University of Southampton
Faculty of Engineering, Science and Mathematics
School of Electronics and Computer Science

Non-orthogonal Random Waveform Based Multiuser Communications

by

Rong Zhang
B.Eng., M.Sc.

A thesis for the award of Doctor of Philosophy
at the University of Southampton

June 2009

SUPERVISOR: *Professor Lajos Hanzo*
M.Sc., Ph.D, FREng, DSc, FIEEE, FIET
Chair of Telecommunications
School of Electronics and Computer Science
University of Southampton
Southampton SO17 1BJ
United Kingdom

Dedicated to my family

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

Faculty of Engineering, Science and Mathematics

School of Electronics and Computer Science

A thesis for the award of Doctor of Philosophy

Non-orthogonal Random Waveform Based Multiuser Communications

by Rong Zhang

This thesis is aimed at providing a unified treatment of non-orthogonal random waveform based multiuser communications, commencing from the underlying theory and providing the way to sophisticated applications. In Chapter 2, we introduce the general framework of non-orthogonal random waveform based multiuser communications, where we focus our attention on the fundamental theory behind this topic. We then pay particular attention to three general application aspects, where the principle of non-orthogonal random waveform based multiuser communications may be exploited. These are applied to cellular systems in Chapter 3, to cooperative systems in Chapter 4 and they are combined with a practical Hybrid Automatic Repeat reQuest (HARQ) system in Chapter 5.

In the cellular system scenario presented in Chapter 3, we discuss three design tradeoffs of Multi-Carrier Interleave Division Multiple Access (MC-IDMA) with the aid of EXtrinsic Information Transfer (EXIT) charts. Based on the analysis of IDMA, we generalise the principle of non-orthogonal random waveform based multiuser communications, leading to the so-called Interleaved Random Code Division Multiple Access (IR-CDMA) concept. We also consider an interference-limited cooperative system in Chapter 4, where we propose a novel Interleaved Random Space Time Code (IR-STC) scheme, which is specifically designed for a Multi-Source Cooperation (MSC) scenario. Furthermore, the network coding concept is also employed, when designing a coded MSC arrangement for the sake of improving the attainable energy efficiency. More specifically, the proposed Physical-layer Algebraic Network Coding (PANC) scheme implicitly inherits the random coding principle applied in a distributed manner. In Chapter 5, we proceed by proposing a novel SuperPosition Coding (SPC) aided multiplexed HARQ scheme for substantially improving the overall end-to-end transmission efficiency.

In parallel to presenting our general framework of non-orthogonal random waveform based multiuser communications, we bear in mind that employing a powerful Multi-User Detector (MUD) for maintaining a near-single-user performance is of high importance. Hence, in Chapter 6 we propose a novel so-called Harmony Search (HS) algorithm for joint iterative channel estimation, data detection and channel decoding. The algorithm proposed can be readily employed to replace all detection algorithms used throughout Chapters 3 to 5, when a high throughput is desired.

Declaration of Authorship

I, **Rong Zhang**, declare that the thesis entitled **Non-orthogonal Random Waveform Based Multiuser Communications** and the work presented in it are my own and has been generated by me as the result of my own original research. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University;
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- Where I have consulted the published work of others, this is always clearly attributed;
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- Parts of this work have been published.

Signed: ..... Date: .....

Acknowledgement

I would like to express my gratitude to many people, without their support this work could not be done.

First of all, I would like to express my gratitude to my Professor, Lajos Hanzo, for giving me the opportunity to join the Comms Laboratory and whose excellent guidance and encouragement have helped me to achieve the seemingly impossible.

Especially, my grateful thanks go to the Virtual Centre of Excellence in Mobile and Personal Communications Ltd. (Mobile VCE) whose funding support, including that of Engineering and Physical Sciences Research Council (EPSRC), is gratefully acknowledged.

I am grateful to my colleges, who enriched my experience for the past three years. Special thanks goes to Dr. Soon Xin Ng and Dr. Jos Akhtman for their generous support. Also thanks to Mr Nicolas Bonello, Dr. Lei Xu, Dr. Robert Maunder, Dr. Lie-Liang Yang and Professor Sheng Chen etc for many interesting discussions.

I want to express my immense and never-ending gratefulness to my admirable parents Chengyu Zhang and Meifang Ren for their love and unconditional support.

List of Journals

1. **R. Zhang** and L. Hanzo, "Harmony Search Aided Iterative Channel Estimation, Multiuser Detection and Channel Decoding for DS-CDMA", submitted to *IEEE Transaction on Wireless Communications*

2. **R. Zhang** and L. Hanzo, "Iterative Multiuser Detection and Channel Decoding for DS-CDMA Using Harmony Search", accepted by *IEEE Signal Processing Letters*

3. **R. Zhang** and L. Hanzo, "Superposition-Coding Aided Multiplexed Hybrid ARQ Scheme for Improved End-to-End Transmission Efficiency", accepted by *IEEE Transaction on Vehicular Technology*

4. **R. Zhang** and L. Hanzo, "Coding Schemes for Energy Efficient Multi-Source Cooperation Aided Uplink Transmission", *IEEE Signal Processing Letters*, May. 2009, Volume 16, pp:438 - 441

5. **R. Zhang** and L. Hanzo, "Interleaved Random Space-Time Coding for Multisource Cooperations", *IEEE Transaction on Vehicular Technology*, May. 2009, Volume 58, pp:2120 - 2125

6. **R. Zhang** and L. Hanzo, "Three Design Aspects of Multicarrier Interleave Division Multiple Access", *IEEE Transaction on Vehicular Technology*, Dec. 2008, Volume 57, pp:1671-1680

7. **R. Zhang** and L. Hanzo, "Space-Time Coding for High-Throughput Interleave Division Multiplexing Aided Multi-Source Co-Operation", *IET Electronics Letters*, Feb. 2008, Volume 44. pp. 367-368.

8. **R. Zhang**, L. Xu, S. Chen and L. Hanzo, "EXIT-chart Aided Hybrid Multiuser Detector Design for Frequency-Domain-Spread Chip-Interleaved MC-CDMA", submitted to *IEEE Transaction on Vehicular Technology*

9. N. Bonello, **R. Zhang**, S. Chen and L. Hanzo, "Reconfigurable Rateless Code", submitted to *IEEE Transaction on Communications*

10. N. Bonello, **R. Zhang**, S. Chen and L. Hanzo, "Channel code division multiple access and its multilevel structured LDPC based instantiation", *IEEE Transaction on Vehicular Technology*, June. 2009, Volume 58, pp:2549 - 2553

List of Conferences

1. **R. Zhang** and L. Hanzo, "Harmony Search Aided Iterative Multiuser Detection and Channel Decoding for DS-CDMA", submitted to *IEEE Globecom 2009*

2. **R. Zhang** and L. Hanzo, "Superposition-Coding Aided Multiplexed Hybrid ARQ Scheme for Improved Link-layer Transmission Efficiency", accepted by *IEEE ICC 2009*

3. **R. Zhang** and L. Hanzo, "Physical Layer Algebraic Network Coding and Superposition Coding for Multisource Cooperation Uplink", accepted by *IEEE VTC 2009-Spring*

4. **R. Zhang** and L. Hanzo, "High-Throughput Non-Orthogonal Interleaved Random Space-Time Coding for Multi-Source Cooperation", in *Proceedings of IEEE Globecom 2008*, New Orleans, USA, Nov. 30-Dec. 04, pp:1-5

5. **R. Zhang** and L. Hanzo, "Interleave Division Multiplexing Aided Space-Time Coding for High-Throughput Uplink Cooperative Communications", in *Proceedings of IEEE WCNC 2008*, Las Vegas, USA, Mar. 31-Apr. 3, pp:465-469

6. **R. Zhang** and L. Hanzo, "EXIT Chart Based Joint Code-Rate and Spreading-Factor Optimisation of Single-Carrier Interleave Division Multiple Access", in *Proceedings of IEEE WCNC 2007*, HongKong, China, 11-15 Mar, pp:735-739

7. **R. Zhang** and L. Hanzo, "Iteratively Detected Multi-Carrier Interleave Division Multiple Access (invited)", in *Proceedings of IEEE MICROCOLL 2007*, Budapest, Hungary, 14-16 May, pp:735-739

8. **R. Zhang**, L. Xu, S. Chen and L. Hanzo, "Repeat Accumulate Code Division Multiple Access and its Hybrid Detection", in *Proceedings of IEEE ICC 2008*, Beijing, China, 19-23 May, pp:4790-4794

9. N. Bonello, **R. Zhang**, S. Chen and L. Hanzo, "Reconfigurable Rateless Code", accepted by *IEEE VTC 2009-Spring*

10. N. Bonello, **R. Zhang**, S. Chen and L. Hanzo, "Channel code division multiple access and its multilevel structured LDPC based instantiation", in *Proceedings of IEEE VTC 2008-Fall*, Calgary, Canada, 21-24 May, pp:1-5

11. L. Xu, **R. Zhang**, S. Chen and L. Hanzo, "EXIT-Chart Aided Hybrid Multiuser Detector Design for Frequency-Domain-Spread Chip-Interleaved MC-CDMA", in *Proceedings of IEEE VTC 2008-Spring*, Marina Bay, Singapore, 11-14 May, pp:1816-1820

Contents

Chapter 1

Introduction

Wireless communications is becoming all-pervasive, moving from being an adjunct to wireline communications to being an essential embedded enabler for multi-service ubiquitous communications. This is due to a confluence of several factors. First of all, there has been an explosive increase in demand for connectivity, driven by wireless data applications. Second, the dramatic progress in new materials and micro-electronics technology has enabled the low-power implementation of sophisticated signal processing algorithms and coding techniques on a small chip-area. Third, the successful Second-Generation (2G) [1] digital wireless standards and the well-established High Speed Packet Access (HSPA) [2] wireless standards provide a concrete demonstration that bridges communications theory and practice.

Although substantial advances have been made, efficient information delivery over wireless links is becoming more and more crucial, requiring sophisticated enabling techniques to satisfy the ever-evolving service requirements. It has been widely recognised that there is no single enabling technique, which is capable of improving the achievable system performance in all scenarios. This is the consequence of having to satisfy inherent design tradeoffs, leading to sophisticated cross-layer interactions. Just to mention a few of the associated challenges: the design of robust point-to-point communications links; efficient resource sharing and user scheduling mechanisms; the design of combined wireless cellular / ad hoc networks and the improvement of inter-operatability with other networks.

Despite these challenges, from the stand-alone physical layer perspective, Multiple Input Multiple Output (MIMO) [3] and Orthogonal Frequency Division Multiplexing (OFDM) [4] constitute promising techniques, especially, when combined with powerful link-adaptation [5] and Hybrid Automatic Repeat reQuest (HARQ) [6] in the link layer, which requires sophisticated Digital Signal Processing (DSP) algorithms. At the time of writing, the Third-Generation Partnership Project's (3GPP) Long Term Evolution (LTE) [7] proposed to employ OFDMA in the Down-Link (DL) and Single Carrier Frequency Division Multiple Access (SC-FDMA) combined with Frequency Domain Equalisation (FDE) in the UpLink (UL). Although substantial data-rate improve-

ments have been achieved by the 3GPP LTE system, it should be recognised that this improvement is mostly attributed to the MIMO technique employed and to its flexible bandwidth allocation. Future wireless systems are expected to become data-centric and IP-based. Massive data rates are required both in the UL and in the DL, leading to the key problem of minimising the *cost-per-bit*. These requirements could not be readily fulfilled, unless further substantial advances are made. In addition to the above-mentioned mature techniques, Interference Cancellation (IC) aided receivers are under consideration in the Worldwide Interoperability for Microwave Access (WIMAX) [8]. Furthermore, the potential employment of relays also attracts a lot of interests [9].

Despite the above-mentioned great strides in technology, there is a lack of a unified treatment of the philosophy of advanced multiuser communications and its applications. Hence, in this thesis **we aim for providing a unified treatment of near-capacity non-orthogonal random waveform based multiuser communications. With this aim, we amalgamate several beneficial wireless communications techniques, namely SuperPosition Coding (SPC), IC, iterative receivers and the employment of quasi-random codes.**

This chapter is organised as follows. We first briefly discuss the necessity of employing SPC, IC, iterative receiver and quasi-random codes in Section 1.1. These discussions are followed by a discourse on the topic of orthogonal versus non-orthogonal multiuser communications in Section 1.2. We then present the inter-dependencies of our main chapters and outline the novel contributions of our work in Section 1.3.

1.1 Core Techniques for Point-to-Point Links

We first briefly discuss the benefits of employing SPC, IC, iterative receivers and quasi-random codes for a transceiver operating in a point-to-point wireless link. Two key design requirements for a point-to-point wireless link are achieving *high bit rates* and *low error rates*, which constitutes a well-known tradeoff. Fundamentally, given a fixed degree of freedom for a system, each related 'subchannel' can either transmit new independent information for the sake of supporting high bit rates or may be used to convey correlated redundancy in terms of channel coding and/or related information replicas over independently faded diversity channels for the sake of achieving low error rates. For example, when communicating over a Binary Input Additive White Gaussian Noise (BI-AWGN) channel, one can either transmit at 1 bit/symbol at an error probability of $Pe \approx 10^{-5}$ at around $E_b/N_0 = 9$dB using BPSK modulation, or opt for achieving the same performance at a lower transmit power using channel codes by sacrificing the throughput and imposing more sophisticated DSP.

Achieving a high throughput. There are several ways of achieving high bit rates. The most straightforward way is to increase the bandwidth by transmitting at higher rates, i.e. using a shorter symbol duration. However, due to the multipath nature of wireless channel, this leads to pro-

longed Inter Symbol Interference (ISI). Alternatively, one can employ the now classic OFDM technique, where a high bit rate is achieved by mapping low-rate data symbols onto multiple parallel subcarriers, where the ISI is reduced owing to the high subcarrier symbol-duration, but the total bandwidth is largely increased. The classic technique of increasing the throughput is to use high-order modulation schemes transmitting several bits per symbol without bandwidth expansion, while requiring an increased Signal-to-Noise-Ratio (SNR), since the phasers are densely packed in the constellation [10]. Additionally, high bit rates can also be attained by multiplexing several streams using multiple antennas (Bell-Labs Layered Space Time Architecture (BLAST)) or using multiple spreading codes.

SPC schemes can be viewed as a specific modulation techniques, where the complex-valued phaser constellation may be viewed as being Gaussian distributed, rather than obeying a predefined ordered structure, as in conventional modulation schemes. On the other hand, a SPC scheme may also be viewed as being a multiplexing technique, where instead of simultaneously transmitting from multiple antennas or spreading codes, a high throughput is achieved by simultaneously transmitting in the form of multiple superimposed streams or layers. This Gaussian distributed channel input of SPC is desirable from an information theoretic point of view, because it is capable of approaching the ergodic channel capacity [11]. However, it renders the receiver more complex, since there is no clearly defined decision boundary as in the context of conventional modulation schemes. Hence, classic stochastic estimation theory plays a crucial role in detecting such a composite signal, such as for example Bayesian Inference [12]. On the other hand, IC techniques provide a low-complexity design alternative, which in theory approach the ultimate Bayesian performance. We will provide a more comprehensive discussion on these two issues, namely on SPC and on IC in Chapter 2.

In fact, the concept of SPC has already been adopted implicitly in many modern communications system designs. The most important one may be the family of Linear Dispersion Codes (LDC) [13], where each antenna transmits a weighted sum of many input symbols at a time and the weighting factor of each antenna's stream is typically found subject to a predefined design criterion. Dirty Paper Coding (DPC) of [14] is also reminiscent of the SPC principle, where in addition to simply superimposing multiple layers, we also subtract the recognisable sources of interference prior to transmission. This may be referred to as using IC at the transmitter.

Achieving low error rates. Low error rates may be achieved by providing correlated redundancy with the aid of channel coding [15] and/or by transmitting independently faded signal replicas in order to achieve a diversity gain [1]. Shannon's channel coding theorem was first implemented by Gallager using random channel coding [16]. Random codes constitute an ensemble of codes having long codewords having a high Hamming distance independently distributed over the 2^N-element codeword space, where N is the length of the codewords. Conventional convolutional codes having a low memory and linear block codes having a short block-length are short codes having a limited diversity of codewords, hence they exhibit moderate coding gain [6]. Low Den-

sity Parity Check (LDPC) codes [17] and interleaved concatenated codes [18] are quasi-random codes, since they have a long input information block length. The randomness of their codewords is guaranteed by their inherently random design.

The employment of quasi-random codes in a K-user scenario is even more important than in single-user channel-coded links. It provides another dimension for separating users, but at the same time potentially increases the complexity of the detection process due to having an enlarged 2^{KN}-element codeword space. Again, as a result of having a Gaussian input owing to the employment of SPC, powerful transmit structure emerge. The turbo-principle [19] inspired a whole host of powerful iterative detection and decoding schemes, which make the receiver trackable. It exploits the fact that the interleaver embedded in the transmitter randomises and decorrelates the information to be transmitted, hence the detection and decoding process may be decoupled. The intricate details of the random coding principle and iterative receiver algorithms will be discussed in Chapter 2.

Exploiting the fact that the probability of simultaneously encountering deep fades for several independent channels is low, diversity techniques are capable of maintaining low error rates by mitigating the fading effects with the aid of multiple independently fading replicas of the transmitted signal. There are various ways of creating independently faded channels in wireless systems [20]. *Time diversity* is achieved by transmitting replicas of the same signal at different time instants, which span several coherence time intervals. For example, the classic idea of channel coding and interleaving, such as the well-known Bit Interleaved Coded Modulation (BICM) scheme [21] constitutes an attractive manifestation of this time-diversity family. *Frequency diversity* is achieved by transmitting replicas of the same signal over different carrier frequencies, where the carriers are separated by at least the coherence bandwidth. A typical example of providing frequency diversity is constituted by a channel-coded and interleaved OFDM system [4]. Additionally, a spread-spectrum system using a RAKE receiver is also capable of providing diversity, where each finger of the RAKE receiver constitutes the realization of an independently faded channel. The realization of independently faded channels in the time and frequency domain are capable of directly exploiting the inherent diversity of the wireless channel. By employing multiple antennas (also referred to as an antenna array) at either the transmitter or the receiver, *spatial diversity* can be achieved. The elements of the antenna array should be sufficiently separated in space, so as to provide independently faded paths. The signal correlations imposed by the insufficient spacing of the antenna elements will degrade the achievable spatial diversity gain.

1.2 Orthogonal versus Non-orthogonal Communications

The SPC concept is closely related to non-orthogonal multiuser communications, hence below we briefly review the classic topic of orthogonal versus non-orthogonal multiuser communications in the context of cellular systems.

The allocation of signalling dimensions to multiple users engaged in continuous transmissions, while obeying certain delay constraints typically requires dedicated channels for maintaining a good performance in order to ensure that their transmissions are unperturbed. Dedicated channels can be created in the system's signal space using a so-called channelisation method such as time-division, frequency-division, code-division, spatial-division or a hybrid combination of these techniques. This is also known as *multiple access*. The UL is also often termed as a multiple access channel, while the DL as a broadcast channel. Without confusion, we will use the terminology of multiuser communications, which includes both the UL and DL scenario [20, 22, 23].

Multiuser communications techniques partition the signalling dimensions into channels and then assign these channels to different users in either an orthogonal or a non-orthogonal approach [24]. The well-known Time Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA) constitute inherently orthogonal methods, whereas a Generalized Multi-Carrier Direct Sequence Code Division Multiple Access (MC-DS-CDMA) can be either orthogonal or non-orthogonal, depending on the specific spreading code design employed [25–28]. In fact, the non-orthogonal approach can be realized in a Generalized code domain having typical instantiations such as for instance Trellis Code Multiple Access (TCMA) [29, 30] and Interleave Division Multiple Access (IDMA) [31, 32].

Orthogonal multiuser communications techniques are equivalent in the sense that they orthogonally partition the signalling dimensions. In particular, given signals having a bandwidth of B and a time duration of T occupying a signal space of dimension $2BT$, K orthogonal channels of dimension $2BT/K$ can be created. By contrast, non-orthogonal multiuser communications allow each user to exploit the entire signal space of dimension $2BT$, but inevitably imposes interference among the users. The performance of non-orthogonal multiuser communications depends on whether UL or DL are considered [20]. A fundamental difference between the UL and the DL is that in the latter, both the signal and the interference are affected by the same channel, while in the UL the signals arriving from different users encounter different channels.

The employment of orthogonal versus non-orthogonal multiuser communications raises numerous pertinent research problems. In fact, in the UL, unless TDMA or FDMA is employed, maintaining orthogonality in the spreading code-domain is challenging, unless accurately controlled adaptive timing-advance alignment is used, as in the emergent Time-Division Synchronised CDMA (TD-SCDMA) standard in China [33]. On the other hand, orthogonality is easier to be realized in the synchronous DL, although the orthogonality of the spreading codes in CDMA may still be destroyed by the multipath channel. In all, the orthogonal approach is more attractive in the DL so as to mitigate the intra-cell interference, since the limited complexity low cost DL receiver may not be capable of removing the multiuser interference imposed in case of a non-orthogonal scenario. However, a non-orthogonal interference-limited scenario is encountered, when the inter-cell interference of multiple cells is taken into account from a system-level point of view, where the DL receiver should eliminate the contamination imposed by a few dominant intercell interferers

imposed by frequency reuse pattern employed.

In all, the performance of today's cellular networks' performance is more constrained by the interference than by any other single effect. Both the UL and the DL of a cellular system exhibit a non-orthogonal interference-limited nature. It is widely recognised that non-orthogonal multiuser communications is superior to its orthogonal counterpart in terms of the maximum achievable capacity [34]. This justifies that the full understanding of non-orthogonal multiuser communications problems is crucial, which is the motivation of this thesis. What is worth emphasising furthermore that for future high-data-rate wireless systems, a combination of orthogonal and non-orthogonal techniques may be necessary.

1.3 Thesis Dependencies, Contributions and Outline

This thesis is aimed at providing a unified treatment of non-orthogonal random waveform based multiuser communications, commencing from the underlying theory and providing the way to sophisticated applications. Let us now first reveal the essential dependencies amongst the various chapters. In Chapter 2, we introduce the general framework of non-orthogonal random waveform based multiuser communications, where we focus our attention on the fundamental theory behind this topic. Three key topics are emphasised, namely the multiuser channel capacity, the random coding principle and the iterative receiver architecture.

Following a comprehensive portrayal of the required background knowledge, we pay particular attention to three general application aspects, where the principle of non-orthogonal random waveform based multiuser communications may be exploited. These are applied to cellular systems in Chapter 3, to cooperative system in Chapter 4 and they are combined with a practical HARQ system in Chapter 5.

In the cellular system scenario presented in Chapter 3, we discuss three design tradeoffs of Multi-Carrier Interleave Division Multiple Access (MC-IDMA) [35] with the aid of EXtrinsic Information Transfer (EXIT) charts [36]. Based on the analysis of IDMA, we generalise the principle of non-orthogonal random waveform based multiuser communications, leading to the so-called Interleaved Random Code Division Multiple Access (IR-CDMA) concept [37]. We also consider an interference-limited cooperative system in Chapter 4, where we propose a novel Interleaved Random Space Time Code (IR-STC) [38–41] scheme, which is specifically designed for a Multi-Source Cooperation (MSC) scenario. Furthermore, the network coding concept of [42] is also employed, when designing a coded MSC arrangement for the sake of improving the attainable energy efficiency. More specifically, the proposed Physical-layer Algebraic Network Coding (PANC) scheme implicitly inherits the random coding principle applied in a distributed manner [43, 44]. In Chapter 5, we proceed by proposing a novel SPC aided multiplexed HARQ scheme for substantially improving the overall end-to-end transmission efficiency [45, 46]. This idea may be viewed as an

extension of the network coding concept or be considered as an adoption of the non-orthogonal random waveform based multiuser communications principle.

In parallel to presenting our general framework of non-orthogonal random waveform based multiuser communications, we bear in mind that employing a powerful Multi-User Detector (MUD) for maintaining a near-single-user performance is of high importance. Hence, in Chapter 6 we propose a novel so-called Harmony Search (HS) algorithm [47] for joint iterative channel estimation, data detection and channel decoding. The algorithm proposed can be readily employed to replace all detection algorithms used throughout Chapters 3 to 5, when a high throughput is desired.

The main contributions and the outline of the thesis are highlighted as follows:

In **Chapter** 2, we introduce the fundamental background knowledge required for designing non-orthogonal random waveform based multiuser communications systems. Based on the general linear Gaussian vector model introduced, our framework becomes applicable to a range of communications scenarios. We then focus our attention on three key topics. First and foremost, the multiuser channel capacity is quantified for both a continuous channel input and for a multi-layer mapping aided Gaussian-like channel input. Secondly, the principles of random coding are reviewed, where we take a further step by considering the associated decoding aspects, where a graphical representation is provided and the sum-product algorithm are introduced. With the aid of factor graphs, we are able to classify two important practical quasi-random codes, namely the family of LDPC codes using implicit interleavers and concatenated interleaved random codes combined with explicit interleavers. Finally, we concentrate our attention on the philosophy of iterative receivers, where two typical detection algorithms are reviewed, namely the Bayesian detector and the IC detector. These three key topics cover the main ideas of the thesis and are essential for the deep understanding of the thesis, although a range of diverse further techniques are also required in the following chapters. In summary, the contributions of the chapter are:

Contribution 1 *An unified treatment of the principle of the non-orthogonal random waveform based multiuser communications is offered. Three key topics are revealed, namely the importance of employing SPC aided Gaussian-like channel inputs, the rationale of employing quasi-random codes in multiuser communications scenarios and the feasibility of employing iterative receivers.*

In **Chapter** 3, we apply the principle of non-orthogonal random waveform based multiuser communications in the context of cellular systems using a novel IDMA system. Following the literature review of various aspects of IDMA, we propose the multicarrier version of IDMA, where instead of offering the conventional Bit-Error-Ratio (BER) based performance characterisation, we provide comprehensive design guidelines concerning three aspects, namely the coding versus spreading tradeoff, the multiplexing versus diversity gain tradeoff and the performance versus complexity tradeoff, where a novel EXIT chart based hybrid detector design is employed. In addition to these semi-analytical techniques, we also make a comparison between OFDMA, MC-CDMA and MC-IDMA. The employment of IDMA in the context of HSPA [2] is also characterised. We

continue by proposing a generalisation of the idea, leading to the concept of IR-CDMA, where the system employing implicit interleaver based random codes for uniquely and unambiguously differentiating the users is introduced and a novel memory efficient interleaver generation technique is proposed. In all, the contributions of the chapter are [35, 37, 48]:

Contribution 2 *Three design aspects of MC-IDMA are discussed, where we focus our attention on the coding versus spreading tradeoff, on the multiplexing versus diversity gain tradeoff and on the performance versus complexity tradeoff. In addition, guidelines for designing low-rate codes for equal-power multiuser transceiver are provided and a hybrid detector is proposed. Furthermore, the concept of IR-CDMA is presented with the aid of specifically designed quasi-random codes for the sake of requiring a low memory in multiuser transceivers.*

In **Chapter** 4, we apply the principle of non-orthogonal random waveform based multiuser communications in a cooperative multiple access scenario. Motivated by the concept of distributed space time coding, we propose the so-called IR-STC scheme for achieving both a high diversity gain as well as a high multiplexing gain. More specifically, achieving a diversity gain is guaranteed by the relaying of replicas of the original source signal through geographically distributed channels, while attaining a multiplexing gain is promised by the combination of the SPC concept and with an iterative receiver. More explicitly, we compare the slot efficiency of MSC and single-source cooperation. Then the detailed design of two-phase constructions is discussed. The IR-STC designed is further analysed with the aid of a matrix representation and the achievable performance is compared to the conventional orthogonal STBC and investigated over the Nakagami-m fading channel with the aid of various relaying techniques. Moreover, we also propose two energy efficient coding schemes for MSC, namely a SPC scheme and the PANC scheme. We will show that SPC may be viewed as a direct extension of the conventional IR-STC concept from the uncoded regime to the channel coded regime, while the PANC design inherits the concept of classic multiplexed codes, where multiple information flows are jointly encoded to a large random code ensemble. We thus present the encoding and decoding procedures, as well as the performance comparison of these two coding schemes, respectively. In all, the contributions of the chapter are [38–41, 43, 44]:

Contribution 3 *A distributed IR-STC is designed for cooperative multiple access, which exhibits a high throughput, a high error resilience and a non-orthogonal nature. Furthermore, different relaying techniques are also investigated under the framework of IR-STC. In addition to designing a distributed system architecture, a novel energy efficient PANC scheme is proposed and its achievable performance is compared to that of the conventional SPC scheme used as a benchmarker.*

In **Chapter** 5, the principle of non-orthogonal random waveform based multiuser communications is adopted in a realistic scenario, namely in a practical HARQ type transceiver. We carefully transplant the SPC concept into the design of a HARQ scheme, leading to the so-called SPC aided multiplexed HARQ arrangement. This is motivated by reducing the delay in HARQ aided wireless Transport Control Protocol (TCP) assisted packet based transmissions. After a re-

view of the existing HARQ schemes, we outline the detailed design methodology of our novel M-HARQ scheme. Importantly, the achievable performance is examined from a cross-layer perspective, where the physical layer packet error ratio performance is quantified, which is followed by the associated link layer throughput investigations and TCP layer mean frame arrival rate analysis, where the latter is discussed in detail with the aid of a Markov model. In all, the contributions of the chapter are [45, 46]:

Contribution 4 *A novel SPC aided multiplexed HARQ scheme is proposed for the sake of reducing the delay of HARQ aided wireless TCP packet based transmission. The ultimate end-to-end throughput is substantially improved with the aid of our design. Furthermore, our scheme may be directly implemented in the context of existing systems.*

In **Chapter** 6, we propose a novel so-called HS aided MUD employing an iterative receiver. Inspired by research advances in applied mathematics, this recently emerged meta-heuristic optimisation algorithm is first briefly introduced and then employed in the classic wireless communications MUD problem of a DS-CDMA system. We highlight the original HS algorithm and outline its detailed design in the context of the proposed A Posteriori Probability (APP)-based HS pitch adjustment step of the algorithm. Furthermore, we extend the algorithm to the entire receiver, which includes channel estimation, data detection and channel decoding. In addition, the pseudo-code of our algorithm is provided and a step-by-step example is included. Again, although independent of our discussions concerning the basic principles of non-orthogonal random waveform based multiuser communications, this novel detection algorithm may be employed for solving diverse problems encountered in communications, especially when the problem considered has a large correlated search space, which is indeed the case in non-orthogonal random waveform based multiuser communications. In all, the contributions of the chapter are [49–51]:

Contribution 5 *A novel HS algorithm aided MUD is proposed and extended to the entire iterative receiver chain. We propose the APP-based pitch adjustment technique for improving the original algorithm, which facilitates the convergence of the algorithm in challenging communications problems. Furthermore, our HS algorithm may be seamlessly combined with the powerful Expectation Maximisation (EM) based channel estimation algorithm.*

In Chapter 7, we draw quantitative conclusions, along with suggestions for future research.

Chapter 2

Non-orthogonal Random Waveform Based Multiuser Communications

2.1 Introduction

2.1.1 Three Key Topics and Outline

We commence by introducing the proposed generalized non-orthogonal random waveform based multiuser communications model. The model of our general binary baseband communications system is shown in Fig 2.1. The transmitter uses a codebook $\mathcal{C} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{|\mathcal{C}|}\}$ of codeword length N and cardinality $|\mathcal{C}|$ to transmit the mth codeword $\mathbf{x}_m$ across the channel subject to the channel transition probability $p(\mathbf{y}|\mathbf{x})$, which quantifies the probability of receiving $\mathbf{y}$, when transmitting $\mathbf{x}$. Let $M = \log_2|\mathcal{C}|$ denotes the number of information bits conveyed by $\mathbf{b}_m$, thus the rate of the code is $R = M/N$. Based on the received signal $\mathbf{y}$, the decoder generates an estimate $\hat{\mathbf{x}}_m$ of the correct message. The channel's codeword error probability is $P_e = P\{\hat{\mathbf{x}}_m \neq \mathbf{x}_m\}$.

The discrete time linear vector-based complex-value baseband system model can be written as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{2.1}$$

where the vectors $\mathbf{x} \in \mathbb{C}^{D_i \times 1}$ and $\mathbf{y} \in \mathbb{C}^{D_o \times 1}$ are the channel's input and output of dimensions D_i

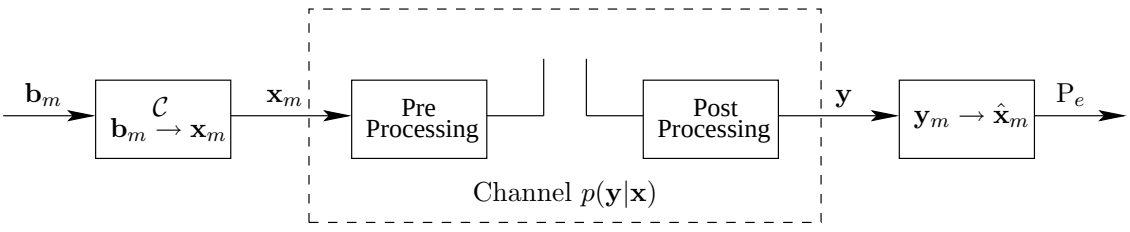


Figure 2.1: The block diagram of an abstract model describing a discrete input memoryless system, where the pre- and post-processing in conjunction with the channel can be considered as an effective channel, which jointly determine the system's transfer function.

and D_o, respectively. Furthermore, $\mathbf{H} \in \mathbb{C}^{D_o \times D_i}$ is the equivalent channel matrix having entries, which are not necessarily i.i.d. Moreover, $\mathbf{n} \in \mathbb{C}^{D_o \times 1}$ is a circularly-symmetric complex-valued Gaussian noise, i.e. we have $n_i \sim \mathcal{CN}(0, N_0)$, which is assumed to be independent of the channel's input $\mathbf{x}$ as well as of $\mathbf{H}$. It is important to emphasise again that the multiplicative factor $\mathbf{H}$ includes any linear pre- or post-processing invoked at the transmitter or receiver in conjunction with the conventional 'channel', hence it may be considered as an *effective* channel[1]. Furthermore, $\mathbf{x}$ may represent a linear or non-linear transformation of the original information sequence $\mathbf{b}$, i.e. we have $\mathbf{x} = \mathcal{F}(\mathbf{b})$. In fact, the transformation $\mathcal{F}$ may be a concatenation of channel coding, DS spreading, modulation etc, we consider only channel coding.

Based on Eq (2.1), the following three key topics are addressed in this chapter:

1. *Capacity bound*: what is the maximal mutual information $I(\mathbf{x}; \mathbf{y})$ between the input and the output of the system? We adopt a fundamental information-theoretic perspective of the multiuser communications channel capacity both for the UL and DL, which allows us to quantify their capacity demonstrating the optimality of non-orthogonal schemes [34]. As another relevant topic, the signal constellation constrained channel capacity is also discussed and beneficial design methods are suggested.

2. *Transmitter side*: what are the desired properties of the channel coding function $\mathcal{F}$ of the system? We introduce the so-called random-coding principle, which was originally used to prove the Shannon-coding theorem [34, 52]. We exemplify several practical quasi-random coding approaches suitable for practical systems with the aid of factor graphs [53].

3. *Receiver side*: what is the best receiver architecture in terms of striking an attractive tradeoff between the computational-complexity imposed and the error-performance attained, while maximising $P(\mathbf{x}|\mathbf{y})$ and $P(\mathbf{b}|\mathbf{y})$. We focus our attention on iterative receiver architectures inspired by the turbo concept [19], where by decomposing the optimum Maximum A posterior Probability (MAP) algorithm introduced-complexity receiver components, a low complexity IC aided detector as well as the optimum Bayesian detector are investigated.

The inter-relations of the above-mentioned three key aspects with Eq (2.1) are portrayed in the stylised illustration of Fig 2.2.

2.1.2 Generalized Linear System Model

In this subsection, we will demonstrate that our system model of Eq (2.1) is generally applicable to a combination of non-orthogonal random waveform based multiuser communications systems, as well as to the MIMO systems and the so-called circulant systems to be described in Section 2.1.2.1,

[1]For instance, in OFDM systems, the *effective* channel may be considered as the combination of the actual channel in conjunction with the IFFT as well as FFT employed at the transmitter and receiver, respectively.

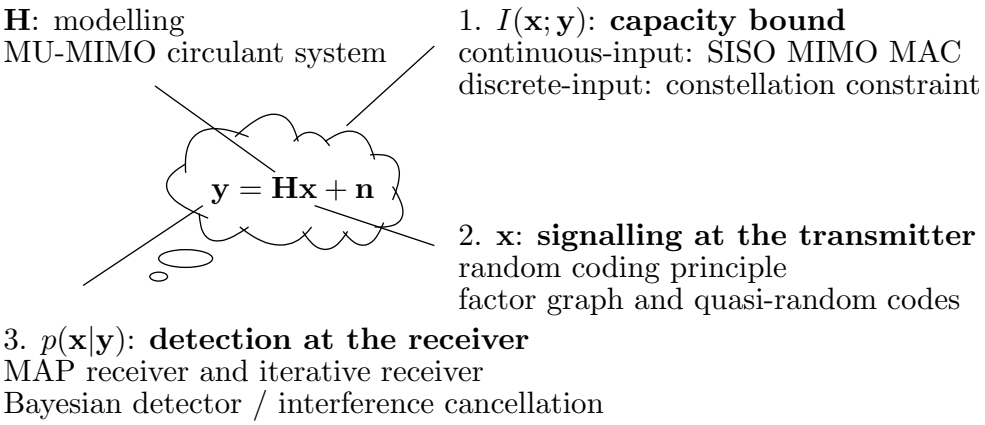


Figure 2.2: The key topics involved in the discussion of non-orthogonal random waveform based multiuser communications, where four aspects are illustrated, namely the system model, the mutual information, the channel coding and the iterative receiver.

leading for example to the well-established Multi-User (MU) MIMO-OFDM system model. This is achieved by appropriately choosing the input and output dimensions D_i and D_o in Eq (2.1). Before further detailing this generalisation, we commence by introducing the circulant system in the context of a Single-User (SU) Single Input Single Output (SISO) system.

2.1.2.1 SU-SISO Circulant System

A square matrix $\mathbf{H}$ of $(M \times M)$ elements is referred to as a circulant matrix, if its ith row / column is constituted by the cyclic shift of its first row / column applied i times, which can be represented as:

$$\mathbf{H} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{M-1} \\ h_{M-1} & h_0 & \cdots & h_{M-2} \\ \vdots & \ddots & \ddots & \vdots \\ h_1 & h_2 & \cdots & h_0 \end{bmatrix}_{M \times M} . \tag{2.2}$$

The square matrix has an EigenValue Decomposition (EVD) expressed as:

$$\mathbf{H} = \mathbf{U}^{\mathrm{H}} \mathbf{\Lambda} \mathbf{U}, \tag{2.3}$$

where $\mathbf{\Lambda}$ is a diagonal matrix hosting the eigenvalues of $\mathbf{H}$ and $\mathbf{U}$ is a unitary matrix whose rows comprise the eigenvectors of $\mathbf{H}$. It is often exploited that the Discrete Fourier Transform (DFT) matrix $\mathbf{F}$ having entries of $w_M = e^{-j2\pi/M}$:

$$\mathbf{F} = \frac{1}{\sqrt{M}} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & w_M & w_M^2 & \cdots & w_M^{M-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 1 & w_M^{M-1} & w_M^{2(M-1)} & \cdots & w_M^{(M-1)^2} \end{bmatrix}_{M \times M} \tag{2.4}$$

is a unitary matrix whose rows comprise the eigenvectors of $\mathbf{H}$, which implies that we have $\mathbf{F} = \mathbf{U}$ and $\mathbf{G} = \mathbf{F}\mathbf{H}\mathbf{F}^{\mathrm{H}} = \mathbf{\Lambda}$.

In communications, the circulant channel matrix $\mathbf{H}$ is encountered in a quasi-static ISI channel after padding the time domain transmitted symbol $\mathbf{x}$ with a Cyclic Prefix (CP) that consists of at least $(L-1)$ quasi-periodically repeated entries of $\mathbf{x}$, where L is the channel's memory length. More specifically, consider the system model of an ISI channel, where the transmitted signal includes a CP. Then the received signal may be represented in the form of:

$$\mathbf{y} = \mathbf{H}_c\mathbf{x}_c + \mathbf{n}, \tag{2.5}$$

where we have $\mathbf{y} \in \mathbb{C}^{M\times 1}$, $\mathbf{n} \in \mathbb{C}^{M\times 1}$ and

$$\begin{aligned} \mathbf{x}_c &= [x_{M-1}, x_{M-2}, \ldots, x_0, x_{M-1}, \ldots, x_{M-L+1}]^{\mathrm{T}}_{(M+L-1)\times 1}, \\ \mathbf{H}_c &= \begin{bmatrix} h_0 & h_1 & \cdots & h_{L-1} & 0 & \cdots & 0 \\ 0 & h_0 & \cdots & h_{L-2} & h_{L-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_0 & \cdots & h_{L-2} & h_{L-1} \end{bmatrix}_{M\times(M+L-1)}, \end{aligned} \tag{2.6}$$

where $\mathbf{x}_c$ hosts M information symbols and $(L-1)$ quasi-periodically repeated CP symbols, while $\mathbf{H}_c$ is the matrix representation of the classic Tapped Delay Line (TDL) model as an instance of a Finite Impulse Response (FIR) filter. At the receiver, after removing the CP, Eq (2.5) becomes equivalent to:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{2.7}$$

where we have $\mathbf{y} \in \mathbb{C}^{M\times 1}$, $\mathbf{n} \in \mathbb{C}^{M\times 1}$ and

$$\begin{aligned} \mathbf{x} &= [x_{M-1}, x_{M-2}, \ldots, x_0]^{\mathrm{T}}_{M\times 1}, \\ \mathbf{H} &= \begin{bmatrix} h_0 & h_1 & \cdots & h_{L-1} & 0 & \cdots & 0 \\ 0 & h_0 & \cdots & h_{L-2} & h_{L-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_0 & \cdots & h_{L-2} & h_{L-1} \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ h_2 & h_3 & \cdots & h_{L-3} & \cdots & h_0 & h_1 \\ h_1 & h_2 & \cdots & h_{L-2} & \cdots & 0 & h_0 \end{bmatrix}_{M\times M}, \end{aligned} \tag{2.8}$$

where $\mathbf{x}$ hosts M number of transmitted symbols only and $\mathbf{H}$ is a circulant matrix, as introduced in Eq (2.2).

2.1.2.2 MU-MIMO Circulant System

Let M denote the total number of subcarriers, N_t and N_r denote the number of transmit and receiver antennas, respectively. As a natural extension of the above-mentioned SU-SISO circulant system, a vector-based circulant K-user MIMO system can also be described as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n},$$

where we have $\mathbf{y} \in \mathbb{C}^{MN_r \times 1}$, $\mathbf{n} \in \mathbb{C}^{MN_r \times 1}$ and,

$$\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K]^{\mathrm{T}}_{MN_tK \times 1}, \tag{2.9}$$

$$\mathbf{x}_k = [\mathbf{x}_{k,M-1}, \mathbf{x}_{k,M-2}, \ldots, \mathbf{x}_{k,0}]^{\mathrm{T}}_{MN_t \times 1}, \tag{2.10}$$

$$\mathbf{x}_{k,m} = [x_{k,m,0}, x_{k,m,1}, \ldots, x_{k,m,N_t-1}]^{\mathrm{T}}_{N_t \times 1}, \tag{2.11}$$

where $\mathbf{x}_{k,m}$ hosts a total of N_t data symbols from the mth subcarrier of user k and $\mathbf{x}_k$ is a vector containing the kth user's data from all subcarriers and all transmitted antennas associated with user k. Furthermore, the effective channel $\mathbf{H}$ becomes:

$$\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_K]_{MN_r \times MN_tK}, \tag{2.12}$$

$$\mathbf{H}_k = \begin{bmatrix} \mathbf{H}_{k,0} & \mathbf{H}_{k,1} & \cdots & \mathbf{H}_{k,M-1} \\ \mathbf{H}_{k,M-1} & \mathbf{H}_{k,0} & \cdots & \mathbf{H}_{k,M-2} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{H}_{k,1} & \mathbf{H}_{k,2} & \cdots & \mathbf{H}_{k,0} \end{bmatrix}_{MN_r \times MN_t}, \tag{2.13}$$

$$\mathbf{H}_{k,m} = \begin{bmatrix} h_{k,m,0,0} & h_{k,m,0,1} & \cdots & h_{k,m,0,N_t-1} \\ h_{k,m,1,0} & h_{k,m,1,1} & \cdots & h_{k,m,1,N_t-1} \\ \vdots & \ddots & \ddots & \vdots \\ h_{k,m,N_r-1,0} & h_{k,m,N_r-1,1} & \cdots & h_{k,m,N_r-1,N_t-1} \end{bmatrix}_{N_r \times N_t}, \tag{2.14}$$

where $\mathbf{H}_{k,m}$ represents the $(N_r \times N_t)$-element MIMO channels of the mth subcarrier of user k, while $\mathbf{H}_k$ represents the kth user's $(N_r \times N_t)$-element MIMO channels of all the M subcarriers.

