact that 3G systems were designed to use their transmitting power efficiently and they are highly adaptive.

Another aspect that has to be considered in this context is the coverage area. Wide-range radio networks influence or might be influenced by a great number of factors. These RANs normally use dedicated frequency bands below ≈ 2 GHz, due to the propagation conditions (satellites do not fit into this classification). On the

other hand, short-range radio networks cover only the range of a few individuals and, thus, the possibility of interfering with or being the victim of interference is drastically less than for wide-range systems. Nevertheless, because various RATs are often located in the same device, interference is not automatically prevented.

Short-range systems typically use frequencies above 2 GHz and often operate in shared frequency bands. Due to the limited propagation, interference issues are limited in space. This reasoning leads to the distinction between wide-range and short-range RANs and the tendency to operate wide-range RANs in dedicated spectrum at low frequencies and short-range RANs in shared spectrum at high frequencies.

2.1.5 Multiband Transmissions

Researchers suspect a single air interface mode is not sufficient for achieving the high spectral efficiency and peak bit rate targets set for next generation wireless systems with power consumption characteristics suitable for portable user terminals [28]. A dual-bandwidth approach for Beyond 3G systems was proposed in [29] to cope with these somewhat conflicting demands. In the proposal, peak bit rates were provided on a wideband channel, but due to challenges in, for example, arranging cellular reuse, UL range, mobile terminal power consumption, and scheduling of a wide array of data rates, a more narrowband channel format is also defined. The dual-bandwidth system is seamlessly linked through flexible scheduling capable of apportioning capacity across the two bands.

A similar multiband approach was also considered for the system concept proposed within the scope of the IST project WINNER [8]. Reasons to have more than one air interface mode include the following:

- *Range and terminal power consumption:* A combination of long range with good error performance and high spectral efficiency (high data rates) is very difficult to achieve with a single, wide-bandwidth carrier.
- *Optimization of cells:* One mode can be more optimized for small cells (hot spots), whereas another mode may be more robust and suitable for wide-area (vehicular) communication.
- *Support for a wide range of data rates:* Filling a single wideband carrier with small packets can be difficult in practice and can lead to very inefficient use of resources.
- *Flexibility in resource allocation:* Smaller amounts of data can be sent on a mode specifically designed for that purpose.
- *Varying services:* A variety of different services can require otherwise different air interfaces.

Based on the preceding reasoning, three air interface modes were defined for the next generation candidate air interface designed within the IST WINNER project to provide coverage in the three main deployment scenarios: wide-area (WAN), metropolitan-area (MAN), and local-area (LAN) networks.

2.1.5.1 Multiband Signal Structure for a Single User Occupying Several Disjoint Frequency Subbands Within a Wide Frequency Band

The need for multiband transmissions arises from the need for highly efficient usage by multiple users and even multiple services, which must share a common wide frequency band. A given user wishing to transmit a broadband signal may find that the common band currently has several vacant portions, none of which is wide enough to accommodate his entire signal, but whose aggregate width is larger than his current required bandwidth. The concept of multiband transmission is to allow the user to split his signal into several subbands, which can be transmitted in the vacant portions of the system band. Here, it is assumed that the transmitters and receivers of the users have RF front ends and analog-to-digital and digital-to-analog converters (ADCs and DACs, respectively) that can accommodate up to the entire common band. In this case, the main transmission and reception operations can be implemented with DSP. If this assumption is not true, for example, when the available transmission bands include several portions separated by more than several hundred megahertz, then multiple RF front ends, ADCs and DACs, and analog filtering operations are necessary [23].

Let's assume a system, in which data is transmitted in fixed-length blocks, separated by cyclic prefixes, which facilitate efficient transmitter and receiver frequency domain processing, using a generalized frequency-domain approach [12]. Assume a discrete Fourier transform (DFT) block size of N_c , and that $M_N (< N_c)$ data symbols must be transmitted using S separate, disjoint frequency “chunks,” or subbands, labeled for $i = 1, 2, \dots, S$, as shown in Figure 2.7, for $S = 2$.

Breaking up a transmitted spectrum into several subbands would be necessary if no sufficiently wide single span of M_N frequencies is available because of the occupancy of much of the available frequency band by other users. Initial spectral analysis of the frequency band would be used as a basis to determine the number and allocation of the subbands. Frequency-domain generation and reception of multiband signals is very flexible and efficient. Receivers will include wideband DFT operations on incoming signal blocks. These operations also serve as a spectral analysis tool for observing the current occupancy of the overall frequency band.

The number of subbands necessary to accommodate a given block of data would depend on how crowded the band is with other users, as well as on the number of data symbols, M_N , to be transmitted in the block. We anticipate that the number of subbands will typically be small—on, say, the order of $S = 2$ or 3. The total bandwidth occupied by the transmitted signal is equal to the bandwidth of the M_N subbands, but the span of frequencies, including gaps, may be larger than this, but less than N_c . In the initial WINNER assumptions, $N_c = 2,048$,

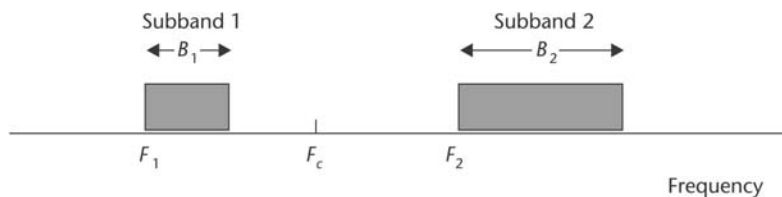


Figure 2.7 Required RF spectrum (F_c is the chosen carrier frequency) [23].

corresponding to a bandwidth of about 100 MHz, and an interfrequency spacing of about 50 kHz. The value of M_N could range from perhaps about 16 to 1,024.

Figure 2.7 shows a scenario in which the data is to be transported over two disjoint subbands of widths B_1 and B_2 Hz, using a carrier frequency F_c . The choice of F_c can be governed by RF implementation considerations. For example, choosing it in the empty region between the two subbands may reduce dc bias problems in a zero-IF receiver.

Figure 2.8 shows the baseband equivalent spectrum, along with the frequency response of an antialiasing filter for DAC converting from a digitally generated oversampled complex waveform, with a sampling rate of B . A practical value for the fixed sampling rate B would be twice the overall system bandwidth. The total range of frequencies spanned by the multiband spectrum is assumed to be within the capability of system ADC and DAC.

Figure 2.9 shows the equivalent periodic baseband spectrum, with frequencies represented by integers, for use with DFTs and their inverses. The DFT size N_c corresponds to the sampling rate B . The integer frequencies shown are as given by (2.1):

$$f_1 = N_c \left\lceil \frac{F_2 - F_c}{B} \right\rceil, f_2 = N_c \left\lceil \frac{F_1 - F_c}{B} \right\rceil, M_1 = N \frac{B_2}{B}, M_2 = N \frac{B_1}{B} \quad (2.1)$$

For $S = 2$ subbands, $M_1 + M_2 = M_N < N/2$. For a general number S of subbands, the M_N data symbols are transmitted on a set of M_N discrete frequencies denoted by

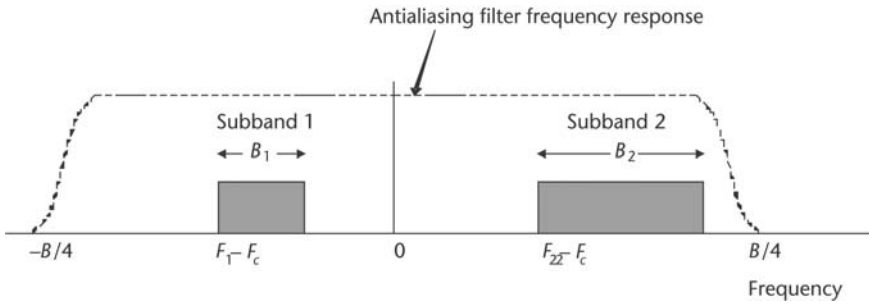


Figure 2.8 Baseband equivalent spectrum [23].

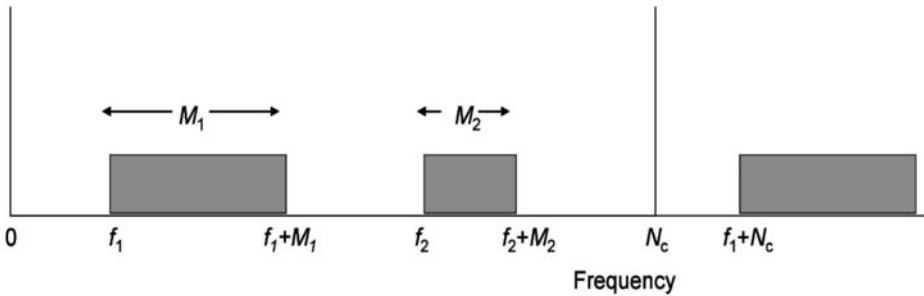


Figure 2.9 Periodic baseband equivalent spectrum, with frequencies represented by integers [23].

$$\{f_1, f_1 + 1, \dots, f_1 + M_1 - 1, f_2, f_2 + 1, \dots, f_2 + M_2 - 1, \dots, f_S, f_S + 1, \dots, f_S + M_S - 1\}$$

where the set $\mathfrak{F}_i = \{f_i, f_i + 1, \dots, f_i + M_i - 1\}$ forms the i th subband, and transports M_i data symbols, and

$$\sum_{i=1}^S M_i = M_N \quad (2.2)$$

Each block will be preceded by a cyclic prefix, and may also be preceded and followed by training symbol sequences, of length M_T symbols, where both the cyclic prefix and M_T exceed the maximum expected channel impulse response span, measured in symbol intervals. A typical block format with training sequences, which also act as cyclic prefixes, is shown in Figure 2.10.

In Figure 2.10, the second training sequence acts as a cyclic prefix for the M_N -length block consisting of data plus the third training sequence. The first training sequence acts as the cyclic prefix for the second training sequence. If data is being transmitted continuously, as a succession of blocks, the first training sequence shown in Figure 2.10 would be the final training sequence of the preceding block. Training sequences are generally chosen to have flat power spectra, and constant envelopes. Examples are found in [25] and [30].

The block diagram of the generalized frequency-domain transmitter used in this concept is shown in Figure 2.11. Figure 2.12 shows the details of the block labeled “matrix time-frequency space selector” in Figure 2.11. The space-frequency selector matrix maps M_N input samples onto frequencies in the sets of data $\{\alpha_i\}$, and, if so desired, also onto different antennas in a MIMO system. The M_N data symbols are represented by $\{A(\ell) = a(\ell), \text{ for } \ell = 0, 1, \dots, M_N - 1\}$. For OFDMA, the $\{A(\ell)\}$ are the coded data symbols themselves. This set of data is partitioned

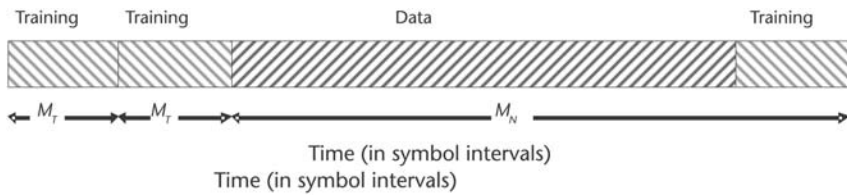


Figure 2.10 Typical transmitted block format [23].

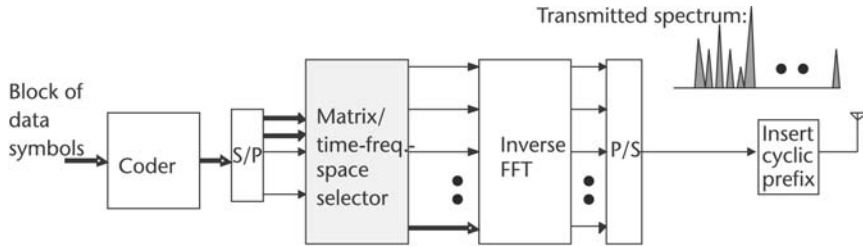


Figure 2.11 General frequency-domain transmitter [23].

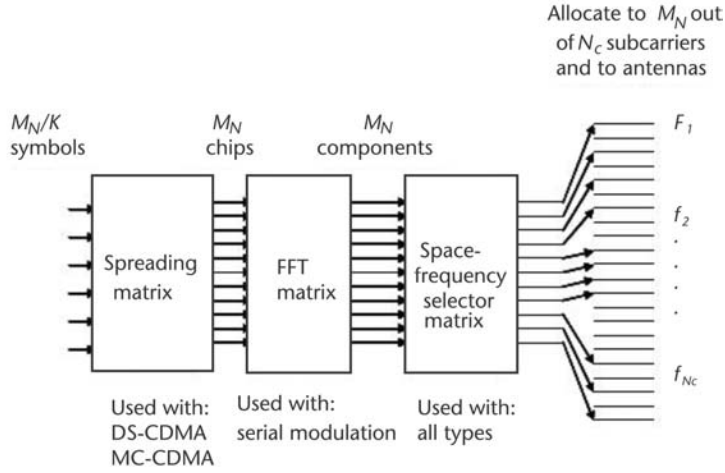


Figure 2.12 Details of the matrix time-frequency selector of Figure 2.11.

into two sets $\alpha_1 = \{A(\ell)\}_{\ell=0}^{M_1-1}$ and $\alpha_2 = \{A(\ell)\}_{\ell=M_1}^{M_N}$, where α_1 is mapped onto frequency set \mathfrak{F}_1 , and α_2 is mapped onto frequency set \mathfrak{F}_2 . In the case of $S > 2$, α_i would be mapped onto frequency set \mathfrak{F}_1 . Zeros are mapped onto the remaining $N_c - M_N$ frequencies constituting the block. The signal waveform is then generated from the N_c -point inverse DFT of the resulting block of data symbols and zeros where we have defined $M_0 = 0$. If more than one user's signal is to be frequency multiplexed in the downlink, each user's signal would occupy and be generated in this way for a disjoint set of frequency subbands.

The FFT matrix is omitted for OFDMA, but is used for serial modulation. If spreading is used, for MC-CDMA or DS-CDMA, the spreading matrix is used to map a sequence of M_N/K data symbols into a larger number, M_N spreading chips, where K is the spreading gain [23]. The corresponding receiver structure is shown Figure 2.13.

For the case of OFDMA, if the signal, with cyclic prefix, is considered as periodic, with period N_c , it will have a line spectrum

$$S(f) = \begin{cases} A(f - f_i) & \text{for } f \in \mathfrak{F}_i \\ 0 & \text{otherwise} \end{cases} \quad (2.3)$$

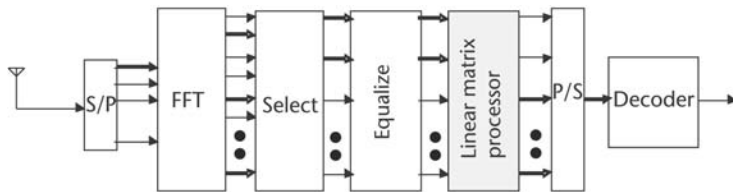


Figure 2.13 General frequency-domain receiver structure [23].

The actual transmitted spectrum will be this line spectrum convolved with the $\frac{\sin(\pi f)}{\sin(\frac{\pi f}{N_c})}$ spectrum of the rectangular time window of N_c samples. The resulting

relatively high sidelobes can be suppressed by using a nonrectangular time window on the block plus its cyclic prefix, such as a raised cosine window. Figure 2.14 shows an example of the line spectrum of a multiband signal with S subbands. For S subbands, the k th sample of the transmitted multiband waveform is then as follows:

$$s(k) = \frac{1}{N_c} \sum_{i=1}^S \sum_{n=0}^{M_i-1} A \left(n + \sum_{i'=0}^{i-1} M_{i'} \right) \exp \left(j \frac{2\pi k(f_i + n)}{N_c} \right) \quad \text{for } k = 0, 1, \dots, N_c - 1 \quad (2.4)$$

Because of the oversampling, there will be at least two samples per data symbol and, thus, a relatively simple fixed lowpass analog filter can be used for DAC conversion.

At the receiver, after downconversion and ADC conversion, the received block, not including the cyclic prefix, is processed by an N_c -point DFT, resulting in

$$R(f) = \begin{cases} A(f - f_i)H(f) + N(f) & \text{for } f \in \mathcal{F}_i \\ N(f) & \text{otherwise} \end{cases} \quad (2.5)$$

In (2.5), $H(f)$ is the channel frequency response at frequency f , and $N(f)$ is the DFT of noise and other unwanted signals at frequency f . The term $R(f)$ is then sampled in the frequency sets $\{\mathcal{F}_i\}$; this results in M_N frequency-domain samples, which are concatenated to form $\{Y(\ell)$, for $\ell = 0, 1, \dots, M_N\}$, which is given by

$$Y(\ell) = A(\ell)H(\ell + f_i) + N(\ell + f_i) \quad \text{for } \ell = \sum_{i'=0}^{i-1} M_{i'}, \dots, \sum_{i'=0}^i M_{i'} \quad \text{and } i = 1, \dots, S \quad (2.6)$$

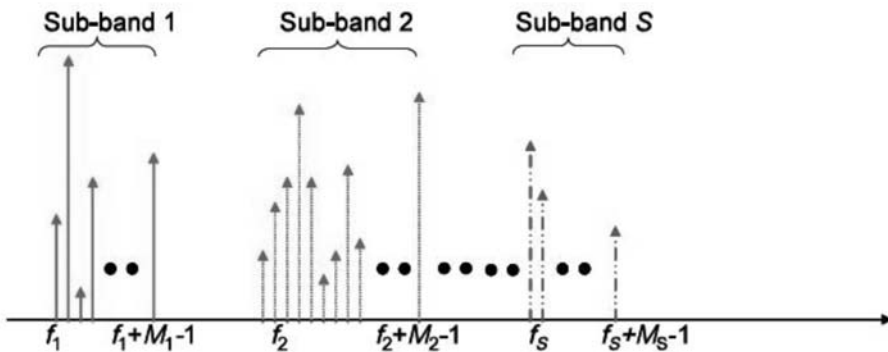


Figure 2.14 Example of a multiband line spectrum [23].

These samples are equalized by the appropriate scalar multiplication of each frequency component and are then passed on to the detector and decoder.

OFDMA is a multiplexing/multiple-access scheme that allows different user signals to coexist within a common DFT block. Therefore, it is very flexible and well suited to the multiband concept, as long as the common frequency band can be processed by each terminal's RF and data conversion components. As with any OFDM or OFDMA system, the local oscillators of transmitters and receivers must be stable and accurate to within a small fraction of the intersubcarrier frequency spacing to minimize ICI between adjacent user spectra.

For serial modulation, the data symbols are in the time domain, and so the frequency-domain symbols $\{A(\ell), \text{ for } \ell = 0, 1, \dots, M_N - 1\}$, are the M_N -point DFT coefficients of the data symbol sequence, or

$$A(\ell) = \sum_{m=0}^{M_N-1} a(m) \exp\left(-j \frac{2\pi m \ell}{M_N}\right), \quad \text{for } \ell = 0, 1, \dots, M_N - 1 \text{ for serial modulation} \quad (2.7)$$

where the $\{a(m)\}$ represent the data symbols. The other operations at the transmitter are identical to those described for OFDMA. The receiver operations are also the same, except that, after equalization and before detection and decoding, an M_N -point inverse DFT operation is performed on the equalized frequency-domain samples to recover the time-domain data.

This approach of transporting the DFT of a data symbol sequence on an arbitrary set of frequencies has been called *carrier interferometry* [31].

By substituting the FFT of (2.7) into (2.4), we see that the sampled waveform can be expressed as

$$s(k) = \frac{1}{N_c} \sum_{i=1}^S \exp\left(j \frac{2\pi k f_i}{N_c}\right) \sum_{m=0}^{M_N-1} d_i(m) g_i\left(k - m \frac{N_c}{M_N}\right) \quad (2.8)$$

where

$$d_i(m) = a(m) \exp\left(-j \frac{2\pi m}{M_N} \sum_{i'=0}^{i-1} M_{i'}\right) \quad \text{for } m = 0, 1, 2, \dots, M_N - 1 \quad (2.9)$$

and

$$g_i(k) = \sum_{n=0}^{M_i-1} \exp\left(j \frac{2\pi n k}{N_c}\right) = \exp\left(j \frac{\pi(M_i-1)k}{N_c}\right) \frac{\sin\left(\frac{\pi M_i k}{N_c}\right)}{\sin\left(\frac{\pi k}{N_c}\right)} \quad (2.10)$$

for $k = 0, 1, 2, \dots, N_c - 1$

Equation (2.9) represents a rotated complex data symbol. Equation (2.10) represents a sampled pulse waveform, similar to a modulated sinc pulse. Thus,

the waveform expression of (2.8) represents the sum of S complex single-carrier waveforms, upconverted by the frequencies $\{f_i\}$. For relatively small values of S , say, 2 or 3, its peak-to-average power ratio (PAPR) will be significantly less than that of a corresponding OFDMA signal, which is the sum of M_N complex exponential waveforms. The smaller dynamic range and consequent lower transmitting power back-off requirement is expected to be the main advantage of this serial modulation approach. This is especially important for MTs, because the peak power requirement (which can be considered as the average power required plus the back-off) is a major factor in determining the cost of portable wireless terminals.

The transmitted serial modulated signal is still in the form of Equation (2.4) can also be considered a “multicarrier” signal. However, using the DFT as in (2.7) reduces the dynamic range of the signal. Thus, the DFT operation can be considered to be a linear preprocessing operation on an OFDMA signal, which effectively reduces its PAPR. Furthermore, it is a very simple operation when compared to alternative linear operations such as selective mapping and partial transmit sequences [12, 23] and requires no side information to be transmitted. If the number of subbands S is relatively large, the composite serial modulated waveform will be the sum of a large number of serial modulated waveforms, and its dynamic range advantage relative to OFDMA will be lost. Thus, OFDMA may be advantageous to use for large S . It may also be advantageous for transmission from BSs to MTs, because a base station will in any case be transmitting many signals in parallel to multiple users, and its power amplifier costs, shared over many users, are not as critical as those of the user terminals.

For both OFDMA and serial modulation, essentially the same multiband average power spectra will be produced. Figure 2.15 shows the average power spectra for serial modulation and OFDMA for two subbands, with $M_N = 256$ data symbols

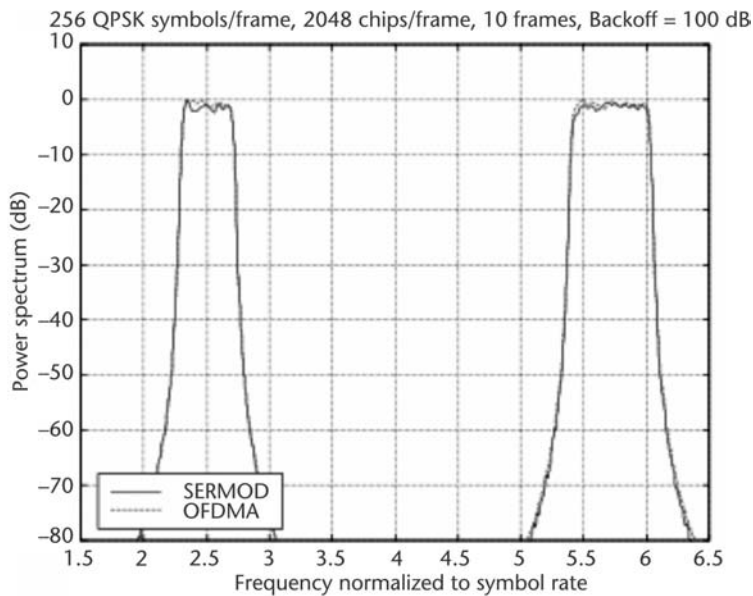


Figure 2.15 Power spectra for $S = 2$, $N_c = 2048$, $M_N = 256$ for OFDMA and serial modulation [23].

within a total bandwidth of $N_c = 2,048$. The other parameters for this figure are $f_1 = 589$, $M_1 = 103$, $f_2 = 1,390$, and $M_2 = M_N - M_1 = 153$.

2.1.5.1.1 Accommodation of Narrowband Terminals.

Individual user terminals with smaller bandwidth capabilities can also be accommodated. For example, suppose the system's common bandwidth is 100 MHz, but a certain terminal A has a bandwidth capability of only 10 MHz. Then the base station constrains its transmitted spectrum to A to have a total bandwidth (e.g., $F_2 + B_2 - F_1$ in Figure 2.7) of less than 10 MHz. Terminal A's receiver samples the received block at a rate of 20 MHz, resulting in $1/10$ the number of samples that would have resulted from sampling the entire 100-MHz band at 200 MHz. The receiver takes the $(N_c/10)$ -point DFT of the sampled block, samples the result at the frequency sets \tilde{f}_1 and \tilde{f}_2 used by the transmitter, and equalizes these samples, resulting in recovery of the transmitted data sequence. For uplink transmission from the narrowband terminal A to the base station, the process is reversed: The terminal generates a block of $(N_c/10)$ samples at the 20-MHz sampling rate, and the base station receiver, sampling at 200 MHz, takes in and processes (by DFT, sampling and equalization) a corresponding oversampled block of N_c samples. The same approach can be used for OFDMA and for serial modulated signals.

2.1.5.1.2 RF Implementation Issues.

The desired signal spectrum may share a wide band with adjacent unwanted signal spectra with higher power. The receiver ADC must then have sufficient dynamic range to preserve the desired signal waveform. For example, suppose the total power of the unwanted adjacent signals is 40 dB above that of the wanted signal. Then the ADC must have roughly $40/(6 \text{ dB/bit}) \cong 7$ extra bits of accuracy beyond what is required for the desired signal received without adjacent channel interference [23].

2.1.5.1.3 MC-CDMA and DS-CDMA.

Multicarrier CDMA (MC-CDMA) and direct sequence CDMA (DS-CDMA) are spread-spectrum signaling systems with advantageous robustness against multipath and multiuser interference. The former is related to OFDM/OFDMA and the latter to serial (single-carrier) modulation. The multiband transmitter and receiver processing described here can be applied with no essential modification to these spread-spectrum schemes, if the data symbols $[A(m)]$ for OFDMA and $a(m)$ for serial modulation] are replaced by *chips*. It makes no difference to the transmitter operation whether it is transmitting a block of data symbols or a block of chips, as long as both appear to be uncorrelated zero-mean sequences. The DFT, frequency sampling, and equalization processes at the receiver produce equalized chips instead of data, ready for subsequent despreading. As in conventional single-band transmission, minimum mean squared error multiuser detection can be employed in both MC-CDMA and DS-CDMA by applying a linear matrix operation on the frequency-domain samples.

Frequency-domain orthogonal signature sequences (FDOSS) constitute another form of spectrum spreading of serial modulated signals [32]. For a single-band system ($S = 1$), it can be generated as described for serial modulation but with the

frequency set \mathcal{F}_1 consisting of frequencies equally spaced at intervals of K , where K is the spreading gain. This equal-spaced frequency constraint restricts the flexibility of the transmitter to generate several subbands with arbitrary bandwidths and center frequencies. As a result, FDOSS was considered of less importance for the multiband concept [6].

2.1.5.1.4 OFDM/IOTA.

OFDM/IOTA is an alternative to classical OFDM and OFDMA, which does not require the use of a cyclic prefix, therefore, reducing overhead [33, 34]. The data symbols are carried using offset QAM modulation, with filtering by a Gaussian-like waveform following the inverse DFT operation.

OFDM/IOTA was also considered in the design of the WINNER air interface [2, 11, 12].

Figure 2.16 shows a detail of the multicarrier spectra for $N_c = 512$. Only the edge of the useful band is visible. Note that the IOTA spectrum decreases dramatically compared to the classical OFDM one. This particularity of OFDM/IOTA is a consequence of the very high localization of the IOTA filter. The IOTA spectrum is not decreasing linearly on the whole bandwidth; this can be explained by the extreme low level of the out-of-band power (around -100 dB).

The IOTA principle can be applied to multiband signals, as shown in Figure 2.17. The subband parameters are the same as those listed for Figure 2.14. The spectrum is essentially the same as that of the OFDMA signal in Figure 2.14 apart from some low-level artifacts, which are likely due to numerical round-off effects. In contrast to the raised cosine time windowing employed in the signals of Figure

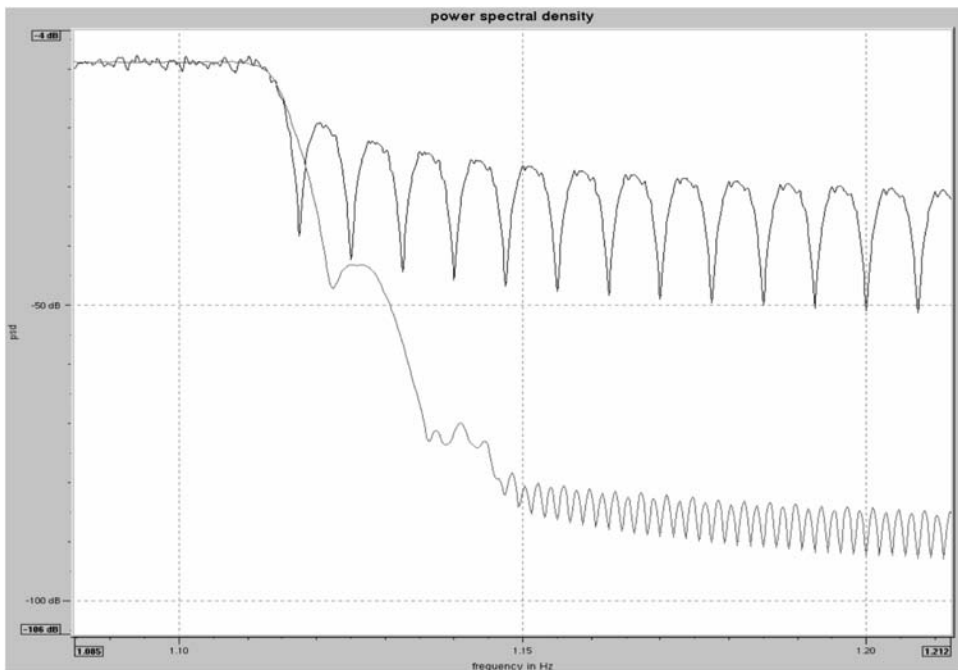


Figure 2.16 Detail of the classical OFDM and OFDM/IOTA spectra ($N_c = 512$) [23].

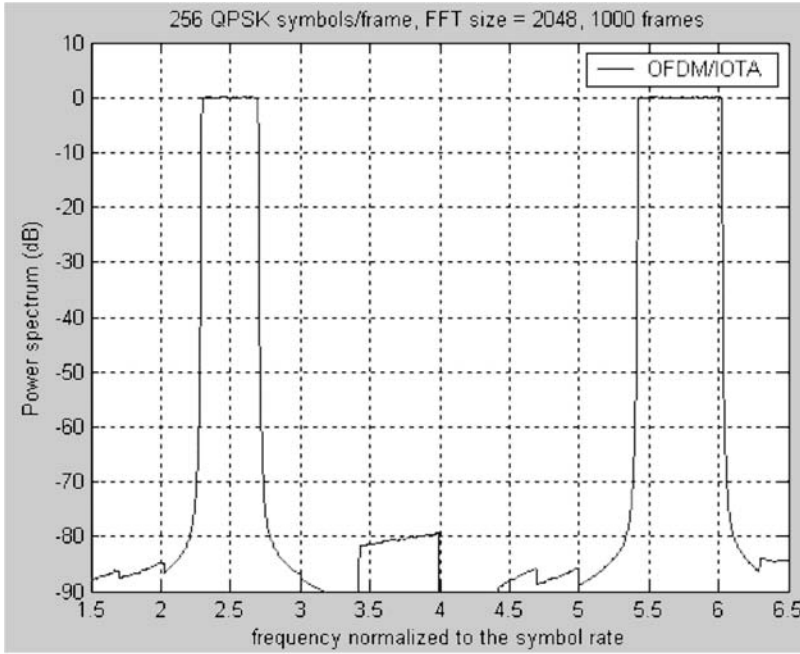


Figure 2.17 Power spectra for $S = 2$, $N_c = 2,048$, and $M_N = 256$ for OFDM/IOTA [23].

2.14, no additional filtering or windowing was employed to generate the IOTA signal.

2.1.5.1.5 Frequency Offset and Phase Noise.

Frequency offset and phase noise are impairments to the received signal that result in ICI for OFDM, OFDMA, and MC-CDMA systems [35, 36]. These impairments are especially significant for uplink signals from multiple users with different local oscillators. For OFDM, OFDMA, and MC-CDMA, the degradation from a frequency offset Δf is approximately proportional to $(N_c \Delta f)^2$, where N_c is the number of subcarriers. Receiver signal processing techniques for OFDM systems exist that can counter frequency offset and phase noise, but they are relatively complicated, since they involve mitigation of frequency-domain ISI [37, 38].

In the absence of sophisticated receiver signal processing techniques to remove frequency offset and phase noise, the intersubcarrier spacing should be about 100 kHz or more; lower values for OFDM systems may require unreasonably expensive or power-hungry local oscillator (LO) subsystems [23]. For a 100-MHz system bandwidth, a 100-kHz subcarrier spacing results in an FFT block duration of about $10 \mu\text{s}$. This may compromise the robustness of large-bandwidth systems to multipath delay spreads of this order of magnitude.

For serial modulation a constant frequency offset just produces a slowly linearly increasing or decreasing phase shift over the received sequential data symbols. There is negligible additional ISI. The slow phase-shift variation is easily estimated by simple decision-directed techniques, and it is removed at the receiver prior to detection. These frequency offset sensitivity properties of OFDMA and serial modulation are shown in Figure 2.18. The signals have the split-band transmitted

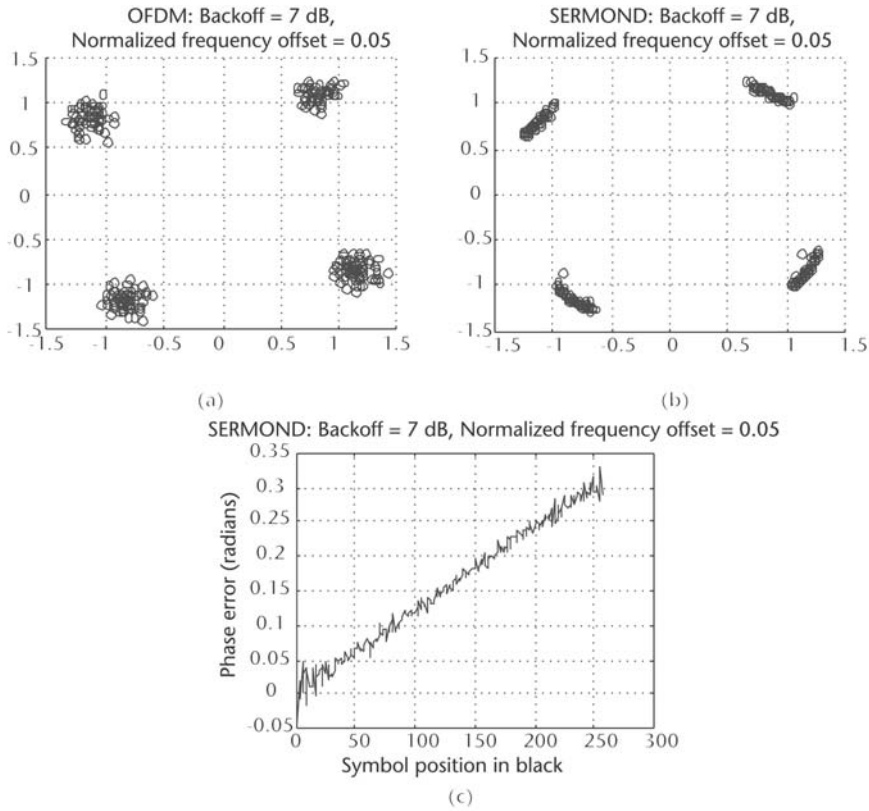


Figure 2.18 (a) Output signal constellation for two-band OFDMA signal, (b) output signal constellation for two-band serial modulated signal, and (c) output phase error variation over a 256-symbol block for serial modulation.

spectra of Figure 2.15, and are subject to frequency offset equal to 5% of intersubcarrier spacing.

Figure 2.18(a) shows the received QPSK signal constellation, measured over one DFT block, for a two-band OFDMA signal after it passes through a Rapp model nonlinear power amplifier with a 7-dB back-off and undergoes a frequency offset equal to 5% of the intersubcarrier spacing. The effect of ISI resulting from the frequency offset is noticeable in the relatively large variations of the “clouds” of received points from their nominal values. It is worth noting that most of the degradation here is due to frequency offset rather than nonlinearity; with no nonlinearity the picture is very similar. The main effect of the nonlinearity at this back-off level is to raise the power spectral sidelobes of the transmitted signal [23]. Figure 2.18(b) shows the corresponding output signal constellation for the serial modulated waveform under the same conditions. The errors between the received points and their nominal values are mostly angular. Figure 2.18(c) shows the angular error over the 256-symbol serial modulated block. It is an almost linear variation in time, with a slope proportional to the frequency offset. There are outliers on the ends whose effects could be eliminated by not transmitting data on those few symbols. A simple decision-directed technique can easily predict this slow variation of the phase error, enabling it to be corrected.

This preliminary study was carried out in the early phase of the design of the WINNER air interface and indicated that serial modulation may have an advantage over parallel modulation (OFDM, OFDMA, MC-CDMA, and so on) in the simplicity of mitigation of frequency offset and phase noise. This can be of importance for the uplink of cellular systems, in which low-cost terminals, with inexpensive local oscillators will be transmitting to a base station.

Another form of frequency multiplexing of different users would be *frequency interleaving*, in which the spectral lines of each user may be closely interleaved with those of its neighbors. In this case there will be *interuser interference* as well as ISI [36]. Interuser interference is potentially more damaging in the uplink because different received user signals may be at different power levels. FDOSS for serial modulated systems [32] also interleaves different users' spectra, and has the same interuser interference problem as interleaved OFDMA [32]. These considerations diminish the attractiveness of spectrum-interleaved OFDMA and FDOSS for the uplink of multiband, multiuser systems.

2.1.5.2 Channel Estimation and Training

2.1.5.2.1 Use of Training Sequences.

A training sequence, for channel estimation and equalizer adaptation, is assumed to be a sequence of M_T symbols, such that the ratio M_N/M_T is an integer, I_T . It is preceded by a copy of itself, which acts as a cyclic prefix (see Figure 2.10). The transmitted training sequence must be generated to occupy the same disjoint frequency subbands as for the data. Furthermore, the block of training symbols must consist of N_c/I_T samples. Let $\{A_T(\ell), \ell = 0, 1, \dots, M_T - 1\}$ represent the DFT coefficients of the training sequence. These M_T coefficients are mapped onto an equally spaced grid of M_T frequencies obtained by decimating the frequency sets $\{\mathcal{F}_i, i = 1, 2, \dots, S\}$ by a factor of I_T . The training waveform is generated by an N_c/I_T -point inverse DFT operation to produce the sampled waveform of (2.11):

$$s_T(k) = \frac{1}{\frac{N_c}{I_T}} \left[\sum_{i=1}^S \sum_{n=0}^{M_i/I_T-1} A_T \left(n + \sum_{i'=0}^{i-1} \frac{M_{i'}}{I_T} \right) \exp \left(j \frac{2\pi k \left(\frac{f_i}{I_T + n} \right)}{\frac{N_c}{I_T}} \right) \right] \quad (2.11)$$

for $k = 0, 1, \dots, \frac{N_c}{I_T} - 1$

The received training signal is subjected to an N_c/I_T -point DFT operation. The resulting frequency-domain sequence is decimated according to the aforementioned frequency grid. This results in S sets of frequency-domain coefficients. The i th set has M_i/I_T coefficients. Because each set corresponds to a disjoint subband that may be far from its neighbors, interpolation is done individually on each subband set to produce M_i frequency-domain samples for the i th set. Then the channel response and/or equalizer coefficients are estimated for each of the S subbands [12].

2.1.5.2.2 Use of In-Band Pilots in OFDMA.

In OFDMA signaling, some of the data symbols can be replaced by known pilot symbols, resulting in a slight reduction in net data rate. The pilot symbols can be used in addition to or instead of training sequences located at a regular grid of frequencies in each subband, and the channel frequency response in each subband can be estimated by interpolation within the subband.

2.1.5.2.3 Use of Pseudorandom Postfix (PRP).

In the pseudorandom postfix (PRP) approach, the cyclic prefix at the beginning and end of each block is replaced by a fixed known sequence weighted by a pseudorandom scalar sequence [39]. The advantage is that the cyclic prefix and channel estimation functions are fulfilled with less overhead than with the separate cyclic prefix and training sequence approach. The purpose of the pseudorandom weighting sequence is to avoid spectral lines resulting from the periodicity of the fixed sequence. The channel response is estimated from frequency-domain processing of the PRP portion of the received block as detailed in [39]. The effect of interference from data symbols is mitigated by averaging over many blocks. Once the channel is estimated, the interfering effect of the inserted known PRP sequences is computed and subtracted, leaving the equivalent of zero-padded block transmissions, which can be equalized by a variety of techniques [40].

2.1.5.3 Impact of Nonlinear Distortions on the Spectrum Mask

Neighboring desired and undesired user spectra may be received with large power variability due to differing path losses. Avoidance of adjacent channel interference then requires low transmitted power spectral sidelobes and rather stringent spectral masks. For example, allowable interference to adjacent-frequency receivers is usually specified in terms of maximum interference power at a certain distance and at a certain frequency offset from the interferer's carrier. Under typical transmitted power and path-loss conditions, this may imply spectral masks with as much as 40 to 60 dB of attenuation of transmitted spectrum sidelobes at a normalized frequency offset of 1. Nonlinear distortion caused by the transmitter's high power amplifier (HPA) is a major factor influencing spectral sidelobe levels. Power back-off and transmitter signal processing techniques designed to reduce the dynamic range of the transmitted waveform can be used to reduce spectrum sidelobe levels. In addition, RF transmitter predistortion techniques can be used.

Figure 2.19 shows the complementary cumulative distribution density function (CCDF) of the PAPR for OFDM with various numbers of subcarriers. The PAPR is a measure of the distribution of the highest peak within a given block, which is one OFDM symbol. If the number of subcarriers exceeds 1,000 or so, the PAPR does not significantly increase. This can be explained intuitively by the fact that excessively high peaks, which are caused by constructive interference of most subcarriers, only happen with diminishing probability. On average, the number of subcarriers, which add up constructively, remains stable for high N_c .

Also shown in Figure 2.19 is the instantaneous to average power ratio (IAPR), which could be a more meaningful measure for nonlinear distortions of the HPA than the PAPR. The PAPR is a measure for the distribution of the highest peak within a given block, typically one OFDM symbol, T_s , defined by

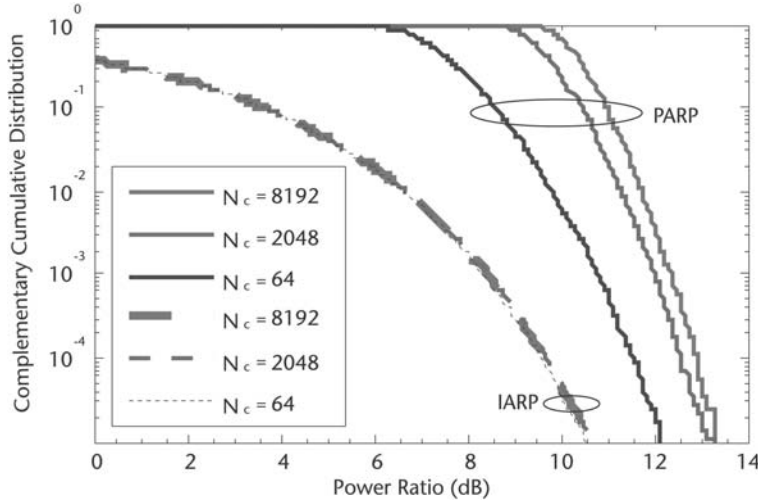


Figure 2.19 Statistical distribution of an OFDM signal's IAPR and PAPR with different number of subcarriers N_c .

$$\text{PAPR}(k)|_{\text{dB}} = 10 \cdot \log_{10} \left(\frac{\max |s(t)|^2}{E\{|s(t)|^2\}} \right) \Big|_{kT_s \leq t < (k+1)T_s} \quad (2.12)$$

However, the nonlinear distortions are not only caused by the highest peak, but by all peaks, which come close to or exceed the back-off of the HPA. This fact is taken into account by the definition of the IAPR:

$$\text{IAPR}(t)|_{\text{dB}} = 10 \cdot \log_{10} \left(\frac{|s(t)|^2}{E\{|s(t)|^2\}} \right) \quad (2.13)$$

It is interesting to note that the cumulative distribution of the IAPR is independent of the number of subcarriers, N_c . Studies of the power spectral density (PSD) verify that *for OFDM with HPA the out-of-band radiation is virtually independent of the number of subcarriers, N_c* . So, in terms of out-of-band radiation, the number of subcarriers is not a critical parameter.

In Figure 2.20 the power spectral density (PSD) is plotted against the normalized frequency f/B , where B is the system bandwidth, and the number of subcarriers is set to $N_c = 2,048$. For comparison purposes the spectrum mask of the HIPERLAN/2 standard, appropriately scaled to the system bandwidth B , is also plotted. Note that in the multiband scenario, with terminals in various locations coexisting with closely neighboring subbands, the requirements for the spectral mask will have to be more stringent than that of HIPERLAN/2. In the context of multiple-bandwidth transmission the appropriate choice of the spectrum mask will be an important issue.

Figure 2.20 shows the spectrum for OFDM without nonlinear distortion (i.e., without HPA). In this graph the impact of the time-domain windowing on the PSD

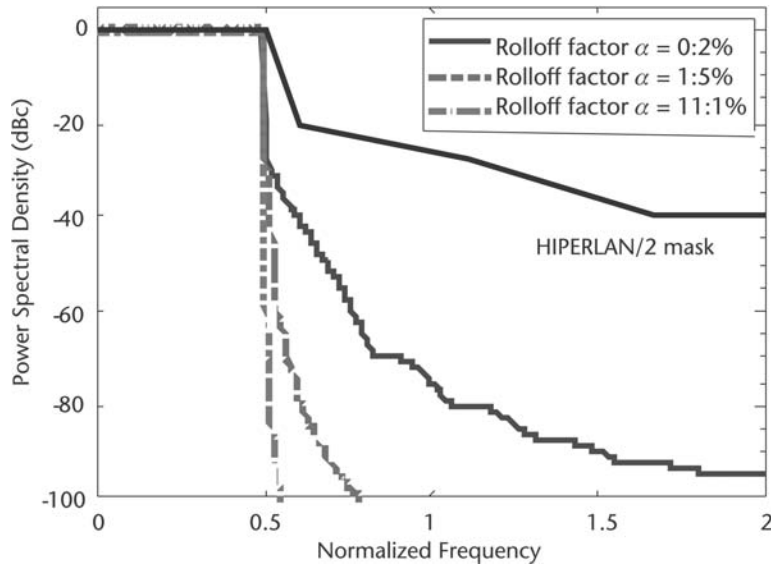


Figure 2.20 PSD versus the normalized frequency, f/B , for OFDM with raised cosine pulse shaping having a roll-off factor α . No HPA was used, that is, without nonlinear distortion.

is examined. As a window function the slopes of the rectangular time-domain window $g(t)$ have been smoothed by a raised cosine with roll-off factor α . We can see that for the HIPERLAN/2 power mask, a roll-off factor as small as 0.2% is sufficient. Even for multiband transmission, with more stringent spectrum mask requirements, a roll-off factor of 1% to 2% should be sufficient. This is particularly true when taking into account HPA nonlinearities, which cause significantly more out-of-band radiation. This fact will be studied in more detail in the figures that follow.

Figure 2.21 shows the PSD against f/B for various input back-offs (IBOs) of the HPA. An ideal HPA is used, which is linear up to saturation; that is, the amplifier distorts those parts of the input signal whose power exceeds the chosen back-off. The figure demonstrates that an IBO of 3 dB does not even meet the relatively mild requirements of the HIPERLAN/2 spectrum mask. Increasing the IBO leads to significantly lower out-of-band radiation, at the expense of reduced power efficiency of the transmitter. Especially for the uplink, power consumption of the mobile terminal is a critical issue. To reduce out-of-band radiation, peak reduction techniques have been proposed in the literature. The target of peak reduction algorithms is to get as close as possible to the PSD without HPA from Figure 2.20.

Although peak reduction techniques can reduce the effects of nonlinearities of the HPA, they require additional signal processing at the transmitter. Most peak reduction techniques require a few additional FFTs, and the computational complexity is proportional to the complexity of an FFT, which is $N * \log N$. To reliably detect peaks, oversampling is applied. In conclusion, the FFT length, N , is higher than the number of subcarriers, N_c , typically by a factor of 2 or higher. There are also techniques of complexity order N and of polynomial order.

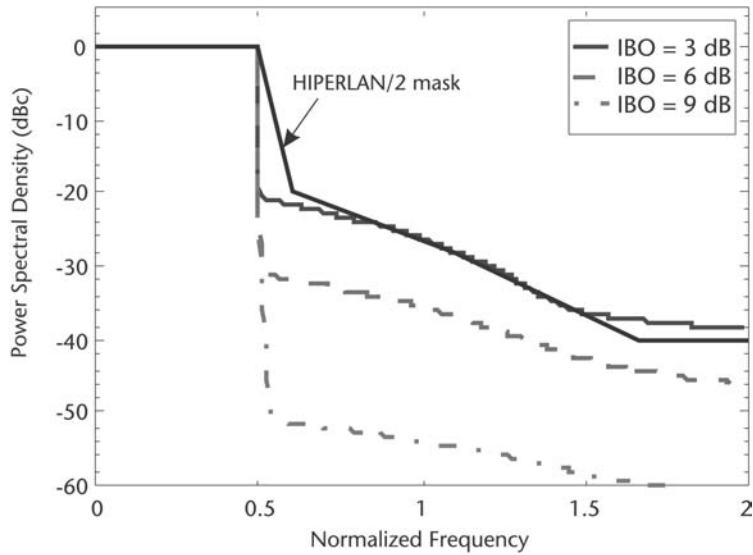


Figure 2.21 PSD against the normalized frequency, f/B , for various IBOs of the HPA. An ideal HPA is used, which is linear up to saturation.

Figure 2.22 shows the PSD against f/B for different amplifier models: Rapp’s model with parameter $p = 2$, to characterize a solid-state power amplifier (SSPA), is compared with an ideal HPA having an idealized linear characteristic below saturation. The comparison between the SSPA model and the ideal HPA is interesting because the ideal HPA plot is the best that can be achieved with predistortion techniques.

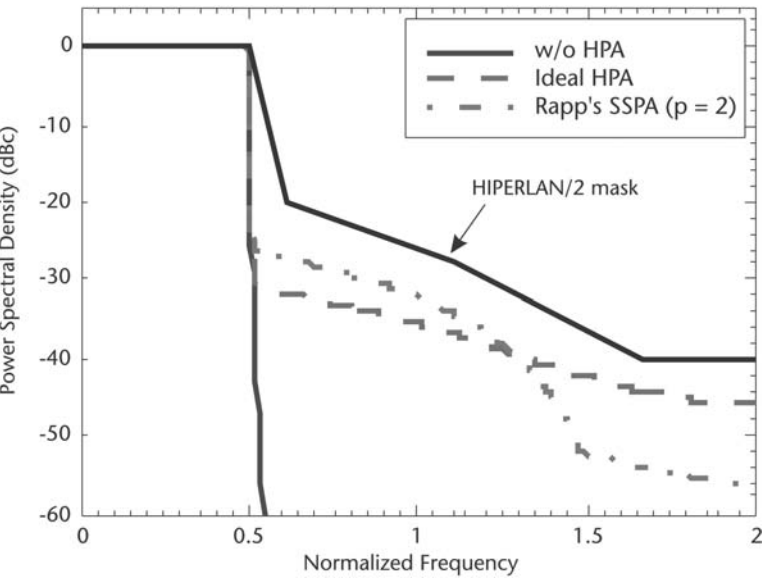


Figure 2.22 Spectrum of OFDM for an ideal HPA compared to an SSPA, with IBO = 6 dB.

Whereas the more realistic SSPA model has higher out-of-band radiation at frequencies below 150% excess bandwidth, at higher frequencies the SSPA model has an even a lower PSD. Both HPA models meet the HIPERLAN/2 power mask at an IBO of 6 dB. However, in the context of multiband transmission, the out-of-band radiation caused by the nonlinear HPA may not be acceptable. Compared to the nondistorted OFDM signal (without HPA), the out-of-band radiation is massively increased due to clipping of high signal peaks. The choice of the HPA model appears to have less impact compared to the selected IBO. However, provided powerful peak reduction techniques the situation may change; and in order to approach the PSD of the nondistorted OFDM signal, both peak reduction and predistortion may be necessary.

2.1.5.4 Use of MIMO Technologies

In packet transmission, the maximum transmission data rate typically corresponds to the peak throughput per cell. It is, therefore, important to investigate the maximum achievable transmission data rate because raising the maximum transmission data rate for one user increases the number of users who can be accommodated by high-speed transmission. In a completely isolated cell environment, in which interference from neighboring cells is nonexistent, the maximum transmission data rate can be raised without limit by increasing the number of transmit/receive antennas and the transmission power.

The underlying mathematical nature of MIMO, where data is transmitted over a matrix rather than a vector channel, creates new and enormous capacity increases. The first results hinting at the capacity gains of MIMO were published in [41]. In [42], the author shows how one can, under certain conditions, transmit multiple independent data streams simultaneously over the eigenmodes of a matrix channel created by transmitting and receiving antennas. Essentially, it was shown in [43, 44] that under certain conditions the capacity increases linearly with $\min(M_T, M_R)$, where M_T and M_R are the number of transmitting antennas and receiving antennas, respectively.

The increase of capacity provides the user with high-data-rate transmission. A good example is the V-BLAST scheme [42, 45], which is more generically called *spatial multiplexing*. This kind of spatial transmission can provide the highest number of data streams at the expense of diversity gains. The signals from different transmitting antennas, however, will be mixed up. To perform symbol detection, the receiver must unmix the channel in one of several various possible ways. The complexity of the optimum maximum likelihood (ML) method is high, even prohibitive when the number of antennas is large or high-order modulations are used. Enhanced variants of this include sphere decoding (SD) [46], lattice reduction (LR) [47, 48], and well-known nulling and canceling [45], which give reasonable trade-offs between complexity and performance. A QRD-M detection algorithm was proposed in [49, 50] to apply to MIMO-OFDM systems. This algorithm demonstrates performance improvements and a flexible configuration.

Another powerful effect of multiple antennas lies in the concept of spatial diversity. In the presence of random fading caused by multipath propagation, the probability of losing the signal vanishes exponentially with the number of

decorrelated antenna elements being used. The typical schemes are space-time coding (STC) techniques in which user data is mapped according to specific principles within the space dimension and time and/or frequency dimension.

MIMO is an indispensable part of the broadband air interface [12]. Therefore, the IST project WINNER [6, 51] investigated several different multiple-antenna techniques, which show good performance, manageable complexity, and flexible application scenarios. Matrix modulation, parallel concatenated transmit diversity, QRD-M detection for MIMO-OFDM systems, and advanced detection methods for MIMO systems can be beneficial for MIMO-OFDM systems.

Linear nonorthogonal matrix modulation can provide transmission of multiple data streams simultaneously with still considerable diversity gain. With this property, it can lead to good link reliability with a high data rate. In principle, it is an open-loop scheme, which allows for application of scenarios such as wide area where full channel information is not available. Thus, the load of signaling is rather limited. It is universal for single-carrier or multiple-carrier systems.

Parallel concatenated transmit diversity systematically approaches the design of coding and modulation schemes for MIMO channels. It can embed diversity into redundancy introduced by the encoder via novel use of constellation dimensions to optimize the encoding process.

The QRD-M detection algorithm for a MIMO-OFDM system combines the QR decomposition of the channel matrix and M-algorithm. It is a suboptimal method, which reduces the complexity of ML detection. Compared with other interference cancelers, the QRD-M algorithm is more attractive because it provides flexible performances with a different value of M . Therefore, replacing the interference canceler with the M-algorithm should greatly enhance performance. With adaptive QRD-M, the performance can be further improved with reasonable computational complexity.

To avoid the exponential complexity of the MLD problem for MIMO systems, sphere detection was proposed [51] to restrict the search for the closest lattice point to include only vector constellation points that fall within a certain search sphere. This approach allows for finding the ML solution with only polynomial complexity, for sufficiently high signal-to-noise ratios (SNR). Lattice reduction is used to find new and better basis vectors that define the same lattice. This approach is especially attractive for environments with a low number of channel variations in time and frequency.

Given the limitations in simulator capabilities, spectral efficiency values cannot be obtained with high accuracy. In the following discussion, one possibility to obtain spectral efficiency is presented [51] under the constraint of full queue simulations consisting of pure snapshots of uncorrelated drops of long-term parameters. In particular, the definition of the satisfied user criterion becomes difficult in this case. Reference [51] provides a pragmatic approach for obtaining spectral efficiency values based on the cumulative distribution function (CDF) of the user throughput obtained from simulations including short-term evolution.

2.1.5.4.1 Basic Spectral Efficiency Calculations.

The spectral efficiency is defined similar to [51] and 3GPPP. A single service scenario is assumed. Therefore, the average normalized throughput can be expressed as

$$\nu_{\text{site}} = \frac{\sum_{k=1}^{N_c} \bar{T}_c(k)}{N_s B} \quad (2.14)$$

where the average cell throughput for all cells k is summed up and divided by the number of sites N_s and the bandwidth B . The average normalized throughput is equivalent to the term *system load* in [52].

Note that in the case of sectorization, the number of cells N_c is a multiple of N_s . The site is defined as physical colocation of hardware serving the set of antennas. Users may be connected to a site either directly or through relay stations. Both types should be taken into account when calculating the spectral efficiency of a site. The site spectral efficiency $\nu_{\text{site}, 98\%}$ is the average normalized throughput for which 98% of the users are satisfied according to the satisfied user criterion below. The area spectral efficiency $\nu_{\text{area}, 98\%}$ is obtained by multiplying the site spectral efficiency by the number of sites N_s per covered area A :

$$\nu_{\text{area}, 98\%} = \nu_{\text{site}, 98\%} \frac{N_s}{A} \quad (2.15)$$

For single-cell or single-site simulations in an area, the spectral efficiency cannot be predicted with sufficient accuracy due to the simplified intercell interference modeling. In this case only site spectral efficiency is used. The spectrum efficiency is given separately for the uplink and downlink.

2.1.5.4.2 Satisfied User Criterion.

The satisfied user criterion is normally based on a threshold of the active session throughput (or packet call throughput); that is, it is based on an average user throughput. In snapshot simulations each user exists only for one time instant and therefore no averaging over time can be performed. Any TDMA component of scheduling will not be considered and any user who is not scheduled instantaneously is necessarily a dissatisfied user. Therefore, a straightforward application of the common satisfied user criterion can be used as in [52, 53]. To overcome this shortcoming of snapshot simulations, the average user throughput with respect to the distance should be used. Assuming a homogeneous user density, 98% of the users will be located within a sector having the radius of $\sqrt{0.98}$ times the nominal cell radius. Assuming a monotonically decreasing function of average user throughput versus normalized distance, we can define a satisfied user criterion as follows. Assume that 98% of users are satisfied if the average user throughput versus the normalized distance is above 500 Kbps for all normalized distances less or equal than $\sqrt{0.98}$. Note, that the average user throughput of 500 Kbps, corresponding to a minimum sustained bit rate as in [52], is also used as the threshold for Internet or FTP traffic models.

2.1.5.4.3 Spectral Efficiency Calculation for Simulations Using Short-Term Evolution.

As a pragmatic approach and, in particular, as a basis for relative comparisons of different spatial processing techniques or variants, spectral efficiency can be based on the CDF of the average user throughput for simulations including short-term

evolution [51]. Here the site spectral efficiency is obtained based on the load for which 90% of users have an average throughput greater than or equal to 500 Kbps. The corresponding average cell throughput is normalized by the bandwidth and further to the number of cells per site or to the area, depending on whether site spectral efficiency or area spectral efficiency is presented.

2.1.5.4.4 Achievable Spectral Efficiency.

In Europe, the FP6 IST project WINNER chose a MIMO-OFDM-based system and set the target value for peak spectral efficiency in an isolated cell environment as 25 bit/s Hz^{-1} as a system requirement for next generation mobile communications [2]. In Japan, NTT DoCoMo proposed a radio transmission method based on variable spreading factor (VSF)-Spread OFDM as one of the candidates for increasing system capacity in the downlink cellular environment and in hot-spot and indoor office environments for next generation mobile communication systems [19]. In field experiments using VSF-Spread OFDM radio access with a channel bandwidth of 100 MHz, ultra-high-speed packet transmission with a throughput of 1 Gbps (spectral efficiency of 10 bit/s Hz^{-1}) was achieved [20]. In those experiments, MIMO consisting of four transmitting/receiving antennas at distances of up to 300m from the BS under mostly non-LOS conditions was used.

In a multicell environment, however, increasing the number of antennas and transmission power also increases the amount of interference from neighboring cells. The maximum achievable transmission data rate is therefore determined by the received signal-to-interference-plus-noise-power ratio (SINR), which takes intercell interference into account. In [19], a multicell environment with 19 cells (3 sectors per cell) with BS transmission power of 20W and a distance between base stations of 500m with the volume of traffic in neighboring cells (channel load) as a parameter was investigated. The cumulative distribution function (cdf) of the received SINR showed that the distribution of received SINR improves as channel load becomes smaller.

Results given in [19] show that the received SINR at nearly all locations within a cell is no more than approximately 30 dB even in the case of a small channel load. We can therefore consider the maximum received SINR in a cellular environment to be approximately 30 dB, assuming a low-channel-load case. By performing computer simulations under the condition of a received SINR of 30 dB, it was found in [20] that spectral efficiency of approximately 50 bit/s Hz^{-1} can be achieved.

2.2 Radio Interfaces Optimized for PANs

Short-range communications, and in particular, personal area networks would require specific air interfaces capable of handling the high data rates demanded by the PAN services and applications [3]. Conceptual personal network (PN) and PAN usage scenarios were described in the FP5 IST project PACWOMAN [41, 54, 55]. Those concepts were later investigated further in the IST projects MAGNET and MAGNET Beyond, which also proposed highly adaptive and scalable PAN air interfaces as building blocks for 4G communication systems [3, 56, 57]. Several PHY layer schemes, broadly categorized into multiple-carrier

techniques, UWB techniques, and hybrid solutions were investigated, in relation to a matching device capability as well as user and device requirements and QoS issues.

Although short-range wireless applications are demanding ever higher data rates, an increase has also been observed in the number of low-rate devices, such as transceivers embedded in a variety of common appliance, exchanging different information types, including their geographic or local positions. The prospect of a shortage in worldwide available spectrum resources and the growing number of active wireless devices, combined with their potential to further aggravate interference and electromagnetic exposure issues, favor the introduction of license-free systems based on ultrawide spectrum overlay concepts. This idea was further elaborated within the scope of the IST project PULSERS, which investigated the possibilities offered by UWB radio technology to enable the reuse of existing spectrum bands [58]. By considering the trade-offs between reduced transmitted power and increased spectrum efficiency when applying short-range technologies supporting dynamic ad hoc networking concepts, spectral reuse can be increased and thus higher capacity per area can be achieved [59]. For example, frequency reuse leads naturally to increased spectral efficiencies when measured in terms of spatial capacity (ratio of a cell's aggregate data rate and the coverage area in bits per second per square meter). Further, besides the ability to potentially operate across bands occupied by existing narrowband systems, UWB radio systems offer an additional flexibility in that they can maintain a cell's spatial capacity by adapting to either a large number of low-rate nodes or a smaller number of high-rate nodes depending on the application.

The frequency band between 3.1 and 10.6 GHz was identified for communication and measurement systems with a tight spectral mask by the Federal Communications System (FCC). Under the FCC UWB regulatory regime, for example, the amount of spectral power that can be transmitted is limited by the mask. This is shown in Figure 2.23 [60].

The European Commission has finally issued details of the licensing regulations for UWB networking in Europe [61]. UWB data transmission is set to be used for cable replacement in consumer goods enabling large files to be transmitted over short distances—typically within a room—at 480 Mbps and above. Such a decision is a first step toward overcoming a potential spectrum “bottleneck.”

The EC has chosen to make use of only part of the spectrum that was approved for use in the United States by the FCC in 2002 [62]. The EIRP value of -41.3 dBm/MHz is described as the *magic figure*. In Europe this will be applied over the 6.0- to 8.5-GHz frequency range, whereas in the United States the FCC applies this over a broader spectrum from a much lower frequency. It is also to be applied provisionally until the end of 2010 in the 4.2- to 4.8-GHz range [61].

IEEE 802.15.3a intends to standardize a high data rate WPAN, up to 10m initially in the 3.1- to 10.6-GHz band, which is one of the applications of UWB [63]. Because wide-area communication systems (e.g., IMT-A candidates) can also be allocated spectrum in this band, coexistence issues need to be carefully taken into account to guarantee the specified QoS for all the existing radio services sharing the same spectrum. With a bandwidth of at least 500 MHz, the overlapping of UWB signals with other RF systems is evident. Coexistence issues have been

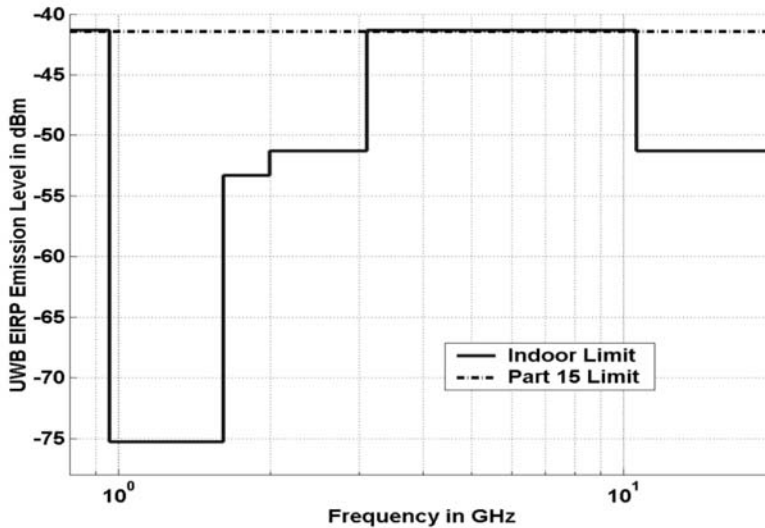


Figure 2.23 FCC mask for spectral power transmission.

studied by various institutions and industrial players and reports can be found from public sources. Measurements have been done using high-power UWB transmitters that exceed the radiation limits given by the FCC. Using these interference sources, the cumulative impact of a high density of UWB devices can be demonstrated. From these studies, UWB seems to be a promising technology that could coexist with the present WLAN and RF systems and standards.

In the case of coexistence with IMT-A system candidates, the magnitude of the impact might be orders of magnitudes higher for a case in which the UWB devices emit the higher power spectral density in this band that is allowed under the FCC regime (see Figure 2.23) [11].

2.2.1 Scenarios and Radio Propagation Models for PANs

New, reliable air interfaces and adaptive MAC schemes for PANs cannot be successfully implemented without a throughout analysis of the specific radio propagation channels.

PAN end-user scenarios considered in the research community of the FP6 projects MAGNET and MAGNET Beyond spanned from very short-range indoor or body-worn networks (1 to 2m) to access-point like scenarios (<10m) in typical environments (i.e., office, home, hospitals, shopping malls). In typical scenarios (e.g., for small body-worn and handheld devices), the PAN radio propagation channel exhibited some interesting characteristics such as highly scattered paths and antenna near-field effects, especially due to the human body proximity conditions. To support the PN concept, the PAN air interface must cover a wide range of data rates and system capabilities and requirements optimized for low-power consumption and cost effectiveness. Therefore, one characteristic of the PAN scenarios is that they are dynamic. Adaptivity and scalability are achieved by optimizing

the air interfaces for the PAN-specific scenarios. The identified PAN-specific scenarios are described in Table 2.4.

The PAN defined in MAGNET Beyond consists of both personal devices carried by a person and stationary devices, which are sharable with other PANs (see Table 2.4). Differently from cellular and wireless LAN systems, peer-to-peer communications (especially from data traffic point of view) are key in this communication environment. In addition, power-efficient and spectrally efficient technologies are the key requirements to enable the wide spread of personal communication systems.

Figure 2.24 shows the relationship between the defined scenarios, the possible services that can be delivered by a PAN air interface, and the candidate transmission technology. Only one technology would not be sufficient to accommodate LDR and HDR PAN usages. This objective is not realistic in terms of the very particular requirements regarding costs and power consumption of LDR communications compared to HDR communications. The HDR air interface that does not have

Table 2.4 Classification of PAN Scenarios

<i>Very Low Data Rate (VLDR)</i>	<i>Low Data Rate (LDR)</i>	<i>Medium Data Rate (MDR)</i>	<i>High Data Rate (HDR)</i>	<i>Very High Data Rate (VHDR)</i>
Energy efficiency; small portable devices; disposable sensors	Energy efficiency; small portable devices and sensors; body-worn devices	Trade-off area balance between energy efficiency and data rate, QoS	Data-rate- and QoS-driven medium-size portable devices	Data-rate-driven high-speed, point-to-point connection
<100 Kbps	<250 Kbps	<11 Mbps	<150 Mbps	>150 Mbps

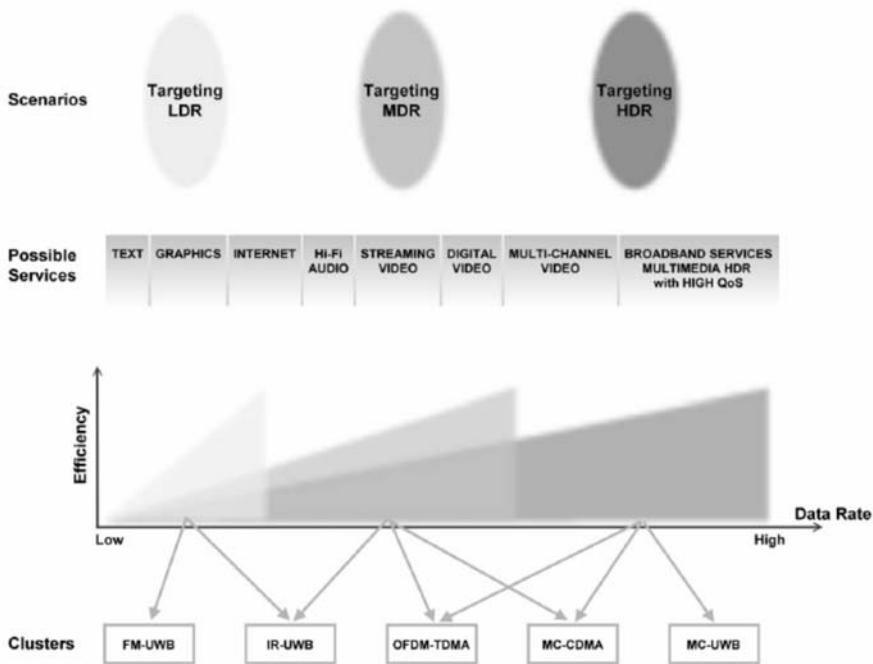


Figure 2.24 PAN air interface technology candidates, scenarios, and possible services.

such constraints on cost and power consumption can be fully scalable and flexible, although it is challenging enough to accommodate operations that require MDR up to HDR communications.

Additionally, seeing the purpose of the future wearable and mobile HDR PAN devices, these latest versions will have to present an acceptable cost, complexity, and power consumption. Therefore, the following technologies have been targeted for the PAN air interfaces [64]:

- HDR: MC-CDMA, OFDM-TDMA for the wideband systems, and MC-UWB for the UWB systems;
- LDR: FM-UWB and IR-UWB, which are both UWB systems.

The HDR approach uses an MC-SS PHY layer along with an IEEE 802.15.3-based MAC layer [65]. The LDR solution is called frequency modulation (FM)-UWB with a bandwidth of at least 500 MHz and transmits data by means of an FM-UWB signal. The utilized MAC scheme for LDR transmissions is based on IEEE 802.15.4. To demonstrate the new air interfaces, prototypes were also developed.

UWB impulse radio (IR-UWB) provides a huge potential for simultaneous data transmission because the information is spread over time, which is due to the duty cycling at the symbol level. UWB radio is a wireless technology for transmitting large amounts of digital data over a wide spectrum of frequency bands with very low power. According to the FCC definition, UWB uses an ultrawide fractional bandwidth of at least 20% or, alternatively, a bandwidth of 500 MHz or more. But the main characteristic is that UWB transmits ultra-low-power spectral density so that the signal in a narrow band is below the noise floor.

The areas of application of UWB are systems with VHDR requirements in the short range, and precise localization and tracking features that exploit the inherent properties of very wideband signaling systems. For example, UWB can rate up to 100 Mbps in 10m with a localization accuracy of less than 1m. The possible uses are as follows [28]:

- Hot-spot WPAN;
- Intelligent wireless area network (IWAN), for example, an in-house network, such as wireless high-definition television (HDTV);
- HDR in peer-to-peer networking (e.g., wireless HDTV);
- Sensor positioning and identification network (SPIN);
- Military applications (because of the low probability of intercept and detection);
- Radar, for examples, obstacle avoidance in unmanned vehicles.

UWB technology can use many different types of modulations, which can also be combined: pulse position modulation, pulse phase modulation, pulse amplitude modulation, frequency sweep, direct sequence, and OFDM. The typical pulse used for UWB is known as a Gaussian doublet. This is sent in a train of pulses where the start of each pulse is pseudorandomly modified for avoiding the frequency

peak at the repetition period. Therefore, the error probability due to collisions is decreased significantly [66]. This also shows that IR-UWB is well suited for medical body area network (BAN) applications.

A BAN consists of a number of nodes and units located on the human body or in proximity such as on everyday clothing. Because low-power transmission is required for body-worn wireless devices, the human body influences the communications channel between wireless wearable devices to form a wireless network.

UWB low transmitted power requirements allow longer battery life for body-worn units. UWB sensors are characterized by low complexity and very low power consumption, compliant with the mandatory UWB physical layer of IEEE 802.15.4a, and therefore are capable of interoperating with other IEEE communication standards.

OFDM-TDMA is usually employed with time-division duplexing. The most prominent OFDM-TDMA/TDD scheme was defined in HIPERLAN/2 [41], which differed from the carrier sensing multiple access with collision avoidance (CSMA/CA) of IEEE 802.11a (e.g., by using fixed-length packets). While the HIPERLAN/2 scheme [13, 67] was not flexible enough in the PAN context with respect to supported data rates or bandwidth efficiency, its general simple structure facilitated the addition of advanced signaling techniques in an almost modular fashion. These techniques included, in particular, the use of link level adaptation in case of partial or full CSI at the transmitter and the use of bit-interleaved coded modulation (BICM) as well as the cross-layer optimization of the corresponding parameters in adaptive MAC schemes. Strategies have been developed in support of OFDM-TDMA/TDD for 4G communications that allow a maximization of the spectral efficiency of this technology with 100% frequency reuse [68].

The HDR system proposed in [69–71] works at the 5.2 GHz unlicensed frequency band with a 40- or 20-MHz bandwidth. The corresponding antenna system design is also directly related to these configurations. PAN (e.g., MAGNET) devices should coexist with other legacy and newly emerging systems. Interference mitigation, therefore, is also of importance for the successful deployment of the developed air interfaces. Interference mitigation is further discussed in Chapter 7 of this book.

Outside the IST projects MAGNET and MAGNET Beyond, other air interfaces for PANs have been developed. The IST project PULSERS contributed to IEEE 802.15.4a channel models. Major candidates in this context are WiMedia, Bluetooth, ZigBee, and WiBree. WiMedia is an HDR air interface with data rates up to 480 Mbps targeting short-range applications in the personal operating space. ZigBee is an LDR solution mainly addressing sensor applications with ranges extending the personal operating space. WiBree is an extension to Bluetooth allowing for LDR communications in a 10m range.

2.2.1.1 General Characteristics and Requirements of an LDR WPAN Air Interface

The LDR system that was proposed within the IST project MAGNET Beyond implemented the FM-UWB technique at the PHY layer. FM-UWB is a scalable air interface technology aimed at short-range (<10m) LDR (up to 100 Kbps) applications characterized by a low power consumption and ease of implementation on an integrated circuit [69]. It constitutes a low-power solution for the PAN. The

FM-UWB system was designed to have a 1-GHz bandwidth centered at 4 GHz. In MAGNET Beyond, a new configuration with a 500-MHz bandwidth centered at 4.5 GHz was proposed as an alternative. For both configurations, an FM-UWB transmitter would not produce any energy below 3.1 GHz and above 5.0 GHz. FM-UWB uses double FM: a low modulation index digital FSK followed by a high modulation index analogue FM to create a constant envelope UWB signal. The system is intended for low-bit-rate, short-range WPAN systems. Table 2.5 summarizes the LDR radio characteristics.

RF center frequency, bandwidth, and TX power are dictated by the UWB regulations. Switching times and latency are switching times of the hardware components and the group delay occurring in the various filters in the transmitter and receiver. The current consumption values are estimations for an optimized IC version in appropriate IC technology. The expected power consumption of this first generation is twice that value. The effective data rate equals half the raw data rate due to the use of Manchester encoding.

Ideally power consumption should be determined by the auto discharge time of the battery rather than the radio power consumption. A typical lithium 50-mA/H capacity button cell has an auto discharge time equal to 10 years.

At a data rate of 100 Kbps and a message of 400 bits sent every minute, this means 2,102 seconds of transmission time per year. With a sensor plus transmit power consumption equal to 5 mA, this means 10,000 mA/s = 2.8 mA/h per year, which is lower than the 5-mAh auto discharge.

The LDR-related user scenarios are highlighted by the concept of the PAN where body-centric transmission is the dominant form of communications. Most devices in this context are body-worn, belt-mounted, or handheld devices. Therefore, the radio environments of such systems are characterized by the user proximity, user dynamics, and realistic terminal/device handling and antenna configurations. The traditional channel models underestimate these channel effects, thus the corresponding PHY layer performances presented here are only upper bounds. Novel, dynamic, UWB indoor radio propagation investigations were carried out for the IST project MAGNET [69, 70]. Four user scenarios were investigated:

- PAN-FD (PAN with fixed device);
- PAN-PD (PAN with portable device);

Table 2.5 FM-UWB Radio Characteristics

<i>RF Center Frequency</i>	<i>LB HB</i>	<i>4.5 GHz 6.4–8.7 GHz</i>
RF bandwidth		500 MHz
RF output power		–14 dBm
Subcarrier frequency		1–2 MHz
Subcarrier modulation		FSK, $\beta = 1$
Raw bit rate		≤125 Kbps
Receiver sensitivity		–80 dBm
TX/RX switching time		≤10 μ s
Latency (at PHY level)		<1 ms @ 100 Kbps
RX synchronization time		<50 bits
Current consumption RX		7 mA
Current consumption TX		4 mA

- PAN-MD (PAN with mobile device);
- BAN.

All measurements were performed in the 3- to 6-GHz band, which included the working frequency band of the FM-UWB system. Characteristics of indoor UWB, body-worn, and mobile-to-mobile radio channels were disclosed.

In all PAN scenarios (with a fixed device, portable device, or mobile device) the wideband power was found to be log-normal distributed, with standard deviation values varying depending on the actual environment. The power levels of the main signal clusters were independent log-normal distributed with typically higher standard deviation in the PAN-MD scenarios compared to the PAN-FD/PD scenarios.

In terms of time-domain channel characteristics, wideband power (signal shadowing) and signal clustering stationarity time values on the order of 0.5 and 0.25 second were determined in the PAN-FD and PAN-PD/MD scenarios, respectively, with walking user speeds between 0.5 m/ps and 1 m/ps. This was one of the novel characteristics of the proposed UWB PAN channel model.

The PAN scenario investigations also showed different time-delay signal clustering depending on the environment and user scenarios. Typically, in PAN-FD scenarios a dual exponential average power delay profile has been observed with the first cluster region extending up to 50-ns excess delay and with decreasing cluster power decay factors. In the PAN-MD scenarios, due to the dual dynamics of the channel (mobile-to-mobile communication scenarios), the average power delay profile exhibited dual exponential power decay with typically low decay in the first delay region up to 25-ns excess delay. Cluster and ray arrival rates were also found to be different for the different PAN scenarios with fixed device, portable device, or mobile device configurations. These time-delay domain signal clustering characteristics are reflected in the average root mean square (rms) delay spread and 90% energy delay window values determined.

The BAN investigation results showed a large spread of the wideband power levels that were, on average, 10 dB higher than the corresponding free-space power level for average “around the body” distances and a 4.5-GHz carrier. Furthermore, significant deviations/differences could be observed between the two considered environments, in both walking user and seated user scenarios. It was concluded that these phenomena were due mainly to the directional antennas used at the BW devices and the dense radio scattering environment in the near range of the user.

Wideband power (signal shadowing) and signal clustering stationarity time values on the order of 0.25 second were determined in the BAN scenarios with user moving speeds between 0.1 and 0.5 m/s. Typically, in all BAN scenarios a dual exponential average power delay profile was observed with very low power decay in the first cluster region extending up to 25-ns excess delay.

The main UWB channel parameters for the investigated PAN/BAN user-proximity scenarios are listed in Table 2.6.

2.2.1.2 General Characteristics and Requirements of an HDR WPAN Air Interface

The MC-SS technology can support the HDR WPAN air interface [3] because it is capable of transmitting data from a low to a high data rate for PAN environments

Table 2.6 Channel Parameters for the Investigated UWB PAN/BAN User-Proximity Scenarios

<i>User-Proximity Scenario Parameters</i>	<i>PAN-FD</i>	<i>PAN-PD/MD</i>	<i>BAN*</i>
Log-normal wideband power shadowing std dev [dB]	3.0–4.5	2.1–3.6	1.5–3.0
Log-normal cluster fading std dev [dB]	4.0–5.6	2.2–6.9	4.3–5.4
Cluster power decay regions	<50 and >50 ns	<25 and >25 ns	<25 and >25 ns
Cluster power decay factor [dB/ns]	0.23 and 0.17	0.13 and 0.20	0.08 and 0.13
Average rms delay spread [ns]	13–35	11–25	12–18
Average 90% energy delay window [ns]	25–88	20–57	24–49
Wideband power and signal clustering channel stationarity time	0.5	0.25	0.25

*In these BAN scenarios, four body-worn devices with a directional antenna and one belt-mounted handset with omnidirectional antennas were used.

where multiple devices will set up multiple peer-to-peer communication links at the same time. The range of data rate envisaged is from tens to hundreds of megabits per second.

Due to propagation characteristics and spectrum availability, the 5-GHz band is appropriate for the MC-SS MAC interface. Bandwidths of 20 and 40 MHz were considered in the MAGNET Beyond air interface and a homogeneous system design approach was applied to fulfill the objective of flexibility.

2.2.2 Increased Capacity and Throughput for a MIMO UWB System

Studies and real channel measurements report a linear increase in capacity of narrowband and broadband MIMO systems also for UWB environments [58, 72, 73]. At 20 dB the achievable spectral efficiency of a simulated MIMO UWB channel for outage probabilities of less than 10% is considerably higher than that of single-input/single-output (SISO) configurations. The calculated ergodic MIMO UWB channel capacity is based on a MIMO UWB channel model including spatial correlation. Table 2.7 summarizes the mean capacities for the different antenna configurations.

The maximum achievable throughput of the system takes into account the increased overhead of a MIMO VHDR multiband OFDM (MBO) transmission due to an increased total preamble size. If a NO-ACK scheme and collision-free channel access according to the Distributed Reservation Protocol (DRP) are

Table 2.7 Spectral Efficiency of MIMO UWB

<i>Antenna Configuration</i>	<i>Capacity (bit/s Hz⁻¹): LOS at SNR = 20 dB</i>	<i>Reference results: LOS at SNR = 30 dB</i>
1 × 1	7.1	8.8
1 × 2	8.5	NA
2 × 2	12.9	16.5
2 × 3	14.5	NA
3 × 3	18.2	23.9

assumed, the system parameters and values for the theoretical maximum throughput can be summarized as shown in Table 2.8.

It can be concluded that for all MIMO transmission modes (data rates), the achievable throughput is maximized for the largest payload size due to the reduced overhead in the system. Moreover, the relative throughput gain compared to SISO decreases slightly with the increasing modulation order and for weaker channel coding.

The values for these calculations refer to the time-orthogonal MIMO preamble design, which is suboptimal in terms of overhead. Better orthogonal preamble sequences may thus be needed for more efficient MIMO transmission.

2.3 Accurate Channel Modeling for Adaptive and Scalable Air Interfaces

For the evaluation of the key questions related to spectrum-efficient air interfaces, a set of reliable and measurement-based channel models is needed. Channel models have to be accurate due to the fact that radio propagation has a significant impact on the performance of future broadband systems. This is especially true with future MIMO radio communication systems because more of the radio channel degrees of freedom in space, time, frequency, and polarization may be exploited to meet the demands on bit rate, spectrum efficiency, and cost. Channel models are needed in performance evaluation of wireless systems, when choosing modulation and coding schemes, in multiantenna system design, in the selection of the channel estimation method, for channel equalization and other baseband algorithm design, and for network planning. It is important to use common and uniform channel models for evaluation, comparison, and selection of technologies. In this context it is clear that realistic and reliable multidimensional channel models are an important part of performance evaluation of next generation communication systems (e.g., IMT-A). Table 2.9 shows the requirements of different propagation scenarios envisioned for next generation communication systems.

Technical selection criteria for accurate channel models that should be considered during the design of new air interfaces are the applicability of the model to the envisioned scenarios, the frequency range, bandwidth, supported antenna configurations, mobility, path loss, short-term and long-term fading, correlation characteristics, and possibility of modifying the antenna pattern. Other selection criteria are simplicity, model availability, and available software implementations.

From a modeling point of view, the mobility speed determines the sampling of the channel. The model can, in principle, handle any number of samples. It is just a calculation issue. The algorithm testing or receiver design engineer can choose the required speed for performance testing, which determines the exact sampling of the channel. Note that the transitions between different scenarios should also be supported by the channel models.

In general, radio channel research has been carried out within the IST FP6 projects to understand the complex MIMO radio channel characteristics. Radio channel research can be divided into three research areas: multidimensional channel measurement and estimation methods, channel modeling, and channel simulation.

Table 2.8 Maximum Throughput of MIMO VHDR MBO for a Packet Size of 4 KB [73]

Sim Mode	A		B		C		D		E		F	
Data Rate per Antenna (Mbps)	640	640	800	800	960	960	1280	1280	1600	1600	1920	1920
Nt	1	2	1	2	1	2	1	2	1	2	1	2
Mod			QPSK						16-QAM			
Bits per tone			2						4			
Payload size (B)							4095					
Tsvm (ns)							312.5					
Ninfo per symbol	200	400	250	500	300	600	400	800	500	1000	600	1200
Rc		1/2		5/8		3/4		1/2		5/8		0.75
Ncoded per symbol	400	800	400	800	400	800	800	1600	800	1600	800	1600
Payload duration (μ s)	51.2	51.2	40.96	40.96	34.13	34.13	25.6	25.6	20.48	20.48	17.07	17.07
OFDM symbols per frame	163	163	131	131	109	109	81	81	65	65	54	54
Frame duration (μ s)	60.57	69.95	50.33	59.71	43.51	52.88	34.97	44.35	29.85	39.23	26.44	35.82
Throughput	523	908	628	1064	721	1195	890	1402	1025	1582	1145	1720
Throughput gain (%)	—	73.7	—	69.5	—	65.8	—	59.4	—	54.4	—	50.3

Table 2.9 Propagation Scenarios

<i>Scenario</i>		<i>Mobility</i>	<i>Frequency</i>
Wide-Area	Rural	200 km/h	2 and 5 GHz
	Suburban	100 km/h	2 and 5 GHz
	Urban	50 km/h	2 and 5 GHz
	Outdoor to indoor	5 km/h	2 and 5 GHz
Short Range	Urban	[5] km/h [50]	2 and 5 GHz
	Indoor	5 km/h	2 and 5 GHz
Multiple link		To be determined	2 and 5 GHz

2.3.1 Existing MIMO Channel Models

This section introduces widely known radio channel models. Existing channel models can be studied and exploited for early assumptions about the channel design for next generation systems. Because MIMO techniques have been very important in the research work for systems beyond IMT-2000, the models that are introduced here as state of the art are the IEEE 802.11n, METRA, and 3GPP SCM. METRA-based IEEE 802.11n models are suitable for short-range scenarios, whereas 3GPP SCM models are able to serve for wide-area scenarios. Unfortunately, both models have essential limitations and drawbacks. More advanced models need to be developed for next generation systems.

2.3.1.1 IEEE 802.11n Channel Model

Channel models of IEEE 802.11n were designed to serve indoor environments for a MIMO WLAN application. The modeled environments are from small offices and residential homes to large offices and open spaces with LOS and non-LOS (NLOS) conditions. The rms delay spreads are from 15 to 150 ns. The propagation channel is directional by nature and dispersed in time, frequency, and space.

A clustered structure is assumed for the propagation environment. In these models this means that multipath components with a variety of delays and incidence angles are grouped into clusters where each cluster has a number of multipath components with different delays and powers but the same angular characteristics. The number of clusters and multipath components varies from model to model, but the maximum number of distinct delay components (channel taps) is 18. WLAN applications operate at 2- and 5-GHz bands. Experimental results from these two frequency bands are averaged to compose the models, thus a single set of models should be applicable on both bands. Channel taps are separated in delay at a minimum of 10 ns, so the bandwidth of the model is 100 MHz. The principle in the generation of channel realizations out of IEEE 802.11n channel models is stochastic. A reference implementation of the model exists for MATLAB [75].

Although the model parameters are geometric, such as angles and array orientations, the creation of channel impulse response coefficients is based on spatial correlation matrices. Temporal correlation and spatial correlation are carried out by separate and independent filtering operations. Spatial correlation between two elements in an antenna array is analytical expression and a function of the power

azimuth spectrum (PAS) and antenna geometry. Transmitter array and receiver array correlation matrices are combined with the MIMO channel correlation matrix by the Kronecker product. This approach assumes that the transmitter and receiver PASs of each channel tap are separable, which is not valid in general [76]. Actually, the model does not state the generation of the coefficients explicitly (e.g., correlation matrix based), but it is included in the model implicitly by giving, for example, the Doppler power spectrum shapes.

The actual model includes tables of parameters for seven indoor environments. Two different kinds of Doppler power spectrums and descriptions of some special features are given. The parameter tables contain cluster structures and excess delay, power, angle of arrival (AoA), angle of departure (AoD), and angular spread (AS) of departure and incidence angles for each multipath component. Features of the models are path loss, shadow fading, deterministic LOS component, and Doppler components due to fluorescent lights. Elevation AS is excluded from the model, thus antenna array orientation in a horizontal plane only applies. Cross-polarization discrimination (XPD) for LOS and NLOS conditions is defined, but not an antenna correlation model for cross-polarized antennas. Thus, the polarization model is not fully described and applicable. The following model types are described in order to list the model mapping to a propagation environment:

- Model A (optional, should not be used for system performance comparisons), flat fading model with 0-ns rms delay spread (one tap at 0-ns delay model); this model can be used for stressing system performance; occurs small percentage of time (locations);
- Model B with 15-ns rms delay spread;
- Model C with 30-ns rms delay spread;
- Model D with 50-ns rms delay spread;
- Model E with 100-ns rms delay spread;
- Model F with 150-ns rms delay spread.

The model mapping to a particular environment is shown in Table 2.10.

The characteristics of the IEEE 802.11n models described here are taken from several measurement results referred to in [77]. The PASs have a truncated Laplacian shape. Means to use uniform and truncated Gaussian PASs are also given in [78].

Table 2.10 Mapping of Channel Models to a Propagation Environment

<i>Environment</i>	<i>Condition</i>	<i>Model</i>
Residential	LOS	B – LOS
	NLOS	B – NLOS
Residential/small office	LOS	B – LOS
	NLOS	C – NLOS
Typical office	LOS	C – LOS
	NLOS	D – NLOS
Large office	LOS	D – LOS
	NLOS	E – NLOS
Large space (indoors and outdoors)	LOS	E – LOS
	NLOS	F – NLOS

Doppler spectra are caused by the so-called environmental scatterer speed. There are two Doppler spectrum shapes: One is called the *bell shape* and has a very small Doppler frequency; the other is called *bell with spike shape*, which in addition to the very small Doppler frequency also has a nonsymmetric peak frequency caused by a bypassing vehicle. The path-loss function has two slopes and displays the so-called break-point distance to separate the slopes. Shadow fading is drawn randomly from log-normal distribution and is fixed for the single use of the channel model.

One special effect in the model are the Doppler components due to fluorescent lights. This takes effect by modulating several channel taps in order to artificially introduce an amplitude modulation (AM) in the received signal. The interferer-to-carrier energy (I/C) ratio is drawn randomly to scale the modulation. The I/C ratio is constant for the single use of the channel model.

The calculation of the channel coefficients is computationally effective. For example, the correlation matrices for the fixed-antenna array geometry need to be calculated only once for each model. Most of the parameters are fixed and only few of them (e.g., the shadow fading effect and the I/C ratio of the fluorescent light effect) are drawn randomly for every realization. Drawbacks of the models are the limited capability to model the antenna arrays and the assumption of separable Tx and Rx PASs for each tap. In principle, all antenna geometries in the horizontal plane are possible. However, it is not possible to model the antenna radiation patterns; all antennas are assumed to be strictly omnidirectional. System level features are not supported in this model. All parameters are given for a single link.

The parameter tables, which are the essence of the IEEE 802.11n models, can also be used with parametric ray-based modeling such as 3GPP SCM. Because parameters are geometrical, they can be fitted to the channel coefficient calculation of SCM, if necessary. This would increase the complexity of the calculations to some extent because every delay has several multipath components and in the SCM type model each multipath should be calculated separately.

2.3.1.2 3GPP Spatial Channel Model

A drive towards MIMO channel modeling between 3GPP and 3GPP2 is done by a joint group, called 3GPP-3GPP2 SCM AHG. The joint group was formed with the purpose of harmonization of the spatial channel modeling assumptions between the 3GPP2 SCM Ad Hoc Group and the 3GPP MIMO Ad Hoc Group. The scope of the 3GPP-3GPP2 SCM AHG was to develop and specify the parameters and methods associated with the spatial channel model that are common to the needs of the 3GPP and 3GPP2 organizations. The goal was to develop a model as a common reference for evaluating different MIMO concepts.

The 3GPP-3GPP2 SCM consists of two “submodels,” the link level channel model and the system level channel model. In the following subsections, the SCM modeling approach is described briefly for both submodels, their main parameters, and key features. The technical details can be found in [79].

2.3.1.2.1 Link Level SCM.

The SCM link level channel model considers a MIMO link in which a single BS transmits to a single mobile station (MS). It is an extension (to the spatial domain) of the fixed tap-delay channel models specified in ITU-R Recommendation M.1225. The difference from the ITU power-delay profiles is that an optional LOS component has been added. All paths are assumed independent and each is characterized by its own spatial channel parameters, PAS, AS, and mean AoA. Reference [79] mentions that the link level model is for calibration purposes, which might refer only to the 3GPP related work. The described link level assumptions define a set of spatial parameters that corresponds to static channel conditions. A variety of SCM parameters examined for a link level analysis are shown in Table 4.1 in [79]. Each channel scenario has its distinctive parameter values. Four different model cases termed “Case I” through “Case IV” are suggested.

The assumptions of the link level simulation do not define a specific array for either the BS or the MS. The proposed antenna patterns used for each sector, reverse link and forward link, are plotted in Figures 4.1 and 4.3 in [79] for three-sector and six-sector cells, respectively. These antenna patterns are targeted for diversity-oriented implementations (i.e., large interelement spacing); for other applications, different antenna patterns can be used.

For each antenna element at the MS, the antenna pattern is assumed omnidirectional with an antenna gain of -1 dBi. Because an omnidirectional MS antenna is assumed, the received PAS per path will remain either Laplacian or uniform. The per-path Doppler spectrum is not defined explicitly; it is implicitly determined by the direction of travel (DoT), the per-path PAS and AoA, and, to some extent, by the implementation of the model (correlation or ray based).

2.3.1.2.2 System Level SCM.

The principle of the SCM system-level channel follows some of the COST 259 recommendations. No specific antenna topologies are enforced and linear array geometry is discussed. The model is a ray-based model where one subset of the parameters contains the stochastic parameters. The center frequency and bandwidth are 2 and 5 MHz, respectively. As a geometric model its uplink and downlink are reciprocal in TDD operation, that is, uplink AoDs are also the downlink AoAs, and vice versa. For frequency-division duplex (FDD) systems, the phases between uplink and downlink are uncorrelated but are fully correlated for TDD systems. In the SCM the simulation is carried out as a sequence of “drops.” During a drop, the channel undergoes fast fading according to the speed of the mobile station. The parameters are regenerated randomly after each drop. Large-scale channel parameters, such as angle spread, delay spread, and shadowing stay constant during a drop; hence, there is no transition behavior in the model.

A key parameter in a geometric channel model is the number of rays along with their spatial, amplitude, and temporal properties. In the SCM each path is formed by summing a number of rays (called subpaths). Hence, each path consists of a superposition of several subrays (or subpaths). The parameters of the subrays have been predefined to produce the desired angular spread. In the SCM model, six paths each with 20 subrays are used. The physical interpretation is that each path is the last interaction with a cluster of 20 scatterers. The principle is illustrated

in [78], where only one cluster, or path, is shown for simplicity. In the beginning of a drop, for each MS, the antenna orientation and the gain parameters are fixed. Then the composite rms delay spread, angle spread, and shadowing parameters are drawn from the distribution functions with predefined parameters. The shadowing between MSs is uncorrelated. However, for a single MS, the shadowing, rms delay spread, and rms angle spread are correlated. Shadowing at the MS between different BSs is correlated.

The center positions of the clusters are determined randomly based on these rms parameters. There are two levels of randomness: (1) random rms parameters (delay/angle/shadowing deviations) drawn from predefined distributions, and (2) random path parameters (delay/angle/shadowing) drawn from distributions defined by the rms parameters. Note that, although the cluster positions are random, the positions of scatterers within a cluster are fixed; this produces the fixed per-path AS as defined in the SCM. Each cluster has 20 scatterers, which causes the 20 subrays corresponding to one of the six rays. Each subray has the same delay and identical power but has different angle of arrivals and departure, which are predefined as a relative offset to the AoA and AoD of the corresponding ray. The subrays at the MS and BS are then randomly paired.

One of the characteristics of the model includes a narrow AS per path with specific mobile station AoA and base station AoD models. The relationship between the temporal and spatial properties has been described by using the results obtained from [80]. In [80] a description of the relationship between the azimuth spread and the standard deviation of the AoA at the BS, as well as a description of the relationship between the distribution of ray delays and the delay spread, is given. The distribution of the occurrence of ray delays is given by an exponential distribution. Corresponding to the increasing delay is a reduction in power also having an exponential decay. The σ_{AS} and σ_{DS} are power weighted terms, and related to the σ_{AoA} and σ_{delay} , which describe the distribution of the angle and delay occurrence, respectively. The relationship describing the power concentration in the angle and the delay (i.e., $r_{AS} = \sigma_{AoD}/\sigma_{AS}$ and $r_{DS} = \sigma_{delays}/\sigma_{DS}$) is used for the generation of the channel realizations. The values of r are specified to describe the distribution of powers in delay and in angle domains to represent the behavior of the channel. The SCM model follows these relationships. It uses a different value for the suburban macro and urban macro scenarios but is not used for urban micro scenarios.

The overall procedure for generating the channel matrices consists of three basic steps:

1. Specify an environment.
2. Obtain parameters to be used in simulation.
3. Generate channel coefficients based on the parameters.

In the first step, selecting the scenario leads to a specific path-loss model. For suburban macrocell and urban macrocell environments, the modified COST 231 Hata urban propagation model is used in the calculations. For the microcell environment, the COST 231 Walfisch-Ikegami model is used for the path-loss calculations. Specific antennas height and environment parameters are assumed. The detailed generation steps of the user parameters are given in [79].

An overview of the SCM channel model is shown in Figure 2.25. Only one cluster (out of six) is shown for simplicity. The pairing between the AoA and AoD subpaths within a path is random.

In addition to the characteristics described above, the following options exist:

- *Polarized arrays*: Cross polarization discrimination is defined. The current model can cover different types of polarized antennas.
- *Far scatterer cluster model urban macro*: This models a bad environment where additional clusters are seen in the environment.
- *LOS model*: The Rician factor is a function of the distance from the BS. The LOS component appears with linearly decreasing probability with respect to distance, with a cut-off distance at 300m from the BS.
- *Urban canyon model*: When switched on, the model modifies the AoA of the paths arriving at the MS.

Based on ranking of the path loss and shadowing of all sectors, strong interferers are modeled as spatially correlated processes, and the remaining sectors are modeled as spatially white Gaussian noise processes based on a flat Rayleigh fading process. To minimize the complexity of the model, the spatial characteristics of the weak interferers are considered spatially white, whereas that of the strong interferers is modeled spatially.

The SCM 3GPP-3GPP2 system level channel model considers three cases of channel scenarios:

- *Suburban macro*: approximately 3-km distance from BS to BS;
- *Urban macro*: approximately 3-km distance from BS to BS);
- *Urban micro*: less than 1-km distance from BS to BS.

The macrocell definition applies when the BS antennas are above the rooftop height. The microcell definition applies otherwise. The SCM model can be extended

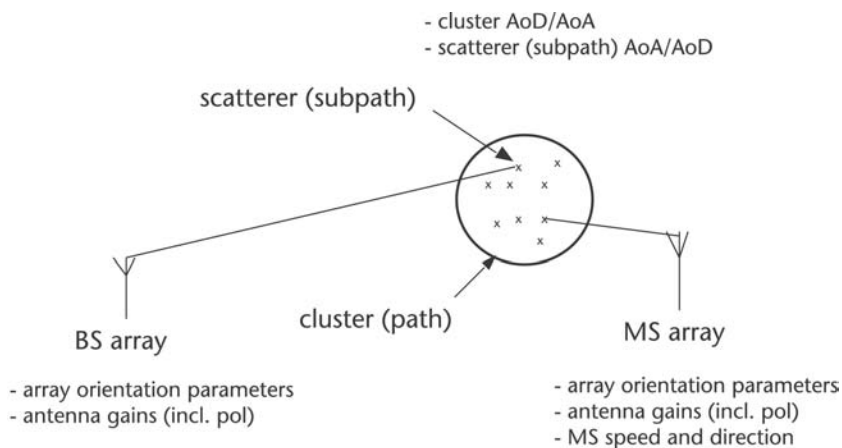


Figure 2.25 Overview of the SCM channel model [81].

to other environments, bandwidths, and center frequencies by redefining the key parameters, such as the minimum path delay difference, the number of paths as proposed in IEEE 802.20 MBWA, the number of subpaths per path, the path loss, the distributions of the rms delay spread, the rms angle spread, the shadowing deviation, and their dependencies. The parameters r_{AS} and r_{DS} , defined earlier, are very important to consider for further improvements of the model for the defined scenarios or other new scenarios such as outdoor to indoor as well as the indoor case.

2.3.1.3 METRA Channel Model

The METRA channel model was developed within the FP5 IST project METRA and the FP5 IST project I-METRA [54]. The main objective of the IST METRA project was to analyze the feasibility and evaluate the performance of introducing multiple-element adaptive antennas into mobile terminals in combination with adaptive BS antenna arrays for UMTS.

The purpose of the channel characterization work in I-METRA was to exploit the measurement data collected during the FP4 (e.g., ACTS Program) project SUN-BEAM (Smart Universal Beamforming) [82] and the METRA projects, in order to gain better insight into the properties of MIMO radio channels. Influences of a large set of parameters and contexts, such as correlation, polarization, directional information, power imbalance, and presence of a LOS component were studied in I-METRA.

The latest version of METRA implements the model and a link level simulation based on it. However, it does not mention how to base a system level simulation on the METRA channel model. The model supported a one-link system, that is, one BS and one MS only. Thus, multihop is not supported. Outputs of the I-METRA project in [83] included the software link level channel model implemented in MATLAB.

The METRA model is a stochastic channel model. Its main strength lies in the fact that it condenses the correlation between the antenna arrays at the receiver and at the transmitter, in two correlation matrices. This is also its weak point. The characterization of the setup is conventionally defined with a power delay spectrum and Doppler spectrum. The PDS determines the realizations of the channel impulse response (CIR) that can be modeled as a tapped delay line. The taps of the CIR and their statistics in turn determine the temporal channel transfer functions. The temporal evolution of the CIR is determined by the Doppler spectrum.

The PDS is modeled by a one-sided decaying exponential. However, in some environments, the power is not monotonically decaying with delay time, as waves arrive at the termination point gathered in clusters. This clustering effect has been modeled in the so-called “hilly” outdoor environments of [84]. A similar clustering process has been identified in [85] for indoor environments. Several tap-based models are proposed to model different environments. They differ in the number, spacing, and relative power of the taps. As a consequence, different terrain profiles are characterized by different delay spreads (DSs). The DS is defined as the root second moment of the PDS.

The time-domain characterization of the setup is defined by the Doppler spectrum. The fading characteristics of the taps are defined by shaping an oversampled Doppler spectrum in the spatial frequency domain. The inverse Fourier transform of this Doppler spectrum defines the complex random fading coefficients in the space domain. Then, it is simple to convert them into the time domain, by taking into account the speed of the mobile. The predefined shapes of the Doppler spectrum are classical (Clarke's spectrum) and flat. It is possible to have any user-defined Doppler spectrum as an input.

The spatial correlation of the antenna arrays is defined by two correlation matrices, namely, \mathbf{R}_{MS} and \mathbf{R}_{BS} , which characterize the correlation at the MS and BS, respectively. The basic assumption for this approach is that the correlation properties of the two link ends can be treated separately; that is, the BS and the MS are not spatially correlated. As a result of this assumption, the joint spatial correlation matrix \mathbf{R} is approximated as the Kronecker product of the \mathbf{R}_{MS} and the \mathbf{R}_{BS} .

The resulting matrix is then properly normalized, multiplying by the inverse of the total channel energy. The \mathbf{R}_{MS} and \mathbf{R}_{BS} matrices are calculated from the PAS and azimuth spread. The AS is defined as the root second moment of the PAS. Different shapes for the PAS are defined for the different scenarios shown in Table 2.11.

The various scenarios have different azimuth spreads: low in rural environments with a tendency to increase as scattering becomes more significant, higher in micro- and picocells. AS increases with a decreasing distance between the MS and BS. The larger (lower) values of the AS have as a consequence lower (larger) correlation between antenna elements. A table with values of AS for certain environments is provided in [83]. Having the DoA angles and choosing an AS, the PAS can be formed, depending on a chosen distribution. The \mathbf{R}_{MS} and the \mathbf{R}_{BS} can be calculated for any antenna topology and pattern.

In principle, any scenario can be modeled with the METRA channel model, given the proper correlation matrices \mathbf{R}_{MS} and \mathbf{R}_{BS} and the description of the channel behavior in time and frequency. This is due to the fact that the parameters of the models have not been given, but have to be given during the realization of the channel coefficients. In this respect the METRA model differs from the other candidate models. In particular, picocellular, microcellular, and macrocellular environments were investigated and evaluated. Calculation of the channel coefficients is comparable in complexity to that of the IEEE 802.11n model.

The METRA model can be evolved to other environments by redefining the key parameters. For example, it is possible to change the center frequency and to increase the bandwidth.

Table 2.11 Different PAS Shapes

		BS	MS
Outdoor	Macrocell	<ul style="list-style-type: none"> • Laplacian • Power of a cosine function • Truncated Gaussian • Uniform 	Uniform
	Microcell		
	Picocell		
Indoor		Almost uniform	
		Uniform	

2.3.1.4 Other Models

Other well-known standard models for WLAN and cellular environments cannot be used directly for MIMO simulations; they either do not include the MIMO characteristics, or they are not complete.

Some important channel models should be mentioned, such as the following MIMO models: IEEE C802.20 [86] and the model developed by the FP5 IST project FITNESS [54]. In addition, development work was done for MIMO models through the COST 273 project. The COST 259 model is another well-known directional reference model [87], the ideas of which have been applied in many subsequent models. Then there are a number of 3G (SISO) models, such as 3GPP, 3GPP2, ITU-R 3G, and UMTS 30.03 models [88], GSM models [89], some WLAN models, such as ETSI BRAN, and various other models. IEEE 802.11n models have been developed on the basis of the ETSI BRAN models. Thus, ETSI BRAN can be considered a simplified SISO case of IEEE 802.11n.

2.3.1.5 Comparison of the Existing Channel Models

The generality requirement suggests the use of ray-based (geometry-based) models. The only ray-based model available is the SCM model. The SCM model cannot as such be applied to WLAN-type subsystems, because its bandwidth is considered too small. The link level SCM model is considered restricted and cannot be used as an initial link level model for next generation air interfaces. When using the stochastic models IEEE 802.11n and METRA, note that the Kronecker product principle assumed in both models oversimplifies the models.

In the single Kronecker product model, the spatial correlation of the antenna arrays is defined by two correlation matrices, \mathbf{R}_{MS} and \mathbf{R}_{BS} , which characterize the mobile station and base station, respectively. The basic assumption for this approach is that the correlation properties of the two link ends can be treated separately; that is, the BS and the MS are not spatially correlated. Separability offers independent optimization of transmit and receive arrays, leading to reduced complexity and simplicity in such a model. As a result of this assumption, the joint spatial correlation matrix \mathbf{R} is approximated as the Kronecker product of the \mathbf{R}_{MS} and the \mathbf{R}_{BS} . The resulting matrix is then properly normalized, multiplying by the inverse of the total channel energy.

In [90] and [91] it was shown that the Kronecker assumption does not hold, in general, for realistic MIMO channels. The Kronecker model fits well only in a scenario where the scatterers are assumed to be uniformly distributed around the antenna arrays at both links. When using the Kronecker product principle, the receiver is not affected by any kind of spatial filtering of the transmitter. Whatever the transmitter is doing, the Rx array always experiences the same spatial distribution of signal energy. Due to the assumed separability, all directions of departure (DoD) are coupled to the same directions of arrival (DoA) at the receiver. The Kronecker model underestimates correlation between the paths for the most part. The reason is that the basic assumption neglects the interdependence of DoAs and DoDs.

All the preceding models were defined either in the vicinity of 2 or 5 GHz. The IEEE 802.11n model was defined for both of them [77]. Therefore, it is assumed

that all of the models can be used in the frequency range from 2 to 5 GHz, or even up to 6 GHz. It seems that there is no major difference in the models in this respect, except for the difference in path loss.

All of the considered models are meant for systems with considerably lower bandwidth than needed for next generation radio systems. The IEEE 802.11n model was planned for the bandwidth of approximately 20 MHz. The SCM and METRA models were planned for the UMTS bandwidth of approximately 5 MHz. Next generation radio systems assume a 100-MHz bandwidth, which implies the need for different channel models.

The IEEE 802.11n model uses mostly more than 10 taps (14 to 18) with minimum tap spacing of 10 ns. From these figures it can be assumed that it would support the 100-MHz bandwidth, at least quite nearly. The SCM model uses only six path delays. It was also planned for the 5-MHz bandwidth only. For this reason the model is probably too narrowband for the widest bandwidth required of next generation radio systems. In the METRA model the tap numbers and spacings have not been defined explicitly.

All three models support link level simulations. However, only 3GPP SCM can also support system level simulations, which is its main goal. In fact, the link level models of the 3GPP SCM are meant to be used only for calibration purposes. None of the models supports ad hoc or multihop networks.

Although the METRA and IEEE 802.11n channel models support short-range scenarios, the 3GPP SCM model does not. The IEEE 802.11n channel model is especially suited for indoor MIMO WLAN systems. The supported environments are residential, residential/small office, typical office, large office, and large space. Every environment can be with or without the LOS component. Additionally it supports outdoor hot-spot scenarios. The modeled value for the mobile speed passing by is 40 kmph.

The METRA channel model also supports different environments: small office, typical office, large space, and modern open office.

Wide-area communications of interest to next generation radio systems include rural, suburban, urban, and outdoor-to-indoor scenarios. The IEEE 802.11n model was intended for indoor and short-range outdoor environments with small delay spreads (i.e., picocells and small microcells). Thus, it is clear that it is not a candidate for wide-area communications. The METRA model, in principle, can be used for any scenario if the proper correlation matrices are found or given. However, the model itself is suitable for link level simulation. It has no system level features. As a result, it cannot be used for MIMO system level simulations when designing a new air interface.

The 3GPP-3GPP2 SCM model was intended for outdoor propagation. The SCM model is defined for urban macro and urban micro as well as suburban scenarios. It is not defined for suburban outdoor-to-indoor scenarios. The SCM has both link level and system level channel model features. If the parameters required for SCM for undefined outdoor scenarios (e.g., suburban outdoor to indoor) are known, then the model can also be considered for this scenario. Of the three candidates, the most suitable model for wide-area communications for next generation radio system investigations is the 3GPP-3GPP2 SCM model. It

is the only model that specifies model parameters for a wide range of outdoor environments for both link level and system level.

2.3.1.6 Selection of State-of-the-Art Channel Models for Next Generation Systems

The IST project WINNER evaluated the state-of-the-art channel models for an initial selection to their suitability for the WINNER air interface design. The proposed WINNER initial model for the short-range simulations was the IEEE 802.11n model. It was assumed short-range scenarios include WLAN-type systems. Then the different networks can probably be simulated separately in the initial simulations. The proposed initial WINNER short-range scenarios are shown in Table 2.12.

The proposed WINNER initial model for the wide-area simulations was the 3GPP SCM model, more specifically, the system level part of the model. The model was proposed for the narrowband models (5-MHz bandwidth). The wide-area systems are assumed to be mainly of the cellular type. The system level simulations with the proposed model are assumed to be restricted to a limited number of BSs and MSs. If extensive numbers of base stations and terminals should be simulated, the simulation should use a higher abstraction level.

2.3.2 Channel Modeling

Development of new channel models requires two steps [92]: (1) setting up a generic channel model and identifying the parameters that have to be determined for its description, and (2) actually performing the measurement campaigns and extracting statistics and distributions of the parameters.

2.3.2.1 Generic Channel Model Statistics

The term *statistical modeling* is typically used to describe a modeling approach that directly models the observed channel in a communication system.

Under the assumption of limited signal bandwidth, limited Doppler bandwidth, finite delay spread, and linearity of the propagation channel, a communication system observes the propagation channel as a discrete-time finite-length filter. The effective (observed) channel thus has only limited degrees of freedom. From a propagation point of view, the channel has infinite bandwidth (and, hence, resolution). Because scatter contributions are often physically clustered together and because of the limited resolution of the observation, it makes sense to define a

Table 2.12 IEEE 802.11n Model Suitability for Short-Range Scenarios in the Scope of WINNER

Scenario		IEEE 802.11n Model
Short Range	Urban (outdoor to outdoor)	F
	Indoor (indoor to indoor)	A, B, C, D, E, F
	Infostations	Not available

channel model by concentrating on these clustered and dominant scattering effects. That can be done by simply dropping the discrete delay times. It is important to note that this is merely a model and the mapping of parameters from a specific propagation channel to this model is not unambiguous (e.g., depends on bandwidth). Due to the large variability in propagation environments and resulting channels, only a statistical description makes sense. With respect to the previous model, we thus get random variables for the delays, tap coefficients, and number of taps/clusters. Under a time-variant assumption, these are modeled as random processes.

Assuming at least wide-sense stationarity, the model can be fully described by the joint distribution of all parameters at all time differences. If it is further assumed that the variables at different delays are independent of each other, we arrive at the classic wide-sense stationary uncorrelated scattering (WSSUS) assumption. From the channel observations, certain properties of the physical (geometric) propagation effects can be derived if additional assumptions are employed. With the assumption of only one reflection per path, for example, a path delay can be mapped to the potential locations of that scatterer, which is described by an ellipse around transmitter and receiver. On the other hand, scatterers can be distributed on such an ellipse, and the statistics of the resulting filter coefficients will match that of the observations. This is referred to as an elliptical model.

The use of multiple-antenna systems adds another dimension to the channel. Individual channels of the type described above can be attributed to each pair of transmitting and receiving antennas. Due to the close distance between the antenna elements, most parameters such as number of paths and delays of the individual paths are equivalent across antenna combinations. In this case we can simply replace the tap coefficients by a vector or matrix.

Again, it is possible to map the channel observations to the physical propagation effects under additional assumptions. For example, with the assumption of plane waves, small relative bandwidth, and the knowledge of the antenna response, the angle of an impinging wave can be estimated. This estimation process is, however, confined by an equivalent resolution limit to the time domain. Specifically, for each delay resolution bin, only $N - 1$ waves can be estimated, if N is the number of antennas.

Similar to the instantaneous filter coefficients, statistical parameters can be related. For a certain delay resolution bin, the correlation between two spatially separated antennas is directly dependent on the angular power spectrum (APS or PAS) of the contributions for that bin.

2.3.2.2 Ray-Based Modeling

In radio-wave propagation at high frequency, multipath propagation wireless channels can be understood and modeled by applying ray theory. A uniform plane wave can be seen as a ray. Hence, each radio path is approximated by a ray, where energy propagates between antennas of the transmitter and receiver within ellipsoid-shaped tubes defined by first Fresnel zone. The higher the frequency, the more accurately the fields, which are called *ray fields*, can be approximated by the zeroth order of the Luneberg-Kline expansion. The interaction of radio waves (or rays)

with the environment determines the characterizing parameters of each ray. Once the detailed database of the environment is known in addition to the locations of the transmitter and receiver, the channel properties are derived from the positions of the scatterers by applying the fundamental laws of the propagation mechanism of electromagnetic waves; that is, rays can be traced to find the coupling paths. This is known as *ray tracing* (i.e., deterministic modeling).

The rays can be characterized from the propagation environment by their delay (or length), angle of arrival, angle of departure, Doppler shift, polarization, and amplitude. Ray-based modeling is also the heart of the geometry-based stochastic modeling approach, which is the most commonly used approach in directional channel modeling for testing the performance of adaptive antenna or MIMO systems. The approach assumes a statistical distribution of discrete scatterers around the two (or one) ends of the wireless link. With the knowledge of position of the scatterers of each channel realization, the rays can be characterized by their parameters because each scatterer usually corresponds to one ray.

The positions of the scatterers are drawn from the probability density functions (PDFs) of multipath delays and the direction of arrivals and departures by the ray-based approach. Depending on the distribution of the obstacles, the interaction of the transmitted signal with the environment changes and consequently different signal components appear superpositioned at the receiver, having a multipath channel. The scatterer distributions depend on the environments. In an indoor environment, the scatterers can be around both of the wireless link ends; in the case of highly elevated antennas (i.e., macrocells), the scatterers can be considered only around the mobile station. Each scatterer has its own direction of departure, DoA, and time of arrival (ToA), which can be determined by the ray-based approach. The distribution of scatterers has a clear influence on the distribution of the direction of arrival and direction of departures and time of arrival. Once the channel parameters of each ray (i.e., delay, azimuth angle, elevation angle at the MS and BS, Doppler frequency, and complex amplitudes) have been characterized, the channel behavior is identified in multiple dimensions.

In general, the understanding of the physical propagation phenomena is based on a *ray-optical* approach. Paths are modeled by planar, narrowband wavefronts. This is motivated by the idea of specular reflections at smooth surfaces. To model the influence of receiver noise, a white noise component is usually added. It is, however, well known that wave propagation phenomena may also comprise diffuse scattered components [93]. Their contribution varies depending on the complexity of the propagation environment. It can be almost negligible in macrocell LOS scenarios, but might even dominate in complicated propagation environments such as factory halls [94]. A data model comprising two components can also be introduced [95]. The first part is considered deterministic and results from a limited number of specular-like reflections. This can be called the structural part of the model because it has a clear geometric interpretation. The second part is observed as a dense, diffuse part that is stochastic in nature and cannot be resolved by the measurement device. It results from distributed diffuse scattering as it occurs in a complicated, multipath rich environment. For example, a sounder having a measurement bandwidth of 120 MHz [96] gives excellent possibilities to resolve a number of specular components. Even though, the spatial resolution is only about

2.5m, which corresponds to 43 wavelengths at 5.2 GHz. Hence, in a “microscopic” sense a big number of superimposed diffused components in any delay bin can be expected. This can be called a *dense multipath model*. The resulting CIR part is therefore adequately modeled by a complex circular normal distribution.

It might be argued that diffuse components can be negligible in the presence of specular paths. This is, however, not consistent with experience. One explanation could be that specular paths can contribute to the received power only for very distinct angular constellations. On the other hand, diffused power has the chance to reach the receiver within a large (almost continuous) variety of propagation angles if there is a large number of widely distributed scatters. Note that modeling of diffuse scattering for the purposes of parameter estimation does not need to model the individual scatterers. Rather, a model is needed that describes the superimposed contributions at the receiver.

2.3.2.3 Modeling Based on Spatial Correlation Matrix

With correlation-based channel models, the spatial correlation is explicitly defined and generated by means of a spatial correlation matrix, whereas with ray-based models the correlation is implicitly present in the channel matrix generation process. The generation of MIMO channel matrices based on a channel correlation matrix is defined as follows:

$$\mathbf{R} = \mathbf{E}[\text{vec}(\mathbf{H})\text{vec}(\mathbf{H})^H] \quad (2.16)$$

where $\text{vec}(\cdot)$ operator is the vectorization operator, which stacks all elements of a matrix in one column. The channel matrix \mathbf{H} is for a single “delay tap;” in general, the correlation matrix \mathbf{R} can be different for each channel tap. The matrix \mathbf{R} captures the spatial correlation properties of the MIMO channel at the transmitter and receiver end. Generation of Rayleigh fading channel matrices with the correlation structure defined by \mathbf{R} can be done according to

$$\mathbf{H} = \text{vec}^{-1}(\mathbf{R}^{1/2}\mathbf{g}) \quad (2.17)$$

where \mathbf{g} is a zero-mean complex circularly symmetric Gaussian vector with independent unit-variance entries, and $\text{vec}^{-1}(\cdot)$ is the inverse vectorization operation. Note that the desired time correlation (i.e., Doppler spectrum) has to be induced into the sequence of channel matrices by filtering the elements of \mathbf{H} over time. A deterministic LOS signal component with a predefined K factor can also be added to the signal to simulate Rician fading.

The spatial correlation matrix, which basically forms the model, can be estimated from measured channel matrices, matrices generated using a ray-based model or (in some cases) derived from analytical calculations. The correlation matrix is antenna array dependent and hence has to be reestimated for different arrays. To simplify analysis and the model itself, a Kronecker structure is often imposed on the correlation matrix; that is, it is assumed that the correlation matrix can be written as a Kronecker product $\mathbf{R} = \mathbf{R}_{\text{Tx}} \otimes \mathbf{R}_{\text{Rx}}$. In this case the channel matrices are generated using

$$\mathbf{H} = \mathbf{R}_{\text{Rx}}^{1/2} \mathbf{G}_w \mathbf{R}_{\text{Tx}}^{1/2} \quad (2.18)$$

where \mathbf{G}_w is an $M \times N$ matrix with zero-mean complex circularly symmetric Gaussian entries. The advantage of the Kronecker assumption is that (2.18) is a computationally simpler operation than (2.17). The underlying assumption is that the directional properties of the radio propagation channel are independent at the receiver and the transmitter. Methods for finding the “optimum” Kronecker approximation for a general correlation matrix \mathbf{R} have been reported in [97–99]. Often the Kronecker approximation is sufficient [98], although it has also been shown to have certain deficiencies [100, 101].

The main advantages of the correlation-based approach are its computational and modeling simplicity; a spatial correlation matrix and a Doppler filter essentially define the model. The main drawbacks are that, because the antennas affect the spatial correlation, the correlation matrix must usually be reestimated for all array geometries, and that the model parameterization describes only the second-order statistics of the channel without any physical interpretation of the propagation medium. Also, with the Kronecker assumption, the modeling accuracy may not be enough for all channel scenarios. On the other hand, using the full correlation matrix \mathbf{R} increases the computational complexity. MIMO channel models using the correlation matrix and the Doppler filtering approach include METRA, I-METRA, and the IEEE P802.11 model.

2.3.2.4 Antenna Interdependency of the Propagation Models

A radio channel model is called *antenna independent* if the antennas and propagation channel are separable in the model formulation. The antenna-independent model supports arbitrary antenna array geometries and arbitrary radiation patterns for the antenna elements. The antenna information can be embedded in the generation procedure of channel model realization.

Reference [102] provided a mathematical notation for a MIMO system antenna-independent channel model proposed for the WINNER air interface. The transfer matrix \mathbf{H} of the MIMO channel with L paths is given by:

$$\mathbf{H}(t, \tau) = \sum_{l=1}^L \mathbf{H}_l(t; \theta_l) = \sum_{l=1}^L \exp(j2\pi\nu_l t) \mathbf{C}_2(\Omega_{2,l}) \mathbf{A}_l \mathbf{C}_l^T(\Omega_{1,l}) \delta(t - \tau_l) \quad (2.19)$$

where the components included are the Doppler frequency ν , the Rx array response matrix \mathbf{C}_2 , the polarization matrix \mathbf{A} , the Tx array response matrix \mathbf{C}_1 , and the delay τ . The parameter vector is

$$\theta_l = (\Omega_{1,l}, \Omega_{2,l}, \tau_l, \nu_l, \mathbf{A}_l) \quad (2.20)$$

The polarization matrix \mathbf{A} of path l is a 2×2 matrix, whose elements $\alpha_{l,i,j}$ are the complex gains from Tx polarization i to Rx polarization j of path l , here $i, j \in \{1, 2\}$ referring to two polarizations:

$$\mathbf{A}_l = \begin{bmatrix} \boldsymbol{\alpha}_{l,1,1} & \boldsymbol{\alpha}_{l,1,2} \\ \boldsymbol{\alpha}_{l,2,1} & \boldsymbol{\alpha}_{l,2,2} \end{bmatrix} \quad (2.21)$$

The array response matrices are composed of array response vectors of two polarizations; for the Tx array:

$$\mathbf{C}_1(\Omega_{1,l}) = [\mathbf{c}_{1,1}(\Omega_{1,l}) \mathbf{c}_{1,2}(\Omega_{1,l})] \quad (2.22)$$

and for the Rx array:

$$\mathbf{C}_2(\Omega_{2,l}) = [\mathbf{c}_{2,1}(\Omega_{2,l}) \mathbf{c}_{2,2}(\Omega_{2,l})] \quad (2.23)$$

The array response vector $\mathbf{c}_{i,j}$, where $i = 1$ denotes Tx, and $i = 2$ Rx, j stands for j th polarization. On the vectors below $\mathbf{r}_{i,m,j}$ is a vector to the m th antenna element with j th polarization. For example,

$$\mathbf{c}_{1,1}(\Omega_{1,l}) = \begin{bmatrix} G_{1,1,1}(\Omega_{1,l}) \exp(j2\pi\lambda_0^{-1}(\Omega_{1,l} \cdot \mathbf{r}_{1,1,1})) \\ G_{1,2,1}(\Omega_{1,l}) \exp(j2\pi\lambda_0^{-1}(\Omega_{1,l} \cdot \mathbf{r}_{1,2,1})) \\ \vdots \\ G_{1,M_1,1}(\Omega_{1,l}) \exp(j2\pi\lambda_0^{-1}(\Omega_{1,l} \cdot \mathbf{r}_{1,M_1,1})) \end{bmatrix} \quad (2.24)$$

where $G_{i,m,j}(\Omega)$ is the antenna field pattern of the Tx/Rx ($i = 1/2$) m th antenna element on the j th polarization and direction Ω .

The right-hand expression in (2.24) is valid if coupling is neglected. However, (2.19) is still valid for any arrays $\mathbf{C}_1(\Omega)$ and $\mathbf{C}_2(\Omega)$ and the array responses (e.g., obtained from calibration).

2.3.3 Simulation of Radio Systems

The performance of transmission techniques for new advanced multiple-antenna systems requires both *link level* and *system level* simulations. Different types of simulations were specified for the WINNER air interface design and were reported in [103]. In addition, the simulations can be static, quasistatic, or dynamic.

The link level simulation focuses on the performance of one radio link between the transmitter and the receiver. If a repeater is involved, the link level simulations include the two hops and the repeater function [103]. The link level simulations may include several mobiles if a multiuser detection receiver is used and also includes several base stations if soft handover is modeled [104]. The link level simulations simulate performance of a single link with a high resolution in time, that is, work in the bit rate (or chip frequency) scale, which is required for an accurate receiver performance evaluation. The performance evaluation may be measured, for example, by the BER, frame error ratio (FER), and so forth.

The scope of a link level simulation is to model the physical layer functionalities. For example, the performance of channel coding and decoding algorithms and

modulation and demodulation algorithms can be compared based on link level simulations. The link level simulations include the necessary parts of transmitting and receiving structures in a system and the proper fading channel model in order to compare the performance of different candidate algorithms.

From a channel modeling perspective, a channel model for link level simulation is the channel between two terminals with antenna arrays for the MIMO case. It is well known that the performance gain achieved from a single link could not be interpreted as a similar gain in system level. At the system level, multiple base stations communicate with multiple mobile stations. The link level simulations of all links in the system level are clearly prohibitive at the time resolution scale of the link level simulation. On the other hand, the time resolution for the system level simulations need to be of the same order as the time constant of the fastest simulated control loop. The link level and system level simulations are linked, so that in most cases the system level simulations use results calculated in the link level simulations as tabulated input parameters [103].

Mobile radio channels can be narrowband (i.e., flat fading channels) or broadband (i.e., frequency selective fading channels). Therefore, different channel models have to be developed. In mobile radio channels, the high mobility causes rapid variations across the time dimension, large multipath delay spread causes severe frequency-selective fading, and large multipath angular spread causes significant variations in the spatial channel responses. For best performance, the transmitter and receiver algorithms must accurately track all dimensions of the channel responses (space, time, and frequency). The algorithms that can be considered optimal for additive white Gaussian noise (AWGN) and flat fading channels, will neither maximize system performance nor guarantee robustness in applications to frequency-selective channels. Channel models for SISO systems provide information about the distributions of envelope and Doppler shifts of the received signals [105–109]. The link level simulation results for SISO systems are highly dependent on these two parameters. Channel models of single-input/multiple-output (SIMO) and multiple-output/single-input (MISO) for link level simulation were studied and different models were proposed [110]. The MIMO channel models that are based on the classical understanding of multipath fading and Doppler spread incorporate additional concepts by including a spatial dimension at the transmitter and receiver ends (i.e., double directional), which is presented in terms of AS, AoA, PAS, and the antenna array correlation matrices at the transmitting and receiving ends. These different domain phenomena and their correlations have a strong influence on the results of the link level simulations. A large body of publications in the open literature deals with channel modeling for single link scenarios.

The link throughput, transmitter, and receiver processing were extensively investigated for MIMO transmission at the link level [111–116]. The channel models used initially started with uncorrelated flat fading channels and successfully took into account more effects of the real environment, that is, correlations of a MIMO channel or the frequency selectivity of the channel.

A review of MIMO channel models that can be used with link level simulations can be traced in [117]. The review covered the MIMO channels of METRA, the two-ring models, the Von Mises angular distribution model [118], the distributed scattering model [119], the extended Saleh-Valenzuela MIMO model [120], the

COST 259 directional channel model, which is extended to the MIMO model [121], the electromagnetic scattering model [122], and the virtual channel model [123, 124]. These are just some examples of physical MIMO channel models. In [125], the 3GPP/3GPP2 link level models are based on the ITU models. The IEEE 802.11n model [126] is based on the ETSI BRAN models for indoor channels [127, 128], extending it to include different clusters and adopting the METRA principle to make it applicable to MIMO systems. The final report of COST 259 contains an extensive review of papers related to channel measurements and modeling [129] for link level simulations. A space-time MIMO channel model that uses a different distribution in different domains is described in [130].

2.3.3.1 Large-Scale Parameters

The large-scale fading in a MIMO radio channel can be characterized in terms of mean amplitude and mean power variation of multipath components. Both mean amplitude and power stay approximately constant within a certain interval of time and space (i.e., few wavelengths). A local-stationary area is a sufficiently small area, within which all of the positions that an MS might occupy are characterized by constant scalar measures such as path loss, shadowing, rms delay spread, rms azimuth spread, antenna correlation, channel rank, mutual information, and number of paths. The constant measures are needed in order to characterize matrix-valued impulse responses for the MIMO channel; they are called bulk parameters because they relate to all of the scalar subchannels and delay components of a full channel impulse response, and represent local averages over 20 to 100 wavelengths of movement of the mobile. These parameters are often used as inputs to the subsequent generation of the detailed impulse responses. With large-scale movements of the MS (i.e., hundreds of wavelengths), the bulk parameters are going to change considerably.

2.3.3.1.1 Path Loss.

Many approaches are available for calculating path loss either for outdoor or indoor environments. The computational complexity is a very important parameter for system level simulations. The path-loss model has to be as computationally efficient as possible without degrading its reliability at the frequency of interest. Therefore, the most favorable path-loss models are *empirical models* with a minimum number of parameters. The empirical models are usually single or double slope models extracted from measurements. Usually, the empirical path-loss formulas are defined in terms of an intercept and path-loss exponent. Then, the path loss can be calculated as a function of distance for specific values of the exponent. There are many values of the exponent for either indoor or outdoor environments. The exponent has a value of 2 in free space. If it is less than 2, that indicates the possible presence of a waveguiding effect as in LOS of indoor environments.

The intercept is determined by the free space path loss to the reference distance and an environment-dependent constant. The general formulations of the empirical path-loss models for outdoor environment can be given as follows:

$$PL(d, f) = PL_{FS,1}(f) + 10n \log_{10}(d) + C \quad [\text{dB}] \quad (2.25)$$

where

$PL_{FS,1}(f) = 20 \log_{10}(4\pi f/c)$ is the free-space path-loss at 1m distance
 d is the distance between the transmitter and the receiver
 f is the center frequency of the signal
 c is the velocity of light in vacuum
 n is the path-loss exponent that depends on the environment, and
 C is an environment dependent constant.

This model can be used in a simplified form:

$$PL(d) = A \log_{10}(d) + B \quad (2.26)$$

where $A = 10n$ and $B = PL_{FS,1}(f) + C$.

For indoor environments the well-known Keenan-Motley path-loss model [131] describes the signal path loss as follows:

$$PL(d, f) = PL_{FS}(d, f) + n_w L_w + n_f L_f \quad [\text{dB}] \quad (2.27)$$

where $PL_{FS}(d, f) = 20 \log(4\pi df/c)$ is the loss in free space for an isotropically radiating antenna, L_w is the attenuation per wall, L_f is the attenuation per floor, n_w is the number of traversed walls, and n_f is the number of traversed floors. This model is widely preferred because it models the number of walls and floors explicitly. However, its validity to the 5-GHz range is questionable and many of the published path-loss models for the indoor environment at the 5-GHz range are given in the form of (2.26) because it seems that they do not fit with (2.27).

A generalization to (2.25) is a multislope model, where different distances are covered with different slopes. An example of such is the two-slope formula used in the WLAN indoor model of [132], which defines an abrupt transition from LOS at the short range to the NLOS condition at longer range after a certain breakpoint.

2.3.3.2 Shadowing

Shadowing is widely accepted to be modeled by the well-known log-normal attenuation model, where the median value for the shadowing attenuation is zero, corresponding to a situation in which the attenuation is given by the path-loss formula used. The decibel value of the excess shadowing fading, X , has normal distribution with zero mean and standard deviation σ . The values of σ depend on the environment. Typical values, for example, in an urban environment vary from 3 to 10 dB.

Assuming a stationary environment the autocorrelation of the shadowing attenuation can be described by an exponential formula defined by the correlation distance d_c [133]. Then the autocorrelation can be assumed symmetric in the xy plane (or in practice on the earth surface):

$$R_X(\Delta r) = \sigma^2 \exp\left(-\frac{\Delta r}{d_c}\right) \quad (2.28)$$

where $\Delta r = \sqrt{\Delta x^2 + \Delta y^2}$ is the distance between the two correlated positions, Δx and Δy are the displacements in the geographical coordinates x and y between the correlated positions, σ^2 is the variance of the random variable X , and d_c is the correlation distance (where the correlation is e^{-1} times the maximum one).

2.3.3.3 Interference Modeling

For advanced broadband MIMO system evaluation, the realistic spatial as well as the temporal properties of multiple received signals, which are affected by the channel characteristics, from the desired sectors/users as well as from the interfering sectors/users are important. The accurate modeling of those intercell and intracell interferences is a crucial (e.g., for precision vs. computational complexity) and still open point. In general for link level simulations no interferers are taken in account. Hence, this is allocated as the task of the channel modeling for the system level simulations. In 3GPP [125], relatively weak interferers were considered to be spatially white; the accurate spatial characteristics were ignored, whereas the strongest interferers are treated as spatially correlated based on their covariance matrixes. With this a Gaussian noise process is colored and added. This concept is a low complexity approach, so it is not able to reflect a realistic interference situation. Therefore, it can be insufficient for sophisticated MIMO receiver testing, in particular for those who aim to suppress and cancel interferers mainly based on the spatial-temporal characteristics of all received signals.

In addition to this, a straightforward approach can be applied where the spatial-temporal characteristics of the interferers' influence is created based on the channel models for link level simulations. Here the simulation tool at the link level has to generate a new channel for each single interferer. By applying superposition and relative power differences for the interferers to the desired signals, the received signal can be created. But due to the multiple call of the link level model, this approach lacks one major issue: correlation (space, time, and frequency) between the multiple interferers, possibly other multiple BSs, and the user is not included. Furthermore, it will also increase the computational complexity, at least multiplied by the number of links.

Interference between different frequency channels is not really of consideration for air interfaces targeting a 100-MHz band and can be considered out of the scope of propagation and channel modeling [102].

2.3.4 Classifying Channel Characteristics

Link level and system level simulations require channel models, which generate channels that display specific characteristics that have a clear influence on the particular problem under study. Some characteristics can be classified as those that are of concern to link level simulations and others are of concern to system level simulations. In dynamic system level simulation, channel parameters such as directions, delays, number of paths, shadow fading loss, and so forth evolve in time.

However, in quasistatic simulations, these parameters are fixed for each *drop* but differ from *drop* to *drop*. As a general level, the channel characteristics that generally have a clear influence on small-scale fading or large-scale fading are of concern and interest to link level simulation or system level simulations, respectively. Table 2.13 shows how different channel features can be classified with respect to link level (LL) and system level (SL) simulations [102].

Realistic link level simulations require specific information about radio channels. The channel models for link level simulations are mainly for statistical characterization of small-scale fading of the mobile channels. The MIMO channel model has to describe correlation properties in the temporal, spatial, and frequency domains. A full statistical description of the channel is desirable to generate the entries of the MIMO channel matrices. One approach is the joint distribution in angle, delay, and Doppler domains [134].

For MIMO channels, the distribution of angles includes both AoAs and AoDs in both azimuth and elevation. The joint angle of arrival and departure distributions requires double directional measurements. Deriving multidimensional joint distribution is intractable mathematically and is computationally intensive. One approach, described in [92] is to use a huge file for the multidimensional distribution. To overcome such a situation, the commonly used assumption for the MIMO link level model is the independence of spatial correlation properties at the transmitter and receiver, commonly known as the METRA principle. This assumption is based on the Kronecker product approach, which neglects the joint spatial structures at the transmitter and receiver and characterizes the MIMO channel assuming that the angles of arrival are independent from the angles of departure. This technique

Table 2.13 Classification of Channel Characteristics: Link Level Versus System Level Simulations

<i>Characteristic</i>	<i>User Side Link Level Simulations</i>	<i>System Level Simulations</i>
Path loss	Fixed value	Yes
Shadowing	Fixed value	Yes
Shadowing correlation distance	—	Yes
Correlation of shadow fading between MSs	—	Yes
Correlation of shadow fading between BSs	—	Yes
Small-scale fading	Yes	Possible
Evolution of large-scale parameters	—	Yes
Cross-correlation of large-scale parameters	—	Yes
Cross-correlation between LL parameters	Yes	—
Evolution of LL parameters in time (with only small movement in space), for example, fixed-wireless and quasistatic Doppler	Yes	—
Distribution of SL parameters (shadowing, interference, LL parameters involving multiple users, LL parameters involving multiple BSs)	—	Yes
Evolution of SL parameters and LL parameters in space (averaging over small movements), for examples, birth and death of paths	—	Yes
Coexistence interference	—	Yes
Nonstationary effects in LL and SL parameters, for example, scenario transitions	—	Yes

does not describe the general spatial mechanisms of MIMO channels. It merely provides a rough estimate of its spatial properties [97].

The spatial features displayed by the MIMO channel are greatly influenced by the antenna arrays, whose characteristics are determined by the positions of their elements, as well as their beam patterns and coupling. The following list summarizes some design considerations for antenna deployment, channel measurements, and identification:

- Planar antenna arrays such as uniform linear arrays (ULAs) or uniform rectangular arrays (URAs) always have a limited viewing angle. They are useful to represent a base station's view of the channel. Circular antennas have a full field of view. They can be conveniently used to represent the mobile station.
- Double directional modeling requires antenna arrays at both ends of the link and MIMO operation of the sounder. For cellular system considerations, a combination of planar and circular arrays is adequate, whereas for ad hoc peer-to-peer networks, identical circular arrays are most preferable.
- Mainly for micro- and picocell scenarios, estimation of the elevation is desired in addition to the azimuth. This requires URA, cylindrical, or spherical arrays. Three-dimensional wave analyses (azimuth and elevation) should always include polarization resolution.
- Spherical antenna arrays may be applied for full azimuth and elevation coverage. However, no geometric solution exists that can arrange more than 20 patch antenna elements on a spherical surface with identical interelement distances. Therefore, the design of spherical arrays will be complicated by nonuniform interelement distances and various relative polarization orientations of adjacent elements.
- Full polarimetric analysis of the radio channel requires not only polarimetric reception but also polarimetric excitation of the channel. This is even true for omnidirectional excitation in which a two-port antenna is needed to send both orthogonal polarized waves with an omnidirectional characteristics and, thus, doubling the required sounder output ports.
- High and reliable resolution in terms of separation capability of closely spaced paths and low probability of outliers requires an antenna architecture that offers a maximum of antenna array aperture size in the respective spatial dimension, including a minimum number of antenna elements, low antenna element coupling, and precise calibration. This also has to include the antenna switches and feeder cables. The characteristics of the antenna elements depend on the basic element design (dipoles, patches, slots, and so on). It has a significant influence on high-resolution performance, estimation ambiguities, probability of outliers, and polarization resolution capability. Furthermore, when modeling the antenna characteristics, the accuracy and computational complexity are important; therefore, an efficient description of the polarimetric beam pattern can be applied, such as the effective aperture distribution function (EADF).

To allow antenna de-embedding/embedding in the channel model, special attention must be paid to antennas used in MIMO channel measurements. Multielement spherical, circular, linear, planar, or cylindrical arrays are required in the measurements to characterize propagation effects, such as DoA and DoD in addition to delay and path gain, and to separate them from antenna effects. Therefore, the antennas used in the channel measurements are often different from those used in the real radio systems.

In the antenna steering vector [see (2.24)] for the channel transfer matrix in (2.19), the antenna locations $r_{i,m,j}$ are freely settable. Additionally, the radiation patterns $G_{i,m,j}(\Omega)$ are freely settable for each antenna element for the two polarizations. Radiation patterns can be given as a closed-form function of the direction or as tabulated values. If available, the closed-form function would be computationally effective. Mutual coupling of array elements can be embedded in the tables of radiation patterns.

The system level aspects of channel models are statistical properties and dependencies that become important in system level simulations, that is, where, as opposed to link level simulations, the location of a user in space is included. Classic dependency parameters include propagation scenario, antenna heights, and distance between transmitter and receiver. Typical dependent parameters are path loss, shadowing, delay spread, and Rician K -factor. Most available models to date consider the distribution of these parameters (with a deterministic dependence) only. More recent publications have examined the cross-correlation between these parameters and their autocorrelation function versus space.

With the consideration of space, not only can new parameters be studied, but previous link level parameters can be reexamined in this context. For example, the evolution of taps/clusters in the delay domain has been previously rendered with a birth/death model. Now the same effect can be modeled in terms of evolution in space of the number and location of scattering clusters. This new area of spatial modeling, however, has not yet been studied very extensively.

Further interest and complexity arise if the interaction between multiple users and/or multiple BSs is considered in detail. This is particularly important for hand-over scenarios, interference evaluation, ad hoc networks, and networks featuring multihop/relaying techniques. Again the correlation and evolution in space of system level and link level parameters concerning these multiples links are of interest. A special emphasis is also on interference modeling. Because modern communication systems are able to deal with interference much more intelligently, its modeling needs to be treated in more detail than has been done previously. Besides other BSs and users, interference also arises from other communication systems (coexistence) in case the medium is shared.

Due to the ubiquitous nature of future communication systems and the move from circuit-switched to packet-switched data transmission, the importance of modeling effects in a nonstationary way has risen. Examples are the modeling of scenario transitions, and allowing for abrupt changes in shadowing or the K -factor.

With the increasing complexity of modeling, the extraction of underlying parameters becomes more crucial as well. Emphasis thus needs to be put on establishing a solid base of different estimation techniques including the understanding of implicit assumptions and limitations.

2.3.5 General Descriptions of Basic Channel Models

The influence of physical mobile radio channels has to be considered carefully within both link level and system level simulations. During transmission in mobile channels some phenomena cannot be neglected (e.g., changing delays, changing direction of arrivals, availability of multipath components). These will affect with varying BER as well as burst error structures. To evaluate the performance of MIMO algorithms, the emulation of the variations of the channel at the link level and system level are required. The test of only fixed channel parameters is not adequate because the fixed parameters only represent snapshots of the channel, which are isolated cases that do not tell the whole scenario and propagation conditions. With these physical channel features and their influence on channel properties, which affect the transmitted data, a channel model has to be as accurate as possible. On the other hand, it has to be as computationally efficient as possible. Because the required details of the channel model depend on the application, several sets of channel models should be defined. For example, the IST project WINNER defined four sets; two sets for each type of simulation: (1) basic channel model for link level simulations, (2) advanced channel model for link level simulations, (3) basic channel model for system level simulations, and (4) advanced channel model for system level simulations. Each set has different models for different scenarios, as discussed in the following sections.

Because different problems will be investigated in link level and system level simulations, the complexity of the channel model will depend on the problem under study. The channel models should be scalable and organized in components. The user can add up or reduce the complexity of the channel model based on the specific simulation of a certain problem for the application of interest. The proposed WINNER channel models are antenna independent, so different antennas configurations with different field patterns can be embedded and de-embedded. General definitions for the four sets of channel models are given next.