Thus, it can be seen that the generalized system model of Eq (2.1) is applicable for the circulant complex-valued MU-MIMO system. This implies that we can design a general detection method that can be applied in any system represented by Eq (2.1), such as in SU-MIMO, MU-MIMO, MU-SISO arrangements. In other words, regardless of the specific system considered, the detector aims for determining each entry x_i of the vector $\mathbf{x}$ of Eq (2.9) by maximising $P(x_i|\mathbf{y})$. The detector is capable of operating, while remaining unconscious of the source of x_i, which may represent a specific user's information in MU-SISO system or antenna's information in a SU-MIMO scenario. The receiver's architecture will be discussed in more detail in Section 2.4.

2.2 Channel Capacity of Multiuser Communications

The theoretical optimality of non-orthogonal random waveform based multiuser communications was demonstrated by information theory [34]. This section provides an insight into the ultimate multiuser communications channel capacity and recommends a range of beneficial techniques, such as IC and SPC [22] etc. This section closely follows classic information theory, summarising several important formulas and conclusions. More detailed expositions of information theory can be found in standard textbooks, such as [34].

For the sake of better understanding this section, some preliminary knowledge on entropy and mutual information is necessary, which are the two most important statistical measures of the information contents of messages, as detailed for example in Chapter 1 of [54].

Let x be a random variable with realizations $\tilde{x}$ drawn from a continuous-valued set $\tilde{x} \in \mathcal{S}$ with a Probability Density Function (PDF) of p_x, the so-called *differential entropy* is expressed as:

$$h(x) = -\int_{\mathcal{S}} p_x(\tilde{x}) \log_b[p_x(\tilde{x})] \mathrm{d}\tilde{x}. \tag{2.15}$$

The unit of $h(x)$ is bits, when $b = 2$ is used. It is a measure of the uncertainty associated with a random variable and represents the minimum average message length, in bits, that must be sent to communicate the true value of the random variable to a recipient. A Gaussian distributed process $x \sim \mathcal{N}(\bar{x}, \sigma_x^2)$ results in the maximum differential entropy, where no constraint is imposed on the maximum value of x, which is given by [34]:

$$\max_{p_x} h(x) = \frac{1}{2}\log_2(2\pi e\sigma_x^2) \qquad -\infty < x < +\infty. \tag{2.16}$$

The *mutual information* $I(x;y)$ between two random variables x and y measures the mutual dependence of the two variables, which can be defined as:

$$I(x;y) = h(y) - h(y|x) = h(x) - h(x|y), \tag{2.17}$$

where $h(x)$ and $h(y)$ are the *marginal* entropies, $h(x|y)$ and $h(y|x)$ are the *conditional* entropies. More explicitly, Eq (2.17) measures, how much the explicit knowledge of say x reduces our uncertainty about the value of y, or vice-versa. Using Bayes' rule, we have the following important *chain-rule* for the mutual information:

$$\begin{aligned} I(x_1, x_2; y) &= h(x_1, x_2) - h(x_1, x_2|y) \\ &= h(x_1) + h(x_2|x_1) - h(x_1|y) - h(x_2|x_1, y) \\ &= I(x_1; y) + I(x_2; y|x_1). \end{aligned} \tag{2.18}$$

Eq (2.18) may be interpretated as the information that x_1 and x_2 jointly provide about y, which is equal to the sum of the information x_1 provides about y plus the additional information x_2 provides about y after observing x_1.

2.2.1 Continuous-Input Channel Capacity

2.2.1.1 Capacity of the Scalar AWGN Channel

A real-valued scalar Additive White Gaussian Noise (AWGN) channel having continuous-valued input and output can be simply characterised as:

$$y = x + n, \tag{2.19}$$

where $x \sim p_x$ and $n \sim p_n$ are two *independent* random variables, denoting the channel input and the AWGN process, respectively. The channel capacity C is defined as the maximal mutual information (maximal uncertainty reduction) between the channel's random input variable x and the channel's random output variable y, which is defined as:

$$C = \max_{p_x} I(x;y), \tag{2.20}$$

where we have $I(x;y) = h(y) - h(y|x) = h(y) - h(n)$ and $h(y)$ as well as $h(n)$ represent the entropies corresponding to the observation y and the AWGN process n.

As stated above, the Gaussian distribution maximises the differential entropy (i.e. p_y should be Gaussian) and the sum of two Gaussian random variables (y and n) is a new Gaussian variable ($x = y - n$) with a variance equal to the difference of the two variances ($\sigma_x^2 = \sigma_y^2 - \sigma_n^2$). Then the PDF p_x of the channel input x should also be Gaussian, i.e. we have $x \sim \mathcal{N}(\bar{x}, \sigma_x^2)$. Therefore, the channel capacity of Eq (2.20) can be written as:

$$\begin{aligned} C &= h(y) - h(n) \\ &= \frac{1}{2}\log_2\left[2\pi e(\sigma_x^2 + \sigma_n^2)\right] - \frac{1}{2}\log_2\left[2\pi e\sigma_n^2\right] \\ &= \frac{1}{2}\log_2\left[1 + \frac{\sigma_x^2}{\sigma_n^2}\right], \end{aligned} \tag{2.21}$$

where σ_x^2 is the energy per dimension of the channel's random input variable x (i.e. $E_s = N_D\sigma_x^2$) and $\sigma_n^2 = N_0/2$ is the energy per dimension of the AWGN process n (i.e. $E_n = N_D\sigma_n^2$), where N_D is the dimension of x and n.

Exploiting the fact that the entropy of a multi-dimensional Gaussian PDF having identical variances in all dimensions is the number of dimensions times the entropy of the one-dimensional Gaussian PDF, thus the multi-dimensional AWGN channel's capacity is given by:

$$\begin{aligned} C &= \frac{N_D}{2}\log_2\left[1 + \frac{\sigma_x^2}{\sigma_n^2}\right] \\ &= \frac{N_D}{2}\log_2\left[1 + \frac{2}{N_D}\frac{E_s}{N_0}\right] \\ &= \frac{N_D}{2}\log_2\left[1 + \frac{2}{N_D}\gamma_s\right], \end{aligned} \tag{2.22}$$

where we define the SNR as $\gamma_s = E_s/N_0$. Thus, the corresponding one-dimensional AWGN channel capacity (C_{1D}) and two-dimensional AWGN channel capacity (C_{2D}) are formulated as:

$$C_{1D} = \frac{1}{2}\log_2\left[1 + 2\gamma_s\right], \tag{2.23}$$

$$C_{2D} = \log_2\left[1 + \gamma_s\right]. \tag{2.24}$$

2.2.1.2 Capacity of the MIMO Fading Channel

A MIMO system equipped with N_t transmit antennas and N_r receiver antennas can be characterised as:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{2.25}$$

where $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$ and $\mathbf{x} \in \mathbb{C}^{N_t \times 1}$ are the corresponding channel output and channel input, respectively. Furthermore, $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$ is the effective channel matrix. Both $\mathbf{H}$ and $\mathbf{x}$ are independent of the noise vector $\mathbf{n} \in \mathbb{C}^{N_r \times 1}$ having entries, which are assumed to be i.i.d samples of zero-mean, spatial-and-temporally white circularly-symmetric complex-valued Gaussian random variables, $n_i \sim \mathcal{CN}(0, N_0), i = 1, \ldots, N_r$.

The capacity is again the maximum mutual information $I(\mathbf{x}; \mathbf{y}|\mathbf{H})$ subject to the total transmit power constraint P:

$$\begin{aligned} C &= \max_{p_\mathbf{x}\ s.t. \mathrm{tr}(\tilde{\mathbf{Q}}_\mathbf{x}) \leq P} I(\mathbf{x}; \mathbf{y}|\mathbf{H}) \\ &= \max_{p_\mathbf{x}\ s.t. \mathrm{tr}(\tilde{\mathbf{Q}}_\mathbf{x}) \leq P} h(\mathbf{y}|\mathbf{H}) - h(\mathbf{n}), \end{aligned} \tag{2.26}$$

with $\tilde{\mathbf{Q}}_\mathbf{x} = \bar{\mathbf{x}}\bar{\mathbf{x}}^\mathrm{H} + \mathbf{Q}_\mathbf{x}$, where $\bar{\mathbf{x}} = [\bar{x}_1, \ldots, \bar{x}_{N_t}]^\mathrm{T}$ denotes the mean vector and $\mathbf{Q}_\mathbf{x}$ denotes the covariance matrix. Again, it is widely recognised that the multivariate complex Gaussian random vectors maximise the differential entropy. Hence, a sufficient condition for maximising the mutual information is to use a *zero-mean* multivariate complex Gaussian transmit signal vector $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_\mathbf{x})$, which yields a *zero-mean* multivariate complex Gaussian received signal vector $\mathbf{y} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_\mathbf{y})$. Exploiting the fact that the entropy of any complex random vectors $\mathbf{z} \in \mathbb{C}^N$ with given covariance matrix $\mathbf{Q}_\mathbf{z}$ is maximised by complex circular-symmetry jointly Gaussian and its corresponding maximum value is $\log_2 [\det(\pi e \mathbf{Q}_\mathbf{z})]$, we arrive at:

$$\begin{aligned} h(\mathbf{n}) &= \log_2 \left[(\pi e)^{N_r} \det(N_0 \mathbf{I}_{N_r}) \right] \\ &= \log_2 \left[(\pi e N_0)^{N_r} \right], \end{aligned} \tag{2.27}$$

$$h(\mathbf{y}|\mathbf{H}) = \log_2 \left[(\pi e)^{N_r} \det(\mathbf{Q}_\mathbf{y}) \right], \tag{2.28}$$

where the covariance matrix $\mathbf{Q}_\mathbf{y}$ is expressed as:

$$\mathbf{Q}_\mathbf{y} = \mathbb{E}\left\{\mathbf{y}\mathbf{y}^\mathrm{H}\right\} = \mathbb{E}\left\{\mathbf{H}\mathbf{x}\mathbf{x}^\mathrm{H}\mathbf{H}^\mathrm{H}\right\} + \mathbb{E}\left\{\mathbf{n}\mathbf{n}^\mathrm{H}\right\} = \mathbf{H}\mathbf{Q}_\mathbf{x}\mathbf{H}^\mathrm{H} + N_0\mathbf{I}_{N_r}. \tag{2.29}$$

Thus, by substituting Eq (2.27) and Eq (2.28) into Eq (2.26), the channel capacity is given as:

$$C = \max_{\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_\mathbf{x})\ s.t. \mathrm{tr}(\mathbf{Q}_\mathbf{x}) \leq P} \log_2 \left[\det(\mathbf{I}_{N_r} + \mathbf{H}\mathbf{Q}_\mathbf{x}\mathbf{H}^\mathrm{H}/N_0) \right], \tag{2.30}$$

which means that the MIMO capacity is obtained by solving the optimisation problem of Eq (2.30) subject to the total power constraint.

Let us now differentiate between two capacity definitions stipulated for fading channel conditions. The capacity of the *ergodic* fading MIMO channel is found by averaging the capacity expression in Eq (2.30) with respect to the stationary distribution of $\mathbf{H}$ and then choosing the appropriate covariance matrix to maximise the mutual information subject to the power constraint [2] P, yielding:

$$C = \max_{\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q}_\mathbf{x})\ s.t. \mathrm{tr}(\mathbf{Q}_\mathbf{x}) \leq P} \mathbb{E}_\mathbf{H} \left\{ \log_2 \left[\det(\mathbf{I}_{N_r} + \mathbf{H}\mathbf{Q}_\mathbf{x}\mathbf{H}^\mathrm{H}/N_0) \right] \right\}, \tag{2.31}$$

[2] In our case, a short-term power constraint is considered, while a long-term constraint could also be imposed by taking the expectation over the trace of the selected input covariance matrix.

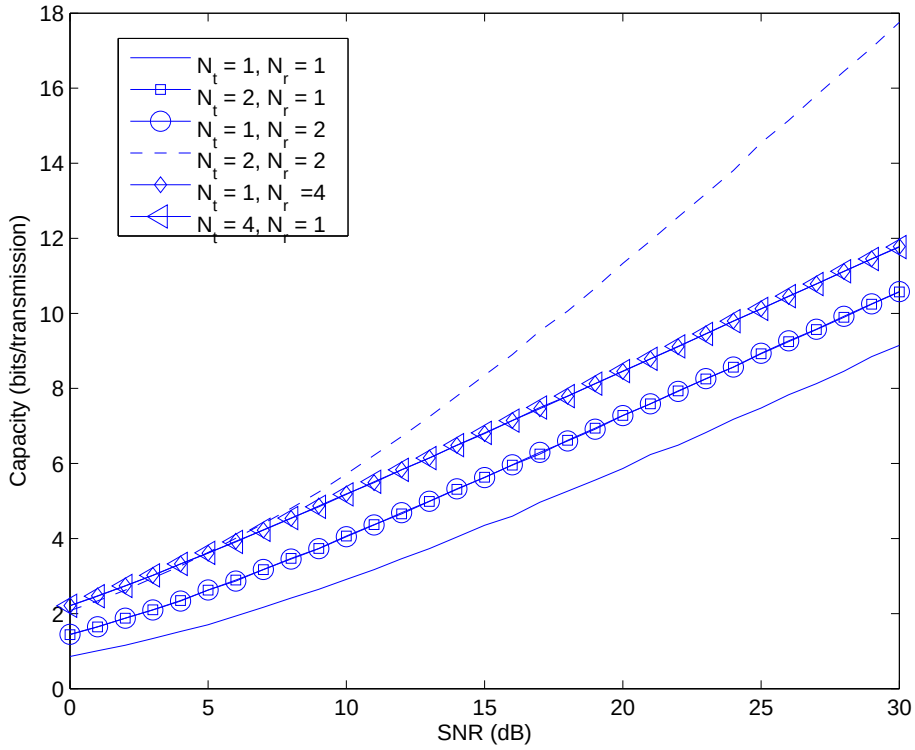


Figure 2.3: Capacity versus SNR for MIMO fading channel with full channel information at both the transmitter and the receiver employing water filling procedure for various combinations of the number of antennas.

where the only difference with respect to Eq (2.30) is the averaging according to the stationary distribution of $\mathbf{H}$. Secondly, the outage probability of the (quasi) *static* fading MIMO channel, which imposes a constant fading envelope for the duration of an entire transmission frame considered [20]. When communicating at a rate of R bits/s/Hz, this outage probability is defined as the probability having a capacity lower than the target value of R, which is formulated as:

$$p_{out} = P\left\{\log_2\left[\det(\mathbf{I}_{N_r} + \mathbf{H}\mathbf{Q}_\mathbf{x}\mathbf{H}^{\mathrm{H}}/N_0)\right] < R\right\}. \tag{2.32}$$

By appropriately adjusting the number of antennas, the above derivation also characterises the capacity of the MISO or SIMO fading channels. The capacity of MIMO fading channel depends on the availability of channel information at the transmitter and/or the receiver. When the channel information is available at both the transmitter and the receiver, the maximum achievable capacity is attained with the aid of the so-called water-filling procedure [22]. Fig 2.3 illustrates the MIMO fading capacity with full channel information at the transmitter as well as at the receiver, again when employing the water filling procedure.

2.2.1.3 Multiuser Channel Capacity

Having introduced the AWGN channel's and the MIMO channel's capacity, let us now discuss the multiuser channel capacity. Without loss of generality, we consider a two-user static channel

characterised by:

$$y = \rho_1 x_1 + \rho_2 x_2 + n, \tag{2.33}$$

where $n \sim \mathcal{CN}(0, N_0)$ is an i.i.d zero-mean circularly-symmetric complex-valued AWGN process and ρ_k is the effective channel gain. Given the independent channel input distributions of p_{x_1} and p_{x_2} of the two users, the capacity of a two-user (UL or DL) channel is described by the so-called *capacity region*, which characterises the optimal tradeoff achievable by *any* multiuser communications scheme that may be defined as the set of all user-rate pairs, satisfying:

$$\begin{aligned} R_1 &< I(x_1; y|x_2), \\ R_2 &< I(x_2; y|x_1), \\ R_1 + R_2 &< I(x_1, x_2; y). \end{aligned}$$

Orthogonal Access. Let us first consider an orthogonal scheme that allocates a fraction α_1 of the degrees of freedom to user 1 and $\alpha_2 = 1 - \alpha_1$, to user 2. Since the effective channel gain of user k is ρ_k , the amount of received energy is $|\rho_k|^2/\alpha_k$ per degree of freedom. More particularly, the following rate pair is achieved by an orthogonal scheme [20]:

$$(R_1, R_2) = \left\{ \alpha_1 \log_2 \left(1 + \frac{|\rho_1|^2}{\alpha_1 N_0}\right), \alpha_2 \log_2 \left(1 + \frac{|\rho_2|^2}{\alpha_2 N_0}\right) \right\}. \tag{2.34}$$

Non-orthogonal Uplink. In the UL, the parameter $\rho_k, k = 1, 2$ of Eq (2.33) represents the effective channel gain of the two independent users. More specifically, we rewrite Eq (2.33) as

$$y = h_1 \sqrt{P_1} x_1 + h_2 \sqrt{P_2} x_2 + n, \tag{2.35}$$

where P_k is the *individual* transmit power constraint for user k, while h_k is the independent channel gain of user k. The typical two-user UL capacity is depicted in Fig 2.4, where the points A and B correspond to a user-rate pair of (R_1, R_2), which are represented as:

$$A \quad (R_1, R_2) = \left\{ \log_2 \left(1 + \frac{P_1 |h_1|^2}{N_0}\right), \log_2 \left(1 + \frac{P_2 |h_2|^2}{P_1 |h_1|^2 + N_0}\right) \right\}, \tag{2.36}$$

$$B \quad (R_1, R_2) = \left\{ \log_2 \left(1 + \frac{P_1 |h_1|^2}{P_2 |h_2|^2 + N_0}\right), \log_2 \left(1 + \frac{P_2 |h_2|^2}{N_0}\right) \right\}. \tag{2.37}$$

The orthogonal and non-orthogonal capacity curves were evaluated from Eq (2.34), Eq (2.36) and Eq (2.37) for $SNR_1 = 5dB$ and $SNR_2 = 10dB$, respectively in Fig 2.4. The sum-rate segment of AB represents the set of optimal operating points of the two-user UL channel. Consider the corner point B for example. At this point, user 1 operates at the rate $I(x_1; y)$. Using the chain rule for the mutual information, we can write $I(x_1, x_2; y) = I(x_1; y) + I(x_2; y|x_1)$. Since the sum rate constraint $I(x_1, x_2; y)$ is tight at the corner point B, user 2 achieves its highest rate of $I(x_2; y|x_1)$. This rate pair can be achieved by a Successive Interference Cancellation (SIC) receiver, which detects user 1 first, treating the signal from user 2 as interference. Next, the SIC detects user 2

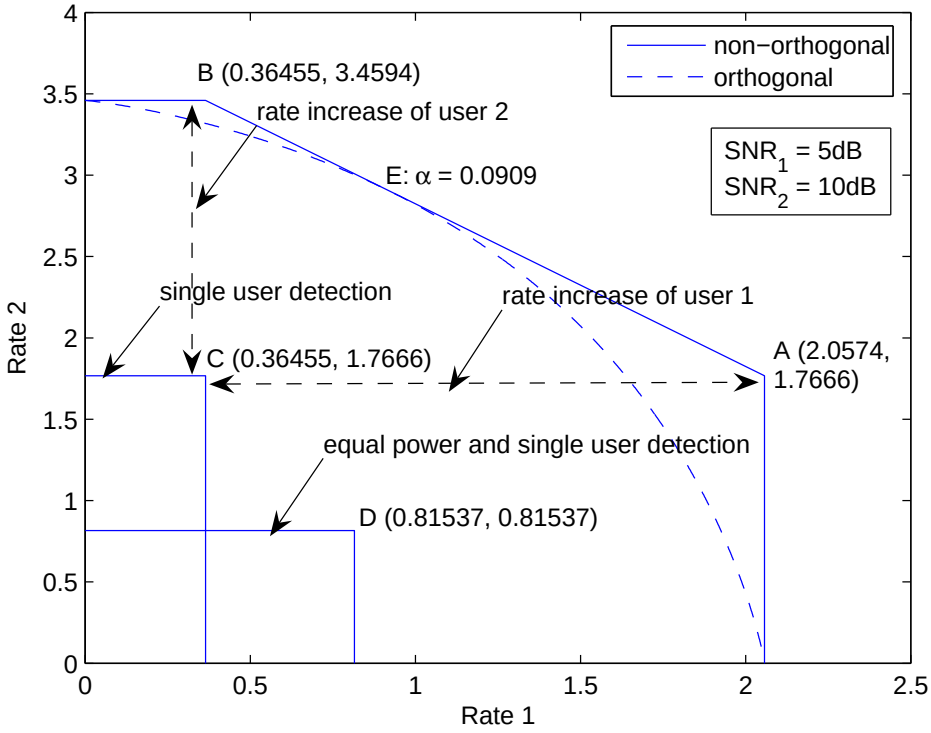


Figure 2.4: The two-user UL channel's capacity. The two users have different receive SNRs of $SNR_1 = \rho_1^2/N_0$ and $SNR_2 = \rho_2^2/N_0$. The points A and B may be approached with the aid of SIC employing non-orthogonal transmission, while the capacity of orthogonal transmission is marginally lower except at point E, where we have $\alpha = 0.0909$. Operation at point C is achieved by conventional single-user detection, while operation at point D is achieved by conventional equal-power allocation employing single-user detection.

conditioned on the already detected information of user 1. Again, the underlined principle is the so-called *chain-rule* of probability, which is often used in information theory.

On the other hand, point C is achieved by the single-user detector, where each user treats the other user's signal as interference. The conventional solution in CDMA represented by point D is to accurately control the power of the users so that we have $\rho_1 = \rho_2$ in Fig 2.4. The corresponding user-rate pairs of the above two points are:

$$C \quad (R_1, R_2) = \left\{ \log_2\left(1 + \frac{P_1|h_1|^2}{P_2|h_2|^2 + N_0}\right), \log_2\left(1 + \frac{P_2|h_2|^2}{P_1|h_1|^2 + N_0}\right) \right\}, \quad (2.38)$$

$$D \quad (R_1, R_2) = \left\{ \log_2\left(1 + \frac{P_2|h_2|^2}{P_2|h_2|^2 + N_0}\right), \log_2\left(1 + \frac{P_2|h_2|^2}{P_2|h_2|^2 + N_0}\right) \right\}. \quad (2.39)$$

Comparing these rates to the capacity region achieved by orthogonal scheme, one can see in Fig 2.4 that the orthogonal scheme is in general suboptimal, except for the single point in the UL, where we have $\alpha_1 = |\rho_1|^2/(|\rho_1|^2 + |\rho_2|^2)$, i.e., where the number of degrees of freedom allocated to each user is proportional to his/her effective received power.

Non-orthogonal Downlink. By contrast, in the DL, the fundamental difference is that both the signal x_1 and x_2 are subjected to the same channel gain, i.e. h_1 or h_2. In order to achieve the best attainable two-user DL channel capacity, an appropriate power splitting is necessary, subject to the

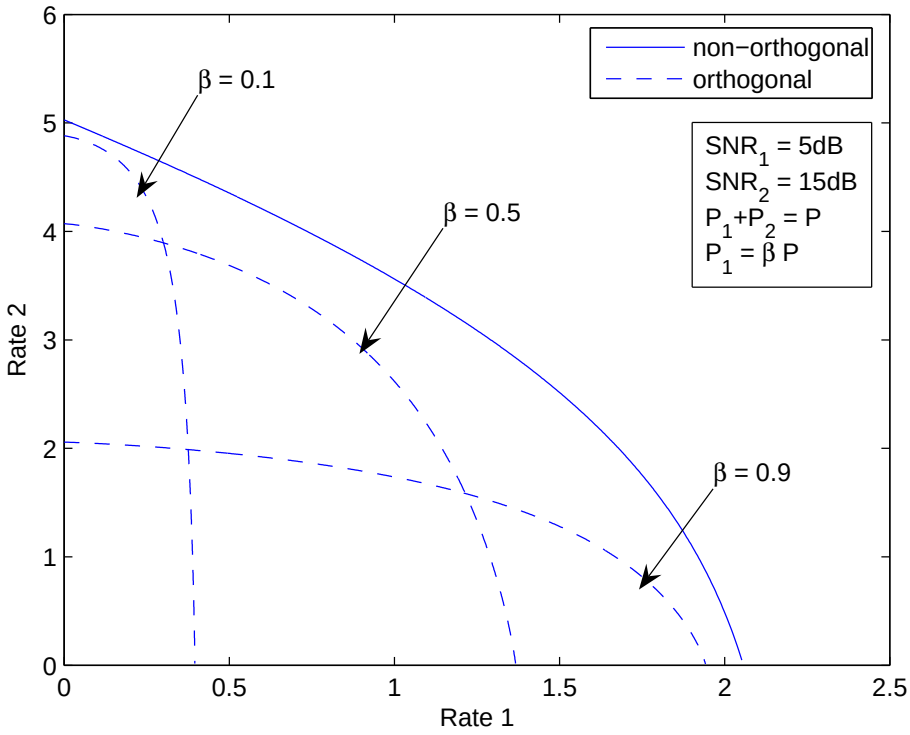


Figure 2.5: The two-user DL channel capacity. Two users have different receive SNRs, $SNR_1 = P|h_1|^2/N_0$ and $SNR_2 = P|h_2|^2/N_0$. The capacity of non-orthogonal transmission is achieved by superposition coding and SIC, while the capacity of orthogonal transmission is suboptimum for all three typical power splitting factor β.

total transmit power constraint of $P = P_1 + P_2$. More specifically, the two channel outputs are:

$$y_1 = h_1(\sqrt{P_1}x_1 + \sqrt{P_2}x_2) + n_1, \tag{2.40}$$

$$y_2 = h_2(\sqrt{P_1}x_1 + \sqrt{P_2}x_2) + n_2. \tag{2.41}$$

As in the UL, we use the chain rule of mutual information and allocate P_1 to user 1, yielding the rate $R_1(P_1) = I(x_1; y)$. Then user 2 achieves his/her highest rate of $R_2(P_2) = I(x_2; y|x_1)$. This linear superposition with an appropriate power splitting used in the DL leads to the concept of SPC.

In the DL, user 1 and user 2 detect his/her data from the observation of y_1 and y_2, respectively. More specifically, if we assume $|h_1| < |h_2|$, user 2 has a better channel than user 1. For user 1, it treats the signal of user 2 as interference and detects his/her data by processing y_1. For user 2, which has a better channel and hence can also detect user 1's data successfully. It detects the data of user 1 and then proceeds to subtract the transmit signal of user 1 from y_2, followed by detecting his/her own data. As before, the following user-rate pair is achieved:

$$(R_1, R_2) = \left\{ \log_2\left(1 + \frac{P_1|h_1|^2}{P_2|h_1|^2 + N_0}\right), \log_2\left(1 + \frac{P_2|h_2|^2}{N_0}\right) \right\}. \tag{2.42}$$

2.2.2 Discrete-Input Channel Capacity

For the sake of approaching the continuous input channel's capacity, it turns out that p_x should obey the continuous Gaussian distribution in order to maximise the differential entropy, as discussed in Section 2.2.1. However, this is generally not the case in real systems, where the alphabet of the channel input $x \in \mathcal{A}$ is a finite discrete symbol alphabet. Hence, the resultant capacity is referred to as the *constellation-constrained* capacity, a terminology, which was originally used by Ungerboeck to motivate the employment of Trellis Coded Modulation (TCM) [55].

2.2.2.1 Constellation-Constrained Capacity

Recall from Eq (2.19), that the constellation-constrained capacity C_c is the mutual information between the *discrete*-input random variable x and the *continuous*-output random variable y, which is defined as:

$$\begin{aligned} C_c &= I(x;y) \\ &= \mathbb{E}\left\{\log_2 \frac{p_{x,y}(\tilde{x},\tilde{y})}{p_x(\tilde{x})p_y(\tilde{y})}\right\} \\ &= \sum_{\tilde{x}\in\mathcal{A}} \int_{-\infty}^{+\infty} \log_2\left(\frac{p_{x,y}(\tilde{x},\tilde{y})}{p_x(\tilde{x})p_y(\tilde{y})}\right) p_{x,y}(\tilde{x},\tilde{y})\mathrm{d}\tilde{y} \\ &= p_x(\tilde{x}_1) \sum_{\tilde{x}_1\in\mathcal{A}} \int_{-\infty}^{+\infty} \log_2\left(\frac{p_{y|x}(\tilde{y}|\tilde{x}_1)}{\sum_{\tilde{x}_2\in\mathcal{A}} p_{y|x}(\tilde{y}|\tilde{x}_2)p_x(\tilde{x}_2)}\right) p_{y|x}(\tilde{y}|\tilde{x}_1)\mathrm{d}\tilde{y} \end{aligned} \tag{2.43}$$

where $\tilde{x}_1$ and $\tilde{x}_2$ are two realizations of x and Eq (2.43) exploits the fact that:

$$\begin{aligned} p_y(\tilde{y}) &= \sum_{\tilde{x}\in\mathcal{A}} p_{x,y}(\tilde{x},\tilde{y}), \\ p_{x,y}(\tilde{x},\tilde{y}) &= p_{y|x}(\tilde{y}|\tilde{x})p_x(\tilde{x}). \end{aligned}$$

Since x and n in Eq (2.19) are independent, the conditional PDF $p_{y|x}(\tilde{y}|\tilde{x})$ can be expressed as the PDF of the noise:

$$p_{y|x}(\tilde{y}|\tilde{x}) = p_n(\tilde{y}-\tilde{x}) = p_n(\tilde{n}). \tag{2.44}$$

where $p_n(\tilde{n})$ represents a one-dimensional ($N_D = 1$) or two-dimensional ($N_D = 2$) Gaussian distributed PDF associated with $\sigma_n^2 = N_0/2$ per dimension:

$$p_n(\tilde{n}) = \frac{1}{(\pi N_0)^{\frac{N_D}{2}}} e^{-\frac{|\tilde{n}|^2}{N_0}}. \tag{2.45}$$

Consider a constellation associated with equal-probability symbols. Then the input x is uniformly distributed in $\mathcal{A}$, i.e. we have $p_x(\tilde{x}) = 1/|\mathcal{A}|$. Let $m = \log_2(|\mathcal{A}|)$ denote the number of bits per symbol. Then Eq (2.43) becomes:

$$C_c = m - \frac{1}{|\mathcal{A}|} \sum_{\tilde{x}_1\in\mathcal{A}} \int_{-\infty}^{+\infty} \log_2\left(\frac{\sum_{\tilde{x}_2\in\mathcal{A}} p_{y|x}(\tilde{y}|\tilde{x}_2)}{p_{y|x}(\tilde{y}|\tilde{x}_1)}\right) p_{y|x}(\tilde{y}|\tilde{x}_1)\mathrm{d}\tilde{y}. \tag{2.46}$$

The integral in Eq (2.46) is identical for all values of $\tilde{x}_1 \in \mathcal{A} = \{x_1, \ldots, x_{|\mathcal{A}|}\}$, thus when opting for $\tilde{x}_1 = x_1$ the constellation-constrained capacity can be expressed as:

$$\begin{aligned} C_c &= m - \int_{-\infty}^{+\infty} \log_2 \left(\frac{\sum_{\tilde{x}_2 \in \mathcal{A}} p_{y|x}(\tilde{y}|\tilde{x}_2)}{p_{y|x}(\tilde{y}|x_1)} \right) p_{y|x}(\tilde{y}|x_1) \mathrm{d}\tilde{y} \\ &= m - \int_{-\infty}^{+\infty} \log_2 \left(1 + \sum_{\tilde{x}_2 \in \mathcal{A}_{(x_1)}} \frac{p_{y|x}(\tilde{y}|\tilde{x}_2)}{p_{y|x}(\tilde{y}|x_1)} \right) p_{y|x}(\tilde{y}|x_1) \mathrm{d}\tilde{y}, \end{aligned} \tag{2.47}$$

where $\mathcal{A}_{(x_1)} = \{x_2, \ldots, x_{|\mathcal{A}|}\}$ is the subset of $\mathcal{A}$ excluding element x_1.

2.2.2.2 EXtrinsic Information Transfer Charts

Let us now introduce the so-called EXIT charts introduced by ten Brink [36], which is widely used in the analysis of turbo-like iterative systems. This technique relies on computing the mutual information of inner constituent component and outer constituent component, namely the I^a_{inner}, I^e_{inner}, I^a_{outer}, I^e_{outer}, where the superscript a and e denotes the *a priori* information and *extrinsic* information, respectively.

We now relate the mutual information I associated with a specific Log-Likelihood Ratio (LLR) by considering again $y = x + n$, where $x \in \mathcal{A} = \{\pm 1\}$ and $n \sim \mathcal{N}(0, 2\sigma_n^2)$ [3]. The LLR of y is given by:

$$\begin{aligned} \mathcal{L}_y &= \log_2 \frac{p(y|x=+1)}{p(y|x=-1)} \\ &= \frac{2}{\sigma_n^2}(x+n) \\ &= \mu x + w, \end{aligned} \tag{2.48}$$

where we have $\mu = 2/\sigma_n^2$, $w \sim \mathcal{N}(0, 2\sigma_w^2)$ with $\sigma_w^2 = 4/\sigma_n^2$. It was suggested in [36] that the LLR values may be assumed to be Gaussian distributed and may be modelled in general by:

$$\mathcal{L} = \mu x + w, \tag{2.49}$$

where $x \in \mathcal{A} = \{\pm 1\}$ and we have the relationship $\mu = \sigma_w^2/2$ and the conditional probability of

$$p(\mathcal{L}|x=\alpha) = \frac{1}{\sqrt{2\pi}\sigma_w} \exp\left[-\frac{(\mathcal{L} - \mu\alpha)^2}{2\sigma_w^2} \right], \tag{2.50}$$

where $\alpha \in \mathcal{A}$. Hence, the mutual information $I(\mathcal{L}, x)$ is derived from Eq (2.47) as follows:

$$\begin{aligned} I(\mathcal{L}, x) &= 1 - \int_{-\infty}^{+\infty} \log_2 \left[1 + \frac{p(\mathcal{L}|x=-1)}{p(\mathcal{L}|x=1)} \right] p(\mathcal{L}|x=1) \mathrm{d}\mathcal{L}, \\ &= 1 - \frac{1}{\sqrt{2\pi}\sigma_w} \int_{-\infty}^{+\infty} \log_2 \left[1 + \exp\left(-\frac{2}{\sigma_w^2}\mu\mathcal{L} \right) \right] \exp\left[-\frac{(\mathcal{L}-\mu)^2}{2\sigma_w^2} \right] \mathrm{d}\mathcal{L}, \\ &= 1 - \frac{1}{\sqrt{2\pi}\sigma_w} \int_{-\infty}^{+\infty} \log_2 \left[1 + \exp\left(-\mathcal{L} \right) \right] \exp\left[-\frac{(\mathcal{L}-\mu)^2}{2\sigma_w^2} \right] \mathrm{d}\mathcal{L}. \end{aligned} \tag{2.51}$$

[3] Here, we do not differentiate between *a priori* information and *extrinsic* information at this stage of the derivation.

For the sake of compactness, we introduce the monotonically increasing J-function and its unique inverse to represent Eq (2.51) as follows:

$$\begin{aligned} I &= J(\sigma = \sigma_w), \\ \sigma &= J^{-1}(I), \end{aligned}$$

where the J-function can be closely approximately by numerical optimisation, yielding $J(\sigma) \approx (1 - 2^{-H_1\sigma^{2H_2}})^{H_3}$ and $J^{-1}(I) \approx [-\log_2(1 - I^{1/H_3})/H_1]^{1/2H_2}$, where we have $H_1 = 0.3073$, $H_2 = 0.8935$ and $H_3 = 1.1064$.

The EXIT charts evaluate the nonlinear function $I^e = \chi(I^a)$, which maps the input *a priori* mutual information $I^a \in [0, \cdots, 1]$ to the output extrinsic mutual information $I^e \in [0, \cdots, 1]$, where the amount of output extrinsic mutual information I^e gleaned from the input *a priori* mutual information determines the convergence behaviour of this soft component. Since the extrinsic information generated by the first component acts as the *a priori* information for the second component and vice versa, in the EXIT charts we alternately swap the abscissa and ordinate axes, depending on which of the two components acts as the source of *a priori* information, corresponding to the abscissa, as discussed in [36]. To elaborate a little further, EXIT charts provide us with an insight into the mutual information exchange between the constituent soft components. More detailed discussions about EXIT charts in in Chapter 3

2.2.2.3 Constellation-Constrained Capacity of Multilayer Mapping

Apart from classic modulation schemes, there is a less conventional modulation scheme, which is of much interest in this thesis, namely the so-called *multilayer-mapping* defined as:

$$x = \sum_{l=1}^{L} x_l \tag{2.52}$$

where $x_l \in \mathcal{A}_l$ is referred to as a symbol and $x \in \mathcal{A}$ is termed as a super-symbol. Without loss of generality, we assume that the symbol alphabet is the same for all the L layers, i.e. $\mathcal{A}_l = \mathcal{A}_0, \forall l = 1, \ldots, L$.

The channel capacity of multilayer mapping is also given by Eq (2.43). There are two important features of multilayer mapping:

1. The *non-equiprobable* nature of the signal points of the super-symbol $x \in \mathcal{A}$. Since the entropy is maximised by equiprobable symbols, this property is not beneficial from the entropy maximisation point of view. At high SNR, where the noise can be neglected, the mutual information is equal to the entropy of the channel input, i.e. we have $I = h(x)$, where $h(x)$ can be expressed as:

$$h(x) = -\sum_{\tilde{x} \in \mathcal{A}} p_x(\tilde{x}) \log_2[p_x(\tilde{x})], \tag{2.53}$$

   Eq (2.53) exhibits its maximum value of $h(x) = \log_2 |\mathcal{A}|$, when we have $p_x(\tilde{x}) = 1/|\mathcal{A}|$.

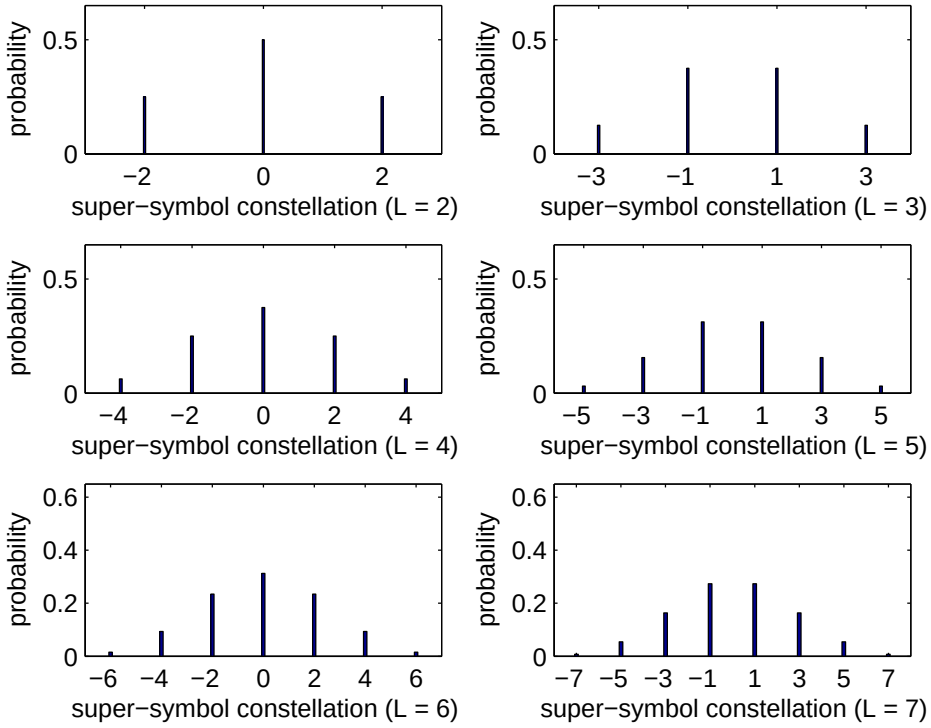


Figure 2.6: The super-symbol alphabet $\mathcal{A} = \left\{u_1, \ldots, u_{|\mathcal{A}|}\right\}$ and its corresponding probability for different number of layers, where each layer employs BPSK modulation $|\mathcal{A}_0| = \{\pm 1\}$.

2. Multilayer mapping has a *reduced cardinality*, where the super-symbol set obeys $|\mathcal{A}| \leq 2^L$, when we assume $\mathcal{A}_0 \in \{\pm 1\}$. The second property will make the bit-to-symbol mapping ambiguous, where several combinations of $x_l, l = 1, \ldots, L$ result in the same constellation x. These two undesired properties of multilayer mapping are further exemplified below.

Example. Consider a BPSK modulated multilayer mapping scheme having L layers, where $\mathcal{A}_0 = \{\pm 1\}$. Then the super-symbol alphabet $\mathcal{A} = \left\{u_1, \ldots, u_{|\mathcal{A}|}\right\}$ in this case has $|\mathcal{A}| = L + 1$ distinct constellation points with its ith entry being $u_i = (2i - 2 - L)$ with a probability of [56]:

$$P(x = u_i) \quad = \quad 2^{-L} \begin{pmatrix} L \\ i-1 \end{pmatrix}. \tag{2.54}$$

Fig 2.6 shows the super-symbol alphabet and the corresponding probability. When we have $L = 3$, there are $L + 1 = 4$ distinct points with unequal probabilities. In this non-equiprobable scenario, the corresponding maximum entropy at high SNRs according to Eq (2.53) is given by:

$$I = L - 2^{-L} \sum_{i=1}^{L-1} \begin{pmatrix} L \\ i \end{pmatrix} \log_2 \begin{pmatrix} L \\ i \end{pmatrix}. \tag{2.55}$$

Fig 2.7 shows the maximum possible entropy at high SNRs as a function of the number of layers L. It can be seen that $L = 3$ layers can at most transmit about 1.8 bits/super-symbol, since only four distinct constellation points are observed. This is strictly lower than the sum of the individual layers' rate, which is 3 bits/super-symbol, corresponding to eight distinct signal constellation points.

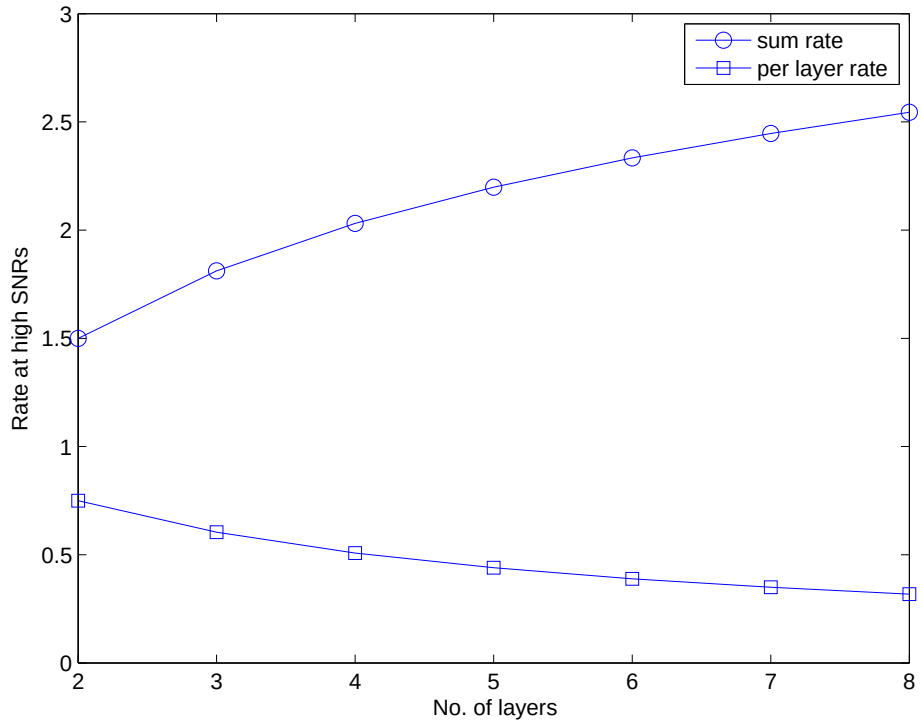


Figure 2.7: The multilayer mapping constellation constraint capacity at high SNR as a function of number of layers, where each layer employs BPSK modulation.

Based on the above discussions, we conclude that the *multilayer mapping scheme's constellation-constrained capacity is bounded by the sum of the individual rates of each layer. The actual capacity depends on the number of distinct signal constellation points in $\mathcal{A}$ and its associated probability of occurrence.*

2.2.2.4 Comments and Remarks on Multilayer Mapping

Although the multilayer mapping scheme's constellation-constrained capacity is reduced owing to the above-mentioned unfavourable properties discussed in the previous section, it also has a particularly attractive property, namely that the resultant super-symbol x is approximately Gaussian distributed, when the number of layers L becomes sufficiently high, which is a consequence of the central limited theorem. For the sake of achieving a high capacity for a multilayer mapping scheme, firstly we have to increase its cardinality $|\mathcal{A}|$ and then make the supersymbols equiprobable. Two simple operations are thus necessary before multilayer mapping takes place, namely layer-specific amplitude scaling factor $\rho_l, l = 1, \ldots, L$ and/or layer-specific signal rotation $\theta_l, l = 1, \ldots, L$. More explicitly, Eq (2.52) becomes:

$$x = \sum_{l=1}^{L} \rho_l e^{j\theta_l} x_l. \tag{2.56}$$

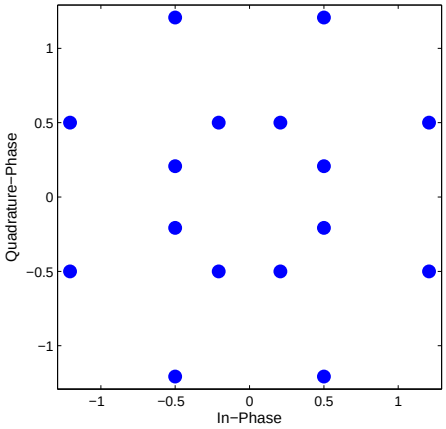
Let us now reconsider the BPSK example of Fig 2.6. If we assign a unique layer-specific phase rotation θ_l to each layer, then a total of $|\mathcal{A}| = 8$ equiprobable distinct signal constellations can

be created, thus a sum-rate of 3 bits/super-symbol is achieved as a maximum. Fig 2.8 illustrates the resultant multilayer mapping constellation having $L = 4$, when we have different amplitude scaling factors $\rho_l, l = 1, \ldots, L$ and phase rotations $\theta_l, l = 1, \ldots, L$.

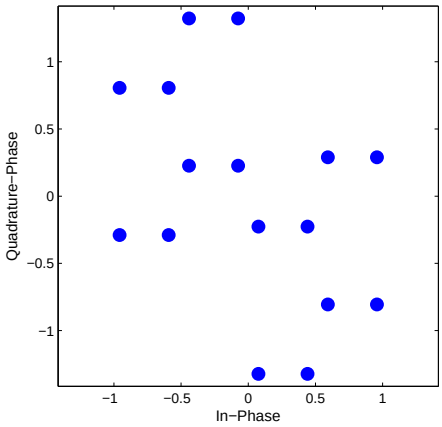
Multilayer mapping, which is also known as *superposition modulation* [57] and *sigma mapping* [58], packs multiple bits (or symbols) into a single symbol (or super-symbol) by weighting and superimposing the multiple constituent bits (or symbols), where we refer to each bit (or symbol) as a *layer*. Again, by superimposing multiple layers, the resultant symbol (super-symbol) becomes near-Gaussian distributed, when the number of layers is sufficiently high. This Gaussian distribution is the key to approaching the Shannon capacity, as discovered by Shannon and discussed before. This signalling scheme achieves a so-called *shaping gain* [11] in contrast to the conventional QAM/PSK modulation.

By allocating the layer-specific amplitude ρ_l and phase-rotation parameter θ_l, the multilayer mapping in conjuction with the coding and interleaving employed becomes analogous to the so-called MultiLevel Coding (MLC) [59, 60] concept as seen in Fig. 2.9(a). In MLC, each bit level employs a different code-rate, such as C_1, C_2 and C_3 in Fig 2.9(a) and is then modulated using a conventional QAM/PSK arrangement combined with a particular bit-to-symbol mapping scheme, such as for example set-partitioning. By contrast, in multilayer mapping, each layer is superimposed employing a different amplitude and phase-rotation. The outer channel code can be employed on a per-layer basis as seen in Fig. 2.9(b) or across layers as seen in Fig. 2.9(d). A similarity of these two schemes is constituted by the unequal error protection capability of the different levels/layers, where unequal protection is achieved by assigning the appropriate code-rate of C_1, C_2 and C_3 as in MLC, while it is achieved by appropriate amplitude scaling in multilayer mapping. In addition, in order to fully exploit the benefits of unequal error protection, while still maintaining a sufficiently high performance even for the least-protected levels/layers, both schemes rely on an iterative detection and decoding aided receiver, where the successive detection of the individual layers obeys the so-called chain-rule of mutual information exchange, which is another similarity. On the other hand, the so-called BICM [21, 61] scheme of Fig. 2.9(c) is similar to the multilayer mapping type II portrayed in Fig. 2.9(d), when the amplitude scaling factor ρ_l is identical.

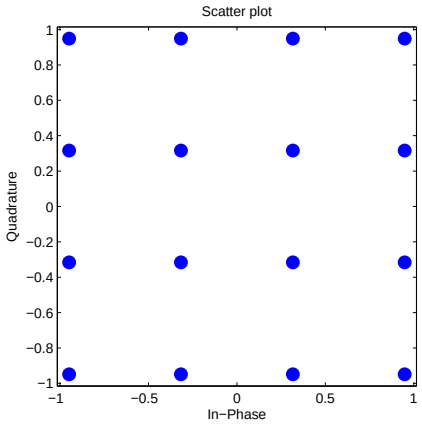
Remarks: It is interesting to note that Eq (2.56) effectively describes an ensemble of systems, which are channel-division, e.g. Spatial Division Multiple Access (SDMA), where the weighting factor $\rho_l e^{j\theta_l}$ can be considered as a unique *virtual* channel. By contrast, Eq (2.52) describes an ensemble of systems, which are code-division, e.g. IDMA, where the differentiation of users and/or signalling layers is based on different interleavers. In other words, *introducing a 'channel-division' element to the 'code-division' is beneficial.* This argument suggests that achieving the maximum capacity is promised by Gaussian-like signalling waveforms, but no explicit coding and detection schemes were considered. When practical iterative detection and decoding techniques are considered, *combining the principle of 'code-division' with that of 'channel-division' is also beneficial.* This is because, at the first stage of the receiver, having distinct *channel* information is essential in



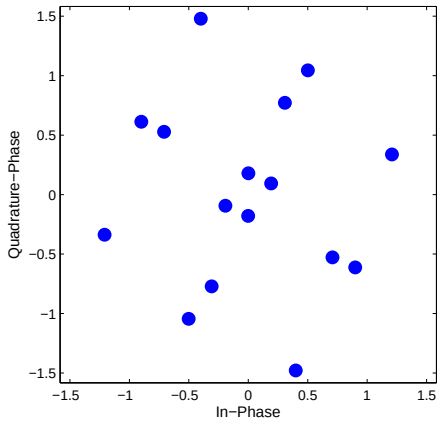
(a) equal ρ and uniform θ



(c) unequal ρ and uniform θ

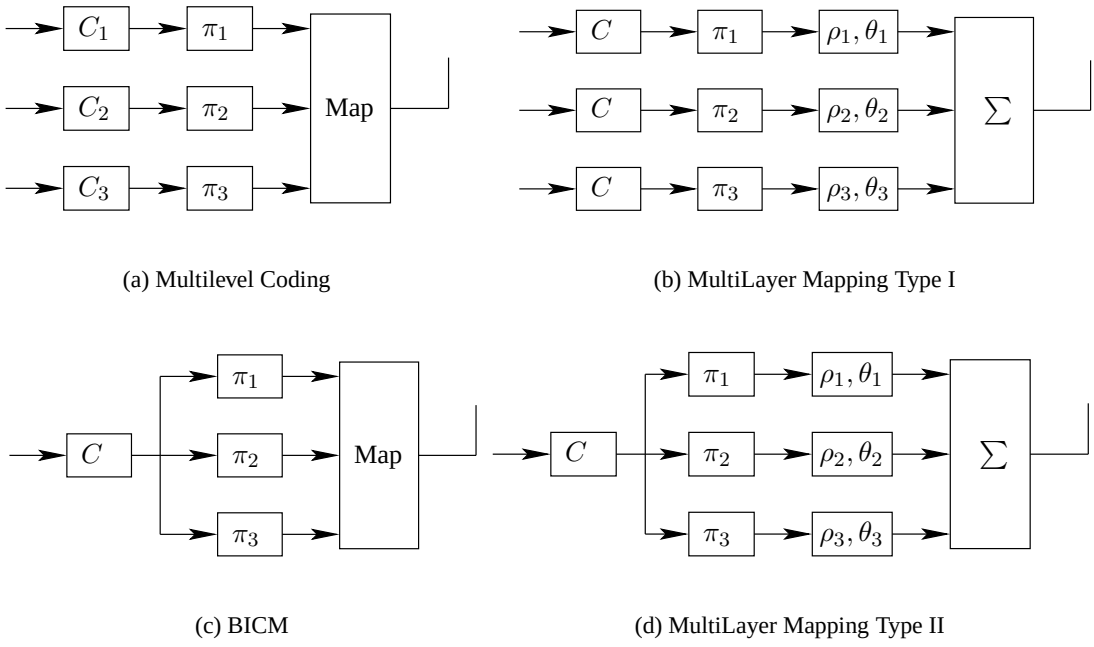


(b) designed ρ and designed θ



(d) equal ρ and random θ

Figure 2.8: Constellation of multilayer mapping associated with different ρ and θ values. (a) equally allocated amplitude ρ and uniformly allocated phase θ; (b) $\rho = [\rho_b, \rho_b/2, \rho_b, \rho_b/2]$ and $\theta = [0, \pi/2, 0, \pi/2]$, where $\rho_b = 2/\sqrt{(16-1)\times 2/3}$; (c) uniformly allocated phase θ and unequally allocated amplitude $\rho = [1/\sqrt{30}, 2/\sqrt{30}, 3/\sqrt{30}, 4/\sqrt{30}]$; (d) equally allocated amplitude ρ and randomly allocated phase $\theta = [2\pi/5, 3\pi/4, 6\pi/7, \pi/3]$;



(a) Multilevel Coding

(b) MultiLayer Mapping Type I

(c) BICM

(d) MultiLayer Mapping Type II

Figure 2.9: The block diagram of the transmitter of (a) Multilevel Coding, using per-layer codes and conventional modulation; (b) Multilayer Mapping Type I, using per-layer codes and superposition modulation; (c) BICM, joint coding and conventional modulation; (d) Multilayer Mapping Type II, joint coding and superposition modulation;

order to facilitate a unique mapping of $y \rightarrow x_l$, when the full alphabet has 2^L legitimate signalling combinations. At the second receiver stage, having a distinct *code* ensures that a unique mapping from the coded bits to the original information bits can take place. The construction of the codes will be discussed in the next section.

2.3 Factor Graph Interpretation of Random Coding

The design of near-capacity transmission arrangements typically relies on channel-coding schemes. A probabilistic approach is that of using a technique referred to as random coding, which was used to prove the well-known Shannon-coding theorem [34, 62]. The Shannon-coding theorem states that for any rate $R < C$, there exits at least one sequence $(2^{NR}, N)$ for the memoryless channel such that the probability of error $P_e \rightarrow 0$ as the block length $N \rightarrow \infty$. Conversely, any sequence of $(2^{NR}, N)$ associated with $P_e \rightarrow 0$ must have $R < C$. The detailed proof of Shannon's coding theorem can be found in various standard textbooks [34, 62]. As a result of this proof, there is a high chance that this randomly generated codebook is indeed that of a near-capacity code, provided that we have long block length N.

However, this random coding principle does not specify an explicit way of constructing an efficient coding and decoding procedure. Hence, other types of codes, which have attractive practical implementations have been proposed, such as the family of algebraic codes [6]. It is important

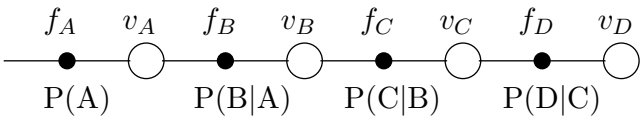


Figure 2.10: Graphical representation of a four-state Markov chain.

to note that two popular channel coding schemes, namely LDPC codes and concatenated codes, are both based on this random coding principle, where random interleavers are either implicitly or explicitly embedded in the code's construction.

In this section, we introduce two practical quasi-random coding approaches, namely LDPC codes and interleaved random codes, which are presented with the aid of the so-called factor graph [53].

2.3.1 Factor Graph and Sum-Product Algorithm

Belief or probability propagation has applications in diverse fields, where probabilistic inference is used, hence also in error control decoding, where the probabilistic value of the transmitted information symbols has to be inferred from noisy received symbols. It has been shown that the APP algorithm based turbo decoding algorithm [63, 64] is an application of belief propagation. In [53], Kschischang *et al.* introduced the factor graph formalism, showing that belief propagation and many other algorithms used in digital communications and signal processing are all representations of a more general message-passing algorithm, namely the *sum-product* algorithm operating on *factor graphs*.

In this subsection, we firstly show the rationale of factor graphs using a simple graphical representation of the Markov chain example of [65] and then introduce the general factor graph philosophy as well as the associated sum-product algorithm. Furthermore, we introduce binary factor graphs in the particular context of channel decoding.

2.3.1.1 General Factor Graph and Sum-Product Algorithm

Factor graphs are capable of characterising how variables of a global function are grouped together locally. The sum-product algorithm computes the so-called marginal probability of a global probabilistic function in terms of local functions, which will be demonstrated in the Markov model example. Normally, factor graphs are bipartite graphs; that is, their nodes fall into just two groups and there are no connections between the nodes within a specific group, only between nodes of the different groups. These groups are referred to as *variable* nodes and *function* nodes.

Example Consider a four-state Markov chain, where we have:

$$P(A,B,C,D) = P(A)P(B|A)P(C|B)P(D|C). \tag{2.57}$$

The corresponding factor graph is depicted in Fig 2.10, where the variable nodes are v_A,v_B,v_C,v_D, which are denoted by circles, while the function nodes representing the probabilistic dependence are $f_A = P(A)$, $f_B = P(B|A)$, $f_C = P(C|B)$ and $f_D = P(D|C)$ and are denoted by dots. Assume that we wish to calculate the marginal probability of $P(D)$, given by:

$$\begin{aligned} P(D) &= \sum_A \sum_B \sum_C P(A,B,C,D) \\ &= \sum_A \sum_B \sum_C P(A)P(B|A)P(C|B)P(D|C), \end{aligned} \tag{2.58}$$

where the 'brute-force' evaluation requires $|\mathcal{A}_A| \times |\mathcal{A}_B| \times |\mathcal{A}_C| \times |\mathcal{A}_D|$ summations, where $|\mathcal{A}_k|$ is the cardinality of variable k's alphabet $\mathcal{A}_k, k = A, B, C, D$. The factorisation of the probabilities now allows us to execute the summations separately according to:

$$\begin{aligned} P(D) &= \sum_C P(D|C)P(C), \\ P(C) &= \sum_B P(C|B)P(B), \\ P(B) &= \sum_A P(B|A)P(A). \end{aligned}$$

where this factorised evaluation has only $|\mathcal{A}_A| \times |\mathcal{A}_B| + |\mathcal{A}_B| \times |\mathcal{A}_C| + |\mathcal{A}_C| \times |\mathcal{A}_D|$ summations, which is computationally more efficient. The *variable* nodes, having two connecting edges, simply pass messages from the input to the output, while the *function* nodes subject the incoming messages to their specific function and sum over the variable representing the incoming message.

We now introduce a general factor graph, which is depicted in Fig 2.11, where the circles denote variable nodes and the squares denote function nodes. Let $\mathbf{x} = [x_1, \ldots, x_n]$ be a set of i.i.d. random variables. Then a typical statistical inference problem is to compute the marginal Probability Mass Function (PMF) of x_i:

$$P(x_i) = \sum_{\mathbf{x}:x_i} P(\mathbf{x}), \tag{2.59}$$

where the sum over $\mathbf{x} : x_i$ indicates that we sum over all elements in $\mathbf{x}$, given that its ith entry was fixed to x_i. This summation can be carried out by the sum-product algorithm in terms of passing messages along the edges of the factor graph seen in Fig 2.11. We have two sets of nodes, which are the function nodes f_j and variable nodes v_i. The set of edges connected to the jth function node is denoted by S_j^f and the set of edges connected to the ith variable node is denoted by S_i^v. The cardinality $d_j^f = |S_j^f|$ of the edges connected to the jth function node and $d_i^v = |S_i^v|$ of the edges connected to the ith variable node is called the *degree* of function node j and of variable node i, respectively. The jth function node carried out summation over all elements of $\mathbf{x}^j$, while the i variable node applies similar local processing on x_i. Then the following general rule applies:

$$\begin{aligned} m(f_j \to v_i) &\propto \sum_{\mathbf{x}^j:x_i} f_j(\mathbf{x}^j) \prod_{k \in S_{j(i)}^f} m(v_k \to f_j), \\ m(v_i \to f_j) &\propto v_i(x_i) \prod_{k \in S_{i(j)}^v} m(f_k \to v_i), \end{aligned}$$

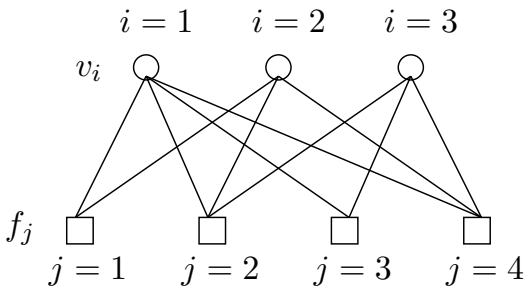


Figure 2.11: General factor graph, which has three variable nodes (circles) and four function nodes (squares), where the corresponding parameters are summarised in Table 2.3.1.1.

v_i	$i = 1$	$i = 2$	$i = 3$
d_i^v	$d_1^v = 4$	$d_2^v = 3$	$d_3^v = 3$
S_i^v	$S_1^v = \{f_1, f_2, f_3, f_4\}$	$S_2^v = \{f_1, f_2, f_4\}$	$S_3^v = \{f_2, f_3, f_4\}$

f_j	$j = 1$	$j = 2$	$j = 3$	$j = 4$
d_j^f	$d_1^f = 2$	$d_2^f = 3$	$d_3^f = 2$	$d_4^f = 3$
S_j^f	$S_1^f = \{v_1, v_2\}$	$S_2^f = \{v_1, v_2, v_3\}$	$S_3^f = \{v_1, v_3\}$	$S_4^f = \{v_1, v_2, v_3\}$
$\mathbf{x}^j$	$\mathbf{x}^1 = [x_1, x_2]$	$\mathbf{x}^2 = [x_1, x_2, x_3]$	$\mathbf{x}^3 = [x_1, x_3]$	$\mathbf{x}^4 = [x_1, x_2, x_3]$

Table 2.1: Parameters of Fig. 2.11.

where $\propto$ indicates proportionality. The sum over $\mathbf{x}^j : x_i$ denotes that we sum over all variables in $\mathbf{x}^j$ where the ith entry is x_i. The set $S_{j(i)}^f$ of edges is given by all edges in set S_j^f except for the element i, which hence represents the *extrinsic* information. Finally, the marginal PMFs at variable node i and function node j are:

$$\begin{aligned} m(x_i) &\propto v_i(x_i) \prod_{j \in S_i^v} m(f_j \to v_i), \\ m(\mathbf{x}^j) &\propto f_j(\mathbf{x}^j) \prod_{i \in S_j^f} m(v_i \to f_j). \end{aligned}$$

2.3.1.2 Binary Factor Graph

In 1981, Tanner [66] characterised both block and convolutional error-correcting codes in terms of their bipartite graphs, which were used as an effective model of their code description, construction and decoding. The fundamentals of iterative decoding on graphs were also presented. This was then extended to trellis and turbo codes by Wiberg [67] by incorporating so-called hidden state variables in Tanner graphs.

In binary error control coding, the symbols are defined generally in GF(2), hence binary factor graphs are of interest, where all variables involved in the factor graph are binary variables. The corresponding algorithms are straightforward instantiations of the above general rules. Furthermore, the function nodes are constituted by simple parity-checks, which are termed as *parity-check* nodes. For a binary random variable $x \in [0, 1]$, instead of using the probabilities, a commonly used alter-

native representation is constituted by the LLR defined as $\mathcal{L} = \log_2[P(x=0|y)/P(x=1|y)]$.

Let us examine the messages calculated both at the variable nodes as well as at the parity-check nodes in a binary format. Initially, we consider nodes of degree 3 as seen in Fig 2.12. For the variable node, the outgoing message x_3 will be the product of the incoming messages from x_1 and x_2, thus we have:

$$\begin{aligned} P(x_3=0|y_3) &= P(x_1=0|y_1)P(x_2=0|y_2), \\ P(x_3=1|y_3) &= P(x_1=1|y_1)P(x_2=1|y_2), \\ \mathcal{L}_3 &= \mathcal{L}_1 + \mathcal{L}_2. \end{aligned} \tag{2.60}$$

For the parity-check node, where the incoming messages x_1 and x_2 imposes the parity-check equation of $x_1 \oplus x_2 \oplus x_3 = 0$. The outgoing messages x_3 are:

$$\begin{aligned} P(x_3=0|y_3) &= P(x_1=0|y_1)P(x_2=0|y_2) + P(x_1=1|y_1)P(x_2=1|y_2), \\ P(x_3=1|y_3) &= P(x_1=0|y_1)P(x_2=1|y_2) + P(x_1=1|y_1)P(x_2=0|y_2), \\ \mathcal{L}_3 &= \log_2 \frac{P(x_1=0|y_1)P(x_2=0|y_2) + P(x_1=1|y_1)P(x_2=1|y_2)}{P(x_1=0|y_1)P(x_2=1|y_2) + P(x_1=1|y_1)P(x_2=0|y_2)}. \end{aligned}$$

Exploiting the fact that $P(x=0|y)P(x=1|y) = e^{\mathcal{L}}$ and dividing both the numerator and denominator by $P(x_1=1|y_1)P(x_2=1|y_2)$, we arrive at:

$$\begin{aligned} \mathcal{L}_3 &= \log_2 \frac{e^{\mathcal{L}_1}e^{\mathcal{L}_2}+1}{e^{\mathcal{L}_1}+e^{\mathcal{L}_2}} \\ &= \log_2 \frac{\cosh(\mathcal{L}_1/2+\mathcal{L}_2/2)}{\cosh(\mathcal{L}_1/2-\mathcal{L}_2/2)}, \end{aligned}$$

where the above equation is obtained by factoring out $e^{\mathcal{L}_1/2}e^{\mathcal{L}_2/2}$ from both the numerator and denominator. Using the expansion of $\cosh(x \pm y)$ and dividing both the numerator and the denominator by $\cosh(\mathcal{L}_1/2)\cosh(\mathcal{L}_2/2)$ leads to the variable node's outgoing message in the form of:

$$\begin{aligned} \mathcal{L}_3 &= \log_2 \frac{1+\tanh(\mathcal{L}_1/2)\tanh(\mathcal{L}_2/2)}{1-\tanh(\mathcal{L}_1/2)\tanh(\mathcal{L}_2/2)} \\ &= 2\tanh^{-1}[\tanh(\mathcal{L}_1/2)\tanh(\mathcal{L}_2/2)], \end{aligned} \tag{2.61}$$

where in Eq (2.61), we exploited the relation of

$$\tanh^{-1}(x) = \frac{1}{2}\log_2 \frac{1+x}{1-x}.$$

2.3.2 Practical Quasi-Random Coding Approach

In this subsection, we introduce two practical quasi-random coding approaches and characterise them with the aid of the factor graphs discussed above, which are constituted by the classic LDPC codes and interleaved random codes, respectively. Moreover, we justify the necessity of employing a distinct code description in the context of non-orthogonal random waveform based multiuser communications scenarios.

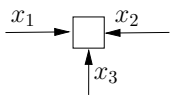
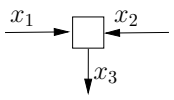
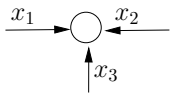
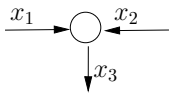
Encoding	Decoding
 $x_1 \oplus x_2 \oplus x_3 = 0$	 $P(x_3 = 0 \mid y_3) = P(x_1 = 0 \mid y_1)P(x_2 = 0 \mid y_2) + P(x_1 = 1 \mid y_1)P(x_2 = 1 \mid y_2)$ $L_3 = L_1 \boxplus L_2 = 2\mathbf{tanh}^{-1}(\mathbf{tanh}(L_1/2)\mathbf{tanh}(L_2/2))$
 $x_1 = x_2 = x_3$	 $P(x_3 = 0 \mid y_3) = P(x_1 = 0 \mid y_1)P(x_2 = 0 \mid y_2)$ $L_3 = L_1 + L_2$

Figure 2.12: Check node and variable nodes in a factor graph, as well as their sum-product algorithm

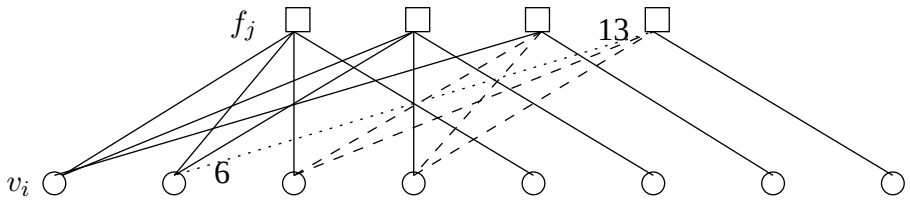


Figure 2.13: Factor graph representation of a Hamming (8,4) code, where eight variable nodes (columns) and four function nodes (rows) are denoted by circles and squares, respectively.

2.3.2.1 LDPC Code Aided Approach

A linear block code of rate $R = M/N$ can be defined in terms of a $(N - M) \times N$ parity-check matrix $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_n]$, where M and N denote the length of the information bit and coded bits, respectively. Each entry $h_{i,j}$ of $\mathbf{H}$ is an element of a finite Galois Field 2, which is either a binary zero or a one and all operations are modulo 2 based. The code is the set of all vectors $\mathbf{x}$ that lie in the (right) null space of $\mathbf{H}$. i.e. we have $\mathbf{Hx} = 0$. Given a parity-check matrix $\mathbf{H}$, we can find a corresponding $(M \times N)$ element generator matrix $\mathbf{G}$, which satisfies $\mathbf{GH}^{\mathrm{T}} = 0$. Then the generator matrix can be used as an encoder according to $\mathbf{x}^{\mathrm{T}} = \mathbf{u}^{\mathrm{T}}\mathbf{G}$.

Fig. 2.13 shows the factor graph of a (8,4) Hamming code having the parity check matrix $\mathbf{H}$:

$$\mathbf{H}_{8,4} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{2.62}$$

In this figure, there is one variable node for each of the $N = 8$ coded bits (columns), and there is one function node for each of the $N - M = 4$ parity-check equations of $\mathbf{H}$ (rows). An edge exists between the ith variable node and the jth parity-check node, if and only if $h_{i,j} = 1$. Furthermore, the function nodes are regular, since their degrees are the same, i.e. we have $d_i^f = 4, i = 1, \ldots, 4$, while the variable nodes are irregular, since their degrees fall into two groups, namely $d_i^v = 3, i = 1, \ldots, 4$ and $d_i^v = 1, i = 5, \ldots, 8$. Observe by comparing Eq (2.62) and Fig 2.13 that the factor graph has a single connection between the four original information bits, directly

copying them to the output according to the systematic encoding part. Furthermore, a binary one in the first column indicates the presence of a parity check link based on the mod 2 connection of the first three input bits, etc. There are a total of 16 edges, which coincide with the total number of ones in the parity-check matrix. The term 'socket' refers to a point on a node to which an edge may be attached. There are possible 16 variable node sockets and 16 function node sockets. We note furthermore that an edge connection can be viewed as an interleaver. To elaborate a little further, viewing the ordering of edges form the variable node sockets side, this interleaver pattern is $\pi = \{1, 5, 9, 2, 6, 13, 3, 10, 14, 7, 11, 15, 4, 8, 12, 16\}$, which states that the variable node socket i is connected to the function node socket $\pi(i)$. For example, as shown by the dotted line, the 6th variable node socket connects to the 13th function node socket, which means that we have $\pi(6) = 13$ according to the above-mentioned interleaver pattern π.

However, since the block length of the above Hamming code is short and the number of ones in it is high, potentially a large number of short cycles may be encountered, one of which is depicted by the dashed line seen in Fig 2.13. It was shown that these short cycles represent an inter-dependence amongst the nodes preventing effective message passing and iterative information updating on the basis of independent extrinsic information, thus resulting in early local convergence before a sufficiently diverse set of codewords was considered. Hence, from a message passing point of view, short block codes like the Hamming code introduced above often result in a poor performance. In order to improve the achievable coding gain, invoking the following principles are recommended:

1. Employ codes having a large block length;
2. Use a low-density distribution of ones in the parity-check matrix;
3. Optimise the location of ones in the parity-check matrix, namely the interleaver pattern.

These requirements are satisfied by the family of the well-known LDPC codes. An LDPC code is a linear block code having a parity-check matrix that is sparse, hence it has a small number of binary one entries. A regular LDPC code construction was proposed by Gallager [17], where the randomly placed binary ones and zeros in the parity-check matrix $\mathbf{H}$ is subjected to the constraint that each row of $\mathbf{H}$ had the same number d^f of binary ones and similarly, each column of $\mathbf{H}$ had

the same number d^v of ones. For example, a 15×20 parity-check matrix may be formulated as:

$$\mathbf{H}_{m,n} = \begin{bmatrix}
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}, \quad (2.63)$$

which has $d^f = 4$ as well as $d^v = 3$ and defines an LDPC code with a length of $N = 20$. It may be referred to as a regular (d^f, d^v, N) LDPC code. In such a regular (d^f, d^v, N) LDPC code ensemble, each information bit is involved in d^v parity checks and each parity-check bit protects d^f information bits. The fraction of ones in the parity-check matrix of a regular LDPC code is d^f / N, which approaches zero as the block length N becomes large and hence leads to the terminology of low-density parity-check codes.

In fact, LDPC codes constitute quasi-random codes in the sense that the interleaver pattern employed generates the resultant 2^M codewords by a random selection from the entire codebook of size 2^N, when the block length N is large. This is an essential assumption in the context of the random coding principle. However, having a large block length N in general results in a complex decoder and an impossible full search. Finally, the message passing algorithm is employed to circumvent the full-search, which ensures convergence towards a near optimum solution. The resultant performance depends on the code construction of the parity check matrix. Having introduced the random LDPC coding approach, let us now consider its interleaved random code based counterpart in the next section.

2.3.2.2 Interleaved Random Code Aided Approach

Apart from the LDPC codes, another class of quasi-random codes is constituted by the family of interleaved random codes. Consider for example transmitting $M = 100$ information bits employing short codes, for instance repetition codes having a repetition factor of 4, which have only two codewords (the all-one and the all-zero). The total number of coded bits becomes $N = 400$. We then

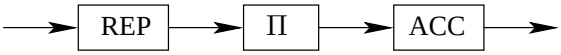


Figure 2.14: The block diagram of repeat accumulate code.

invoke a random interleaver, which scrambles these identical consecutive coded bits, resulting in an interleaved repetition codes having 2^M legitimate codewords, which are randomly selected from a 2^N-entry codebook. However, this interleaving-aided permutation does not change the code's properties, since one can deinterleave, i.e. descramble the randomised codewords in order to recover the original codewords. In order to prevent this decomposition and improve the efficiency of extrinsic information exchange and hence improve the achievable performance, another code acting as an inner code is usually concatenated with the original code acting as an outer code, resulting a serial concatenated code. It is known that the inner code should have a recursive property, i.e. an Infinite Impulse Response (IIR) so as to provide a good performance by ensuring that the interleaved outer codewords can not be decomposed. When a simple unity-rate code acting as an accumulator is incorporated, the resultant serial concatenated code becomes a powerful, yet low-complexity Repeat Accumulate (RA) code [68], which is seen in Fig 2.14.

RA codes can be viewed both as a serial concatenated code described above as well as a special type of LDPC code. Consider the regular non-systematic RA code depicted in Fig 2.15, where we have 3 information bits and a length-4, i.e. rate-1/4 repetition code. The interleaver pattern viewed from the repetition code side's is $\pi = \{3,5,1,6,4,9,11,2,7,12,8,10\}$. The encoding procedure of RA codes can be represented by a generator matrix as follows:

$$\mathbf{G} = \mathbf{G}_{\text{REP}} \times \mathbf{G}_{\pi} \times \mathbf{G}_{\text{ACC}}, \tag{2.64}$$

where the repetition matrix $\mathbf{G}_{\text{REP}}$, permutation matrix $\mathbf{G}_{\pi}$ and accumulation matrix $\mathbf{G}_{\text{ACC}}$ are

given as:

$$\mathbf{G}_{\text{REP}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{2.65}$$

$$\mathbf{G}_{\pi} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \tag{2.66}$$

$$\mathbf{G}_{\text{ACC}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.67}$$

The major drawback of LDPC codes is that the conversion of the large parity-check matrix $\mathbf{H}$ to the generator matrix $\mathbf{G}$ is complex and requires a large memory. The family of RA codes is of particular interest, since it circumvents the problem and it has an encoding complexity which only linearly increases with N. Furthermore, they are capable of approaching the capacity at a moderate complexity. The decoding of RA codes is typically carried out by the sum-product algorithm discussed in Section 2.3.1.2.

The inner code of a serial concatenated code may be constituted not only by a channel code, but also by the data detector, leading to the concept of a serially concatenated iterative receiver. Consider a non-orthogonal UL random waveform based multiuser communications system transmitting two independent data streams generated by the following two different configurations:

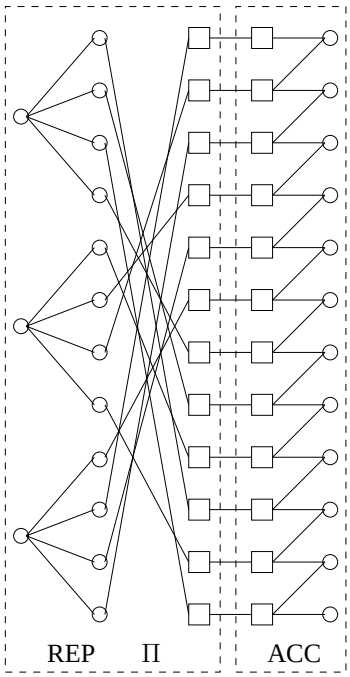


Figure 2.15: Factor graph representation of a regular RA code, where we have 3 information nodes, a rate-1/4 repetition code and an interleaver pattern of $\pi = \{10, 8, 12, 7, 2, 11, 9, 4, 6, 1, 5, 3\}$. Hence we have a total of 12 check nodes.

1. Employing a different outer channel code but the same interleavers;

2. Employing the same inner channel code and but a different interleaver.

The receiver firstly separates these two data streams and then employs channel decoding for each one. The corresponding factor graphs are depicted in Fig 2.16 and Fig 2.17. According to the first configuration, when the same interleavers are employed, the graph can be decomposed into subgraphs, as seen in Fig 2.16. This is not beneficial from a message passing point of view, since this results in short cycles between these subgraphs and introduces a large correlation between the intrinsic and extrinsic information, regardless whether the outer channel code employed is the same or different. However, for the second configuration, although the same channel code is employed, the differently interleaved codewords effectively obey the random coding principle. Hence, the resultant factor graph cannot be decomposed into subgraphs, which in turn enlarges the decoder's full-search-based codebook size. Furthermore, when LDPC codes are employed, the interleavers are redundant, since they have already been embedded in the generator matrix. Unfortunately, an optimum interleaver pattern design, which is capable of effectively lengthening the cycles and hence improving the iterative process remains an open problem, especially for a large number of data streams [69]. More explicitly, the problem is tractable, when the block length is short, while this optimisation becomes unrealistic for a long block length, however simply random interleavers are capable of promising a good performance based on the random coding principle.

In addition to the above-mentioned random coding arrangement, we may take another look from the classic coding point of view. We also considered convolutional codes. The free distance

of convolutional codes is relatively low for a low code memory, while convolutional codes with a large memory value are generally too complex and thus unattractive. The weakness of convolutional codes with a low memory is that they have a large number of low weight codewords, which may result in persistent errors even at high SNRs. By combining them with appropriately designed interleavers, an additional memory is introduced and the resultant code becomes effectively a long random block code. Thus the distance spectrum substantially improves due to interleaving, which is the fundamental benefit of the parallel concatenated turbo codes. To elaborate a little further, turbo coding employs simple component codes and achieves a near-optimum performance by exchanging extrinsic information between the component decoders separated by interleavers during each iteration. In summary, the above mentioned facts underline the significance of the quasi-random coding principle used in non-orthogonal random waveform based multiuser communications, when employing iterative receivers.

2.4 Iterative Turbo Receivers

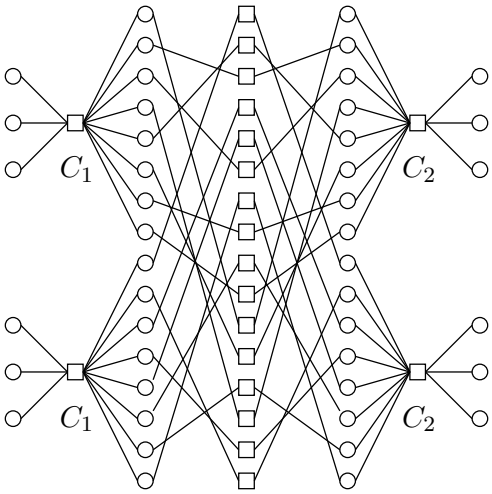
In the previous section, we have provided a rudimentary overview of the achievable capacity of non-orthogonal random waveform based multiuser communications systems and of a range of the practical Forward Error Correction (FEC) coding methods. The factor graph representation introduced in the previous section also suggests that the receiver can be viewed as a concatenated system, where message passing between the receiver components can be used. In this section, we introduce the iterative data detection and channel decoding inherited from the turbo detection concepts [19].

2.4.1 Maximum A posteriori Probability Criterion

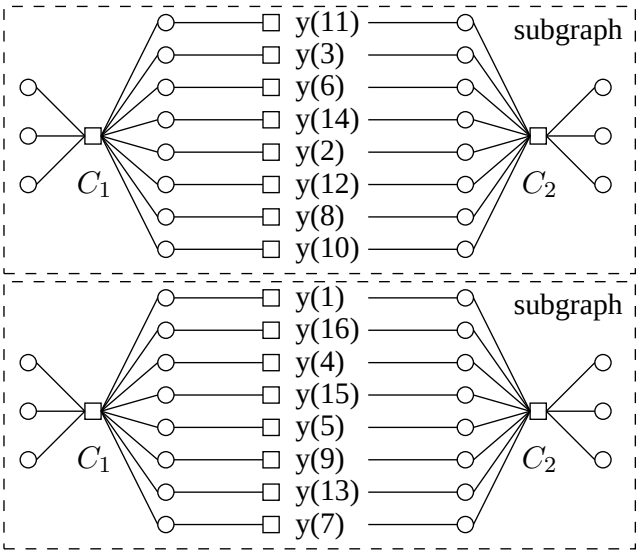
The optimum receiver may employ diverse optimisation criteria, including the so-called MAP criterion. Suppose that we want to estimate the transmitted signal vector $\mathbf{x}$ based on the received signal vector $\mathbf{y}$, so that the estimate of the transmitted vector has the minimal error probability. This is equivalent to maximising the APP, which can be formulated with the aid of Bayes' rule as follows:

$$\begin{aligned}
\mathbf{x}^* &= arg \max_{\mathbf{x}^h \in \mathcal{A}_\mathbf{x}} P(\mathbf{x} = \mathbf{x}^h | \mathbf{y}) && (2.68)\\
&= arg \max_{\mathbf{x}^h \in \mathcal{A}_\mathbf{x}} p(\mathbf{y}|\mathbf{x} = \mathbf{x}^h) P(\mathbf{x} = \mathbf{x}^h) / p(\mathbf{y}) && (2.69)\\
&\propto arg \max_{\mathbf{x}^h \in \mathcal{A}_\mathbf{x}} \underbrace{p(\mathbf{y}|\mathbf{x} = \mathbf{x}^h)}_{extrinsic} \underbrace{P(\mathbf{x} = \mathbf{x}^h)}_{apriori}, && (2.70)
\end{aligned}$$

where $\mathbf{x}^h$ is a hypothesis of $\mathbf{x}$ and $\mathcal{A}_\mathbf{x}$ is the set consisting of all possible observations of $\mathbf{x}$. Again, Bayes' rule was used in this derivation and in the last equation, the denominator $p(\mathbf{y})$ is dropped, since it is independent of the decision. The term $p(\mathbf{y}|\mathbf{x} = \mathbf{x}^h)$ represents the conditional PDF of the received signal vector $\mathbf{y}$ given $\mathbf{x} = \mathbf{x}^h$, denoting the so-called *extrinsic* information and $P(\mathbf{x} = \mathbf{x}^h)$ is the *a priori* information concerning the transmitted signal vector $\mathbf{x}$.

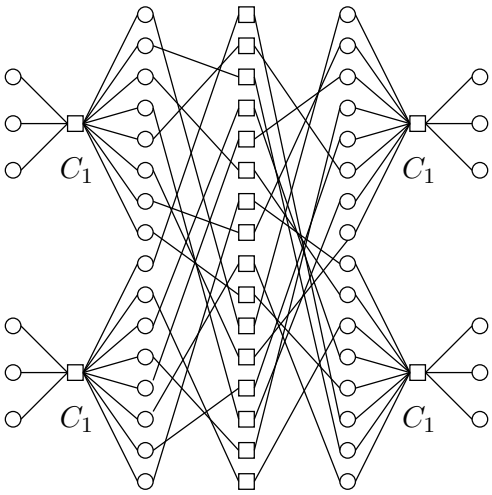


(a) Configuration 1

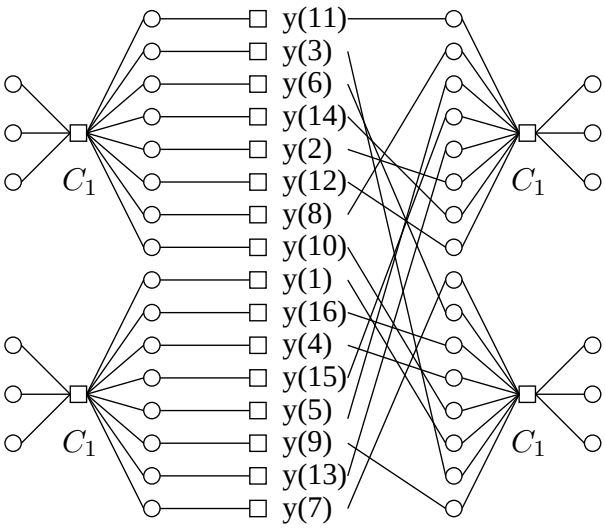


(b) Configuration 1

Figure 2.16: (a) Factor graph representation of iterative receiver employing configuration 1, where we employ different channel code C_1 and C_2 and the same interleaver. The interleaver pattern is $\pi = \{11, 3, 6, 14, 2, 12, 8, 10, 1, 16, 4, 15, 5, 9, 13, 7\}$. (b) In configuration 1, the original graph is decomposed into subgraphs.



(a) Configuration 2



(b) Configuration 2

Figure 2.17: (a) Factor graph representation of iterative receiver employing configuration 2, where we employ the same channel code and different interleavers. The interleaver pattern of the left data stream is $\pi = \{11, 3, 6, 14, 2, 12, 8, 10, 1, 16, 4, 15, 5, 9, 13, 7\}$, the interleaver pattern of the right data stream is $\pi = \{11, 8, 5, 15, 13, 2, 14, 12, 7, 6, 16, 4, 10, 1, 3, 9\}$. (b) In configuration 2, where the original graph can not be decomposed into subgraphs.

An element-wise MAP algorithm aims to estimate the ith element x^* of the true transmitted symbol x_i of the transmitted signal vector $\mathbf{x}$ based on the received signal vector $\mathbf{y}$, so that we have the minimal symbol error probability:

$$x^* = arg \max_{x_i^h \in \mathcal{A}_{x_i}} P(x_i = x_i^h | \mathbf{y}) \quad (2.71)$$

$$\propto arg \max_{x_i^h \in \mathcal{A}_{x_i}} \sum_{\mathbf{x}^h \in \mathcal{A}_{\mathbf{x}:x_i^h}} p(\mathbf{y}|\mathbf{x} = \mathbf{x}^h) P(\mathbf{x} = \mathbf{x}^h), \quad (2.72)$$

where x_i^h is a hypothesis value of x_i and $\mathbf{x}^h$ is a hypothesis value of the vector $\mathbf{x}$. Furthermore, $\mathcal{A}_{x_i}$ represents the set consisting of all possible legitimate observations of x_i and $\mathcal{A}_{\mathbf{x}:x_i^h}$ is the set consisting of all possible legitimate observations of $\mathbf{x}$, where its ith entry is x_i^h.

When the input *a priori* probability $P(\mathbf{x} = \mathbf{x}^h)$ is omitted or assumed to have equal probability, then the MAP criterion becomes synonymous to the Maximum Likelihood (ML) criterion. The optimum MAP / ML detector exhibits high computation complexity, hence numerous reduced-complexity decision methods were proposed, many of which employ the so-called *dynamic programming* [70]. An efficient *dynamic programming* technique is constituted by the Viterbi Algorithm (VA) [71] in the case of ML detection / decoding and by the BCJR algorithm [72] in the case of MAP detection / decoding.

2.4.2 Iterative Data Detection and Channel Decoding

The system model of a channel coded K-user MIMO system employing N_t transmit and N_r receiver antennas is seen in Fig 2.18, which is described by:

$$\mathbf{y}(i) = \mathbf{H}(i)\mathbf{x}(i) + \mathbf{n}(i) \quad (2.73)$$

where we have $\mathbf{y}(i) \in \mathbb{C}^{N_r \times 1}$, $\mathbf{H}(i) \in \mathbb{C}^{N_r \times KN_t}$, $\mathbf{x}(i) \in \mathbb{C}^{KN_t \times 1}$ and $\mathbf{n} \in \mathbb{C}^{N_r \times 1}$ is a circularly symmetric complex-valued Gaussian distribution, i.e. we have $n_i \sim \mathcal{CN}(0, N_0)$.