2.3.5.1 Basic Channel Model for Link Level Simulations

This model represents the physical channel variations for small-scale movements (i.e., a few wavelengths). In small-scale movement, the basic channel model may be sufficient but only a small set of the infinite number of possible channels is covered, because only a small number of PDPs are defined. The basic channel model for LL simulation offers characterization of the small-scale fading considering a double directional, geometry-based dispersion on different domains. Small-scale fading is the effect that is visible over short distances, whereas most other parameters tend to change rather slowly. The path loss and the shadowing have constant values since the basic model neglects a detailed characterization of large-scale fading and characteristics for small-scale MS movements. The model will be a MIMO tapped delay line (TDL) model, which will degenerate into a SISO TDL model when single antennas are used at both ends and can also be used for SIMO or MISO simulations when one antenna is used at one end and multiple antennas at the other end.

2.3.5.2 Advanced Channel Model for Link Level Simulations

This model represents the physical channel variations for large-scale movements of the MS for one link. In such scale of movements, the large-scale channel parameters (i.e., mean direction of arrival, mean direction of departure, number of paths, angular spread, delay spread, path loss, shadowing parameters, and relative mean power of multipath components) change considerably, so the assumption of having them fixed does not hold any more. To model the radio channel accurately, its dynamic has to be modeled carefully. This model can be thought of as a sequence of local stationary channels that the MS crosses along some trajectory for large-scale movements while keeping large-scale parameters changing continuously between the local areas. Such a model might be useful for investigating subspace-tracking algorithms in detail or for exploiting long-term effects on channel estimation or to improve Rx operation. This model may have limited applications in link level simulations.

2.3.5.3 Basic Channel Model for System Level Simulations

In this model multiple cells and multiple MS are considered. The basic channel for SL simulations is based on the “drop” concept in the context of 3GPP, normally in quasistatic system simulations. In this concept the successive channel responses and their parameters are independent. There is no time evolution of large-scale parameters within each “drop.” Parameters generated in each “drop” are independent from the previous one.

In this model set, the correlation of shadow fading is modeled for different base stations but not between different users. Path parameters of each ray (i.e., delay and angles) are fixed during each “drop” but small-scale fading is generated via Doppler effect by selecting the velocity vector of the MS and updating the channel, assuming that the delay, angles, and so forth are not changing. For each MS the delay spread and angular spread in addition to the shadow fading are correlated but there is no correlation between different MSs. The channels of the interferers will also be modeled. The basic channel for SL simulations that will be provided by WP5 is based on the SCM model and extended to the 100-MHz bandwidth at a 5-GHz frequency range with proper path-loss and shadowing models.

2.3.5.4 Advanced Channel Model for System Level Simulations

This model is suitable for dynamic system simulations. The main feature of this channel model is to model the dynamic features of the channel. In this model successive channel realizations or impulse responses are dependent.

2.3.5.5 Validation

Validation within the channel modeling effort occurs on two levels: the verification of the model description in the model documentation, and the verification of the software reference implementation. A model validation should verify the following:

- The models display backward compatibility. For example, MIMO models should retain SISO model characteristics when setting the antenna configuration to 1×1 . Similarly, they should match statistics of parameters in classic narrowband models.
- The models accurately reproduce a newly included propagation effect.
- The models show increased accuracy for a certain parameter or parameter extension to justify the increased complexity.

Validation is based on comparing model output channel realizations with that of measured channels. Independent parameters can be used to show improvement (e.g., eigenvalue statistics that are typically modeled only indirectly). A third level of validation is connected with the channel measurements and parameter extraction. Here, care has to be taken to verify the applicability and accuracy of parameter estimation and extraction algorithms.

2.4 Conclusions

The design of spectrally efficient air interfaces is a complex and multifaceted process. Propagation scenarios specify environment, mobility, frequency, and bandwidth for each scenario. Channel model requirements should be defined for LL and SL models. These requirements include path-loss, large-scale fading, small-scale fading, noise and interference, transitions, and antenna effects.

Capabilities for flexible spectrum use and sharing are widely seen to be a novel and important feature in the Beyond 3G communication systems promoting reconfigurability and scalability of the system.

Nomadic/local-area wireless access systems usually operate in noise-limited, short-range, limited mobility environments. In general, no cochannel interference has to be considered. Spectral efficiencies in the range of 5 to 10 bit/s Hz⁻¹ per cell have been seen as feasible for these systems. Considering ITU-R Recommendation M.1645 for a wireless access with target throughput up to approximately 1 Gbps, and considering the maximum possible future system spectral efficiency (10 bit/s Hz⁻¹ per cell) leads to the conclusion that the minimum needed bandwidth will be typically around 100 MHz.

Adding the spectrum demands for wide-area coverage (depending on deployment option, it may add up to two carriers of 20 and 40 MHz) and nomadic/local-area wireless access (i.e., a 100-MHz carrier) leads to a demand of 160 MHz per operator, without considering the needed guard bands.

The presented spectrum demand is only a rough estimate, and a reduction in spectrum demand may be expected if the performance of the radio interface can be enhanced. When the achievable radio interface throughput approaches the Shannon capacity bound, the cell throughput increases, which in turn results in lower spectrum demand. Spectrum demand is also heavily affected by the propagation path-loss and coverage requirements, which can be solved by introducing deployment concepts such as relaying. Adaptive transmission techniques provide a significant improvement of spectrum usage for data services by better utilization of the actually available C/I ratio at the physical layer. Several other methods also exist for the

reduction of cochannel interference, ranging from cell sectorization to interference cancellation methods at the receivers. These will be discussed in subsequent chapters in this book.

Multicarrier systems are known to offer good spectral occupancy properties, especially when the number of carriers is large: The *power spectrum density* is very steep. Those systems easily fit transmit spectral masks. To do so, band guards are set up; some carriers, on the edges of the spectrum, are not used. The required size of band guards to fit a particular spectrum mask is a point of comparison. The PAPR problem is another parameter that directly impacts the spectrum occupancy.

The mentioned techniques can be considered to be quite mature, although in some cases challenging from the complexity viewpoint. They are basically linked to the air interface development and definitely bounded by theoretical limits. However, due to the large amount of spectrum required, all means for improving the overall use of spectrum and possibly reducing spectrum demand must be researched.

References

- [1] International Telecommunication Union, www.itu.int.
- [2] FP6 IST Project Wireless Interface New Radio, WINNER and WINNER II, www.ist-winner.org.
- [3] FP6 IST Project My Personal Adaptive Global Net (MAGNET) and MAGNET Beyond, www.ist-magnet.org.
- [4] IST Project SURFACE, www.ist-surface.org/deliverables.htm.
- [5] Radio Communications Group, "Background on IMT-Advanced," Report IMT-ADV/1-E, March 2008, www.itu.int.
- [6] IST Project WINNER, "Intermediate Requirements per Scenario," Deliverable 1.2, February 2005.
- [7] Recommendation ITU-R M.1645, "Framework and Overall Objectives of the Future Development of IMT 2000 and Systems Beyond IMT 2000," www.itu.int.
- [8] IST Project WINNER, "Intermediate Concept Proposal and Evaluation," Concept Group WA, D6.13.1, www.ist-winner.org.
- [9] IST Project WINNER, "Identification of Radio-Link Technologies," Deliverable D2.1, June 2004, www.ist-winner.org.
- [10] IST Project WINNER, "Feasibility of Multibandwidth Transmissions" Deliverable D2.2, October 2004, www.ist-winner.org.
- [11] IST Project WINNER, "WINNER Spectrum Aspects: Methods for Efficient Sharing, Flexible Spectrum Use and Coexistence," Deliverable D 6.1, October 2004, www.ist-winner.org.
- [12] IST Project WINNER, "Identification of Radio Link Technologies," Deliverable D2.1, July 2004, www.ist-winner.org.
- [13] HIPERLAN Standards, www.etsi.org.
- [14] IEEE Standards, www.standards.ieee.org.
- [15] Berno, A., and N. Laurenti, "Time and Frequency Synchronization for HIPERLAN/2," *Proc. of the IFIP-TC6 Conference on Networking*, Pisa, Italy, 2002.
- [16] Jankiraman, M., and R. Prasad, "Algorithm Assisted Synchronization of OFDM Systems for Fading Channels," *Proc. of the 5th OFDM Workshop*, Hamburg, Germany, September 2000.

- [17] Sandell, M., J. J. van de Beek, and P. O. Börjesson, "Timing and Frequency Synchronization in OFDM Systems Using the Cyclic Prefix," *Proc. of the 1995 IEEE International Symposium on Synchronization*, Essen, Germany, December 1995.
- [18] Tufvesson, F., M. Faulkner, and O. Edfors, "Time and Frequency Synchronization for OFDM Using PN-Sequence Preambles," *Proc. of IEEE Vehicular Technology Conference*, Amsterdam, The Netherlands, September 1999.
- [19] Taoka, H., and K. Higuchi, "Fourth Generation Mobile Communication Using MIMO High-Speed Packet Transmission," *NTT Technical Review*, Vol. 6, No. 11, November 2008.
- [20] Taoka, H., K. Higuchi, and M. Sawahashi, "Field Experiments on Real-Time 1-Gbps High-Speed Packet Transmission in MIMO-OFDM Broadband Packet Radio Access," *Proc. of the 63rd IEEE VTC2006-Spring*, Melbourne, Australia, 2006, Vol. 4, pp. 1812–1816.
- [21] "Spectrum for Future Mobile and Wireless Communications," WWRF White Paper, December 2003.
- [22] Federal Communications Commission Web site on Cognitive Radio, www.fcc.gov/oet/cognitiveradio.
- [23] IST Project WINNER, "Feasibility of Multi-Band Transmissions," Deliverable D2.2, July 2004, www.ist-winner.org.
- [24] Prasad, R., and A. Mihovska, *New Horizons in Mobile and Wireless Communications, Volume 3: Reconfigurability*, Norwood, MA: Artech House, 2009.
- [25] Chu, D. C., "Polyphase Codes with Good Periodic Correlation Properties," *IEEE Trans. on Information Theory*, July 1972, pp. 531–532.
- [26] Huschke, J., et al., "Downlink Capacity of UMTS Coexisting with DVB-T MFNs and Regional SFNs," *Proc. of IEE Conference 'Getting the Most Out of the Radio Spectrum'*, London, October 2002.
- [27] Shared Spectrum Company, "Measurements," www.sharedspectrum.com.
- [28] "Definition of UWB Scenarios–Initial," IST-2003-506897 PULSERS, Deliverable D2a2, March 2004.
- [29] Rinne, M., et al., "Dual Bandwidth Approach to New Air Interface," *Proc. of 11th Wireless World Research Forum Meeting*, Oslo, Norway, June 2004.
- [30] Frank, R. L., and S. A. Zadoff, "Phase Shift Codes with Good Periodic Correlation Properties," *IRE Trans. on Information Theory*, October 1962, pp. 381–382.
- [31] Hijazi, S., et al., "Flexible Spectrum Use and Better Coexistence at the Physical Layer of Future Wireless Systems via a Multicarrier Platform," *IEEE Wireless Communications*, April 2004, pp. 64–71.
- [32] Dinis, R., et al., "A Multiple Access Scheme for the Uplink of Broadband Wireless Systems," *Proc. of Globecom 2004*, Dallas, TX, November 2004.
- [33] Le Floch, B., M. Alard, and C. Berrou, "Coded Orthogonal Frequency Division Multiplex," *Proc. of IEEE*, Vol. 83, No. 6, June 1995.
- [34] Lacroix, D., J. P. Javaudin, and N. Goudard, "IOTA, an Advanced OFDM Modulation for Future Broadband Physical Layers," *Proc. of 7th Wireless World Research Forum Meeting*, December 2002.
- [35] Pollet, T., M. Van Bladel, and M. Moeneclaey, "BER Sensitivity of OFDM Systems to Carrier Frequency Offset and Wiener Phase Noise," *IEEE Trans. on Communications*, Vol. 43, Nos. 2–4, February–April 1995, pp. 191–193.
- [36] Tonello, A. M., Laurenti, N., and S., Pupolin, "Analysis of the Uplink of an Asynchronous Multiuser DMT OFDMA System Impaired by Time Offset, Frequency Offset, and Multipath Fading," *Proc. of IEEE VTC 2000 Fall*, pp. 1094–1099.
- [37] Leus, G., et al., "Direct Semi-Blind Design of Serial Linear Equalizers for Doubly-Selective Channels," *Proc. of 2004 IEEE International Conference on Communications*, Vol. 5, June 20–24, 2004, pp. 2626–2630.

- [38] Armstrong, J., "Analysis of New and Existing Methods of Reducing Inter-carrier Interference Due to Carrier Frequency Offset in OFDM," *IEEE Trans. on Communications*, Vol. 47, No. 3, March 1999, pp. 365–369.
- [39] Muck, M., "A Pseudo Random Postfix OFDM Modulator and Inherent Channel Estimation Techniques," *Proc. of IEEE Globecom 2003*, San Francisco, CA, December 2003.
- [40] Muquet, B., et al., "Cyclic Prefixing or Zero Padding for Wireless Multicarrier Transmissions?" *IEEE Trans. on Communications*, Vol. 50, No. 12, December 2002, pp. 2136–2148.
- [41] Winters, J. H., "On the Capacity of Radio Communication Systems with Diversity in a Rayleigh Fading Environment," *IEEE J. on Selected Areas of Communications*, Vol. SAC-5, June 1987, pp. 871–878.
- [42] Foschini, G. J. "Layered Space-Time Architecture for Wireless Communication in a Fading Environment when Using Multielement Antennas," *Bell Labs Technical J.*, Autumn 1996, pp. 41–59.
- [43] Foschini, G., J., and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment When Using Multiple Antennas," *J. on Wireless Personal Communications*, Vol. 6, No. 3, March 1998, pp. 311–335.
- [44] Telatar, E., "Capacity of Multiantenna Gaussian Channels," AT&T Bell Laboratories Technical Memo, June 1995.
- [45] Golden, G. D., et al., "Detection Algorithm and Initial Laboratory Results Using the V-BLAST Space-Time Communication Architecture," *IEEE Electronic Letters*, Vol. 35, No. 1, 1999, pp. 14–15.
- [46] Damen, M. O., A. Chkeif, and J. C. Belfiore, "Lattice Codes Decoder for Space-Time Codes," *IEEE Communications Letters*, Vol. 4, May 2000, pp. 161–163.
- [47] Yao, H., and G. W. Wornell, "Lattice-Reduction-Aided Detectors for MIMO Communication Systems," *Proc. of IEEE Globecom 2002*, Taipei, Taiwan, November 2002.
- [48] Wuebben, D., et al., "Near-Maximum-Likelihood Detection of MIMO Systems Using R-MMSE-Based Lattice Reduction," *Proc. of IEEE International Conference on Communications (ICC'2004)*, Paris, France, June 2004.
- [49] Kim, K. J., and R. A. Iltis, "Joint Detection and Channel Estimation Algorithms for QS-CDMA Signals over Time-Varying Channels," *IEEE Trans. on Communications*, Vol. 50, May 2002.
- [50] Kim, K. J., "A New Joint Data Detection and Channel Estimation Algorithms for MIMO-OFDM Systems," Nokia Internal Report, 2002.
- [51] IST Project WINNER, "Assessment of Advanced Beamforming and MIMO Technologies," Deliverable 2.7, February 2005, www.ist-winner.org.
- [52] IST Project WINNER, "Preliminary WINNER System Concept," Deliverable 7.2, www.ist-winner.org.
- [53] "Universal Mobile Telecommunications System (UMTS); Selection Procedures for the Choice of Radio Transmission Technologies of the UMTS (UMTS 30.03, Version 3.2.0), TR 101 112 V 3.2.0 (1998).
- [54] Prasad, R., (ed.), *Towards the Wireless Information Society*, Vols. I and Vol. II, Norwood, MA: Artech House, 2006.
- [55] Louagie, F., L. Munoz, and S. Kyriazakos, "Paving the Way for Fourth Generation: A New Family of Wireless Personal Area Networks," *Proc. of IST Mobile Summit 2003*, Aveiro, Portugal, June 15–18, 2003.
- [56] FP6 Overview, <http://cordis.europa.eu/ist/so/mobile-wireless/home.html>.
- [57] Olsen, R. L., et al., "My Personal Adaptive Global NET—MAGNET," *Proc. of 8th Wireless World Research Forum Meeting*, Beijing, China, February 2004.
- [58] IST Project PULSERS, Document EOI-FP6-2002, www.pulsers.eu.

- [59] Hirt, W., and D. Porcino, "Pervasive Ultra-Wide Band Low Spectral Energy Radio Systems," White Paper, *Proc. of Wireless World Research Forum Meeting*, November 18, 2002.
- [60] Federal Communications Commission, "The First Report and Order," ET Docket No. 98-153, 2002.
- [61] European Commission, Article ID 197800214, www.europa.eu.
- [62] Federal Communications Commission, "Revision of Part 15 of the Commission's Rules Regarding Ultra-Wideband Transmission Systems," First Report and Order, ET Docket 98-153, FCC 02-48; Adopted February 14, 2002; Released April 22, 2002.
- [63] Ghavami, M., L. Michael, and R. Kohno, *Ultra Wideband Signals and Systems in Communication Engineering*, New York: John Wiley & Sons, 2004.
- [64] "Survey on Worldwide Research on 4G Air Interface Concepts," IST-2002 507039 Project 4MORE, Deliverable 8.2, June 2004, <http://4more.av.it.pt/docs/D8.2.pdf>.
- [65] IST Project MAGNET Beyond, "PHY/MAC Benchmarking of the Target MAGNET FM-UWB and MC-SS Air Interfaces," Deliverable 3.2.2, June 2007, www.ist-magnet.org.
- [66] Foerster, J., et al., "Ultra-Wideband Technology for Short- or Medium-Range Wireless Communications," *Intel Technology J.*, May 2001.
- [67] Khun-Jush, J., et al., "HIPERLAN TYPE 2 for Broadband Wireless Communications," www.ericsson.com/ericsson/corpinfo/publications/review/2000_02/109.shtml.
- [68] Omiyi, P. E., and H. Haas, "Maximising Spectral Efficiency in 4th Generation OFDM/TDMA TDD Hybrid Cellular Mobile/Ad-Hoc Wireless Communications," *Proc. of IEEE VTC 2004*, Milan, Italy, May 2004.
- [69] IST Project MAGNET Beyond, "Prototype Specification of Antenna and Radio Front-End Schemes for PAN Devices," Deliverable 3.1.1, June 2006, www.ist-magnet.org.
- [70] IST Project 507102 MAGNET, "PAN Channel Characterisation (Part I)," Deliverable D.3.1.2a October 2004, www.ist-magnet.org.
- [71] IST Project 507102 MAGNET, "PAN Channel Characterisation (Part II)," Deliverable D.3.1.2a, June 2005, www.ist-magnet.org.
- [72] Boelskei, H., D. Gesbert, and A. Paulraj, "On the Capacity of Wireless Systems Employing OFDM-based Spatial Multiplexing," *IEEE Trans. on Communication*, February 2002, pp. 225–234.
- [73] Malik, W., et al., "Performance Analysis of Ultra-Wideband Spatial MIMO Communications Systems," *Proc. of the IST Mobile Summit 2006*, Dresden, Germany, June 2005.
- [74] IST Project PULSERS, "Study of MIMO VHDR MBO Specific System Concepts," Deliverable 2a, www.pulsers.eu.
- [75] Schumacher, L., "WLAN MIMO Channel Matlab Program," www.info.fundp.ac.be/~lsc/Research/IEEE_80211-HTSG_CMSC/distribution_terms.html.
- [76] Özcelik, H., et al., "Deficiencies of Kronecker MIMO Radio Channel Model," *IEE Electronics Letters*, Vol. 39, No. 16, August 2003.
- [77] "IEEE P802.11 Wireless LANs, TGn Channel Models," IEEE 802.11-03/940r2, January 9, 2004.
- [78] Schumacher, L., K. I. Pedersen, and P. E. Mogensen, "From Antenna Spacings to Theoretical Capacities—Guidelines for Simulating MIMO Systems," *Proc. of PIMRC'02*, September 2002, Vol. 2, pp. 587–592.
- [79] "Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations, Release 6," 3GPP TR25.996 V6.1.0 (2003-09), 2003.
- [80] Algans, A., K. I. Pedersen, and P. E. Mogensen, "Experimental Analysis of the Joint Statistical Properties of Azimuth Spread, Delay Spread, and Shadow Fading," *IEEE J. on Selected Areas in Communications*, Vol. 20, No. 3, April 2002, pp. 523–531.
- [81] IST Project WINNER, "Selection of Channel Models," Deliverable 5.1. February 2005, www.ist-winner.org.

- [82] Prasad, R., *Towards a Global 3G System*, Vol. I, Norwood, MA: Artech House, 2004.
- [83] Schumacher, L., et al., "MIMO Channel Characterization," D2, Version 1.1, IST-1999-11729 METRA, February 2001, www.metra.org.
- [84] Commission of the European Communities, "Digital Land Mobile Radio Communications-CPST 207 Final Report," ECSC-EEC-EAEC, 1989.
- [85] Saleh, A., and R., Valenzuela, "A Statistical Model for Indoor Multipath Propagation," *IEEE J. on Selected Areas in Communications*, Vol. 5, No. 2, February 1987, pp. 128–137.
- [86] Kogiantis, A., and A. Moustakas, "Channel Modeling for MBWA," IEEE C802.20-03/42, May 12, 2003.
- [87] Correia, L. M., "Wireless Flexible Personalised Communications," in *COST259: European Co-operation in Mobile Radio Research*, New York: John Wiley & Sons, 2001.
- [88] "Selection Procedures for the Choice of Radio Transmission Technologies of the UMTS," ETSI UMTS 30.03 Technical Report, 1998.
- [89] 3GPP Technical Specification Group GSM/EDGE, "Radio Access Network; Radio Transmission and Reception," (Release 1999), 3GPP TS 05.05 V8.16.0 (2003-08), www.3gpp.org/ftp/Specs/html-info/0505.htm.
- [90] Özcelik, H., et al., "Deficiencies of Kronecker MIMO Radio Channel Model," *IEE Electronics Letters*, Vol. 39, No. 16, August 2003.
- [91] Herdin, M., et al., "Variation of Measured Indoor MIMO Capacity with Receive Direction and Position at 5.2 GHz," *IEEE Electronics Letters*, Vol. 38, No. 21, October 2002, pp. 1283–1285.
- [92] Molisch, A., "A Generic Model for MIMO Wireless Propagation Channels in Macro- and Microcells," *IEEE Trans. on Signal Processing*, Vol. 52, No. 1 January 2004.
- [93] Vaughan, R., and J. B. Andersen, "Channel Propagation and Antennas for Mobile Communications," *IEE Electromagnetic Waves Series*, No. 50, 2003.
- [94] Richter, A., M. Landmann, and R. S. Thomä, "On the Relevance of Dense Multipath Components in a Micro Cell Scenario," COST 273 TD(04)137, June 2004.
- [95] Richter, A., "Parametric Modeling and Estimation of Distributed Diffuse Scattering Components of Radio Channels," COST 273 TD(03)198, September 2003.
- [96] Thoma, R. S., et al., "MIMO Vector Channel Sounder Measurement for Smart Antenna System Evaluation," *European Trans. on Telecommunications*, Special Issue on Smart Antennas, Vol. 12, No. 5, September–October 2001, pp. 427–438.
- [97] Weichselberger, W., "Spatial Structure of Multiple Antenna Radio Channels, A Signal Processing Viewpoint," Technical University of Wein, 2003.
- [98] Kermoal, J., P., et al., "A Stochastic MIMO Radio Channel Model with Experimental Validation," *IEEE J. on Selected Areas in Communications*, Vol. 20, No. 6, August 2002, pp. 1211–1226.
- [99] Yu, K., et al., "Second-Order Statistics of NLOS Indoor MIMO Channels Based on 5.2-GHz Measurements," *Proc. of the Global Telecommunications Conference (GLOBECOM)*, November 2001, Vol. 1, pp. 156–160.
- [100] Özcelik, H., et al., "Deficiencies of 'Kronecker' MIMO Radio Channel Model," *IEE Electronics Letters*, Vol. 39, No. 16, August 2003.
- [101] Algans, A., K. I. Pedersen, and P. E. Mogensen, "Experimental Analysis of the Joint Statistical Properties of Azimuth Spread, Delay Spread, and Shadow Fading," *IEEE Journal on Selected Areas in Communication*, Vol. 20, No. 3, April 2002.
- [102] IST Project WINNER, "Determination of Propagation Scenarios," Deliverable 5.2, June 2004, at <http://www.ist-winner.org>.
- [103] Alexiou, A. "Link to System Interface Methodology," WINNER Internal Document, May 2004.
- [104] Holma, H., "A Study of UMTS Terrestrial Radio Access Performance," Ph.D. Thesis, Helsinki University of Technology, 2003.

- [105] Bello, P. A., "Characterization of Randomly Time-Variant Linear Channels," *IEEE Trans. on Communications Systems*, Vol. CS-11, December 1963, pp. 360–393.
- [106] Clarke, R., H., "A Statistical Theory of Mobile-Radio Reception," *Bell System Technical J.*, Vol. 47, No. 6, July–August 1968, pp. 957–000.
- [107] Jakes, W. C., (ed.), *Microwave Mobile Communications*, New York: John Wiley & Sons, 1974.
- [108] Rappaport, T. S., S. Y. Seidel, and K. Takamizawa, "Statistical Channel Impulse Response Models for Factory and Open Plan Building Radio Communication System Design," *IEEE Trans. on Communications*, Vol. 39, No. 5, May 1991, pp. 794–807.
- [109] Hashemi, H., "Impulse Response Modeling of Indoor Radio Propagation Channels," *IEEE J. on Selected Areas in Communications*, Vol. 11, No. 7, September 1993, pp. 967–978.
- [110] Ertel, R. B., et al., "Overview of Spatial Channel Models for Antenna Array Communication Systems," *IEEE Personal Communications*, Vol. 5, No. 1, February 1998, pp. 10–22.
- [111] Foschini, G. J., and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment," *Wireless Personal Communications*, Vol. 6, 1998, pp. 311–335.
- [112] Raleigh, G. G., and J. M. Ciofi, "Spatio-Temporal Coding for Wireless Communication," *IEEE Trans. on Communications*, Vol. 46, No. 3, March 1998, pp. 357–366.
- [113] Bach Andersen, J., "Array Gain and Capacity for Known Random Channels with Multiple Element Arrays at Both Ends," *IEEE Journal on Selected Areas in Communication*, November 2000, pp. 2172–2178.
- [114] Farrokhi, F. R., et al., "Link Optimal Space-Time Processing with Multiple Transmit and Receive Antennas," *IEEE Communications Letters*, March 2001, pp. 85–87.
- [115] Ivrlac, M. T., et al., "Efficient Use of Fading Correlations in MIMO Systems," *Proc. IEEE Vehicular Technology Conference*, October 2001.
- [116] Wong, K. K., et al., "Adaptive Spatial Subcarrier Trellis Coded MQAM and Power Optimization for OFDM Transmission," *Proc. IEEE Vehicular Technology Conference*, May 2000.
- [117] Yu, K., and B. Ottersten, "Models for MIMO Propagation Channels: A Review," Special Issue on Adaptive Antennas and MIMO Systems, *J. on Wireless Communications and Mobile Computing*, Vol. 2, No. 7, November 2002, pp. 553–666.
- [118] Abdi, A., and M. Kaveh "A Versatile Spatio-Temporal Correlation Function for Mobile Fading Channels with Non-Isotropic Scattering" *Proc. of the IEEE Workshop on Statistical Signal Array Processing*, 2000, pp. 58–62.
- [119] Gesbert, D., et al., "Outdoor MIMO Wireless Channels: Models and Performance Prediction," *IEEE Trans. on Communications*, December 2002.
- [120] Wallace, J. W., and M. A. Jensen. "Modeling the Indoor MIMO Wireless Channel," *IEEE Trans. on Antennas and Propagation*, Vol. 50, No. 5, May 2002, pp. 591–599.
- [121] Molisch, A. F., "A Generic Model for MIMO Wireless Propagation Channels," *Proc. of the IEEE International Conference on Communications (ICC'2002)*, Vol. 1, 2002, pp. 277–282.
- [122] Svantesson, T., "A Physical MIMO Radio Channel Model for Multi-Element Multi-Polarized Antenna Systems," *Proc. IEEE Vehicular Technology Conference Fall*, 2001, Vol. 2, pp. 1083–1087.
- [123] Stege, M., et al., "A Multiple Input-Multiple Output Channel Model for Simulation of Tx- and Rx-Diversity Wireless Systems," *Proc. of the IEEE Vehicular Technology Conference Fall*, 2000, Vol. 2, pp. 833–839.
- [124] Sayeed, A. M., "Modeling and Capacity of Realistic Spatial MIMO Channels," *Proc. of the 2001 International Conference on Acoustics, Speech and Signal Processing (ICASSP 2001)*, May 2001.
- [125] "Spatial Channel Model for Multiple Input Multiple Output (MIMO) Simulations, Release 6," 3GPP TR25.996 V6.1.0 (2003-09).

- [126] “IEEE P802.11 Wireless LANs, TGn Channel Models,” IEEE 802.11-03/940r2, January 9, 2004.
- [127] Medbo, J., and P. Schramm, “Channel Models for HIPERLAN/2,” ETSI/BRAN 3ERI085B, 1998, www.etsi.org.
- [128] Medbo, J., and J.-E. Berg, “Measured Radiowave Propagation Characteristics at 5 GHz for Typical HIPERLAN/2 Scenarios,” ETSI/BRAN 3ERI084A, 1998, www.etsi.org.
- [129] Correia, L. M. (Ed.), “Wireless Flexible Personalised Communications,” in *COST 259: European Co-operation in Mobile Radio Research*, New York: John Wiley & Sons, 2001.
- [130] Xu, H., et al., “A Generalized Space-Time Multiple-Input Multiple-Output (MIMO) Channel Model,” *IEEE Trans. on Wireless Communications*, Vol. 3, No. 3, May 2004, pp. 966–975.
- [131] Keenen, A. J., and M. J. Motley, “Radio Coverage in Buildings,” *British Telecom Technical J.*, Vol. 8, No. 1, January 1990, pp. 19–24.
- [132] Rhodes, V. J., “Path Loss Proposal for the IEEE 802.11 HTSG Channel Model Ad Hoc Group,” April 22, 2003, www.ieee.org/802.11.
- [133] Gudmundson, M., “Correlation Model for Shadow Fading in Mobile Radio Systems,” *Electronics Letters*, Vol. 27, No. 23, November 7, 1991.
- [134] Fleury, B. H., “First- and Second-Order Characterization of Direction Dispersion and Space Selectivity in the Radio Channel,” *IEEE Trans. on Information Theory*, Vol. 46, No. 6, September 2000, pp. 2027–2044.

Coding and Modulation

Transmission over a wireless channel is generally subject to strong degradations such as thermal background noise, multipath fading, and noise in RF front ends. To enable reliable transmission of data and speech, therefore, it is necessary to introduce techniques that allow for the efficient detection and correction of errors due to a channel outage. *Channel coding* is a procedure used to find codes with good error correction performance that at the same time are adaptable to channel conditions and have a reasonably low encoding/decoding complexity such that they can be included in wireless handheld devices.

Modulation is used to transform the transmitted data into a format suitable for transmission using a carrier frequency. Modulation methods incorporate different approaches of embedding information into channel waveforms. Correspondingly, *demodulation* methods are utilized to extract the information from received waveforms that have been propagated through the channel.

This chapter describes the different coding and modulation methods used in the design of next generation radio interfaces within the EU that were funded under the FP6 R&D effort [1]. Major work was performed by the IST projects WINNER and WINNER II [2], MAGNET and MAGNET Beyond [3], SURFACE [4], PULSERS [5], E2R and E2R II [6], and the IST Project NEWCOM [7], to mention a few.

This chapter is organized as follows. Section 3.1 gives the state of the art in the analysis and design of coding and modulation for next generation systems. Comparative analyses and results regarding the advantages and disadvantages of parallel concatenated and low-density parity-check codes that were obtained by work performed within the scope of the IST project WINNER are given. In a similar manner, the advantages and disadvantages of different modulation schemes are treated with respect to their eligibility for a next generation air interface.

Section 3.2 describes the design methodology of multiuser space-time codes. Work performed for the IST project MASCOT led to major achievements in this area. The methodology for the characterization of the error event regions is given, as well as code design criteria.

Section 3.3 describes contributions in the area of advanced modulation and coding schemes, achieved during the final design phase of the WINNER air interface. These contributions are insightful about the approach toward the selection of the most suitable schemes when considering the targeted performance requirements for next generation radio interfaces. Section 3.4 concludes the chapter.

3.1 Introduction

Major advances in reliable communications research have had a direct impact on the volume and quality of services offered in all wireless systems. The signal processing subsystems, on which wireless communications are built, are key to these advances [8]. Despite notable recent gains in coding theory, multiple-access techniques, algorithm design, and processing efficiency, the ever-increasing demand for higher data rates pushes current generation designs to their practical limits. To ensure that future systems will not be confined to permanent saturation, major cooperative research efforts into the fundamental capabilities and limitations of wireless systems must be pursued, addressing the present knowledge gaps in a number of areas: information-theoretic capacity limits in time-varying multiuser channels, short packet code design for two-way communications, optimal resource allocation strategies under QoS constraints, and increased mobility requirements, to name just a few. Although the basic signal processing techniques exploited in wireless systems have become increasingly mature and specialized during the past few decades, joint optimization of the various signal processing subsystems is a necessary ingredient in solutions that are to meet future demands in wireless communications [8].

3.1.1 Analysis and Design of Coding Algorithms for Next Generation Systems

To keep up with the constantly increasing demand for higher and higher bit rates with better and better QoS, future wireless communications will have to be designed to approach the capacity of the propagation channel when any kind of dispersion effects in the radio channel, such as dispersion in delay, direction at the transmitter and at the receiver, and Doppler frequency, are exploited by the system [7]. Theoretical works have shown that by deploying multiple-element antennas at both the transmitter and the receiver of a communication system (i.e., by exploiting the dispersion in direction at both transmitter and receiver sites), a tremendous increase in capacity can be achieved compared to systems equipped with single antennas [2, 4, 9]. However, recent works have also shown that the capacity of these MIMO systems strongly depends on the underlying propagation conditions [10]. For instance, in the so-called keyhole channels, which are interesting because they reveal how well a system performs in conditions different than the Rayleigh channel, no gain in capacity can be obtained by using multiple-element antennas [10]. Results strongly point out the need to develop a better understanding of the relationship between the underlying propagation conditions and the capacity of the propagation channel and, more generally, the features of the channel critical to communications systems operating in it.

3.1.1.1 Introduction to Channel Coding

Channel coding adds redundancy to the information stream, which can be used later to detect or correct errors due to transmission distortion. This redundancy can be determined by using two different classes. The first class comprises block codes, where independent blocks (codewords) of symbols are constructed with a

constant mapping between information and redundancy. The second method is convolutional encoding, where the redundancy is continuously calculated for the code sequence using a combination (convolution) of successive information symbols. Also, convolutional codes are possible for short block length services.

In his “Mathematical Theory of Communications,” Claude E. Shannon showed in 1948 [11] that there existed codes that would permit transmission across noisy channels with an arbitrarily small probability of error, as long as the information data rate was inferior to the capacity of the channel (i.e., if an unlimited codeword length was allowed). Shannon, however, did not point out how to find a code that will allow the error-free transmission [12]. Since then, a lot of research has been done to find codes, which perform close to the Shannon theoretical limit.

The basic theory of block codes was developed by Golay [13] and Hamming [14] with the introduction of the binary multiple-error correcting Golay code and the one-error correcting Hamming codes, respectively. Other important binary block codes are the BCH codes [15, 16], the Reed-Muller codes [17, 18], and the quadratic residue codes [19, 20]. The low-density, parity-check (LDPC) codes were introduced by Gallager in 1963 [21] and belong to the class of binary block codes. The most famous nonbinary block codes are the Reed-Solomon (RS) codes, invented by Reed and Solomon in 1960 [22]. These codes are still used in many practical systems (e.g., in storage devices such as compact discs or digital versatile discs, wireless or mobile communications, satellite communications, digital television, and high-speed modems.) RS codes are used primarily as concatenated codes together with convolutional codes. This is done, for example, in the well-known ESA/NASA standard for satellite data transmission.

In addition to block codes, convolutional codes are the other major class of codes for error correction [23]. Convolutional codes are based on a linear mapping of a set of information words to a set of codeword words. Conceptually, information and codewords are of infinite length. For practical applications, codewords with finite lengths are used almost exclusively. Therefore termination, truncation, or tail-biting is necessary. The existence of a maximum-likelihood (ML) decoding procedure that can be implemented with reasonable complexity is the reason for their widespread use [24]. Reference [25] demonstrated later that this algorithm performs an ML decoding of convolutional codes. Another important fact is that soft input can be used and soft output can be created when decoding convolutional codes. This can be done using either the Bahl, Cocke, Jelinek, and Raviv (BCJR) algorithm [26] or the SOVA algorithm from [27].

Due to the possibility of a SISO decoding technique, convolutional codes can be used as component codes for many different concatenated codes. This is done for the parallel-concatenated convolutional codes, the so-called “turbo” codes [28, 29], and for serially concatenated convolutional codes, the so-called Woven codes [12]. A concatenation between convolutional codes and block codes can be constructed.

The advantages of convolutional codes are that they provide very good performance for short block lengths, they provide very easy encoding using only shift registers, and they use practical ML decoding algorithms. Unfortunately, for larger block lengths the error correction performance of convolutional codes is signifi-

cantly lower than that of turbo codes or LDPC codes because the minimum distance of convolutional codes is quite low [12].

The field of error-correcting codes has been subject to a new revolution in the last decade with the emergence of two classes of powerful error-correcting codes: the parallel-concatenated codes (e.g., turbo codes) and the LDPC codes. Turbo codes are constructed from the concatenation of two error-correcting codes, separated by an interleaver, and their emergence represents one of the most significant breakthroughs in the coding area. The two families of advanced error-correcting codes were studied in detail in [12]. The same reference describes other classes of codes (e.g., Woven codes, serially concatenated codes, and repeat accumulate codes).

When mapping a codeword onto a set of transmission resources with differing signal-to-interference ratios, a tool is needed for predicting/estimating the resulting codeword error rate from the set of SINRs. Such a tool can be used for several purposes, including these two:

- As link-to-system interfaces in system level simulators, which use these predicted codeword error rates and do not need to implement the complete decoding;
- As a component of a link adaptation scheme, in which a codeword is mapped onto a set of chunk layers, with varying individual link adaptations and SINRs [30].

Conventional turbo codes and LDPC codes offer impressive error correction capabilities for very long blocks, but such block lengths impose latency delays that prove unacceptable for two-way communications or other mobile applications. The error-correction capabilities for more modest block lengths are more nuanced: Certain configurations work well in controlled laboratory conditions, but robustness issues once such codes are exposed to hostile real-world environments still present many open questions. Extending the gains of high-performance codes to moderate block lengths is of interest in all mobile and interactive applications. Extending the performance advantages of iterative decoding to interconnected receiver components has previously been evidenced by turbo equalizers, turbo synchronizers, turbo detectors, and so forth [8].

Turbo codes are now a mature technology, incorporated into several standards and into many commercial systems [12]. Their high flexibility enables an easy adjustment to all kinds of code rates and block sizes. During the past 10 years, intensive research efforts have focused on enhancements of these codes, simplification of the decoding algorithms, and on the related techniques: turbo-coded modulations, turbo multiuser detection (MUD), and so forth. The resulting advances alleviate their main drawbacks: The complexity of the decoding algorithm can be lowered thanks to suboptimal versions without scarifying performance degradation. The asymptotic performance has been improved to moderate and even eliminated the so-called error floor. For these reasons, turbo codes are considered likely candidates for future air interfaces.

LDPC codes are the main competitors of turbo codes because of their low complexity and good performance. LDPC codes were shown to be capable of

reaching the channel capacity very closely at low complexity because of their iterative decoding algorithms (or message-passing decoders) [21].

The performance of LDPC codes can be improved by using irregular sparse parity-check matrices. To design good irregular codes, the so-called *density evolution* was introduced in [31]. Reference [32] showed that a carefully designed irregular LDPC code of rate 1/2 can perform within 0.0045 dB of the Shannon limit by transmission over an additive wide noise Gaussian channel (AWGN) channel.

A random construction of the parity-check matrix does not necessarily result in a code with “good” performance because of small cycles in the Tanner graph [33] representation of the parity-check matrix. Reference [21] had already introduced a construction method to avoid such small cycles; [21] shows that the minimum distance of LDPC codes could increase linearly with the length of the code when the number of nonzero elements in each column j is at least three. Many other research groups worldwide investigated different construction methods for the parity-check matrices of LDPC codes without small cycles. It was shown that irregular LDPC codes in particular have a very poor performance when the parity-check matrices are constructed randomly. Therefore, the so-called progressive edge-growth (PEG) algorithm was introduced, which is a randomly based construction method for regular and irregular LDPC codes without small cycles [34].

3.1.1.1.1 Encoding LDPC Codes by Parity Check Matrix.

Although the decoding complexity of an LDPC code under message-passing decoding increases linearly with the codeword length n because of the sparseness of the parity-check matrix, the encoding complexity increases quadratically with n for conventional encoders based on a (nonsparse) generator matrix [12]. Also, the memory to store the generator matrix increases quadratically with n and, therefore, an implementation of an encoder for a long LDPC code needs a huge amount of memory. Reference [12] described a straight-forward method for encoding LDPC codes by the parity-check matrix. When the rows in the parity-check matrix are linearly independent, an equivalent code can always be found, where Gaussian elimination can bring the parity-check matrix into a lower triangular form. This is shown in Figure 3.1.

By definition the associated code consists of the set of n -tuples \mathbf{c} , so that $\mathbf{H}\mathbf{c}^T = \mathbf{0}^T$. We can split the codeword vector \mathbf{c} into a systematic part \mathbf{s} and a parity part \mathbf{p} , so that $\mathbf{c} = (\mathbf{s}, \mathbf{p})$. Then a systematic encoder is constructed as follows:

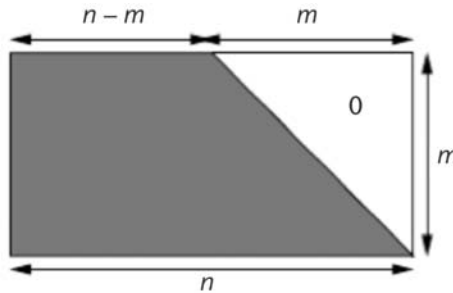


Figure 3.1 An equivalent parity-check matrix in lower triangular form [12].

1. Fill \mathbf{s} with the $(n - m)$ desired information symbols.
2. Determine the m parity-check symbols using back-substitution. More precisely, for $l \in \{0, \dots, m - 1\}$, calculate (3.1):

$$p_l = \sum_{j=1}^{n-m} H_{l,j} s_j + \sum_{j=1}^{l-1} H_{l,j+n-m} p_j \quad (3.1)$$

The parity-check matrix loses its sparseness, which is indicated by the dark area in Figure 3.1, which is caused by the use of Gaussian elimination. Hence, the encoding time and the memory used increase quadratically with the codeword length n . An efficient encoder would change the parity-check matrix into a lower triangular form by keeping the sparseness of the matrix. This is not possible for most LDPC codes.

An algorithm described in [35] tries to minimize g and can change the parity-check matrix into an approximate lower triangular form by keeping the sparseness of the matrix. For the remaining g rows, Gaussian elimination can be used to get the parity-check matrix shown in Figure 3.2.

This parity-check matrix can be used for the efficient encoding of the corresponding LDPC code as follows. Again the codeword vector \mathbf{c} is split into a systematic part \mathbf{s} , a first parity part \mathbf{g} , and a second parity part \mathbf{p} , so that $\mathbf{c} = (\mathbf{s}, \mathbf{g}, \mathbf{p})$. Then the code bits can be calculated in the following way:

1. Fill \mathbf{s} with the $(n - m)$ desired information symbols.
2. Determine the g parity-check symbols using back-substitution. More precisely, for $l \in \{0, \dots, g - 1\}$, calculate (3.2):

$$g_l = \sum_{j=1}^{n-m} H_{l,j} s_j + \sum_{j=l-1}^1 H_{l,j+n-m} g_j \quad (3.2)$$

3. For $l \in \{0, \dots, m - g - 1\}$ calculate by back-substitution:

$$p_l = \sum_{j=1}^{n-m} H_{l,j} s_j + \sum_{j=n-m+1}^{n-m-g} H_{l,j} g_j + \sum_{j=1}^{l-1} H_{l,j+n-m+g} p_j \quad (3.3)$$

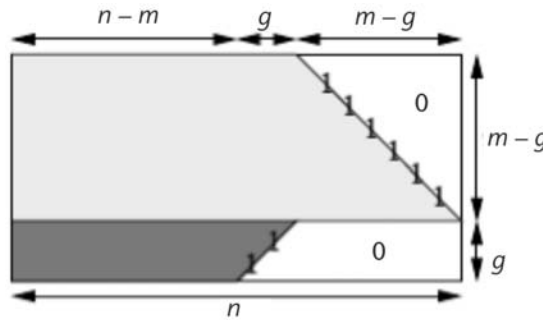


Figure 3.2 Parity-check matrix used for encoding [12].

The upper part of the matrix is still sparse and, therefore, the complexity of encoding increases only linearly with n . Only the lower part—the last g rows—is not sparse and, therefore, the encoding complexity of the g bits increases quadratically with n .

It was shown by simulation and also theoretically in [35] that for good irregular codes the number of dense rows in the parity-check matrix g is normally equal to one, which can be reduced to zero by deleting one nonzero element in a column of the parity-check matrix with two nonzero elements. This results in nearly no performance loss.

Another approach to reduce the complexity of encoding is done by means of the algebraic or geometric construction methods of LDPC codes [12]. Note that these construction methods are hardly flexible in choosing the rate or the number of ones in the rows and the columns.

3.1.1.1.2 Turbo Encoding.

A turbo encoder relies on the concatenation in parallel of two recursive systematic convolutional (RSC) encoders separated by an interleaver. This is shown in Figure 3.3. The description here is restricted to the concatenation of two codes. The parallel nature is related to the fact that the two constituent encoders operate simultaneously on the same set of input bits (in their natural order for the first encoder and in a permuted order for the second one) rather than performing the encoding processes one after the other. Because the encoders are systematic, the first output of the first RSC encoder is the information data stream. The second output corresponds to the parity information produced by the first RSC encoder. Regarding the second RSC encoder, only the parity information generated by this encoder is emitted, corresponding to the third output of the turbo encoder. Its natural code rate is therefore one-third. Puncturing patterns can be applied to the parity bits, in order to provide different coding rates.

The distance spectrum interpretation and the union bound performance analysis over AWGN and fading channels in [29, 36, 37] show the importance of multiplici-

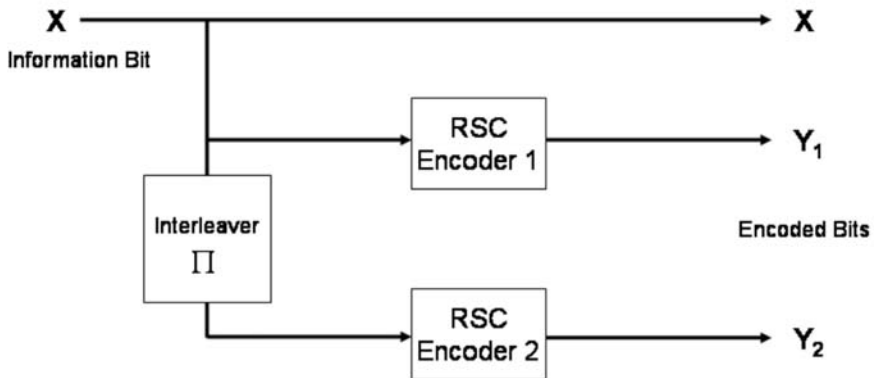


Figure 3.3 A turbo encoder [12].

ties and error coefficients in the turbo codes' distance spectrum. The interleaver maps an information sequence generating low-weight codewords from the first convolutional code with high probability into a sequence generating high-weight codewords from the second convolutional code and thus plays a key role. Very large information interleavers increase such probability, resulting in a distance spectrum with a few low-weight codewords and a large number of average-weight codewords similar to the distance spectrum of random-like codes. The *free distance* and the first several spectral lines above it are solely determined by the weight 2 information sequences. The minimum weight among constituent codes codewords caused by the weight 2 information sequences is denoted as the *effective free distance*. The maximization of the *effective free distance* was the main design criteria for constituent code optimization in [29, 38, 39], which was different from the optimization criteria used for stand-alone convolutional codes. In [40], the optimization considered information sequences for weights up to 6, whereas in [41] constituent codes were designed to optimize the distance spectrum of the turbo codes in the area of low to medium distances (i.e., significant spectral lines).

3.1.1.1.3 LDPC Decoding.

The decoding algorithms have the task of finding the codeword with minimum Euclidean distance from the received signal. The different algorithms used for LDPC decoding (belief propagation, sum-product decoding, and variations on the topic) all iteratively approximate the ML solution of this problem. Given that the code itself is “good,” that is, has good distance properties, these decoding algorithms allow for virtually error-free transmission at SNRs almost arbitrarily close to the Shannon bound [12]. In order for a code to achieve this performance, however, the required block length is very large. Appropriate bounds for limited block lengths were derived in [42].

LDPC decoding is done by exchanging messages on the edges between the nodes of the bipartite Tanner graph, an equivalent representation of the parity-check matrix \mathbf{H} of the LDPC code (see Figure 3.4).

The Tanner graph of an LDPC code contains two types of nodes: the variable nodes $v_i, i = 1, \dots, n$ on the left-hand side, representing the bits of the codeword, and the check nodes $c_j, j = 1, \dots, m$ on the right-hand side, representing the constraints imposed by the parity checks. Two nodes v_i and c_j are connected if and only if the parity-check matrix \mathbf{H} contains a nonzero entry at position (i, j) , that is, the variable v_i is checked by the check c_j .

At the outset of the decoding process, each variable node is initialized with the corresponding soft output $f_i^{(a)}$ from the channel detector, where $f_i^{(a)}$ denotes the a posteriori probability that v_i is in state a , conditioned on the observed channel output. Decoding is done by consecutively passing messages from one side of the graph to the other. During the first half-iteration, each variable node v_i sends messages $Q_{v_i, c_j}^{(a)}$ —its “belief” of being in state a given the input from all adjacent check nodes except c_j to check node c_j . During the second half-iteration, each check node c_j sends the messages $R_{c_j, v_i}^{(a)}$, representing its probability of being satisfied, given v_i is in state a and taking into account the beliefs of all other adjacent variable nodes except v_i , to variable node v_i .

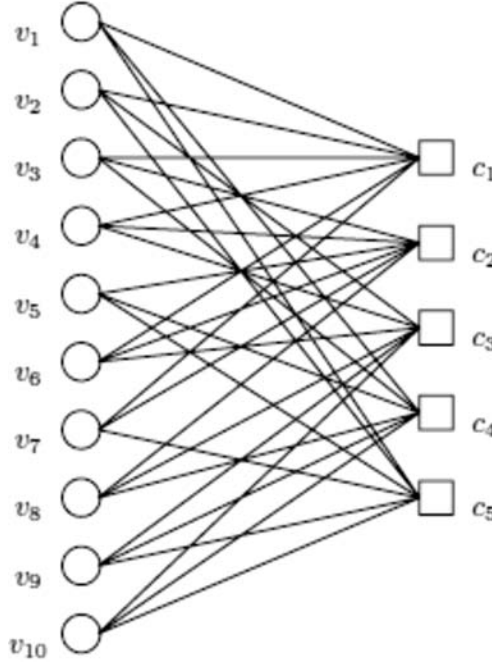


Figure 3.4 Tanner graph representation of a block length 10 (3, 6) regular LDPC code [12].

The probability domain messages $Q_{v_i, c_j}^{(a)}$ and $R_{c_j, v_i}^{(a)}$ for a binary LDPC are defined as follows:

$$Q_{v_i, c_j}^{(a)} = k^{(a)} f_i^{(a)} \prod_{c_k \in M(v_i), c_k \neq c_j} R_{c_k, v_i}^{(a)}, \quad \text{where } k^{(a)} = \left(\sum_a Q_{v_i, c_j}^{(a)} \right)^{-1} \quad (3.4)$$

and

$$R_{c_j, v_i}^{(0)} = \frac{1}{2} + \frac{1}{2} \prod_{v_k \in M(c_j), v_k \neq v_i} \left(1 - 2Q_{v_k, c_j}^{(1)} \right), \quad R_{c_j, v_i}^{(1)} = 1 - R_{c_j, v_i}^{(0)} \quad (3.5)$$

These calculations may equivalently be defined for the log domain, transforming multiplications and divisions into additions and subtractions, which is usually desirable for implementation purposes (in terms of processor cycles as well as the required precision).

This process is repeated until all checks are satisfied or a maximum number of iterations is reached. The first case corresponds to successful decoding, whereas in the latter case, a decoding error is declared. It is important to state that for sufficiently large block lengths, the number of undetected decoding errors is almost negligible. Hence, one may relax the requirements for an additional error detection code (CRC). However, the number of decoder iterations is not fixed, in contrast to turbo decoding. In fact, the *average* number of decoder iterations is usually quite low in comparison to the *maximum* number of iterations. This variable

decoding delay is one of the main disadvantages of LDPC decoders for real-time implementations, because it requires large buffers as every decoder will be designed for the average throughput and needs to store information for the next blocks whenever the current block has a large decoding delay.

The implementation complexity of LDPC decoders is comparatively low, because every variable/check node only needs to perform simple arithmetic operations on a very small set of values (i.e., the number of connecting edges of the node). Because all nodes in one half-iteration operate independently, the algorithm lends itself to use in parallel architectures. In contrast to turbo decoders, where implementation of the BCJR algorithm on parallel architectures is far from being trivial, this offers a good advantage. However, the implementation of the connection grid (also known as the *edge interleaver*) of the LDPC decoder can be a problem on parallel architectures, because it may create collisions in memory access [43].

The decoding complexity decreases with a rising SNR as the message-passing process converges faster to a final solution. Similar properties are known for sequential decoders and sphere detectors [12].

3.1.1.1.4 Turbo Decoding.

Optimum decoding of turbo codes is impossible because the overall *trellis* is time varying and has a number of states, which grow exponentially with the size of the interleaver [12]. Turbo codes are consequently decoded through an iterative technique. The turbo decoder consists of the two component decoders, concatenated in parallel and separated by the same interleaver as at the encoder side. The component codes are decoded iteratively, each one after the other, using SISO algorithms. These algorithms produce reliability information for the input sequences, which are then exchanged between the two decoders iteratively. This process is shown in Figure 3.5.

The first decoder operates on the received sequence corresponding to the systematic part of the stream (x_k) and to the parity information from the first

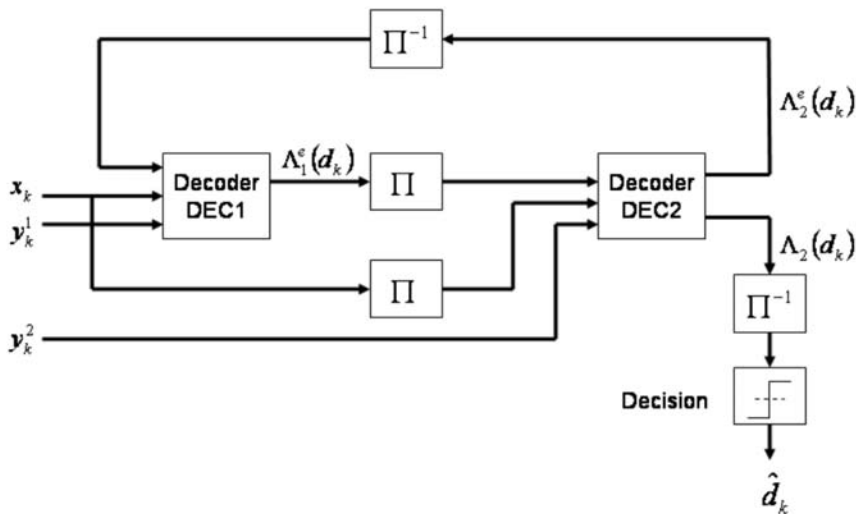


Figure 3.5 Iterative turbo decoder [12].

encoder (y_k^1), and then produces a soft output representing the reliability of the decoded bits.

Extrinsic information $\Lambda_1^e(d_k)$, which depends on the redundant information introduced by the encoder and does not contain the information at the input of the decoder, is determined based on these metrics and is fed in an interleaved order to the second decoder. This extrinsic information together with the received systematic sequence is then exploited by the second decoder to decode the received sequence corresponding to the second encoder (y_k^2) and produces new extrinsic information $\Lambda_2^e(d_k)$, which is de-interleaved and fed back to the first decoder for a new iteration, and so forth.

Each decoder builds on each other's performance to improve the reliability of the decoding. After a certain number of iterations, the hard decision at the output of the second decoder determines the decoded data bit \hat{d}_k . Thus the decoding algorithm has an iterative nature.

3.1.1.1.5 Design and Performance of Codes

3.1.1.1.5.1 Turbo Codes. The major parameter involved in the performance bound related to the effective free distance, together with the design of turbo codes, can be expressed as shown in (3.6):

$$d_{\text{free, eff}} = 2 + 2 \cdot z_{\min} \quad (3.6)$$

where z_{\min} corresponds to the lowest error weight generated by an information sequence of weight 2 [44]. At high SNRs, the free distance of the code influences the BER probability of turbo codes. This parameter depends on how the interleaver affects the slope of the error curve. A low free distance implies a flattening of the BER curve, termed the *error floor*.

For turbo codes associated with random interleavers, by maximizing the effective free distance $d_{\text{free, eff}}$, thus z_{\min} , a criteria can be defined that can be used in the design of interleavers that would operate at high SNR values. In this case, the BER probability for a given coding rate R , can be evaluated as

$$P_b(e) \approx B_{\text{free, eff}} \cdot Q\left(\sqrt{2 \cdot d_{\text{free, eff}} \cdot R \cdot \frac{E_b}{N_0}}\right) \quad (3.7)$$

where $B_{\text{free, eff}}$ corresponds to the error coefficient related to the effective free distance $d_{\text{free, eff}}$, and Q represents the complementary error function defined as follows:

$$Q(x) = \frac{1}{\sqrt{2 \cdot \pi}} \cdot \int_x^{+\infty} e^{-t^2/2} \cdot dt \quad (3.8)$$

The upper bound of bit error probability of a turbo code is given by

$$P_b(e) \leq \sum_{d=d_{\text{free}}} B_d \cdot Q\left(\sqrt{2 \cdot d \cdot R \cdot \frac{E_b}{N_0}}\right) \quad (3.9)$$

with (d, B_d) representing the code distance spectrum. For low SNR regions, the distance spectrum has to be taken as the design criteria.

The interleaver mainly affects performance through its length. An increasing interleaver size randomizes the information exchanged between the two decoders, resulting in a better decorrelation of the inputs of the two decoders and, consequently, in better asymptotic performance [12].

In the area of low and moderate SNRs, where performance is mainly determined by the large number of average-weight codewords, the choice of a particular information interleaver realization is not critical as opposed to its length, which is the main reason for the good performance of turbo codes. A well-designed information interleaver plays an important role in determining the first several spectral lines that limit the asymptotic behavior of any code at high SNR [42]. The information interleaving design for simultaneous termination of both constituent codes has been studied in [45]. The performance with short information interleavers and the effects of puncturing to the distance spectrum has been extensively studied [46, 47].

Increasing the number of iterations noticeably impacts the performance for a low number of iterations. Suboptimal iterative decoding closely approaches the optimal performance for a growing number of iterations [12]. However, after a certain number of iterations, the performance exhibits very small improvements when the number of iterations gets larger, at the expense of an increasing complexity and decoding delay. An optimal value for the number of iterations should therefore be selected in order to achieve a trade-off between complexity and performance. A stopping criterion can be applied to stop the iterative process and prevent unnecessary computations [48].

When dealing with generalized multicarrier (MC) transmission schemes, such as MC-SS, diverse combining methods are well known to combat channel fading, by means of diversity order. For the turbo-coded version of any MC-SS transmission scheme, it is therefore interesting to study the impact of a simple single-user detection (SUD) scheme, such as maximal ratio combining (MRC), equal gain combining (EGC), or selective combining (SC) on the performance of the turbo code, and for which modifications must be done at the branch metric level [12]. SNR estimation is of principal interest with some decoding algorithms such as log-MAP detectors. This requires the design of new algorithms to take the diversity aspect into account.

3.1.1.1.5.2 LDPC Codes. The performance of LDPC codes can in general be improved for low E_b/N_0 by allowing nodes of various degrees and by choosing the proportion of the various degrees carefully [49]. Therefore, an irregular LDPC code can be defined [12] for which the degrees of each set of nodes are chosen according to some distribution, such as the variable node degree distribution $\lambda(x)$ and the check node degree distribution $\rho(x)$.

To find good variable node and check node degree distributions, the density evolution method is introduced [50]. This algorithm calculates the channel parameter (e.g., the noise power for the AWGN channel) for which an infinite long

LDPC code can be transmitted without errors. The probability distribution of log-likelihood ratios (LLRs) after the transmission, and after each iteration of the sum-product or message-passing algorithm, is evaluated. Also, the error probability after each iteration is calculated. To determine the threshold, the channel parameter is iteratively increased and the algorithm is run until the error probability converges to zero or to a value larger than zero when the number of iterations tends to infinity. The first time the error probability does not converge to zero, the threshold has been found.

In communication systems with LDPC codes, iterative receiver and decoder techniques can improve the performance by passing the LLRs generated in each separate function between the receiver/decoder pairs. The receiver and decoder (i.e., for an LDPC code) are designed separately. Furthermore, LDPC codes and also the decoders are usually designed for AWGN channels. Any statistical characteristics of the fading are usually not used explicitly in the key step (e.g., probability density evolution) when designing an irregular LDPC code. The design process is more complex for MIMO channels due to the multiple channels existing between the receiver and transmitter.

Some work has used turbo-coded structures to “optimally” combine LLRs, because there is usually an interleaver between the MIMO receiver and the LDPC decoder [12]. The designs of each component of the receiver, however, are still independent. The common component is *likelihood*, which when used iteratively should also improve the BER performance over any single component. One evolutionary step for combining iterative soft likelihoods is to describe and connect the MIMO receiver and LDPC codes in a more general framework (e.g., factor graphs). This allows for incorporating the LDPC design to incorporate the statistics of the MIMO channel explicitly in the “density evolution” step in the LDPC design procedure.

In the following discussion, the design method of LDPC for the MIMO channel proposed in [51] and considered in [12] is described.

In MIMO-OFDM systems with M_T transmit and M_R receive antennas, respectively, the received signal at each frequency bin k after FFT has the form $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$ where $\mathbf{y} = [y_1(k) y_2(k) \dots y_{M_R}(k)]^T$ is the received signal at frequency bin k and $\mathbf{x} = [c_1(k) c_2(k) \dots c_{M_T}(k)]^T$ is the data symbol vector at frequency bin k from all transmitting antennas. The channel matrix \mathbf{H} can be written as $\mathbf{H} = \mathbf{Q}\mathbf{R}$ where \mathbf{Q} is an orthonormal matrix and \mathbf{R} is usually an upper triangular matrix. This matrix decomposition is the QRD decomposition. Applying \mathbf{Q}^H to the received vector \mathbf{y} yields:

$$\begin{bmatrix} y'_1(k) \\ y'_2(k) \\ \vdots \\ y'_{M_R}(k) \end{bmatrix} = \begin{bmatrix} r_{1,1}(k) & r_{2,1}(k) & \dots & r_{M_T,1}(k) \\ 0 & r_{2,2}(k) & \dots & r_{M_T,2}(k) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{M_T,M_T}(k) \\ 0 & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} c_1(k) \\ c_2(k) \\ \vdots \\ c_{M_T}(k) \end{bmatrix} + \begin{bmatrix} n'_1(k) \\ n'_2(k) \\ \vdots \\ n'_{M_R}(k) \end{bmatrix} \quad (3.10)$$

when $M_R > M_T$. This decomposition allows the symbols to be estimated iteratively using successive interference cancellation starting with symbol $c_{M_T}(k)$ and progressively upward using previous symbol estimates. The QRD form has other attractive properties. Each element in the matrix \mathbf{R} has a pdf that is some variant of the χ pdf [52] (e.g., $r_{M_T, M_R}(k) \sim \sqrt{\chi_2}$, $r_{M_T-1, M_R}(k) \sim \sqrt{\chi_4}$). Furthermore, all off-diagonal entries have χ_2 pdf. All matrix entries are also independent. These properties can be derived by using Bartlett's decomposition [53]. These properties make the probability calculation easier when defining edge probabilities in a factor graph.

Figure 3.6 shows a Tanner graph representation for a 4×4 MIMO channel receiver using the QRD algorithm. For frequency bin k , a subset of codeword nodes of an LDPC code for a MIMO-OFDM system would be $[c_1(k) \ c_2(k) \ \dots \ c_{M_T}(k)]$. LDPC codes require a statistical description of input channel signals that are connected to codeword nodes. Therefore, the factor graph shows edges from received signals $[y'_1(k) \ y'_2(k) \ \dots \ y'_{M_R}(k)]$ to codeword nodes $[c_1(k) \ c_2(k) \ \dots \ c_{M_T}(k)]$ based on the QRD algorithm. The edges between the codeword nodes and the parity check nodes are determined by the parity-check matrix defined for an LDPC. The edges in the factor graph are message-passing probabilities.

The factor graph representation for the LDPC component shows the expected relationships between codeword nodes and parity-check nodes, but codeword nodes are also related to each other through transition probabilities depending on the entries in matrix \mathbf{R} .

A full cycle of the LDPC recursion step for both “density evolution” and decoding could be as follows:

- Likelihoods from codeword nodes \rightarrow likelihoods to parity-check nodes;
- Likelihoods from parity-check nodes \rightarrow likelihoods to codeword nodes;
- Likelihoods from codeword nodes \rightarrow likelihoods to codeword nodes.

Figure 3.7 shows an alternative QRD/LDPC design with the addition of a lower diagonal QRD decomposition to provide for another set of channel observables. In this case, the channel matrix is decomposed through two QRDs: $\mathbf{H} = \mathbf{Q}_1 \mathbf{R}_1$, $\mathbf{H} = \mathbf{Q}_2 \mathbf{R}_2$, where \mathbf{R}_1 and \mathbf{R}_2 are an upper and lower triangular matrix, respectively.

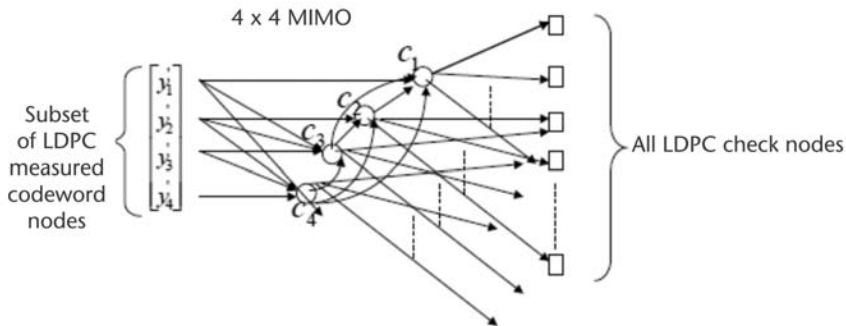


Figure 3.6 A factor graph representation of QRD/LDPC [12].

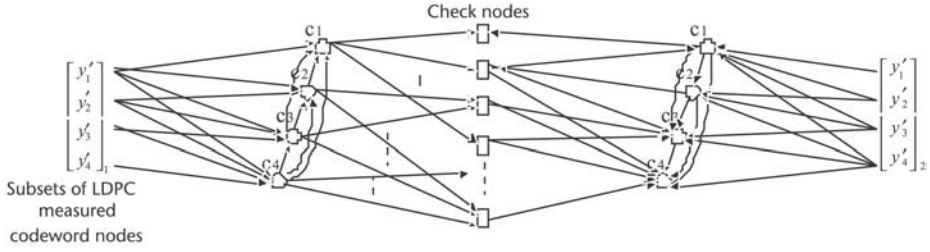


Figure 3.7 A QRD/LDPC symmetric design [12].

These could have been used for SIC in the traditional QRD-M algorithm to provide an LDPC code with codeword inputs. In this configuration, the recursion is defined to increase the “girth” (i.e., length) of the cycles in the factor graph. Let C_1 and C_2 be defined as a set of codeword nodes for upper triangular and lower triangular matrices, respectively. Let P be defined as the set of parity-check nodes; then a full-cycle recursion for both density evolution (DE) and decoding could be defined as follows:

- $C_1 \rightarrow P$
- $P \rightarrow C_2$
- $C_2 \rightarrow C_2$
- $C_2 \rightarrow P$
- $P \rightarrow C_1$
- $C_1 \rightarrow C_1$

In both design cases, it is necessary to determine the pdf of each $r_{i,j}$ in the upper or lower triangular matrices. The second design method of LDPC codes for MIMO channels is very suitable for MIMO-OFDM systems. It explores the statistical characteristics of MIMO fading channels so that the receiver and decoder are designed jointly and then the performance is improved.

In summary, LDPC codes have the following advantages:

- High flexibility in terms of code rate and length;
- Decoding errors are usually detectable;
- Low complexity constituent decoders (variable node decoder, check node decoder);
- Easy to make parallel;
- Decoding complexity lowers with rising SNR;
- No need for outer interleaver.

3.1.1.1.6 Summary.

The design of a radio interface requires careful investigation of the best suited coding techniques. Such techniques should exhibit good performance and manageable encoding and decoding complexity.

Turbo codes and irregular LDPC codes show a very good performance for low E_b/N_0 , but a flattening of the BER and FER occurs for high E_b/N_0 , the so-called error floor. Nevertheless, this error floor decreases with the latest generation of turbo codes, duo-binary turbo codes (DBTC). The main advantage of turbo codes compared with LDPC codes is the easy encoding, done by two shift registers and an interleaver. The rate adaptation by puncturing is an advantage of turbo codes. In contrast to a turbo decoder, an LDPC decoder can easily be parallelized in hardware implementations, which shortens the decoding delay. This is a big advantage of LDPC code for applications, where a small delay is necessary. For both codes an iterative decoding algorithm is used. Here, one iteration of an LDPC decoder is less complex than one iteration of a turbo decoder, but more iterations are needed for the decoding of LDPC codes.

The BER and FER of Woven codes, serially concatenated convolutional codes, and regular LDPC codes are worse than the BER and FER of turbo codes and irregular LDPC codes for low E_b/N_0 , but the flattening of the error rates occurs at a much higher E_b/N_0 . Also for these codes a complex iterative decoding algorithm must be used. While the encoding of irregular LDPC codes is not very complex, because the sparse parity-check matrix can be used, the encoding of arbitrary regular LDPC codes can be a very hard task.

Woven codes, serially concatenated convolutional codes (SCCCs), DBTCs, and regular LDPC codes should be used when very low error rates must be guaranteed. For applications that do not need a very small BER and FER, turbo codes and irregular LDPC codes should be used. All of these codes work very well when the block length is quite large. For shorter block lengths convolutional codes are a good choice. Also the encoding and, especially, the decoding of convolutional codes are less complex.

3.1.2 Analysis and Design of Modulation Schemes for Next Generation Systems

Modulation methods incorporate different approaches for embedding information in channel waveforms [12]. Correspondingly, demodulation methods are utilized to extract the information from the received waveforms, which have been propagated through the channel. Modulation and demodulation methods for transmission considered in the design phase of the WINNER air interface [2] included multi-carrier modulations: OFDM, MC-CDMA, and filterbank modulations including generic filtered multitone and IOTA modulation. Other modulation methods considered include general single-carrier methods, conventional serial transmission SC methods, and frequency-domain orthogonal spreading.

A common modular representation framework consists of an optional DFT operation, a frequency-domain scheduling/allocation operation, an inverse discrete Fourier transform (IDFT) operation, prefix/postfix addition, and transmitter pulse shaping operation.

3.1.2.1 Common Modular Representation

3.1.2.1.1 DFT.

A sequence of M complex data symbols $\{\mathbf{a} \in C^{M \times 1}\}$ is transformed with a DFT operation by a DFT matrix \mathbf{F} :

$$[\mathbf{F}]_{k,l} = \frac{1}{\sqrt{M}} e^{-j(2\pi kl/M)} \quad \text{for } k, l = 0, 1, \dots, M-1 \quad (3.11)$$

into a frequency-domain data vector \mathbf{b} . Multicarrier modulators can neglect this block, since they directly excite the frequency-domain signal.

3.1.2.1.2 Frequency Assignment and Scheduling.

The resulting frequency-domain sequence \mathbf{b} is assigned to the available N_m frequencies by the scheduling/allocation matrix \mathbf{M} , resulting in the vector \mathbf{c} . For orthogonal frequency allocation \mathbf{M} is an orthogonal $N \times M$ matrix of zeros and ones, with each column having a single 1 indicating the frequency allocation. Some examples of the allocation matrix shape are given in Table 3.1.

3.1.2.1.3 IDFT Waveform Filtering.

The allocation is followed by the waveform filtering matrix, described with an $N_c \times N_c$ matrix \mathbf{V} . For single- and multiple-carrier modulations this is the IDFT matrix \mathbf{F}^H . If waveform filtering is desired in order to realize alternative waveforms, these can be produced with an alternative filtering operation. The waveform filter produces the time-domain vector \mathbf{d} .

3.1.2.1.4 Prefix/Postfix.

The insertion of a prefix is described by the prefix insertion, operation, which produces the vector \mathbf{e} . The operation is given by

$$\mathbf{e} = \bar{\mathbf{Z}}_P \mathbf{d} + \mathbf{Z}_P \quad (3.12)$$

where the prefix matrix is given by

Table 3.1 Examples of an Allocation Matrix

Single-Carrier Transmission	$\mathbf{M} = \mathbf{I}$
OFDM	$\mathbf{M} = \mathbf{I}$
OFDMA (sparse band)	$\mathbf{M} = \begin{bmatrix} 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ 0 & 1 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 1 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix}$
FDMA and OFDMA (continuous band)	$\mathbf{M} = \begin{bmatrix} \mathbf{0}_o \\ \mathbf{I}_M \\ \mathbf{0}_{N_m - M - o} \end{bmatrix}$ <p>Parameter o gives the frequency offset.</p>
FDOSS	$\mathbf{M} = \mathbf{I}_M \otimes \tilde{\mathbf{e}}_{p,K}$, $\tilde{\mathbf{e}}_{p,K}$ is the p th column of \mathbf{I}_K , the identity matrix of dimension K and p is the user index
MC-CDMA	$\mathbf{M} = \mathbf{W}$ (code matrix)

$$\mathbf{Z}_P = \begin{bmatrix} \mathbf{z}_P \\ 0 \end{bmatrix} \quad (3.13)$$

and

$$\bar{\mathbf{Z}}_{Pt} = \begin{bmatrix} 0 \\ \mathbf{I} \end{bmatrix} \quad (3.14)$$

A postfix is attached with a postfix insertion matrix as in

$$\mathbf{e} = \bar{\mathbf{Z}}_{Po} \mathbf{d} + \mathbf{Z}_P \quad (3.15)$$

where the postfix matrix is given by

$$\bar{\mathbf{Z}}_P = \begin{bmatrix} 0 \\ \mathbf{z}_P \end{bmatrix} \quad (3.16)$$

and

$$\bar{\mathbf{Z}}_{Po} = \begin{bmatrix} \mathbf{I} \\ 0 \end{bmatrix} \quad (3.17)$$

whichever is desired, and \mathbf{Z}_P denotes the prefix/postfix sequence. If the modulation does not utilize prefixes or postfixes the matrix can be removed.

3.1.2.1.5 Transmitter Filtering.

The transmitted signal is filtered with a transmission filter with a pulse waveform $g(t - \tau)$ that needs to be sampled at the receiver-sampling rate. The filtering operation is described as a multiplication with a Toeplitz matrix \mathbf{G} , whose columns are defined by the samples of $g(n + n_s \tau_s)$, where n is the index of the transmitted symbol, n_s is the sample index, and τ_s is the inverse sampling rate. The filtering can also be described by a frequency-domain filter placed before the IDFT, in which case the filter g becomes a simple lowpass filter. This is, however, mathematically equivalent to the preceding description.

3.1.2.1.6 Transmitted Signal.

Based on the preceding discussion, the transmitted signal (in this case utilizing a postfix) can be given as the results of the matrix operation:

$$\mathbf{x} = \mathbf{G} \bar{\mathbf{Z}}_{Po} \mathbf{V} \mathbf{M} \mathbf{F} \mathbf{a} + \mathbf{G} \mathbf{z}_p \quad (3.18)$$

A generalized frequency-domain transmitter is shown in Figure 3.8.

3.1.2.2 Multicarrier Modulation

Multicarrier modulation (MCM) is a technique for transmitting data by dividing a high bit rate stream into several lower data rate substreams, and then using these

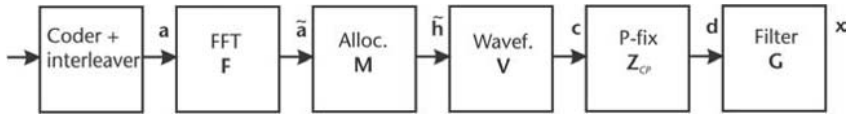


Figure 3.8 Generalized frequency-domain transmitter.

substreams to modulate several subcarriers. As a special form of MCM, orthogonal frequency-division multiplexing has been extensively studied and become the most popular MCM technique.

OFDM as a multiplexing technique is similar to FDM in the sense that the multiple-user access is achieved by subdividing the available bandwidth into multiple frequency subcarriers, which are then allocated to users. In addition, OFDM can be considered to be a special kind multicarrier transmission technique that uses the spectrum efficiently by spacing the subcarriers much closer together, actually overlapping them. This is achieved by making all of the carriers orthogonal to one another and thereby preventing interference between the closely spaced carriers.

The orthogonality of the subcarriers means that each carrier has an integer number of cycles over a symbol time period. Due to this, the spectrum of each carrier has a null at the center frequency of each of the other carriers in the overall OFDM signal. This results in no interference between the carriers, allowing them to be spaced very close. In the case of multipath channels, the orthogonality can be maintained better by introducing a cyclic prefix to the OFDM symbols. A CP is a copy of the last part of the OFDM symbol that is appended to the transmitted symbol. This makes the transmitted signal periodic and decreases the transceiver complexity by enabling the use of DFT, at the expense of additional overhead or SNR loss.

Each subcarrier in an OFDM signal has a very narrow bandwidth (less than the coherence bandwidth) and the resulting subcarrier symbol rate is low. This results in the signal having a high tolerance to multipath delay spread, because the delay spread must be very long to cause significant ISI. OFDM has been particularly successful in numerous wireless applications, where its superior performance in multipath environments is desirable. OFDM in conjunction with proper coding and interleaving (COFDM = coded OFDM) takes full advantage of frequency diversity, and results in a powerful technique for combating wireless channel impairments. A very efficient way to modulate an OFDM signal is to use an IDFT of size N_{FFT} , N_c for modulation. A detailed overview of OFDM modulation and demodulation as well as the signal system concept, including synchronization, is available in [54, 55].

One major drawback of OFDM modulation is its sensitivity to synchronization errors; in particular, to carrier frequency synchronization errors. Two destructive effects are caused by a carrier frequency offset in OFDM systems. One is the reduction of signal amplitude (the sinc functions are shifted and no longer sampled at the peak) and the other is the introduction of ICI from the other carriers.

Carrier phase noise is caused by imperfections in transmitter and receiver oscillators. In an OFDM system, the carrier phase noise manifests itself as phase rotation (so-called common phase error) and as ICI.

ICI can be suppressed after FFT with linear or decision feedback equalizers (DFE). Limited frequency offset can be compensated for in the time domain with a linear predictor (and extra FFT). In some situations the OFDM guard period is too short to avoid ISI totally. Restrictions for guard period length are set by the required system capacity and by channel coherence time and FFT complexity restricting the total OFDM symbol duration. By allowing a shorter guard period and limited ISI, higher system capacity and mobility can be achieved.

When designing an OFDM mobile system [12], both ISI and ICI can be avoided, or at least be kept at an acceptable level. That is, a sufficient guard interval prevents ISI and ICI due to frequency selective fading; and a good choice of the symbol duration T provides a constant CIR over at least one OFDM symbol.

The loss of spectral efficiency due to the guard interval is given by $T_{GI}/T = N_{GI}/N_{FFT}$. Thus, to increase the spectral efficiency, the guard interval should be as short as possible. This can be achieved by increasing the number of subcarriers. Increasing the symbol duration, however, makes the system more sensitive to the time selectivity of the channel due to the Doppler spread. Because the time selectivity affects the orthogonality of the carriers, a larger symbol duration gives rise to more ICI. The lengthening of the symbol duration, introduced to combat the frequency selectivity is limited by the time-variant nature of the mobile radio channel.

For highly mobile environments, especially in a macrocell environment, a trade-off is observed. In this case the channel will be extremely frequency selective *and* time variant. As a rule, a normalized Doppler frequency $\zeta_{\max}T = 10\%$ will not result in significant ICI due to the Doppler spread. In this case, the degradation due to imperfect channel estimation will be far more significant. On the other hand, the guard interval should be less than 20% of the symbol duration. An analysis of the performance and parameter optimization of N_c and N_{GI} for OFDM and for MC-CDMA is given in [12, 56, 57]. The approach is to maximize the average useful signal energy of the received signal, or in other words, to minimize the leakage through ICI and ISI. In [56] it was shown that the optimum values for N_c and N_{GI} not only depend on the channel parameters but also on the SNR.

Although these optimization techniques are important tools for the design of an OFDM system, the problem is that a practical system needs to cope with various operating scenarios having many different channel parameters. One solution to this problem is to optimize the system parameters for the worst-case scenario, which in terms of the channel parameters is a macrocell with high-mobility users. Another solution may be to optimize the system parameters for different scenarios (i.e., indoor or outdoor, pico-, micro- or macrocells), which requires a highly flexible transceiver.

For a next generation mobile radio system, the downlink bandwidth will be about $B = 100$ MHz at a carrier frequency of 4 to 6 GHz [58]. The Doppler spread at a velocity of 300 kmph is then about $\zeta_{\max}T \approx 1$ to 2 kHz. A normalized Doppler frequency of $\zeta_{\max}T = 10\%$ results in a minimum subcarrier frequency separation of $1/T > 20$ kHz, which limits the maximum number of subcarriers to $N_{FFT} = BT < 5,000$. If the guard interval satisfies $T_{GI}/T < 20\%$, the maximum tolerable delay of the channel becomes $\tau_{\max} < T_{GI} = 10 \mu s$.

If either the carrier frequency or the maximum delay of the channel is increased, orthogonality at the receiver can no longer be maintained [55].

3.1.2.2.1 Pseudorandom Postfix OFDM (PRP-OFDM) for SISO Antenna Case.

The basic transceiver chain of a PRP-OFDM system is shown in Figure 3.9. The general structure of the modulator is similar to the standard cyclic-prefix OFDM (CP-OFDM) unit. The difference is that instead of a cyclic extension of the time-domain OFDM symbol, a pseudorandomly weighted sequence is appended to the symbol. This sequence is a priori known to the transmitter and the receiver.

In the given baseband discrete-time block equivalent model of an N_c carrier PRP-OFDM system, the i th $N_c \times 1$ input digital vector $\tilde{\mathbf{x}}_{N_c}(i)$ is first modulated by the IFFT matrix

$$\mathbf{F}_{N_c}^H := \frac{1}{\sqrt{N_c}} (W_{N_c}^{ij})^H, \quad 0 \leq i < N_c, \quad 0 \leq j < N_c \quad \text{and} \quad W_{N_c} := e^{-j2\pi/N_c}$$

Then, a deterministic postfix vector $\mathbf{c}_D := (c_0, \dots, c_{D-1})^T$ weighted by a pseudorandom value $\alpha(i) \in \mathbb{C}$ is appended to the IFFT output $\mathbf{x}_{N_c}(i)$. With $P := N_c + D$, the corresponding $P \times 1$ transmitted vector is $\mathbf{x}_P(i) := \mathbf{F}_{N_c}^H \tilde{\mathbf{x}}_{N_c}(i) + \alpha(i)\mathbf{c}_P$, where:

$$\mathbf{F}_{N_c}^H := \begin{bmatrix} \mathbf{I}_{N_c} \\ \mathbf{0}_{D, N_c} \end{bmatrix}_{P \times N_c} \mathbf{F}_{N_c}^H \quad \text{and} \quad \mathbf{c}_P := (0_{1, N_c} \quad \mathbf{c}_D^T)^T \quad (3.19)$$

The elements of $\mathbf{x}_{N_c}(i)$ can be assumed to be *i.i.d.* and *zero-mean* random variables of variance $\sigma_x^2 = 1$, which are independent of $\alpha(i) \in \mathbb{C}_D$. The samples of $\mathbf{x}_P(i)$ are then sent sequentially through the channel modeled here as an FIR filter of order L , $H(z) := \sum_{n=0}^{L-1} h_n z^{-n}$. The OFDM system is designed such that the postfix duration exceeds the channel memory $L \leq D + 1$.

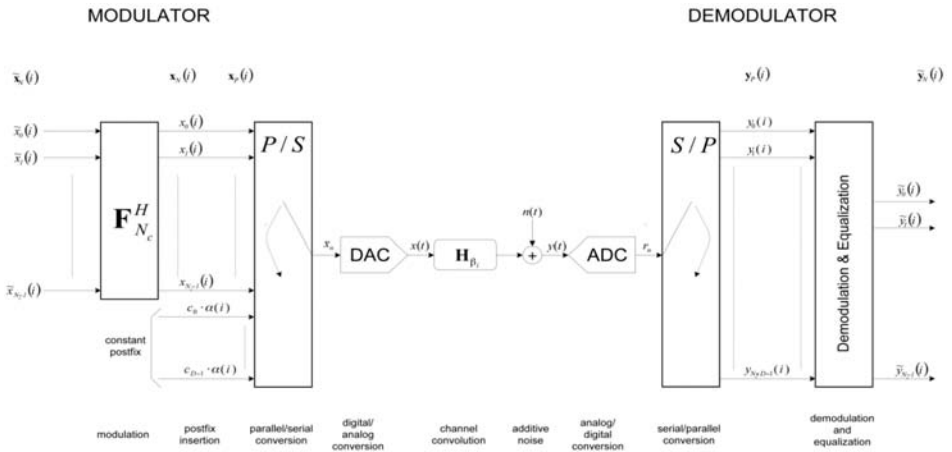


Figure 3.9 Basic PRP-OFDM transceiver chain [12].

Let $\mathbf{H}_{\text{ISI}}(P)$ and $\mathbf{H}_{\text{IBI}}(P)$, respectively, be the size P Toeplitz inferior and superior triangular matrices of first column $[h_0, h_2, \dots, h_{L-1}, 0, \rightarrow, 0]^T$ and first row $[0, \rightarrow, 0, h_{L-1}, \dots, h_1]$. The channel convolution can be modeled by $\mathbf{y}_P(i) := \mathbf{H}_{\text{ISI}}(P)\mathbf{x}_P(i) + \mathbf{H}_{\text{IBI}}(P)\mathbf{x}_P(i-1) + \mathbf{n}_P(i)$. The $\mathbf{H}_{\text{ISI}}(P)$ and $\mathbf{H}_{\text{IBI}}(P)$ matrices represent, respectively, the intrablock and interblock interference, and $\mathbf{n}_P(i)$ is the i th AWGN vector of element variance σ_n^2 . Since $\mathbf{x}_P(i) = \mathbf{F}_{\text{ZP}}^H \mathbf{x}_{N_c}(i) + \alpha(i)\mathbf{c}_P$, (3.20) holds:

$$\mathbf{y}_P(i) = (\mathbf{H}_{\text{ISI}}(P) + \beta_i \mathbf{H}_{\text{IBI}}(P))\mathbf{x}_P(i) + \mathbf{n}_P(i) \quad (3.20)$$

where $\beta_i := \frac{\alpha(i-1)}{\alpha(i)}$ and $\mathbf{H}_{\beta_i} := (\mathbf{H}_{\text{ISI}}(P) + \beta_i \mathbf{H}_{\text{IBI}}(P))$ is pseudocirculant; a circular matrix whose $(D-1) \times (D-1)$ upper triangular part is weighted by β_i . The three channel matrices part of (3.20) as well as the combined matrix are shown in Figures 3.10 to 3.13, respectively.

In Figure 3.9 the time-domain samples of an OFDM symbol are split into four parts of identical size. The pseudorandomly weighted postfix is assumed to have a length of one-fourth of the OFDM symbol length.

The expression of the received block is thus:

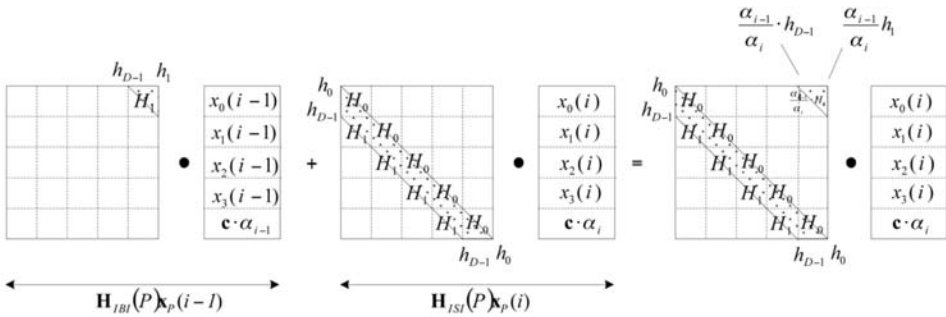


Figure 3.10 Channel convolution of one received block [12].

$$\mathbf{H}_{\text{IBI}} = \begin{bmatrix} 0 & \rightarrow & 0 & h_{D-1} & h_{D-2} & \rightarrow & h_1 \\ \downarrow & \mu & \mu & \mu & h_{D-1} & \mu & h_2 \\ \downarrow & \mu & \mu & \mu & \mu & \mu & \downarrow \\ \downarrow & \mu & \mu & \mu & \mu & \mu & h_{D-1} \\ \downarrow & \mu & \mu & \mu & \mu & \mu & 0 \\ \downarrow & \mu & \mu & \mu & \mu & \mu & \downarrow \\ 0 & \mu & \mu & \mu & \mu & \mu & \rightarrow \end{bmatrix}$$

Figure 3.11 The inter block interference (IBI) channel matrix [12].

$$= \mathbf{H}_{ISI} = \begin{pmatrix} h_0 & 0 & \rightarrow & \rightarrow & \rightarrow & \rightarrow & 0 \\ h_1 & \mu & \mu & \mu & \mu & \mu & \downarrow \\ \downarrow & \mu & \mu & \mu & \mu & \mu & \downarrow \\ h_{D-2} & \mu & \mu & \mu & \mu & \mu & \downarrow \\ h_{D-1} & \mu & \mu & \mu & \mu & \mu & \downarrow \\ 0 & \mu & \mu & \mu & \mu & \mu & \downarrow \\ \downarrow & \mu & \mu & \mu & \mu & \mu & 0 \\ 0 & \mu & 0 & h_{D-1} & h_{D-2} & \rightarrow & h_0 \end{pmatrix}$$

Figure 3.12 The ISI channel matrix [12].

$$= \mathbf{H}_{ISI} + \mathbf{H}_{IBI} = \begin{pmatrix} h_0 & 0 & \rightarrow & h_{D-1} & h_{D-2} & \rightarrow & h_1 \\ h_1 & \mu & \mu & \mu & h_{D-1} & \mu & h_2 \\ \downarrow & \mu & \mu & \mu & \mu & \mu & \downarrow \\ h_{D-2} & \mu & \mu & \mu & \mu & \mu & h_{D-1} \\ h_{D-1} & \mu & \mu & \mu & \mu & \mu & 0 \\ 0 & \mu & \mu & \mu & \mu & \mu & \downarrow \\ \downarrow & \mu & \mu & \mu & \mu & \mu & 0 \\ 0 & \mu & 0 & h_{D-1} & h_{D-2} & \mu & h_0 \end{pmatrix}$$

Figure 3.13 The combined IBI and ISI channel matrix [12].

$$\mathbf{y}_P(i) = \mathbf{H}_{\beta_1} \left(\mathbf{F}_{ZP}^H \mathbf{x}_{N_c}(i) + \alpha(i) \mathbf{c}_P \right) + \mathbf{n}_P(i) = \mathbf{H}_{\beta_1} \begin{pmatrix} \mathbf{F}_{N_c}^H \mathbf{x}_{N_c}(i) \\ \alpha(i) \mathbf{c}_D \end{pmatrix} + \mathbf{n}_P(i) \quad (3.21)$$

Equation (3.21) is quite generic and also captures the CP and ZP modulation schemes. Indeed ZP-OFDM corresponds to $\alpha(i) = 0$ and CP-OFDM is achieved for $\alpha(i) = 0$, $\beta_i = 1$, $\forall i$ when \mathbf{F}_{ZP}^H is replaced by \mathbf{F}_{CP}^H :

$$\mathbf{F}_{CP}^H := \begin{bmatrix} \mathbf{0}_{D, N_c-D} & \mathbf{I}_D \\ \mathbf{I}_{N_c} & \end{bmatrix}_{P \times N_c} \mathbf{F}_{N_c}^H \quad (3.22)$$

The term $\mathbf{H}_{\beta_1} \mathbf{c}_P$ can be retrieved by a simple averaging (i.e., mean value calculation) of the received samples if the OFDM data symbols $\mathbf{F}_{N_c}^H \mathbf{x}_{N_c}(i)$ are assumed to be zero mean. The channel can be extracted afterwards by deconvolution. (These issues were discussed in detail in [12].) Based on these definitions, the general frame structure is as shown in Figure 3.14.

The introduction of pseudorandom postfixes leads to the possible ability to estimate the channel impulse response at a very low computational complexity. This was shown in full detail in [12]. Decoding of the symbols is still possible even if the channel nulls fall onto data carriers (at high SNR). In addition, several



Figure 3.14 General frame structure [12].

performance and complexity trade-offs are possible in the receiver architecture. This scheme keeps the same throughput and spectral efficiency as standard CP-OFDM if the length of the CP-OFDM guard interval is identical to the length of the pseudorandomly weighted postfix of PRP-OFDM. Receiver complexity is slightly increased compared to CP-OFDM (concerning low-complexity OLA-based implementation).

3.1.2.2.2 PRP-OFDM for MIMO Context.

The SISO concept can be extended to MIMO as shown in Figure 3.15, where the baseband discrete-time block is an equivalent model of an N_c carrier MIMO PRP-OFDM transceiver with M_T transmit and M_R receive antennas. The proposed scheme in Figure 3.15 is generic and the analysis is limited to space-time block coding (STBC).

The initial serial stream of constellation symbols $x(jN_c), \dots, x(jN_c + N_c - 1)$ is serial-to-parallel converted; the j th $N_c \times 1$ input digital vector $\tilde{\mathbf{x}}(j)$ is then modulated by the IFFT matrix $\mathbf{F}_{N_c}^H$ with $[\mathbf{F}_{N_c}]_{k,l} = (1/\sqrt{N_c}) W_{N_c}^{kl}$, $0 \leq k < N_c$, $0 \leq l < N_c$, where $W_{N_c} = e^{-j2\pi/N_c}$.

The resulting $N_c \times 1$ time-domain vector $\tilde{\mathbf{x}}(j)$ is processed by any suitable ST encoder \mathbf{M} creating the outputs

$$\bar{\mathbf{X}}(i) = M(\mathbf{x}(iM_T), \dots, \mathbf{x}(iM_T + M_T - 1)) = \{\bar{\mathbf{x}}_l(iM + k), 1 \leq l \leq M_T, 0 \leq k < M\}$$

with i being the block number and $n = iM + k$ indexing the outputs in Figure 3.15. Note that in the context of STBCs, \mathbf{M} can differ from M_T [59]. In the MIMO case, the deterministic $D \times 1$ postfix vector \mathbf{c} is treated by a specific ST encoder \mathbf{W} , which outputs the $D \times 1$ vectors $\bar{\mathbf{c}}_l(n)$, $1 \leq l \leq M_T$. In this case, \mathbf{W} ensures the identification of the complete MIMO channel. To avoid unpleasant spectrum prop-

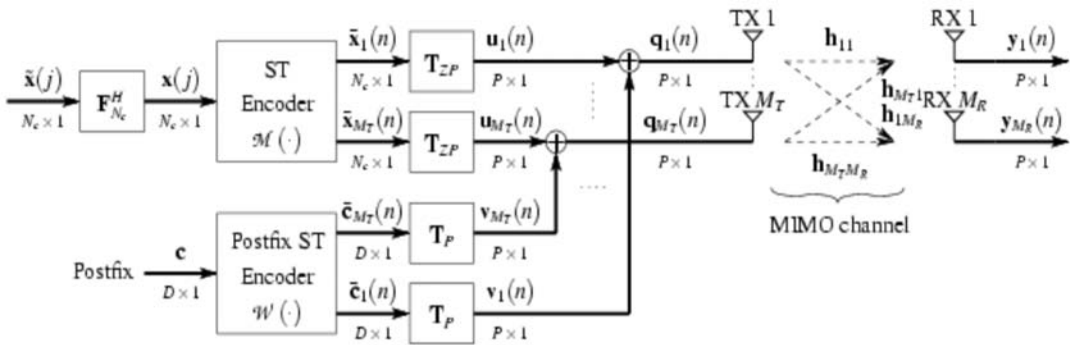


Figure 3.15 A MIMO-PRP-OFDM scheme [12].

erties, the postfix vector is weighted by a scalar pseudorandom sequence. The postfix vectors $\bar{\mathbf{c}}_l(n)$ are then linearly precoded by the matrix \mathbf{T}_P :

$$\mathbf{T}_P = \begin{bmatrix} 0_{N_c, D} \\ \mathbf{I}_D \end{bmatrix}_{P \times D} \quad (3.23)$$

The resulting $\mathbf{v}_l(n)$ are finally added to the data symbols $\mathbf{u}_l(n) : \mathbf{q}_l(n) = \mathbf{u}_l(n) + \mathbf{v}_l(n)$, $1 \leq l \leq M_T$. Let \mathbf{H}_{lm} be a $P \times P$ circulant matrix whose first row is given by $[h_{lm}(0), 0 \rightarrow 0, h_{lm}(L-1), \dots, h_{lm}(1)]$, where $\mathbf{h}_{lm} = [h_{lm}(0), \dots, h_{lm}(L-1), 0 \rightarrow 0]^T$ is the $P \times 1$ channel impulse response between the l th transmitting and the m th receiving antenna; D is chosen such that $D \geq L-1$. Define $\mathbf{H}_{lm}^{\text{ISI}}$ as the lower triangular part of \mathbf{H}_{lm} including the main diagonal, which represents the ISI; $\mathbf{H}_{lm}^{\text{IBI}}$ shall contain the upper triangular part of \mathbf{H}_{lm} representing the IBI, such that $\mathbf{H}_{lm} = \mathbf{H}_{lm}^{\text{ISI}} + \mathbf{H}_{lm}^{\text{IBI}}$. Therefore, the received vector on the m th antenna, $1 \leq m \leq M_R$ is given by

$$\mathbf{y}_m(n) = \sum_{i=1}^{N_t} \left[\mathbf{H}_{lm}^{\text{ISI}} \mathbf{q}_l(n) + \mathbf{H}_{lm}^{\text{IBI}} \mathbf{q}_l(n-1) \right] + \mathbf{n}_m(n) \quad (3.24)$$

where $\mathbf{n}_m(n)$ is an zero-mean additive white i.i.d. Gaussian noise term.

A choice of the pseudorandom postfix ST encoder \mathbf{W} as discussed in [12] can allow for simple identification of all channels \mathbf{h}_{lm} , $1 \leq l \leq M_T$, $1 \leq m \leq M_R$.

3.1.2.2.3 Example of a 2×1 MIMO Modulation Scheme.

It is possible to apply at the transmitter the 2×1 ST encoder \mathbf{M} proposed by [60] in the ZP single-carrier context, which takes two consecutive OFDM symbols $\mathbf{x}(2i)$ and $\mathbf{x}(2i+1)$ to form the following coded matrix:

$$\begin{bmatrix} \bar{\mathbf{x}}_1(2i) & \bar{\mathbf{x}}_1(2i+1) \\ \bar{\mathbf{x}}_2(2i) & \bar{\mathbf{x}}_2(2i+1) \end{bmatrix} = \begin{bmatrix} \mathbf{x}(2i) & -\mathbf{P}_{N_c}^0 \mathbf{x}^*(2i+1) \\ \mathbf{x}(2i+1) & \mathbf{P}_{N_c}^0 \mathbf{x}^*(2i) \end{bmatrix} \quad (3.25)$$

where the permutation matrices \mathbf{P}_J^n are such that, for a $J \times 1$ vector $\mathbf{a} = [a(0), \dots, a(J-1)]^T$, $\{\mathbf{P}_J^n \mathbf{a}\}_p = a((J-1-p+n) \bmod J)$, with $0 \leq p \leq J-1$. Equation (3.25) reduces to the Alamouti ST block code if $N_c = 1$ [12].

Because the channel is known, it is always possible to retrieve the MIMO ZP-OFDM signals from (3.25) by subtracting from the received signal the known PRP contribution:

$$\mathbf{y}^{\text{ZP}}(n) = \mathbf{y}(n) - \sum_{l=1}^2 \left[\mathbf{H}_l^{\text{IBI}} \mathbf{v}_l(n-1) + \mathbf{H}_l^{\text{ISI}} \mathbf{v}_l(n) \right] \quad (3.26)$$

which leads to $\mathbf{y}^{\text{ZP}}(n) = \sum_{l=1}^2 \mathbf{H}_l \mathbf{T}_{\text{ZP}} \bar{\mathbf{x}}_l(n) + \mathbf{n}(n)$. Note that (1) no constraint has to be set on \mathbf{W} for the symbol recovery, and (2) the PRP interference cancellation procedure proposed is generic and can be applied to any ST encoder \mathbf{M} [12].

3.1.2.3 Equalization Techniques

3.1.2.3.1 SISO PRP-OFDM.

3.1.2.3.1.1 Demodulation Technique I: Overlap-Add-Based Low-Complexity Demodulation/Equalization. The idea is to convert the received PRP-OFDM block vector into a form corresponding to a standard CP-OFDM symbol after the suppression of the guard interval. For this purpose, the postfix sequence is subtracted and an overlap-add (OLA) operation is performed:

$$\mathbf{y}_{N_c}^{\text{OLA}}(i) := \underbrace{\left[\frac{\mathbf{I}_{D \times D} \mathbf{0}_{D \times (N_c - D)} (\beta_i \mathbf{I}_{D \times D})}{\mathbf{0}_{(N_c - D) \times D} \mathbf{I}_{(N_c - D) \times (N_c - D)} \mathbf{0}_{(N_c - D) \times D}} \right]_{N_c \times P}}_{\text{OLA operation}} (\mathbf{y}_P(i) - \alpha(i) \hat{\mathbf{H}}_{\beta_i} \mathbf{c}_P) \quad (3.27)$$

The resulting vector $\mathbf{y}_{N_c}^{\text{OLA}}(i)$ is treated afterwards in the receiver as a standard CP-OFDM symbol; the only difference is that the noise variance is slightly impacted by the OLA operation. The corresponding low-complexity architecture is shown in Figure 3.16.

3.1.2.3.1.2 Demodulation Technique II: MMSE-Based or Pseudochannel-Inverse-Based ZP-OFDM Demodulation/Equalization. One can first reduce (3.21) to the ZP-OFDM case by a simple subtraction of the known postfix convolved by the pseudocirculant channel matrix: $\mathbf{y}_P^{\text{ZP}}(i) := \mathbf{y}_P(i) - \alpha(i) \hat{\mathbf{H}}_{\beta_i} \mathbf{c}_P$, where $\hat{\mathbf{H}}_{\beta_i}$ is derived from the current channel estimate. In that case all known methods related to the ZP-OFDM can be applied. With an OLA approach, ZP-OFDM-OLA can attain almost the same performance and complexity as CP-OFDM [12].

3.1.2.3.1.3 Demodulation Technique III: MMSE Equalization Based on Pseudocirculant Channel Matrices/Equalization. It is also possible to directly equalize (3.21) by relying on the diagonalization properties of pseudocirculant matrices applied to \mathbf{H}_{β_i} . According to [61]

$$\mathbf{H}_{\beta_i} := \mathbf{V}_P^{-1}(i) \mathbf{D}_i \mathbf{V}_P(i) \quad (3.28)$$

where

$$\mathbf{D}_i := \text{diag}\{H(\beta_i^{-1/P}), \dots, H(\beta_i^{-1/P} e^{j2\pi(P-1/P)})\}$$

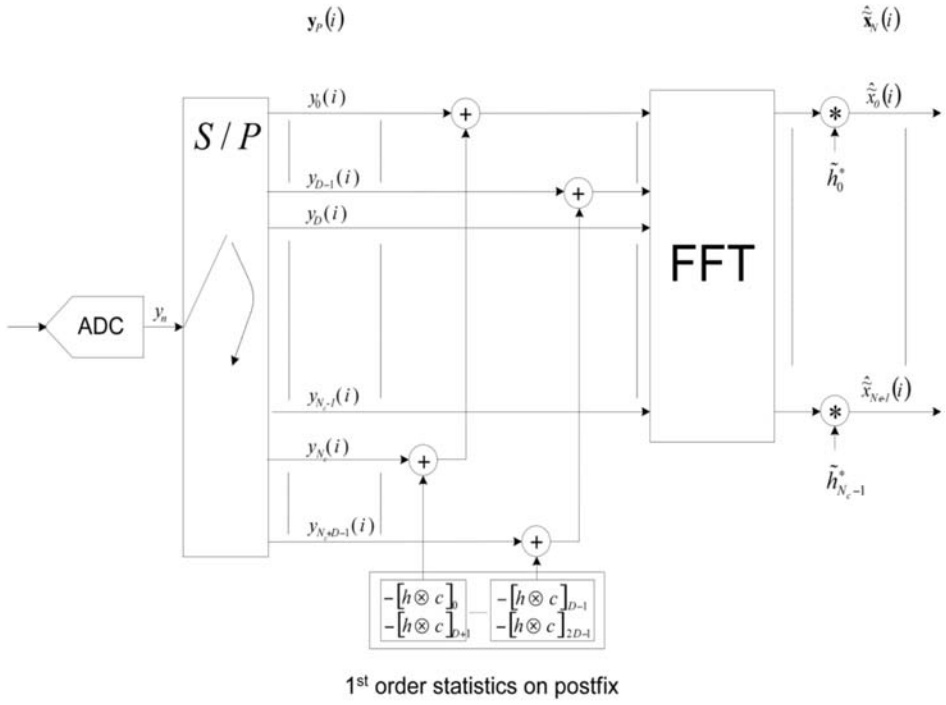


Figure 3.16 OLA decoding architecture [12].

and

$$\mathbf{v}_P(i) := \left[\frac{1}{P} \sum_{n=0}^{P-1} |\beta_i|^{-2n/P} \right]^{1/2} \mathbf{F}_P \text{diag}\{1, \beta_i^{1/P}, \dots, \beta_i^{P-1/P}\}$$

The following theorem was derived in [12] by means of standard mathematical tools.

Theorem 1. The biased equalization matrix in the MMSE sense for retrieving an estimate of $\mathbf{x}_{N_c}(i)$ in (3.21) is given by

$$\mathbf{G}_{\text{MMSE}}^{\text{PRP}} := \mathbf{F}_{N_c} [\mathbf{I}_{N_c} \mathbf{0}_{N_c, D}] \mathbf{R}_{\mathbf{n}_P} \mathbf{H}_{\beta_1}^H \mathbf{Q}^{-1} = \mathbf{F}_{N_c} [\mathbf{I}_{N_c} \mathbf{0}_{N_c, D}] \mathbf{R}_{\mathbf{n}_P} \mathbf{V}_P^H(i) \mathbf{D}_i^H \hat{\mathbf{Q}}^{-1} \mathbf{V}_P(i) \quad (3.29)$$

with

$$\begin{aligned}
\mathbf{Q} &:= \mathbf{R}_{\mathbf{n}_p} + \mathbf{H}_{\beta_1} \mathbf{R}_{s_p} \mathbf{H}_{\beta_1}^H \\
\hat{\mathbf{Q}} &:= \mathbf{R}_{\mathbf{n}_p} + \mathbf{D}_i \hat{\mathbf{R}}_{s_p} \mathbf{D}_i^H \\
\mathbf{R}_{\mathbf{n}_p} &:= \mathbb{E}[\mathbf{n}_p(i) \mathbf{n}_p^H(i)] = \sigma_n^2 \mathbf{I}_P \\
\mathbf{R}_{s_p} &:= \mathbb{E}[\mathbf{x}_p(i) \mathbf{x}_p^H(i)] \\
\hat{\mathbf{R}}_{s_p} &:= \mathbf{V}_p(i) \mathbf{R}_{s_p} \mathbf{V}_p^H(i)
\end{aligned}$$

Contrary to the noise that has a diagonal autocorrelation in the \mathbf{V}_p domain: $\mathbf{R}_{\mathbf{n}_p} = \sigma_n^2 \mathbf{I}_P$. This is no longer the case for the time-domain vector $\mathbf{y}_p(i)$ since it contains the deterministic postfix. Thus, the expression of $\mathbf{G}_{\text{MMSE}}^{\text{PRP}}$ does not allow for easy hardware implementation. To overcome this issue, the following assumption can be made resulting in a suboptimal equalizer:

$$\mathbf{G}_{\text{MMSE}}^{\text{PRP}}(i) \approx \mathbf{F}_{N_c} [\mathbf{I}_{N_c} \mathbf{0}_{N_c, D}] \mathbf{V}_p^H(i) \mathbf{D}_i^H (\sigma_n^2 \mathbf{I} + \mathbf{D}_i \mathbf{D}_i^H)^{-1} \mathbf{V}_p(i)$$

This amounts to an approximate of $\mathbb{E}[\mathbf{x}_p(i) \mathbf{x}_p^H(i)]$ by $\sigma_s^2 \mathbf{I}_P$. In the IEEE 802.11a context one can check that with QPSK constellations this yields almost identical results up to 10^{-3} BER [12].

3.1.2.3.1.4 ML Decoding Techniques for PRP-OFDM. From (3.21), after equalization by any of the $N_c \times P$ matrices \mathbf{G} presented earlier, the vector to be decoded can generally be expressed by $\hat{\mathbf{x}} := \mathbf{G} \mathbf{y}_p(i) = \mathbf{G}_d \tilde{\mathbf{x}}_N(i) + \hat{\mathbf{n}}_N$ where \mathbf{G}_d is a diagonal weighting matrix and $\hat{\mathbf{n}}_N$ the total noise plus interference contribution, which for simplicity's sake is approximated here as Gaussian and zero mean.

For ML decoding, usually a log-likelihood approach is chosen based on a multivariate Gaussian law leading to the following expression:

$$\hat{\mathbf{d}} := \arg \max_{\hat{\mathbf{d}}} \left\{ - \sum_{i=0}^{S-1} (\mathbf{G}_d m_{N_c}(\hat{\mathbf{d}}(i)) - \hat{\mathbf{x}}(i))^H \mathbf{R}_{\hat{\mathbf{n}}_N}^{-1} (\mathbf{G}_d m_{N_c}(\hat{\mathbf{d}}(i)) - \hat{\mathbf{x}}(i)) \right\} \quad (3.30)$$

where vector $\hat{\mathbf{d}}$ contains an estimation of the original uncoded information bits, $\hat{\mathbf{d}}(i)$ gathers the corresponding bits after encoding, puncturing, and so forth, within the i th OFDM symbol. S is the number of OFDM symbols in the sequence to be decoded, $m_{N_c}(\cdot)$ is an operator representing the mapping of encoded information bits onto the N_c constellations, one for each carrier of the OFDM symbol.

Thus, all that is needed to perform the decoding is an estimation of the noise covariance matrix $\mathbf{R}_{\hat{\mathbf{n}}_N}$, which requires the following derivations:

$$\hat{\mathbf{x}} = \mathbf{G} \mathbf{y}_p(i) = \mathbf{G}_d \tilde{\mathbf{x}}_{N_c}(i) + \mathbf{G}_f \tilde{\mathbf{x}}_{N_c}(i) + \alpha(i) \mathbf{G}_p \mathbf{c}_D + \mathbf{G} \mathbf{n}_p(i) \quad (3.31)$$

where \mathbf{G}_d is a $N_c \times N_c$ diagonal matrix and \mathbf{G}_f an $N_c \times N_c$ full matrix with the main diagonal being zero such that

$$\mathbf{G}_d + \mathbf{G}_f := \mathbf{G}\mathbf{H}_{\beta_i} [(\mathbf{F}_{N_c}^H)^T \mathbf{0}_{D, N_c}^T]^T = \mathbf{G}\mathbf{H}_{\beta_i} [\mathbf{I}_{N_c} \mathbf{0}_{D, N_c}^T]^T \mathbf{F}_{N_c}^H$$

where \mathbf{G}_d is an $N_c \times D$ matrix containing the last D columns of the matrix $\mathbf{G}\mathbf{H}_{\beta_i}$. Thus, $\mathbf{G}_f \tilde{\mathbf{x}}_{N_c}(i)$ represents the ISI. The total noise plus interference vector is $\hat{\mathbf{n}}_N = \mathbf{G}_f \tilde{\mathbf{x}}_{N_c}(i) + \mathbf{G}\mathbf{n}_P(i) + \alpha(i)\mathbf{G}_p \mathbf{c}_D$ and its covariance is

$$\mathbf{R}_{\hat{\mathbf{n}}_N} = \sigma_x^2 \mathbf{G}_f \mathbf{G}_f^H + \sigma_n^2 \mathbf{G}\mathbf{G}^H + \mathbf{G}_p \mathbf{c}_D \mathbf{c}_D^H \mathbf{G}_p^H \quad (3.32)$$

To deal with the non-Gaussian nature of the term $\mathbf{G}_p \mathbf{c}_D$, a nonlinear equalization scheme can be used that would suppress its contribution by substituting into (3.31) the following expression:

$$\hat{\tilde{\mathbf{x}}}^{(2)} := \mathbf{G}\mathbf{y}_P(i) - \alpha(i)\mathbf{G}_p \mathbf{c}_D \quad (3.33)$$

The overall noise covariance presented above is not diagonal, which yields to a very high complexity decoding scheme if no approximation is applied. One way to achieve a reasonable decoding complexity is to approximate $\mathbf{R}_{\hat{\mathbf{n}}_N}$ by a matrix containing only its main diagonal elements and use standard OFDM Viterbi decoding. In that case (3.30) reduces to the classical weighted summation of the Euclidian distances by the inverse noise variances. In the following we call these weighted Euclidian distances the Viterbi metrics. Several further simplifications are discussed in the sequel.

3.1.2.3.2 Equalization Techniques for MIMO PRP-OFDM.

Based on the 2×1 MIMO modulator given as an example in Section 3.1.2.2.3, it is always possible to retrieve the MIMO ZP-OFDM signals by subtracting from the received signal the known PRP contribution:

$$\mathbf{y}^{ZP}(n) = \mathbf{y}(n) - \sum_{i=1}^2 [\mathbf{H}_l^{\text{IBI}} \mathbf{v}_l(n-1) + \mathbf{H}_l^{\text{ISI}} \mathbf{v}_l(n)] \quad (3.34)$$

which leads to

$$\mathbf{y}^{ZP}(n) = \sum_{i=1}^2 \mathbf{H}_l \mathbf{T}_{ZP} \bar{\mathbf{x}}_1(n) + \mathbf{n}(n)$$

Note that (1) no constraint has to be set on \mathbf{W} for the symbol recovery, and (2) the PRP interference cancellation procedure proposed is generic and can be applied to any ST encoder \mathbf{M} .

The same detection algorithm as presented for the SISO case can be applied to the signal in (3.34). As $\mathbf{P}_P^N \mathbf{T}_{ZP} = \mathbf{T}_{ZP} \mathbf{P}_{N_c}^0$, then $\mathbf{D}_1 = \text{diag}\{\mathbf{h}_1\}$,

$\mathbf{D}_2 = \text{diag}\{\mathbf{h}_2\}$, $\mathbf{n}(2i) = \mathbf{F}_P \mathbf{n}(2i)$ and $\mathbf{n}(2i+1) = \mathbf{F}_P \mathbf{P}_P^N \mathbf{n}^*(2i+1)$. If it is switched to the frequency domain by computing $\mathbf{y}(2i) = \mathbf{F}_P \mathbf{y}^{ZP}(2i)$ and $\mathbf{y}(2i+1) = \mathbf{F}_P (\mathbf{P}_P^N \mathbf{y}^{ZP}(2i+1))^*$, and exploiting the fact that $\mathbf{H}_l = \mathbf{F}_P^H \mathbf{D}_l \mathbf{F}_P$, $1 \leq l \leq 2$, then:

$$\begin{bmatrix} \mathbf{y}(2i) \\ \mathbf{y}(2i+1) \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{D}_1 & \mathbf{D}_2 \\ \mathbf{D}_2^* & -\mathbf{D}_1^* \end{bmatrix}}_{=\mathbf{D}} \begin{bmatrix} \mathbf{F}_P \mathbf{T}_{ZP} \mathbf{x}(2i) \\ \mathbf{F}_P \mathbf{T}_{ZP} \mathbf{x}(2i+1) \end{bmatrix} + \begin{bmatrix} \mathbf{n}(2i) \\ \mathbf{n}(2i+1) \end{bmatrix} \quad (3.35)$$

where \mathbf{D} is an orthogonal channel matrix. Thus multiplying $[\mathbf{y}(2i)^T, \mathbf{y}(2i+1)^T]^T$ by \mathbf{D}^H achieves the separation of the transmitted signals $\mathbf{x}(2i)$ and $\mathbf{x}(2i+1)$ and it can be shown [60] that full transmit diversity is achieved. The separation of the signals allows for the same equalization scheme to be used as in the single-antenna case.

3.1.2.3.3 Design Issues on the PRP-OFDM Postfix Sequence.

Reference [12] provided recommendations for the design of the PRP-OFDM postfix and the choice of the pseudorandom weighting sequence. First it is desirable for the introduction of the pseudorandom postfix to result in a flat spectrum of the signal sent onto the channel. To analyze the spectral properties of the PRP-OFDM signal and because the signal is obviously not stationary but cyclostationary with periodicity P (duration of the OFDM block), the order zero cyclo spectrum of the transmitted time-domain sequence $s(k)$, $k \in \mathbf{N}$ has to be calculated:

$$S_{s,s}^{(0)}(z) = \sum_{k \in \mathbf{Z}} z^{-k} \frac{1}{P} \sum_{l=0}^{P-1} R_{s,s}(l, k) \quad (3.36)$$

with $R_{s,s}(l, k) = E[s_{l+k} s_l^*]$. Hereby, $R_{s,s}(l, k)$ is given for the symbol $s(k = 0 \dots P-1)$ as

$$R_{s,s}(l, k) = \begin{cases} E[s_{l+k} s_l^*] & \text{for } k+l \geq 0 \text{ and } k+l < P \\ s_{l+k} s_l^* E_\alpha & \text{for } k+l \geq mP \text{ and } k+l < mP+D, m \in \mathbf{Z}/\{0\} \\ 0 & \text{otherwise} \end{cases} \quad (3.37)$$

with

$$E_\alpha = E \left[\alpha \left(\left\lfloor \frac{l+n}{P} \right\rfloor \right) \alpha^* \left(\left\lfloor \frac{l}{P} \right\rfloor \right) \right]$$

Now it is clear that it is desirable to choose $\alpha(i)$, $i \in \mathbf{Z}$ such that $E_\alpha = 0$ in order to clear all influence of the deterministic postfix in the second order statistics of the transmitted signal. This is achievable by choosing $\alpha(i)$ as a pseudorandom, zero-mean value.

To specify the content of D samples composing the postfix, the following criteria can be considered:

- Minimize the time-domain PAPR.
- Minimize out-of-band radiations; that is, concentrate signal power on useful carriers.
- Maximize spectral flatness over useful carriers since the channel is not known at the transmitter (do not privilege certain carriers).

The resulting postfix is obtained through a multidimensional optimization involving a complex cost function.

3.1.2.3.4 Assessment Results.

Simulation results for the BER were obtained by [2, 12] for a MIMO system with two transmitting antennas and one receiving antenna in the context of the 5.2-GHz IEEE 802.11a WLAN standard. All simulations have been performed for QPSK symbols, a code rate of $R = 1/2$ and uncorrelated ETSI BRAN-A channels with mean unit power. The frame length is set to 432 bytes of uncoded data, each transmitted over 2,000 channel realizations. Different mobility conditions have been considered between 0 and 30 meters per second (mps). At the receiver side, the channel power delay profile, the Doppler frequency, and the noise variance are assumed to be known. The following two systems were compared:

- Modified IEEE 802.11a standard using Alamouti STBC (CP-OFDM), with zero-forcing (ZF) equalization and MMSE MIMO channel estimation based on preambles;
- MIMO PRP-OFDM modulator with ZP-OFDM decoding, with MMSE equalization, channel estimation based on 41 received Alamouti blocks of 2 PRP-OFDM symbols (from 20 before to 20 after the latest block). No preambles are introduced.

Without mobility and knowing the channel power delay profile and noise variance in the receiver, CP-OFDM and PRP-OFDM using MIMO channel estimation show approximately the same performance as CP-OFDM with a perfectly known MIMO channel. This is shown in Figure 3.17.

In the presence of mobility and without tracking, CP-OFDM experiences an error floor at a BER of $2 \cdot 10^{-4}$ and 10^{-2} for a speed of 10 and 20 mps, respectively. In turn, at 30 mps, PRP-OFDM shows practically no performance loss compared to the static case up to a BER of $5 \cdot 10^{-4}$. At a BER of 10^{-4} , the loss is limited to approximately 2 dB. This robustness is obtained thanks to the PRP-OFDM modulation, for which the MIMO channel tracking method applies.

3.1.2.4 Filtered Multitone Modulation

Although the OFDM system is a promising candidate for a next generation air interface, it displays two main drawbacks, which are particularly critical in uplink applications, namely, the sensitivity to frequency synchronization errors and the

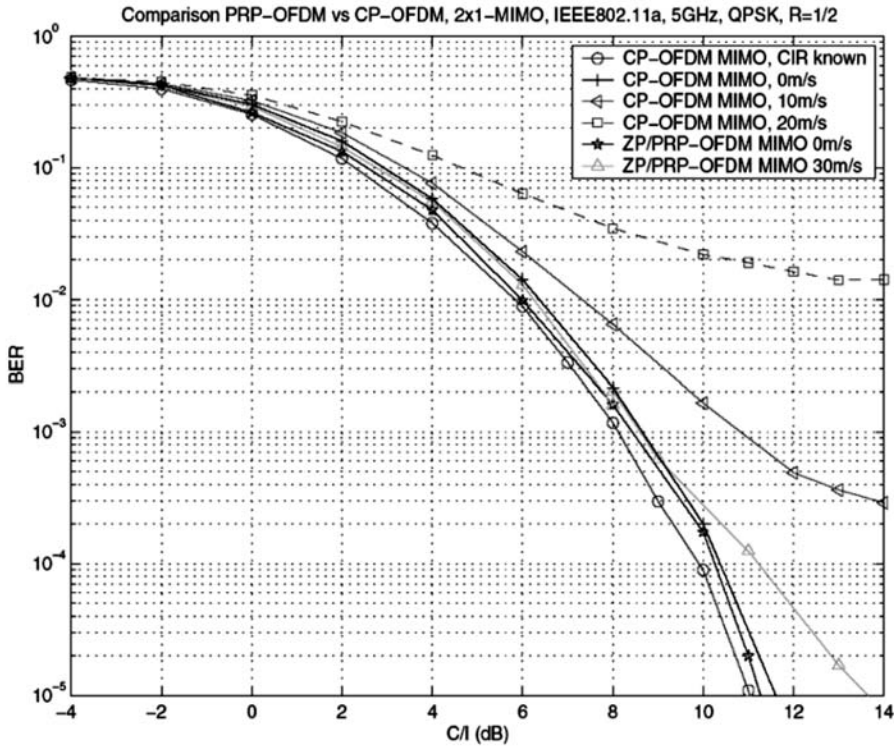


Figure 3.17 MIMO-PRP-OFDM simulation results [12].

high PAPR. OFDM belongs to the general family of filter-bank multicarrier (FB-MC) modulation systems [62], in which parallel data streams are filtered by interpolating filters with proper spectral characteristics, prior to subcarrier modulation. If the filters are all frequency-shifted versions of one prototype filter and the subcarriers are equally spaced, the FB-MC scheme can be efficiently implemented through a bank of polyphase filters and FFT.

In OFDM, the prototype filter has a rectangular shape in the time domain of length equal to the OFDM symbol duration T . Therefore, the polyphase filtering disappears in the digital implementation, thus enabling a certain complexity reduction. On the other hand, the resulting sinc-shaped subcarrier spectra strongly overlap. This is not an issue as long as the subcarrier spacing is exactly equal to $1/T$ [12]. In the presence, however, of frequency dispersion arising from time and frequency synchronization errors as well as Doppler effects and phase noise, ICI occurs. This may cause severe performance degradation especially in asynchronous multiuser applications (e.g., in the uplink of OFDM-based multiple-access systems, such as OFDMA and MC-CDMA).

In a general model for FB-MC modulation systems, ideally, the selected prototype filter should be well localized in both the time and frequency domains to guarantee the complete elimination of both ISI and ICI [63]. Suitable quality measures can be introduced, which, besides the most traditional BER, allow for assessing conveniently the trade-off between the system robustness to the time and

frequency dispersion of the radio channel and the amount of transmitted data (i.e., the achievable throughput), for different choices of the prototype filter [12]. The advantages envisaged in a special FB-MC scheme are referred in [12] as filtered multitone (FMT), whose prototype filter has a well-localized shape in the frequency domain. FMT had been traditionally discussed for very high data rate digital subscriber line (VDSL) wireline communications [64], and also has recently received attention for wireless applications [65, 66]. The high spectral containment of the adjacent subcarriers guarantees negligible ICI after transmission through the channel, without the insertion of a cyclic prefix. Preliminary studies carried out in [67] and [68] had shown that FMT could provide higher robustness to time and frequency offsets than OFDM and, as a consequence, higher spectral efficiency. The IST project WINNER [2] introduced in the design phase of the WINNER air interface a general time-continuous FB-MC system model and framed OFDM as a special case [12]. Then the equivalent discrete time model was derived, as shown in Figure 3.18.

Figure 3.19 shows the transmitting signal spectrum in case of root-raised cosine (rrcos) with roll-off factor $\rho = 0.17$ for both critical sampling and noncritical sampling with $K/N_c = 18/16$, where K is the oversampling factor, and N_c is the number of subcarriers.

We can observe how the level of crossover of adjacent subchannel spectra is reduced by means of the noncritical sampling. The high spectral containment of FMT systems makes them particularly robust to time and frequency synchronization offsets. It might be argued that this is achieved at the expense of a loss in bandwidth

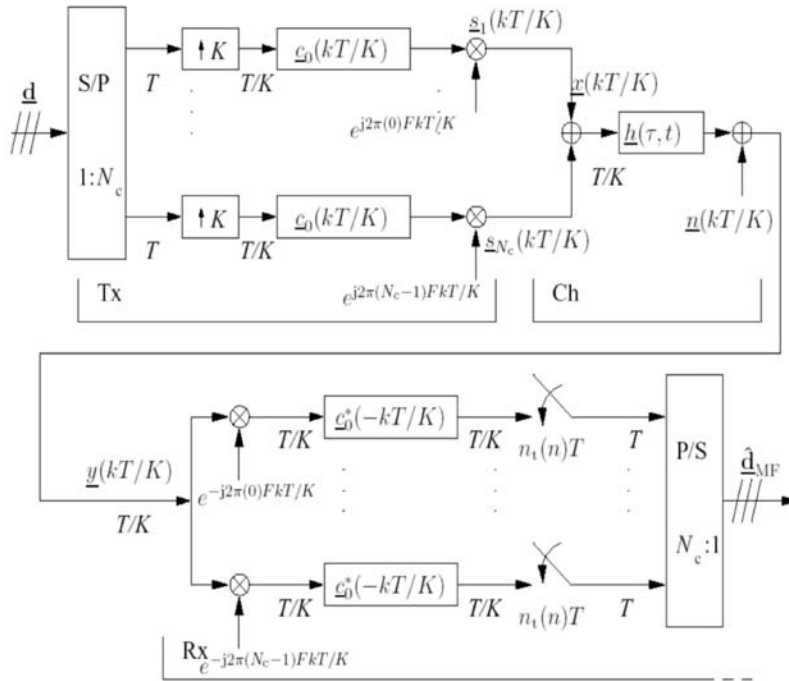


Figure 3.18 Discrete time system model [12].

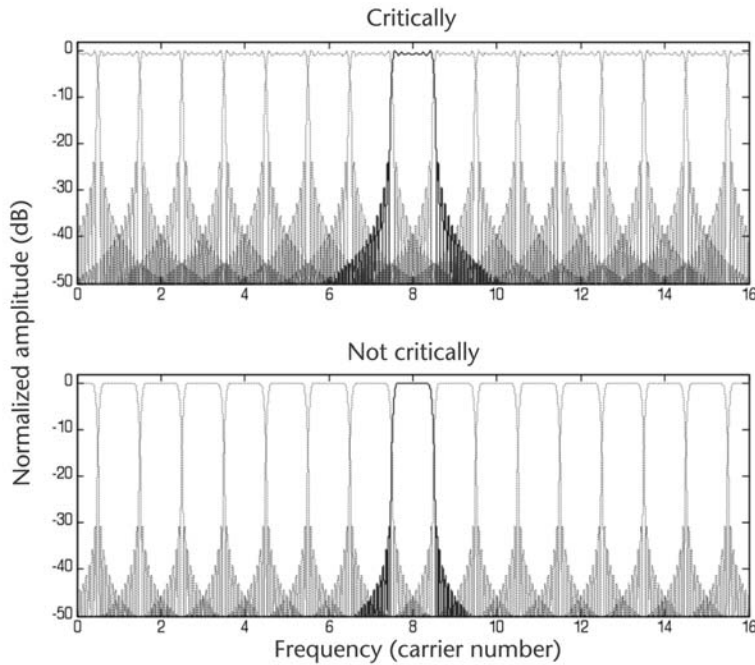


Figure 3.19 Spectrum shape in FMT system with *rrcos* prototype filter [12].

efficiency caused by the practical transition region of the filter frequency response and by the noncritical sampling. Indeed, by letting $K \rightarrow N_c$, the penalty in bandwidth efficiency becomes vanishingly small at the price of an increase in the filter implementation complexity, whose spectral roll-off has to become sharper and sharper [69]. Researchers have shown, however, that for the same redundancy in bandwidth, introduced in the form of a CP in OFDM systems, and an excess bandwidth in FMT systems, FMT-FDMA results in higher achievable bit rate than OFDMA in the presence of frequency dispersion [67, 68].

In the design of the prototype filter described here, the objective of large spectral containment is more easily achieved by relaxing the perfect reconstruction constraint, which may imply the need for equalization to remove ISI [65]. In this context, *rrcos* was considered a good choice [12], because it satisfied the Nyquist criterion in the time domain. Hence, at the receiver the transmitting filter $c_0(t)$ does not need to be equalized.

Thanks to the high spectral containment, spectral efficiency is gained by FMT in comparison to OFDM also because a lower number of virtual subcarriers has to be considered [70].

An efficient digital implementation can be obtained for FMT systems in terms of FFT as well as for OFDM systems, for both critical and noncritical sampling, under the following assumptions [12]:

- The filter on the n th subchannel is obtained through a frequency shift of the prototype filter.
- The subcarriers are equally spaced across the bandwidth.

Figure 3.20 shows the digital implementation of an FMT in the noncritical sampled case.

The transmit signal sample at time kT_c is obtained by convolving the signal at the output of the IDFT on the $(k \bmod N_c)$ th branch with the $(k \bmod N)$ th polyphase component, with respect to K , of the prototype filter. By assuming the prototype filter $c_0(t)$ to have length L over the period T_c , the single polyphase filter has length $\lceil L/K \rceil$. It follows that, in addition to the N_c -points FFT operation, $2K\lceil L/K \rceil$ real multiplications and $2K(\lceil L/K \rceil - 1)$ real additions are required for the implementation of the polyphase filtering.

3.1.2.5 Comparisons of Different FB-MC Solutions

The performance of the different FB-MC solutions was evaluated in [12] by proposing a receiver reference model as shown in Figure 3.21 and an expression of the quality measures based on two quality criteria, the rate R and the SNR degradation δ .

At the receiver $\underline{y}(t)$ is

$$\underline{y}(t) = \int_{-\infty}^{+\infty} \underline{x}(t - \tau) \underline{h}(\tau, t) d\tau + \underline{n}(t) = \sum_{n=1}^N \underline{d}_n \underline{c}_{\tau, n}(t) + \underline{n}(t) \quad (3.38)$$

where $n(t)$ represents the AWGN with psd N_0 , which is first fed into a bank of N parallel filters with each of these filters being matched to one of N data symbol specific receiving signatures $\underline{c}_{\tau, n}(t)$ as defined here:

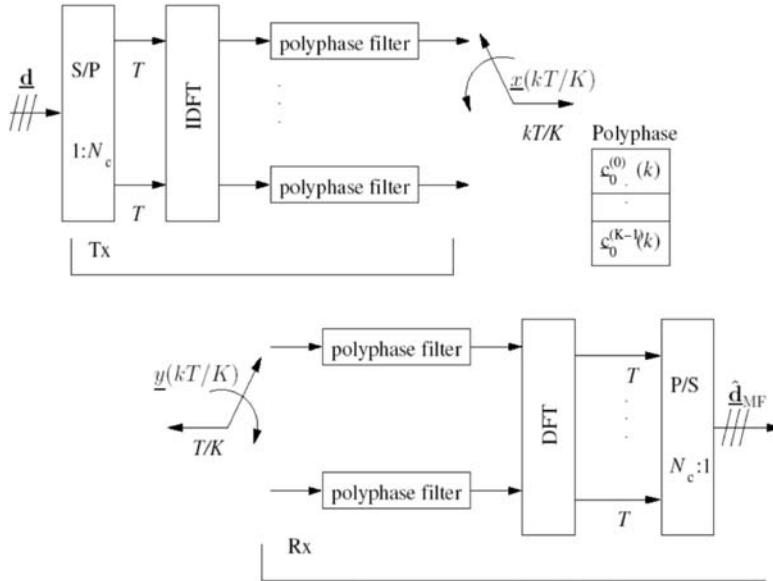


Figure 3.20 FMT digital implementation in the noncritical sampled case [12].

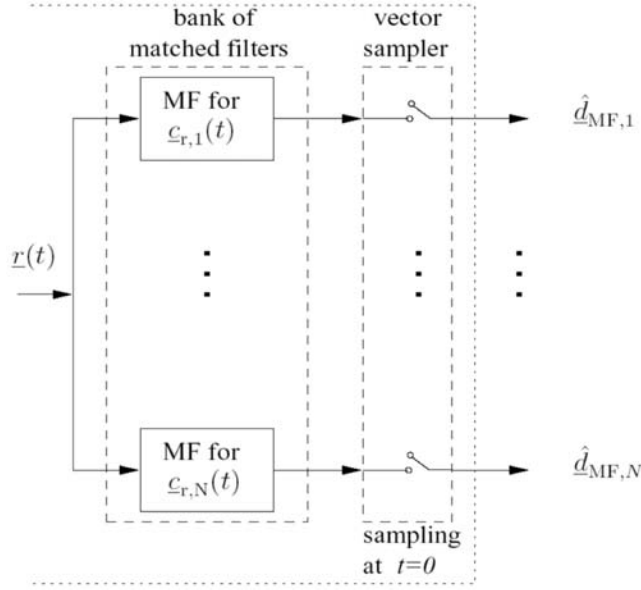


Figure 3.21 Matched filter (MF) receiver front end.

$$\underline{c}_{\tau,n}(t) \int_{-\infty}^{+\infty} \underline{c}_n(t-\tau) \underline{h}(\tau, t) d\tau \quad (3.39)$$

The filter bank plus a vector sampler from the matched filter receiver (MF Rx) front end are shown in Figure 3.21.

The impulse response of the MF for $\underline{c}_{\tau,n}(t)$ is given by

$$\underline{g}_n(\tau) = \underline{c}_{\tau,n}^*(-\tau) \quad (3.40)$$

The outputs of the N MFs are sampled at $t = 0$ by a vector sampler, thus yielding the MF estimate of a sufficient set of statistics of the data vector $\underline{\mathbf{d}}$ of the transmitted data symbols defined as $\underline{\mathbf{d}} = (\underline{d}_1 \dots \underline{d}_N)^T$:

$$\hat{\underline{\mathbf{d}}}_{MF} = (\hat{\underline{d}}_{MF,1} \dots \hat{\underline{d}}_{MF,N})^T \quad (3.41)$$

Therefore, no information on these data symbols is lost on the way from the Rx input to the outputs of the data vector sampler. In (3.41) $\hat{\underline{\mathbf{d}}}_{MF}$ is the sum of a useful part originating in the received desired signal and of a noise part $\underline{\mathbf{n}}_{MF} = (\underline{n}_{MF,1} \dots \underline{n}_{MF,N})^T$ stemming from the received noise signal $\underline{n}(t)$ of (3.38).

Under consideration of (3.40), the contribution of the transmitted data symbol $\underline{d}_{n'}$, $n' = 1 \dots N$, to the MF estimate $\hat{\underline{d}}_{MF,n}$ of \underline{d}_n can be expressed as

$$\underline{d}_{n'} \int_{-\infty}^{+\infty} \underline{c}_{\tau, n'}(t - \tau) \underline{g}_n(\tau) d\tau \Big|_{t=0} = \underline{d}_{n'} \int_{-\infty}^{+\infty} \underline{c}_{\tau, n'}(\tau) \underline{c}_{\tau, n}^*(\tau) d\tau \quad (3.42)$$

Consequently, with the noise contribution

$$\underline{n}_{\text{MF}, n} = \int_{-\infty}^{+\infty} \underline{n}(t - \tau) \underline{g}_n(\tau) d\tau \Big|_{t=0} = \int_{-\infty}^{+\infty} \underline{n}(\tau) \underline{c}_{\tau, n}^*(\tau) d\tau \quad (3.43)$$

to $\hat{\underline{d}}_{\text{MF}, n}$, it is modified to

$$\hat{\underline{d}}_{\text{MF}, n} = \sum_{n'=1}^N \underline{d}_{n'} \int_{-\infty}^{+\infty} \underline{c}_{\tau, n'}(t) \underline{c}_{\tau, n}^*(t) dt + \underline{n}_{\text{MF}, n} \quad (3.44)$$

The correlation coefficients can be defined as follows:

$$\underline{A}_{n, n'} = \int_{-\infty}^{+\infty} \underline{c}_{\tau, n'}(t) \underline{c}_{\tau, n}^*(t) dt \quad (3.45)$$

of the N data symbol specific receive signatures $\underline{c}_{\tau, n}(t)$ of (3.39). Then, (3.44) can be rewritten as

$$\hat{\underline{d}}_{\text{MF}, n} = \sum_{n'=1}^N \underline{d}_{n'} \underline{A}_{n, n'} + \underline{n}_{\text{MF}, n} \quad (3.46)$$

with $\underline{A}_{n, n'}$ of (3.45) being the $N \times N$ correlation system matrix:

$$\underline{A} = \begin{pmatrix} \underline{A}_{1,1} & \dots & \underline{A}_{1,N} \\ \vdots & \ddots & \vdots \\ \underline{A}_{N,1} & \dots & \underline{A}_{N,N} \end{pmatrix} \quad (3.47)$$

Based on (3.47), the data estimate $\hat{\underline{d}}_{\text{MF}}$ can now be expressed as

$$\hat{\underline{d}}_{\text{MF}} = \underline{A} \underline{d} + \underline{n}_{\text{MF}} \quad (3.48)$$

The element $\underline{A}_{n, n'}$ in position (n, n') of \underline{A} of (3.47) represents the contribution of the transmitted data symbol $\underline{d}_{n'}$ to the received data symbol estimate $\hat{\underline{d}}_{\text{MF}, n}$. For $n = n'$, $\underline{A}_{n, n'}$ represents the useful part. For $n \neq n'$, $\underline{A}_{n, n'}$ represents the ISI [71] generated by the presence of the transmitted symbols different from the symbol to be recovered.

From (3.43), the covariance coefficients of the noise at the output of the MF front end are determined by (3.49) also has the roman E that I queried earlier. If it is supposed to be bold or italic, please fix.)

$$E\{\underline{n}_{MF,n} \underline{n}_{MF,n'}^*\} = N_0 \int_{-\infty}^{+\infty} \underline{c}_{\tau,n'}(t) \underline{c}_{\tau,n}^*(t) dt \quad (3.49)$$

and the noise covariance matrix becomes

$$\underline{\mathbf{R}}_n = E\{\underline{\mathbf{n}}_{MF} \underline{\mathbf{n}}_{MF}^H\} = N_0 \underline{\mathbf{A}} \quad (3.50)$$

The noise covariance matrix $\underline{\mathbf{R}}_n$ is given by the system matrix $\underline{\mathbf{A}}$ scaled by the noise power spectral density N_0 . This particular structure is determined by the chosen model in which the noise contribution is colored by the matched filter with impulse response $\underline{g}_n(\tau)$.

The rate R and the SNR degradation δ represent a trade-off in the system design and, therefore, are chosen as the performance criteria [12]. By increasing the rate R , the system conveys a larger number of data symbols in less time. By increasing the SNR degradation, the system performs worse in terms of QoS. Ideally, the goal is to have a high rate R and a low SNR degradation δ . The two parameters are strictly related to each other; with increasing R , the SNR degradation δ increases. Hence, the goal of the system design is to find the optimum trade-off between the two quantities.

The parameter δ was derived for the ZF block linear equalizer [72], the rate R was independent of the receiver. Rate R represents the ratio between the number of transmitted data symbols and the total capital invested in the transmission in terms of bandwidth and time. The total occupied transmission bandwidth can be defined as

$$B_{\text{tot}} = (N_c - 1)F + B_0 \quad (3.51)$$

and the total transmit duration as

$$T_{\text{tot}} = (N_t - 1)T + T_0 \quad (3.52)$$

Because $R = \frac{N}{T_{\text{tot}} B_{\text{tot}}}$, then from (3.51) and (3.52), we obtain

$$R = \frac{N_t N_c}{[(N_t - 1)T + T_0][(N_c - 1)F + B_0]} \quad (3.53)$$

By assuming N_t and N_c large enough, (3.53) becomes $R = 1/TF$.

If the sampling theorem should be addressed, to transmit with a period T , a bandwidth of at least $1/T$ would be needed. As a consequence, to avoid interference

among adjacent subcarriers, the subcarrier spacing F should be $F \geq 1/T$. It follows that rate R would be $R \leq 1$. If this condition is not verified, a significant increment of the distortion in the received signal is expected.

The SNR degradation is defined as the ratio between the SNR at the MF output under an ideal condition (i.e., over an AWGN channel), and the SNR at the postdetector output in the presence of a multipath channel. In an ideal AWGN channel, the MF is the best possible solution with respect to the SNR [71]; therefore, the SNR degradation provides a measure of how far the detection algorithm is from the best solution in the absence of interference. From [72], if the detection algorithm is a ZF-BLE, then the mean SNR degradation with $\underline{\mathbf{A}}$ of (3.47) and $\underline{\mathbf{R}}_n$ of (3.50) is

$$\delta_{\text{ZF-BLE}} = \frac{1}{N} \sum_{n=1}^N [\underline{\mathbf{A}}^H \underline{\mathbf{R}}_n^{-1} \underline{\mathbf{A}}]_{n,n} [(\underline{\mathbf{A}}^H \underline{\mathbf{R}}_n^{-1} \underline{\mathbf{A}})^{-1}]_{n,n} \quad (3.54)$$

which is valid only for the proposed system modeling. In the OFDM case, the simplification of (3.54), given by

$$\tilde{\epsilon}_{\tau,n}(t) = \underline{H}_{n_c(n)} \frac{1}{\sqrt{T}} \text{rect}\left(\frac{t - (n_t(n) - 1)T}{T - \tau_{\max}}\right) \exp[j2\pi(n_c(n) - 1)Ft] \quad (3.55)$$

is also valid, although it might require slight modifications [12].

To measure the degradation of the conventional OFDM systems due to the deployment of the guard interval τ_{\max} , the SNR degradation $\tilde{\delta}$ can be defined as the ratio between the SNR γ_{MF} obtained by the optimum solution of the proposed modeling, and the SNR $\tilde{\gamma}_{\text{MF}}$ obtained by the conventional OFDM system. With this characterization the degradation of the conventional OFDM system when deploying the MF can be defined as

$$\tilde{\delta}_{\text{MF}} = \frac{\gamma_{\text{MF}}}{\tilde{\gamma}_{\text{MF}}} \quad (3.56)$$

The degradation of conventional OFDM when deploying the ZF-BLE can be defined as [12]

$$\tilde{\delta}_{\text{ZF-BLE}} = \frac{\gamma_{\text{MF}}}{\tilde{\gamma}_{\text{ZF-BLE}}} \quad (3.57)$$

The diagonal structure of the system matrix implies that the degradation is always the same for the different detection schemes so that

$$\tilde{\delta}_{\text{MF}} = \tilde{\delta}_{\text{ZF-BLE}} \quad (3.58)$$

The degradation of conventional OFDM does not depend on the detection algorithm because the deployment of the guard interval avoids a priori the presence of ISI and ICI.

FMT is much more robust than OFDM in the presence of frequency dispersion, while still enabling an efficient DFT-based digital implementation. However, the PAPR issue is not solved by FMT in case the number of subcarriers is chosen high enough to guarantee flat fading on each subchannel as in OFDM. Note that, in this respect, FMT represents a valid alternative also to the single-carrier frequency-domain equalization (SC-FDE) approach that has been proposed for uplink applications in order to avoid high PAPR at the mobile terminal transmitters. The number of subcarriers may be chosen in such a way as to subdivide the broadband channel in many wideband-to-narrowband subchannels, in which many signals are transmitted in parallel in FMT-based FDMA. The use of FMT enables a good separation of adjacent subchannel signals, so that at the receiver single-user per subchannel detection can be carried out. The signals on the different subchannels may be in turn either SC signals, for which FDE can be used at the receiver in the case of frequency-selective fading on the single subchannel, or OFDM signals, depending on the capability and the complexity affordable by the mobile terminal.

In comparison with the SC-FDE approach, the FMT-FDMA approach allows for much more flexible usage of the available broadband channel, by guaranteeing higher granularity and independence of the different users. Moreover, it has to be noticed that in each subchannel CDMA and TDMA can be applied as well [12].

Another interesting alternative to classical OFDM considered in the design of the WINNER air interface [12] was the OFDM/OffsetQAM (OQAM), because it does not require the use of a guard interval (i.e., CP), which increases its spectral efficiency. The main difference between OFDM/OQAM and classical OFDM signal generation lays in the filtering by the prototype function $g(t)$ after the IFFT. Thanks to the inverse Fourier transform (IFT), the prototype function $g(t)$ is implemented in its polyphase form, which strongly reduces the complexity of the filtering.