An receiver consists of a data DETector (DET) and a bank of K individual single-user channel DECoders (DEC), where the DET acts as the inner decoder and DEC acts as the outer decoder, as seen in Fig. 2.18. These two components are separated by random interleavers, which play two primary roles in the system:

1. We may consider the interleavers in conjunction with the outer channel codes as random codes, since they effectively enlarge the memory of the outer channel codes, hence increasing their minimum distance.

2. The interleaver employed makes the consecutive coded bits uncorrelated, since any existing statistical dependencies between the intrinsic and extrinsic information degrades the achieved iterative decoding performance.

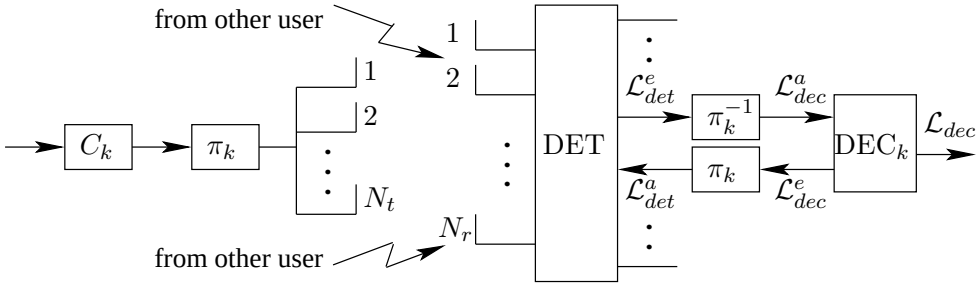


Figure 2.18: The block diagram of K-user MIMO system employing iterative data detection and channel decoding.

Let us now explore these two components in more detail.

At the DET, we maximise the APP $P[x_{k,u}(i)|\mathbf{y}(i)]$ of interest, where $x_{k,u}(i)$ denotes the ith coded bit of the kth user's uth transmit antenna. For simplicity, we omit the coded bit time index i, since the consecutive coded bits are uncorrelated. Since an binary code is employed, where we have $x_{k,u} \in \{\pm 1\}$, the corresponding APP can be expressed in terms of LLRs as:

$$\mathcal{L}_{det}(x_{k,u}) = \log_2 \frac{P(x_{k,u}=+1|\mathbf{y})}{P(x_{k,u}=-1|\mathbf{y})} \tag{2.74}$$

$$\propto \underbrace{\log_2 \frac{p(\mathbf{y}|x_{k,u}=+1)}{p(\mathbf{y}|x_{k,u}=-1)}}_{\mathcal{L}^e_{det}(x_{k,u})} + \underbrace{\log_2 \frac{P(x_{k,u}=+1)}{P(x_{k,u}=-1)}}_{\mathcal{L}^a_{det}(x_{k,u})}, \tag{2.75}$$

where $\mathcal{L}^e_{det}(x_{k,u})$ denotes the *extrinsic* LLR and $\mathcal{L}^a_{det}(x_{k,u})$ denotes the *a priori* LLR calculated by the DEC, as seen in Fig 2.18. The *extrinsic* information, which is independent of the *a priori* information is then deinterleaved and forwarded as the *a priori* information $\mathcal{L}^a_{dec}(x_{k,u})$ to the DEC of Fig 2.18.

The DEC invokes the classic soft decoding algorithm, such as the BCJR [72] algorithm and delivers two *a posteriori* LLR values, namely 1) $\mathcal{L}_{dec}(b_k)$ acting as the final soft decision of the transmitted information bit of user k and 2) $\mathcal{L}_{dec}(x_{k,u})$, which is the soft LLR of the codewords. By subtracting the *a priori* LLR $\mathcal{L}^a_{dec}(x_{k,u})$ from $\mathcal{L}_{dec}(x_{k,u})$, we generate the extrinsic information $\mathcal{L}^e_{dec}(x_{k,u})$ of the codeword, which acts as the *a priori* LLR $\mathcal{L}^a_{det}(x_{k,u})$ of Fig 2.18 input to the DET after interleaving. These iterations are terminated, when a predefined termination criterion is satisfied.

The iterative data detection and channel decoding substantially reduces the complexity of the receiver in comparison to using their joint non-iterative detection and decoding. For example, when considering a convolutional coded K-user MIMO system, the complexity of such a system using a joint MAP receiver is prohibitive, yielding a dynamic programming complexity on the order of $\mathcal{O}(2^{N_t N_r K + Kv})$ [70], where v is the constraint length of the convolutional code employed. By contrast, when decomposing them into two separate components, the associated complexity is partitioned according to $\mathcal{O}(2^{N_t N_r K} + 2^v)$. A further complexity reduction can be achieved within both the DET component and the DEC component. We will focus on the first component in this

thesis, because this component dominates the complexity.

There are numerous contributions on iterative receivers in the literature designed for a K-user CDMA system. Moher [73] demonstrated by simulation that the combination of FEC and random interleaving constitutes a powerful scheme even when we have highly correlated signals for the different users. In [74], Alexander *et al*. presented a MAP based DET algorithm, which exhibited an exponentially increasing complexity as a function the number of users. In [75], Wang and Poor replaced the MAP based DET by a simple interference canceller followed by a single-user Minimum Mean Square Error (MMSE) detector in order to perform interference suppression at a reduced-complexity as approximation of the MAP based DET, yielding a complexity on order of $O(K^3)$ as a function the number of users. The system proposed by Alexander *et al.* [76] then employed an interference canceller, which subtracts the interference estimates and has a linearly increasing complexity as a function the number of users. Simulation results provided in [76] showed that the achievable BER performance is close to the single-user performance even for heavily loaded scenarios. In the following section, we will introduce two DET algorithms, namely the optimum Bayesian detector and a low-complexity IC aided detector.

2.4.2.1 Optimum Bayesian Detection

Let us now discuss the *extrinsic* information $\mathcal{L}^e_{det}(x_{k,u})$ delivered by the optimum Bayesian DET, where the conditional PDF $p(\mathbf{y}|x_{k,u}=\alpha)$ can be computed by subtracting the *a priori* probability $P(x_{k,u}=\alpha)$ of the particular symbol from the joint PDF $p(\mathbf{y}, x_{k,u}=\alpha)$, where we have $\alpha \in \{\pm 1\}$ [74]:

$$p(\mathbf{y}|x_{k,u}=\alpha) = \frac{p(\mathbf{y}, x_{k,u}=\alpha)}{P(x_{k,u}=\alpha)} \tag{2.76}$$

$$= \sum_{\mathbf{x}:x_{k,u}=\alpha} \frac{p(\mathbf{y}|\mathbf{x})P(\mathbf{x})}{P(x_{k,u}=\alpha)} \tag{2.77}$$

$$= \sum_{\mathbf{x}:x_{k,u}=\alpha} p(\mathbf{y}|\mathbf{x}) \prod_{(i,j)=(1,1),(i,j)\neq(k,u)}^{K,N_t} P(x_{i,j}). \tag{2.78}$$

The *a priori* probability $P(x_{i,j})$ can be obtained from the LLR $L^e_{dec}(x_{i,j})$ fedback to the DET by the DEC, which acts as the *a priori* LLR $L^a_{det}(x_{i,j})$ for the DET after interleaving. Since we have $L^a_{det}(x_{i,j}) = \log_2[P(x_{i,j}=+1)/P(x_{i,j}=-1)]$, we can obtain [75]:

$$P(x_{i,j}=\alpha) = \frac{\exp[\alpha\mathcal{L}^a_{det}(x_{i,j})]}{1+\exp[\alpha\mathcal{L}^a_{det}(x_{i,j})]} \tag{2.79}$$

$$= \frac{\exp[\alpha\mathcal{L}^a_{det}(x_{i,j})/2]}{\exp[-\alpha\mathcal{L}^a_{det}(x_{i,j})/2]+\exp[\alpha\mathcal{L}^a_{det}(x_{i,j})/2]} \tag{2.80}$$

$$= \cosh[\mathcal{L}^a_{det}(x_{i,j})/2]\frac{1+\alpha\tanh[\mathcal{L}^a_{det}(x_{i,j})/2]}{2\cosh[\mathcal{L}^a_{det}(x_{i,j})/2]} \tag{2.81}$$

$$= \frac{1}{2}+\frac{\alpha}{2}\tanh[\mathcal{L}^a_{det}(x_{i,j})/2]. \tag{2.82}$$

The conditional PDF $p(\mathbf{y}|\mathbf{x})$ is calculated based on the hypothesis that $\mathbf{x}$ having the (k,u)th entry $x_{k,u}$ was transmitted, which can be expressed as:

$$p(\mathbf{y}|\mathbf{x}) \propto \exp\left(-\frac{||\mathbf{y}-\mathbf{H}\mathbf{x}||^2}{2\sigma_n^2}\right). \tag{2.83}$$

Thus the extrinsic information $\mathcal{L}_{det}^{e}(x_{k,u})$, which is calculated with the aid of the conditional PDF $p(\mathbf{y}|x_{k,u}=\alpha)$, is given by:

$$\mathcal{L}_{det}^{e}(x_{k,u}) = \log_2 \frac{\sum_{\mathbf{x}:x_{k,u}=+1} \exp\left(-||\mathbf{y}-\mathbf{H}\mathbf{x}||^2/2\sigma_n^2\right) \prod_{(i,j)=(1,1),(i,j)\neq(k,u)}^{K,N_t} P(x_{i,j})}{\sum_{\mathbf{x}:x_{k,u}=-1} \exp\left(-||\mathbf{y}-\mathbf{H}\mathbf{x}||^2/2\sigma_n^2\right) \prod_{(i,j)=(1,1),(i,j)\neq(k,u)}^{K,N_t} P(x_{i,j})}, \tag{2.84}$$

where $P(x_{i,j})$ is given by Eq (2.82). The extrinsic information $\mathcal{L}_{det}^{e}(x_{k,u})$ is then deinterleaved and fed to the DEC of Fig 2.18.

2.4.2.2 Low-Complexity Interference Cancellation

The optimum Bayesian DET algorithm of Section 2.4.2.1 exhibits a prohibitive complexity, thus as a design alternative we also introduce a reduced complexity IC type detector. Let us rewrite Eq. (2.73) as:

$$\mathbf{y} = \mathbf{h}_{k,u}x_{k,u} + \xi, \tag{2.85}$$

where $\mathbf{h}_{k,u}$ is the (k,u)th column of $\mathbf{H}$ and ξ denotes the interference plus noise formulated as:

$$\xi = \sum_{(i,j)=(1,1),(i,j)\neq(k,u)}^{K,N_t} \mathbf{h}_{i,j}x_{i,j} + \mathbf{n}. \tag{2.86}$$

By approximating ξ as a joint Gaussian random vector, which can be justified by the central limit theorem provided that the number of users supported and the number of transmit antennas employed is sufficiently high, we have the *extrinsic* information as:

$$\begin{aligned}
\mathcal{L}_{det}^{e}(x_{k,u}) &= \log_2 \frac{p(\mathbf{y}|x_{k,u}=+1)}{p(\mathbf{y}|x_{k,u}=-1)} \\
&= \log_2 \frac{\exp\left(||\mathbf{y}-\hat{\xi}-\mathbf{h}_{k,u}||^2/2\mathbf{R}_{\xi}\right)}{\exp\left(||\mathbf{y}-\hat{\xi}+\mathbf{h}_{k,u}||^2/2\mathbf{R}_{\xi}\right)} \\
&= 2\mathbf{h}_{k,u}^{\mathrm{T}}\mathbf{R}_{\xi}^{-1}(\mathbf{y}-\hat{\xi}),
\end{aligned} \tag{2.87}$$

where $\hat{\xi}$ denotes an estimate of the parameter ξ, while $\mathbf{R}_{\xi} = \mathrm{Cov}(\xi,\xi)$ is the covariance of matrix ξ formulated as:

$$\begin{aligned}
\hat{\xi} &= \hat{\mathbf{y}} - \mathbf{h}_{k,u}\hat{x}_{k,u} \\
&= \sum_{(i,j)=(1,1)}^{K,N_t} \mathbf{h}_{i,j}\hat{x}_{i,j} - \mathbf{h}_{k,u}\hat{x}_{k,u}, \qquad (2.88) \\
\mathbf{R}_{\xi} &= \mathbf{R}_{\mathbf{y}} - v_{k,u}\mathbf{h}_{k,u}\mathbf{h}_{k,u}^{\mathrm{T}} \\
&= \mathbf{H}\mathbf{V}\mathbf{H}^{\mathrm{T}} + \sigma_n^2\mathbf{I} - v_{k,u}\mathbf{h}_{k,u}\mathbf{h}_{k,u}^{\mathrm{T}}, \qquad (2.89)
\end{aligned}$$

where the soft estimate $\hat{x}_{i,j}$ and the instantaneous variance $v_{i,j}$, which is the (i,j) diagonal entry of $\mathbf{V} = \mathrm{Cov}(\mathbf{x},\mathbf{x}) = \mathrm{diag}[v_{1,1},\ldots,v_{K,N_t}]$, are given by:

$$\begin{aligned}
\hat{x}_{i,j} &= \sum_{\alpha\in\pm 1} \alpha P(\hat{x}_{i,j}=\alpha) \\
&= \sum_{\alpha\in\pm 1} \frac{\alpha}{2}\left(1+\alpha\tanh\left[\frac{\mathcal{L}^a_{det}(x_{i,j})}{2}\right]\right), \\
&= \tanh\left[\frac{\mathcal{L}^a_{det}(x_{i,j})}{2}\right] \qquad (2.90)\\
v_{i,j} &= \mathbb{E}\left\{\hat{x}^2_{i,j}\right\} - \mathbb{E}\left\{\hat{x}_{i,j}\right\}^2 \\
&= 1-\hat{x}^2_{i,j}. \qquad (2.91)
\end{aligned}$$

It can be seen in above derivation that only the knowledge of $\hat{x}_{i,j}$ is needed for the *extrinsic* information $\mathcal{L}^e_{det}(x_{k,u})$. Instead of the exponential complexity of the Bayesian DET as a function of the number of users K and the number of transmit antennas N_t, the low-complexity IC aided DET has a linearly increasing complexity as a function of the dimension of $\mathbf{x}$.

2.5 Conclusion

After introducing the necessary background and highlighting the motivation of employing non-orthogonal random waveform based multiuser communications in future wireless systems, our discussions were focused on three key topics, namely, on the generation of the maximal mutual information, on the design of ideal as well as practical coding schemes and on various receiver architectures.

We constructed a generalized linear system model, which is capable of describing diverse systems, including multiuser systems as well as MIMO systems. Our information theory based analysis suggested that non-orthogonal random waveform based multiuser communications is superior in comparison to the traditional orthogonal approach. This capacity improvement can be practically realised, when employing SIC as well as the SPC. We also investigated the multilayer mapping schemes. It was revealed that the conventional multilayer mapping scheme suffers from two undesired properties, namely, from its reduced cardinality and from having a non-equiprobable phaser constellation. In order to eliminate this capacity penalty imposed by these properties as well as for the sake of maintaining a Gaussian-like channel input constellation, it was argued in Section 2.2.2.4 that layer-specific phase rotations and power allocation can be applied.

The artificially imposed phase rotation and power allocation stressed the importance of combining the channel division philosophy with the aid of code division. The employment of code division was also suggested to be crucial from an information theory point of view in the context of random coding principle, which has been widely used in contemporary FEC code design. In Section 2.3, we provided a comprehensive introduction on factor graphs, and discussed two practical quasi-random coding approaches, namely LDPC codes and interleaved random codes. It may be concluded that

interleaved codewords essentially constitute random codewords, which are differently interleaved codes invoked in the context of non-orthogonal random waveform based multiuser communications. We argued in Section 2.3.2.2 that from the receiver's point of view, differently interleaved codewords make the graph fully connected and prevent the graph from being decomposed into local subgraphs.

In Section 2.4.2, the decomposition of a MAP based optimum receiver was established, which led to the so-called iterative data detection and channel decoding principle. This was justified by the interleavers employed between the DEC and DET blocks of Fig 2.18. Two primary benefits of the interleavers were also emphasised. Firstly, as argued in Section 2.4.2, the interleavers made the consecutive coded bits uncorrelated. Secondly, the interleaved codewords may be considered as those of quasi-random codes. After discussing the rationale of decomposing the receiver into its DET and DEC components, we focused our attention on the algorithmic part of DET. Both the optimum Bayesian detector and the low-complexity IC detector were derived in Sections 2.4.2.1 and 2.4.2.2, respectively.

In summary, this chapter included the essential fundamental knowledge for the forthcoming chapters, where we will provide more details about the various instantiations of non-orthogonal random waveform based multiuser communications systems. The IDMA arrangement of Chapter 3 is based on the quasi-random coding principle, which employs interleaved codes. The IR-CDMA scheme of Chapter 3 employs RA codes, which obeys another type of quasi-random coding principle. This non-orthogonal principle can be applied not only in conventional multiuser communications, but also in the context of cooperative multiuser communications scenarios. This is covered in our IR-STC aided MSC system and the PANC scheme detailed in Chapter 4. Furthermore, the principle of non-orthogonal random waveform based multiuser communications may also be applicable to the practical HARQ system of Chapter 5.

Chapter 3

Three Design Aspects of Multicarrier Interleave Division Multiple Access and Its Generalisation - Cellular System Application

An important instantiation of non-orthogonal random waveform based multiuser communications is the so-called IDMA technique. We first present the rationale of employing chip-interleaving by reviewing the transition from traditional DS-CDMA to its chip-interleaved relative, where the literature review of IDMA is offered and the benefits of IDMA are highlighted in Section 3.1.

A novel generalised transceiver architecture, namely that of Multi-Carrier Interleave Division Multiplexing aided Interleave Division Multiple Access (MC-IDM-IDMA) is introduced in Section 3.2. A Parallel Interference Cancellation (PIC) assisted iterative MUD algorithm is invoked in Section 3.2.2 and the convergence behaviour of the system is evaluated with the aid of EXIT charts under the assumption of an equal-power multiuser scenario in Section 3.2.3. Three interesting design aspects of the MC-IDM-IDMA system are analysed in Section 3.2.4, namely 1) the multiplexing versus diversity tradeoffs; 2) the coding versus spreading tradeoffs, where the achievable bandwidth efficiency of both Multi-Carrier Interleave Division Multiplexing (MC-IDM) and MC-IDMA are characterised, leading to general design guidelines for low-rate codes; 3) the complexity versus performance tradeoffs and the EXIT characteristics of two different MUD algorithms are investigated, leading to a reduced-complexity hybrid MUD.

Furthermore in Section 3.3, we propose the concept of Generalised Code Division Multiple Access (GCDMA), employing channel codes from the family of quasi-random codes. The resultant scheme is hence also referred to as IR-CDMA. Apart from the IDMA scheme introduced above, which employs an *explicit* interleaver, we propose Structured Embedded (SE) interleavers in Section 3.3.3.1, which are *implicitly* used in RA code aided IR-CDMA systems and show their

equivalence with the the family of random interleavers, invoked in the same context. More explicitly, the SE interleavers combined with RA codes are designed for the sake of reducing the memory storage requirements in the multiuser transceiver. We then characterise the achievable performance of the hybrid detector proposed in Section 3.3.3.2 for our IR-CDMA system with the aid of EXIT charts in an equal-power scenario. Furthermore in Section 3.3.3.3, we discuss a power allocation technique for this family of near-capacity codes. It is shown that for a large interleaver length, IR-CDMA is capable of approaching the channel capacity.

3.1 State-of-the-Art in IDMA

3.1.1 From DS-CDMA to IDMA

In wireless spread spectrum communications [28], the employment of channel coding is crucial. In the early work of Hui [77], it was shown that the system's effective information throughput was maximised by low-rate channel coding, where a bandwidth expansion was imposed purely due to channel coding. However, their conclusion was based on single-user Matched Filter (MF) aided detection, where a high residual multiuser interference had to be mitigated by the channel codec. A misunderstanding was revealed in the early work of Viterbi [78] that channel coding at a rate lower than unity may reduce the effective processing gain of DS-spreading, when communicating in a fixed bandwidth. However, Viterbi's classic work [79] suggests that bandwidth expansion dedicated to low-rate channel coding has the potential of fully exploiting the achievable processing gain, while simultaneously offering a high coding gain as well as approaching the capacity of the Gaussian multiuser channel with the aid of SIC and single-user decoding. Further insightful theoretical analysis based on MUD was then provided by Verdu in [80], confirming the findings of [77] that indeed bandwidth expansion entirely dedicated to channel coding is favoured for single-user MF aided detection as well as optimum joint detection and decoding. These considerations motivate to two significant improvements, namely *code-spread* and *chip-interleaving*.

The concept of *code-spread* CDMA was suggested by Frenger *et al.* in [81], where the authors proposed the employment of so-called maximum-free-distance low-rate convolutional codes found by computer search, which combine channel coding and DS-spreading. More specifically, these maximum-free-distance low-rate convolutional codes are designed by first finding the 'best' generator polynomials specifying the most meritorious convolutional code for a specific constraint length having the maximum free distance. The resultant codes are then concatenated an appropriate number of times, according to the specific code rate required. The resultant performance was shown to be better than that of conventional CDMA and of the low-rate orthogonal code-spread CDMA scheme of [79].

The system concept of [81] is essentially that of interleaving the chips after DS-spreading, which may be referred to as *chip-interleaving*, rather than using the classic interleaver after channel

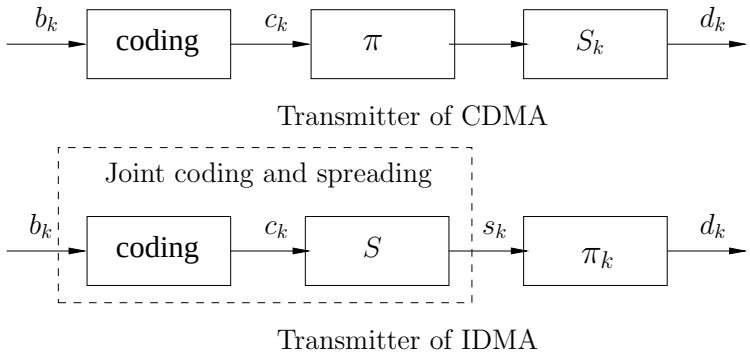


Figure 3.1: The transmitter of IDMA and CDMA.

coding but before DS-spreading. A comprehensive performance evaluation of chip-interleaving in the context of turbo coded CDMA was provided in [82]. In [83] the author further explicitly augmented the concept of chip-interleaved CDMA, which is capable of mitigating the effect of both ISI and Multi-User Interference (MUI). It was demonstrated that the employment of chip-interleaving simplifies iterative data detection and channel decoding. The performance of chip-interleaved CDMA and code-spread CDMA was also compared by Frenger *et al.* in [81]. The efficiency of chip-interleaving in the presence of wideband jamming has also been investigated in [84].

Based on the code-spread chip-interleaved CDMA philosophy of Frenger *et al.* [81] and Proakis *et al.* [83], the IDMA philosophy was further developed by Ping and his team [31] [85] as well as by Höher and Schöneich [32] [86]. IDMA entails reversing the classic position of DS-spreading (S_k) and interleaving (π) employed in traditional CDMA system [28], leading to chip-interleaving instead of bit-interleaving, where the different users are distinguished by their unique user-specific chip-interleavers π_k, as seen in Fig. 3.1. The benefits of IDMA are:

1. *Near-capacity random coding*: The user-specific chip-interleavers effectively enlarge the free-distance of the channel code employed and the resultant codewords of the different users constitute unique noise-like random signatures, since the resultant system satisfies the conditions to be met by Shannonian near-capacity systems.

2. *Time diversity*: Chip-interleaving is capable of increasing the achievable time diversity in time-selective channels, where the burst errors are dispersed and hence the achievable decoding capability is improved.

3. *Joint coding and spreading design*: IDMA is capable of striking the best tradeoff between the channel-coding rate and DS-spreading factor, hence improving the attainable error-resilience.

4. *Near-capacity iterative receiver design*. The chip interleaver employed is capable of substantially improving the reception quality, when an iterative receiver is used, provided that the consecutive chips of the randomly coded sequences are sufficiently uncorrelated, which is analogous to having a sufficiently long turbo-interleaver in serially concatenated codes.

3.1.2 Literature Review of IDMA

The discussions on IDMA found in the open literature were concentrated on the overall system analysis [31, 85–87], on channel code design [48, 88, 89], on power allocation [85, 90] and on their diverse practical implementation aspects [91–97].

System Analysis. The concept of IDMA was first proposed by Frenger *et al.* [81] in the context of code-spread CDMA and was termed as IDMA by Ping in [31]. The system was further analysed in [85], where both the system architecture and a low-complexity receiver algorithm were detailed. A so-called SNR evolution technique was used for tracking the iterative detection convergence of the receiver and for evaluating the attainable performance using a semi-analytical technique. An optimised power allocation scheme was studied with the aid of linear-programming in the context of an iterative receiver. An information theoretical analysis has been provided for example in [87], where IDMA was demonstrated to be capable of asymptotically approaching the channel capacity with the advent of optimum power control. Similarly, in Höher's work [86], the multiuser capacity was analysed, yielding similar conclusions to those of [85]. Later on, the authors of [98] studied the IDMA system in an equal user-power scenario, where an EXIT chart was used for the associated asymptotic performance assessment. Furthermore, the unequal user-power scenario was also considered in [98], where the spectral efficiency was maximised by invoking the so-called differential evolution technique to optimise the associated user-power level. An important conclusion of this paper is that after invoking a sufficiently high number of iterations between the chip-based interference canceller and the despreading components, the resultant interference variance was approximately the same as that, when employing an MMSE type interference canceller proposed for DS-CDMA system [75]. This conclusion was further justified by the authors in [99]. The authors of [100] adopted a matrix-based approach and underlined the importance of incorporating a scrambler after the DS-spreading. As a further advance, in [101], the advantages of IDMA compared to those of CDMA were further justified.

Code Design. Low-rate turbo-Hadamard codes [88] and zig-zag codes [89] were employed in IDMA arrangements designed for time-hopping aided Ultra WideBand (UWB) systems by Ping and Wang in [102]. It was shown that with the aid of powerful low-rate FEC codes, a performance within a fraction of a dB from the Shannon capacity was achieved. Instead of employing a low-rate code, the issue of joint channel coding and DS-spreading was investigated with the aid of EXIT chart in [48].

Power Allocation. The importance of power allocation in IDMA systems was justified by the argument that instead of using random coding and optimum joint detection and decoding, the maximum sum-capacity in a Gaussian Multiple Access Channel (GMAC) can be approached by SIC, followed by single-user decoding, when employing optimum power allocation. The power allocation was optimised in [85], where a linear programming based power allocation scheme was used. Further improvements achieved by power allocation when combined with a practical coding

scheme were reported in [90]. A power allocation scheme suitable for different target BER constraints was proposed in [103], while a practical power allocation scheme having discrete power levels was considered in [104].

Practical Issues. In addition to these theoretical studies, a practical channel estimation scheme was proposed in [91]. The achievable near-far resistance of IDMA was compared to that of DS-CDMA in [92] in the context of asynchronous uplink transmission, where a similar performance was observed for both. The employment of a HARQ aided transmission technique designed for IDMA was presented in [93]. A cross-layer operation aided IDMA system was designed in [94]. Diverse data detection algorithms were characterised in [95–97].

Owing to its meritorious properties, IDMA has been proposed for numerous applications, such as next-generation cellular systems [32, 105] and UWB systems [102] as well as for ad hoc networks [106].

3.2 Three Design Aspects of Multicarrier IDMA

3.2.1 Introduction

Multi-carrier techniques [4, 28] constitute promising enablers for employment in next-generation wireless communications and have been used for example in the IEEE 802.11 Wireless Local Area Network (WLAN) modem family. OFDMA [4] and Generalized MC-DS-CDMA [26] constitute promising multiple access schemes for employment in the downlink of future wireless communications systems, as detailed in [28], [4]. A specific benefit is that upon allocating a subcarrier bandwidth, which is lower than the channel's coherent bandwidth, multi-carrier techniques become capable of transforming a frequency selective fading channel into a frequency-flat fading channel for each subcarrier. This may be achieved even in case of high data rates and highly dispersive channels. Upon concatenating a CP, the effects of ISI between consecutive OFDM symbols may be avoided [22]. Thus, the further development of single-carrier IDMA to multi-carrier IDMA is beneficial, when communicating over highly dispersive wideband channels.

The multi-carrier version of IDMA was proposed in the earlier conference papers [107] and [108], which may be viewed as a MC-CDMA system having both Time-Domain (TD) and Frequency-Domain (FD) chip interleaving or Frequency Hopping (FH). Thus MC-IDMA is more resilient to correlated subcarrier fading routinely experienced in low-dispersion channels. Furthermore, since spreading is used before interleaving, the resultant chips are randomly mapped to different time- and frequency-slots, resulting both TD and FD spreading. Thus MC-IDMA benefits from both the TD and FD diversity. An enlightening tutorial paper about the concept of MC-IDMA can be found in [109], where the principles of MC-IDMA as well as its comparison to MC-CDMA and OFDMA were outlined. Later on, inspired by concept of Frequency Domain Equalisation (FDE), which is particularly suitable for the cellular uplink as an alternative to the OFDMA technique,

IDMA combined with FDE was proposed in [110] and its performance was investigated when employing Zero-Padding (ZP) as well as CP. Furthermore, the employment of the interleave-division concept to both multiplexing as well as to multiple access was characterised in [111], which was termed as Multi-Carrier Multilayer IDMA (MC-ML-IDMA).

In this section, we generalise the philosophy of MC-IDMA, leading to the concept of MC-IDM-IDMA, where each user of the original MC-IDMA system [108] transmits multiple streams differentiated by stream-specific chip-interleavers. As a special case of our general concept, MC-IDM can be considered as a DL broadcast system scenario, while MC-IDMA may be considered as an UL multiple access system. We analyse the system with the aid of EXIT charts [36, 112] and investigate three interesting design aspects of our MC-IDM-IDMA system, namely the associated *multiplexing versus diversity tradeoffs*, the *coding versus spreading tradeoffs* and the *complexity versus performance tradeoffs*.

The novelty and contribution of this section can be summarised as follows:

1. *Multiplexing versus diversity tradeoffs*. We compare the difference between a MC-IDM and a MC-IDMA system, explicitly quantifying the multiplexing gain versus diversity gain tradeoffs in our proposed MC-IDM-IDMA system.

2. *Coding versus spreading tradeoffs*. The bandwidth efficiency of both MC-IDM and MC-IDMA systems can be improved by jointly designing the channel coding rate and the Spreading Factor (SF). Hence we will discuss the coding versus spreading tradeoffs in conjunction with various channel codes and provide design guidelines for choosing a low-rate code for both systems.

   The joint code-rate and SF design was discussed in the context of CDMA systems in [113, 114] and in the references therein. It was found that the conclusions of joint code-rate and SF design depend not only on the choice of the specific FEC codes, but also on the properties of the MUD algorithms employed [80]. By contrast, in both the MC-IDM and MC-IDMA systems, the MUD is carried out on a chip-by-chip basis, which implies that the joint code-rate and SF design is equivalent to the design of an amalgamated low-rate channel coding scheme.

3. *Complexity versus performance tradeoffs*. In order to enhance the achievable system load, we investigate two different MUDs using EXIT charts, namely the optimum ML MUD and the suboptimum PIC scheme. Their different characteristics are discussed and a reduced-complexity hybrid MUD concept is proposed, which strikes an attractive tradeoff between the complexity imposed and the performance attained.

This section is organised as follows. In Subsection 3.2.2, a brief system overview is presented, while in Subsection 3.2.3, the EXIT chart analysis of MC-IDMA is provided. In Subsection 3.2.4, we discuss the above-mentioned three tradeoffs involved in the design of our MC-IDM-IDMA

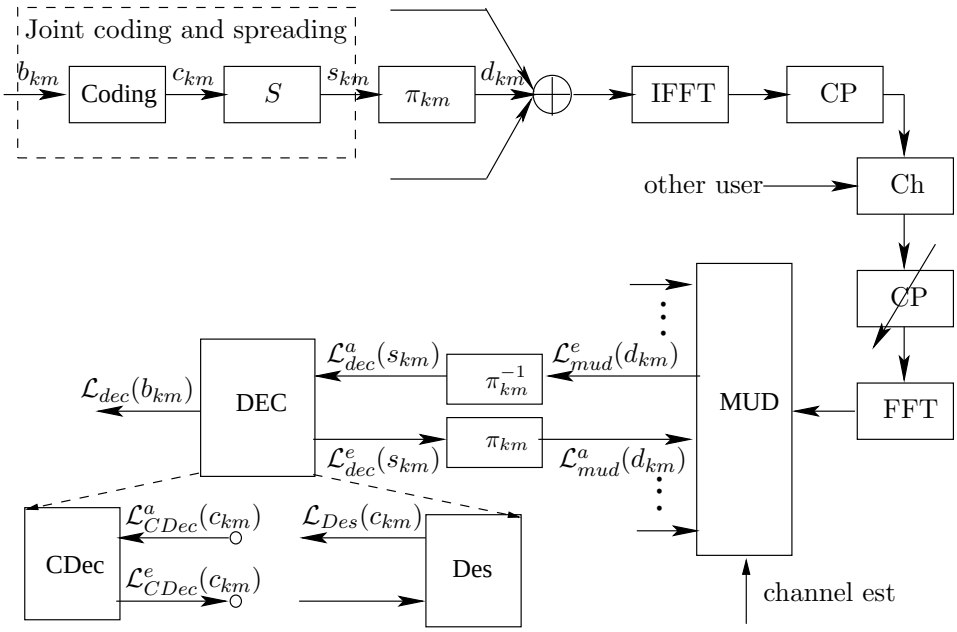


Figure 3.2: Block diagram of the MC-IDM-IDMA transceiver, corresponding to kth user's mth stream.

system, namely the multiplexing versus diversity tradeoffs, the coding versus spreading tradeoffs and the complexity versus performance tradeoffs. Furthermore, practical considerations of IDMA is provided in Subsection 3.2.5. Finally, we conclude our discourse in Section 3.4.

3.2.2 System Overview

We consider a general MC-IDM-IDMA UL system, where each of the K users has M_k channel-coded and DS-spread multiplexed chip-streams distinguished by stream-specific chip-interleavers π_{km}, as seen in Fig. 3.2. The system employs BPSK modulation for communicating over K independent dispersive Rayleigh fading channels. Following OFDM [115], each subcarrier experiences flat fading and when a sufficiently long random chip-interleaver is employed in both the TD as well as in the FD, the corresponding Frequency Domain Channel Transfer Function (FDCTF) of each subcarrier may be assumed to be uncorrelated. We use the simplifying assumption of perfect FDCTF knowledge at the receiver employing a single antenna. The received signal can be expressed as:

$$y[n,i] \quad = \quad \sum_{k=1}^{K} H_k[n,i] \sum_{m=1}^{M_k} d_{km}[n,i] + \nu[n,i], \tag{3.1}$$

where $d_{km}[n,i]$ denotes the kth user's mth stream's ith subcarrier in the nth OFDM symbol, $H_k[n,i]$ is the FDCTF corresponding to user k at the ith subcarrier of the nth OFDM symbol and ν is the AWGN having a zero mean and a variance of $2\sigma_n^2$. It is noted that our generalised system model represents a DL MC-IDM system when we have $K = 1$, where all streams experience the same FDCTF. By contrast an UL MC-IDMA system is modelled, which we have $M_k = 1, \forall k$. Furthermore, we assume that each user has the same number of streams, i.e. $M_k = M, \forall k$.

The receiver of MC-IDM-IDMA consists of a soft in soft out MUD and a bank of KM individual soft in soft out decoders DEC, which are constituted by the combined Channel Decoder (CDec) and De-spreader (Des), as seen in Fig. 3.2. The soft information exchanged between the receiver components is constituted by the extrinsic LLR [72]. The soft MUD employs MF based soft PIC on a chip-by-chip basis [31].

We now consider the ith subcarrier of the nth OFDM symbol. Hereafter we omit the index i and n for notational simplicity. At the lth iteration, the MUD outputs its extrinsic information $\mathcal{L}^e_{mud}(d_{km})$ based on the channel input y combined with the *a priori* information $\mathcal{L}^e_{dec}(d_{km})$ provided by the DEC. When considering the kth user's mth stream, we have

$$y = H_k d_{km} + \xi, \tag{3.2}$$

where $\xi = \sum_{j\neq k} H_j \sum_{m=1}^{M} d_{jm} + H_k \sum_{p\neq m} d_{kp} + \nu$ represents the interference plus noise. Let Im and Re represent the imaginary and real part of a complex number, respectively. In case of BPSK modulation the real part of $H_k^* y$ constitutes sufficient statistics for estimating d_{km}, where $(\cdot)^*$ denotes conjugation, resulting in:

$$Re(H_k^* y) = |H_k|^2 d_{km} + Re(H_k^* \xi). \tag{3.3}$$

We denote the soft estimate of a variable a by $(\hat{a})$. Then, $Re(H_k^* \hat{\xi})$ and its instantaneous power $\mathrm{V}[Re(H_k^* \hat{\xi})]$ are given by:

$$Re(H_k^* \hat{\xi}) = H_k^{Re} \hat{y}^{Re} + H_k^{Im} \hat{y}^{Im} - |H_k|^2 \hat{d}_{km} \tag{3.4}$$

$$\mathrm{V}[Re(H_k^* \hat{\xi})] = (H_k^{Re})^2 \mathrm{V}(\hat{y}^{Re}) + (H_k^{Im})^2 \mathrm{V}(\hat{y}^{Im}) \tag{3.5}$$

$$-|H_k|^4 \mathrm{V}(\hat{d}_{km}) + 2 H_k^{Re} H_k^{Im} \phi, \tag{3.6}$$

where $\phi = \sum_{k=1}^{K} H_k^{Re} H_k^{Im} \sum_{m=1}^{M} \mathrm{V}(\hat{d}_{km})$, the soft estimate $\hat{y}$ and its instantaneous power is thus expressed as:

$$\hat{y}^{Re} = \sum_{k=1}^{K} H_k^{Re} \sum_{m=1}^{M} \hat{d}_{km} \tag{3.7}$$

$$\hat{y}^{Im} = \sum_{k=1}^{K} H_k^{Im} \sum_{m=1}^{M} \hat{d}_{km} \tag{3.8}$$

$$\mathrm{V}(\hat{y}^{Re}) = \sum_{k=1}^{K} (H_k^{Re})^2 \sum_{m=1}^{M} \mathrm{V}(\hat{d}_{km}) + \sigma_n^2 \tag{3.9}$$

$$\mathrm{V}(\hat{y}^{Im}) = \sum_{k=1}^{K} (H_k^{Im})^2 \sum_{m=1}^{M} \mathrm{V}(\hat{d}_{km}) + \sigma_n^2. \tag{3.10}$$

The soft estimate $\hat{d}_{km}$ can be represented as $\hat{d}_{km} = \tanh[\mathcal{L}^e_{dec}(d_{km})/2]$, while its instantaneous power is given by $\mathrm{V}(\hat{d}_{km}) = 1 - \hat{d}_{km}^2$. Assuming ξ is Gaussian distributed, the extrinsic information $\mathcal{L}^e_{MUD}(d_{km})$ is given by:

$$\mathcal{L}^e_{mud}(d_{km}) = 2|H_k|^2 \frac{Re(H_k^* y) - Re(H_k^* \hat{\xi})}{\mathrm{V}[Re(H_k^* \hat{\xi})]}. \tag{3.11}$$

Then the extrinsic information gleaned from the MUD is forwarded as *a priori* information to the DEC, which computes a more reliable extrinsic information $\mathcal{L}^e_{dec}(d_{km})$ for the next iteration. The iterations are terminated, when a predefined termination criterion is satisfied. Finally, the LLR $\mathcal{L}_{dec}(b_{km})$ of Fig. 3.2, which represents the original information bits is subjected to a soft/hard decision.

3.2.3 EXIT Chart Based Semi-Analytical Characterisation

The performance analysis of conventional bit-interleaved coded DS-CDMA using iterative receiver is based on the assumption of 1) having a block length of $N \rightarrow \infty$, 2) employing random spreading codes having a spreading gain of $G \rightarrow \infty$ and 3) that the number of users obeys $K \rightarrow \infty$, while maintaining $K/G = \alpha$. Several analysis techniques were proposed for this task, such as DE [116], Variance Transfer (VT) functions [117] and EXIT chart analysis [112].

The MUD of an IDMA system is simpler than that of classic CDMA systems. Since the MUD processes the detection is chip-basis, no spreading matrix analysis is necessary, which would be necessary for example when an MMSE based MUD is employed in a DS-CDMA system. The analysis of an iterative IDMA receiver was carried out using the so-called *SNR evolution* technique in [31], where the kth user's input/output SNR relation of $SNR^{out}_{mud,k} = v(SNR^{in}_{mud,k})$ was observed at the MUD. The output SNR $SNR^{out}_{mud,k}$ of the MUD was fed into the common channel decoder of all users, which obeyed the function $SNR^{out}_{dec} = f(SNR^{out}_{mud,k})$. Then SNR^{out}_{dec} is fed back to the MUD in the next iteration, resulting $SNR^{out,l}_{mud,k} = v[f(SNR^{out,l-1}_{mud,k})]$, where l denotes the iteration index. The function $f(*)$ was acquired by simulation for a specific channel code, while the function $v(*)$ was different for the different users and depends on the particular MUD algorithm employed.

In this subsection, we consider binary modulation and carry out the iterative receiver's analysis for both a SC-IDMA system in a constant equal-power AWGN channel environment and in a MC-IDMA system context based on EXIT charts [36], where the MC-IDM/IDMA communicating over a dispersive uncorrelated Rayleigh fading channel having a 5-path chip-spaced negative exponentially decaying Channel Impulse Response (CIR). Before carrying out the PIC aided iterative receiver's EXIT-chart analysis, we first introduce the EXIT functions of a general iterative receiver and describe the two component EXIT functions of the IDMA system.

3.2.3.1 EXIT Functions for Iterative Receiver

EXIT charts were introduced by ten Brink [36] and were used to analyse iterative receivers designed for CDMA systems by Wang in [112]. This technique relies on computing the mutual information of two constituent components, namely that of the soft MUD and the soft DEC, which are denoted by I^a_{mud}, I^e_{mud}, I^a_{dec}, I^e_{dec}, where the derivation of mutual information is given in Chapter 2.

Next we evaluate the nonlinear function $I^e = \chi(I^a)$, which maps the input *a priori* mutual

information $I^a \in [0, \cdots, 1]$ to the output extrinsic mutual information $I^e \in [0, \cdots, 1]$, where the amount of output extrinsic mutual information $I^e \in [0, \cdots, 1]$ gleaned from the input *a priori* mutual information determines the convergence behaviour of this soft component. Finally, since the extrinsic information generated by the first soft component acts as the *a priori* information for the second soft component and vice versa, in the EXIT charts we alternately swap the abscissa and ordinate axes, depending on which of the two components acts as the source of *a priori* information, corresponding to the abscissa, as discussed in [36]. To elaborate a little further, EXIT charts provide us with an insight into the mutual information exchange between the constituent soft components, where the chip-interleaver allows us to decouple the constituent soft components, which can be hence analysed separately.

The accuracy of this technique relies on the assumption of using a sufficiently *high transmission block length* and an *infinite-duration random chip-interleaver* for supporting a large number of users K or streams M [1], when the resultant extrinsic information provided by the soft MUD for the kth soft DEC $\mathcal{L}^e_{mud,k}$ converges to the Gaussian distribution described as [116]:

$$\mathcal{L}^e_{mud,k} \sim \mathcal{N}(2\rho^e_{mud,k}, 4\rho^e_{mud,k}), \tag{3.12}$$

where $\rho^e_{mud,k}$ is the asymptotic output SNR of the soft MUD. The available extrinsic information $\mathcal{L}^e_{mud,k}$, which may be of limited benefit in highly interference-limited scenarios acts as the input of the soft DEC of Fig. 3.2. The extrinsic information $\mathcal{L}^e_{dec,k}$ gleaned from the soft DEC can also be approximated by a Gaussian distribution formulated as [116]:

$$\mathcal{L}^e_{dec,k} \sim \mathcal{N}(2\rho^e_{dec,k}, 4\rho^e_{dec,k}), \tag{3.13}$$

where $\rho^e_{dec,k}$ is the asymptotic output SNR of the soft DEC. The corresponding mutual information $I(\mathcal{L}; X)$ in terms of $\mathcal{L}^e_{mud,k}$ and $\mathcal{L}^e_{dec,k}$ are calculated as:

$$I(\mathcal{L}; X) = \frac{1}{2} \sum_{x \in (\pm 1)} \int_{-\infty}^{+\infty} p(l|x) \log_2 \frac{2p(l|x)}{p(l|+1) + p(l|-1)} dl, \tag{3.14}$$

where $I(\mathcal{L}; X) \in [0, 1]$ and $p(l|x)$ is the distribution of the extrinsic LLRs. The EXIT function of the soft MUD is denoted by T_{mud}, while that of the soft DEC by T_{dec}. The flow of mutual information is portrayed in Fig. 3.3, which can be written as:

$$\begin{aligned} I^{e,l}_{mud,k} &= T_{mud}(I^{a,l}_{mud,j}, \forall j \neq k) \\ &= T_{mud}(I^{e,l-1}_{dec,j}, \forall j \neq k) \\ &= T_{mud}(T_{dec}(I^{a,l-1}_{dec,j}), \forall j \neq k) \\ &= T_{mud}(T_{dec}(I^{e,l-1}_{mud,j}), \forall j \neq k), \end{aligned} \tag{3.15}$$

where l denotes the iteration index, while j and k denotes the user index. It is clear from Eq. (3.15) that the mutual information $I^{e,l}_{mud,k}$ of user k depends on its previous value $I^{e,l-1}_{mud,j}$ encountered for *all*

[1] These assumptions are explained further in Section 3.2.3.4 and are justified with the aid of examples.

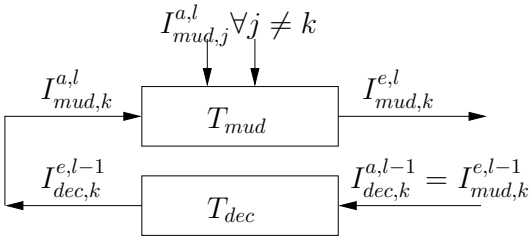


Figure 3.3: The flow of mutual information in the system

other users, as suggested by the non-linear functions T_{dec} and T_{mud}. The mutual information $I^{e,l}_{mud,k}$ is a function of $\rho^{e,l}_{mud,k}$ according to Eq (3.12), which obeys:

$$\rho^{e,l}_{mud,k} = \frac{P_k}{\sigma^2 + P^l_I}, \tag{3.16}$$

where P_k and P^l_I denotes the signal power of the desired user k and the residual interference power at iteration l, respectively. It is worth noting that in the current system context all the individual users' signals affect the shape of the EXIT chart and hence a $(K+1)$-dimensional EXIT chart would be necessary, if the K users had different received signal powers. However, provided that all users' signal is received at the same power, their mutual information may be replaced by the corresponding average, which allows us to use a conventional two-dimensional EXIT chart.

Thus, the employment of two-dimensional EXIT charts is appropriate when we study PIC and in the context of a system communicating over an AWGN channel or over a fading channel subjected to uncorrelated fading in either the FD or the TD, provided that all users' signals are received at an equal power, i.e. when accurate power control is possible.

3.2.3.2 Investigation of the EXIT Function T_{mud}

Let us now investigate the EXIT chart of the iterative IDMA receiver for transmission over an AWGN channel. First of all, let us investigate the MUD's EXIT function T_{mud} seen in Fig. 3.4. The curves converge to the points A, B and C at the operational point of $E_b/N_0 = 10dB$, where E_b is the bit-energy and N_0 is the power spectral density of the noise. Let us denote the number of users by K and the bandwidth expansion factor by Ω, yielding a chip energy of $E_c = E_b/\Omega$. We can then draw the following conclusions from Fig. 3.4.

1. Let us first investigate the left hand-side area of Fig. 3.4, where I^a_{mud} is low. In this area, we have an unreliable estimate of the MUI and hence the residual interference power after PIC remains high, resulting in a low SINR or equivalently, in low extrinsic mutual information I^e_{mud}, which is highly *interference limited*. This can be verified by the curves corresponding to $K = 5$, which represent a lower I^e_{mud} value than the curves recorded for $K = 4$ users, as seen in Fig. 3.4.

2. Let us now consider the right hand-side area of Fig. 3.4, where I^a_{mud} is high. In this area, we benefit from a reliable estimation of the MUI, since the interference was essentially cancelled, resulting in a high SINR or equivalently, in a high extrinsic mutual information I^e_{mud}. Consequently, this scenario may be viewed as *noise limited*.
   Let us now consider the two curves having $E_b/N_0 = 10dB$ and using $\Omega = 4$. Although they correspond to different number of users, namely $K = 4$ and $K = 5$, they both converge to the point $A(I^a_{mud}, I^e_{mud}) = (1, 0.95)$ upon increasing I^a_{mud}. At this point, the interference was entirely cancelled and the resultant extrinsic mutual information I^e_{mud} only depends on the SNR. The same conclusions may be drawn from the curves grouped by $\Omega = 8$.
   Let us finally consider the curves of Fig. 3.4 grouped by $\Omega = 16$ and supporting the same number of users at the different E_b/N_0 values of $5dB$ and $10dB$. At the left hand-side of Fig. 3.4, we can only observe a small difference between the two curves, since this area represents an *interference limited* scenario. However, at the right hand-side of Fig. 3.4, the difference between the $5dB$ and $10dB$ curves becomes quite significant and as expected, the curve corresponding to $E_b/N_0 = 5dB$ remains below the curve associated with $E_b/N_0 = 10dB$.

3. Since the extrinsic mutual information I^e_{mud} defined in Eq. (3.15) is a monotonically increasing function of $\rho^e_{mud,k}$ in Eq. (3.12) of the corresponding Gaussian distributed LLRs [116], increasing the MUI or reducing the operating E_b/N_0 value results in a reduced $\rho^e_{mud,k}$ and hence in a reduced extrinsic mutual information I^e_{mud} over the whole range of the *a priori* mutual information I^a_{mud}. Hence the shape of T_{mud} is maintained. Besides, by increasing the MUI, the difference of extrinsic mutual information I^e_{mud} is more pronounced in the *interference-limited* region, while reducing the operating E_b/N_0 value, the difference of extrinsic mutual information I^e_{mud} is more pronounced in the *noise-limited* region as seen in Fig. 3.4.

4. A final interesting point may be observed by comparing the three extrinsic mutual information points A, B and C in Fig. 3.4, corresponding to $\Omega = 4$, $\Omega = 8$ and $\Omega = 16$ in conjunction with the perfect *a priori* mutual information of $I^a_{mud} = 1$ at $E_b/N_0 = 10dB$. Since the extrinsic mutual information is measured on a chip-by-chip basis, given the same E_b/N_0 value, the larger the bandwidth expansion factor, the lower the output E_c/N_0 value.

3.2.3.3 Investigation of the EXIT Function T_{dec}

We now investigate the EXIT curve of T_{dec} shown in Fig. 3.5, where three different codes are considered in conjunction with a total bandwidth expansion factor of $\Omega = 8$. More specifically, we consider three combinations: 1) Code 1 - the concatenation of a rate-1/4 RA [68] code and a rate-1/2 repetition code, 2) Code 2 - the concatenation of a rate-1/2 convolutional code and a rate-1/4

repetition code and finally, 3) Code 3 - a stand-alone rate-1/8 repetition code [2]. The MUD's EXIT curves are also shown in Fig. 3.5, corresponding to $K = 5,7,8$ at $E_b/N_0 = 10dB$. Based on Fig. 3.5, we arrive at the following conclusions.

1. As discussed before, the left hand-side area of Fig. 3.5 represents an *interference limited* region. Hence the higher the number of users, the lower the extrinsic mutual information I^e_{mud}. Upon increasing the number of users and hence the MUI, the MUD's EXIT curve would first touch Code 1 in Fig. 3.5, then Code 2 and finally Code 3, indicating that the latter scenario of Code 3 is capable of supporting the highest number of users at a given SNR. Hence, we refer to this area of the EXIT chart as the region of high *multiple access capability*.

2. By contrast, the right hand-side area of Fig. 3.5 is again referred to as being *noise-limited*, where the MUI was essentially cancelled. As expected, the near-capacity RA code achieves $I^e_{dec} = 1$ for the lowest *a priori* mutual information of $I^a_{dec} = 0.2$, as seen in Fig. 3.5, which implies a low E_b/N_0 value. Given the *a priori* mutual information of $I^a_{dec} = 0.4$, the convolutional code represented by curve 2 also becomes capable of achieving an infinitesimally low BER. Hence, we refer to this area as a region of high *error correction capability*.

3.2.3.4 Investigation of the Accuracy of EXIT Charts

Let us now consider the accuracy of EXIT charts in analysing the iterative IDMA receiver's performance in an AWGN channel scenario as well as MC-IDMA receiver's performance in dispersive fading channel scenario.

Accuracy of IDMA. Let us now consider the accuracy of EXIT charts in analysing the iterative IDMA receiver's performance in an AWGN channel scenario. We investigate the system's performance by recording the decoding trajectory as well as by Monte Carlo simulation based BER evaluation in various scenarios. The decoding trajectory represents the actual mutual information exchange between two turbo components. It may failed to accurately follow the EXIT chart analysis prediction because the initial assumptions of having Gaussian distributed LLRs, which is a prerequisite of accurate EXIT chart analysis may not be fulfilled.

Fig. 3.6 shows actual decoding trajectory of the RA coded iterative IDMA system in four different scenarios, where Code 1 of subsection 3.2.3.3 was used. The left handside of Fig. 3.6 characterises two different block length scenarios, namely 1024 and 10 000 at $E_b/N_0 = 2dB$ while supporting a low user-load of $K = 4$. Since the RA code's performance substantially depends on the block length, we can see that at a short block length, the system is unable to approach the single-user performance because the decoding trajectory does not follow the EXIT curves and is

[2] We chose these three codes for our analysis, as they represent three typical codes, namely the near-capacity RA code, the convolutional code having a modest coding gain and the repetition code having no coding gain.

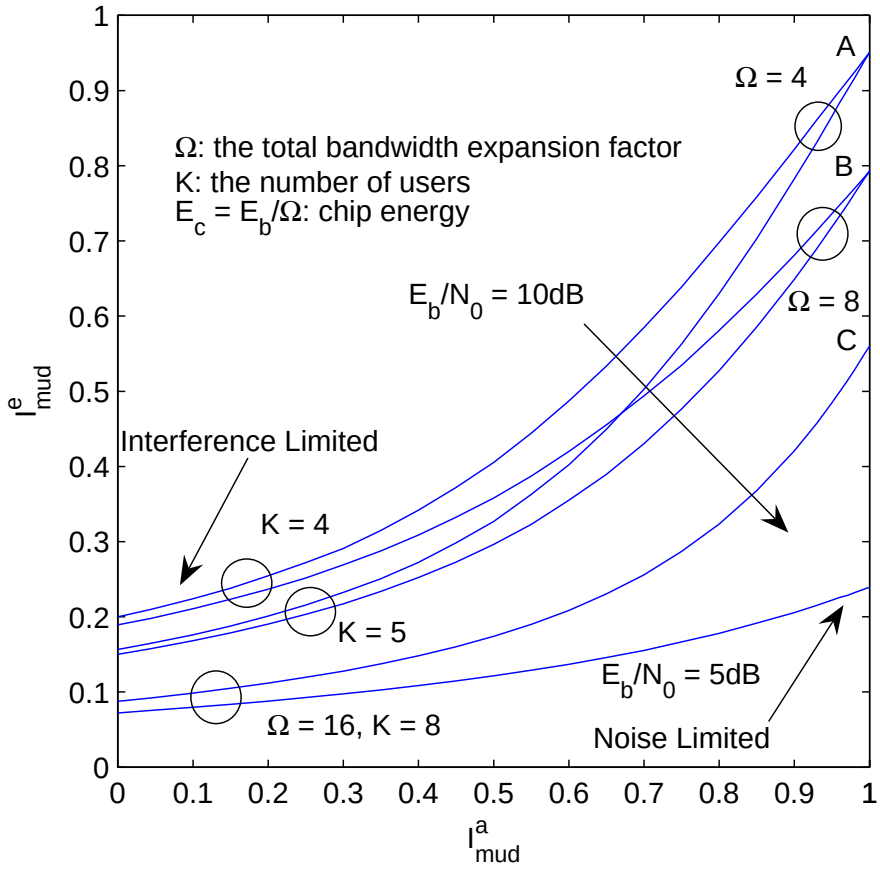


Figure 3.4: The MUD's EXIT charts.

pre-converged, while at a high block length, a near single-user performance is attained. The explicit deficiency of having a short block length is that a reduced extrinsic information fed from the DEC to the MUD.

This observation becomes more grave in high user-load scenarios, as seen at the right handside of Fig. 3.6, where the scenario of $K = 8$ users and $E_b/N_0 = 12dB$ was investigated. The performance of RA codes having a short block length is dominated by the inadequacy of the short interleavers or more explicitly, of the generator matrix. Thus the different interleavers employed by the different users results in a different performance. This is equivalent to a system, which has a different channel code for each user. Thus the MUD may experience different input *a priori* information for each user. This becomes particularly noticeable when the *a priori* information is relatively high, where the degraded performance of the short-block-length RA codes employed by the different users results in catastrophic error propagation effects across the different users. Hence, the EXIT chart fails to adequately characterise the express of the system's behaviour. By contrast, when having a high block length, the system becomes capable of matching the EXIT chart prediction quite accurately.

The corresponding BER results provided by our Monte Carlo simulations are portrayed in Figs. 3.7 and 3.8. As seen in these two figures, the short-block-length RA coded IDMA exhibits a

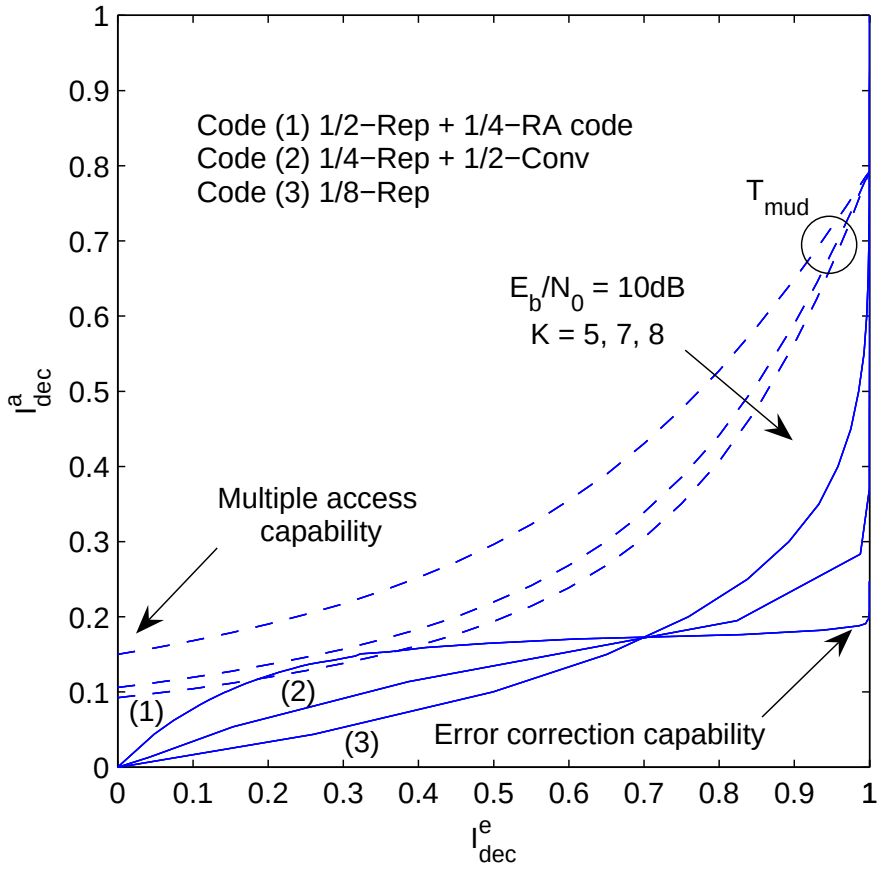


Figure 3.5: EXIT charts of the various channel decoders.

degraded performance as expected. This BER performance degradation becomes more severe, when the user-load becomes high. When a moderate user-load of $K = 4$ was encountered, even a short block length was capable of performing adequately at a slightly higher E_b/N_0 value. By contrast, when full user-load of $K = 8$ was encountered, the short-block-length RA coded system would necessitate a significantly increased E_b/N_0 value and a higher number of iterations as well.

Accuracy of MC-IDMA. Let us now consider the EXIT chart analysis of MC-IDMA having 128 subcarriers, where the CIR encountered in this paper was an independently generated 5-path chip-spaced negative exponentially decaying dispersive channel. Each path experienced uncorrelated Rayleigh fading. The difference in comparison to the previously considered AWGN channel manifests itself in terms of a different EXIT function T_{mud}, which takes into account of the CIR knowledge. The resultant EXIT curve T_{dec} of the outer decoder can be modelled as in the previously considered AWGN scenario, since the outer channel DEC is fed with the near-Gaussian distributed output extrinsic information of the MUD.

Let us now consider the accuracy of our EXIT chart analysis in a MC-IDMA scenario. Observe in Fig. 3.9 that a maximum of $K = 28$ users were supported by Code 3 of subsection 3.2.3.3 at $E_b/N_0 = 12dB$. By contrast, $K = 9$ users were serviced by Code 1 of subsection 3.2.3.3 at $E_b/N_0 = 3dB$, where the maximum number of users was found by identifying the EXIT curves

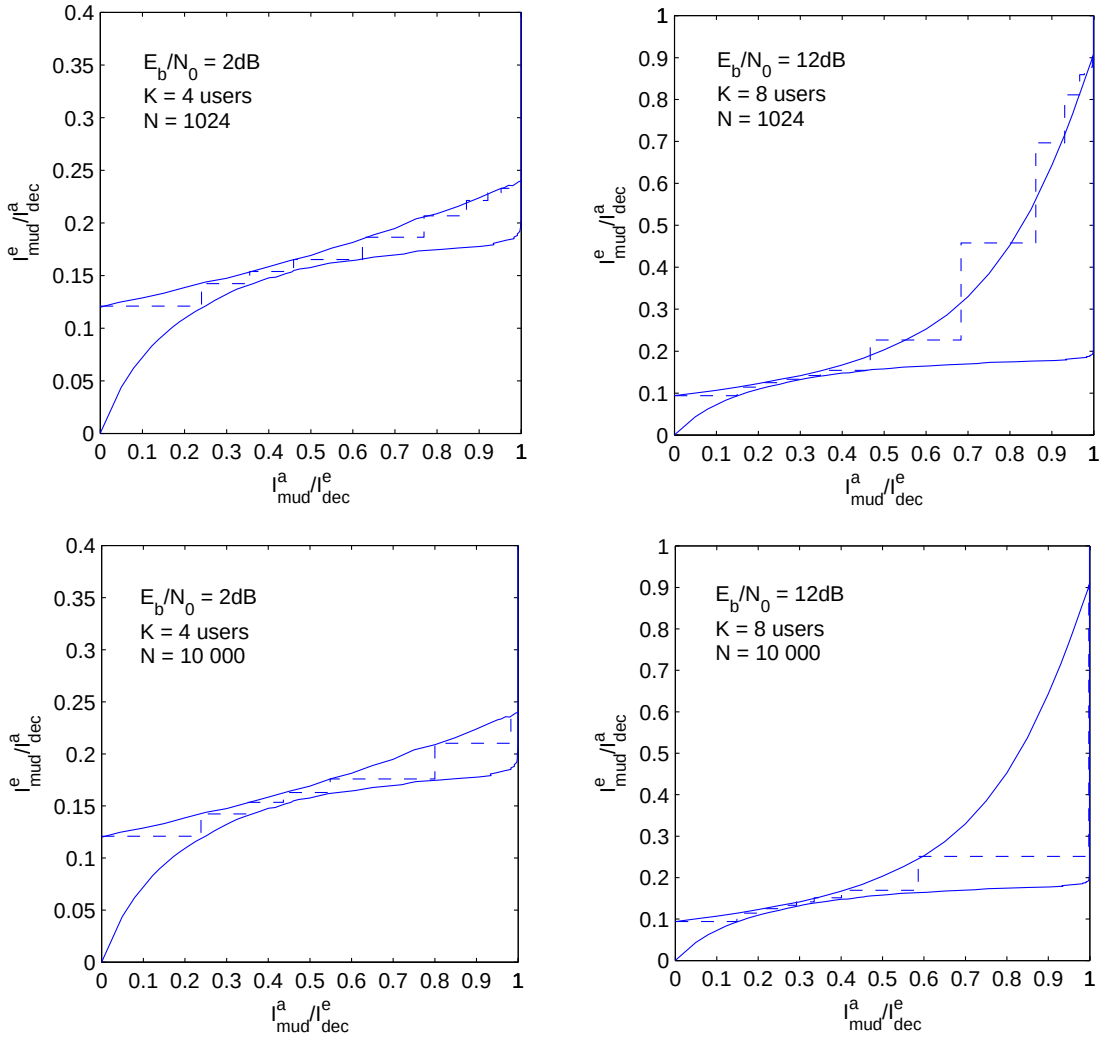


Figure 3.6: Decoding trajectory of the RA coded system having different block lengths and user-loads.

of the MUD and of the DEC, where they "just" touch each other. The corresponding simulation trajectory of Fig. 3.9 confirms the convergence and accuracy of our EXIT chart analysis. Furthermore, the corresponding BER simulation results were portrayed in Fig. 3.10, where as many as $K = 28$ users were supported by Code 3 at $E_b/N_0 = 12dB$. By contrast, Code 1 supported $K = 9$ users at $E_b/N_0 = 3dB$ and $K = 16$ users at $E_b/N_0 = 12dB$, confirming a good agreement with our EXIT chart analysis. The corresponding number of PIC iterations was fixed to $\mathcal{I}_{PIC} = 30$ and the number of iterations within the code was $\mathcal{I}_{RA} = 20$.

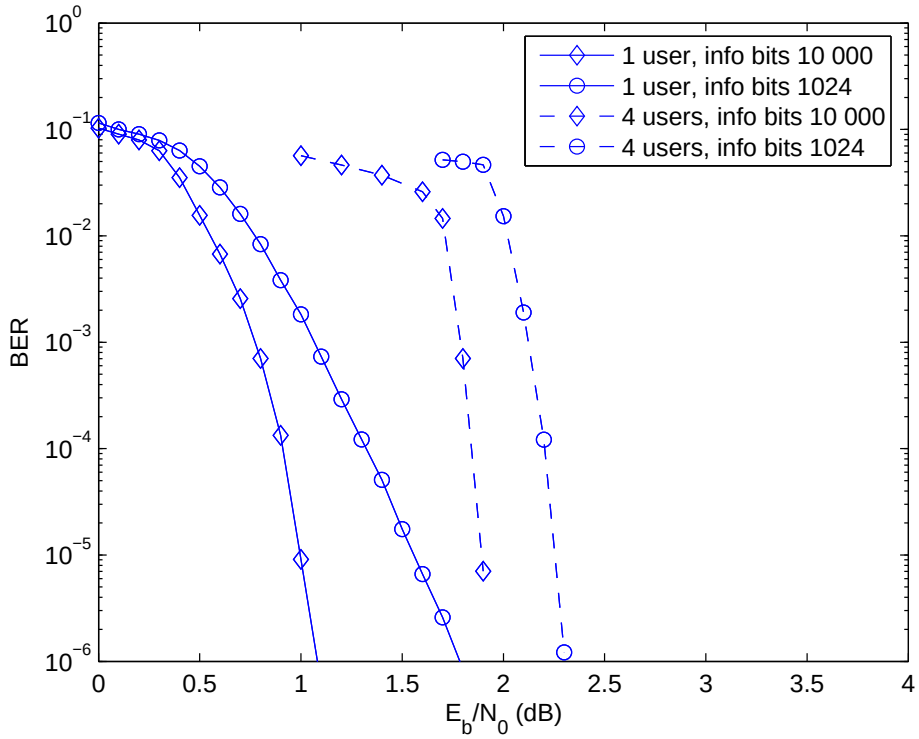


Figure 3.7: Performance of the RA coded SC-IDMA system for transmission over an AWGN channel having a low user-load of $K = 4$.

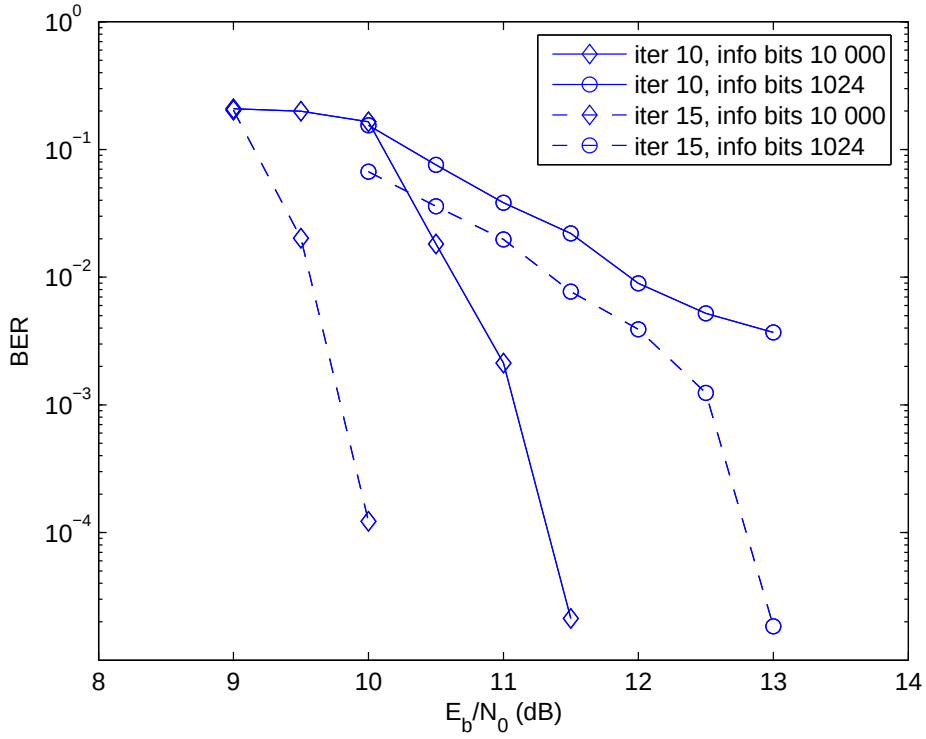


Figure 3.8: Performance of the RA coded SC-IDMA system for an AWGN channel when supporting a high user-load of $K = 8$.

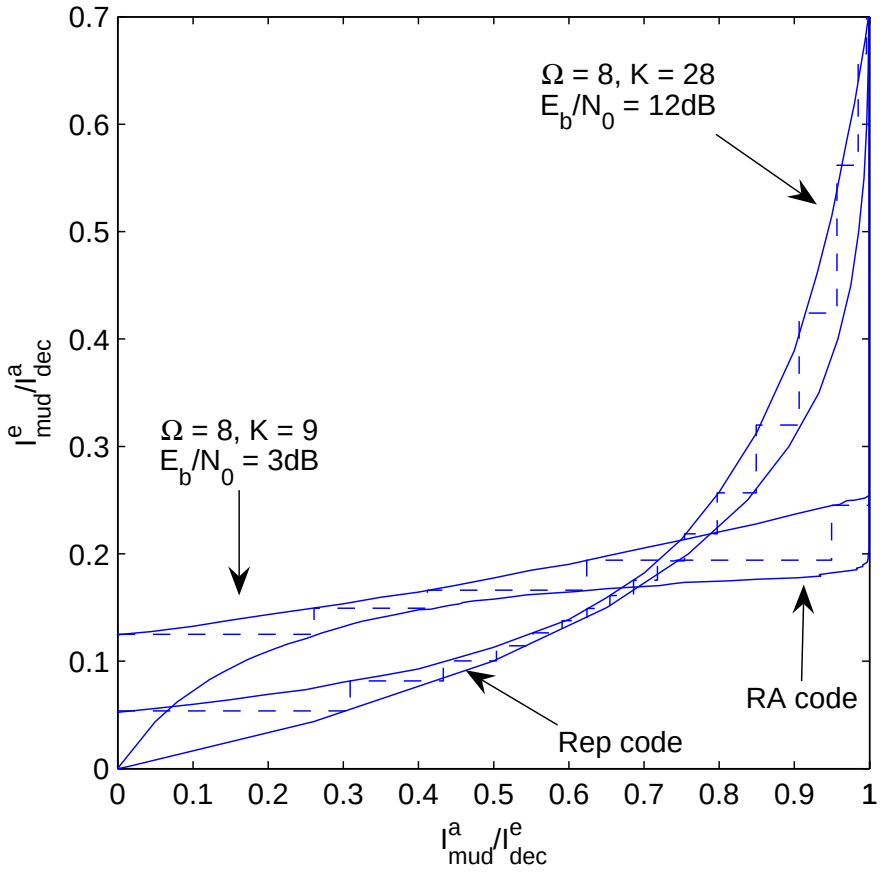


Figure 3.9: Maximum number of users supported by the 128-subcarrier MC-IDMA system in a 5-path chip-spaced negative exponentially decaying uncorrelated Rayleigh fading channel.

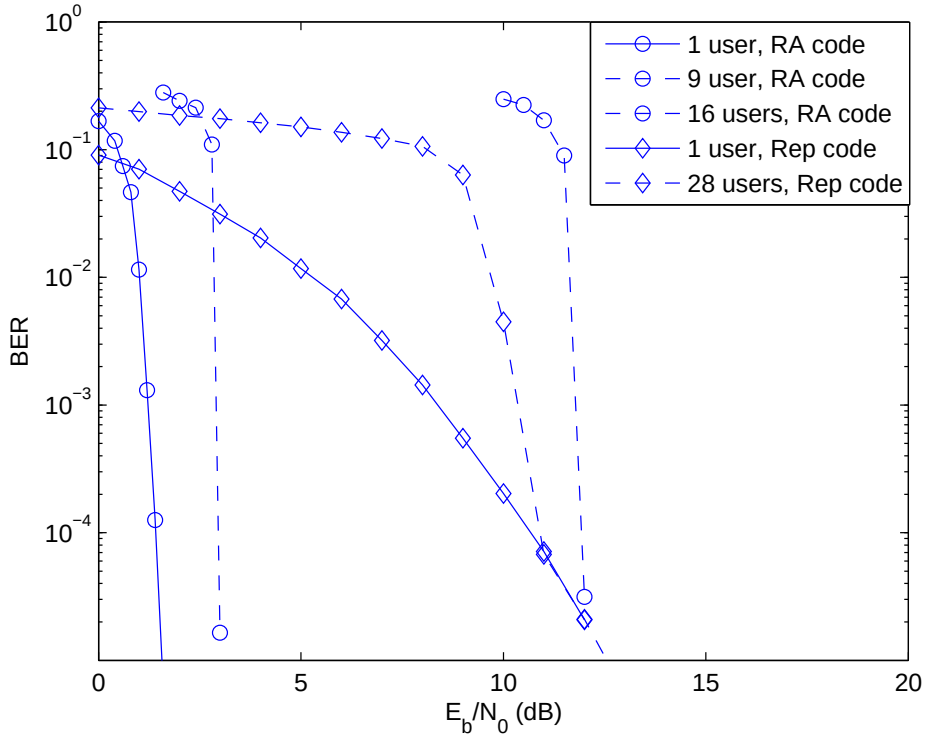


Figure 3.10: Performance of the RA and repetition coded 128-subcarrier MC-IDMA system in a 5-path chip-spaced negative exponentially decaying uncorrelated Rayleigh fading channel.

3.2.4 Design Trade-offs in MC-IDMA Systems

3.2.4.1 Coding versus Spreading Trade-offs

In traditional CDMA systems, the main objective of spreading is that of supporting a multiplicity of users. When non-orthogonal spreading sequences are used, or the orthogonality of the spreading sequences is destroyed by the dispersive channel, the MUI may be mitigated with the aid of various MUD algorithms. Hence, the conclusions of joint code-rate and SF design depend not only on the choice of the specific FEC codes, but also on the properties of the MUD algorithms employed [80]. In the early work of Hui [77], it was shown that the system's effective information throughput was maximised by low-rate channel coding, where a bandwidth expansion was imposed purely due to channel coding. However, their conclusion was based on single-user matched filtering aided detection, where a high residual multiuser interference had to be mitigated by the channel codec. Further insightful theoretical analysis based on multiuser detection was then provided by Verdu in [80], confirming the findings of [77] that indeed bandwidth expansion entirely dedicated to channel coding is favoured for single-user matched filter aided detection and the optimal joint detection and decoding. By contrast, there is a beneficial tradeoff between the channel coding rate and SF, when linear MUD, such as decorrelating or MMSE MUD is used. These concepts were then further developed in a multicell context in [118]. In [113], Yue and Wang analysed the associated coding and spreading tradeoffs in the context of an iteratively detected CDMA system using LDPC codes with the aid of density evolution analysis. Additionally, in [119] Tang and Ryan studied the coding versus spreading tradeoffs in the context of a synchronous CDMA system employing both linear and non-linear detectors in both finite-size system and large-system.

Based on this backdrop, these investigations underline the importance of joint code-rate and spreading-factor optimisation in the context of another system, namely IDMA. Although IDMA does not constitute necessarily a spread-spectrum system, an important advantage of IDMA is that it facilitates the flexible joint optimisation of the channel coding rate and the SF using for example low-rate RA codes or diverse Serially Concatenated Code (SCC) [72]. A specific example is constituted by the concatenation of high-rate FEC codes and low-complexity repetition codes, namely DS-spreading. In IDMA systems, the sole purpose of using repetition coding is to shape the EXIT curve of the SCC. Furthermore, MUD is carried out on a chip-by-chip basis, which implies that the joint code-rate and SF design is not dependent on the choice of MUD algorithms, while it is equivalent to the design of an amalgamated low-rate channel coding scheme.

Definition and Methodology. In a SCC aided IDMA system, given a fixed total bandwidth expansion factor of Ω, we may jointly quantify the bandwidth expansion imposed by FEC coding having a code-rate of R and that by DS-spreading having a SF of G, yielding $\Omega = G/R$. For the sake of characterising the achievable throughput of the IDMA system, we define the normalised

user-load as:

$$\eta = \frac{K_{max}}{\Omega}, \tag{3.17}$$

where K_{max} is the maximum number of users supported, given a fixed bandwidth expansion factor Ω and E_b/N_0 value. Note that for a fixed bandwidth expansion factor Ω and a channel code-rate R, the affordable SF is determined. Hence, K_{max} is a function of Ω, R and E_b/N_0. In case of an IDM system, we may accordingly define the normalised *per-user throughput* as $\eta = M_{max}/\Omega$, where M_{max} is the maximum number of bit-streams supported.

Let us now continue our investigations by fixing the bandwidth expansion factor to Ω and find the maximum number of users supported by different channel coding combinations at various E_b/N_0 values, while maintaining a given target BER of 10^{-5} with the aid of two functions provided by EXIT charts.

1. *Estimating BER*. The EXIT chart can be used to obtain an estimate on the bit error probability after an arbitrary number of iterations. The soft bit values output by the MUD after a given number of iterations can be expressed using the a *posteriori* LLRs, which is the sum of the *a priori* LLRs and the extrinsic LLRs denoted by $\mathcal{L}_{mud,k} = \mathcal{L}^a_{mud,k} + \mathcal{L}^e_{mud,k}$. It is Gaussian distributed since both $\mathcal{L}^a_{mud,k}$ and $\mathcal{L}^e_{mud,k}$ are Gaussian. The coded bit error probability is $Pe \approx Q(\sigma/2)$, where σ is the variance of $\mathcal{L}_{mud,k}$ and Q is the well-known Gaussian Q-function. Since $\mathcal{L}^a_{mud,k}$ and $\mathcal{L}^e_{mud,k}$ are independent, then $\sigma^2 = \sigma_a^2 + \sigma_e^2$, where σ_a and σ_e represents the variance of the *a priori* LLRs $\mathcal{L}^a_{mud,k}$ and the extrinsic LLRs $\mathcal{L}^e_{mud,k}$ respectively. The values of σ_a and σ_e can be obtained from the corresponding mutual information $I^a_{mud,k}$ and $I^e_{mud,k}$, since the mutual information is a monotonically increasing function of the LLRs.

2. *Identifying system load*. We can solve the joint code-rate and SF design by generating the EXIT curves of both the T_{mud} and T_{dec} and then finding the specific number of users, for which an open EXIT-tunnel exists. More explicitly, we plot the EXIT curve T_{mud} for various number of users and the maximum number of users is found, where there is no intersection between the inner and outer codes' EXIT curves. In Fig. 3.11, the bandwidth expansion factor was fixed to $\Omega = 8$. A rate-1/4 RA code concatenated with a rate-1/2 repetition code was used and $E_b/N_0 = 10dB$ was assumed. We found that the maximum number of user supported by this SCC scheme was $K_{max} = 7$, corresponding to a normalised *user-load* of $\eta = K/\Omega = 7/8$, where the inner and outer codes' EXIT curves 'just' touch each other.

Joint Coding and Spreading Design for IDMA. As a first step, we study the joint code-rate and spreading-factor optimisation in a SCC aided IDMA system, where the convolutional code is employed as the outer code and repetition code is employed as the inner code. Let us now proceed by investigating the joint code-rate and SF design by employing the range of different-rate convolutional codes summarised in Table 3.1, which have similar constraint lengths and were

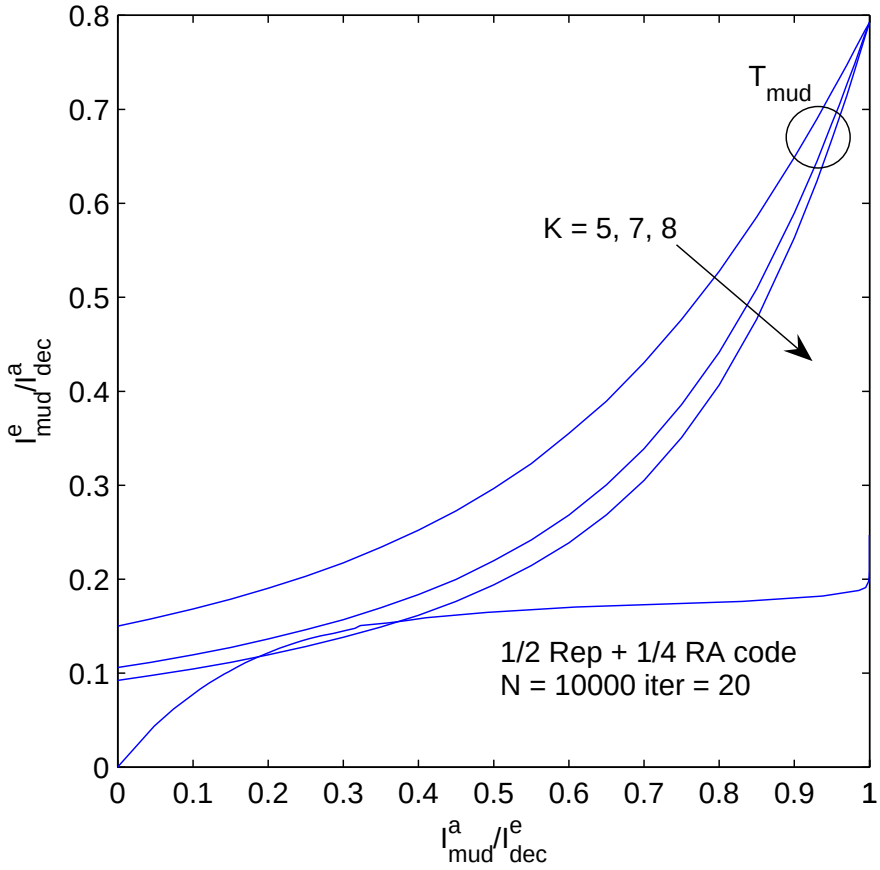


Figure 3.11: EXIT charts for different number of users supported in coded IDMA at a bandwidth expansion factor of $\Omega = 8$ and E_b/N_0 = 10dB, using $R = 1/4$ RA code

selected from [6]. Importantly, please also observe in Table 3.1 that the detection complexity of all the different scenarios are adjusted to be the same in terms of the total number of trellis states, which was determined by the product of the number of trellis states and the number of iterations. We set the target BER to $P_e = 10^{-5}$.

Fig. 3.12 shows the normalised *user-load* supported versus various code-rate and SF combinations for two different E_b/N_0 scenarios. It can be seen in Fig. 3.12 that at $E_b/N_0 = 5$ dB, the normalised user-load was maximised by a combination of the rate-5/8 code and the SF of $G = 5$ of Table 3.1, when supporting $K_{max} = 9$ users. By contrast, the IDMA system favours the employment of pure spreading without FEC coding at E_b/N_0 = 10dB, as can be seen from Fig. 3.12, reaching $\eta = 1.75$ at a rate of $R = 1$.

Fig. 3.13 portrays the E_b/N_0 value required for supporting a normalised *user-load* of both $\eta = 1.00$ and $\eta = 1.25$. It can be observed that having a high code-rate results in a low coding gain and hence a high transmit power is required for supporting the target *user-load*. The optimum code-rate is $R = 5/8$ in this scenario. However, having a code-rate R below the optimum also requires an increased E_b/N_0 for supporting the target *user-load*.

At high E_b/N_0 values, where the noise is insignificant, the soft MUD and the soft de-spreader

Code Rate	Octal Generator	Constraint length	SF (G)	No. of Iterations / Trellis States (Complexity)
$R = 1/8$	[7 7 5 5]	3	-	10 / 4 (40)
$R = 2/8$	[5 7 7 7]	3	2	10 / 4 (40)
$R = 1/3$	[5 7 7]	3	3	10 / 4 (40)
$R = 4/8$	[5 7]	3	4	10 / 4 (40)
$R = 2/3$	[4 2 6] [1 4 7]	3	5	10 / 4 (40)
$R = 3/4$	[6 2 2 6] [1 6 0 7] [0 2 5 5]	5	6	10 / 4 (40)

Table 3.1: Table of convolutional codes used in joint code-rate and spreading-factor optimisation

Scheme	1/64-Rep	1/2-Cov+1/32-Rep	1/64-RA	1/4-RA+1/16-Rep
Type of System	MC-IDMA / MC-IDM			
Type of Detection	PIC			
Ω	64			
$(E_b/N_0)_t$	12	6	3	3
η^*	3.625 / 2	2.75 / 1.375	1.5 / 0.625	2 / 0.75

Table 3.2: Summary of the parameters and main results of coding versus spreading tradeoffs. $(E_b/N_0)_t$ represents the minimum E_b/N_0 value in dB at which the channel code operates and η^* denotes the maximum throughput expressed in bits/chip

become capable of almost completely eliminating the MUI when allocation the bandwidth to spreading, while at low E_b/N_0 values the output extrinsic information of the soft de-spreader is gravely noise contaminated, which can only be improved by the relatively low-rate and hence powerful soft channel decoder, this in turn reduces the processing gain.

Joint Coding and Spreading Design for MC-IDMA. Let us now continue our investigations by fixing the bandwidth expansion factor to $\Omega = 64$ and find the maximum number of users supported by different channel coding combinations at various E_b/N_0 values, while maintaining a given target BER of 10^{-5} with the aid of EXIT charts. Four typical channel coding combinations are investigated as seen in Table 3.2, namely 1) Scheme 1 - rate-1/64 repetition code 2) Scheme 2 - the concatenation of a rate-1/2 convolutional code and rate-1/32 repetition code 3) Scheme 3 - rate-1/64 RA code 4) Scheme 4 - the concatenation of a rate-1/4 RA code and rate-1/16 repetition code.

In Fig. 3.14 and Fig. 3.15, the rate-1/64 pure repetition coded MC-IDM and MC-IDMA Scheme 1 shows the highest *multiple access capability* at high E_b/N_0 values for all the codes considered. In a channel coded system scenario, as expected, the convolutional coded Scheme 2 supports a higher number of users than the RA coded Scheme 3 and Scheme 4 for SNR values in excess of $E_b/N_0 = 5dB$, since it has a higher *multiple access capability*, as seen in Fig. 3.5.