3.1.2.6 Comparisons of Parallel and Serial Modulation

The IST project WINNER considered further the advantages and disadvantages of other types of modulation schemes, such as serial modulation, because the frequency-domain implementation of the serial modulation transmitter can also be useful for the filtering process applied to the transmitted signal, especially if the transmitter needs the flexibility to generate an arbitrary spectrum. Equalization processing at the receiver is carried out in the frequency domain, and the resulting system can be called single-carrier with frequency domain equalization (SC-FDE). Similar to OFDM systems, SC-FDE systems exploit frequency diversity, but through the equalization process rather than relying entirely on coding. If the channel remains stationary during a block transmission, the linearly equalized channel appears to the decoder as an AWGN channel with constant SNR. The SNR is approximately averaged over the frequency band. Intrablock interleaving is unnecessary. A nonblock version, without inserted cyclic prefixes, can also be implemented by using overlap-save or overlap-add filtering at the receiver [73].

A spread-spectrum CDMA multiple-access scheme with a spreading factor K is a particular version of spread-spectrum serial modulation, which also bears a relationship to OFDMA. The prematrix operation consists of a DFT and a frequency assignment process. This scheme has a relationship with a carrier interferometry

(CI) approach, in which block-by-block serial modulation is used, but the individual pulses that transport the data symbols are multicarrier signals [12]. FDMA can be considered a special case of the above-mentioned scheme, with a spreading factor of $G = 1$, DFT block length $N_M P$, and the set of frequencies for the p th user chosen, for example, as follows:

$$F^{(p)} = \{f_p + \ell; \ell = 0, 1, \dots, N_M - 1; 0 \leq f_p \leq N_M(P - 1); 0 \leq p \leq P - 1\} \quad (3.59)$$

An FDMA signal has bandwidth N_M , plus guard bands, with no spreading. It may be located anywhere within the frequency band by an appropriate choice of f_p . Different users can have arbitrary, nonoverlapping bandwidths. Furthermore, the spectrum of a given user does not need to be contiguous. In this case, the signal bears a resemblance to an OFDMA signal; the difference is that each frequency in the FDMA signal carries a DFT component of the data symbol sequence. At the receiver, the DFT output is sampled at the appropriate frequencies' to isolate the signal spectrum of interest, FDE is applied; and the serial data stream is recovered by the application of an inverse DFT of length N_M . Figure 3.22 shows several coexisting FDMA signals, including one ($p = 2$) with a split spectrum.

Comparisons of coded linear SC-FDE systems with nonadaptively-loaded coded OFDM systems have shown that the two systems offer similar BER performance in frequency selective channels [74, 75], for the same average received SNR. For code rates of about 1/2 or less, nonadaptively-loaded coded OFDM shows a 0.5- to 1-dB average SNR gain over coded linearly equalized SC-FDE. Code rates closer to 1 tend to favor SC-FDE. The performance of SC-FDE can be further enhanced to exceed that of coded OFDM (without adaptive loading), by the addition of full-length or sparse [76] time-domain decision feedback equalization (DFE), iterative equalization [77], or turbo equalization [78]. Frequency-domain turbo processing can yield further performance improvement [79].

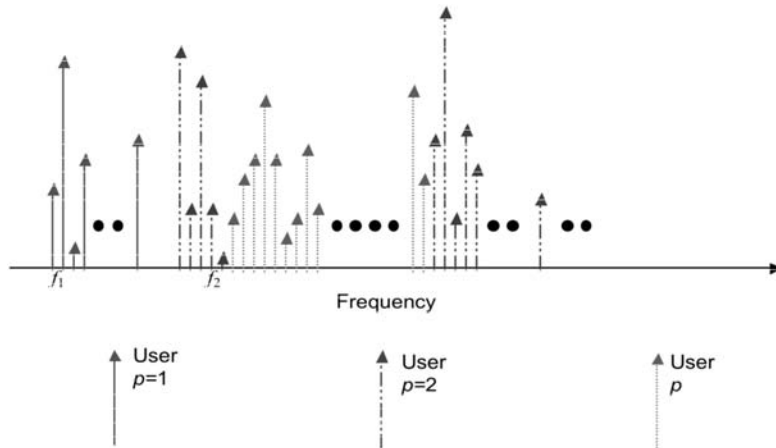


Figure 3.22 FDMA signals, including one ($p = 2$) with a split spectrum [12].

If BER performance is evaluated as a function of the *peak* bit energy to noise ratio averaged over an ensemble of channels [12], coded SC-FDE is consistently superior to coded OFDM over a wide range of code rates and QAM constellation sizes [75]. In effect, this comparison is between parallel and serial modulated systems, which are using the same power amplifier with a fixed peak transmitted power output.

Signal processing and clipping techniques are available that can be used to reduce the PAPR and spectral regrowth of parallel-modulated systems. These include selective mapping and partial transmit sequences, reference signal subtraction, coding and clipping, and filtering [12]. All of these approaches entail extra signal processing complexity, especially at the transmitter; they may also require extra overhead signaling or redundant symbols, and can also degrade performance and bandwidth efficiency.

The back-off and power amplifier efficiency are less important issues for the downlink since the cost of the base station power amplifiers is shared among many terminals. Also, because the base station usually has to transmit many signals simultaneously through a common power amplifier, the resulting PAPR will be high regardless of whether individual downlink signals are serial or parallel modulated. The power amplification issues, however, can significantly influence the cost of the wireless subscriber unit. Thus, a very sound air interface approach in cellular systems is to use *serial modulation in the uplink* and *parallel modulation in the downlink* [12, 76]. The main advantage of serial modulation for the terminal to base link is the lower peak power of its transmitted waveform and lower required power back-off.

The peak transmitter power requirement of a mobile terminal is also influenced by the multiple access method. Possible uplink multiple-access methods, with comparable spectral efficiencies, include TDMA, CDMA, and FDMA, or combinations of them. Pure TDMA requires that all terminals transmit at the same high aggregate bit rate and with the same high peak power. This places a severe cost penalty on terminals that only need to transmit uplink at very low bit rates. CDMA and FDMA are, therefore, *preferable* to pure TDMA because they allow the peak power of each terminal, as well as its average power, to be proportional to its bit rate [12].

The spread-spectrum serial-modulation signals mentioned earlier have the serial modulation, low PAPR and DS-CDMA properties, and the advantage that multiuser interference is eliminated simply by appropriate sampling of the DFT of the received signal blocks at the receiver [12]. They have two disadvantages relative to conventional serial DS-CDMA:

1. They have a lower diversity order, since such a signal occupies only the N_M subcarriers in a block, whereas a DS-CDMA signal occupies up to $N_M G$, where N_M is the number of data symbols per block, and G is the spreading factor.
2. These signals have no spreading code protection from similar signals in adjacent cells.

In an interference environment, the multiuser interference robustness of these signals more than compensates for their diversity loss. The problem of interference

from adjacent cell uplink FDOSS signals can be solved by optimal spatial combining at the base station and/or by applying direct sequence spreading to the data symbols before transmission [12].

Cognitive radio systems that share spectrum efficiently among different users and different systems should be able to transmit signals with spectra that avoid interference with the spectra of other users. Among parallel-modulated systems, OFDMA signals are the most flexible in this regard. Among serial modulated systems, the spread-spectrum signals mentioned earlier have only a limited capability in this regard, because their spectra consist of lines at equally spaced intervals. Removal of this equal-spacing constraint, however, results in FDMA signals like those shown in Figure 3.22. In particular, a spectrum such as that of user $p = 2$ in Figure 3.22 can be split into several disjoint sections, if necessary. If each such section carries the DFT of its own data sequence, the result is the parallel transmission of several serial modulated signals—as many as the number of disjoint spectral segments. For example, the $p = 2$ user in Figure 3.22 transmits two parallel signals, each serially modulated. The resulting composite signal will have a PAPR higher than that of a corresponding single serial signal, but lower than that of an OFDMA signal with more than two subcarriers [12].

The spectrum flexibility inherent in serial modulated FDMA signals is also of value for a system in which the user terminals may be of several different types, each with different bandwidth. The bandwidths of the signals transmitted and received by these terminals are then tailored to their capabilities.

One obvious drawback of FDMA signals is that they do not contain inherent spectrum spreading, and they generally occupy a smaller total bandwidth than DS-CDMA or the spread-spectrum signals mentioned here. Therefore, they must rely on coding, equalization, and/or space diversity against fading.

3.2 Design of Multiuser Space-Time Codes

Multiuser space-time code design was studied elaborately by researchers with the FP7 IST project MASCOT [9]. The results achieved allowed for the identification of the rate regions where and for each user it is optimal to employ codes designed for the single-user case. The number of receiving antennas played an important role in the code design criteria. The design of space-time/frequency codes for single-user multiantenna channels had been studied in great detail [80–82]. Past work had focused mostly on employing single-user space-time codes for each of the users and separating the users in signal space [83] or on canceling of the multiuser interference [84], which in turn can be inefficient in regard to the achievable transmission rate or performance.

The analysis carried out within the MASCOT project was based on an idea by Gallager [85] that was used to characterize the error mechanisms in two-user AWGN multiple-access channels. Depending on the transmission rate tuple, it was shown in [85] that the dominant error event is either one of the two users or both users being in error. This meant that the rate regions where single-user error events dominated could be dealt with using space-time/frequency codes designed for the single-user case, such as those given in [80–82]. The rate region in which the event

of both users (or a subset of the users in the case of more than two users) being in error dominates required new design criteria.

An important conceptual difference between the setup in [85] and the case described here originates from the fading nature of the channel, which results in two sources of errors, namely, errors due to additive noise (also present in the AWGN case) and errors due to the channel being in outage [86]. Throughout the description here, it is assumed that the block lengths are large enough for errors due to outages to dominate the error performance. For simplicity, the design procedure was restricted to the case of two users and a receiver with perfect channel state information.

3.2.1 Characterization of the Error Event Regions

We assume that each of the two users is equipped with M_T transmitting antennas and the receiver employs M_R antennas. The matrix-valued fading channel between the two users and the receiver is assumed to be frequency-selective fading with the $M_R \times 2 M_T$ transfer function given by

$$\begin{aligned} \bar{\mathbf{H}}(e^{j2\pi\theta}) &= [\mathbf{H}_1(e^{j2\pi\theta}) \mathbf{H}_2(e^{j2\pi\theta})] \\ &= \sum_{l=0}^{L-1} \bar{\mathbf{H}}[l] e^{-j2\pi l\theta}, \quad 0 \leq \theta < 1 \end{aligned} \quad (3.60)$$

with

$$\mathbf{H}_i(e^{j2\pi\theta}) = \sum_{l=0}^{L-1} \mathbf{H}_i[l] e^{-j2\pi l\theta} \quad (i = 1, 2) \quad (3.61)$$

the channel between $M_R \times M_T$ between user i and the receiver and for a perfect Rayleigh channel [86]. The receiver is assumed to have perfect knowledge of the channels of both users, whereas the transmitters do not have any CSI.

For simplicity, an N -periodic signal model is assumed, which means that the impact of the channel on the transmitted signal is described by circular convolution rather than linear convolution. Such a signal model is obtained when each of the user's OFDM and the CP (guard interval) length exceeds the channel order (see Section 3.1.2.2). If the number of OFDM tones is N (where $N \geq 2 M_T L$ is assumed throughout), the received vector signal on the n th tone is given by

$$\mathbf{y}^{(n)} = \sqrt{\frac{P}{2}} \bar{\mathbf{H}}\left(e^{j2\pi \frac{n}{N}}\right) \bar{\mathbf{c}}^{(n)} + \mathbf{z}^{(n)}, \quad n = 0, 1, \dots, N-1 \quad (3.62)$$

where

$$\bar{\mathbf{c}}^{(n)} = [(\mathbf{c}_1^{(n)})^T (\mathbf{c}_2^{(n)})^T]^T$$

and

$$\mathbf{C}_i = [\mathbf{c}_i^{(0)} \mathbf{c}_i^{(1)} \dots \mathbf{c}_i^{(N-1)}], \quad i = (1, 2)$$

are the codewords transmitted by user i . The codewords are chosen according to some probability assignment. Although the users cannot cooperate when selecting the codewords, their codebooks can still be designed jointly; that is, the design can take the multiuser nature of the problem explicitly into account [86]. This is particularly important when dealing with rate tuples that lie in the region where the dominant error event corresponds to both users being in error.

The space-time/frequency code design criteria will depend on the individual data rates R_i ($i = 1, 2$) of the two users. Therefore, the error probability is first related to the rate tuple (R_1, R_2) . The receiver employs joint ML decoding according to

$$\hat{\bar{\mathbf{C}}} = \arg \min_{\bar{\mathbf{C}}} \sum_{n=0}^{N-1} \left\| \mathbf{y}^{(n)} - \sqrt{\frac{P}{2}} \bar{\mathbf{H}} \left(e^{j2\pi \frac{n}{N}} \right) \bar{\mathbf{c}}^{(n)} \right\|^2 \quad (3.63)$$

with $\bar{\mathbf{C}} = [\mathbf{C}_1^T \mathbf{C}_2^T]^T$ and $\hat{\bar{\mathbf{C}}} = [\hat{\mathbf{C}}_1^T \hat{\mathbf{C}}_2^T]^T$. An error occurs, whenever the decision is made in favor of a codeword tuple $(\hat{\mathbf{C}}_1, \hat{\mathbf{C}}_2) \neq (\mathbf{C}_1, \mathbf{C}_2)$. In line with the reasoning in [85], three types of error events can be identified: errors of type 1 and 2, where only the codeword of user 1 or 2, respectively, is in error, and of type 3, where the codewords of both users are decoded erroneously. Denoting the corresponding error probabilities as $P_{ek|H}$ ($k = 1, 2, 3$), the total average (with respect to the random channel) error probability is given by

$$P_e = P_{e1} + P_{e2} + P_{e3} \quad (3.64)$$

where $P_{ek} = \mathcal{E}_H[P_{ek|H}]$. Depending on the desired transmission rate tuple (R_1, R_2) , one of the three terms in (3.64) dominates the total error probability P_e , leading to the dominant error event regions shown in Figure 3.23 and defined as follows: P_{ek} dominates in region $k = 1, 2, 3$. If P_{e1} or P_{e2} dominate, for each of the two users, a choice of codes designed for single-user channels is sufficient. Moreover, these codes can be chosen independently of each other. If P_{e3} dominates, a joint design of the codebooks of the two users is required.

Exact expressions for the error probability in (3.64) as a function of (R_1, R_2) are difficult to obtain and it is recommended to refer to the standard upper bound in terms of error exponents [86]. Setting $R_3 = R_1 + R_2$ yields

$$P_{ek|H} \leq e^{-NE_{rk}(R_k, \bar{\mathbf{H}})}, \quad k = 1, 2, 3 \quad (3.65)$$

where the random coding exponent $E_{rk}(R_k, \bar{\mathbf{H}})$ ($k = 1, 2, 3$) is given by

$$E_{rk}(R_k, \bar{\mathbf{H}}) = \max_{0 \leq \tau \leq 1} \max_Q (E_{0k}(\tau, \mathbf{Q}, \bar{\mathbf{H}}) - \tau R_k) \quad (3.66)$$

where \mathbf{Q} denotes the probability assignment on the transmitted codeword matrices.

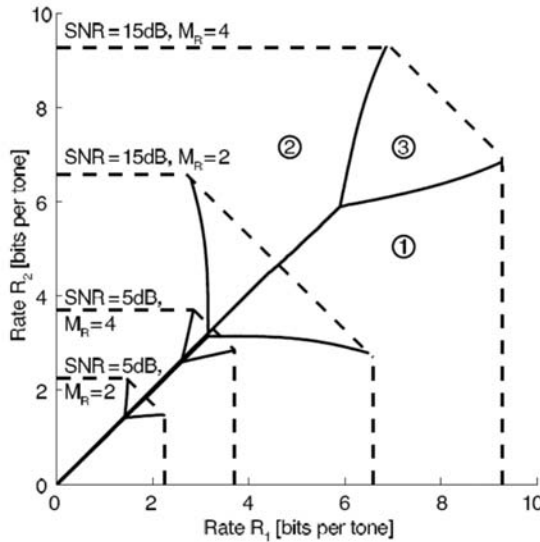


Figure 3.23 Dominant error event regions for the two-user multiple-access channel with $M_T = 2$ and $L = 1$. The dashed lines represent the ergodic capacity regions [86].

The probability density function of $E_{rk}(R_k, \bar{\mathbf{H}})$ has a nonzero probability mass at $E_{rk}(R_k, \bar{\mathbf{H}}) = 0$ because the random coding exponent becomes zero for all channel realizations that do not support the desired rate R_k (due to outage). This yields for the outage probability

$$P_{ek} = P(E_{rk}(R_k, \bar{\mathbf{H}}) = 0) \quad (3.67)$$

Figure 3.23 shows further the dominant error event regions for Gaussian codebooks (which are capacity achieving), with $M_T = 2$ and $L = 1$, and the various choices of M_R and SNR. For fixed SNR, increasing M_R results in a reduction of the relative size of the region in which both users are in error (region 3). This is due to the fact that for large M_R there are more spatial degrees of freedom and, hence, imposing “separation” through an appropriate joint code design is required only for a small set of (high) rates. From Figure 3.23, we can see that this effect is much less pronounced for low SNR, where the concept of spatial separation is not relevant [86, 87]. In summary, [86] concluded that increasing M_R results in choosing single-user codes for each of the two users being optimal in a larger fraction of the entire rate region. In the case of a general number of users, error events corresponding to any subset of users being in error can dominate and the code design has to account for that. Finally, L has hardly any impact on the shape of the dominant error event regions.

3.2.2 Code Design Criteria

The code design criteria depend on the transmission rate tuple (R_1, R_2) . Reference [86] established the general design guideline by identifying first the corresponding rate-dependent code design criteria. For this purpose, it was assumed in [86] that

the codeword matrices are composed of elements drawn from finite scalar constellations. Defining just the outage probability [see (3.67)] would not lead to any analytically tractable design criteria.

By defining $M_1 = |\mathcal{C}_1|$, $M_2 = |\mathcal{C}_2|$, and $M_3 = M_1 M_2$, the corresponding maximum transmission rates R_k^0 are given by $R_k^0 = (1/N) \log M_k$, $k = (1, 2, 3)$. The cutoff rates can be defined by choosing $\tau = 1$ in (3.66). This gives

$$E_{rk}(R_k, \bar{\mathbf{H}}) \geq E_{0k}(1, \bar{\mathbf{Q}}, \bar{\mathbf{H}}) - R_k \quad (3.68)$$

for arbitrary $\bar{\mathbf{Q}}$. Using (3.68) in (3.67) gives the following for the upper bound:

$$P_{ek} \leq \mathbb{P} \left(\frac{1}{M_k} \sum_{\substack{C_k = E_b \\ C_k \neq E_b}} e^{\frac{1}{2} \sum_{n=0}^{N-1} |\mathbf{H}_k(e^{j2\pi(n/N)})(e_k^{(n)} - e_k^{(n)})|^2} \geq e^{N(R_k^0 - R_k)} \right), \quad k = 1, 2, 3 \quad (3.69)$$

where $\rho = P/2N_0$, $\bar{M}_1 = M_1 - 1$, $\bar{M}_2 = M_2 - 1$, $\bar{M}_3 = \bar{M}_1 \bar{M}_2$, $\mathbf{H}_3 = \bar{\mathbf{H}}$, $\mathbf{C}_3 = \bar{\mathbf{C}}$, and $\mathbf{E}_3 = \bar{\mathbf{E}} = [\mathbf{E}_1^T \mathbf{E}_2^T]^T$. The notation $\mathbf{C}_3 \neq \mathbf{E}_3$ means that both $\mathbf{C}_1 \neq \mathbf{E}_1$ and $\mathbf{C}_2 \neq \mathbf{E}_2$. The upper bound in (3.69) requires tightening of the corresponding bound, which is achieved by a proper choice of $\bar{\mathbf{Q}}$.

It was shown in [86] that by keeping R_k fixed the following bound condition can be defined:

$$\lim_{\rho \rightarrow \infty} \frac{\log P_{ek}(\mathbf{R})}{\log \rho} = -\text{rank}\{\mathbf{R}\} M_R \quad (3.70)$$

which implies that the high-SNR error probability P_{ek} is dominated by the codeword difference matrices R_k with minimum rank. In the high-SNR regime, the space-time/frequency code design criteria minimizing (3.64) can be summarized as follows:

1. Given the rate tuple (R_1, R_2) , determine the type ($k = 1, 2, 3$) of an error event dominating the overall error probability P_e .
2. If the dominant error event is of type k ($k = 1, 2, 3$), the corresponding error probability P_{ek} is minimized by codes that fulfill the following criteria:
 - a. *Rank criterion:* For every codeword pair $(\mathbf{C}_k, \mathbf{E}_k)$ with $\mathbf{C}_k \neq \mathbf{E}_k$ the rank of the corresponding codeword difference matrix R_k should be maximized.
 - b. *Eigenvalue criterion:* For every codeword pair $(\mathbf{C}_k, \mathbf{E}_k)$ with $\mathbf{C}_k \neq \mathbf{E}_k$, the vector of eigenvalues of the corresponding matrix R_k should be majorized by any other possible choice of eigenvalues.

It is interesting to observe that a completely different derivation (as compared to the one in [80–82]), aimed at minimizing the probability of encountering a bad effective channel realization, would result in essentially the same criteria as those based on the pairwise error probability [82].

The joint code design would use the criteria as follows. If the rate tuple (R_1, R_2) lies in the dominant error event region 1 or 2, codes would be designed according to criteria 2a and 2b for each of the two users. The two codebooks can furthermore be chosen independently of each other.

If (R_1, R_2) lies in region 3, design criteria 2a and 2b have to be applied to the sum of the two codeword difference matrices $\mathbf{R}_3 = \mathbf{R}_1 + \mathbf{R}_2$ (joint design). The users, however, will not cooperate in selecting their codewords. In the case of a general number of users, the design rule proposed by [86] is to apply design criteria 2a and 2b to the sum of the codeword difference matrices corresponding to the subset of users leading to the dominant error event.

In a similar way, the design criteria for low SNR can be defined as well.

The importance of the joint code design for high SNR was highlighted in [86] by a numerical example. The corresponding error probabilities P_{e3} that can be obtained are shown in Figure 3.24.

The presented design criteria show that depending on the transmission rate tuple, joint code designs may or may not be necessary. It was shown that joint designs essentially require that the classical design criteria be applied to a sum of codeword difference matrices, with the specific sum depending on the subset of users leading to the dominant error event for the given transmission rate tuple. Systematic joint code designs constitute an interesting area of research. The classical (based on pairwise error probability) code design criteria were recovered in the preceding description from a criterion that essentially aims at minimizing the probability of encountering a bad effective channel realization.

3.2.3 Multiuser MIMO Downlink Space-Time Block Coded Transmission Scheme

Space-time coding was proposed in [80] by combining channel coding and transmit diversity techniques in order to achieve diversity and coding gains.

In the downlink of a MIMO multiuser system, a base station simultaneously transmits signals for many different users. Multiple antennas at both transmitters

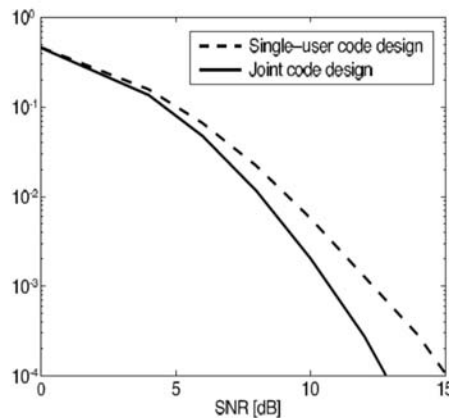


Figure 3.24 Error probability P_{e3} for single-user and joint code design as a function of SNR [86].

and receivers ensure high data rates, which are required in the next generation wireless communication systems. The FP6 IST project MASCOT focused on combining space-time block codes (STBCs) with a downlink transmission scheme for a MIMO multiuser system. The project considered a CDMA multiplexing scheme and showed how to use TAST codes [86] and perfect STBCs in the downlink of a MIMO multiuser system. Different orthogonal spreading matrices were used to separate the data streams of multiple users. At the receiver of each user, after despreading the received signal sequence, the maximum likelihood decoding was obtained by a lattice decoder. The performance of the downlink system was evaluated by simulation in terms of block error rate (BLER). Due to multipath propagation, different time-delayed versions of the transmitted signal may result in multiple access interference (MAI). The impact of MAI on system performance was analyzed by means of a simulation. The downlink of a K -user MIMO multiuser system is shown in Figure 3.25.

The base station is assumed to have n_T transmitting antennas where $n_T = \sum_{k=1}^K n_R^{(k)}$ and $n_R^{(k)}$ denotes the number of receiving antennas for the k th user. This enables support for terminals with different receiver complexity.

Because the encoder and decoder structures for any user are similar, only the general code construction was considered and lattice decoding for an arbitrary k th user. The scenario also considered square linear dispersion STBCs with full rate and full diversity. The QAM information symbols were encoded by such STBCs into an $n_T \times n_T$ codeword matrix $\mathbf{X} = \{x_{i,l}\} \in \mathbb{C}$, $i, l = 1, \dots, n_T$. The codeword matrices for each user were separated by orthogonal spreading matrices.

For the k th user, it was assumed that a codeword \mathbf{X} was transmitted. The ML receiver might decide erroneously in favor of another codeword $\hat{\mathbf{X}}$. If r denotes the rank of the codeword difference matrix $\mathbf{C} = \mathbf{X} - \hat{\mathbf{X}}$, and λ_i , $i = 1, \dots, r$ denotes the eigenvalues of the codeword distance matrix $\mathbf{A} = \mathbf{C}\mathbf{C}^+$ [86], then the pairwise error probability (PWE) is upper bounded by

$$P(\mathbf{X} - \hat{\mathbf{X}}) \leq \left(\prod_{i=1}^r \lambda_i \right)^{-n_R} \left(\frac{E_s}{N_0} \right)^{-rn_R} \quad (3.71)$$

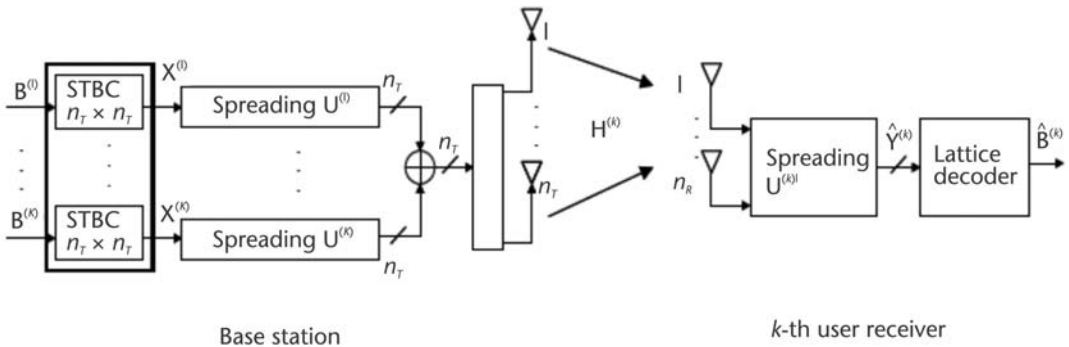


Figure 3.25 A MIMO multiuser scheme with STBC [86].

where E_s is the average energy of QAM information symbols and E_s/N_0 denotes the SNR per transmitting antenna. The minimum value of rnR in (3.71) is the

diversity gain and the minimum value of $\left(\prod_{i=1}^r \lambda_i\right)^{-n_R}$ is the coding gain. The

following criteria were used to minimize the PWEF:

- *Rank criterion:* The minimum rank r of the codeword difference matrix should be maximized.
- *Determinant criterion:* The minimum determinant of the coding gain should be maximized.

The performance was evaluated in terms of BLER by Monte Carlo simulations. The SNR per user was defined as $\text{SNR} = nRE_s/N_0$. The simulations considered two different systems.

System 1 assumed that the base station had four transmitting antennas. There are two users, each of which has two receiving antennas and 16-QAM symbols are used. It was assumed that each STBC block consists of 8 symbols per user. Perfect 4×4 STBCs were chosen using two layers for a full diversity gain. The spreading matrices were chosen to be Hadamard matrices.

Figure 3.26 shows the performance of system 1 together with the diversity gain slope. Due to multipath propagation, different time-delayed versions of the

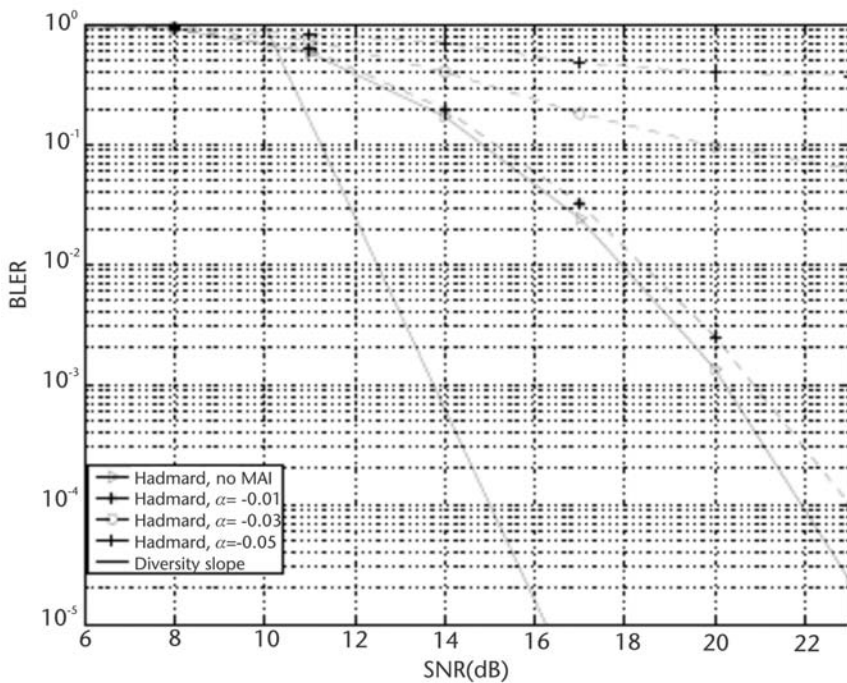


Figure 3.26 Performance of the downlink of a MIMO multiuser system with 4×4 STBCs using two layers [86].

transmitted signal may result in MAI, which destroys the orthogonality of the spreading codes.

Figure 3.27 shows the performance of a system for two users: User 1 has one receiving antenna, user 2 has three receiving antennas, and 16-QAM symbols are used. The total transmission rate per block is 16 symbols. The transmitted codeword consists of a one-layer perfect STBCs for user 1 (4 QAM symbols), and a three-layer perfect STBC for user 2 (12 QAM symbols). User 2 achieves better performance at higher SNRs thanks to its receiving antenna diversity, even if its rate is three times the rate of user 1.

The proposed MASCOT scheme is suitable to a network supporting different types of terminals due to its flexibility. The performance of STBCs depends on the amount of MAI in the system. Further, MASCOT had investigated spatially multiplexed LDPC codes, and the performance of these codes was optimized in function of the degree distributions of the bipartite graph associated with the LDPC code [86]. The process also involved comparisons of systems with different parameters, for which the spectral efficiency was computed. The maximum achievable spectral efficiency is shown in Figure 3.28 for different system loads of 1, 2, 4, and 8 as a function of E_b/N_0 .

It can be observed, that the spectral efficiency increases with increasing system load. Therefore, from a theoretical point of view, all of the redundancy should be spent for coding and no spreading should be performed. From a practical point of view, spreading can be used to keep the system load approximately constant and design an LDPC code for this load.

To obtain LDPC codes with practical parameters, MASCOT restricted the variable node degree distribution to three nonzero parameters (2, 3, and 10) and

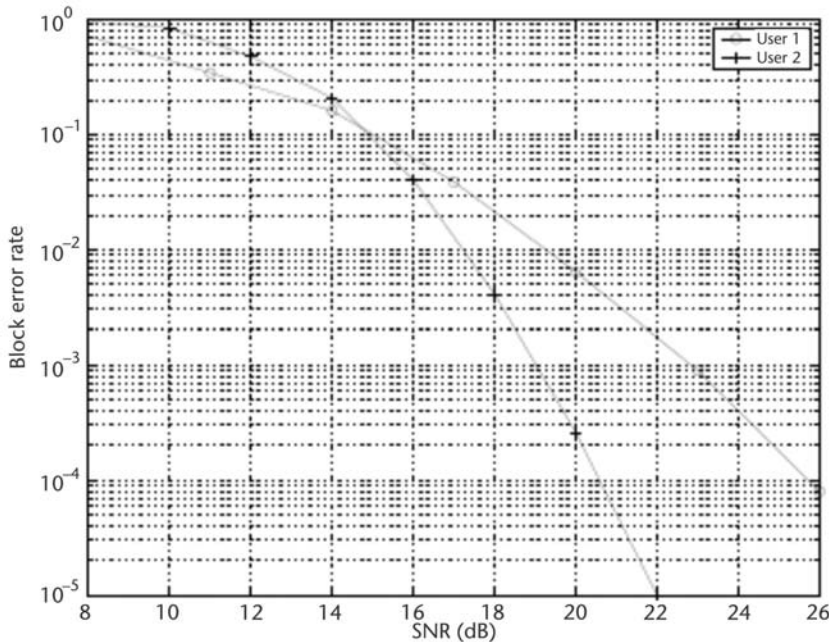


Figure 3.27 Performance of the downlink for two users with different numbers of antennas [86].

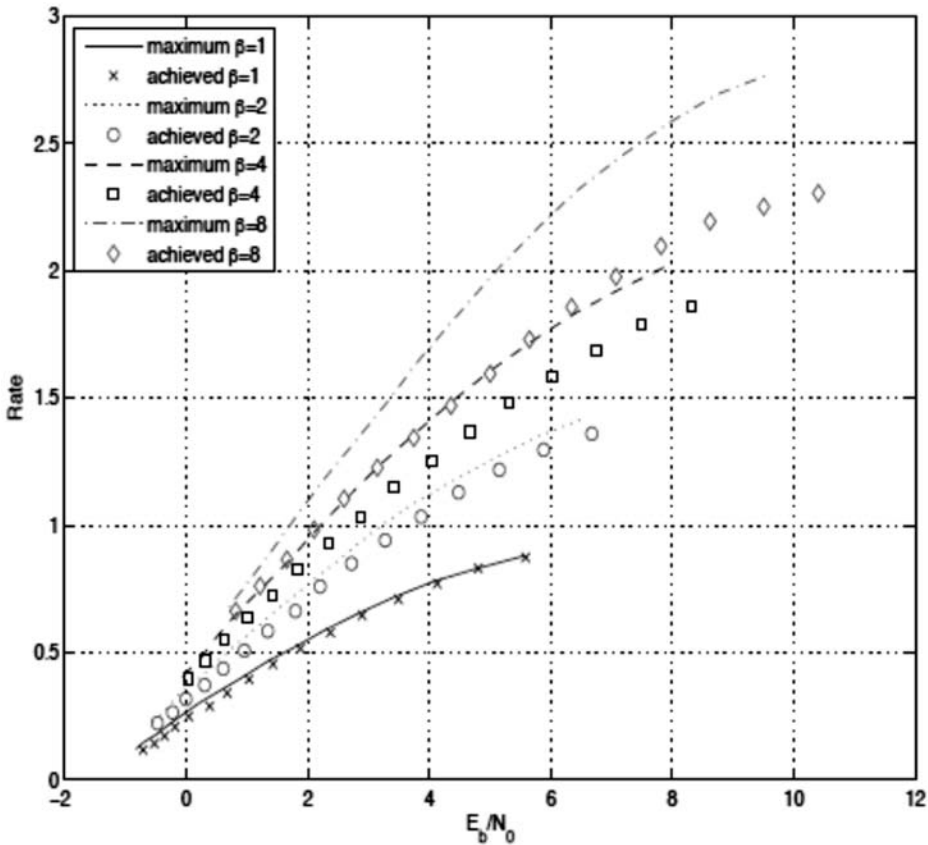


Figure 3.28 Spectral efficiency and results of LDPC code optimization [86].

set the maximum check node degree to 30. The optimization of the variable node degree distribution had been performed by means of an exhaustive search, whereas for every distribution, the check node degree distribution had been optimized using linear programming. Systems with higher load always outperform those with smaller load. For small loads the achieved spectral efficiency (see Figure 3.28) of the designed systems is close to the maximum achievable spectral efficiency. For increasing load and increasing SNR, the gap increases. This is because the shape of the EXIT function of the multiuser detector [86] cannot be matched by the LDPC code with the given parameters.

The EXIT functions of the multiuser detector and the optimized LDPC code for a system with a load of $\beta = 4$ and a spectral efficiency of 1.0, leading to a code rate $R = 0.25$, are shown in Figure 3.29.

From the EXIT chart analysis, the threshold of this system is at $E_b/N_0 = 2.75$ dB. The LDPC code in this case was optimized for an AWGN channel. This system cannot converge for the given load because the EXIT functions always intersect (even for the case of very large SNRs).

By using the area property of EXIT charts, the achievable spectral efficiency of systems of different loads can be compared. For the case where multiuser inter-

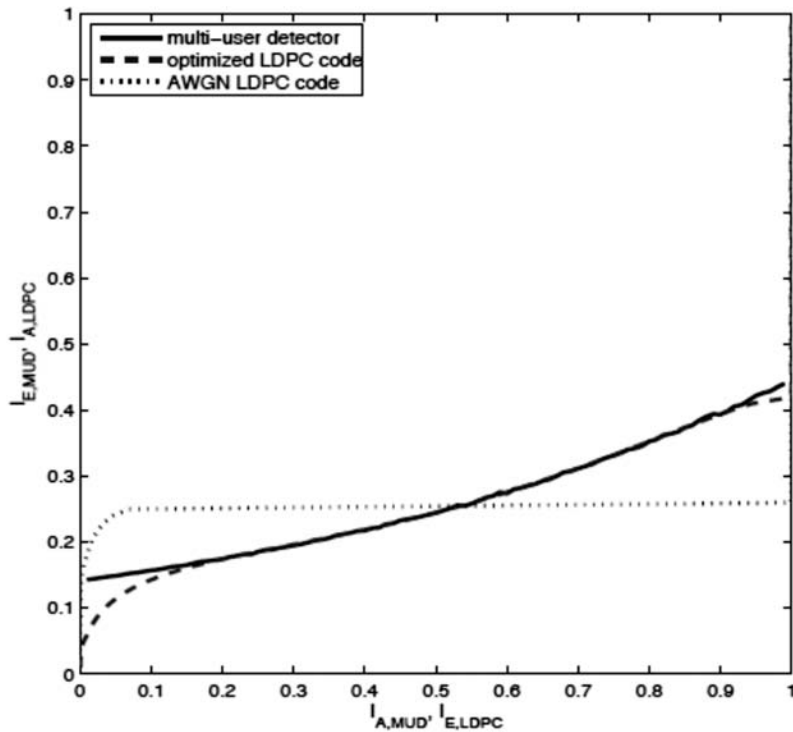


Figure 3.29 EXIT function of multiuser detector and LDPC decoder for $\beta = 4$ and $E_b/N_0 = 2.75$ dB.

ference is stronger than the interference caused by Gaussian noise, an LDPC code optimized for this scenario can significantly outperform a standard LDPC code that is optimized for an AWGN channel.

3.3 Advanced Modulation and Coding Techniques

In the scope of research on next generation communication systems, the IST project WINNER studied state-of-the-art channel coding schemes (see Sections 3.1.1 and 3.1.2) and optimized them for performance, complexity, and flexibility. As a final choice for support of an IMT-A candidate interface parallel concatenated convolutional codes (PCCC), duo-binary PCCC, low-density parity check codes (LDPC) and zig-zag LDPC were studied [88]. A practical FEC scheme suitable for a flexible and adaptive wireless system should be able to integrate well with modulation formats having varying bandwidth efficiency, as well as offer good protection to information blocks of varying sizes, and also to provide parity symbol sets suited for incremental redundancy automatic repeat request (ARQ) [89].

Two candidates for FEC coding for medium and large packet lengths were proposed in [88]; DBTCs and quasicyclic block-low-density parity-check (QC-BLDPC) codes. Refinements and optimizations of these advanced channel coding candidates can enable key features such as rate-compatibility through punc-

turing (RCP) for making full use of the advantages of an incremental redundancy (IR) hybrid-ARQ scheme (type II), together with targeting higher codeword lengths (lifting of LDPC codes) [88].

Both coding schemes yield superior performance at packet lengths above 200 information bits and can be implemented efficiently. For smaller packets (e.g., needed for broadcast control information), however, these are not applicable and a low-rate convolutional code (CC), which in [88] was considered for information lengths down to 25 information bits, was proposed instead.

The WINNER baseline design described in [90] uses rate-compatible codes with a base rate of $1/2$, and provides performance results for two block lengths, 288 and 1,152 bits. In the final system concept [91], the base code rate was changed to a rate of $1/3$, to improve the performance of retransmission schemes. Codeword error rate results were obtained for these codes.

3.3.1 Modulation and Coding Schemes with DTBCs

Figure 3.30 shows a parallel concatenated convolutional encoder and the corresponding iterative decoder. The information message u is encoded twice: directly by the encoder X_1 and a permuted version of the message by the encoder X_2 . Both encoded bitstreams, as well as the message itself, are transmitted. At the receiver side, each coded bitstream is decoded separately by a SISO decoder and the obtained information is used by the other decoder, which in turn returns new *extrinsic* information to the first decoder. After several iterations, the obtained a posteriori L -values are mapped to an estimate of the message u by hard decoding.

In Figure 3.30, the received channel symbols are scaled appropriately and are demultiplexed into the L -values corresponding to the systematic bits L_u , the L -values corresponding to the coded bits of encoder X_1 , L_{c1} , and the L -values associated with the encoded bits of X_2 , L_{c2} . The E_1 term denotes the extrinsic information of the first decoder, which becomes the a priori information A_2 for the second decoder and vice versa for E_2 and A_1 .

DBTCs are used in several standards [92, 93], and have been found to offer very good performance in conjunction with higher order modulation [94].

The main enhancement of DBTCs compared to the original turbo codes are the *component codes*, which encode two bits at a time. Both component codes are

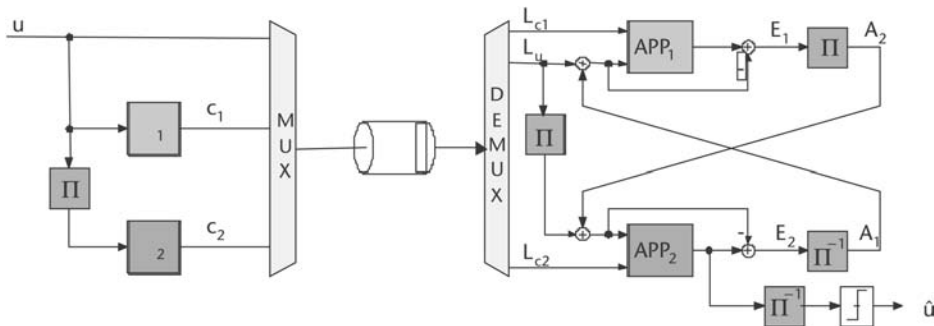


Figure 3.30 Block diagram of a parallel-concatenated turbo code [88].

identical. The term *duo-binary* refers more to the number of input bits per transition (e.g., $k_0 = 2$) but component codes are still binary convolutional codes and all operations are carried out in the binary field $\text{GF}(2)$ [88].

Figure 3.31 shows an encoder implementation that has the following transfer function matrix:

$$\mathbf{G}(D) = \begin{pmatrix} 1 & 0 & \frac{1 + D^2 + D^3}{1 + D + D^3} & \frac{1 + D^3}{1 + D + D^3} \\ 0 & 1 & \frac{1 + D + D^2 + D^3}{1 + D + D^3} & \frac{1 + D^2}{1 + D + D^3} \end{pmatrix} \quad (3.72)$$

The component codes are *recursive*, which is valid for all parallel turbo codes. As a characteristic, turbo codes have relatively simple codes with low memory component codes. This is also valid in the description here, because the component codes defined by $\mathbf{G}(D)$ have only $S = 8$ states.

Figure 3.32 shows the state transition table and the trellis diagram that were used for the implementation of the DBTC.

The turbo encoder comprises two component encoders as shown in Figure 3.31, the mother code rate of the turbo encoder is $1/3$, because for each input bit couple $(u^{(1)}, u^{(2)})$, which is also transmitted, two encoded bits couples $(c_1^{(1)}, c_1^{(2)}), (c_2^{(1)}, c_2^{(2)})$ are produced.

To make the DTBC rate compatible, the code rates of $1/3$, $2/5$, and $4/7$ were added to the initial DTBC analysis results reported in [95]. The puncturing patterns, which were applied to both component encoders, are shown in Table 3.2. The rows of the puncturing matrix correspond to the outputs $c^{(1)}$ and $c^{(2)}$ of Figure 3.32.

Note that other patterns might lead to better performance; however, no significant differences are expected [88].

For the termination of convolutional codes, three possibilities were reported in [88]:

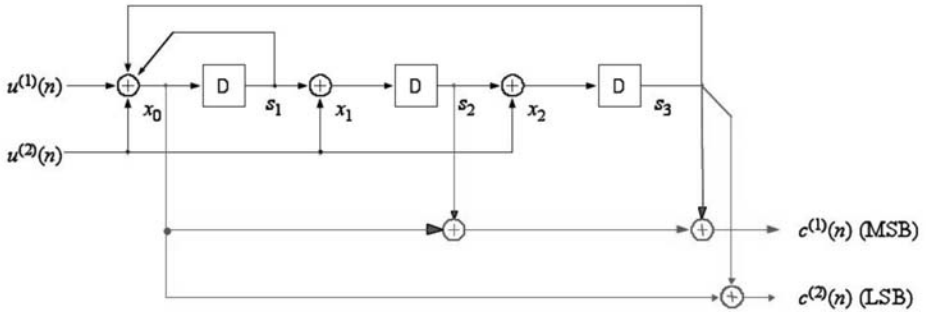


Figure 3.31 Example of an encoder implementation of a component code for DTBC [88].

1. *Truncation*: Encoding is terminated with the last bit of the message and the encoder terminates in a state that is not known at the receiver side.
2. *Zero-termination*: After encoding of the message, some tails bits are appended that drive the encoder state to zero.
3. *Tail-biting*: The initial state is chosen such that it coincides with the final state.

Truncation leads to poor protection of the last bits of the message, which leads to an error floor and should, therefore, be avoided. Zero-termination offers the same protection for all bits of the message, but requires some extra bits to drive the encoder states to zero. Tail-biting combines the advantages of the previous ones and results in equal protection for all bits, while maintaining the code rate. The encoding and decoding complexity increase only to a small extent: Although the encoding procedure has to be performed twice, this has no significant impact because turbo encoding is very simple with respect to decoding. At the decoder, tail-biting can be considered by updating and storing the forward and backward metrics of the first trellis section after each iteration.

The initial state for each component encoder is determined by the following procedure:

1. Initialize the encoder with state $s_0 = 0$ and encode the sequence u . The encoder arrives at state $s_{K/2}$. Discard the encoder output.
2. Initialize the encoder with state $s_0 = S_{\text{circ}}(n, s_{K/2})$ according to Table 3.3 where $n = \text{mod}(K/2, 7)$. The table contains no entry for $n = 0$, which means that K must not be a multiple of 14.

The circular state $s_0 = s_{K/2}$ is unknown at the receiver. This, however, does not lead to a noticeable increase of the decoder complexity. The simplest method to consider tail-biting in the decoder is to adequately initialize the metrics in the BCJR decoder and to update them after each iteration [88].

The interleaver is defined in two steps: first the bit couples $(u^{(1)}(n), u^{(2)}(n))$, are considered as one symbol and are permuted; then every second couple is reversed. Alternatively, the interleaver can be treated as any other bitwise interleaver by the following definition [88]: Consider initially the sequence $(p_0, p_1, \dots, p_{K/2-1})$, which is a permutation of $(0, 1, \dots, K/2 - 1)$ and is given by

Table 3.3 Determination of the Initial State

$s_{K/2}$								
n	0	1	2	3	4	5	6	7
1	0	6	4	2	7	1	3	5
2	0	3	7	4	5	6	2	1
3	0	5	3	6	2	7	1	4
4	0	4	1	5	6	2	7	3
5	0	2	5	7	1	3	4	6
6	0	7	6	1	3	4	5	2

$$p_k = \text{mod} \left(\left\{ \begin{array}{ll} kP_0 + 1 & \text{if } \text{mod}(k, 4) = 0 \\ kP_0 + 1 + K/4 + P_1 & \text{if } \text{mod}(k, 4) = 1 \\ kP_0 + 1 + P_2 & \text{if } \text{mod}(k, 4) = 2 \\ kP_0 + 1 + K/4 + P_3 & \text{if } \text{mod}(k, 4) = 3 \end{array} \right\}, \frac{K}{2} \right), \quad k = 0, 1, \dots, \frac{K}{2} - 1 \quad (3.73)$$

Based on this sequence, define

$$\mathbf{q}_k = \begin{pmatrix} q_{k,1} \\ q_{k,2} \end{pmatrix} = 2p_k + \text{mod}_2 \left(\begin{matrix} k+1 \\ k \end{matrix} \right), \quad k = 0, 1, \dots, \frac{K}{2} - 1 \quad (3.74)$$

Obtain the interleaver $\boldsymbol{\pi}$ as

$$\boldsymbol{\pi} = (q_{0,1}, q_{0,2}, q_{1,1}, q_{1,2}, \dots, q_{(K/2)-1,1}, q_{(K/2)-1,2}) \quad (3.75)$$

which is a permutation of $(0, 1, \dots, K - 1)$.

The parameters for different FEC block lengths are given in Table 3.4.

Figure 3.33 shows the DBTC decoder. The a posteriori probability (APP) decoders are implemented with the BCJR algorithm [26] and compute the APP L -value corresponding to each information bit. After subtracting the corresponding

Table 3.4 Interleaver Parameters for Different Codeword Lengths

K	P_0	P_1	P_2	P_3
48	5	0	0	0
72	11	18	0	18
96	13	24	0	24
144	11	6	0	6
192	7	48	24	72
216	11	54	56	2
240	13	60	0	60
288	17	74	72	2
360	11	90	0	90
384	11	96	48	144
432	13	108	0	108
480	13	120	60	180
576	23	44	212	112
768	29	74	136	86
864	13	0	4	8
960	13	240	120	360
1152	31	42	232	18
1296	41	38	196	22
1536	41	90	124	70
1728	37	62	28	94
1920	13	480	240	720
2304	49	8	196	24
2880	17	720	360	540
3072	55	36	160	32
3456	59	48	168	76
3840	17	960	480	1440
4800	17	1200	600	1800

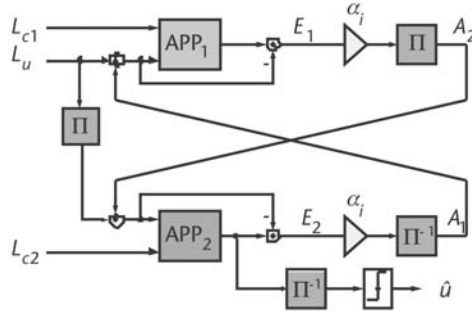


Figure 3.33 DTBC decoder [88].

a priori information, this value is passed as extrinsic information via the interleaver/deinterleaver to the other decoder. The current implementation uses 8 iterations according to the trade-off complexity/performance reported in [96].

The APP decoders are implemented with the max-log approximation reported in [97]; the extrinsic L -values E_1, E_2 are scaled by the factor α_i , which depends on the following iteration:

$$\alpha_i = \begin{cases} 0.5 & \text{for } i = 1 \\ 0.75 & \text{for } i = 2, \dots, 7 \\ 1 & \text{for } i = 8 \end{cases} \quad (3.76)$$

A slightly better performance can be achieved by replacing the max-log approximation with the exact APP calculation and by increasing the number of iterations. This is, however, at the cost of a higher decoder complexity [88].

Figure 3.34 shows the codeword error rate (CWER) for DBTC coding with QPSK modulation for the message sizes $K = 288$ and $K = 1,152$ over an AWGN channel. Figures 3.35 and 3.36 show the CWER for DTBC coding with 16-QAM and 256-QAM modulation, respectively, and the same message sizes. The curves for BPSK can be derived by a simple SNR shift, as follows. Let $\text{CWER}_b\left(\frac{E_S}{N_0}\right)$ denote the word error probability for 2^b -QAM. Then, write for BPSK: $\text{CWER}_1\left(\frac{E_S}{N_0}\right) = \text{CWER}_2\left(2\frac{E_S}{N_0}\right)$, which corresponds to a 3-dB shift of the curves.

If the simulated SNRs and corresponding CWERs are denoted by $\gamma_1, \gamma_2, \dots, \gamma_n$ and p_1, p_2, \dots, p_n , then the word error probability is approximated by an exponential function with two parameters as follows:

$$\tilde{p}_w^{(c)}(\gamma) = \begin{cases} 1 & \text{for } \gamma \leq \gamma_0^{(c)} \\ \exp(-\alpha^{(c)} \gamma - \gamma_0^{(c)}) & \text{for } \gamma \geq \gamma_0^{(c)} \end{cases} \quad (3.77)$$

where c is the index that identifies the modulation and coding scheme (MCS).

The two parameters $\alpha^{(c)}$ and $\gamma_0^{(c)}$ are calculated such that the relative quadratic error

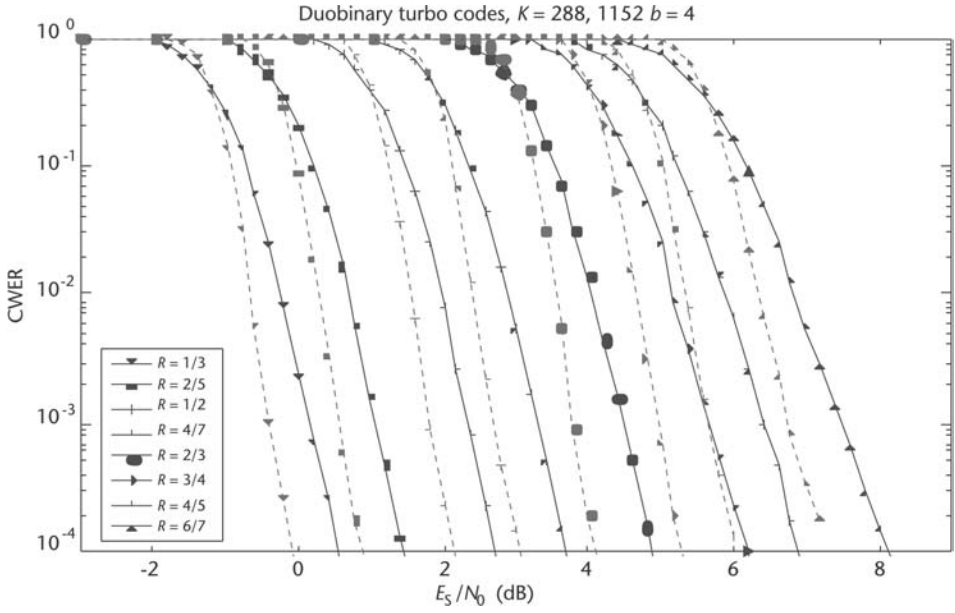


Figure 3.34 CWER curves for QPSK. Continuous lines: $K = 288$, dashed lines: $K = 1,152$ [88].

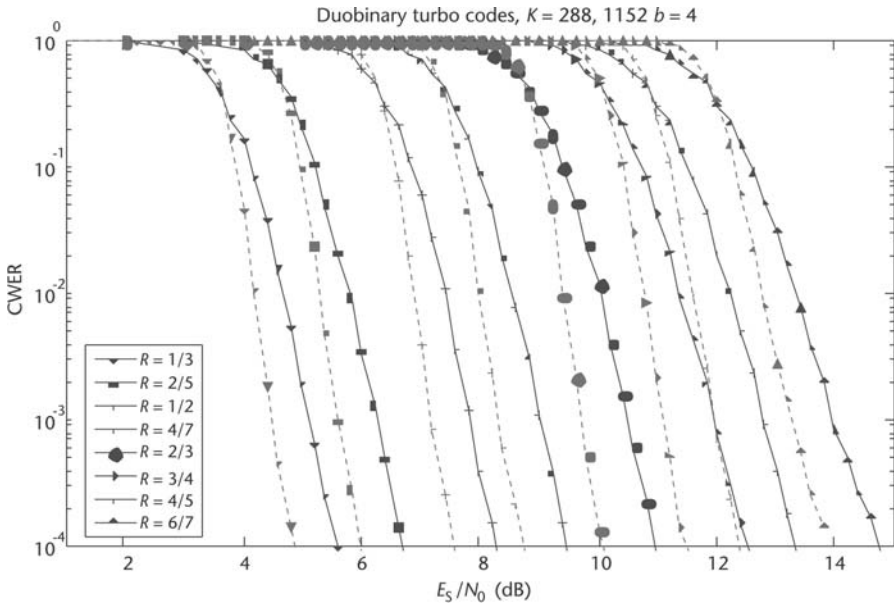


Figure 3.35 CWER curves for 16-QAM. Continuous lines: $K = 288$, dashed lines: $K = 1,152$ [88].

$$\sum_{i=1}^n \frac{(p_i - p_w^{(c)}(\gamma_i))^2}{p_i^2} \quad (3.78)$$

is minimized. Figure 3.37 shows the approximation for the simulation results in Figure 3.34. Only for very high CWER is the approximation rather coarse. The parameters for the approximation of the preceding simulation are available in [88].

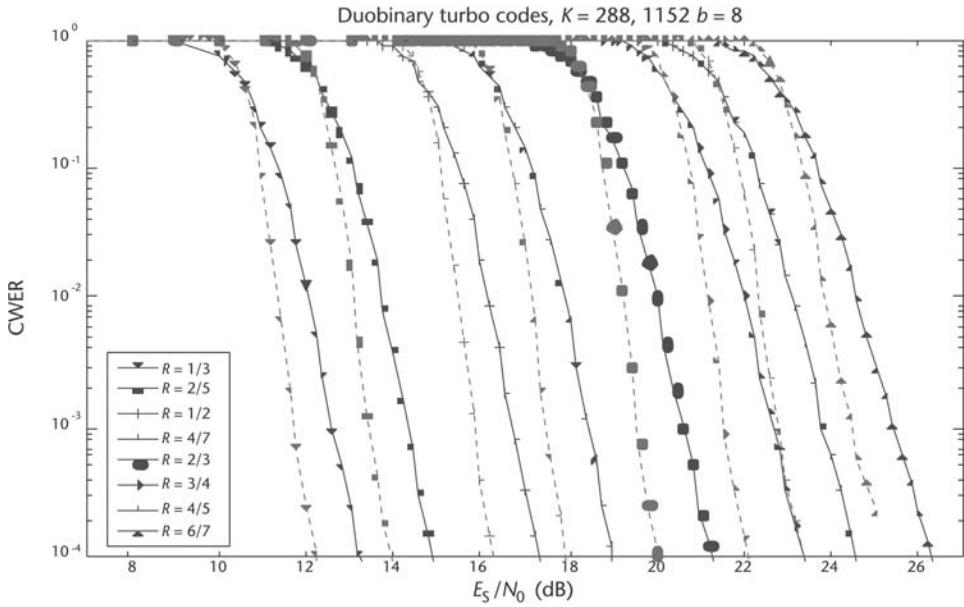


Figure 3.36 CWER curves for 256-QAM. Continuous lines: $K = 288$, dashed lines: $K = 1,152$ [88].

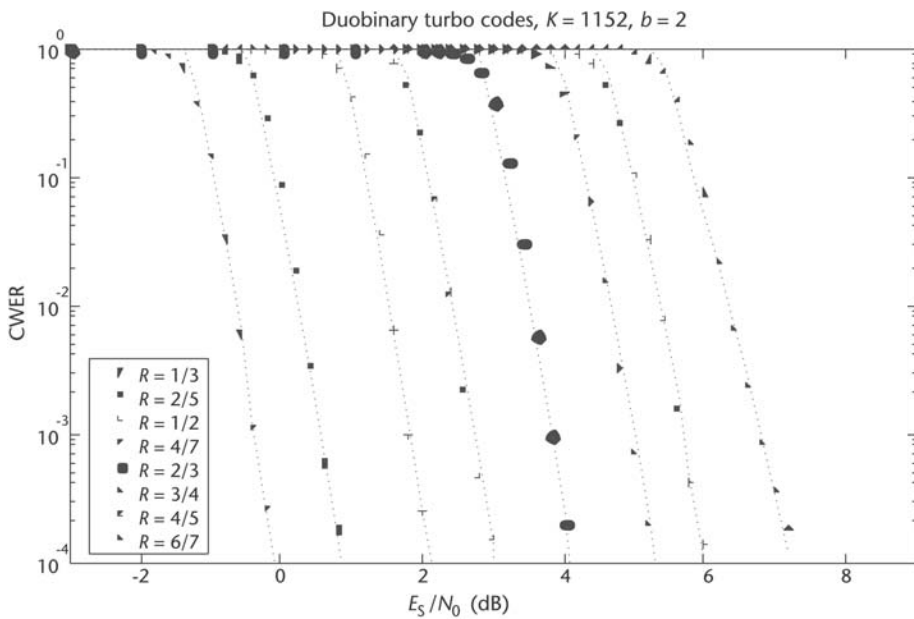


Figure 3.37 Approximated CWER curves (dotted lines) in comparison with simulation results (black markers) [88].

The CWER curves in Figures 3.34 through 3.36 were plotted as a function of the SNR $= E_s/N_0$, but they can also be plotted as a throughput over the SNR, where the throughput can be defined as $b \cdot R \cdot (1 - P_w)$. Figure 3.38 shows such a plot for 40 different MCSs. Such a plot allows for the exclusion of the schemes that require higher SNR for achieving the same or lower throughput than others.

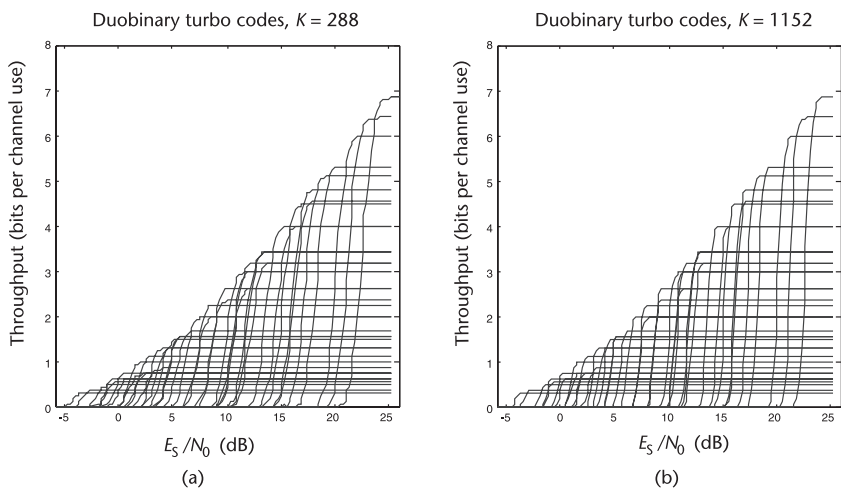


Figure 3.38 (a, b) Throughputs over SNR plots for 40 MCS [88].

The envelope curves also can be approximated by a modified capacity formula of the following type:

$$R = \alpha_1 \cdot 1d \left(1 + \alpha_2 \cdot \frac{E_S}{N_0} \right) \quad (3.79)$$

Then, from all possible MCS, a set of eight can be selected that provides the SNR thresholds as shown in Table 3.5. The selected MCSs are shown in Figure 3.39.

Table 3.5 SNR Thresholds for Eight Selected MCSs with a Target CWER of 0.01

MCS	1	2	3	4	5	6	7	8
b	1		2		4		6	8
R	1/3	1/2	3/4	1/2	3/4	2/3	6/7	4/5
$b \cdot R$	0.33	1.00	1.50	2.00	3.00	4.00	5.14	6.40
SNR threshold, γ_{\min} , $K = 288$	-3.2	2.0	5.2	7.6	11.4	15.2	19.2	23.2
MI threshold, $I_b(\gamma_{\min})$, γ_{\min} , $K = 288$	0.47	1.28	1.74	2.59	3.47	4.74	5.69	7.20
SNR threshold, γ_{\min} , $K = 1,152$	-3.6	1.6	4.8	7.0	10.8	14.6	18.4	22.4
MI threshold, $I_b(\gamma_{\min})$, γ_{\min} , $K = 1,152$	0.44	1.22	1.69	2.44	3.34	4.57	5.54	6.98

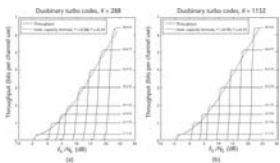


Figure 3.39 Throughput for eight selected MCSs [88].

3.3.2 Quasicyclic Block LDPC Codes

The QC-BLDPC codes are defined by sparse parity-check matrices of size $M \times N$ consisting of square submatrices (subblocks) of size $Z \times Z$ that are either zero or contain a cyclic-shifted identity matrix. M is the number of rows in the parity-check matrix, N is the code-length (number of columns) and the information size K is given by $K = N - M$.

The parity-check matrices are derived from the so-called base matrix \mathbf{H}_b of size $m \times n$ and the expansion factor Z , which determines the subblock size and, thus, the size of the derived code. That means, from one base matrix different code lengths can be constructed using different expansion factors:

$$N = Z \cdot n \quad (3.80)$$

One base matrix is specified per mother code rate: $R = K/N = 1 - m/n$.

The encoding and decoding processes are quite similar to the ones described in Section 3.1.1. Normally, the whole encoding process can be implemented through a high-throughput pipeline structure [88]. In the case of optimizing the codes for the type of air interface as developed within the IST project WINNER, some of the complexity of the procedure can be avoided.

Block LDPC codes are quasicyclic (i.e., a cyclicshift by a number smaller than the subblock size Z of a codeword yields again a codeword). From the symmetry of the codes, it follows that each bit within one subblock is equally important for the decoder and, hence, equally suitable for puncturing. It is therefore reasonable to define the puncturing pattern “subblock-wise.”

For $R = 1/2$ of the base matrix with parameters given in [88], a set of puncturing patterns was optimized for the code rates in region $\frac{24}{26} \leq R \leq \frac{24}{48}$. The puncturing patterns then were described by the following priority vector \mathbf{P} :

$$\mathbf{P} = [1, 2, 3, 4, 5, 6, 7, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24, 30, 34, 38, 41, 42, \dots, 13, 29, 46, 8, 32, 25, 37, 40, 44, 48, 27, 45, 33, 35, 36, 47, 31, 28, 26, 39, 9, 43, 22] \quad (3.81)$$

The priority vector \mathbf{P} gives the order in which subblocks of the codeword should be sent in a H-ARQ process. It can be used to define an interleaver in order to implement arbitrary punctured code rates [88].

Performance was evaluated by Monte Carlo simulations for the RCP BLDPC for a combination of the following parameters:

$$b = \{1, 2, 4, 6, 8\} \quad (3.82)$$

$$R = \frac{24}{48 - P}, \quad P = 0, 2, 4, \dots, 22 \quad (3.83)$$

where $b = \log_2(M)$ is the number of bits per constellation symbol, where M is the constellation size, R is the code rate, and P is the number of punctured subblocks from the codeword for mother code rate $R = 1/2$.

The Monte Carlo simulation used the following simulation chain: In the transmitter, each information packet of $K = 288$ or $K = 1,152$ random bits¹ is encoded with the BLDPC code encoder, then rate-compatible punctured, interleaved using a pseudorandom bit interleaver, and finally mapped into constellation symbols of b bits. Such a block of symbols is then transmitted through an AWGN channel. In the receiver, a soft demodulation is performed for each symbol of a block to obtain the LLR. The demodulator assumed a max-log-MAP approximation. Next, such an LLR block has been deinterleaved, depunctured, and sent to the BLDPC code decoder. The decoder employs a standard belief propagation algorithm in the LLR domain in parallel fashion (flooding schedule); that is, all variable \rightarrow check node messages updated in one sweep, then all check \rightarrow variable node messages updated in another sweep. The maximum number of decoding iterations was set to 50. The results are shown in Figure 3.40.

If a need for base code rates lower than $1/2$ arises [e.g., due to the need to serve low SINR users (cell edge) or to improve the coding gain obtainable with hybrid ARQ retransmission schemes when using multiple retransmissions], the QC-BLDPC would have to be redesigned to end up with at least a rate of $1/3$ for an LDPC code [88].

An evaluation of such a new rate for the modulation formats QPSK, 16-QAM, 64-QAM and 256-QAM, chosen in the context of the IST project WINNER [2] and for $K = 288$ information data bits (expansion factor = $Z_f = 18$), is shown in Figure 3.41.

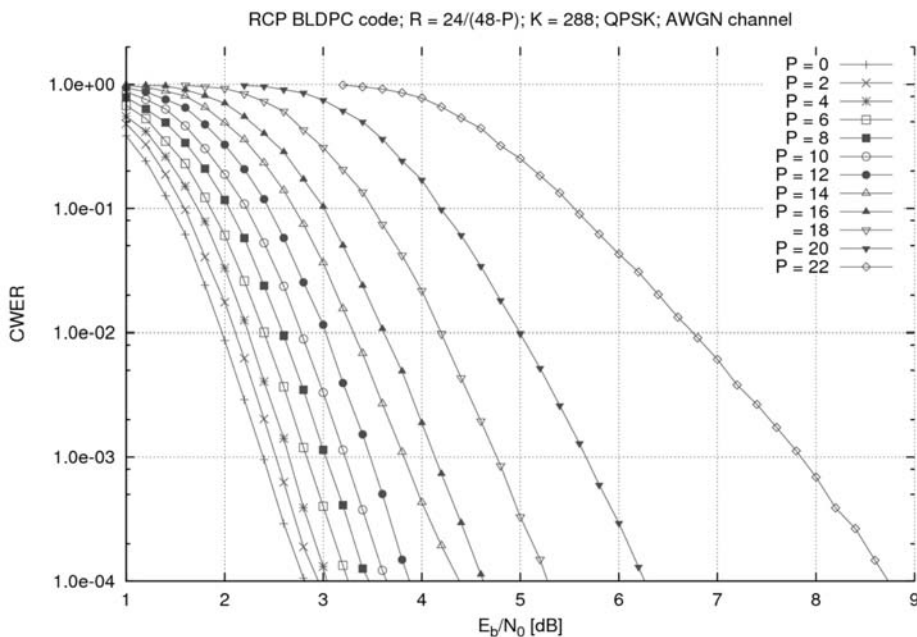


Figure 3.40 CWER curves for QPSK (and BPSK) and $K = 288$ [88].

1. These two particular sizes of information bits are taken from the baseline design assumptions.

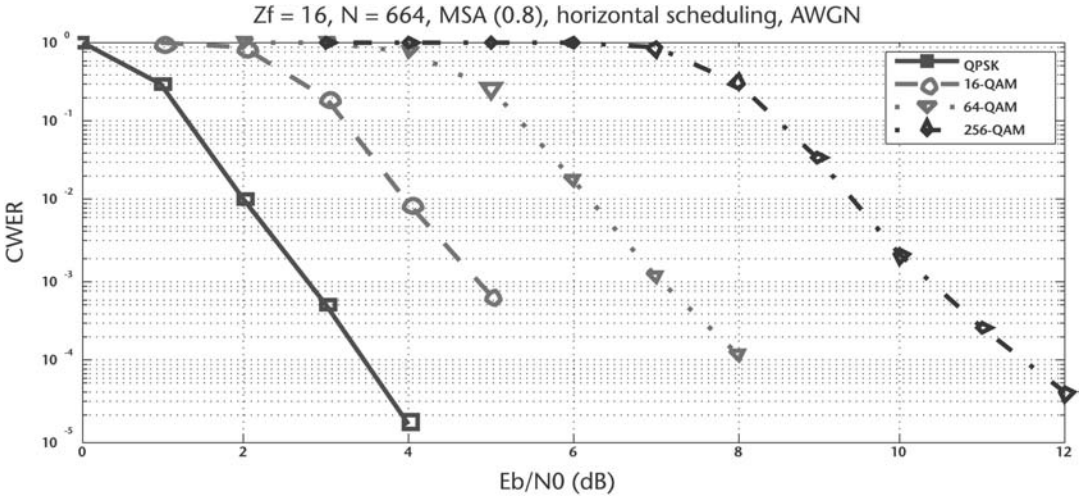


Figure 3.41 CWER performance for $K = 288$ bits ($Z_f = 18$) [88].

Note that the codeword length is a very important parameter related to the achievable throughput. In the following, the codeword length has been increased to 27,000 bits [88] for the purpose of complying with very high throughput requirements of the air interface. To ensure backward compatibility with already developed codes (e.g., $R_c = 1/2, 2/3, 3/4$), together with the lowest coding rate available (e.g., $R_c = 1/3$), the so-called *lifting* method can be applied on the original parity-check matrix. References [98] and [99] apply *lifting* to existing LDPC codes in order to increase the maximum allowable codeword length, while keeping the same performance for the previous range of codeword lengths (backward compatibility).

The constraints of the example described here are as follows:

- $N_b = 48 \rightarrow$ codeword length is a multiple of 48 (dimension of the base-model matrix):
 - Current maximum codeword length = $4,608 = 96 \times 48$;
 - Current maximum expansion factor = $Z_{f\max} = 96$;
 - \rightarrow New maximum codeword length = $27,648 = 576 \times 48$;
 - \rightarrow New maximum expansion factor = $Z_{f\max} = 576$.
- According to [99], the modulo lifting procedure refers to the resulting exponents $E(\mathbf{H}_k)$ of the parity check matrix \mathbf{H}_k corresponding to an expansion factor L_k as given by

$$E(\mathbf{H}_k) = E(\mathbf{H}) \bmod (L_k) \quad (3.84)$$

By applying the modulo lifting procedure step by step as described in [99], new parity-check matrices can be produced for the following coding rates: $R_c = 1/3, 1/2, 2/3$, and $3/4$, leading to the improved performances for the longer codewords shown in Figure 3.42.

Further, multiple impairments might result from the system itself in which such coding techniques are used [88]. As a result, even optimal decoding algorithms such as log-MAP for turbo codes, or LLR-BP for LDPC, even though those reach

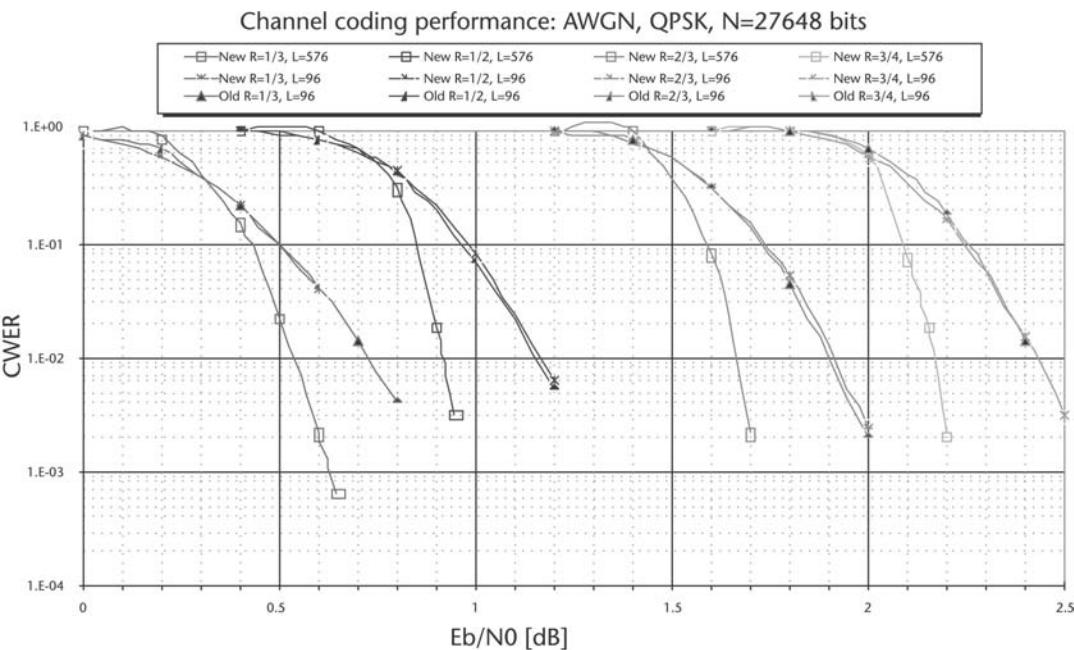


Figure 3.42 CWER performance results with lifted LDPC codes [88].

close to Shannon capacity performances, might experience severe degradations due to external impairments.

One of the key parameters common to decoders is the SNR estimation [100]. Therefore, the accuracy requested by the SNR estimation algorithms (and impacted by the channel estimation) must be evaluated in order to avoid prohibitive performance degradations.

The example here is restricted only to LDPC and considers different modulations, namely, QPSK, 16-QAM, and 64-QAM, with a half-rate $R_c = 1/2$ LDPC codes, as defined in [96]. Depending on the acceptable degradation in terms of performance (BER or CWER), the curves in Figures 3.43 to 3.45 can be used for checking the suitability of the channel estimation algorithms through their impact on the SNR estimation. For example, with QPSK for an operating point of $E_b/N_0 = 3$ dB, the SNR offset can be in the range of $[-3; +3]$ dB, if a BER above 10^{-5} should be avoided. An offset of -5 dB (i.e., underestimation) will force such a QPSK transmission to be degraded up to a BER of ~ 0.1 . On the contrary, even after a $+10$ -dB offset (i.e., overestimation), a BER of 10^{-2} can be maintained.

Figure 3.44 displays similar behavior for 16-QAM: There is a need for only a -3 -dB offset with a true $E_b/N_0 = 4$ dB to be above $\text{BER} = 10^{-2}$ but an overestimation of $+6$ dB is necessary.

In summary, the optimal performance of log-BP decoding of LDPC codes can be decreased because of SNR mismatch [88].

3.3.3 Low-Rate Convolutional Codes for Broadcast Information

The modulation and coding requirements for control channel signaling are different than the ones for user data transmission. The information sent through the control

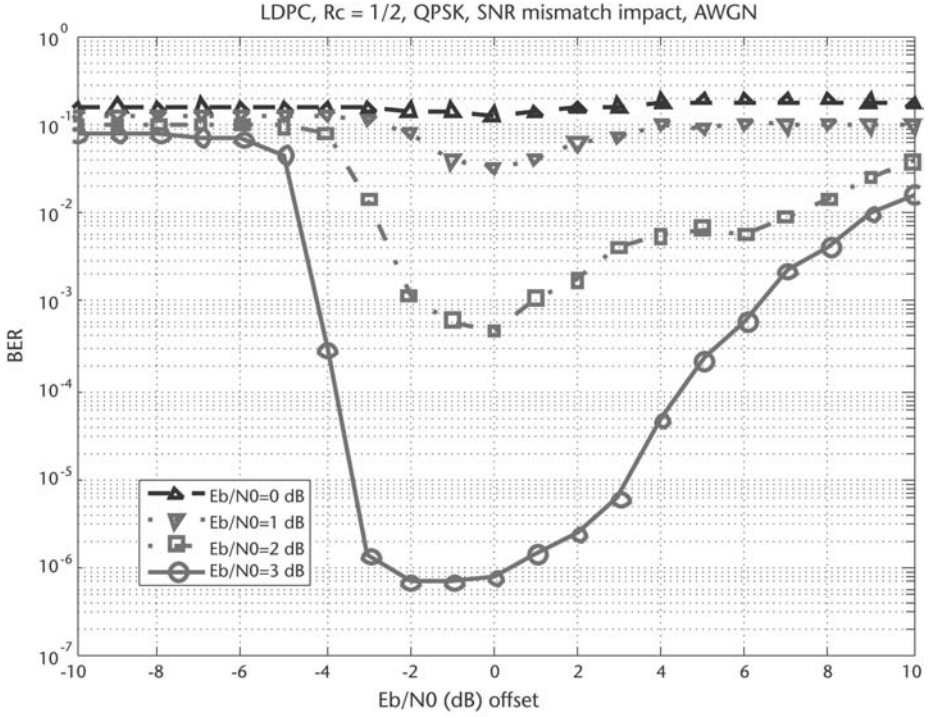


Figure 3.43 SNR mismatch impact on LDPC codes, $R = 1/2$, QPSK [88].

channel is very important for proper functioning of the advanced protocols of a next generation air interface concept. Although, the BLDPCC and DBTC codes proposed earlier for the WINNER system provide excellent coding performance as shown in [96], they cannot be used for encoding control information due to the very short packet sizes considered (e.g., 25 information bits). Therefore, low-rate convolutional codes, which can be used for encoding of such short packets by choosing a tail-biting algorithm, are an option.

For this purpose, instead of the maximum free distance (MFD) convolutional code defined in [96] with $R = 1/3$ and $G_A = [575, 623, 727]_{\text{oct}}$, an optimum distance spectrum (ODS) convolutional code [101] with $R = 1/4$ can be used. According to [101], an optimum distance spectrum convolutional code is a code generated by a feed-forward encoder with a superior distance spectrum compared to all other like encoders with the same rate R and constraint length L .

The superior distance spectrum is defined as follows [88].

A feedforward encoder with error weights c_d , giving a code free distance d_f has superior distance spectrum to an encoder with error weights \tilde{c}_d , giving a code free distance of \tilde{d}_f if one of the following conditions is fulfilled:

1. $d_f > \tilde{d}_f$
2. $d_f = \tilde{d}_f$ and there exists an integer $l \geq 0$ such that
 - a. $c_d = \tilde{c}_d$ for $d = d_f, d_f + 1, \dots, d_f + l - 1$
 - b. $c_d < \tilde{c}_d$ for $d = d_f + l$

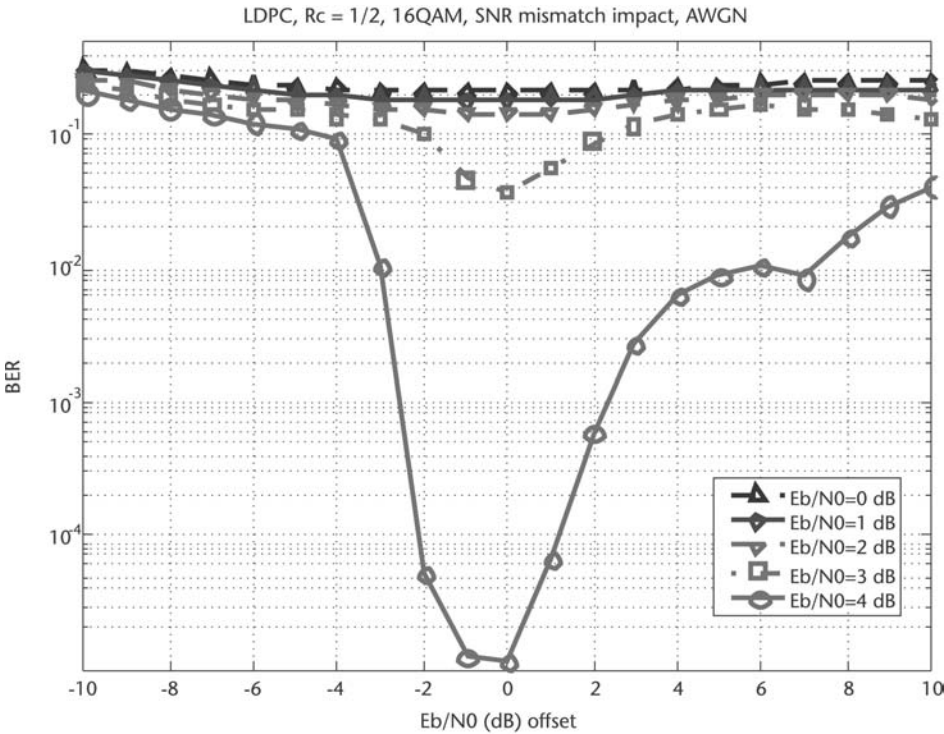


Figure 3.44 SNR mismatch impact on LDPC codes, $R = 1/2$, 16-QAM [88].

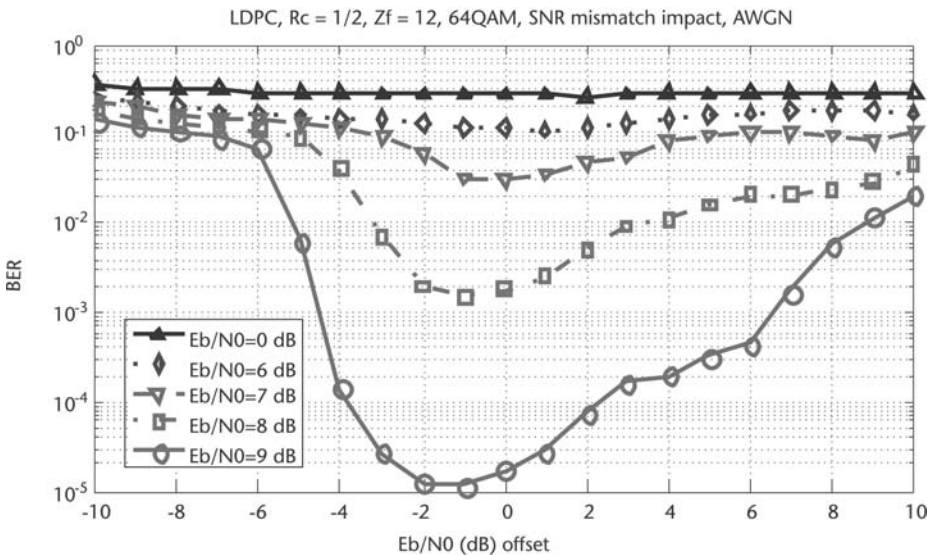


Figure 3.45 SNR mismatch impact on LDPC codes, $R = 1/2$, 64-QAM [88].

This means that for the same code rate and constraint length, an ODS code has the same free distance as an MFD code, but a lower or equal information error weight spectrum.

The BER and CWER performance results shown in Figures 3.46 and 3.47 were obtained for the convolutional code with the following generator polynomials: $G_B = [473, 513, 671, 765]_{oct}$. These results are compared with the results for the convolutional code from phase I.

Related to tail-biting Viterbi decoding, the “brute-force” tail-biting algorithm is $2^{k(L-1)}$ times more complex than a standard Viterbi decoding with a known tail, where k represents the number of inputs of the convolutional code ($R = k/n$) and L is the constraint length. For a convolutional code with $L = 9$, this means an additional complexity factor of 256 [88]. Therefore, other convolutional codes with shorter constraint lengths might be a good compromise between the decoding complexity and performance figures. Figure 3.48 compares CWER and BER results of a few $R = 1/4$ ODS convolutional codes with different constraint lengths ($L = \{6, 7, 8, 9\}$). The CWER performance of the shortest code in this group (i.e., with constraint length $L = 6$) is about 0.5 dB worse than the code with $L = 9$. On the other hand, the decoding complexity of this shortest code is $2^{9-6} = 8$ times lower than the longest one, so it might be a good candidate for a next generation air interface.

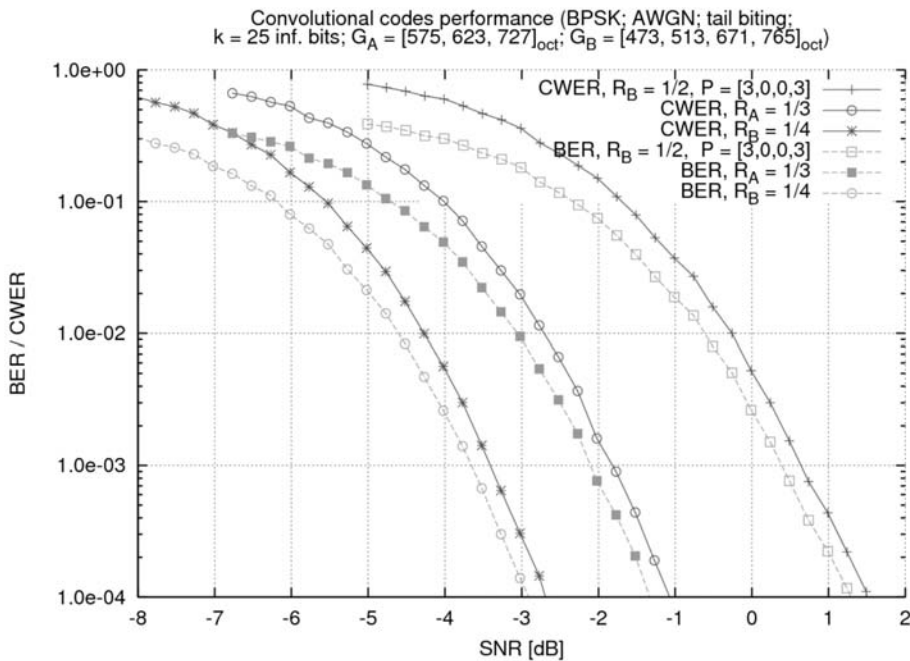


Figure 3.46 BER and CWER versus SNR results of $R = 1/4$ (ODS), $R = 1/3$ (MFD), and $R = 1/2$ (ODS, punctured) convolutional codes for $K = 25$ information bits (BPSK, AWGN, tail-biting) [88].

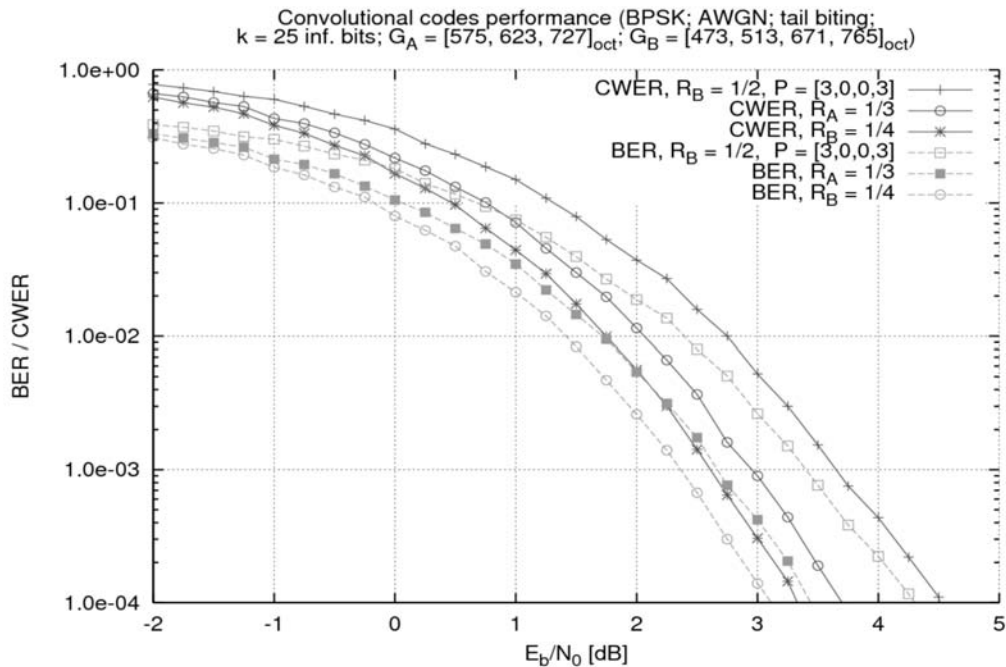


Figure 3.47 BER and CWER versus E_b/N_0 results of $R = 1/4$ (ODS), $R = 1/3$ (MFD), and $R = 1/2$ (ODS, punctured) convolutional codes for $K = 25$ information bits (BPSK, AWGN, tail-biting) [88].

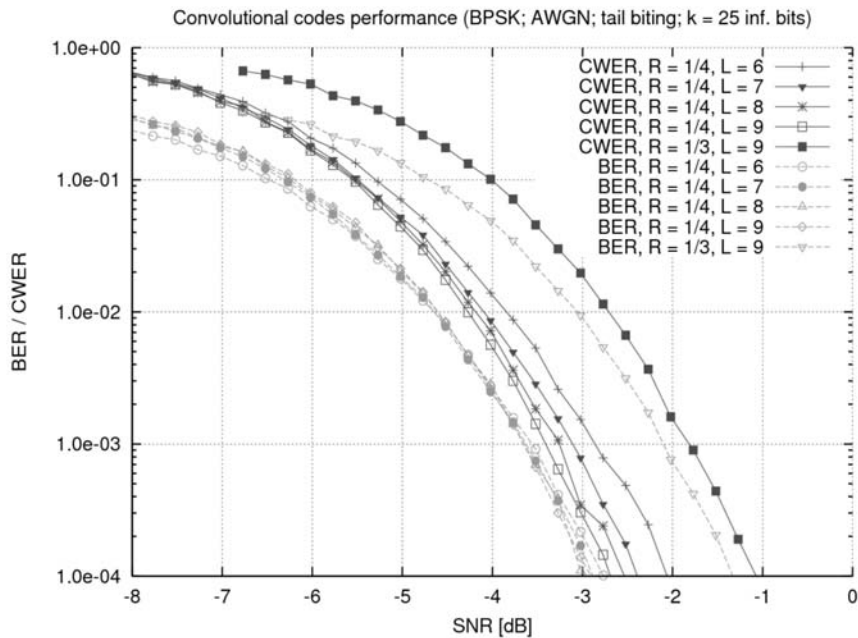


Figure 3.48 BER and CWER versus SNR results of $R = 1/4$ ODS convolutional codes for various constraint lengths L and $K = 25$ information bits (BPSK, AWGN, tail-biting) [88].

3.4 Conclusions

Modulation and coding are fundamental for any radio interface system. When choosing modulation and coding schemes for an air interface, many factors must be taken into consideration but, ultimately, the choice will be closely linked to the air interface baseline and reference design and system concept requirements.

Normally, the choice will be made by creating a domain of suitability of the candidate schemes, rather than opting for a clear-cut choice.

Based on the latest requirements in terms of information block length, interest in higher lengths has been increasing, including for extremely high lengths above 27 Kbits for the codeword. This interest in a higher data block length is being shared by different next generation air interface system candidates, such as 3GPP-LTE (6,144 bits), IEEE 802.20 (8,192 bits), UMB (7,680 bits), and IEEE 802.16m (around 8,000 bits), that are targeting IMT-A requirements. As a consequence, even though DBTCs are offering an optimal medium packet length, the LDPC solution becomes attractive because of the beneficial trade-off between performance and complexity, while going for large or extremely large data packets.

To ease the implementation of the coding schemes into system level simulations, the codeword error rate should be analytically approximated by rather simple calculations as described in this chapter. Recent developments reported by the FP6 IST projects, show that a B-LDPC code with mother code rate below 1/2, namely, with a rate of 1/3 is required for an IMT-A candidate air interface system. For example, the QC-BLDPC codes were chosen for the WINNER air interface, complemented by convolutional codes for small packets.

Practical, robust techniques for attaining signaling rates above 1 bit/s Hz⁻¹ of channel bandwidth can be based on studies of coset codes, multilevel codes, and bit-interleaved codes over Galois fields, which transmit many bits per symbol, and the development of optimum bit mappings for high-performance systems, for which conventional Gray mappings may prove suboptimal.

Modulation and coding are adapted to changing environments using link adaptation, also called adaptive coding and modulation (ACM). Such schemes are described in Chapter 4 of this book.

References

- [1] FP6 Overview, <http://cordis.europa.eu/ist/so/mobile-wireless/home.html>.
- [2] FP6 IST Project Wireless Interface New Radio, WINNER and WINNER II, www.ist-winner.org.
- [3] FP6 IST Project My Personal Adaptive Global Net (MAGNET) and MAGNET Beyond, www.ist-magnet.org.
- [4] FP6 IST Project SURFACE, www.ist-surface.org/deliverables.htm.
- [5] FP6 IST Project PULSERS, Document EOI-FP6-2002, www.pulsers.eu.
- [6] FP6 IST Project E2R, www.e2r.motlabs.com
- [7] FP6 IST Project NEWCOM, <http://newcom.ismb.it>.
- [8] FP6 IST Project NEWCOM, Deliverable D1.1, "Analysis and Design of Algorithms for Signal Processing at Large in Wireless Systems," March 2005, <http://newcom.ismb.it>.
- [9] FP6 IST Project MASCOT, www.ist-mascot.org/deliverables.

- [10] Foschini, G. J., and M. J. Gans, "On Limits of Wireless Communications in Fading Environments When Using Multiple Antennas," *International J. of Wireless Personal Communications*, Vol. 6, No. 3, March 1998, pp. 311–335.
- [11] Shannon, C. E., "A Mathematical Theory of Communications," *Bell Systems Technical J.*, Vol. 27, 1948.
- [12] FP6 IST Project WINNER, "Identification of Radio Link Technologies," Deliverable 2.1, July 2004, www.ist-winner.org.
- [13] Golay, M. J. E., "Notes on Digital Coding," *Proc. of IEEE*, Vol. 37, 1949, p. 657.
- [14] Hamming, R., "Error Detecting and Error Correcting Codes," *Bell Systems Technical J.*, Vol. 29, 1950, pp. 147–160.
- [15] Bose, R. C., and D. K. Ray-Chaudhuri, "On a Class of Error Correcting Binary Group Codes," *Information and Control*, Vol. 3, 1960, pp. 68–79.
- [16] Bose, R. C., and D. K. Ray-Chaudhuri, "Further Results on Error Correcting Binary Group Codes," *Information and Control*, Vol. 3, 1960, pp. 279–290.
- [17] Reed, I. S. "A Class of Multiple-Error-Correcting Codes and the Decoding Scheme," *IEEE Trans. on Information Theory*, Vol. 4, 1954, pp. 38–49.
- [18] Muller, D. E., "Application of Boolean Algebra to Switching Circuit Design and to Error Detection," *IEEE Trans. on Computing*, Vol. 3, 1954, pp. 6–12.
- [19] Assmus, E. F., Jr., and H. F. Mattson, Jr., "Error-Correcting Codes: An Axiomatic Approach," *Information and Control*, Vol. 6, 1963, pp. 315–330.
- [20] Bossert, M., *Channel Coding for Telecommunications*, New York: John Wiley & Sons, 1999.
- [21] Gallager, R. G., *Low-Density Parity-Check Codes*, Cambridge, MA: The MIT Press, 1963.
- [22] Reed, I. S., and G. Solomon, "Polynomial Codes over Certain Finite Fields," *J. of Social Industrial Applied Mathematics*, Vol. 8, 1960, pp. 300–304.
- [23] Elias, P., "Coding for Noisy Channels," *IRE Convention Record*, Part 4, 1955, pp. 37–47.
- [24] Viterbi, A. J., "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm," *IEEE Trans. on Information Theory*, Vol. IT-13, April 1967, pp. 260–269.
- [25] Forney, G. D., Jr., "Convolutional Codes II: Maximum Likelihood Decoding," *Information and Control*, Vol. 25, July 1974, pp. 222–266.
- [26] Bahl, L. R., et al., "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate," *IEEE Trans. on Information Theory*, March 1974.
- [27] Hagenauer, J., and P. Hoeher, "A Viterbi Algorithm with Soft-Decision Outputs and Its Applications," *Proc. of Globecom*, 1989.
- [28] Berrou, C., A. Glavieux, and P. Thitimajshima, "Near Shannon Limit Error Correcting Coding and Decoding," *Proc. of ICC*, 1993.
- [29] Benedetto, S., and G. Montorsi, "Role of Recursive Convolutional Codes in Turbo Codes," *IEE Electronic Letters*, 1995.
- [30] Brüningshaus, K., et al., "Link Performance Models for System Level Simulations of Broadband Radio Access Systems," *Proc. of PIMRC*, Berlin, Germany, September 2005.
- [31] Richardson, T. J., M. A. Shokrollahi, and R. Urbanke, "Design of Capacity-Approaching Irregular Low-Density Parity-Check Codes," *IEEE Trans. on Information Theory*, Vol. 47, No. 2, February 2001, pp. 617–637.
- [32] Chung, S. Y., et al., "On the Design of Low-Density Parity-Check Codes Within 0.0045 dB of the Shannon Limit," *IEEE Letters on Communication*, Vol. 5, February 2001, pp. 58–60.
- [33] Tanner, R. M., "A Recursive Approach to Low Complexity Codes," *IEEE Trans. on Information Theory*, Vol. IT-27, September 1981, pp. 533–547.

- [34] Hu, X. Y., E. Eleftheriou, and D. M. Arnold, "Regular and Irregular Progressive Edge-Growth Tanner Graphs," *IEEE Trans. on Information Theory*, 2003.
- [35] Richardson, T. J., and R. L. Urbanke, "Efficient Encoding of Low-Density Parity-Check Codes," *IEEE Trans. on Information Theory*, Vol. 47, No. 2, February 2001, pp. 638–656.
- [36] Perez, L., J. Seghers, and D. J. Costello, Jr., "A Distance Spectrum Interpretation of Turbo Codes," *IEEE Trans. on Information Theory*, November 1996.
- [37] Sason, I., and S. Shamai, "Improved Upper Bounds on the (ML) Decoding Error Probability of Parallel and Serial Concatenated Turbo Codes via Ensemble Distance Spectrum," *IEEE Trans. on Communications*, 2000.
- [38] Divsalar, D., and F. Pollara, "On the Design of Turbo Codes," TDA Progress Report 42-123, Pasadena, CA: Jet Propulsion Laboratory, 1995.
- [39] Hall, E. K., and S. G. Wilson, "Design and Analysis of Turbo Codes on Rayleigh Fading Channels," *IEEE J. on Selected Areas of Communications*, February 1998.
- [40] Benedetto, S., R. Garello, and G. Montorsi, "A Search for Good Convolutional Codes to Be Used in the Construction of Turbo Codes," *IEEE Trans. on Communications*, September 1998.
- [41] Yuan, J., B. Vucetic, and W. Feng, "Combined Turbo Codes and Interleaving Design," *IEEE Trans. on Communications*, April 1999.
- [42] Dolinar, S., D. Divsalar, and F. Pollara, "Code Performance as a Function of Code Block Size," TMO Progress Report 42-133, Pasadena, CA, May 1998.
- [43] Zhang, T., and K. K. Parhi, "VLSI-Implementation Oriented (3,K)-Regular LDPC Codes," *Proc. of IEEE Workshop on Signal Processing Systems (SiPS '01)*, September 2001, pp. 25–26.
- [44] Vucetic, B., and J. Yuan, *Turbo Codes: Principles and Applications*, Boston: Kluwer Academic Publishers, 2000.
- [45] Hokfelt, J., O. Edfors, and T. Maseng, "On the Theory and Performance of Trellis Termination Methods for Turbo Codes," *IEEE J. on Selected Areas in Communications*, May 2001.
- [46] Jung, P., and M. Nashan, "Performance Evaluation of Turbo Codes for Short Frame Transmission Systems," *IEE Electronic Letters*, January 1994.
- [47] Salah, M. M., et al., "Approach for Deriving Performance Bounds of Punctured Turbo Codes," *IEE Electronic Letters*, December 1999.
- [48] Hagenauer, J., E. Offer, and L. Papke, "Iterative Decoding of Binary Block and Convolutional Codes," *IEEE Trans. on Information Theory*, Vol. 42, No. 2, March 1996, pp. 429–445.
- [49] MacKay, D. J. C., S. T. Wilson, and M. C. Davey, "Comparison of Constructions of Irregular Codes," *Proc. of the 36th Allerton Conference on Communication, Control, and Computing*, September 1998.
- [50] Richardson, T. J., M. A. Shokrollahi, and R. Urbanke, "Design of Capacity-Approaching Irregular Low-Density Parity-Check Codes," *IEEE Trans. on Information Theory*, Vol. 47, No. 2, February 2001, pp. 617–637.
- [51] Reid, A., K. J. Kim, and V. Stolzmann, "Design of Irregular LDPC Codes for Time-Varying OFDM MIMO Channels," Nokia Internal Report, 2004.
- [52] Choi, W., R. Negi, and J. M. Cioffi, "Combined ML and DFE Decoding for the V-BLAST System," *Proc. of ICC*, June 2000, Vol. 3, pp. 1243–1248.
- [53] Muirhead, R. J., *Aspects of Multivariate Statistical Theory*, New York: John Wiley & Sons, 1982.
- [54] Prasad, R., and U. S. Jha, *OFDM Towards Fixed and Mobile Broadband Wireless Access*, Norwood, MA: Artech House, 2007.
- [55] van Nee, R., and R. Prasad, *OFDM for Multimedia Communications*, Norwood, MA: Artech House, 2000.

- [56] Steendam, H., and M. Moeneclaey, "Analysis and Optimization of the Performance of OFDM on Frequency-Selective Time-Selective Fading Channels," *IEEE Trans. on Communications*, Vol. 47, December 1999, pp. 1811–1819.
- [57] Hara, S., and R. Prasad, "Design and Performance of Multicarrier CDMA System in Frequency-Selective Rayleigh Fading Channels," *IEEE Trans. on Vehicular Technology*, Vol. 48, September 1999, pp. 1584–1595.
- [58] "Framework and Overall Objectives of the Future Development of IMT 2000 and Systems Beyond IMT 2000," Recommendation ITU-R M.1645, www.itu.int.
- [59] Tarokh, V., H. Jafarkhani, and A. R. Calderbank, "Space-Time Block Codes from Orthogonal Designs," *IEEE Trans. on Information Theory*, Vol. 45, July 1999, pp. 1456–1467.
- [60] Zhou, S., and G. B. Giannakis, "Space-Time Coding with Maximum Diversity Gains over Frequency-Selective Fading Channels," *IEEE Signal Processing Letters*, Vol. 8, October 2001, pp. 269–272.
- [61] Muck, M., M. de Courville, and P. Duhamel, "A Pseudo Random Postfix OFDM Modulator. Part I: Semi-Blind Channel Estimation and Equalization in the Static Case," *IEEE Trans. on Signal Processing*, 2004.
- [62] Vaidjanathan, P. P., *Multirate Systems and Filter Banks*, Englewood Cliffs, NJ: Prentice Hall, 1993.
- [63] Bolcskei, H., "Efficient Design of Pulse Shaping Filters for OFDM Systems," *Proc. of SPIE*, Denver, Colorado, July 1999, Vol. 3813, pp. 625–636.
- [64] Cherubini, G., E. Elefthriou, and S. Oelcer, "Filtered Multi-Tone Modulation for Very High Speed Digital Subscriber Lines," *IEEE J. on Selected Areas of Communications*, Vol. 20, No. 5, June 2002, pp. 1016–1028.
- [65] Benvenuto, N., S. Tomasin, and L. Tomba, "Equalization Methods in OFDM and FMT Systems for Broadband Wireless Communications," *IEEE Trans. on Communications*, Vol. 50, September 2002, pp. 1413–1418.
- [66] Berenguer, I., and I. J. Wassel, "FMT Modulation: Receiver Filter Bank Definition for the Derivation of an Efficient Implementation," *Proc. of the 7th International OFDM Workshop (InOWo)*, Hamburg, Germany, September 2002.
- [67] Costa, E., et al., "Comparison of FMT-FDMA and OFDMA in Asynchronous Uplink Applications," *Proc. of the 8th International OFDM-Workshop (InOWo)*, Hamburg, Germany, September 2003.
- [68] Costa, E., et al., "Asynchronous Uplink Transmission in FMT-FDMA Systems," *Proc. of the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Beijing, China, 2003.
- [69] Cherubini, G., et al., "Filter Bank Modulation Techniques for Very High-Speed Digital Subscriber Lines," *IEEE Communications Magazine*, May 2000.
- [70] Tomba, L., "On the Use of DMT or FMT for MC-CDMA Systems," *Proc. of the Multi-Carrier Spread Spectrum Workshop*, Oberpfaffenhofen, Germany, September 2001.
- [71] Proakis, J. G., *Digital Communications*, 3rd ed., New York: McGraw-Hill, 1995.
- [72] Klein, A., "Multi-User Detection of CDMA Signals- Algorithms and Their Application to Cellular Mobile Radio," *Fortschrittberichte VDI*, 1996, Series 10, No. 423, Dusseldorf, Germany: VDI-Verlag.
- [73] Shynk, J., "Frequency Domain and Multirate Adaptive Filtering," *IEEE Signal Processing Magazine*, January 1992, pp. 14–37.
- [74] Sari, H., G. Karam, and I. Jeanclaude, "Frequency-Domain Equalization of Mobile Radio and Terrestrial Broadcast Channels," *Proc. of Globecom*, San Francisco, December 1994, pp. 1–5.
- [75] Montezuma, P., and A. Gusmão, "A Pragmatic Coded Modulation Choice for Future Broadband Wireless Communications," *Proc. of VTC 2001 Spring*, Rhodes, Greece, May 2001, Vol. 2, pp. 1324–1328.

- [76] Falconer, D., et al., "Frequency Domain Equalization for Single-Carrier Broadband Wireless Systems," *IEEE Communications Magazine*, Vol. 40, No. 4, April 2002, pp. 58–66.
- [77] Dinis, R., A. Gusmão, and N. Esteves, "On Broadband Block Transmission over Strongly Frequency-Selective Fading Channels," *Proc. of Wireless 2003*, Calgary, Canada, July 2003.
- [78] Abe, T., S. Tomisato, and T. Matsumoto, "Performance Evaluations of a Space-Time Turbo Equalizer in Frequency Selective MIMO Channels Using Field Measurement Data," *Proc. of the IEEE Workshop on MIMO Communication Systems*, London, December 2001, pp. 21/1–21/5.
- [79] Yee, M. S., M. Sandell, and Y. Sun, "Comparison Study of Single-Carrier and Multi-Carrier Modulation Using Iterative Based Receiver for MIMO System," *Proc. of IEEE VTC-2004 Spring*, Milan, May 2004.
- [80] Tarokh, V., N. Seshadri, and A. R. Calderbank, "Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction," *IEEE Trans. on Information Theory*, Vol. 44, No. 2, March 1998, pp. 744–765.
- [81] Bölcskei, H., and A. J. Paulraj, "Space-Frequency Coded Broadband OFDM Systems," *Proc. of the IEEE Wireless Communications Networking Conference (WCNC)*, Chicago, September 2000, Vol. 1, pp. 1–6.
- [82] Lu, B., and X. Wang, "Space-Time Code Design in OFDM Systems," *Proc. of Globecom*, San Francisco, CA, November 2000, Vol. 2, pp. 1000–1004.
- [83] Ng, B. K., and E. S. Sousa, "On Bandwidth-Efficient Multiuser-Space-Time Signal Design and Detection," *IEEE J. on Selected Areas of Communications*, Vol. 20, No. 2, February 2002, pp. 320–329.
- [84] Naguib, A. F., N. Seshadri, and A. R. Calderbank, "Applications of Space-Time Block Codes and Interference Suppression for High Capacity and High Data Rate Wireless Systems," *Proc. of Asilomar Conference of Signals, Systems, and Computing*, Pacific Grove, CA, November 1998, Vol. 2, pp. 1803–1810.
- [85] Gallager, R. G., "A Perspective on Multiaccess Channels," *IEEE Trans. on Information Theory*, Vol. 31, No. 2, March 1985, pp. 124–142.
- [86] FP6 IST Project MASCOT, "Multiuser Space-Time Code Design and Database of Specific Code Constructions," Deliverable 1.2.1, June 2007, www.ist-mascot.org/deliverables.
- [87] Visuri, S., and H. Bölcskei, "Multiple-Access Strategies for Frequency-Selective MIMO Channels," *IEEE Trans. on Information Theory*, Vol. 52, No. 9, September 2006, pp. 3980–3993.
- [88] FP6 IST Project WINNER II, "Modulation and Coding Schemes for the WINNER II System Concept," Deliverable D2.2.3, November 2007, www.ist-winner.org.
- [89] FP6 IST Project WINNER, "Performance Assessment of the WINNER System Concept," Deliverable D7.8, December 2005, www.ist-winner.org.
- [90] FP6 IST Project WINNER II, "D6.13.7 WINNER Test Scenarios and Calibration Cases, Issue 2," Deliverable 6.13.7, December 2006, www.ist-winner.org.
- [91] FP6 IST Project WINNER II, "D6.13.14 WINNER Final System Concept," Deliverable 6.13.14, December 2007, www.ist-winner.org.
- [92] "Digital Video Broadcasting (DVB); Interaction Channel for Digital Terrestrial Television (RCT) Incorporating Multiple Access OFDM," ETSI EN 301 958 V1.1.1, March 2002. {AU: publisher or website?}
- [93] IEEE Std 802.16-2004, "Part 16: Air Interface for Fixed Broadband Wireless Access Systems," IEEE Standards Institution, New York, 2004.
- [94] Berrou, C., et al., "Duo-Binary Turbo Codes Associated with High-Order Modulation," *Proc. of ESA TTC 2001*, Noordwijk, the Netherlands, October 2001.
- [95] FP6 IST Project WINNER II, "Joint Modulation and Coding Procedures for the WINNER II System Concept," Deliverable D2.2.1, October 2006, www.ist-winner.org.

- [96] FP6 IST Project WINNER, “Final Report on Identified RI Key Technologies, System Concepts, and Their Assessment,” Deliverable D2.10 December 2005, www.ist-winner.org.
- [97] Robertson, P., E. Villebrun, and P. Höher, “A Comparison of Optimal and Sub-Optimal MAP Decoding Algorithms Operating in the Log Domain,” *Proc. of IEEE ICC '95*, Seattle, June 1995.
- [98] Myung, S., and K. Yang, “Extension of Quasi-Cyclic LDPC Codes by Lifting,” *Proc. of IEEE International Symposium on Information Theory*, 2005, pp. 2305–2309.
- [99] Myung, S., K. Yang, and J. Kim, “Quasi-Cyclic LDPC Codes for Fast Encoding,” *IEEE Trans. on Information Theory*, Vol. 51, August 2005, pp. 2894–2901.
- [100] Summers, T. A., and S. G. Wilson, “SNR Mismatch and Online Estimation in Turbo Decoding,” *IEEE Trans. on Communications*, Vol. 46, No. 4, April 1998.
- [101] Frenger, P., et al., “Multi-Rate Convolutional Codes,” Technical Report No. 21, Göteborg: Chalmers University, Department of Signals and Systems, Communications Systems Group, April 1998.

Multiple-Access Schemes

To be successfully designed and adopted, next generation radio interfaces must employ technologies that allow for exploiting frequency-selective channel properties (e.g., multicarrier transmissions), for attaining ultrahigh spectral efficiencies (e.g., MIMO technologies), for minimizing the required SNR at the reception side (e.g., turbo principle), for being adaptive and flexible (e.g., modulation, spatial processing, cooperative relaying), and so forth. The multiple-access scheme is part of the toolbox of radio system elements that can be used in the most appropriate combination to efficiently support any given situation.

The choice of a multiple-access scheme is one of the crucial decisions made in the design of a radio communication system. It enables the spectrum resources available to the system to be shared among many users. Thus, the signals generated by them can be effectively separated at the receivers. In a mobile communication system the radio interface determines only a small part of the overall system cost; however, its design largely influences the total system design including the fixed part of the network and to a large extent it determines the cost and quality of the system operation [1].

The performance and quality achieved by the choice of a multiple-access scheme has to be considered in a particular radio environment. For example, the results of the comparison of several multiple-access schemes can be different when considered for a single- or multiple-cell environment. Propagation conditions also play an important role. In addition, optimized access schemes for PAN communications have also gained a momentum. Peer-to-peer communication, especially from data traffic point of view, will happen differently in a PAN communication environment than in a cellular or wireless LAN system.

A great number of the achievements that have been fed into the standardization groups and partnership projects have resulted from the R&D efforts of the projects under the FP6 IST umbrella [2]. Chapter 4 builds on the contributions of the following major IST projects: WINNER and WINNER II [3], MAGNET and MAGNET Beyond [4], MASCOT [5], SURFACE [6], 4MORE [7], FIREWORKS [8], and others, all of which are aimed at achieving less complexity and better performance through the proper choice of multiple-access scheme or the enhancement of an existing multiple-access scheme.

4.1 Introduction

The choice of an optimal multiple-access scheme or a combination of schemes is very dependent on the performance requirements set up for the radio interface

system. Channel state adaptations, user requirements, and changing traffic patterns are only some of the considerations surrounding the final choice. Other important considerations are that the radio interface system under design be able to interwork successfully with already existing or yet to be developed systems. According to the recommendations provided by the ITU, the capabilities of next generation radio interfaces must agree with the following requirements for peak aggregate user data rates [9]:

- Up to approximately 100 Mbps for mobile access;
- Up to approximately 1 Gbps for nomadic/LAN access.

Whether the choice is made to enhance an existing wireless access scheme or to develop a new one depends on the abilities of the system to comply with the identified work requirements. For example, the enhancements of existing systems, such as HSDPA for WCDMA in UMTS, or WLAN, are limited by system-inherent restrictions, such as the allocated carrier bandwidth. This, in turn, negatively affects the requirement for high throughput or ubiquitous coverage, respectively. Next generation systems will be driven by and must comply with individual user requirements. This imposes the need for adaptive and user-centric solutions. On the other hand, one serious requirement is to reduce the complexity of the proposed solutions.

Part of the research efforts have focused on how existing access techniques can be foreseen with matching efficiency, capacity, and scalability. New access schemes [10, 11], such as OFDMA, SC-FDMA, interleaved FDMA, and multicarrier code-division, multiple-access (MC-CDMA) are gaining more importance. WiMax [12, 13] is using OFDMA in the downlink and in the uplink, and OFDMA is being considered for the downlink evolving from UMTS technologies, such as LTE and LTE-Advanced [13–15]. IFDMA is being considered as suitable for the uplink because it provides fewer power fluctuations and, thus, avoids amplifier issues. MC-CDMA has been proposed for the IEEE 802.20 standard. These access schemes offer the same efficiencies as the classical technologies, such as FDMA, TDMA, and CDMA, but achieve scalability and higher data rates. Classical multiple-access schemes have been analyzed and discussed in detail in numerous books and publications [16–22].

It has been shown [1, 3, 14] that classical schemes by themselves will not be sufficient to meet the technical requirements of next generation systems (e.g., IMT-A). Among some of the additional concepts studied in detail by researchers are the use of OFDM as a means for multiple access, providing different users with different parts of the total resource through the use of the OFDMA and MC-CDMA techniques, as well as SDMA. Compared to classical FDMA, OFDMA and MC-CDMA have the potential to provide a significant increase in flexibility and spectral efficiency [23, 24]. The development of suitable combinations of SDMA and other access techniques has proven to have high potential for significant synergies and performance increases.

Multiple-access schemes are accompanied primarily by duplexing techniques, which allow the users to send and receive signals simultaneously, or at least quasi-simultaneously. They are categorized into two basic techniques: frequency-division duplex and time-division duplex. In the first method both transmission directions

are separated on the frequency axis by assigning different bands to them, whereas in TDD the direction of transmission is periodically reverted using the same frequency band. In wired systems echo cancellation turns out to be the most effective duplexing method [23]. Multiple-access schemes can be applied jointly with several modulations used to carry user information. Both single- and multiple-carrier modulations can be applied.

4.1.1 Selection Criteria for a Multiple-Access Scheme

The main selection criteria for the evaluation of access methods for a next generation radio interface can be summarized as follows:

- *High performance*: The system must offer higher performance than current systems.
- *Adaptability and flexibility*: The system should ideally reach the performance targets in all specified scenarios and environments. This means it should have a high degree of adaptation to varying channel, interference, and traffic conditions. Also the suitability of the access methods for different network deployment concepts, supported services, and resource allocation types needs to be taken into account.
- *Efficiency of resources used*: To avoid problems with spectrum allocation, the total bandwidth required for air interface candidates for next generation systems should be minimized. This will lead to requirements for high spectral efficiency for all types of services and data rates. Efficient adaptation methods and advanced radio resource management will be needed.
- *Complexity and cost*: The complexity of the whole system, ease of implementation, and cost need to be kept in mind. A sensible balance between complexity and adaptability has to be maintained in order to make the whole concept feasible.

The problem of multiple access is a problem of resource allocation under uncertainty and under economic constraints [11]. There are two fundamentally different ways of handling the problem of uncertainty, in particular, uncertain channel states and interference environments, which are the following:

1. *Averaging*: In the case of time-varying channels, averaging of time-varying propagation properties to combat fading is accomplished by coding and interleaving. Multiple antennas may also be utilized to reduce the channel variations by means of different techniques of space-time coding at transmitters and diversity combining at receivers. Spreading and frequency hopping are classical techniques for making the interference properties more predictable by averaging their properties over a large bandwidth.
2. *Adaptation*: If the channel/interference states are at least partially known at the transmitter, link adaptation can be used to adjust the data rate to/from each terminal to the instantaneous quality of the link. Adaptive beamforming may also be used. Channels and interference levels for different users, in general, vary independently; therefore, the resources can be

allocated to users who at the moment can utilize them best. It is more that the variations are utilized, rather than averaged away, which potentially provides a large increase in the throughput and in the spectral efficiency [23]. Furthermore, if transmission can be coordinated over a wider area, then the most significant interferers can be avoided by means of advanced radio resource management. Interference avoidance by coordinated transmission has the potential to significantly increase the spectral efficiency of the system.

Averaging provides robustness, while adaptive transmission and reception provide a potentially large increase in performance.

Another general aspect is the question of *orthogonality* versus *contention*. Multiple-access schemes may be based on orthogonal (nonoverlapping) resource allocation, such as TDMA, FDMA, or OFDMA, together with intercell frequency resource partitioning or perhaps coordinated scheduling to avoid intercell interference (ICI). Orthogonal schemes offer the potential to obtain a high spectral efficiency within each link. Methods that are not designed based on orthogonality, and that allow for high levels of interference or packet collisions, lead to a lower spectral efficiency but allow for a more flexible deployment.

The feasibility of adaptive transmission is of central importance for the choice of a multiple-access scheme, and for the efficiency and flexibility of the resulting scheme. This question can be regarded as central in all schemes, whether they are based on multicarrier or single-carrier transmission, or on TDMA, OFDMA, MC-CDMA, or SDMA.

The resource allocation that a multiple-access scheme provides is closely related to *capacity* [23]. Depending on the point of view of comparison, capacity can have different interpretations. On one side, capacity can be considered from an information-theoretic approach, which allows for evaluation of performance limits rather than estimation of a real performance. This tool allows for drawing valuable conclusions on multiple-access strategies. The capacity is measured in bit/s Hz^{-1} and shows the maximum amount of information per spectrum unit that can be sent by all active users. In cellular network design, capacity refers to the maximum number of users simultaneously served by the system located on the unit area, or alternatively the maximum value of traffic served by the system per unit area. Multiple-access capacity has also been defined as the number of users who can be supported at a given error performance level [25].

4.1.2 Single-Carrier Versus Multicarrier Schemes

The choice of a single-carrier or multicarrier scheme is an important design consideration. Single-carrier schemes have the advantage of a low signal PAPR, which leads to higher power efficiency at transmission, whereas multicarrier schemes offer the potential of adaptation of transmission parameters in the frequency domain, which leads to improved resource utilization [23]. A comparison between these two schemes was performed within the scope of the IST project WINNER when searching for a multiple-access scheme to comply with the requirements identified for an IMT-A air interface candidate [9, 23]. Regarding possible impacts on the multiple-

access and medium-access schemes, no fundamental difference was identified. In the downlink, the low PAPR of single-carrier signals was not considered a major advantage and, accordingly, no significant argument advocating single-carrier access in the downlink could be identified. Hence, in downlinks, the focus of the work and the evaluations on multiple-access methods of projects, such as the IST project WINNER, were based on multicarrier transmission schemes. In the uplink, however, a transmitted signal with low PAPR was considered favorable, especially for wide-area coverage. On the other hand, for short-range communications the PAPR is less important and it seemed attractive [3] to employ multicarrier transmission in the uplink also. Both single-carrier and multicarrier access schemes can be seen as viable uplink alternatives [23].

4.1.2.1 Single-Carrier Solutions

Single-carrier solutions for the uplink of a wideband cellular system, as proposed by the IST project WINNER, included spread and nonspread solutions, namely, DS-CDMA and SC-TDMA [23]. It was shown that the main challenge for single-carrier systems is how to deal with a channel response spanning many symbols, resulting from the combination of wide system bandwidth and channel delay spread.

For DS-CDMA, it was observed that such a long channel response has two main effects: First, the implementation of the Rake receiver becomes more complex because there are more paths to be combined; second, ISI increases and spans several symbols, which increases the complexity of multiuser detection or interference cancellation methods, such as the MMSE receiver. Taking into account the receiver complexity, two system configurations were considered: *continuous* and *block* transmission. Their corresponding Rake receiver implementations, time- and frequency-domain Rake, were shown. In addition, the implementation of MMSE receivers for both transmission methods was also derived. It was shown that complexity may be a limiting issue; therefore, a method to reduce complexity by introducing an FDMA component was proposed. Furthermore, a similar method was also proposed to deal with dual-rate users. Finally, the feasibility of time- and frequency-domain Rake receivers and MMSE receivers was shown through simulation.

Figures 4.1 and 4.2 show the raw BER performance for continuous transmission and block transmission, respectively, with both matched filter and MMSE receivers, under different system loads. The set of parameters used for these simulations is described in Table 4.1.

From the simulation results in Figures 4.1 and 4.2, we can conclude that time- and frequency-domain processing showed similar performance, and that MMSE receivers significantly outperform matched filter receivers. The error rate of the MMSE receiver for continuous transmission at 50% load is lower than that of the matched filter at 25% load (and much lower in the case of block transmission), which is basically more than double the system capacity. However, this improved performance comes at the expense of much increased complexity, which requires the consideration of complexity reduction schemes.

In TDMA, there is no (intracell) multiple-access interference as long as the nodes in the network are well synchronized. The challenge in the TDMA system

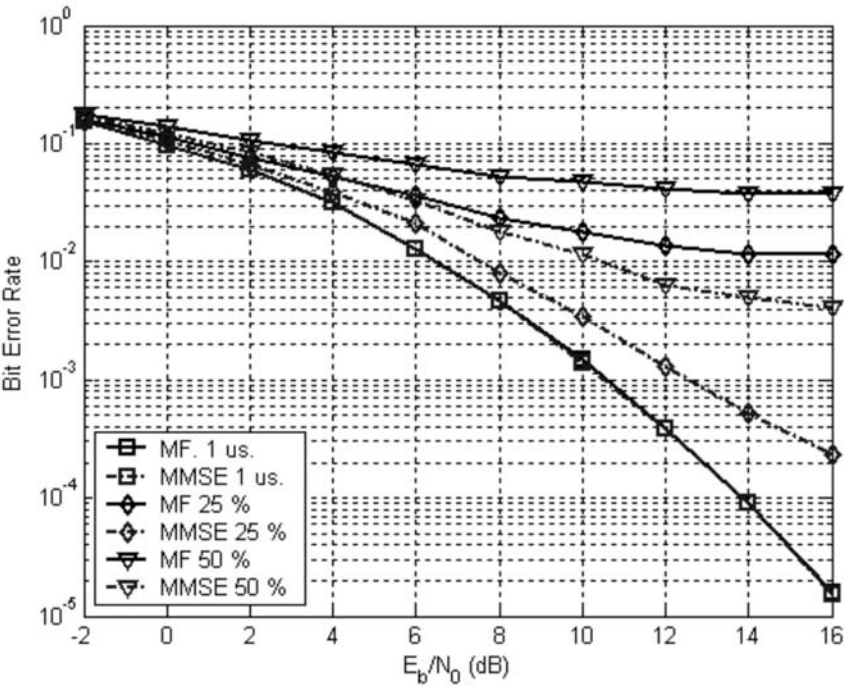


Figure 4.1 Raw BER for continuous transmission for matched filter and MMSE receivers [23].

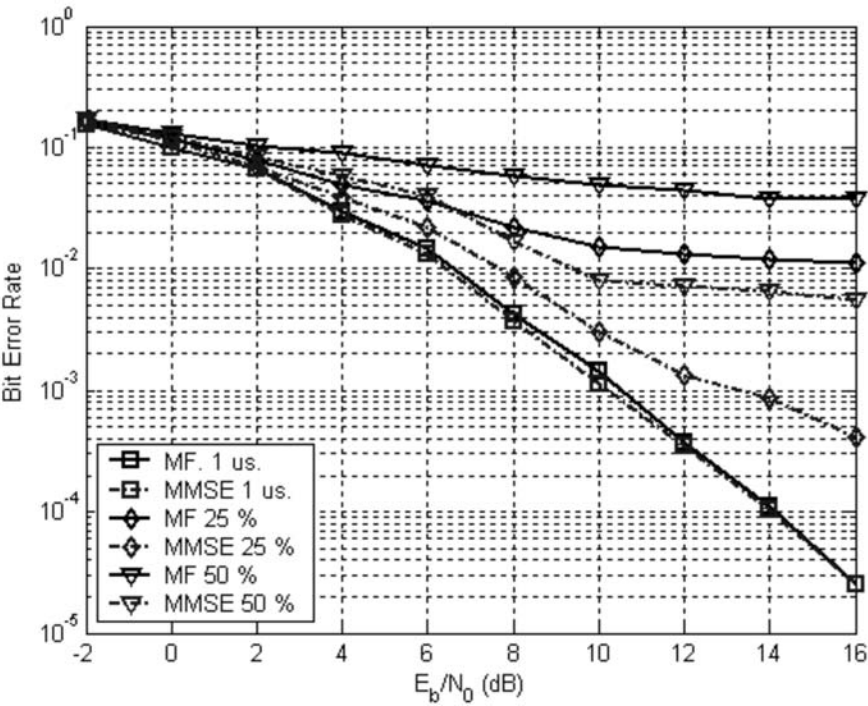


Figure 4.2 Raw BER comparison for block transmission for matched filter and MMSE receivers [23].

Table 4.1 Parameters for MMSE Receiver Simulation [23]

<i>Parameter</i>	<i>Value</i>
Channel model	Eight-tap, independent Rayleigh-faded
Power control	Average (slow) power control
Mobility	Static users
Multicell	No
Modulation	BPSK
Block size (BT)	64
Prefix length (BT)	32
Spreading gain	16
Spreading sequences	Binary random sequences
Number of users	1, 4, and 8

is to limit the ISI caused by channel time dispersion. A basic multiple-access scheme such as TDMA is mostly modulation agnostic [23]. Single-carrier schemes employing basic waveforms, however, require a contiguous frequency band for transmission, and cannot utilize sparse frequencies. Some frequency-domain spreading methods (e.g., FDOSS) can be considered to have single-carrier transmission in that their transmitters can be implemented with serial time-domain processing only. For highly bandwidth-efficient transmissions, however, these suffer from overhead due to the spreading operation.

In MIMO systems with multiuser transmission, the frequency-selective channel can also effectively separate users. This effect can be exploited by spatial filtering at the receiver. Thus, given multipath-rich propagation conditions resulting in high angular spread at the receiver (or low angular spread with different angles of arrival of users), the transmission can provide SDMA-type properties.

An initial evaluation of the link level performance of single-carrier, nonspread TDMA was performed for two receiver structures: the linear MMSE frequency-domain equalizer and the frequency-domain SC-MMSE turbo equalizer. A single user with full-bandwidth block transmission was assumed with the parameters as shown in Table 4.2.

Figures 4.3 and 4.4 show the calibration results for AWGN static single-path channels, and the reported performance is given for the linear MMSE equalizer

Table 4.2 SC-TDMA Simulation Parameters

<i>Parameter</i>	<i>Value</i>
Channel model	TDL
Scenario	Vehicular A
Power control	Average (slow) power control
Mobility	Static users
Multicell	No
Coding	Rate of 1/2, convolutional code with generators 561, 753
Modulation	QPSK, 16-QAM
Demodulator	MAP (not max)
Data symbols	1,536
Prefix length	256
Multiplexed pilots	128
Pulse shaping	Root-raised cosine, roll-off 0.23
Number of users	1
Turbo iterations	2

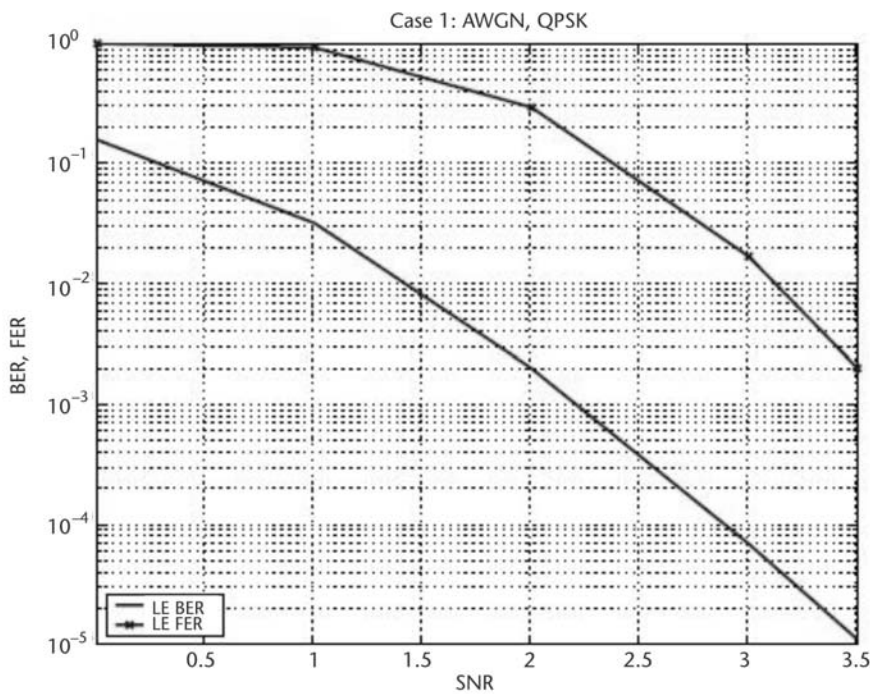


Figure 4.3 Calibration plot for AWGN (case 1) channel [23].

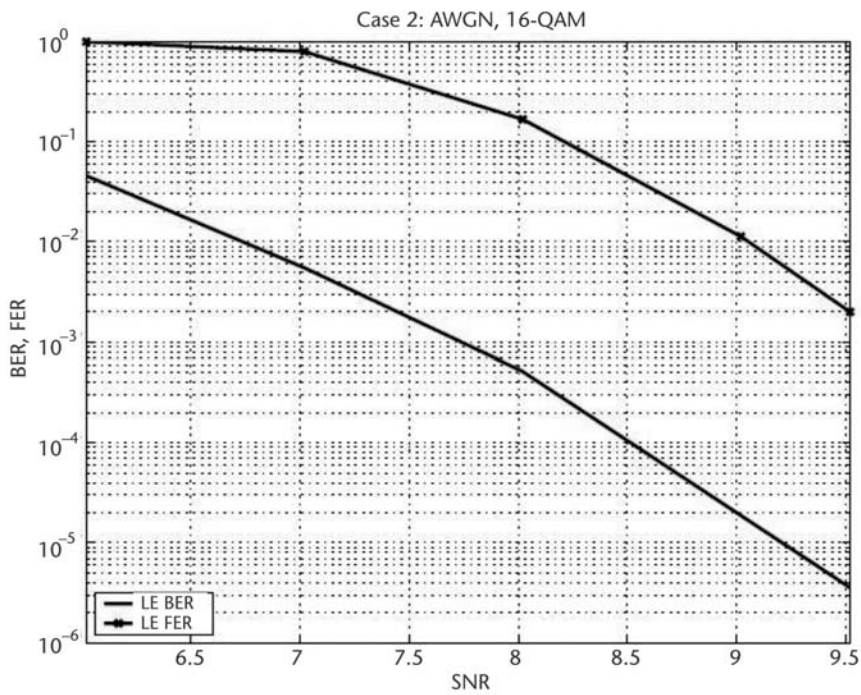


Figure 4.4 Calibration plot for AWGN (case 2) channel [23].

(LE). Figure 4.5 shows the same studies for a frequency-selective channel utilized to improve the performance of the turbo equalizer (TE). The TE performs two iterations of interference cancellation and MMSE filtering, and the performance of the last iteration is reported. In many cases, only one iteration would suffice.

The same studies were performed for an ITU Vehicular A channel, for which both cases 1 and 2 were evaluated. The results are shown in Figures 4.6 and 4.7. The TE now offers a performance gain of 1 to 2 dB over the LE in terms of frame error rate (FER)/BLER. The turbo gain is most visible in case 2 with 16-QAM modulation. Given that effective link throughput is mostly determined by the BLER, the improvement was significant.

Both the linear and the turbo equalizations can be considered to be effective in equalizing the frequency-selective channel. The utilization of frequency-domain methods enables the receivers to capture channel diversity while maintaining low computational complexity.

The trade-off between complexity and performance of different receiver structures for wideband TDMA is an interesting topic for further studies.

4.1.2.2 MC-CDMA

Multicarrier code-division multiple-access is a multiple-access scheme used in OFDM-based telecommunication systems that allows the system to support multiple users at the same time. MC-CDMA spreads each user symbol in the frequency domain. That is, each user symbol is carried over multiple parallel subcarriers, but

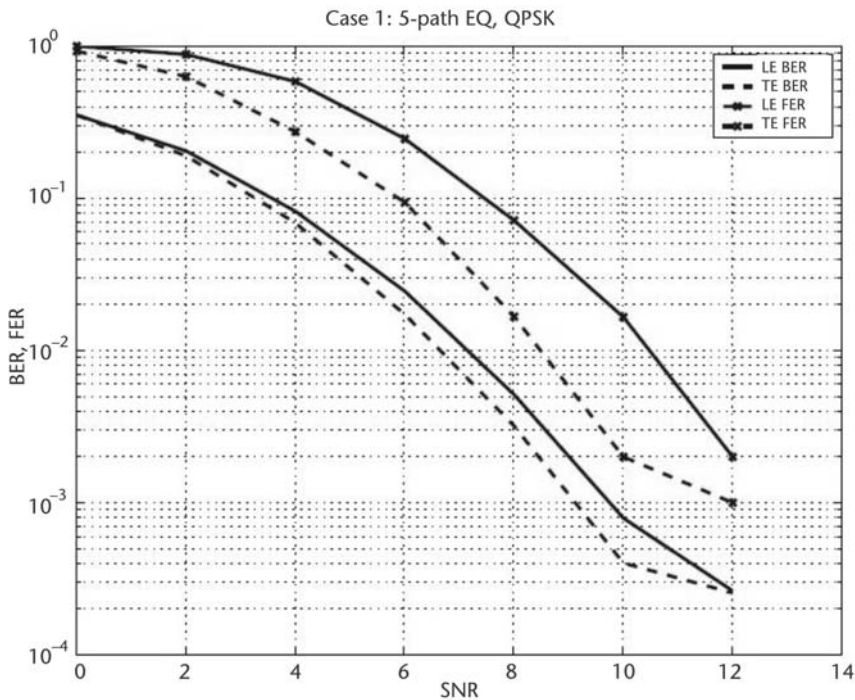


Figure 4.5 Calibration plot for multipath channel with Rayleigh fading (case 1) [23].

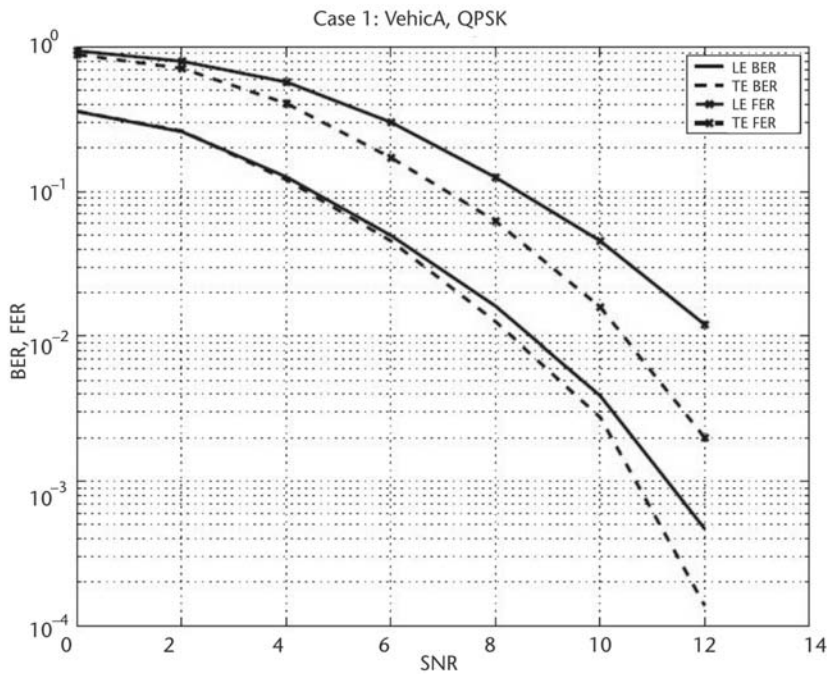


Figure 4.6 Vehicular A channel, case 1.

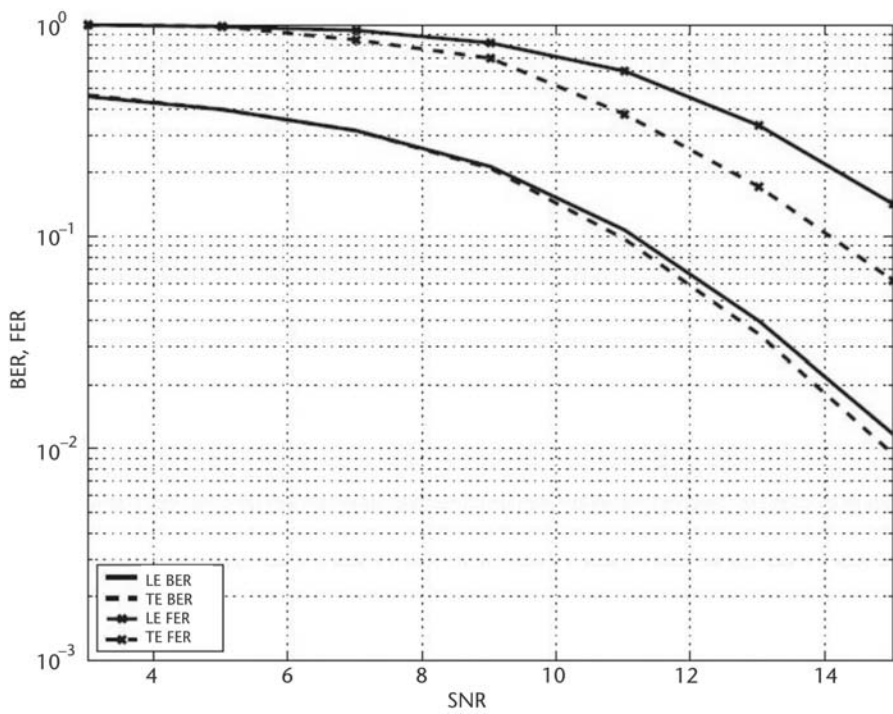


Figure 4.7 Vehicular A channel, case 2.

it is phase shifted (typically 0 or 180 degrees) according to a code value. The code values differ per subcarrier and per user. The receiver combines all subcarrier signals by weighing these to compensate varying signal strengths and undo the code shift. The receiver can separate the signals of different users, because these have different (e.g., orthogonal) code values.

The spreading operation from CDMA in MC-CDMA can be used for multiple-access schemes as well as for exploiting frequency diversity gain. Additionally, it is also possible to use the spreading operation for multiple code transmissions to increase the data rate. As a result, a system operating with MC-CDMA technology can be flexibly changed from single-device transmission with high-data-rate transmission to multiple-device transmission with a low data rate for each device. From each transmission link's point of view, scalability in terms of the data rate can be easily achieved by changing the length and number of spreading codes assigned to the link's modulation and coding scheme (see Chapter 3 of this book). Consequently, spreading and multiplexing are key features to enable flexible and scalable transmission in an MC-CDMA system.

4.1.2.3 OFDMA

In OFDMA, an individual subcarrier or group of subcarriers is assigned to different users. The most widely used methods for allocating users are *grouped subcarriers* and *interleaved spread subcarriers*. In grouped subcarriers each user is assigned a group of contiguous subcarriers. In the comb-spread subcarrier method, subcarriers are allocated in a fixed comb pattern, which causes them to spread over the entire system bandwidth. The grouped subcarriers method minimizes interuser interference, but is susceptible to fading, because some users' whole group of subcarriers may suffer from a null in the spectrum. The use of comb-spread subcarriers can partially overcome fading by frequency diversity, but such subcarriers are more susceptible to interuser interference in the case of imperfectly synchronized users.

In addition to OFDMA, hybrid TDMA/OFDMA can also be used to allocate time-frequency resources to users in the system. The advantages of OFDMA or hybrid TDMA/OFDMA identified by the IST WINNER project [23] are as follows:

- Robustness to frequency selective fading;
- No intercode interference (compared to CDMA-based schemes);
- Easy data rate adjustment (compared to TDMA);
- The possibility of adaptive scheduling;
- Increased spectral efficiency (no guard band needed as compared to FDMA).

The disadvantages are:

- Severe OFDMA performance degradation when the underlying channel undergoes deep fading; FEC (possibly in conjunction with frequency hopping) is imperative for OFDMA;
- Difficulty with subcarrier synchronization;
- Need for coordinated subcarrier assignments;

- Performance degradation in a multicell environment due to cochannel interference from other cells;
- All disadvantages typical for OFDM such as sensitivity to frequency offset and a high PAPR.

OFDMA has become an area of intensive research. Reference [26] considers the uplink performance of an OFDMA system for mobile communications. There the random OFDMA system is described based on multicarrier FDMA, in which each user has a set of randomly selected subchannels. The subchannels assigned to the individual user are analogous with the user's individual number. The subchannels are utilized by the user in any cell in which he may be located. The disadvantage of the proposed system is the problem of corrupted subchannels due to collisions with subchannels used by other users, which results in the loss of capacity. This loss, however, is at least partially compensated for by the fact that no signaling protocols for the assignment of subchannels to the users are needed. The handover is also simplified. The effect of Doppler spread in OFDM and OFDMA mobile radio systems showed the validity of Doppler frequency correction in the detection process [27].

Strictly connected with the OFDMA scheme is the problem of optimization of the OFDM system for many users [23]. Because of diverse channel patterns between the base station and user terminals, power allocation and bit loading become an important factor. In addition, different users may have different QoS requirements, which further complicates the problem of resource allocation. It has been shown that an OFDMA system with a multiuser subcarrier and power and bit allocation can offer a substantially larger capacity over OFDM/TDMA [28].

The problem of finding an optimal subcarrier and power allocation strategy for downlink communications to multiple users in an OFDMA system was analyzed in [29]. The criterion for system optimization was the minimization of total power consumption with constraints on the BER and the transmission rate for users requiring different classes of service. The problem of joint subcarrier and power allocation was divided into two steps. In the first step, the number of subcarriers assigned to each user was determined based on the users' average SNR. In the second stage, the algorithm found the best assignment of subcarriers to users.

Finally, OFDMA is also one of the modes adopted for the IEEE 802.16 wireless access [13]. OFDMA, due to its lower complexity and ability to maintain orthogonality on frequency-selective fading channels, outperforms other multiple-access schemes for higher system loads [23].

4.1.2.4 Adaptive Multiuser TDMA/OFDMA for Mobile Terminals

A powerful way to increase the spectral efficiency of a system is by opportunistic scheduling combined with link adaptation. Channel resources are then given to the user who can utilize them best [23, 30]. When the channels to K users are independently Rayleigh fading and have the same average power, the spectral efficiency (sum-of-rates capacity) will grow approximately as $\log(\log(K))$ if the channel is assigned to the user who has the highest instantaneous SNR.