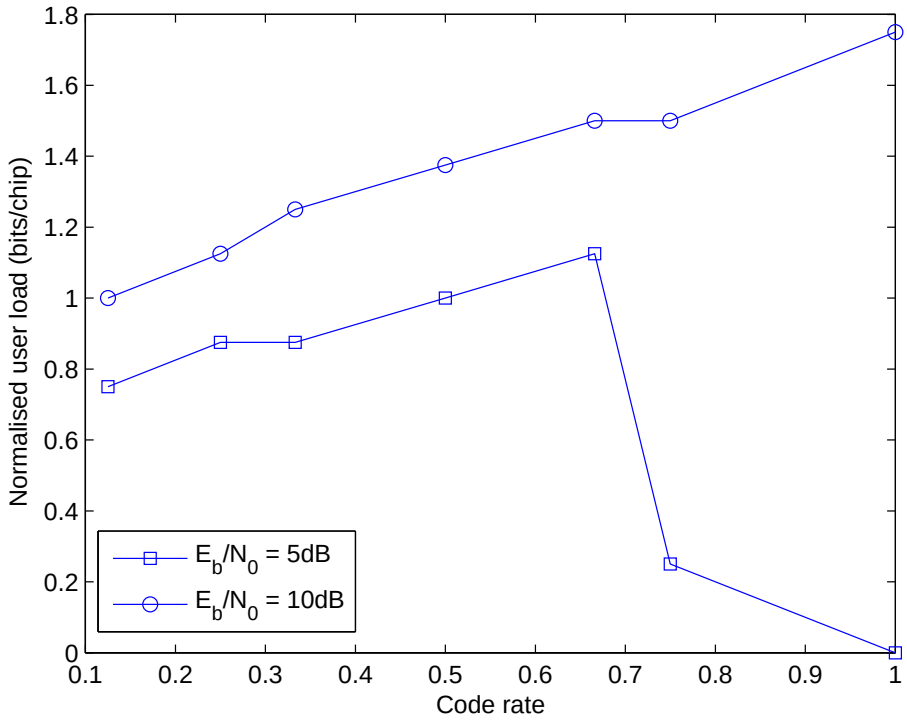


Figure 3.12: Normalised user-load versus code-rate for a fixed bandwidth expansion factor of $\Omega = 8$ and for $P_e = 10^{-5}$ using the convolutional codes of Table 3.1

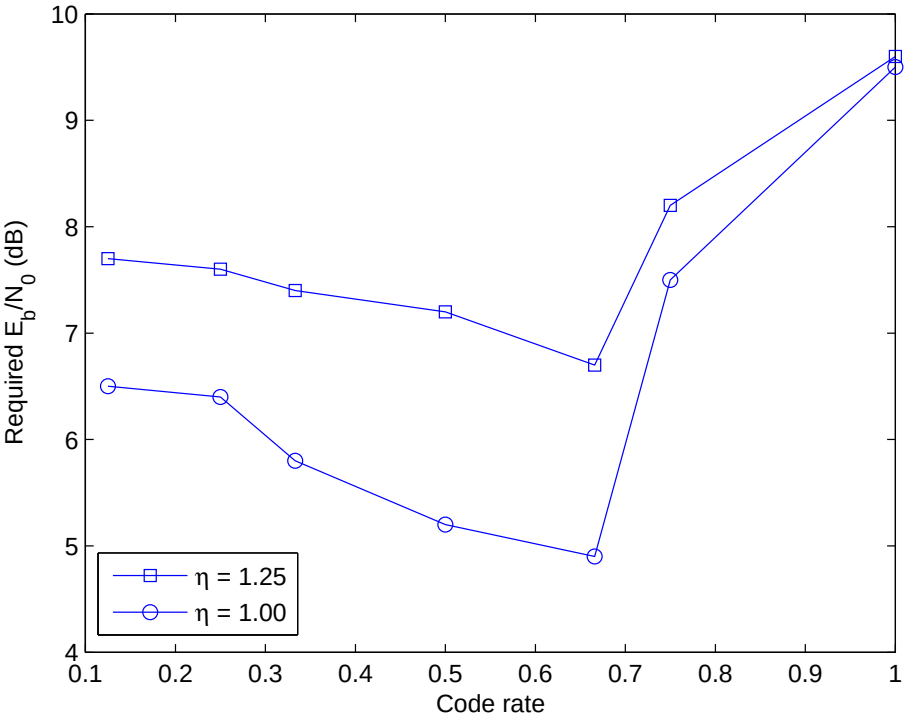


Figure 3.13: The E_b/N_0 values required for supporting a normalised user-load of η versus the code-rate at a fixed bandwidth expansion factor of $\Omega = 8$ and for $P_e = 10^{-5}$ using the convolutional codes of Table 3.1

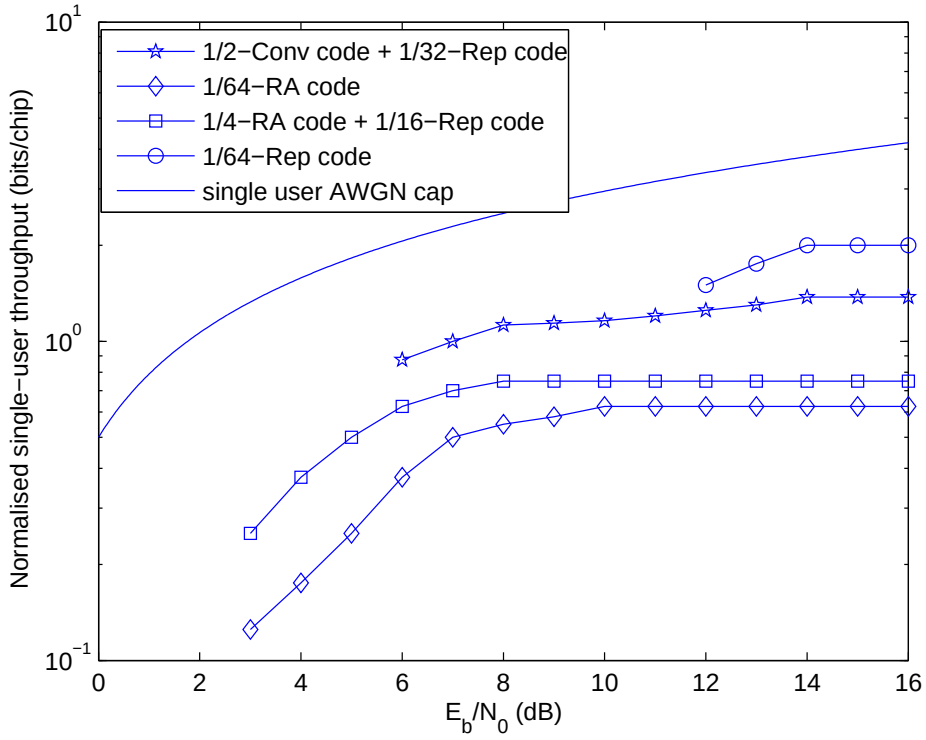


Figure 3.14: Normalised per user throughput versus E_b/N_0 performance of the 128 subcarrier MC-IDM for a fixed bandwidth expansion factor of $\Omega = 64$ and for $P_e = 10^{-5}$ using the various channel coding schemes considered in a 5-path chip-spaced negative exponentially decaying uncorrelated Rayleigh fading channel.

However, the RA coded Schemes 3 and 4 are capable of reliably operating at lower E_b/N_0 values, as a benefit of their higher *error correction capability*.

RA coded Scheme 3 and Scheme 4 were also compared in Fig. 3.14 and Fig. 3.15, where the number of RA-decoding iterations was fixed to $\mathcal{I}_{RA} = 20$. The low-rate RA coded Scheme 3 using an insufficient number of decoding iterations has roughly the same *error correction capability*, but an inferior *multiple access capability* in comparison to the RA coded Scheme 4 represented by the square-shaped legends. The number of iterations required by the rate-1/64 RA coded Scheme 3 was found to be $\mathcal{I}_{RA} = 80$ to have a superior performance in comparison to Scheme 4.

Importantly, the achievable normalised user-load of both the convolutional coded MC-IDMA Scheme 2 and repetition coded MC-IDMA Scheme 1 communicating over fading channels is quite remarkable in comparison to the single-user AWGN scenario, as seen in Fig. 3.15. This is a benefit of the so-called multiuser diversity, which will be discussed later. In fact, the achievable user-load improvements increase without limits upon increasing E_b/N_0, as stated in [120] in the context of DS-CDMA. These benefits can be further increased by using MUDs having a near-single-user performance or optimised unequal power/rate allocation [85].

Remarks. Based on the above analysis, we found that the joint coding and spreading design of the MC-IDM and MC-IDMA systems have to take into account the specific E_b/N_0 values encountered. At high E_b/N_0 values, DS-spreading dispensing with FEC was capable of supporting

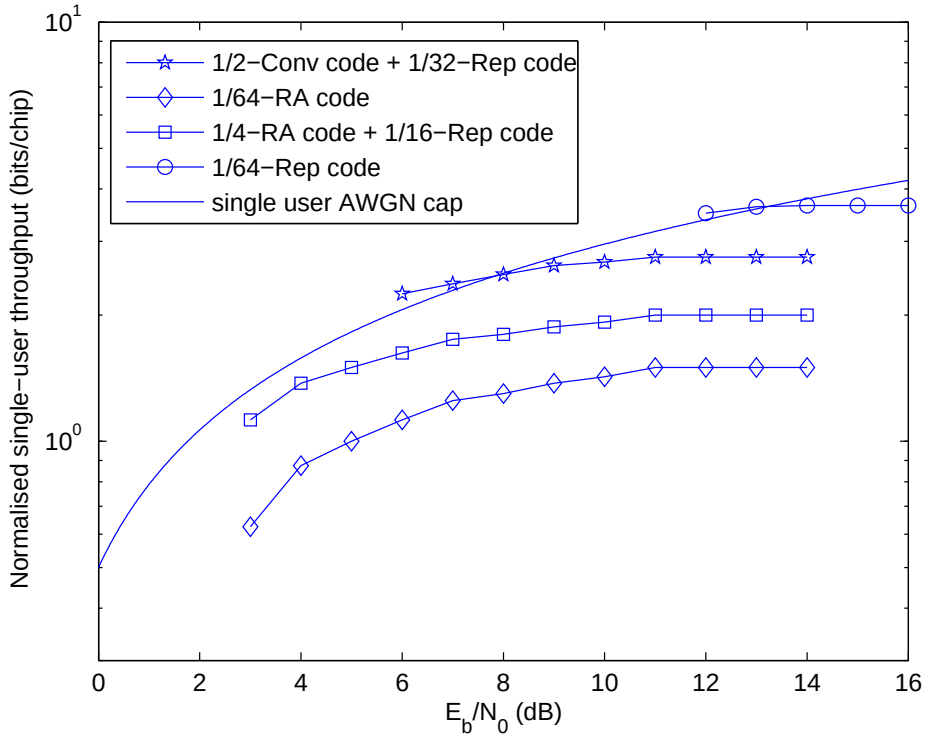


Figure 3.15: Normalised user-load versus E_b/N_0 performance of the 128 subcarrier MC-IDMA for a fixed bandwidth expansion factor of $\Omega = 64$ and for $P_e = 10^{-5}$ using the various channel coding schemes considered in a 5-path chip-spaced negative exponentially decaying uncorrelated Rayleigh fading channel.

the highest normalised throughput. By contrast, at low E_b/N_0 values the employment of explicit channel coding becomes crucial, since it improves the achievable BER at the cost of reducing the attainable multiple access capability.

The authors of [113] found that the employment of channel coding was always beneficial for MF based iterative MUDs in DS-CDMA systems, regardless of the E_b/N_0 values encountered. When considering the MC-IDM and MC-IDMA systems, which also employ MF based iterative MUD, our conclusions stated above were similar to those drawn for a MMSE based iterative MUD in a DS-CDMA system, as reported in [113], which potentially requires high-complexity matrix inversion operations. This suggests that the employment of classic CDMA DS-spreading codes should be replaced by repetition codes in the context of both MC-IDM and MC-IDMA in the interest of benefiting from both its low-complexity MF based detection and an improved bandwidth efficiency.

In classic DS-CDMA, the different MUDs exploit the diverse correlation properties of the DS-spreading codes differently. Hence they result in different channel coding versus spreading tradeoffs, depending on the specific features of the MUDs employed. The MUD of MC-IDM and MC-IDMA operates on a chip-by-chip basis, thus the joint coding and spreading design becomes a single joint low-rate code design problem. *A rule of thumb, when designing MC-IDM and IDMA systems for operating in various channel conditions at a given E_b/N_0 constraint and at an afford-*

Scenario	SC-IDM / IDMA	MC-IDM	MC-IDMA
Channel	AWGN	Rayleigh	
Type of Detection	PIC		
Type of Code	1/8-Rep		
Ω	8		
E_b/N_0	14		
η	2		3.5

Table 3.3: Summary of the parameters and main results of multiplexing versus diversity tradeoffs.

able decoding complexity is that of opting for the specific FEC channel code having the highest rate, which is 'just' capable of meeting the specific target BER requirements and then dedicating the rest of the affordable bandwidth expansion to DS-spreading, i.e. to DS repetition codes. It is also important to notice that appropriate unequal power allocation depending on the specific code [85] is also possible for the sake of achieving high bandwidth efficiency. Then the joint coding and spreading problem becomes irrespective to the code selection. However, this needs a successive interference cancellation aided MUD, which is very much sensitive to error propagation from a practical implementation point of view.

3.2.4.2 Multiplexing versus Diversity Trade-offs

In an MC-IDMA system, at any chip instant, each user's chip may be differentiated by their different FDCTFs, which is essentially equivalent to the concept of SDMA [115] [3]. By contrast, in an MC-IDM system, each stream's signal is differentiated by its unique stream-specific chip-interleaver, noting that there is no way of differentiating among the streams at any chip instant, since their FDCTFs are identical.

In Fig. 3.16, we compare the three typical channel coded/spread system configurations summarised in Table 3.3, all of which have a total bandwidth expansion factor of $\Omega = 8$ and operate at $E_b/N_0 = 14dB$. Explicitly: 1) Scenario 1 - AWGN channel representing a constant equal-power system; 2) Scenario 2 - MC-IDM representing a multiplexing system designed for supporting the highest possible throughput for a single user, which can also be viewed as a DL broadcast system scenario or the scenario of co-located users/antennas; 3) Scenario 3 - MC-IDMA representing an UL multiple access system supporting geographically dispersed users/antennas. Based on Fig. 3.16, we infer the following conclusions:

1. In the noise-limited right-hand-side area of Fig. 3.16, the interference was almost completely cancelled, resulting in a near-single-user performance. Observe in Fig. 3.16 that as expected, Rayleigh fading imposes a detrimental effect on both the MC-IDMA and MC-IDM type

[3] SDMA [115] is a multiple access technique where each user is differentiated by their user-specific channel impulse response.

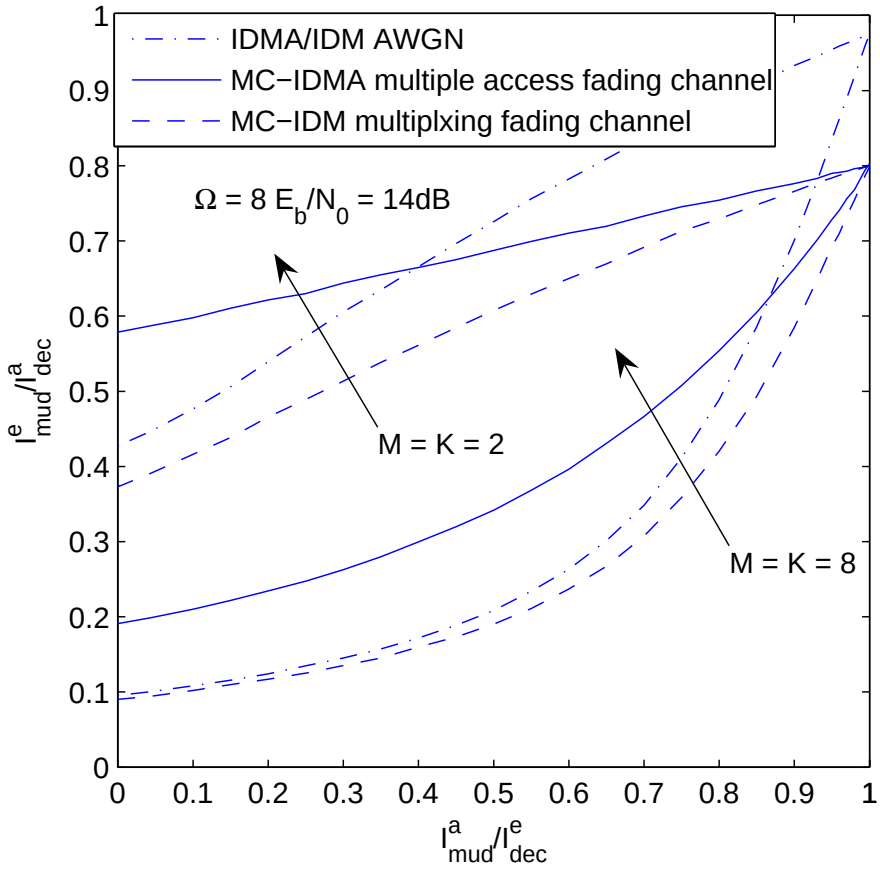


Figure 3.16: The MUD's EXIT curve for both the 128-subcarrier MC-IDMA and MC-IDM systems for transmission over the 5-path uncorrelated Rayleigh fading as well as for an AWGN channel environment.

systems compared to the AWGN channel, when operating at $E_b/N_0 = 14dB$.

2. For the multiplexing Scenario 2 using $M = 2$ and $M = 8$, the extrinsic information recorded for the MC-IDM system under fading channel conditions was upper-bounded by the AWGN channel's curve. This implies that for the MC-IDM Scenario 2, where the same CIR is experienced by all streams of a given user, the fading always exhibits detrimental effects, leading to either a higher BER or requiring a higher E_b/N_0 value. Nonetheless, the increased single-user throughput may be considered to be the explicit benefit of the system's *multiplexing gain*.

3. In contrast to the MC-IDM Scenario 2, when the MC-IDMA Scenario 3 was employed, in the interference-limited region of Fig. 3.16, the system attained an increased output extrinsic information under fading channel conditions in comparison to the AWGN channel scenario, where each user was assumed to experience *independent* fading. This is because each chip can be differentiated by its different FDCTF, resulting in a more reliable estimation of each chip, or equivalently, in a higher extrinsic information. This beneficial effect may be referred to as *multiuser diversity* [120].

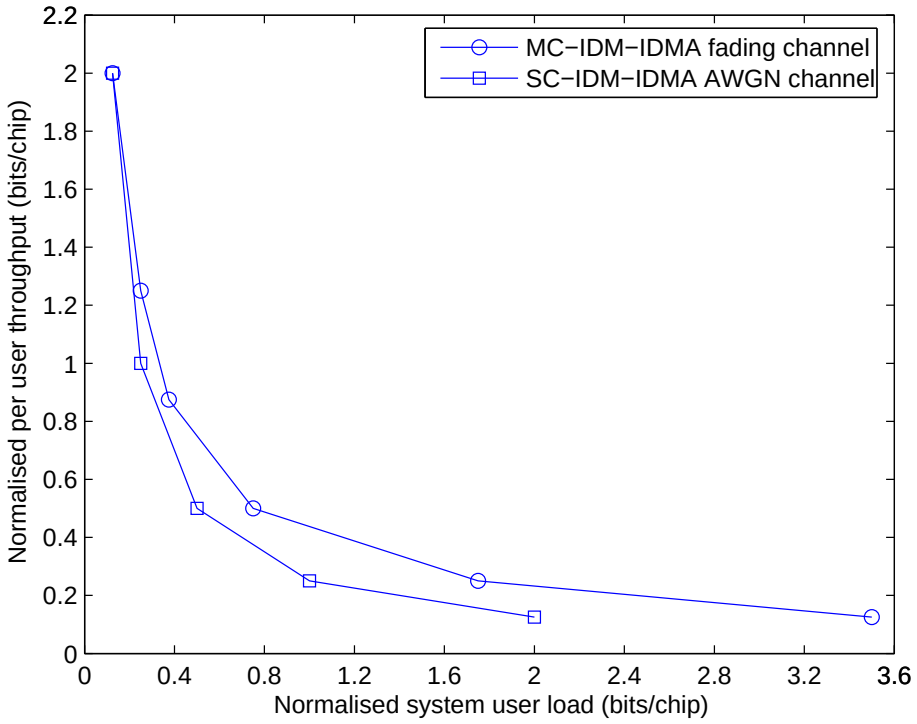


Figure 3.17: The tradeoff between the normalised per user throughput (multiplexing gain) for the 128-subcarrier MC-IDM system and the normalised user-load (diversity gain) of 128-subcarrier MC-IDMA in a 5-path chip-spaced negative exponentially decaying uncorrelated Rayleigh fading channel.

4. Since the *a priori* information was increased during the consecutive PIC MUD iterations, the interference was essentially cancelled, and as a result, a gradually reduced multiuser diversity benefit was observed. Thus, the improvements recorded as a benefit of multiuser diversity in the MC-IDMA Scenario 3 become less and less compared to the equivalent MC-IDM Scenario 2, as seen in Fig. 3.16.

Let us now investigate our proposed MC-IDM-IDMA system's multiplexing versus multiuser diversity gain. Fig. 3.17 shows the associated tradeoffs between the multiplexing gain (per-user throughput) provided by the MC-IDM Scenario 2 and the diversity gain (user-load) supported by the MC-IDMA Scenario 3. The achievable gains were normalised to the total bandwidth expansion factor of $\Omega = 8$, dedicated to length-8 repetition coding at $E_b/N_0 = 14dB$, when communicating over both a dispersive Rayleigh fading channel as well as in AWGN channel conditions. It is worth emphasising that in the MC-IDM-IDMA system, it is the total system's normalised user-load that is higher than that of the AWGN channel scenario, while the individual users' effective throughput is typically low. The maximum normalised per-user throughput provided by the MC-IDM Scenario 2 is seen to be 2 bits/chip in Fig. 3.17, while the maximum normalised user-load provided by the MC-IDMA Scenario 3 is 3.5 bits/chip in Fig. 3.17. Our proposed system is thus flexible in terms of assigning the total system throughput to a single IDM user or providing access for several users.

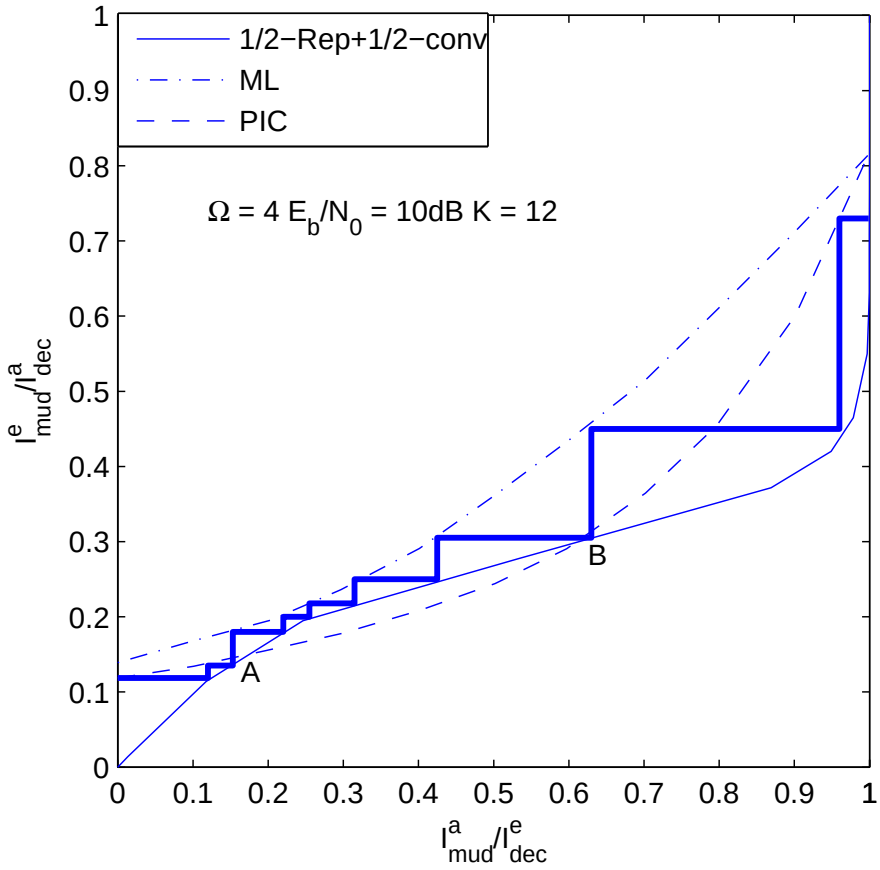


Figure 3.18: EXIT chart analysis of different MUDs.

3.2.4.3 Complexity versus Performance Trade-offs

For the sake of achieving higher user-load, more powerful MUD algorithms are necessary instead of employing sophisticated power allocation schemes. In this subsection, we investigate the EXIT curves of the optimum ML MUD and the currently proposed PIC MUD. Furthermore, a hybrid MUD is proposed, which is capable of approaching the ML MUD's performance at a reduced complexity.

Fig. 3.18 compares the T_{mud} function of the $K = 12$-user MC-IDMA system communicating over a dispersive Rayleigh fading channel scenario having a bandwidth expansion factor of $\Omega = 4$ and $E_b/N_0 = 10dB$, when employing a sub-optimum MF based PIC MUD and ML MUD as listed in Table 3.4. At the left of Fig. 3.18, which corresponds to the interference-limited region, the ML MUD outputs only marginally higher extrinsic information than the PIC MUD. This implies that the achievable multiuser diversity is sufficiently high for almost entirely compensating for the suboptimal nature of the MF based PIC MUD algorithm compared to the optimum ML MUD. As the amount of available *a priori* information increases during the consecutive iterations, the gap between these two T_{mud} curves becomes more substantial owing to the reduced multiuser diversity and the suboptimal nature of the MF based PIC MUD. Ultimately, these two T_{mud} curves tend to the corresponding single-user performance, when the effects of MUI have been more or less

Detection	PIC	ML
Type of System	MC-IDMA	
Type of Code	1/2-Rep+1/2-Conv	
Ω	4	
E_b/N_0	10	
η	3	

Table 3.4: Summary of the parameters and main results of complexity versus performance trade-offs.

eliminated.

However, the complexity order of a ML detector increases exponentially with the number of users according to $\mathcal{O}(2^K)$, while that of a PIC detector increases linearly with the number of users obeying $\mathcal{O}(K)$ according to [31]. Hence, designing a MUD scheme which is capable of approaching the ML performance with reduced complexity is of interest. This may be achieved with the aid of a hybrid MUD [121]. As seen in Fig. 3.18, during the initial few iterations we activate the PIC MUD. Then the ML MUD is activated, in order to overcome the PIC MUD's EXIT curve intersection between points A and point B in Fig. 3.18, where the PIC MUD's EXIT curve T_{mud} is below the EXIT curve T_{dec} of DEC. After this bottleneck, the PIC MUD is reactivated in the interest of reducing the complexity imposed. Additionally, other near-ML MUDs can be employed for the sake of further reducing the complexity imposed, such as Genetic Algorithm and Sphere Decoding aided MUDs [115].

3.2.5 Practical Application of MC-IDMA

3.2.5.1 Comparison of OFDMA, MC-CDMA and MC-IDMA

After a detailed design analysis of MC-IDM-IDMA system, we now compare OFDMA, MC-CDMA and MC-IDMA systems both in the UL seen in Table 3.5 as well as in the DL seen in Table 3.6. In summary, as an non-orthogonal MA scheme, MC-IDMA has all the advantages and disadvantages of non-orthogonal multiple access schemes in comparison to orthogonal ones. More explicitly:

1. It supports a higher user-load than the orthogonal scheme.
2. It does not require coordination of the transmissions of the users, hence it is capable of asynchronous communications..
3. It suffers from near-far effects, although this drawback can be eliminated using a turbo MUD.
4. Apart from the above previous three considerations based on UL, in the DL MC-IDMA is capable of mitigating the intercell interference, while at the same time suppressing the

Uplink (UL)	SC-CDMA	SC-IDMA
Type of MA	non-orthogonal	non-orthogonal
Near-far effect	sensitive	sensitive
Coordination	not required	not required
ISI mitigation	equalisation / MC	equalisation / MC
Receiver Type	arbitrary	turbo MUD
Complexity at BS	flexible	high
Multiple Access	a high user-load by employing MUD	the user-load may be as high as 3.5 bits/chip or even higher, when employing unequal-power allocation or sophisticated MUD
Per-user throughput	Modulation / Multicode	Modulation / IDM as high as 2 bits/chip achieving 3.5 bits/chip by assigning stream-specific phase rotation

Table 3.5: Comparison of SC-CDMA and SC-IDMA in the UL

Downlink (DL)	OFDMA	MC-CDMA	MC-IDMA
orthogonality	orthogonal	non-orthogonal	non-orthogonal
Intracell interference	no	suppressed by MUD	suppressed by MUD
Intercell interference	present	mitigated	mitigated
Receiver Type	SUD	arbitrary	turbo MUD
Complexity at MS	low	flexible	high
Per-user throughput	moderate	moderate / high	high

Table 3.6: Comparison of OFDMA, MC-CDMA and MC-IDMA in the DL

intracell interference using a MUD.

3.2.5.2 Implementation Example of IDMA

Bandwidth expansion allows the IDMA system to accommodate multiple users. The traditional spread-spectrum systems, such as DS-CDMA, provide a low data rate per-user, which is unsuitable for high-rate data services. MC-IDM-IDMA, on the other hand, is flexible in terms of its per-user throughput and multiple access capability. We demonstrated in Fig 3.17 that the per-user throughput can be as high as 2 bits/chip, when BPSK modulation is employed and 3.5 bits/chip total system user-load can be achieved. This is achieved by using a turbo-receiver at the BS, where the detection front-end employs MF based MUD having a complexity which increases linearly with the number of users/streams. Even higher throughput can be achieved by power/rate optimisation or more sophisticated detectors. However, due to its complexity, the turbo-receiver may be more applicable

$\Omega = 8$	CDM	1 code		6 code	
		480kbps		2.88Mbps	
	IDM	$\eta = 1/8$	$\eta = 1$	$\eta = 2$	$\eta = 3.5$
		480kbps	3.84Mbps	7.68Mbps	13.44 Mbps
$\Omega = 16$	CDM	1 code		6 code	
		240kbps		1.44Mbps	
	IDM	$\eta = 1/16$	$\eta = 1$	$\eta = 2$	$\eta = 3.5$
		480kbps	3.84Mbps	7.68Mbps	13.44 Mbps

Table 3.7: Achievable throughput of IDM concept in UTRA FDD UL system.

in the UL than in the DL. Hence, MC-IDM-IDMA is applicable in the UL for both delay-sensitive services, such as interactive voice and video applications as well as for high data-rate applications.

For example, in the UTRA FDD UL, which employs Orthogonal Variable Spreading Factor (OVSF) codes has length of 4-256 chips in the Dedicated Physical Data CHannels (DPDCH). The scrambling is done by Gold codes for a 38 400chips per 10ms frame. Given a 5MHz bandwidth, a bandwidth expansion factor of $\Omega = 8$, this corresponds to a data-rate of 480 kbps. In MC-IDM, where a 2 bits/chip throughput is observed, this results in a 7.68 Mbps per-user throughput. In MC-IDMA or when using stream-specific phase-rotations in the context of MC-IDM, a 3.5 bits/chip per-user throughput can be achieved, corresponding to 13.44Mbps, which is achieved with the aid of 28 IDM streams as seen in Table 3.7.

3.3 Interleaved Random Code Division Multiple Access

In this section, we Generalized the concept of IDMA, leading to the so called GCDMA scheme, which may be defined as a multiple access scheme that separates users in the 'generalised' code domain constituted by either classic spreading codes or by sophisticated channel codes, while sharing the same time-slots and frequency bands.

3.3.1 Introduction

In this generalised spirit, apart from the well-known DS-CDMA [28] scheme, two lesser-known GCDMA schemes were also proposed, namely TCMA [29] and IDMA [31]. In the former, the separation of the users is achieved by the unique combination of user-specific Generator Polynomials (GP) combined with bit-to-symbol mapping schemes and interleavers, whilst the latter employs user-specific interleavers, which may be regarded as rate-one channel codes. These two GCDMA techniques are instantiations of our more general non-orthogonal random waveform based multiuser communications, where the employed interleaved codes obey the random-coding principle. Alternatively, random-coding principle may be achieved by the family of LDPC codes. Hence we

will jointly refer to these schemes using the generic terminology of IR-CDMA.

Typically, a relatively short code constraint length is favoured in TCMA systems in order to attain a reasonably low decoding complexity. Naturally, this reduces both the number of GPs and the number of users supported. It reduces the probability of encountering random-like, low-correlation codewords. It was shown in [29] that in TCMA the employment of a unique user-specific interleaver after the TCM scheme is essential for the sake of attaining an attractive multi-user BER performance, since the interleaved codewords become more random-like and potentially impose a reduced interference owing to their lower correlation. This is due to the fact that the cardinality of the set of interleaved codewords $|\pi(\mathcal{C})|$, becomes significantly larger than that of the $|\mathcal{C}|$, where $\mathcal{C}$ represents the set of all possible TCM codewords and π represents the set of user-specific interleavers. Hence TCMA may be viewed as a special case of IDMA, employing TCM codes as the outer channel code and removing the DS-spreading operation of IDMA, thus resulting in a potentially narrowband multiple access system. A specific feature of TCMA is that each user's transmitted symbol typically contains more than two bits. This potentially renders the receiver complex, when the number of users is high. High-complexity ML detection was proposed in the original contribution on TCMA [29]. Another potential complexity-related issue is the high peak-to-average power ratio of the transmitted signal having a large modulation signal constellation [115]. By contrast, IDMA has the benefit of employing low-complexity binary signalling for each user and yet, it is capable of achieving a high overall spectral efficiency. Naturally, increasing the number of users is achieved at the cost of sacrificing the individual users' throughput. Thus, IDMA is suitable for relatively low-rate UL communications. It was also shown in [31] that purely code-spread IDMA but no spreading combined with appropriate power allocation is capable of approaching the channel capacity, with the aid of using Turbo-Hadamard codes for example [122].

Based on the random coding principle, the family LDPC codes [17] becomes also attractive in the context of [123], since it has a built-in interleaver, which renders the employment of an extra user-specific interleaver after channel coding unnecessary. However, low-rate LDPC codes is hard to design. On the other hand, RA [124] codes having diverse coding rates constitute attractive candidates for employment in IR-CDMA. Similar to LDPC codes, RA codes also have an innate interleaver, while exhibiting a linearly increasing encoding complexity as a function of the codeword length. In this chapter, we propose non-systematic regular RA codes aided IR-CDMA, where different users are distinguished by their unique user-specific RA generator matrices, i.e. their interleavers. We design a SE interleaver generation technique for the sake of reducing the interleaver's storage requirements and investigate its correlation properties compared to those of random interleavers. We also employ the hybrid detection scheme for the sake of supporting a high user-load in an equal user-power scenario. Furthermore, we discuss a low-complexity multi-user power-allocation method designed for our RA codes aided IR-CDMA system. It will be demonstrated that when a sufficiently high interleaver length is employed, the RA codes aided IR-CDMA system becomes capable of approaching the Gaussian channel's capacity.

In all, the novelty and contribution of this section is that

1. We introduce the general concept of IR-CDMA, which employs random-coding principle. The conventional TCMA and IDMA are its subclasses when the random-coding principle is achieved by interleaved codes.

2. We propose low-rate RA codes aided IR-CDMA when the outer code employed is quasi-random codes, where the interleaver is embedded in the GM. In particular, we circumvent the memory problem in the context of multiuser communications by our novel structured design without sacrificing the BER/BLER performance.

3. Furthermore, in order to enhance the user-load of the multiuser communications system, we propose both a hybrid detection scheme in equal-power scenario as well as a low-complexity capacity approaching power-allocation method.

The rest of the section is organised as follows. In Subsection 3.3.2, we present the transceiver architecture of our proposed system. In Subsection 3.3.3, we discuss several design aspects of our RA codes aided IR-CDMA system, namely the SE interleaver generation technique advocated, the hybrid multiuser detector designed for equal-power allocation and finally, we suggest a low-complexity unequal power allocation scheme. Finally, we conclude our discourse in Section 3.4.

3.3.2 System Overview

Consider a K-user non-orthogonal multiuser communication system. The kth user's signal b_k is encoded by its user-unique channel code $\mathcal{C}_k$ at a rate of r, resulting the codeword $x_k = \mathcal{C}_k(b_k)$. In a traditional interleaved codes aided IR-CDMA, e.g. IDMA system, the channel code could be the same for all users and a user-unique interleaver is employed, i.e. $x_k = \pi_k[\mathcal{C}(b_k)]$. The canonical discrete-time real-valued model as seen in Fig. 3.19 for the multiple access channel is given by:

$$y = \sum_{k=1}^{K} H_k x_k + n, \tag{3.18}$$

where $x_k \in \{\pm 1\}$, y and $n \sim \mathcal{CN}(0, N_0)$ denotes the transmitted signal, received signal and AWGN signals, respectively. $H_k = h_k \rho_k$ denotes the equivalent channel of user k, where ρ_k contains the kth user's amplitude subject to power allocation, h_k denotes the identical independently distributed CIR of user k.

An iterative receiver, consisting of a soft detector and a bank of K individual soft decoders is employed for the sake of seeking a tradeoff between the complexity of joint multiuser detection and decoding as well as the performance loss due to separate multiuser detection and single-user decoding. We introduce two detection algorithms, namely the optimal ML detector and a low complexity IC aided detector [31].

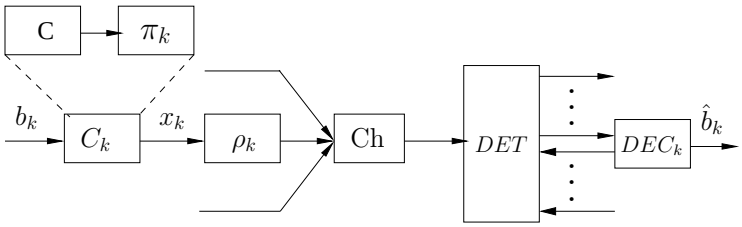


Figure 3.19: Transceiver of the proposed RA-CDMA system

ML detector: The ML detector calculates the extrinsic information $\mathcal{L}^e_{det}(x_k)$ based on the sum of all the K users' 2^K conditional probabilities for the lth legitimate noiseless output $\hat{y}_l$, given all other bits' *a priori* information ϕ corresponding to $x^l_k = \pm 1$, where $l \in 2^K$:

$$\mathcal{L}^e_{dec}(x_k) = \log_2 \frac{\sum_{\forall l: x^l_k = +1} \exp\left(-\frac{(\|y-\hat{y}_l\|)^2}{2\sigma_n^2} + \phi\right)}{\sum_{\forall l: x^l_k = -1} \exp\left(-\frac{(\|y-\hat{y}_l\|)^2}{2\sigma_n^2} + \phi\right)}, \tag{3.19}$$

where

$$\phi = \frac{1}{2} \sum_{j \neq k} x^l_j \mathcal{L}^a_{det}(x^l_j). \tag{3.20}$$

IC aided detector: Let us now rewrite Eq. (3.18) as

$$y = H_k x_k + \xi, \tag{3.21}$$

where $H_k = h_k \rho_k$ represents the composite CIR of user k, while $\xi = \sum_{j \neq k}^{K} H_j x_j + n$ represents the interference plus noise. In the case of binary modulation, the real part (Re) of $H^*_k y$ constitutes sufficient statistics for estimating x_k, resulting in

$$Re(H^*_k y) = |H_k|^2 x_k + Re(H^*_k \xi), \tag{3.22}$$

where $(\cdot)^*$ denotes the conjugate complex value of a variable. We denote the soft estimate of a variable a by $(\hat{a})$. Then, $Re(H^*_k \hat{\xi})$ and its variance $V[Re(H^*_k \hat{\xi})]$ are given by:

$$Re(H^*_k \hat{\xi}) = H^{Re}_k \hat{y}^{Re} + H^{Im}_k \hat{y}^{Im} - |H_k|^2 \hat{x}_k \tag{3.23}$$

$$V[Re(H^*_k \hat{\xi})] = (H^{Re}_k)^2 V(\hat{y}^{Re}) + (H^{Im}_k)^2 V(\hat{y}^{Im}) - |H_k|^4 V(\hat{x}_k) + 2H^{Re}_k H^{Im}_k \phi, \tag{3.24}$$

where $\phi = \sum_{k=1}^{K} H^{Re}_k H^{Im}_k V(\hat{x}_k)$ and $Im(\cdot)$ represents the imaginary part of a complex number. The soft estimate $\hat{y}$ and its variance can be expressed by:

$$\hat{y}^{Re} = \sum_{k=1}^{K} H^{Re}_k \hat{x}_k, \tag{3.25}$$

$$V(\hat{y}^{Re}) = \sum_{k=1}^{K} (H^{Re}_k)^2 V(\hat{x}_k) + \sigma_n^2. \tag{3.26}$$

We remark that the above two equations would be equally valid for the imaginary counterpart. The soft bit $\hat{x}_k$ can be represented as $\hat{x}_k = \tanh[\mathcal{L}^e_{dec}(x_k)/2]$, while its variance is given by $V(\hat{x}_k) = 1 - \hat{x}_k^2$. Assuming ξ is Gaussian distributed, the extrinsic information $\mathcal{L}^e_{det}(x_k)$ is given by:

$$\mathcal{L}^e_{det}(x_k) \quad = \quad 2|H_k|^2 \frac{Re(H_k^* y) - Re(H_k^* \hat{\xi})}{V[Re(H_k^* \hat{\xi})]}. \tag{3.27}$$

Then the extrinsic information forwarded from the detector is used as *a priori* information of the RA decoder, which in turn generates more reliable extrinsic information $\mathcal{L}^e_{dec}(x_k)$ for the next iteration.

3.3.3 Design of RA Code Aided IR-CDMA

RA [124] codes having diverse coding rates constitute attractive candidates for employment in IR-CDMA. Similar to LDPC codes, RA codes also have an innate interleaver, while exhibiting a linearly increasing encoding complexity as a function of the codeword length. We now propose non-systematic regular RA codes aided IR-CDMA, where different users are distinguished by their unique user-specific RA generator matrices, i.e. their interleavers.

3.3.3.1 Interleaver Generation

Consider non-systematic regular RA codes, which have a repetition depth of L and information packets length of Q. The generator of an RA code can be designed directly in a structured way, which is represented by $\mathbf{G} = \mathbf{G}_1 \times \mathbf{G}_2$, where $\mathbf{G}_1$ is a $(Q \times QL)$-element matrix specified by the interleaver and $\mathbf{G}_2$ is a $(QL \times QL)$-element matrix specified by the accumulator:

$$\mathbf{G}_2 = \begin{bmatrix} 1 & 1 & 1 & & 1 & 1 \\ 1 & 1 & 1 & & 1 & 1 \\ 0 & 1 & 1 & & 1 & 1 \\ & & & \ddots & & \\ 0 & 0 & 0 & & 1 & 1 \\ 0 & 0 & 0 & & 1 & 1 \end{bmatrix}.$$

The user-specific RA codes become unique by appropriately designing their inherent user-specific interleavers π^u_k. Inspired by the efficient Nested interleaver concept of [125], here we propose novel user-specific SE interleavers.

Construction of SE Interleavers. The proposed SE interleavers are constructed from a system-specific base interleaver, a user-specific base interleaver, a so-called constituent interleaver set, which are then subjected to a position sorting operation, all of which are detailed below.

The *system-specific base interleaver* π^b is a randomly generated interleaver of length-Q. Each user has a distinct *user-specific base interleaver* $\pi^b_k, k = 1, \ldots, K$ having the same length-Q as the system-specific base interleaver π^b. The $(k+1)$st user-specific base interleaver is an interleaved

version of the kth user-specific base interleaver rearranged by the system-specific base interleaver π^b, as follows: $\pi^b_{k+1} = \pi^b(\pi^b_k)$ and $\pi^b_1 = \pi^b$.

The *constituent interleaver set* of user k is represented by L number / level of length-Q interleavers, which is formulated as $\pi^s_k = \{\pi_1, \pi_2, \ldots, \pi_L\}$. Each element $\pi_l \in \pi^s_k, l = 1, \ldots, L$ of the constituent interleaver set is a distinct length-Q interleaver, having the same length as the system-specific base interleaver π^b. The $(l+1)$st constituent interleaver is an interleaved version of the lth constituent interleaver rearranged by the user-specific base interleaver π^b_k, according to $\pi_{l+1} = \pi^b_k(\pi_l)$ and $\pi_1 = \pi^b_k$.

Finally, the L number of length-Q interleavers are concatenated to form a unique length-LQ interleaver. This is carried out by the constituent interleaver set *position sorting operation*, as defined by the position mapping function f, which maps the index $q_l = 1, \ldots, Q$ within all the L number of length-Q constituent interleavers $\pi_l \in \pi^s_k, l = 1, \ldots, L$ into a single user-specific interleaver $\pi^u_k = f(\pi^s_k)$ unambiguously mapping the LQ number of input bit positions to the interleaved positions $q = 1, \ldots, LQ$. More specifically, the index $q_l = 1, \ldots, Q$ within any of the L number of length-Q constituent interleaver $\pi_l, l = 1, \ldots, L$ is mapped to $q = (q_l - 1)L + l$.

Cross-correlation Evaluation. Let us now demonstrate the equivalence of our proposed SE interleavers to random interleavers in terms of the correlation metric introduced in [125]. In other words, our goal is to demonstrate that despite its significantly reduced memory requirements, the proposed interleaver generation technique does not increase the correlation between the pairs of interleaved information sequences in comparison to using random interleavers.