Efficient use of opportunistic scheduling requires a new approach to radio system design, because channel variability is utilized, rather than counteracted, by, for example, interleaving. Traditional approaches for improving radio links by reducing their variability by diversity combining would in fact decrease the possibility of improving spectral efficiency by opportunistic scheduling.

There are two potential practical difficulties that might reduce the gains promised by opportunistic scheduling combined with link adaptation:

- Some channel state information is required at the transmitter. Due to time delays in the transmission feedback control loop, this information could be outdated and unreliable. Because the link adaptation will be based on outdated channel state information, the performance of channel predictors is an important determining factor of these adaptation schemes.
- The feedback overhead grows with the number of active users and with the speed of the time variability that is to be utilized. A feedback data rate that is too high could negate the potential gains attained by the adaptive scheme.

Opportunistic scheduling can be applied to time slots in TDMA systems. However, when the transmission bandwidth is increased above the coherence bandwidth, the channel variability between time slots will be reduced, due to independent fading within different parts of the total bandwidth. The opportunity for scheduling gains is then reduced. In wideband fading channels, it is attractive to allocate time-frequency resources (here denoted bins), each of which has an approximately flat channel. Adaptation over frequency as well as time may further potentially reduce the time delays encountered by the data streams to slow-moving users. For example, in an adaptive multiuser OFDM/TDMA downlink, predictions of the channel quality in each time-frequency bin would be generated for the time instance when these slots can be utilized for transmission. Based on these predictions and the requirements of different data streams, a scheduling algorithm would allocate the time-frequency bins to the different users. Adaptive modulation and possibly also power control by multiuser water filling is used to adapt the transmission to the link quality within each bin.

The most challenging environment in which to evaluate multiple-access schemes for next generation systems would be characterized by the following:

- *Wide-area coverage*: This type of coverage gives rise to large delay spreads and small correlation bandwidths. This would limit the useful frequency extent of time-frequency resources allocated to different users. The number of bins utilized within a given bandwidth thus becomes large, which increases the required feedback data rate.
- *High speed user mobility*: Users moving at high speeds, for instance, traveling on a train, have unique needs that increase the difficulty of predicting the channel sufficiently far ahead, and increase the required feedback data rate.
- *Large bandwidth*: If the whole bandwidth is to be potentially allocated to all users, the required feedback data rate would increase in proportion to the total bandwidth.

Adaptive systems that are shown to work acceptably under the conditions just listed would work better in the less challenging conditions of shorter range communication, lower expected terminal velocities, and smaller bandwidths.

In [31, 32], the downlink of an adaptive OFDM FDD system was adjusted for use in existing 3G bands of 5 MHz at 2 GHz. With 512 subcarriers, a subcarrier spacing of 10 kHz, and a guard time/cyclic prefix of 11 μ s, it was designed for macrocellular environments. The design vehicle speed was 100 kmph, and the assumed total feedback delay required channel prediction 2 ms ahead in time. At 2 GHz and 100 kmph, this corresponded to a required prediction range of 0.3 wavelength. The resulting spectral efficiency will depend crucially on the choice of scheduling strategy, on the deployment scenario, and on the possibility of reducing the ICI [32].

4.1.2.5 SDMA

Space-division multiple-access is a powerful multiple-access scheme that can support any of the previously described multiple-access schemes. It adds an additional dimension to a multiple-access scheme in the form of terminal angular separation. It is achieved by directing the beam toward the selected terminal. Depending on the advancement of the applied technique, beamforming can be applied in the uplink or in both the uplink and downlink. The efficiency of SDMA strongly depends on the angular distribution of terminals located in the cell or the range of an access point.

The key bottleneck for the application of SDMA is spatial separability [23]. Spatial separability means that for the successful transmission by two users sharing the same resource, certain beamforming vectors exist for which a certain signal-to-interference ratio (SIR) threshold will not be underrun. For a given channel, spatial separability depends on the number of users and on their spatial channel properties. In addition, beamforming vectors and transmission powers affect interference levels and SIRs at all receivers and thus affect spatial separability. This brings up the need for interference balancing. The joint optimization of antenna weights and decisions about spatial separability is a complex nonlinear problem. In [33] the second order statistics of the channel are used to minimize transmitting power with the proposed “power minimizer” algorithm or its linearized or simplified versions.

In a downlink multiuser environment, the achievable SIRs of all users are limited by the amount of mutual cross-talk. In [34] the beamforming weights also are to be determined in a joint optimization under the given channel conditions aimed at balancing the relative SIR levels.

A joint treatment of the MAC and PHY layers for SDMA promises to optimize throughput with the help of adaptive modulation and resource allocation. In [35] a combination of SDMA together with three different multiple-access schemes (TDMA, CDMA, and OFDMA) was investigated. The proposed heuristic algorithms are based on the greedy assignment strategies. Even with the loss of orthogonality caused by beamforming, the system load can be considerably increased by the combination of SDMA and CDMA [23].

All the above-mentioned algorithms have a nonnegligible computational cost. A cheap and simple alternative, which is suitable for macrocellular environments with small angular spreads, is to use a grid of fixed beams with reuse of resources of nonneighboring beams.

SDMA techniques reported in the literature use multiple antennas at the transmitter side, typically the BS or access point (AP) that applies to single-antenna terminals. Whereas in TDMA or FDMA the frame duration is limited by length or by the number of frequencies, in SDMA the low number of antennas severely limits the performance of the system [36].

Extensions to the case in which both transmitter and receiver have multiple antennas have been proposed in [37]. There two constrained solutions are described, namely, block diagonalization and successive optimization, which yield closed-form expressions to maximize the sum capacity and provide a reasonable trade-off between performance and complexity. In addition, they provide extensions if these two methods cannot be supported. In [38], the multiuser MIMO channel is decomposed into several parallel independent conventional single-user MIMO channels, in which, if a transmitting antenna is added, the number of spatial channels for each user is also increased by one. Finally, iterative algorithms are considered for uplink MIMO systems in [39] to find the optimum transmitting and receiving linear filters according to the mean-squared-error criterion, assuming an error-free, low-delay feedback channel.

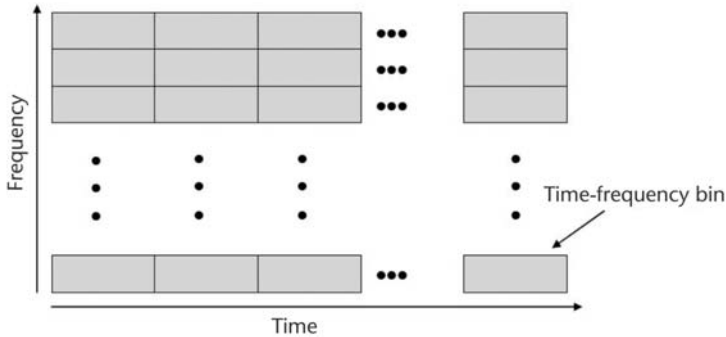
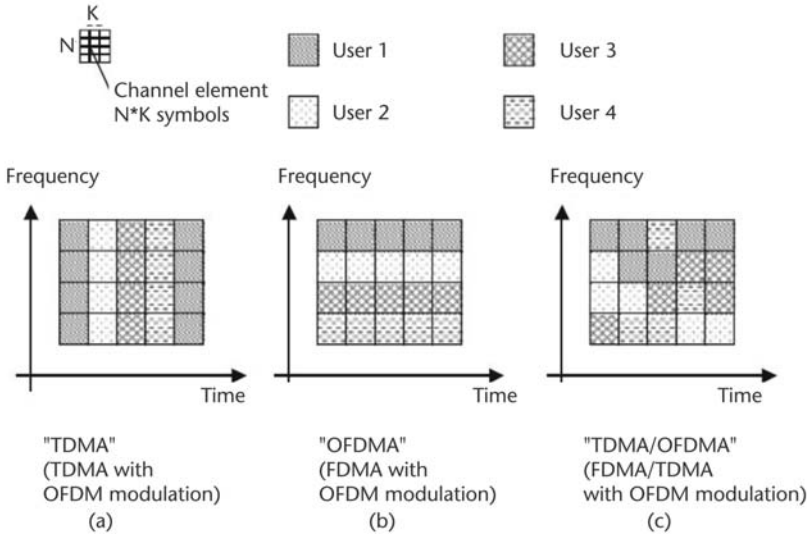
4.1.2.6 Definition and Evaluation Scenario of Multicarrier Solutions

The following scenario for the definition of the various multiple-access schemes was proposed in [23]. In the downlink the BS performs a multiplexing operation on the user's data streams onto the downlink channel, and in the uplink the channel access rules form a multiple-access scheme. The frequency adaptive multiple-access/multiplexing with OFDM modulation is denoted OFDMA, the time adaptive multiple-access/multiplexing with OFDM modulation as TDMA, and the time-frequency adaptive multiple-access/multiplexing with OFDM modulation as TDMA/OFDMA. These schemes are shown in Figure 4.8.

Figure 4.8 assumes that the minimum transmission unit is a channel element (bin) comprising N subcarriers in frequency and extending K symbols in time (so in total, there are $N \times K$ symbols within a channel element). In all cases, provided that $N > 1$, multicarrier transmission is achieved.

In the TDMA/OFDMA system, time-frequency resources, called bins, are allocated to users. The allocation is assumed to be exclusive in the single-antenna systems studied here. With later generalization to multiple antennas within cells, it becomes possible to let several users share a time-frequency bin. The bins are assumed to be of equal size (number of adjacent subcarriers times number of symbols) over the whole system bandwidth. The appropriate bin size in the radio propagation scenarios defined for next generation systems (e.g., WINNER) and user traffic scenarios is an important design parameter. A frame of time-frequency bins is shown in Figure 4.9.

The various issues that arise when designing TDMA/OFDMA multiple-access systems—utilizing link adaptation, applying multiuser scheduling, and balancing



multiple antenna resources—should not be studied in isolation. This is true, in particular, for a radio interface design that takes all of these aspects into account.

Both channel predictability aspects and control signalling overhead aspects are affected in nontrivial ways by the choice of duplex scheme. A detailed TDMA/OFDMA system design should consider and analyze proposals with both TDD and FDD.

In OFDMA systems, typically bins at a certain frequency are allocated statically or semistatically to the individual users. The subcarriers in a frequency bin can be either contiguous or picked up from interleaved positions over the entire available subcarriers. The former can be referred to as block OFDMA (BOFDMA), and the latter can be referred to as interleaved OFDMA (IOFDMA). That is, the BOFDMA is OFDMA without subcarrier interleaving and IOFDMA is OFDMA with subcarrier interleaving.

Fast link adaptation in an OFDMA system can be attractive because it can increase the spectral efficiency of the system. The difference with TDMA/OFDMA

is that fast user scheduling between bins at different frequencies is not performed. If link adaptation with adaptive frequency resource allocation is going to be used with an OFDMA scheme, BOFDMA is apparently the right choice. This is because in IOFDMA the subcarrier interleaving provides an averaging effect, thus the property of each time-frequency bin will tend to be similar. If link adaptation with fixed frequency resource allocation is used, either BOFDMA or IOFDMA can be adopted, but the performance depends on the channel conditions.

TDMA combined with OFDM (denoted earlier as TDMA in Figure 4.8) is a special case of TDMA/OFDMA. In TDMA the bin, or time slot, consists of all subcarriers and one or many OFDM symbol(s).

4.2 Analysis of TDMA/OFDMA for an IMT-A Candidate System

4.2.1 Key Design Aspects

4.2.1.1 Link Adaptation

With link adaptation the transmitted data rate is adjusted according to the measured channel quality in both time and frequency. By using a flexible resource allocation algorithm, QoS targets can be fulfilled and the overall system capacity can be increased without the need for fast power control. The feasibility of using a fast link adaptation was in the downlink, considering the typical channel propagation conditions, the transmission technology, and the transmission parameters defined in the context of the WINNER air interface [40]. The conservative bounds on the maximum vehicle speed for adaptive transmission based on channel prediction were derived based on input from basic assumptions in the literature on loop delay and needed feedback information.

The main requirement for beneficial link adaptation is accurate knowledge of the channel state information at the transmitter [23]. In the fast fading case, the CSI in the future transmission period will be different from that of the current period. Therefore, channel estimation and/or prediction must be performed and the CSI must be available at the transmitter as soon as possible in order to avoid mismatch between the CSI used in the link adaptation and the actual channel state. This can be realized in the following two ways [23]:

- For slow fading channels, once the channel estimate has provided the CSI, it is sufficient to keep the adaptation update period much shorter than the channel coherence time, for example, $\leq 0.1/f_{D\max}$, (where $f_{D\max}$ is the maximum Doppler spread of the channel), so that the channel can be regarded as constant during the period. In this case, a new channel estimation/prediction is not necessary. For fast fading channels, however, this restriction on the update period will become unrealistic.
- For fast fading channels (i.e., large Doppler spread or short coherence time), the channel can no longer be treated as time invariant. In this case, channel estimation/prediction must be performed for the future transmission period.

The link adaptation procedure is as follows [23]:

- In time slot j , all active terminals predict the CSI (i.e., SINR) at time slot $j + n_{\text{delay}}$ for all or part of the time-frequency bins, where n_{delay} is the reaction delay in terms of number of time slots.
- The terminals send back to the BS their predicted CSI. In [23] it was assumed that 3 bits are enough to indicate eight levels of channel state.
- Scheduling is then performed by the BS, and allocation decisions (including resource allocation, modulation, and coding) for time slot $j + n_{\text{delay}}$ are broadcast.

For effective implementation of link adaptation, the system design should meet the following three requirements:

$$T_{\text{slot}} \leq \frac{0.1}{f_D} \quad (4.1)$$

$$T_{\text{delay}} \leq \frac{0.1}{3f_D} \quad (4.2)$$

$$\Delta f_b \leq \frac{1}{5\tau} \quad (4.3)$$

where f_D and τ are channel Doppler spread and channel delay spread, respectively, of a particular user; T_{slot} is the time slot duration; and T_{delay} is the reaction delay.

Regardless of which channel prediction/estimation method (MMSE, LS, etc.) is employed, to obtain a usable (i.e., accurate enough) channel prediction, the prediction step should be shorter than [40–42]:

$$\frac{1}{3f_{D_{\text{max}}}} \quad (4.4)$$

(i.e., $\frac{1}{3} \lambda$ of terminal moving range), where $f_{D_{\text{max}}}$ is the maximum Doppler spread, and λ is the carrier wavelength.

The prediction of the channel state is done over a time horizon D , which covers the total feedback delay. Different designs for TDMA/OFDMA systems along with different duplex schemes will result in different feedback delays. The feedback delay limits the largest terminal velocity v for which fast link adaptation can be used. The key property is the required normalized prediction horizon in space, expressed as a fraction of the carrier wavelength λ and prediction horizon in time D :

$$\frac{vD}{\lambda} \quad (4.5)$$

Prediction over significantly more than $1/3$ carrier wavelength will not be feasible in most propagation environments under realistic assumptions for the SNR and the utilized prediction algorithm [41]. Equation (4.5) also deals with the

dependence of the carrier frequency. The lower the carrier frequency, the larger the maximal user speed that can be supported for a predictor with a certain prediction horizon.

The frame structure for an adaptive TDMA/OFDMA TDD scheme based on channel prediction and investigated for vehicular velocities at carrier frequency of 5 GHz is shown in Figure 4.10.

Each TDD frame contains an uplink slot, a downlink slot, and two time guard symbols. The total frame length is always 32 OFDM symbols, that is, $720 \mu\text{s}$. In symmetric UL/DL operation, the UL bin and DL bin have a default size of 15 OFDM symbols, but in 1:2 and 2:1 operation, the UL and DL bin lengths are 10:20 and 20:10 OFDM symbols, respectively.

The total system bandwidth may be up to 100 MHz, but may also be smaller (narrowband case). The total bandwidth is split into subbands, called *contention bands*, to reduce the required feedback data rate. Each terminal is in competition for bins in one or several contention bands. The contention bandwidth is set to 52 bin widths, that is, $10.4 \text{ MHz} = 208 \text{ subcarriers} = 1/8$ of the maximum 1,664 available subcarriers within the maximum 84-MHz useful bandwidth in the 100-MHz system bandwidth. The contention bandwidth is a parameter that needs to be considered with different traffic models [23].

To make a fair comparison of the systems using TDD and FDD, the total useful bandwidth for the two modes should be the same. Thus, parameters for the FDD mode are based on the TDD mode with proper adjustment for the bandwidth [23]. The duplex distance is assumed to be $15 \times \text{BW}$, but the channel characteristics for the UL and DL carriers are assumed to be described by the same channel model as for the 5-GHz TDD channel model. Figures 4.11 and 4.12 show the FDD frame structures for the downlink and uplink, respectively.

Each FDD DL/UL frame consists of two slots. The total frame length is always 30 OFDM symbols, that is, $675 \mu\text{s}$, and the UL bin and DL bin have a default size of 15 OFDM symbols. This makes the FDD system comparable with the TDD system with respect to slot size, but the frame length could be changed to 32 OFDM symbols in order to make coexistence with the TDD system possible, for example, in the form of a hybrid TDD/FDD system.

Total system bandwidth may be up to 50 MHz for the UL and DL, respectively, but may also be smaller for the UL and/or DL. However, the total bandwidth is

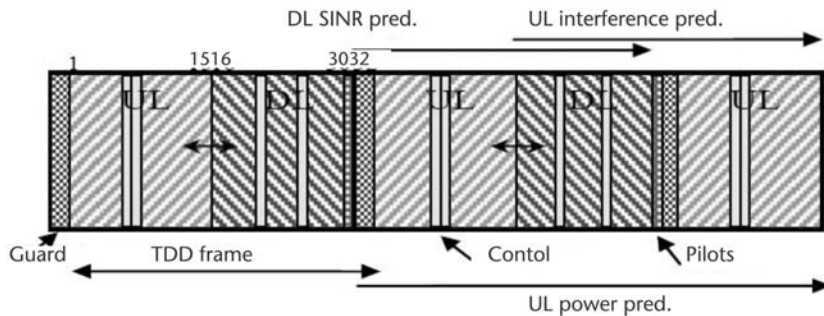


Figure 4.10 Frame structure for an adaptive TDMA/OFDMA TDD scheme based on channel prediction [23].

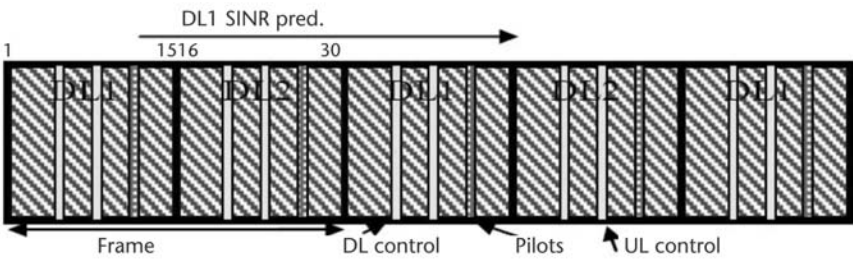


Figure 4.11 Downlink frame structure for adaptive TDMA/OFDMA FDD scheme based on channel prediction [23].

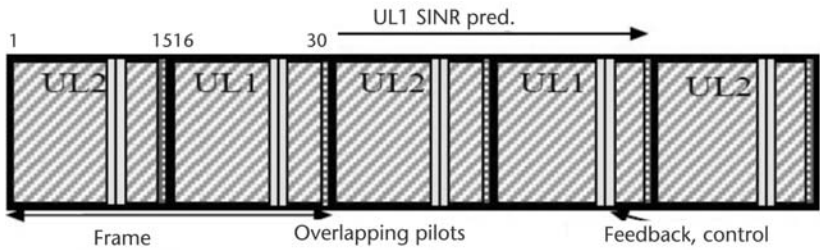


Figure 4.12 Uplink frame structure for adaptive TDMA/OFDMA FDD scheme based on channel prediction [23].

split into contention bands. As for the TDD system above, each terminal is in competition for bins in one or several contention bands. The contention bandwidth is set to 52 bin widths, that is, $10.4 \text{ MHz} = 208 \text{ subcarriers} = 1/4$ of the maximum 832 available subcarriers within the maximum 42-MHz useful bandwidth in the 50-MHz UL/DL bandwidth. The contention bandwidth is a parameter that needs further study, for example, in conjunction with traffic models.

UL/DL system level symmetry is static and the asymmetry ratio depends on the available bandwidth around the UL and DL carrier, with default asymmetry ratio 1:1, and can be adapted at the time of deployment to the available bandwidth.

The disadvantages for a BS to work in full-duplex FDD are acceptable, but for the terminals a half-duplex mode is very beneficial with respect to cost and power consumption. The full-duplex mode is feasible, however, for moving APs serving moving networks, homogeneous (i.e., same radio interface on both sides) relays regarded as terminals from the BS side, and possibly high-end terminals. Thus, both full-duplex and half-duplex terminals should be supported.

Half-duplex terminals are divided into two sets. In one of the sets, the terminals are allowed to transmit/receive in the UL1/DL1 slots, and terminals in the other set are allowed to transmit/receive in the UL2/DL2 slots. Full-duplex terminals may transmit/receive in both of the UL/DL slots and can be regarded as terminals consisting of one terminal from both sets. The prediction horizon in the UL and DL is the same for full-duplex terminals as for half-duplex terminals. The main difference is that the channel predictors in the BS and full-duplex terminals can be updated with pilot information from each UL/DL slot, making it possible to improve

the prediction accuracy. This allows for a longer normalized prediction horizon in terms of wavelengths. Thus, full-duplex terminals could be supported with a higher maximal speed and/or larger throughput due to the more robust choice of modulation and coding scheme used in the presence of a smaller prediction error (in addition to the basic doubling of throughput compared to half-duplex terminals).

4.2.1.2 Scheduling

Multiuser scheduling provides an additional scheduling gain, called *multiuser diversity*. The reason for this is that the variations of channels to/from different terminals will be uncorrelated. The spectral efficiency of the system may be then increased by allocating resources to the users with the instantaneously best channels, under the QoS constraints of the active calls per session. The total throughput then increases with the number of active users in a cell, in contrast to CDMA systems, where it decreases with the number of users due to system self-interference. In a spectrally efficient adaptive TDMA/OFDMA system, the scheduling algorithm should try to maximize the system throughput and at the same time provide a certain degree of fairness between the users and their user data flows. In general, the fairness criteria between users and system capacity with respect to different QoS classes should be tunable in order to satisfy operators having different business models regarding offered services and user subscription models.

4.2.1.3 ICI Mitigation Under Small Frequency Reuse Factor

Intercell interference avoidance is another important application for scheduling to reduce the frequency reuse factor in cellular systems, by coordinating users at the cell borders of adjacent cells. This is especially important in TDD systems. The SINR within cells must be balanced with the resource reuse between cells. To reduce interference, scheduling can easily be coordinated between sectors/beams of a radio AP.

To reduce interference between opposing sectors, a very simple semistatic solution is to use a separate time-frequency allocation in high-interference areas, which is reused with a reuse factor of 3 [32]. This results in a system with good SIR in most parts of sectors within a regular hexagonal pattern. The total reuse factor is between 1 and 2. A more flexible approach, with potentially better performance in less than fully loaded systems, is to coordinate the use of time-frequency bins, with respect to the opposing sector, in the high-interference outer parts of the sectors. Another tool is to use relays in outer parts of sectors. However, both of these methods imply increased control signaling between base stations/relays and additional signaling delays.

4.2.2 Performance Modeling and Evaluation

4.2.2.1 Prediction Error Model

The scheduler needs to know the predicted channel power. For simulations aimed at investigating only the adaptive TDMA/OFDMA approach, channel prediction

is not performed. Instead a prediction error can be added, generated by a prediction error model, to the actual simulated channel.

Assume that the complex scalar channel $h(t)$ is predicted by a linear MMSE estimator, for example, a Kalman predictor in the frequency domain, designed to minimize

$$\sigma_{\epsilon_b}^2 = E|h(t) - \hat{h}(t | t - L)|^2 \quad (4.6)$$

The received power is assumed to be predicted by an unbiased quadratic power predictor [41]. For a given prediction of the complex channel, this operation can be expressed by

$$\hat{p}(t | t - L) = |\hat{h}(t | t - L)|^2 + \sigma_{\epsilon_b}^2 \quad (4.7)$$

In multilink and system level simulators, a model is needed for generating predicted channels in a simple way, without having to run numerous prediction algorithms for each active terminal.

Such a prediction error model, using a complex Gaussian prediction error with appropriate variance, can then be defined as

$$\hat{h}(t | t - L) = ch(t) + \Delta h(t) \quad (4.8)$$

where

$$c = 1 - (\sigma_{\epsilon_b}^2 / \sigma_h^2) \quad (4.9)$$

and

$$\Delta h(t) \in N(0, c(1 - c)\sigma_h^2) \quad (4.10)$$

with σ_h^2 being the channel power. The complex-valued error $\Delta h(t)$ is selected as constant within a bin. This model will reproduce not only the correct variance, but also the correct PDF of the prediction error. For a given vehicle velocity, SNIR, and predictor implementation, the complex prediction error variance in (4.6) is assumed to be known. Predicted channel samples, with appropriate statistical properties of the prediction error, can then be reproduced within a multilink simulator by drawing one complex-valued Gaussian random number per bin [23]. Link adaptation and scheduling routines in a multilink simulator utilize these predicted channel values.

4.2.2.2 Prediction Performance in TDD and FDD Downlinks and Uplinks

The unbiased quadratic predictor in (4.7) is used to predict the channel power, whereas the Kalman algorithm described in [42] has been utilized to predict the

complex channel for each 50-kHz subcarrier. The Kalman predictor utilizes the correlation of the channel in the frequency domain by predicting eight subcarriers (400 kHz) in parallel. Autoregressive models of order 4 are used to model the channel correlation properties in the time domain. They are adjusted to Jakes' fading statistics, with the maximal Doppler frequency assumed to be known. These algorithms have been applied on channels with a Vehicular A power delay profile and noise with known power. The predictors have been applied to the following cases:

- FDD downlinks, in which the channel state estimation is performed in the terminal. It is based on the pilot symbols and also on decision-directed estimation of the DL control symbols.
- FDD uplinks, in which channel prediction is performed at the access point, using overlapping pilots from the terminals. A generalization of the Kalman algorithm, described in [40] is utilized. Its performance is investigated for the cases of two, four, and eight simultaneous users in each contention band, all with the same received power.
- TDD downlinks and uplinks, in which the channel power is predicted by the terminals, and reports are transmitted to the access point via the UL control symbols. The prediction is performed based on the DL pilots and control symbols that are used for decision-directed estimation.

Figure 4.13 shows the prediction error, as a function of the prediction horizon scaled in wavelengths, for different values of the SNR. The prediction accuracy depends on the horizon in wavelength, which in turn depends on the vehicle velocity and the carrier wavelength via (4.1). It also depends on the SNR. For a given SNR, there is therefore a maximum value of the velocity that allows reliable adaptive transmission. For combinations of velocities and SNRs beyond such a boundary, nonadaptive transmission must be used. The normalized prediction error variance is given by $\sigma_{\epsilon_h}^2 / \sigma_h^2$.

Figures 4.14 and 4.15 show the same results but for FDD downlinks, where two, four, and eight users simultaneously transmit overlapping pilots, all with the same average received power. In a Kalman estimator based on overlapping pilots, separate sets of states are used for describing the channel of each user. The autoregressive models that describe the fading statistics of each user are adjusted individually to the velocity of the users [42]. Uplink control information could be used for improving the estimate by decision-directed methods. However, it is not utilized here, because the exact form of coding and signaling for the FDD uplink control data has not yet been specified. In the results presented here, all users have the same velocity and travel through the same type of propagation environment, but their channels were generated as independent realizations from this channel statistics. Furthermore, the average received power is the same for all users (slow power control is assumed to be used).

The results indicate that prediction based on overlapping pilots will decrease in accuracy with an increasing number of users, but that this decrease is rather modest. Channel predictions in FDD uplinks in which not too many users occupy

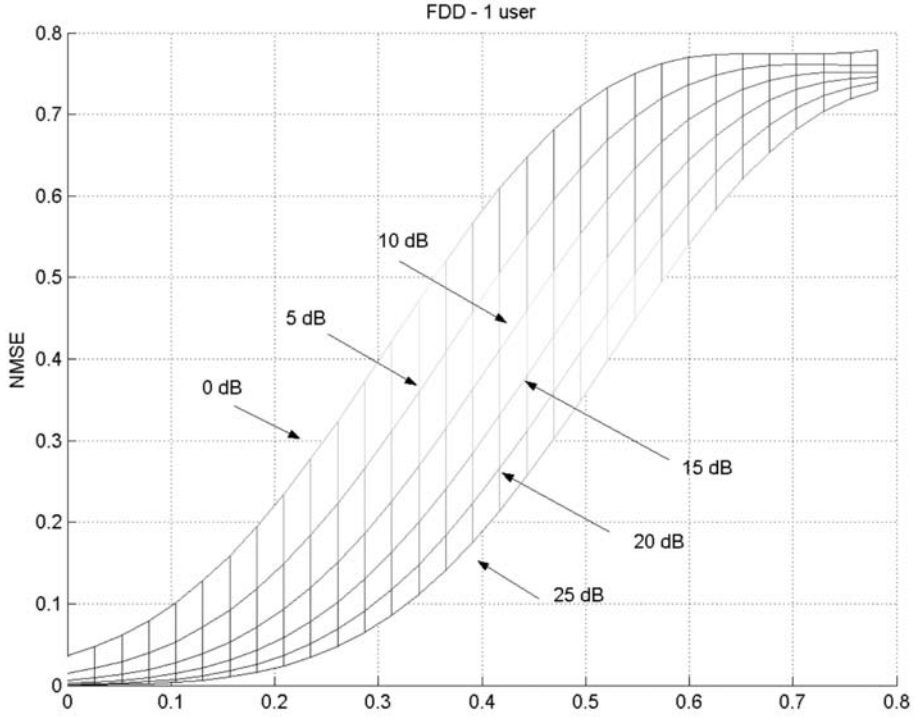


Figure 4.13 Prediction accuracy in terms of the complex normalized prediction error, as a function of the prediction horizon scaled in carrier wavelengths, and as function of the SNR. Results for FDD downlink over a Vehicular A channel, using a Kalman algorithm that utilizes eight subcarriers in parallel [23].

each contention band of the total bandwidth thus seem feasible. For more details, see [23]. Figure 4.16 shows the corresponding results for TDD downlinks and uplinks.

In the proposed TDD system, both UL and DL channel prediction is performed at the terminal, based on the common pilot symbols that are transmitted from the access point. Thus, the performance degradation due to use of overlapping pilots is avoided. Compared to the FDD downlink, the prediction accuracy is somewhat worse because the TDD duty cycle, which halves the rate of downlink pilots. Note also that the required prediction horizon is higher for the TDD uplink than for any of the other cases. Combined, this results in the TDD uplink having the worst performance for a given terminal velocity and SNR.

Table 4.3 summarizes the required prediction horizons scaled in wavelengths for the TDD and FDD systems under consideration operating at 5 GHz. Equation (4.5) gives the horizon as a fraction of the wavelength:

$$\frac{\nu \text{ [km/h]} \times D \text{ [ms]}}{216} \quad (4.11)$$

Adaptive transmission to/from a terminal will be feasible up to a maximal velocity, for a given SNIR, or, equivalently, down to a limiting SNIR at a given

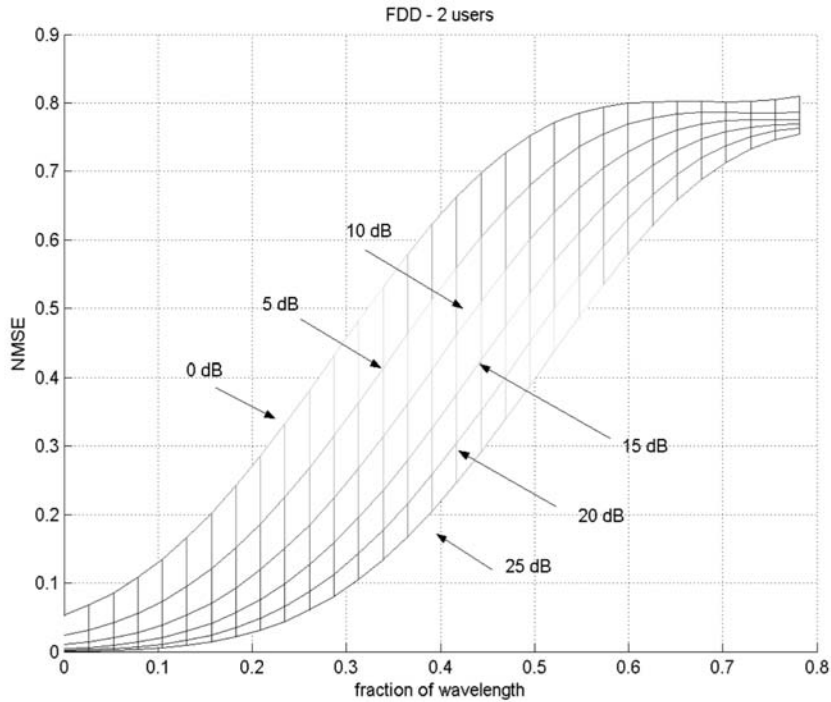


Figure 4.14 Prediction accuracy in terms of the complex normalized prediction error, as a function of the prediction horizon scaled in carrier wavelengths and as a function of the SNR. Results are for an FDD uplink over a Vehicular A channel, using a Kalman algorithm for overlapping uplink pilots that utilizes eight subcarriers in parallel. Average result is for two simultaneous UL users in the contention band [23].

velocity. A preliminary estimate of the limiting SNIR is based on the results in Figures 4.13 through 4.16.

From earlier investigations of the sensitivity of uncoded and Trellis-coded adaptive modulation to prediction errors, it has been found [23] that if the rate limits are adjusted to take the prediction uncertainty into account, a power prediction normalized mean square error (NMSE) of 0.1 leads to only a minor degradation in the attained spectral efficiency [43–45].

A limit of 0.20 for the normalized variance $\sigma_{\epsilon_b}^2 / \sigma_b^2$ of the complex prediction error was assumed in [23]. For higher values, a nonadaptive fallback mode is to be utilized instead of adaptive transmission. For Rayleigh fading channels, this corresponds to a power prediction NMSE of 0.18 [42]. The resulting limits for the SNIR are shown in Table 4.4. The corresponding prediction horizons from Table 4.3 are indicated within parentheses.

The results presented in Table 4.4 are of interest for qualitative comparisons among the links. These show that adaptive transmission works in the widest variety of situations in the FDD downlinks, whereas it works only in the narrowest range of circumstances in the TDD uplink [23].

These results are based on the assumptions and scalings of the presented uplinks and downlinks and were performed with the WINNER air interface as a reference

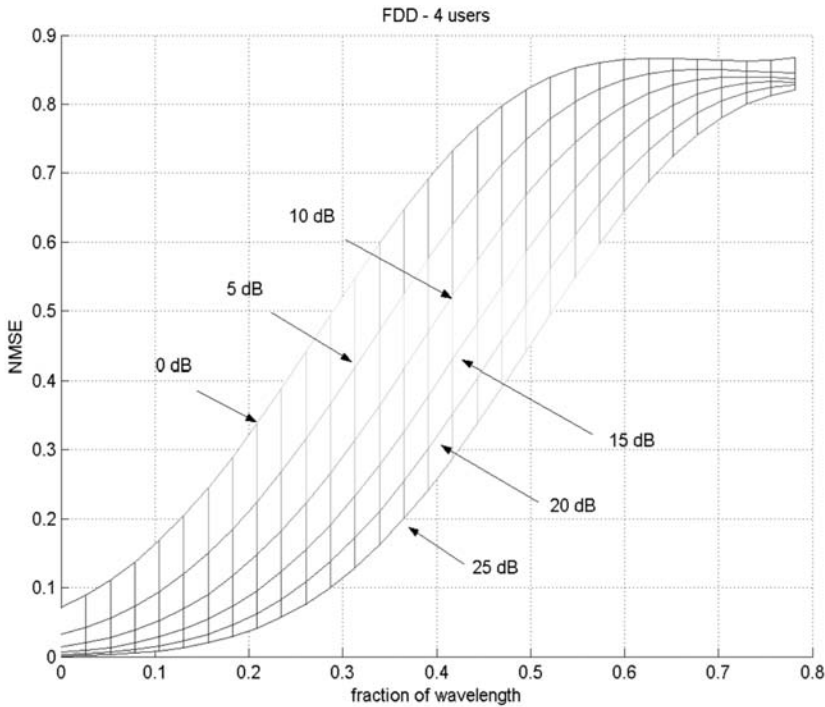


Figure 4.15 Prediction accuracy in terms of the complex normalized prediction error, as a function of the prediction horizon scaled in carrier wavelengths and as a function of the SNR. Results are for an FDD uplink over a Vehicular A channel, using a Kalman algorithm for overlapping uplink pilots, that utilizes eight subcarriers in parallel. Average result is for four simultaneous UL users in the contention band [23].

system. The results are based on the assumption that tight feedback loops are indeed possible. They take only channel prediction into account and neglect the added uncertainty that would be generated by interference prediction. On the other hand, the results are somewhat conservative in that the prediction is performed at the far end of the bin to be allocated. The numerical results can be modified further based on continued investigations into the sensitivity of the attained spectral efficiency and packet error rates to prediction errors. Follow-up studies and results are available in [23].

Potential gain in total AP throughput, measured in terms of bits per payload symbol within an FDD downlink, can be obtained when combining proportional fair scheduling with link adaptation. In Figure 4.17, the users are placed at the same distance, so they have the same average SNR in Vehicular A channels.

All users in Figure 4.17 move at 35 kmph. Link adaptation and scheduling are based on the predicted power in the scheduled bins. The required prediction horizon in the FDD downlink is 0.72 ms, which for 35 kmph at 5 GHz corresponds to a distance of 0.12 carrier wavelength. Link adaptation is here performed on a single-user basis, using eight modulation-coding rates, from BPSK rate 1/2 to 64-QAM rate 5/6. The rate limits are adjusted taking the prediction inaccuracy into account. The spectral efficiency is maximized for a single-user link, under a maximal BER constraint of 0.0001. All active terminals predict the channels for all bins, and

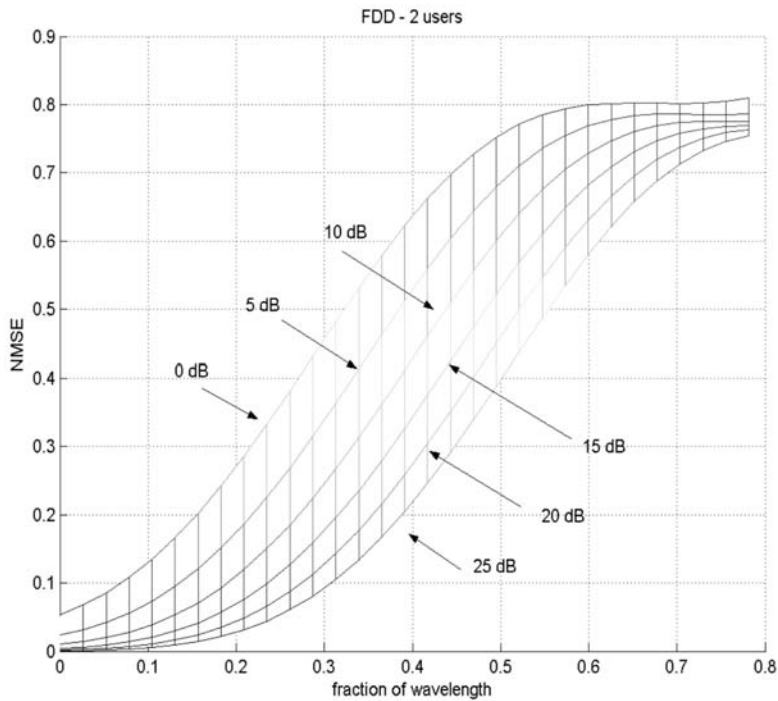


Figure 4.16 Prediction accuracy in terms of the complex normalized prediction error, as a function of the prediction horizon scaled in carrier wavelengths and as a function of the SNR. Results are for a TDD uplink and downlink over a Vehicular A channel, using a Kalman algorithm based on using eight subcarriers in parallel [23].

Table 4.3 Maximal Prediction Horizons in the TDD and FDD Adaptive Multiple-Access Schemes [23]

	Maximal Horizon D (ms)	Fraction of λ at 35 kmph	Fraction of λ at 50 kmph	Fraction of λ at 70 kmph
TDD downlink	0.70	0.11	0.16	0.23
TDD uplink (1:1)	1.10	0.18	0.25	0.35
FDD downlink	0.72	0.12	0.17	0.23
FDD uplink	0.65	0.11	0.15	0.21

Table 4.4 Hypothetical SNR Limits Delineating the Use of Adaptive Transmission Exemplified for Three Terminal Velocities for a 5-GHz Carrier Frequency [23]

	35 kmph	50 kmph	70 kmph
TDD downlink	<0 dB (0.11 λ)	1 dB (0.16 λ)	8 dB (0.23 λ)
TDD uplink	5 dB (0.18 λ)	9 dB (0.25 λ)	24 dB (0.35 λ)
FDD downlink	<0 dB (0.12 λ)	0 dB (0.17 λ)	4 dB (0.23 λ)
FDD uplink, two users	<0 dB (0.11 λ)	0 dB (0.15 λ)	5 dB (0.21 λ)
FDD uplink, four users	<0 dB (0.11 λ)	2 dB (0.15 λ)	8 dB (0.21 λ)
FDD uplink, eight users	1 dB (0.11 λ)	4 dB (0.15 λ)	14 dB (0.21 λ)

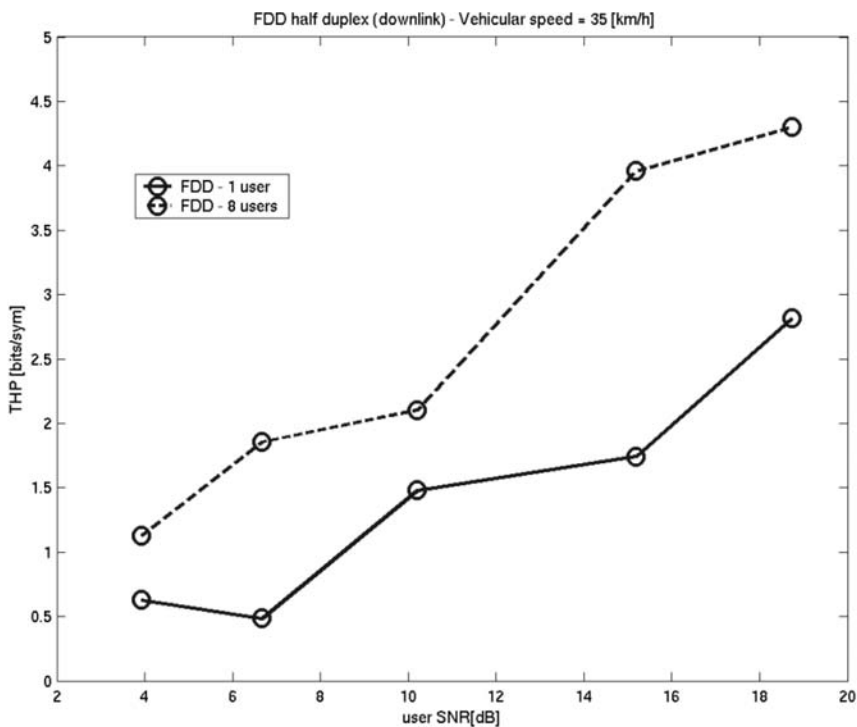


Figure 4.17 Throughput in bits per payload symbols, as a function of the average SNR, for an adaptive FDD half-duplex downlink with one user (solid) and for eight users who obtain resources by proportional fair scheduling (dashed) [23].

report the resulting rates via the uplink feedback channel. The scheduler allocates each bin to the user who reports the highest rate, relative to its own average. Vehicular A channel statistics and the transmission parameters are shown in Table 4.5 [23]. As an additional comparison, the simulation results for an adaptive OFDMA system and the WINNER parameters are listed in Table 4.6.

In an OFDMA uplink transmission, several parameters can be adapted as a function of the channel conditions and the traffic type. To tackle this highly complex optimization problem, it can be separated in two main steps:

Table 4.5 Parameters at 5 GHz for FDD Mode

Parameter	Value	Unit and Notes
Center frequency	5 ± 0.384	GHz
Number of subcarriers in OFDM	1,024	Equals the length of FFT
FFT BW	51.2	MHz; fits with the 50-kHz subcarrier separation
Signal BW	41.6	MHz; due to implementation reasons, only about 80% of the channel BW can be used
Number of subcarriers in use	832	Subcarriers [-416: 416]; 0 is not used
Subcarrier spacing	50	kHz
OFDM symbol length (excluding cyclic prefix)	20	μs
Cyclic prefix length	2.5	μs (corresponds to 256 samples)

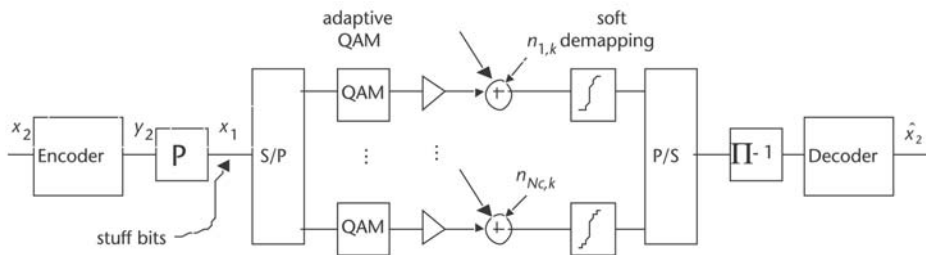
Table 4.6 Simulation Parameters for an OFDMA Uplink [3, 23, 40]

Parameter	Value
Channel model	SCM; implementation of WP5
Winner scenario	C2
Sampling time	$T_A = 5/512 \mu s$
Number of subcarriers	$N_C = 1,664$
FFT size	$N_{FFT} = 2,048$
SNR gap	$\Gamma = 2.6 \text{ dB}$
Channel code	convolutional code (561,753)
Detection	Soft demodulation, soft Viterbi decoding
<i>SCM Parameters</i>	
Number of BS antennas	scmpar.NumBsElements = 1
Number of MS antennas	scmpar.NumMsElements = 1
	scmpar.Scenario = 'urban_macro'
	scmpar.NumTimeSamples = 1
	scmpar.UniformTimeSampling = 'yes'
	scmpar.CenterFrequency = 5e9
	scmpar.DelaySamplingInterval = T_A
	scmpar.PathLossModelUsed = 'yes'
	scmpar.ShadowingModelUsed = 'yes'
	linkpar.MsVelocity = 0.01*ones(1,U);
	linkpar.MsBsDistance = [35 88 116 138 156 172 186 200]

1. *Subcarrier allocation*: This step determines which subcarrier is allocated to which user.
2. *Bit and power loading*: After the subcarriers have been assigned to the users, the problem of adaptively assigning the appropriate QAM constellation and power for each subcarrier can be solved for each user individually. The simulation chain for one user in an OFDMA uplink is shown in Figure 4.18, and the frequency responses for eight users are shown in Figure 4.19.

The corresponding subcarrier and bit allocation has been derived with the multiuser, water-filling algorithm [23], which is shown in Figure 4.20 and Table 4.7. Note that the convergence of this algorithm is not guaranteed.

For the given channel parameters and the corresponding subcarrier and bit and power allocation, which has been obtained by the multiuser water-filling algorithm, a Monte Carlo simulation was set up to estimate the BER and PER for all eight users [23]. In a more realistic scenario, each user was given a minimum

**Figure 4.18** Simulation chain for one user in OFDMA uplink.

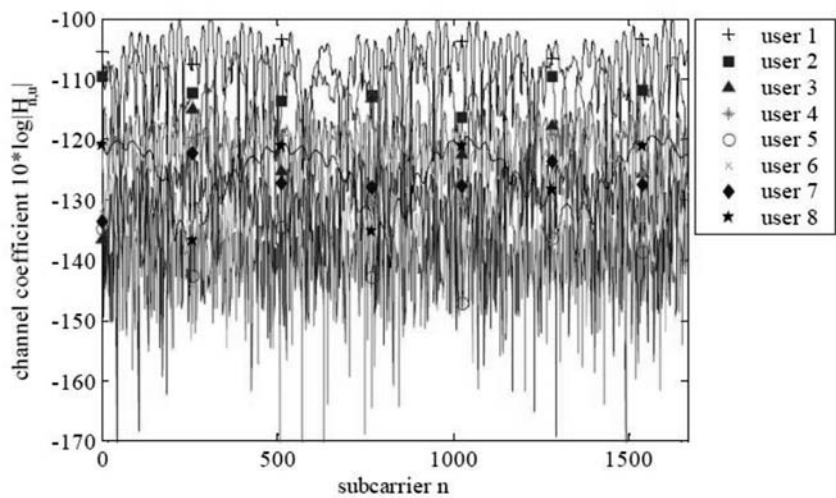


Figure 4.19 Channel transfer functions for all eight users [23].

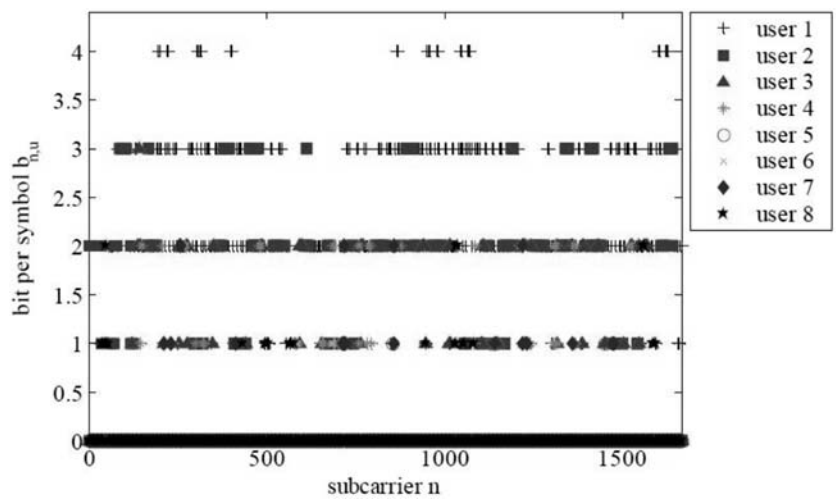


Figure 4.20 Subcarrier and bit allocation [23].

Table 4.7 Bit and Power Allocations for an OFDMA Uplink by Means of the Multiuser Water-Filling Algorithm

Users	1	2	3	4	5	6	7	8
Assigned transmit power p_u (dB)	23	23	23	23	21.9	22.8	23	23
Coded bits per OFDM symbol b_u	2,110	980	228	138	4	18	36	74
Information bit rate (Mbps)	46.9.8	21.8	5.1	3.1	0.088	0.4	0.8	1.6

bit rate and transmitting power and the overall transmitting power was minimized. If a two-step algorithm is applied (see [23] for a description), the bit rate for user 5 can maximally correspond to 2 bits/OFDM symbol; otherwise, the transmitting power constraint would be violated. The minimum bit rates have been chosen as $\hat{b}_u = [20, 20, 20, 20, 2, 8, 20, 20]$, and the resulting subcarrier and bit allocation is shown in Figure 4.21.

The total throughput per base station is rather low for the chosen transmission parameters. For users at the border of the cell (which number is already unrealistically small), the situation worsens drastically. This can be mainly attributed to the high attenuation, the given transmitting power, and the receiver noise levels. A more mature channel model can improve the preceding simulation results.

Finally, a rudimentary performance assessment of an OFDM/TDMA TDD system as a special case of TDMA/OFDMA for a single-cell and a multicell deployment was given. For this scheme, the bin, or time slot, consists of all subcarriers and one or many OFDM symbol(s). The simulations assume a channel bandwidth of 100 MHz. Such a bandwidth is recommended by the ITU [9] to meet the data rate requirements of the next generation air interface even for MIMO systems employing multiple transmitting and receiving antennas.

The evaluation is performed for the following assumptions [23]: Access points are equipped with omnidirectional antennas and if multiple cells are deployed, APs (sites) are located in a regular, hexagonal pattern. A single antenna is used for transmission and reception in both the access points and the user terminals. Moreover, the network is assumed to be perfectly synchronized, that is, we assume that all nodes in the system, both access points and user terminals, share a common time and frequency reference. The multicell network is deployed with a frequency reuse of 1.

Users are static and located at random positions within the coverage area of the network. In the single-cell deployment, eight users are located in the cell while the load in the multicell deployment is, on average, eight users/cell. A user terminal

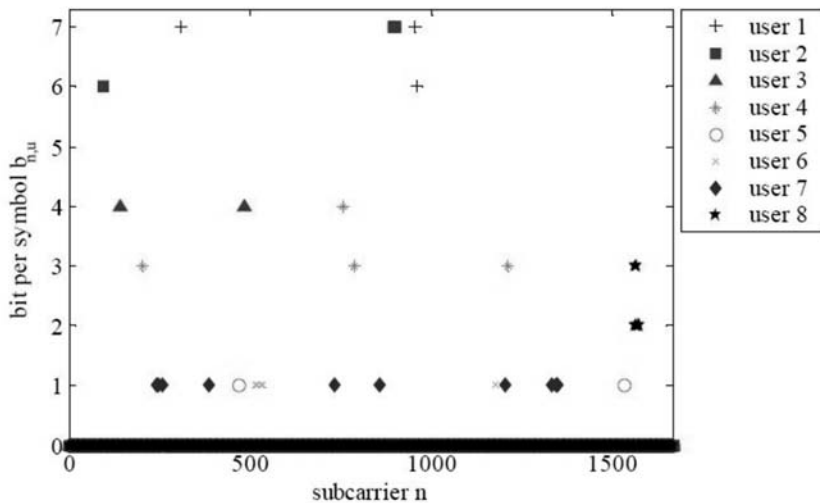


Figure 4.21 Subcarrier and bit allocation of two-step algorithm [23].

connects to the access point in the network to which it measures the lowest (long-term) path loss. OFDM transmission is employed in both the uplink and downlink. Moreover, in both links scheduling of users is performed periodically in a round-robin fashion. That is, within a cell pure time sharing of the channel is employed (TDMA) and the transmission scheduling is not channel dependent. When scheduled for transmission, a user has access to the whole bandwidth. UL and DL transmissions are separated by the time-division duplex scheme. The employed TDD TDMA frame structure is shown in Figure 4.22.

In Figure 4.22, there are 24 slots (OFDM symbols) in a frame and the frame length is 0.54 ms. The number of slots within a frame and the frame length are, however, not important for the outcome of this evaluation. As long as the number of UL and DL slots within a frame is equal, the performance is independent of the frame length. Accordingly, the frame configuration chosen here should be seen as an example rather than a proposed solution.

The transmission power is fixed and equally distributed over the subcarriers. Moreover, among the modulation and coding schemes (see Table 4.8) providing an estimated packet error rate below 10%, the scheme with the highest data rate is selected for transmission.

Figure 4.23 shows the performance of the different schemes in an AWGN channel. For a transmission, one codeword covers the complete data packet and a single modulation scheme is used over the whole band. The adaptive modulation and coding is based on ideal measurements, that is, it is assumed that measurement data is available at the transmitter side without errors or delays. Every data block is mapped to 1 OFDM time symbol over all 1,664 subcarriers.

The OFDM transmission is modeled as a multitude of parallel channels. The channel, both of the direct (useful) link and the interfering links, is in general frequency selective, but within a subcarrier the channel is considered as flat fading. To limit the computational complexity, the quality of the link is calculated only every 400 kHz, that is, every 16th subcarrier. Using the fact that the Vehicular A channel has an rms delay spread of 0.37 μ s and defining the coherence bandwidth as the frequency offset for which the correlation has decreased to 0.5 ($\rho = 0.5$),

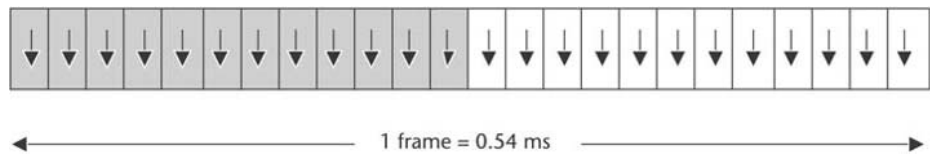


Figure 4.22 The studied OFDM/TDMA TDD frame structure [23].

Table 4.8 Selected Modulation and Coding Schemes (MCS) [23]

	<i>Modulation</i>	<i>Code Rate</i>	<i>Data Block Length (bits)</i>	<i>Data Rate (Mbps)</i>
MCS1	BPSK	1/2	824	36.6
MCS2	QPSK	1/2	1,656	73.6
MCS3	16QAM	1/2	3,320	147.6
MCS4	16QAM	5/6	5,542	246.3

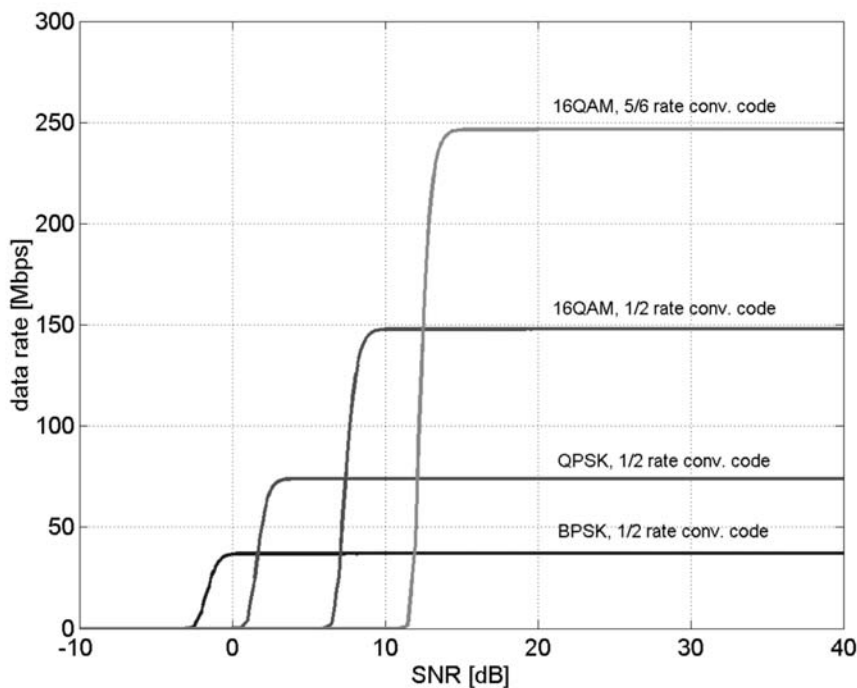


Figure 4.23 Achievable data rates for the different modulation and coding schemes in an AWGN channel.

the coherence bandwidth of the Vehicular A channel profile is at least 450 kHz:

$$b_{\text{COH}} \geq \frac{\arccos(\rho)}{2 \cdot \pi \cdot \tau_{\text{RMS}}}$$

(4.12)

The effects of ICI and ISI due to time dispersion are not modeled. Similarly, the effect of ICI caused by Doppler spread is ignored (users are static). Hence, within a single cell there is no interference, and in the multicell deployment there exists only ICI. The performance of the demodulator and decoder is modeled according to the OFDM link quality model described in [23]. The most important simulation parameters used for the single-cell and multicell deployment are summarized in Table 4.9.

Table 4.9 Simulation Parameters

	Single-Cell Deployment	Multicell Deployment
Number of cells	1	16
Load	8 users	8 users/cell (on average)
	0.54 ms (24 OFDM symbols)	0.54 ms (24 OFDM symbols)
Frame length		
Number of frames per snapshot	32	16
Number of simulated snapshots	50	25

The effective SINR (SINR_{eff}) of a packet comprising P symbols all modulated by the modulation scheme M is calculated as follows:

$$\text{SINR}_{\text{eff}} = I_M^{-1} \left(\frac{1}{P} \sum_{p=1}^P I_M \left(\frac{\text{SINR}_p}{\beta} \right) \right) \cdot \beta \quad (4.13)$$

where I_M is the BICM capacity for an AWGN channel using the modulation scheme M , I_M^{-1} is the inverse of I_M , and β is a model parameter whose value depends on the modulation and coding scheme used.

For each modulation and coding scheme, the model was trained and verified for many different channel realizations [23]. For each channel realization, link simulations were performed for several different noise powers and in this way the average block error rate is determined for some different SNRs. An operating region of a decade above and below a target BLER of 3% is considered and, therefore, only simulated points in the region of 0.3% to 30% are kept. In this way, for each channel realization $l = 1, \dots, L$, N_l simulated BLER values are used to train the adjustment factor β . More specifically, the adjustment factor is chosen to be

$$\hat{\beta} = \arg \min_{\beta} \sum_{l=1}^L \sum_{n=1}^{N_l} (p_{\text{pred},l,n}(\beta) - p_{\text{sim},l,n})^2 \quad (4.14)$$

where $p_{\text{pred},l,n}(\beta)$ is the predicted BLEP for a candidate value of the adjustment factor β and $p_{\text{sim},l,n}$ is the corresponding simulated BLER. Figure 4.24 show the validation of the link quality model for MCS1. In a similar way, the validations for the other modulation and coding schemes can be obtained. In the figure, crosses joined with lines are AWGN simulations and circles joined with lines represent fading independent from subcarrier to subcarrier. Other dots are results for realizations generated according to the 3GPP SCM Suburban Macro with six paths and Urban Macro extrapolated with 60 paths. In Figure 4.24, the BER versus the average (packet) SNR is plotted in the upper left plot and the upper right plot depicts the block error rate versus the effective SNR. In the lower plots, a scatter plot of the error versus the block error rate and a histogram of the model error are given to the left and the right, respectively. The error is calculated as the difference in average SNR between the AWGN link level simulation and the simulation of the trained link quality model for which the estimated block error probability (of the trained link quality model) coincides with the block error rate (of the AWGN link simulation). When using the trained values of β the model error is typically below 0.2 dB [23].

In a single-cell deployment, there is virtually no interference in an OFDM/TDMA system. Hence, the performance is simply determined by the channel in between the transmitter and the receiver, the transmitting power, and the noise. The left plot in Figure 4.25 shows the average user packet error rate. For most users, the packet error rate is relatively low. However, one must keep in mind that in the studied scenario, no measurement errors or delays are modeled in the link

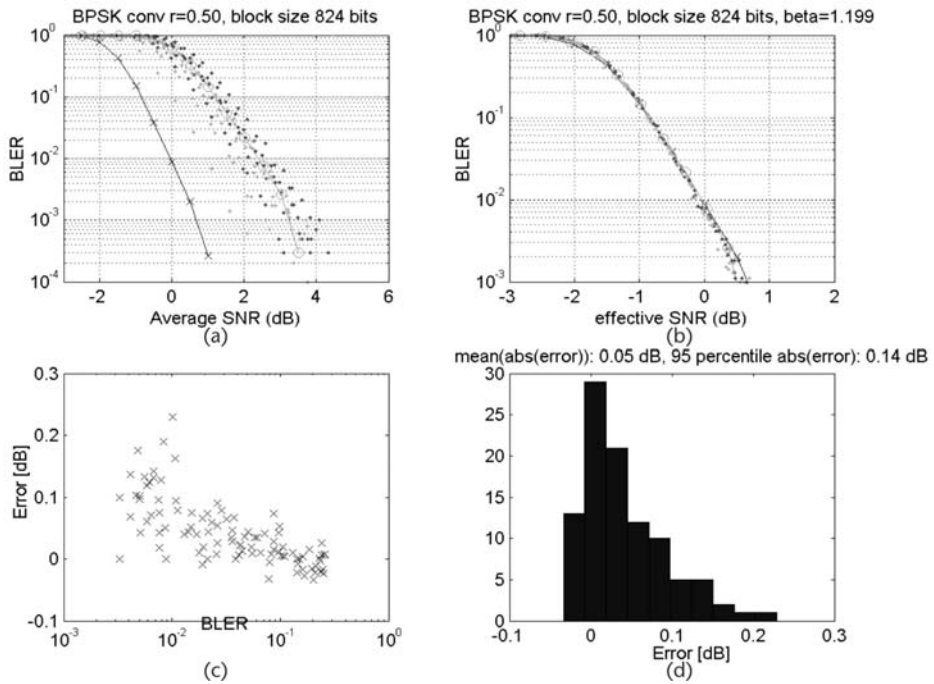


Figure 4.24 (a–d) Validation of link quality model for MCS1 (BPSK, 1/2 rate convolutional code).

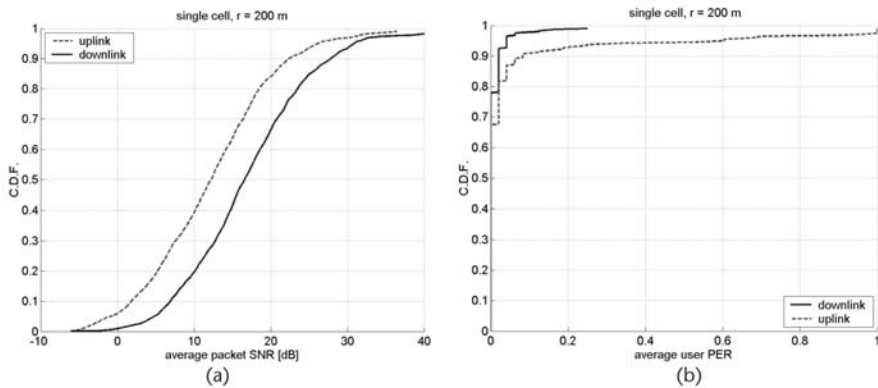


Figure 4.25 (a) Average packet SNR and (b) average user packet error rate in a single-cell deployment [23].

adaptation procedure. Such impairments would, if considered, increase the packet error rate in the system.

Figure 4.26 shows the average user data rate (on the physical layer) and the user throughput. The user data rate distribution indicates that there are quite a few users, roughly 30% in the downlink and 15% in the uplink, whose data rates are limited by the available modulation and coding schemes. Hence, by introducing higher order modulation or higher code rates than what is currently employed, the performance for these users would improve. Moreover, the fact that the DL

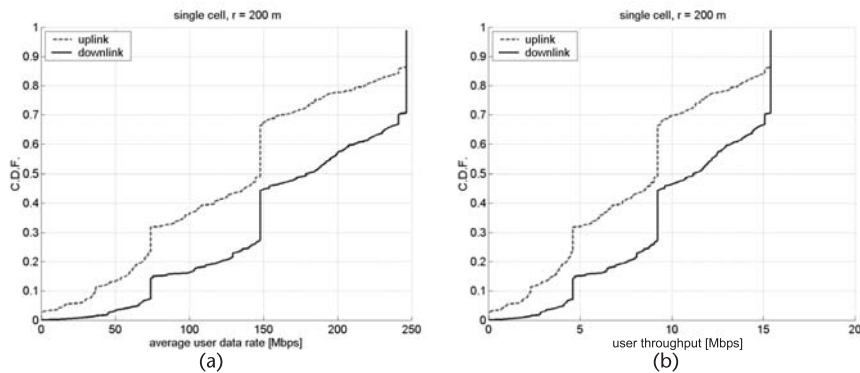


Figure 4.26 (a) Average user data rate and (b) user throughput in a single-cell deployment [23].

performance is slightly better compared to the UL performance can be explained by the difference in output power. The right plot in Figure 4.26 shows the user throughput, that is, the average user data rate during a session. Considering that there are eight users sharing the resource (in uplink and downlink) and that half of the time is used for UL and DL transmissions, the relation between the user data rate distribution and the user throughput distribution is, in this particular case, a factor of 16.

Figure 4.27 shows the average user data rates (on the physical layer) for the same cell sizes. In this case, the 10th, 50th, and 90th percentiles of the average user data rates are depicted. Note that the performance decreases rapidly with the cell size. For example, at a cell radius of 800m, the total throughput is less than 1/3 of the total throughput achieved when the cell radius is 200m.

In a network comprising multiple cells, ICI affects the performance of the OFDM/TDMA system. In the situation considered here, in which the system virtually is fully loaded all of the time, the DL interference experienced by the user terminals will be relatively static, because it only varies according to the movement of the user terminal itself. In the uplink, however, the interference can typically change in between different transmission attempts because the interference experi-

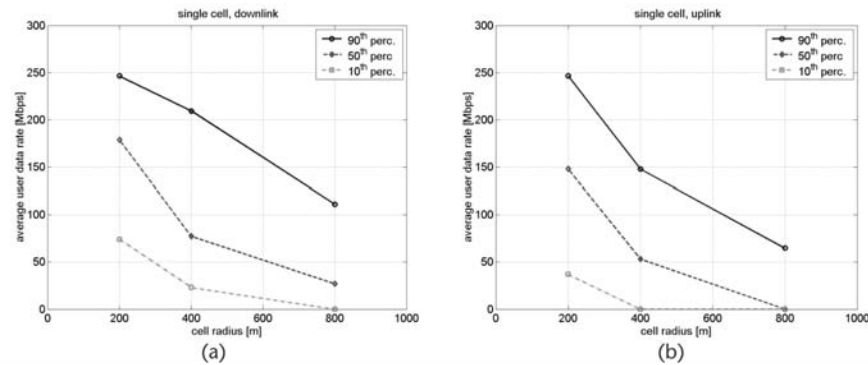


Figure 4.27 User data rates in (a) downlink and (b) uplink [23].

enced at the AP receiver depends on the positions of the user terminals in neighboring cells that are currently transmitting.

The left side of Figure 4.28 shows the average packet SINR for the case when the cell radius is 200m. The average user packet error rate (PER) is shown to the right in the same figure. Compared to the corresponding single-cell case we notice that, as expected, the interference from other cells degrades the performance of the system. For example, the fraction of packets received with an SINR above 20 dB is now much smaller compared to the single-cell deployment and, likewise, the fraction of packets received with a SINR below 0 dB has increased significantly. Furthermore, in this particular case (200m cell radius), the UL and DL performances are fairly balanced compared to the single-cell case in which the downlink outperformed the uplink due to the higher available transmitting power. In a multicell deployment, however, it is not only the transmitting power that determines performance, but also the interference conditions.

Figure 4.29 shows the 10th, 50th, and 90th percentiles of the user data rates (on the physical layer) in the downlink and uplink. Note that the users close to the cell border (i.e., the ones experiencing a high path loss to the serving access point and high DL interference from adjacent cells) are not supported.

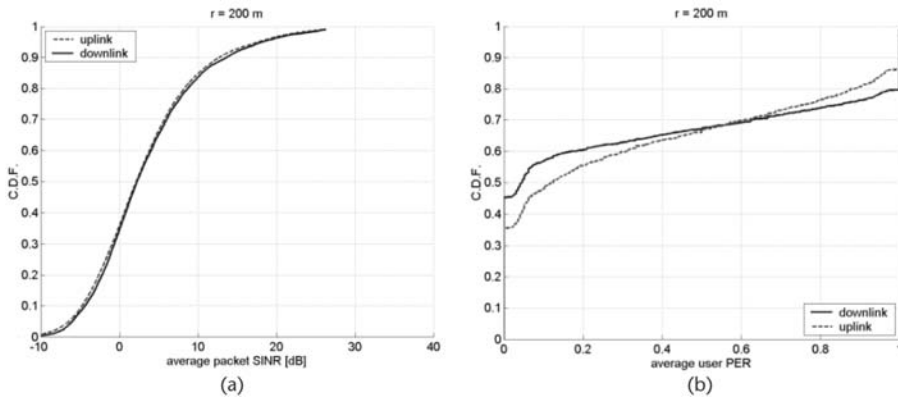


Figure 4.28 (a) Average packet SINR and (b) average user PER in a multicell deployment [23].

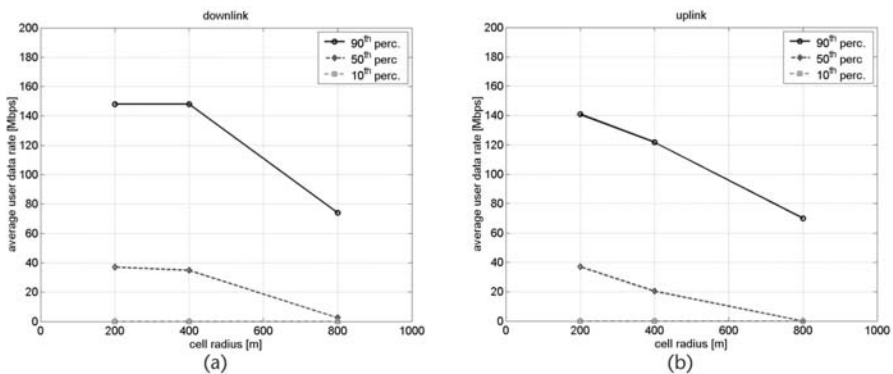


Figure 4.29 User data rates in (a) downlink and (b) uplink for multicell deployment [23].

The results show that an OFDM/TDMA TDD system has the potential to provide high data rates to network subscribers. However, some problems and limitations of the studied system can also be identified. Perhaps the most obvious limitation is the restricted coverage, especially for multicell deployments. To serve also users far from the access point, options that can provide more processing gain than what is currently available would be needed. This might be achieved, for example, by using channel codes with lower code rates, spreading or hybrid ARQ with soft combining, or incremental redundancy. Moreover, a higher transmitting power would improve performance in coverage-limited scenarios, but it is currently unclear what the output power constraints will be for this type of network. The use of multiple antennas is also a viable option for improving performance. For example, in the uplink, receiver diversity at the access point is a straightforward way to improve performance, and perhaps such a solution is also possible in the downlink for some terminal types. Multiple antennas at the access point used for beamforming is also an attractive option. Likewise, deploying the system with a reuse larger than 1 is a possible way to improve coverage in a multicell network. The simple fact that the network operation point typically will be (significantly) below what is studied here, namely, a fully loaded system, is also something that should be considered when performance is analyzed.

Within the employed OFDM/TDMA framework, it would further be possible to improve performance by using an adaptive modulation and coding procedure that adjusts transmissions according to the frequency selectivity of the channel. Furthermore, channel-dependent scheduling could potentially be used as well.

4.3 Conclusions

The allocation of transmission resources to multiple users for transmission of packet data is a topic closely related to scheduling. A wide variety of values for the access point throughput and the spectral efficiency could be obtained by varying the algorithm that controls the scheduling to different users for a given choice of multiple-access technology. It is far from clear that the problem of investigating multiple access can, or should, be studied separately from the problem of selecting appropriate scheduling and MAC algorithms.

Furthermore, it is hard to separate a multiple-access scheme from the numerous choices that must be made in the design of the underlying link: synchronization, channel estimation, coding, and receiver algorithms. What is really compared when comparing schemes are the very specific combinations of the multiple-access schemes and physical layer implementations under a set of predefined requirements.

References

- [1] Baier, P. W., P. Jung, and A. Klein, "Taking Challenge of Multiple Access for Third-Generation Cellular Mobile Radio Systems—A European View," *IEEE Communications Magazine*, February 1996, pp. 82–89.
- [2] FP6 Overview, <http://cordis.europa.eu/ist/so/mobile-wireless/home.html>.

- [3] FP6 IST Project Wireless Interface New Radio, WINNER and WINNER II, www.ist-winner.org.
- [4] FP6 IST Project My Personal Adaptive Global Net (MAGNET) and MAGNET Beyond, www.ist-magnet.org.
- [5] FP6 IST Project MASCOT, www.ist-mascot.org/deliverables.
- [6] FP6 IST Project SURFACE, www.ist-surface.org/deliverables.htm.
- [7] FP 6 IST Project 4MORE. www.ist-winner.org.
- [8] FP6 IST Project Flexible Relay Wireless OFDM-Based Networks (FIREWORKS), www.fireworks.intranet.gr.
- [9] "Framework and Overall Objectives of the Future Development of IMT 2000 and Systems Beyond IMT 2000," Recommendation ITU-R M.1645, www.itu.int.
- [10] International Telecommunications Union, www.itu.int.
- [11] "Broadband Radio Access Networks (BRAN); HIPERLAN Type2; Requirements and Architectures for Interworking Between HIPERLAN/2 and 3rd Generation Cellular Systems," ETSI TR 101 957, V1.1.1 (2001-08).
- [12] WiMAX Forum, www.wimaxforum.org.
- [13] IEEE 802.16e, www.ieee802.org/16/tge.
- [14] Long Term Evolution, www.3gpp.org/Highlights/LTE/LTE.htm.
- [15] Third Generation Partnership Project, www.3gpp.org.
- [16] Prasad, R., *CDMA for Mobile Communications*, Norwood, MA: Artech House, 1996.
- [17] Prasad, R., W. Mohr, and W. Konhauser, *Third Generation Mobile Communications*, Norwood, MA: Artech House, 2000.
- [18] Rappaport, T. S., *Wireless Communications: Principles and Practice*, Upper Saddle River, NJ: Prentice Hall, 1996.
- [19] Proakis, J. G., *Digital Communications*, New York: McGraw-Hill, 1995.
- [20] Sklar, B., *Digital Communications: Fundamentals and Applications*, Englewood Cliffs, NJ: Prentice Hall, 1988.
- [21] Wesolowski, K., *Mobile Communication Systems*, New York: John Wiley & Sons, 2002.
- [22] Pahlavan, K., and A. H. Levesque, *Wireless Information Networks*, New York: John Wiley & Sons, 1995.
- [23] FP6 IST Project WINNER, "D2.6 Assessment of Multiple-access Technologies," Deliverable 2.6, October 2004, www.ist-winner.org.
- [24] Lee, E. A., and D. G. Messerschmitt, *Digital Communication*, Boston: Kluwer Academic Publishers, 1988.
- [25] Medard, M., "The Effect upon Channel Capacity in Wireless Communications of Perfect and Imperfect Knowledge of the Channel," *IEEE Trans. on Information Theory*, Vol. 46, No. 3, May 2000, pp. 933–946.
- [26] Nogueroles, R., et al., "Performance of a Random OFDMA System for Mobile Communications," *Proc. of 1998 International Zurich Seminar on Broadband Communications*, February 1998, pp. 37–43.
- [27] Robertson, P., and S. Kaiser, "The Effects of Doppler Spreads in OFDM(A) Mobile Radio Systems," *Proc. of IEEE VTC '99*, Amsterdam, The Netherlands, September 1999, pp. 329–333.
- [28] Wong, C. Y., et al., "Multiuser OFDM with Adaptive Subcarrier, Bit, and Power Allocation," *IEEE J. of Selected Areas of Communications*, Vol. 17, No. 10, October 1999, pp. 1747–1757.
- [29] Kivanc, D., G. Li, and H. Liu, "Computationally Efficient Bandwidth Allocation and Power Control for OFDMA," *IEEE Trans. on Communications*, Vol. 2, No. 6, November 2003, pp. 1150–1158.
- [30] Knopp, R., and P. A. Humbelt, "Multiple-Accessing over Frequency-Selective Channels," *Proc. of IEEE PIMRC*, Toronto, September 1995.

- [31] Wang, W., et al., "Impact of Multiuser Diversity and Channel Variability on Adaptive OFDM," *Proc. of IEEE VTC-Fall*, Orlando, FL, October 2003.
- [32] Sternad, M., et al., "Attaining Both Coverage and High Spectral Efficiency with Adaptive OFDM Downlinks," *Proc. of IEEE VTC 2003—Fall*, Orlando, FL, October 2003.
- [33] Farsakh, C., and J. A. Nossek, "Spatial Covariance Based Downlink Beamforming in an SDMA Mobile Radio System," *IEEE Trans. on Communications*, Vol. 46, November 1998, pp. 1497–1506.
- [34] Boche, H., and M. Schubert, "Optimal Multi-User Interference Balancing Using Transmit Beamforming," *International J. on Wireless Personal Communications*, Vol. 26, No. 4, September 2003, pp. 305–324.
- [35] Koutsopoulos, I., T. Ren, and L. Tassiulas, "The Impact of Space Division Multiplexing on Resource Allocation: A Unified Approach," *IEEE INFOCOM*, New York, 2003.
- [36] Bartolomé, D., and A. I. Pérez-Neira, "Performance Analysis of Scheduling and Admission Control for Multiuser Downlink SDMA," *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004)*, Montreal, May 17–21, 2004.
- [37] Spencer, Q. H., A. L. Swindlehurst, and M. Haardt, "Zero-Forcing Methods for Downlink Spatial Multiplexing in Multiuser MIMO Channels," *IEEE Trans. on Signal Processing*, Vol. 52, No. 2, February 2004, pp. 461–471.
- [38] Choi, L. U., and R. D. Murch, "A Transmit Preprocessing Technique for Multiuser MIMO Systems Using a Decomposition Approach," *IEEE Trans. on Wireless Communications*, Vol. 3, No. 1, January 2004, pp. 20–24.
- [39] Serbetli, S., and A. Yener, "Transceiver Optimization for Multiuser MIMO Systems," *IEEE Trans. on Signal Processing*, Vol. 52, No. 1, January 2004, pp. 214–226.
- [40] IST Project WINNER, "Feasibility of Multi-Bandwidth Transmissions," Deliverable D2.2, October 2004, www.ist-winner.org.
- [41] Ekman, T., M. Sternad, and A. Ahlen, "Unbiased Power Prediction on Broadband Channels," *Proc. of IEEE VTC 2002-Fall*, Vancouver, Canada, September 2002.
- [42] Sternad, M., and D. Aronsson, "Channel Estimation and Prediction for Adaptive OFDMA/TDMA Uplinks, Based on Overlapping Pilots," *Proc. of the International Conference on Acoustics, Speech and Signal Processing (ICASSP 2005)*, Philadelphia, March 2005.
- [43] Ottosson, T., et al., "Towards 4G IP-Based Wireless Systems: A Proposal for the Uplink," *5th Wireless World Research Forum Meeting*, Tempe, AZ, March 2002.
- [44] Sternad, M., and D. Aronsson, "Channel Estimation and Prediction for Adaptive OFDM Downlinks," *Proc. of IEEE VTC-Fall*, Orlando, FL, October 2003.
- [45] Falahati, S., et al., "Adaptive Modulation Systems for Predicted Wireless Channels," *IEEE Trans. on Communications*, Vol. 52, February 2004, pp. 307–316.

Smart Antennas and Related Technologies

The performance of a next generation wireless communication system depends on the interaction of various components, including antennas, channel estimation, modulation schemes, error-protection coding, security algorithms, and dedicated hardware. The antenna is part of the physical side of the challenge to provide very high data rates through a radio interface. Antennas must be considered with respect to the other components such as circuitry, coding, modulation, and so forth.

Smart antennas and enabling technologies are important for improving device and network performance. Results can be measured in improved coverage and capacity gains. Smart antennas use a combination of an array of multiple antennas and appropriate signal processing to produce desirable antenna patterns. MIMO methods are a subset of smart-antenna techniques that exploit multipath propagation by taking advantage of random channel fading and multipath delay spread and combining the spatial dimension brought by the use of multiple antennas with the time dimension in a way that exploits the diversity present to improve the quality of the link or, on the other hand, to improve the spectral efficiency by creating parallel and independent data lines in the system (i.e., multiplexing gain).

This chapter presents the results in the area of advanced antenna concepts and related technologies achieved by the FP6 IST projects [1], which investigated and proposed multiple-antenna technology for IMT-A candidate systems (e.g., the IST WINNER and WINNER II projects [2]), in the OFDM WLAN systems (e.g., the IST project OBAN [3]), for short-range high- and low-data-rate communications, for exploring the potential of UWB and MC-SS radio technologies for personal area communications (e.g., the IST projects MAGNET and MAGNET Beyond [4] and the IST project PULSERS [5]), reconfigurable antenna structures (e.g., in the IST project MEMBRANE [6]), and others.

This chapter is organized as follows. Section 5.1 introduces into the subject and provides an overview of multiple-antenna techniques, models, and concepts. Section 5.2 describes the methodology for employing suitable antenna techniques for the improvement of the DL throughput for WLAN systems. Section 5.3 describes the design and implementation concerns of antenna arrays for next generation radio interfaces that provide ubiquitous coverage. Section 5.4 describes the design and implementation considerations for antennas used for UWB systems, and Section 5.5 concludes the chapter.

5.1 Introduction

The use of several antennas at the transmitter, at the receiver, or on both sides is known to significantly increase the theoretically achievable capacity of a wireless system [7]. Multiple transmitting and receiving antennas provide spatial diversity to improve range and spatial multiplexing for higher data rates as well as interference cancellation ability [8–14]. By including the spatial dimension, that is, the use of multiple antennas at both the transmitting and receiving sides, MIMO systems can effectively turn multipath propagation, generally considered as a drawback in wireless communications, into an advantage and linearly increase the capacity and, thus, the throughput and/or improve the quality of transmission in terms of BER.

Major requirements for next generation (e.g., IMT-A) multiple-antenna concepts include high performance, robustness, adaptability to a wide range of scenarios and terminal classes, and efficient interworking with the basic transmission and network concepts defined for the system [15].

To identify a versatile multiple-antenna concept, different spatial processing techniques need to be investigated and compared with respect to their performance in different usage scenarios, and their compatibility to form a component of the overall multiple-antenna concept [15]. Due to the manifold world of spatial processing techniques with different underlying ideas, objectives, and use cases, comparing and assessing the techniques with respect to performance, complexity, scalability, application scenarios, or required signaling and network functionality can be a difficult task, especially if no common simulation environment is available. The available channel knowledge at the transmitter strongly depends on the user terminal velocities, which means that it would vary in different scenarios. The spatial channel properties are scenario dependent and, therefore, the most promising spatial processing techniques for wide-area communications and short-range scenarios would, in general, be different, even if the same transmission technology is employed [2]. Therefore, the antenna structures for systems of ubiquitous coverage and those for short-range communication systems display differences forthcoming from the initial scenario requirements.

In a next generation communication system, ideally, a single overall concept should be able to adapt to all envisioned scenarios via an adaptive use of channel knowledge [15].

5.1.1 Multiple-Antenna Methods and Techniques

A multiple-antenna method is the pure algorithm that fosters the specific benefits of spatial processing (e.g., gains of beamforming, spatial multiplexing, spatial diversity) [15]. A multiple-antenna technique is a multiple-antenna method integrated in a radio access network (RAN). This means that required measurements, signaling procedures, protocols, and supporting radio interface functions must be provided in addition to the pure spatial processing algorithm when designing a multiple-antenna concept.

Different multiple-antenna methods exploit different properties of the radio channel and have different performance objectives (e.g., optimization of cell throughput, peak data rate, delay). For example, the ubiquitous WINNER radio

access concept [2] and, therefore, also the WINNER multiple-antenna concept must be able to operate under a wide range of radio access conditions, characterized by different deployments (e.g., cell radius, AP antenna configuration) and usage scenarios (with respect to number of users, terminal capabilities, user velocity, service type). Because the properties of the channels are expected to be different, the WINNER multiple-antenna concept consists of a set of multiple-antenna techniques. In the case of such an air interface, it is important to identify a set of multiple-antenna methods that provides flexibility and high performance in many different conditions, while minimizing the overall required enablers, complexity, and overhead [15].

A complete multiple-antenna concept should provide the means to ensure the following:

- An adequate feedback of measurements from the receivers, in addition to the measurements done on a return link to facilitate scheduling, link adaptation, and the selection of transmit parameters that control the spatial processing. This channel knowledge makes it possible for the transmitter to assess the suitability of different resource allocations in terms of users, streams, and spatial processing techniques in a multiuser setting. Examples from WCDMA include feedback bits for closed-loop transmit diversity and channel quality measurements for fast scheduling and link adaptation.
- An adequate signaling process, procedures, and architecture to ensure that the transmitter can adapt the transmission and resource allocation according to the current conditions. Thus, multiplexing of dedicated and common channels, the control of phase references for demodulation, the procedures for mobility management, and the mechanisms for (re)configuring the transmission format of the radio links should not prevent the use of efficient spatial processing.

For short-range technologies, the antenna design must cope with a set of different challenges. Some of these involve a solution to the problem of how to efficiently exchange large quantities of data at such a short range. For example, UWB antennas must have good performance over very large bandwidths, but such antennas must be designed as small and low-cost products [16]. The targeted applications for such antennas extend from low-data-rate (LDR) systems, such as sensor networks, to high-data-rate (HDR) systems, such as BAN, PAN, home entertainment, or WLAN systems. Small- to medium-size antennas and arrays would be of concern to a system for point-to-point communications.

Figure 5.1 shows the two antenna elements of commercially available folded monopole antennas used in the IST project MAGNET for UWB PAN and BAN initial channel measurements and investigations [4]. Figure 5.2 shows custom-designed UWB planar printed monopole-type antennas employed in the initial UWB PAN measurements.

5.1.2 Benefits of Multiple Transmission Schemes

The major benefits to be obtained by applying multiple-antenna transmission schemes are summarized as follows [19]:

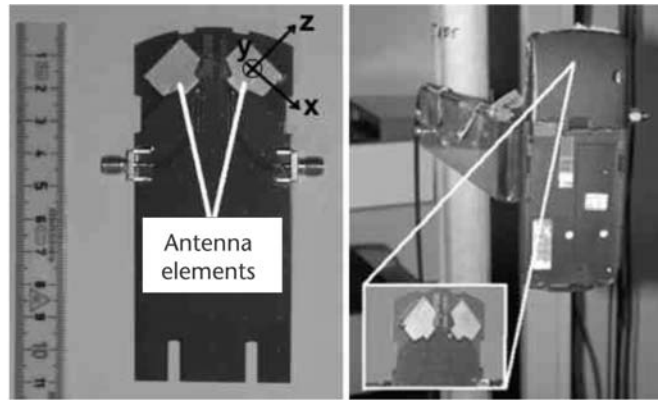


Figure 5.1 A handheld device under test with two folded monopole antenna elements used in the PAN-MD (PAN-mobile device) and BAN MAGNET scenarios [17].

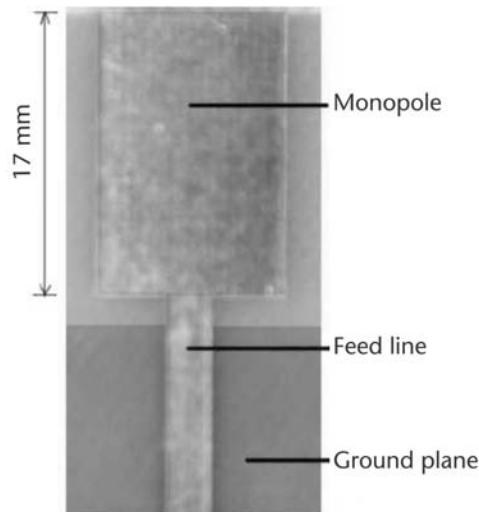


Figure 5.2 A planar printed monopole antenna element used in the MAGNET UWB-PAN measurements [18].

- *Array gain:* This is the average increase in SNR at the receiver due to the coherent combining effect of using multiple antennas at the Tx and/or Rx side. The resulting enhancement of signal power leads to increased range coverage and quality of reception. To achieve an array gain, MIMO systems typically require perfect channel knowledge either at the transmitter or the receiver, or at both.
- *Diversity gain:* By introducing redundancy to the system in the form of additional antennas, various gains from one or several of the temporal, frequency, or spatial dimensions added might be exploitable. The extent to which this gain enables the system to achieve higher capacity or reliability of reception is typically characterized by the number of independent fading paths in a channel (also referred to as *diversity order*). In the case of *temporal*

diversity, the transmitted symbols are repeated across time by a combination of channel coding and interleaving schemes (e.g., STTC/STBC) [10, 20]. To be able to extract this gain, the channel must provide sufficient variations in time; that is, the channel coherence time should be smaller than the duration of the coded interleaved symbol, which will essentially turn into an independent replica of its original version. In a short-range very HDR (VHDR) application scenario (i.e., very high audio/video transmissions over very short distances and pure LOS environments), such as the scenario investigated in the IST project PULSERS [5], however, the expected diversity provided by such schemes might be only marginal. Frequency diversity, on the other hand, may provide replicas of the transmitted signal in the frequency domain, given that the coherence bandwidth of the channel is smaller than that of the signal. Because in this case different parts of the assigned spectrum will suffer independent fades, the probability that all signal copies will be in a deep fade simultaneously is rather small. Although achieving frequency diversity for OFDM-based UWB systems appears to be quite reasonable [19], little focus on these is required provided that the VHDR system employs no frequency hopping and thus only intrasymbol space-frequency coding becomes feasible.

- *Spatial multiplexing gain:* With spatial multiplexing, multiple antennas at both the transmitting and receiving side are used to send independent data streams in parallel. In indoor environments with rich scattering, this technique allows for the creation of independent spatial data pipes within the same frequency band, which significantly increases the data rate and spectral efficiency. The channel capacity increases linearly with the minimum number of transmitting/receiving antennas for the same signal bandwidth and at no additional power consumption [21–23].
- *Interference suppression gain:* This is achieved by setting the antenna weights based on some sort of knowledge about the interference situation. In this way, it is possible to shape the effective antenna beam and to transmit and/or receive in different directions. The result is that the interference spread to and from users in other directions is reduced; that is, a spatial interference suppression, or interference avoidance, is realized [15].

5.1.3 Summary of Multiple-Antenna Methods

5.1.3.1 Beamforming

Beamforming techniques use multiple antennas to focus beams in certain spatial directions to leverage array and interference rejection gains. In the single-link case, the transmitter and/or receiver adapts its antenna weights according to the amount of channel knowledge and can thus achieve array gain. In the multiuser case, beamforming enables spatial selectivity by the allocation of different antenna weights to different users, thereby allowing for SDMA and achieving interference rejection and multiuser diversity gains [15].

5.1.3.2 Diversity Techniques

Diversity techniques aim at increasing the link reliability. This increase often comes at the expense of transmission rate.

5.1.3.2.1 Receive Diversity.

A receiver equipped with multiple antennas receives multiple copies of the transmitted signal (or signals) that it can combine to increase reliability. This combination forms the basis of *receive diversity* and it is achieved by summing up the weighting output of each receiving antenna. The optimum weights are chosen to maximize the output SINR, and they depend on the instantaneous fading state when the receiver knows this state, or on the receiver correlation matrix when it only has statistical channel information. In particular, maximum ratio combining (MRC) is optimal for spatially white noise plus interference, and MMSE combining for the spatially colored case [15, 24, 25].

5.1.3.2.2 Transmit Diversity.

A transmitter equipped with multiple antennas can achieve transmit diversity by spreading the transmitted symbols over time and space (space-time coding). The classical theory of space-time coding as developed in [26] specifies the so-called rank and determinant criteria to construct space-time codes: The amount of transmit diversity is equal to the smallest rank of codeword-difference matrices, and the coding gain is related to the product of the nonzero singular values of these matrices. The specific design and implementation of the spreading depends on the level of channel knowledge at the transmitter and receiver.

High channel correlation can significantly degrade performance [27]. Linear precoding techniques [28] assume knowledge of the (slowly varying) antenna correlations at the transmitter and can, thus, force transmission on the nonzero eigenmodes of the transmitting antenna correlation matrix [29]. Such a technique is shown in Figure 5.3.

The linear transformation F is determined to optimize a given criterion, for example, to minimize an upper bound on the pairwise error probability [27, 30], which focuses specifically on correlation-based CSI, or maximizes the expected SINR at the receiver. The optimal design in the latter case is to transmit in the direction of the strongest correlation matrix eigenvector, that is, an invariant beam-former pointing to one direction.

Further performance improvements can be leveraged if, in addition to the correlation-based CSI, the instantaneous and quantized CSI is fed back from the receiver. This is shown in Figure 5.4.

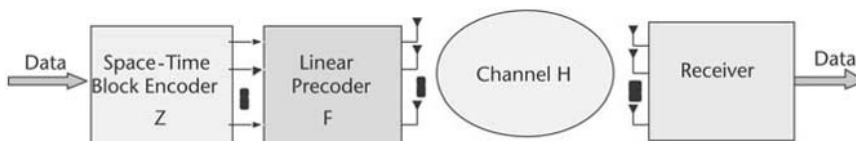


Figure 5.3 Reconfigurable transmission scheme combining space-time block codes and a linear transformation designed with respect to the channel knowledge available at the transmitter [15].

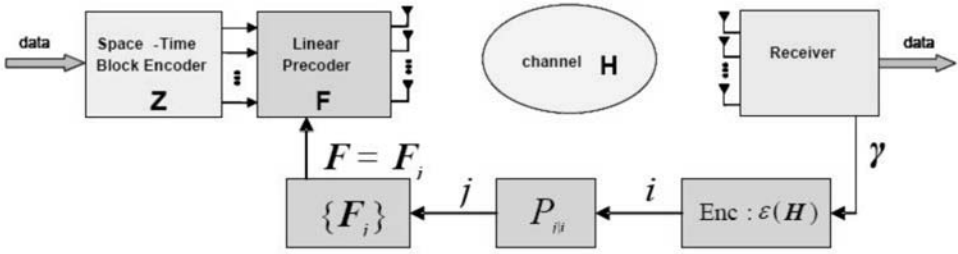


Figure 5.4 Weighted space-time coding where quantized CSI from a feedback link is used to choose a suitable precoder matrix [15].

A joint method for designing the quantizer and the precoding matrices that minimize a certain upper bound on the symbol error probability can be derived for orthogonal and quasiorthogonal STBCs [30]. The computation of the optimal precoder is not trackable in real time anymore, but it can be easily done offline because the CSI is quantized.

In certain scenarios (e.g., high mobility) channel parameters often cannot be acquired even at the receiver [15]. This is especially true for frequency-selective channels, where the number of channel parameters to estimate can be very large. For such cases, space-frequency code (SFC) designs for the noncoherent MIMO-OFDM channel can be based on a fully noncoherent approach [31] or a differential coding approach [32]. The case for noncoherent SFCs specifically tailored to frequency-selective channels is demonstrated by the existence of codes designed for noncoherent frequency-flat channels, whose performance breaks down in a frequency-selective environment (this is in stark contrast to the coherent case).

5.1.3.2.3 Spatial Multiplexing.

The simultaneous use of multiple transmitting and receiving antennas shows, at least theoretically, a large increase in data rates and, hence, spectral efficiency (see Chapter 2 of this book). This increase is achieved by spatial-division multiplexing (SMUX) techniques that transmit simultaneously independent data streams from different antennas. This architecture is shown for four transmitting and MR receiving antennas in Figure 5.5. It effectively creates spatial data pipes between the transmitter and receiver.

The input data streams are spatially multiplexed and, under favorable channel conditions, the spatial signatures of these transmitted streams are well separated. The receiver, having knowledge of the channel, can differentiate between these cochannel streams and extract each. The condition for spatial signature separation is that the channel matrix should have full rank, in which case the increase in transmission rate is linear in the number of transmitting antenna pairs $Q = \min(MT, MR)$. A rank-deficient channel matrix (e.g., such as would arise in a limited scattering environment) reduces the effective number of data pipes Q [15].

At the receiver, several approaches can be used to perform detection with different levels of complexity and thus error performance. When the transmitter does not have channel state information, it distributes rate and power equally among the transmitting antennas. This architecture is sufficient to provide full spatial multiplexing gain.

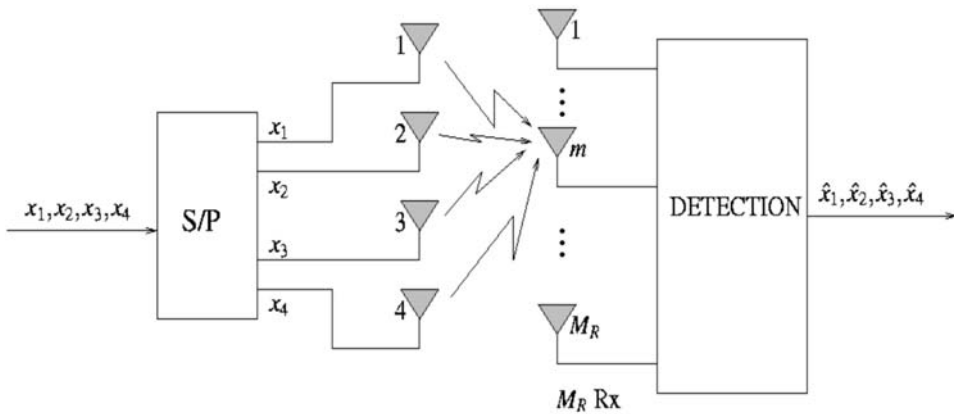


Figure 5.5 Spatial division multiplexing scheme [15].

With full knowledge of the channel (and of the noise plus interference covariance matrix), the transmitter may apply precoding to effectively create Q orthogonal subchannels. In addition, it can apply different coding and modulation schemes on each data stream [33]. This “bit loading is done to reflect the different capacities of each subchannel. Similarly, the available transmitting power may be distributed unequally among the subchannels in order to improve the throughput. Typically, this is done according to the subchannel singular values by using a water-filling (WF) algorithm [34]. Closed-loop SMUX does not increase the effective number of data pipes (because open-loop SMUX already achieves this number), but it can improve the throughput on these data pipes [15].

The optimal signaling strategy for closed-loop SMUX theoretically requires a Gaussian codebook with continuous rate allocation among eigenmodes, an approach that is not feasible in practice. Therefore, a near-optimal solution is required that takes into account the finite rate allocation granularity imposed by the limited number of modulation/coding schemes.

Closed-loop SMUX as described earlier requires instantaneous channel knowledge at the transmitter and usually complex bit/power loading algorithms. When only limited CSI is available, simpler schemes have been proposed to improve the basic throughput gains of open-loop spatial multiplexing. These include per-antenna rate control (PARC) and selective PARC.

Instead of antennas, beams also can be selected (without the receiver being aware of the use of antennas or beams). In the case of a single beam, this corresponds to a fixed beam approach; in the case of antenna set selection combined with vector modulation (BLAST), this corresponds to spatial multiplexing in the beam domain [15].

The diversity and SMUX techniques described earlier leverage competing MIMO gains that cannot be fully achieved simultaneously. Transmit diversity increases link reliability by introducing redundancy in multiple dimensions (space, time, frequency, and so on); therefore, it often comes at the price of sending a unique data stream (or reduced number of streams) over all of the available channels. SMUX techniques achieve high spectral efficiency by transmitting independent

symbol streams and, hence, do not directly provide reliability. Techniques that ensure a trade-off between these two types of MIMO gain have been reported [15].

A natural trade-off between diversity and multiplexing arises in the context of frequency-selective fading MIMO channels as described in [35]. For such channels, space-frequency coded OFDM is capable of realizing both spatial- and frequency-diversity gains. The presence of multipath propagation and the availability of frequency diversity generally allows for increasing the “spatial signaling rate,” that is, for using the multiple antennas to multiplex independent data streams rather than to provide spatial diversity. It is, therefore, desirable to be able to allocate the channel’s degrees of freedom in space and frequency in a flexible way in order to achieve a variable allocation to multiplexing and diversity transmission modes. A simple construction is proposed in [35] that consists of a linear outer code that guarantees a certain amount of frequency diversity, and an inner (independently designed) space-time code that realizes spatial diversity. The amount of SMUX is controlled by the inner code, whereas the outer code guarantees a prescribed diversity order. In general, the amount of spatial multiplexing available should be fed back from the receiver.

“Hard switching” between transmitter modes, also called *spatial domain link adaptation*, can provide for a trade-off between the two gains by exploiting the channel conditions fed back by the receiver.

5.1.3.2.4 Receiver Architecture for Coded MIMO.

The main classes of receiver architecture for MIMO techniques are as follows:

- Full-search (or joint) ML receivers;
- Sphere decoder as a low-complexity implementation of a (near) ML receiver;
- Linear zero forcing (ZF) or minimum mean square error (MMSE) receivers;
- (Ordered) successive interference cancellation (SIC)-type receivers (e.g., V-BLAST).

Joint ML detection yields the best performance at the cost of a high computational cost that can indeed get prohibitive in the MIMO context. Sphere decoding (via the Fincke-Pohst algorithm) may be a computationally attractive alternative with close to ML performance [36]. The idea is to estimate the transmitted symbols by searching only over those symbols that belong to a hyperdimensional sphere centered at the received signal. This algorithm has been adapted to decode lattice codes for the Gaussian and Rayleigh channels in [37], and extended to the MIMO case in [38]. A list sphere decoding algorithm is proposed in [39] to further reduce the complexity of the receiver.

Less expensive techniques are based on linear detection using the ZF or MMSE criteria to perform the detection of the transmitted signals. ZF linear processing nulls out completely unwanted signal components, disregarding the noise, and results in noise enhancement. MMSE processing takes into account noise components and is usually preferred in cases when significant unwanted signal components are to be suppressed. The performance of linear receivers depends on the dimensionality of the system, that is, on the number of transmitting and receiving antennas, and on the correlation characteristics of the received signals.

The performance of linear receivers can be improved by combining them with successive cancellation. The key idea is layer peeling in which the individual data streams are successively decoded and stripped away layer by layer. The algorithm starts by detecting an arbitrarily chosen data symbol (using ZF or MMSE in the front end) assuming that the other symbols are interference. After the detection of the chosen symbol, its contribution from the received vector is subtracted and the procedure is repeated until all symbols are detected. To minimize the effects of error propagation, it is advantageous to perform successive canceling from the “strongest” to the “weakest” signal as proposed for V-BLAST [40]. (Specifically, the rule is to detect streams in the decreasing order of their postprocessing SNR.)

A similar family of techniques suited to coded transmission uses prior knowledge provided by a SISO decoder of the used channel code. Such a “turbo” approach considers the transmission to be a serial concatenation of a channel code (outer code) and the channel (inner code), usually separated by an interleaver. The SISO processing of the MIMO channel must usually be suboptimal due to the dimensionality of the channel.

5.1.3.2.5 Antenna Hopping for Multiuser MISO/MIMO Downlink.

Beamforming with random beamforming weights in combination with channel-dependent scheduling is often referred to as *opportunistic beamforming* and has been proposed in [41]. A multiuser downlink hopping system has also been considered that operates in a manner similar to that of opportunistic beamforming, the difference being that hopping operates in the antenna rather than the beam domain [15]. The gain is achieved by multiuser scheduling while hopping over the transmission antennas. This operation is particularly attractive in the case of a large angular spread (i.e., rich scattering) when thinking of spatially separated antennas because beamforming is less beneficial. Another application is in the polarization domain. Applying multiuser antenna hopping over cross-polarized antennas can even be combined with traditional opportunistic beamforming.

5.1.3.2.6 Sum Capacity Optimization of a MIMO Multiple-Access Channel.

The sum capacity of a MIMO multiple access channel is given by

$$C(\mathbf{Q}, \mathbf{H}) = \log \det \left(\mathbf{I} + \text{SNR} \sum_{k=1}^K \mathbf{H}_k \mathbf{Q}_k \mathbf{H}_k^H \right) \quad (5.1)$$

where \mathbf{H}_k is the channel matrix and \mathbf{Q}_k the covariance matrix of the data of user k . Here, it has been assumed that the receiver knows the channel realizations perfectly. If in addition the transmitters have full CSI, the sum capacity over the covariance matrices \mathbf{Q}_k can be optimized given the following sum power constraint:

$$\sum_{k=1}^K p_k = \sum_{k=1}^K \text{trace}(\mathbf{Q}_k) \leq P \quad (5.2)$$

The sum power constraint is important if the power can be distributed across the users in one cell but the cell sum power is limited in order to control the ICI.

For small SNR, this algorithm yields only one active user, in which case signal processing at the receiver is easy and scheduling is very simple [15].

Interference from intercell or intracell users degrades the performance of a point-to-point MIMO link. If some information about this interference is known, it can be mitigated (or even avoided) often via optimized joint user transmission [42, 43]. When this is not possible, this degradation can be quantified in terms of capacity (in bits per channel use) by modeling the link as corrupted by a worst-case colored noise plus interference ($n + i$) term. The desired signal and the $n + i$ terms are modeled according to the capacity achieving distribution (i.e., Gaussian) with covariance matrices \mathbf{Q} and \mathbf{Z} , respectively. This choice of distribution for the $n + i$ term yields the worst-case interference, but additionally, it is necessary to solve it for the worst $n + i$ covariance matrix \mathbf{Z} . The spatial signature of the interfering users determines the matrix \mathbf{Z} whose eigenvectors are the directions of the interfering signals and whose eigenvalues are the average powers. We distinguish three cases: (1) fixed sum noise power, (2) fixed interference power, and (3) fixed per-receiving-antenna noise power.

For case 1 the vector MIMO channel with perfect CSI at the transmitter transforms into a MIMO channel without CSI and white Gaussian noise as shown in Figure 5.6.

In case 2, the worst case noise directions are the right eigenvectors of the channel matrix times an arbitrary permutation matrix. In other words, with perfect CSI the worst-case $n + i$ decomposes the MIMO link into a set of parallel channels as shown in Figure 5.7.

In case 3, the worst-case noise color reduces the achievable performance of a MIMO system with n cooperating transmitting antennas to n users who perform only power control, as shown in Figure 5.8.

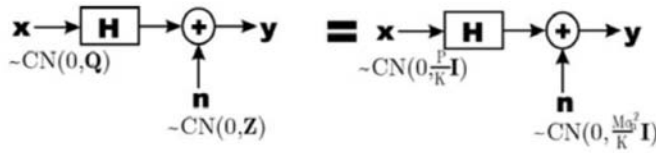


Figure 5.6 Vector MIMO channel capacity with worst-case noise with trace constraint equals capacity of open-loop MIMO system and white noise (no interference) [15].

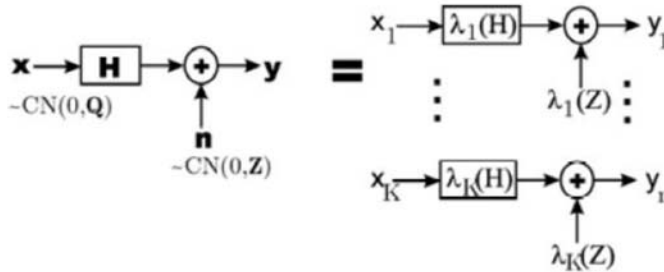


Figure 5.7 Worst-case noise directions: the vector MIMO channel capacity equals capacity of orthogonalized parallel channels with different noise powers [15].

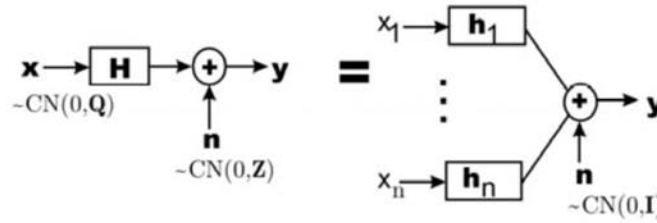


Figure 5.8 Worst-case colored noise: the vector MIMO channel capacity equals the sum of the capacity of the corresponding SIMO MAC [15].

5.1.3.3 Characterization of Multiple-Antenna Elements

The selection of a candidate multiple-antenna technique should be performed as a function of the scenario, because it critically determines the amount of CSI available and, therefore, the techniques that are applicable.

For example, the prevailing propagation conditions in the wide-area scenario are low angular spread at the BS, large angular spread at the mobile terminal (except for LOS conditions), and relatively high velocities. Therefore, it is more realistic to assume that only long-term CSI and/or highly quantized short-term CSI (e.g., single CQI) can be made available to the transmitter. This essentially precludes closed-loop techniques that require very accurate short-term CSI. It allows, however, for techniques that limit the instantaneous knowledge to a reduced set of quantities (e.g., PARC, grid-of-beams beamforming (BF), opportunistic BF, antenna selection), closed-loop techniques that rely only on long-term channel knowledge (e.g., long-term eigen beamforming, multiuser beamforming based on second order statistics for all users), open-loop techniques that do not require any channel knowledge apart from maybe a low-rate feedback to switch between transmission modes (e.g., open-loop T_x diversity, general open-loop matrix modulation, or antenna hopping), and noncoherent transmission. In addition, the proper combination of techniques that exploit long-term CSI with open-loop signal design (e.g., long-term eigen beamforming plus STBC, vector modulation, or matrix modulation across beams and/or polarizations) is interesting.

By contrast to the wide-area case, propagation conditions in short-range scenarios are characterized by an increased angular spread at the BS (therefore, beamforming can be less effective, but SMUX and/or diversity is applicable), increased temporal dispersion (which allows for frequency diversity), and reduced time variability of the channel. Schemes that require high-quality, short-term channel knowledge may be feasible in this case, whereas open-loop signals and noncoherent designs are less relevant. Promising techniques for hotspot scenarios include multiuser linear (e.g., ZF, block diagonalization) and nonlinear precoding (e.g., Tomlinson-Harashima precoding (THP), sphere encoding), linear/nonlinear closed-loop MIMO (e.g., SVD-based SMUX), open-loop vector/matrix modulation for above-pedestrian velocities, and general techniques discussed for wide-area coverage but with scheduling or link adaptations taking into account the spatial-temporal channel properties and Doppler spread typical for hot-spot scenarios.

In a packet-switched MIMO VHDR multiband orthogonal (MBO) system with random arrival times and, especially, very high data rates, acquisition and band-

tracking must be completed shortly after a packet has been received. To achieve this, it is common to precede the payload data with a carefully chosen sequence of known preamble and training symbols, in order to provide enough information to the receiver for correct synchronization, channel estimation, successive equalization, and frequency tracking [19]. To estimate the MIMO channel and ensure that the separate channels between each antenna pair can be uniquely identified, the preamble symbols to be transmitted over the different antennas should be orthogonal and shift-orthogonal [44]. Such considerations would also affect the design of the transmitter and receiver architectures.

5.1.4 Spatial Modes

For the assessment of different multiple-antenna techniques, it is important to define spatial modes, which represent the combination of enabling requirements and spatial processing gains leveraged [15]. Based on the concept of spatial modes, we can differentiate between *competing* techniques and *complementary* techniques. Ideally, *competing* techniques are those that allow implementing identical spatial modes, whereas *complementary* techniques implement different spatial modes. In reality, there is a gradual transition between these two ideal cases and almost any two techniques will contain complementary aspects; also, because almost all techniques differ in their detailed requirements regarding support functions and control information.

5.2 Downlink Capacity Enhancement of IEEE 802.11a/g

The problem of downlink throughput improvement for IEEE 802.11a/g systems was addressed by using a modified access point equipped with multiple antennas within the IST project OBAN [3]. The main restriction was that the standard terminals should not be modified in any way. An alternating time-offset SDMA solution was proposed for a conference room scenario. A simulation based on the channel models approved by the IEEE 802.11 Standards Group was performed that took into account the main features of the IEEE 802.11a/g standard. It was demonstrated that a near doubling of the downlink capacity can be achieved in a conference room environment. The results presented in this section are based on investigations carried out within the OBAN project. A general multiple-antenna WLAN structure is shown in Figure 5.9.

An access point is equipped with $N > 1$ antennas, and L terminals are equipped with single or multiple $M \geq 1$ antennas. When spatial processing is implemented, the RF channel between the transmitting and receiving antenna array elements is commonly referred to as a MIMO channel. When sufficient RF signal scattering exists, the MIMO channel can support communications on $N_s = \min(N, M)$ spatial channels. Techniques employing both full or partial CSI have been reported as possible techniques to use for such a scenario [10, 21, 46, 47]. Multiple-antenna systems with single-antenna terminals were chosen for the investigations in OBAN because multiple-antenna terminals were not available at that time.

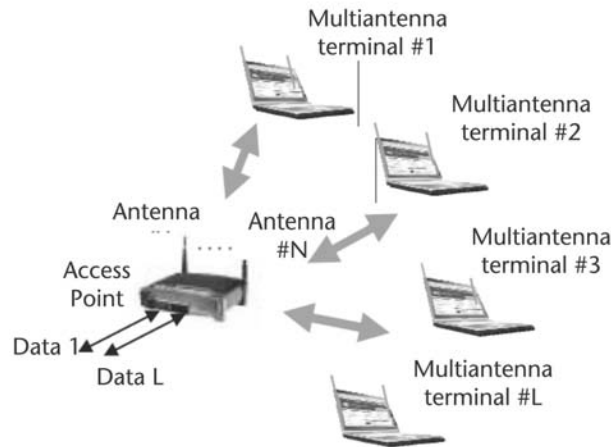


Figure 5.9 A general MIMO WLAN structure [45].

In this case, advanced multiple-antenna APs are required to serve legacy (e.g., IEEE 802.11) terminals. The main techniques applicable in such a case are as follows [45]:

- Downlink beamforming and STBC for range extension is a this special MISO case in which multiple AP antennas are used at the transmitter to increase the SNR at the receiver.
- Downlink SDMA is a special case of transmit beamforming for MIMO systems with noncooperative receivers. SDMA can simultaneously serve a number of users in the same frequency band leading to capacity improvements.
- Uplink interference cancellation at the AP allows range extension and capacity improvements under intrasystem and intersystem interference in a SIMO system. In [48, 49], this option was considered relevant to the OBAN scenario because of the coexistence problem.

Some of the multiple-antenna techniques are already implemented in commercially available products [50, 51] and some of them were under consideration for the IEEE 802.11n standard [8, 9], which was the advanced version of the current IEEE 802.11 standard.

At the time of the investigations, in 802.11a, selection diversity could be applied at the AP by selecting between two antennas, of which one experiences the best SNR. This very simple scheme improves the coverage but still relies on a single RF module and a switch for the antenna selection. With the addition of a complete RF stage per available antenna, more complex digital signal processing can be envisaged in order to fully exploit the multiple-antenna capabilities in terms of diversity or multiplexing gains. With the objective of increasing the peak bit rate, spatial diversity can be used to increase the robustness of the transmission of higher order modulations that are costly in terms of RF requirements, whereas spatial multiplexing enables simultaneous transmission of several parallel data flows with

lower modulation order, but with additional interantenna interference. For instance, a two-transmitting-antenna wireless link with two or more receiving antennas can be designed to either maximize spatial diversity (e.g., via an STBC technique) or perform spatial multiplexing [45]. For a symmetric antenna configuration, however, when the number of transmitting and receiving antennas is equal to two, it appears that spatial multiplexing gain is very difficult to exploit, whereas STBC techniques combined with coherent receive processing can achieve significant improvements in BER performance with low computational complexity [48]. This gain can translate into a larger range and, therefore, larger throughput at a given distance. It can also be traded off against reduction of the transmitting power and, hence, an increase of battery life for portable devices, or in order to support higher order modulations to increase the bit rate.

For the case in which the number of transmitting antennas is greater than two and the number of receiving antennas is greater than one, transmit diversity and spatial multiplexing gains can be obtained simultaneously, but there is a fundamental trade-off in the level of transmit diversity and multiplexing gains that can be achieved in any transmission scheme. Thus, the performance of specific transmission and reception strategies for achieving both diversity and multiplexing gains is an interesting area of research that needs to be carefully considered.

OBAN focused on the 802.11a/g physical layer, which is OFDM based. Among the various advantages of OFDM transmissions over the wireless medium are the ability to counteract multipath (frequency-selective) fading and the possibility of performing efficient FFT-based equalization (the equalizer is just a scalar multiplier at each OFDM subcarrier in the frequency domain). 802.11a supports a “data payload communication capability” of 6, 9, 12, 18, 24, 36, 48, and 54 Mbps. These transmission rates are obtained by varying the modulation type or the channel encoding rates. For the study here, the term *data payload* and the actual data throughput is a function of both the transmission mode and the channel conditions. For example, the 6-Mbps mode is more robust to adverse channel conditions (such as noise, interference, and severe multipath) than the 24-Mbps mode [45]. Consequently, the 6-Mbps mode can be actually faster than the 24-Mbps mode because of potentially fewer lost packets. This relationship between the transmission rate at the PHY layer and PHY throughput has a fundamental impact on how the system should be optimized for data throughput maximization.

Because the 802.11 standard clearly dictates that control frames should be understood by all nodes, and in order not to raise any backward compatibility issues with any existing equipment, development was constrained [45] in the transmission-reception of data packets only. (Note that results can be extended without undue complications to control-management packets as well.) Here, it is assumed that once a connection has been established between mobile nodes (or between mobile nodes and the AP), then STBC may be used to boost network performance. For this, the communicating nodes need to have STBC capabilities: The sender node should have more than one transmitting antenna and an STBC encoder, and the receiver should be able to decode STBC signals.

Given the power/size limitations of mobile nodes, the nodes with multiple transmitting antennas would be the APs. Hence, by equipping the AP with more than one transmitting antenna and using STBC to transmit data packets, the quality

of the wireless link can be significantly improved. Furthermore, the small decoding complexity of STBC codes makes it possible to enhance the mobile nodes with STBC decoding capabilities as well. Such a scheme improves mainly the downlink performance.

It is envisioned that in many scenarios, the downlink will be the bottleneck (e.g., the streaming or download of a digital movie in a home entertainment system). On the other hand, as hardware costs decrease, it will become economically feasible to manufacture mobile nodes with STBC transmitters.

The basic rule for splitting a spatial stream across multiple antennas is to transmit two related streams on different antennas [8, 9]. It is known that the transmitted radio wave is composed of in-phase and quadrature components, where the quadrature wave is a quarter cycle out of phase with the in-phase component. Phase shifts are represented mathematically by the imaginary part of the complex number in the constellation. The complex conjugate of a complex number has the same real part, but flips the sign on the imaginary part.

Physically, the radio wave from the complex conjugate will have the same in-phase component, but the quadrature component will have the opposite/negated phase shift. When there are extra antennas, the WWiSE proposal [8] mandates that a spatial stream and its complex conjugate be transmitted on an antenna pair. The rules for splitting spatial streams are independent of the channel bandwidth, although 40-MHz spatial streams will carry more bits. The WWiSE proposal is essentially a MIMO evolution of the 802.11a PHY. It requires mandatory support for 2×1 and 2×2 transmission modes, where both the transmitter and receiver have at most two transceivers active. However, it is likely that most products that are based on the eventual 802.11n standard (which may or may not resemble either of the proposals) will support at least some of the optional modes, which are 3×2 , 3×3 , 4×2 , 4×3 , and 4×4 transmission modes. It is likely that cost constraints on client devices will restrict them to operating in a one- or two-transceiver mode, whereas APs will have more transceivers.

Basic APs might have only two, while the most expensive enterprise-class APs might have three or four transceivers. TGnSync's proposed peak rates are higher, although at the cost of more aggressive coding. WWiSE's 135-Mbps data rate is accomplished by using 64-QAM at $R = 5/6$; TGnSync gets to 140 Mbps by using a $7/8$ rate code and cutting the guard interval to half. (Without the short guard interval, TGnSync's data rate is only 126 Mbps.) The advanced beamforming mode uses the much larger 256-QAM constellation.

From the preceding discussion, it is clear that the most important transmission configurations are the mandatory options, that is, the 2×1 and 2×2 transmission modes. As a consequence, OBAN focused on these two configurations, addressing the improvements guaranteed by space-time codes on the PER performance with respect to point-to-point communications in different channel conditions.

5.2.1 Space and Time Diversity in the OBAN Scenario

This section describes the results of the PHY system simulations that validated the effectiveness of employing space and time diversity coding for the assumptions used in the OBAN scenario. The project designed STBC codes and compared their

performances with the ones proposed for IEEE 802.11a/g for various operational conditions. As a performance metric used in connection of performance evaluation used was the PER, which statistically measures the percentage of correct packets transmitted with the proposed codes. Code design for various scenarios was discussed in detail in Chapter 3 of this book.

Figure 5.10 shows the architecture of the proposed space-time bit-interleaved coded, OFDM-based WLAN transceiver for the 2Tx-1Rx case, which was one of the most important system configurations within the OBAN scenario, along with the 2Tx-2Rx case [45].

Information bits are encoded through a convolutional encoder whose polynomial generators are of the form shown here [45]:

$$G(D) = [1 + D^3 + D^4 + D^5 + D^6, 1 + D + D^3 + D^4 + D^6] \quad (5.3)$$

After puncturing, coded bits are first bit interleaved with a variable-length interleaver designed to break the symbol correlation introduced by WLAN fading channels, and then modulated with the modulation format chosen according to one of the WLAN transmission modes shown in Table 5.1.

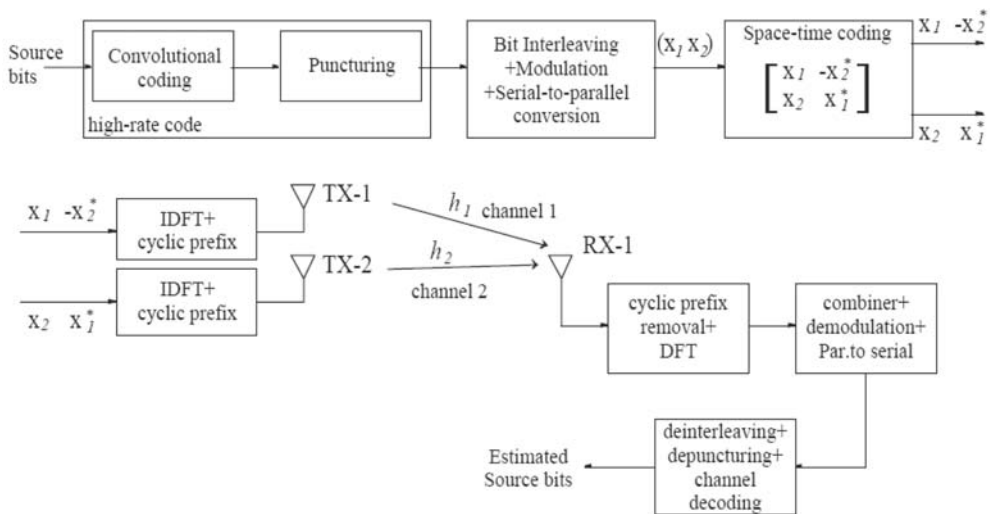


Figure 5.10 Block diagram of a 2Tx-1Rx space-time coded OFDM system [45].

Table 5.1 Parameters of the IEEE 802.11a/g Standards [45]

R_a -Mbit/s	Code Rate	Modulation	$N m_1$ Bits	$N m_1 R_a$ Bits
6.0	1/2	BPSK	48	24
9.0	1/2	BPSK	48	36
12.0	1/2	QPSK	96	48
18.0	1/2	QPSK	96	72
24.0	1/2	16-QAM	192	96
36.0	1/2	16-QAM	192	144
48.0	1/2	64-QAM	288	192
54.0	1/2	64-QAM	288	216

Modulated data is first converted in a parallel format, and then coded with an STB code employing a two-level degree of diversity. STB-coded data are then OFDM modulated and transmitted over the channel. The receiver accomplishes the inverse operations done from the transmitter. The next section addresses performance comparisons of the proposed architecture with respect to existing WLAN IEEE802.11 transceivers.

The results presented here are based on the WLAN channel models B, C, and D [45]. In particular, channels B and C are the core models within the OBAN scenario. Both WLAN channels represent an indoor environment where there is no line of sight and the multipath follows a Rayleigh distribution. RMS delay spreads are 15 ns for channel model B, 30 ns for channel model C, and 50 ns for channel model D. The maximum excess delay spread of these channels is less than 800 ns, which is the guard interval duration specified by IEEE 802.11a/g standards in order to protect data from ICI. Figure 5.11 shows the results for a 3/4 rate convolutional encoded QPSK with STB coding (2Tx-1Rx and 2Tx-2Rx) compared to the SISO case (1Tx-1Rx), and channel model D.

There is noticeable performance improvement as the number of antennas at the receiver increases. As a reference example, for a target PER of 10^{-2} , QPSK 2Tx-2Rx achieves a gain of 8 dB over the SISO transceiver. For the same target PER, 2Tx-1Rx achieves a roughly 3.5-dB gain over the SISO case. However, it can be observed that at a large PER (50%) the gain reduces to 0.2 and 3.6 dB, respectively, compared to the SISO case. The gain of STB-coded systems is, therefore, maximized for error-sensitive applications requiring low values of the PER.

5.2.2 Downlink Beamforming Under EIRP Constraints in WLAN OFDM Systems

By comparing the conventional maximum ratio (MR) combining solution under the total power (TP) and EIRP constraints in typical IEEE 802.11 propagation

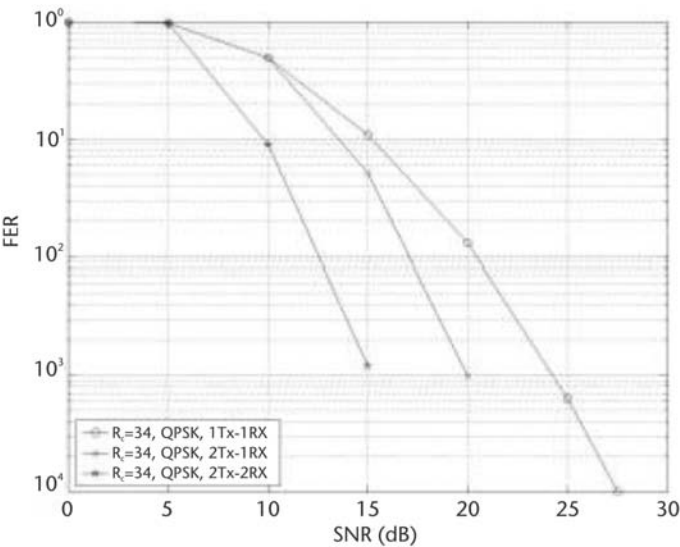


Figure 5.11 PER performance of STB-coded IEEE802.11a standard under channel model D [45].

conditions [35, 36], it was shown [45] that there is a significant performance gap compared to the TP restricted MR beamformer that can be filled with some advanced EIRP-based solutions.

Figure 5.12 shows a system model consisting of an AP with N transmitting antennas and a single antenna terminal. A transmitted signal $\mathbf{x}(f)$ in the antenna of N elements is given by

$$\mathbf{x}(f) = \mathbf{w}(f)s(f) \quad (5.4)$$

where $\mathbf{w}(f)$ is the $N \times 1$ weight vector, $s(f)$ is the transmitted symbol, $f = 1, K, F$ is the subcarrier number, and F is the total number of subcarriers.

A conventional MR combining solution subject to the TP constraint can be formulated as follows:

$$\mathbf{w}_{\text{MR}}(f) = \tilde{\mathbf{h}}^*(f) \quad (5.5)$$

$$\mathbf{w}_{\text{TR}}(f) = \alpha(f)\sqrt{P_T}\mathbf{w}_{\text{MR}}(f), \sum_{f=1}^F \alpha^2(f) = 1 \quad (5.6)$$

where $\mathbf{w}_{\text{MR}}(f)$ and $\mathbf{w}_{\text{TP}}(f)$ are the $N \times 1$ weight vectors representing the MR and TP solutions, $\mathbf{h}(f)$ is the $1 \times N$ vector representing the propagation channel at the f th subcarrier, $\tilde{\mathbf{h}}(f) = \mathbf{h}(f)/\|\mathbf{h}\|$ is the normalized channel vector, P_T is the total power constraint and $\alpha(f)$ is a power loading function between subcarriers.

The EIRP constraint for an OFDM system is formulated as follows:

$$\max_{\Theta} \sum_{f=1}^F |\mathbf{w}^*(f)\mathbf{a}(\Theta, f)|^2 < \text{EIRP}_a \quad (5.7)$$

$$\max_{\Theta} \sum_{\Phi_i} |\mathbf{w}^*(f)\mathbf{a}(\Theta, f)|^2 < \text{EIRP}_d$$

where $\mathbf{a}(\Theta, f)$ is the $N \times 1$ vector of array manifold depending on the antenna configuration; Θ is a direction-of-arrival parameter; EIRP_a and EIRP_d are the average and spectral density constraints, respectively; and Φ_i are subbands (groups of adjacent subcarriers) over which the spectral density restriction is defined, for example, it can be defined over each subcarrier separately.

The simplest EIRP-restricted solution can be obtained by means of scaling the conventional TP beamformer in (5.6) according to the EIRP constraint:

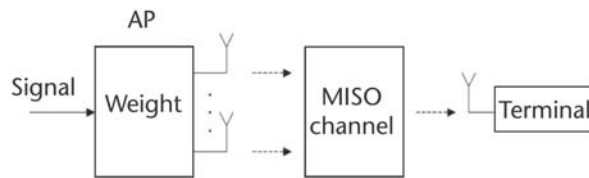


Figure 5.12 System model [45].

$$\mathbf{w}_{\text{SMR}}(f) = \sqrt{\frac{\text{EIRP}_s}{\max_{\Theta} \sum_{f=1}^F |\mathbf{w}_{\text{TP}}^*(f) \mathbf{a}(\Theta, f)|^2}} \mathbf{w}_{\text{TP}}(f), f = 1, \dots, F \quad (5.8)$$

If required, additional scaling according to the spectral density constraint can be introduced similar to (5.8) for the subcarriers that exceed the spectral density constraint:

$$\mathbf{w}_{\text{SMR}_d}(f) = \sqrt{\frac{\text{EIRP}_d}{\max_{\Theta} |\mathbf{w}_{\text{SMR}}^*(f) \mathbf{a}(\Theta, f)|^2}} \mathbf{w}_{\text{SMR}}(f) \quad (5.9)$$

for all f , where

$$\max_{\Theta} |\mathbf{w}_{\text{SMR}}^*(f) \mathbf{a}(\Theta, f)|^2 > \text{EIRP}_d$$

The spectral density constraint in (5.9) is formulated over each subcarrier for simplicity. Both these solutions are scaled MR (SMR) beamformers. The main SMR disadvantage is that it is based on TP that is optimized without taking into account the EIRP constraint. This means that in some propagation conditions the antenna pattern may have sharp peaks to some directions leading to a corresponding reduction in the total power and significant performance degradation.

In the simple narrowband scenario, direct optimization of the beamforming weights subject to the EIRP constraint was introduced [54]:

$$\mathbf{w}_{\text{EIRP}} = \arg \max_{\mathbf{w}} \text{Re}(\mathbf{h}' \mathbf{w}) \quad (5.10)$$

$$|\mathbf{w}^* \mathbf{a}(\Theta_l)|^2 < \text{EIRP}, l = 1, \dots, L \quad (5.11)$$

where L is the number of controllable directions. According to [54], (5.10) and (5.11) are a convex second-order cone programming (SOCP) problem that can be efficiently solved by the available numerical techniques such as the interior point algorithm [55]. An even more computationally efficient approach is proposed in [56] for a similar problem that allows approximation of the quadric constraint in (5.11) by means of a set of linear equations and, hence, reduction of the (5.10) and (5.11) constraint maximization to linear programming.

In the OFDM case considered, the optimization problem in (5.10) and (5.11) can be formulated over each subcarrier separately with consecutive scaling of the overall solution similar to the SMR case in (5.9):

$$\mathbf{w}_{\text{SE}}(f) = \sqrt{\frac{\text{EIRP}_s}{\max_{\Theta} \sum_{f=1}^F |\alpha(f) \mathbf{w}_{\text{EIRP}}^*(f) \mathbf{a}(\Theta, f)|^2}} \mathbf{w}_{\text{EIRP}}(f), f = 1, \dots, F \quad (5.12)$$

$$\mathbf{w}_{SE_d}(f) = \sqrt{\frac{\text{EIRP}_d}{\max_{\Theta} |\mathbf{w}_{SE}^*(f) \mathbf{a}(\Theta, f)|^2}} \mathbf{w}_{SE}(f) \quad (5.13)$$

for all f where

$$\max_{\Theta} |\mathbf{w}_{SE}^*(f) \mathbf{a}(\Theta, f)|^2 > \text{EIRP}_d$$

This solution is the scaled EIRP (SE) beamformer [45]. The main SE disadvantage is that narrowband EIRP-restricted optimization may be a very poor approximation of the general solution especially for a low number of antennas, which is the most practically important case.

Figures 5.13 through 5.15 show all 52 working subcarriers and total antenna patterns for two, three, and four omnidirectional AP antennas with one wavelength distance between them in the 2.4-GHz E-channel case (100-ns delay spread, 1.2-kmph Doppler speed) [52]. The distance between the AP and terminal is $D = 140\text{m}$ in this simulation. Following the European regulation for the 2.4-GHz band [21], only the average EIRP constraint $\text{EIRP}_a = 100 \text{ mW}$ is used. In each case, the total power is reduced according to (5.8).

For the CSI, the channels were estimated from the ACK burst on the uplink assuming an ideal channel reciprocity and $6 \times 10^{-10} \text{ mW}$ noise power and 100-mW transmit power at the terminal. A uniform power loading between subcarriers had been used in all simulations for simplicity.

The following observations can be made for Figures 5.13 through 5.15:

- The total power is less than 100 mW in all three cases. This means that the conventional TP restricted MR combining solution with $\text{TP} = 100 \text{ mW}$ constraint outperforms the SMR beamformer in all of these cases as long

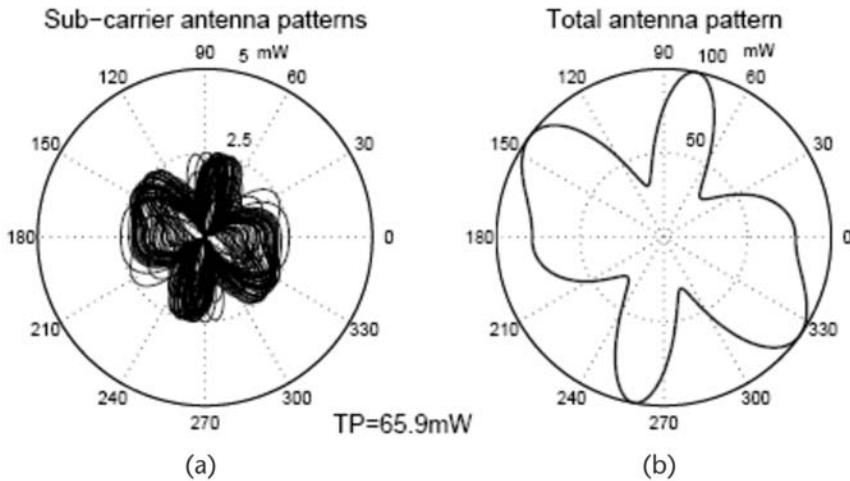


Figure 5.13 (a, b) Typical example of the SMR antenna patterns: $N = 2$, E-channel [45].

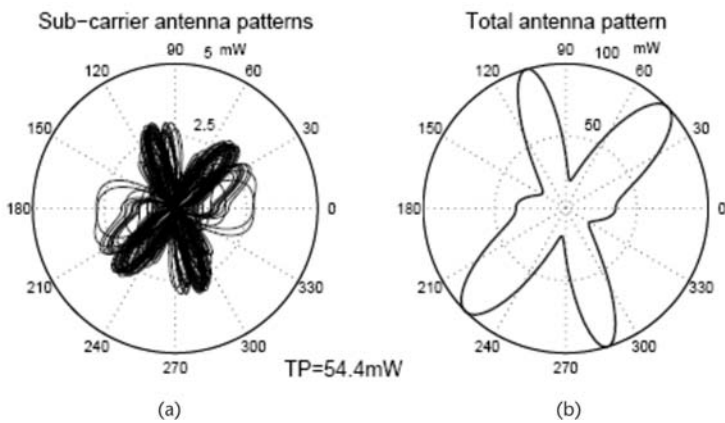


Figure 5.14 (a, b) Typical example of the SMR antenna patterns: $N = 3$, E-channel [45].

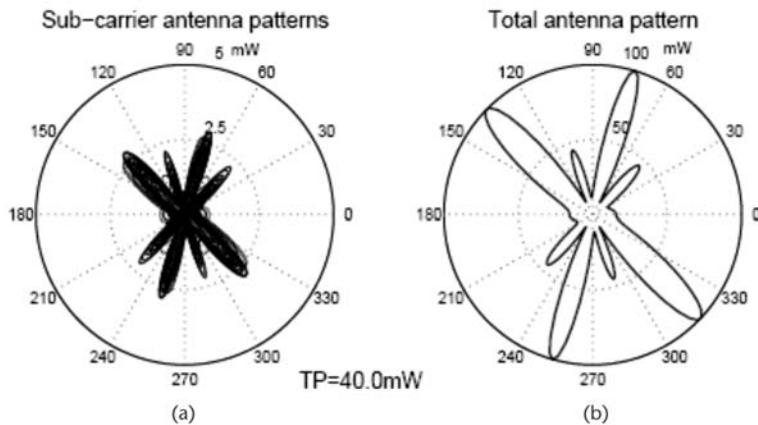


Figure 5.15 (a, b) Typical example of the SMR antenna patterns: $N = 4$, E-channel [45].

- as the shape of the antenna patterns is exactly the same for both MR and SMR.
- The only possibility for TP to equal $EIRP_a$ is to use an omnidirectional total antenna pattern. It never happens in the SMR case with $N > 1$.
 - An increased number of antenna elements leads to sharper antenna patterns and higher reduction of the total power.

The cumulative distribution function for SNR gains at the receiver for both TP and SMR are shown in Figures 5.16 through 5.18 for the same environment as in Figures 5.13 through 5.15. The SNR gains are calculated compared to the single antenna AP for each subcarrier separately in 500 independent trials.

A comparison of the CDF curves in Figures 5.16 through 5.18 show the following:

- The TP beamformer demonstrates the best results in all three cases with the better results for the higher number of antennas (e.g., 70% and 30% of

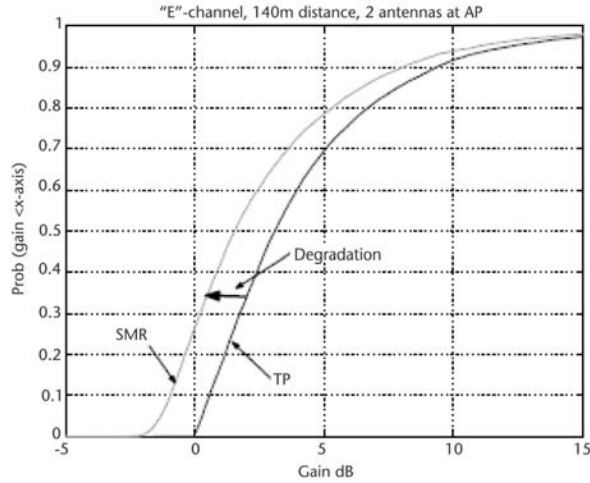


Figure 5.16 CDF of the SNR gain: $N = 2$, E-channel [45].

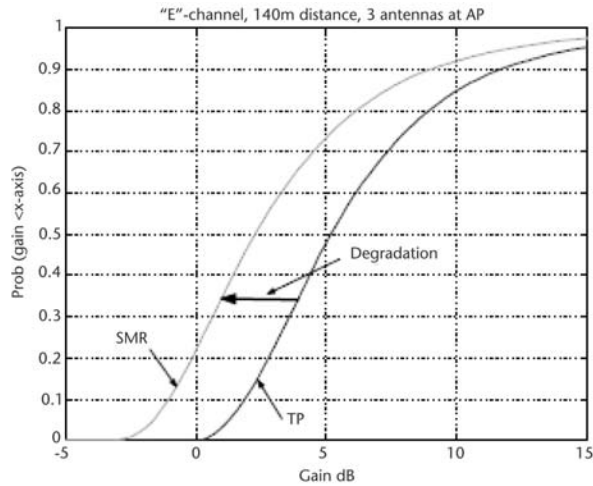


Figure 5.17 CDF of the SNR gain: $N = 3$, E-channel [45].

trials have SNR gains higher than 5 dB for $N = 4$ and $N = 2$, respectively). This is in line with the theoretical considerations.

- SMR shows a performance degradation of a few decibels, which increases with the number of antennas (e.g., the maximal difference is 1.5 dB and 3 dB for $N = 2$ and $N = 3$, respectively).
- Some subcarriers may experience a negative SNR gain in the SMR case that is not possible in the conventional TP case. For example, this happens in almost 30% and 10% of trials for $N = 2$ and $N = 4$, respectively.

The main conclusion from this comparison is that there is a gap of a few decibels between the conventional TP beamformer and the simple SMR solution

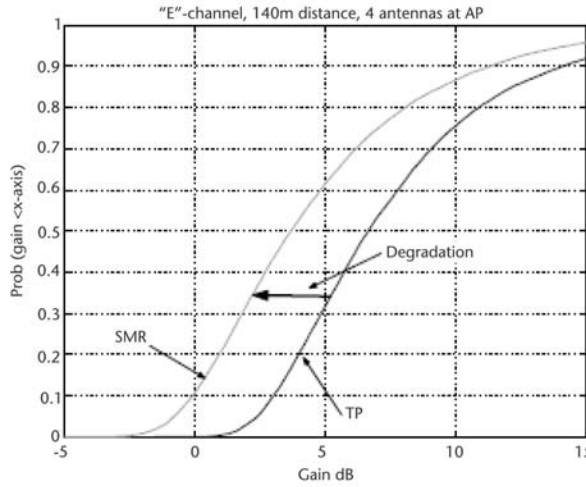


Figure 5.18 CDF of the SNR gain: $N = 4$, E-channel [45].

in the typical broadband multipath propagation scenario. What follows is to find solutions that could improve the performance under the EIRP constraint.

Replacing the quadratic constraint in (5.11) with a set of linear restrictions yields [45]:

$$\text{Re}[\mathbf{w}^* \mathbf{a}(\Theta_i)] \cos\left(\frac{2\pi}{G} + g \frac{2\pi}{G}\right) + \text{Im}[\mathbf{w}^* \mathbf{a}(\Theta_i)] \sin\left(\frac{2\pi}{G} + g \frac{2\pi}{G}\right) < \sqrt{\text{EIRP}} \cos\left(\frac{\pi}{G}\right) \quad (5.14)$$

where $l = 1, \dots, L$, $g = 1, \dots, G-1$, and G is the approximation order that controls its accuracy.

The total number of linear constraints in (5.14) equals LG . For the simulations, high values for these parameters were used: $L = 50$ and $G = 16$ to study the SE potential. For the constraint maximization in (5.11) and (5.14) the “linprog” routine from the Matlab Optimization Toolbox was applied. Typical antenna patterns for the optimized weights after normalization in (5.14) are shown in Figure 5.19 for two antennas at the AP.

It can be seen that all of the patterns are close to the omnidirectional case in this example. This is because the maximization in SE is performed over each subcarrier separately. This does not mean that the SNR at the receiver is always higher than the SNR for the single transmitting antenna case (i.e., it is omnidirectional as well). It appears that many of the separate subcarrier weights are very close to the single antenna solutions with one nonzero weight corresponding to the better channel. This situation is shown in Figure 5.20, which shows the absolute values of the weights in the example in Figure 5.19.

For the higher number of antennas, more complicated patterns can be formed. This is shown in Figures 5.21 and 5.22 for $N = 4$. In this case some performance improvement can be achieved for the OFDM SE algorithm.