The correlation χ between two independently generated random information bit sequences s_1 and s_2 interleaved by two different interleavers π_1 and π_2 is given by the scalar product $\circ$ between $\pi_1(s_1)$ and $\pi_2(s_2)$, which can be written as:

$$\chi = |\pi_1(s_1) \circ \pi_2(s_2)|. \tag{3.28}$$

Since evaluating the correlation amongst all possible pairs of random sequences s_1 and s_2 has a high computational cost, we seek a lower-complexity alternative [125]. We represent s_1 as $s_1 = \sum_{i=1}^{N} \alpha_i \mathbf{b}_i$, where $\alpha_i \in \{\pm 1\}$ and the vector $\{\mathbf{b}_i : \mathbf{b}_i(i) = 1, \mathbf{b}_i(j) = 0\}$ of length N is a vector within the basis set $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N] \subset \mathbb{R}^N$. On the other hand, s_2 can be replaced by generating a set $\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, \ldots, \mathbf{g}_N] \subset \mathbb{R}^N$, where each vector $\mathbf{g}_i$ of length N has an entry of $\mathbf{g}_i(j) = -1$ when we have $j < i$ and $\mathbf{g}_i(j) = 1$ for $j \geq i$. Thus the correlation of Eq. (3.28) becomes the so-called upper-bounded basis correlation vector $\chi^b = [\chi^b_1, \chi^b_2, \ldots, \chi^b_N]$ defined in [125], where each entry $\chi^b_j, j = 1, \ldots, N$ is represented as:

$$\chi^b_j = \sum_{i=1}^{N} |\pi_1(\mathbf{b}_i) \circ \pi_2(\mathbf{g}_j)|. \tag{3.29}$$

Fig. 3.20 demonstrates the normalised average histogram of the correlations recorded for both random interleavers and for our proposed SE interleavers. The total interleaver length was set to

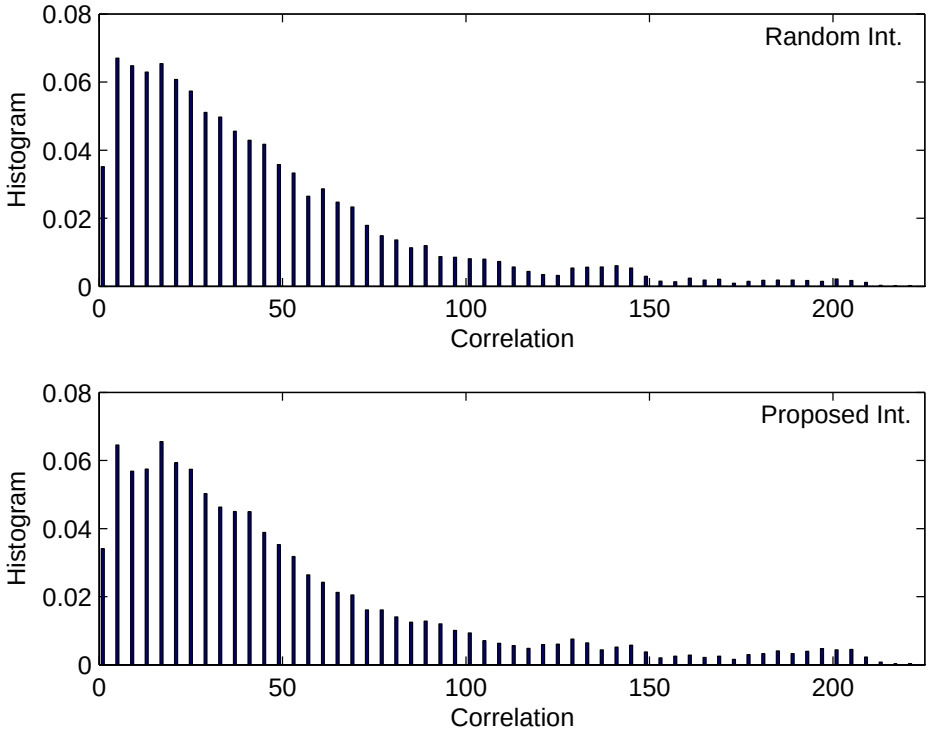


Figure 3.20: The normalised average histogram of the correlation for both the random interleavers and our proposed SE embed interleavers. The total interleaver length was set to $N = 1024$ and we divided it into $L = 4$ constituent interleavers, each having length $Q = 256$. A total number of 100 pairs of interleavers of both random interleavers and our proposed interleavers were averaged.

$N = 1024$ and we divided them into $L = 4$ constituent interleavers, each having a length of $Q = 256$. We averaged the correlations over both 100 pairs of random interleavers and 100 pairs of our proposed SE interleavers. Fig. 3.20 suggests that statistically speaking, both interleaver families exhibit similar correlations.

Furthermore, the system-specific base interleaver π^b can be generated by a shift register based m-sequence generator. Thus, our proposed scheme, which exhibits similar correlation properties to those of random interleavers, does not require the storage of all the K-users' interleavers. Only m bits have to be stored, where m in our system is related to the length $Q = 2^m$ of the source information packets, rather than to the length LQ of the RA aided IR-CDMA coded packets. This is lower than the storage requirements of the basic Nest interleavers proposed in [125], especially when the RA aided IR-CDMA coded packets are long.

Simulation Results Fig. 3.21 compared the BER performance of our proposed RA codes aided IR-CDMA scheme, where the user-separation is achieved by user-specific SE interleavers to that of a conventional randomly generated RA coded IDMA configuration [31], where the user separation is achieved by user-specific interleavers after channel coding. The simulation parameters are summarised in Table 3.8. Fig. 3.21 shows that our proposed system does not suffer from any BER performance loss, which implies that having an extra interleaver after channel coding is redundant in conventional RA coded IDMA systems and hence the employment of 'stand-alone' unique user-

Modulation Type	BPSK
Channel Type	AWGN (Fig.3.21) and UR (Fig.3.23)
Rate of RA Code	0.25
No. of SE levels	4
Info. Packet Length (Q)	1024
Iter No. (Det and Dec)	5 (4 users), 10 (6 users), 20 (8 users)
Iter No. (SPA)	20

Table 3.8: Summary of the simulation parameters.

specific SE generator matrices is sufficient in our RA codes aided IR-CDMA system. When RA code is employed, only a base interleaver is needed, making the memory requirements independent of the number of user K. Furthermore, we note that the difference in the memory requirements of the two systems will become more pronounced upon increasing the number of users Q or the block length N. The proposed system will be applicable in situations, where low-delay requirements are an absolute necessity, for example in speech and video communications.

3.3.3.2 Hybrid Detection

We propose a novel hybrid detection scheme for the sake of supporting large user-load under equal power scenario.

EXIT chart Analysis. Let us firstly investigate the convergence behaviour of the iterative multiuser receiver in an equal user-power scenario, i.e. when we have $\rho_k = \rho, \forall k$ using EXIT charts [36]. Fig. 3.22 compares two detector's EXIT curves for a $K = 8$-user RA-CDMA system communicating over an uncorrelated non-dispersive Rayleigh fading channel, when employing a code-rate of $r = 1/4$ at $E_b/N_0 = 11dB$, where we employed a sub-optimum PIC detector and the optimum ML detector.

At the left of Fig. 3.22, which corresponds to the *interference-limited* region, the ML detector outputs only marginally higher extrinsic information than the PIC detector. As the amount of available *a priori* information increases, the discrepancy between these two EXIT curves becomes more substantial owing to the suboptimal nature of the PIC detector. Ultimately, these two curves tend to the corresponding single-user performance, when the effects of interference have been more or less eliminated, corresponding to the *noise-limited* region at the right of Fig. 3.22.

Hybrid Detector. As seen in Fig. 3.22, the PIC detector has a more limited ability to exploit the extrinsic information during the consecutive iterations compared to the ML detector and hence its performance is suboptimum. Our novel proposition is to use the PIC detector during the first few iterations and then subsequently activate the higher-complexity, but more powerful ML detector for a few further iterations to avoid encountering the "bottleneck" region in the EXIT chart of the PIC detector. Beyond the EXIT-chart's bottle-neck region we may then safely activate again the

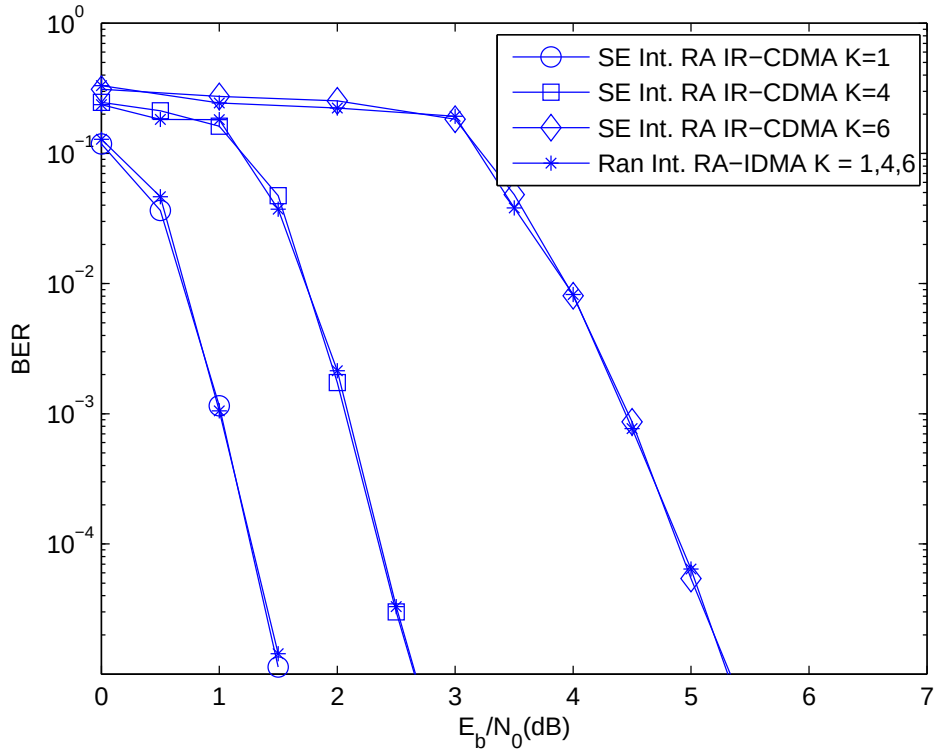


Figure 3.21: Comparison of the BER performance of our proposed SE RA codes aided IR-CDMA to that of conventional randomly generated RA coded IDMA configuration. The simulation parameters are given in Table 3.8.

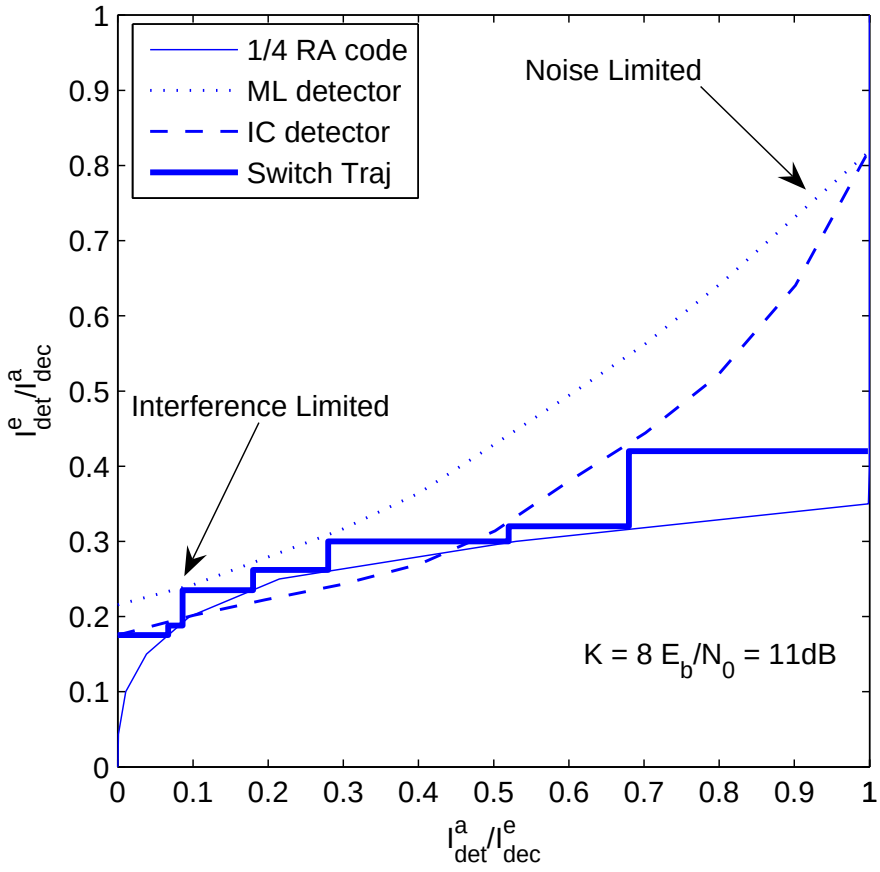


Figure 3.22: EXIT chart analysis of RA codes aided IR-CDMA system.

lower-complexity PIC detector.

In order to determine the activation instant of the different detectors, the average of the extrinsic information's absolute value $|\mathcal{L}_{det}^{e}|$ generated by the detector is invoked, which is supposed to be monotonically increasing as the number of iterations is increased. The value of $|\mathcal{L}_{det}^{e}|$ is recorded after each detector iteration and then compared to that of the previous iteration, resulting in the incremental value Δ as:

$$\Delta = |\mathcal{L}_{det}^{e,l}| - |\mathcal{L}_{det}^{e,l-1}|. \tag{3.30}$$

More explicitly, we determine the detector activation order as follows:

Step 1: Initially, we activate the PIC detector.

Step 2: At iteration l, if $\Delta < \Delta_t$, the ML detector is activated, where Δ_t is a small experimentally predefined positive value.

Step 3: At iteration l, we determine:
1) the actual extrinsic mutual information value $I_{dec}^{e,l}$ from the decoder based on the recorded $|\mathcal{L}_{det}^{e,l}|$.
2) the required *a priori* mutual information $I_{det}^{a,l}$ for the PIC detector to get the same $|\mathcal{L}_{det}^{e,l}|$.
If $I_{dec}^{e,l} - I_{det}^{a,l} > \delta$, where δ is a small positive predefined value, we switch to the PIC detector.

As seen in Fig.3.22, the simulation-based stair-case-shaped detection trajectory closely follows the EXIT curves of the receiver components employing our algorithm. During the initial few iterations, the PIC detector is activated. As the incremental iteration gain of each additional iteration is reduced, the ML detector becomes activated, in order to overcome the PIC detector's EXIT tunnel constriction. When the consecutive iteration gains of the ML detector become sufficiently high again, the PIC detector is reactivated in the interest of reducing the complexity imposed. In order to avoid any potential failure to converge, the threshold of δ used for switching to the PIC detector should be appropriately adjusted.

Simulation Results. We now show that our proposed hybrid detection is capable of approaching the ML performance thus increasing the user-load. Fig.3.23 portrays the attainable BER performance of both the PIC and of the proposed hybrid detectors as a function of the E_b/N_0 and as well as of the number of users supported. The simulation parameters are given in Table 3.8. Observe in Fig.3.23 that at a multiuser sum-rate of $R = K \times r = 2$, the PIC detector fails to converge, while the proposed hybrid detector achieves an infinitesimally low BER at $E_b/N_0 \approx 11dB$. This E_b/N_0 value is only about 3dB away from the ergodic Rayleigh fading channel's capacity [20].

As an explicit benefit, the proposed hybrid detector becomes capable of achieving a similar performance to that of the potentially excessive-complexity ML detector at a moderate complexity, since it is only activated during the critical detector iterations. Moreover, the ML detector can be

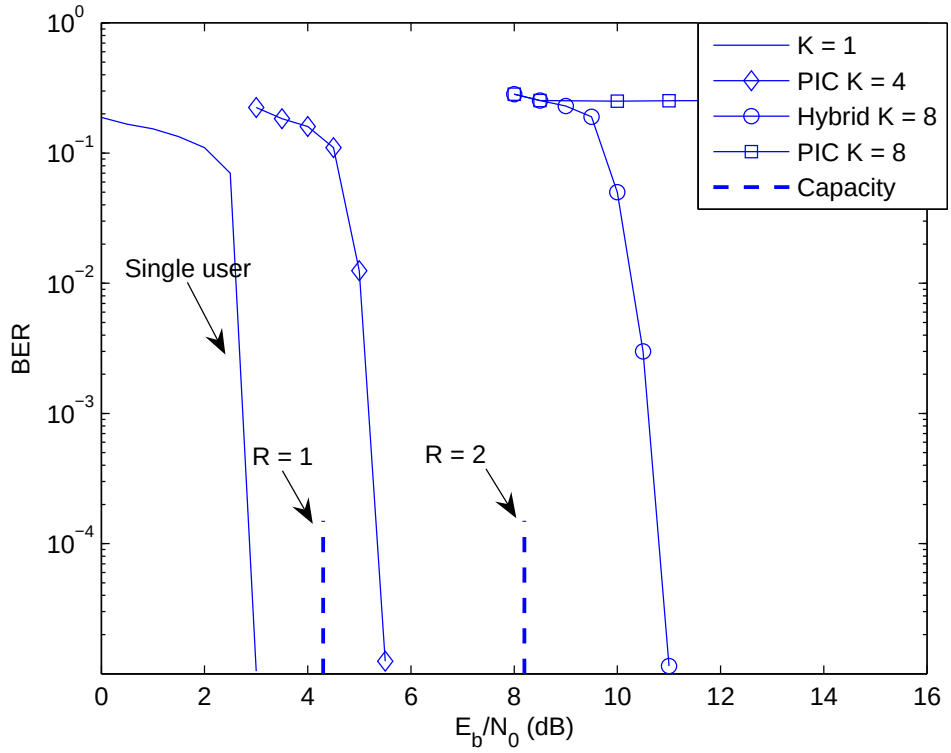


Figure 3.23: Performance of the rate $r = 1/4$ RA codes aided IR-CDMA system for transmission over a non-dispersive, uncorrelated Rayleigh fading channel using the proposed hybrid detector. The simulation parameters are given in Table 3.8.

replaced by any low-complexity near-ML type detector [115], in order to further reduce the overall complexity.

3.3.3.3 Power Allocation

The capacity of the scalar Gaussian multiple access channel can be approached by SIC and decoding combined with optimum power- or rate-allocation [34]. Motivated by this, we propose a low-complexity power allocation scheme for our system employing binary near-capacity RA codes, which are capable of approaching the capacity for sufficiently long information block lengths, given a total multi-user power constraint of $P = \sum_{k=1}^{K} P_k$. We assume having an equal user-rate scenario, i.e. that we have $R = Kr$, where R and r denotes the sum of the user rates and each user's rate, respectively.

The minimum required SNR SNR^m at a given multiuser sum-rate R is given by [34], $\text{SNR}^m = (2^{2R} - 1)$. Our proposed power allocation scheme operates as follows. We set a small incremental value τ for representing the difference between the minimum SNR^m and the required SNR^r at the multiuser sum-rate R as follows:

$$\tau = (\text{SNR}^r)_{dB} - (\text{SNR}^m)_{dB}\,. \tag{3.31}$$

According to the individual RA decoder of user k, we have a threshold value SNR^t, at which the

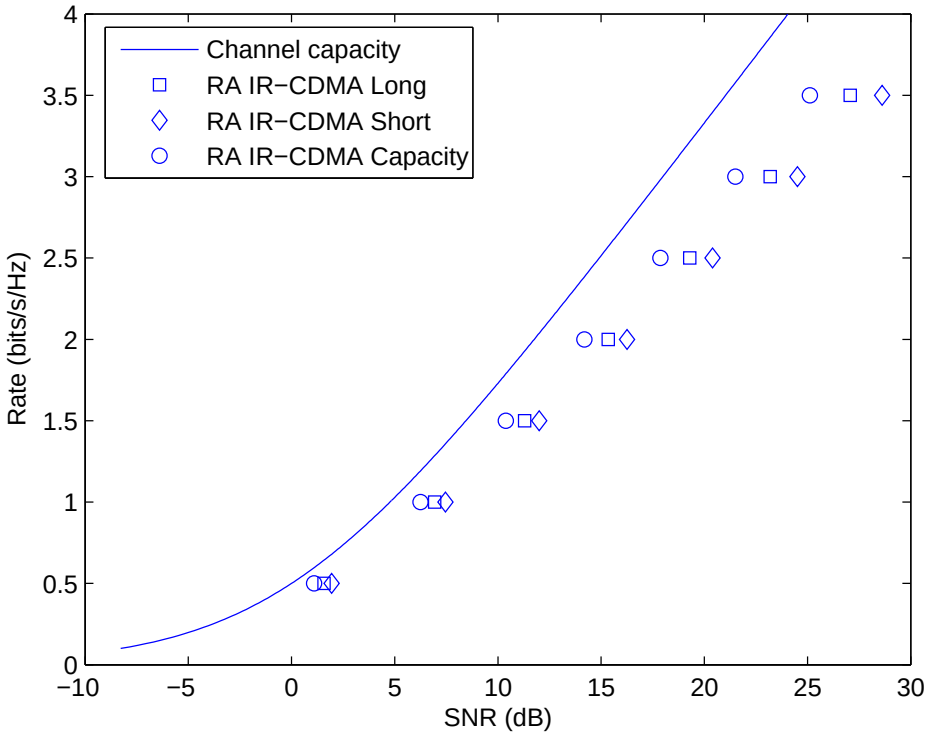


Figure 3.24: The achievable rate of RA codes aided IR-CDMA using the proposed simple power allocation scheme and SE embedded interleavers for both short information bit length of 4096 and long information bit length of 10 000.

RA decoder exhibits an infinitesimally low BER. This threshold may be acquired by simulations or by using a semi-analytic EXIT chart based prediction. Then the power allocated to user k is given by:

$$\hat{P}_k = \mathrm{SNR}^t \left(\sum_{j=k+1}^{K} \hat{P}_j + \frac{P}{10^{\frac{\tau}{10}} \mathrm{SNR}^m} \right). \tag{3.32}$$

Furthermore, the difference between the minimum SNR and the required SNR of user k is given by:

$$\tau_k = \left(2^{2r} - 1\right)_{dB} - \left(\mathrm{SNR}^t\right)_{dB}. \tag{3.33}$$

Following the aforementioned power allocation procedure, we calculate the total power consumption according to $\hat{P} = \sum_{k=1}^{K} \hat{P}_k$. If we have $\hat{P} > P$, we set a larger incremental value τ and reallocate the individual users' power. The power allocation is achieved when we have $\hat{P} = P$ and the resultant threshold value τ quantifies the distance from capacity for the multiuser sum-rate R.

We now show that the RA codes with our proposed simple power allocation scheme and SE embedded interleavers are capable of achieving the Gaussian multiple access channel capacity. Fig 3.24 shows the achievable rate of the RA aided IR-CDMA system using the proposed power allocation regime for both a short information sequence length of 4096 and for a long information sequence length of 10 000 bits. Although not explicitly shown here due to space limitations, when

a rate $r = 1/4$ RA code was employed, the actual E_b/N_0 values required at these two information sequence lengths for achieving an infinitesimally low BER were observed to be at 0.8dB and 0.5dB, when using either bit-by-bit simulation or EXIT chart predictions, respectively, in conjunction with $I_{RA} = 50$ RA decoding iterations. The minimum decoding threshold E_b/N_0 value for a rate $r = 1/4$ RA code having an infinite information sequence length was reported to be 0.1dB when using the sum-product decoding algorithm [124]. Hence we may refer to the corresponding achievable rate as the approximate RA aided IR-CDMA capacity.

Observe in Fig 3.24 that RA aided IR-CDMA system is capable of approaching the Gaussian channel's capacity, when employing the proposed power allocation scheme. When τ_k is small, the overall discrepancy τ of Eq. (3.31) will also be small and as a benefit of the power allocation algorithm contrived, the system operates close to the achievable channel capacity, as seen in Fig. 3.24. However, the overall difference τ formulated in Eq. (3.31) increases further, when the number of users K becomes high, especially when a short information sequence length is used. This is because our power allocation scheme does not take into account the iterative receiver's operation, which results in a small τ at a high multiuser sum-rate. This drawback can be eliminated by the simulation based power allocation scheme of [85], which does take into account the iterative receiver's operation at the cost of a higher complexity. Thus our low-complexity power allocation is suitable in moderate multiuser sum-rate scenario, e.g. $R \leq 3$.

3.4 Conclusion

In this chapter, a specific instantiation of non-orthogonal random waveform based multiuser communications, namely IDMA was introduced, where different users are separated by means of their unique, user-specific explicit chip-interleavers. Based on a detailed literature review concerning the rationale of the transition from DS-CDMA to IDMA and motivated by the concept of chip-interleaving, four important benefits of employing IDMA were considered, namely the benefits of random coding, its improved time diversity, its powerful iterative receiver architecture and its flexible joint coding and spreading design.

In Section 3.2, a novel generalized MC-IDM-IDMA structure was proposed and its convergence behaviour was analysed with the aid of EXIT charts. The differences between MC-IDM and MC-IDMA were highlighted and the multiuser diversity benefits of MC-IDMA were revealed in Section 3.2.4.2. The proposed system is flexible in terms of either assigning the total system throughput to a single user or providing a multiple access capability for several users. Furthermore, in Section 3.2.4.1, the associated coding versus DS-spreading tradeoffs were discussed and our findings were compared to the characteristics of traditional DS-CDMA. It was found that in contrast to the separate DS-spreading and channel coding design of traditional DS-CDMA systems, the DS-spreading operation should be replaced by a repetition code in the context of IDMA so as to be jointly designed with the FEC code for the sake of improving the achievable bandwidth efficiency.

Hence, a general design rule was provided for the low-rate codes of our systems. The associated complexity versus performance tradeoffs were discussed in Section 3.2.4.3 by investigating two MUDs' EXIT curves and a reduced-complexity hybrid MUD concept was provided. Based on these theoretical studies, in Section 3.2.5, we compared MC-IDMA to two mature techniques, namely to OFDMA and to MC-CDMA both in the UL and DL scenario. The potential employment of IDMA in a practical system as a variant of multicode transmission was also advocated. Although these tradeoffs are based on semi-analytical EXIT chart aided techniques, pure analytical investigations may be pursued in our future work.

As a natural generalisation of the random coding principle, we proposed the concept of IR-CDMA, where different users are separated by their unique quasi-random codes. Apart from the IDMA family, user-specific SE interleaver aided RA codes were proposed in the context of IR-CDMA. The specially designed SE interleavers significantly reduce the memory requirements in a multiuser context, while maintaining the target BER as well as BLER performance. This is an explicit benefit of using a structured interleaver design. Since RA codes constitute a class of near-capacity codes, it was also demonstrated that our low-rate RA code aided IR-CDMA system employing the SE interleavers is capable of approaching the Gaussian channel's capacity with the aid of the low-complexity unequal power allocation scheme of Section 3.3.3.3. More explicitly, the discrepancy between the capacity and the achievable throughput of our IR-CDMA is a function of both the block length as well as of the number of users. As a result, the IR-CDMA system becomes more suitable for employment in a moderate multiuser sum-rate scenario.

Chapter 4

Interleaved Random Space-Time Coding and Energy Efficient Network Coding for Multi-Source Cooperation - Cooperative System Application

Apart from the infrastructure based wireless networks discussed in Chapter 3, the concept of non-orthogonal random waveform based multiuser communications can also be applied in wireless *ad-hoc* networks. In Section 4.1, we propose a novel distributed IR-STC scheme designed for MSC employing various relaying techniques, namely Amplify-Forward (AF), Decode-Forward (DF), Soft-Decode-Forward (SDF) and Differential-Decode-Forward (DDF). We characterise the achievable slot utilisation efficiency and introduce a two-phase communication regime for our IR-STC aided MSC in Section 4.1.2. A matrix based formalism is introduced in Section 4.1.3 for describing our IR-STC scheme and the SE random interleaver contrived in Chapter 3 is employed. We characterise the achievable performance of the proposed IR-STC design and compare it to that of the traditional G_2 and G_4 Orthogonal Space Time Block Code (OSTBC) invoked for MSC in Section 4.1.4.1. The performance of our IR-STC is then characterised in conjunction with various relaying techniques under different inter-source Nakagami-m fading channel condition in Section 4.1.4.2.

In Section 4.2, we continue by considering coding schemes designed for energy efficient MSC. More explicitly, in Section 4.2.2, we propose both a powerful SPC scheme and a PANC scheme. The concept of generalised network coding is introduced and the relationship between SPC and Network Coding (NC) is revealed. Our simulation results provided in Section 4.2.4 demonstrate that both of the proposed schemes are capable of performing close to the outage probability bound at relatively low transmit powers. Moreover, compared to the SPC scheme considered, the proposed PANC arrangement imposes a lower complexity at the cost of a slight performance degradation, while maintaining the same throughput and delay.

4.1 Cooperative Interleaved Random Space-Time Coding Scheme

4.1.1 Introduction

MIMO systems [2] are capable of providing both diversity and coding gains in the context of Space-Time Codes (STC) [126] as well as of supporting a high multiplexing gain, when using for example BLAST [127]. However, at the Mobile Station (MS), it may be impractical to accommodate multiple antennas. Apart perhaps from a pair of $\lambda/2$-spaced beamforming elements separated by 7.5cm at a carrier frequency of 2GHz. Even if multiple antennas are accommodated by a laptop's backplane, their signals are often exposed to shadow fading, which imposes simultaneous correlated fading on the elements. In this case, the MIMO-related benefits erode, unless they are combined with HSPA style adaptive modulation [2]. Alternatively, the novel concept of cooperative communications allows us to assign the MIMO elements to geographically separated cooperating MSs, which are no longer prone to shadowing-induced correlated fading, leading to the concept of Virtual MIMO (VMIMO) [128], [129].

Cooperative diversity [130] relying on a distributed MIMO system [115] is capable of eliminating the correlated fading induced diversity-gain erosion of co-located MIMO elements imposed by the omni-present shadow fading. Hence this novel technique is capable of improving the BER performance, while supporting a high throughput as well as providing an improved cell-edge coverage [131]. It has the potential of beneficially combining the traditional infrastructure based wireless networks and the ad-hoc wireless network philosophy [132]. Most research in the literature was dedicated to the aspects of information-theoretic analysis, to creating practical relaying techniques and to the investigation of distributed STC designs [133–135]. Recently, the Cooperative Multiple Access (CMA) channel has attracted substantial research interests [135], where multiple sources forming a cluster of cooperating nodes communicate with the destination, which is also known as MSC [136, 137].

The main objective of cooperative communications is that of simultaneously achieving both a high multiplexing gain and a high diversity gain [138]. Attractive coded modulation schemes were designed for both the spatial and time domain, using for example BICM [21, 61], MLC [59] and Turbo-BLAST schemes [139]. Similar to turbo-BLAST, recently, a multilayer Interleave Division Multiplexing (IDM) aided STC was proposed by Wu and Ping in [140], and its potential applicability in cooperative communications was also alluded to. The resultant IDM-STC was then investigated and analysed in [141], where a similar performance was reported to that attained by Alamouti's STBC.

Against this background, we propose an error-resilient yet high-throughput IR-STC scheme, which is contrived for MSC. Our IR-STC designed for MSC exhibits several beneficial properties: 1) It achieves a *high-throughput* as a benefit of its high slot utilisation efficiency with the aid of the superposition coding concept [57, 58]. 2) It has a *low BER* thanks to the powerful iterative receiver

employed [75]. 3) Our IR-STC design constitutes a *non-orthogonal* scheme, which is capable of approaching the (cooperative) multiple access channel's capacity [135]. 4) The IR-STC proposed has the benefit of operating with the aid of using autonomously generated random interleavers, which facilitates their cooperation without any central controller, even without knowing the number of nodes. This is in contrast to the conventional distributed OSTBC, where each node emulates a specific virtual array element of a structured STC. Thus, our design avoids the requirement of a centralised code allocation procedure. Given the above-mentioned attributes, we refer to the proposed scheme as a *decentralised* non-orthogonal IR-STC.

Our goal is to quantify the slot utilisation efficiency achieved by MSC compared to conventional Single Source Cooperation (SSC), which allows us to simultaneously achieve full diversity and a high throughput. We analyse the achievable rate, power-efficiency and the network-architecture's flexibility and contrast the benefits of IR-STC to those of the traditional G_2 and G_4 OSTBC design [72]. Furthermore, we investigate several relaying techniques in the context of our IR-STC aided MSC, such as the AF, DF, SDF and DDF technique. In a nutshell, the novel contribution of this section is that *we design a decentralised error-resilient, yet high-throughput non-orthogonal IR-STC scheme suitable for MSC and characterise its achievable performance, when employing various relaying techniques and encountering Nakagami-m fading inter-source channels.*

The rest of this section is organised as follows. In Subsection 4.1.2, we describe the MSC scenarios considered, highlight its slot utilisation efficiency and introduce our novel IR-STC architecture designed for MSC. In Subsection 4.1.3, we analyse our IR-STC with the aid of matrix representation. In Subsection 4.1.4, we outline the achievable benefits compared to the traditional G_2 and G_4 OSTBC design and investigate the performance of our IR-STC scheme employing different relaying techniques using simulations. Finally, we conclude our discourse in Section 4.3.

4.1.2 Construction of IR-STC Aided MSC

4.1.2.1 Cooperation Scenario

Consider a cluster of single-antenna sources cooperatively communicating with a destination employing a single receive antenna resulting in a Virtual Multiple Input Single Output (VMISO) system. In this VMISO cluster, we assume having a total of N Cooperating Sources (CS), K Active Sources (AS) and $(N-K)$ Relaying Sources (RS). An example of this is illustrated in Fig. 4.1 having 5 CSs, 4 ASs and 1 RS.

Cooperative communications typically entail two phases. In *Phase-I cooperation*, the source information emanating from all K ASs is broadcast to all N CSs in a Time Division Duplex (TDD) manner under the assumption of perfect synchronisation. By contrast, *Phase-II cooperation* is defined as the joint transmission of a combined IR-STC signal by the concerted action of all the

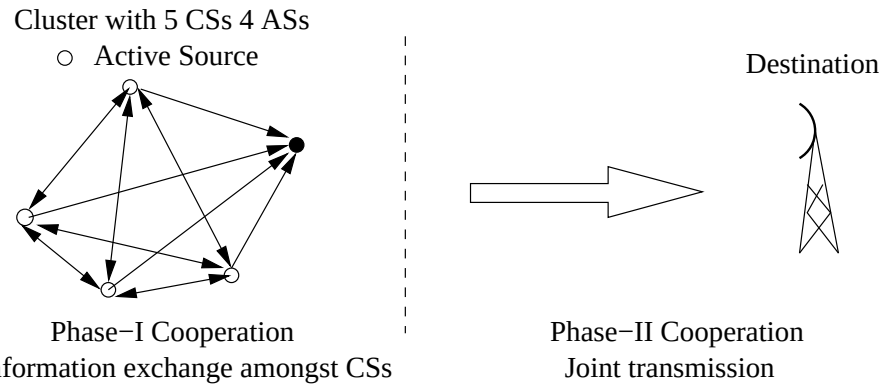


Figure 4.1: A two-phase VMISO cooperation cluster having 5 CSs, 4 ASs and 1 RS.

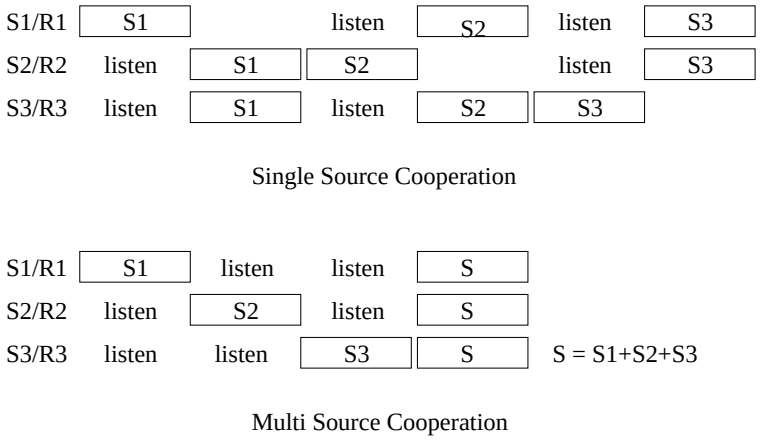


Figure 4.2: Slot utilisation efficiency of Multi-Source Cooperation and Single Source Cooperation.

N CSs. In this Chapter, the superscript $(\cdot)^{(1)}$ and $(\cdot)^{(2)}$ denotes *Phase-I cooperation* and *Phase-II cooperation*.

We assume that all inter-source channels denoted as $h_{s,s}$ and the source-destination channel denoted as $h_{s,d}$ experience i.i.d Nakagami-m fading obeying the PDF of [1]:

$$p_Z(x) = \frac{2m^m}{\Gamma(m)\Omega^m} x^{2m-1} \exp\left\{-\frac{mx^2}{\Omega}\right\}, x \geq 0 \tag{4.1}$$

where $m \geq 0.5$ is the Nakagami-m fading parameter, $\Omega = \mathrm{E}\left\{x^2\right\}$ is the variance of x and $\Gamma(\cdot)$ is the Gamma function. In this section, the inter-source channel $h_{s,s}$ is assumed to be asymmetric, i.e. we have $h_{k,n} \neq h_{n,k}$, where $h_{k,n}$ represents the inter-source channel between source k and source n, which tend to be in close proximity of each other. Since each of the K ASs transmits its information in a *Phase-I cooperation* slot, it is reasonable to assume that the fading envelope is constant in that slot and fades independently for the different slots. We also assume that the inter-source channels benefit from a higher effective SNR, i.e. we have $\gamma_{s,s} > \gamma_{s,d}$. Furthermore, $m > 1$ is used only for the inter-source channels, where an SNR-based node pre-selection scheme may be used for spotting the specific CSs benefiting from a high-quality channel.

4.1.2.2 Slot Utilisation Efficiency

As seen in Fig. 4.2, where $N = K = 3$, in conventional SSC, each source broadcasts its information to all $(N-1)$ CSs during the *Phase-I cooperation*, which is followed by a joint relaying of their information to the destination by the concerted action of the $(N-1)$ CSs in *Phase-II cooperation*. An entire cooperative transmission phase is concluded, when all K ASs completed their cooperation. By contrast, MSC is constituted by a full cycle of information broadcasting from all K ASs to all N CSs during *Phase-I cooperation*, followed by their joint transmission to the destination during *Phase-II cooperation*, where each CS transmits all K ASs' information. Therefore, each CS simultaneously transmits multiple sources' information with the aid of their superposition, resulting in a high throughput. This implies that each source is served simultaneously by multiple CSs, which are chosen to be those that experience a high-quality inter-source channel and hence the entire set of ASs benefits from a high diversity gain.

Let us define the *slot utilisation efficiency* of a cooperative scheme η_s as the ratio between the time duration required for transmitting all K ASs' information in a non-cooperative manner and that necessitated in a cooperative manner. Assume that the information broadcast phase of each source requires a time duration of T_1, while the joint transmission in *Phase-II* takes a time duration of T_2. Thus the slot utilisation efficiency of SSC is $\eta_{s-ssc} = T/(KT_1 + KT_2)$, while that of MSC is $\eta_{s-msc} = T/(KT_1 + T_2)$, where T is the time duration required for transmitting all K ASs' information in a non-cooperative manner, which is $T = KT_1$ when TDMA is used. Therefore, MSC is preferable to SSC in terms of its higher slot utilisation efficiency.

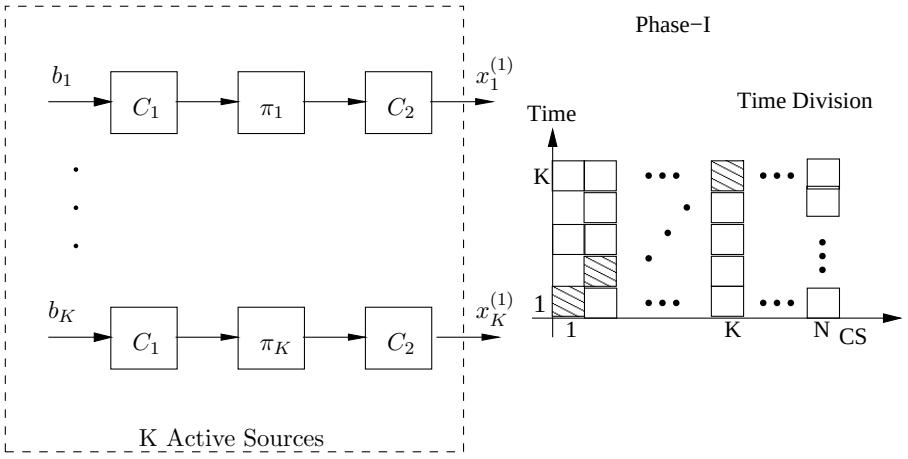
4.1.2.3 Two Phase IR-STC Construction

Phase-I Cooperation. As seen in Fig. 4.3, we assume that each BPSC modulated AS employs two repetition codes C_1 of rate r_1 and C_2 of rate r_2, which are separated by a AS-specific interleaver π_k. During *Phase-I cooperation*, the kth AS transmits a repetition coded and randomly interleaved bit-stream $x_k^{(1)} = \mathrm{C}_2\left\{\pi_k[\mathrm{C}_1(b_k)]\right\}, k \in [1, K]$. Then, depending on whether the inter-source channel $h_{s,s}$ is known at all CSs, two different transmission modes can be employed, namely coherent modulation and non-coherent modulation.

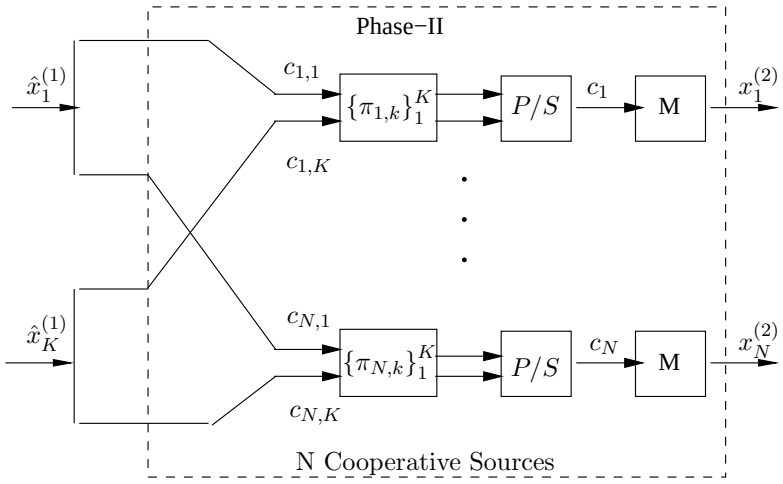
Coherent Modulation. During the kth of the K number of available TDD time-slots, the nth CS receives the signal transmitted from the kth AS, yielding the received signal $y_{k,n}^{(1)} = h_{k,n}x_k^{(1)} + n_0, k \in [1, K], n \in [1, N]_{(k)}$, where $n_0 \sim \mathcal{CN}(0, N_0)$ denotes the complex-valued Additive White Gaussian Noise (AWGN). In this scenario, three relaying techniques are considered.

1. *Amplify Forward.* The signal $y_{k,n}^{(1)}$ received by the nth CS is scaled to meet the average power constraint, yielding:

$$\hat{x}_{k,n}^{(1)} = \frac{y_{k,n}^{(1)}}{\sqrt{N_0 + |h_{k,n}|^2}}. \tag{4.2}$$



(a) Phase I Cooperation



(b) Phase II Cooperation

Figure 4.3: Block diagram of Interleaved Random Space-Time Code aided Multi-Source Cooperation

2. *Soft Decode Forward*. The soft value ($\mathcal{L}$) of the signal $y_{k,n}^{(1)}$ received at the nth CS is calculated as [36] $\mathcal{L} = 4(|h_{k,n}|^2 x_k^{(1)} + \mathcal{R}\left\{h_{k,n}^* n\right\})/N_0$. This is then scaled to meet the average power constraint:

$$\hat{x}_{k,n}^{(1)} = \frac{\mathcal{L}N_0}{4|h_{k,n}|\sqrt{N_0 + |h_{k,n}|^2}}. \tag{4.3}$$

   Eq (4.3) essentially describes an AF technique in an uncoded or repetition coded system, because the soft value $\mathcal{L}$ can be viewed as an equivalent analogue-valued received signal.

3. *Decode Forward*. The signal $y_{k,n}^{(1)}$ received by the nth CS is subject to BPSK hard detection, resulting in

$$\hat{x}_{k,n}^{(1)} = \text{sign}\left(\mathcal{R}\left\{h_{k,n}^* y_{k,n}^{(1)}\right\}\right), \tag{4.4}$$

where the superscripts $(\cdot)^*$ denotes complex conjugation.

Non-coherent Modulation. When $h_{s,s}$ is unknown at the CSs, non-coherently detected differentially encoded BPSK (DBPSK) modulation can be employed. Then the transmitted bit-stream is expressed as $s_k^{(1)}(i) = s_k^{(1)}(i-1)x_k^{(1)}(i), i \in [1, M]$, where M is the length of bit-stream $x_k^{(1)}$ and $s_k^{(1)}(0) = 1$ is a dummy bit used by the DBPSK detector as a reference. Thus, we have $y_{k,n}^{(1)} = h_{k,n}s_k^{(1)} + n_0$.

Let us assume the presence of slow fading. Then $h_{s,s}$ may be considered to be constant over two consecutive bits, hence non-coherent detection is performed according to:

$$\begin{aligned} \hat{x}_{k,n}^{(1)}(i) &= \text{sign}\left(\mathcal{R}\left\{y_{k,n}^{(1)*}(i-1)y_{k,n}^{(1)}(i)\right\}\right) \\ &= \text{sign}\left(\mathcal{R}\left\{|h_{k,n}|^2 x_k^{(1)}(i) + v\right\}\right), \end{aligned} \tag{4.5}$$

where $v \sim \mathcal{CN}(0, 2|h_{k,n}|^2 N_0)$ is a complex-valued AWGN component having a doubled noise variance in comparison to coherent detection, where the latter relies on accurate channel knowledge.

When comparing these four relaying techniques, Eq. (4.2) and Eq. (4.3) retain the original signal, but scale both the signal and the noise component, while Eq. (4.4) and Eq. (4.5) assume first detecting and then reconstructing the signal, depending on the channel quality. We refer to the first two techniques as *non-regenerative* relay techniques and to the latter two as *regenerative* relay techniques.

Phase-II Cooperation Following *Phase-I cooperation*, each of the N CSs detects/scales all the K ASs' bit-streams according to the the above four relaying techniques characterised by Eq. (4.2)-Eq. (4.5). When considering the nth of the N CSs, the joint IR-STC codeword is constructed as follows

1. *Codeword generation*: The nth CS forms K parallel streams

$$c_{n,k}(i) = \hat{x}_{k,n}^{(1)}[N(i-1) + n], \tag{4.6}$$

   where $i \in [1, M/N], k \in [1, K]$. These K streams are interleaved by K distinct interleavers of the CS-specific interleaver set $\{\pi_{n,k}\}_{k=1}^{K}$ and then Parallel-to-Serial (P/S) converted to c_n.