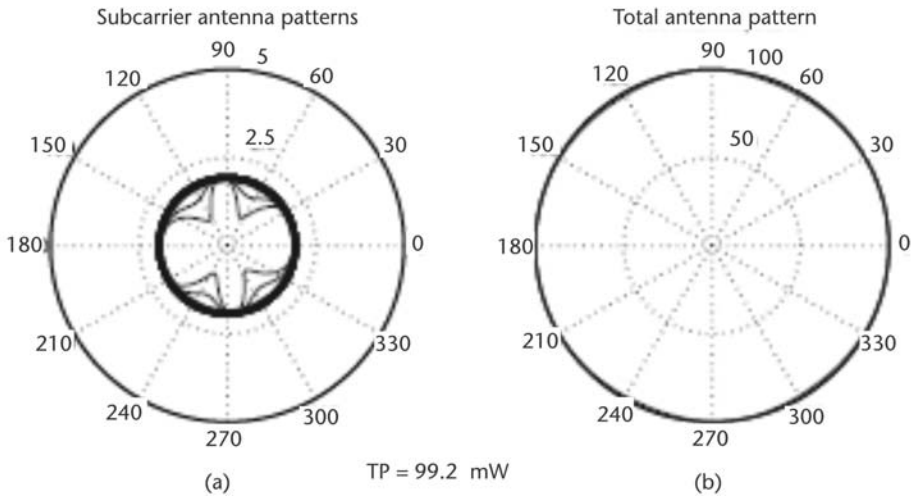


Figure 5.19 Typical example of the SE antenna patterns: $N = 2$, E-channel [45].

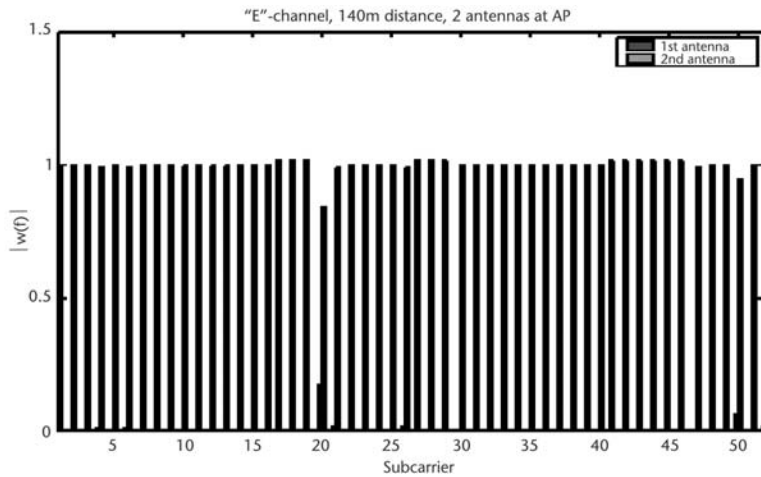


Figure 5.20 Absolute values of the weights [45].

A statistical SE performance in terms of the CDF of the SNR gains is shown in Figures 5.23 through 5.25 together with the SMR and TP curves.

A trivial omnidirectional solution was found in most of the trials, especially, for the low number of AP antennas (e.g., about 50% and 10% for $N = 2$ and $N = 4$, respectively). Taking into account that for two antennas at the AP, the SNR gain for SE is almost always lower than in the SMR case (see Figure 5.23), it can be concluded that there is no reason to expect any range extension possibility for SE compared to SMR. In the four-antenna case in Figure 5.25, the situation is the opposite, but the SE performance is still approaching that for the SMR. It can be concluded that direct OFDM extension of narrowband EIRP-based beamforming cannot significantly improve the simplest SMR beamformer. Thus, a better solution for WLAN OFDM under the EIRP constraint must be sought.

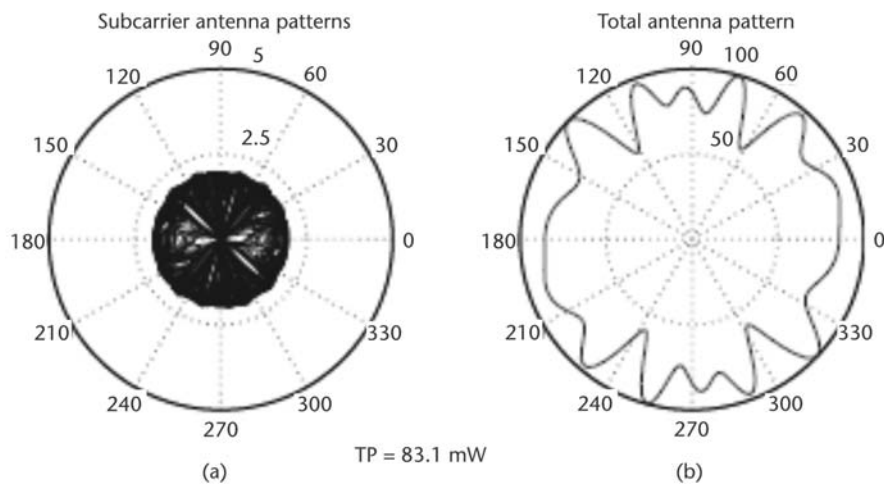


Figure 5.21 Typical example of the SE antenna patterns: $N = 4$, E-channel [45].

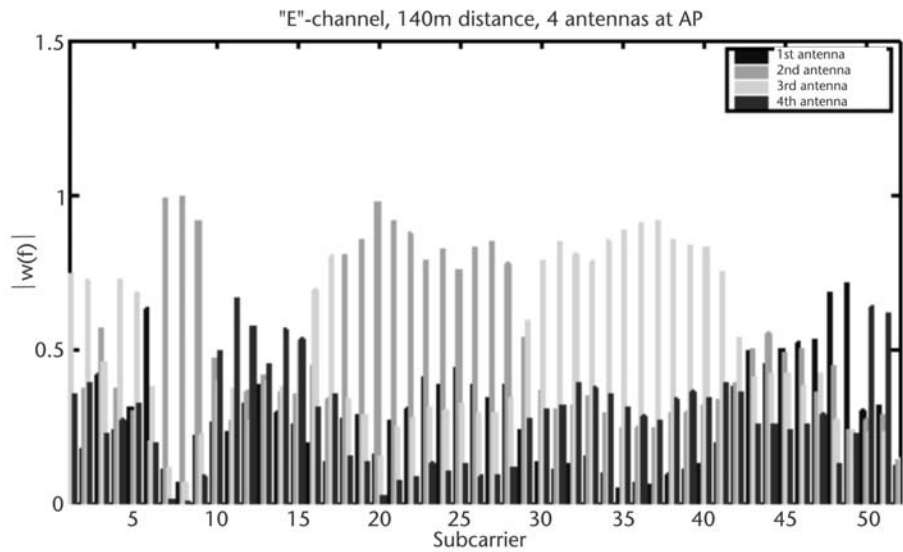


Figure 5.22 Absolute values of the weights [45].

Such a solution can be achieved by defining a general optimization problem subject to the EIRP constraint [45]. Different optimization criteria similar to the OFDM power allocation criteria can be exploited. Two examples of such criteria are maximization of the average SNR or maximization of the minimum SNR over all of the subcarriers at the receiver. Their efficiency may be different for different signaling schemes, coding schemes, and propagation conditions.

Note that a very high number of variables and constraints might make this solution impractical. For example, 52 working subcarriers in IEEE 802.11a/g lead to 208 and 416 real value optimization parameters in the two- and four-antenna cases, respectively. Some complexity reduction can be achieved by means of the

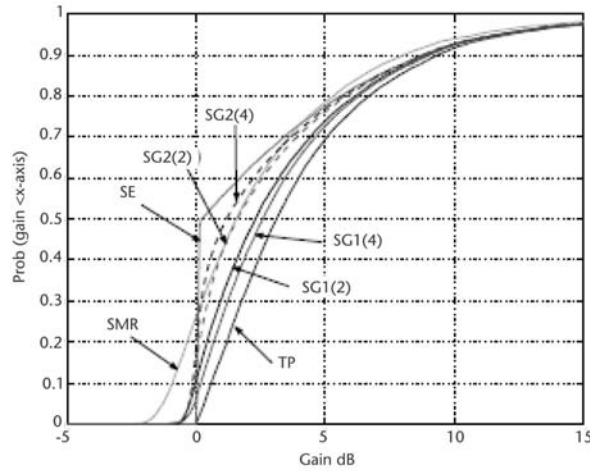


Figure 5.23 CDF of the SNR gain: $N = 2$, E-channel [45].

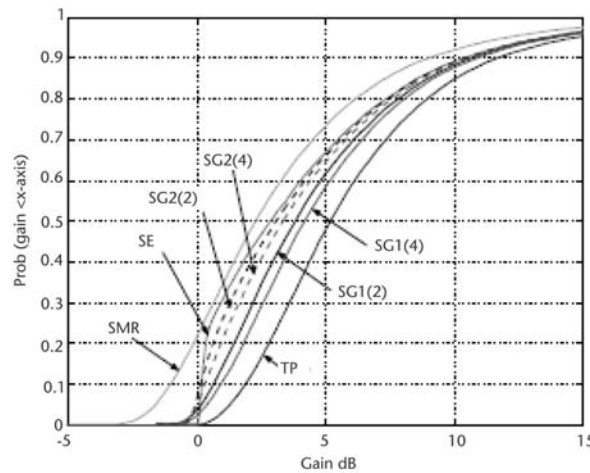


Figure 5.24 CDF of the SNR gain: $N = 3$, E-channel [45].

conventional grouping of adjacent subcarriers with highly correlated channels in some propagation conditions, but this still may be too complicated for online implementation.

Instead, a suboptimal solution based on subcarrier grouping can be applied as reported in [57], which used an EIRP-based optimization over a group of nonadjacent subcarriers with low channel correlation for decomposition of the joint optimization problem as described in [45]. The antenna patterns achievable through this solution are given in Figures 5.26 and 5.27.

Figure 5.26 shows the antenna patterns and gains in one particular optimization group $\Phi_{13} = [18, 45]$ in the two-antenna AP case. Both scaled EIRP beamformer (SE) and scaled group beamformer (SG) have the same omnidirectional total

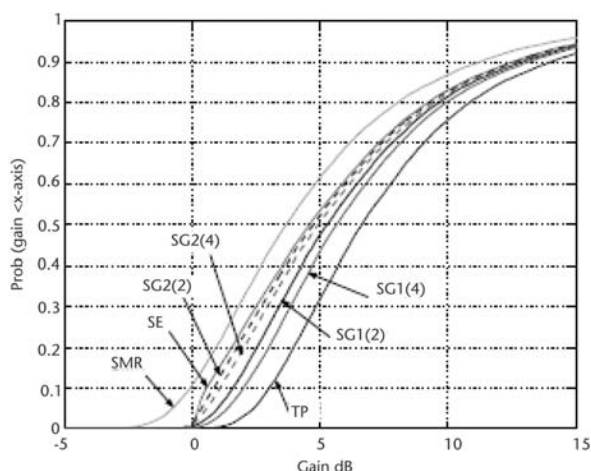


Figure 5.25 CDF of the SNR gain: $N = 4$, E-channel [45].

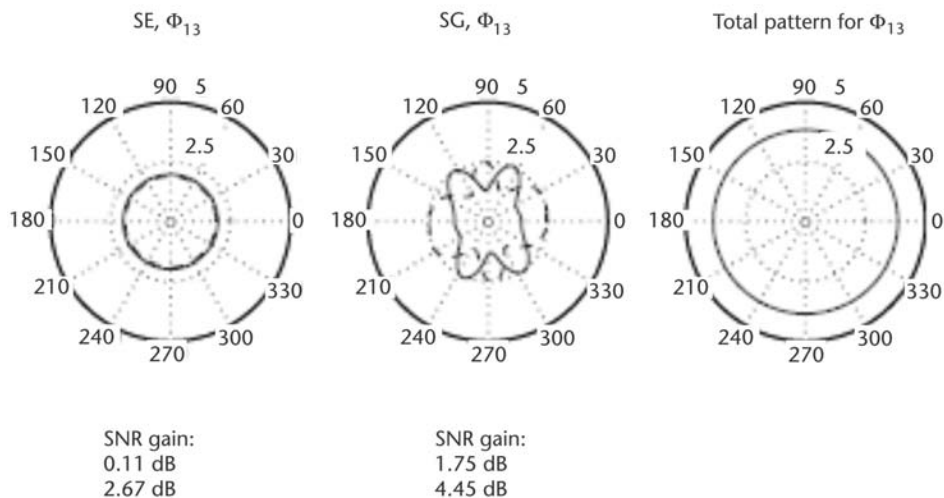


Figure 5.26 Typical example of the antenna patterns for one group of subcarriers: $N = 2$, E-channel [45].

antenna pattern over this group of subcarriers. The difference is that SE finds weights that are close to the single antenna case with very low antenna gains compared to the conventional single antenna AP.

Figure 5.27 shows an example of all 52 subcarriers and the total antenna patterns for SMR and SG based on the average EIRP restriction in the three-antenna AP case. Sharp colocated subcarrier beams are formed in the SMR case, leading to a significant reduction in total power (TP = 47.1 mW) because of the EIRP constraint. The proposed solution forms spatially distributed beams leading to the omnidirectional total pattern and TP \approx 100 mW. The CDF for the SNR gain for the SG algorithm was shown earlier in Figures 5.23 through 5.25. The

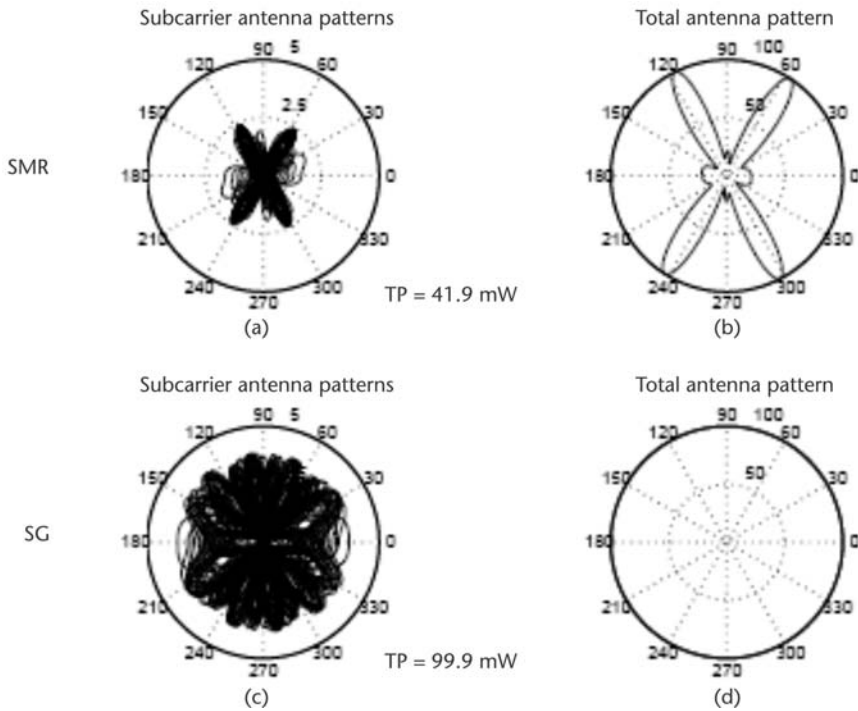


Figure 5.27 (a–d) Typical example of the antenna patterns for all the subcarriers: $N = 3$, E-channel [45].

results show that the SG grouping is very important. SG2 does not demonstrate any significant improvement compared to SE because highly correlated channels in a group do not help with the lack of dimensionality. In contrast, the proposed subcarrier grouping in SG1 shows significant performance improvement in this scenario.

An additional EIRP spectral density constraint may be defined in some frequency bands for OFDM systems, for example, in the 5.2-GHz band [58]. Scaling the weights on some subcarriers with a violated spectral density restriction is a suboptimal solution that may cause some additional performance degradation especially for a high number of antennas when sharp beams can be formed [45].

Figure 5.28 shows the CDF for the four-antenna case with the spectral density constraint. Four subcarriers in a group lead to some performance degradation compared to the case of two subcarriers in a group (not shown in the figure). An explanation is that the spectral density constraints make the problem suboptimal when sharp beams can be formed as in the case with four AP antennas and four subcarriers in an optimization group. This situation is probably not critical because for some important bands like the 2.4-GHz band, spectral density constraints are not defined for WLAN OFDM systems.

Figure 5.29 shows the estimated PER performance for two and four antennas at the AP.

All beamforming algorithms demonstrate significant range extension compared to the single-antenna case. In the four-antenna case, the EIRP-based SMR solution

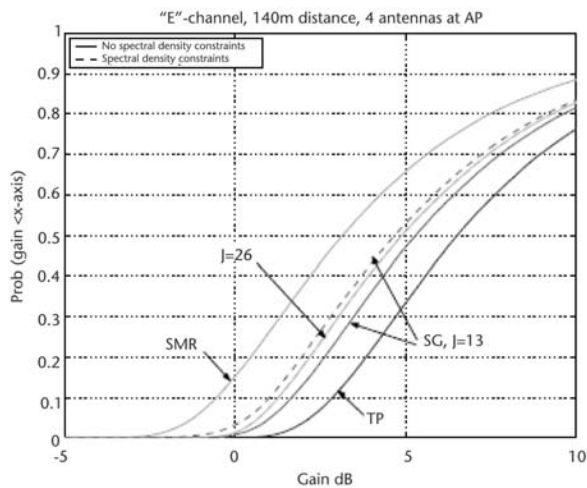


Figure 5.28 CDF for SNR gains with and without EIRP spectral density restrictions: $N = 4$, E-channel [45].

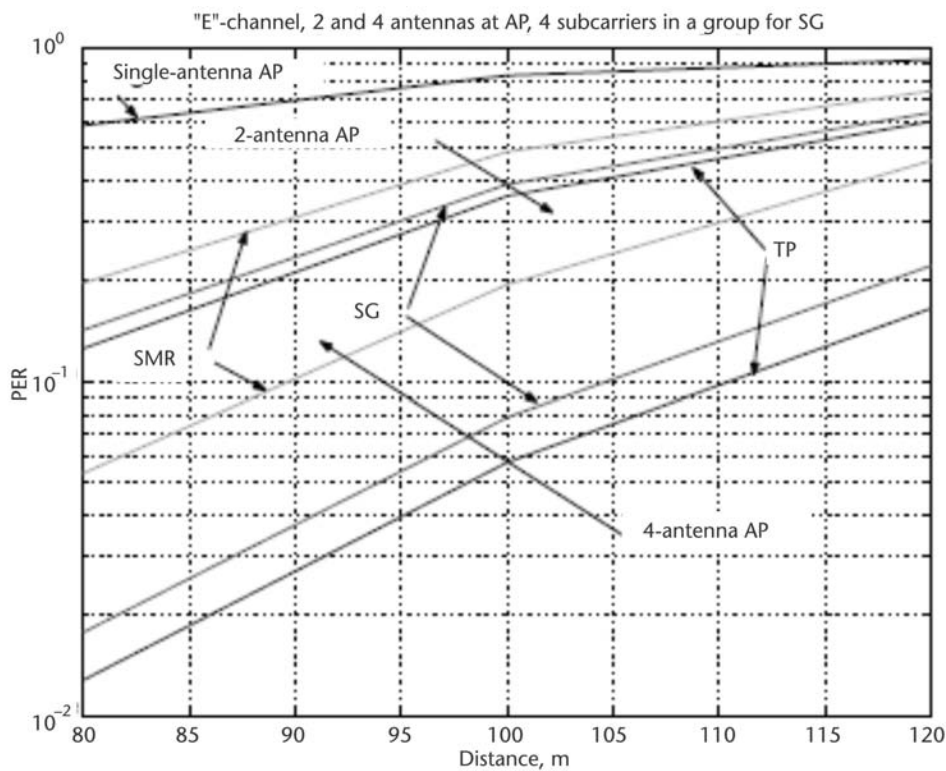


Figure 5.29 PER performance: $N = 2$ and $N = 4$, E-channel.

shows about a 22% performance degradation at 1% PER compared to the basic beamformer with the total power restriction. At the same PER level, the proposed SG solution demonstrates about a 17% improvement compared to the TP case.

5.2.3 Downlink Capacity Enhancement of IEEE 802.11a/g Using SDMA with a Multiple-Antenna Access Point

The basic DL scenario assumes single-antenna terminals waiting for data from an AP equipped with N antennas. In the general case, the number of terminals may be higher or even much higher than the number of transmitting antennas, that is, $L \gg N$ (see Figure 5.9 earlier in the chapter). The conventional solution is to serve all of the users sequentially with single or multiple transmitting antennas using conventional, beamforming, or transmit diversity schemes. An SDMA solution attempts to serve some subset of $M < L$ users, where $M \leq N$. Then, the rest of the users can be served sequentially on a group-by-group basis. DL SDMA is a special case of an $N \times M$ MIMO system with no cooperation between users. This restricts the potential SDMA performance, first of all, because effective multiuser receiving techniques cannot be applied in this case.

The narrowband (e.g., for each separate subcarrier of an OFDM system) transmitted and received signals for the basic DL SDMA operation for a group of M users can be expressed as follows:

$$\mathbf{X} = \mathbf{W}\mathbf{S} \quad (5.15)$$

$$\mathbf{y}_m = \mathbf{h}_m^T \mathbf{X} + \mathbf{z}_m, \quad m = 1, \dots, M \quad (5.16)$$

where \mathbf{S} is the $M \times T$ matrix of the transmitted signals ($E\{\mathbf{S}\mathbf{S}^*\}/T = P\mathbf{I}_M$), T is the number of symbols, P is the signal power, \mathbf{I}_M is the $M \times M$ identity matrix, \mathbf{W} is the $N \times M$ weight matrix, \mathbf{X} is the $N \times T$ matrix of the combined transmitted signal, \mathbf{y}_m is the $1 \times T$ vector of the signal received by the m th terminal, \mathbf{h}_m is an $N \times 1$ vector representing the propagation channel from the N AP antennas to the single antenna at the m th terminal, and \mathbf{z}_m is a $1 \times T$ AWGN noise vector with variance σ^2 at the m th terminal.

A key feature of wireless LANs is that the radio conditions are different at the transmitter and receiver. As a result, the transmitter has no way of knowing whether the transmitted data was received correctly at the receiver. Upon successful reception of a data burst [i.e., a MAC packet data unit (MPDU)], the receiver should send an ACK frame to the transmitter as confirmation. Should the transmitter not receive an ACK frame, it will assume a lost MPDU and will attempt retransmission. The time interval between the last symbols of the MPDU and the first symbol of the ACK frame is referred to as a short interframe space (SIFS) interval and is fixed at $16 \mu\text{s}$ [52]. The duration of a MPDU is arbitrary, and the duration of an ACK frame is between 24 and $32 \mu\text{s}$, depending on the modulation and coding PHY parameters. The IEEE 802.11 MAC protocol is based on CSMA/CA. This essentially describes a “listen before you talk” access mechanism, whereby an IEEE 802.11 radio (terminal or access point) listens to the medium before starting a transmission.

For the SDMA scenario, it had been assumed that the AP first selects M out of L terminals within its coverage area for purposes of DL SDMA transmissions. The selection may be arbitrary, based on a scheduling policy, or based on their suitability for supporting SDMA (i.e., colinearity of their DL channels).

For the simulations, a simple grouping strategy can be assumed, in which a group of SDMA users is formed from randomly selected closely spaced terminals at some specified distance from the AP. A randomly selected group of terminals is used for the whole 10-ms SDMA session. Independent randomly selected groups are used for different SDMA sessions. This strategy allows us to investigate the expected performance defined by the standard IEEE 802.11 environment itself without smart grouping techniques. This scenario is shown in Figure 5.30. Note that a direct application of the conventional simultaneous DL SDMA transmission scheme to an IEEE 802.11a/g AP is difficult because of the IEEE 802.11a/g standard requirements for ACKs. The particular case of two SDMA users in a group, that is, $M = 2$, can be addressed by means of a time-shifted SDMA transmission. The proposed solution is to impose a time offset between the SDMA-transmitted MPDUs, causing a similar time offset in the ACK responses of the two terminals, as shown in Figure 5.31.

A time offset reduces the interference between the ACKs, in particular during the critical synchronization and pilot intervals of ACK1. In principle, the maximum value of this offset is $16\text{ }\mu\text{s}$ because simultaneous transmission and reception are not allowed at the AP or terminal. It is important for the ACK2 symbols interfering with ACK1 (i.e., the synchronization and pilot symbols) to be known a priori. This implies that the minimum allowed offset is restricted as well.

The main difficulty in recovering the shifted ACKs is that both of them cannot be sampled synchronously because the SIFS intervals are maintained independently at each terminal only within some finite accuracy. The idea is to sample the received

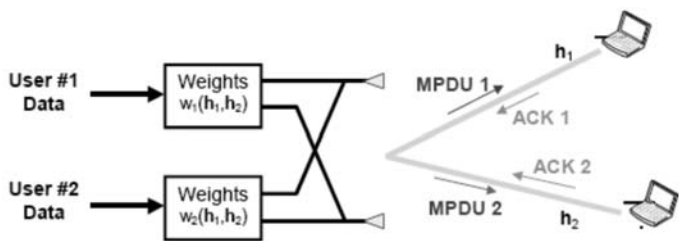


Figure 5.30 SDMA for a group of two terminals [45].

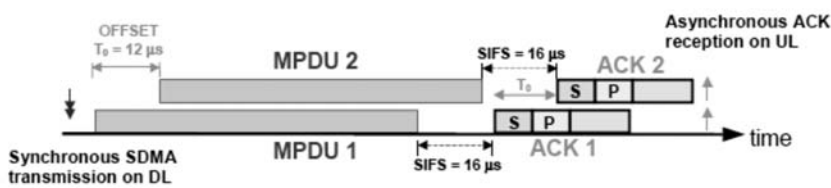


Figure 5.31 Time offset SDMA slot.

additive mixture of the ACKs synchronously for the nondelayed ACK (ACK1 in Figure 5.31) and use the channel estimate and the known symbols of the shifted ACK (ACK2 in Figure 5.30) that interfere with ACK1 to form an oversampled replica of the received ACK2. This oversampled replica is then subtracted from the received signal after a fine synchronization to obtain a cleaned up version of the received signal for ACK1 recovery. When all of the symbols of ACK1 are recovered, then a similar procedure is implemented to obtain a cleaned up version of the received signal for ACK2 recovery.

With this in mind and the detailed calculations reported in [45], the IEEE 802.11g system had been simulated for two users in a conference room scenario. The main simulation assumptions are as follows:

- 2.4-GHz frequency range;
- 16-QAM, convolutional encoding with 3/4 code rate;
- 4,320 information bit packets (35 OFDM symbols or 140- μ s total slot duration);
- 10-ms SDMA session (initialization is not included);
- 2λ separation between AP antennas, 4λ separation between terminals;
- 3m to 15m distance range;
- 16-dBm transmit power;
- -92-dBm noise power;
- Asynchronous ACK arrivals with $(16 \pm 0.5) \mu$ s delay;
- B-channel with 15-ns RMS delay spread without fluorescent light effects;
- D-channel with 50-ns RMS delay spread with fluorescent light effects.

Figure 5.32 shows the synchronization step of the employed ACK reception algorithm. It shows the three-dimensional minimization function for one typical SDMA slot in the D-channel case. The global minimum of this function corresponds to the synchronization parameters \hat{t}_1 and \hat{t}_2 used in the algorithm.

Figure 5.33 shows a CDF of the overall throughput performance for a two-antenna AP $N = 2$ in the B-channel case. Benchmarks 1 and 2 demonstrate similar results because of the high SNR in a conference room environment. Benchmark 3 suggests that in the case of an ideal ACK receiver, the throughput can be doubled in the considered environment for almost all randomly selected channels. Neverthe-

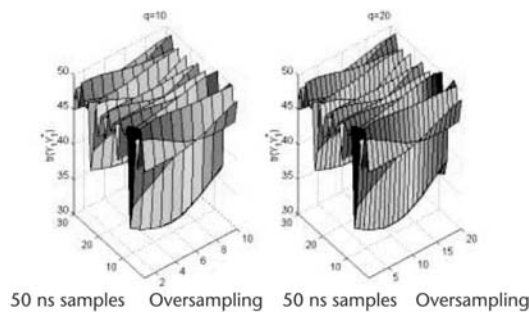


Figure 5.32 Illustration of the synchronization step in the ACK reception algorithm [45].

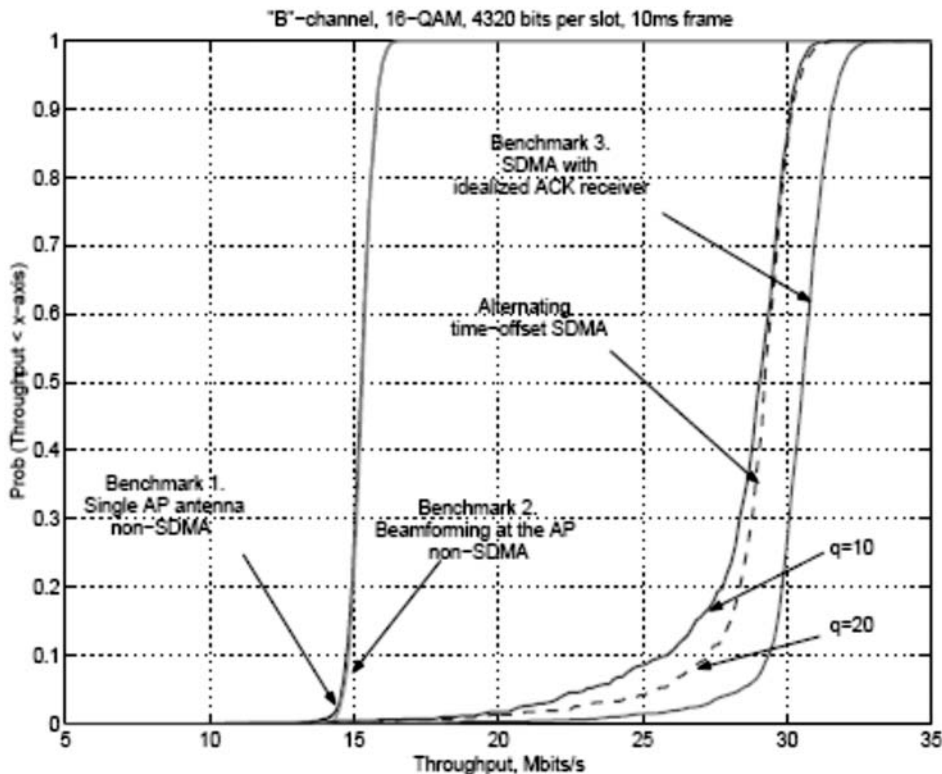


Figure 5.33 Throughput CDF for two SDMA users and two-antenna AP: B-channel, 3m to 9m distance.

less, one can see that a small portion of “poor” channels still exists in this environment. The proposed SDMA solution demonstrates results close to benchmark 3 for most of the trials. The constant CDF shift for the alternating time offset solution corresponds to the introduced time offset, which is equal to $12\ \mu\text{s}$ in this simulation. Additional performance degradation compared to benchmark 3 is connected with the nonideal ACK receiver.

Similar results in the D-channel environment are shown in Figure 5.34, where we can see that benchmark 3 and especially the proposed solution suffer some performance degradation in this case. This behavior reflects the more complicated channel conditions in terms of the delay spread and time-variation effects (fluorescent light effects are included in the D-model).

An increased number of AP antennas can significantly improve the performance for the D-channel case. The estimated throughput performance for three AP antennas ($N = 3$) is shown in Figure 5.35, whereas Figure 5.36 shows the throughput results in the same scenario, but for longer distances (5m to 30m) and $q = 20$.

Benchmark 2 outperforms benchmark 1 with and without channel updates. This is because of the low sensitivity of the main beam to channel errors caused by the low Doppler speed in the D-channel model in the non-SDMA system.

In contrast to non-SDMA beamforming, very significant performance degradation can be observed for both benchmark 3 and the proposed solution without

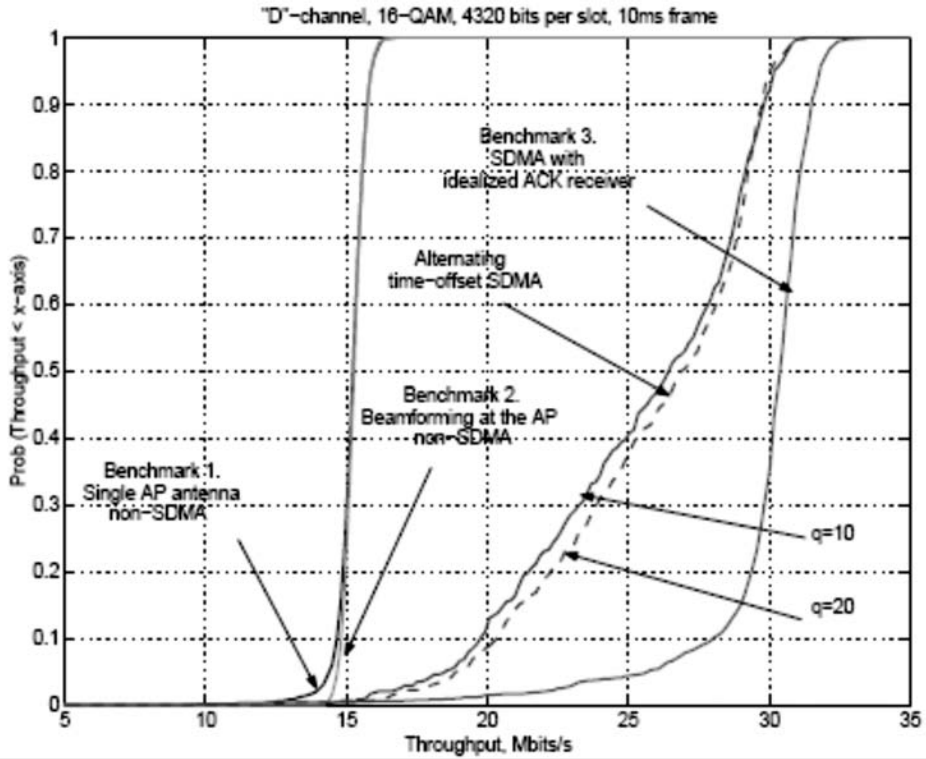


Figure 5.34 Throughput CDF for two SDMA users and two-antenna AP: D-channel, 5m to 15m distance.

channel updates during a 10-ms SDMA session. This is because of the very high sensitivity of the beamforming zeros toward the SDMA users even in the low Doppler speed environment. This justifies the need for regular channel updates, and the alternating channel estimation procedure is effective in the considered scenario.

The proposed alternating time offset SDMA solution supports legacy terminals but requires some modifications to the transmission protocol at the AP. The main features of the IEEE 802.11a/g standard, such as channel conditions, transmission protocol, and data and ACK slot structures, have been taken into account. The downlink capacity in a conference room environment is almost doubled, which was demonstrated by simulations based on the channel models approved by the IEEE 802.11 Standard Group.

5.3 Modeling of Two- and Three-Dimensional Antenna Arrays for Channel Model Measurements

The spatial features displayed by the MIMO channel are greatly influenced by the antenna arrays, whose characteristics are determined by the positions of their elements, as well as their beam patterns and coupling [59]. This section summarizes

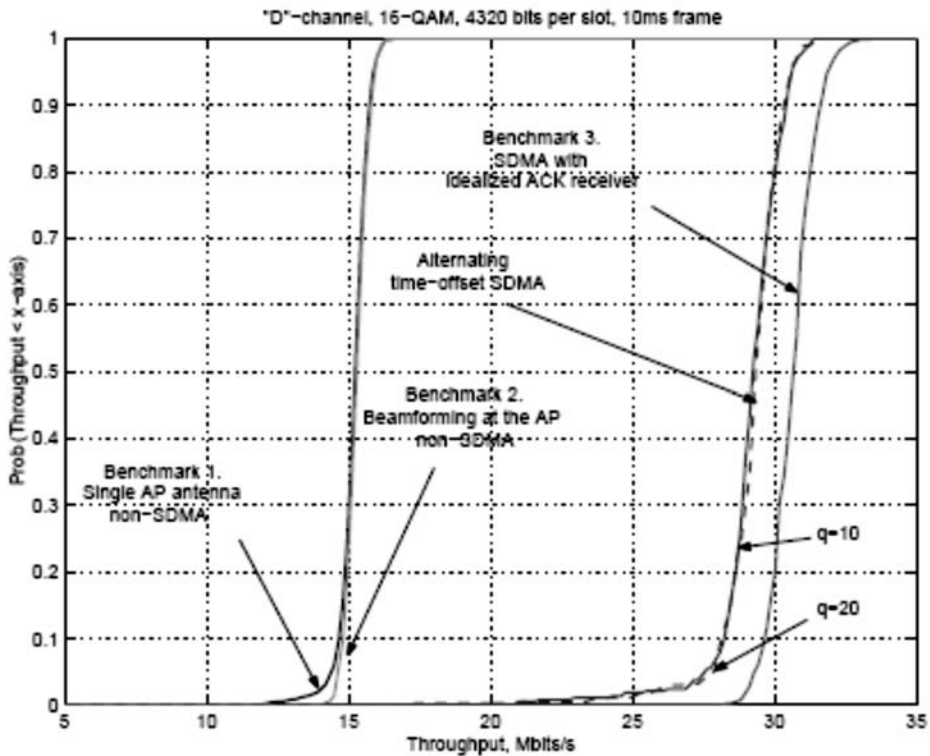


Figure 5.35 Throughput CDF for two SDMA users and three-antenna AP: D-channel, 5m to 15m distance.

some design considerations for antenna deployment, channel measurement, and identification that were proposed and implemented by the IST projects WINNER and WINNER II [2]. The considerations were made in support of the design of the WINNER air interface, which is a scalable concept in terms of carrier bandwidth and carrier frequency range. The system concept supports a wide range of radio environments providing a significant improvement in performance and QoS. The WINNER radio interface optimizes the use of spectral resources (e.g., through the exploitation of actual channel conditions and multiple-antenna technology).

5.3.1 Overview of Antenna Models

Planar antenna arrays such as uniform linear arrays or uniform rectangular arrays always have a limited viewing angle. They are useful to represent a BS view of the channel. Circular antennas have a full field of view. They can be conveniently used to represent the MS.

Double directional modeling requires antenna arrays at both ends of the link and MIMO operation of the sounder. For cellular system consideration, a combination of planar and circular arrays is adequate, whereas for ad hoc peer-to-peer networks identical circular arrays are most preferable.

Mainly for microcell and picocell scenarios, estimation of the elevation is desired in addition to the azimuth. This requires URA, cylindrical, or spherical

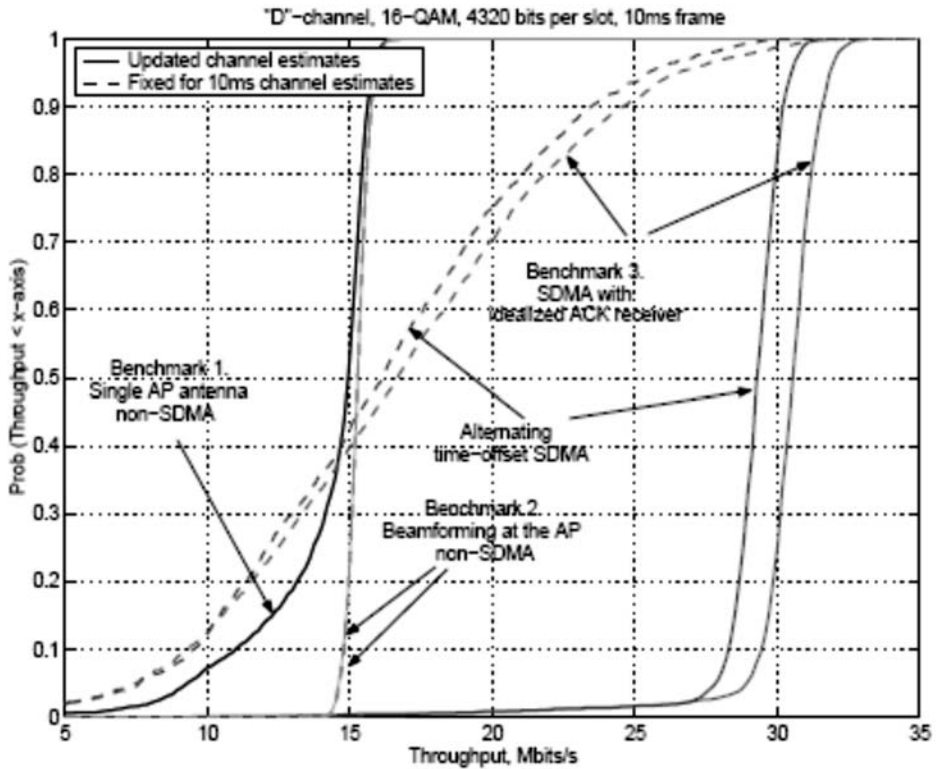


Figure 5.36 Throughput CDF for two SDMA users and three-antenna AP: D-channel, 5m to 30m distance. Dashed curves correspond to channels fixed over 10-ms intervals [45].

arrays. Three-dimensional wave analysis (azimuth and elevation) should always include polarization resolution.

Spherical antenna arrays may be applied for full azimuth and elevation coverage. However, there exists no geometric solution to arrange more than 20 patch antenna elements on a spherical surface with identical interelement distances. Therefore, the design of spherical arrays might be complicated by nonuniform interelement distances and various relative polarization orientations of adjacent elements.

Full polarimetric analysis of the radio channel requires not only polarimetric reception but also polarimetric excitation of the channel. This is even true for omnidirectional excitation, in which a two-port antenna is needed to send orthogonal polarized waves with omnidirectional characteristics, thus doubling the required sounder output ports.

High and reliable resolution in terms of separation capability of closely spaced paths and low probability of outliers requires an antenna architecture that offers a maximum antenna array aperture size in the respective spatial dimension, including a minimum number of antenna elements, low antenna element coupling, and precise calibration. This also has to include the antenna switches and feeder cables. The characteristics of the antenna elements depend on the basic element design (dipoles, patches, slots, and so on). The basic design has a significant influence on high-resolution performance, estimation ambiguities, probability of outliers, and

polarization resolution capability. Furthermore, when modeling the antenna characteristics the accuracy and the computational complexity are important, therefore an efficient description of the polarimetric beam pattern can be applied, such as the effective aperture distribution function.

To allow antenna de-embedding/embedding in the channel model, special attention must be paid to antennas used in MIMO channel measurements. Multielement spherical, circular, linear, planar, or cylindrical arrays are required in the measurements to characterize propagation effects, such as DoA and DoD in addition to the delay and path gain, and to separate them from the antenna effects. Therefore, the antennas used in channel measurements are often different from those used in the real radio systems [59].

The following general assumptions were made for the WINNER channel models (see also Chapter 2 of this book) [59]. It was assumed that both computational complexity and number of parameters of the channel models should be minimized without losing accuracy and reliability, that the models should support both software and hardware simulations and testing, and that the models for the link level algorithm testing would also be defined in hardware. The propagation models should consist of an antenna-independent part and a part that defines the antenna characteristics. Further, the models should support arbitrary antenna patterns, polarizations, and antenna array configurations. A set of predefined antenna arrays should also be delivered.

5.3.2 Modeling of Antenna Characteristics

Transfer matrix \mathbf{H} of the MIMO channel with L paths can be given by [60–63]:

$$\mathbf{H}(t, \tau) = \sum_{l=1}^L \mathbf{H}_l(t; \theta_l) = \sum_{l=1}^L \exp(j2\pi\nu_l t) \mathbf{C}_2(\Omega_{2,l}) \mathbf{A}_l \mathbf{C}_l^T(\Omega_{1,l}) \delta(t - \tau_l) \quad (5.17)$$

where the components included are the Doppler frequency ν , the Rx array response matrix \mathbf{C}_2 , the polarization matrix \mathbf{A} , the Tx array response matrix \mathbf{C}_1 , and the delay τ . The parameter vector is

$$\theta_l = (\Omega_{1,l}, \Omega_{2,l}, \tau_l, \nu_l, \mathbf{A}_l) \quad (5.18)$$

The polarization matrix \mathbf{A} of path l is a 2×2 matrix, of which elements $\alpha_{l,i,j}$ are the complex gains from Tx polarization i to Rx polarization j of path l , where $i, j \in \{1,2\}$ refers to two polarizations:

$$\mathbf{A}_l = \begin{bmatrix} \alpha_{l,1,1} & \alpha_{l,1,2} \\ \alpha_{l,2,1} & \alpha_{l,2,2} \end{bmatrix} \quad (5.19)$$

The array response matrices are composed of array response vectors of two polarizations. For the Tx array,

$$\mathbf{C}_1(\Omega_{1,l}) = [\mathbf{c}_{1,1}(\Omega_{1,l}) \quad \mathbf{c}_{1,2}(\Omega_{1,l})] \quad (5.20)$$

and for the Rx array,

$$\mathbf{C}_2(\Omega_{2,l}) = [\mathbf{c}_{2,1}(\Omega_{2,l}) \quad \mathbf{c}_{2,2}(\Omega_{2,l})] \quad (5.21)$$

The array response vector is $\mathbf{c}_{i,j}$ where $i = 1$ denotes Tx, $i = 2$ Rx, and j stands for the j th polarization. On the vectors below $\mathbf{r}_{i,m,j}$ is the vector to the m th antenna element with j th polarization. For example,

$$\mathbf{c}_{1,1}(\Omega_{1,l}) = \begin{bmatrix} G_{1,1,1}(\Omega_{1,l}) \exp(j2\pi\lambda_0^{-1}(\Omega_{1,l} \cdot \mathbf{r}_{1,1,1})) \\ G_{1,2,1}(\Omega_{1,l}) \exp(j2\pi\lambda_0^{-1}(\Omega_{1,l} \cdot \mathbf{r}_{1,2,1})) \\ \vdots \\ G_{1,M_1,1}(\Omega_{1,l}) \exp(j2\pi\lambda_0^{-1}(\Omega_{1,l} \cdot \mathbf{r}_{1,M_1,1})) \end{bmatrix} \quad (5.22)$$

where $G_{i,m,j}(\Omega)$ is the antenna field pattern of the Tx/Rx ($i = 1/2$) m th antenna element on the j th polarization and direction Ω .

Note that the right-hand expression in (5.22) is valid if coupling is neglected. However, (5.17) is still valid for any arrays $\mathbf{C}_1(\Omega)$ and $\mathbf{C}_2(\Omega)$ and the array responses (e.g., obtained from calibration).

In the antenna steering vector of (5.22) for the channel transfer matrix in (5.17), the antenna locations $\mathbf{r}_{i,m,j}$ are freely settable [59]. Additionally, the radiation patterns $G_{i,m,j}(\Omega)$ are freely settable for each antenna element for the two polarizations. Radiation patterns can be given as a closed-form function of the direction or as tabulated values. If available, the closed-form function would be computationally effective. A table of radiation pattern values on azimuth and elevation directions is more general, but it requires interpolation. Mutual coupling of array elements can be embedded in the tables of radiation patterns.

The radio channel model is termed *antenna independent* if the antennas and propagation channel are separable in the model formulation [e.g., (5.17)]. An antenna-independent model supports arbitrary antenna array geometries and arbitrary radiation patterns for the antenna elements. The antenna information can be embedded in the generation procedure of the channel model realization.

The cell structure and topology of the antennas have to be carefully considered for system level simulations. Different BS sides could have different antenna designs/characteristics as well as the MS antennas. Furthermore, for path-loss modeling issues, the specific different antenna gains also play a role.

An antenna array is described as a collection of antenna elements, where the array itself can be placed anywhere in space. The coordinate system where this placement takes place is called the global-coordinate-system (GCS). It is possible to rotate the array by any axis of the GCS. The GCS is used to define the radio network system layout and as a reference system for polarization.

5.3.2.1 Three-Dimensional Antenna Arrays

A channel model of the type developed for the WINNER radio interface needs as inputs general information such as the channel scenario and the MIMO setup, the

antenna configurations (e.g., radiation patterns and array geometries), and the system layout information (e.g., relative distances and orientations of the transceivers). An output of the model is a set of discrete channel impulse responses with matrix coefficients (see Chapter 2 of this book or refer to [64] for the exact matrix). The entries of the matrices are complex channel coefficients for each of the transmitter/receiver antenna element pairs. The channel impulse responses are the realizations of the radio channel for discrete time instants and for different radio links.

The channel sampling frequency ultimately has to be equal to the simulation system sampling frequency. To obtain feasible computational complexity, it is not possible to generate channel realizations on the sampling frequency of the system to be simulated. The channel realizations have to be generated on some lower sampling frequency and then interpolated to the desired frequency. A practical solution is to generate the channel samples with a sample density (oversampling factor) of 2, interpolate them accurately to a sample density 64, and then apply zero order hold interpolation to the system sampling frequency [64]. The channel impulse responses can be generated during the simulation or stored on a file before the simulation on a low sample density. Interpolation can be done during the system simulation.

Figure 5.37 shows an example of a system layout in the Cartesian coordinates:

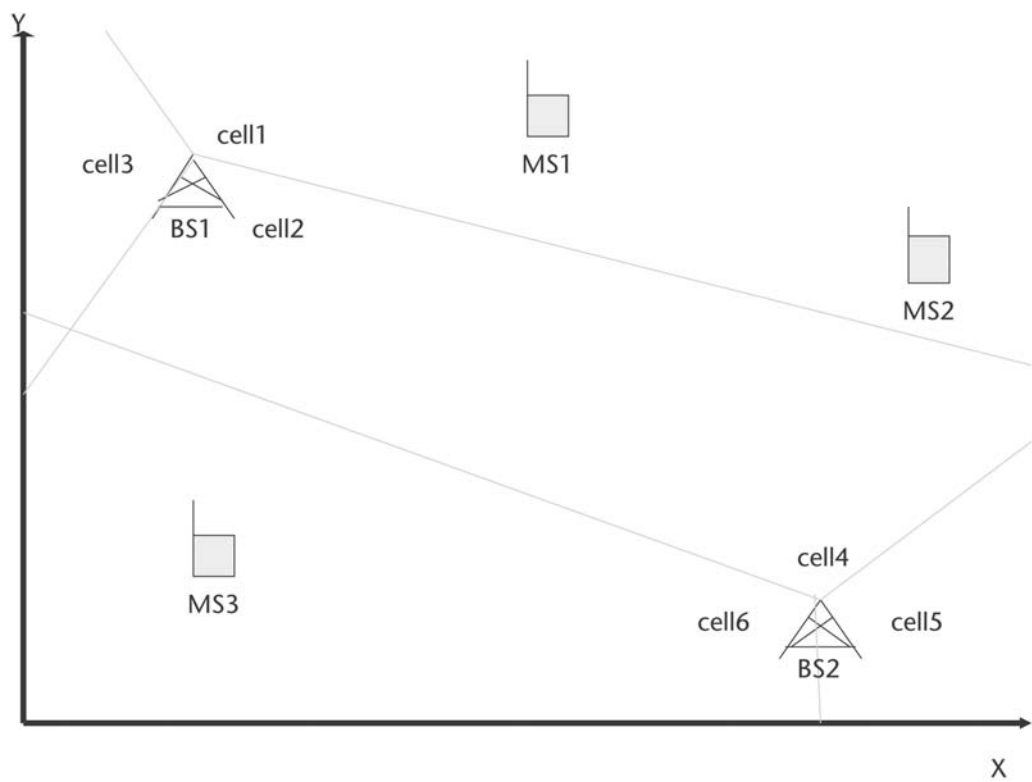


Figure 5.37 A system layout of multiple BSs and MSs [64].

All of the BSs and MSs have (x, y) coordinates. The MSs and cells (sectors) also have an array broadside orientation, where north (up) is the zero angle. The positive direction of the angles is the clockwise direction.

Both the distance and LOS direction information of the radio links are calculated for the input of the model. The distance between BS_i and MS_k is

$$d_{BS_i, MS_k} = \sqrt{(x_{BS_i} - x_{MS_k})^2 + (y_{BS_i} - y_{MS_k})^2} \quad (5.23)$$

The LOS direction from BS_i to MS_k with respect to the BS antenna array broadside is

$$\theta_{BS_i, MS_k} = \begin{cases} -\arctan\left(\frac{y_{MS_k} - y_{BS_i}}{x_{MS_k} - x_{BS_i}}\right) + 90^\circ - \Omega_{BS_i} & \text{when } x_{MS_k} \geq x_{BS_i} \\ -\arctan\left(\frac{y_{MS_k} - y_{BS_i}}{x_{MS_k} - x_{BS_i}}\right) - 90^\circ - \Omega_{BS_i} & \text{when } x_{MS_k} < x_{BS_i} \end{cases} \quad (5.24)$$

The angles and orientations are shown in Figure 5.38.

In the case of Figure 5.38, pairing matrix \mathbf{A} [see (5.25)] is a 3×6 matrix with values $\{0,1\}$. Value 0 standing for link MS_m to cell $_n$ is not modeled.

$$\mathbf{A} = \begin{bmatrix} \chi_{ms1, c1} & \chi_{ms1, c2} & \cdots & \chi_{ms1, c6} \\ \chi_{ms2, c1} & \chi_{ms2, c2} & \cdots & \chi_{ms2, c6} \\ \vdots & \vdots & \ddots & \vdots \\ \chi_{ms3, c1} & \chi_{ms3, c2} & \cdots & \chi_{ms3, c6} \end{bmatrix} \quad (5.25)$$

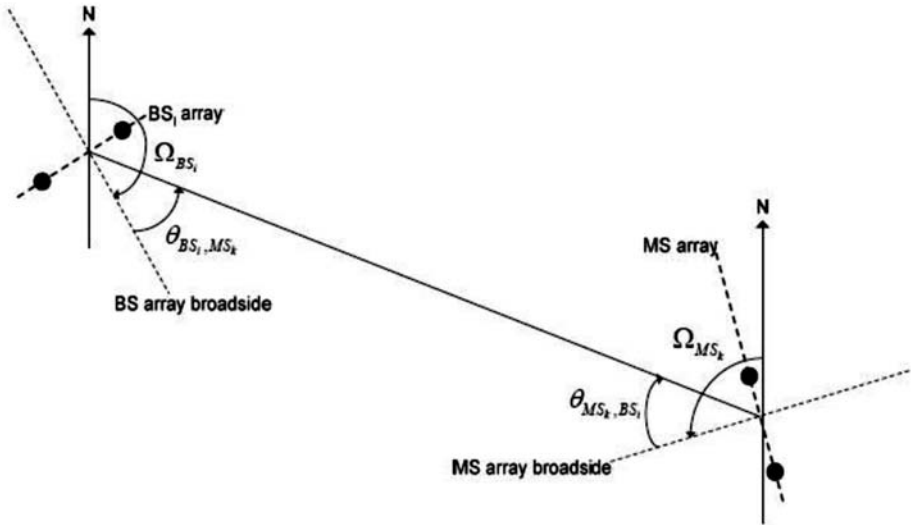


Figure 5.38 BS and MS antenna array orientations [64].

The pairing matrix can be applied to select which radio links will be generated and which will not.

In the following, ULAs [59] are extended to three-dimensional antenna arrays in support of the two-dimensional WINNER channel model. In ULAs, only one dimension is used for the definition of the antenna arrays; in this case, this was the distance between the ULA elements [65].

To introduce three-dimensional antenna arrays, it is necessary to introduce all three spatial dimensions for the definition of the antenna positions inside the array [65, 66]. This is shown in Figure 5.39.

Two coordinate systems should be defined: a *primary* array coordinate system (ACS), which defines the position [Cartesian (x, y, z) or spherical (φ, θ, r) CS] and rotation ($RotX, RotY, RotZ$) of the secondary coordinate system that is assigned to every single antenna element. The *secondary* system, namely, the *element* coordinate system (ECS), is used to define the antenna field patterns. In summary, the field pattern of each antenna element is defined in the ECS, whereas its position and rotation are described in the ACS with six parameters: $x, y, z, RotX, RotY, RotZ$. This is the most general definition. The rotation angles ($RotX, RotY, RotZ$) are defining the sequence of rotations over the x, y , and z axes of the ACS,

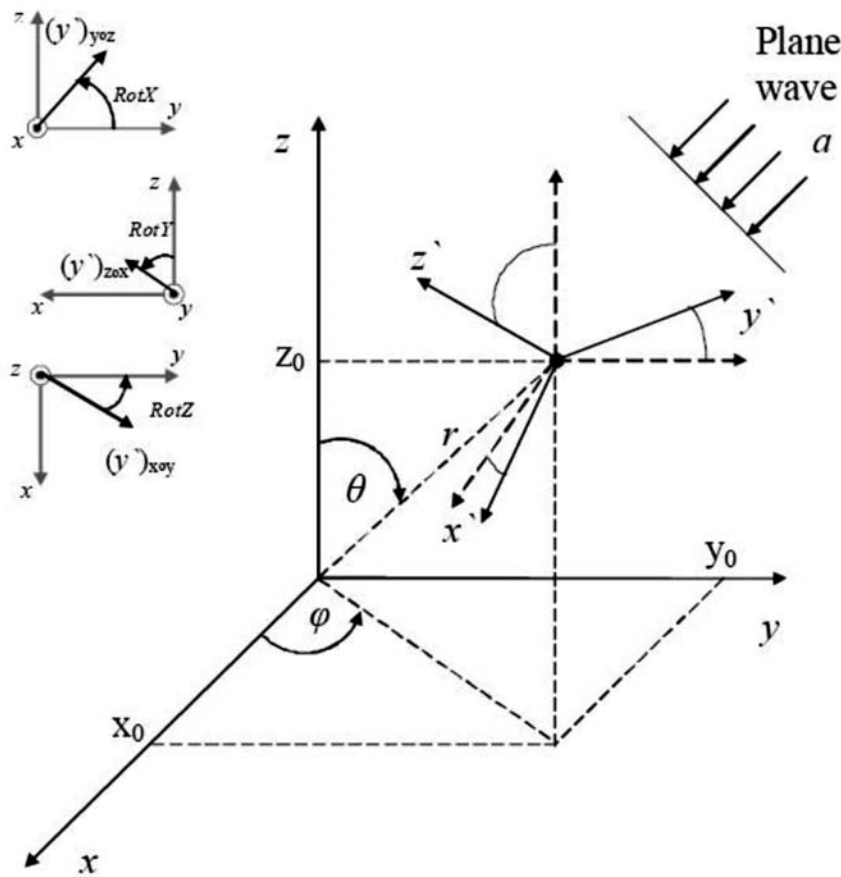


Figure 5.39 A three-dimensional antenna coordinate system [65].

respectively. The sequence of rotations necessary to get the y' axis from the original y axis is also shown in Figure 5.39.

As an example, at the beginning let us assume $(\alpha, \beta, \gamma) = (0^\circ, 0^\circ, 0^\circ)$, which means that there is no tilt of the beam pattern. Further, it is assumed that the waves are planar. The planar wave propagates in the direction \mathbf{a} , as shown in Figure 5.39. If \mathbf{a} is a unit vector and its coordinates in the spherical coordinate system are $(\phi', \theta', 1)$, then in the Cartesian coordinate system \mathbf{a} can be written as follows:

$$\mathbf{a} = \begin{bmatrix} -\sin \theta' \cos \phi' \\ -\sin \theta' \sin \phi' \\ -\cos \theta' \end{bmatrix} \quad (5.26)$$

The “minus” comes from the definition of the directions in Figure 5.39.

The time delay difference (i.e., the difference in arrival time of the received signal at the different antenna elements) defines the phase difference in the received envelopes from one planar wave at different receiving antennas. It is directly proportional to the difference in traveled paths (Δs) of the wave to the different antennas. If the position of one antenna element is $(x_0, y_0, z_0, 0, 0, 0)$, then Δs corresponding to the antenna positioned at $(x_1, y_1, z_1, 0, 0, 0)$ is

$$\Delta s = -(\sin \theta' \cos \phi' \cdot (x_0 - x_1) + \sin \theta' \sin \phi' \cdot (y_0 - y_1) + \cos \theta' \cdot (z_0 - z_1)) \quad (5.27)$$

If the second antenna is placed in the beginning of the coordinate system (the reference point) so that $(x_1, y_1, z_1) = (0, 0, 0)$, then

$$\Delta s = -(\sin \theta' \cos \phi' \cdot x_0 + \sin \theta' \sin \phi' \cdot y_0 + \cos \theta' \cdot z_0) \quad (5.28)$$

In general, for the antenna element at position $\mathbf{r} = [x_0, y_0, z_0]^T$, the following holds:

$$\Delta s = \mathbf{a} \cdot \mathbf{r} \quad (5.29)$$

By using the center of the ACS as a *phase reference*, the phase difference can be calculated for all antenna elements [65].

5.3.2.2 Two-Dimensional Antenna Arrays

In the two-dimensional case, there is one dimension less, which decreases a lot of the complexity of the implementation. For a two-dimensional array, any antenna can be represented with three parameters: parameters x and y (ϕ and r), which define the position of an antenna in the planar array, and one additional parameter $\gamma(\text{RotZ})$, which defines the antenna direction, as shown in Figure 5.40.

In the following, the ULA and two-dimensional array [the uniform circular arrays (UCA) is shown as a possible two-dimensional structure in Figure 5.40] are

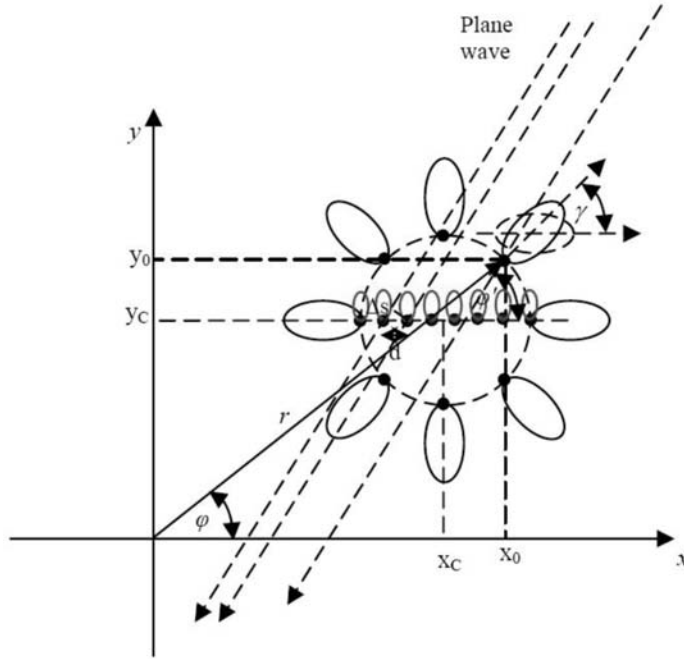


Figure 5.40 A two-dimensional antenna coordinate system [65].

compared. If d is the distance between the adjacent ULA antennas, and ϕ' is the angle between the wavefront direction and the line connecting the elements of the ULA as shown in Figure 5.40, for two adjacent ULA elements the following holds:

$$\Delta s = d * \cos(\phi') \quad (5.30)$$

On the other hand, the traveled distance for two adjacent elements in the two-dimensional space is different. If (x_0, y_0) and (x_1, y_1) are the positions of two receiving antennas, and ϕ' is the wave departure angle as defined in Figure 5.40, then

$$\Delta s = -((x_0 - x_1) * \cos(\phi') + (y_0 - y_1) * \sin(\phi')) \quad (5.31)$$

This result is the same as (5.27) for $\theta' = \pi/2$.

In the implementation, based on the array structure, the referent point must be selected (e.g., for the UCA that would be the center of the circle where the UCA elements are positioned). The relative phase difference to the referent point can be calculated for all of the antenna elements, as shown in Figure 5.41.

The antenna element position can be presented in the polar coordinate system as follows:

$$\mathbf{r} = \begin{bmatrix} r * \cos(\phi) \\ r * \sin(\phi) \end{bmatrix} \quad (5.32)$$

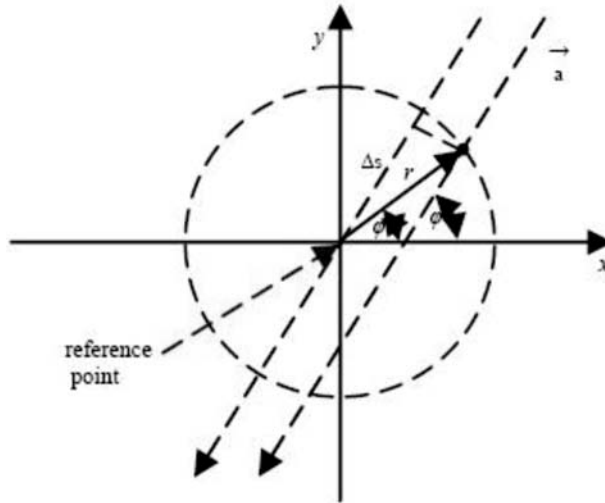


Figure 5.41 Two-dimensional coordinate system with defined local center (0,0) [65].

The unitary vector representing the wave direction can be written

$$\mathbf{a} = - \begin{bmatrix} \cos(\phi') \\ \sin(\phi') \end{bmatrix} \quad (5.33)$$

Then Δs can be calculated as a projection of the vector \mathbf{r} onto the vector \mathbf{a} (similar to the three-dimensional case):

$$\Delta s = \mathbf{r} \cdot \mathbf{a} = -r * \cos(\phi) \cos(\phi') - r * \sin(\phi) \sin(\phi') = -r * \cos(\phi' - \phi) \quad (5.34)$$

In addition to the change in the expression for Δs , the field patterns have to be rotated respectively by $-\gamma$ to represent the real element orientation (see Figure 5.40). This step is the same as in the three-dimensional case.

If the arrays are defined as described earlier, any field pattern can be supported [65]. Also, different field patterns for different antennas can be used. It is realistic for the same antenna elements (the same field patterns) to be used within one antenna array.

5.3.2.3 Special Two-Dimensional Structures

It is possible to implement specific two-dimensional structures (ULA, UCA, URA, and so on). The end user could use these structures much easier and faster than the general structures, as defined earlier. The user interfaces for some structures are defined in the following subsections.

5.3.2.3.1 N-ULA.

For the representation of an N -element ULA (N -ULA), we can assume that all elements are positioned on the x axes. The center of the array is positioned at the

coordinate $(0, 0)$. Therefore, two parameters are required to fully define the ULA: the number of antennas N and the distance between them d , as shown in Figure 5.42.

If N is an odd number, then the center antenna can be placed at the beginning of the coordinate system as shown in Figure 5.42; otherwise, the center of the coordinate system will be between the two central antennas. Based on the N and d , the positions of all antennas $[x_1, x_2, \dots, x_N]$ ($y = \text{zeros}(1, n)$) can be calculated and can become the general input structure for an implementation.

5.3.2.3.2 N -UCA.

The N -element UCA (N -UCA), shown in Figure 5.43, can be defined by the number of antenna elements N , the radius of the circle r , and the offset angle of the first element from the x axis: ϕ_0 .

As in ULA case, we assume that the center of the UCA is placed at the beginning of the coordinate system. Because the array is uniform, the distances and angles between the antenna elements should be the same: $\Delta\phi = 2\pi/N$.

5.3.2.3.3 $M \times N$ -URA.

The $M \times N$ URA is shown in Figure 5.44. The center of the array is at the beginning of the coordinate system.

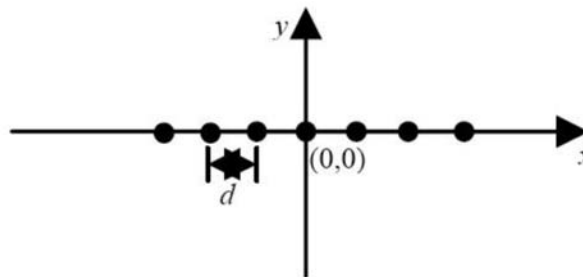


Figure 5.42 N -ULA representation [65].

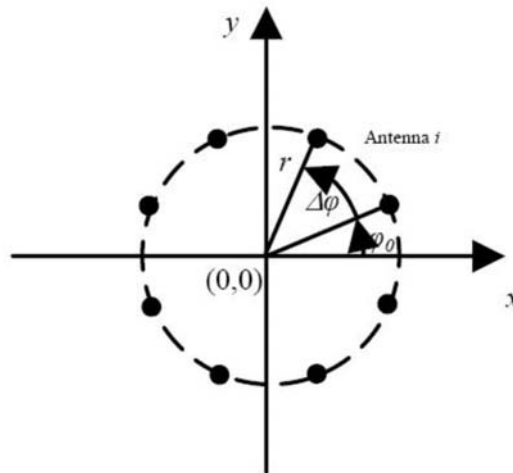


Figure 5.43 N -UCA representation [65].

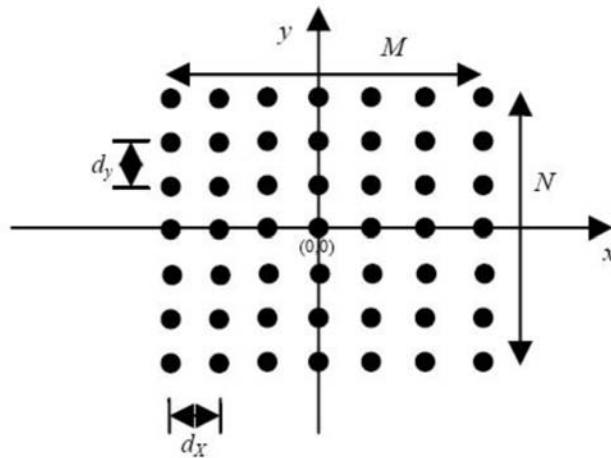


Figure 5.44 The $M \times N$ -URA representation [65].

There are two cases: with an odd number and with an even number of elements per array dimension. The array can be fully defined with M , N , d_x , and d_y .

5.3.2.4 Construction of an Antenna Array

Figure 5.45 shows the blocks for the construction of an antenna array. The example given here is for the three-dimensional case.

The geometry can be defined explicitly using *Pos* and *Rot* arguments followed by an $N \times 3$ matrix, where N is the number of elements. Alternatively, common array types, such as UCA and ULA can be defined with just a few parameters. The way and order in which the arguments defining the geometry are calculated is as follows:

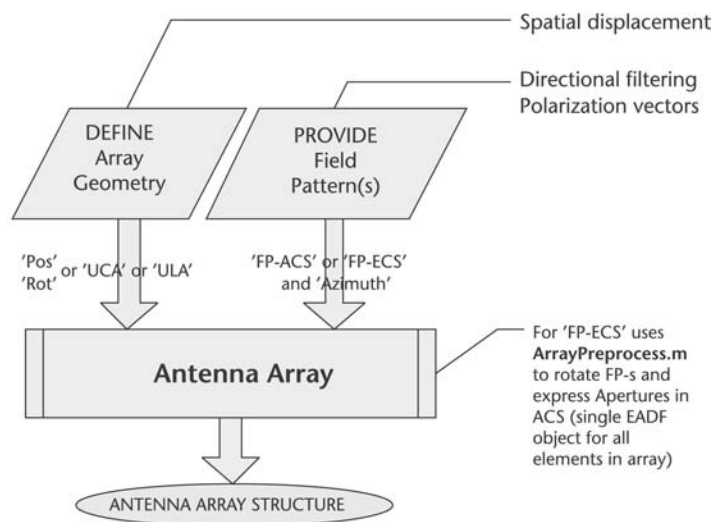


Figure 5.45 Construction of an antenna array.

1. *Pos* has the highest importance and it would be processed first.
2. *Rot* will be taken into account only if *Pos* exists as well. If *Rot* parameters are not provided, the default *Rotation* for the array elements is 0.
3. *UCA* parameters will be processed if *Pos* is not defined.
 - a. The first parameter is the number of antennas (N).
 - b. The second parameter is the radius of the UCA in meters. This parameter is optional, the default value is 1m.
 - c. The elements are placed starting from the x axis ($\phi = 0$) every $\Delta\phi = 2\pi/N$, and the n th element is rotated for $(n - 1)\Delta\phi$ in the positive mathematical direction.
4. *ULA* parameters will be processed if neither *Pos* nor *UCA* is defined.
 - a. The first parameter is the number of antennas (N).
 - b. The second parameter is the distance between antennas in meters. This parameter is optional, the default is $1/N$.
 - c. The elements are placed along the x axis in such a way that the center of the array is at $[0; 0; 0]$ (for an even N there is no antenna element at $[0; 0; 0]$).
5. If there are no parameters defining the geometry, the default is a single antenna positioned at the center of the ACS, without rotation.

The number of elements is also defined implicitly in this phase. The following is an example of the construction of a UCA with 16 elements, and a radius of 5 cm. The field pattern for each element is defined in its own ECS. This requires that each array element be measured separately or that an analytical description (i.e., mathematical expression/function) be used to define the patterns:

```
Array=AntennaArray('UCA',16,0.05,'FP-ECS',FP);
```

The field pattern (FP) should have the following dimensions:

$$\begin{bmatrix} ELNUM & POL & EL & AZ \\ 16 & 2 & 1 & \#AS \end{bmatrix}$$

- The first dimension should fit to the number of elements required for the UCA geometry; alternatively, it can be 1, which means that the same FP is applied to all array elements.
- Both polarizations must be provided for in the FP. If only a single polarization is available, the other dimension must be filled with zeros.
- If the model is two dimensional, only a single (zero) elevation is used, and the third dimension of the FP is equal to 1.
- Typically, the FP is sampled once per angular degree, giving a total number of azimuth samples $\#AS = 360$. We assume that FP samples are taken from the uniform grid. If an optional *Azimuth* argument is provided it should have $\#AS$ elements.

5.3.3 Broadband RF Module Implementation

A broadband RF module capable of laboratory and field testing and implementing the above-described antenna structures was developed within the IST project WINNER. It is shown in Figures 5.46 and 5.47 for the Rx and Tx implementations, respectively [64, 65]. The wideband RF front end makes it possible to use the module with various wireless applications.

The intermediate frequency (IF) center frequency can be tuned to match the different analog-to-digital (ADC) and digital-to-analog converter (DAC) interface frequencies (i.e., baseband to 55 MHz). IF filtering is also tunable. A configuration

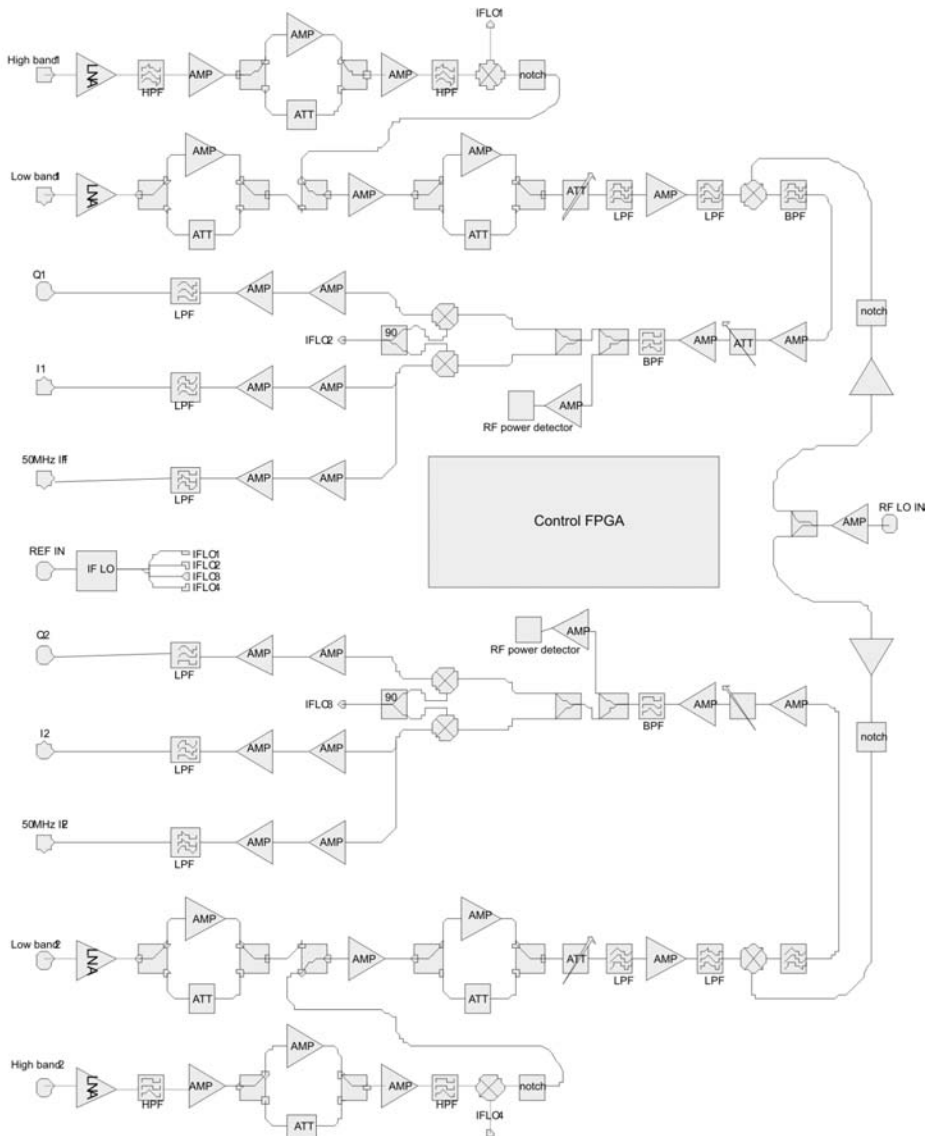


Figure 5.46 RF module Rx diagram [65].

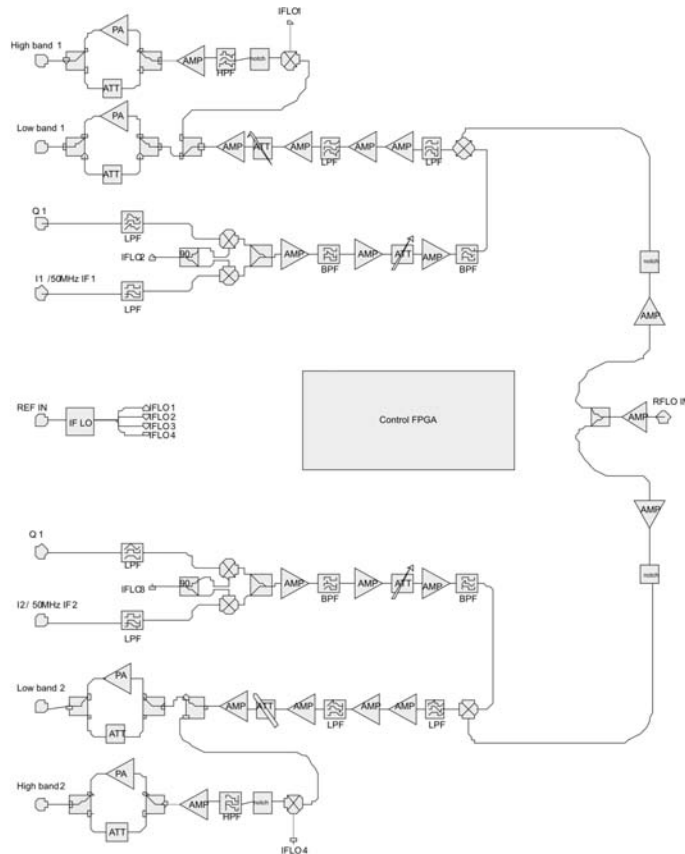


Figure 5.47 RF module Tx diagram [65].

of two channels in one module easily enables small MIMO configurations, because the same LO signal can be fed to both RF chains to avoid phase differences between channels.

The characteristics of the RF interface are as follows:

- Rx Module
 - Two-channel broadband RF receiver;
 - Frequency bands: low band at 300 to 2,700 MHz; high band at 3,500 to 5,800 MHz;
 - IF output frequency: 50 MHz (default);
 - IF BW: 80 MHz;
 - Noise figure (NF): 2.5 dB (low band: 1 GHz);
 - Maximum gain: 75 dB (low band: 1 GHz);
 - External AGC.
- Tx Module
 - Two-channel broadband RF transmitter;
 - Frequency bands: low band of 300 to 2,700 MHz; high band of 3,500 to 5,800 MHz;

- IF input frequency: 50 MHz (default);
- IF BW: 80 MHz;
- Output power: -40 to 0 dBm;
- Power amplifier needed if used for field tests.

A snapshot of the implemented module is shown in Figure 5.48.

5.4 Antenna Architectures for UWB Systems

Antenna architectures for UWB systems were researched and developed within the FP6 IST projects PULSERS and PULSERS II [5]. PULSERS concluded that for AP applications, antennas with high directivity and gain were beneficial. A low-cost array of two elements with planar and three-dimensional reflectors were developed. The maximum dimensions were 89 mm wide and 63 mm long, and the height of the reflector was 40 mm. The gain of the reflector was measured as 9.6 dB, and the reflection coefficient in the frequency range between 3 and 10 GHz was below -10 dB. The planar reflector was found to have similar matching properties but lower gain.

Packaged and size constrained antennas were also investigated. It was concluded that the antenna should be optimized for its environment. The packaging influence on the performance was found to be greater in smaller housings—even a few millimeters around the antenna were found to lessen considerably the impact; also low dielectric constants materials were found to have less impact on the performance than higher ones.

The project designed small quasi-omnidirectional planar antennas for BAN applications, as well as semidirectional ones for BPANs. Their behavior was shown to be promising in terms of their size. A significant BAN measurement campaign was achieved by focusing on the antenna influence and considering real, imperfect

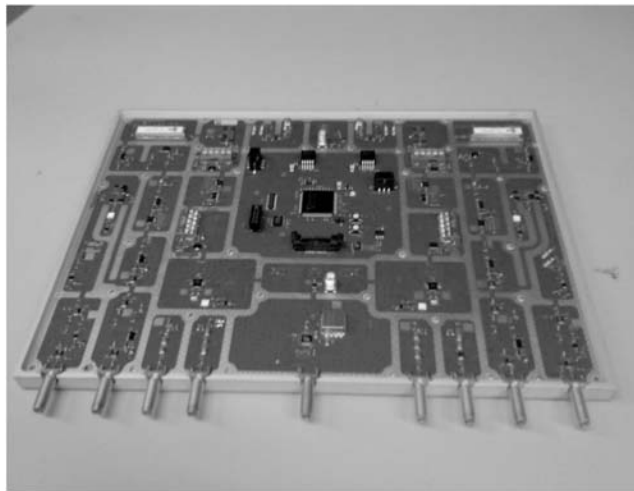


Figure 5.48 The implemented RF module [65].

radiators. Researchers determined in several cases that the matching level or input bandwidth can be improved by antennas able to couple significantly close to the body. Small planar antennas were also affected by body proximity even if these showed good performance in isolated conditions.

All results were used in a statistical antenna design in which the parameters of the antennas were chosen to ensure the best results for the PULSERS scenarios.

5.4.1 Main Transceiver Principles

Considering the operating conditions envisioned for UWB signals (SNR, in particular), the loss on the link budget caused by this extreme way of sampling the incoming signal was found to be asymptotically marginal [67, 68]. In the PULSERS II solution, this idea was extended by adding the presence of signal information (called an “energy bit”) to the already available polarity information [69]. The main purpose was to deliver final specifications and implementation results with respect to the PHY and MAC layers integrated in an impulse radio–ultrawide-band (IR-UWB) low data rate–location and tracking (LDR-LT) system.

Researchers have shown that hybrid transceiver architectures could benefit from both coherent and noncoherent approaches [68]. More particularly, the described DBPSK-oriented Rx architecture advantageously combined differential demodulation and preamble detection schemes with a direct 1-bit subsampling of the incoming RF signal. The first assumption driving the idea for the addition of the presence of signal information was that the main contribution to the noise comes from the electronics itself instead of the channel. Thus, the objective was to design a circuit that would be as immune as possible to its own noise. The principle adopted in [69] was to consider that if the same signal went through the circuit following separate paths, it would be affected by uncorrelated noises.

Secondly, the gain required to sample UWB signals using energy detection techniques was so high that it would be very difficult to design stable linear amplification chains. Thus, the idea relied on performing energy detection on signals extremely distorted by the amplification chain. The principle was to perform the first part of the amplification using an analog amplification chain by the use of low noise signals (LNAs), and the remaining gain was obtained by saturating the signal over 1 bit using cascaded digital inverters.

Thus, the energy detection module was a digital comparison between two signals taking value -1 or 1 . When -1 is mapped to logic value 0 and 1 is mapped to logic value 1 , the resulting logic operator is an XNOR gate. If each branch is considered separately, the output signal is the polarity of the incoming signal. After remapping and multiplication of one branch output with the XNOR output, a three-level signal $\{-1; 0; 1\}$, also called a “1.5-bit” signal, can be obtained. This is shown in Figure 5.49 where a differential antenna has been employed.

This general principle enables two different kinds of modulation: noncoherent N-PPM and differentially coherent DBPSK. The noncoherent N-PPM demodulation can be performed using the *Energy* output alone. Assuming synchronization (i.e., the knowledge of the beginning and end of each PRP), the XNOR output is split into N distinct pulse repetition intervals (PRIs). For each PRI, the 1-bit signal is integrated within a duration T_i . Once integrated, the N energetic values are com-

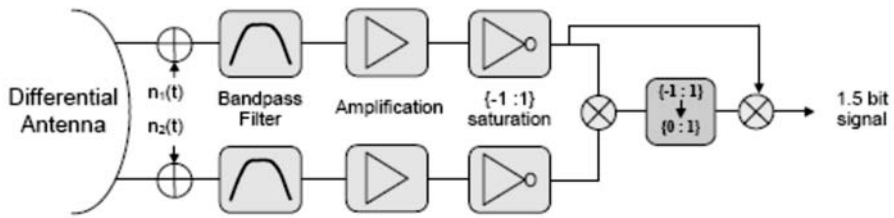


Figure 5.49 The 1.5-bit conversion principle [69].

pared. The decision is made by considering the PPM temporal position that provides the most significant amount of collected energy among all possible positions. This is shown in Figure 5.50.

The proposed principle can also handle a DBPSK demodulation scheme by combining the XNOR output with the *Sign* output, as mentioned previously. After mapping, the demodulation is a classical DBPSK demodulation scheme. Assuming synchronization, the incoming signal sample is multiplied by the corresponding sample in the previous PRP, and the result of the product is then integrated over a duration T_i . If the result is positive, the decision is “1;” otherwise, it is “-1.”

The implemented pulse generator architecture is shown in Figure 5.51.

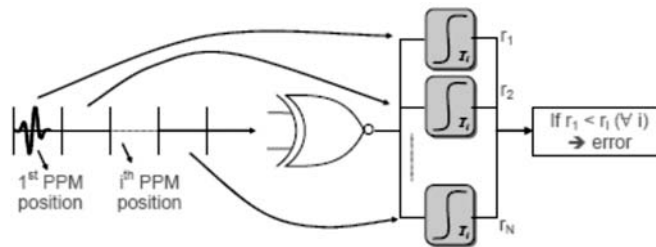


Figure 5.50 N-PPM noncoherent demodulation [69].

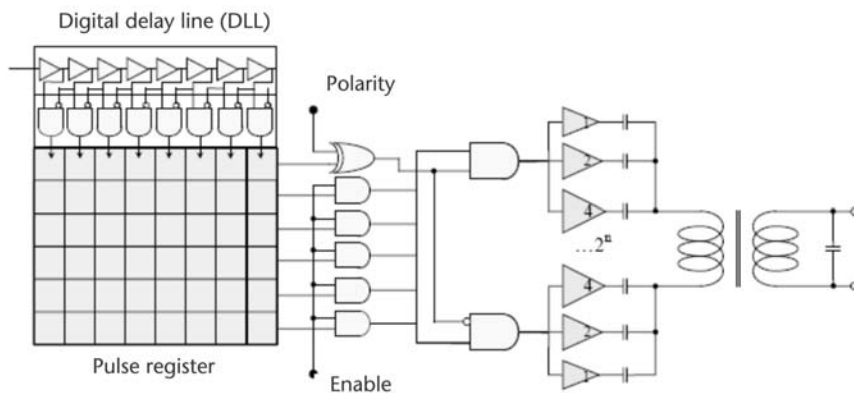


Figure 5.51 Pulse generator architecture [69].

Although the implemented RF architecture handled both energy detection/PPM and (D)BPSK modulations, the modulation scheme chosen for the PULSERS II demonstrator was based on DBPSK. This modulation scheme was interesting because it provided good immunity to clock drift and did not require any channel estimating. Demodulation of the incoming signal was performed using differential correlation between the incoming signal corresponding to the current data symbol and the previous one.

To improve the SNR before making a decision about the current symbol, redundancy was added to the current symbol by repeating it several times, either through the coherent integration of sampled signals before differential correlations (IC), or through the posterior accumulation of correlation results afterwards (ACC). To avoid spectral spikes, the repeated pulses energy was also spread using a PN code [69].

The transmitted signal used in the performance analysis of PULSERS was a pulse of 1-GHz bandwidth (-10 -dB bandwidth) centered on 4 GHz. The reference channel models used were IEEE 15.3a channels 1 and 4 [70], respectively, depicted as CM1 and CM4 in Figure 5.51. Channel model 1 is a good representative of indoor configurations, for which the transmitted signal has limited interaction with the propagation environment. CM1 channels are defined as LOS, with a limited transmission range (less than 10m), and the corresponding CIR is spread over 20 ns. The integration window duration for this kind of channel was chosen as 20 ns. Channel model 4, on the other hand, is a good representative of indoor transmission channels, for which the transmitted signal is blocked and spread by the environment. This results in a CIR where the first path is severely attenuated or even undetectable. In addition, the time spread of this CIR is generally significant. A common value for the useful part of the CIR is 60 ns. This value was used in PULSERS II as the length of the integration window for CM4 channels.

The performance analysis considered, among others, the effect induced by the XNOR gate. A real gate has a limited bandwidth. This bandwidth limitation can be modeled as the combination of a pure lowpass filter and a hysteresis factor. The effect of such modeling is shown in Figure 5.52.

The resulting degradation observed on 1.5-bit signals leads to further degradations of the achievable BER performance. This is shown in Figure 5.53.

Another aspect to consider is the amplification chain. Figure 5.54 shows how the noise figure of the LNA is reflected by an additional performance loss of around 5 dB.

The analog front-end operating principle relies on a signal split over two separate paths that are isolated as much as possible from each other in order to

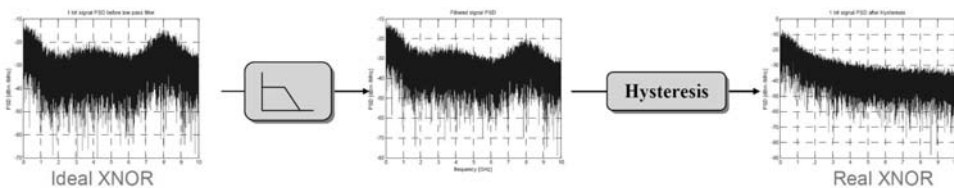


Figure 5.52 Effect of a realistic gate model on the output spectrum [69].

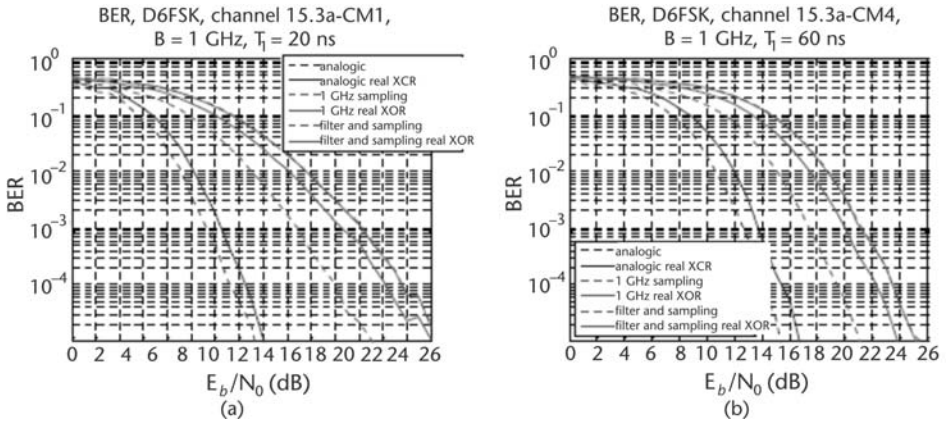


Figure 5.53 (a, b) Influence of the XNOR gate [69].

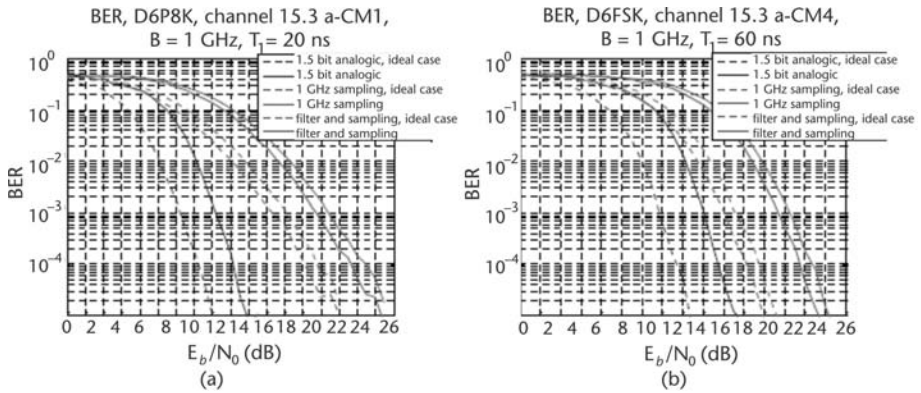


Figure 5.54 (a, b) Influence of the amplification chain [69].

be unaffected by uncorrelated noises. Following this idea, the receiver can be connected to a differential input antenna.

The little performance improvement introduced by the use of such an antenna in the ideal case is completely masked when all previous imperfections are considered. This is shown in Figure 5.55. Nevertheless, differential antennas have other advantages for miniaturization purposes (e.g., no need for a ground plane and better isotropy). In the PULSERS II real demonstrator, real differential antennas were used.

5.4.2 RF Front End and Antenna Implementation and Experiments

5.4.2.1 Differential Antennas

A UWB differential antenna was designed for the 4.25- to 4.75-GHz frequency band to be tested in the PULSERS II project [69]. The first serious requirement was to preserve the differential access so that the concepts described previously would still apply. Another constraint concerned the radiating pattern that was expected to be as omnidirectional as possible, despite the vicinity of the RF circuit

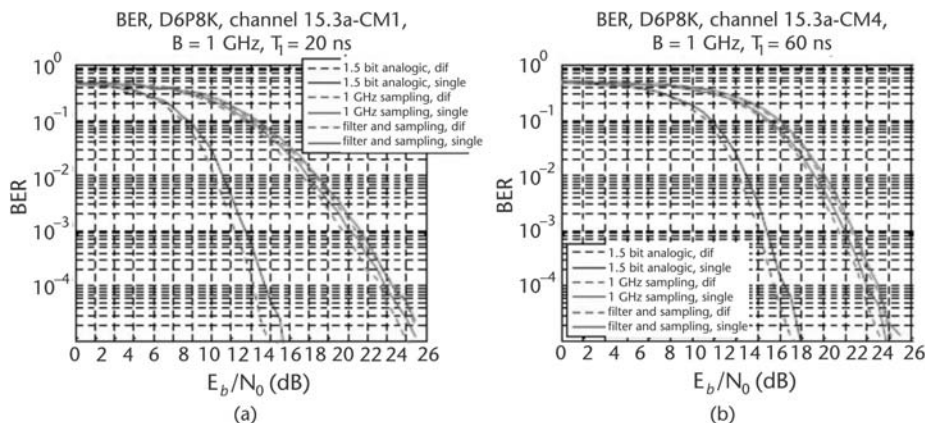


Figure 5.55 (a, b) Influence of the antennas [69].

board and other metallic objects (e.g., the battery) in the reactive near-field area (i.e., 1-cm radius sphere) of the antenna. Note that a filtering version of this differential antenna might help to enhance the link budget.

As a first step in the investigation, the differential antenna was designed for a free-space environment. A structure equivalent to a flat bifilar transmission line was figured out to drive the differential mode with a 50Ω characteristic impedance, from the RF circuit to the radiating structure (see Figure 5.56).

A microwave, low-loss RO4003 substrate was also used to guarantee a good radiating efficiency.

In the second phase of the study, researchers decided to deal with the impact of the vicinity of the circuit, which could not respect the symmetry of the differential antenna any more. The differential mode could be perturbed by the asymmetrical distribution of surface currents on the antenna structure. Thus, a parametric study was carried out with an appropriate electromagnetic (EM) simulator to evaluate the minimal distance separating the antenna from the circuit board. A separation distance of 1 cm was then chosen.

The process used to measure the reflection coefficient on a differential 50Ω antenna with a single mode vector network analyzer is described in [71]. Figure

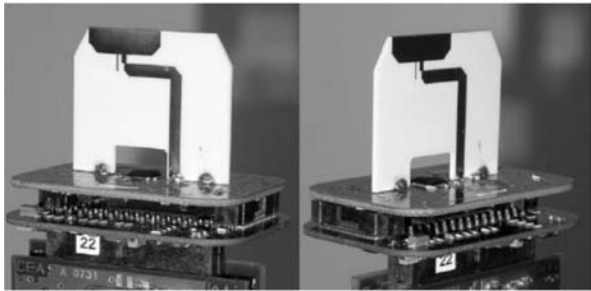


Figure 5.56 Pictures of the differential antenna mounted on top of the PHY board of a PULSERS demonstration platform [69].

5.57 shows a satisfying behavior ($|S_{11}| \text{ dB} < -9.5 \text{ dB}$) over the specified 4.25- to 4.75-GHz band when the antenna is measured in a free-space scenario.

When the circuit board is positioned at a distance of 1 cm from the bottom of the antenna, the impedance matching is degraded in a worst case around 4.3 GHz, for which $|S_{11}|$ is up to -5 dB .

The vicinity perturbation is more important for the lower frequency components. Beyond 4.5 GHz, the metallic board is no longer in the reactive zone and its influence is negligible. The observed $|S_{11}|$ perturbation, however, only results in a worst case mismatch loss of about 1.7 dB.

The results on impedance matching, the use of low-loss substrate, and the size of the antenna imply that the radiating efficiency should be better than 60%. The radiating pattern was not characterized. However, in relation to the $|S_{11}|$ perturbation, some ripples can be expected on the omnidirectional pattern in the azimuth plan. The effect of this discrepancy on the radio link can be smoothed by the presence of numerous multipath components in typical indoor environments.

5.4.2.2 Analog RF Front End

The proposed PULSERS II RF front end implemented a complete impulse radio transceiver. On the Rx side, it was composed of an amplification chain followed by a high-speed ADC. The proposed approach consisted of sampling the incoming signal on 1 bit so as to reduce the power consumption and use the redundancy and the UWB channel temporal diversity to compensate for the loss of amplitude information.

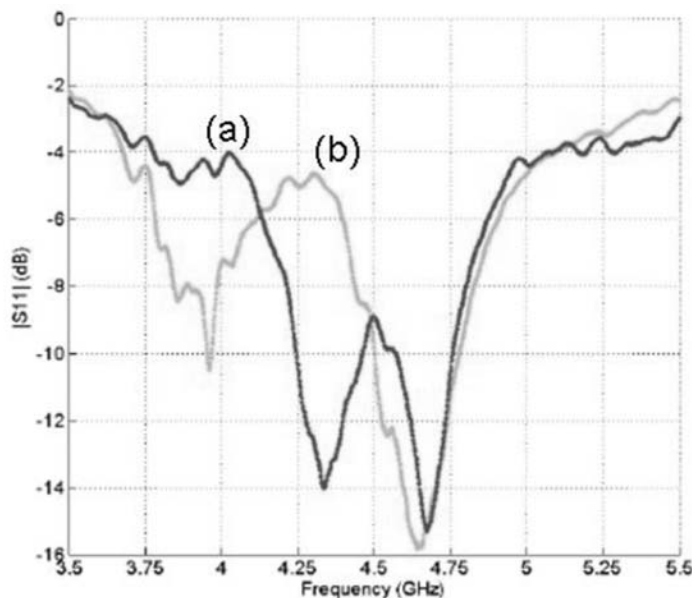


Figure 5.57 Reflection coefficient of the matching impedance measured in (a) free space and (b) with metals close by [69].

On the Tx side, a pulse memory followed by a high-speed DAC was used. This enables the generation of any arbitrary pulse shape within the desired RF bandwidth.

The signal energy information was obtained by duplicating the amplification chain over two different paths, thus helping to discriminate samples affected by the noise generated by the circuit itself.

The RF source was a differential antenna driving the circuit inputs by a 50Ω differential signal as described in Section 5.4.2.1. Note that such an antenna is not absolutely required but it can help to improve the link budget by 1 or 2 additional decibels. Furthermore, differential antennas are more isotropic than single-ended ones [69].

The input interface acts as a signal splitter, allowing the RF signal to be applied on two identical amplification channels. It is also in charge of input impedance matching with the antenna. The amplification chain was composed of a first LNA followed by several amplification stages optimized for gain and common-mode rejection. This amplification chain provided 80 dB of gain per channel. The last stages of amplification saturated the amplified signal in such a way that the resulting signal is quantified on 1 bit.

The inputs of the XNOR gate are driven by two 1-bit signals representing the sign of the RF signal affected by uncorrelated noise signals. The XNOR gate operates as a multiplier. Its operands are two-level numbers (−1, +1) respectively mapped to logic values 0 and 1. The result accounts for the instantaneous “power” of the signal on 1 bit (0, +1).

The analog RF implemented architecture is shown in Figure 5.58.

This circuitry has exactly the same effects on the spectrum of the outgoing 1-bit signal as a classical mixer. In the frequency domain, the spectrum has a lobe located near 0 GHz and another one near 9 GHz. If a pure energy detection scheme (e.g., an N-PPM modulation scheme) is envisioned with this front end, an antialiasing filter can be activated in order to remove this high-frequency compo-

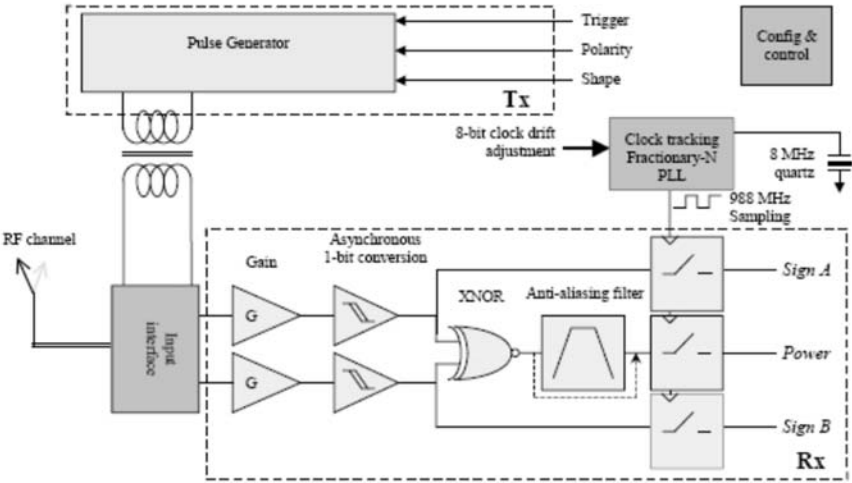


Figure 5.58 Implemented analog RF front end architecture [69].

nent. In the case described here, this filter is not used because it would degrade the performance with the 1.5-bit operating mode where the transposition of the signal to the baseband frequency domain is achieved thanks to the spectrum folding caused by subsampling.

In the implementation described here, the sampling clock is provided by a fractionary- N phase-locked loop (PLL). This PLL enables clock drift compensation with a precision down to 1 ppm and can be directly locked on a quartz reference.

This analog front end can be used for energy detection purposes as well. The described prototype chip initially had been designed to enable both operating modes and to perform preliminary tests for further developments aiming at a more strict compliance with IEEE 802.15.4a demodulation schemes. In particular, it could help to evaluate the use of 1-bit pure energy detection and thus open the door to extremely low-cost, low-power implementations. For this purpose, additional test modules were included within the PULSERS circuit. Full details can be found in [67–69]. The final implemented transceiver architecture of the complete chip is shown in Figure 5.59.

5.5 Conclusions

This chapter described theoretical concepts of spatial processing and some practical questions related to the implementation of antenna techniques in a radio system. Questions concerning system complexity, measurements, feedback, control overhead, robustness against impairments, and flexibility/scalability (with respect to

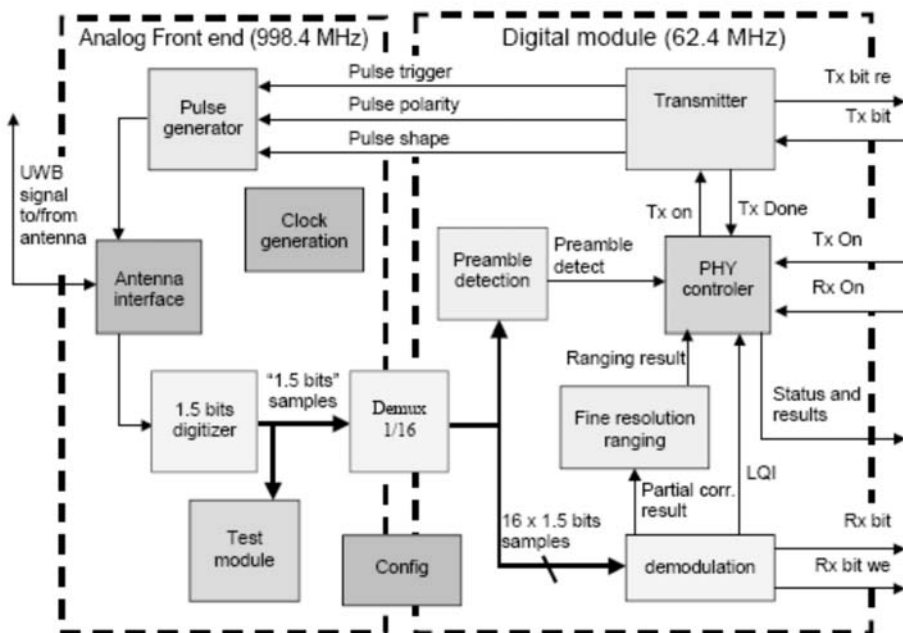


Figure 5.59 Overall transceiver architecture including test modules [69].

the terminal capability, number of transmitting and receiving antennas, and propagation conditions) were addressed and reflected.

Radio propagation has a significant impact on the performance of wireless communication systems. The impact on future broadband systems is even more important due to the increased data rate, bandwidth, mobility, adaptivity, QoS, and so forth. Because of the major influence on system performance and complexity, future radio channel models and simulations have to be more versatile and accurate than earlier systems. Antenna implementations and techniques are crucial to ensure the correct channel modeling and estimations.

The FP6 IST projects provided significant results in the area of multiple antennas and enabling technologies. Research and implementation spanned not only the single-link propagation environments, but also multihop scenarios. Significant results were achieved for the implementation and usage of antennas suitable for short-range technologies and VHDRs.

The design of UWB antennas is particularly challenging. On one hand, UWB antennas must provide good performance over very large bandwidths; on the other hand, these are limited by size constraints. Antenna array measurements in an anechoic chamber can provide relevant information for the design of UWB arrays.

Novel link adaptation, medium access control, and routing algorithms are able to exploit the benefits of multiple-antenna usage. Such approaches will be discussed in Chapters 6 and 7 of this book. For particular studies in the area of short-range technologies and ad hoc networks, the reader is referred to the volume in this series on ad hoc networks.

Employment of smart antennas techniques without a thorough understanding and consideration of their interaction with other layers can prove to be counterproductive with respect to the overall radio interface system concept.

References

- [1] FP6 Overview, <http://cordis.europa.eu/ist/so/mobile-wireless/home.html>.
- [2] FP6 IST Projects Wireless Interface New Radio, WINNER and WINNER II, www.ist-winner.org.
- [3] FP6 IST Project OBAN, oban.tubit.tu-berlin.de.
- [4] FP6 IST Projects MAGNET and MAGNET Beyond, www.ist-magnet.ist.org.
- [5] FP6 IST project PULSERS, www.pulsers.eu.
- [6] FP6 Project MEMBRANE, www.imperial.ac.uk/membrane.
- [7] FP6 IST Project OBAN, "Emerging Technologies," Deliverable 2.1, May 2006, oban.tubit.tu-berlin.de.
- [8] IEEE Standards, "WWiSE Proposal: High Throughput Extension of the 802.11 Standard," IEEE 802.11-05/0149r1, January 2005.
- [9] IEEE Standards, "TGn Sync Proposal Technical Specification," IEEE 802.11-04/0889r6, May 2005.
- [10] Alamouti, S. M., "A Simple Transmit Diversity Technique for Wireless Communications," *IEEE J. on Selected Areas in Communications*, Vol. 16, No. 8, October 1998, pp. 1451–1458.
- [11] Jankiraman, M., *Space-Time Codes and MIMO Systems*, Norwood, MA: Artech House 2004.

- [12] Lehne, P. H., and M. Pettersen, "An Overview of Smart Antenna Technology for Mobile Communication Systems," *IEEE Communications Surveys*, Vol. 2, No. 4, Fourth Quarter 1999, pp. 2–13.
- [13] Alexiou, A., and M. Haardt, "Smart Antenna Technologies for Future Wireless Systems: Trends and Challenges," *IEEE Communications Magazine*, September 2004, pp. 90–97.
- [14] Winters, J. H., "On the Capacity of Radio Communication Systems with Diversity in a Rayleigh Fading Environment," *IEEE J. on Selected Areas in Communications*, Vol. SAC-5, June 1987, pp. 871–878.
- [15] FP6 IST Project WINNER, "Assessment of Advanced Beamforming and MIMO Technologies," Deliverable 2.7, February 2005, www.ist-winner.org.
- [16] FP6 Project PULSERS, "Antenna Design Issues," Deliverable 4.2, June 2005, www.pulsers.eu.
- [17] FP6 Project MAGNET, "Specifications and Requirements for PAN Channel Measurements and Modeling," Deliverable 3.1.1, March 2004, www.ist-magnet.org.
- [18] FP6 Project MAGNET, "PAN Channel Characterization," Deliverable 3.1.2a, October 2004, www.ist-magnet.org.
- [19] FP6 Project PULSERS Phase II, "Study of MIMO VHDR MB-OFDM Specific System Concepts," Deliverable2a-3.1, December 2007, www.pulsers.eu.
- [20] Molisch, A., M. Win, and J. Winters, "Space-Time-Frequency (STF) Coding for MIMO-OFDM Systems," *IEEE Communication Letters*, September 2002, pp. 370–373.
- [21] Foschini, G., and M. Gans, "On Limits of Wireless Communications in a Fading Environment When Using Multiple Antennas," *Wireless Personal Communications*, March 1998, pp. 311–335.
- [22] Telatar, I., "Capacity of Multi-Antenna Gaussian Channels," Bell Labs Technical Memorandum, 1995.
- [23] Boelskei, H., D. Gesbert, and A. Paulraj, "On the Capacity of Wireless Systems Employing OFDM-Based Spatial Multiplexing," *IEEE Trans. on Communications*, February 2002, pp. 225–234.
- [24] Liberti, J. C., and T. S. Rappaport, *Smart Antennas for Wireless Communications*, Englewood Cliffs, NJ: Prentice Hall, 1999.
- [25] Paulraj, A. J., and C. B. Papadias, "Space-Time Processing for Wireless Communications," *IEEE Signal Processing Magazine*, November 1997, pp. 49–83.
- [26] Tarokh, V., N. Seshadri, and A. R. Calderbank, "Space-Time Codes for High Data Rate Wireless Communication: Performance Criterion and Code Construction," *IEEE Trans. on Information Theory*, Vol. 44, March 1998, pp. 744–765.
- [27] Shiu, D., et al., "Fading Correlation and Its Effect on the Capacity of Multi-Element Antenna Systems," *IEEE Trans. on Communication*, Vol. 48, March 2000, pp. 502–513.
- [28] Sampath, H., and A. Paulraj, "Linear Precoding for Space-Time Coded Systems with Known Fading Correlations," *IEEE Communication Letters*, June 2002, pp. 239–241.
- [29] Brunner, C., J. Hammerschmidt, and J. Nossek, "Downlink Eigenbeam-Forming in WCDMA," *Proc. of European Wireless 2000*, Dresden, September 2000.
- [30] Jöngren, G., M. Skoglund, and B. Ottersten, "Combining Beamforming and Orthogonal Space-Time Block Coding," *IEEE Trans. on Information Theory*, March 2002, pp. 611–627.
- [31] Borgmann, M., and H. Bölcskei, "Noncoherent Space-Frequency Coded MIMO-OFDM," *IEEE J. on Selected Areas in Communications*, April 2004.
- [32] Giese, J., and M. Skoglund, "Space-Time Constellation Design for Unknown Frequency-Selective Channels," *Proc. of the International Conference on Communications*, Anchorage, AK, May 2003.
- [33] Goldsmith, A. J., and S. G. Chua, "Adaptive Coded Modulation for Fading Channels," *IEEE Trans. on Communications*, Vol. 46, No. 5, May 1998.

- [34] Goldsmith, A., et al., "Capacity Limits of MIMO Channels," *IEEE J. on Selected Areas in Communications*, Vol. 21, No. 5, June 2003.
- [35] Boelcskei, H., M. Borgamn, and A. J. Paulraj, "Space-Frequency Coded MIMO-OFDM with Variable Multiplexing-Diversity Tradeoff," *Proc. of IEEE International Conference on Communications*, Anchorage, AK, May 2003, pp. 2837–2841.
- [36] Jaldén, J., and B. Ottersten, "On the Complexity of Sphere Decoding in Digital Communications," *IEEE Trans. on Signal Processing*, 2004.
- [37] Viterbo, E., and J. Boutros, "A Universal Lattice Code Decoder for Fading Channels," *IEEE Trans. on Information Theory*, July 1999, pp. 1639–1642.
- [38] Damen, O., A. Chkeif, and J. C. Belfiore, "Lattice-Code Decoder for Space-Time Codes," *IEEE Communication Letters*, May 2000, pp. 161–163.
- [39] Hochwald, B. M., and S. Ten Brink, "Achieving Near-Capacity on a Multiple-Antenna Channel," *IEEE Trans. on Communications*, March 2003, pp. 389–399.
- [40] Wolniansky, P. W., et al., "V-BLAST: An Architecture for Realizing Very High Data Rates over the Rich-Scattering Wireless Channel," *Proc. of ISSSE 98*, Pisa, Italy, September 1998.
- [41] Viswanath, P., D. N. C. Tse, and R. Laroia, "Opportunistic Beamforming Using Dumb Antennas," *IEEE Trans. on Information Theory*, June 2002.
- [42] Rose, C., S. Ulukus, and R. D. Yates, "Wireless Systems and Interference Avoidance," *IEEE Trans. on Wireless Communications*, Vol. 1, No. 3, July 2002, pp. 415–428.
- [43] Yu, W., et al., "Iterative Water-Filling for Gaussian Vector Multiple Access Channels," *IEEE Trans. on Information Theory*, January 2004.
- [44] Zelst, A., and T. Schenk, "Implementation of a MIMO OFDM-Based Wireless LAN System," *IEEE Trans. on Signal Processing*, February 2004, pp. 483–494.
- [45] FP6 IST Project OBAN, "Emerging Technologies," Deliverable 2.1, May 2006, oban.tubit.tu-berlin.de.
- [46] Foschini, G. I., "Layered Space-Time Architecture for Wireless Communications in Fading Environments When Using Multi-Element Antennas," *Bell Labs Technical J.*, pp. 41–59, 1996.
- [47] Godara, L. C., "Application of Antenna Arrays to Mobile Communications. Part 1: Performance Improvement, Feasibility, and System Considerations," *Proc. of IEEE*, Vol. 85, July 1997, pp. 1031–1060.
- [48] FP6 IST Project OBAN, "Scenarios and Wireless Performance and Coverage," Deliverable 8, May 2005, oban.tubit.tu-berlin.de.
- [49] Kuzminskiy, A. M., et al., "Interference Scenarios in Future Wireless Open Access Networks," *Proc. of 11th WWRF*, June 2004.
- [50] Belkin, www.belkin.com.
- [51] Motia, www.motia.com.
- [52] Erceg, V., et al., "TGn Channel Models (IEEE 802.11-03/940r2)," High Throughput Task Group, IEEE P802.11, March 2004.
- [53] Schumacher, L., "WLAN MIMO Channel MATLAB Program," www.info.fundp.ac.be/~lsc/Research/IEEE_80211-HTSG-CMSC/distribution_terms.html.
- [54] Zettenberg, P., et al., "Performance of Multiple Receive Multiple-Transmit Beamforming in WLAN-type Systems Under Power or EIRP Constraints with Delayed Channel Estimates," *Proc. of VTC*, 2002, pp. 1906–1910.
- [55] Boyd, S., and L. Vandenberghe, *Convex Optimization*, New York: Cambridge University Press, 2004.
- [56] Abramovich, Y. I., and M. B. Sverdlik, "Some Optimization Problems Based on the Minimax Criterion for Antenna Array Beamforming," *Radio Engineering Electronics and Physics*, 1974, No. 10, pp. 2176–2178.

- [57] Kuzminskiy, A. M., "Increasing the Range of Access Point Cells for a Given Throughput in a Downlink of a Wireless Local Area Network," Application for U.S. Letters Patent, 2100.021900, filed September 30, 2005.
- [58] "UK Interface Requirement 2006 Short Range, Broadband, Data Services Operating in the Frequency Range 5150–5725 MHz," 98/34/EC, 2002/0254/UK, October 2002.
- [59] FP6 IST Project WINNER, "Determination of Propagation Scenarios," Deliverable 5.2, June 2004, www.ist-winner.org.
- [60] Fleury, B., X. Yin, and A. Kocian, "Impact of the Propagation Conditions on the Properties of MIMO Channels," *Proc. of the International Conference on Electromagnetics in Advanced Applications (ICEAA 2003)*, Turin, Italy, September 2003.
- [61] Heddergott, R., U. Bernhard, and B. Fleury, "Stochastic Radio Channel Model for Advanced Indoor Mobile Communications Systems," *Proc. of IEEE PIMRC*, Helsinki, September 1997, pp. 140–144.
- [62] Pedersen, K., P. Mogensen, and B. Fleury, "A Stochastic Model of the Temporal and Azimuthal Dispersion Seen at the Base Station in Outdoor Environments," *IEEE Trans. on Vehicular Technology*, Vol. VT-49, No. 2, March 2000, pp. 437–447.
- [63] Fleury, B., et al., "High-Resolution Channel Parameter Estimation for Communication Systems Equipped with Antenna Arrays," *Proc. of 13th IFAC Symposium on System Identification ((SYSID 2003))*, Rotterdam, August 2003.
- [64] FP6 IST Project WINNER, "Final Report on Link Level and System Level Channel Models," Deliverable 5.4, November 2005, www.ist-winner.org.
- [65] FP6 IST Project WINNER II, "Report on Validation and Implementation of the WINNER Physical Layer," Deliverable 6.12.2, November 2007, www.ist-winner.org.
- [66] FP6 IST Project WINNER II, "WINNER II Channel Models," Deliverable 1.1.2, September 2007, www.ist-winner.org.
- [67] IST PULSERS Phase II, "LDR-LT Detailed Architecture Specifications—PHY and MAC Layers," Deliverable 3a-3.2, March 2008, www.pulsers.eu.
- [68] IST PULSERS Phase II, "LDR-LT Concept Specifications—PHY and MAC Layers," Deliverable 3a-3.1, January 2007, www.pulsers.eu.
- [69] IST PULSERS Phase II, "LDR-LT Implementation and Test Report—PHY and MAC Layers," Deliverable 3a-3.3, October 2008, www.pulsers.eu.
- [70] Foerster, J., et al., "IEEE 802.15.3a Channel Modeling Subcommittee Report," Technical Report IEEE P802.15-02/490r1-SG3a, February 2003.
- [71] Meys, R., and F. Janssens, "Measuring the Impedance of Balanced Antennas by an S Parameter Method," *IEEE Antennas and Propagation Magazine*, Vol. 40, December 1998, pp. 62–65.

Coverage Enhancement Technologies

A fundamental challenge in radio interface design is to introduce functionalities that allow for maximum possible coverage while keeping deployment costs down. Technologies to increase coverage, in addition to the ones discussed in the previous chapters (e.g., smart-antenna techniques, adaptive transmission techniques), make use of multihop capabilities by means of flexible relay-based deployments, mesh network architectures, and cooperative communications.

The concept of flexible relay-based cell deployment distributes the high capacity available at the inner cell area so as to provide a capacity boost in the outer areas. Advanced PHY layer techniques, such as cooperative relaying, that leverage the relay-enhanced deployment can additionally increase the performance of such relay-enhanced cells.

A number of FP6 IST projects [1] undertook research and investigation of solutions for multihop communication systems. The FP6 IST projects WINNER and WINNER II [2] investigated relay technologies as a means of ensuring successful deployment at low costs, while supporting the ubiquitous coverage capabilities of the novel WINNER radio interface. The FP6 IST project FIREWORKS [3] introduced flexible relay-based deployments to OFDMA-based WMAN/WLAN technologies. The FP6 project MEMBRANE [4] developed intelligent antenna techniques in support of multihop networks. Some of these results were discussed in Chapter 5 of this book.

This chapter describes advances in the area of relay technologies as a way to enhance the capabilities of different radio interfaces. It builds on achievements of the WINNER, WINNER II, and FIREWORKS projects.

Section 6.1 introduces the motivation behind the use of relay technologies. Section 6.2 describes proposed RRM schemes suitable for relay-based deployments. Section 6.3 investigates the benefits of adaptive antenna techniques in relay-based scenarios. Section 6.4 concludes the chapter.

6.1 Introduction

Broadband radio interface technologies with high multiplexing bit rates are characterized by unevenly distributed QoS characteristics, depending on the distance between the terminal and the BSs, and very limited range due to their high sensitivity against interference, attenuation, shadowing, the possibly high spectrum bands in the scope of next generation radio systems (2.8 to 5 GHz), and limited transmission power [5].

One of the most important factors for HDR communication is the existence of low path loss. In existing cellular systems, obstacles such as buildings and hilly terrains introduce high path loss, which affects the communication robustness and throughput. An improvement would be possible by using advanced antennas and setting up a high number of picocellular BSs, but that would result in high deployment costs. A more cost-efficient and innovative solution is to trade capacity against range where the capacity of a BS is partially utilized by a number of relay stations acting as wireless BSs in some desirable configuration to divide the path into segments and then communicate around or over any obscuring object. Compared to picocellular BSs, the relay stations do not need a wired network connection. The lack of a direct backbone connection makes the relay stations not only cheaper, but also more flexible in positioning because they do not rely on fixed network access, but merely on a power supply. The relays might be either fixed or mobile and serve to cover otherwise shadowed areas and to enlarge the limited coverage range of a BS. This solution is seen as very attractive by the relevant research community, because it is considered unlikely that the high traffic capacity of a broadband BS will be used up by the user terminals roaming in its cell.

Further, the offered capacity distribution of a single AP is highly unfair against the UT. Figure 6.1(a) shows that only users next to the AP can profit from the high multiplexing bit rates, but the user density increases with increasing distance due to the increasing area. Figure 6.1(b) shows the optimal solution, and Figure 6.1(c) shows that the employment of relay stations helps with the optimal deployment of capacity.

A fixed relay stations (FRS) is a cheaper solution to deploy because it builds on the fixed backbone network access and requires only some additional radio resources and hardware. These are also two factors to consider when judging the economical efficiency of relay-based deployment.

A methodology called *weighted spectral efficiency* allows a fair measure of deployment costs without having final knowledge about exact hardware costs. The IST project WINNER proposed, investigated, and evaluated novel multihop concepts building on this concept.

Adaptive antenna systems (see Chapter 5) allow for simultaneous communication between stations. SDMA techniques can be used to simultaneously serve

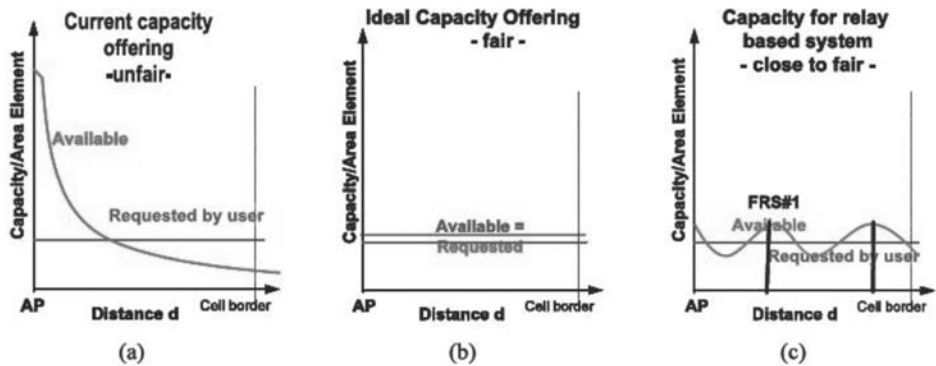


Figure 6.1 (a–c) Capacity offering per area element, with and without relays [5].

several relay stations with only one BS. The low mobility expected for relays in the infrastructure network privileges the use of predistortion or joint detection techniques. Furthermore, different subscriber stations are scheduled in SDMA mode to be concurrently served by a BS or a relay. Relay stations and/or subscriber stations, which are sufficiently spaced, are served simultaneously.

Cooperative communication relies on several links established between one receiver (subscriber station) and different transmitters (base and/or relay stations). On one hand, transmitting the same signal via several independent links simultaneously increases the received signal strength. On the other hand, transmitting independent signals simultaneously increases the throughput. Finally, transmitting the same signal consecutively via different links decreases the BER of the decoded data.

6.1.1 Standards Supporting Multihop Communications

6.1.1.1 IEEE 802.16 Mesh Mode

Mesh networks based on 802.16 [6] (mesh WMAN) are intended to be used for backhaul networks that span urban infrastructure. Mesh WMANs are easy to deploy and, therefore, of interest to operators. With mesh WMANs, the wireless spectrum that has been allocated to a provider can be used not only to connect customers but also to interconnect several BSs. No cabling or leased lines are needed and, hence, the operational expenditures (OPEX) are reduced [6].

The IEEE 802.16 mesh mode is an optional feature of the IEEE 802.16 standard [6]. In contrast to the mandatory point-to-multipoint (PMP) configuration, where traffic only occurs between the BS and the subscriber stations (SSs), in the mesh mode, traffic can be routed through other SSs and can occur directly between SSs. This is shown in Figure 6.2. Such SSs are called *mesh SSs*. The BS that is connected to the backhaul network is called a *mesh BS*. All data and control messages are forwarded in the time domain by the mesh SSs.

6.1.1.2 IEEE 802.16e

WiMAX [8] handover is the process in which a mobile station migrates from the air interface of one BS, called the *serving BS* (i.e., with this BS, the MS has most

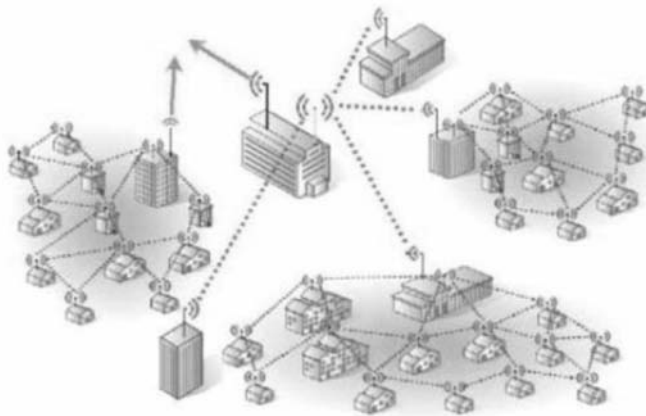


Figure 6.2 WMAN mesh deployment [7].

recently completed the registration at an initial network entry or during a handover), to another air interface provided by another BS, called the *target BS* (the BS whose downlink transmission can be received by the MS). The decision about handover execution can come either from the MS or the BS. Two basic types of handover exist in WiMAX [8]:

- Hard handover;
- Soft handover.

Hard handover is a mandatory type of handover. This handover is applied when the MS communicates with just one BS. Connection with the old BS is broken before the new connection is established.

Hard handover cannot be used when high-speed user mobility is required. For high-speed user mobility, soft handover is used. During soft handover the MS simultaneously communicates with two or more BSs. Two optional types of soft handover are defined in 802.16e:

- Fast base station switching (FBSS);
- Macrodiversity handover (MDHO).

In FBSS, the MS and BS maintain a list of BSs that are involved in the FBSS. This list of BSs is called a *diversity set*, and it is continuously monitored by the MS. The MS monitors all BSs, from which it can receive a signal. These BSs are called *neighbor BSs*. The BS can be added to (or dropped from) the diversity set if the long-term carrier interference noise ratio (CINR) of this BS is higher or less than the CINR of the serving BS within a defined threshold. For every moment an anchor BS is defined by the MS; this anchor BS can be chosen from the diversity set only. The anchor BS is the BS where the MS is registered and synchronized, and it performs ranging and monitors the downlink channel for control information. All traffic communication and management messages are exchanged between the MS and the anchor BS. An anchor BS can be changed from frame to frame, meaning that every frame can be sent by a different anchor BS. The changing of an anchor BS depends on a resolution of the MS. The data transmission process in FBSS is shown in Figure 6.3.

MDHO also uses the diversity set of BSs. The communication proceeds among a MS and all BSs in the diversity set. For downlink MDHO, two or more BSs transmit data to the MS such that diversity combining can be performed at the MS. For uplink MDHO, the transmission of a MS is received by multiple BSs, where selection diversity of the received information is performed. This process is shown in Figure 6.4.

For WiMAX networks, the IEEE began work in 2006 on 802.16j “Mobile Multihop Relay” (MMR). The basic idea behind MMR is to allow WiMAX BSs, which do not have a backhaul connection, to communicate with BSs that do. One of the main OPEX drivers in mobile networks today is the requirement to connect each base station directly to the network. In most cases this means installing a cable or microwave connection at each BS site. For fixed line connections, UMTS/HSPA base stations primarily use a number of E-1 or T-1 links today, each

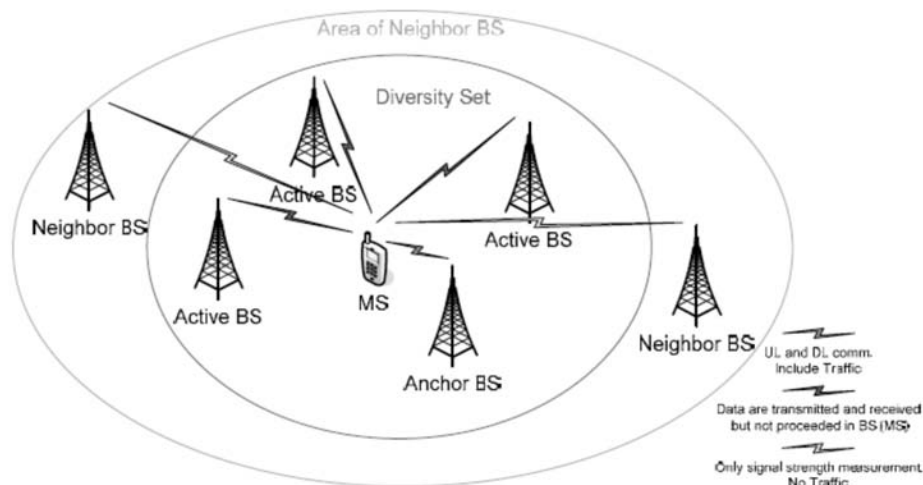


Figure 6.3 Data transmission in FBSS [7].

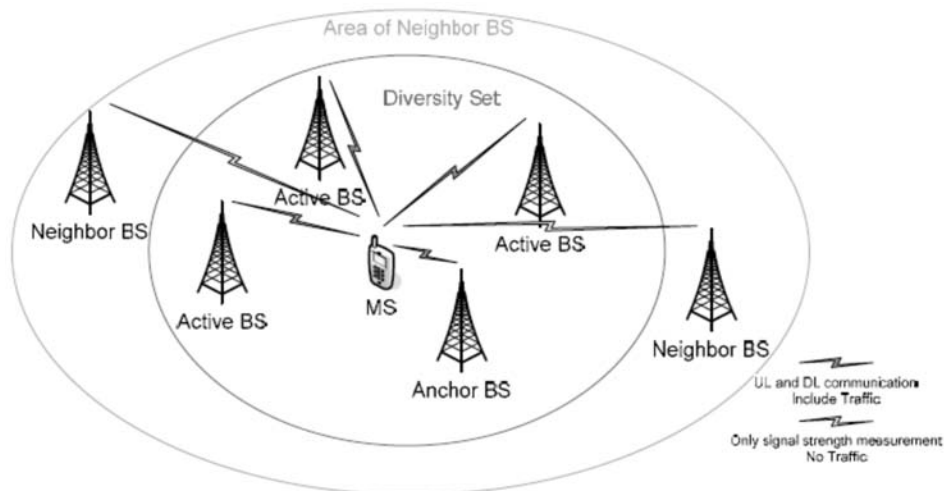


Figure 6.4 Data transmission in MDHO [7].

being capable of transmitting about 2 Mbps. With rising bandwidths, however, using several E-1 links for each BS can become difficult in terms of both the cost and availability of wires. Alternatives are fiber or use of a transmission technology that can get more out of a pair of copper cables, which can still remain expensive. On one hand, this will reduce the bandwidth available to users in the cells involved in relaying packets. On the other hand, this is one way to save costs and extend the network coverage into areas where connecting a BS directly to the network via a fixed line connection is economically or technically not feasible. The topology of such a mesh architecture is shown in Figure 6.5.

Generally speaking, backhaul is used to indicate the service of forwarding traffic from the originator node to an access point from which it can be distributed

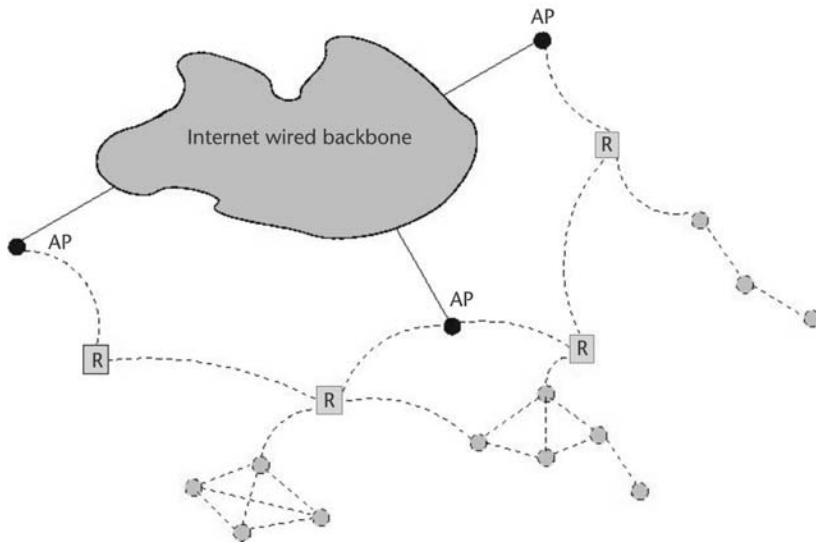


Figure 6.5 Mesh network topology [9].

over an external network. Specifically in the mesh case, the traffic is originated in the users' devices, traverses the wireless backbone, and is distributed over the Internet network.

Modifications to enhance the coverage, throughput, and system capacity of 802.16 networks by specifying 802.16 multihop relay capabilities and functionalities of interoperable relay stations and BSs were consequently made by a special working group within the 802.16e standard [8]. It is expected that the complexity of relay stations will be considerably less than the complexity of legacy IEEE 802.16 BSs. The gains in coverage and throughput can be leveraged to reduce total deployment costs for a given system performance requirement and thereby improve the economic viability of IEEE 802.16 systems. The main features offered by the IEEE 802.16j [10] MMR system are as follows:

- Spectrally efficient architectures and spatial frequency reuse, which is the result of a flexible placement of cell sites (i.e., relays). Access limitations are also reduced.
- Replacement of low-rate, unreliable links between mobile BSs (MBSs) and mobile SSs (MSSs) with multiple high-rate, reliable links between MBS and mobile relay stations (MRSs) and/or MRSs and MSSs.
- Load sharing among RSs and fault tolerance via multipath redundancy.

6.1.1.3 IEEE 802.11s

Standard 802.11s [11] specifies an extension to the IEEE 802.11 MAC at the link layer to solve the interoperability problem by defining an architecture and a protocol that support both broadcast/multicast and unicast delivery. To achieve the above-mentioned solution, it uses “radio-aware” metrics over self-configuring multihop topologies. It adds extra functions to allow wireless nodes to discover each other,

authenticate and establish connections, and work out the most efficient route for a particular task.

All APs in 802.11 are usually directly connected to the wired network and provide wireless connectivity to client devices. Client devices, on the other hand, are typically implemented as simple 802.11 stations that must associate with an AP in order to gain access to the network. The 802.11s specification suggests that APs will be able to establish peer-to-peer wireless links with neighboring APs and in this way to establish a mesh infrastructure, without the need for a wired network connection to each AP. Moreover all other devices, today categorized only as clients, will be able to establish and forward peer-to-peer connections with other clients and APs.

The nodes that support mesh services are referred to as mesh points (MPs). Some mesh points can simply provide AP services; these devices are called mesh APs (MAPs). In addition, MAPs can also support the 802.11 nonmesh stations (STA) (see Figure 6.6).

Simple STAs do not participate in WLAN mesh services, such as path selection and forwarding. Mesh services are implemented as logical MAC interfaces independently on the 802.11 MAC interface. It follows that a single device can play the role of MP and AP (or MAP) for nonmesh stations at the same time. A congestion control mechanism implemented at each MP sustains the ratio of complexity and performance of transmitted data at an appropriate level. The simplest form of the WLAN mesh is the one in which all nodes operate on one radio channel, as shown in Figure 6.7(a).

For multichannel operations, multiple radios or channel switching capability is required. When channel switching is not supported, each radio of the mesh points operates on one channel at a time [see Figures 6.7(b, c)]. The channel may change during the lifetime of the mesh network according to dynamic frequency selection (DFS).

Mesh points are connected to each other using a common channel. A set of mesh points, which are interconnected to each other via a common channel, is referred to as a unified channel graph (UCG). Figure 6.8 shows that nodes that support multichannel radio may belong to several UCGs.

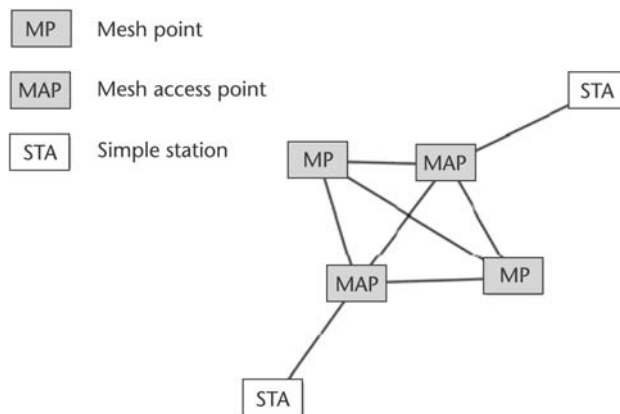


Figure 6.6 WLAN 802.11s mesh architecture [7].

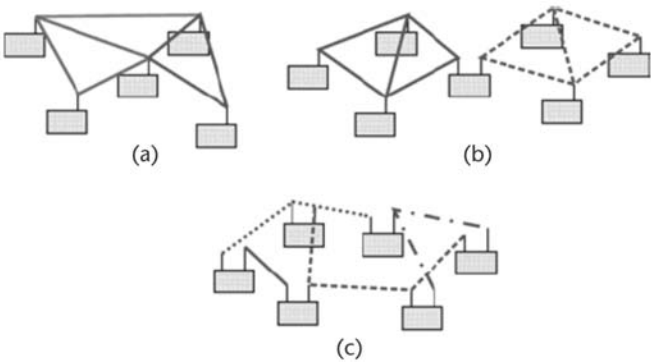


Figure 6.7 (a–c) Example of channel configurations in a WLAN mesh configuration [7].

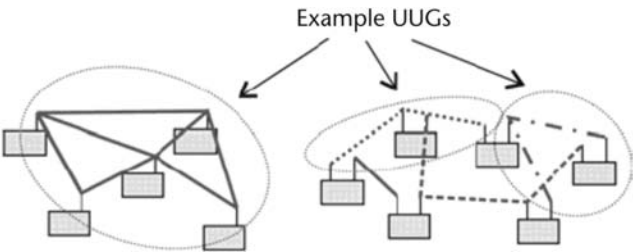


Figure 6.8 UCGs in a WLAN mesh [7].

6.1.2 Relay Technologies in Support of Ubiquitous Radio Access

The WINNER system [2] is a totally new concept in radio access. It is built on the recognition that developing disparate systems for different purposes (cellular, WLAN, short-range access) will no longer be sufficient in the future converged wireless world [7]. Thus, it provides wireless access for a wide range of services and applications across all environments, from short range to wide area, with one single adaptive system concept.

Both RSs with homogeneous radio interfaces and RSs with different radio interface modes are considered in the WINNER concept. Figure 6.9 shows an example of a WINNER deployment scenario, that combines these two types of RSs for coverage in an urban hot-spot area scenario supplemented by wide-area radio coverage.

Several functions of a wireless communication system have to be adapted in order to work well in relay-based deployments [12]. A proper RRM scheme is essential to achieve the full benefits of a relay-based deployment. For that reason an RRM framework for relay-enhanced cells (RECs) was introduced in WINNER. The proposed flexible resource partitioning for RECs allows the use of relay nodes (RNs) for coverage extension, to cover shadowed areas, and to increase the capacity in the cell. Moreover, cooperative relaying was integrated in the RRM framework. The RRM framework was applied to a metropolitan-area and wide-area deployment scenario to verify its flexibility. Through extensive simulations, the proposed

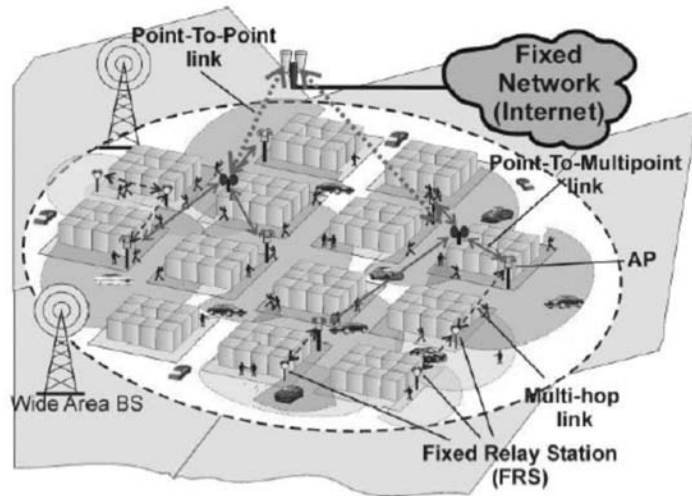


Figure 6.9 Deployment of relays in a ubiquitous system concept [7].

flexible resource partitioning was shown to outperform static resource partitioning [12].

6.1.2.1 Relay Deployment Costs

Deployment costs for a radio access point consist mainly of two parts, the hardware and the site acquisition costs [12]. Hardware costs are driven primarily by the maximum output power and the complexity of the AP (e.g., multiple antennas). Hence, a cheap AP is characterized by a low output power, a single antenna, and potentially a low complexity (which, however, could mean lower capability). The site acquisition cost of an AP is driven primarily by the size of the equipment and by whether cooling, shelter, and backhaul transmission connections are required. Also, in many industrial countries the density of deployed APs is already high, meaning that finding new candidate site positions, especially for macrostation equipment, is difficult in many areas.

The IST project WINNER compared different types of APs and their total deployment costs. Studies [5, 9, 12, 13] have shown that a small size and low complexity are beneficial from a cost perspective. This is valid for an AP using a backhaul connection, as well as relays. Hence, the most important factor that differs between a relay and a wired AP is the *cost of transmission*. Potentially, the fact that relays do not require wired transmission implies that the choices for site positioning are more numerous, which might potentially lead to a reduction of the deployment cost of relays compared to a wired access point.

Total deployment costs are often differentiated between capital expenditures (CAPEX) and OPEX. CAPEX include the once-only cost invested in acquiring the necessary equipment, and are to be paid at the time of deployment. These expenses are usually connected with an approximate usage time during which the equipment is in service. Costs that appear periodically in time are classified as OPEX. Typical costs that can be characterized as OPEX are power consumption, operation and

maintenance, and transmission. In addition, deployment sites can be rented for a timely fee. It is worth pointing out in this context that OPEX and CAPEX can easily be transferred to one another by assuming a certain life for the equipment and an interest rate. This implies that a comparison of costs can be made in CAPEX or OPEX.

Some examples of CAPEX and OPEX costs are summarized in Table 6.1 and Table 6.2, respectively. These figures are merely examples provided by the IST project WINNER and partially based on the available literature [14, 15] and on estimations. Figures can vary depending on the deployment scenario and should be understood only as a demonstration basis.

As a general rule, considering a 10-year depreciation value, OPEX should be higher than the annualized CAPEX, that is, CAPEX divided by 10 (years). The

Table 6.1 Example of CAPEX Cost Elements for Outdoor Scenarios [13]

<i>Cost Element</i>	<i>Unitary Cost (K€)</i>	<i>Assumptions/Remarks</i>
Macro BS equipment	50	Three sectors include air conditioning, batteries, container antennas, cables, etc.
Micro BS equipment	5	
Relay equipment	7	More expensive than macro BS due to technological innovation required
Macro BS Site acquisition and deployment	60	
Micro BS site acquisition and deployment	6	Small footprint
Relay site acquisition and deployment	4	Small footprint, no backhaul
Gateways	100	
RRM servers	100	
Macro fixed line connections	0.15	Connection to fiber optic access network
Micro fixed line connections	0.05	Connection to mass-market ADSL line

Table 6.2 Example of OPEX Cost Elements for Outdoor Scenarios [13]

<i>Cost Element</i>	<i>Unitary Cost (K€/yr)</i>	<i>Net Present Value (K€) for 10 Years</i>	<i>Assumptions/Remarks</i>
Macro BS site rent, maintenance, power, and operation	17	132.7	Includes software upgrades
Micro BS site rent, maintenance, and power	3	23.4	Small footprint, no back-up batteries
Relay site rent, maintenance, and power	2	15.6	33 dBm relay; no back-up batteries
Macro BS fixed line connections	20	156.0	Professional high-capacity backhaul line failsafe
Micro BS fixed line connections	0.8	6.24	Mass-market DSL

OPEX costs are represented by their net present value, assuming a lifetime of 10 years and a discount rate of 6%.

To make a fair comparison between the performance of relay-based and traditional single-hop deployments, a system level class III simulator was implemented. The simulations were carried out in a particularly dense urban zone, corresponding to the microcellular scenario or metropolitan area contemplated as one of the three basic scenarios in the WINNER system [13]. The simulator performed simulations at a system level, but from a static or quasistatic point of view, the study was focused on the analysis of a short snapshot of the system and did not deal with the dynamic behavior of the system.

Two different types of simulations were performed. First, simulations of a conventional single-hop deployment using exclusively bisectorial BSs located uniformly over a certain scenario and then simulations in the same scenario with the BSs located in the same sites but including two relay nodes per site. The performance increase was determined in terms of spectral efficiency and coverage. The real scenario for the simulations focused on an area with a regular distribution of blocks similar to the Manhattan example [12]. Figure 6.10 shows the scenario in which the simulations were carried out, indicating the name of each site included in both

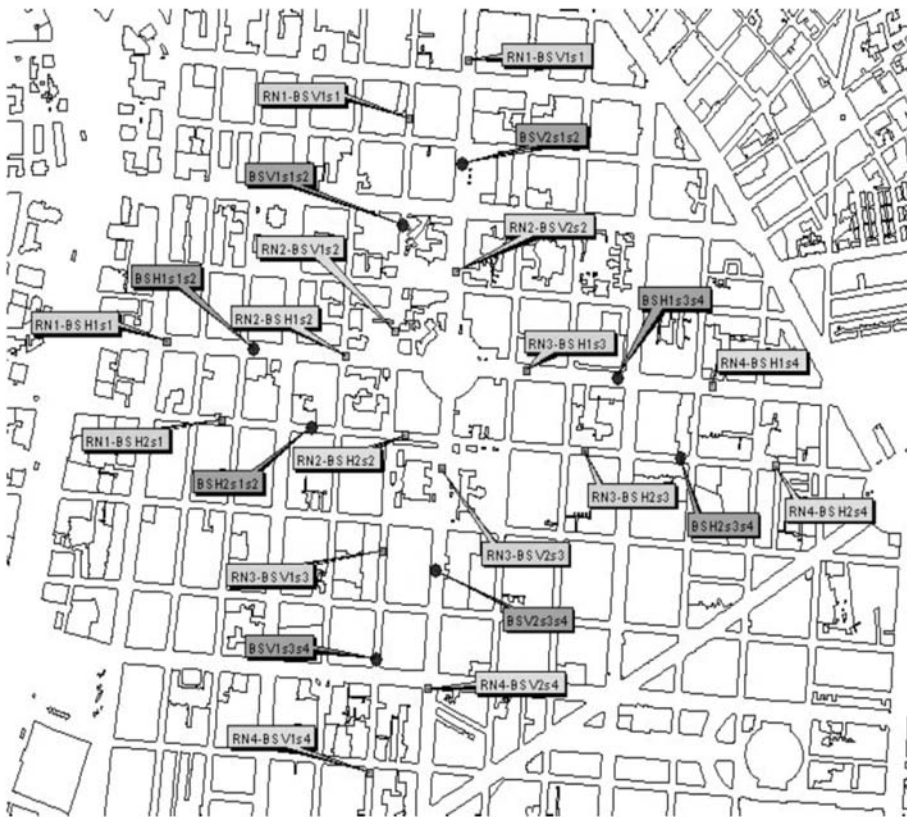


Figure 6.10 Real cartography of the city of Madrid, Spain, indicating the sites used for the simulations [12].

types of deployments: relay based and traditional. In the initial simulations two horizontal and two vertical streets were chosen, located uniformly for each one, two bisectorial BSs. In this way, for the traditional deployment, a total of 8 BSs was used, or 16 sectors with a frequency reuse of one. On the other hand, for the relay-based deployment a relay node per sector was added, so that 16 relays were included for this kind of deployment, splitting them into two groups (odd and even relays) in order to use different radio resources and to avoid interference between relays of different groups. By defining a certain grid with the location and direction of movement toward the relay APs (RAPs), the path loss can be calculated for all relay APs.

The distance between the points of the grid shown in Figure 6.10 determines the assumed user density in the simulations.

The grid used in the simulations is shown in Figure 6.11 with an interspacing of 10m between the points of the grid.

Although the chosen scenario is quite regular and similar to the Manhattan case, the size of the buildings and the width of the streets are not equal; therefore, the results could be different not only for the relative position of the sites in the grid (centered sites are exposed to a higher interference level than border sites), but also by the particular characteristics (composition and geometry) of the buildings around the site.

Basically, once the grid and the position of the RAPs has been defined for the scenario, the steps of the simulation are as follows:

- 1. Compute for each RAP (sector or relay node) the received power in each point of the grid.

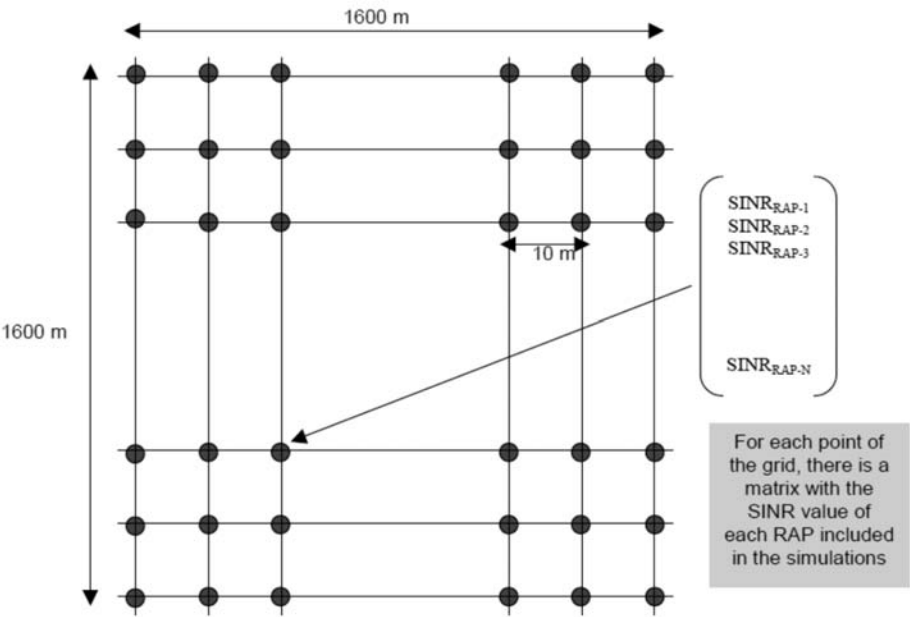


Figure 6.11 Size and interspacing of the grid used in the simulations [12].

2. Calculate the SINR and spectral efficiency for each point and for each RAP. The spectral efficiency was calculated according to the Shannon law and taking into account a penalty of 4 dB.
3. Determine the coverage areas for each RAP. The points of the grid included in the coverage area of a certain RAP are those that display an SINR greater than the SINR of the other RAPs.
4. Calculate the cell throughput for each single sector (single-hop conventional deployment) or enhanced sector (relay-based deployment), assuming a certain limit for the minimum spectral efficiency, which can be used for defining the cell service area of a given RAP.

Figure 6.12 shows the different phases of the simulations. The following assumptions were made:

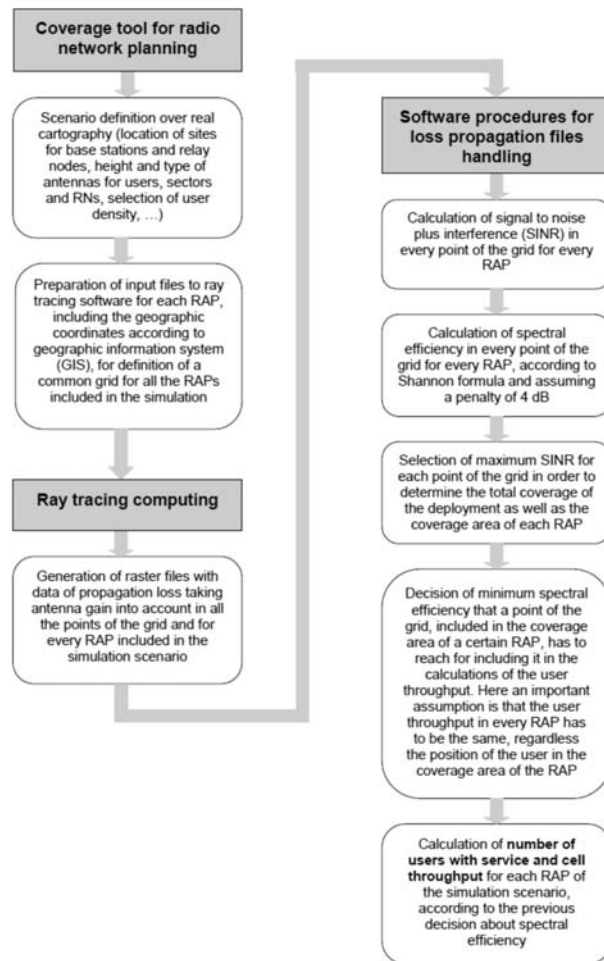


Figure 6.12 Flowchart of the simulations for performance comparison of conventional and relay-based deployments [12].

- Duplexing scheme and asymmetry: TDD (1:1);
- Carrier central frequency: 3,950 MHz;
- Channel bandwidth: 100 MHz;
- Location and height of base stations: below rooftop at 10m from the street floor;
- Maximum transmit power per sector: 37 dBm (5.012W);
- Number of antennas and type for sector: 1 antenna K733337XD with a radiation pattern as shown in Figure 6.13.

The employed resource partitioning is shown in Figure 6.14.

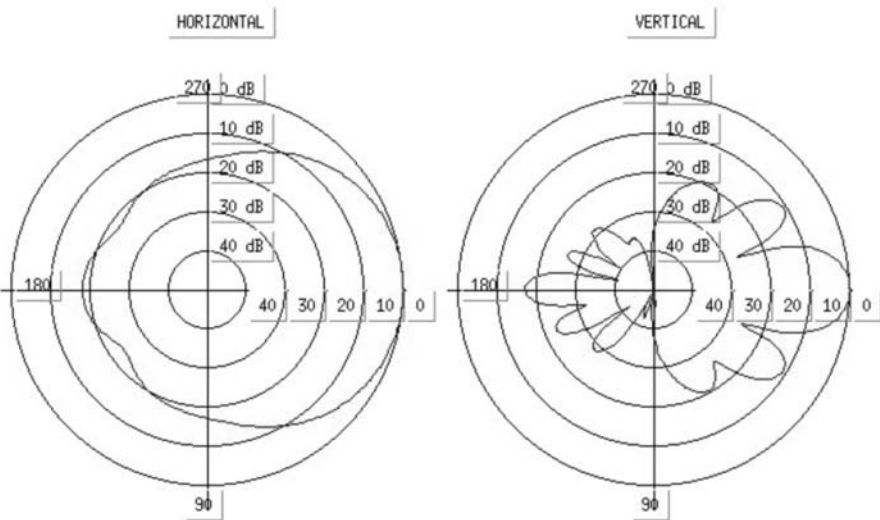


Figure 6.13 Radiation pattern for the bisectorial antenna of BSs used in the simulations [12].

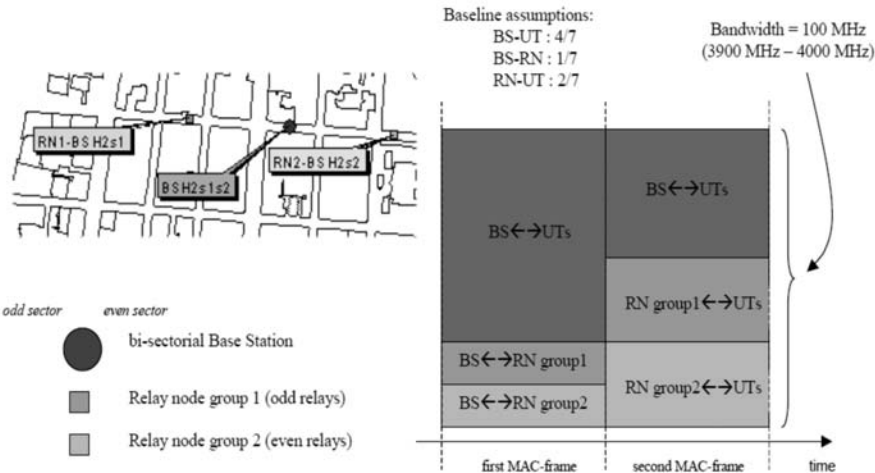


Figure 6.14 Resource partitioning for relay-based deployment simulations [12].

The computation of the propagation loss is based on a ray-tracing model instead of a propagation model because the simulations are done over a real and particular scenario with known characteristics (e.g., materials and heights of buildings). Ray tracing, as a deterministic model, can yield more realistic results than a propagation model. The calculation of the propagation loss includes the antenna gain effect for every point of the defined grid and for all RAPs (BSs or RNs) included in the simulation. The SINR is calculated according to the following equation:

$$\text{SINR}_{\text{RAP}-i} \text{ (dB)} = S_{\text{RAP}-i} \text{ (dBm)} - [N + I] \text{ (dBm)} \quad (6.1)$$

where the signal value is given by the expression

$$S_{\text{RAP}-i} \text{ (dBm)} = Ptx_{\text{RAP}-i} \text{ (dBm)} - L_{\text{RAP}-i} \text{ (dB)} \quad (6.2)$$

Here $Ptx_{\text{RAP}-i}$ is the transmitted power by RAP number i in dBm, and $L_{\text{RAP}-i}$ is the propagation loss for the RAP number i taking its antenna gain into account. Note that a uniform power distribution was assumed over the whole bandwidth allocated to each RAP, so that the SINR level for a RAP using all of the bandwidth is calculated assuming full power. For the calculation of the noise N , the receiver noise figure NF (7 dB for the UT case and 5 dB for the RN case), and the thermal noise (kTB) in terms of noise power spectral density (-174 dBm/Hz) and receiver bandwidth (B), are considered as follows:

$$N \text{ (dBm)} = kT \text{ (dBm/Hz)} + B \text{ (dBHz)} + NF \text{ (dB)} \quad (6.3)$$

Interferences influencing in RAP number i are calculated according to the following expression:

$$I_{\text{RAP}-i} \text{ (dBm)} = \sum_{j=i}^M [Ptx_{\text{RAP}-j} \text{ (dBm)} - L_{\text{RAP}-j} \text{ (dB)}] \quad (6.4)$$

The simulations focused only on DL transmissions, so the interference for a certain RAP is calculated according to the contribution of the rest of the transmitters radiating in the same time and in the same spectrum. For the estimation of the spectral efficiency in each point of the grid and for each RAP, and assuming Shannon's theorem approach including a penalty of 4 dB, the new SINR becomes:

$$\text{SINR}_{\text{RAP}-i} \text{ (dB)} - \text{SINR}_{\text{RAP}-i} \text{ (dB)} - 4 \text{ (dB)} \quad (6.5)$$

Using the SINR' value, the spectral efficiency of each RAP can be calculated for all the points of the grid in bit/s Hz⁻¹ ($\text{sEff}_{\text{RAP}-i}$) by the following expression:

$$\text{sEff}_{\text{RAP}-i} = \log_2(1 + 10^{\text{SINR}_{\text{RAP}-i}/10}) \quad (6.6)$$

The coverage area of every RAP inside the common grid used in the simulated network deployment can be calculated by simply including those points of the grid

showing a better SINR. To delimit this coverage area, the user throughput is decided in terms of the points, in which the spectral efficiency is higher than a certain value. In other words, the issue is to decide what value of this parameter has to be achieved by the user or point of the grid for giving it service. In addition, a fixed throughput for all the users served by a given RAP is assumed, so a trade-off between the number of users served by the RAP and the radio resources (bandwidth) allocated to each user is achieved to account for fair resource partitioning. A minimum spectral efficiency of 1 bit/s Hz^{-1} was assumed for all the users. The user target throughput of a certain RAP is obtained once the number of users of this RAP is known (the points of the grid with spectral efficiency higher than 1 bit/s Hz^{-1}), and taking the allocated bandwidth into account. Hence, for each RAP the throughput of those users with spectral efficiency exceeding this threshold is given in terms of its allocated bandwidth ($Bw_{\text{user}-j}$) by the expression

$$T_{\text{user}-j} = Bw_{\text{user}-j} \cdot s\text{Eff}_{\text{user}-j} \quad (6.7)$$

Assuming a fixed throughput for all the users of a certain RAP according to

$$T_{\text{user}} = T_{\text{user}-j} (\forall j \in \text{CoverageArea}_{\text{RAP}-i} \text{ and } s\text{Eff} \geq 1) \quad (6.8)$$

the calculation of the user throughput of a certain RAP is obtained rearranging the terms of the following expression:

$$T_{\text{user}} = \frac{\sum_{j=1}^M Bw_{\text{user}-j}}{\sum_{j=1}^M \frac{1}{s\text{Eff}_{\text{user}-j}}} = \frac{Bw_{\text{TotalRAP}-i}}{\sum_{j=1}^M \frac{1}{s\text{Eff}_{\text{user}-j}}} \quad (6.9)$$

where $Bw_{\text{TotalRAP}-i}$ is the total bandwidth available by the RAP number i for serving the user terminals, and M is the total amount of users served by the RAP $-i$ (sector of a BS or RN). The available bandwidth of a sector in the conventional deployment case is 100 MHz, whereas in a relay-based deployment, this depends on the proportion of bandwidth devoted to the BS-RN link. In the same way, the bandwidth allocated to the RN-UT links will be a certain fraction (baseline assumption 2/7) of the total bandwidth of the system.

Finally, the total throughput of a given cell RAP $-i$ is immediate and obtained through the summation of each user throughput, according to the following equation:

$$T_{\text{cell}} = M \cdot T_{\text{user}} \quad (6.10)$$

Figure 6.15 shows the results for the total service area covered for each of the two deployments.

Table 6.3 shows the comparison of the aggregate throughput and the total service area provided for both deployments, as well as the percentage of wasted

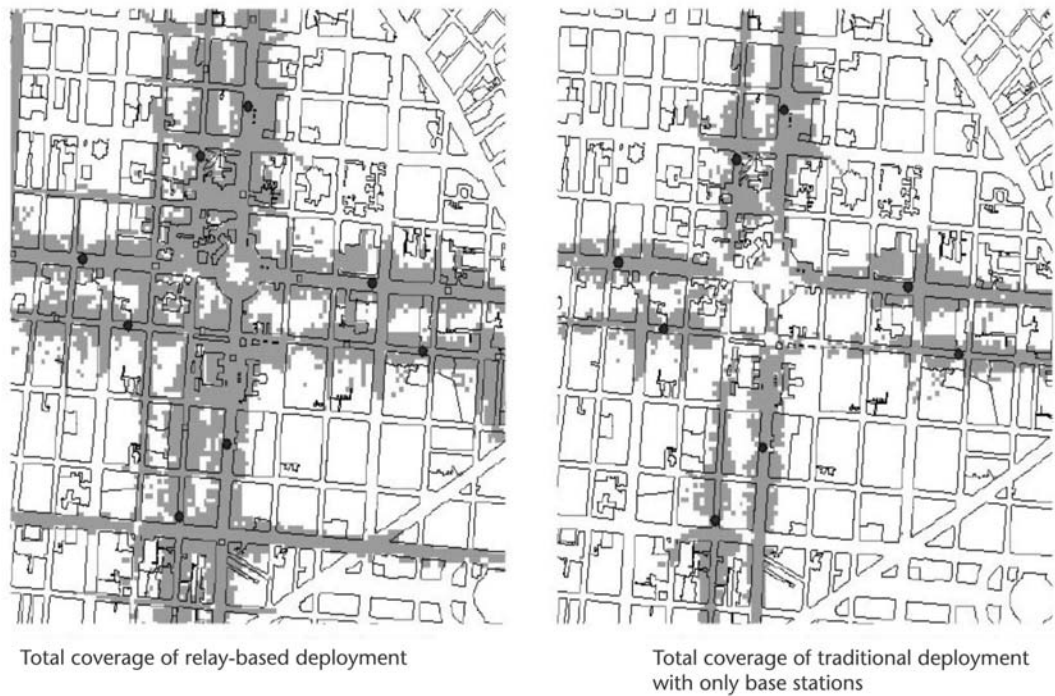


Figure 6.15 Global comparison of the coverage areas (gray color) obtained in the simulations for relay-based and traditional deployments [12].

Table 6.3 Performance Comparison Results of Relay-Based and Traditional Deployments [12]

Comparison Site	Cell Throughput Ratio (Relay- Based/ Traditional)	Users Number Ratio (Relay- Based/ Traditional)	Users Number Ratio for	
			Relay-Based Deployment (Relay Node/ Sector)	Percentage of Wasted Resources 1*hop (%)
s2+RN2 vs. s2 (BSH1)	0.75	3.21	6.25	42.85
s3+RN3 vs. s3 (BSH1)	0.86	1.59	2.54	49.65
s2+RN2 vs. s2 (BSH2)	0.89	1.58	2.40	51.49
s3+RN2 vs. s3 (BSH2)	0.89	2.13	4.23	36.34
s2+RN2 vs. s2 (BSV1)	0.87	1.51	2.45	44.55
s3+RN3 vs. s3 (BSV1)	0.84	1.93	4.21	38.98
s2+RN2 vs. s2 (BSV2)	0.87	2.04	3.77	33.83
s3+RN3 vs. s3 (BSV2)	0.97	1.80	4.29	18.78
Mean values	0.87	1.98	3.77	39.56

resources in the first hop for the communications through relays. The high percentage of wasted resources in the BS-RN link, close to 40%, shows an inefficient resource-partitioning scheme.

The main observation is that the cell service area in the relay-based deployment is almost doubled compared to the traditional single-hop deployment giving service to more users (a coverage area mean ratio of 1.98). The coverage enlargement comes at a cost of a lower average user throughput in the REC than is observed

in the conventional single-hop deployment. For the particular scenario (urban microcellular) and the particular assumptions considered in the simulations under discussion, the total cell capacity point decreased in the relay-based deployment compared to the conventional single hop using only BSs. The mean ratio between the cell throughputs of the REC and the conventional deployment is slightly lower than 1 (a total capacity mean ratio of 0.87). Use of a flexible and efficient resource-partitioning scheme could have a big potential to increase the performance of the relay-based deployment [13].

6.1.2.2 Cooperative Relaying and Intelligent Deployment

In cooperative relaying, a receiving node may combine signals from more than one transmitting node. Multihop diversity is one example in which a receiving node combines the signals received from previous nodes in the path. In a two-hop DL path, the UT combines the signal from an RN with the signal from the BS.

Intelligent deployment describes any method for reducing the probability that an RN has a poor link to the BS and has the effect of reducing the mean path length. One such method is to move an RN to an alternative site if the signal quality between the RN and the BS at the first site is below a specified threshold. An extreme example is the perfect deployment, which implies that all RNs have a good link to the BS and, therefore, that the maximum path length is two hops.

Cooperative relaying and intelligent deployment reduce the total cost of a multihop network by reducing the number of relays needed for a given performance. This is evidenced by increased curvature of the indifference or isoperformance curve (see Figure 6.16), with the result that the least costly system is found closer to the origin, representing an overall reduction in network costs due to the introduc-

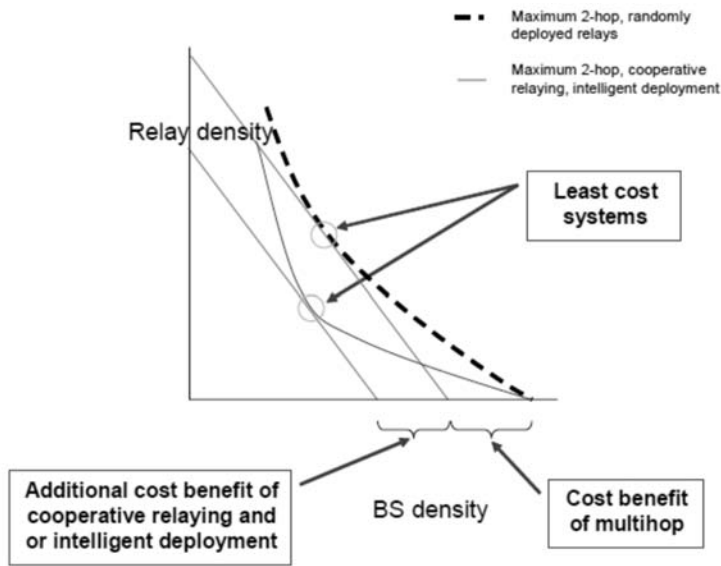


Figure 6.16 Effect of cooperative relaying and intelligent deployment on the indifference map [12].

tion of these techniques. Note that the indifference curves for both systems meet on the x axis.

6.1.2.3 Spatial Processing

Spatial processing techniques make use of the additional degrees of freedom provided by multiple antennas at the transmitter and/or the receiver. MIMO techniques may be configured to provide improved link capacity and SINR. The 2×2 space time transmit diversity (STTD) is a form of MIMO, in which two transmitting antennas and two receiving antennas are configured to provide an improvement in SINR at the receiver of ~ 3 dB relative to a SISO system. To compare systems employing spatial processing (MIMO) with more conventional systems (SISO), two additional parameters must be introduced: α , the cost ratio of a MIMO BS to a SISO BS; and β , the cost ratio of a MIMO RN to a SISO RN. An indifference curve for a MIMO system can be plotted on the same indifference map as an indifference curve for a SISO system if the density of the MIMO BSs is mapped to an equivalent density of SISO base stations by multiplying by α .

Similarly, the density of MIMO relays should be mapped to an equivalent density of SISO relays by multiplying by β . For example, if a MIMO BS costs 1.1 times more than a SISO BS, then $\alpha = 1.1$. Similarly, if a MIMO RN costs 1.2 times more than a SISO RN, then $\beta = 1.2$. The equivalent SISO network corresponding to a MIMO network containing 1 MIMO BS per square kilometer and 5 MIMO RNs per square kilometer would comprise 1.1 SISO BSs per square kilometer and 6 SISO RNs per square kilometer.

MIMO reduces the number of required BSs, resulting in a steeper indifference or isoperformance curve. Furthermore, the indifference curves no longer meet on the x axis. Although the incremental benefit of multihop may be lower than in a SISO system, the combined benefit of MIMO and relaying is typically greater than relaying alone. The overall effect is to produce a lower cost network when spatial processing is combined with relaying compared to a SISO multihop system or a MIMO cellular system. This is shown by the proximity of the indifference curve to the origin as shown in Figure 6.17.

The methodology presented here offers the benefit of allowing a direct comparison of different systems and deployments at a common level of performance and with a common metric, namely, relative cost. It forms the basis for the evaluation of relay-based systems and techniques within the various scenarios defined for the WINNER system [9].

6.2 RRM Schemes for Relay-Based Deployments

Radio resource management schemes are necessary in order to minimize interference in relay-based systems by the coordination of radio resources across the RAPs that might be either BSs or RNs. Section 6.1.2.1 demonstrated that resource partitioning might be beneficial in order to enhance the coverage and spectral efficiency of the RECs. Different schemes and scenarios were investigated [12, 13] in the WINNER project to ensure that the final system concept becomes a strong candidate in the

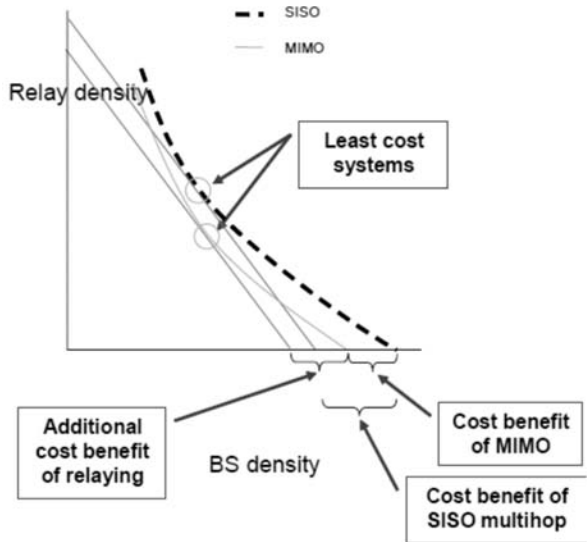


Figure 6.17 Effect of 2×2 STTD MIMO spatial processing on the indifference map [13].

IMT-A competition [16]. In the following subsections some of these investigations are described.

6.2.1 Spectrum Partitioning

This investigation considered a cellular system consisting of 19 hexagonal RECs. Each REC is assumed to have a BS, located at the cell center, and 6 fixed relay nodes (FRNs), located on the lines connecting the BS, and the six cell vertices. This is shown in Figure 6.18.

The distance between the BS and any cell vertex is assumed to be D and the distance between the BS and any FRN is d_{BF} . The ratio between d_{BF} and D

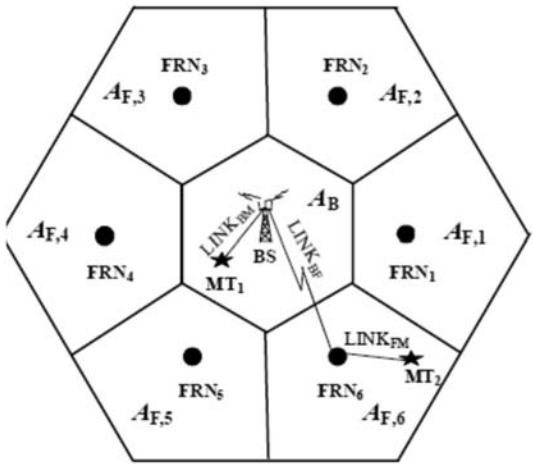


Figure 6.18 Cellular system enhanced with two-hop FRNs [13].

determines the positioning of the FRNs. In each cell, an MT can communicate directly with the BS over a one-hop link or alternatively establish a two-hop link through an FRN, under the control of a *shortest-distance-based* routing controller. The described study is restricted to the uplink and assumes that a homogeneous air interface is employed.

The spectrum partitioning procedure partitions the bandwidth available in each REC. If this bandwidth is supposed to be of B Hz, among the set of links connecting the BS and the FRNs, the ones of importance are links between the FRNs and the MTs and the BS and MTs.

Reference [13] proposed partitioning the available spectrum into two parts: part 1 of B_1 Hz, which is assigned to the MTs communicating over the one-hop links; and part 2 of $B_2 = (B - B_1)$ Hz, which is used by the MTs communicating over the two-hop links. The part 2 spectrum is further partitioned into two subparts: subpart 1 of B_{21} Hz, which is assigned to links connecting the FRNs and the MTs, and subpart 2 of $B_{22} = (B_2 - B_{21})$ Hz, which is used on links connecting the BS and the FRNs. The subpart 1 spectrum is used to define N_2 equal-bandwidth channels, denoted as C_{FM} ; whereas the subpart 2 spectrum is used to define the same number of equal-bandwidth channels, denoted as C_{BF} . Both C_{FM} and C_{BF} are further divided into three equal-number channel sets: One set is shared by FRN_1 and FRN_2 , another is shared by FRN_3 and FRN_4 , and the remaining set is shared by FRN_5 and FRN_6 . Three “virtual sectors” are thus created with the above assignment scheme although omnidirectional antennas are deployed.

The spectrum partitioning problem can be solved with a channel-dependent solution, of which the SIRs on the respective link sets are taken into account. In particular, the ratio of $B_1:B_{21}:B_{22}$ is obtained as the solution of three linear equations. In addition, the obtained ratio depends heavily on where the FRNs are placed [13]. The ratio of $B_1:B_{21}:B_{22}$, under the constraint

$$B_1 + B_{21} + B_{22} = B \quad (6.11)$$

is the key issue for the spectrum partitioning framework discussed here. This ratio should be closely related to the ratio of the *mean aggregate data rate* of one-hop MTs, denoted as $\bar{R}_{A,1}$, to that of two-hop MTs, denoted as $\bar{R}_{A,2}$. Assuming that the MTs are requesting to transmit with the same data rate on the uplink, then the following holds:

$$\frac{\bar{R}_{A,1}}{\bar{R}_{A,2}} = \frac{\bar{N}_1}{\bar{N}_2} \quad (6.12)$$

where \bar{N}_1 and \bar{N}_2 represent the mean number of one-hop MTs and that of two-hop UTs in each cell, respectively. Making the assumption that the MTs follow a spatially uniform distribution, then

$$\frac{\bar{N}_1}{\bar{N}_2} = \frac{A_B}{\sum_{i=1}^6 A_{F,i}} \quad (6.13)$$

where A_B is the area of the region where any MT will communicate over a one-hop link, and $A_{F,i}$ is the region where any MT establishes a two-hop link through the j th FRN. Given a fixed FRN positioning (i.e., d_{BF}/D), the right-hand-side of (6.13) can be transformed through simple geometric manipulation into

$$\frac{A_B}{\sum_{i=1}^6 A_{F,i}} = \frac{d_{BF}^2}{3D^2 - d_{BF}^2} = \frac{\left(\frac{d_{BF}}{D}\right)^2}{\left[3 - \left(\frac{d_{BF}}{D}\right)^2\right]} \triangleq \delta\left(\frac{d_{BF}}{D}\right) \quad (6.14)$$

where $\delta(d_{BF}/D)$ is defined for notational convenience.

If the average data rate achieved by one-hop UTs is denoted as \bar{R}_1 , then the mean aggregate data rate for one-hop MTs can be approximated as follows:

$$\bar{R}_{A,1} \approx \bar{N}_1 \bar{R}_1 \quad (6.15)$$

Referring to the Shannon capacity formula,

$$\bar{R}_1 \approx \frac{B_1}{N_1} E[\log(1 + \text{SINR}_{BM})] \approx \frac{B_1}{N_1} E[\log(1 + \text{SIR}_{BM})] \quad (6.16)$$

where $E[x]$ represents the expectation of random variable x , SINR_{BM} and SIR_{BM} are the signal-to-noise-and-interference ratio and signal-to-interference ratio on a link connecting a BS and a MT, respectively. The approximation in (6.16) considers an interference-limited environment.

The mean aggregate data rate of two-hop MTs is calculated considering the spectrum partitioning between the first-hop links and the second-hop links. It has been established that for a given data rate requirement on a point-to-point link, the minimum bandwidth can be achieved by balancing the data rates on different hops [17]. Heuristically, this allows for balancing the mean aggregate data rate on the first-hop links and that on the second-hop links or

$$B_{21} E[\log(1 + \text{SIR}_{FM})] = B_{22} E[\log(1 + \text{SIR}_{BF})] \quad (6.17)$$

where SIR_{FM} and SIR_{BF} represent the SIR on a link connecting an FRN and an MT and that connecting a BS and an FRN. Equation (6.17) also defines the mean aggregate data rate for the set of two-hop MTs, or

$$\bar{R}_{A,2} = B_{21} E[\log(1 + \text{SIR}_{FM})] = B_{22} E[\log(1 + \text{SIR}_{BF})] \quad (6.18)$$

From (6.15), (6.16), and (6.18):

$$\frac{\bar{R}_{A,1}}{\bar{R}_{A,2}} = \frac{B_1 E[\log(1 + \text{SIR}_{BM})]}{B_{21} E[\log(1 + \text{SIR}_{FM})]} \quad (6.19)$$

which, together with (6.12), (6.13), and (6.14), defines a linear relationship between B_1 and B_{21} :

$$B_1 E[\log(1 + \text{SIR}_{BM})] = \delta \left(\frac{d_{BF}}{D} \right) B_{21} E[\log(1 + \text{SIR}_{FM})] \quad (6.20)$$

For a given FRN positioning, it is possible to analytically derive $E[\log(1 + \text{SIR}_{BM})]$, $E[\log(1 + \text{SIR}_{FM})]$, and $E[\log(1 + \text{SIR}_{BF})]$, yet the procedure is tedious. Instead, through computer simulations it is possible to obtain experimental values, denoted as E_{BM} , E_{FM} , and E_{BF} , respectively. By jointly solving (6.11), (6.17), and (6.20), the following closed-form solution can be obtained:

$$B_1 = \frac{\left(\frac{E_{FM}}{E_{BM}} \right) \delta \left(\frac{d_{BF}}{D} \right)}{\left(\frac{E_{FM}}{E_{BM}} \right) \delta \left(\frac{d_{BF}}{D} \right) + \left(\frac{E_{FM}}{E_{BF}} \right) + 1} B \quad (6.21)$$

$$B_{21} = \frac{1}{\left(\frac{E_{FM}}{E_{BM}} \right) \delta \left(\frac{d_{BF}}{D} \right) + \left(\frac{E_{FM}}{E_{BF}} \right) + 1} B \quad (6.22)$$

$$B_{22} = \frac{\left(\frac{E_{FM}}{E_{BM}} \right)}{\left(\frac{E_{FM}}{E_{BM}} \right) \delta \left(\frac{d_{BF}}{D} \right) + \left(\frac{E_{FM}}{E_{BF}} \right) + 1} B \quad (6.23)$$

These equations demonstrate the dependence on the FRN positioning of the proposed spectrum partitioning scheme. Such dependence has motivated the investigations on relay positioning.

FRN positioning was investigated [13] through snapshot-based computer simulations of a 19-cell cellular environment. In each independent trial, the data rate achieved by each MT in the central base station is calculated by use of the Shannon capacity formula. By performing independent trials the CDF of the user data rate can be obtained. The unreliability on links connecting a BS and the associated FRNs is determined by assuming different path-loss exponents. In particular, $\gamma = 2.0$ is used to model the LOS case; $\gamma = 3.5$ is used to model the NLOS case; and $\gamma = 2.5$ and $\gamma = 3.0$ are two cases in between. All other links are assumed to have NLOS channel conditions. In addition, $B = 25.6$ MHz, $f_c = 5$ GHz, and $D = 500$ m.

Figure 6.19 shows the dependence on d_{BF}/D of both the mean user data rate and the value corresponding to the 5% point on the CDF curve. This figure shows the significant impact that FRN positioning has on system performance. With the above-defined criterion, the optimal FRN site is at a position where $d_{BF}/D = 0.2$ when the BS and the FRNs are communicating with NLOS links, at a position where $d_{BF}/D = 0.4$ when $\gamma = 3.0$, and at a position where $d_{BF}/D = 0.6$ for the remaining two cases. This observation shows that the better the quality of the links

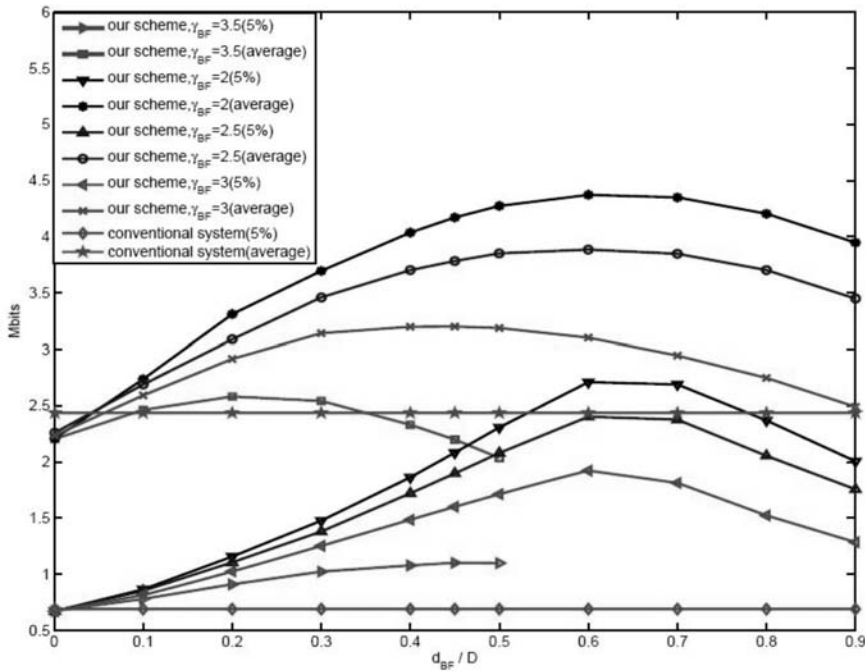


Figure 6.19 Dependence of average per-user data rate on FRN positioning, denoted as “average,” and per-user data rate at the 5% point of CDF, denoted as “5%” [13].

connecting the BS and the FRNs, the larger the distance between the FRN and the BS.

In addition, as shown in Figure 6.19, when a BS can communicate with FRNs over LOS links, a REC with the proposed scheme can offer significantly better performance over a conventional cell without relaying. Figure 6.19 also shows that the proposed scheme offers promising robustness against the channel unreliability on the BS-FRN links. In particular, in all simulated channel conditions, a REC can outperform a conventional cell as long as the proposed spectrum partitioning scheme is employed and the FRNs are placed with proper locations.

Figure 6.20 shows the dependence on the FRN positioning of the uplink average user throughput obtained through system level simulations [12, 13].

This figure shows both a heavy-load case and a light-load scenario. Note that FRN positioning has a more important impact on system performance when the system has a heavier loading level.

6.2.2 Adaptive Soft Resource Partitioning

Soft resource partitioning enables frequency reuse 1 and at the same time each RAP has high power resources with reduced interference available to schedule UTs in the border area with other RAPs because, instead of frequency reuse, power masks in the time or frequency domain are assigned to each RAP [13].

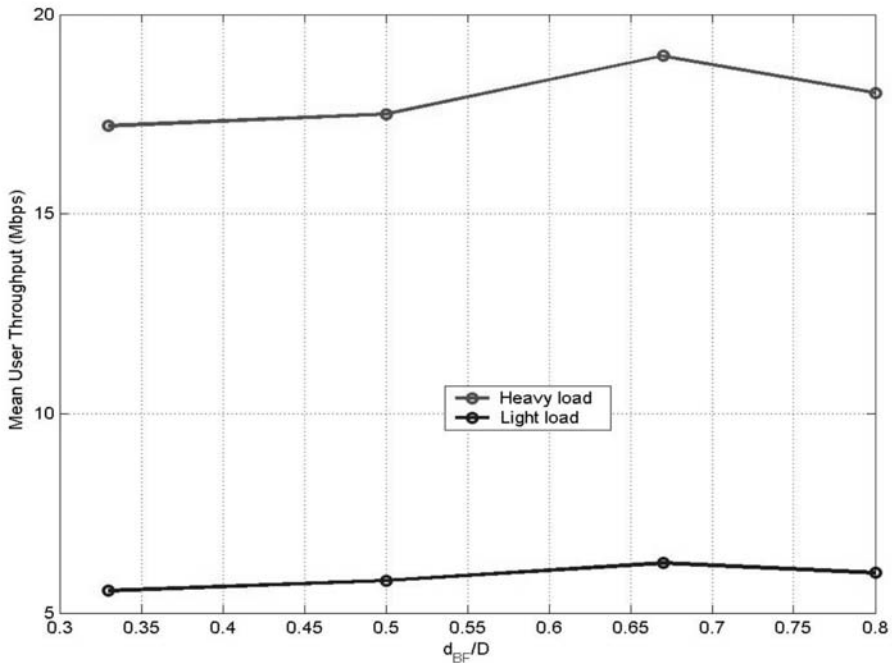


Figure 6.20 Average user throughput [13].

A REC can have multiple RNs, each providing different traffic loads to the network due to the cell size and the user density variations. Therefore, a soft-frequency reuse pattern with a single power mask for the whole cell, in which the high-power resources are then split between the RNs, does not make efficient use of the available radio resources. Thus, a local power mask adaptation is desirable. The IST Project WINNER [2] developed a mechanism for adapting the power mask based on the traffic load situation within the REC [5, 9, 12, 13].

The scheme is described in Figure 6.21 from the view point of RN1.

After receiving the messages from the neighboring RNs, RN1 decides in this example to request resources from RN6. The figure shows the power mask update; RN6 lowers its power level for the requested resource block (first block) to the power level that RN1 was using before the request, and RN1 uses the highest power level for the requested resource block. It is assumed that an initial power mask exists. It can be set by a network planning tool or it can be found by other means (e.g., the network learns its topology and sets the power masks accordingly). The operation is starting from this initial power mask and the RAPs will try to operate within this power mask.

Figure 6.22 shows the steps of the mechanism of the distributed radio resource partitioning scheme. The mechanism is explained here for RN1 in the scenario of Figure 6.22.

RN1 monitors its buffer level. There are three (one is optional) predefined thresholds:

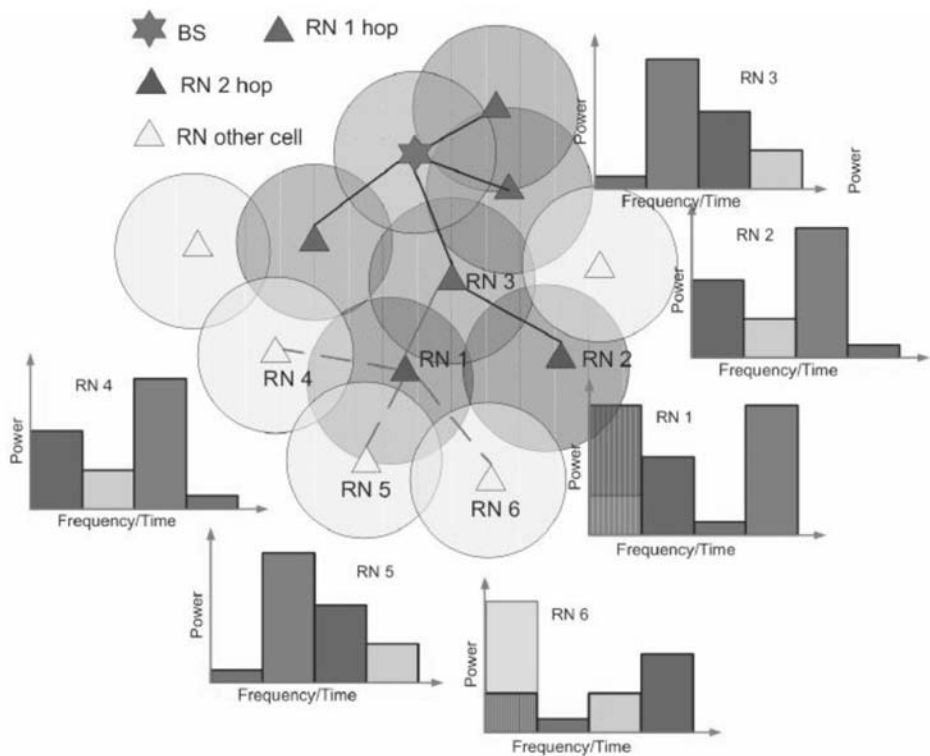


Figure 6.21 Scenario for adaptive soft resource partitioning [13].

- A threshold that if the exceeded triggers resource request = *max threshold*;
- A threshold, below which the operation is normal = *normal threshold*;
- A threshold that if the exceeded causes drop of flows (optional) = *drop threshold*.

The difference between the thresholds is large enough so that the resource requests happen only at a relatively slow timescale, (e.g., every second). In this way, the amount of signaling between the nodes involved in the resource request is reduced.

If the buffer level exceeds the *max* threshold, then RN1 sends a *resource request* message to neighbors RN3, RN4, RN5, and RN6. These RNs will answer to this message with a *resource status* message containing their buffer level and the available resources, (i.e., their current power mask).

RN1 will check the buffer levels from the other RNs and decide if it should take high-power resources from one of the other RNs. It will base the decision on the following criteria:

- Only resources from nodes where the buffer level is below the normal threshold for normal operation will be used;
- Resources from the node that has the lowest buffer level and for which the highest power gain (the difference of the new high power level and the

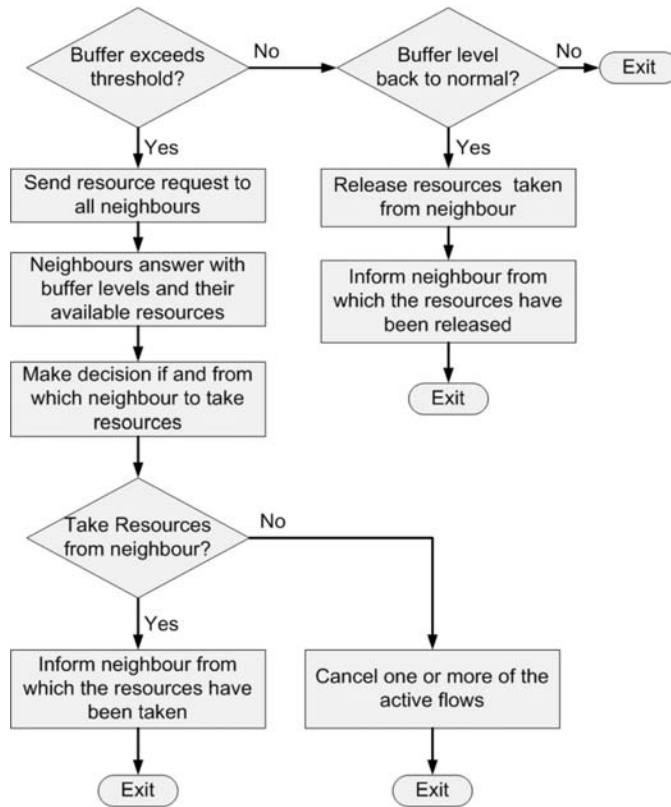


Figure 6.22 Flowchart of distributed radio resource partitioning scheme [13].

currently used power level) can be achieved. The resources are taken from the node with the highest u :

$$u = a * (\text{normal_threshold} - \text{buffer_level}) + b * \text{power_difference}$$

RN 1 will then signal to the node, from which it takes the resources that it should now use a lower power level (i.e., the power level that RN1 was using before). RN6 acknowledges the resource update.

Figure 6.23 shows such an updated power mask. In this case RN1 decides to use the high power level of RN6 in the first resource block. After signaling the updated power mask to RN6, RN1 uses the highest power level and RN6 uses the power level that RN1 was using in the first resource block.

If no resource can be taken from the neighboring nodes, RN1 will drop one of the active flows it is serving. RN1 drops the flow that has the lowest average throughput per transmission resource and the lowest recent throughput per high-power transmission resource (e.g., in the last 5 seconds).

$$v = a * \text{avg_throughput_per_resource} + b * \text{recent_throughput_per_resource}$$

If the buffer level drops below the normal operational threshold, RN1 returns to the initial power mask. It will signal all of the RNs from which it has taken

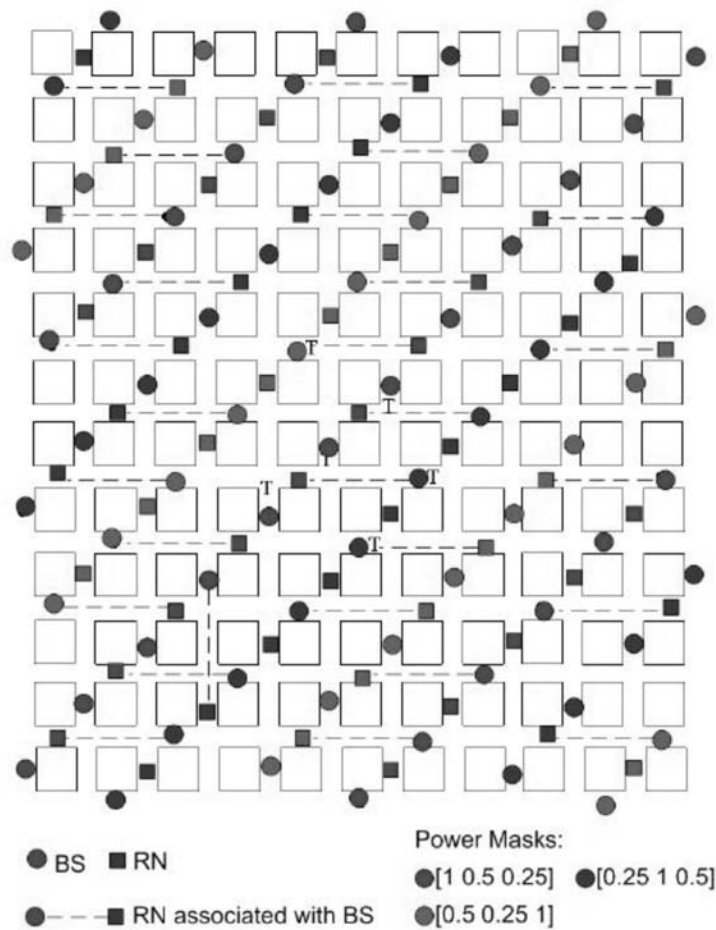


Figure 6.23 Power mask assignment during soft resource partitioning [13].

power resources, and they will update their power mask to the initial values for these resources.

Figure 6.23 shows the power mask assignment of the simulation scenario that assumes metropolitan area [18].

Both the BS and the RNs are placed below the rooftop at a height of 10m. An OFDM system with an exponential effective SIR mapping (EESM) link to system mapping has been simulated by a dynamic event-driven simulator that simulates UL and DL directions simultaneously with OFDM symbol resolution [13].

Figure 6.24 shows the user throughput as a function of distance from the BS. The user throughput for the hard-frequency reuse is much less than that of the user throughput for the soft-frequency reuse. In addition, the user throughput as a function of distance is more evenly distributed for the adaptive mask than it is for the fixed mask. Especially, users further away from the BS benefit from the power mask adaptation. In the simulated Manhattan grid scenario, user throughput will not only depend on the distance from the BS but also on other criteria, such as whether the user is in the same street as the serving RAP or in an interference

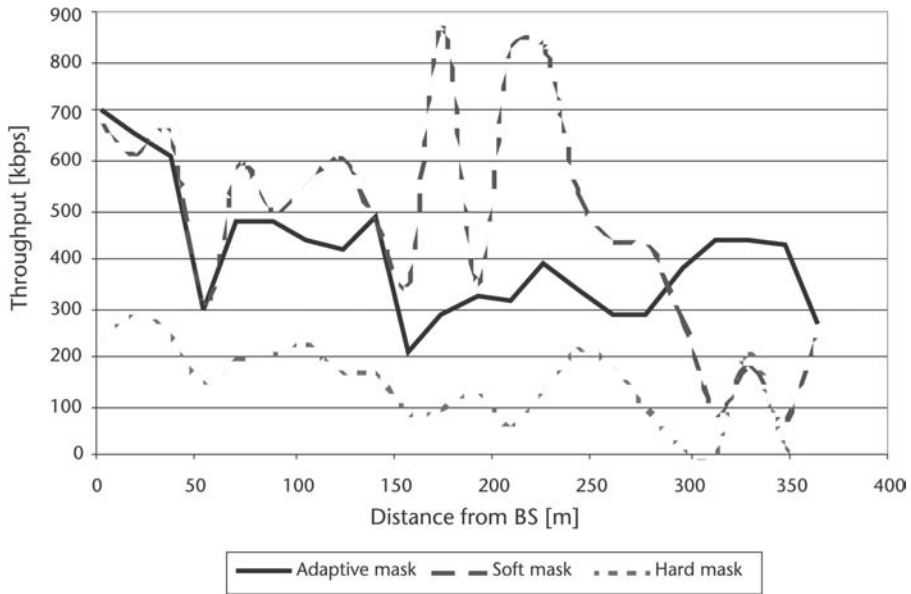


Figure 6.24 User throughput as a function of distance from the BS [13].

situation. Nevertheless, the results provide some insight into the properties of the power mask adaptation.

Figure 6.25 shows the SINR of the data packets received by the MTs. Hard-frequency reuse significantly improves the SINR of the received packets compared to the soft-frequency reuse strategies. Secondly, the power mask adaptation can improve the SINR of the received packets compared to the fixed power mask.

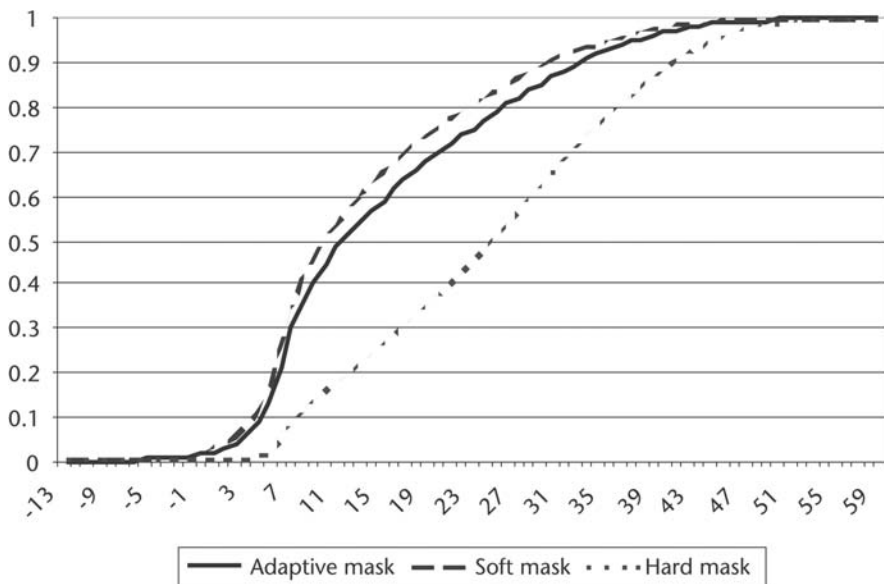


Figure 6.25 SINR of the received packets during soft resource partitioning [13].

Even though hard-frequency reuse can significantly improve the SINR of the received packets, system performance is still significantly worse. This is due to the fact that only one-third of the overall resources are available to each RAP. Soft-frequency reuse enables frequency reuse 1, while still allowing for interference coordination by using power masks instead of hard-frequency reuse. Further, power mask adaptation can improve the throughput, especially for low-throughput users. The user throughput is more evenly distributed within the cell when power mask adaptation is used.

Other ways to increase the cell capacity are through enabling intra-REC reuse of resources by use of smart-antenna technologies.

This can be done by allowing the BS only to enable parallel transmissions on identical resources (e.g., SDMA with beamforming) between the BS and the RNs, thereby reducing the overhead of the relay links. The fixed nature of the RNs helps a smart-antenna-equipped BS to deploy relatively precise beams with negligible error in the DoA/DoD estimation. The channel is considered relatively stable; that is, channel feedback is required less often. An example of the beam pattern assumed for a wide-area deployment with three FRNs is shown in Figure 6.26.

The patterns were generated by a nine-element uniform circular array, and a nonsectorized setup has been used. Reference [19] suggests a four-element uniform linear array (see Chapter 5) for sectorized deployments.

In another concept, smart antennas can be assumed to be present at both BSs and RNs. Beamforming is used and spatially compatible (sets of) beams at different RAPs are identified and earmarked for concurrent parallel usage of the same resources. Full details on the described concepts are available in [13].

6.2.3 Multihop Relay Scenarios

A multihop relay (n -hop, $n \geq 2$) scenario is an extension of the two-hop scenario that combines a linear deployment of RNs, from a logical viewpoint, along different branches; and PMP last-hop connections toward UTs around each RN. This is shown in Figure 6.27.

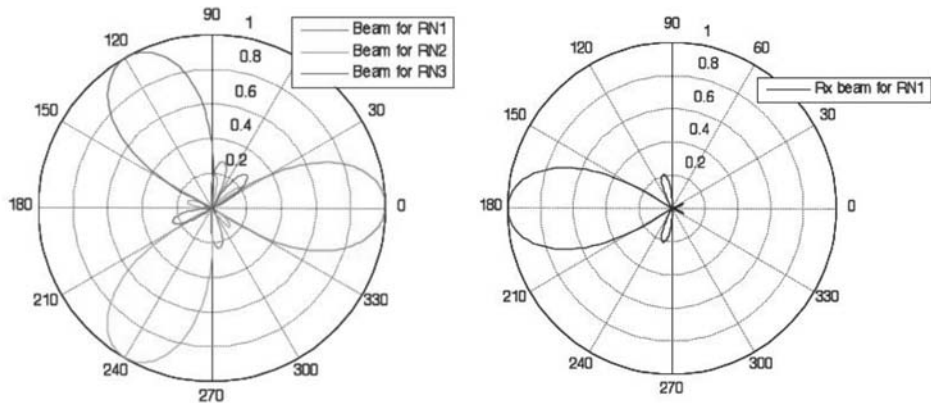


Figure 6.26 (Left) Tx beam patterns for the BS when serving the FRNs and (right) Rx beam pattern for one FRN.

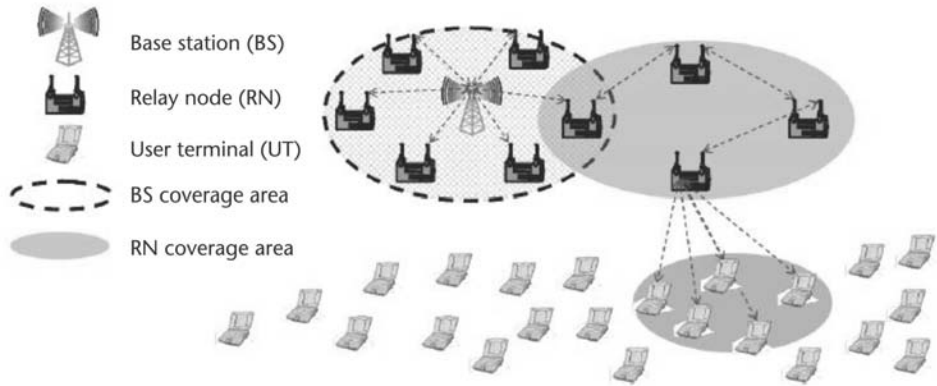


Figure 6.27 Example of a multihop relay scenario [13].

The logical topology can represent a wide variety of physical deployments, taking into account that the linear deployment of RNs is merely a logical requirement, not a physical one. Multihop traffic is transmitted among one BS, which is connected to a fixed backbone network, and fixed RNs that are strategically placed. The last-hop traffic takes place between the RN and a variable number of UTs. Note that multihop and last-hop links are discriminated in the frequency domain.

For typical multihop scenarios the overall system capacity, cell coverage, end-to-end delay, and fairness with respect to an increasing number of hops have to be addressed.

A novel resource request and allocation scheme in the multihop relaying scenario (n -hop, $n \geq 2$), named *connection based scheduling with differentiated service support* (CbS-DSS) has been proposed [20]. The CbS-DSS mechanism is able to satisfy different QoS requirements for different service classes (e.g., nonreal-time and real-time traffic). Requests for different service classes are distinguished in order to allow the BS to assign grants according to QoS constraints.

Figure 6.28 shows how real-time and nonreal-time traffic efficiently share the available radio resources by use of CbS-DSS if the system is not overloaded. In overloading conditions, the real-time traffic exploits radio resources initially

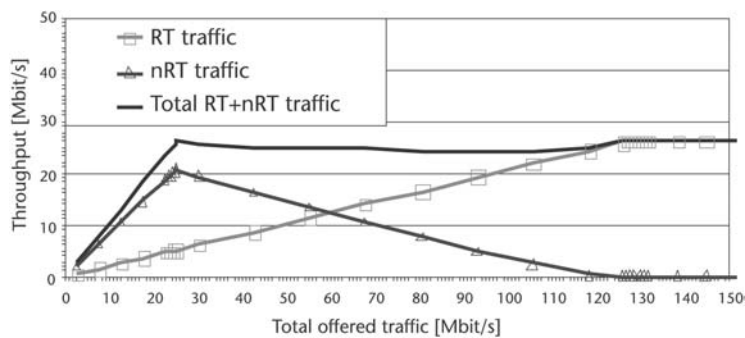


Figure 6.28 Uplink throughput versus total offered traffic, whereas the load at each RN (RN1, RN2, RN3, and RN4) is composed of 20% real-time traffic and 80% nonreal-time traffic [13].

assigned to nonreal-time traffic. The number of hops does not affect the delay and fairness with respect to the number of hops between the BS and RN when employing CbS-DSS.

6.3 Advanced Antenna Techniques for Relaying

6.3.1 Cooperative Relaying

Cooperative transmission is based on pioneering work on the relay channel [21]. The main results were obtained for a degraded channel taking into account that the relay can work in full-duplex mode (i.e., the relay is receiving data and transmitting it at the same time). However, this assumption requires reconsideration in view of the device evolution [7]. Therefore, the original work had been extended to half duplex (i.e., the relay receives the data and transmits in an orthogonal way, e.g., TDD) [22], showing that it is possible to obtain some gains in terms of diversity.

The different possibilities that the source and relay collaborate to transmit to the destination can be grouped in three protocols [7] as shown in Figure 6.29:

- *Protocol I:* The source communicates with the relay and the destination during the first time slot (solid lines in Figure 6.29). In the second time slot, the relay terminal communicates with the destination (dashed line). This protocol shows the same structure as the broadcast channel (BC) and can be considered for DL transmissions in a centralized wireless networks.
- *Protocol II:* In the first slot the source transmits to the relay (solid line). Afterwards, in the second time slot, both transmit to the destination (dashed lines). In this case the channel becomes a MAC with the user and relay as transmitters and one destination. This protocol is useful for UL transmissions in a cellular network.
- *Protocol III:* This protocol is a mixture of the previous protocols. During the first slot, the source transmits to the relay and the destination (solid lines), whereas in the second time slot, the source and the relay transmit to the destination (dashed lines). This protocol can achieve better spectral

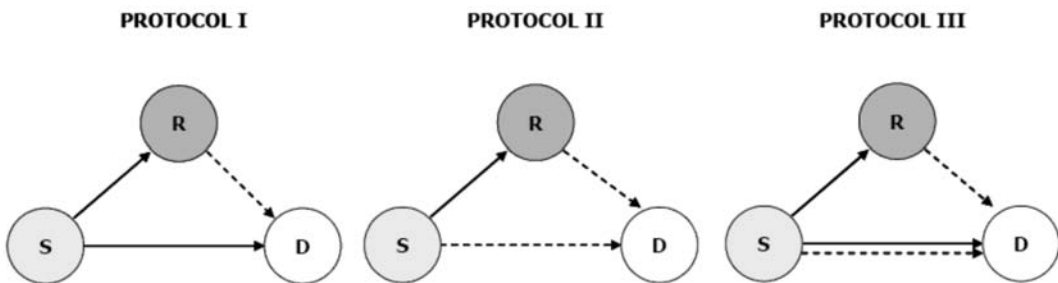


Figure 6.29 Cooperative transmission protocols in orthogonal access. In pure relaying protocols, there is no direct link between the S and D terminals [7].

efficiency in outage and ergodic capacity than the previous protocols and can be considered for ad hoc networks.

In the following sections, capacity equations are derived only for the first protocol, although the equations for the other two can be easily derived.

6.3.1.1 Cooperative Amplify and Forward

In this case, the relay is using the same number of frequency bands and the same time-slot duration as the source. By stacking the two received vectors at the destination, through both the direct link (S - D link) and the second hop R - D link, the received signal is obtained as follows:

$$\begin{bmatrix} \mathbf{y}_{S-D}^k \\ \mathbf{y}_{R-D}^k \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{H}_{S-D}^k \\ \mathbf{H}_{R-D}^k \mathbf{G}^k \mathbf{H}_{S-R}^k \end{bmatrix}}_{\mathbf{H}_{RC}^k} \mathbf{x}^k + \begin{bmatrix} \mathbf{I}_N & 0 & 0 \\ 0 & \mathbf{H}_{R-D}^k \mathbf{G}^k & \mathbf{I}_N \end{bmatrix} \begin{bmatrix} \mathbf{w}_{S-D}^k \\ \mathbf{w}_{S-R}^k \\ \mathbf{w}_{R-D}^k \end{bmatrix} \quad (6.24)$$

If the covariance matrix of the noise of the amplify and forward (AF) is defined as

$$\mathbf{R}_w^k = \begin{bmatrix} \mathbf{R}_{w,S-D}^k & 0 \\ 0 & \mathbf{H}_{S-D}^k \mathbf{G}^k \mathbf{R}_{w,S-R}^k \mathbf{G}^k (\mathbf{H}_{R-D}^k)^H + \mathbf{R}_{w,R-D}^k \end{bmatrix} \quad (6.25)$$

then the amplify and forward mutual information is

$$I_{AF} = \frac{1}{2} \frac{1}{N_C} \sum_{k=1}^{N_C} \log_2 |\mathbf{I}_{2N} + \mathbf{H}_{AP}^k \mathbf{R}_x^k (\mathbf{H}_{AP}^k)^H (\mathbf{R}_w^k)^{-1}| \quad [\text{bits/s Hz}^{-1}] \quad (6.26)$$

Note that the equivalent channel matrix \mathbf{H}_{AP}^k has dimension N : $N_S \times 2N_D$. When $N_S > 2N_D$, an $N_S \times 2N_D$ MIMO system is able to double the capacity of an $N_S \times N_D$ system (except for the fact that there is a penalty term 1/2 that comes from the capacity losses of using orthogonal relaying). Therefore, under suitable coding and channel conditions, substantial capacity improvement can be expected for the AF cooperative scheme compared to an AF pure relaying approach [7].

6.3.1.2 Cooperative Decode and Forward

Depending on the signal transmitted by the relay, two possible cooperative techniques can be applied: the repetition code (RC), in which the relay retransmits the same signal (possibly using the same STBC), or the unconstrained code (UC), in which the relay retransmits an uncorrelated codeword using a different STBC and/or puncturing matrices when convolutional codes are considered.

6.3.1.2.1 Decode and Forward Transmission with Repetition Coding (DF-RC).

The signal received by the repetition code is given by:

$$\begin{bmatrix} \mathbf{y}_{S-D}^k \\ \mathbf{y}_{R-D}^k \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{H}_{S-D}^k \\ \mathbf{H}_{R-D}^k \sqrt{\frac{P_R^k}{P_S^k}} \end{bmatrix}}_{\mathbf{H}_{RC}^k} \mathbf{x}^k + \begin{bmatrix} \mathbf{w}_{S-D}^k \\ \mathbf{w}_{R-D}^k \end{bmatrix} \quad (6.27)$$

For the DF-RC scheme, the time devoted to the relay transmission has to be the same as the transmission from the source (and also the number of frequency bands) [7].

From (6.27), the mutual information on the cooperative link is the sum of the mutual informations on the S - D and R - D direct links:

$$I_{C-RC} = \frac{1}{2} \frac{1}{N_C} \sum_{k=1}^{N_C} \log_2 |\mathbf{I}_{2N_D} + \mathbf{H}_{RC}^k \mathbf{R}_x^k (\mathbf{H}_{RC}^k)^H (\mathbf{R}_w^k)^{-1}| \quad (6.28)$$

The equivalent channel matrix \mathbf{H}_{RC}^k has dimension $N_S \times 2N_D$. There is a limitation, however, due to the fact that the relay needs to recover the symbols transmitted from the source. The overall scheme achieves the following mutual information:

$$I_{RC} = \min(I_{S-R}^{DiL}, I_{C-RC}) \quad (6.29)$$

with I_{S-R}^{DiL} being the mutual information of the source-relay link, and with dimension $N_S \times N_D$ given by

$$I_{S-R}^{DiL} = \frac{1}{2} \frac{1}{N_C} \sum_{k=1}^{N_C} \log_2 |\mathbf{I}_{N_R} + \mathbf{H}_{S-R}^k \mathbf{R}_x^k (\mathbf{H}_{S-R}^k)^H (\mathbf{R}_{w,S-R}^k)^{-1}| \quad (6.30)$$

6.3.1.2.2 Decode and Forward Transmission with Unconstrained Coding (DF-UC).

An alternate transmission scheme for DF is obtained when the signal transmitted from the relay station adopts the most suitable unconstrained channel coding scheme. For the DF-UC, the signal model is given by

$$\begin{bmatrix} \mathbf{y}_{S-D}^k \\ \mathbf{y}_{R-D}^k \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{S-D}^k & 0 \\ 0 & \mathbf{H}_{R-D}^k \sqrt{\frac{P_R^k}{P_S^k}} \end{bmatrix} \begin{bmatrix} \mathbf{x}_S \\ \mathbf{x}_R \end{bmatrix} + \begin{bmatrix} \mathbf{w}_{S-D}^k \\ \mathbf{w}_{R-D}^k \end{bmatrix} \quad (6.31)$$

From (6.30) and (6.31) the mutual information on the cooperative link is the sum of the mutual information on the S - D and R - D direct links as in

$$I_{C-VC} = T_1 \sum_{k=1}^{N_{C1}} \log_2 \left| \mathbf{I}_N + \mathbf{H}_{S-D}^k \mathbf{R}_{x,S}^k (\mathbf{H}_{S-D}^k)^H (\mathbf{R}_{w,S-D}^k)^{-1} \right| \quad (6.32)$$

$$+ T_2 \sum_{k=1}^{N_{C2}} \log_2 \left| \mathbf{I}_N + \mathbf{H}_{R-D}^k \mathbf{R}_{x,R}^k (\mathbf{H}_{R-D}^k)^H (\mathbf{R}_{w,R-D}^k)^{-1} \right|$$

Based on (6.32), the overall scheme achieves the following mutual information:

$$I_{VC} = \min(I_{S-R}^{DiL}, I_{C-VC}) \quad (6.33)$$

where

$$I_{S-R}^{DiL} = T_1 \sum_{k=1}^{N_{C1}} \log_2 \left| \mathbf{I}_{N_R} + \mathbf{H}_{S-R}^k \mathbf{R}_{x,S}^k (\mathbf{H}_{S-R}^k)^H (\mathbf{R}_{w,S-R}^k)^{-1} \right| \quad (6.34)$$

In this case it is possible to adapt the time duration and the frequency bands devoted to both transmissions (i.e., relay transmission and transmission from the source). The total number of radio resource units employed is $T_1 \cdot N_{C1} + T_2 \cdot N_{C2}$, whereas for the AF and the DF with RC it is $2 \cdot T \cdot N_C$.

6.3.1.2.3 Cell Capacity Gains.

The example investigated in [7] considers the performance of a single-carrier system with a relay placed halfway between the BS and the cell edge. Figure 6.30 shows the ergodic capacity in bit/s Hz⁻¹ for a single-carrier system, considering a direct

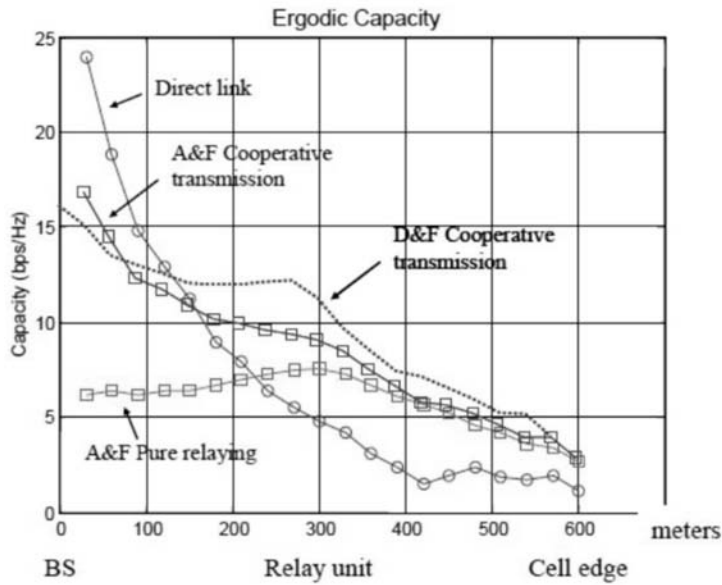


Figure 6.30 Ergodic capacity (bit/s Hz⁻¹) versus distance from the BS, for DL transmissions under different forwarding and cooperative strategies [7].

transmission, AF pure relaying, AF cooperative transmission, and DF-UC cooperative transmission. For a fair comparison, the same overall power is considered for all of the schemes.

The number of antennas is two for the source BS and one for the subscriber station. The number of antennas in the relay station is one for the AF approach and two for the DF. These numbers were chosen in order to prevent the source-RS link from limiting the performance of the DF cooperative scheme when used in the downlink. The relay unit is placed at a distance to the BS equal to half the maximum radius of the cell (i.e., 300m from the BS). The relay station is a lamppost in LOS with the BS. In Figure 6.30, eight possible RSs around the BS were considered, and the SS selects the nearest lamppost as the relay terminal.

In relaying/cooperative schemes the effect of having less capacity than the *DiL* when the subscriber stations are close to the BS is due to: (1) the 1/2 factor (two transmissions instances are required), and (2) the fact that the RS offers limited help if it is not close to the destination, compared to the capacity obtained from a direct link. Note that the capacity improvement due to the cooperation compared with pure relaying compensates partly for the efficiency loss when distances between the source and the destination are close. Furthermore, when the destination is far from the source, cooperation achieves a substantial improvement in capacity compared to direct transmission. Moreover, AF offers an acceptable performance with the lowest complexity (it amplifies and forwards the received signal).

Cooperative techniques behave, therefore, like a MIMO system with twice the number of receiving antennas. Therefore, for those systems with a number of transmitting antennas equal to or greater than $2N_D$, capacity can be increased at a maximum of twice the capacity without cooperation [7]. Nevertheless, due to the efficiency loss with 1/2, for AF no multiplexing gain can be obtained unless other considerations are considered, for example, the spatial reuse of the relaying slots when the relays are other users instead of fixed relays.

Spatial reuse allows capacity improvements under the following rationale. In the following, the downlink, with the BS serving K users is considered. Initially, the K DL transmissions are allocated. Eventually, K' simultaneous retransmissions from the corresponding relays are allocated in a single slot, with $K' \leq K$. Therefore, the effective capacity of a single cooperative connection has to be multiplied by a factor $K/(K + 1)$, instead of a factor 1/2, corresponding to the relay slot nonreuse case. This amounts to saying that, in a wireless communication scenario, if the RRM strategy allows K to be high enough, the term $K/(K + 1)$ approaches 1.

When reuse of the relay slot by multiple cooperating users is considered, the interfering power from other relays is also considered. Therefore, the design power control algorithms must minimize the interference at the relay link and maximize the total mutual information at the same time [7].

Figure 6.31 considers a scenario in which other subscriber stations may act as a relay. The user density is 10^{-4} m^{-2} .

This case considers spatial reuse of the relay slot: The relay and destination are close and the interference in the relay slot is negligible. The power in the relay slot is adjusted under a distributed game theory-based algorithm to minimize the interference to other relay destination pairs.

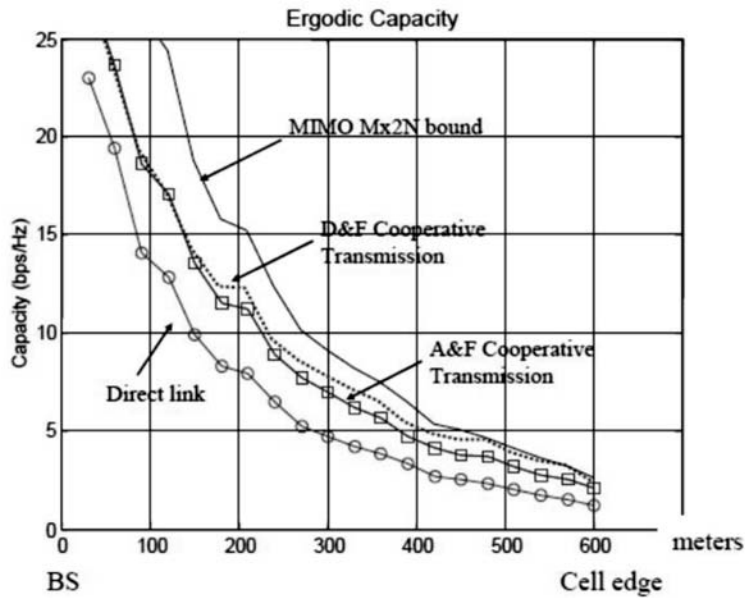


Figure 6.31 Ergodic capacity (bit/s Hz^{-1}) versus distance from the BS, for DL transmissions under different forwarding and cooperative strategies. The relay is an idle subscriber station close to the destination [7].

Because the efficiency loss is smaller than $1/2$, the cooperation offers a capacity improvement for all positions in the cell. In this case, cooperative AF offers a capacity very close to that of cooperative DF-UC, despite the fact that the AF uses only one antenna at the RS, while the DF requires two antennas.

6.3.1.2.4 Geometry-Dependent Capacity Gains.

Figures 6.32, 6.33, and 6.34 study the outage cooperative capacity in a multicarrier system for different relative positions between the source-relay-destination in three particular cases. A high spatial reuse is considered with negligible interference in the relay slot.

In all cases the same assignment of resources for the first and second hop was chosen. SNR_0 denotes the quality of the direct link between source and destination, SNR_1 is the quality of the first hop (between source and relay), and SNR is the quality of the second hop.

For a high-SNR source-relay link (as can be the case if the relay is close to the source), the DF seems to be the best cooperative technique, with its potential for larger capacity gains than MIMO direct transmission. On the other hand, when the relay is close to the destination, AF provide better results.

For the DF-UC, an alternate way to control the efficiency loss for a single communication is by adjusting the time and frequency between the first transmission (source to destination and relay) and the second transmission (relay-destination) for the DL cooperation. In such a case, some multiplexing gains can be obtained even if no spatial reuse is considered.

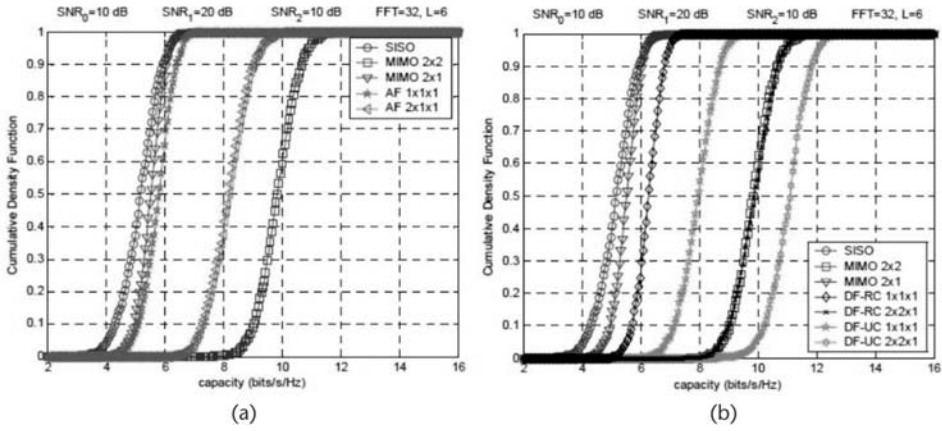


Figure 6.32 Case 1: relay close to the source; mutual information for the (a) AF and (b) DF-UC. $SNR_0 = 10$ dB, $SNR_1 = 20$ dB, $SNR_2 = 10$ dB (relay close to the source). $L = 6$ channel taps and $N_c = 32$ [7].

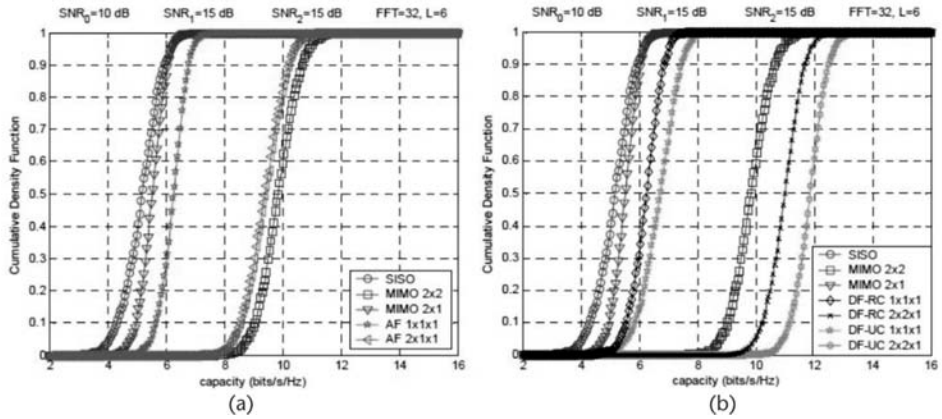


Figure 6.33 Case 2: relay halfway between the source and destination; mutual information for the (a) AF and (b) DF-UC. $SNR_0 = 10$ dB, $SNR_1 = 15$ dB, $SNR_2 = 15$ dB (relay between source and destination). $L = 6$ channel taps and $N_c = 32$ [7].

6.3.1.2.5 Application to the Downlink of an 802.16-Like System.

The potential improvement brought about by applying cooperative DF relaying in the context of the DL of an 802.16-like system has been assessed [7]. The improvement was measured in terms of the increase of ergodic capacity, assuming a single SS and a typical urban cellular deployment. Then, the performance of the various time-domain cooperative relaying protocols of type I, II, and III was investigated for a deployment consisting of a single cell of radius r_{cell} with a BS surrounded by Nr_{RS} . The BS-RS distance is denoted d_{BS-RS} . The angular separation between relays is fixed to $2\pi N_r$. In a first step, the cell is assumed isolated and a single SS is considered. Figure 6.35 considers cooperative transmissions involving a single relay.

The source transmits with the same power through different carriers. Then, the mutual information of the direct link is given by

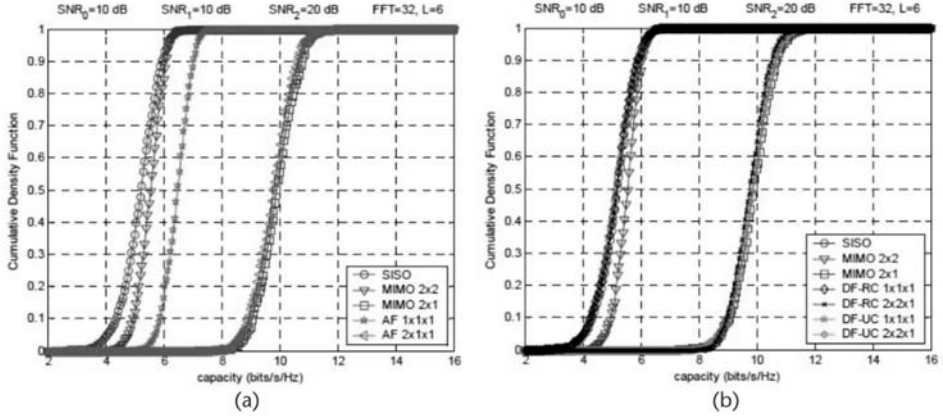


Figure 6.34 Case 3: relay close to the destination; mutual information for the (a) AF and (b) DF-UC. $\text{SNR}_0 = 10$ dB, $\text{SNR}_1 = 10$ dB, $\text{SNR}_2 = 20$ dB (relay close to the destination). $L = 6$ channel taps and $N_c = 32$ [7].

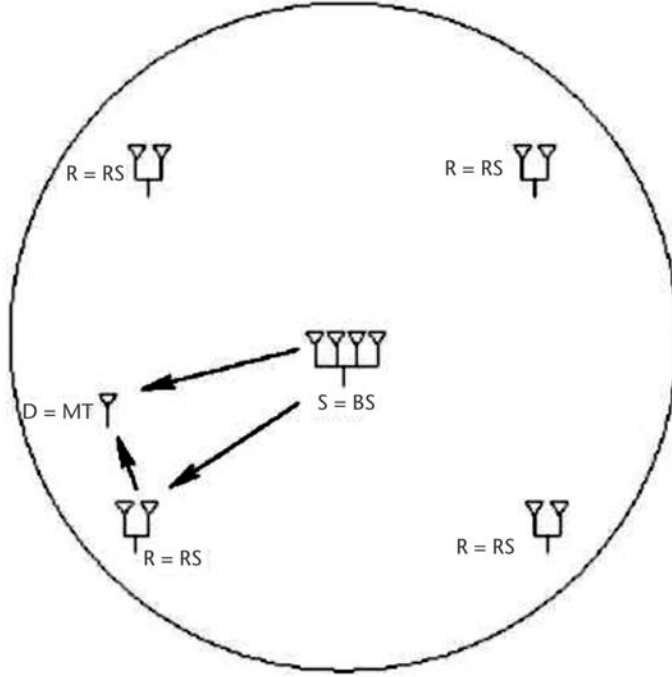


Figure 6.35 Deployment of cooperative relaying in the downlink [7].

$$I_{S-D}^{DiL} = \frac{1}{T_S} \sum_{k=1}^{N_C} \log_2 \left| I_{N_D} + \frac{P_S}{N_S} \mathbf{H}_{S-D}^k (\mathbf{H}_{S-D}^k)^H \mathbf{R}_{w,S-D}^{-1} \right| \quad (6.35)$$

In the downlink, the same number of subcarriers is assigned for both hops, as in an OFDM system. A central scheduler selects the time sharing between the

two hops, which maximizes the capacity. In this case the conventional two-hop forwarding capacity is given by

$$C_{\text{conv}} = \max_{0 < t < 1} \left(\min(tI_{S-R}^{DtL}, (1-t)I_{R-D}^{DtL}) \right) \quad (6.36)$$

For Protocol I, the transmissions from the BS and RS are orthogonal. The relay retransmits an uncorrelated codeword, which maximizes the mutual information without transmitting the CSI. An example of an implementation that is similar to this scheme is the distributed turbocodes scheme, in which the source encodes the packet with a first constituent encoder, and the relay re-interleaves the decoded information word and re-encodes it with a new constituent encoder [23].

The overall scheme achieves the following mutual information:

$$C_{\text{coopDF},1} = \max_{0 < t < 1} \left(\min(tI_{S-D}^{DtL} + (1-t)I_{R-D}^{DtL}, tI_{S-R}^{DtL}) \right) \quad (6.37)$$

From (6.36) and (6.37), we can see that the mutual information of this scheme is always greater than that of noncooperative DF and direct link (the latter being achieved when $t = 1$). However, both cooperative and noncooperative DF are limited by the capacity of the S-R link.

For Protocol II, the transmissions from the BS and the RS are simultaneous during the second time slot, and the destination does not listen to the first slot. This scheme would be the only one applicable to IEEE802.16j, because it does not require modification of the legacy SS. The capacity is given by

$$C_{\text{coopDF},2} = \max_{0 < t < 1} \left(\min((1-t)I_{R-D}^{DtL}, tI_{S-R}^{DtL}) \right) \quad (6.38)$$

Finally, Protocol III outperforms both Protocols I and II by allowing simultaneous transmission during the second slot and by letting the SS also receive during the first slot. The resulting capacity is

$$C_{\text{coopDF},3} = \max_{0 < t < 1} \left(\min(tI_{S-D}^{DtL} + (1-t)I_{R-D}^{DtL}, tI_{S-R}^{DtL} + (1-t)I_{R-D}^{DtL}) \right) \quad (6.39)$$

In a typical city center or suburban scenario, an expensive BS is equipped with a large number of antenna elements (e.g., eight antenna elements per BS sector). A few RSs are deployed in the cell on rooftops or on top of lampposts, with LOS propagation to the BS. Finally, the SS is typically equipped with one (e.g., handset) or two (e.g., PDA) antennas. Such assumptions can be exploited to obtain an approximation of the capacity formulas versus the average SNR on each link and to better understand the dependence of each scheme on the link SNRs.

Assuming $N_{BS} \gg N_{RS} \gg N_{SS}$, the capacity of Protocol I can be approximated as follows:

$$C_{\text{coopDF}1} = \max_{0 < t < 1} \left(\min(tN_{RS} \log_2(1 + \text{SNR}_{S-R}) \cdot tN_{MT} \log_2(1 + \text{SNR}_{S-D}) + (1 - t)N_{MT} \log_2(1 + \text{SNR}_{R-D})) \right) \quad (6.40)$$

Therefore, if SNR_{S-R} is large enough, the BS-RS link becomes much better than the rest of the links, so the capacity is limited by the right-hand term of (6.40) and becomes

$$C_{\text{coopDF}1} \approx N_{MT} \max_{0 < t < 1} \left(t \log_2(1 + \text{SNR}_{S-D}) + (1 - t) \log_2(1 + \text{SNR}_{R-D}) \right) \quad (6.41)$$

Likewise, the capacity of the conventional relaying scheme becomes

$$C_{\text{conv}} \approx \max_{0 < t < 1} \left(\min(tN_{RS} \log_2(1 + \text{SNR}_{S-R}), (1 - t)N_{MT} \log_2(1 + \text{SNR}_{R-D})) \right) \quad (6.42)$$

Similar to the cooperative relaying scheme, the capacity is limited by the right-hand term of (6.42), which only depends on the R-D (RS→SS) link. Therefore, the optimization results in a value of the parameter t close to zero, and the following capacity:

$$C_{\text{conv}} \approx N_{MT} \log_2(1 + \text{SNR}_{R-D}) \quad (6.43)$$

The capacity of the direct link is approximately

$$C_{DL} \approx N_{MT} \log_2(1 + \text{SNR}_{S-D}) \quad (6.44)$$

Equations (6.42), (6.43), and (6.44) can be used to understand the relative capacity of the cooperative protocols. This is shown in Figure 6.36.

When the SNR on the S-D (BS→SS) link is much higher than the R-D (RS→SS) link (lower-right corner), direct link communication is the best conventional technique, and cooperative DF cannot outperform it. When the SNR on the R-D (RS→SS) link is much higher than the S-D (BS→SS) link (upper-left corner), conventional DF relaying is the best technique, and cooperative transmission does not improve the capacity. When the SNR on the R-D (RS→SS) link and on the S-D (BS→SS) links are equivalent (along the diagonal), a capacity gain of up to 20% can be achieved.

The impact of cooperative relaying in the DL was further studied in a realistic deployment scenario, where the simulations investigated the maximum average mutual information that can be achieved as a function of the position of the SS in the cell. The path-loss model that was used was the WINNER SCME urban micro-cell model [18]. The BS and RS are both equipped with four antenna elements, but the BS has a three-sector antenna gain of 15 dB and a 3-dB beamwidth of 70°, whereas the RS is assumed to be equipped with an omnidirectional antenna having a 5-dB gain (thanks to limited elevation). A single RS was positioned per sector at

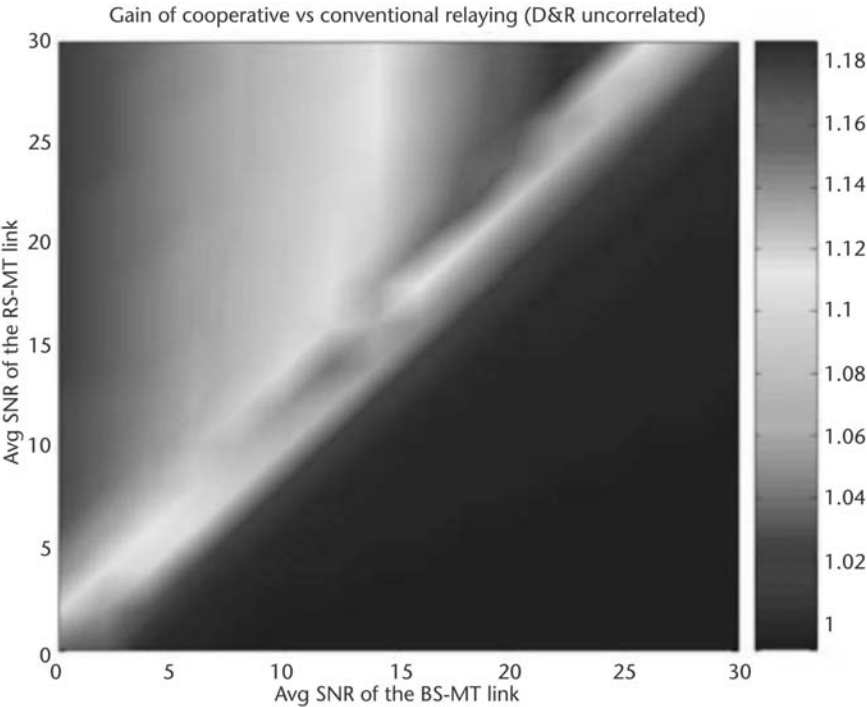


Figure 6.36 Relative capacity gain of cooperative DF Protocol I versus conventional techniques (best of direct link and conventional relaying) in an $8 \text{ BS} \times 4 \text{ RS} \times 1 \text{ SS}$ antenna configuration, with a high SNR = 30 dB on the BS-RS link and for different values of average SNR on the BS-SS and RS-SS links [7].

a distance equal to 60% of the cell range. The BS-RS link is assumed to be LOS, but the BS-SS and RS-SS links are NLOS. The SS has two antenna elements (it is difficult to envision more antennas in handsets) and only a 1-dB antenna gain. Also, the SNR did not exceed 30 dB in order to account for the RF impairments (e.g., phase noise, nonlinearities) of any realistic transceiver.

The objective was to understand the maximum average mutual information as a function of the position of the SS in the cell. Therefore, all subcarriers were allocated to the SS. A minimum SNR requirement of -6 dB was assumed thanks to cyclic delay diversity at the BS, MRC at the SS, the use of robust coding (e.g., rate 1/3 turbocode), and modulation (e.g., BPSK). This requirement determines the cell range, which is computed to ensure 75% coverage probability at the cell edge.

Figure 6.37 shows the mutual information for the direct link, averaged over 200 independent shadowing realizations.

This figure shows the problem encountered by modern cellular systems, which offer a wide range of bit rates selected by an AMC algorithm: A very high “peak throughput” greater than 100 Mbps is achievable only in a very small area around the BS, but in the greater part of the cell less than 50 Mbps (down to 20 Mbps at cell edge) is achievable.

Conventional DF relaying creates areas around the RS where a relatively high throughput (>60 Mbps) is achievable, as shown in Figure 6.38. The range of these

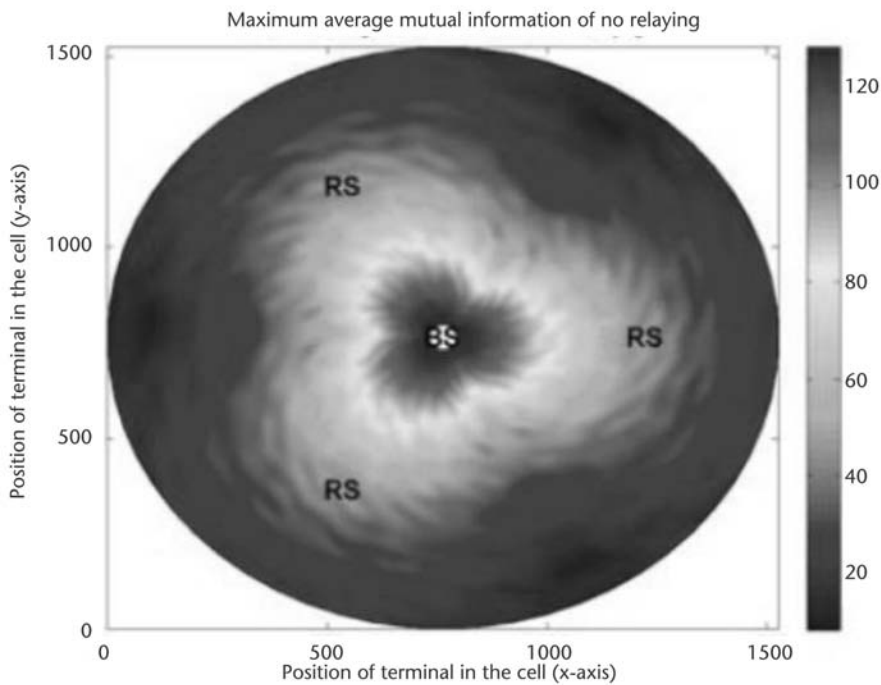


Figure 6.37 WMAN maximum DL capacity versus SS position, direct link [7].

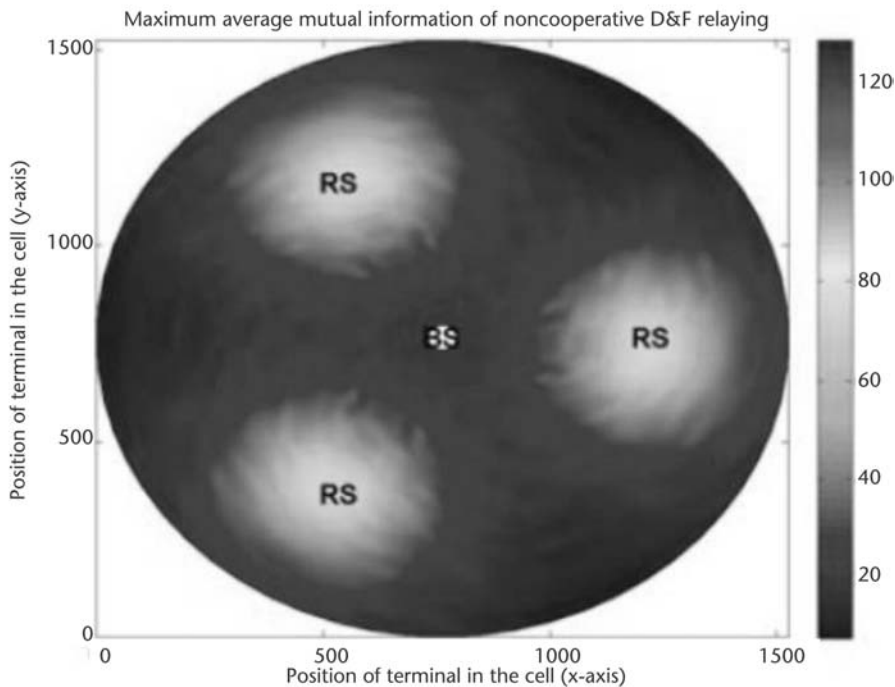


Figure 6.38 WMAN maximum DL capacity versus SS position, two-hop DF [7].

“hot spots” is very limited, and in the assumed simulation scenario, probably up to 10 relays per cell would be required to offer more than 50 Mbps of capacity at every location.

The mutual information of cooperative DF is always greater than that of noncooperative DF and that of direct link. Figure 6.39 shows the capacity of Protocols I, II, and III as a function of the average SNR on the RS-SS link, for a large number of fast fading realizations.

To match the conditions of the scenario, $\text{SNR} = 30$ dB was assumed. The $\text{SNR}_{\text{BS-SS}} = 5$ dB, which means the SS is close to the cell border. Note that Protocol II yields an improvement only when the SS is also far away from the RS and the two SNRs are equivalent. If the SS gets close to RS, then $I_{\text{SR-D}} \approx I_{\text{R-D}}$ and the performance of Protocol II reduces to that of noncooperative two-hop transmission. Protocol III outperforms both Protocols I and II at every SNR.

Based on the results in Figure 6.39, Figure 6.40 shows the relative mutual information for Protocol III with respect to the maximum of direct link and noncooperative DF.

Note that there is no significant gain in areas close to the BS where direct link is the best choice. With Protocol III, an increase of up to 25% of the average capacity can be observed except in areas around the BS. This shows how cooperative relaying can alleviate the problem of nonhomogeneous capacity. Figure 6.41 shows the performance of the various protocols.

In the simulated city-center scenario, the deployment of fixed cooperative relays can increase the DL throughput by 20% on the average in areas that are far enough

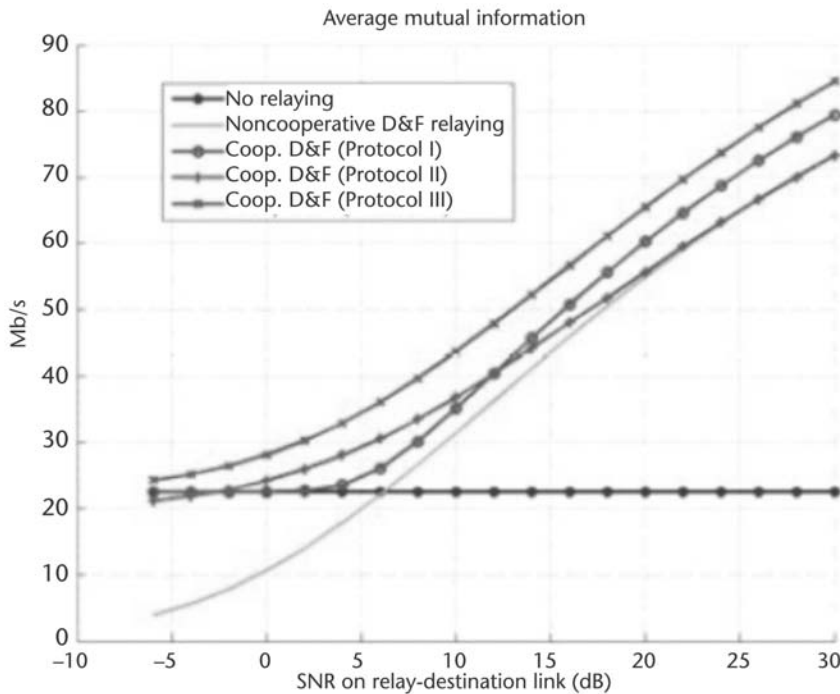


Figure 6.39 Achievable rate with the various cooperative DF protocols versus $\text{SNR}_{\text{RS-SS}}$, for $\text{SNR}_{\text{BS-SS}} = 5$ dB and $\text{SNR}_{\text{BS-RS}} = 30$ dB [7].

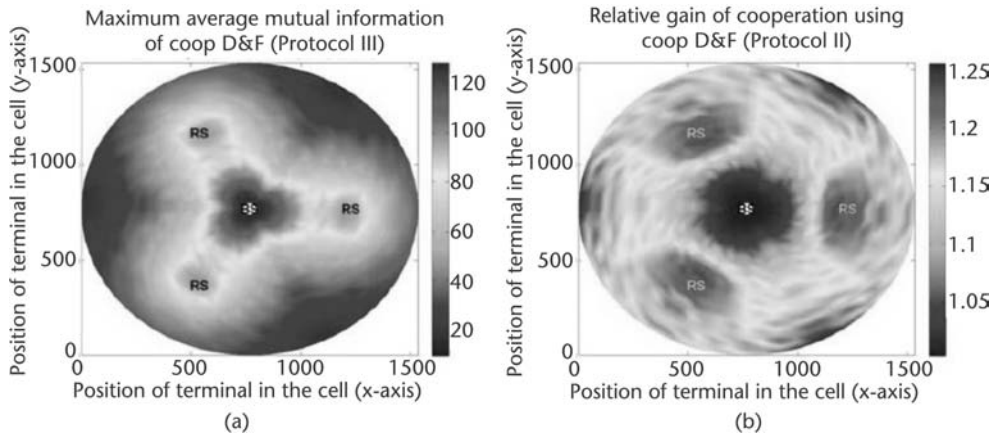


Figure 6.40 WMAN maximum DL capacity improvement by cooperative DF relaying, Protocol III. (a) Absolute value and (b) relative gain versus best of noncooperative techniques [7].

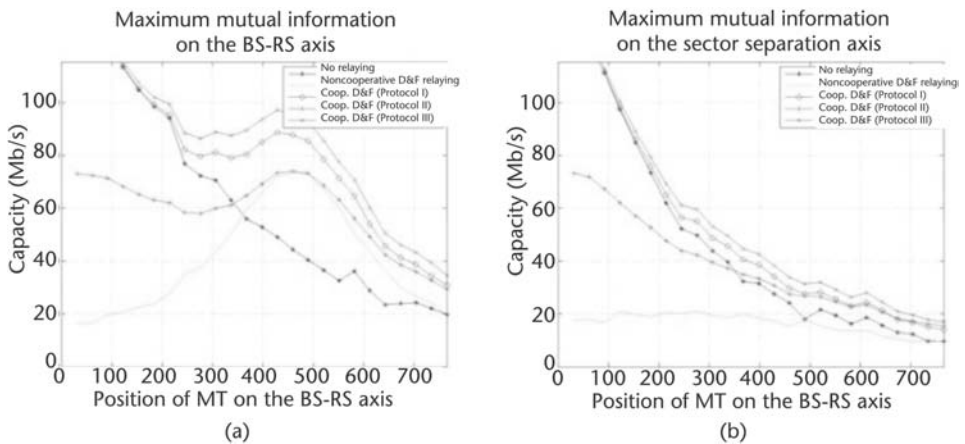


Figure 6.41 Maximum DL capacity of direct link, noncooperative and cooperative DF. (a) Along the BS-RS axis and (b) along the sector separation axis [7].

from the BS, resulting in a more homogeneous coverage of the cell. However, the relative gains remain limited, even when the most complex and best performing cooperative relaying protocol is used.

To further improve the coverage, CSI should be exploited at the transmitting side in the BS and RS, for improving the beamforming gain at low SNR, which was not considered in the simulations. Another potential means to significantly increase the cell throughput would also be the use of a dedicated large bandwidth channel between the BS and RS, as long as they are in LOS with each other and therefore benefit from a high SNR. This idea is shown in Figure 6.42, where an 80-MHz bandwidth is assumed on the BS-RS link.

The number of antennas is $N_{BS} = N_{RS} = 2$ and $N_{SS} = 4$ and $10 \text{ SNR}_{BS-SS} = 10 \text{ dB}$. In this case, there is enough capacity on the BS-RS link to create a 4×4 virtual MIMO channel and benefit from a large capacity increase.

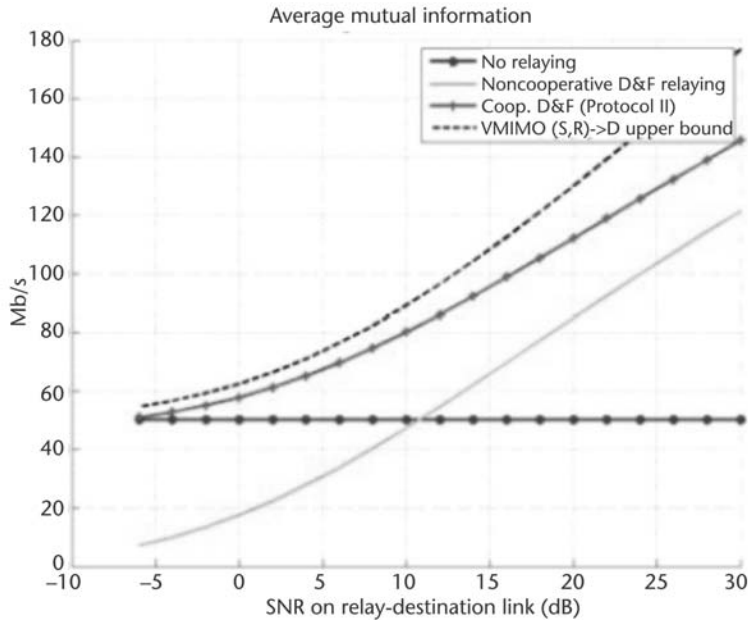


Figure 6.42 Impact of a dedicated high-bandwidth BS-RS link on the cooperative DF gain (Protocol II) [7].

6.3.2 Spatial Filtering and SDMA

The 802.16 standard [6] includes means to integrate adaptive antenna techniques. Comparable approaches are also considered in 3GPP [24, 25] and IEEE 802.11n [26]. These advanced antenna techniques have a significant impact on the capacity and service quality provided by wireless links and the efficient use of the available spectrum [27]. The simultaneous transmission can be accomplished by predistortion or beamforming techniques. The concurrent reception is known as *joint detection*. In general, the concurrent transmission/reception of data to/from different spatially separated channels is called space-division multiple access (SDMA). It provides another degree of freedom to conventional TDMA-, FDMA-, or CDMA-based medium access.

Systems that operate in SDMA must fulfill several requirements (see Chapters 3 and 4). First, the PHY layer must be able to dynamically adapt its receiving and transmitting characteristics. Therefore, an antenna array and a proper algorithm to calculate the antenna weights are necessary. Second, the enhanced PHY layer has to offer its SDMA services to the MAC layer so that the MAC can leverage the new features. This is done by extending the services offered at the PHY service access point (SAP). Third, the MAC protocol has to cope with the spatial domain. Different peer MAC entities communicate by means of standardized MAC protocol data units (PDUs). Beside the time and frequency domains, these PDUs have to be able to coordinate and utilize the spatial domain. Lastly, MAC behavior, which is not standardized, is affected by the spatial domain as well. Vendor-specific algorithms, such as scheduling, have to efficiently handle the new spatial domain.

In a multicell environment the challenge is to allocate subchannel resources in an efficient manner so that intercell interference (ICI) effects are minimized.

Methods that require smart antennas at the BS and only a single antenna element at the SS were investigated by the FP6 IST project FIREWORKS [3]. Such an approach makes the SS devices cheaper and simpler.

Two methods are used to increase system capacity by use of beamforming: spatial filtering for interference reduction (SFIR) and SDMA. These are shown in Figure 6.43.

Employing SFIR with smart antennas reduces the interference produced during transmission and experienced during reception. For this reason, cells with the same frequency can be placed closer together, thus reducing the cluster size or the reuse distance. If the cell capacity is not affected, this leads to an improvement of the total spectral efficiency. Thus, less spectrum is needed to operate a network or more bandwidth is available in each cell.

In SDMA operation the number of users within a given cell is increased. Multiple users can be served on the same time/frequency slot because the BS distinguishes them by means of their different spatial signatures. These basic techniques using smart antennas are employed in the following simulations and compared to a scenario consisting of BS with omnidirectional antennas.

The evaluated scenario consists of seven cells, each with a central base station and 25 subscriber stations. Measurements are only performed in the central cell. The stations in the surrounding six cells only produce interference for the central

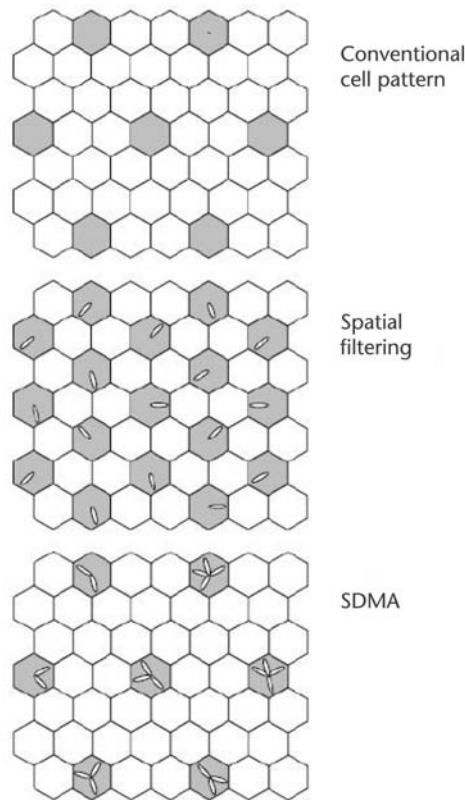


Figure 6.43 Principles of SFIR and SDMA [7].

cell and are not evaluated. Nevertheless, the same event-driven stochastic simulation, with identical average traffic loads and with the same degree of detail, is conducted at all 182 stations. The cells have a radius of $R = 1,750\text{m}$ and an $N = 7$ cell cluster order. Each BS is equipped with a nine-element uniform circular antenna array used to serve the whole cell without sectorization. The SSs are equipped with standard omnidirectional antennas. For both station types, the transmit power is 1W and no further power control is performed. A bandwidth of 20 MHz with a midfrequency of 5.470 GHz is used. All stations are assumed to be fixed. Because roof-top deployment for the SS antennas is envisioned, the path-loss model presumes LOS conditions. The “C1 LOS” path-loss model for a suburban environment is derived by the WINNER project in [9].

The simulator implements the detailed 802.16 MAC protocol for both the BSs and SSs. A stochastic traffic model generates a well-defined traffic load with a fixed packet size of $1,024\text{ bits}$ and a negative-exponential interarrival time. The interarrival time is adapted to generate various amounts of offered traffic. A radio interference simulation engine (RISE) models the PHY layer as well as the wireless channel. It is based on look-up tables that are generated by a sophisticated link layer simulation chain, which was developed by the FP5 IST project STRIKE [28]. A detailed description of the simulator and the presented performance measures can be found in [7].

Figure 6.44 shows the DL cell throughput versus the total offered DL traffic. The diagonal line marks the total offered traffic.

The variation parameter inside the graph is the transmission mode. In the (nonbeamforming) reference scenario, all transmissions are performed using omnidirectional antennas. In the beamforming scenario with a maximum number of one concurrent data stream, the BS uses smart antennas only to reduce the interference. It

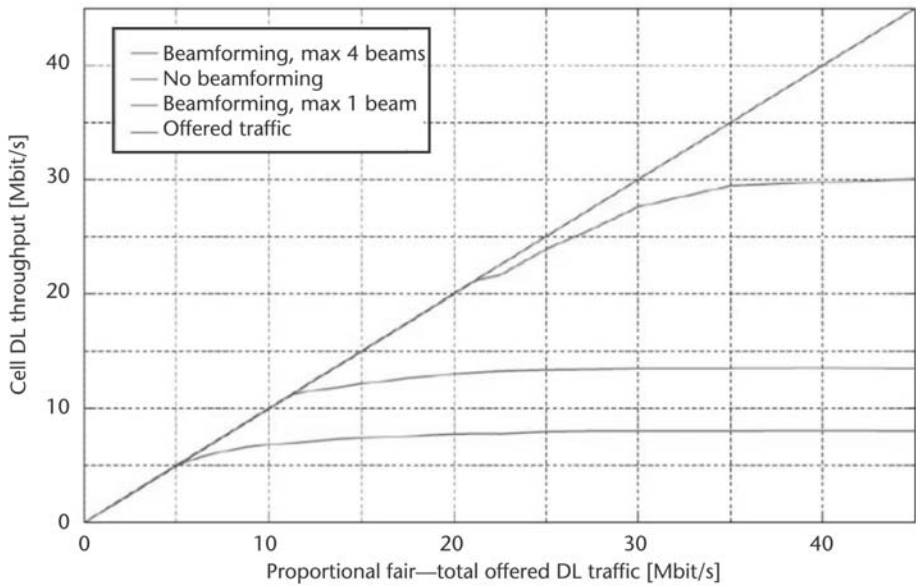


Figure 6.44 DL cell throughput for no beamforming, spatial filtering, and SDMA [7].

calculates optimized antenna patterns for the transmission and reception of signals, but one station is served at a time. This SFIR technique is often proposed to reduce cluster sizes in cellular systems. In the beamforming scenario with a maximum number of four data streams, the smart antenna at the BS is used to serve up to four stations in parallel. The concurrent data streams are spatially separated by optimized antenna patterns (SDMA).

Figure 6.44 shows that the nonbeamforming scenario can carry up to 5 Mbps offered DL traffic. Because a symmetric partition into DL and UL subframes is assumed, this results in a total cell throughput (DL and UL traffic) of approximately 10 Mbps. If the traffic load increases further, the system runs into an overload situation in which packets have to be buffered and finally discarded. Using SFIR, the ICI is reduced and higher order PHY modes are applicable. The increased capacity is reflected by a saturation level of approximately 12 Mbps.

By using SDMA, the BS can serve up to four stations concurrently. Although SDMA introduces intracell interference and increases the ICI, the concurrent data streams further increase the cell capacity up to a saturation level of 22 Mbps. Taking the saturation levels as a reference, SFIR transmission achieves a gain of 240% compared to the nonbeamforming case. The SDMA transmission with up to four beams achieves more than 80% gain compared to the SFIR transmission. When compared with the nonbeamforming transmission, it reaches an even higher gain of 440%.

Thus, in a conventional single-hop SDMA-capable 802.16 system, when the BSs operate in SDMA, capacity can be increased. SDMA operation is also beneficial in RECs. In a multihop scenario, the SDMA operation on the first hop hardly differs from the single-hop scenario. BSs can schedule and serve SSs simultaneously. If fixed relay stations are assumed, the first-hop SDMA operation is even less complex. Fixed relays are positioned apart from each other. If the BS leverages the knowledge about their position, it can group the relays into one spatial group. Without mutual interference between relays, their capacity on the first hop can be increased. Because these links have to carry a lot of data, this enhancement is one key aspect in the design of a high-capacity multihop system [7].

If a relay station has smart antennas that allow for beamforming in the downlink or joint detection in the uplink, SDMA operation is also possible in the subcell of the relay. If the relay has the control about its subcell, it can simply schedule and serve the SSs simultaneously. The relays will have the same knowledge about the channel as the BS. Thus, the SDMA operation of the relay on the second hop does not differ from the operation in SDMA of the BS on the first hop.

If the subcell is under the control of the BS, it schedules the SSs even if these are on the second or third hops. To build up a proper SDMA schedule, the BS needs to get all required information about the spatial separability of the SSs on the remote hops. This information is collected by the relay and afterwards signaled to the BS. Here signaling overhead must be accounted for.

In highly shadowed scenarios, such as the Manhattan scenario, relays do not even have LOS connections to each other. This gives the opportunity to let the relays operate simultaneously.

This simultaneous operation of subcells promises a significant increase in capacity. Only a fraction of the 802.16 MAC frame has to be dedicated to the second

hop operation. Care, however, has to be taken that the mutual interference between simultaneously operating relays stays within reasonable limits.

6.4 Conclusions

The main motivation for deploying relays is to decrease overall network costs while maintaining a required service level. Use of the indifference curves is a tool to compare the cost efficiency of different deployment and technology options.

A proper RRM scheme is essential to achieve the full benefits of a relay-based deployment. This chapter demonstrated the benefits of use of spectrum and soft resource partitioning in a REC scenario. Flexible resource partitioning outperforms static resource partitioning.

Throughput enhancement via multihop deployments (in which relays are placed inside the BS coverage area) is possible in Manhattan-like scenarios. In wide-area scenarios the mean cell capacity decreases compared to the single-hop case. Smart antennas could increase the capacity of multihop deployments. However, single-hop deployments could also use smart antennas such that the general advantages and disadvantages in the throughput scenario might remain unchanged.

By placing relays at the original cell edge, the coverage of the BS can be extended. Using three relays, the coverage area is extended by a factor of 3. Additionally, the achievable link distances increase due to the increased cochannel distance. Thus, in rural scenarios the gain in coverage is larger.

The overhead due to data forwarding and due to additional signaling decreases the capacity of a relay-enhanced cell significantly. To regain the lost capacity, relays can operate simultaneously. This mode of operation exploits the characteristics of multihop deployments. The geometric distance of relays and the shadowing due to obstacles limit the mutual interference between simultaneously operating relay stations. This allows the cell capacity to be regained.

Cooperative communication leverages relays to retransmit the original signal to the destination. The receiver can combine both the original signal and the retransmitted signal to increase capacity and/or reliability. This technique has a huge potential to increase cell capacity. Different cooperative forwarding techniques can be employed.

Spatiotemporal processing techniques can meet the requirements for an IMT-A system together with relaying and enable cost-efficient deployments. The MIMO cooperative relaying technique can increase the average user throughput by 50% compared to single-path MIMO relaying.

Multihopping (i.e., the use of relays between the end user and the end node) is primarily motivated by the low power and the low heights of the BS and RNs. In low-power transmissions, multihopping helps increase the range. Moreover, because low-height BSs are likely to be surrounded by several obstacles, multihopping helps avoid the problem via multiple links that are more likely to have LOS between them.

The challenge in multihop networks is determining how to organize the network so as to enable efficient and reliable communication and cooperation among the different nodes. A natural approach to this problem is to create and maintain

distributed structures that store relevant control information about the network. Such structures, which include spanning trees, dominating sets, independent sets, clusterings, and other network decompositions, play an important role in a variety of schemes for frequency and bandwidth allocation, routing, location management, and resource discovery.

References

- [1] FP6 Overview, <http://cordis.europa.eu/ist/so/mobile-wireless/home.html>.
- [2] FP6 IST Projects Wireless Interface New Radio, WINNER and WINNER II, www.ist-winner.org.
- [3] FP6 IST project FIREWORKS, <http://fireworks.intranet.gr>.
- [4] FP6 Project MEMBRANE, www.imperial.ac.uk/membrane.
- [5] FP6 IST Project WINNER, "Description of Identified New Relay-Based Radio Network Deployment Concepts," Deliverable 3.1, October 2004, www.ist-winner.org.
- [6] IEEE 802.16 Mesh Networking Standard, www.ieee802.org/16.
- [7] FP6 IST Project FIREWORKS, "Cellular Deployment Concepts for Relay-Based Systems," Deliverable 2.1, January 2007, <http://fireworks.intranet.gr>.
- [8] IEEE 802.16e, www.ieee802.org/16/tge.
- [9] FP6 IST Project WINNER, "Definition and Assessment of Relay-Based Cellular Deployment Concepts for Future Radio Scenarios Considering 1st Protocol Characteristics," Deliverable 3.2, June 2005, www.ist-winner.org.
- [10] IEEE 802.16j, www.ieee802.org/16/pubs/80216j.html.
- [11] IEEE 802.11s, www.ieee802.org/11/Reports/tgs_update.htm.
- [12] FP6 IST project WINNER II, "Assessment of Relay-Based Deployment Concepts and Detailed Description of Multi-Hop Capable RAN Protocols," Deliverable 3.5.2, June 2007, www.ist-winner.org.
- [13] FP6 IST Project WINNER II, "Final Assessment of Relaying Concepts for All CGs Scenarios Under Consideration of Related WINNER L1 and L2 Protocol Functions," Deliverable 3.5.1, September 2007. www.ist-winner.org.
- [14] Johansson, K., et al., "Relation Between Base Station Characteristics and Cost Structure in a Cellular System," *15th Intl. Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC)*, September 2004.
- [15] FP6 IST Project WINNER II, "Cost Models for Different Deployment Scenarios," Deliverable D6.13.6, November 2006, www.ist-winner.org.
- [16] "Framework and Overall Objectives of the Future Development of IMT 2000 and Systems Beyond IMT 2000," Recommendation ITU-R M.1645, www.itu.int.
- [17] Liang, Y., and V. V. Veeravalli, "Gaussian Orthogonal Relay Channels: Optimal Resource Allocation and Capacity," *IEEE Trans. on Information Theory*, Vol. 51, No. 9, September 2005, pp. 3284–3289.
- [18] FP6 IST Project WINNER II, "Metropolitan Deployment Scenario," Deliverable 6.13.8, November 2007, www.ist-winner.org.
- [19] FP6 IST Project WINNER II, "Wide Area Deployment Scenario," Deliverable 6.13.1, December 2006, www.ist-winner.org.
- [20] FP6 IST Project WINNER, "Definition and Assessment of Relay-Based Cellular Deployment Concepts for Future Radio Scenarios Considering 1st Protocol Characteristics," Deliverable 3.4, June 2005, www.ist-winner.org.
- [21] Cover, T., and A. A. El Gamal, "Capacity Theorems for the Relay Channel," *IEEE Trans. on Information Theory*, Vol. 25, No. 5, September 1979.

- [22] Host-Madsen, A., "On the Capacity of Cooperative Diversity in Slow Fading Channels," *Proc. of the Allerton Conference on Communications, Control and Computing*, Monticello, IL, October 2002.
- [23] FP5 IST Project ROMANTIK, "Resource Management and Advanced Transceiver Algorithms for Multihop Networks," Deliverable 4.4.1, August 2004, www.ist-world.org.
- [24] Long-Term Evolution, www.3gpp.org/Highlights/LTE/LTE.htm.
- [25] Third Generation Partnership Project, www.3gpp.org.
- [26] IEEE 802.11n, http://grouper.ieee.org/groups/802/11/Reports/tgn_update.htm.
- [27] Ghosh, W., and A. Chen, "Broadband Wireless Access with WiMax/802.16: Current Performance Benchmarks and Future Potential," *IEEE Communications Magazine*, February 2005, pp. 129–135.
- [28] Prasad, R., (ed.), *Towards the Wireless Information Society, Vols. I and II*, Norwood, MA: Artech House, 2005.

Interference Management

Interference management is based on the concept that use of the radio spectrum is controlled through managing the radio interference instead of directly controlling its causes such as transmitters and antennas. Managing the radio spectrum and managing interference have always been linked. The use of the radio spectrum has the potential to cause interference and interference can impede the operation of radio equipment. Therefore, radio transmissions are the cause and interference is the effect. However, most spectrum management techniques aim to control the causes of radio interference rather than the effects.

Technology plays an important part in attempting to deal with the radio environment from the interference perspective; systems have been developed that mitigate interference or allow operation of radio systems in a situation where potentially harmful interference exists. Restrictions on transmitter powers, or antenna directionality, and many other techniques are typically employed in order to try to ensure that radio transmitters do not cause undue interference.

Interference management was studied extensively in the FP6 IST projects WINNER and WINNER II [1–3] in support of the ubiquitous radio system concept designed for next generation systems [4]. Interference issues were also studied in the scope of the FP6 IST project FIREWORKS [5] in support of the relay enhanced OFDMA concept. Interference cancellation techniques were studied in support of the open broadband access network proposed by the FP6 IST project OBAN [6]. An interference analysis of MIMO relay networks was performed within the scope of the FP6 IST project MEMBRANE [7]. Interference issues were also studied as part of the reconfigurable radio interface within the FP6 IST project SURFACE [8].

Interference was also studied in the context of a PAN scenario within the FP6 IST projects MAGNET and MAGNET II [9] and PULSERS and PULSERS II [10]. These projects, as well as projects such as the IST projects E2R [11] and RESOLUTION [12], studied hardware implementation effects on interference in reconfigurable environments.

This chapter builds on the results of the FP6 projects and describes solutions for various scenarios. Section 7.1 introduces into the topic of interference management and mitigation. Section 7.2 describes proposals for intercell interference (ICI) modeling. Section 7.3 describes interference models proposed for PAN-optimized radio systems. Section 7.4 concludes the chapter.

7.1 Introduction

Defining interference is not straightforward: A signal that causes problems for one technology or user may cause a disturbance of little consequence to another. To understand interference and make comparisons between different interference mitigation schemes and management techniques, it is first necessary to define what is meant by interference. There are several definitions of interference, different both on axes of severity (i.e., to what extent the interference causes a problem) and of definition (i.e., the measure of performance that is affected) [12].

Interference can be defined as an undesirable signal superimposed at the receiver antenna element, coming from the same or a different source than the desired signal. This definition implies that interference is viewed from the physical point of radio-wave propagation (e.g., the antenna element). Simulation studies in this context should include the geometry of the antenna element and its angle-dependent ability, if any, to capture the energy of an approaching radio wave [9]. For the case of an antenna array, this angle dependency can be achieved by proper weighting of the individual receiving signals and combining. Because of this combining, the receiver may be able to notch individual interferers and the interference may be canceled. But this implies a certain level of intelligence regarding interference at the receiver and, thus, implies that interference should be defined on a per-antenna-element basis in order to not depend on receiver intelligence. For broadband communication over frequency-selective channels, the power density of the interferer over the frequency axis may be either constant (i.e., the interference is said to be *white*) or different (i.e., the interference is *colored*).

In general, radio communication links may experience several kinds of interference, such as intersymbol interference (ISI), multiple access interference (MAI), and interference from external sources. ISI may be caused by a time-dispersive radio channel, whereas MAI appears when the limited set of radio resources is reused throughout the system [9]. To reach high spectrum efficiency, it is important that the radio interface, in particular, be able to handle MAI. The MAI can originate either from its own cell, so-called *intracell interference*, or from other cells, *intercell interference*. In a system such as the WINNER radio concept, for which the multiple access schemes were designed to be orthogonal in the cell both for the uplink and downlink, the user would not experience any intracell interference. Therefore, only ICI was investigated in this project.

In a radio system where relaying is considered an integral part of the air interface technology for coverage and capacity improvements, ICI must also be accounted for. This was discussed briefly in Chapter 6. ICI within a REC is controlled by the BS through enforcing a MAC mechanism [10]. When a centralized routing strategy is employed with the relay technology, from the interference perspective, one major advantage is that *cross-optimization* can be performed in order to avoid extra interference to highly loaded cells and, hence, improve system performance. Having relays as a part of the radio interface opens up a set of opportunities to develop more effective interference management mechanisms in both intracell and intercell scenarios [13–16].

For assessing ICI issues, system level simulations need to be conducted. Different approaches and concepts need to be compared against each other. This can be

done either by a relative comparison or comparison toward a reference system design. The latter has the implication that tight alignment on the models and assumptions needs to be assured.

In a BAN only self-interference (i.e., body proximity and intra-PAN interference) is considered. Such interference requires a different approach to interference because no dedicated frequency bands are reserved for PAN systems [17]. Multi-PAN interference will be considered in a situation where several PANs operate in proximity. PANs from different users can connect and use the same (shared) devices. Consequently, common channel resources have to be shared between all PANs and control information has to be exchanged either between all or clusters of PANs. Major problems arise as a result of intra-PAN, inter-PAN, and non-PAN interference as well as sharing of devices and spectrum resources. Spectrum-efficient and robust multiple-access schemes are to be used, and it is envisioned that devices could support multihop in order to extend the network coordination influence or even data communication.

7.1.1 Methods to Assess the Severity of Interference

A number of different methods have been defined [18] to assess the severity of interference:

- Interference area (IA);
- Interference scenario occurrence probability (ISOP);
- Monte Carlo (MC);
- Worst case (WC).

In fixed wireless systems, the occurrence and level of interference depends mainly on the placement of the terminal stations (TSs) and the central stations (CSs), as well as on the antenna gains. The IA is defined as the area in which unacceptable interference occurs, relative to the area of the cell or sector.

The ISOP is defined as the probability that at least one terminal is placed in the IA. This measure is related to the number of terminals deployed in a cell, and to the cell planning methodology. In [18], this method is used to evaluate guard bands between frequency bands belonging to different operators.

The Monte Carlo method is used in [18] to evaluate the interference probabilities between terminals, because the placement of terminals is rather random. The worst-case method derives system deployment parameters to ensure that the interference is in all cases below a given threshold.

7.1.1.1 The Interference Scenario Occurrence Probability

The ISOP is the probability that a TS or CS experiences unacceptable interference. In the scenario where the CS is affected by harmful interference, an assessment based on the ISOP might not be adequate because in this case a whole cell could be blocked [19]. Therefore, for the CS a WC analysis is preferred. In the CS-to-TS interference case, an ISOP of 1% means that, on average, 1 terminal out of 100 faces interference problems.

7.1.1.2 The Interference Area

The IA is the quotient of the area in which interference exceeds a set threshold relative to the area of the whole cell or sector. This is equivalent to the probability that a single terminal placed anywhere in the cell will suffer from unacceptable interference. The IA is computed for specific antenna patterns, channel assignments, and guard bands. This measure is related to the ISOP, however, it does not depend on the number of deployed terminals.

7.1.1.3 An Example of Rudimentary Coexistence Analysis for Fixed Radio Systems

The fixed radio systems designed according to ETSI standards do not specify the equipment in such detail that terminals from different manufacturers are interoperable; that is, there is no common air interface [19]. This implies that the susceptibility for interference will vary among systems from different manufacturers. In the following subsections, an example of the standard EN 301 021, which applies TDMA as the multiple-access scheme, is given [20, 21]. The technical parameters of the other ETSI standards [22] do not differ substantially, so that the main conclusions drawn for this particular standard can be generalized to standards with other multiple-access schemes.

7.1.1.3.1 Channel Arrangements.

For the channel arrangements, that is, the center frequencies, bandwidths, and duplex arrangements, various solutions exist. Some possible channel spacings are recommended in CEPT/ERC/REC 12-08 E [23], which in turn refers to ITU-R F.635 and ITU-R F.382. Annex A of [23] recommends a channel spacing of {15, 20, 30, 40} MHz, whereas Annex B recommends a spacing of $N \cdot 0.25$ MHz, with $N \in \{1, 2, \dots, 200\}$. For the channel arrangements, several possibilities that depend on the channel spacing are also recommended.

7.1.1.3.2 Transmitting Power Ranges.

The maximum transmitting power at point C' of the block diagram in Figure 7.1 is 35 dBm [20]. No specifications about EIRP or antenna gains are given, thus any

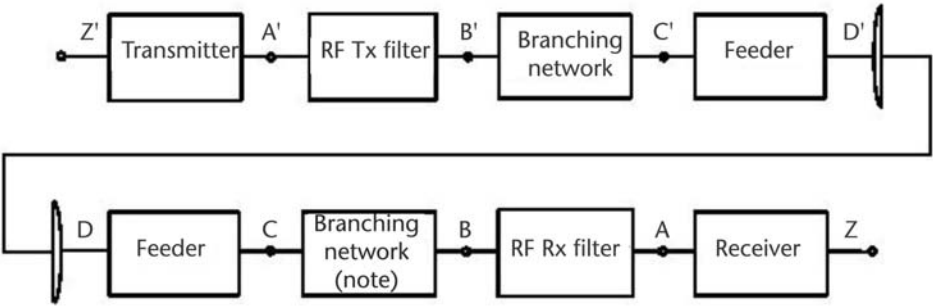


Figure 7.1 RF block diagram of EN 301 021 [20].

calculation including the radiated power is necessarily based on some assumptions. Automatic power control is optional but not mandatory.

7.1.1.3.3 Spectrum Density Masks.

The EN 301 021 standard [20] defines eight “system types,” which represent different bit rates and modulation types. The spectrum density masks (an example of which is shown in Figure 7.2) depend on this system type and on the chosen channel spacing.

7.1.1.3.4 Interference Sensitivity.

The EN 301 021 standard [20] distinguishes the following three types of interference:

- Adjacent channel interference;
- Cochannel interference;
- Continuous wave (CW) interference.

For issues of spectrum sharing, the sensitivity to cochannel interference is the most relevant one. Cochannel sensitivity is defined as the resistance of the receiver with respect to an unwanted signal of the same type at the same frequency, but with less power. The standard defines the cochannel interference sensitivity with the following procedure: Two transmitters are connected to the receiver under test, as shown in Figure 7.3.

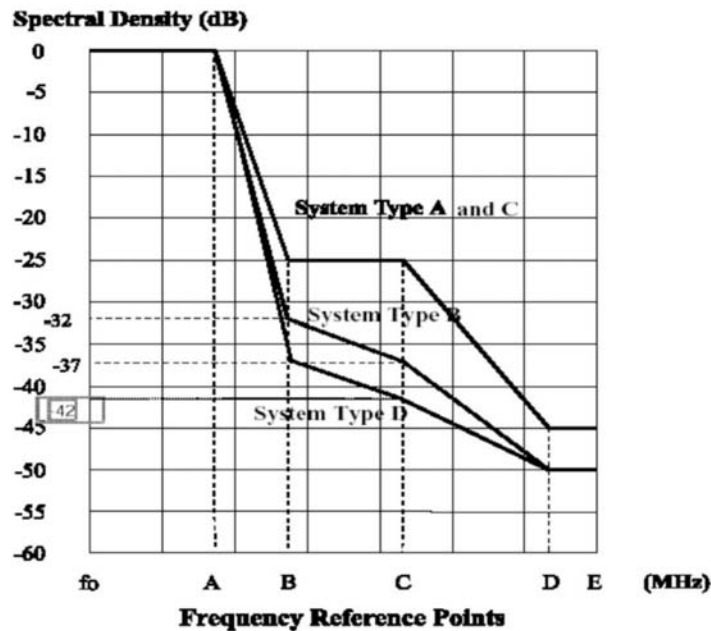


Figure 7.2 Power spectrum masks for system types A, B, C, and D. f_0 denotes the actual carrier frequency, and the frequency reference points are further specified in tables given in [20].

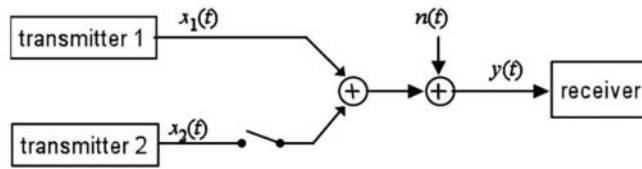


Figure 7.3 Setting for test of interference sensitivity [19].

In the first pass, the power of transmitter 1, $S_1 = E[|x_1(t)|^2]$, is adjusted such that the SNR γ_1 at the receiver corresponds to the BER $P_b(\gamma_1) = 10^{-6}$. The SNR is thus given by $\gamma_1 = S_1/N$, with $N = E[n(t)|^2]$. Then, the signal of transmitter 2 is injected with power $N_2 = \beta \cdot S_1$, where $\beta = 10^{-23 \text{ dB}/10}$ for system types A, C, and E. The SNR at the receiver is now

$$\gamma_2 = \frac{S_1}{N + N_2} \quad (7.1)$$

At this SNR, the BER shall not be greater than 10^{-5} , that is, $P_b(\gamma_2) \leq 10^{-5}$.

Figure 7.4 shows that this method for defining the sensitivity against interference is very specific to the considered system because it depends on the slope of the $P_b(\gamma)$ curve. A system that employs strong channel coding would probably not fulfill this requirement because the BER curves for coded systems are much steeper.

One way to estimate coarsely the permitted interference level is via Table 8 of EN 301 021 [20]. There, the receiver signal level for a performance at $P_b = 10^{-6}$ is defined to lie between -90 dBm and -60 dBm. The interfering signal has to be at a level of 23, 30, or 37 dB below, depending on the system type. To fulfill this requirement in all cases, the interference power at the receiver must not exceed -127 dBm.

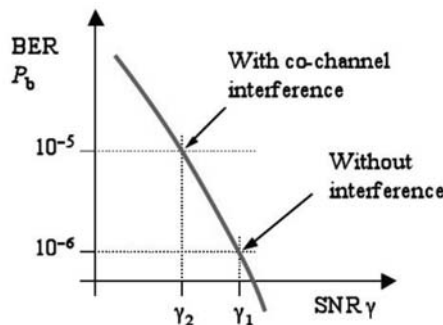


Figure 7.4 SNRs and BERs for testing the sensitivity to cochannel interference [19].

7.1.1.3.5 Rudimentary Coexistence Analysis.

The interference limit for FWA systems is estimated in [24] as $P_{r,i} = -146$ dBW/MHz. This is the maximum interference PSD that is allowed at the receiver's antenna. As an orientation level for the transmitting power, the maximum power of a UMTS data terminal can be selected. This power is limited to 0.25W and the bandwidth is assumed to be 5 MHz, which gives

$$P_t = \frac{0.25\text{W}}{5\text{ MHz}} \triangleq -13 \frac{\text{dBW}}{\text{MHz}} \quad (7.2)$$

Assuming one interfering UMTS terminal transmitting at its maximum power, the required attenuation between the UMTS terminal and the FWA receiver is then 133 dB. This allows us to estimate roughly the necessary distance d_{\min} . The antenna gain of the UMTS terminal is assumed to be 0 dBi and the boresight gain of the FWA antenna according to Table 7 of [24] is 18 dBi. Assuming that the attenuation is only due to free-space loss and spherical diffraction, the following holds:

$$A_{fal} + A_{sp} + A_{fm} = P_t - P_r + G_t + G_c = 133\text{ dB} + G(\varphi) \quad (7.3)$$

where $G(\varphi)$ is the FWA antenna gain as a function of the azimuth angle, with $G(0) = 18$ dBi. For this gain, the RPE defined in ETSI EN 302 085, range 1, class TS2, is assumed. A further assumption is $A_{fm} = 10$ dB; A_{fsl} is the free-space loss; A_{bm} is the basic median path loss, for which [20] provides extensive experimental data for frequencies up to 3 GHz, as a function of the antenna heights; and G_c and G_t are the central and terminal station height factors, which are also based on experimental data. Figure 7.5 shows the necessary distance required from an FWA terminal to avoid unacceptable interference. The attenuation consists of free-space loss and spherical diffraction.

Instead of considering only free-space loss and spherical diffraction, a possibly better estimate for the path loss is given by the extrapolated Okumura-Hata model:

$$A_{50} + A_{fm} = 133\text{ dB} + G(\varphi) \quad (7.4)$$

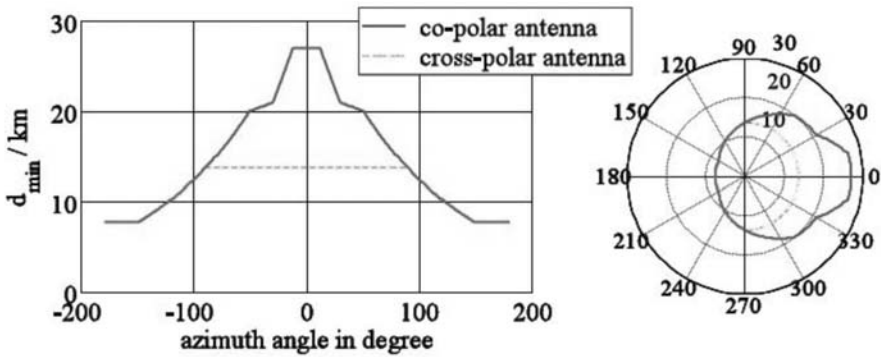


Figure 7.5 Necessary distance required from FWA terminal to avoid unacceptable interference [19].

The attenuation is calculated with the extrapolated Okumura model. Then the minimum distance can be obtained, which is shown for this case in Figure 7.6.

The results in Figures 7.5 and 7.6 differ considerably, which demonstrates the strong dependence on the chosen propagation model. Additionally, note that all channel models are a more or less coarse approximation of the real-world phenomena in radio propagation and can never capture the actual propagation conditions in a particular case.

7.1.2 Interference Management

Existing interference management methods are based on static transmitting power masks and allocation of a frequency band to a service with a predefined RAT or RAT group. This approach successfully prevents interference; however, in many situations it does not lead to the most efficient use of the spectrum, because it does not take into account the temporal and spatial variations of the actual radio environment. A fundamental observation about the interference potential of any radio emission is that its impact on any radio service depends on the interference level at the *victim receiver's antenna*. The impact also depends on the capabilities of the receiver to “suppress” or cope with interference. However, the ability to suppress interference is limited by fundamental limits [19], which cannot be overcome by any measures taken at the receiver. Completely unpredictable interference (i.e., interference resembling white Gaussian noise) cannot be suppressed at all. It has the same effect as thermal noise. It lowers the SNR at the receiver and, thus, the limits shown in Figure 7.7 apply. Reference [19] reports that this also holds for “smart” or “cognitive” radios.

The effect of interference on a point-to-point radio link can be explained directly with the curves in Figure 7.7.

Interference that stems from an unknown source or from several unknown sources, appears to the receiver like *noise*. Because interference is normally neither Gaussian nor white, its effect is possibly more severe than that of AWGN. For a coarse estimation, interference might be regarded as AWGN, which adds to the noise power N_0 . The SNR in Figure 7.7 is thus reduced by interference, decreasing

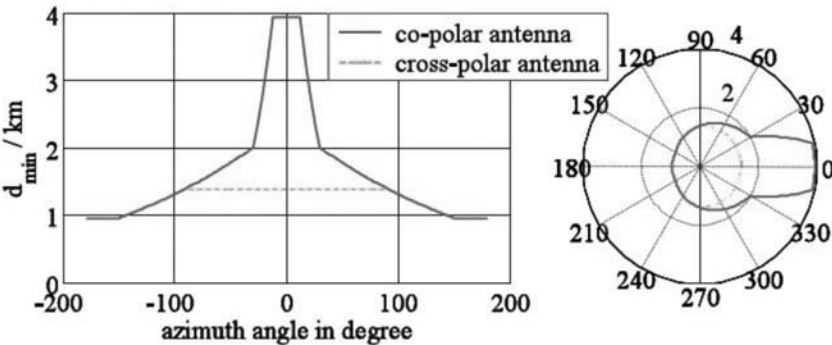


Figure 7.6 Minimum distance required from FWA terminal to avoid unacceptable interference [19].

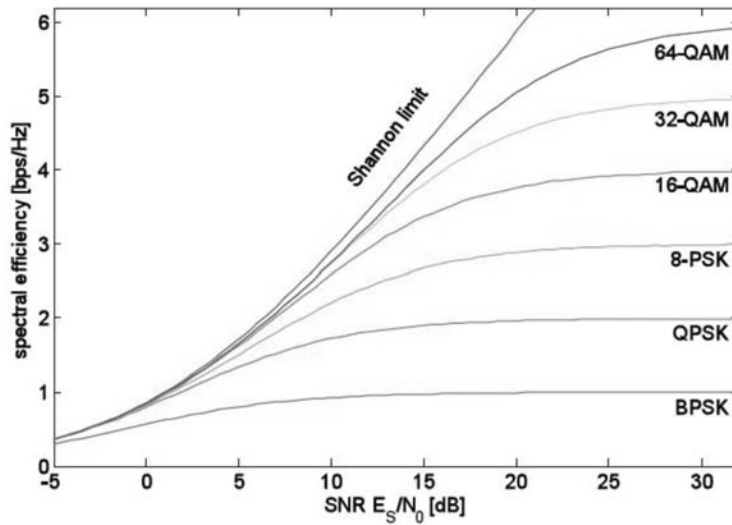


Figure 7.7 Channel capacity for Rayleigh-fading channel for continuous input (Shannon limit) and QAM input signals [19].

the possible spectral efficiency according to the capacity curves. For a nonadaptive transmission system, which operates at a fixed link spectral efficiency, this simply means that the transmission stops at a certain interference level. Adaptive coding and modulation schemes can react to increased interference by lowering their data rate, that is, by selecting an operating point with less spectral efficiency.

The only way to make a system immune against interference is to provide sufficient SNR margins (i.e., to increase the transmitting power). However, this leads to higher power consumption.

Provided that the influence interference can be described in a quantitative way and can be measured, a more flexible concept of interference measurement could be based on real-time measurements taking into account the actual RF environment. This concept would allow a secondary user to assess—before transmitting—whether his signal would cause “harmful” interference to the primary system. The idea for this concept is motivated by the desire to allow spectrum to be utilized by secondary users while guaranteeing that no harmful interference is generated that affects the licensee. One of the main ideas expressed in an FCC report [25, 26] was that the influence of interference was determined at the receiver and not at the transmitter. Thus, interference management that is based on signal and interference levels at the receiver input can be more exact and realistic. On one hand, it would present opportunities for secondary users, who would be allowed to transmit as long as they do not cause “harmful” interference. On the other hand, it can give more certainty to primary users that the preestablished limits will be met than in the case of an unlicensed secondary user that is just based on regulation of the rather static transmitting parameters of the unlicensed devices.

The question is how and where the interference power can be measured. The only relevant location, as the FCC proposal suggested, is the antenna outputs of all receivers belonging to the primary system. It is not clear how a secondary

system can obtain this information [19]. Another question is the frequency of measurements. There might be *temporal holes* in a certain frequency band, but it is difficult to predict when the primary system is going to access these frequencies.

A secondary system that is to be deployed in the same frequency range as an established system can only take measurements in its own devices using its own antennas, whereas interference happens at the primary system’s receivers. This discrepancy causes a plethora of problems, some of which have been described in the comments made in reply to the FCC report [27–29]. Among these are issues such as the hidden terminal problem, the estimation of path loss, and equal right access.

7.1.2.1 Interference Modeling

The general situation of interference is investigated using an example scenario in which a mobile station (where V is the victim receiver) with links to its base station is interfered with by MSs from a system in an adjacent frequency band [19]. This scenario is shown in Figure 7.8.

The unwanted emissions of the interfering stations (I_i) and their impacts at the victim receiver (V) need first to be assessed. Depending on the distances between the interferers and the victim, appropriate propagation models have to be chosen to assess the attenuation on the respective paths. The interfering station with the minimum distance to the victim receiver is the *dominant interferer*. As under special circumstances, interfering stations with greater distances will contribute noticeable interference; the number of interferers taken into account can be varied.

The carrier-to-interference ratio (C/I) measured at the victim receiver depends on both the user path length and the interference path lengths. To estimate this, the distance between the transmitter and the receiver on the user link and as well as between the interfering stations and the victim receiver and density of active

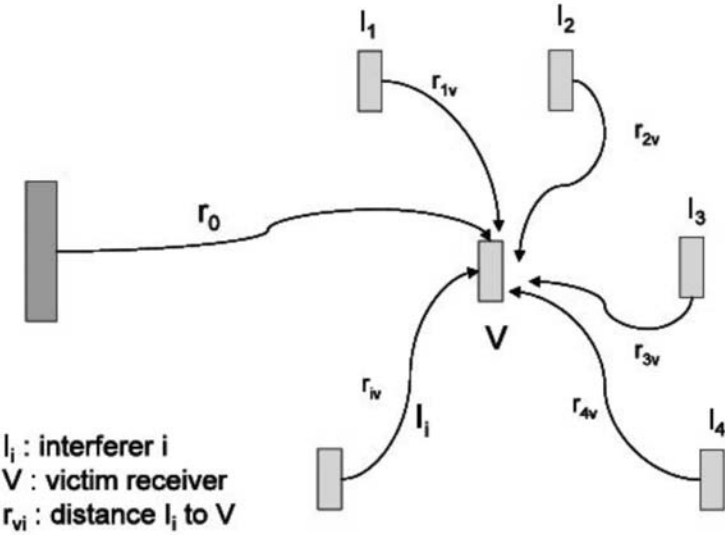


Figure 7.8 Example of an interference scenario [19].

users are important parameters. The user path length can be determined assuming that there is exactly one receiver within the coverage radius of the victim link BS. All mobiles are uniformly distributed within the considered scenario with the victim link base station in the center. More than one interfering mobile station can be taken into account depending on the active user's density. In the case when the interferer is a BS, it is assumed by approximation that the BSs of a single system operator are equally operated and it can be assumed that only one interfering BS is located within the interference scenario.

To assess the influence of the interference power and the desired power at the receiver and thus C/I , the propagation losses on the interference path and the user path have to be calculated. Methods have been developed to determine the properties of radio channels, which take the main physical effects into account in the form of models. They simulate various characteristics of a channel such as the propagation coefficients and fading behavior. In particular, fading needs to be considered in mobile communications scenarios. In view of the permissible level of unwanted emissions, it can be sufficient to take the fading due to shadowing into account and to calculate the median path fading value. This slowly varying signal length can be described by a log-normal distribution. Respective values for variances of the distribution are known from measurements and are dependent on topography and morphology.

For the efficient use of radio spectrum, it is essential to know the *power density spectrum* emitted by interfering stations. The impact of the emissions at the receiver depends on the interfering frequency, receiving band, interference power, and receiver characteristics. Apart from that, the duty-cycle pattern of transmitter activity as well as the amplitude probability distribution during active transmitting periods, which is influenced by the modulation method, need to be considered. Therefore, the transmitter and receiver characteristics of the different systems have to be taken into account. For this purpose the permissible interference power and acceptable received signal levels of each system as defined in standards needs to be considered when specifying spectrum masks.

The mask for the interfering transmitter represents the maximum permissible emission levels as a function of the frequency. To define a mask for the emissions, the different sources of out-of-band emissions are combined in one mask. The most important sources of out-of-band energy are as follows;

- Effects of the modulation process;
- Rise and fall times of the transmitted signals (switching transients);
- Intermodulation products;
- Wideband noise.

To assess the receiver characteristics, appropriate masks need to be developed that model the effects of adjacent and cochannel interferences while taking into account intermodulations and blocking characteristics [19].

7.1.3 Interference Mitigation

Different methods are available to mitigate ICI, such as interference avoidance, interference averaging, and smart antennas.

7.1.3.1 Interference Avoidance

In general, ICI avoidance techniques are based on restrictions or specific allocation schemes applied to the downlink resources available for transmission toward the UTs. They usually apply in a coordinated way between cells so that ICI avoidance is often mentioned together with ICI coordination.

Reference [30] used these two classifications for the avoidance techniques investigated for the WINNER system concept [2]: (1) resource management and partitioning and (2) resource allocation and scheduling. Resource management and partitioning are techniques based on restrictions on the transmitting power over the resources where predefined frequency reuse schemes apply as well as techniques for flexible or even self-adaptive reuse partitioning where the frequency reuse itself varies inside the cells.

Resource allocation and scheduling techniques consider intrinsically a frequency reuse value of 1; *dynamic chunk allocation* using cell coordination is one of them, *decentralized dynamic channel allocation* for the TDD mode [14] is another one. Scheduling schemes based on a cost function or QoS provisioning and aware of ICI are also part of this group.

Furthermore, the following three types of coordination may apply:

- *Static coordination*: This means that reconfiguration of the restrictions is done on a timescale corresponding to days. Internode communication is very limited (set up of restrictions), basically with a rate on the order of days.
- *Semistatic coordination*: With this type of coordination the reconfiguration of the restrictions is done on a timescale corresponding to seconds or longer. The signaling rate is generally of the order of tens of seconds to minutes.
- *Dynamic interference coordination*: In this case the timescale usually corresponds to several tens to hundreds of milliseconds.

7.1.3.2 Interference Averaging

Reference [31] investigated how interference averaging can be used to mitigate ICI. Interference averaging techniques are based on averaging the interference over users, thereby reducing the probability that some users experience very high interference. Three basic strategies for averaging the ICI were proposed in [31]: *frequency hopping*, *interference cancellation*, and combinations of *dynamic channel allocation and scheduling techniques*.

Frequency hopping is based on the use of random patterns of frequencies by different users in different cells. It may be useful for narrowband transmissions (small packet payloads) and in scenarios where other diversity means (e.g., spatial diversity) are scarce. The benefits of frequency hopping have been studied for nonfrequency-adaptive WINNER uplinks by comparing the baseline B-IFDMA uplink multiple-access scheme with an OFDMA/TDMA alternative with and without frequency hopping, and referred to as FH-OFDMA/TDMA [31] (see also Chapter 4). The performed studies assumed ideal link adaptation.

Interference cancellation reduces interference by subtracting it from the received signal before detection. Use of random channel allocation techniques coupled with

several scheduling algorithms is a relatively simple method to average interference by random allocation of resources among the users. A processing gain has to be present in the system in order to be able to separate the interference from the signal of interest. This processing gain can be obtained through spreading (as in MC-CDMA or the SS-MC-MA scheme [32]) or by using low-rate, error-correcting codes (e.g., in OFDMA). The spatial dimension (i.e., multiple-antenna reception) also provides processing gains for signal separation. Well-known techniques are successive interference cancellation (SIC) [33] and parallel interference cancellation (PIC) [34], which can be applied iteratively, possibly taking advantage of the channel coding through a turbo-equalization process [35].

Random and minimum interference channel allocation techniques coupled with several scheduling algorithms represent another way to average the ICI. These *combinations of allocation and scheduling algorithms* may lead to a different level of interference. Besides the interference averaging ability, the implementation complexity, signalization, and measurement overhead are also important. For example, random dynamic channel allocation due to its simplicity cannot attain the performance of an optimal channel allocation algorithm; however, under certain conditions it may provide sufficient performance with minimum signalization and measurement overhead. Therefore, when assessing the most promising combination, it is imperative to take all of these aspects into account.

The performance of the methods can be evaluated by means of link level and system level simulations. Link level investigations permit ICI to be studied from the one perspective of one link, and the effectiveness of interference mitigation techniques to be assessed in terms of BER or BLER improvement at the receiver. In link level simulations, the performance of a single UT is measured, and for a given set of parameters this generates a single value for each performance measure. This is in contrast to system level simulations, in which large numbers of UTs are simulated in order to generate probability distributions for the performance measures (averaged over the UT locations). Link level studies should vary one parameter in order to investigate that parameter's effect on the simulated performance. The choice of this independent variable depends on the investigation [31].

7.1.3.3 Smart-Antenna-Based Interference Mitigation

Chapter 5 of this book explained the benefits of smart-antenna techniques in support of next generation radio communication systems. These techniques offer one way to mitigate ICI by employing smart antennas at the transmitters and/or receivers in a wireless communication system [4]. A smart antenna can be defined as an antenna arrangement consisting of multiple antenna elements that can be colocated or distributed, for which the transmission and/or reception is controlled by advanced signal processing functionality.

Smart antennas can be used in a wireless communication system in any of several ways to mitigate ICI. They typically use the directivity and/or diversity properties of the spatial processing that can be achieved with the multiple-antenna concept.

Beamforming is an efficient means to combat ICI and, in particular, to protect the user at the cell border. By transmitting in a narrow beam directed toward the

desired user instead of a sector-wide beam, it is possible to significantly reduce the interference spread to other cells in the system. Beamforming can be carried out in different ways; at the highest level we can distinguish between *adaptive* and *fixed* beamforming. In adaptive beamforming the antenna weights are set to optimize the antenna pattern. This can be done according to several different optimization criteria and based on different amounts of channel knowledge. In fixed beamforming, also called *grid of beams (GoB)*, the transmitter has a certain number of preformed beams, among which it can select the best beam for transmission.

Beamforming can also be used for reception where it is possible to direct the receiving beam in the direction of the desired transmission. When adaptive beamforming is used, it is also possible to put nulls in certain directions (e.g., where strong interferers are located).

The amount of channel knowledge required for beamforming varies depending on the technique; adaptive techniques may require rather complete CSI to be reported, whereas other techniques just require channel quality indicators (CQIs), and some techniques can use uplink information to select the most suitable beam and thereby require no information to be reported. More details are available in [36]. If the beamforming in different cells does need to be coordinated, additional inter-BS signaling might be required [4].

By transmitting the same signal from different antennas, reception at the other end of the link can be improved due to enhanced *diversity*. This allows reduction of the transmitting power and, thereby, the interference spread in the system. This is particularly useful at cell borders where the ICI level typically is high, and several BSs can cooperate (i.e., macrodiversity).

Transmit diversity as such (e.g., *open-loop transmit diversity*) has no specific requirements on measurements and signaling. More advanced schemes, such as *closed-loop transmit diversity*, may require channel knowledge to be reported. When macrodiversity is applied (e.g., in the case where two or several BSs cooperate), additional inter-BS signaling and time synchronization are required.

The use of multiple antennas at receivers facilitates establishment of spatial diversity branches, which can be used for implementation of *receive diversity* and/or interference suppression techniques in the receive processing [4].

The most well-known method for receive diversity is traditional maximum ratio combining, in which the combining weights are selected accounting for the radio channel (of the desired signal), the noise power, and the interference power at the different receiving antennas. In that context, MRC can be considered a type of receive beamforming. With interference suppression techniques, the spatial or spatiotemporal properties of the interference are also taken into account. One example is interference rejection combining (IRC) (i.e., MMSE or optimal combining [37]), which determines the combining weights based on the channel and the (spatial) noise and interference covariance matrix (i.e., not only is the interference power considered, but also the spatial coloring of the interference). This means that IRC puts nulls in the direction of the interferer(s). Receive diversity and interference suppression are applicable both at the BS and the UT [4].

Interference suppression techniques require accurate interference measurements, which places requirements on the pilot design. In addition, such techniques are advantageous with a time-synchronized network, but not a requirement; the

interference suppression techniques will still work but the gain will not be maximized.

The smart-antenna-based interference mitigation methods introduced earlier can also be combined with each other, as well as with interference averaging and interference avoidance methods. One example is the combination of transmit beamforming at the BS and the use of interference-suppressing UTs. The two methods used in this combination have been demonstrated to complement each other very well and provide almost additive gains for GSM [38]. The reason for this is that most baseband interference suppression methods (e.g., IRC) are designed to be very efficient for cases with one or few strongly dominating interfering sources, and that beamforming creates an interference environment that very often is characterized by this.

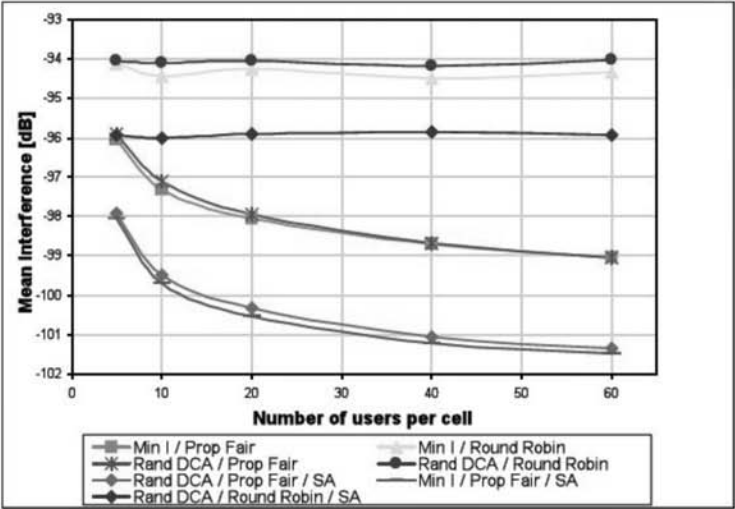
Most of the smart-antenna-based interference mitigation methods can, in principle, also be used in conjunction with interference averaging [31] and interference avoidance [30] methods to provide additional gains. For example, UT receivers with interference suppression capabilities will work in any network and provide additional robustness to interference, even though the level of interference might be lower due to an avoidance scheme [4].

The impact of transmit beamforming on interference in the downlink and in the presence of some interference averaging techniques, such as random dynamic channel allocation (DCA) [31], or interference avoidance techniques, such as minimum interference DCA [30], were investigated in [4] for the WINNER system concept. Furthermore, the interplay between beamforming and scheduling strategies and their relative impact on interference and system performance were studied. The investigations were performed with a Class II fully dynamic system level simulator [39] in a base coverage urban scenario and downlink according to the WINNER baseline design [40], and an adaptive beamforming technique was considered. The results are shown in Figure 7.9.

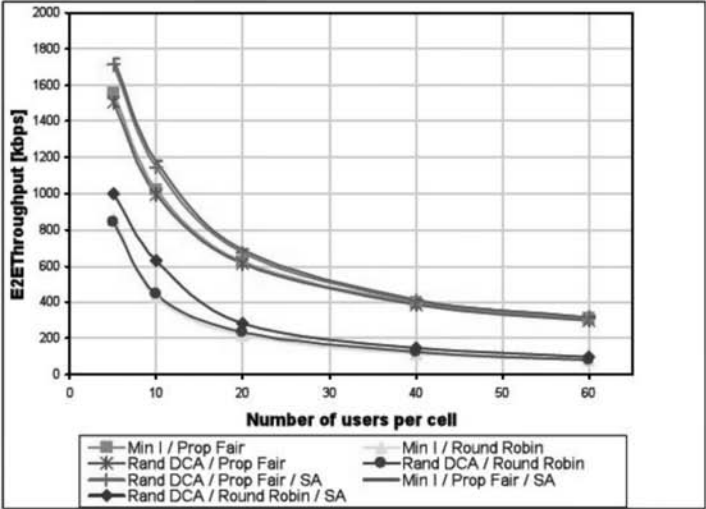
This figure shows that the highest interference reduction and throughput gain comes from *proportional fair* scheduling and then from beamforming (in Figure 7.9 denoted as smart antennas, SA), whereas the gain from minimum interference DCA is relatively low. The gains of DCA, scheduling, and beamforming are not additive regarding the interference or the user throughput. The beamforming gain in interference reduction is almost load independent, but it improves the throughput significantly, mainly for lower loads. The main gain in throughput comes from scheduling, which effectively exploits channel variations at the UTs.

7.2 ICI Modeling for an Ubiquitous Radio System Concept

When analyzing or assessing the performance of a radio network, it is important to assess the overall network performance, accounting for scenarios when several pairs of communicating nodes must share a common radio resource [30]. For example, in a cellular network, the designer must consider the fact that the resources in a cell are shared among all UTs associated with the cell and must also account for interference from neighboring cells.



(a)



(b)

Figure 7.9 (a, b) Results for different combinations of interference mitigation methods [4].

Multicell evaluations of cellular networks are often performed by computer simulations (e.g., system level simulations). System level simulations include deployment models, user behavior models, and channel models. They further comprise models of the network functionality and the radio network algorithms. The deployment models consider the BS density, the BS equipment, and the BS antenna positions (e.g., below or above roof-tops, for example). User behavior models include models that take into account user position, user mobility, and the traffic generated by the user (in the uplink and downlink). Channel models model the radio channel of desired and interfering links.

In a system level simulation, there is typically no explicit modeling of PHY layer procedures such as modulation and coding. Instead, less complex link performance models are used to estimate the performance of single links. Such a link performance model is often referred to as a link-to-system interface. The link-to-system interface needs as input a measure of the radio link quality and delivers as output an estimate of the packet error probability. Often, the SINR is used as a measure of the radio link quality, which means that system level simulations must include the calculation of the received SINR. It can be distinguished between *actual value* interfaces, where the SINR depends on the fast fading realizations of the channel [41], and *average value* interfaces, where channel quality metrics are an average of the instantaneous SINRs over the fast fading.

This section describes different methods and models that may be used in multicell system level simulations. Two different options for setting up a multicell system simulation that accounts for interference from neighboring cells are discussed based on work performed in support of the WINNER radio system design [30].

7.2.1 Multicell Simulation Setup

Real-world radio networks may comprise a very large number of cells. For example, some cellular networks have basically nationwide coverage and comprise hundreds or thousands of base stations. Cells in such radio networks often experience interference from other cells and it is essential to capture the impact of such ICI in radio network analysis and assessment work.

In practice, when analyzing radio networks by means of computer simulations, it is only possible to model a limited number of cells. That is, the considered system layout is limited and the cells in the outer parts of the modeled network are not surrounded by cells on all sides. The interference situation in these cells differs significantly from the interference experienced in the central part of the network, where cells are surrounded by interfering cells on all sides. Hence, the performance of such a network is not representative of a large real-world network in which basically all cells experience interference from all directions.

Two popular techniques exist to account for the impact of surrounding cells and ICI: the *central cell* technique and the *wraparound* technique.

7.2.1.1 Central Cell Technique

In the central cell technique, the system's statistics are collected only in the central part of the simulated multicell layout. This is typically the central cell in an omni-

directional layout, or the three sectors of the central cell in a trisected layout. Here, only the system functions of the central cell and the associated terminals are simulated in detail; the remaining cells are accounted for via simplified models.

Compared to a multicell simulation in which UTs are generated in multiple cells, for example, using the wraparound technique, the central cell technique allows considerable savings in memory resources, because only a reduced number of links needs to be monitored and managed at the same time. From the simulation duration perspective, however, the central cell technique may be equivalent to a multicell simulation using the wraparound technique since more snapshots of UT locations are needed to collect the same number of statistics on UTs within the central cell, compared to collecting data from UTs studied simultaneously within a larger number of cells.

The central cell technique is particularly suited to the downlink, because it avoids the explicit simulation of the UTs in the neighboring cells. The interference situation can be generated accurately by simulating the transmitted signal from the neighboring BSs only. In the uplink, however, where the interference is created by the signals transmitted from the UTs in neighboring cells, an accurate modeling of their positions, transmitting powers, and channels in terms of the BS of interest cannot be avoided in most cases. Hence, the savings in simulator complexity may be much lower in the uplink case, which may decrease interest in this method with respect to a full multicell simulation using the wraparound technique.

7.2.1.2 Wraparound Technique

The use of the wraparound technique [42] is one way to overcome the limitations of a limited system. With the wraparound method the cell layout is folded so that cells on the right side of the network are connected with cells on the left side and, similarly, cells in the upper part of the network get connected to cells in the lower part. The created area may be seen as borderless, but with a finite surface, and it may be visualized as a torus [42].

In practice, wraparound may be implemented by creating six copies of the studied network deployment and placing these copies around the original network [43]. Figure 7.10 shows an example of the wraparound technique in which a hexagonal cell layout comprising 19 BSs with 120-degree cell sectors is considered.

In this example, each BS is located at the center of one hexagon with one transceiver per sector. In an ICI scenario each transceiver is considered to interfere with other victims. There is a one-to-one mapping between cells/sectors of the center network and cells/sectors of each copy, so that every cell in the extended network is identified with one of the cells in the central (original) hexagonal network. Those corresponding cells thus have the same antenna configuration, traffic, fading, and so forth, as shown in Figure 7.10. Hence, every BS in the original network has one and only one corresponding BS in each of the copies, and the channel between a UT and a BS can now be determined by calculating the channel between the UT of interest and all seven BSs—the BS in the original cell layout and one BS in each of the six copies. Among the seven calculated channels, the one with the lowest attenuation is selected for further use. Typically, the selection is performed based on long-term channel information including distance

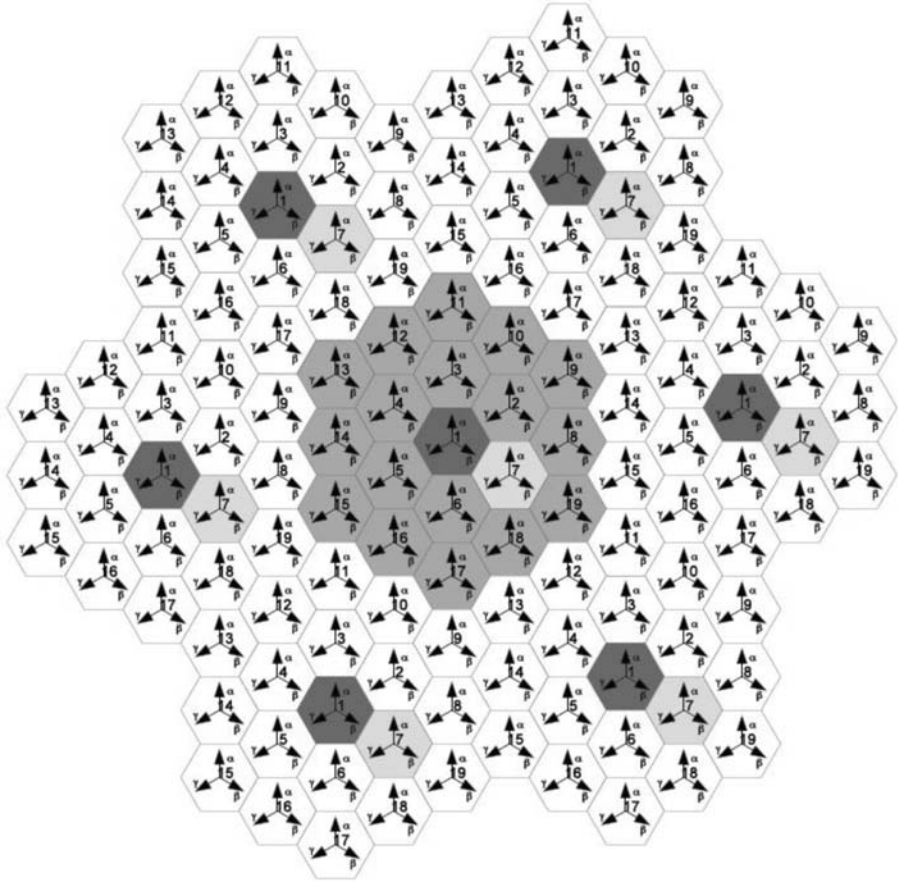


Figure 7.10 Principle of the wraparound technique [30].

attenuation, shadow fading, and antenna gains. Short-term channel variations caused by multipath propagation and fast fading may be added later to obtain a full representation of the frequency-selective channel. Figure 7.10 shows transceivers equipped with sector antennas, but they may also be equipped with beamforming antennas.

The wraparound technique is suitable for both DL and UL simulations. An advantage of this technique compared to the central cell technique is that simulation data can be collected from all cells, which may reduce the simulation time required to collect sufficient statistics.

7.2.2 Evaluation of ICI

Here, an evaluation of the ICI is done based on interference avoidance methods proposed by [30] for the WINNER radio system.

The defined scenarios for the evaluation are base coverage urban, microcellular coverage, and indoor scenarios [40]. The FDD physical layer mode is used for the base coverage urban scenario, whereas TDD is used for the microcellular and indoor scenarios. Emphasis is placed on DL interference avoidance for the base

coverage urban FDD scenarios. Both the downlink and uplink are investigated in the schemes proposed for the microcellular and indoor scenarios. According to deployment-specific baseline parameters, an intersite distance of 1 km is used for the base coverage urban (macrocellular) FDD scenario and the “Manhattan grid” with defined building and street dimensions is considered for the microcellular TDD scenario.

Although a full-buffer traffic model is assumed mostly as defined in the baseline design, in order to obtain realistic queue and delay statistics, realistic traffic models for VoIP, streaming video, FTP, and so forth are implemented in some of the schemes investigated. Detailed descriptions of the various traffic models can be found in [40]. In modeling ICI, both central cell and wraparound techniques are used.

Based on the performance parameters to be observed in different proposed schemes, both Class II and III simulators were used [39, 40]. The Class II and III simulators are system level simulators defined to capture the dynamic and quasi-static aspects of system behavior, respectively.

The studied interference avoidance schemes differ in design (e.g., traffic modeling, scheduling algorithms) because they were applied to different scenarios. As a result, they do not face the same challenges and thus evaluation criteria may vary from one case to another. Therefore, the common criteria must be identified, which show the performance as well as the technical interpretations of the relevant requirements. These are available in [44].

The mandatory criteria for performance evaluation identified based on [40] are as follows:

- CDF of the user throughput;
- CDF of the (chunk) SINR;
- The average sector throughput.

Of these criteria, CDF of the user throughput seems to be the most important one, because it can be used for deriving standard performance measure criteria such as fairness, average throughput, and spectrum efficiency, as well as the cell edge user throughput considering a defined percentile value.

The gain of interference avoidance techniques that were the focus of the investigations can be also directly read from the CDF of the SINR (the gain in mean SINR, decreasing the percentage of users having SINR below certain threshold, and so forth). The SINR measure is defined as the SINR after receiver processing but prior to decoding and geometrically averaged over the symbols (i.e., subcarrier) involved in a code block, for both the frequency-adaptive and the nonadaptive modes. In other words, this SINR represents the average channel quality of the scheduled resources.

Because network providers are often interested in maximizing the average performance per cell or sector, the average cell/sector throughput is also an important performance criterion.

Delay distribution is also an important performance criterion for end-to-end evaluations, because the delay experienced by a user has a great impact on the

user's satisfaction, especially in the case of real-time services. Delay is defined in [40] as follows:

- *Packet delay* is the time interval from when the packet enters the transmission queue to when the packet is received successfully. If a packet is not successfully delivered by the end of a simulation run, none of the information bits of the packet will be counted.
- *Packet call delay* is the time interval from when the first packet of a packet call enters the transmission queue to when the last packet of the packet call is received successfully. If a packet call is not successfully delivered by the end of a simulation run, the packet call will not be counted in the performance statistics.

Note, however, that the packet delay criterion is not applicable to the full queue traffic model. In addition, the packet call delay criterion may not be easily applicable in Class III simulators, in which only a fraction of the packet call is generally considered. This actually depends on the traffic model. For an HTTP traffic model, the packet call represents the download of one Web page (made up of several packets), whose duration (less than 1 sec) can be encompassed in one snapshot. Considering a mean reading time of 30 sec, it is likely that one snapshot will not be able to capture one complete packet call for each user, except at the price of very long simulation durations.

The delay criterion, therefore, may not be the most relevant for investigations of the cell or sector and user throughput. On the other hand, the CDF of the delay is important for scheduling investigations, because scheduling can make a trade-off between throughput and delay. Therefore, in the case of scheduling simulations, the delay CDF along with the throughput CDF are needed to evaluate the performance. Furthermore, delay could be indirectly estimated from the SINR CDF, which is an important performance criterion.

Performance of the studied schemes based on the assessment criteria is sensitive to the overall design and assumptions related to the environment and the deployment. For example, the performance of an interference avoidance scheme may depend on the resource management/scheduling strategy used. Accordingly, when analyzing the performance of a specific strategy, the other assumptions should be well described and the impact of changing these other strategies/assumptions should be investigated as much as possible.

Figure 7.11 shows the behavior of the CDF for average user throughput for different frequency reuse values when an interference coordination (IFC) scheme is employed as an interference measure.

The IFC scheme is evaluated for the frequency reuse values of $7/6$, $3/2$, and $3/1$ leading to a power reduction per cell of 10 dB in $1/7$, $2/3$, and $1/3$ of the resources. Both UT velocities of 3 and 30 kmph have been considered.

Figure 7.11 shows that the IFC scheme performs better for reuse $7/6$ over reuse $3/2$ and reuse $3/1$, in that order, whereby the overall performance degradation occurs as expected when the UT velocity increases. This holds for the average user throughput as well as for the cell edge user throughput (determined by the 5th percentile of the average user throughput). These results reflect the greater amount

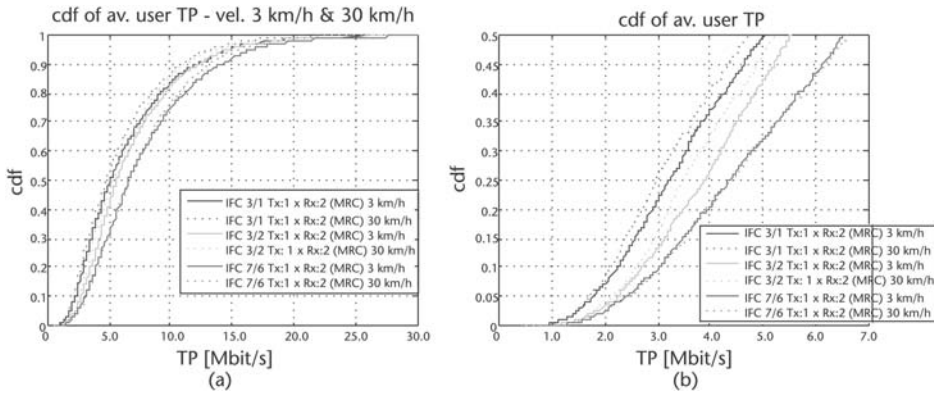


Figure 7.11 (a, b) CDF for average user throughput for different reuse values [30].

of resources made available by using the IFC scheme with the reuse 7/6 because in that case only 1/7 of the resources are reduced in power. An improvement of the situation at the cell edge border is indicated by the 5th percentile of the CDF of the average user throughput [30].

In the case of UT velocity of 3 kmph, the investigated IFC scheme used with the reuse 7/6 outperforms the cases of reuse values 3/2 and 3/1, respectively, by 15% and 23%, for the average sector throughput, and by 7% and 36%, respectively, for the cell edge throughput.

Besides these results obtained for the different frequency reuse values for the IFC scheme considered, investigations were conducted to compare to the configuration using one single transmitting antenna at the BS side yet still two receiving antennas at the UT side, keeping the MRC diversity scheme at the UT side but without IFC. This latter case corresponds to a frequency reuse value of 1 over the whole bandwidth for all cells.

Under the simulation assumptions stated earlier, the degradation of performance that was observed for the average sector throughput for all reuse values investigated for the IFC scheme revealed that the worst case is reuse 3/1. For reuse 7/6 the degradation observed amounts to -8.1% and for reuse values 3/2 and 3/1, it corresponds to -20.3% and -25.5%, respectively. The cell edge use throughput improves slightly by +0.4% for the IFC scheme with 7/6 reuse compared to the case without IFC, but it decreases for the 3/2 and 3/1 reuse values, in that order, by -6.2% and -25.7%, respectively.

This measured behavior, which indicates that a frequency reuse of 1 without any coordination addressing the ICI performs, on average, better than the IFC scheme investigated here, is very dependent on the simulation assumptions [30]. For example, the number of UTs simulated per sector is 10, whereby the intersite distance is set to 1,000m in the deployment scenario considered. This corresponds to a rather low density of users, which does not ensure a uniform distribution, which in turn would be beneficial for schemes based on restrictions in the cell. Furthermore, the type of scheduler used in combination with the IFC scheme would also have an impact on the performance achieved. Preliminary investigations conducted with a scheduler based on a minimum target throughput per served

user, and services such as VoIP, are given in [30]. The results shown there indicate that significant improvements at the cell border could be achieved by the IFC scheme with 7/6 reuse over the case without IFC.

The CDFs of the average SINR at UT receiver output as further assessment criteria are shown in Figures 7.12 and 7.13 for all reuse values investigated at both UT velocities.

Further evaluations of the behavior of the IFC scheme under the same simulation assumptions combine this scheme with transmit diversity at the BS side, with the

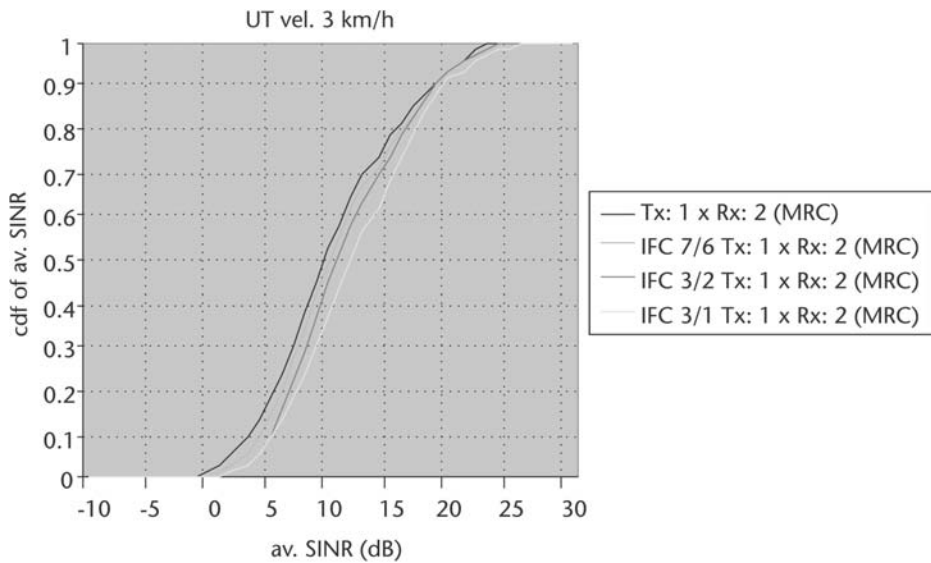


Figure 7.12 CDF of the average SINR (3 kmph) [30].

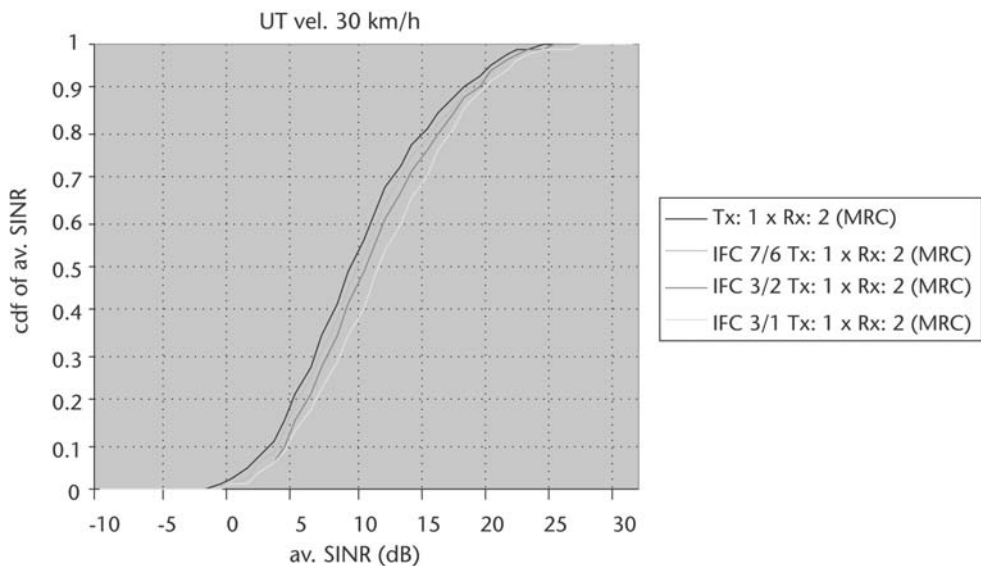


Figure 7.13 CDF of the average SINR (30 kmph) [30].

diversity at the UT side being maintained by the MRC scheme. For this, the fixed grid of beams was evaluated for the IFC scheme using the reuse value of 7/6; two and four transmitting antennas at the BS side were considered for a UT velocity of 3 kmph. The results are shown in Figure 7.14.

Improvements for the assessment criteria are achieved when the single antenna at the transmitting side is replaced by a fixed GoB configuration in all cases; for the combination with the IFC scheme, however, the same behavior as before is observed: The cell edge throughput degrades compared to the reuse 1 case by -6.5% and -2.7% , respectively, for two and four transmitting antennas used in the fixed GoB scheme. Degradation for the average sector throughput amounts to -4.8% and -6.8% , respectively, in those cases compared to the reuse 1 scheme.

The CDFs of the average SINR at the UT receiver output are shown for the cases of combination with GoB for a UT velocity of 3 kmph in Figure 7.15.

The observations made show that degradation of performance occurs when the IFC scheme is used in combination with GoB in all cases. Furthermore, the impact of the GoB scheme is predominant so that the IFC scheme in this combination behaves like the frequency reuse 1 case.

Figure 7.16 shows a multicell simulation setup with the central cell technique for trisector cells used for the evaluation of interference cancellation techniques [31]. The scenario considers a hexagonal deployment of 57 sectors with an intersite distance of 1,000m. The multicell environment is accounted for using the central cell technique, which means that the UTs and system functions are effectively modeled in the three sectors of the central cell (the dark ones in Figure 7.16); the remaining sectors are accounted for via simplified models. The ICI modeling involves the 56 interfering sectors. The channel of the 8 dominant interferers in a long-term basis is accurately modeled, whereas the other interferer channels are modeled as SISO and single-path ones. All of the interfering BSs are assumed to be transmitting at full power. Moreover, all of the resources are permanently occupied in the interfering cells, which means that the system operates at full load.

Both frequency-adaptive and nonadaptive modes are considered, using OFDMA with chunk-wise resource allocation for both modes. The UT velocity is 3 kmph in frequency-adaptive mode and 50 kmph in nonadaptive mode. The *score*-

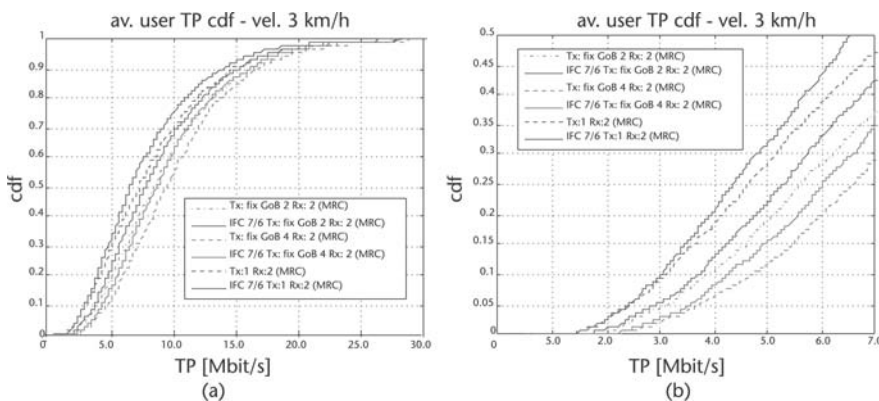


Figure 7.14 (a, b) CDF of the average user throughput (IFC 7/6 with GoB) [30].

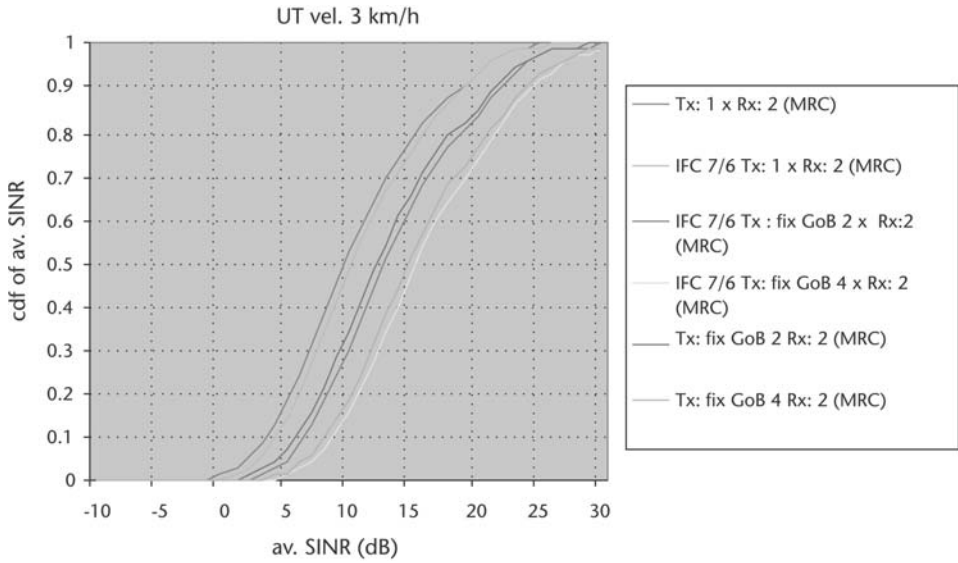


Figure 7.15 CDF of the average SINR (IFC 7/6 with GoB) [30].

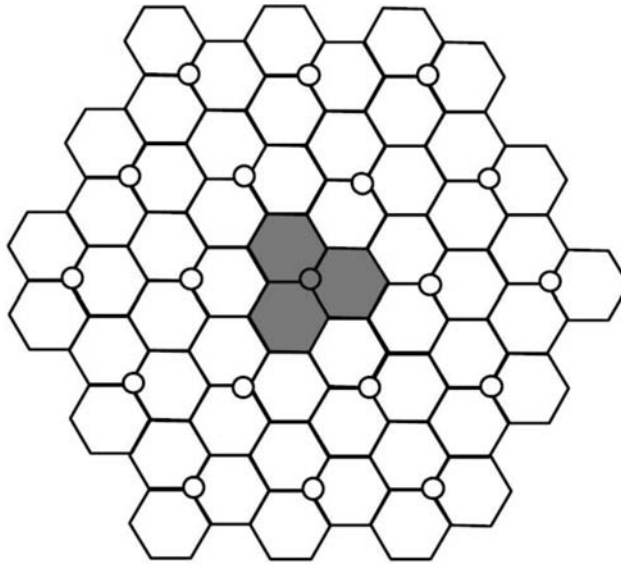


Figure 7.16 Multicell simulation setup for interference cancellation evaluation [31].

based algorithm [31] is used for scheduling in the adaptive mode. Nine users are scheduled simultaneously in the 45-MHz bandwidth, which leads to 16 chunks per user, which are not necessarily colocalized. The total number of users in each sector is 24 on average, which implies 12 users in the downlink direction per half frame using the FDD half-duplex scheme. A full-buffer traffic model is used for all users, that is, the users always have data waiting for transmission. HARQ with chase combining is fully simulated with four retransmissions and explicit feedback

of ACK/NACK messages. Link adaptation is simulated with a one-frame delay for the CQI feedback, the latter of which is assumed to be perfectly received. Ten modulation and coding schemes are used, with BPSK QPSK, 16QAM, and 64 QAM modulations and DBTC (see Chapter 3) with rates 1/2, 2/3, and 3/4.

Only single-antenna transmission is employed at the BS transmitter, whereas the UT always uses two receiving antennas. Moreover, perfect estimation of the correlation matrix of the interference + noise term is assumed, as well as that of the channel of the user of interest.

In this study, only the cancellation of the two dominant interferers was evaluated, because for complexity reasons it may be difficult to estimate the signal coming from several interferers at the same time.

In addition, the focus is on the cancellation of the dominant interferers on average (i.e., identified from its long-term averaged received power and evaluated in simulations from the path loss, shadowing, and antenna gains), because it may be difficult to identify the dominant interferer on an instantaneous basis in practice.

The limited system performance that can be reached with such an interference cancellation algorithm was also investigated. In other words, the ICI coming from the dominant interferer(s) has been assumed to be perfectly estimated and subtracted from the received signal, without taking into account concerns about the feasibility of the scheme (sufficient processing gain or efficient implementation). Therefore, the intercell interference cancellation (IIC) scheme is referred to as “ideal IIC” in the results. As a consequence, the results provided have to be considered as an upper bound of the performance attainable in practice.

In Figure 7.17, the IIC receiver canceling the dominant interferer is denoted by IIC(1), whereas the receiver canceling the two dominant interferers is denoted by IIC(1+2). All the results presented are obtained in the nonadaptive mode; the results for the frequency-adaptive mode can be found in [31]. The CDFs of the average user throughput for the different schemes are also shown in Figure 7.17, together with an enlargement of the cell edge region.

In addition, the CDF of the (geometrically) average SINR measured over the allocated chunks is shown in Figure 7.18 for the different mitigation techniques considered. We can see that the SINR improvement is not constant; it is lower in the low-SINR region and higher in the high-SINR region.

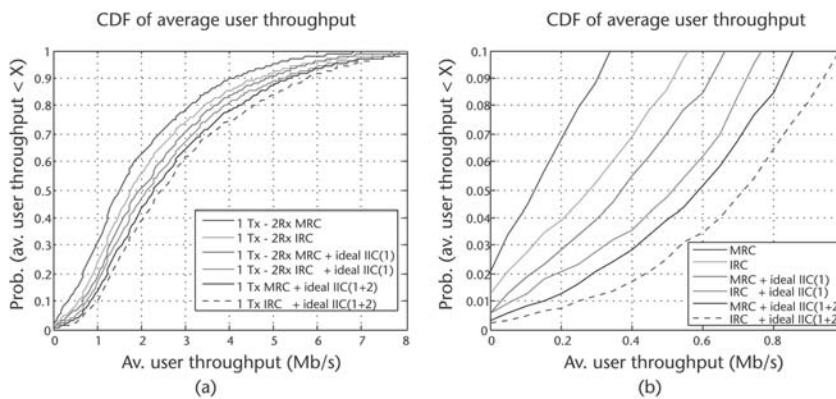


Figure 7.17 (a, b) CDF of average user throughput with interference mitigation technique [31].

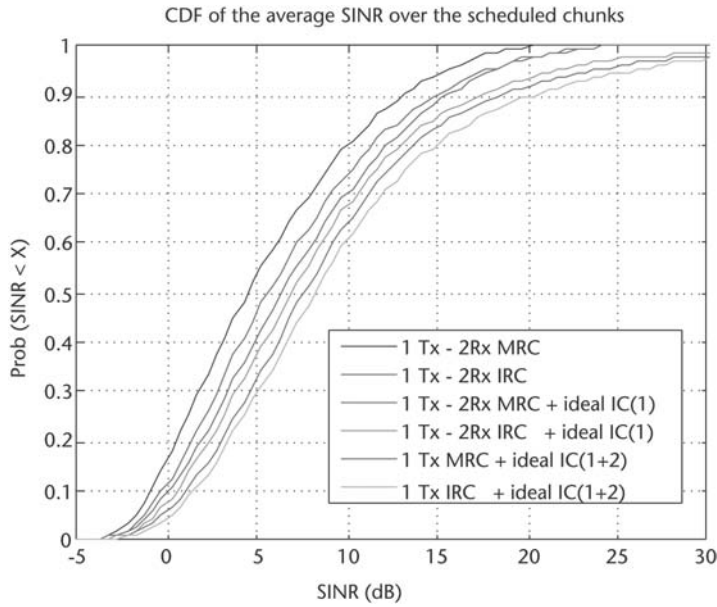


Figure 7.18 CDF of the average SINR measured over the allocated chunks.

For all of the considered figures of merits, the MRC is significantly outperformed by the interference rejection combining (IRC), which is itself significantly outperformed by the IIC schemes. The greatest gains are observed on the cell edge throughput, where factors 2.5, 3, and >4 can be achieved over the MRC receiver for the IRC, MRC+IIC(1), and MRC+IIC(1+2) schemes, respectively. In terms of average SINR (observed at 50% of the SINR CDF), the cancellation of the dominant and then of the two dominant interferers brings 1.8- and 3-dB gains, respectively. Note that the additional gain brought by canceling the second dominant interferer is lower than the initial gain provided by the cancellation of the dominant interferer, but is still significant (approximately +30% in terms of cell edge throughput).

At last, the relative gains brought by the IIC schemes to the IRC receiver are lower than those obtained with the MRC receiver. This is explained by the interference mitigation already performed by the IRC, which reduces the benefits of the additional cancellation.

A nonideal estimation will result in a significant processing gain in the system (which reduces the service throughput) and high computational complexity because a channel decoding step has to be performed in the interference detection process. Therefore, it is likely that the first practical devices implementing IIC features will achieve much lower gains [31]. Nevertheless, the results obtained here provide upper bounds on the performance that can be reached with the described interference mitigation schemes.

The use of IIC schemes with beamforming techniques at the transmitter such as the GoB would help to reduce the large amount of interference received when the UT is hit by an interfering beam. In particular, it is likely that only one strong interferer needs to be canceled in that case. However, reliable estimation of the interferer's channel may require the beam of the interferers to be estimated, which may impose additional complexity to the UT.

In a similar way the use of smart antennas can be evaluated. Transmit beam-forming, for example, in the form of the WINNER wide-area baseline scheme fixed GoB, was shown to be an efficient means to reduce the interference spread in the system and in particular to protect users at the cell border from ICI [4]. With SDMA on top of GoB, the system performance can be improved at the expense of slightly less protection of the cell edge users. Also, the use of multiple-antenna receivers at both the BS and the UTs was shown to be an efficient means to mitigate interference. This allows implementation of receive diversity combining schemes such as MRC and spatial interference suppression schemes such as IRC. The impact of such implementations is shown in Figure 7.19, which shows that the best performance is reached by the SDMA technique combined with the tapered fixed GoB built by four antennas, which gives a significant gain of approximately 47% when compared to the closed-loop transmit diversity method.

MRC provides considerable improvements in interference tolerance both when used at user terminals in the downlink and at BSs in the uplink. Additional improvements are achieved with IRC, in particular, for DL reception at user terminals. In this context it should be mentioned that IRC is a more complex method than MRC, but studies on how this complexity can be reduced, thereby saving, for example, terminal battery life, have also been conducted. For the downlink, the combination of transmit beamforming at the BSs and multiple-antenna reception with IRC at UTs has been identified as an attractive combination. Furthermore, different aspects of macrodiversity (transmit diversity from several BSs) were also studied. For example, different receive combining methods for SFN and MFN networks in conjunction with MBMS were investigated. In addition, one macrodiversity method based on cyclic delay diversity (CDD) was shown to have potential to further improve the ICI situation at cell border areas.

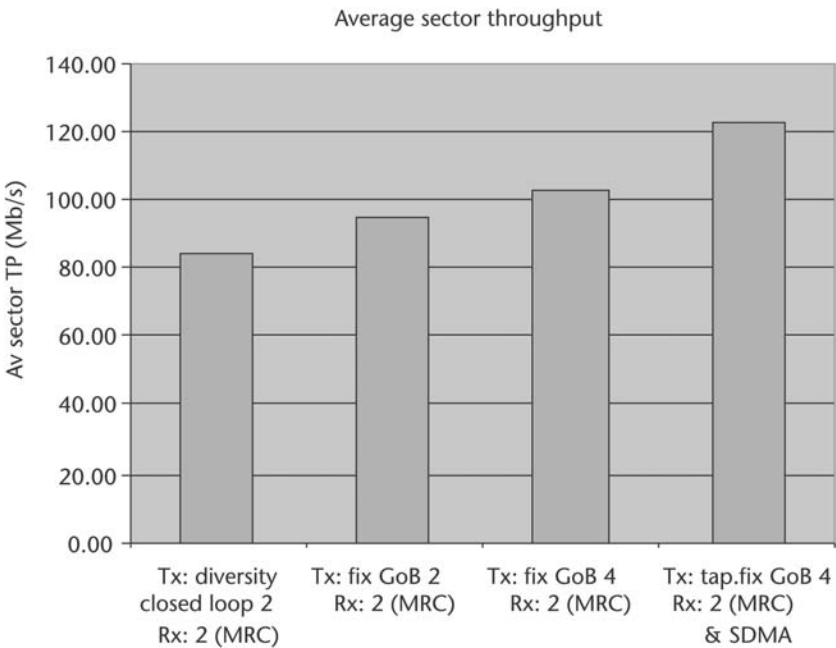


Figure 7.19 Average sector throughput achieved by different diversity techniques [4].

A combination of beamforming and resource management methods (e.g., scheduling) provides further improvements in system performance. Such combinations were extensively studied in [4], and some results were given in Figure 7.9.

7.3 Interference Scenarios for PAN-Optimized Radio Interfaces

The FP6 IST project MAGNET Beyond [9] proposed and optimized two air interfaces for wireless PANs targeting high-data-rate and low-data-rate applications. The HDR approach was based on an MC-SS PHY layer along with an IEEE 802.15.3-based MAC layer. Research work developed PHY, MAC, and cross-layer enhancements to the core MAGNET HDR and LDR protocols that improve the spectrum efficiency and mitigate the interference suffered by the PAN (e.g., MAGNET) wireless links [45]. It is expected that such techniques will be indispensable for the delivery of streaming or real-time audio and video over wireless links and networks. This necessity is dictated by the time-varying nature of the wireless channel capacity, the higher BER and lower bandwidth compared to the wired links, and the stringent jitter, delay, and BER requirements of audio/video transmission.

The following interference scenarios were identified as important cases that should be modeled and studied in order to cover the most current and future anticipated causes of interference in a PAN radio environment:

- MAGNET MC-SS + IEEE 802.11a-based WLAN;
- MAGNET MC-SS + (noncollaborative) MAGNET MC-SS;
- MAGNET MC-SS + generic partial-band interferer.

Advanced cross-layer optimization techniques were developed in order to mitigate the adverse effects of interference to the MC-SS communication link [45]. The design and testing (through simulations) of the appropriate interference mitigation schemes included a working model of the interference that the actual system was expected to experience in operating conditions. Here, a simple model of interference that incorporates PHY and MAC characteristics of the interfering signals in order to capture the temporal and frequency properties of these unwanted signals is described.

Figure 7.20 shows a wireless link with data transmission between a desired transmitter and a receiver through a radio channel model. In addition to the studied link, a number of interfering transmitters are present. These are, in contrast to the studied link, controlled by interference model parameters rather than by data transmitted over a certain physical layer.

The blocks in Figure 7.20, including the “Interfering transmitter” block, were specified in the MAGNET simulation chain. For specification of the latter, the existing MAGNET radio channel models needed to specify the detailed time correlation properties as well as correlations between the desired link and the interfering link. Such correlation properties are important when developing and evaluating interference mitigation techniques.

By defining the interference scenarios described earlier, answers to the following basic questions regarding the characteristics of the interference were sought:

7.3.1.1 Measurements

The measured data was obtained using a MIMO channel sounder based on the correlation principle. Up to 16 Tx channels could be measured truly simultaneously, each branch with 1W of transmitting power. On the Rx side 8 channels were measured in parallel and, using switching, each branch could be extended so that in total 32 Rx channels were measured. Additional information about the sounder is available in [46]. The work presented here used 4 Tx and 32 Rx channels and the full-complex 4×32 MIMO channel was measured in a time-triggered way at a rate of 60 Hz. In a postmeasurement processing procedure, the measurements were compensated for the sounder system response and the bandwidth was limited to about 100 MHz. The carrier frequency was 5.5 GHz, which was chosen to fit the available antennas. Note that the MAGNET HDR carrier is 5.2 GHz, but the difference is believed to be unimportant.

Three users were involved in the measurements, Rx-A, Rx-B, and the Tx. In the scenario, each user carries four identical devices located on the right shoulder, the left part of the chest, on the right hip, and on the left ankle. Each of the four identical Tx devices has one directional stacked patch antenna of size about 3×3 cm. Each of the Rx-A devices has four planar inverted-F antennas arranged at the corners of a programmable circuit board (PCB) with dimensions of about 3×3 cm. Each of the Rx-B devices has four monopole-like antennas arranged at the corners of a PCB. These arrays are rather large, about 4×11 cm, but they could be integrated into, for example, clothing. More details of the body-worn devices can be found in [45].

An overview of the indoor measurement site is shown in Figure 7.21, where each of the seven locations are shown. The environment can be described as a laboratory area connected to a corridor through large openings. During a measurement, the three users were distributed in the environment as specified in Table 7.1, where each line in the figure represents three measurements: two where the users walk along a rectangular path, and one where the users walk along a figure-eight-shaped path. For the rectangular-shaped path, the users keep the same orientation by walking forwards, sideways, and backwards along the path. Two different sets of orientations are used, and at each location the users remain inside a rectangular area of approximately $2.4\text{m} \times 0.8\text{m}$.

During the 90-second measurement period, the users are instructed to walk at their own desired speed, thus resulting in different speeds and many different combinations of distances between users and relative orientations of the users for the figure-eight-type of measurements.

7.3.1.2 Measurement Analysis

After compensation for the response of the measurement system, the result of each measurement is a set of complex samples $h(d, t, r, p)$, where d is an integer used to index the samples in delay and similarly p is an index to the 60-Hz measurement samples. The transmitter and receiver branch indices are t and r , respectively. In the following only the wideband power is considered, computed as follows:

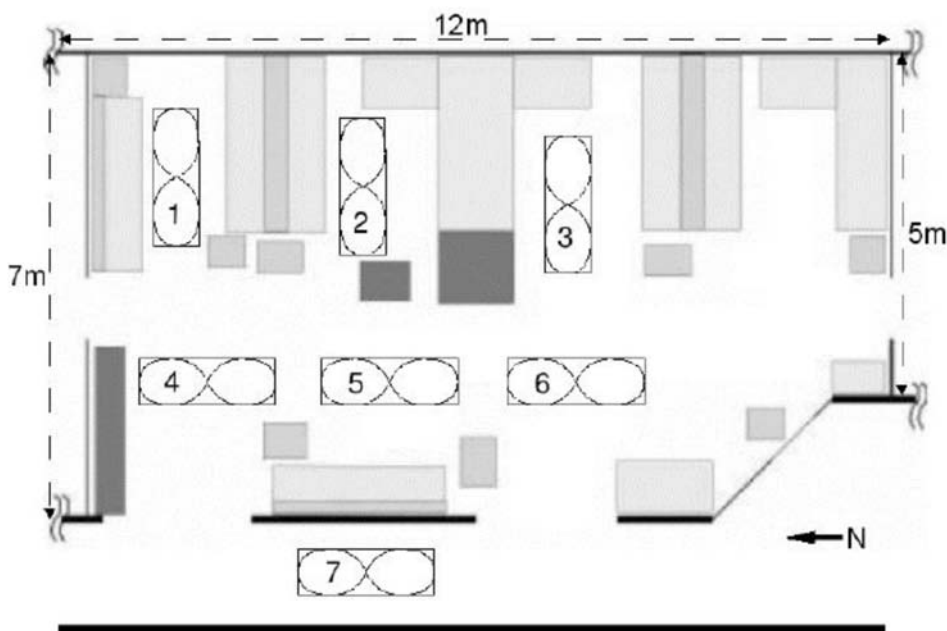


Figure 7.21 Measurement site and locations used during the measurements [45].

Table 7.1 Overview of PAN Measurement Site with Location Numbers and Orientations

2 (E, W)	3 E	1 E
2 (E, W)	2 E	1 E
4 (N, S)	5 N	6 N
5 (N, S)	4 N	6 N
6 (N, S)	4 N	5 N
7 (N, S)	5 N	3 (N, W)

$$s(t, r, p) = \sum_{d=0}^{D-1} |h(d, t, r, p)|^2 \tag{7.5}$$

where D is the number of samples in the delay dimension. For example, this corresponds to a system where an equalizer is employed to collect the power versus frequency. In the analysis below the power is normalized according to

$$s_n(t, r, p) = \frac{s(t, r, p)}{\frac{1}{P} \sum_{p=0}^{P-1} s(t, r, p)} \tag{7.6}$$

or independently for each Tx-Rx branch, where $P = 5,400$ is the number of channel samples in the measurement. The normalization removes differences such as those due to antenna efficiency, average polarization mismatch, and average path loss. The log-power is defined as

$$s_{\log}(t, r, p) = \log_{10}[s_n(t, r, p) - s_{\min}] \quad (7.7)$$

where

$$s_{\min} = \min_{t,r,p} [p_n(t, r, p)] \quad (7.8)$$

Equation (7.8) shifts the linear power values to ensure the correct mapping of the logarithm function and $\alpha < 1$ is a constant chosen as 0.99. The log-power computed from the measured data using (7.7) is the starting point for the analyses given below. To allow the combination of the data acquired for the different Tx-Rx branch combinations, the mean of the log-power in (7.7) is subtracted and the data is normalized individually for the branch combinations. An example of the measured log-power is given in Figure 7.22.

Before analyzing the properties of the measured data, the stationarity of the underlying random processes should also be investigated.

The observed log-power $s_{\log}(\cdot)$ is a random variable resulting from a dynamic combination of variations in path loss due to changing distances, body shadowing due to changes in the relative orientation of the Tx and Rx devices, and fading because of minor changes in propagation path lengths leading to phase changes.

Figure 7.23 shows an example of a CDF estimated from the data of different Tx and Rx combinations of a single measurement run.

Figure 7.23 further shows the CDF of Gaussian data with mean and variance estimated from the measured data. Disregarding the tail with very low probability, the match between the experimental CDF and the Gaussian CDF is significant.

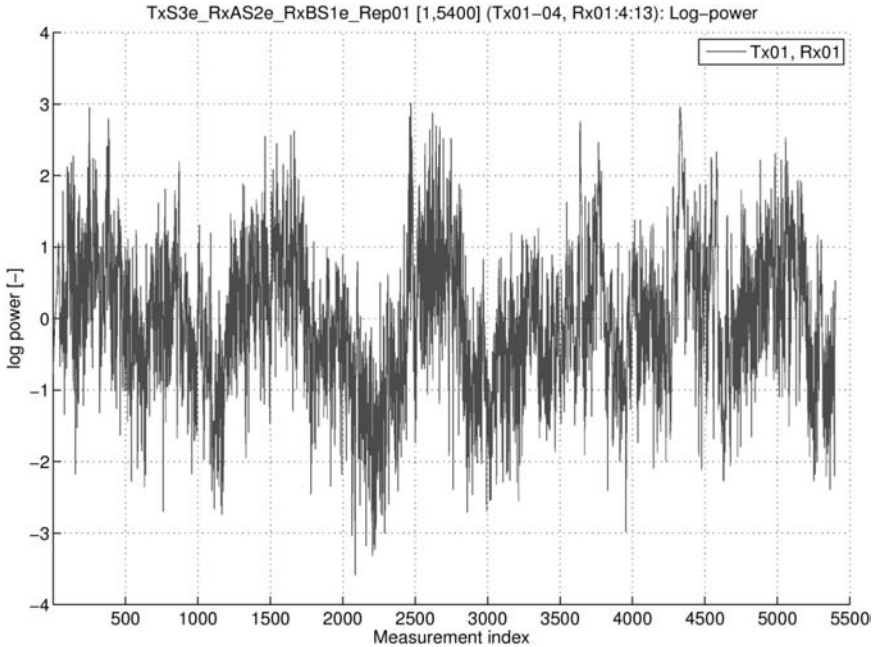


Figure 7.22 Example of measured log-power data for a single Tx-Rx branch combination [45].

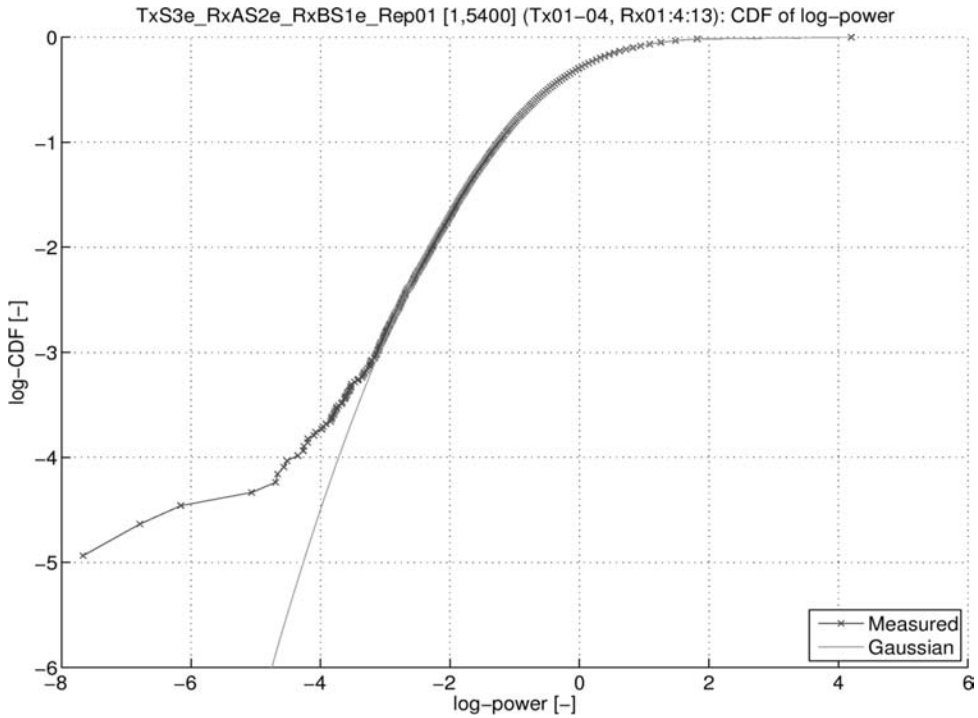


Figure 7.23 CDF of the log-power for a specific measurement [45].

The match is a general tendency among the different measurements. This is shown in Figure 7.24.

The autospectrum was computed for all combinations of Tx-Rx branches and for each measurement scenario. The average of the spectra obtained with different Tx-Rx combinations is used to represent the spectrum for every scenario. This is shown in Figure 7.25.

A sample of the derived stochastic model based on the autospectra obtained from measurements is shown in Figure 7.26.

Judging from the estimated spectra, the sampling rate is indeed sufficient and other sampling rates may be obtained by interpolation of the model.

The modeling of the temporal correlation functions of MC-SS wireless channels in MAGNET was also based on measurements reported in [47, 48]. Most multipath components are not affected by the movement in the chosen indoor scenario, and the time correlation of the channel is quite strong for several hundred milliseconds. The MC-SS channels were modeled according to these measurements and considerations. The channel impulse response from transmitter to receiver is of the form:

$$h(t, \tau) = h_{nf}(t, \tau) + h_f(t, \tau) \quad (7.9)$$

where $h_{nf}(t, \tau)$ is a nonfading part, resulting from multipath components not affected by movements in the environment, and $h_f(t, \tau)$ is a fading part, resulting from multipath components affected by movements in the environment. These are in themselves assumed to be on the form:

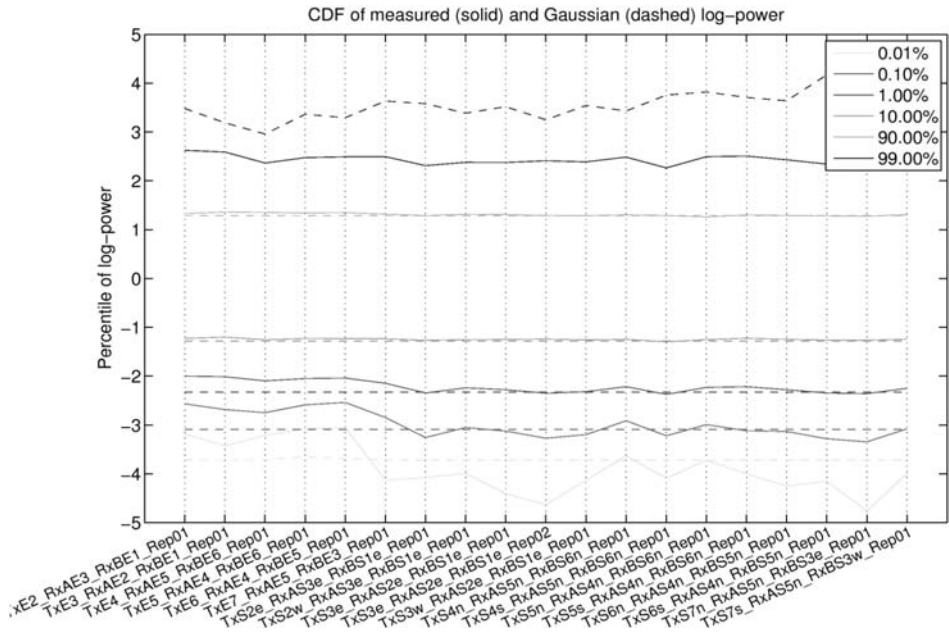


Figure 7.24 Percentiles for the different measurements compared to percentiles of matched log-normal distributions [45].

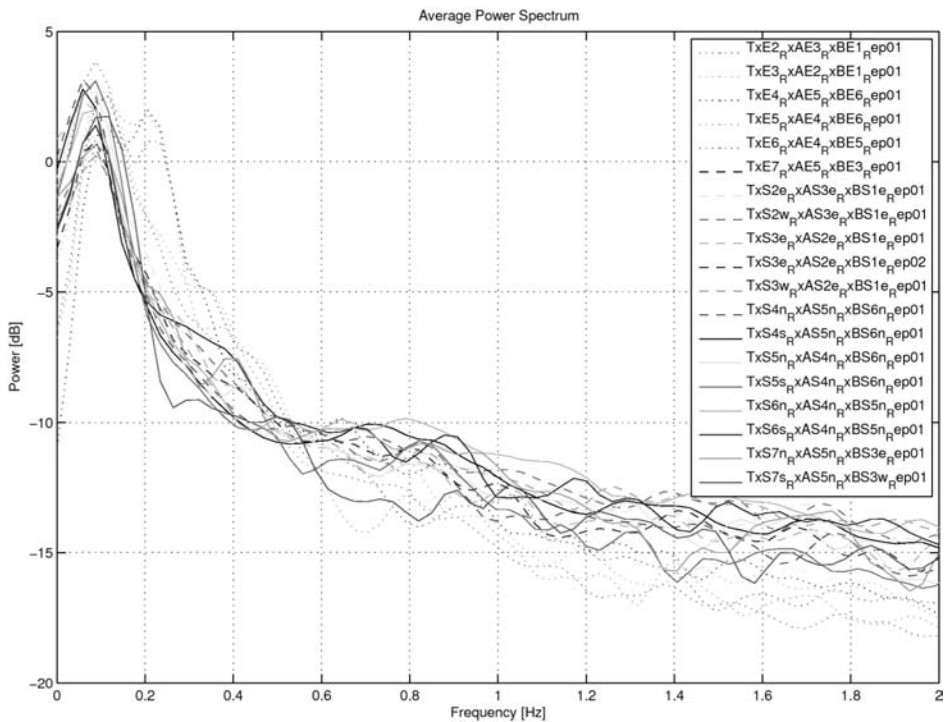


Figure 7.25 Autospectra for different measurement scenarios, computed as an average over the different Tx-Rx branch combinations [45].

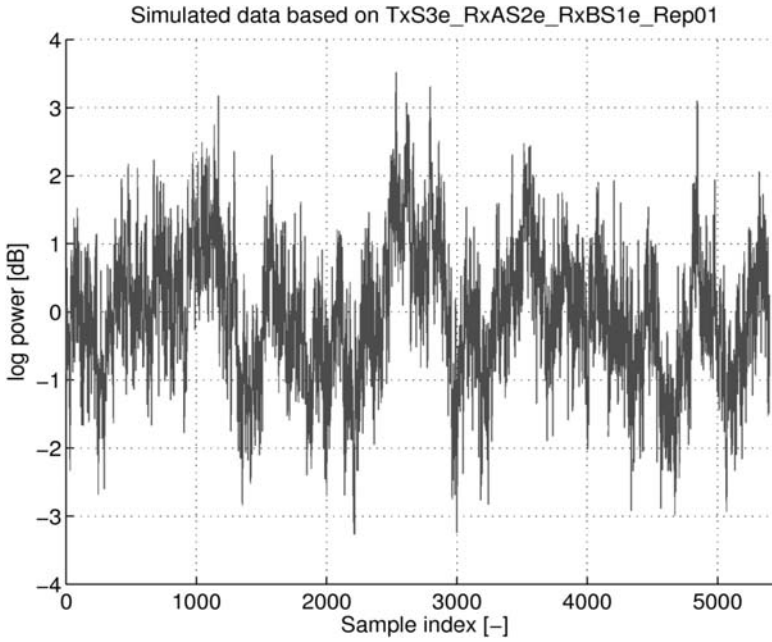


Figure 7.26 A sample path of the derived stochastic model [45].

$$h_{nf}(t, \tau) = \sum_k a_k \delta(\tau - \tau_k) \quad (7.10)$$

and

$$h_f(t, \tau) = \sum_k b_k(t) \delta(\tau - \tau_k) \quad (7.11)$$

where a_k and $b_k(t)$ are the complex fixed/fading tap coefficients and τ_k the corresponding tap delays.

The power-delay profiles were chosen from the extensive measurements and parameterizations found in [47, 48]; the creation of the time correlation of the fading part of the channel, however, requires more consideration. A standard way of creating a signal with a certain time correlation is to feed white noise to a filter with impulse response. This, however, makes it difficult to generate correlated impulse responses that are well separated in time, since it requires filtering of the white signals for as long as required. This is essentially equivalent to generating many impulse responses, which are close in time, and then using only a small fraction of them in the simulations. Because interferers are expected to transmit in bursts, it is of interest to use a model where correlated channels in corresponding bursts are more easily generated (i.e., the channel values between transmit bursts are not required). A more versatile method than filtering, which has proven useful and has good convergence properties, was developed in [49]. Following this approach wide-sense stationary uncorrelated scattering channels can be generated, where the fading coefficients $b_k[m]$ have arbitrary correlation/Doppler spectrum.

Compared to [49], [45] made some changes reflecting the fact that *sample-spaced* channels were being generated rather than *continuous-time delay* channels.

With modeling, usually the channels from different interferers will be completely independent. In reality, however, a user with a handheld terminal rotating while receiving data, when there are two interferers on opposite sides, will in most cases create a strong (negative) correlation between the received powers from the two interferers. Models that allow for generation of correlated channels from different interferers should be preferred for such interference investigations. Models where channels from different interferers are correlated, however, will be much more involved and require many more parameters and processing power when performing simulations. Note that at the time of the investigations reported in [45], there were no in-depth studies available of correlation between channels from different interferers under different MAGNET scenarios.

7.3.2 Temporal Properties of Interferers

Many existing and future interference sources will be the different wireless interfaces transmitting packet data. A certain amount of the interference may, however, come from more continuous sources, which are switched on and off at times. In all of these cases, high activity is likely to be followed by high activity, and vice versa. This suggests a two-stage model that uses: (1) a stochastic model for the activity level and (2) a separate stochastic model translating the activity level into packet/burst transmission times and packet lengths. This model is shown in Figure 7.27.

Reference [45] suggests the use of a finite state machine (FSM) to model the activity level of a terminal/interferer. The FSM has two or more levels of activity, each represented by a state, for a particular interference source. The state transitions can be modeled as a discrete or continuous-time Markov chain. A continuous-time FSM model is shown in Figure 7.28.

The continuous-time Markov chain works as follows:

- The amount of time the interferer spends in state i before making a transition into a different state is exponentially distributed with mean $1/u_i$.
- When the process leaves state i , it next enters state j with some probability P_{ij} , which satisfies:

$$P_{ii} = 0, \text{ for all } i;$$

$$\sum_j P_{ij} = 1, \text{ for all } i;$$

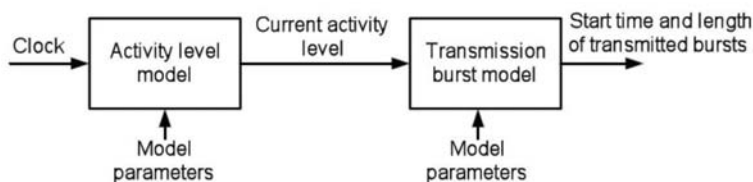


Figure 7.27 Two-stage model for the generation of start times and lengths of transmitted bursts of interference [45].

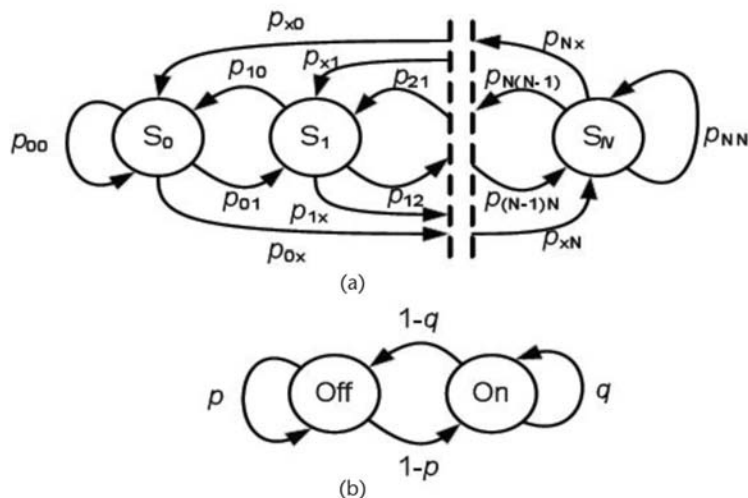


Figure 7.28 FSMs modeling of an interferer activity level: (A) generic model and (B) simplest form of on/off model [45].

Initialization

- N Number of states in FSM
- \mathbf{P} $N \times N$ matrix of state transition probabilities
- \mathbf{u} $1 \times N$ vector of transition rates u_i out of state i
- $S[0]$ Initial state (randomized according to steady-state distribution in most cases)

Input

- Nothing (except maybe a clock/time reference triggering state transitions)

Output

- S Activity level (current state)

Internal State

- S Current state

The current activity level of the interference transmitter is translated into an activity level to bursts of transmitted energy (packets), determined by both their start times and their lengths. This can be done with parameterized processes for start times and lengths of bursts (packets), where the current activity level determines the process properties. Possible choices include a Poisson process for burst start times and a Pareto or an empirical measurement-based distribution for the burst lengths [45].

Initialization

- $D \times N \times M$ parameter matrix with one row of M parameters for each activity level

Input

- S Current activity level (state)

Output

- t_start Start time (in sample/s) of next transmitted burst
 t_len Length (in sample/s) of next transmitted burst

Internal State

- None or t_start/t_len of previous burst, depending on the type of processes chosen (Memoryless processes, for example, the Poisson process, do not need an internal state.)

A number of scenarios in the form of parameter sets for the interference model can be defined. The choice of distributions and parameters is rather arbitrary; they should be simple enough for implementation and generic enough to incorporate a number of different interferers. However, the aim of the following is to have a common basis for future system performance analysis.

In each activity level state, the frame interarrival times are exponentially distributed at a rate determined by the activity level state. The frame lengths are distributed according to an empirical distribution developed based on actual measurements of frame lengths in 802.11 WLANs. In the following, the mean frame length is denoted as $T_{IBmeanLength}$.

More specifically, the difference between the (unadjusted) start times of two consecutive bursts of interference (frames) is exponentially distributed with the following probability density:

$$f_{IBinterv}(x) = \lambda_i \exp(-\lambda_i x), x \geq 0 \quad (7.12)$$

with mean $1/\lambda_i$. The differences between the start times of different pairs of consecutive bursts are independent. The exponential distribution of the interval lengths leads to a Poisson distribution of the number of start times in a given time interval. Note that we assume that all u_i are at least an order of magnitude smaller than all λ_i . This guarantees that a certain number of frames are transmitted before switching over to the next activity level. We consider four activity states:

- State 1: $\lambda_1 = (2T_{IBmeanLength})^{-1}$ (i.e., interference during half of the time in the mean)

- State 2: $\lambda_2 = (4T_{\text{IBmeanLength}})^{-1}$ (i.e., interference during a quarter of the time in the mean)
- State 3: $\lambda_3 = (8T_{\text{IBmeanLength}})^{-1}$ (i.e., interference during an eighth of the time in the mean)
- State 4: $\lambda_4 = 0$ (i.e., no interference)

To generate the starting times of frame transmissions in such a way that consecutive frame transmissions do not overlap, the following procedure is employed: First, a list of *unadjusted start times* is generated. Second, durations are associated with the burst start times and overlaps are eliminated.

The unadjusted start time generation process works as follows:

1. $t = 0.0$; $k = 1$; $S(k)$ (defining the current state): random number between 1 and 4.
2. The time of the next state transition is Y_k , where $Y_k - t$ is generated according to the distribution $f_Y(x) = u_{S(k)} \exp(-u_{S(k)}x)$, $x \geq 0$.
3. The random variable Z_t is subject to the distribution $f_Z(x) = \lambda_{S(k)} \exp(-\lambda_{S(k)}x)$, $x \geq 0$, where $\lambda_{S(k)}$ depends on the current state $S(k)$ as defined earlier.
4. If $t + Z_t < Y_k$:
 $t + Z_t$ is added to the list of unadjusted start times;
 $t := t + Z_t$;
the procedure continues with step 3.
5. Else,
 $t := Y_k$;
 $k := k + 1$;
a state transition is performed according to state transition probabilities P_{ij} ; the procedure continues with step 2.

After completing the unadjusted start times generation, the list is processed entry by entry (starting from the first entry) as follows:

- A burst duration is generated subject to the frame length distribution and associated with the start time.
- If the resulting burst overlaps with earlier bursts of interference, the burst is shifted entirely rightward (adjusted) on the time axis until there is no more overlapping with the preceding burst.

The frame transmission time is generated in a two-step process:

1. First the frame length in bytes is generated according to the CDF shown in Figure 7.29 where A and B are parameters that depend on the type of data transmitted over the WLAN.
2. Then the transmission time is calculated as $t_{tr} = 8 * \text{frame length (in bytes)} / \text{transmission rate (in bps)}$. The transmission rate is selected when performing an activity level state transition with equal probability among all 802.11a

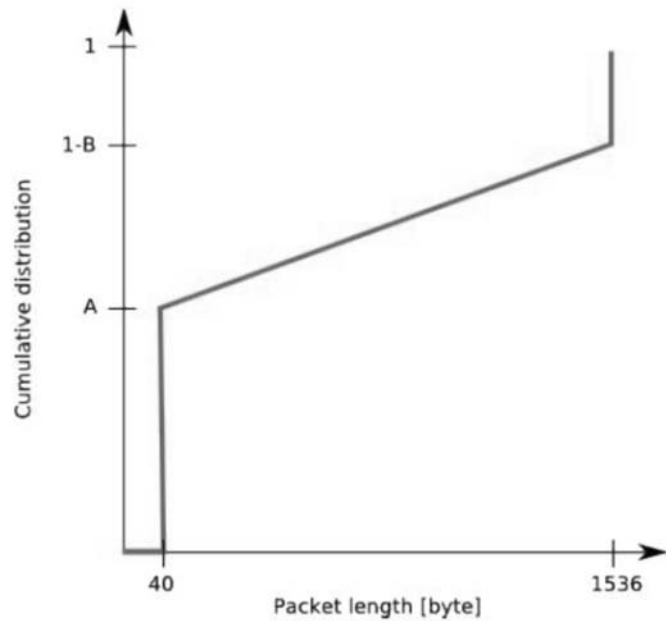


Figure 7.29 Model CDF of frame lengths in 802.11 WLANs [45].

available transmission rates: 6, 9, 12, 18, 24, 36, 48, and 54 Mbps (and it is kept constant for all frames transmitted until the next state transition).¹

Table 7.2 shows two possible scenarios. A scenario definition includes specific values for the above-mentioned parameters except for the mean SIR γ_{mean} , which depends on the distance of the interfering sources from the PAN, and on the transmitting power levels.

The corresponding transmitted signals are created with correct frequency characteristics. To achieve a high accuracy, actual signals representing transmitted packets with the correct modulation formats and coding can be generated. This

Table 7.2 Interference Model Parameters for 802.11a Interferers [45]

Scenario Name	Spectral Mask	N_I	A	B	$1/u_i$ (ms)	P		
IEEE 802.11a, heavy traffic	One of the four specific IEEE 802.11a spectral masks, defined in Appendix 1	4	0.7	0.2	[20]	[0	0.4	0.4
					20]	0.5	0	0.3
					20]	0.2	0.4	0
					20]	0.3	0.3	0.4
IEEE 802.11a, moderate traffic	One of the four specific IEEE 802.11a spectral masks, defined in Appendix 1	2	0.7	0.2	[20]	[0	0.4	0.4
					20]	0.5	0	0.3
					20]	0.2	0.4	0
					20]	0.3	0.3	0.4

1. This simplifies the calculation of $T_{IB\text{meanLength}}$ and thus λ_i for this state.

will, however, increase the complexity of the simulation process, therefore, [45] proposes to use the following strategy:

- Generate white Gaussian noise with unit power/variance, in the time interval $[t_start, t_start + t_len]$;
- If the interfering system has a particular property that we want to include, this is simulated by modifying the white noise (e.g., replacing the first part by a copy of the last part, creating a CP emulation for OFDM-type interference).
- Use a discrete-time filter (FIR/IIR) to create the proper spectral shape (baseband).
- Multiply the signal by the correct carrier factor (within simulated bandwidth).

Initialization

- **s_type** Parameter vector describing the specific type of interference created. First entry of parameter vector:
 - 0 = White noise only
 - 1 = OFDM-type cyclic prefix (second parameter gives CP length)

Input

- **t_start** Start time of burst
- **t_len** Length of burst.
- **shape** Frequency characteristics in terms of filter (FIR/IIR) coefficients
- **f_c** Center frequency of burst
- **sig_ptr** Pointer to the interference signal vector where burst is added

Output

- Nothing; modifies the interference signal vector directly

Internal State

- None

This step yields a unit power interference signal from one interferer.

The MAGNET HDR simulation bandwidth is 40 MHz, which essentially makes room for *two* in-band (cochannel) interferers and *two* adjacent channel interferers to have a significant impact on the HDR performance. These *four* interferer locations, relative to the simulated bandwidth, are shown in Figure 7.30. AL and AU denote the adjacent channels, lower and upper, while CL and CU denote cochannel, lower and upper, interferers.

Removing any components outside the simulated bandwidths gives the four in-band interference spectral masks, which are shown in Figure 7.31.

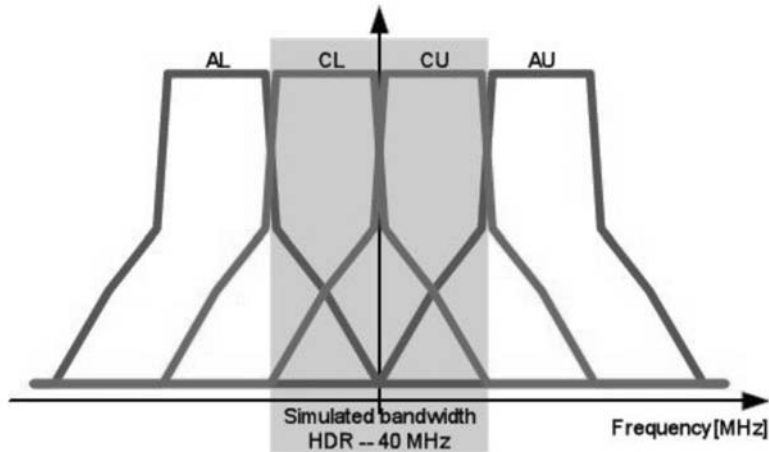


Figure 7.30 The 40-MHz simulated HDR bandwidth and the four major interferer spectral masks [45].

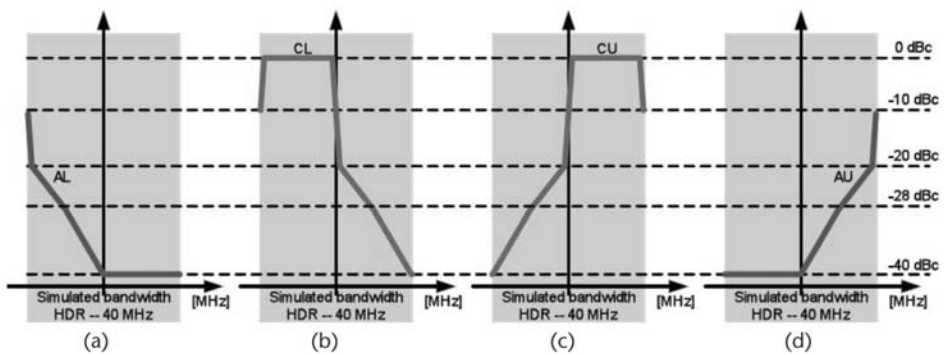


Figure 7.31 MAGNET HDR in-band spectral masks for the four major interferers, two cochannels (CL and CU) and two adjacent channels (AL and AU).

The four transmitting power masks in Figure 7.31 can be specified by the following relative frequencies (f) in MHz and power (p) in dBc breakpoints [plot(f,p) in Matlab to get the curves]:

- **AL**
 $f = [-20 \ -19 \ -10 \ 0 \ 20]$
 $p = [-10 \ -20 \ -28 \ -40 \ -40]$
- **CL**
 $f = [-20 \ -19 \ -1 \ 1 \ 10 \ 20]$
 $p = [-10 \ 0 \ 0 \ -20 \ -28 \ -40]$
- **CU**
 $f = [-20 \ -10 \ -1 \ 1 \ 19 \ 20]$
 $p = [-40 \ -28 \ -20 \ 0 \ 0 \ -10]$

- AU
 $f = [-20 \ 0 \ 10 \ 19 \ 20]$
 $p = [-40 \ -40 \ -28 \ -20 \ -10]$

A quick test of filter design to obtain filters that will simulate the four different interferer frequency locations can be found in [45]. Filter designs are performed both for the assumption of a factor 2 oversampling and the current situation of no oversampling (Nyquist rate simulations). The second case is much more difficult. The required FIR filters for the AL, CL, CU, and AU interferer locations can all be derived from a symmetric (real) FIR filter designed for a higher sampling frequency. The reason for using a higher sampling frequency than 40 MHz is that it substantially simplifies the design. If the sampling frequency is set to a value of at least 80 MHz, we can avoid unwanted aliasing effects when moving the main lobe of the interferers to their correct center frequencies (−30, −10, 10, and 30 MHz) relative to the HDR center frequency as shown in Figure 7.30. This is done by a simple frequency translation of the FIR filter coefficients. Symmetry properties due to the −30/+30 and −10/+10 MHz frequency translations of the same real filter also imply that we only need two basic sets of filter coefficients, for example, the ones at positive frequency shifts. The other two are simply the complex conjugates of the basic sets of filter coefficients. The corresponding frequency responses and the design targets are shown in Figure 7.32.

There is a dependence between different interferers. For example, when a business meeting ends, people change their behavior and their personal wireless devices change their behavior accordingly. For example, several may switch on their mobile phones at almost the same time, creating a lot of interference from several devices simultaneously. It may be of value to be able to simulate such behavior, so that the impact can be predicted.

The current proposal for modeling a single interferer can be used to create such a situation. The activity level model, discussed earlier, has an initialization in which the starting state of the FSM can be set to any desired value. If this state is set to the highest (or a high) activity level for the simulated interferers, they will all start to transmit with a high activity at the start of the simulation. If this correlated behavior is not desired, the initial state should, as indicated earlier, be randomized according to the steady-state distribution of the used FSM.

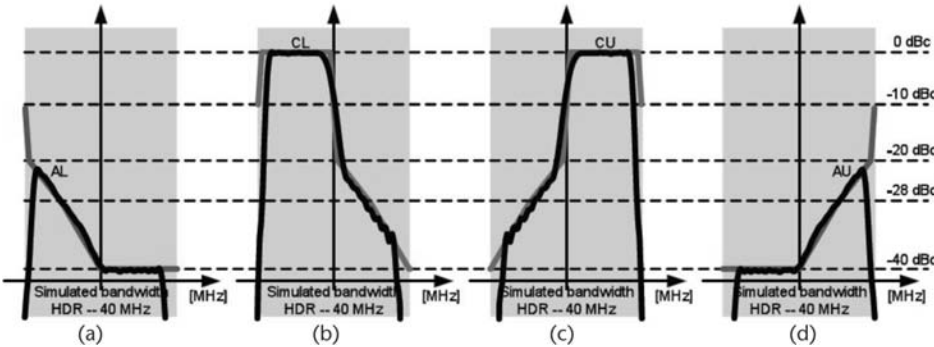


Figure 7.32 (a–d) Filter amplitude spectra when using filter coefficients.

The choice of specific stochastic models and probability distributions is driven by the need for a trade-off between accuracy and simplicity of definition and implementation. The application field of these interference models in developing and evaluating interference mitigation techniques was also an important design consideration in [45]. For example, the continuous-time Markov chain model for the activity levels and the Poisson process for burst starting times are very simple to implement and very well-studied stochastic processes. One could argue that network traffic is of a large-scale-dependent, heavy-tailed nature that cannot be accurately captured by these memoryless processes. If the interference mitigation technique to be adopted relies on using interfering burst (packet) interarrival periods to transmit useful traffic, a more accurate stochastic model of the packet interarrival times would be necessary for testing the algorithm. For example, for an 802.11a interferer, a distribution for the interarrival times that more closely matches the 802.11 back-off mechanism under heavy traffic load should be developed. However, mitigating interference by using interarrival times to transmit is not deemed a suitable solution because it can create unwanted interference from the studied system to the interfering system. A fair approach to avoid interference from non-cooperative sources, once it is detected, would be to use an unaffected part of the spectrum by reshaping the power loading of the MC-SS subcarriers. Such an interference mitigation policy does not require the exact statistical properties of the interarrival times of the interference bursts. What needs to be tested is the ability of the interference mitigation mechanism to accurately detect the start and end of an interference activity period, so that idle periods can be used to transmit at full capacity. Such a mechanism can be adequately tested against the simple stochastic model presented in this document.

For the same reasons, the specific values of the various parameters suggested here are only intended to be used for evaluating the interference mitigation algorithms under a set of reasonable scenarios. The actual values that these parameters can take in various network configurations (number of interfering transmitters and distance from the affected receiver) and transmitted traffic types and loads can vary in a much broader range. For example, the measured frame length CDF shown in Figure 7.29 is averaged over all traffic types. Particular traffic types (e.g., VoIP or remote log-in) exhibit much different packet length distributions.

7.3.3 Unified Interference Modeling Derived from Link Budget Analysis

Reference [45] proposed a macroscopic interference model intended to predict the impact of interferers on the radio coverage assessments of MAGNET devices operating in the same area.

The model considers interference involved by devices using the same transmission scheme (MC-SS air interface) and other WLAN technologies using identical RF bands at the same time.

Several interference models were introduced to validate the concept of the *unified interference model*. First, white Gaussian wideband interferers are encountered. They contribute to an additional white Gaussian noise affecting the received power-to-noise ratio of the device of interest as follows:

$$\text{SNR}' = \text{SNR} \cdot \frac{1}{1 + \alpha} \quad (7.13)$$

The received signal, before demodulation is expressed as follows:

$$y(t) = x(t) \cdot h(t, \tau) = \sum_{i=0}^{N_i} ai(t) \cdot x(t - \tau_i(t)) \quad (7.14)$$

$$h(t, \tau) = \sum_{i=0}^{N_i} ai(t) \cdot \delta(\tau - \tau_i(t))$$

Secondly, colored noise interferers are considered. The system is then affected with an additional colored noise that will be assimilated to a global time-frequency selective filter filtering a white Gaussian noise with a higher power density noise contribution or described by a higher noise temperature [noise figure (NF)]. This approach is derived from an equivalent baseband phase noise model and multipath channel impact on a radio signal. When the coherence bandwidth of the channel is smaller than the transmission bandwidth, the multipath channel is described by a filter. A similar approach for colored noise interference contribution was adopted in [45]. The combination of multipath channel and interferers transforms a single filter into a correlation of two filters affected with additional noise.

The received signal without interference, before demodulation, is expressed as follows:

$$y(t) = x(t) \cdot h(t, \tau) = \sum_{i=0}^{N_i} ai(t) \cdot x(t - \tau_i(t)) \quad (7.15)$$

$$h(t, \tau) = \sum_{i=0}^{N_i} ai(t) \cdot \delta(\tau - \tau_i(t))$$

When considering the colored interference contribution, $b(t) * g(t, \tau)$ corresponds to the interference filter, which is modeled as a selective filter. This filter modifies the received signal without interference in a similar way as the multipath propagation channel. Furthermore, an additional white Gaussian noise should be considered.

From this expression and physical considerations surrounding the filtering process for modeling narrowband interferers, it appears that the combination of the channel and colored interference should be assimilated to a modified multipath propagation channel.

The received signal, before demodulation, is expressed as follows:

$$y(t) = x(t) \cdot h_M(t, \tau) = \sum_{i=0}^{N_i} a'i(t) \cdot x(t - \tau_i(t)) \quad (7.16)$$

$$h_M(t, \tau) = h(t, \tau) \otimes g(t, \tau) + b(t) \otimes g(t, \tau)$$

In these two cases, link level degradations will be translated into an interference multipath-channel margin (IMCM), degrading link level performance. The interference model utilizes the multipath channel margin concept introduced in [48] to quantify the impact of interferers on BER performance.

Figure 7.33 shows a representation of the impact IMCM has on BER performance, which uses the same concept of the multipath channel margin in considering indifferently the modified multipath channel model or AWGN interferers.

It appears that interferers can be modeled as equivalent modified multipath channel models inducing relative BER degradations quantified by IMCM on link level performance. The estimation of the modified multipath channel margin results in a first macroscopic approach, as shown in Figure 7.34. This approach does not require to model $g(t, \tau)$. It can simply be deduced from an assessment of α and α' considering the different natures of interferers and propagation scenarios in link level simulations as detailed in [50].

7.3.4 Cooperative Transmission for WPAN Scenarios

A new approach to interference mitigation using cooperative transmission was proposed for the MAGNET scenarios. For all devices that cannot afford multiple antennas, such as WPAN devices, these cooperative transmission schemes are inter-

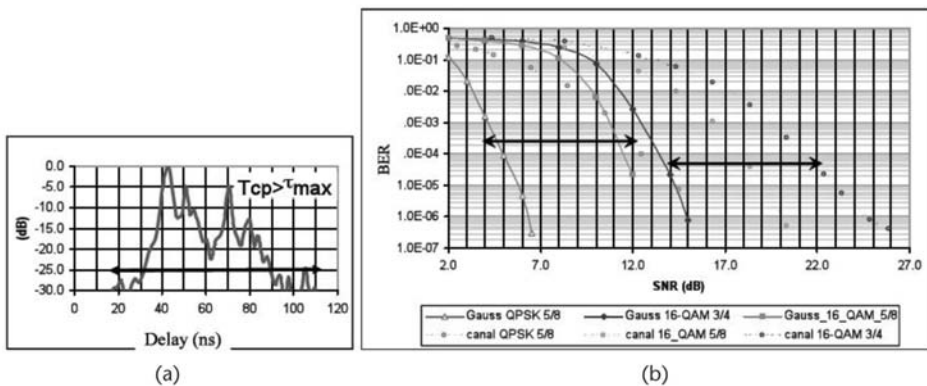


Figure 7.33 (a, b) IMCM/multipath channel margin representation [45].

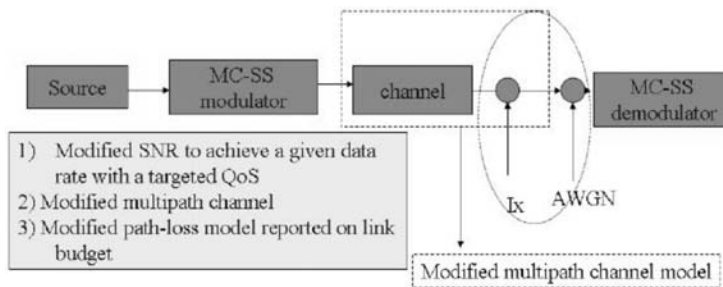


Figure 7.34 Unified interference model for PAN scenarios [45].

esting. Because MAGNET's HDR air interface is working in the unlicensed 5-GHz ISM band, the devices are exposed to cochannel interference (CCI) from other systems, such as WLANs, following the IEEE 802.11a standard. Receiving devices can react to outages due to CCI by using retransmission requests. However, frame retransmission may be unfavorable in view of latency and battery power limitations. Cooperative diversity techniques were proposed in [45] for wireless networks with coordinated peer-to-peer traffic in the presence of randomly arriving CCI. Unused time slots and idle devices with energy reserves are taken advantage of for reducing outage probabilities through amplify-and-forward procedures and linear signal combining at the receiver end. It was shown that the diversity orders associated with peer-to-peer transmissions can be assessed from knowledge about the slot allocation and the numbers of interfering signals per time slot. This knowledge can be obtained by idle devices through channel sensing.

The cooperative diversity scheme can significantly reduce the outage probabilities. Figures 7.35, 7.36, and 7.37 show the obtained outage probabilities versus the amount of interference ν_{interf} for superframes with $N_{\text{SF}} = 8$ time slots, comprising an average 2, 4, and 6 intra-WPAN frames, $\gamma = P_s/P_n$ equal to 20 dB, and γ_{min} equal to 10 dB. The outage probabilities are shown for employing relaying policies 1 and 2 as well as for no relaying.

Such an approach has the following benefits:

- Downsizing of the overhead in terms of control information exchange for handling retransmission requests;
- Reduction of the average decoding delay;
- Balancing of the energy reserves in the WPAN by letting noncritical devices carry out retransmissions.

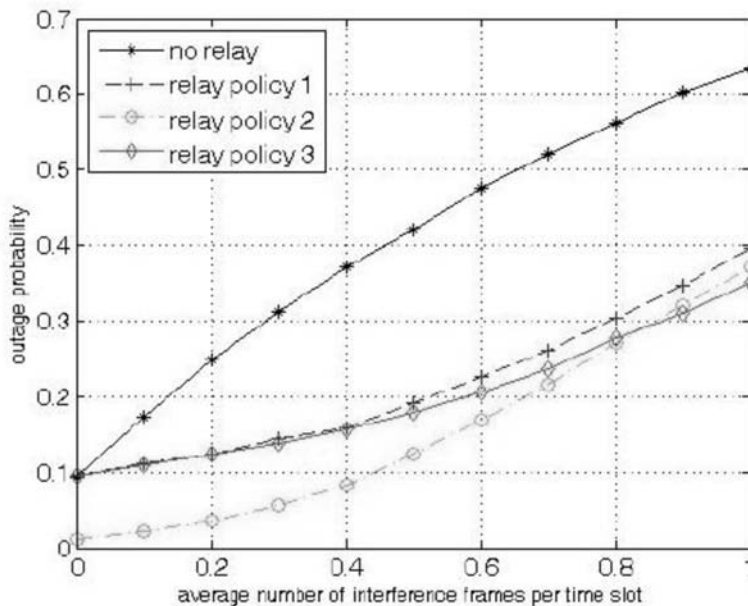


Figure 7.35 Outage probability versus ν_{interf} for superframes comprising $N_{\text{SF}} = 8$ time slots for an average $\nu_{\text{frames}} = 2$ intra-WPAN frames [45].

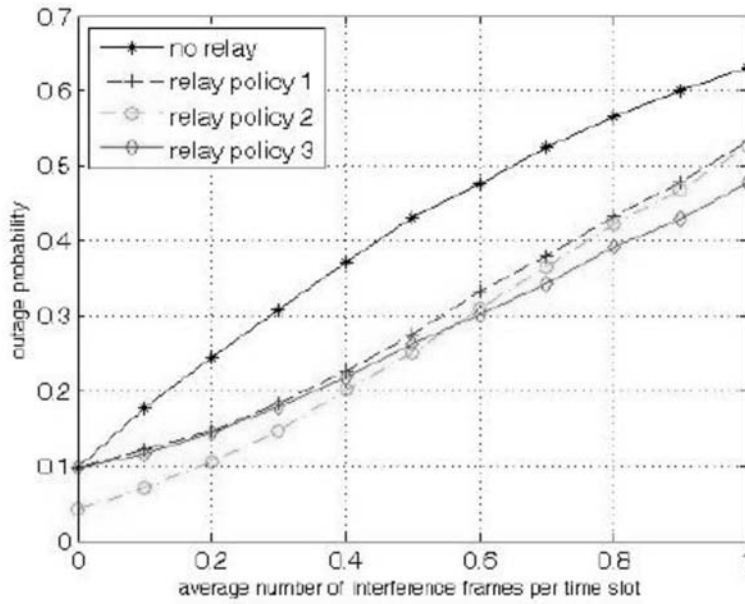


Figure 7.36 Outage probability versus ν_{interf} for superframes comprising NSF = 8 time slots for an average $\nu_{\text{frames}} = 4$ intra-WPAN frames [45].

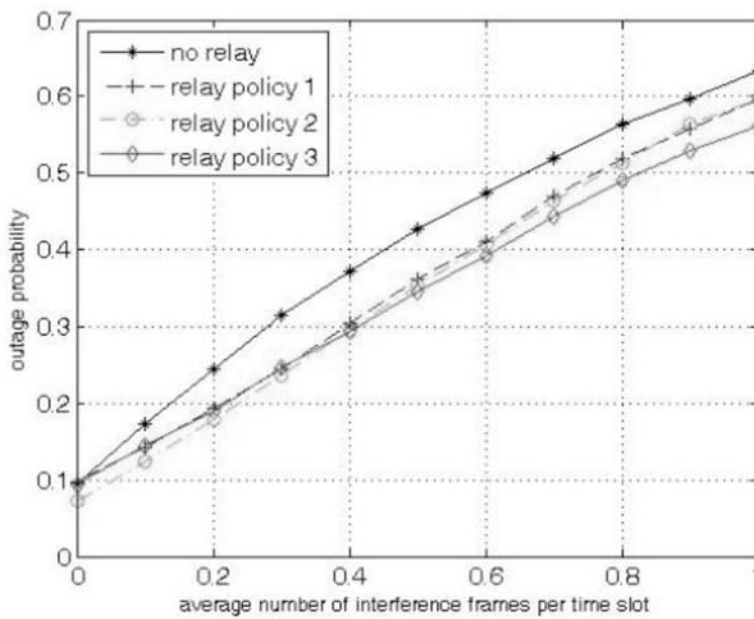


Figure 7.37 Outage probability versus ν_{interf} for superframes comprising NSF = 8 time slots for an average $\nu_{\text{frames}} = 6$ intra-WPAN frames [45].

Besides the derived minimum diversity order, other criteria may be suitable for distributed decision making and facilitate even greater outage reductions. Moreover, remaining battery energies or QoS criteria may be incorporated in the relaying policies. Such more advanced cooperative schemes are worth studying in the future.

7.3.5 Interference Mitigation Through Dynamic Interleaving Codes

Reference [9] proposed a simple *interference diversity technique* for the MAGNET Beyond system that transforms a colored noise distribution into a white Gaussian noise distribution using a dynamic interleaving process to modify the propagation channel impact on the demodulation/decoding process. This transformation significantly improves the link level performance and is less complex than the CDMA process usually used to lead the same target. The transformation consists of carrying out a time-variant interleaving process applied to the data symbols/spread data symbols allocated to the data subcarrier of an OFDM signal [45].

The interleaving design will vary whether interleaving is applied on spread data symbols or on data symbols. The required configurations are shown in Figures 7.38 and 7.39, respectively.

The interleaving considers the algorithm described in [51]. This interleaver allows the following:

- Interleaving spreading maximization on a multilevel scale;
- Interleaving partitioning preservation.

These two properties are efficiently implemented to keep Hadamard codes when their lengths are inferior to the number of data subcarriers per OFDM

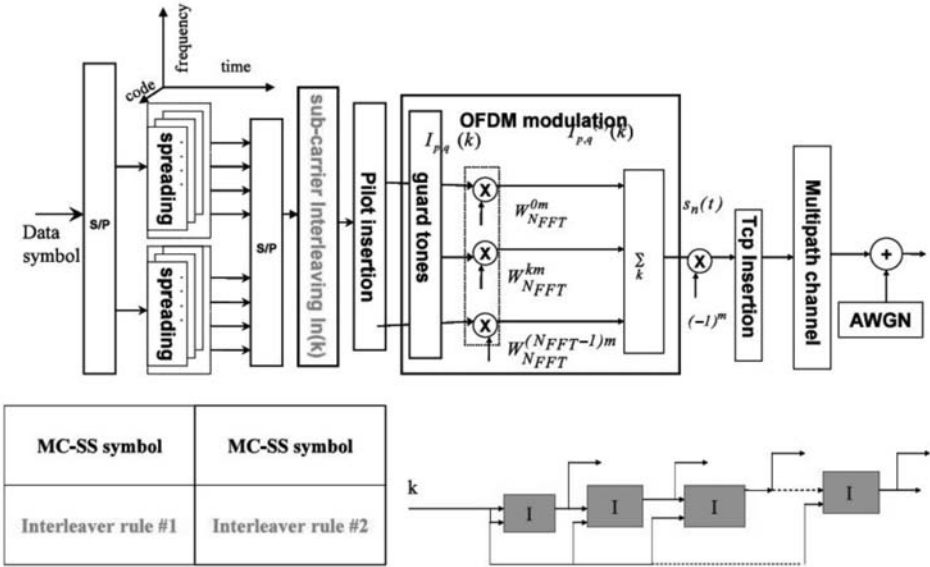


Figure 7.38 Dynamic interleaving process foreseen as an interference diversity technique when interleaving is applied on spread data symbols [45].

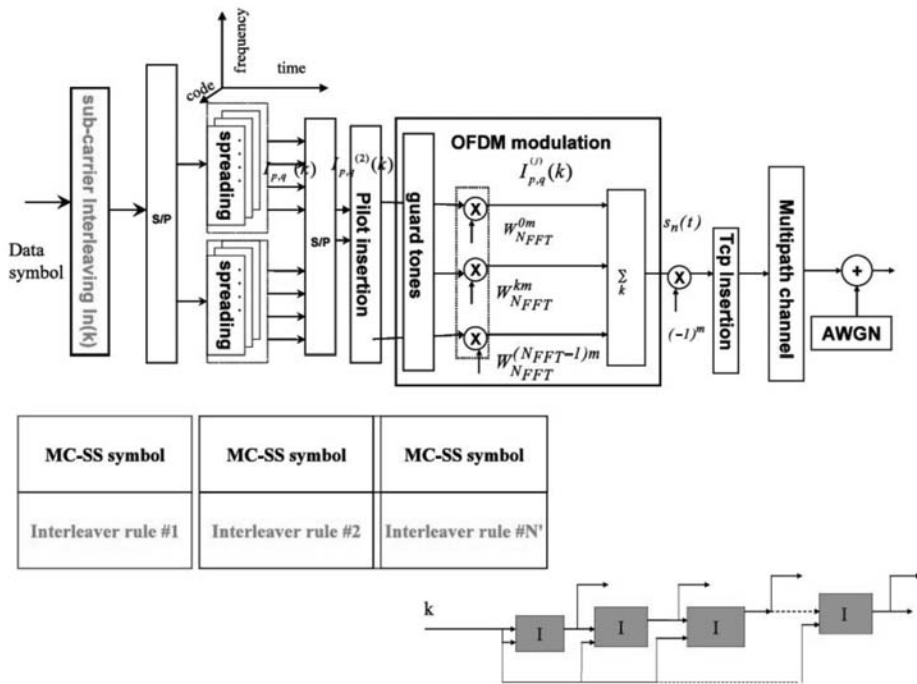


Figure 7.39 Dynamic interleaving process foreseen as an interference diversity technique when interleaving is applied on data symbols [45].

multiplex. This case was developed as a hardware platform [9]. The global principle for the proposed system is shown in Figure 7.40.

A time-variant interleaver is applied on data symbols before the spreading transform. This operation involves a modification of the transfer function of the propagation channel associated with each interleaving pattern rule. The modification of the transfer function of the channel is significant when the propagation channel is multipath and when narrowband interferers damage some subcarriers of the OFDM signal.

The impact of dynamic subcarrier mapping allocation on the OFDM demodulation is shown in Figure 7.41.

The link level performance and dynamic interleaving gain process are shown in Figures 7.42 and 7.43, respectively.

Configuration 0 corresponds to the case in which no interleaving is processed on spread symbols, *configuration 1* corresponds to the case in which a single interleaving pattern rule is applied in a static manner, *configuration 2* corresponds to the case in which two different rules are applied, and *configuration 3* corresponds to the case in which three different interleaving patterns are applied. The gains between the static and dynamic interleaving data symbol mapping allocation are set to 4 dB. When the number of permutation rules is increased, the gains are near 1 dB between two and three permutation rules. These gains allow the radiated power to be reduced and also ensure enhanced coexistence between MC-SS devices and other devices.

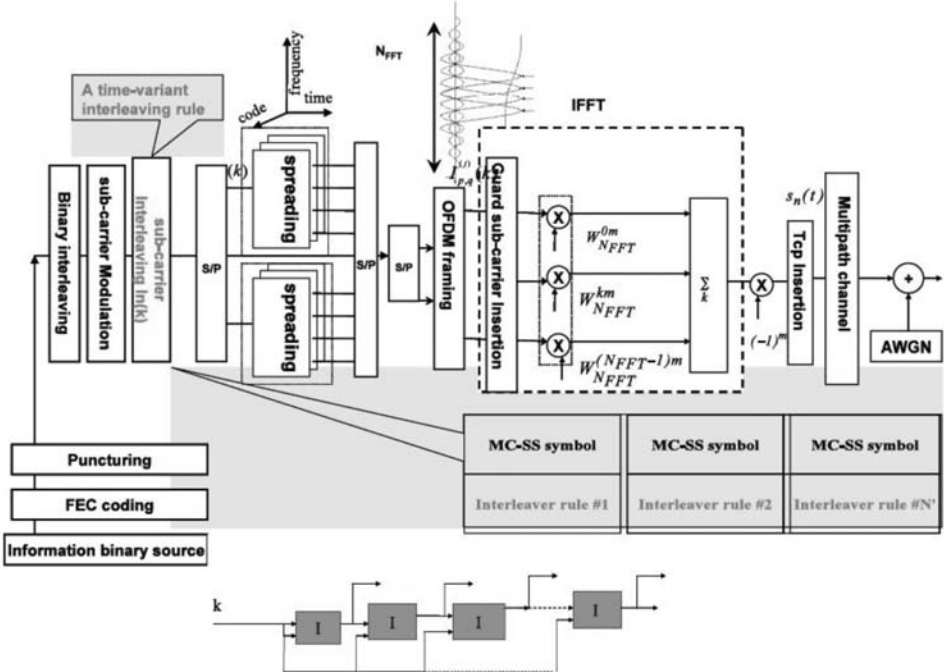


Figure 7.40 Dynamic data symbol allocation process for the MC-SS air interface [45].

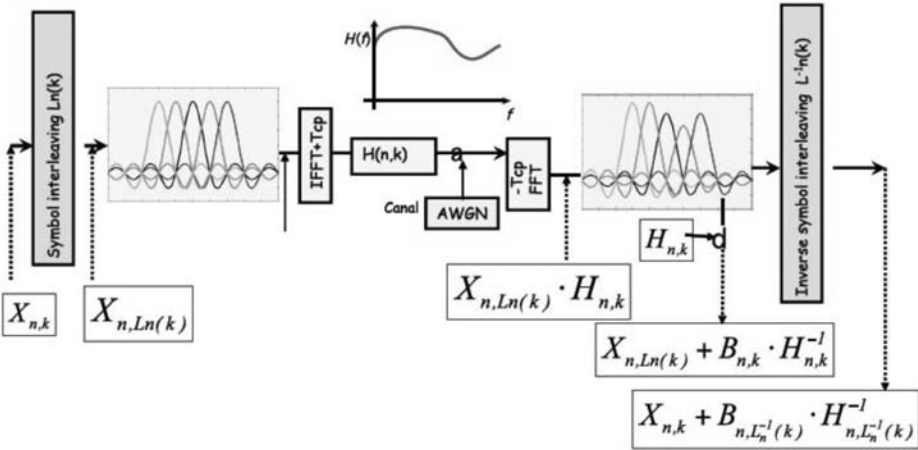


Figure 7.41 Impact of symbol/data subcarrier interleaving process [45].

7.4 Conclusions

Regardless of the type of radio interface, radio system performance can be significantly improved on through the use of proper interference management techniques. Intercell interference is an issue in a system intended to provide ubiquitous coverage and to host users with various mobility. Some of the methods that can mitigate the undesirable effects of interference include smart-antenna-based techniques (e.g.,

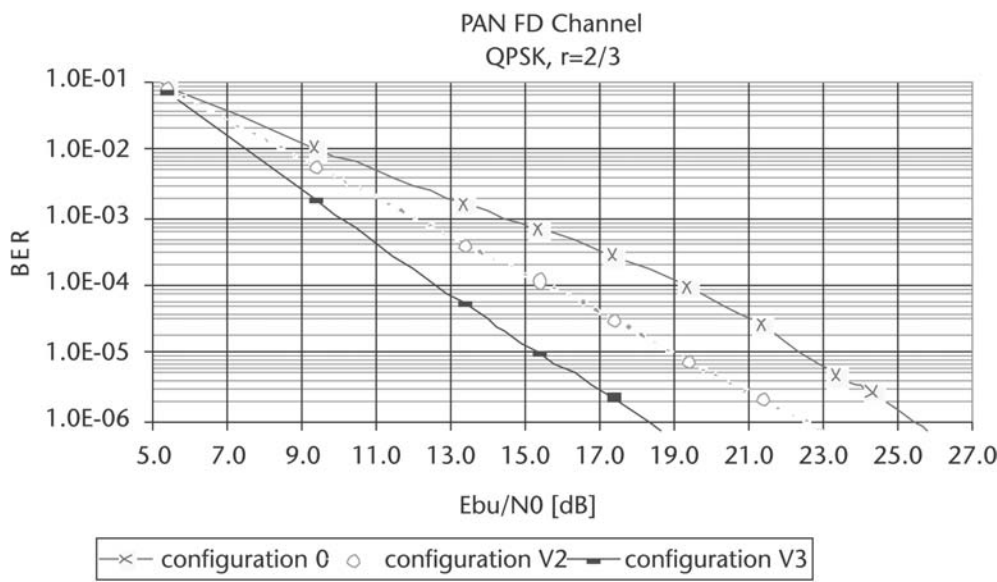


Figure 7.42 Link level performance [45].

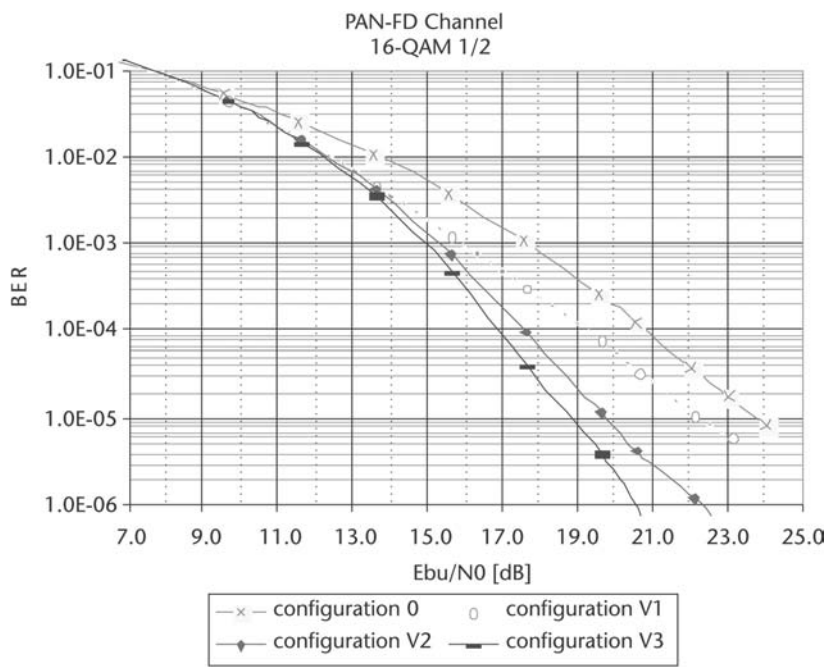


Figure 7.43 Dynamic interleaving gain [45].

beamforming, transmit and receive diversity) and interference avoidance and averaging. Combinations of such schemes can also be effective.

Use of multiple-antenna receivers at both base stations and user terminals can be efficient means to mitigate interference. Such usage allows for implementation

of receive diversity combining schemes such as maximum ratio combining and spatial interference suppression schemes such as interference rejection combining.

Interference cancellation allows for significant performance gains, especially within the cell border areas, at the price of a substantial increase in baseband complexity at the UT receiver, as well as the requirements on the time synchronization of the network and control signaling design. The use of frequency hopping provides the least effective reduction of interference compared to other techniques.

In a WPAN scenario, interference mitigation techniques based on dynamic implementation of an advanced turbo-based interleaving structure can whiten the colored noise contribution of the interferer on signals under a low-complexity interleaving processing. An interleaving process considering time-variant interleaving is less complex than a CDMA process usually implemented to transform the power noise distribution.

A stochastic model of the interference can be used to implement realistic interference generating modules at both the PHY and the MAC simulators for PAN systems. In an alternative approach to model the effects of interference on WPAN systems, a unified interference model based on link budget considerations can be beneficial. The principle consists of translating the interference level into an interference multipath channel margin that affects link level performance. This representation can be used in designing an advanced detection and avoidance (DAA) techniques to detect interferers and carry out interference mitigation

References

- [1] FP6 Overview, <http://cordis.europa.eu/ist/so/mobile-wireless/home.html>.
- [2] FP6 IST Projects Wireless Interface New Radio, WINNER and WINNER II, www.ist-winner.org.
- [3] International Telecommunication Union (ITU), www.itu.int.
- [4] FP6 IST project FIREWORKS, <http://fireworks.intranet.gr>.
- [5] FP6 IST Project OBAN, <http://cordis.europa.eu/ist/so/mobile-wireless/home.html>.
- [6] FP6 Project MEMBRANE, www.imperial.ac.uk/membrane.
- [7] FP6 IST Project SURFACE, <http://cordis.europa.eu/ist/so/mobile-wireless/home.html>.
- [8] FP6 IST Projects MAGNET and MAGNET Beyond, www.ist-magnet.org.
- [9] FP6 IST Projects PULSERS and PULSERS II, www.pulsers.eu.
- [10] FP6 IST Projects E2R and E2R II, <http://cordis.europa.eu/ist/so/mobile-wireless/home.html>.
- [11] FP6 IST Project RESOLUTION, <http://cordis.europa.eu/ist/so/mobile-wireless/home.html>.
- [12] FP6 IST Project WINNER II, "Smart Antenna Based Interference Mitigation," Deliverable 4.7.3, June 2007, www.ist-winner.org
- [13] FP6 IST Project WINNER II, "Final Assessment of Relaying Concepts for All CGs Scenarios under Consideration of Related WINNER L1 and L2 Protocol Functions," Deliverable 3.5.1, September 2007, www.ist-winner.org.
- [14] FP6 IST Project WINNER II, "Concept and Criteria for Coordination across Base Stations to Improve the Mutual Interference Situation," Deliverable 3.3, June 2005, www.ist-winner.org.
- [15] FP6 IST Project WINNER II, "Definition and Assessment of Relay-Based Cellular Deployment Concepts for Future Radio Scenarios Considering 1st Protocol Characteristics," Deliverable 3.4, June 2005, www.ist-winner.org.

- [16] FP6 IST Project WINNER II, "Proposal of the Best Suited Deployment Concepts for the Identified Scenarios and Related RAN Protocols," Deliverable 3.5, December 2005, www.ist-winner.org.
- [17] FP6 IST Project MAGNET, Deliverable 3.3, "Requirement Specification for MAC/RRM."
- [18] European Radiocommunications Committee, European Conference of Postal and Telecommunications Administrations, "The Analysis of the Coexistence of Two FWA Cells in the 24.5–26.5 GHz and 27.5–29.5 GHz Bands," ERC Report 99, October 2000, Edinburgh, Scotland.
- [19] FP6 IST Project WINNER, "WINNER Spectrum Aspects: Methods for Efficient Sharing, Flexible Spectrum Use, and Coexistence," Deliverable 6.1, October 2004, www.ist-winner.org.
- [20] "Fixed Radio Systems; Point-to-Multipoint Equipment; Time Division Multiple Access (TDMA); Point-to-Multipoint Digital Radio Systems in Frequency Bands in the Range 3 GHz to 11 GHz," ETSI EN 301 021, V1.5.1 (2002-02).
- [21] "Broadband Radio Access Networks (BRAN); HIPERLAN Type2; Requirements and Architectures for Interworking Between HIPERLAN/2 and 3rd Generation Cellular Systems," ETSI TR 101 957, V1.1.1 (2001-08).
- [22] European Telecommunication Standardization Institute, www.etsi.org/WebSite/homepage.aspx.
- [23] "Harmonised Radio Frequency Channel Arrangements and Block Allocations for Low, Medium and High Capacity Systems in the Band 3600 MHz to 4200 MHz," CEPT/ERC/REC 12-08 E (Podebrady 1997, Saariselkä 1998), May 14, 1998.
- [24] European Radiocommunications Committee, European Conference of Postal and Telecommunications Administrations, "The Analysis of the Coexistence of FWA Cells in the 3.4–3.8 GHz Band," ERC Report 33, May 2003.
- [25] Federal Communications Commission, www.fcc.gov.
- [26] Federal Communications Commission, "Spectrum Policy Task Force Report," ET Docket No. 02-135, November 2002.
- [27] Motorola, "Reply Comments of Motorola to SPTF Report," ET Docket No. 02-135, submitted to FCC January 27, 2003.
- [28] Shared Spectrum, "Reply Comments of Shared Spectrum to SPTF Report," ET Docket No. 02-135, submitted to FCC January 27, 2003.
- [29] V-Comm, "Reply Comments of V-Comm to SPTF Report," ET Docket No. 02-135, submitted to FCC February 28, 2003.
- [30] FP6 IST Project WINNER II, "Inter-Cell Interference Avoidance," Deliverable 4.7.2, June 2007, www.ist-winner.org.
- [31] FP6 IST Project WINNER II, "Inter-Cell Interference Avoidance," Deliverable 4.7.1, June 2007, www.ist-winner.org.
- [32] Fazel, K., and S. Kaiser, *Multi-Carrier and Spread Spectrum Systems*, New York: John Wiley and Sons, 2003.
- [33] Rasmussen, L. K., T. J. Lim, and A. L. Johansson, "A Matrix-Algebraic Approach to Successive Interference Cancellation in CDMA," *IEEE Trans. on Communications*, Vol. 48, No. 1, January 2000, pp. 145–151.
- [34] Divsalar, D., M. K. Simon, and D. Raphaeli, "Improved Parallel Interference Cancellation for CDMA," *IEEE Trans. on Communications*, Vol. 46, No. 2, February 1998, pp. 258–268.
- [35] Wang, X., and H. V. Poor, "Iterative (Turbo) Soft Interference Cancellation and Decoding for Coded CDMA," *IEEE Trans. on Communications*, Vol. 47, No. 7, July 1999, pp. 1046–1061.
- [36] FP6 IST project WINNER II, "The WINNER II Air Interface: Refined Spatial-Temporal Processing Solutions," Deliverable 3.4.1, November 2006, www.ist-winner.org.

- [37] Winters, J. H., "Optimum Combining in Digital Mobile Radio with Cochannel Interference," *IEEE J. on Selected Areas in Communications*, Vol. 2, No. 4, July 1984.
- [38] Craig, S., et al., "Channel Allocation Tiering (CHAT): Taking GSM/EDGE Networks Beyond One-Reuse," *Proc. of IEEE Vehicular Technology Conference (VTC Spring 2001)*, Rhodes, Greece, May 2001.
- [39] FP6 IST Project WINNER II, "WINNER II Test Scenarios and Calibration Cases Issue 1," Deliverable 6.13.1, June 2006, www.ist-winner.org.
- [40] FP6 IST Project WINNER II, "WINNER II Test Scenarios and Calibration Cases Issue 2," Deliverable 6.13.7, December 2006, www.ist-winner.org.
- [41] Brueninghaus, K., et al., "Link Performance Models for System Level Simulations of Broadband Radio Access Systems," *Proc. of PIMRC'05*, September 2005.
- [42] Zander, J., et al., *Radio Resource Management for Wireless Networks*, Norwood, MA: Artech House, 2001.
- [43] IEEE 802.20 Standard on Mobile Broadband Wireless Access, www.ieee802.org/20.
- [44] FP6 IST Project WINNER II, "WINNER System Requirements," Deliverable 6.11.1, May 2006, www.ist-winner.org.
- [45] FP6 IST Project MAGNET Beyond, "Advanced Interference Mitigation and Spectrum Efficiency Techniques," Deliverable 3.A1, June 2008, www.ist-magnet.org.
- [46] Nielsen, J. Ø., and J. B. Andersen, "Indoor MIMO Channel Measurement and Modeling," *Proc. of Wireless Personal Multimedia Communications (WPMC'05)*, September 2005, pp. 479–483.
- [47] FP6 IST Project MAGNET, "PAN Channel Characterization (Part I)," Deliverable 3.1.2a, October 2004, www.ist-magnet.org.
- [48] FP6 IST Project MAGNET, "PAN Channel Characterization (Part II)," Deliverable 3.1.2b, June 2005, www.ist-magnet.org.
- [49] Höher, P., "A Statistical Discrete-Time Model for the WSSUS Multipath Channel," *IEEE Trans. on Vehicular Technology*, Vol. 41, No. 4, April 1992, pp. 461–468.
- [50] FP6 IST Project MAGNET, "Prototype Specification of Antenna and Radio Front-End Schemes for PAN Devices," Deliverable 3.1.1, June 2006, www.ist-magnet.org.
- [51] FP6 IST Project MAGNET, "Prototype Specification for the FM-UWB and MC-SS RA Schemes," Deliverable 3.2.1, June 2006, www.ist-magnet.org.

Efficient Management of Radio Resources

A major challenge for the future generation mobile communication systems is how to provide services to various types of technologies, terminals, and applications. The fundamental goal is to ensure seamless mobility for the user. This, in turn, translates into the need for strategies that can ensure that resources are always available for both the user and the network.

Efficient management of radio resources is needed in order to provide the required QoS dictated by the large set of required applications that are foreseen [1–4]. The strict layering approach of the radio system architectures presents significant limitations to achieving this goal, especially in a wireless environment.

Radio networks are characterized by a dynamic topology; the link exhibits a time-varying quality due to fading and shadowing, in addition to multiuser interference (see Chapter 7). To cope with this variability of the wireless channel, various techniques have been developed either at the PHY or MAC layer.

This chapter describes novel techniques, such as scheduling and sophisticated resource allocation algorithms, that have been proposed and developed in support of next generation radio system concepts and QoS provisioning. Different protocols and cross-layer strategies are also presented. All contributions are based on work performed by the FP6 IST projects [5] that were active in the area of mobile and wireless Beyond 3G and broadband technologies.

The FP6 IST projects WINNER and WINNER II [6] investigated, proposed, and assessed radio protocols and RRM strategies in support of the ubiquitous radio system concept designed for next generation systems [7] and also suitable for support of multihop communications. The FP6 IST project SURFACE [8] developed suitable resource allocation strategies in support of a generalized radio interface, which should be reconfigurable depending on radio channel conditions and QoS requirements. The FP6 IST project 4MORE focused on cross-layer optimizations for an MC-CDMA-based system, whereas the FP6 IST project MEMBRANE investigated optimized performance in a mesh network scenario [9–11]. The FP6 IST projects MAGNET and MAGNET Beyond [9] focused on radio interface optimizations for WPAN environments.

Section 8.1 introduces the main issues related to resource management. Some challenges related to the medium access layer design are also presented. Section 8.2 introduces the various resource management strategies most commonly applied for resource allocation at lower layers. Section 8.3 describes an approach to cross-layer design and implementation in support of mobile communications for a next

generation UT based on MC-CDMA. Section 8.4 describes a concept proposed to ensure access to resources for users of a multimode terminal in a WPAN environment. Section 8.5 concludes this chapter.

8.1 Introduction

Current radio communication systems are built on a layering approach that has proved to be very successful and allowed the development of complex communication networks [12]. Each layer adds value to the services provided by the set of lower layers in such a way that the highest layer offers the set of services needed to run distributed applications. The independence of each layer remains ensured, and this permits changes within a layer without affecting the other layers, provided they still offer the same service to the next higher layer.

One of the most important problems in networking is the allocation of limited resources among the users of the network. In a traditional layered network architecture, the resources to be allocated at the MAC and at the network layers are based on the use of communication links, which are considered to be “bit pipes” that deliver data at a fixed rate with a certain error probability [13]. This bit pipe is a simple abstraction of the underlying PHY and data link layers. In a way, this abstraction separates the treatments on the MAC layer and the PHY layer, generating two different areas of research, which the authors of [13] refer to as the networking and communication groups. Research in the networking field has focused on allocating these bit pipes among different flows of randomly arriving traffic using approaches such as packet scheduling and collision resolution. The main aim in networking is to use the bit pipes in the most efficient way possible, while providing acceptable QoS in terms of delay and throughput to each user.

On the other hand, the communication problem has focused on creating better bit pipes, that is, on improving the transmission rate or spectral efficiency for a given channel by methods such as advanced detection, modulation, and coding. (These last were discussed in previous chapters of this book.) The random arrivals and departures of traffic are usually ignored and delay is not considered.

For enhanced radio system performance with respect to the PHY layer, the adaptation to the channel characteristics can resort to adaptive transmission techniques that go well beyond the conventional adaptation of transmitted power, whereas for the MAC layer schemes such as appropriate scheduling or sophisticated resource allocation algorithms that try to minimize the number of channels needed to accommodate users are used.

8.1.1 User and Control Planes

In a modern radio communication system, the functionalities are split into user and control plane functionalities [3]. These two planes, however, have different structures. Whereas the user plane tries to hide its complexity by use of some form of layering, it also has a structure that supports interoperability, extensibility, and scalability. The control plane, on the other hand, is designed to cut across that layering, giving visibility and access to all aspects of the network, which must be

monitored and measured. This gives it poor scalability. However, to be able to deliver to a user the minimum amount of QoS for a given application and fulfill its security requirements, the information coming from both planes must be correlated in order to make the correct decision.

In a future radio system (e.g., the WINNER radio system [6]), the control plane functions implement system-wide control of the total radio resources and the user plane functions ensure end-to-end control of the user data.

The behavior of the control plane and user plane functions are defined by the radio resource management algorithms (which could be centralized, decentralized or partly decentralized [14]), and the RRM decisions are distributed to the different nodes and the user terminals by the radio resource control protocol. The RRC protocol is a network protocol and as such should be supported by all nodes implementing the RRM functions in this network. In an ubiquitous radio system concept in which parameterizations (i.e., modes) of the same PHY layer transmission technology are expected to serve different deployment scenarios [6], certain commands are not applicable for all nodes supporting only a specific mode. Common configuration commands could be the scheduling policy between and inside the modes (e.g., fairness criteria between services and users), and specific commands could be spectrum allocation and interference control. Thus, the RRC protocol needs to consist of a generic part RRC-g that is supported by all nodes implementing the RRM functions in the radio access system and a radio mode specific part, denoted RRC-r, for each system mode supported by the nodes in that particular mode [14]. RRC configures the radio link layer (RLL), but it also uses the RLL to transport configuration commands to remote nodes.

The RLL defines services to the upper layer both in the control and in the user plane, and consists of a MAC sublayer and a radio link control (RLC) sublayer. A convergence protocol sublayer may also be included in the user plane if it is needed to support a certain user data flow [14]. One goal of next generation radio system design is to support a wide range of deployment techniques, but at the same time keep the complexity low. For this reason, the standard link control functions positioned in the RLC sublayer should ideally be common to all transmission modes (i.e., RLC-g), such that there is then no need for a specific RLC for the different modes. This type of architecture was proposed initially for the WINNER radio system and is shown in Figure 8.1.

In this concept mode-specific HARQ schemes are seen as a part of the mode-specific MAC sublayer (MAC-r).

The RRM functions configure and reconfigure the generic link layer (GLL) based on radio resource availability, user terminal capability, and admission control. A configuration of the GLL may include room for the GLL to make its own decisions of mapping flows onto the radio modes based on, for example, information on instantaneous mode capacity and capability received directly from the MAC-r entities in a radio access system node or in a UT (originating in the PHY). Hence, the MAC-g sublayer performs intermode scheduling between the modes available for each user data flow based on a generic scheduling metric common to all modes. In summary, the PHY layer defines a set of resources (in time, frequency, and so forth) that may be allocated, the MAC layer controls the usage

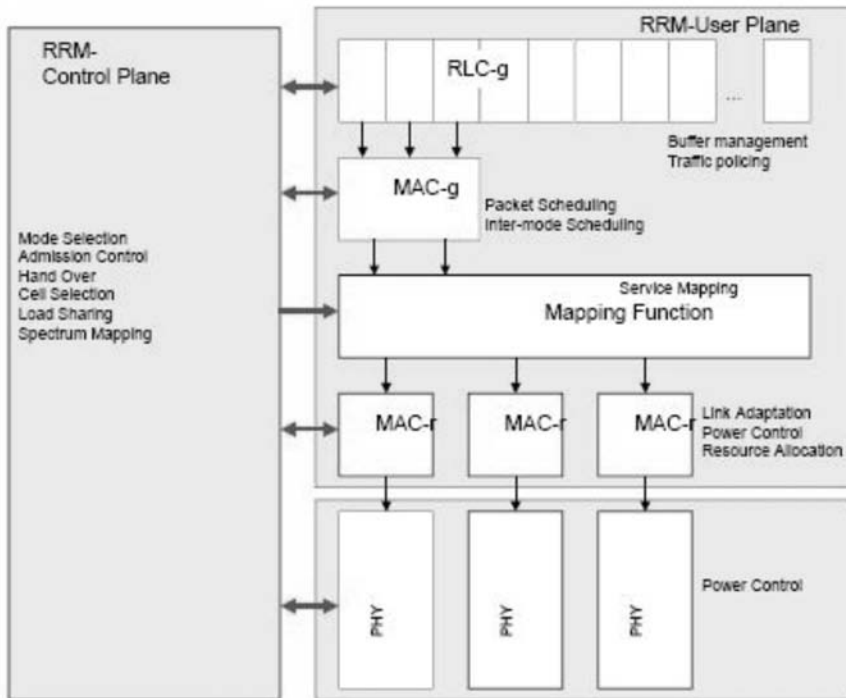


Figure 8.2 Mapping of control plane functionality onto protocols [14].

transmitted individually and according to their specific QoS requirements. For example, a real-time (RT) service and an FTP service may be mapped to different flows, thus allowing a scheduler (e.g., service level controller) to privilege the RT service and to stay within the delay constraints of that service.

The context of a flow is established and released by the *flow establishment/release* function. If a flow should be moved from one base station to another, the context of the flow has to be transferred to the new BS. This might happen in the case of radio handover due to changed radio link conditions or if multiple radio links are available to the UT (e.g., overlay cells) due to shifting of the load between the cells. The transfer could happen within one domain (micromobility) or between two domains (macromobility).

8.1.2 MAC Design Challenges

MAC protocols are based on multiple-access technologies [17], which enable access to the spectrum of resources for multiple users/stations. The underlying resource *radio channel* comprises the dimensions of space, time, code, and frequency.

The MAC has a decisive role in QoS considerations. Radio channel resources have to be assigned to all interested parties in a proper way in order to meet the QoS requirements of users. Therefore, the MAC protocol needs to possess information about the status of the channel in the sense of knowing what resources are available and the requirements of the parties interested in using the resources.

The manner in which the MAC protocol controls access to the medium significantly influences the level of QoS the whole system is able to offer [14].

A major challenge posed in the design of the MAC protocol comes from the design of the PHY layer and, especially, the size of communication that is available. Traditionally, MAC and link control protocols have been designed to minimize the number of bits in each message and have relied on techniques such as handshakes to maintain the consistent state information at both ends of the link. As a more efficient paradigm for protocol design, researchers proposed [14–17] to minimize the number of messages transported back and forth and convey more information in each of them. This is a good approach for a transmission quantum that is large and a significant PHY overhead. Self-correcting protocols that do not suffer from catastrophic failure in inconsistent states also will most likely have to be employed. This would potentially allow for eliminating some of the multitiered handshakes that are often present in communication protocols.

From the point of view of the MAC protocol, a channel arrangement in which a narrower band channel is available in addition to a broadband one could be beneficial because the granularity on the narrowband is finer and communication on that channel is less power consuming, which allows for more frequent terminal status update signaling [14].

Multicarrier modulation schemes also offer a potential solution to overcome the coarse granularity on the PHY layer if multiple transmissions by several stations using a nonoverlapping subset of the carriers are allowed. Theoretically, the capacity in this case could be allocated with the granularity of a single subcarrier, but in real life practical aspects prohibit this because of increases in the signaling overhead of the constellation used. The feasibility of such a scheme has yet to be demonstrated.

For a highly dynamic air interface, the capacity of the link will fluctuate constantly and fast link adaptation will be required. The traffic in a next generation radio system is also going to be dynamic, and flows are likely to be of short duration. This means that the concept of *radio bearers*, in the sense of second and third generation (2G and 3G) mobile systems [3, 4], together with the associated signaling setup, will not be sufficient to support the new radio systems.

UT power consumption is another key design driver when considering the radio protocols of a future radio system, and this must be reflected in all aspects of the MAC protocol design [14]. Because power consumption is determined at the time at which either the receiver or transmitter is turned on, unnecessary reception should be kept to a minimum. This requirement is not unique to a given system but is present in all battery-powered systems.

The multiple schemes for signaling of the resources needs will most likely have to be combined in the design of a MAC protocol. The following mechanisms were identified as candidates to be employed in the eventual protocol design of the WINNER system:

- *Polling* of individual terminals can provide a deterministic channel access delay that is foreseen to be needed by some services. It is, however, relatively resource consuming. Therefore, it cannot be the sole mechanism for resource allocation. If the uplink transmission quantum is large and the resource is

needed only infrequently, the overhead imposed on the system can become prohibitive.

- *Random reservation schemes*, that is, schemes where resource requests are sent over a random access channel, are suitable for handling burst traffic and can—with an appropriate back-off algorithm and parameter choices—offer good link utilization. Coarse uplink granularity will increase the overhead of the random reservation protocol and this aspect should be considered in the design of the physical layer.
- *Multicast polling* or polling of a group of terminals, as defined in the IEEE 802.16 MAC protocol [18], can be considered to be something of a hybrid between a random reservation scheme and polling. In a sense it is a random reservation scheme made available for a subset of the terminal or a subset of the traffic classes.
- *Piggybacking* is a key technique to mitigate the problems at the MAC layer that arise from the coarse granularity of multicarrier systems with a large number of subcarriers. In particular, piggybacking of requests for additional resources and of control signaling on data transmissions has the potential to cut radio link overhead significantly. Piggybacking works especially well when a given terminal is in a state of high communication activity [14].

The overhead introduced by the MAC protocol should be further reduced by the realization of an efficient broadcast/control phase. An issue that might be faced in the design of the MAC protocols for a multihop communication system is the need to support communication between the BS and a UT via a number of intermediate relay nodes. The design of the MAC protocol is one of the crucial elements determining the performance of such an arrangement.

Hidden and *exposed* stations must be taken into account in the design in order to achieve high utilization of resources and keep the interference at reasonable levels.

To exploit the potentials of advanced antenna concepts, the MAC protocol has to support structures for the description of resources in the spatial dimension.

The design of a MAC protocol is a trade-off between throughput and delay. It is often considered beneficial if the protocol can achieve high throughput at the expense of somewhat longer delays. The usual argument is that packet traffic is generally delay insensitive. In other words, the impact of message delays on application performance is traditionally neglected. The behavior of higher layer protocols, such as TCP, is, however, negatively influenced by an increased end-to-end delay. This means that there is a dependency between the air interface delay and the throughput that has not been traditionally taken into account in the development and analysis of MAC protocols. This is particularly true for next generation radio access systems, where the radio access architecture is simplified and buffer control has been removed [1, 2, 6]. The delay becomes even more of an issue if relays are present, that is, if the data suffers multiple instances of the air interface delay before reaching the core network.

Figure 8.3 shows the basic structure of a MAC frame where it is assumed that the medium access is controlled by a master terminal. This master terminal should

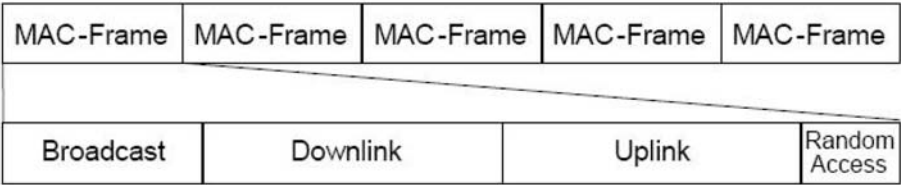


Figure 8.3 Basic MAC frame structure [14].

be able to transmit user data to associated slave terminals, which for their part should be able to transmit user data to the master terminal.

The logical relation of such a data exchange is called a *connection*. In the simplest case one connection exists between a sender and a receiver. It could be possible to establish multiple connections. This provides QoS support because it allows for different treatment of the different connections (e.g., in terms of required error rate). A user data transfer between slave terminals would not constrict the concept. This definition of tasks could easily be mapped onto cellular communication scenarios, where the BS acts as a master and the UTs can be seen as slave terminals. In an ad hoc peer-to-peer communication between mobile terminals, one of the terminals would be able to take over the master role, while the others act as slaves.

A frame is composed of a broadcast phase, a DL phase, an UL phase, and a phase for random access. During the broadcast phase the controlling master terminal sends out at least a beacon and a table of contents (TOC) for the UL and DL phases that follow.

Inside the beacon, information about the controlling terminal, the length of the TOC, and other information irrelevant in this scope are transferred. The TOC is used to describe all transmissions occurring in the DL and UL phases of the current frame. For each transmission the orientation (DL/UL), the starting point, the length, and the receiver are specified.

During the DL and UL phases user data and some additional control information are sent from the master terminal to the slave terminals and vice versa. A specific part of the UL phase is reserved for *contention-based random access*. The random access phase is used for association of the slave terminals to the master terminal, for requesting resources for an established connection, and for setting up a new connection.

Because every connection that is scheduled in a frame has to be described in the TOC, the overhead necessary to describe a frame usually grows with the number of connections scheduled in a frame.

One way to mitigate excessive overhead would be to schedule as few connections as possible in one frame. This scheduling method is known as exhaustive round-robin (EXRR). If there is enough data in the queue of a connection, one connection could take up the whole frame. Applying EXRR has a severe drawback [14]. Depending on the number of connections, and with traffic waiting in queues, the delay, which can be achieved with this scheduling algorithm, can become prohibitively high. This is shown in Figure 8.4.

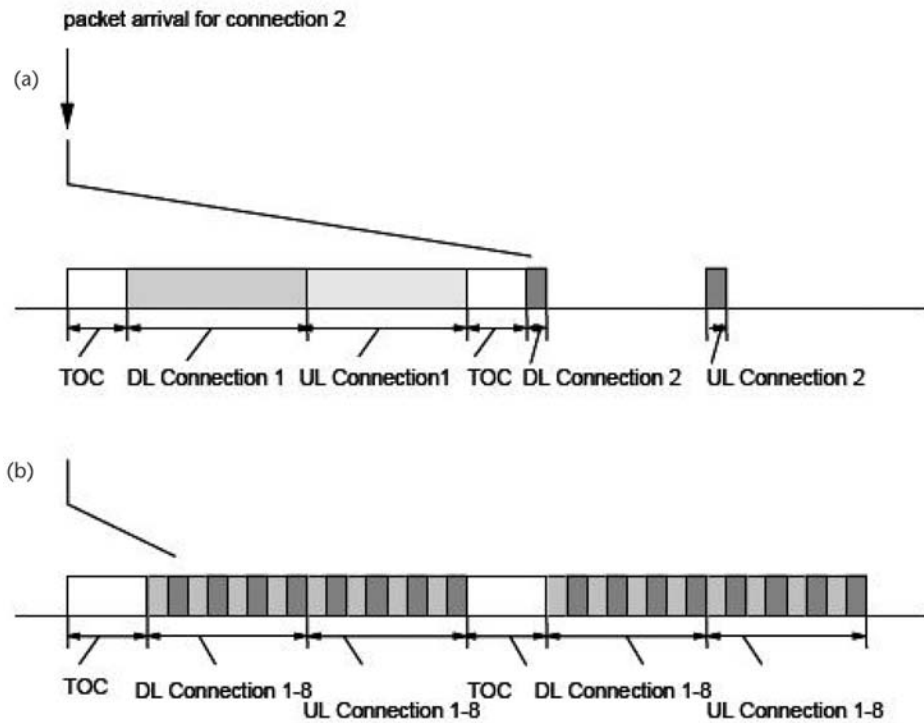


Figure 8.4 Delay constraints of (a) EXRR and (b) round-robin [14].

To prevent unacceptable delays, it may be necessary to allow numerous connections to be scheduled in a frame. This has an advantageous impact on the achievable delay. Depending on the packet size used by the upper layers, the delay may be reduced severely because the response to a packet may arrive as soon as the next frame as opposed to a delay of several frames using EXRR.

The concept of frame descriptor tables (FDTs) allows for reducing the delays in combination with a reduced overhead. The FDT has the same basic function as the TOC. It describes the contents of the UL and DL phase of a frame. Each FDT transmitted has an ID with which it can be identified in the following frames, for example, within the beacon. Therefore, it is easily possible to alternate between two or more FDTs if, for example, there is a certain periodicity in the needs of some connections. The main advantage of this concept is the resulting decrease in overhead. Changes in the layout of a frame can be coded by simply transmitting a number in the beacon.

A classification of connections would help determining where to schedule a new connection. If it is known a priori that a given connection is only needed very infrequently, there should be a means to communicate this to the scheduler.

A further step would be a hierarchical subdivision of the UL and DL parts in the sections of the same PHY mode/burst mode. One FDT could describe the changes of PHY modes in the frame and another could describe the transmissions in each PHY mode. The concept of subdividing frames into subunits would be especially useful if assuming a grouping of transmissions of the same burst mode.

This concept of grouping is used, for example, in IEEE 802.16a [19]. One FDT is used to describe the changes in the burst mode during the frame. Other FDTs are used to describe the contents of each block of the same burst mode. In this way, connections using the same burst modes can be changed without interfering with the scheduling of connections in other burst modes. This can be of importance if different classes of connections use different burst modes. For example, data connections using high bit rates are more likely to have high fluctuations of bandwidth needs (therefore necessitating frequent changes in frame layout) than connections carrying voice traffic.

The concept of FDTs allows for a reduction of overhead by decreasing the amount of control signaling. Multihop solutions proposed for frame-based MAC protocols [20, 21] show that this reduction can be even more important. A drawback of the multihop MAC protocol is the control signaling that is needed for each hop. This results in an increasing overhead with an increasing number of hops. With the help of FDTs, however, this overhead can be kept small. For a multihop solution, which establishes fixed or even partly fixed connections for the relaying of data, implementing the concept of FDT in such MAC protocols is getting even more interesting.

8.2 Resource Management Algorithms

8.2.1 Scheduling Algorithms

In time-domain scheduling, in which a DL transmission slot is exclusively allocated to packet transmission of a single user, system simulations using a detailed PHY layer model and realistic traffic models, show that important multiuser diversity gains are obtained if information about the channel quality is included in the scheduling algorithms [14].

In a realistic cellular system, to be efficient and to achieve fairness among the users, the algorithms have to cope with the fact that the channel qualities are unequally distributed among users. For example, a score-based scheduler achieves considerable gains in bursty traffic, where the number of users competing for the radio resource changes frequently [14].

Table 8.1 shows four scheduling algorithms, which can be classified with respect to the achievable multiuser diversity gain and their fairness into three categories.

The algorithms were classified for a downlink system employing an OFDM-TDMA physical layer. The considered system context is very close to the specifications employed by 3GPP [2] within the OFDM study item for HSDPA [22].

Table 8.1 Basic Classification of Scheduling Algorithms [14]

<i>Scheduler</i>	<i>Acronym</i>	<i>Fairness</i>	<i>Multiuser Diversity Gain</i>	<i>Category</i>
Round-robin	RR	Yes	No	0
Score based	SB	Yes	Yes	1
Proportional fair	PFS			
Maximum throughput	Max Thr.	No	Yes	2

Round-robin scheduling is a classical algorithm that schedules the UTs sequentially, irrespective of the reported CQI. The following exceptions apply:

- A UT having no packets waiting in one of its HARQ processes is skipped.
- If the expected BLER for the transmission toward the scheduled mobile is lower than a given threshold, the UT is only virtually scheduled, that is, no data is transmitted over the air.

This algorithm cannot achieve multiuser diversity but is fair in the sense that all UTs are treated equally regarding the total allocated timeshare of the resources.

Maximum throughput scheduling is designed to select the UT that has the highest achievable throughput. Fairness is not taken into account by this algorithm. At each transit time interval (TTI) n ,

- The link adaptation determines the best MCS for each mobile based on CQI.
- The mobile with the best potentially achievable data rate $DRR_i(n)$ is scheduled.
- The exceptions are handled as for round-robin scheduling.

Here, $DRR_i(n)$ is the data rate potentially achievable for UT_{*i*} at time n with respect to the selected modulation and coding scheme (see Chapter 3). The number of bits in the waiting queue or in the HARQ processes buffers of UT_{*i*} are not taken into account here.

The proportional fair scheduler (PFS) [23, 24] combines the principles of multiuser diversity and fairness by using an adequate metric. The underlying assumption for a fair sharing of resources among the UTs is that the CQIs are identically distributed. At each TTI n , the following steps are performed:

- The scheduling metric $m_i[n]$ is computed for each mobile i .
- The mobile i with packets waiting and with the highest metric $m_i[n]$ is scheduled.
- The exceptions are handled as for round-robin scheduling.

The metric value $m_i[n]$ is computed as follows:

$$m_i[n] = \frac{DRR_i[n]}{T_i[n]} \quad (8.1)$$

where n is the time index (number of TTIs), $DRR_i(n)$ is the data rate potentially achievable for mobile i at time n depending on the selected MCS, and $T_i[n]$ is the average throughput served to mobile i up to time n . The value of $T_i[n]$ is computed as follows:

$$T_i[n] = \lambda T_i[n-1] + (1-\lambda)N_i[n-1] \quad (8.2)$$

where $N_i[n-1]$ is the number of information bits UT has received correctly within TTI $n-1$ and

$$\lambda = 1 - \frac{T_{TTI}}{T_{PFS}} \quad (8.3)$$

where T_{PFS} is the size of the averaging window, and T_{TTI} is the length of the TTI.

The score-based (SB) scheduler was introduced in [25]. This scheduler combines the current CQI and the principle of fairness. The major advantage of this algorithm is that fairness between users is achieved, even if the respective CQI distributions differ, which is generally the case in realistic systems.

At each TTI n the algorithm performs the following steps:

- Sorts the achievable data rates $DRR_i[n], \dots, DRR_i[n - T_{SB}/TTI]$ obtained for a given mobile i within an observation window of length T_{SB} (multiple of the TTI) in descending order.
- Determines the mobile k for which the current data rate $DRR_k[n]$ achieves the highest rank within its own value set.
- Schedules the mobile k .

The exceptions are handled as for the round-robin algorithm.

Figure 8.5 shows the role of the scheduler in the MAC layer. For each TTI, the scheduling algorithm selects an active (with data waiting) HARQ process of the scheduled user. Based on a CQI obtained from feedback of the UT, the scheduling

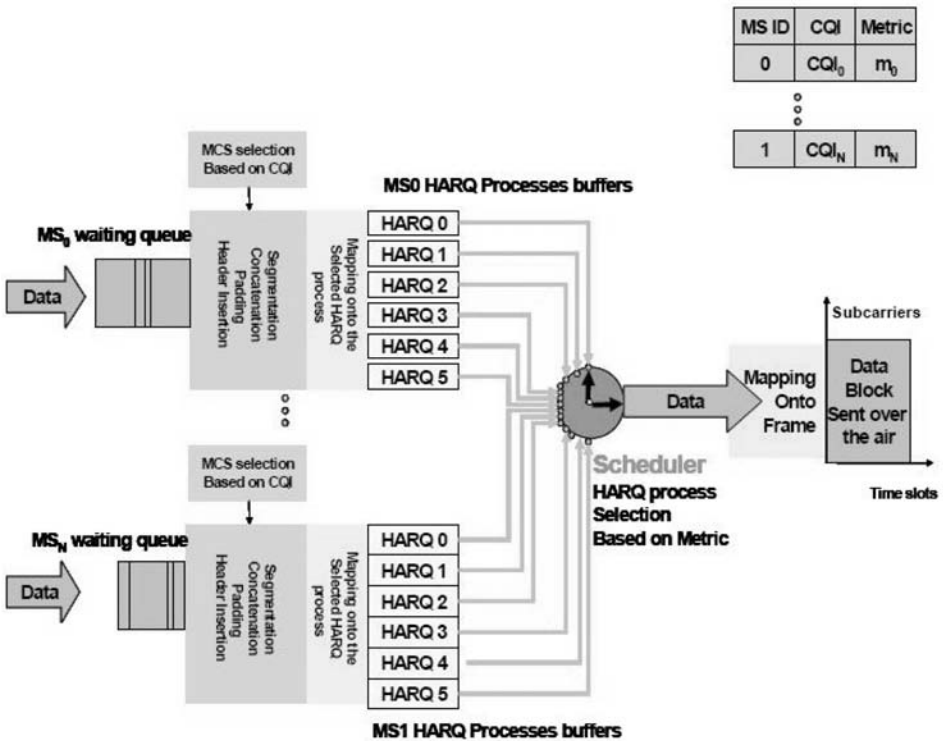


Figure 8.5 Role of the scheduler in the MAC layer [14].

algorithms calculate a metric m_i for each user i and schedule the user with the highest metric.

The layout of the scheduling/MAC interactions shown in Figure 8.5, was implemented as a simulator tool and used to evaluate the above-described scheduling mechanisms in [14]. HARQ is implemented by a bank of totally asynchronous stop-and-wait processes for each user. A type I HARQ protocol using chase combining (i.e., the same data block with the same MCS is retransmitted) was considered. To briefly describe the HARQ protocol, a particular UT MS_i having a connection with a particular BS_j is considered. The BS_j sends data to the MS_i using a HARQ process with a given number n . Then the protocol consists of the following steps for the first transmission of a data block:

- *TTI 1:* BS_j schedules HARQ process number n of MS_i , with a given MCS.
- *TTI 2:* BS_j broadcasts signaling, informs that a first transmission is scheduled for the next subframe and for HARQ process number n of MS_i , and provides all necessary information (e.g., MCS).
- *TTI 3:* BS_j transmits the data block of HARQ process number n of MS_i during the current subframe, and keeps a copy of this data block in its buffer dedicated to HARQ process number n of MS_i .
- *TTI 4:* The MS_i determines if the data block is erroneous by means of a cyclic redundancy check (CRC). If it is not erroneous, then the MS_i sends an ACK message to the BS_j . Otherwise, the MS_i sends a NACK message and waits for the data block retransmission. Meanwhile, it stores this data block as a “first version” in its buffer dedicated to the HARQ process number n . The BS_j receives an ACK/NACK message. If ACK was sent by MS_i , the data block is cancelled from the buffer and the HARQ process n can be used for a new transmission. Otherwise, HARQ process n is again activated and can be scheduled for retransmission.

8.2.1.1 Simulation Example for Assessment of Scheduling Algorithms

The simulations chosen here as an example of an assessment methodology for scheduling algorithms were performed in [6] and reported in [14]. The chosen setup considered a cellular-based OFDM/TDMA system composed of trisected hexagonal cells as shown in Figure 8.6.

The central cell approach has been used (see Chapter 7), in which only UTs within the three sectors of the central cell are actually simulated. The neighboring cells are assumed to be fully loaded during the whole simulation. Thus, these BSs transmit with maximum power.

The simulation mode is a combined snapshot dynamic. This means that one full simulation is composed of a large number of runs of dynamic simulations. At the beginning of each simulation run, the UTs are dropped uniformly over the three central sectors. During a dynamic run, traffic is generated according to given traffic models [14], and the fast-fading channel evolves according to the multipath channel models and the mobile speeds. However, the UTs positions are kept constant during each run.

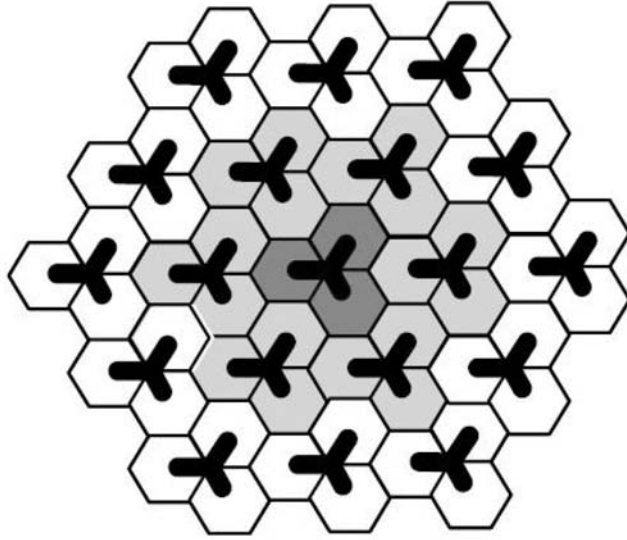


Figure 8.6 Simulation layout of a cellular system [14].

Within the scope of this simulation, at least 50 independent runs were run, each of them having a duration of at least 30 seconds = 15,000 TTIs. Details on the characteristics and the probability distributions employed in the traffic models can be found in [16].

8.2.1.2 Performance Metrics

The *average service throughput per sector* is used to study the network throughput performance and is measured as follows:

$$R_{\text{service}} = \frac{b}{k \cdot T} \quad (8.4)$$

where b is the total number of correctly received data bits by all UTs in the simulated system over the whole simulated time, k is the number of sectors in the simulation, and T is the simulated time.

The *average service throughput per user* is used to study the throughput performance for a given user and is measured as follows:

$$R_{\text{user}} = \frac{b_{\text{user}}}{T} \quad (8.5)$$

where b_{user} is the total number of correctly received data bits by a given UT in the simulated system over the whole simulated time and T is the simulated time.

The *average packet delay per user* is defined as the ratio between the accumulated delay for all packets received by a given UT and the total number of packets during one run. The delay for an individual packet is defined as the time between when the packet enters the queue at the BS and the time when the packet has been

received successively by the UT. If a packet has not been successfully delivered by the end of a run, its ending time is the end of the run.

Figure 8.7 shows the average service throughput per sector obtained with the different scheduling algorithms for full-queue traffic and 10 users per sector.

The maximum throughput scheduling method achieves a tremendous gain compared to round-robin scheduling. This clearly highlights the advantage of CQI usage for scheduling and the achievable multiuser diversity gain. The fair scheduling methods also achieve a multiuser diversity gain, but offer a lower service throughput than the maximum throughput scheduling. This is not surprising in a full-queue scenario, where the service throughput is obviously maximized by serving the users with good CQIs first.

In a full-queue scenario, there is no criterion to evaluate the quality of service achieved for the users individually. It is interesting to observe, however, how the achieved service throughput for each user is distributed in the sector.

Figure 8.8 shows the CDF of the average service throughput per user when various scheduling algorithms are employed.

The fair scheduling algorithms of categories 0 and 1 have CDFs with similar shape that are relatively steep in a limited throughput range. The PFS and the SB scheduling algorithms have similar performances and achieve considerable gains compared to the round-robin scheduler. In contrast, the shape of the CDF for maximum throughput scheduling indicates its very unequal distribution of the throughput among the users. With maximum throughput scheduling the probability for a user to have a lower service throughput than with round-robin scheduling is around 0.55. An important reason for this result is the unequal distribution of the CQIs of the different users in a sector due to their different geometry factors. Generally, the users close to the BS are selected more often by the maximum throughput scheduler than users at the sector edge.

Figure 8.9 shows the average service throughput per sector obtained with the different scheduling algorithms for nonreal-time video traffic and 10, 20, and 40 users per sector, respectively. There are two qualitative differences between the

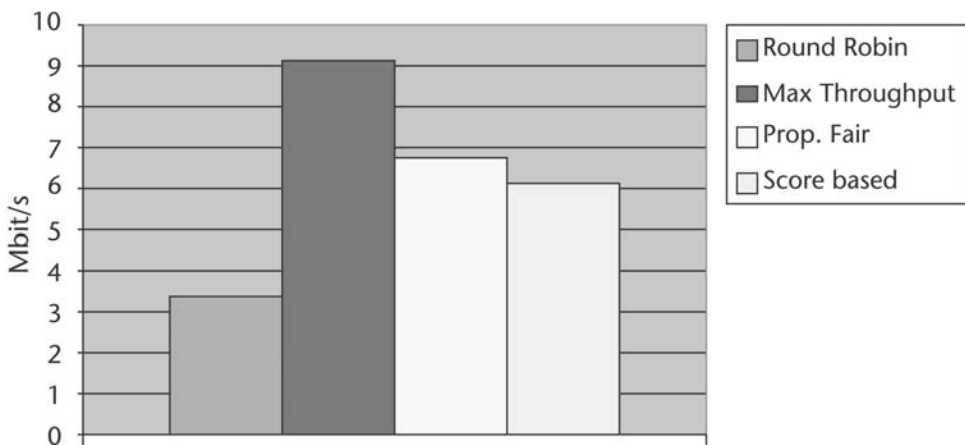


Figure 8.7 Average service throughput per sector obtained with the different scheduling algorithms for full-queue traffic and 10 users per sector [14].

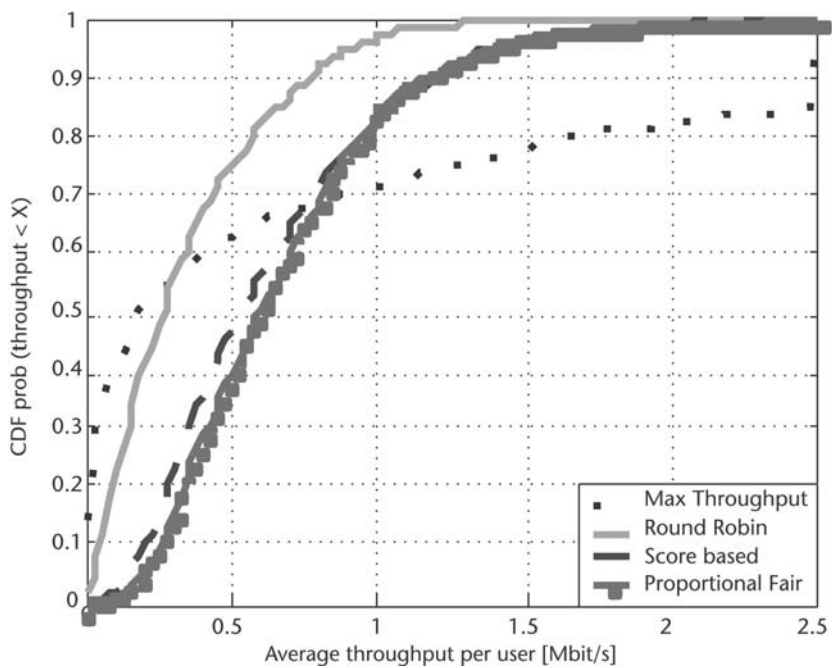


Figure 8.8 Average service throughput per user for full-queue traffic and 10 users per sector [14].

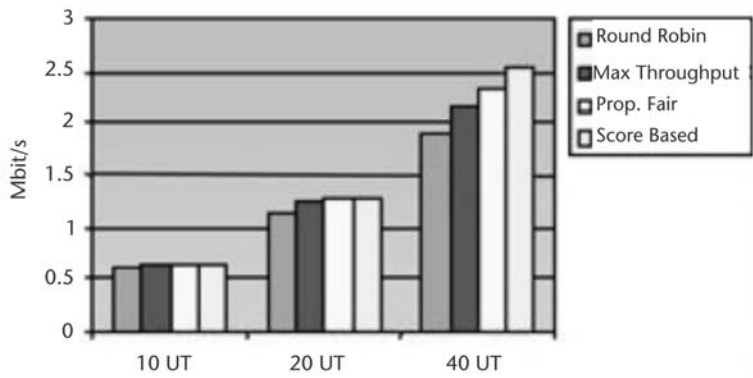


Figure 8.9 Average service throughput per sector for nonreal-time video traffic and different scheduling strategies [14].

results obtained for full queue and nonreal-time video traffic. First, the multiuser diversity gain obtained by the schedulers based on CQI is less important for nonreal-time video than for full queue. The reason for that is simply the fact that the number of users simultaneously competing for the radio resources is lower for nonreal-time video, because not all UTs have packets waiting in their queues. The gains of the CQI-based schedulers clearly increase with the number of UTs per sector.

Second, the fair schedulers of category 1 perform better than maximum throughput scheduling. Two facts are important to the explanation of this result.

First, the UTs have a finite number of packets waiting in their queues and, second, the CQIs of the different UTs are unequally distributed. In the extreme case of two users, a user is close to the BS and another is close to the border of the sector. In a simplistic way, we can assume that the achievable throughput for the user close to the BS is always higher than that for the user at the border. Hence, with the maximum throughput scheduler, the waiting queues of both users are emptied sequentially; first to be emptied is the queue of the user close to the BS, then the queue of the user at the border. In this case, we can say that both users are scheduled independently of their individual CQIs and, thus, achieve an individual throughput, which is close to their average throughput. In contrast, the fair schedulers tend to schedule a UT if its individual CQI is good compared to its average CQI. Thus, each time a user is scheduled, its individual throughput is close to its peak throughput. In this case, the overall service throughput obtained for maximum throughput scheduling is close to the sum of the individual average throughputs of the UT, while it is close to the sum of the individual peak throughputs of the UT for the fair schedulers.

In summary, the fair schedulers of category 1 can achieve a better overall throughput with bursty traffic models and the unequally distributed channel qualities of the UTs. The score-based scheduler was especially designed for unequal CQI distributions and, thus, has a considerably better performance than the proportional fair scheduler.

The QoS achieved by the different schedulers can be assessed through the average packet delay per user. Figure 8.10 shows the CDF of the average packet delay per user for 20 and 40 users per sector. For 20 users per sector, the CQI-based scheduling algorithms achieve considerably shorter delays than the round-robin scheduler. Already at this low system load the score-based scheduler proves its superiority compared to PFS and maximum throughput scheduling. For 40 users per sector, the advantage of the score-based scheduler becomes even more evident.

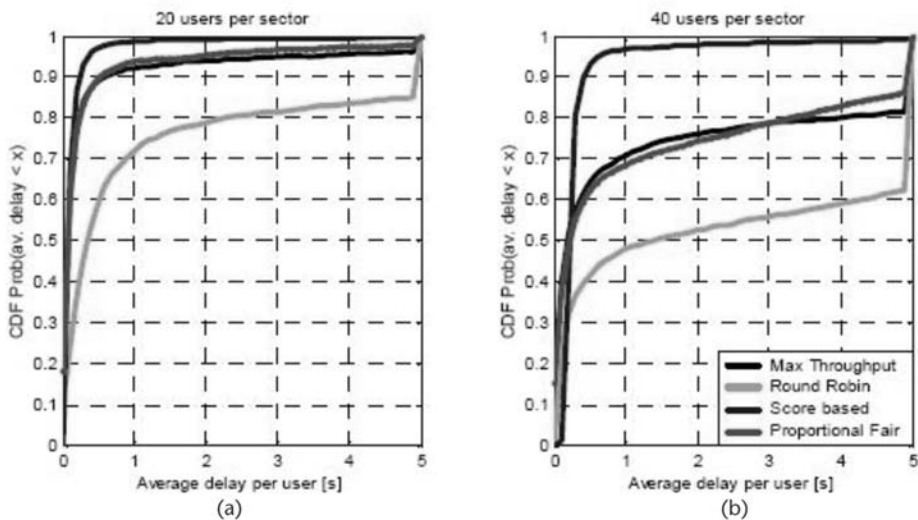


Figure 8.10 (a, b) Average packet delay per user for nonreal-time video [14].

While the increased system load leads to considerably higher delays for the other schedulers, the score-based scheduler only shows a slight increase in the average delay per user.

The results highlight the gains achieved from multiuser diversity, when CQI feedback from the UT is used in the scheduling algorithm. The results presented here were obtained in a scenario with low mobility, where the CQI mismatch due to feedback delays is almost negligible. A more detailed measurement model and an assessment of the impact of higher mobility are recommended for more realistic results.

The unequal distribution of the CQIs of different users due to their different path loss and shadowing has an important impact on the performance of the scheduling algorithms with realistic traffic models. Here, schedulers that include fairness in their metrics are clearly advantageous and can even achieve a higher overall throughput than the maximum throughput scheduler. A score-based scheduling algorithm can be especially designed to cope with unequal CQI distributions.

8.2.2 Link Adaptation

Link adaptation algorithms provide powerful means to achieve the highest possible performance in wireless communication systems by dynamically adapting various transmission parameters of the wireless link to the changing radio channel and interference conditions. A typical transmitting parameter, which is often dynamically adapted in time, is the *transmitting power*. By dynamically adapting the transmitting power to the actual channel conditions, energy in battery-powered devices is saved and the system performance is enhanced due to interference reduction to neighboring cells. Other parameters include modulation order and coding rate, spreading factor, and signaling bandwidth.

Link adaptation can not only improve the performance of channels that vary in time, but allows them to cope with frequency-selective fading, which is caused by the multipath propagation of the radio waves. In multicarrier-based systems, frequency-selective fading causes variations in the channel gains at different subcarriers, leading to severe performance degradations at the subcarriers where a deep fade occurs. This fading problem, however, can be improved by link adaptation strategies. A promising technique for wireless systems using OFDM is adaptive bit and power loading [26], in which a different number of bits and power values are assigned to each subcarrier depending on the measured channel conditions and QoS requirements.

The purpose of link adaptation is to ensure that the most efficient set of transmission parameters (or transmission mode) is always used over the varying channel conditions. In systems without link adaptation, a single mode of operation has to be chosen such that an acceptable performance level is maintained even for poor channel conditions. Consequently, when the channel is good, the system is unable to exploit it fully. In contrast, if several modes of operation are available, higher spectral efficiency can be achieved during favorable conditions, and robustness to errors can be increased during bad conditions. The adaptation period of the link adaptation algorithms has to be much smaller than the coherence time of the channel to obtain the maximum benefit [14].

A sensitive issue in link adaptation algorithms is the definition of a CQI, based on which the adaptation will be performed. Various metrics can be employed to compute the switching threshold between different modes. A typical metric is the average SNR measured at the receiver. The minimum SNR required to achieve a given error performance target is computed for each mode and stored as the switching threshold. Once the average SNR is measured, the optimum mode is determined according to the thresholds and fed back to the transmitter.

This method poses a number of practical problems. First, the delay introduced by the feedback operation and other implementation limitations will not allow an instantaneous adaptation. If this delay is bigger than the channel's coherence time, the new transmission mode will not be optimal anymore. In addition, the performance of the whole scheme might be compromised by SNR estimation errors. A proposed method to avoid this is to estimate the statistics of the SNR instead of a single realization. Then, the threshold determination can be done, for example, based on the moments of the SNR. The estimation accuracy and the computational efficiency have to be balanced.

Another metric that can be used to perform link adaptation consists of tracking reception errors of packets for each candidate mode. The observed packet error rate (PER) provides implicit channel information for each trained mode. The estimation accuracy, however, depends on the number of packets received during the observation window; therefore, this method is traffic dependent. If the user does not send or receive packets, one loses track of the channel quality. Consequently, the reaction time of the algorithm is difficult to control.

In general, measuring the ability of the scheme to perform fast adaptation under realistic traffic and without requiring additional bandwidth is an open research problem. The motivation for using metrics that do not require feedback from the receiver lies not only in the practical problems caused by providing the transmitter with CQI, but also in the fact that, in general, existing systems do not support any protocol means for the receiver to inform the transmitter about the actual channel quality.

8.3 Cross-Layer Optimization and Design

Unlike fixed wired networks in which channels are time invariant, wireless channels are very dynamic and time varying. Hence, the conventional design approach of optimizing the PHY and MAC layers independently fails to exploit the dynamic nature of the PHY layer and is suboptimal in multiuser wireless channels. Therefore, despite the fact that the separation between networking and communication problems has many advantages, both practical and conceptual, there is an increasing awareness that this approach is not the proper one, in particular within the context of modern wireless data networks. This motivates the cross-layer approach, according to which the PHY layer and higher networking layers are jointly optimized.

Cross-layer optimization design presents two great challenges. The MAC layer scheduling algorithm must adapt to the dynamics of the system, which in turn depends on the PHY layer dynamics and the source statistics. In modern packet-

based wireless data networks, the data traffic arrives randomly at transmitters, in the form of variable-size packets. Transmitters then send the data over fading channels, generally in the presence of interference generated by other transmitters. In this scenario, the problem of resource allocation, in terms of power control and rate allocation, becomes of fundamental importance.

Two issues were identified that are central to the resource allocation problem for multiuser communication:

- At the network and MAC layer, QoS issues regarding packet throughput and delay are important
- At the PHY layer, issues of accurate channel modeling, detection, and coding are prominent.

The cross-layer approach combines all of these issues into an integrated framework [13].

The main shortcoming of the strictly based PHY layer approaches is that they do not take into account the impact on the upper layers, whereas the main shortcoming of strict layering MAC layer-based schemes is that they are based on “hard” channels, that is, they use very limited information from the PHY layer [27]. If spectrum or other radio resource usage is at a premium, more efficient management might be possible if cooperation is considered between the MAC and the PHY layers, namely, by using a richer set of PHY layer parameters at the MAC layer. This is particularly true in multiple-transmitter, multiple-receiver (MTMR) systems, in which the existence of SDMA at the PHY layer raises significant issues at the upper layers and whose impact is significantly different according to the access scheme. In fact in SISO wireless systems, the radio resources (e.g., time slots or frequency slots) in the TDMA/FDMA access schemes can be considered without significant error for well-designed orthogonal systems. Even with CDMA schemes, the codes are designed to be orthogonal or quasiorthogonal. Although transmission through dispersive channels may lead to a break in the orthogonality, the departure from orthogonality may be moderate—at least for low-dispersion wireless channels—and thus from a layer 2 viewpoint it is not unreasonable to consider that the radio resources (e.g., codes, time slots, or frequency slots) are available on an on/off basis.

The situation changes radically when considering the spatial dimension associated with the use of MTMR. Unlike the other types of radio resources (i.e., frequency slots, time slots, codes), this one cannot be considered as a set of orthogonal or quasi orthogonal items and, therefore, considerable inefficiencies are to be expected if, from the viewpoint of the MAC layer, following the conventional strict layering approach, the PHY layer is viewed as providing a set of “hard” (on/off) bit pipes. Such a situation motivates a joint design involving cooperation between the two layers, violating the modularity and independence inherent to the strict layering approach [27].

Although the need to break—or the convenience of breaking—the strict layering approach to reduce inefficiencies in wireless networks comes from the dynamic and random nature of the physical channel, the first choice when considering a cross or joint layer design, is to imagine a cross-layer information entity that

contains system state information and can be exploited by higher layers (transport and network layers) to support adaptive tuning of layer mechanisms to improve throughput efficiency and maintain link quality.

8.3.1 Cross-Layer Information and Signaling

A cross-layer architecture encompasses an additional complexity relative to a strict layering approach due to the fact that additional information besides the one that defines the basic service provided by the layer has to be exchanged. This need to exchange additional cross-layer information (CLI) leads to the two fundamental questions related to cross-layer design:

1. What type of information should be exchanged across protocol layers and how frequently should this exchange proceed?
2. Which procedures will be adequate and efficient for conducting this exchange of information?

Cross-layer design is, by definition, a departure from a rule-based procedure defined in a strict layering approach [27], therefore, there is no universal answer to these questions. Basically, the procedure will depend on the specific algorithm requirements of a given layer and on several algorithm-dependent aspects (e.g., a centralized or distributed approach), and also on the specific system architecture.

Although no unique cross-layer architecture yet exists, a general approach would consider the type of information in a broad sense that can be used at a given layer. Between the PHY layer and a generic MAC layer, we can distinguish three types of information flow that could be useful:

- PHY to MAC, which means that PHY layer information can be passed to the MAC layer to improve the collision resolution;
- PHY to data link control (DLC), which means that CSI provides side information to enhance the link reliability or throughput (i.e., by advanced packet scheduling);
- Network to DLC parameters from the higher layers related to the QoS requirements for a given application (e.g., tolerable delay, max-min allowable bit rate, BER/FER requirements) may be made available to the lower layers.

Because the cross-layer design approach to minimize inefficiencies in wireless networks had been mainly motivated by the dynamic underlying nature of the wireless channel, CSI is the most important type of information from the lower layer that is needed for the upper layers [27]. This raises problems of CSI availability and CSI requirements from the upper layers. The latter depend on the specific algorithms to be implemented at the MAC or upper layers, and generally the basic information that is required is the SINR and propagation loss [28]. For MIMO systems, however, a reduced number of parameters related to the spatial characteristics (e.g., AoA, number of spatial signatures [29]) is necessarily sufficient for the currently proposed MAC and routing algorithms employing PHY layer information.

Of a different nature is the problem of CSI availability. The issues are quite different depending on whether a receiver or transmitter function is being performed [27]. For a receiver function, the CSI is generally available from the channel estimation unit of the PHY layer and the problem will be the one of *signaling* this information through the layers. For a transmitter function (e.g., scheduling) the problem is more complicated. In TDD systems, due to the channel reciprocity, channel estimates of the receiving link can be used for the transmitting functions. This is only valid, however, up to a certain extent; due to the variability of the wireless channel, if the delay between the receiving and the transmitting slots is higher than the coherence time of the channel (which may be common with high-speed mobiles), the CSI is not reliable. In such a case, the designer has to resort to long-term statistics or to a reduced set of parameters that do not show such a high variability (e.g., AoA). Even with low speeds, problems might occur (e.g., considering the UL-DL architecture of cellular networks, different access schemes and data formats are employed in the two links).

The problem of information exchange requires intranode, interlayer signaling, and internode signaling and distribution of relevant parameters for cross-layer operation.

For interlayer signaling and transport of CLI, essentially three methods are available [30]:

- *Packet headers:* In IPv6, optional network layer information can be encoded in additional headers where the CLI can then be encapsulated. This method makes use of IP data packets as in-band message carriers with no need to use a dedicated internal message protocol.
- *Internet Control Message Protocol (ICMP) messages:* ICMP is a widely deployed signaling protocol in IP-based networks that can be considered an alternative for conveying the required interlayer CLI.
- *Network service:* In [31], a specific access network service called wireless channel information (WCI) was proposed. In this scheme, channel and link states from the PHY layer and link layer are gathered, abstracted, and managed by third parties, the so-called distributed WCI servers. Interested applications then access the WCI for their required parameters from the lowest two layers.
- *Local profiles:* In [32], local profiles are used to store information that is periodically updated. CLI is abstracted from each necessary layer and stored in separate profiles. Other interested layer(s) can then select the profile(s) to fetch the desired information.

As a comparison, the FP6 IST project 4MORE [33], which investigated the introduction of a broadband component and focused on the design of a system-on-chip (SoC) for a 4G terminal employing multiple antennas and based on MC-CDMA techniques, chose for its cross-layer framework the network service or local profiles approach based on the use of a separate user profile plane [34], in which each layer writes and reads user-specific information. This is shown in Figure 8.11.

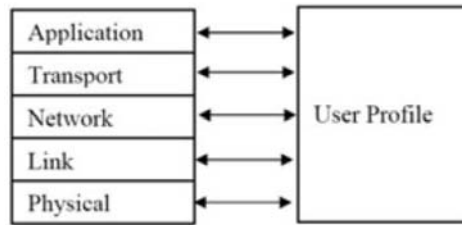


Figure 8.11 Cross-layer signaling architecture based on the existence of user profile plane [27].

The 4MORE project chose a centralized approach [34], in which the resource allocation was decided at the BS and, therefore, the profile associated with each user only needed to be stored at the BS.

Table 8.2 suggests an initial selection of parameters for a cross-layer approach.

Internode distribution of CLI is of crucial importance for the implementation of both distributed algorithms and centralized algorithms. Although there is no specification of any universal architecture, and different solutions are possible according to the system considered, the internode distribution of CLI and resource map assignments can rely on some kind of a broadcast channel, either differentiated by codes or using specific time slots [27].

8.3.2 Scheduling

One of the functions that can benefit the most from a joint PHY-MAC design (even with a SISO channel) is the scheduler for packet transmission. Efficient packet scheduling schemes are required to manage users' access to resources according to the instantaneous traffic requirement and the channel conditions they are experiencing. The most straightforward approach for performing scheduling while taking into account the channel characteristics is to order the users according to the SIR of the respective channel. This is generally termed *opportunistic scheduling* [35], which exploits the multiuser diversity.

Such an approach gives obviously good results when we have homogeneous traffic insensitive to the delay, because this corresponds to a throughput maximization [28]. However, if the delay is an important requirement, or if there are multiple services with different costs, this may lead to inefficiencies because the users with the best channels tend to monopolize the system resources. Therefore, proposals have been made for scheduling algorithms that use the parameters from the PHY layer but also consider the QoS from the upper layers. Within the FP5 IST project MATRICE [36, 37], the prioritization function to schedule the packets, considered

Table 8.2 Selection of Parameters for Cross-Layer Information [27]

<i>Application Related Parameters</i>	<i>CSI-Related Parameters</i>
Maximum allowable delay	SIR
Allowable packet error rate	Orthogonality factor
Minimum bit rate required	Special Signature

a function that weighted parameters from the PHY layer, namely, the SIR and the number of attempted retransmissions considering chase combining at the reception, and the allowable delay for the service considered.

Other approaches have attempted to equalize the average throughput to provide and, in doing so, some kind of “fairness” function has been proposed either for single or multiantenna systems [38]; however, these still fit in the category of PHY layer-based scheduling algorithms, because the fairness is something that is a consequence of the algorithm and not a QoS input from the layers above. Joint selection of the modulation and coding scheme together with the scheduling has also been considered [37].

Another example is if a dynamic resource allocation (DRA) [27] algorithm is used in a combination with a MIMO channel, the introduction of antenna elements at both the BS and MT will create an additional dimension in resource space. Therefore, more effective scheduling is required to exploit the additional “co-slots.” Therefore, when a new user enters the system, given an allocated set of users, the orthogonality between the desired user and each and every other user on the same time slot, and code, needs to be explored.

The aim of a good scheduling algorithm is to allocate a resource in a three-dimensional space (i.e., time, code, slot) where it interferes the least with each and every other user. This requires an assessment of the correlation properties between users, so that a suitable code can be chosen. In a broader sense, the presence of multiantenna arrays will define several cluster groups within the cell, and the DRA algorithm will need to make an estimate about which cluster group each user belongs to within the cell. The users that lie at the cluster boundaries will inherently interfere across several clusters, and, therefore, if it is a low-priority user, the scheduler may decide to defer the user to the queue until the user’s priority increases or its position improves [27].

In the following, an algorithm that is part of the above-described resource allocation policy and that decides how to allocate a user in the spatial domain given a channel state entity matrix is defined [27].

The method is based on a best-fit scheduling technique [35] and applied to an MC-CDMA air interface. It is assumed that a priority function provides a list of scheduled users according to QoS only, and defers the CSI to the packet allocator instead.

It is assumed that a list of scheduled packets is fed to a first-in/first-out (FIFO) buffer, and that with each packet there is information surrounding the minimum SIR requirements and spatial signature. The packet allocator algorithm is given by the following routine.

The algorithm picks a packet from the beginning of the list and tries to assign a suitable resource both in time, code, and co-slot. Given that a single-user code would be allocated, then the problem is redefined to this situation. To allocate a co-slot, a metric must be attained for the correlation that exists between the new user and the existing users who have already been allocated resources. For this, the code in which the least interference results is chosen. One particular method is to attain the channel signature matrix and to calculate the average covariance matrix. From the matrix, the correlation between users on the same code can be

distinguished. The correlation is obtained for all users, and effectively the users that exhibit the least correlation among them are grouped together or

$$T_{f,b} = \min_{i,j} \arg \alpha_{i,j} \quad (8.6)$$

$$T_{f-1,b} = \min_{\substack{i,j \\ i,j \\ i,j}} \arg \alpha_{u,v}$$

where α is the channel autocorrelation value for the s th and the i th users. The α /CSI value reflects the number of MIMO taps resulting in excessive matrix computations. Equation (8.6) shows that the minimum correlation between users must be found. The minimum value is then assigned a code on a co-slot b , and the user is excluded from the next search computation.

For nonreal-time services, all co-slots must be filled, so that effectively the end results are N clusters of a user, where N is the number of antennas. If not all codes are employed, then the users are inserted one by one, until the FIFO is empty. This procedure is shown in Figure 8.12.

Figure 8.13 shows a detailed view of the packet allocation unit. It considers two lists that are reprioritized according to their antenna allocation.

Additionally, extended scheduling can be applied here according to CSI. A HARQ process is applicable to each available code, where there will be four instances of HARQ per process. A rate selection mechanism will allocate the more appropriate modulation and coding scheme to maximize the link throughput; however, it is optional whether additional signaling bits are multiplexed here, for PHY layer control, or if the information is made available to the PHY layer via the user profile.

In summary, an additional functional block would be required to support MIMO-based scheduling, in the form of a *packet space divider*. The packet space divider needs to separate users according to their correlation properties into antenna

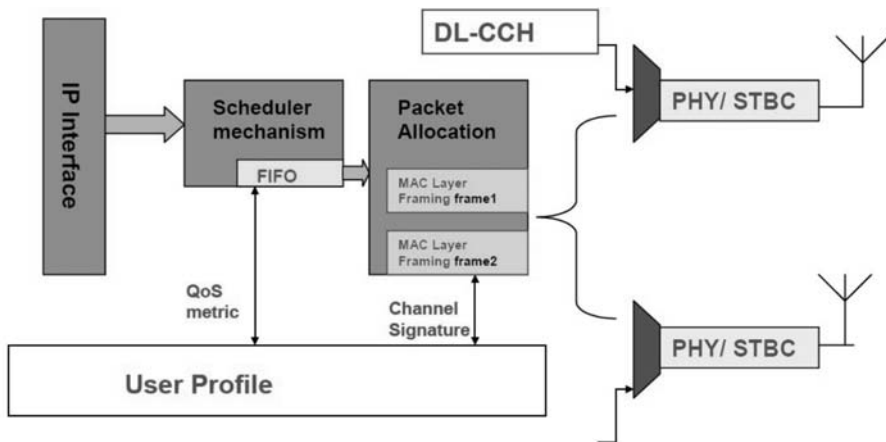


Figure 8.12 MAC layer design in support of DRA [27].

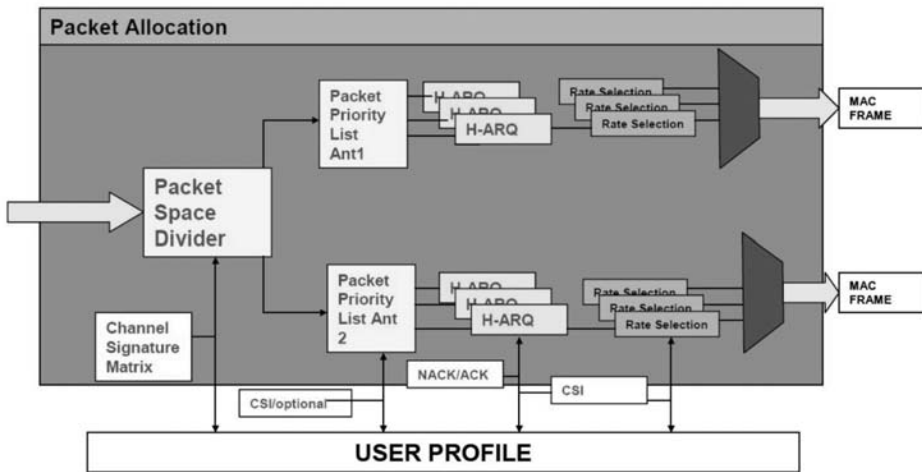


Figure 8.13 Packet allocation unit [27].

clusters. However, the metric that is required to realize this decision has to reflect the correlation between the users.

8.4 Cross-Optimization of FM-UWB and MC-SS Radio Interface Solutions

The new generation of WPAN devices requires dual-mode air interfaces in order to achieve high spectrum efficiency and to span from low data rates to high data rates (from a few bits per second to hundred of megabits per second) [39, 40].

Several scenarios have been proposed that clarify the need for multimode LDR/HDR WPAN radio interfaces. The two radio interfaces, based on FM-UWB and MC-SS transmission technologies, respectively, for the LDR and HDR applications, are outlined in [41], which focuses on the need for high coexistence levels.

Figure 8.14 shows the protocol architecture of a multimode terminal for WPAN applications.

The architecture is composed of the cross-layer database/functionality, the logical link control sublayer, and the bridging functionalities located within the frame convergence sublayer and the service-specific convergence sublayer. Full details are available in [42].

Because the transmitting power of HDR signals is 34 dB above the transmitting power of LDR signals, and because the latter are additionally spread over a large frequency range, simultaneous LDR transmitters hardly have an impact on HDR signal reception, except if an interferer and a receiver reside within one device.

It has been found that LNA filtering, external filtering, and antenna filtering are good solutions for the management of interference of HDR transmission on LDR reception [40]. However, when the two radio interfaces are located close to each other, the MAC and higher layers mechanisms are needed to ensure coexistence

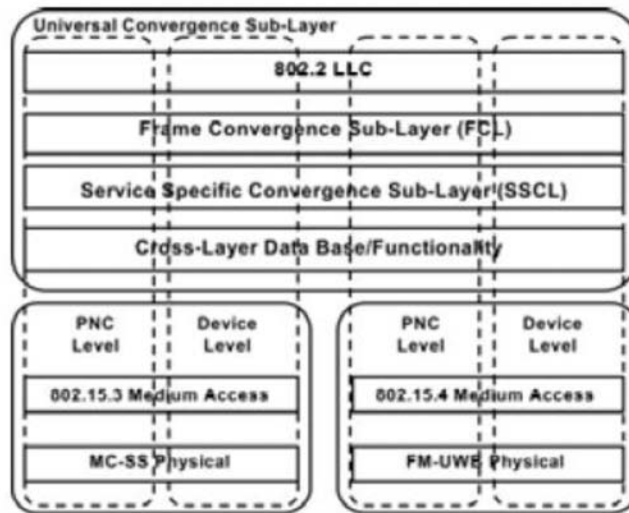


Figure 8.14 Protocol architecture of a multimode terminal [42].

[42]. At the same time also, the MACs should be different because of the different needs in terms of application requirements, duty cycle, and complexity.

In a UWB-based WPAN network, the main objective of the distributed MAC is to use the limited radio resources effectively and provide the coordination function for the multiple accesses [41].

The MAC layer must be able to guarantee the different QoS requests of the network layer, including bit rate, delay, and PER. The different QoS requests from the network layer can be classified into the typical “best-effort” (non-QoS-guaranteed) and QoS-guaranteed traffic. Each services class has a corresponding service policy, which must be defined at the MAC level. The radio resource can be logically divided into two parts:

- One for the QoS-guaranteed services, which is negotiated with the network layer and is only controlled by the network layer;
- One for the best-effort services, which dynamically changes depending on the MAC layer and the PHY layer.

Typical applications for LDR devices (e.g., sensor networks) can run with low duty cycles (under 1%). The simultaneous use of different and/or uncoordinated wireless networks that overlap (at least partially) in range, time, and frequency generates mutual effects of interference, which decreases the performance of such networks.

The two categories of coexistence mechanisms are *collaborative* coexistence mechanisms, in which the two interfering networks exchange information, and *noncollaborative* coexistence mechanisms, in which the exchange of information is not allowed [43]. The possibility of exchanging information is quite easy when the two radio interfaces are colocated in the same dual-mode terminal [42].

Reference [39] proposed a collaborative coexistence mechanism between an LDR and HDR interface, the so-called alternating wireless activity (AWA). It

controls and synchronizes access to the network of the personal devices associated with the LDR and HDR WPANs. Because it relies on time-division alternation of LDR and HDR WPANs, it totally avoids interference even in the worst case that occurs when the LDR (HDR) interferer is located close to the HDR (LDR) receiver. Its functionalities are positioned in a common protocol layer above the LDR and HDR MAC sublayers.

This novel coexistence mechanism was proposed between 802.15.4-based LDR and 802.15.3-based HDR WPANs radio standards and exploits the MAC features of the two, but can be properly used by any kind of PHY layer proposed for the two standards.

Channels conflicts lead to in-band interference, which is one of the strongest types of interference. However, if a channel conflict does not occur, out-of-band interference can be experienced, especially when the interferer is located near to the receiver (see Chapter 7). In a multimode terminal this effect is even more critical, because the two radio interfaces are close, and thus it is necessary to approach the problem in an exclusive way.

Although the PHY layer of the IEEE 802.15.3 standard [44] is quite different from the PHY proposed in the MAGNET HDR interface [9], the MAC layer is almost the same. This MAC includes several mechanisms that allow for flexible resource management and support of QoS, which are key requirements for HDR WPAN devices.

An 802.15.3-based WPAN operates as a centrally controlled ad hoc network called piconet; data exchange in a piconet is performed in a peer-to-peer manner. A piconet consists of a piconet coordinator (PNC) and one or more devices (DEVs) that are synchronized with the PNC. Synchronization is required because the MAC superframe is structured in time slots. Two devices in the piconet can communicate directly by either randomly accessing the time slots in the contention access period (CAP) of the superframe or by accessing the channel in some assigned time slots of the superframe during the channel time allocation period (CTAP).

Depending on the application requirements, a MAGNET LDR WPAN may operate in either a star topology or a peer-to-peer topology. In the star topology, communication is established between devices and a central controller called the PAN coordinator. The peer-to-peer topology also has a PAN controller, however, it differs from the star topology in that any device may communicate with any other device.

The PAN coordinator bounds its channel time by using a superframe structure (the MAGNET LDR MAC allows only for a beacon-enabled PAN). A superframe starts with the transmission of a beacon frame. In a superframe-enabled WPAN, the superframe can have an active and an inactive portion; the active portion is divided into 16 equally sized slots. The general structure of the MAC superframe is based on the 802.15.4 MAC standard [45], and consists of four parts:

- The *beacon frame*, which is used to synchronize the devices associated with the WPAN, identify the WPAN, and describe the structure of the superframe. The beacon frame is transmitted in the first slot of each superframe.
- The *contention access period*, during which devices can communicate using a slotted CSMA/CA mechanism.

- The *contention-free period* (CFP), during which access to the channel is controlled by the PNC, which assigns guaranteed time slots (GTSs) for that communication in response to the request message. The PAN coordinator can allocate up to 7 GTSs, and a GTS can occupy more than one slot.
- The *inactive period*, during which devices may enter a low-power mode.

The coexistence mechanism proposed in [39] exploits the capability of this standard to run with low duty cycles.

8.4.1 AWA Algorithm

The AWA algorithm works by controlling and synchronizing the access to the network of the two radio interfaces. Its functionalities are positioned in a common protocol layer over the two MAC sublayers. It makes use of the 802.15.3 child piconet functionality and the inactive period of the 802.15.4 device.

Because no 802.15.3 devices are transmitting during a private CTA allocated to a child piconet, this CTA can be allocated to a 802.15.4 WPAN that will not be interfered with by any HDR device. On the other hand, because no 802.15.4 devices are transmitting during the inactive portion of the superframe, this inactive portion will be synchronized to overlap the entire 802.15.3 superframe except the i th private CTA that is overlapping in time with the active portion of the 802.15.4 WPAN. In this case the 802.15.3 WPAN will not be interfered with by any LDR device.

Synchronization of the 802.15.3 and 802.15.4 superframes is shown in Figure 8.15.

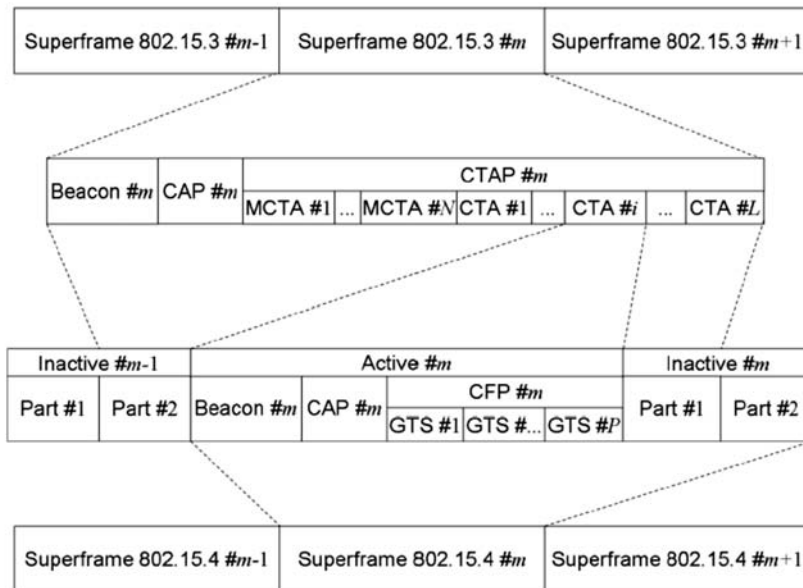


Figure 8.15 Synchronization of the 802.15.3 and 802.15.4 superframes (AWA) [39].

The i th CTA of the m th 802.15.3 superframe is allocated to the active portion of the m th 802.15.4 superframe. The inactive portion is virtually divided into two parts: The m th 802.15.3 superframe starts simultaneously with the second part of the inactive portion of the $(m - 1)$ th 802.15.4 superframe, while the m th 802.15.3 superframe ends simultaneously with the first part of the inactive portion of the m th 802.15.4 superframe.

The synchronization of the two superframe sequences frees all LDR and HDR devices associated with a common LDR and HDR PAN controller from interference.

The duration of the first part of the inactive portion of the m th superframe is denoted as $SID_{LDR}^{m'}$, whereas the second part of the inactive portion is denoted as $SID_{LDR}^{m''}$.

Because the LDR and the HDR WPAN controllers are expected to exchange information for network coordination and synchronization of the respective superframes, the AWA mechanism is a collaborative coexistence mechanism.

There are two restrictions to the exploitation of the AWA coexistence algorithm:

- There must be a dual-mode HDR/LDR WPAN device within the common coverage area of the 802.15.3 and 802.15.4 WPANs. This device shall be the coordinator of both 802.15.3 and 802.15.4 WPANs.
- The 802.15.4 WPAN must be beacon enabled with an active and an inactive period. This restriction is compulsory according to MAGNET LDR MAC.

It is assumed that coordination of the HDR WPAN is performed by the 802.15.3-based radio interface of the LDR/HDR dual-mode device, whereas the coordination of the LDR WPAN is performed by the 802.15.4-based air interface of the LDR/HDR dual-mode device.

The steps of the algorithm are as follows:

1. The coordinator of the 802.15.3 WPAN starts a piconet, allocating a private CTA (the i th CTA) to the 802.15.4 WPAN. If a CTA is allocated to a child piconet, the DestID and SrcID related to this CTA will both be the DEVID of the DEV that is the child piconet's PNC. This means that the 802.15.3 PNC will use its own ID as DestID and SrcID of the 802.15.4 virtual child piconet. The 802.15.3 PNC should use a pseudostatic private CTA.
2. The coordinator of the 802.15.4 WPAN sets the superframe duration equal to the superframe duration of the 802.15.3 WPAN. Therefore, the superframe periodicity is the same for both 802.15.3 and 802.15.4 WPANs.
3. Under the assumption that all of the private CTAs allocated to the 802.15.4 WPAN are pseudostatic, the following equation holds:

$$CTD_{HDR}^{(m)} = CTD_{HDR}^{(m+1)} \quad (8.7)$$

where $m = 1, 2, \dots$, is the transmission index. Furthermore, the position of the pseudostatic time slot (i.e., the i th CTA) of the m th 802.15.3 superframe is equal to the position of the pseudostatic time slot (i.e., the i th CTA) of the $(m + 1)$ th 802.15.3 superframe.

4. For the synchronization of the two networks, the following holds:

$$\begin{aligned}
 \text{SFD}_{\text{HDR}}^{(m)} &= \text{SFD}_{\text{LDR}}^{(m)} \\
 \text{CTD}_{\text{HDR}}^{(m)} &= \text{SAD}_{\text{LDR}}^{(m)} \\
 \text{SID}_{\text{LDR}}^{m'} &= \text{SID}_{\text{LDR}}^{m+1'} \\
 \text{SID}_{\text{LDR}}^{m''} &= \text{SID}_{\text{LDR}}^{m+1''}
 \end{aligned} \tag{8.8}$$

Because of the constraint on SFD_{HDR} , which is $0 < \text{SFD}_{\text{HDR}} < 65535 \cdot 10^{-6}$ sec, and the constraint on SFD_{LDR} where $0 < \text{SO} < \text{BO} < 14$, $\text{SFD}_{\text{HDR}}^{(m)} = \text{SFD}_{\text{LDR}}^{(m)}$ can only be satisfied with $\text{BO} = 1, 2$.

Furthermore, to use the inactive portion, possible values of BO and SO are $(\text{BO}, \text{SO}) = (1, 0)$, $(\text{BO}, \text{SO}) = (2, 0)$, and $(\text{BO}, \text{SO}) = (2, 1)$, which provide 802.15.4 duty cycles of 50%, 25%, and 50%, respectively.

8.4.2 Improved AWA

In an improved version of the AWA coexistence mechanism (IAWA), the private CTA for LDR is not allocated once per superframe; it is instead allocated once per N superframes. In other words, the private CTA is allocated at the superframe m_N (where $m = 1, 2, \dots$, is the transmission index), while it skips the private CTA allocation for the next $N - 1$ HDR superframes.

Use of the improved version of the AWA coexistence mechanism is compulsory when the LDR beacon interval is higher than $65,535 \cdot 10^{-6}$ sec, which is the maximum value of SFD_{HDR} . This happens with $\text{BO} > 2$ or when the data rate is lower than 120 Kbps.

In the basic version of the AWA mechanism, the PNC of the HDR network computes the superframe structure once per time. With this improved version, the PNC computes the superframe structure once per N superframes. For both AWA mechanisms, basic and improved, no modification to the IEEE 802.15.3 and 802.15.4 MAC standards is required [39].

An overview of the structure of the MAGNET LDR superframe is shown in Figure 8.16, which shows that the minimum superframe (SF) duration with the LDR radio interface is equal to 38.4 ms.

The synchronization of the 802.15.3 and 802.15.4 SFs is shown in Figure 8.17.

With the improved version of AWA, it is possible to use more HDR superframes, which leads to a more flexible solution.

Due to timing constraints, the time duration of the HDR beacon plus the HDR CAP, and the HDR CTA cannot be higher than 27.135 ms (i.e., the maximum allowed HDR SF duration is 65.53 ms and the minimum duration for the LDR active part is 38.4 ms). Thus, only the selected combinations in the lower part of Figure 8.17 are allowed.

8.4.3 Performance Analysis

The FM-UWB and MC-SS air interfaces are characterized by the parameters shown in Table 8.3. The interference between the proposed FM-UWB and MC-SS systems

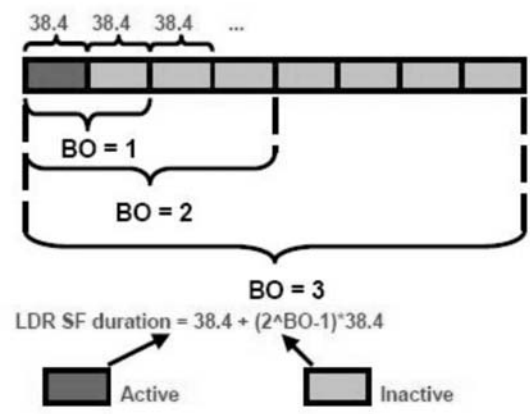


Figure 8.16 LDR superframe structure [39].

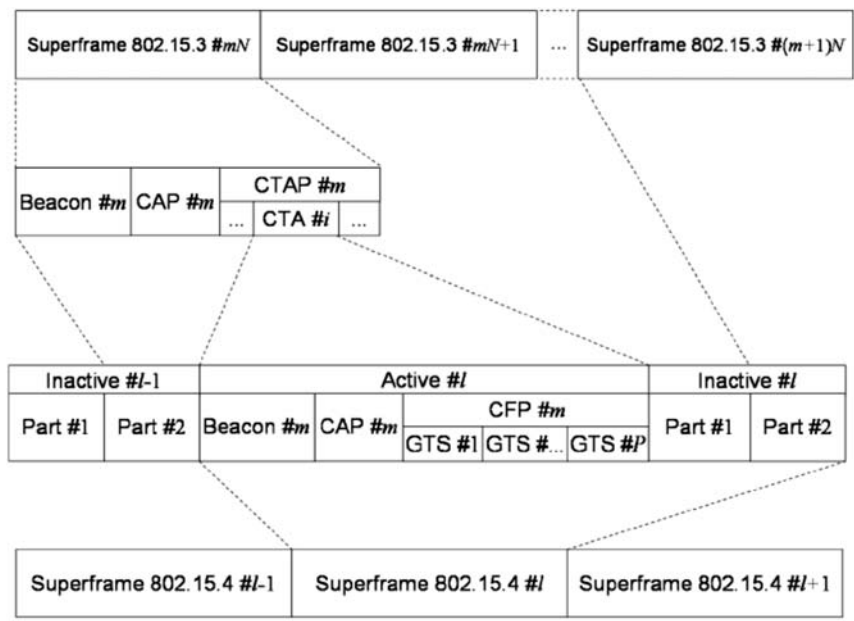


Figure 8.17 Synchronization of the 802.15.3 and 802.15.4 superframes (IAWA) [39].

Table 8.3 Characteristics of the FM-UWB and MC-SS Radio Interfaces

Parameter	FM-UWB	MC-SS
Transmit power	−14 dBm	+20 dBm
RF center frequency	4.5 GHz	5.25 GHz
RF signal bandwidth	500 MHz	36 MHz
RF signal envelope	Constant	Strongly varying
Predominant modulation	FM	AM
Path loss at 1m	45 dB	46.5 dB

can be classified as out-of-band interference. The major differences are the transmission power (34 dB of difference) and signal bandwidth. The signal attenuation as a function of the distance is almost equal, because the two operating frequencies are relatively close to each other.

The center frequency of the MAGNET HDR and LDR air interfaces is different (5.2 GHz for HDR and 4.5 GHz for LDR). This means that the HDR and LDR signal overlapping at the receiver cannot be simulated at the baseband, for which the center frequency is 0 Hz for both radio interfaces. Therefore, a passband multimode simulator was developed [39] taking into account the actual carrier frequencies of the system. This multimode simulator has been developed by using the SIMULINK tool, which is recommended when a high level of system complexity has to be managed and the simulation speed is a concern. The simulator is shown in Figure 8.18.

The *normalization gain block* normalizes the two power levels (LDR and HDR), whereas blocks G1 and G2 take care of the path loss of the HDR link (HDR transmitter to LDR receiver) and LDR link (LDR transmitter and LDR receiver), respectively.

No channel was considered in the assessments nor were the additional gains coming from filtering or other hardware interference attenuation solutions considered. The *transmitter chain block* and relative *receiver chain block* were developed according to the specification explained earlier including channel coding, puncturing, interleaving, spreading, and multicode and OFDM modulation. In particular, 16-QAM and a coding rate of 3/4 were considered.

The SIMULINK model used for the multimode simulator is based on the specification of the respective HDR and LDR PHY layers. While in FM-UWB the analog FM modulator naturally translates the baseband FSK digital modulation

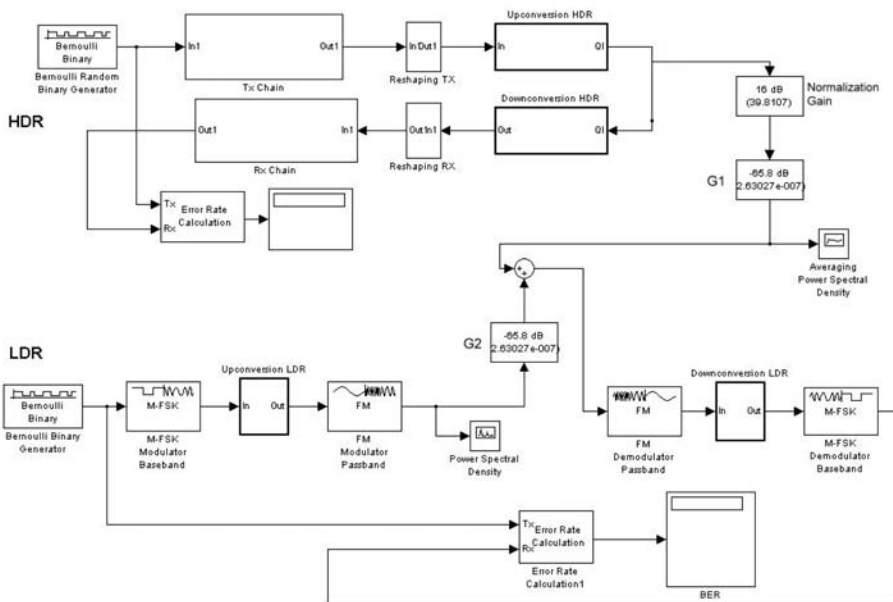


Figure 8.18 SIMULINK-based multimode simulator [39].

to passband at 4.5 GHz; the MC-SS PHY layer requires a quadrature upconverter and a quadrature downconverter in order to translate the center frequency f_c to 5.2 GHz, where $x_I(t)$ is the in-phase component of the bandpass signal and $x_Q(t)$ is the quadrature component of the bandpass signal. This is shown in Figure 8.19.

The proposed PHY layer schemes for LDR and HDR WPAN are subject to mutual out-of-band interference when closely located devices (e.g., dual-mode devices) are considered.

The wideband structure of the one transmission scheme and the relatively close respective center frequency leads to a situation such that the out-of-band emission of the interferer device is added to the desired signal of the receiving device, resulting in a degradation of link performance.

In the simulation model, the RF signals coming from the HDR transmitter and LDR transmitter are added to interfere with each other. The path loss between an LDR transmitter and LDR receiver was initially assumed fixed, whereas the path loss between an HDR transmitter and the LDR receiver varied. Simulations assumed different values of the distance receiver to interferer and the results were obtained with path losses higher than 66.8 dB (distances higher than 10m). Distances below 10m lead to completely corrupted reception, which makes it impossible for two devices (HDR and LDR) to coexist in the same room.

Figure 8.20 shows the BER curves for two distances of the LDR link (path losses of 66.8 and 56.8 dB, respectively, 10m and 5m).

In this situation the PER (calculated assuming that the minimum size is 64 octets) is very high. This is shown in Figure 8.21.

Use of the IAWA mechanism will lower the data rate but there will be no packet corruption. IAWA allows for choosing between different duty cycles according to the beacon order and the number of HDR superframes considered. The data rate achievable with IAWA is shown in Table 8.4. Considering LDR BO = 2 and a number of 3 HDR SFs there is a data rate drop from 100 to 33 Kbps. A performance comparison with and without the IAWA mechanism is shown in Table 8.5.

The collaborative coexistence mechanism, which is applied to HDR and LDR WPAN interfaces, controls and synchronizes access to the network. The synchronization of the two superframe sequences frees all LDR and HDR devices associated

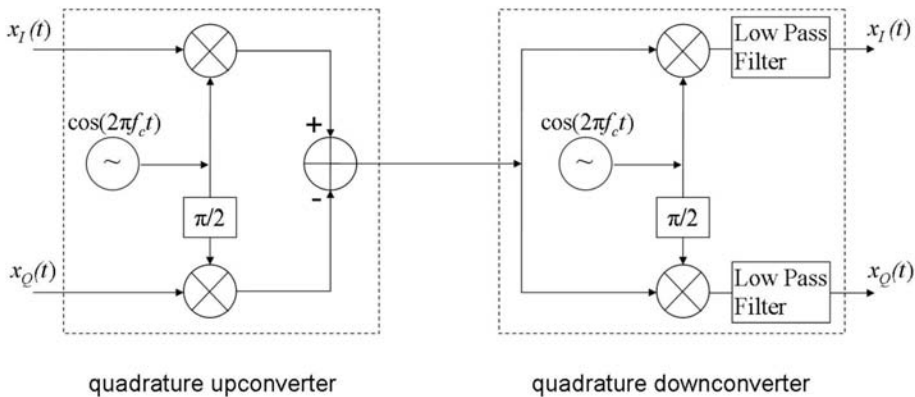


Figure 8.19 Quadrature upconverter and downconverter [39].

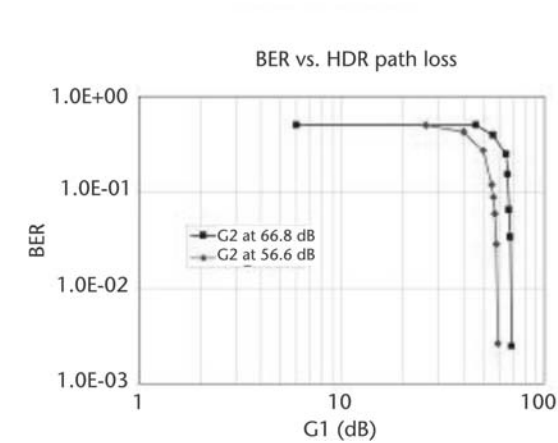


Figure 8.20 BER versus path loss for HDR [39].

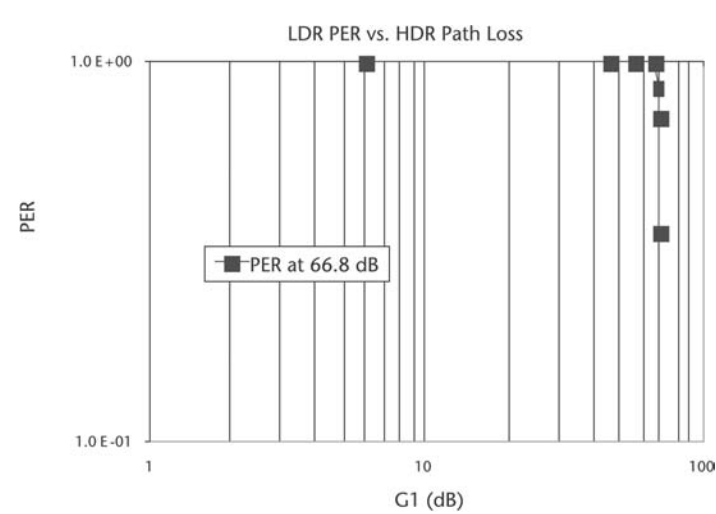


Figure 8.21 PER versus path loss for HDR [39].

Table 8.4 Data Rate Achievable with IAWA [43]

Beacon Order	LDR Duty Cycle	HDR Duty Cycle	LDR Source Data Rate (bps)	HDR Source Data Rate (bps)
2	33.33%	66.67%	33,330	57,749,554
3	14.29%	85.71%	14,290	74,242,002
4	6.67%	93.33%	6,670	80,842,446
5	3.23%	96.77%	3,230	83,822,174

with a common LDR and HDR PAN controller from interference. This mechanism decreases the duty cycle of the two networks; however, it does allow for the coexistence of an LDR and an HDR WPAN, even when the two PHY transmission schemes are overlapping in the frequency band and the devices are closely positioned.

Table 8.5 Performance Comparison with G2 Equal to 66.8 dB [39]

<i>G1</i> (dB)	<i>LDR Source</i> <i>Data Rate</i> <i>w/o IAWA</i> (bps)	<i>LDR Source</i> <i>Data Rate</i> <i>with IAWA</i> (bps)	<i>HDR Source</i> <i>Data Rate</i> <i>w/o IAWA</i> (bps)	<i>HDR Source</i> <i>Data Rate</i> <i>with IAWA</i> (bps)	<i>BER w/o</i> <i>IAWA</i>	<i>BER with</i> <i>IAWA</i>	<i>PER w/o</i> <i>IAWA</i>	<i>PER with</i> <i>IAWA</i>	<i>Throughput</i> <i>w/o IAWA</i>	<i>Throughput</i> <i>with IAWA</i>
6	100,000	33,330	86,620,000	57,749,554	0.5	0	1	0	0	33,330
26	100,000	33,330	86,620,000	57,749,554	0.5	0	1	0	0	33,330
46	100,000	33,330	86,620,000	57,749,554	0.5	0	1	0	0	33,330
56	100,000	33,330	86,620,000	57,749,554	0.4	0	1	0	0	33,330
65	100,000	33,330	86,620,000	57,749,554	0.25	0	1	0	0	33,330
66	100,000	33,330	86,620,000	57,749,554	0.11	0	1	0	0	33,330
67	100,000	33,330	86,620,000	57,749,554	0.065	0	1	0	1.11022E-10	33,330
68	100,000	33,330	86,620,000	57,749,554	0.025	0	0.999997	0	0.234619	33,330
69	100,000	33,330	86,620,000	57,749,554	0.0025	0	0.722407	0	27759.20	33,330
70	100,000	33,330	86,620,000	57,749,554	1.00E-05	0	0.005106	0	99489.30	33,330

8.5 Conclusions

Radio resource management is a vast topic and many strategies exist in support of situations that might occur in different scenarios. This chapter presented some of the major issues related to the management of resources for the lower layers of various radio systems. The goal was to provide an understanding of how resource management should be addressed and what the most commonly used strategies are, as well as some possible scenarios that might arise.

RRM is the combination of mechanisms implemented in the radio network entities to optimize the use of radio resources and maximize system capacity.

With the involvement of multiple radio access technologies and multiple cell layers, these mechanisms can be extended for radio resource sharing and spectrum management.

Even though a number of resources can be managed locally, that is, without involving cooperation among distinct systems, it is necessary to communicate information related to the management of the resources and provide for dynamic control of mechanisms and parameters in a multisystem environment. A large number of radio protocols and network algorithms were studied and proposed for realizing this necessity.

Collaborative coexistence mechanisms enable the use of multiple radio interfaces at limited levels of interference by smart scheduling of the channel access. These mechanisms are employed on the MAC and higher layers and require control information exchange among active peers.

With the evolution of mobile and wireless systems, and the demand for higher data rates, completely new system level implementations might be required that can change the way protocols are organized today. On the other hand, protocols do not disappear, because instances in different points of the system have to communicate, and headings for protocol control are still multiplexed with data, although they are more flexibly included in the data stream.

References

- [1] "Framework and Overall Objectives of the Future Development of IMT 2000 and Systems Beyond IMT 2000," Recommendation ITU-R M.1645, www.itu.int.
- [2] "Long-Term Evolution," www.3gpp.org/Highlights/LTE/LTE.htm.
- [3] Third Generation Partnership Project, www.3gpp.org.
- [4] Universal Mobile Telecommunication System Forum, www.ums-forum.org.
- [5] FP6 Overview, <http://cordis.europa.eu/ist/so/mobile-wireless/home.html>.
- [6] FP6 IST Projects Wireless Interface New Radio, WINNER and WINNER II, www.ist-winner.org.
- [7] International Telecommunications Union (ITU), www.itu.int.
- [8] FP6 IST Project SURFACE, <http://cordis.europa.eu/ist/so/mobile-wireless/home.html>.
- [9] FP6 IST Projects MAGNET and MAGNET II, www.ist-magnet.org.
- [10] FP6 Project MEMBRANE, www.imperial.ac.uk/membrane.
- [11] FP6 IST Projects PULSERS and PULSERS II, www.pulsers.eu.

- [12] Zimmermann, H., "OSI Reference Model: The ISO Model of Architecture for Open Systems Interconnection," *IEEE Trans. on Communications*, Vol. 28, No. 4, April 1980, pp. 425–432.
- [13] FP6 IST Project SURFACE, "Design of the MultiUser Optimal TX and RX Strategy," Deliverable 4.2, June 2007, <http://cordis.europa.eu/ist/so/mobile-wireless/home.html>.
- [14] FP6 IST Project WINNER, "Assessment of Key Enhanced Radio Protocols," Deliverable 2.8, February 2005, www.ist-winner.org.
- [15] FP6 IST Project WINNER, "Description of Deployment Concepts for Future Radio Scenarios Integrating Different Relaying Technologies in a Cellular Infrastructure Including Definition, Assessment, and Performance Comparison of RAN Protocols for Relay-Based Systems," Deliverable 3.2, February 2005, www.ist-winner.org.
- [16] FP6 IST Project WINNER, "Initial System Concept Description," Deliverable 7.3, February 2005, www.ist-winner.org.
- [17] FP6 IST Project WINNER, "Assessment of Multiple Access Technologies," Deliverable D2.6, December 2004, www.ist-winner.org.
- [18] IEEE 802.16e, www.ieee802.org/16/tge.
- [19] IEEE Standard for Local and Metropolitan Area Networks Part 16, "Air Interface for Fixed Broadband Wireless Access Systems. Amendment 2: Medium Access Control Modifications and Additional Physical Layer Specifications for 2–11 GHz," January 2003, www.ieee802.org/16/tge.
- [20] Esseling, N., "Ein Relaiskonzept für das hochbitratige drahtlose lokale Netz HIPERLAN/2," Ph.D. Dissertation, Aachen University, Lehrstuhl für Kommunikationsnetze, July 2004, p. 307.
- [21] Esseling, N., B. Walke, and R. Pabst, "Performance Evaluation of a Fixed Relay Concept for Next Generation Wireless Systems," *Proc. of PIMRC 2004*, Barcelona, September 2004.
- [22] Third Generation Partnership Project, "TR 25.892: Feasibility Study for OFDM for UTRAN Enhancement," Version 1.0.0, February 2004.
- [23] Tse, D., and S. Hanly, "Multi-Access Fading Channels: Part I: Polymatroid Structure, Optimal Resource Allocation and Throughput Capacities," *IEEE Trans. on Information Theory*, Vol. 44, No. 7, November 1998, pp. 2796–2815.
- [24] Third Generation Partnership Project, Nortel, "Update of OFDM SI Simulation Methodology," R1-03-0224, 3GPP TSG-RAN-1 Meeting No. 31, Tokyo, February 2003.
- [25] Bonald, [AU: Author initials?], "A Score-Based Opportunistic Scheduler for Fading Radio Channels," *Proc. of European Wireless Conference*, Barcelona, February 2004.
- [26] Noll Barreto, A., and S. Furrer, "Adaptive Bit Loading for Wireless OFDM Systems," *Proc. PIMRC 2001*, San Diego, September 2001.
- [27] FP6 IST Project 4MORE, "Definition of Framework for Joint Layer Design and Optimization," Deliverable D2.1, June 2004, <http://4more.av.it.pt>.
- [28] Phan Huy, D. T., et al., "System Level Performance Evaluation of the MATRICE Air Interface," *Proc. of the Mobile IST Summit 2004*, Aveiro, Portugal, June 2004.
- [29] Hoshyar, R., et al., "Enhancement of MAC Design in MIMO-OFDM-Based HIPERMAN System to Accommodate Uplink Multi-User Transmission," *Proc. of the Mobile IST Summit 2004*, Aveiro, Portugal, June 2004.
- [30] Wang, Q., and M. A. Abu-Rgheff, "Cross-Layer Signalling for Next-Generation Wireless Systems," *Proc. of WCNC*, 2003.
- [31] Kim, B., "A Network Service Providing Wireless Channel Information for Adaptive Mobile Applications: Part I: Proposal," *Proc. of IEEE ICC*, 2001.
- [32] Chen, K., S. H. Shan, and K. Nahrstedt, "Cross-Layer Design for Data Accessibility in Mobile Ad Hoc Networks," *Wireless Personal Communications*, Vol. 21, No. 1, April 2002, pp. 49–76.
- [33] FP6 IST Project 4MORE, <http://4more.av.it.pt>.

- [34] FP6 IST Project 4MORE, “System Definition and Criteria to Select the Schemes to Be Integrated,” Deliverable 1.2, July 2004, <http://4more.av.it.pt>.
- [35] Liu, X., et al., “Opportunistic Transmission Scheduling with Resource-Sharing Constraints in Wireless Networks,” *IEEE J. on Selected Areas in Communications*, Vol. 19, No. 10, October 2001.
- [36] FP5 IST Project MATRICE, “Layer 2 and 3 Simulation Platform—Final Version,” Deliverable 4.4, May 2004, <http://cordis.europa.eu/ist/ka4>.
- [37] Prasad, R. (Ed.), *Towards the Wireless Information Society*, Vols. I and II, Norwood, MA: Artech House, 2005.
- [38] Avidor, D., J. Ling, and C. Papadias, “Jointly Opportunistic Beamforming and Scheduling (JOBS) for Downlink Packet Access,” *Proc. of ICC*, 2003.
- [39] FP6 IST Project MAGNET Beyond, “Cross Optimization of the FM-UWB and MC-SS Air Interface Solutions,” Deliverable 3.3A, June 2008, www.ist-magnet.org.
- [40] De Sanctis, M., J. Gerrits, and J. Perez Vila, “Coexistence Concept for the Implementation of LDR/HDR WPAN Multimode Devices,” *Telettronikk Journal* (by Telenor), Special Issue on Personal Networks, 1st Quarter 2007, pp. 101–112.
- [41] FP6 IST Project MAGNET, “MAC/RRM Schemes for WPAN,” Deliverable 3.3.2a, October 2004, www.ist-magnet.org.
- [42] FP6 IST Project MAGNET Beyond, “Coexistence Concept for the Implementation of the FM-UWB and MC-SS RA Solutions,” Deliverable 3.3.1, June 2006, www.ist-magnet.org.
- [43] “Coexistence of Wireless Personal Area Networks with Other Wireless Devices Operating in Unlicensed Frequency Bands,” IEEE Standard 802.15.2, August 2003.
- [44] “Standard for Telecommunications and Information Exchange Between Systems Local Area Metropolitan Area Networks Specific Requirements Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for High Rate Wireless Personal Area Networks (WPAN),” IEEE 802.15.3-2003, www.ieee802.org.
- [45] “Standard for Telecommunications and Information Exchange Between Systems Local Area Metropolitan Area Networks Specific Requirements Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low Rate Wireless Personal Area Networks (WPAN),” IEEE 802.15.4-2003, www.ieee802.org.

Conclusions

The radio systems already being deployed are mainly evolutionary technologies (e.g., the cellular mobile systems such as GSM/GPRS, CDMA2000, UMTS, TD-CDMA, HSDPA/HSUPA, and EV-DO). R&D efforts, including standardization, are focused on both the evolutionary and revolutionary approaches (e.g., 3GPP LTE and LTE-Advanced), the evolution of 3GPP2 EV-DO (Revision C or AIE), IEEE 802.16e and 802.16m Mobile WiMax, IEEE 802.20, and proprietary systems such as Flash-OFDM. Much effort has been put into standards (e.g., IEEE 802.15) that allow for efficient use of ICT technologies in our economic and social lives.

Research results from many of the FP6 IST projects feed into the development of some of these technologies (e.g., 3GPP LTE, IMT-Advanced, IEEE 802.15, and IEEE 802.16). This book described the particular achievements related to the provision of spectrum-efficient, adaptive, and flexible radio interface technologies.

The realization of spectrally efficient radio interfaces can be achieved by use of technologies such as MIMO techniques, and in combination with adaptive modulation and coding, cognitive radio technologies, novel ways of channel and propagation modeling, interference modeling, and so forth. Enhancing the performance of radio interface technology can help to reduce spectrum demand, while accommodating more users. The spectral efficiency of MIMO transmission can be significantly increased if some level of channel state information is available at the transmitter, allowing the system to effectively adapt to the radio channel and take full advantage of the available spectrum. The main challenge is to make the CSI available at the transmitter (CSIT). This can be achieved by conveying CSI as feedback information over the reverse link or requesting the CSI at the transmitter during receiving periods.

Adaptive modulation and coding techniques together with MIMO are the principal radio techniques to be considered when developing future radio systems. LDPC and duo-binary turbo codes were studied in adaptive coding and modulation schemes for next generation cellular systems. Such systems show only small performance differences when compared to schemes such as convolutional or BLDP codes, but can be competing FEC schemes for larger block lengths.

The choice of an optimal multiple-access scheme or a combination of schemes is very dependent on the performance requirements set up for the radio interface system. Adaptation to channel states, user requirements, and changing traffic patterns are only some of the considerations for the final choice. Other important considerations are that the radio interface system under design be able to interwork successfully with already existing or yet to be developed systems.

Multiple antennas can improve the single-user and overall cell throughput considerably. The challenge is to select the right multiple-antenna approach for

each traffic/channel/interference situation. A trade-off of multilayer, diversity, beamforming, and multiuser diversity gains has to be found while keeping complexity and feedback overhead low. Multiuser MIMO techniques in the context of cellular networks comprise BSs that employ antenna arrays and MTs with possibly multiple antenna elements. The role of multiple-antenna techniques in multiuser environments is essentially to multiplex users and data streams in the spatial domain by taking advantage of all of the degrees of freedom offered by multiple-antenna processing. To enable ubiquitous broadband wireless access, MIMO transmission must be made robust against multicell interference. Realization of the potential capacity gains of MIMO under such conditions have not yet been demonstrated.

A set of innovative concepts for advanced antenna schemes was proposed by a few FP6 IST projects in the context of conventional cellular networks. The proposed concepts included several detailed SDMA signal processing solutions for improving the capacity of cellular and mesh networks employing multiple-antenna BSs and MTs.

Multihop communication systems can be realized by use of relay technologies to reduce the infrastructure costs and increase the coverage of radio systems. Cost-effective coverage has to ensure QoS in terms of delay and throughput guarantees. The user experience has to be ubiquitous, regardless of whether it is close to the base station or at the cell edge. A special challenge is to provide coverage at higher frequencies with an existing infrastructure. One means to improve coverage is relaying. With relaying, however, issues such as MAC design, RRM, interference, routing, and power saving have to be addressed.

Intelligent scheduling, fractional-frequency reuse, interference avoidance, exploitation of frequency diversity, and interference cancellation with single or multiple antennas are all potential means of addressing interference. The SIR of interference-limited radio deployments can be addressed by the selection of multiple antennas, adaptive modulation, and interference cancellation methods. ARQ and relaying are potential methods for achieving high performance in interference-limited scenarios.

Next generation systems based on OFDMA depend on efficient radio resource management mechanisms to sustain a high level of system performance. The complexity of the system, encompassing time, frequency, and space dimensions, and the increasing need to reduce operational expenditures require further research into efficient optimization and novel resource management algorithms.

From an optimization point of view, scheduling at the MAC layer is a multi-dimensional problem. The resource dimensions available for scheduling are time, subcarriers, modulation, coding rate, power allocation, and a spatial channel (with MIMO in spatial multiplexing mode). Different aspects have to be considered for resource scheduling at the MAC layer. Many techniques can be applied but the choice will depend on the context of the scheduling problem.

In the context of multimode terminals, it is important to establish efficient means for coexistence that guarantee successful access to resources. Cross-layer approaches can proven quite efficient in this context.

After cross-layer design, the possibility of globally allowing relationships between any modules of the protocol design appears natural and leads to the possible introduction of nonlayered (layerless) architectures.

About the Editors

Ramjee Prasad received a B.Sc. in engineering from the Bihar Institute of Technology, Sindri, India, in 1968, and an M.Sc. and a Ph.D. from Birla Institute of Technology (BIT), Ranchi, India, in 1970 and 1979, respectively.

Dr. Prasad has a long path of achievement and rich experience in the academic, managerial, research, and business spheres of the mobile communication areas. He joined BIT as a senior research fellow in 1970 and became an associate professor in 1980. While with BIT, he supervised a number of research projects in the areas of microwave and plasma engineering. From 1983 to 1988, he was with the University of Dar es Salaam (UDSM), Tanzania, where he became a professor of telecommunications in the Department of Electrical Engineering in 1986. At UDSM, he was responsible for the collaborative project Satellite Communications for Rural Zones with Eindhoven University of Technology, the Netherlands. From February 1988 through May 1999, he was with the Telecommunications and Traffic Control Systems Group at Delft University of Technology (DUT), where he was actively involved in the area of wireless personal and multimedia communications (WPMC). He was the founding head and program director of the Centre for Wireless and Personal Communications (CWPC) of International Research Centre for Telecommunications and Radar (IRCTR).

Since June 1999, Dr. Prasad has held the chair of Wireless Information and Multimedia Communications at Aalborg University, Denmark (AAU). He was also the codirector of AAU's Center for PersonKommunikation until January 2004, when he became the founding director of the Center for TeleInfrastruktur (CTIF), established as a large multiarea research center on the premises of Aalborg University.

Dr. Prasad is a worldwide established scientist, which is evident from his many international academic, industrial, and governmental awards and distinctions, more than 25 published books, numerous journal and conference publications, a sizeable amount of graduated Ph.D. students, and an even larger amount of graduated M.Sc. students. Under his initiative, international M.Sc. programs were started with the Birla Institute of Technology in India, the Insititute of Technology Bandung in Indonesia. Recently, a cooperation was established with the Athens Information Technology (AIT) in Greece.

Under Dr. Prasad's successful leadership and extraordinary vision, CTIF currently has more than 150 scientists from different parts of the world and three CTIF branches in other countries: CTIF-Italy (inaugurated in 2006 in Rome), CTIF-India (inaugurated on December 7, 2007, in Kolkata), and CTIF-Japan (inaugurated on October 3, 2008).

Dr. Prasad was a business delegate in the Official Business Delegation led by Her Majesty The Queen of Denmark Margarethe II to South Korea in October 2007. He is a Fellow of the IEE, a Fellow of the IETE, a senior member of the IEEE, and a member of NERG. He was the recipient of the Telenor Nordic Research Award (2005), the Samsung Electronics Advisor Award (2005), the Yearly Aalborg-European Achievements Award (2004), and the IEEE Communication Society Award for Achievements in the area of Personal, Wireless, and Mobile Systems and Networks (2003).

Dr. Prasad is a member of the steering, advisory, and program committees of many IEEE international conferences. He was the founding chairman of the European Centre of Excellence in Telecommunications, known as HERMES, and is now an honorary chair of HERMES. HERMES currently has 10 member organizations from Europe. Dr. Prasad is the founding cochair of the International Symposium on Wireless Personal Multimedia Communications (WPMC), which has taken place annually since 1999.

Dr. Prasad has been strongly involved in European research programs, such as the FP4-ACTS project FRAMES (Future Radio Wideband Multiple Access Systems), which set up the UMTS standard, as a DUT project leader. He was a project coordinator of EU projects during FP5 (CELLO, PRODEMIS) and FP6 (MAGNET and MAGNET Beyond), and is currently involved in FP7. Dr. Prasad was the project leader for several international industrially funded projects with NOKIA, SAMSUNG, Ericsson Telebit, and SIEMENS, to name a few. He is a technical advisor to many industrial international companies, is the founder of the IEEE Symposium on Communications and Vehicular Technoliógy (SCVT) in Benelux, and was the chairman of SCVT in 1993.

Dr. Prasad is the founding editor-in-chief of the Springer International Journal on Wireless Personal Communications. He is a member of the editorial board of other international journals and is the series editor of the Artech House Universal Personal Communications Series.

Albena Mihovska completed a B.Sc. in engineering at the Technical University of Sofia, Bulgaria, in 1990, followed by an M.Sc. in engineering at the Technical University of Delft, the Netherlands, in 1999. Since then, Professor Mihovska has been with Aalborg University, Denmark, where she is currently an associate professor at the Center for TeleInfrastruktur (CTIF).

During her years of employment at Aalborg University, Professor Mihovska gained extensive experience in the administrative and technical management of EU-funded research projects. She also gained experience initializing industrial research cooperation as well as research cooperation funded by the EU. She joined Aalborg University as a research engineer in July 1999 and was appointed to the European Union-funded technical management team within the FP4 ACTS project ASAP until its successful completion in 2001. From September 2001 until April 2005, she was the project coordinator of the European Union-funded FP5 IST project PRODEMIS as a special support action instrument until its successful completion. The project was a main supporting project of the EU IST projects within the mobile and satellite area. The outcome of the project was published as two books by Artech House in 2005, as well as in a number of technical research

publications in peer-reviewed journals and conferences, an e-conference on mobile communications, a joint workshop, and a technology roadmap for the future development of mobile communications.

From January 2004 until December 2005, Professor Mihovska was the research coordinator of the research team within the EU-funded IST FP6 project WINNER, which continued from January 2006 to December 2007 as WINNER II. The main objective of the project was the design of a new air interface that could be a competitive candidate for next generation systems, in the scope of standardization activities within the IMT-Advanced ITU group. The project was a part of the WWI initiative and, as such, had close and required cross-issue collaboration with the rest of the WWI projects. Professor Mihovska was part of the research teams working toward the identification of the system requirements and the design of interworking mechanisms between the newly designed system and other systems. Within the project, she proposed a concept for cooperation between different systems based on an autonomous decision framework. Based on this research idea, the theoretical approach was put forward as a development activity in the second stage of the project and was successfully demonstrated at a number of international events, including the Wireless Radio Communication (WRC) '07 Conference held in Geneva from October through November 2007. The experimental setup is now being considered for use in other projects, such as the CELTIC project WINNER+ and the FP7 project FUTON, working toward an architecture design for converging heterogeneous systems and service provisioning in which AAU is a consortium member. Professor Mihovska was part of the research group within the project consortium who developed the final system concept requirements for the air interface.

From September 2006 to March 2008, Professor Mihovska was the deputy technical manager of MAGNET Beyond. Therein, she contributed to the overall technical work progress and to the finalization of the MAGNET Beyond Platform system requirements. Further, she was involved in AAU-related research activities in the area of security for personal networks (PNs).

Since April 2008, Professor Mihovska has been involved in research activities within the CELTIC project WINNER+, working toward advanced radio system technologies for IMT-Advanced systems. She is conducting research activities within the area of advanced radio resource management, cross-layer optimization, and spectrum aggregation.

The work proposed in pursuit of her Ph.D. degree from Aalborg University is a novel concept for interworking between radio resource management entities in the context of next generation mobile communication systems. It is based on research activities commenced prior to and continued within the frames of the WINNER project. The concepts proposed within her Ph.D. thesis have been successfully implemented in the overall WINNER concept and have resulted in a number of peer-reviewed journal and conference publications, including the demonstration activities mentioned above. In addition, she has a number of project-related publications, presentations, and various international and EU events.

Professor Mihovska is a reviewer for IEEE Communication Letters and The Springer Journal of Telecommunication Systems. She has been part of the organizing

and TPC committees of a number of international conferences, such as WPMC 2002, WCNC 2007, VTC 2008 Spring, the IST Mobile Summits 2002-2007, ATSSMA-NAEC 2009, IEEE Mobile WiMAX 2009 Symposium, and several workshops.

Index

A

- ACE project, 17
- Adaptation, 179–80
- Adaptive beamforming, 346
- Adaptive modulation and coding (AMC), 2, 171
- Adaptive multiuser TDMA/OFDMA, 188–90
 - difficulties, 189
 - downlink, 189
 - environment, 189
 - opportunistic scheduling, 189
 - See also* TDMA/OFDMA
- Adaptive soft resource positioning, 304–10
 - flowchart, 307
 - power mask assignment, 308
 - scenario illustration, 306
 - SINR of received packets, 309
- Adaptive transmission techniques, 92
- Additive white Gaussian noise (AWGN), 83
 - BICM capacity, 210
 - calibration plot, 184
 - channel, 139, 208
 - interference, 340, 379
 - simulations, 210
- Adjacent channel interference, 337
- Air interfaces
 - FM-UWB, 419, 420
 - HDR, 380
 - MC-SS, 419, 420
 - multiple modes, 36
 - WINNER project, 5, 6
- Alternating wireless activity (AWA), 415, 417–19
 - as collaborative coexistence mechanism, 418
 - improved (IAWA), 419
- Analog-to-digital converter (ADC), 265
- Angle of arrival (AoA), 68, 87
- Angle of departure (AoD), 68, 87
- Angular spread (AS), 68
- Antenna arrays
 - construction of, 263–64
 - field pattern (FP), 264
 - linear, 252
 - modeling, 251–67
 - planar, 252
 - rectangular, 252
 - response matrices, 254–55
 - response vector, 255
 - special two-dimensional structures, 261–63
 - spherical, 253
 - three-dimensional, 255–59
 - two-dimensional, 259–61
- Antenna hopping, 226
- Antenna independent channel models, 81
- Antenna models
 - characteristics, 254–64
 - double directional, 252
 - overview, 252–54
- Antennas
 - cell structure, 255
 - differential, 271–73
 - multiple, 429–30
 - smart, 217–76
 - steering vector, 255
 - topology, 255
- Antenna steering vector, 89
- Antenna structures (UWB systems), 267–75

- A posteriori probability (APP), 158, 159
- Array coordinate system (ACS)
 - primary, 258
 - secondary, 258
 - two-dimensional, 261
- Array gain, 220
- Automatic repeat request (ARQ), 153
- Average packet delay per user, 402–3, 405
- Average service throughput per sector, 402, 404
- Average service throughput per user, 402, 404
- Average user throughput, 305, 354
- Averaging, 179, 180
- B**
- Backhaul, 285–86
- Bahl, Cocke, Jelinek, and Raviv (BCJR)
 - algorithm, 103
- BCH codes, 103
- Beamforming, 221
 - adaptive, 346
 - downlink, 234–47
 - fixed, 346
 - grid-of-beams, 228
 - for ICI, 345–46
 - opportunistic, 226
 - resource management and, 361
 - scaled MR (SMR), 236
- Bell shape, 69
- Bell with spike shape, 69
- Bit and power allocations, 206
- Bit and power loading, 205
- Block codes, 103
- Block error rate (BLER), 149
- Block OFDMA (BOFDMA), 192
- Body-to-body channel gain modeling, 362–69
- Broadband RF module, 265–67
- Burst error structures, 90
- C**
- Capacity, 8
- Capital expenditures (CAPEX), 289
- Carrier interference noise ratio (CINR), 284
- Carrier interferometry, 42
- Carrier-to-interference ratio (C/I), 342
- CDMA2000, 14
- Cell capacity gains, 315–17
- Central cell technique, 349–50
- Channel characteristics
 - classification, 86–89
 - design considerations, 88
 - spatial features, 88
- Channel coding
 - advanced techniques, 153–70
 - algorithms, 102–16
 - with DBTCs, 154–62
 - defined, 101
 - introduction to, 102–16
 - redundancy, 102
- Channel estimation
 - in-band pilots, 49
 - MMSE MIMO, 131
 - pseudorandom postfix (PRP), 49
 - training sequences, 48
- Channel gain modeling, 362–69
 - autospectra of measurements, 367
 - measurement analysis, 363–69
 - measurements, 363
 - measurement site and locations, 364
 - percentiles of measurements, 367
- Channel modeling, 77–82
 - antenna interdependence, 81–82
 - channel characteristics classification, 86–89
 - interference, 86
 - ray-based, 78–80
 - spatial correlation matrix-based, 80–81
 - statistical, 77–78
 - validation, 91–92
- Channel models, 65–77
 - 3GPP spatial, 69–73
 - accuracy, 65
 - antenna independent, 81, 255
 - comparison, 75–77
 - evaluation, 77
 - general descriptions, 90–92
 - generic statistics, 77–78
 - geometric, 70
 - IEEE 802.11n, 67–69

- for link level simulations (advanced), 91
- for link level simulations (basic), 90
- mapping to propagation environment, 68
- METRA, 73–74
- MIMO, 67–77, 83
- requirements, 92
- for system level simulations (advanced), 91
- for system level simulations (basic), 91
- technical selection criteria, 65
- virtual, 84
- WINNER project, 254
- Channel transfer functions, 206
- Closed-loop SMUX, 224
- Closed-loop transmit diversity, 346
- Coarse synchronization, 30
- Cochannel interference (CCI), 337–38, 380
- Code division multiple-access (CDMA)
 - direct sequence (DS-CDMA), 44–45, 118
 - multicarrier (MC-CDMA), 44, 178, 185–87
 - wideband (WCDMA), 3
- Codes
 - BCH, 103
 - block, 103
 - convolutional, 103
 - defined, 103
 - design and performance, 111–15
 - error-correcting, 104
 - Golay, 103
 - Hamming, 103
 - LDPC, 103, 104
 - multiuser space-time, 143–53
 - QC-BLDPC, 153
 - quadratic residue, 103
 - Reed-Muller, 103
 - Reed-Solomon (RS), 103
 - space-time block (STBC), 124, 149
 - summary, 115–16
 - turbo, 104
- Codeword error rate (CWER), 159
 - approximated curves, 161
 - coarse approximation, 160
 - curves for 16-QAM, 160
 - curves for 256-QAM, 161
 - curves for QPSK, 160, 164
 - performance, 165, 169
- Coding algorithms, 102–16
- Coexistence, 32
- Cognitive radio, 33
- Collaborative coexistence mechanisms, 425
- Competing techniques, 229
- Complementary techniques, 229
- Contention
 - bands, 195
 - defined, 180
 - random access, 396
- Continuous-time delay channels, 369
- Continuous-time Markov chain, 369
- Continuous wave (CW) interference, 337–38
- Convolutional codes
 - advantages, 103–4
 - defined, 103
 - low-rate, 166–70
 - MFD, 167
- Cooperative communication, 283, 330
- Cooperative diversity, 380
- Cooperative relaying, 298–99, 312–26
 - capacity, 321
 - cooperative amplify and forward, 313
 - cooperative decode and forward, 313–26
 - defined, 312
 - deployment in downlink, 319
 - impact in DL, 321
 - protocols, 312–13
- COST 231 Walfisch-Ikegami model, 71
- COST 259 model, 75, 84
- Coverage
 - area, 8, 297
 - relay-based deployments, 295, 297
 - ubiquitous, 11
 - wide-area, 189
- Coverage enhancement, 281–331
 - introduction, 281–99
 - multihop communications, 283–88
 - relay technologies, 288–99

- Cross-layer information (CLI), 409
 - interlayer signaling and transport, 410
 - internode distribution, 411
 - selection of parameters, 411
- Cross-layer optimization
 - challenges, 407
 - design, 407–14
 - information and signaling, 409–11
 - PHY layer approaches, 408
 - scheduling, 411–14
 - signaling architecture, 411
- Cross-optimization, 334
 - of FM-UWB, 414–24
 - of MC-SS, 414–24
- Cross-polarization discrimination (XPD), 68
- CSI at transmitter (CSIT), 429
- Cyclic delay diversity (CDD), 360
- Cyclic prefix (CP), 29
- Cyclic prefix OFDM (CP-OFDM), 121, 123
 - guard interval, 124
 - receiver complexity, 124
- Cyclic redundancy check (CRC), 401
- D**
- Decision feedback equalizers (DFE), 120, 141
- Decode and forward, 313–26
 - application to downlink, 318–26
 - cell capacity gains, 315–17
 - geometry-dependent capacity gains, 317–18
 - with repetition coding (DF-RC), 314
 - with unconstrained coding (DF-UC), 314–15
- Delay
 - distribution, 352
 - packet, 353
 - packet call, 353
 - spread, 89
- Demodulation, 101
- Dense multipath model, 80
- Density evolution, 105
- Detection and avoidance (DAA)
 - techniques, 386
- Determinant criterion, 150
- Differential antennas, 271–73
 - impedance matching, 273
 - photographs, 272
 - See also* Antennas
- Differential binary phase shift keying (DBPSK), 269, 270
- Digital signal processing (DSP), 33
- Digital-to-analog converter (DAC), 265
- Digital video broadcasting–terrestrial (DVB-T), 34
- Direct digital synthesizers (DDS), 17
- Direct sequence CDMA (DS-CDMA), 44–45
 - channel response, 181
 - MMSE in, 44
- Discrete Fourier transform (DFT), 37, 116–17
 - coefficients, 42
 - common block, 42
 - of data symbol sequence, 42
 - matrix, 116–17
 - of noise, 41
 - size, 38
- Discrete time system model, 133
- Diversity
 - cooperative, 380
 - cyclic delay (CDD), 360
 - gain, 220–21
 - order, 220
 - receive, 222, 346
 - SMUX and, 224–25
 - techniques, 222–27
 - transmit, 222–23, 246
- Dominant interferer, 342
- Doppler frequency, 69
- Doppler spread, 120
- Double directional modeling, 252
- Downlink
 - beamforming, 234–47
 - capacity enhancement, 247–51
 - resource blocks, 12–13
 - system control data, 9–10
 - WCDMA transmission, 16
- Duo-binary PCCC, 153
- Duo-binary turbo codes (DBTCs), 116
 - component code, 155, 156
 - decoder, 158, 159

- defined, 155
- encoder implementation for, 155
- enhancements, 154–55
- modulation and coding schemes with, 154–62
- puncturing patterns, 156
- rate compatible, 155, 156
- recursive codes, 155
- tail-biting, 157
- truncation, 157
- zero-termination, 157
- Dynamic channel allocation (DCA), 347
- Dynamic interleaving, 382, 383, 385
 - gain, 385
 - process, 382, 383
- Dynamic resource allocation (DRA), 412, 413
- Dynamic spectrum allocation (DSA), 34
- E**
- Edge interleaver, 110
- Effective aperture distribution function (EADF), 88
- Effective isotropic radiated power (EIRP), 234
 - average constraint, 237
 - average restriction, 244
 - constraint for OFDM system, 235
 - optimization, 243
 - scaled beamformer, 237, 243
 - spectral density constraint, 245, 246
- Eigenvalue criterion, 147
- Empirical models, 84
- Enabling technologies, 13
- Equal gain combining (EGC), 112
- Equalization techniques, 126–31
 - assessment results, 131
 - MIMO-PRP-OFDM, 129–30
 - overlap-add (OLA), 126
 - PRP-OFDM, 128–29
 - PRP-OFDM postfix sequence, 130–31
 - pseudo-circulant channel matrices/
 - equalization, 126–28
 - ZP-OFDM, 126
- Equivalent isotropic radiated power (EIRP) constraints, 16
- Error-correcting codes, 104

- Estimation techniques, 12–13
- European-funded research, 15–17
- Exhaustive round-robin (EXRR), 396, 397
- EXIT functions, 152, 153
- Exponential effective SIR mapping (EESM), 308
- F**
- Fair scheduling algorithms, 403
- Fast base station switching (FBSS), 284
- Fast Fourier transform (FFT)
 - FMT systems and, 134–35
 - module, 16
- Feedforward encoder, 167
- Filterbank modulation, 116
- Filtered multitone (FMT), 131–35
 - digital implementation, 135
 - FFT and, 134–35
 - robustness, 140
 - spectral efficiency, 134
 - spectrum shape, 134
- Fincke-Pohst algorithm, 225
- Fine synchronization, 30–31
- Finite state machine (FSM) model
 - continuous-time Markov chain, 369
 - illustrated, 370
 - starting state, 376
- FIREWORKS project, 7
 - defined, 16
 - relay enhanced OFDMA, 333
- FITNESS project, 75
- Fixed beamforming, 346
- Fixed cooperative relays, 324
- Fixed radio coexistence analysis, 336–40
 - channel arrangements, 336
 - interference sensitivity, 337–38
 - spectrum density masks, 337
 - transmitting power ranges, 336–37
- Fixed relay nodes (FRNs)
 - MT connection link, 302
 - optimal site, 303
 - positioning, 302, 303, 304
 - two-hop, 300
- Fixed relay stations (FRS), 282
- Flows, 392–93

FM-UWB air interface, 419
 characteristics, 420
 high-band RF IC, 17
 low-band IC, 17
 4MORE project
 cross-layer optimizations, 389
 defined, 16
 system-on-chip (SoC), 410
 Frame descriptor tables (FDTs)
 alternating between, 397
 concept, 397, 398
 function, 397
 Frequency assignment and scheduling, 117
 Frequency-division duplex (FDD)
 downlinks, 199, 200
 maximal prediction horizons, 203
 prediction performance, 198–214
 TDMA/OFDMA, 196
 uplinks, 199
 Frequency-division multiple-access (FDMA), 31
 interleaved, 178
 signals, 141, 143
 single-carrier (SC-FDMA), 178
 Frequency-domain orthogonal signature sequences (FDOSS), 44–45, 48
 Frequency-domain receiver, 40
 Frequency-domain transmitter, 39
 Frequency hopping, 344
 Frequency interleaving, 48
 Frequency offset, 46–48
 OFDMA sensitivity properties, 46
 result, 46
 for serial modulation, 46
 Full-buffer traffic model, 352
G
 Gain
 array, 220
 diversity, 220–21
 dynamic interleaving, 385
 interference suppression, 221
 SNR, 239, 240
 spatial multiplexing, 221
 Gaussian wideband interferers, 377
 General Packet Radio Service (GPRS), 5

Geometry-dependent capacity gains, 317–18
 Global-coordinate-system (GCS), 255
 Golay code, 103
 Grid of beams (GoB), 346, 360
 Grouped subcarriers, 187
H
 Hamming codes, 103
 Hard-frequency reuse, 310
 Hard handover, 284
 Hard switching, 225
 High-data-rate (HDR) systems, 219
 BER versus path loss, 423
 MAGNET, in-band spectral masks, 375
 MAGNET, simulation bandwidth, 374
 PER versus path loss, 423
 signal transmitting power, 414
 very (VHDR), 221
 WPAN radio standards, 416
 High-speed downlink packet access (HSDPA), 2, 3
 Hybrid ARQ (HARQ), 2
 process, 413
 process number, 401
I
 ICI modeling, 348–61
 evaluation, 351–61
 multicell simulation setup, 349–51
 See also Intercell interference (ICI)
 IEEE 802.x standards, 2
 IEEE 802.11, 4
 control frames, 231
 standardization body, 18
 IEEE 802.11a, 234
 IEEE 802.11g, 249
 IEEE 802.11n
 channel model, 67–69, 76
 goal of, 17–18
 model suitability, 77
 IEEE 802.11s, 286–88
 IEEE 802.16
 evolution paths, 14
 mesh mode, 283
 IEEE 802.16e, 283–86
 Impedance matching, 273

- Improved AWA (IAWA), 419
- Impulse radio-ultrawide band (IR-UWB), 268
- IMT-2000, 2
 - environments, 26
 - mobile WiMAX in, 3–4
- In-band pilots, 49
- Intelligent deployment, 298
- Intelligent scheduling, 430
- Inter-cell interference cancellation (IIC), 358
 - ideal, 358
 - schemes, 359
- Inter-cell interference (ICI)
 - averaging, 344
 - beamforming and, 345–46
 - effects, minimizing, 326
 - evaluation, 351–61
 - time dispersion effects, 209
 - See also* ICI modeling
- Interference
 - adjacent channel, 337
 - averaging, 344–45
 - avoidance, 344
 - as AWGN, 340
 - cancellation, 344–45, 386
 - cochannel, 337–38
 - colored, 334
 - continuous wave (CW), 337–38
 - defined, 334
 - inter-PAN, 335
 - intracell, 334
 - intra-PAN, 335
 - limit, 339
 - modeling, 86, 342–43
 - as noise, 340
 - power, 341, 343
 - sensitivity, 337–38
 - severity assessment, 335–40
 - study of, 333
 - at victim receiver's antenna, 340
 - white, 334
 - See also* Inter-cell interference (ICI)
- Interference area (IA), 336
- Interference diversity technique, 382
- Interference management, 333–86
 - based on signal and interference levels, 341
 - basis, 333
 - detection and avoidance (DAA)
 - techniques, 386
 - introduction, 334–47
 - PAN-optimized scenarios, 361–84
 - relays in, 334
 - study of, 333
 - technology, 333
- Interference mitigation, 343–47
 - averaging, 344–45
 - avoidance, 344
 - with dynamic interleaving codes, 382–84
 - methods, 343
 - results, 348
 - smart antenna-based, 345–47
- Interference multipath-channel margin (IMCM), 379
- Interference rejection combining (IRC), 346, 359
- Interference scenario occurrence
 - probability (ISOP), 335
- Interferer temporal properties, 369–77
 - frame transmission time, 372–73
 - initialization, 370, 371, 374
 - input, 370, 371, 374
 - internal state, 370, 371, 374
 - model parameters, 373
 - output, 370, 371, 374
- Interleaved FDMA, 178
- Interleaved OFDMA (IOFDMA), 192
- Interleaved spread subcarriers, 187
- Interleaving
 - dynamic, 382, 383, 385
 - partitioning preservation, 382
 - spreading maximization, 382
 - symbol/data subcarrier, 384
 - time-invariant, 383
- Intermediate frequency (IF), 265
- International Mobile Telecommunications-Advanced (IMT-A), 2, 3
 - “always on” experience, 4
 - global research targets, 18
 - minimum performance capabilities, 27
 - system design and performance, 26
 - technology candidates, 14

- International Telecommunication Union (ITU-R), 1, 14
- Internet Control Message Protocol (ICMP) messages, 410
- Inter-PAN interference, 335
- Intersymbol interference (ISI), 28
 - between adjacent user spectra, 42
 - channel matrix, 123
 - suppression, 120
 - time dispersion effects, 209
- Interuser interface, 48
- Intracell interference, 334
- Intra-PAN interference, 335
- Inverse discrete Fourier transform (IDFT), 116, 117
- IOTA modulation, 116
- Iterative equalization, 141
- J**
- Joint detection, 326
- K**
- Kalman predictor, 198
- K -factor, 89
- Kronecker product model, 75, 80
- L**
- LDPC decoders, 108–10
 - average number of iterations, 109
 - connection grid, 110
 - disadvantages, 110
 - implementation complexity, 110
 - maximum number of iterations, 109
 - See also* Low-density, parity-check (LDPC) codes
- Lifting method, 165
- Linear antenna arrays, 252
- Link adaptation, 13, 193–97, 406–7
 - channel performance improvement, 406
 - CQI and, 407
 - defined, 406
 - procedure, 193–94
 - purpose, 406
- Link budget analysis, 377–79
- Link level simulations, 82
 - advanced channel model, 91
 - basic channel model, 90
- Local profiles, 410
- Log-likelihood ratios (LLRs), 113
- Long Term Evolution (LTE) initiative, 2
- Low-data-rate (LDR) systems
 - antennas extended from, 219
 - MAGNET, superframe, 419
 - superframe structure, 420
 - WPAN, 416
- Low-density, parity-check (LDPC) codes, 103
 - advantages, 115
 - decoding, 108–10
 - design and performance, 112–15
 - encoding by parity check matrix, 105–7
 - error correction, 104
 - geometric construction methods, 107
 - irregular, 112
 - LLRs, 113
 - MASCOT, 151–52
 - minimum distance, 105
 - multiplexed, 151
 - optimization, 152
 - performance, 105
 - QRD, 114–15
 - quasicyclic block (QC-BLDPC), 163–66
 - SNR mismatch impact, 167, 168
 - Tanner graph, 108, 109
 - turbo codes versus, 104–5
- Low-density parity check codes (LDPCC), 153
- Low-rate convolutional codes, 166–70
- Luneberg-Kline expansion, 78
- M**
- MAC layer, 9, 390
 - design in support of DRA, 413
 - QoS guarantee, 415
 - scheduler role in, 400
 - scheduling algorithm, 407
 - superframe, 416–17
- MAC packet data unit (MPDU), 247
- MAC protocols
 - access control, 394
 - channel arrangement, 394
 - design challenges, 393–98

- FDT, 398
- frame structure, 396
- role in QoS, 393
- WINNER project, 394–95
- Macrodiversity handover (MDHO), 284
 - data transmission in, 285
 - diversity set of BSs, 284
- MAGNET Beyond project, 361
- MAGNET II project, 333
- MAGNET project
 - defined, 17
 - HDR air interface, 380
 - HDR in-band spectral masks, 375
 - HDR simulation bandwidth, 374
 - interference, 333
 - LDR MAC, 416
 - LDR superframe, 419
 - LDR WPAN, 416
 - radio channel models, 361
 - simulation chain, 361
 - temporal correlation function modeling, 366
- Market
 - requirements, 13–14
 - services, 14
- MASCOT project, 101
 - analysis, 143
 - LDPC codes, 151–52
 - STBCs with downlink transmission schemes, 149
 - terminal support, 151
- Matched filter (MF), 136
- Maximal ratio combining (MRC), 112
- Maximum free distance (MFD)
 - convolutional code, 167
- Maximum-likelihood (ML) decoding, 103
- Maximum ratio combining (MRC), 222, 234
- Maximum throughput scheduling, 403
- MC-SS air interface, 4, 419, 420
- Medium access control layer. *See* MAC layer
- MEMBRANE project
 - defined, 16–17
 - interference analysis, 333
- Mesh APs (MAPs), 287
- Mesh BS, 283
- Mesh network topology, 285
- Mesh points (MPs), 287
- Mesh SSs, 283
- Mesh WMANs, 283
- METRA channel model, 73–74
 - comparison, 76
 - evolution to environments, 74
 - PAS shapes, 74
 - scenario modeling, 74
 - as stochastic, 73
 - See also* Channel models
- METRA project, 73
- MIMO channels
 - characteristics, 65
 - estimation, 229
 - measurements, 89
 - models, 67–77, 83
 - spatial features, 88
 - sum capacity optimization, 226–28
 - vector, 227–28
 - See also* Multiple input/multiple output (MIMO)
- MIMO PRP-OFDM, 124–25
 - equalization techniques, 129–30
 - example, 125–26
 - illustrated, 124
 - simulation results, 132
 - with ZP-OFDM decoding, 131
 - See also* Pseudorandom postfix OFDM (PRP-OFDM)
- MIMO WLAN structure, 230
- Minimum mean squared error (MMSE), 44
- MMSE receivers, 181
 - block transmission, 182
 - continuous transmission, 182
 - simulation parameters, 183
- Mobile BSs (MBSs), 286
- Mobile SSs (MSSs), 286
- Mobile WiMAX, 2, 3, 4
- Mobile/wireless systems, 1–7
 - as economic growth driver, 1–2
 - flexibility, 25
 - next generation, 5–6, 24
 - overview, 2–5

- Mobile/wireless systems (continued)
 - radio systems planning/optimization, 6–7
 - scalability, 25, 27
 - spectral efficiency, 25
 - technology evolution, 15
- Modular representation, 116–18
 - DFT, 116–17
 - frequency assignment and scheduling, 117
 - IDFT waveform filtering, 117
 - prefix/postfix, 117–18
 - transmitted signal, 118
 - transmitter filtering, 118
- Modulation
 - advanced techniques, 153–70
 - with DBTCs, 154–62
 - defined, 101
 - equalization techniques, 126–31
 - FB-MC solution comparison, 135–36
 - filterbank, 116
 - filtered multitone, 131–35
 - IOTA, 116
 - modular representation, 116–18
 - multicarrier, 116, 118–26
 - parallel versus serial, 140–43
 - schemes, 116–43
 - See also specific modulation*
- Modulation and coding schemes (MCS), 208, 209
- Monte Carlo simulations, 150
 - in BER and PER estimation, 205
 - for RCP BLDP, 163
- Multiband orthogonal (MBO) system, 228
- Multiband signal structure, 37–48
 - DS-CDMA, 44–45
 - frequency offset, 46–48
 - MC-CDMA, 44–45
 - narrowband terminal accommodation, 44
 - OFDM/IOTA, 45–46
 - phase noise, 46–48
 - RF implementation issues, 44
- Multiband transmissions, 36–56
 - channel estimation and training, 48–53
 - line spectrum, 41
 - MIMO technologies, 53–56
 - signal structure, 37–48
- Multicarrier CDMA (MC-CDMA), 44, 178, 185–87
 - defined, 185
 - flexibility, 187
 - MMSE in, 44
 - scalability, 187
 - spreading operation, 187
- Multicarrier modulation (MCM), 116, 118–26
 - defined, 118–19
 - PRP-OFDM (MIMO), 124–26
 - PRP-OFDM (SISO antenna), 121–24
- Multicarrier-spread spectrum (MC-SS), 4, 419, 420
- Multicast polling, 395
- Multicell evaluations, 349
- Multicell simulation
 - central cell technique, 349–50
 - for interference cancellation
 - evaluation, 357
 - setup, 349–51
 - wraparound technique, 350–51
- Multihop communications
 - standards, 283–88
 - systems, 430
- Multihopping, 330
- Multihop relay scenarios, 310–12
 - defined, 310
 - example, 311
 - system capacity, 311
- Multimedia Broadcast Multicast Services (MBMS), 3
- Multipath propagation access
 - interference (MAI), 149
- Multiple-access schemes, 177–214
 - adaptability/flexibility, 179
 - capacity, 180
 - complexity and cost, 179
 - conclusions, 214
 - contention, 180
 - definition of, 192
 - duplexing techniques, 178
 - efficiency, 179
 - high speed user mobility, 189

- interleaved FDMA, 178
- introduction, 177–93
- large bandwidth, 189
- MC-CDMA, 178
- OFDMA, 178
- optimal, choice of, 177–78
- orthogonality, 180
- performance, 179
- resource allocation, 180
- SC-FDMA, 178
- SDMA, 190–91
- selection criteria, 179–80
- single-carrier versus multicarrier, 180–93
- TDMA/OFDMA, 193–214
- uncertainty, 179–80
- wide-area coverage, 189
- Multiple-antenna methods, 218–19
 - array gain, 220
 - beamforming, 221
 - benefits, 219–21
 - defined, 218
 - diversity gain, 220–21
 - diversity techniques, 222–27
 - elements, characterization of, 228–29
 - interference suppression gain, 221
 - property exploitation, 219
 - spatial multiplexing gain, 221
 - summary, 221–29
- Multiple-antenna systems, 429–30
 - benefits, 429
 - challenges, 429–30
 - with single-antenna terminals, 229
- Multiple-antenna techniques, 218
- Multiple input/multiple output (MIMO)
 - antenna techniques, 3
 - multiuser techniques, 430
 - OFDM combination, 13
 - receiver architecture, 225–26
 - receiver testing, 86
 - relay density, 299
 - relay systems, 16
 - RN, 299
 - space diversity techniques, 17
 - technologies, 15
 - transceivers, 17
 - transmission, spectral efficiency, 429
 - See also* MIMO channels; MIMO PRP-OFDM
- Multiple-input/single-output (MISO), 83
- Multiuser detectors, EXIT functions, 152, 153
- Multiuser diversity, 197
- Multiuser scheduling, 197
- Multiuser space-time codes, 143–53
 - code design criteria, 145, 146–48
 - downlink transmission scheme, 148–53
 - error event regions, 144–46
- $M \times N$ -URA, 262–63
- N**
 - N -element UCA (N -UCA), 262
 - N -element ULA (N -ULA), 261–62
 - Network services, 410
 - Nonmesh stations (STA), 287
 - Normalization gain block, 421
 - Normalized mean square error (NMSE), 201
 - Nyquist criterion, 134
- O**
 - OBAN project, 229
 - defined, 16
 - IEEE 802.11a/g physical layer, 231
 - scenario, 230
 - space and time diversity, 232–34
 - OFDM-based radio interfaces, 27–32
 - bandwidth allocation, 31
 - frequency offset sensitivity, 28
 - overview, 28
 - structure illustration, 30
 - synchronization algorithms, 28–32
 - See also* Radio interfaces
 - OFDM/IOTA, 45–46
 - as alternative, 45
 - power spectra, 46
 - spectra, 45
 - OFDM/OQAM, 140
 - Open-loop SMUX, 224
 - Open-loop transmit diversity, 346
 - Operation expenditures (OPEX), 289–90
 - Opportunistic beamforming, 226
 - Opportunistic scheduling, 189, 411

- Organization, this book, 18–19
- Orthogonal frequency-division multiple
 - access (OFDMA), 4, 178, 187–88
 - advantages, 187
 - average power spectra, 43
 - bins, 192
 - block, 192
 - capacity, 188
 - defined, 187
 - disadvantages, 187–88
 - frequency offset sensitivity properties, 46–47
 - in-band pilots, 49
 - interleaved, 192
 - link adaptation, 192–93
 - research, 188
 - signals, 143
 - uplink, bit and power allocations, 206
 - uplink, simulation parameters, 205
- Orthogonal frequency division
 - multiplexing (OFDM), 3
 - carrier phase noise, 119
 - CDMA combination, 13
 - cyclic prefix (CP-OFDM), 121, 123, 124
 - defined, 27, 119
 - drawback, 119
 - MIMO combination, 13
 - prototype filter, 132
 - pseudorandom postfix (PRP), 121–26
 - signal, 119
 - symbol, 119, 120
 - symbol duration, 132
 - technique, 27
 - tones, 144
 - Viterbi decoding, 129
 - See also* OFDM-based radio interfaces
- Orthogonality, 180
- Outage probability, 381
- Overlap-add (OLA), 126, 127
- P**
- Packet allocation unit, 414
- Packet binary convolutional coding (PBCC), 18
- Packet call delay, 353
- Packet delay, 353
- Packet error rate (PER), 213, 407
 - estimated performance, 245
 - performance, 246
- Packet headers, 410
- Pairwise error probability (PWEPP), 149, 150
- PAN-optimized interference scenarios, 361–84
 - body-to-body channel gain modeling, 362–69
 - cooperative transmission, 379–82
 - interference mitigation, 382–84
 - temporal properties of interferers, 369–77
 - unified interference modeling, 377–79
- Parallel concatenated convolutional codes (PCCC), 153
- Parallel modulation, 140–43
- Parity-check codes, 101
- Parity-check matrices, 105
 - encoding LDPC codes by, 105–7
 - illustrated, 106
- Path loss, 84–85, 89
- Peak-to-average power ratio (PAPR), 43
- Phase-locked loop, fractionary- N , 275
- Phase noise, 46–48
 - as impairment, 46
 - in RF oscillators, 29
- PHY layer, 392
 - adaptation to channel characteristics, 390
 - dynamics and source statistics, 407
- Physical channel
 - downlink, 9–10
 - requirements, 9–11
 - uplink, 10–11
- Piggybacking, 395
- Pilot signals, 29
- Planar antenna arrays, 252
- Point-to-multipoint (PMP)
 - configurations, 283
- Polarization matrix, 81
- Polling, 394–95
- Postfix, 118
- Power-delay profiles, 368
- Power spectrum density, 93, 343

Prediction accuracy, 200, 201, 202, 203
Prediction error model, 197–98
Prefix, 117–18
Prescalers, 17
Primary array coordinate system, 258
Probability density functions (PDFs), 79
Progressive edge-growth, 105
Propagation scenarios, 67
Proportional fair scheduling, 347, 399–400
Protocols
 cooperative relaying, 312–13
 MAC, 393–98
 requirements, 11–12
Pseudorandom postfix OFDM (PRP-OFDM), 121–26
 MIMO, 124–26, 129–30
 ML decoding techniques, 128–29
 postfix sequence design issues, 130–31
 SISO antenna, 121–24
 transceiver chain, 121
Pseudorandom postfix (PRP), 49
Pulse repetition intervals (PRIs), 268
PULSERS II project, 270
 differential antennas, 271
 interference, 333
 RF front end, 273–75
PULSERS project
 defined, 16
 interference, 333
 performance analysis, 270
Q
QRD/LDPC, 114–15
 defined, 114
 factor graph, 114
 symmetric design, 115
Quadratic residue codes, 103
Quadrature downconverter, 422
Quadrature upconverter, 422
Quality of service (QoS)
 guarantee, 9
 parameters, 9
 provided service, 8
 requirements, 8
Quasicyclic block LDPC (QC-BLDPC)
 codes, 153, 163–66

R
Radio access networks (RANs)
 short-range, 36
 wide-range, 35, 36
Radio access technologies (RATs)
 generations, 5
 heterogeneity of, 2
 R&D, 6
 transport technologies, 6
Radio bearers, 394
Radio channel research, 65
Radio interface design
 challenges and requirements, 8–14
 dimensions, 8–9
 enabling technologies, 13
 estimation techniques, 12–13
 marketing requirements, 13–14
 physical channel requirements, 9–12
 protocol requirements, 11–12
 resource blocks, 12
 technical requirements, 9–13
Radio interfaces
 coexistence and spectrum sharing, 32–34
 introduction, 23–56
 multiband transmissions, 36–56
 OFDM-based, 27–32
 protocol architecture, 392
 secondary spectrum use opportunities, 34–36
 spectral efficiency, 25
 spectrum-efficient technologies, 23–93
 for ubiquitous communications, 25–27
Radio link control (RLC) layer, 9
Radio resource control (RRC) layer, 9
Radio resource management (RRM), 5, 389–425
 adaptive soft resource positioning, 304–10
 conclusions, 425
 decisions, 391
 efficient, 389
 framework, 288
 functions, 391
 multihop relay scenarios, 310–12
 for relay-based deployments, 299–312

- Radio resource management (RRM)
 - (continued)
 - spectrum partitioning, 300–304
 - WINNER/WINNER II and, 389
- Radio systems
 - large-scale parameters, 84
 - link level simulation, 82
 - path loss, 84–85
 - planning and optimization, 6–7
 - shadowing, 85–86
 - simulation of, 82–86
 - system level simulation, 82
- Rake receivers, 181
- Random reservation schemes, 395
- Rank criterion, 147, 150
- Rate-compatibility through puncturing (RCP), 153–54
- Ray-based modeling, 78–80
 - dense multipath model, 80
 - ray fields, 78
 - ray optical, 79
 - ray tracing, 79
- Rayleigh-fading channel, 341
- Ray tracing, 295
- Receive diversity, 222, 346
- Rectangular antenna arrays, 252
- Recursive systematic convolutional (RSC)
 - encoders, 107
- Reed-Muller codes, 103
- Reed-Solomon (RS) codes, 103
- Relative chain block, 421
- Relay APs (RAPs), 292
 - bandwidth availability, 296
 - coverage area, 295
 - high power resources, 304
 - interferences and, 295
 - number of users served by, 296
 - transmitted power, 295
 - user throughput, 296
- Relay-based deployments
 - coverage areas, 297
 - intelligent, 298–99
 - performance comparison, 293, 297
 - resource partitioning, 294
 - RRM schemes, 299–312
- Relay deployment costs, 289–98
 - hardware, 289
 - site acquisition, 289
- Relay-enhanced cells (RECs), 288, 299
- Relaying
 - advanced techniques, 312–30
 - cooperative, 298–99, 312–26
 - DF, 322
- Relay nodes (RNs), 288, 300
- Relays
 - at cell edge, 330
 - in interference management, 334
 - leveraging, 330
 - simultaneous operation, 330
- Relay slot, spatial reuse, 316
- Relay technologies, 288–99
- RESOLUTION project, 333
- Resource allocation, 180
- Resource blocks, 12–13
- Resource management algorithms,
 - 398–407
 - link adaptation, 406–7
 - performance metrics, 402–6
 - scheduling, 398–406
- Resource partitioning, 294
- RF front end
 - architecture, 274
 - PULSERS II, 273–75
- RF module, 265–67
 - interface characteristics, 266–67
 - Rx diagram, 265
 - Tx diagram, 266
- Rician fading, 80
- Rician K -factor, 89
- Round-robin scheduling, 399
- S**
- Saleh-Valenzuela MIMO model, 83
- Sample-spaced channels, 369
- Scalability, 25, 27
- Scaled EIRP (SE) beamformer, 237, 243
- Scaled group (SG) beamformer, 243–44
- Scaled MR (SMR)
 - antenna patterns, 238
 - beamformers, 236
- Scheduling, 197
- Scheduling algorithms, 398–406
 - classification of, 398
 - maximum throughput, 403
 - performance metrics, 402–6

- proportional fair scheduler (PFS), 399–400
- round-robin scheduling, 399
- score-based (SB) scheduler, 400
- simulation example, 401–2
- SCME urban microcell model, 321
- Score-based algorithm, 356–57
- Score-based (SB) scheduler, 400
- Secondary array coordinate system, 258
- Secondary spectrum use, 34–36
- Second-order cone programming (SOCP)
 - problem, 236
- Selective combining (SC), 112
- Serially concatenated convolutional codes (SCCCs), 116
- Serial modulation, 140–43
- Shadowing, 85–86, 89
 - attenuation, 85–86
 - independent realizations, 322
 - modeling, 85
- Shannon capacity formula, 302
- Shannon's theorem, 295
- Short interframe space (SIFS), 247
- Signal-to-interference-noise ratio (SINR)
 - calculating, 293
 - CDF of, 352, 355, 356, 357
 - effective, 210
 - received packets, soft resource partitioning, 309
- Signal-to-noise ratio (SNR)
 - average packet, 211, 213
 - degradation, 139
 - gain, 239, 240
 - low, design criteria, 148
 - mismatch impact on LDPC codes, 167, 168
 - thresholds, 162
- SIMULINK model, 421
- Single-carrier FDMA (SC-FDMA), 178
- Single-carrier frequency-domain equalization (SC-FDE), 140
- Single-carrier solutions, 181–85
- Single-carrier TDMA (SC-TDMA), 183
- Single-input/multiple-output (SIMO), 83
- Smart antennas, 217–76
 - antenna array modeling, 251–67
 - conclusions, 275–76
 - defined, 217
 - downlink capacity enhancement, 229–51
 - importance, 217
 - interference mitigation, 345–47
 - introduction, 218–29
 - SFIR with, 327
 - UWB antenna architectures, 267–75
 - See also* Antennas
- Software-defined radios (SDRs), 33
- SOVA algorithm, 103
- Space-division multiple access (SDMA), 190–91
 - for a group of terminals, 248
 - antennas, 191
 - bottleneck, 190
 - defined, 190
 - DL cell throughput, 328
 - potential performance, 247
 - principle, 327
 - single-hop, 329
 - spatial filtering and, 326–30
 - throughput CDF for, 250, 251, 252, 253
 - time offset slot, 248
 - time offset solution, 251
 - time-shifted transmission, 248
- Space-frequency code (SFC), 223
- Space-time block coding (STBC), 124, 149
 - decoding capabilities, 232
 - with linear transformation, 222
 - MIMO multiuser scheme with, 149
 - quasiorthogonal, 223
 - signals, 231
 - techniques, 231
 - transmitters, 232
 - weighted, 223
- Space-time coding, 148
- Space time transmit diversity (STTD), 299
- Spatial channel model (SCM), 69–73
 - channel matrices, 71
 - channel scenarios, 72
 - comparison, 75, 76
 - link level, 70
 - overview, 72

- Spatial channel model (SCM) (continued)
 - submodels, 69
 - system level, 70–73
 - WINNER project, 77
 - See also* Channel models
 - Spatial correlation matrix
 - as antenna array dependent, 80
 - estimation, 80
 - modeling based on, 80–81
 - Spatial domain link adaptation, 225
 - Spatial filtering for interference reduction (SFIR), 327
 - DL cell throughput, 328
 - principle, 327
 - with smart antennas, 327
 - Spatial holes, 35
 - Spatial modes, 229
 - Spatial multiplexing, 223–25
 - closed-loop, 224
 - diversity and, 224–25
 - gain, 221
 - open-loop, 224
 - scheme illustration, 224
 - spectral efficiency, 224–25
 - techniques, 223, 224
 - Spatial processing, 299
 - Spatial reuse, 316
 - Spatiotemporal processing, 330
 - Spectral efficiency, 25
 - Spectrum
 - density masks, 337
 - holes, 35
 - multiband line, 41
 - occupancy, 34
 - secondary use, 34–36
 - Spectrum partitioning, 300–304
 - parts, 301
 - problem, 301
 - procedure, 301
 - Spectrum sharing
 - bands, 33
 - cognitive radio and, 33
 - defined, 32
 - efficient, 32
 - Sphere decoding, 225
 - Spherical antenna arrays, 253
 - Spread-spectrum systems, 34
 - Standardization activities, 14–18
 - European-funded research, 15–17
 - IEEE 802.11, 18
 - 3GPP2 Revision, 17–18
 - WPANs/WBANs, 18
 - Statistical modeling, 77–78
 - Stochastic models, 368, 377, 386
 - Subcarrier allocation, 205
 - Successive interference cancellation (SIC), 345
 - Sum capacity optimization, 226–28
 - SURFACE project
 - defined, 17
 - interference issues, 333
 - resource allocation strategies, 389
 - Synchronization
 - algorithms, 28–32
 - coarse, 30
 - cyclic prefix, 29
 - fine, 30–31
 - pilot signals, 29
 - types of, 29
 - System level simulations, 82
 - advanced channel model, 91
 - basic channel model, 91
- ## T
- TDMA/OFDMA
 - adaptive multiuser, 188–90
 - advantages, 187
 - analysis for IMT-A candidate system, 193–214
 - average packet SNR, 212, 213
 - capacity, 188
 - defined, 187
 - design aspects, 193–97
 - disadvantages, 187–88
 - FDD, 196
 - frame of time-frequency bins, 191, 192
 - framework, 214
 - ICI mitigation, 197
 - link adaptation, 193–97
 - performance modeling, 197–214
 - prediction error model, 197–98
 - scheduling, 197
 - single-cell deployment, 210

- system design issues, 191–92
 - TDD, 195, 208, 214
 - user data rates, 212, 213
- Temporal diversity, 220–21
- Temporal holes, 35, 342
- 3GPP spatial channel model, 69–73
 - comparison, 75, 76
 - link level, 70
 - submodels, 69
 - system level, 70–73
 - WINNER project, 77
- 3G (SISO) models, 75
- 3G system enhancements, 2–5
- Three-dimensional antenna arrays, 255–59, 255–61
 - BS/MS orientations, 257
 - illustrated, 258
 - MS/MS system layout, 256
 - See also* Antenna arrays
- Time-division duplex (TDD)
 - downlinks, 199, 200
 - frame structure, 30, 208
 - maximal prediction horizons, 203
 - prediction performance, 198–214
 - TDMA/OFDMA scheme, 195, 214
 - uplinks, 199, 200
- Time-division multiple-access (TDMA), 142
 - challenge, 181–83
 - intracell interference and, 181
 - nonspread, 183
 - single-carrier (SC-TDMA), 183
 - wideband, 185
- Tomlinson-Harashima precoding (THP), 228
- Training sequences, 48
- Transceiver architecture, 275
- Transfer functions, 144, 206
- Transit time interval (TTI), 399, 400
- Transmit diversity, 222–23
 - closed-loop, 346
 - open-loop, 346
 - See also* Diversity
- Transmitter chain block, 421
- Transmitter filtering, 118
- Turbo codes
 - decoding, 110–11
 - defined, 104
 - design and performance, 111–12
 - duo-binary (DBTC), 116
 - encoding, 107–8
 - error correction, 104
 - error floor, 111
 - LDPC codes versus, 104–5
 - as mature technology, 104
 - SNR, 112
- Turbo decoders, 110–11
 - extrinsic information, 111
 - iterative, f, 110
 - operation, 110–11
- Turbo equalizer (TE), 141, 185
- Turbo multiuser detection (MUD), 104
- Two-dimensional antenna arrays, 259–61
 - cell structure and topology, 255
 - global-coordinate-system (GCS), 255
 - illustrated, 260
 - implementation, 260
 - See also* Antenna arrays
- Two-dimensional coordinate system, 261
- U
- Ultrawide band (UWB), 4
 - antenna structures, 267–75
 - arrays, 276
 - differential antenna, 271
 - impulse radio (IR-UWB), 268
 - signals, 268
- Unified interference model, 377–79
 - defined, 377
 - IMCM, 379
 - PAN scenarios, 379
- Uniform linear arrays (ULAs), 88
- Uniform rectangular arrays (URAs), 88
- UNITE project, 17
- Universal Mobile Telecommunications Services (UMTS), 2, 6
- Uplink system control data, 10–11
- URANUS project, 16
- User data rates, 212
- User terminals (UTs), 4
 - capabilities, 8
 - output power limitations, 11
- User throughput
 - average, 305, 354
 - as function of distance, 308, 309

V

- Validation, 91–92
- V-BLAST, 225, 226
- Vector MIMO channels, 227–28
- Very high data rate digital subscriber line (VDSL), 133
- Viterbi decoding, 129
- Voltage controlled oscillators (VCOs), 17

W

- Weighted spectral efficiency, 282
- Wideband code division multiple-access (WCDMA), 3
- Wideband demodulators (WBDM), 17
- Wideband Distribution System (WDS), 3
- Wideband TDMA, 185
- Wide-sense stationary uncorrelated scattering (WSSUS) assumption, 78
- WiMAX
 - handover, 283–84
 - networks, 284–85
- WINDECT project, 17
- WINNER II project
 - defined, 15–16
 - RRM strategies, 389
- WINNER project, 7
 - air interface, 5, 6
 - antenna-independent channel model, 81
 - baseline design, 154

- channel models, 77, 254
- generic model, 15
- MAC protocol, 394–95
- multicell system simulation, 349
- parity-check codes, 101
- power mask adaptation, 305
- radio access concept, 218–19
- RAN architecture, 5
- RECs, 288
- RRM strategies, 389
- SCME urban microcell model, 321
- single-carrier versus multicarrier schemes, 180–81
- ubiquitous radio system concept, 7
- wide-area baseline scheme, 360
- Wireless PANs (WPANs), 10
- Wireless systems. *See* Mobile/wireless systems
- Wireless World Research Forum (WWRF), 1
- WLAN system enhancements, 2–5
- Wraparound technique, 350–51

X

- XNOR gate, 268, 274
- XNOR output, 268, 269

Z

- Zero-forcing (ZF) equalization, 131
- Zero-padding OFDM (ZP-OFDM), 126
- ZF block linear equalizer, 138
- Zig-zag LDPC, 153

The Artech House Universal Personal Communications Series

Ramjee Prasad, Series Editor

4G Roadmap and Emerging Communication Technologies, Young Kyun Kim and
Ramjee Prasad

802.11 WLANs and IP Networking: Security, QoS, and Mobility, Anand R. Prasad and
Neeli R. Prasad

CDMA for Wireless Personal Communications, Ramjee Prasad

From WPANs to Personal Networks: Technologies and Applications, Ramjee Prasad and
Luc Deneire

IPI/ATM Mobile Satellite Networks, John Farserotu and Ramjee Prasad

Multicarrier Techniques for 4G Mobile Communications, Shinsuke Hara and
Ramjee Prasad

New Horizons in Mobile and Wireless Communications, Ramjee Prasad and
Albena Mihovska, editors

Volume 1: Radio Interfaces

Volume 2: Networks, Services, and Applications

Volume 3: Reconfigurability

Volume 4: Ad Hoc Networks and PANs

OFDM Towards Broadband Wireless Access, Uma Shanker and Jha Ramjee Prasad

OFDM for Wireless Communications Systems, Ramjee Prasad

OFDM for Wireless Multimedia Communications, Richard van Nee and Ramjee Prasad

Practical Radio Resource Management in Wireless Systems, Sofoklis A. Kyriazakos and
George T. Karetzos

Radio over Fiber Technologies for Mobile Communications Networks, Hamed
Al-Raweshidy and Shozo Komaki, editors

Simulation and Software Radio for Mobile Communications, Hiroshi Harada and
Ramjee Prasad

Space-Time Codes and MIMO Systems, Mohinder Jankiraman

TDD-CDMA for Wireless Communications, Riaz Esmailzadeh and Masao Nakagawa

Technology Trends in Wireless Communications, Ramjee Prasad and Marina Ruggieri

Third Generation Mobile Communication Systems, Ramjee Prasad, Werner Mohr, and
Walter Konhäuser, editors

Towards a Global 3G System: Advanced Mobile Communications in Europe, Volume 1,
Ramjee Prasad, editor

Towards a Global 3G System: Advanced Mobile Communications in Europe, Volume 2,
Ramjee Prasad, editor

Towards the Wireless Information Society: Heterogeneous Networks,
Ramjee Prasad, editor

Towards the Wireless Information Society: Systems, Services, and Applications,
Ramjee Prasad, editor

Universal Wireless Personal Communications, Ramjee Prasad

WCDMA: Towards IP Mobility and Mobile Internet, Tero Ojanperä and Ramjee Prasad
editors

Wideband CDMA for Third Generation Mobile Communications,
Tero Ojanperä and Ramjee Prasad, editors

Wireless Communications Security, Hideki Imai, Mohammad Ghulam Rahman and
Kazukuni Kobara

Wireless IP and Building the Mobile Internet, Sudhir Dixit and Ramjee Prasad, editors

WLAN Systems and Wireless IP for Next Generation Communications, Neeli Prasad and
Anand Prasad, editors

WLANS and WPANs towards 4G Wireless, Ramjee Prasad and Luis Muñoz

For further information on these and other Artech House titles, including previously considered out-of-print books now available through our In-Print-Forever® (IPF®) program, contact:

Artech House	Artech House
685 Canton Street	16 Sussex Street
Norwood, MA 02062	London SW1V 4RW UK
Phone: 781-769-9750	Phone: +44 (0)20 7596-8750
Fax: 781-769-6334	Fax: +44 (0)20 7630-0166
e-mail: artech@artechhouse.com	e-mail: artech-uk@artechhouse.com

Find us on the World Wide Web at: www.artechhouse.com



EX
LIBRIS
Eugene A.
Katkovsky