2. *Multilayer mapping*: Then the signal transmitted from the nth CS becomes

$$x_n^{(2)}(i) = \frac{1}{\sqrt{L_n}} \sum_{l=1}^{L_n} \rho_{n,l} e^{j\theta_{n,l}} c_n[L_n(i-1) + l], \tag{4.7}$$

   where $i \in [1, MK/NL_n]$ and L_n is referred to as the number of *layers* contributed by the nth CS, while $\rho_{n,l}$ and $\theta_{n,l} \in [0, \pi)$ denotes the layer-specific amplitude and phase rotation.

In this section, we assume $L_n = L, \rho_{n,l} = \rho_l, \theta_{n,l} = \theta_l, \forall n \in [1, N]$. Furthermore, we employ a layer-specific uniform phase rotation so that the L number of layers are uniformly phase-rotated on the two-dimensional signal space.

In contrast to classic mapping and modulation schemes, such as PSK and QAM, multilayer mapping, which are based on the theory of multiuser communications, was designed to generate an approximately Gaussian distributed transmitted signal, which allows the system to approach the Shannon capacity. The rationale of allocating a different power ρ_l to each of the L layers is philosophically similar to that of the multilevel coding concept, where we create a number of different protection levels and detect them by gleaning extrinsic information from the previously decoded levels using multistage decoding. The associated phase rotation has two benefits, namely that of reducing the Peak-to-Average Power Ratio (PAPR) of the transmitted $x_n^{(2)}$ and making $x_n^{(2)}$ having L layers more distinguishable for the detector.

An *iterative receiver* is employed at the destination of *Phase-II cooperation*, where either optimum but complex ML detection or reduced-complexity suboptimum IC may be employed [140]. Employing different relaying techniques requires different amount of inter-source channel knowledge at the destination. For the regenerative relay techniques of Eq. (4.4) and Eq. (4.5), no inter-source channel knowledge is required at the destination, while for the non-regenerative relay techniques of Eq. (4.2) and Eq. (4.3), inter-source channel knowledge is required at the destination. However for SDF, the knowledge of the inter-source channel's magnitude $|h_{s,s}|$ at the destination is sufficient.

4.1.2.4 Effective Throughput of IR-STC

Now, let us discuss the effective throughput of our IR-STC. In this section, we employ repetition codes of code-rate r_1 and r_2 for both C_1 and C_2 , respectively, resulting in a total code-rate of $r = (r_1 \times r_2)$. When considering a cluster of N CSs, the overall rate of the IR-STC scheme becomes $r_{IR-STC} = rN$. When ignoring the throughput reduction imposed by *Phase-I cooperation* for the sake of a simple argument, the effective throughput of the cluster employing multilayer IR-STC may be expressed as:

$$\eta_{IR-STC} \quad = \quad r_{IR-STC} \times L. \tag{4.8}$$

For example, when $r = 1/8$-rate repetition coded $N = K = 4$ sources are in a cluster and $L = 7$ layers are superimposed at each CSs, an aggregate rate as high as $\eta_{IR-STC} = 3.5$ is achievable.

By contrast, consider having $N = K = 4$ sources in a cluster using a traditional G_4 type OSTBC generator matrix (GM) [72]:

$$\mathbf{G}_4 = \begin{bmatrix} -x_4^* & -x_3^* & -x_2^* & x_1^* & -x_4 & -x_3 & -x_2 & x_1 \\ -x_3^* & x_4^* & x_1^* & x_2^* & -x_3 & x_4 & x_1 & x_2 \\ x_2^* & x_1^* & -x_4^* & x_3^* & x_2 & x_1 & -x_4 & x_3 \\ x_1^* & -x_2^* & x_3^* & x_4^* & x_1 & -x_2 & x_3 & x_4 \end{bmatrix}.$$

The traditional STBC used in MSC operates as follows: 1) exchange the information of $x_k, k \in [1, K]$, which requires K time-slots; 2) joint transmission of $x_k, k \in [1, K]$ using all N CSs according

to the above matrices, where the nth CS transmits the nth row of the above GM. For the sake of achieving a high throughput, each symbol x_k can be modulated onto an $\mathcal{M}$-ary constellation. The effective throughput of the cluster excluding the overhead of the *Phase-I cooperation* can thus be defined as

$$\eta_{OSTBC} = r_{OSTBC} \times \log_2 \mathcal{M}, \tag{4.9}$$

where $r_{OSTBC} = 1/2$ is the rate of the G_4 OSTBC.

4.1.3 Analysis and Design of IR-STC Aided MSC

4.1.3.1 Matrix Representation of IR-STC

Our IR-STC scheme employs a random distributed AS-specific interleaver π_k as well as CS-specific interleaver set $\{\pi_{n,k}\}_{k=1}^{K}$ for differentiating the various sources. The 'distributed' nature of the random interleaver emphasises the fact that in contrast to the centrally controlled random interleaver assignment regime of classic cellular IDMA-style systems [140], here the interleavers are assigned autonomously. We adopt a matrix representation for the kth AS's IR-STC $\mathbf{C}_k$. In our scheme, the information source signal $\mathbf{b}_k$ of length P is firstly repetition coded by C_1, randomly interleaved by AS-specific interleaver and further repetition coded by C_2, yielding the sequence $\mathbf{x}_k^{(1)}$ of length $M = P/r$, obeying $\mathbf{x}_k = \mathbf{R}_2[\pi_k(\mathbf{R}_1)]\mathbf{b}_k$, where:

$$\begin{aligned} \mathbf{R}_1 &= \text{diag}\left[\mathbf{1}_{1/r_1}, \ldots, \mathbf{1}_{1/r_1}\right]_{(M_1 \times P)} \\ \mathbf{R}_2 &= \text{diag}\left[\mathbf{1}_{1/r_2}, \ldots, \mathbf{1}_{1/r_2}\right]_{(M \times M_1)} \end{aligned}$$

are the two repetition codes' matrices and $M_1 = P/r_1$ and the all-one vector is denoted as $\mathbf{1}_N = [1, \ldots, 1]^T$. The random AS-specific interleaver π_k of Fig. 4.3 permutes the corresponding rows in the matrix $\mathbf{R}_1$, yielding the matrix $\hat{\mathbf{R}}_1 = \pi_k(\mathbf{R}_1)$. Denote that $\mathbf{G} = \mathbf{R}_2\hat{\mathbf{R}}_1$, the IR-STC matrix $\mathbf{C}_k$ of size $(M/N \times N)$ is constructed according to $\mathbf{C}_k = \hat{\mathbf{\Theta}}_k\mathbf{B}_k$, where we have

$$\begin{aligned} \hat{\mathbf{\Theta}}_k &= \left[\pi_{1,k}\left(\mathbf{\Theta}_k^1\right), \ldots, \pi_{N,k}\left(\mathbf{\Theta}_k^N\right)\right]_{(M/N \times PN)} \\ \mathbf{\Theta}_k &= \left[\mathbf{\Theta}_k^1, \ldots, \mathbf{\Theta}_k^N\right]_{(M/N \times PN)} \\ \mathbf{\Theta}_k^n &= \left[\mathbf{G}_{(n,:)}, \ldots, \mathbf{G}_{(M-N+n,:)}\right]^{\mathrm{T}}_{(M/N \times P)} \\ \mathbf{B}_k &= \text{diag}\left[\mathbf{b}_k, \ldots, \mathbf{b}_k\right]_{(PN \times N)}. \end{aligned}$$

where the subscript $\mathbf{G}_{(k,:)}$ stands for the kth row of matrix $\mathbf{G}$.

Our IR-STC employs the random distributed AS-specific interleaver π_k as well as CS-specific interleaver set $\{\pi_{n,k}\}_{k=1}^{K}$, which results in a distinct IR-STC matrix $\hat{\mathbf{\Theta}}_k$. Then these K distinct IR-STCs are superimposed and transmitted simultaneously for the sake of supporting a high throughput. This random construction is different from the one proposed in [142], where each CS transmits a random linear combination of the columns of an existing OSTBC.

Setting	1	2
$\mathbf{G}$	$\mathbf{G} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}^{\mathrm{T}}$	$\mathbf{G} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}^{\mathrm{T}}$
$\mathbf{\Theta}_k^1$	$\mathbf{\Theta}_k^1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}}$	$\mathbf{\Theta}_k^1 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}^{\mathrm{T}}$
$\mathbf{\Theta}_k^2$	$\mathbf{\Theta}_k^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}^{\mathrm{T}}$	$\mathbf{\Theta}_k^2 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}^{\mathrm{T}}$
$\hat{\mathbf{\Theta}}_k$	$\hat{\mathbf{\Theta}}_k = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$	$\hat{\mathbf{\Theta}}_k = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$
$\mathbf{C}_k$	$\mathbf{C}_k = \begin{bmatrix} b_k(1) & b_k(1) & b_k(1) & b_k(1) \\ b_k(2) & b_k(2) & b_k(2) & b_k(2) \end{bmatrix}^{\mathrm{T}}$	$\mathbf{C}_k = \begin{bmatrix} b_k(1) & b_k(2) & b_k(1) & b_k(2) \\ b_k(2) & b_k(1) & b_k(2) & b_k(1) \end{bmatrix}^{\mathrm{T}}$

Table 4.1: Matrix representation of the difference between Setting 1 and Setting 2.

4.1.3.2 Investigation of IR-STC

We now illustrate the IR-STC employed in MSC having three typical settings of r_1 and r_2 given a fixed total code-rate $r = 1/4$ compared to the Alamouti G_2 OSTBC [126] benchmarker represented by the solid line seen in Fig 4.4, where we have $P = 1024$, $K = N = 2$. The three code-rate settings are 1) $r_1 = 1/4$ and $r_2 = 1$; 2) $r_1 = 1$ and $r_2 = 1/4$; 3) $r_1 = 1/2$ and $r_2 = 1/2$.

As seen in Fig 4.4, both Setting 2 and 3 have a similar BER as the G_2 OSTBC, implying that these two IR-STCs are capable of achieving full transmit diversity, as the G_2 OSTBC does. By contrast, Setting 1 has the worst performance.

The rationale of these findings is exemplified below. Assuming $P = 2$, the transmitted information source signal is expressed as $\mathbf{b}_k = [b_k(1), b_k(2)]^{\mathrm{T}}$ and the repetition matrices of the $r_1 = 1/4$-rate $\mathbf{R}_1$ in Setting 1 and of the $r_2 = 1/4$-rate $\mathbf{R}_2$ in Setting 2 are represented as:

$$\mathbf{R}_1 = \mathbf{R}_2 = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}^{\mathrm{T}}.$$

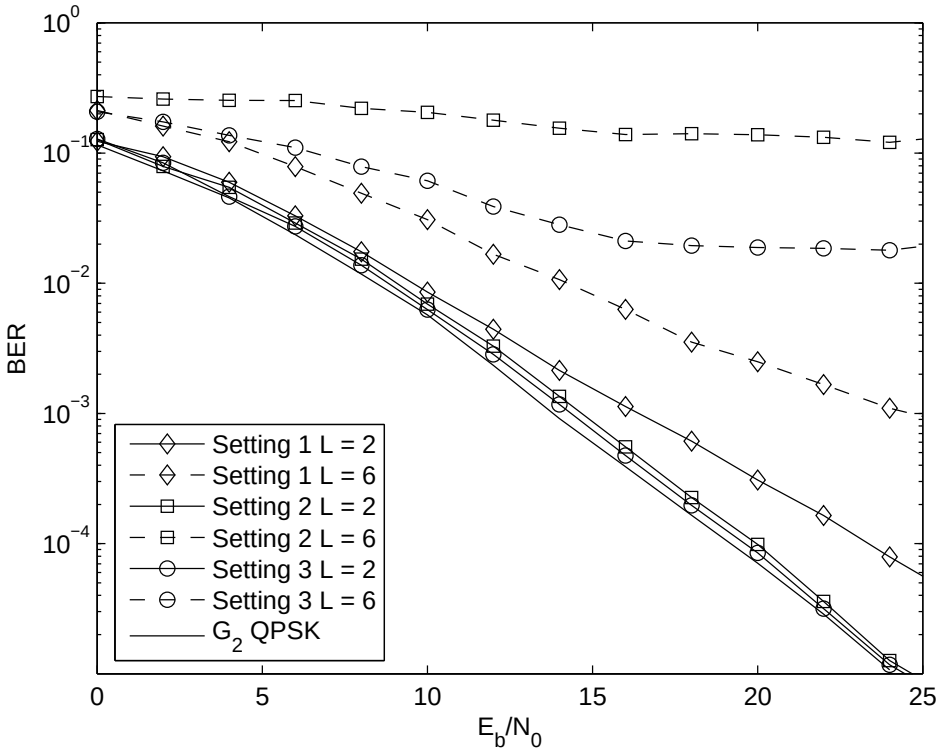


Figure 4.4: Comparison of the three typical code-rate settings of IR-STC to that of G_2 OSTBC in cooperative communications.

In this example, the AS-specific interleaver pattern employed in Setting 1 was $\pi_k = [7,5,1,3,8,2,6,4]$ and the CS-specific interleavers employed in both Settings were $\pi_{1,k} = [3,1,4,2]$ and $\pi_{2,k} = [4,2,1,3]$.

In Setting 1, we have $\mathbf{G} = \pi_k(\mathbf{R}_1)$. Then each row in $\mathbf{R}_1$ is randomly mapped onto $\mathbf{G}_{(i,:)}, i \in [1,M]$ by the AS-specific interleaver π_k and dispersed over $\mathbf{\Theta}_k$. In this scenario, all $1/r_1 = 4$ rows corresponding to the same diagonal element of $\mathbf{R}_1$ may be mapped onto the same $\mathbf{\Theta}_k^n$ thus cannot be spread across all $N = 2$ CSs, which results in a reduced effective number of CSs, i.e. in a reduced spatial diversity order. In Table 4.1, the resultant $b_k(1)$ or $b_k(2)$ is assisted only by a single CS, regardless of the CS-specific interleavers are used.

By contrast, when only C_2 is employed as in Setting 2, we have $\mathbf{G} = \mathbf{R}_2$, and then the ith row in $\mathbf{R}_2$ is directly mapped onto $\mathbf{G}_{(i,:)}$ and dispersed over $\mathbf{\Theta}_k$. In this scenario, the $1/r_2 = 4$ rows corresponding to the same diagonal element of $\mathbf{R}_2$ are mapped across all the $N = 2$ CSs, although a different CS-specific interleaver is used by each CS. As seen in Table 4.1, both $b_k(1)$ and $b_k(2)$ are assisted by $N = 2$ CSs. Hence, Setting 2 benefits from a full diversity.

However, when a sufficiently high number of layers, such as $L = 6$ is employed, Setting 1 was seen to be best in Fig 4.4, while Setting 2 is incapable of supporting the high throughput of Setting 1, where $\eta_{IR-STC} = 3$. This is because Setting 1 has a factor N higher effective interleaver length compared to Setting 2 and hence it is capable of cancelling the interference imposed by a large number of layers, when an iterative receiver is employed. This implies that Setting 1 is a

high-throughput *multiplexing-oriented* configuration, while Setting 2 is a low-throughput *diversity-oriented* configuration.

4.1.3.3 Distributed Interleaver Design

The generation of distributed random interleavers used in our IR-STC aided MSC should be carried out in an efficient manner, while at the same time maintaining their random nature. We again employ the so-called SE interleaver as introduced in Chapter 3. Without loss of generality, we discuss the generation of AS-specific interleavers. The SE interleavers are constructed from three hierarchical layers, namely from a system-specific base interleaver, a AS-specific base interleaver and a so-called constituent interleaver set. These interleavers are then subjected to a position sorting operation, all of which are detailed below.

The *system-specific base interleaver* π^b is a randomly generated interleaver of length-Q. Additionally, each AS has a distinct *AS-specific base interleaver* $\pi^b_k, k \in [1, K]$ having the same length-Q as the system-specific base interleaver π^b. The $(k+1)$st AS-specific base interleaver used in the $(k+1)$st TDD slot is an interleaved version of the kth AS-specific base interleaver used in the TDD slot k, which was rearranged by the system-specific base interleaver π^b, as follows: $\pi^b_{k+1} = \pi^b(\pi^b_k)$ and $\pi^b_1 = \pi^b$.

The *constituent interleaver set* of AS k is represented by U number/level of length-Q interleavers, which is formulated as $\pi^c_k = \{\pi^1, \pi^2, \ldots, \pi^U\}$. Each element $\pi^u \in \pi^c_k, u \in [1, U]$ of the constituent interleaver set is a distinct length-Q interleaver, having the same length as the system-specific base interleaver π^b. The $(u+1)$st constituent interleaver is an interleaved version of the uth constituent interleaver, which was rearranged by the AS-specific base interleaver π^b_k, according to $\pi^{u+1} = \pi^b_k(\pi^u)$ and $\pi^1 = \pi^b_k$.

Finally, the U number of length-Q interleavers are concatenated to form a unique length-UQ interleaver. This is carried out by the constituent interleaver set *position sorting operation*, as defined by the position mapping function f, which maps the index $q^u \in [1, Q]$ within all the U number of length-Q constituent interleavers $\pi^u \in \pi^c_k, u \in [1, U]$ into a single AS-specific interleaver $\pi_k = f(\pi^c_k)$. From a different perspective, this implies unambiguously mapping the UQ number of input bit positions to the interleaved positions $q \in [1, UQ]$. More specifically, the index $q^u \in [1, Q]$ within any of the U number of length-Q constituent interleavers $\pi^u, u \in [1, U]$ is mapped to $q = (q^u - 1)U + u$.

4.1.4 Performance Evaluation

In this subsection, we investigate the performance of our IR-STC aided MSC employing different relaying techniques and stipulating different assumptions concerning $h_{s,s}$ by varying the Nakagami-m fading parameters. Both $h_{s,s}$ and $h_{s,d}$ are assumed to be quasi-static, i.e. constant in every 1024

Simulation Parameter				
Info. bit length	1024			
Modulation Type	BPSK			
Code-rate	r = 1/8			
Cluster Info.	N = 4, K = 4			
Receiver	IC, 30 iterations			
Effective Throughput	G_4	IR-STC	Δ_1	Δ_2
$\eta_{IR-STC} = 2$	$\mathcal{M} = 16$	$L = 4$	6	-
$\eta_{IR-STC} = 2.5$	$\mathcal{M} = 32$	$L = 5$	5	-
$\eta_{IR-STC} = 3$	$\mathcal{M} = 64$	$L = 6$	4	9
$\eta_{IR-STC} = 3.5$	$\mathcal{M} = 128$	$L = 7$	2	7
$\eta_{IR-STC} = 4$	$\mathcal{M} = 256$	$L = 8$	-1	4
Relaying Techniques	AF	SDF	DF	DDF
Cooperating Source	$\lvert h_{s,s}\rvert$	$h_{s,s}$	$h_{s,s}$	-
Destination	$h_{s,s}$	$\lvert h_{s,s}\rvert$	-	-

Table 4.2: Summary of the simulation parameters and the power gain Δ in dB of IR-STC compared to G_4 OSTBC in MSC, where Δ_1, Δ_2 denotes the power gain corresponding to 16-QAM and 64-QAM in G_4 OSTBC, respectively.

information symbol block, but they are independently faded between the consecutive blocks. The value of $h_{s,d}$ is assumed to be perfectly known at the destination, while the knowledge of $h_{s,s}$ is either explicitly required or dispensed with at both the CSs and the destination, depending on the specific relaying technique employed. In all of our simulations, perfect relaying implies that all K ASs' information bits are perfectly known at all N CSs. Our simulation parameters are listed in Table. 4.2.

4.1.4.1 Benefits of IR-STC Aided MSC

We first assume perfect relaying and outline the benefits of our IR-STC aided MSC compared to the traditional G_4 OSTBC.

Power efficiency Fig. 4.5 demonstrates that the achievable BER performance of our IR-STC aided MSC scheme is superior to G_4 OSTBC aided MSC. In this experiment, Setting 1 was used for $L > 5$ and Setting 3 was used for $L < 6$. It can be seen in Fig. 4.5 that the maximum number of layers supported was $L = 7$, which is equivalent to a G_4 OSTBC scheme using a large and hence error-sensitive 128-QAM constellation, and yet, our scheme required a lower power than the 4 bit/symbol G_4 OSTBC aided 16-QAM scheme, as observed at $BER \leq 10^{-5}$.

High throughput In addition to uniform phase rotation, non-uniform power allocation is also considered. In this paper, no attempts were made to formally optimise the power allocation scheme. Instead, the following simple non-uniform power allocation strategy [86] was adopted, where this

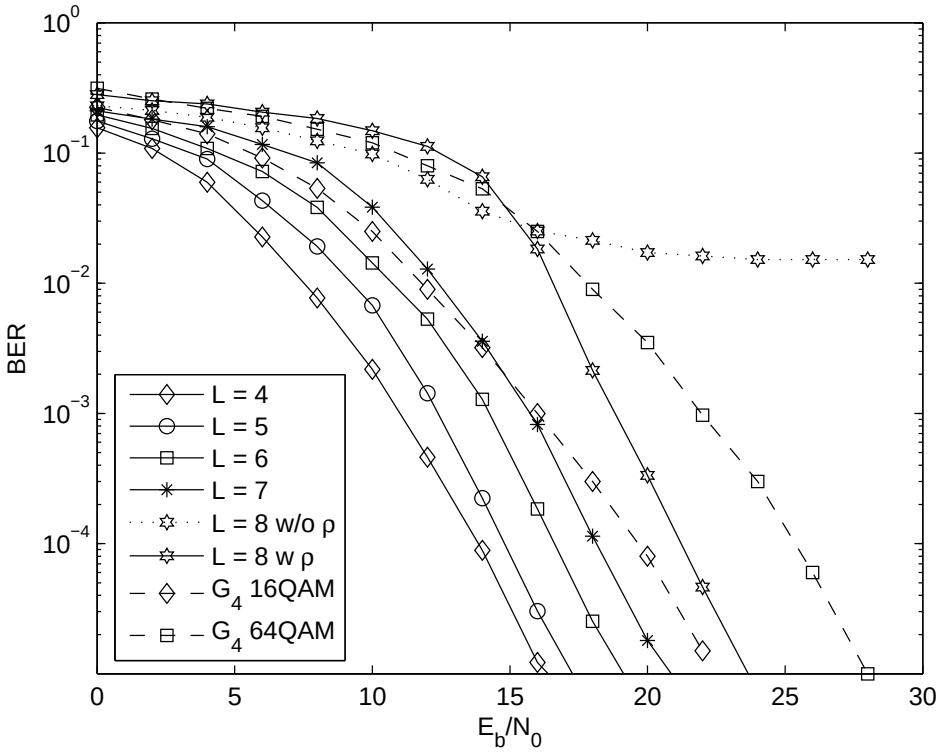


Figure 4.5: Performance of G_4 OSTBC and IR-STC over quasi-static fading in cooperative communications, where $L = 4,5,6,7$ layers corresponding upto $\eta_{IR-STC} = 3.5$ transmission were supported. Additionally, with the aid of non-uniform power allocation, $L = 8$ layers can be supported.

power allocation strategy under consideration was derived assuming capacity-achieving codes. Consider allocating the length-L layer-specific amplitudes, which obeys:

$$\rho_{l+1}^2 = \rho_l^2/\beta, \tag{4.10}$$

while ensuring that $\sum_{l=1}^{L} \rho_l^2 = P_n$, where we refer to $\beta \geq 1$ as the scaling factor and P_n is the maximum total power of the nth CS, which is assumed to be equal for all N CSs.

Returning briefly to Fig. 4.5, it also shows the achievable BER performance of IR-STC aided MSC, when using non-uniform power allocation of each CS. Upon investigating the most appropriate values of β experimentally, $\beta = 1.2$ was found to be adequate and the number of layers for which an adequate BER performance was attainable was found to be as high as $L = 8$, corresponding to a 256-QAM modulated 8 bits/symbol G_4 OSTBC scheme, while requiring a lower power than the 6 bits/symbol G_4 OSTBC aided 64-QAM scheme.

The achievable power gain of IR-STC used in MSC was summarised in Table 4.2, where Δ was the E_b/N_0 gain of IR-STC at $BER = 10^{-4}$ over conventional G_4 OSTBC scheme having identical-throughput, i.e. we had $\Delta = (E_b/N_0)_{OSTBC} - (E_b/N_0)_{IR-STC}$.

Flexibility The design flexibility of IR-STC allows the employment of an arbitrary number of sources as seen in Fig. 4.6. This implies that IR-STC based MSC can be used in diverse cooperative scenarios. More explicitly, when $r = 1/8$ IR-STC was used, the system was capable of supporting

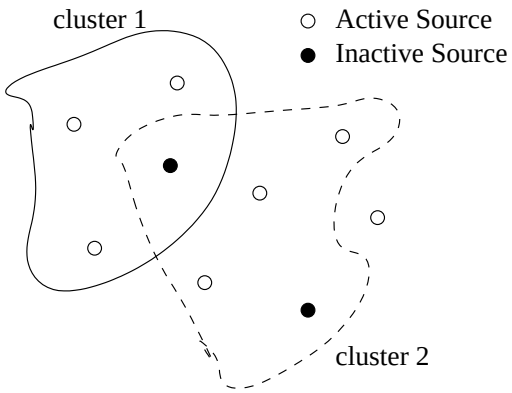


Figure 4.6: Flexible cluster formed by arbitrary number of sources.

$K = 2, 3, 4, 5$ ASs without designing different matrices when the traditional OSTBC code was employed.

Suffice to say, however that in most practical scenarios having a diversity order of more than five attains a near-AWGN BER performance and hence there is limited benefit in further increasing the diversity order, i.e. the number of CSs. This flexibility is beneficial in terms of forming a flexible cluster of CSs, allowing sources to freely join or disjoin the cluster of cooperation.

4.1.4.2 Various Relaying Techniques

Let us now investigate various relaying techniques employed in our IR-STC aided MSC, where the multiplexing-oriented Setting 1 was used.

Fig. 4.7 characterises three different relaying techniques, namely AF, DF and DDF employed in our IR-STC aided MSC scheme, when the inter-source channel SNR was $\gamma_{s,s} = 20dB$. As expected, the higher the value of the Nakagami parameter m, the less hostile the channel fading encountered, which results in an improved BER performance for all the three relaying techniques. For all three m values considered, DF leads to the best BER performance, while the performance attained by DDF is better than that of AF, except for $m = 1$. The worst performance of AF relaying is mainly a consequence of its noise enhancement. To elaborate a little further, the inferior performance of DDF compared to that of DF is a direct consequence of its doubled noise variance, when non-coherent detection was employed. For $m = 1$, the effect of noise enhancement imposed by AF relaying is less severe than that of the doubled noise variance of non-coherent detection encountered by DDF, as evidenced by the results of Fig. 4.7.

Similar conclusions can also be inferred from Fig. 4.8, which characterises the three different relaying techniques, when the inter-source channel SNR was $\gamma_{s,s} = 30dB$. The inferior performance imposed by the doubled noise variance of non-coherent detection in DDF compared to DF was less obvious for $m > 1$, since at $\gamma_{s,s} = 30dB$ both DF and DDF become capable of near-error-

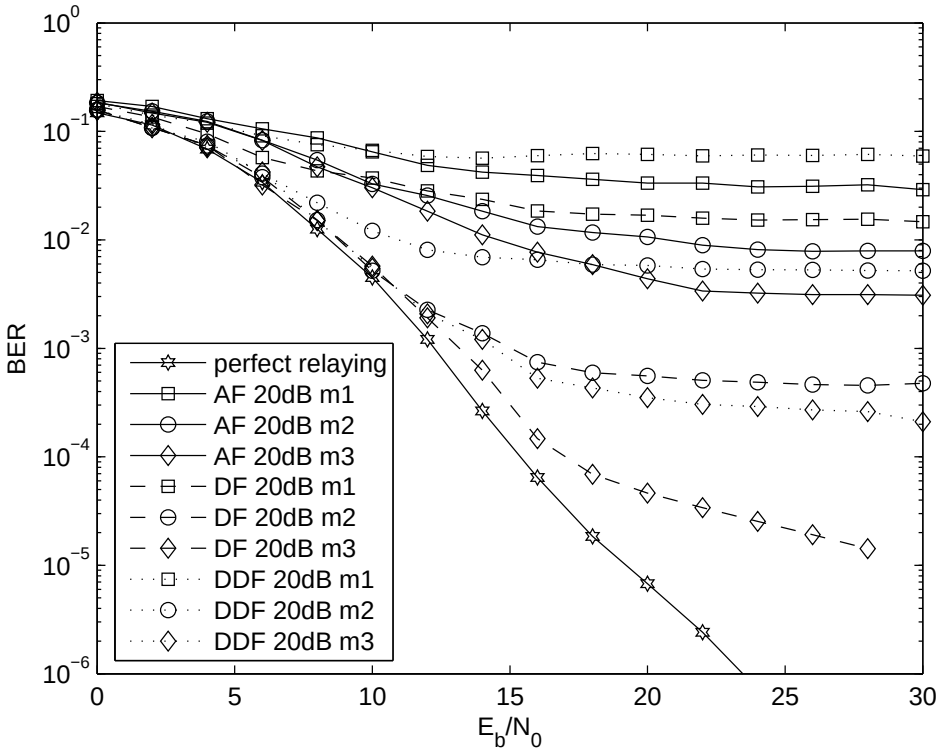


Figure 4.7: Comparison of AF, DF and DDF relaying techniques employed in the 4-layer IR-STC aided MSC scheme, when the inter-source channel SNR was $\gamma_{s,s} = 20dB$.

free detection, while detection errors persist for $m = 1$ in both techniques. Likewise, it can be seen in Fig. 4.8 that the discrepancy between AF and DF becomes small because of the less pronounced noise enhancement, when we have benign fading. However, the difference of AF and DF remains relatively high in Fig. 4.9, which characterises all the three relaying techniques at an inter-source channel SNR $\gamma_{s,s} = 30dB$, when supporting $L = 7$ layers per CS. In this case, the effect of noise enhancement is further aggravated by superimposing more layers for the sake of achieving a high throughput, although each noise contribution itself may be relatively modest.

Fig. 4.10 compares two non-regenerative relaying techniques, namely AF and SDF employed in our IR-STC aided MSC scheme, when the inter-source channel SNR was $\gamma_{s,s} = 30dB$. It can be seen that there is only an insignificant difference between these two techniques. However, when the AF technique is employed, the knowledge of $h_{s,s}$ is required at the destination. By contrast, only the knowledge of $|h_{s,s}|$ is required at the destination, when SDF is employed. It can be seen in Fig. 4.10 that when a carrier phase error of $\phi = \pi/16$ is imposed on $h_{s,s}$ at the destination, a significantly reduced BER performance is observed. This implies that SDF is a better relaying technique compared to AF, provided that the CSs are capable of acquiring accurate knowledge of $h_{s,s}$. However, having the knowledge of $|h_{s,s}|$ at the CS is sufficient for ensuring reliable operation of the AF technique. We note furthermore that in uncoded or repetition coded systems, SDF essentially becomes an AF technique, which is inferior to the DF technique. When a serial concatenated outer channel code is employed, SDF becomes capable of enabling soft channel decoding and the corresponding extrinsic information $\mathcal{L}$ becomes more reliable. This results in a higher mutual infor-

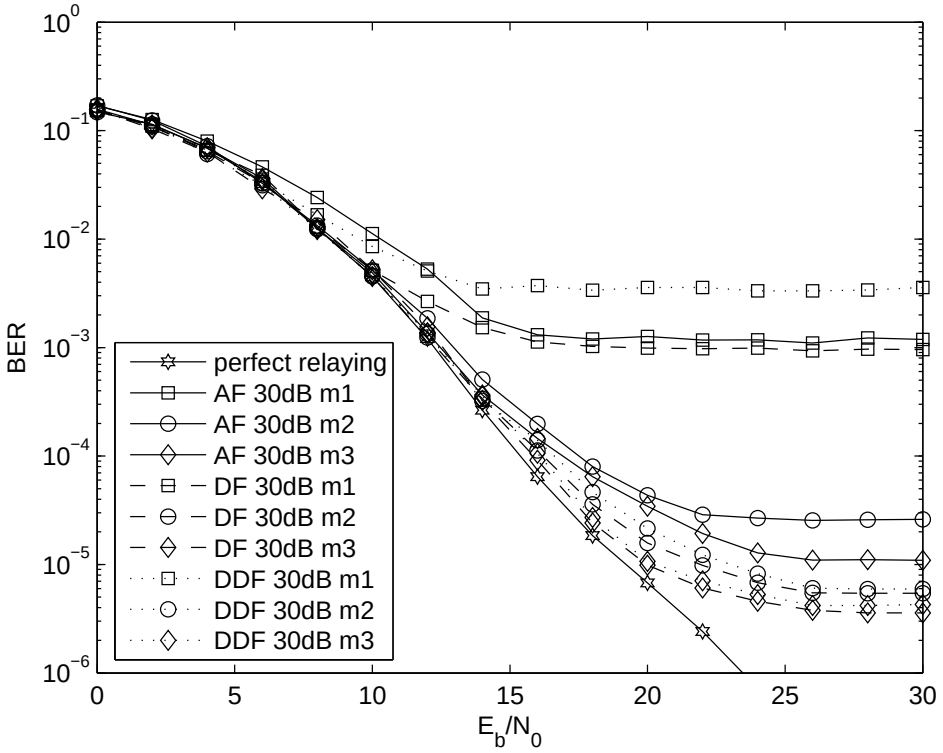


Figure 4.8: Comparison of AF, DF and DDF relaying techniques employed in the 4-layer IR-STC aided MSC scheme, when the inter-source channel SNR was $\gamma_{s,s} = 30dB$.

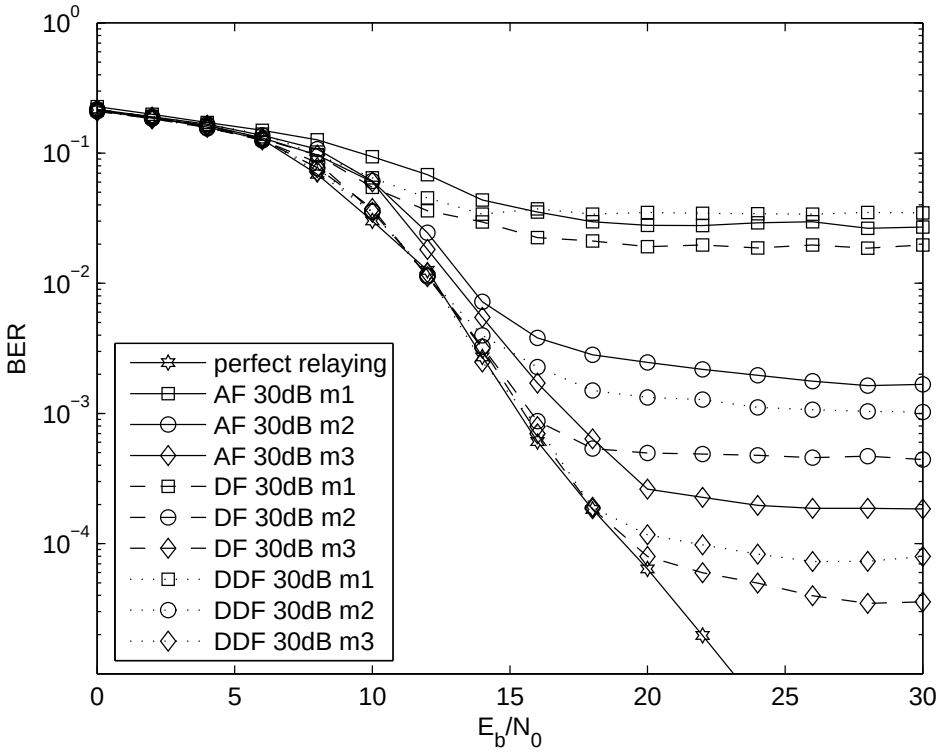


Figure 4.9: Comparison of AF, DF and DDF relaying techniques employed in the 7-layer IR-STC aided MSC scheme, when the inter-source channel SNR was $\gamma_{s,s} = 30dB$.

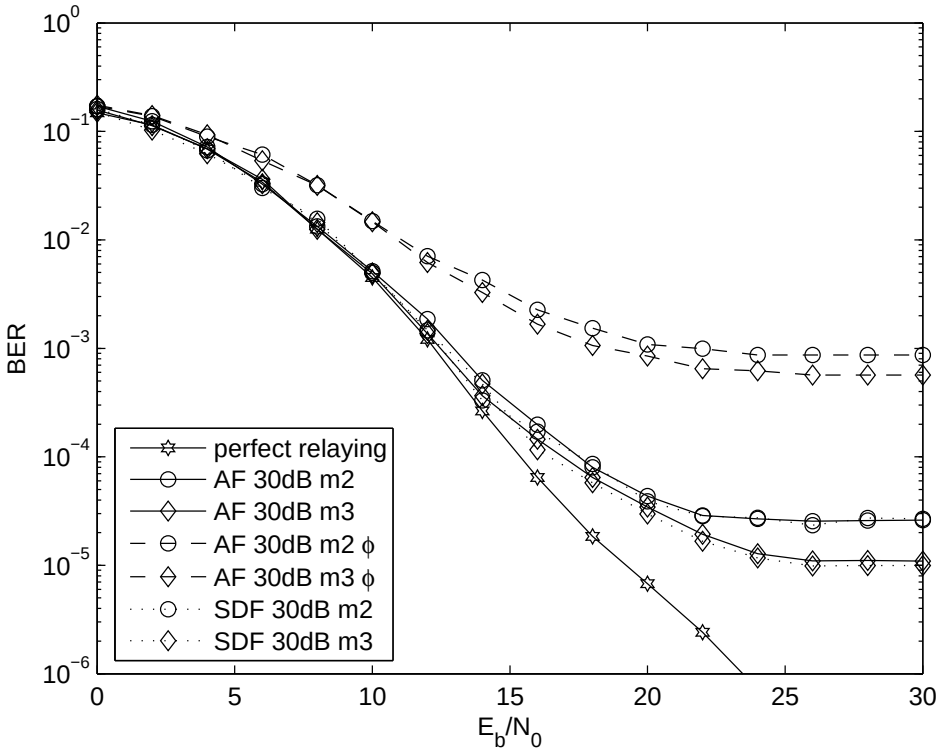


Figure 4.10: Comparison of AF, SDF and AF subject to carrier phase error at the destination employed in the 4-layer IR-STC aided MSC scheme, when the inter-source channel SNR was $\gamma_{s,s} = 30dB$.

mation $I(x_k^{(1)})$, which is equivalent to having a reduced noise variance. Therefore, the achievable performance is expected to be better than that of DF.

All of our previous investigations were based on having a fixed inter-source channel SNR. To expound further, Fig 4.11 shows three different relaying techniques, namely AF, DF and DDF employed in our IR-STC aided MSC, when the channel $h_{s,s}$ experiences different Nakagami-m fading and assuming a consistently higher SNR value than that associated with $h_{s,d}$, i.e. we have $\Delta = \gamma_{s,s}/\gamma_{s,d} > 0dB$. It can be seen in Fig. 4.11 that for $m = 2$ DF performs consistently better than the other two techniques and approaches the perfect relaying performance, namely that of the system, which regenerates the source information without decision errors. Surprisingly, DDF also performs consistently better than the AF technique. However, when severe Rayleigh fading is encountered, i.e. we have $m = 1$, AF has the best performance at high SNR, where the effect of noise enhancement is negligible. By contrast, the performance of both DF and DDF are unacceptable, owing to the effects of Rayleigh fading. Therefore, ideally the specific relaying technique used should be determined according to the specific Nakagami-m fading values encountered. This suggests switching amongst the different relaying modes.

Remarks: We may now conclude that when the SNR of the $h_{s,s}$ channel is better than that of $h_{s,d}$, DF is the best relaying strategy in the presence of benign fading. When a sufficiently high-SNR benign faded $h_{s,s}$-channel is experienced, close-to-perfect relaying performance is attainable. The

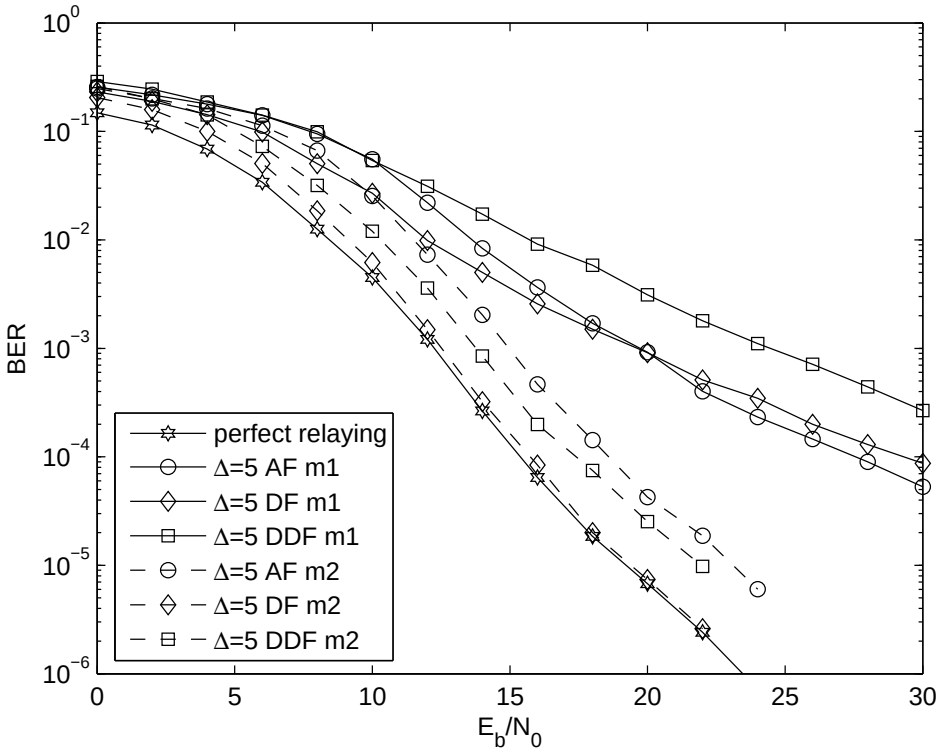


Figure 4.11: Comparison of AF, DF and DDF relaying techniques employed in the 4-layer IR-STC aided MSC scheme having inter-source channel SNR to source-destination channel SNR ratio $\Delta = 5$, when $h_{s,s}$ is subjected to Nakagami fading associated with $m = 1$ and $m = 2$.

AF technique is only preferable at high SNRs when severe fading is encountered. DDF performs consistently worse than DF due to the doubled noise variance of non-coherent detection. Surprisingly, when the fading is benign, non-coherent DDF without the cost of estimating all inter-source channel knowledges outperforms the coherent detected AF technique. Therefore, a pre-selection of the CSs benefiting from a high-SNR $h_{s,s}$ channel-which typically also have high Nakagami-m values-is important in MSC.

4.2 Cooperative Energy Efficient Coding Schemes

In contrast to the uncoded system discussed in the previous section, we now aim for improving the energy efficiency of our proposed MSC framework with the aid of two specifically designed coding schemes, namely SPC and a PANC scheme.

4.2.1 Introduction

Different from classic time-multiplexing, in the SPC scheme the multiple sources' information is code-multiplexed in order to generate the superimposed and appropriately rotated composite signal, which results in a high throughput and low crest-factor. Thus we will introduce an outer channel-coded SPC-aided MSC arrangement, which will be used as the benchmarker of the proposed PANC

scheme.

NC may be viewed as a technique of conveying a linear combination of multiple information streams, rather than using conventional routing or relaying for delivering these information flows individually with the aid of classic resource allocation, such as time-multiplexing or code-multiplexing. The philosophy of network coding was proposed by Yeung [42] for the sake of enhancing the wired channel's capacity. The potential diversity gain facilitated by network coding used in wireless networks was then illustrated for example in [143].

Apart from the original network-layer applications, it has recently been recognised that the physical-layer of wireless networks also benefits from NC. The concept of joint channel coding and network coding was proposed in the context of the classic two-way relay channel [144] and the multiple access relay channel [145], where the concept of distributed channel codes was generalised and the redundancy inherent in the network code was exploited in order to support channel decoding. The employment of network coding was proposed for a two-user cooperation-aided scenario in [146, 147], where a promising performance was observed. However, its extension to MSC is not straightforward, since the unique recovery of the ith information flow s_i from an aggregate of N module 2 superimposed information flows created as $s_1 \oplus s_2, \ldots, s_{N-1} \oplus s_N$ is generally impossible [148]. We therefore generalise the concept of network coding and propose the so-called PANC scheme.

In a nutshell, the novel contribution of this section is that *we propose both a SPC scheme and a PANC scheme, which are capable of performing close to the best possible outage probability bound in the context of MSC. Our numerical results show that compared to SPC, the novel PANC arrangement exhibits a reduced complexity (ι) at the cost of a slight performance degradation (P_e^{bl}), while maintaining the same throughput (η) and delay (τ).*

The rest of this section is organised as follows. In Subsection 4.2.2, we describe our MSC model and propose the SPC and PANC schemes considered. Furthermore, in Subsection 4.2.3, the iterative receiver structure and the soft PANC decoding algorithm advocated are also discussed. In Subsection 4.2.4, the outage probability bound of MSC is analysed and the numerical results characterising both schemes are provided. Finally, we conclude our discourse in Section 4.3.

4.2.2 Cooperative Code Design

We now briefly recall the cooperation scenario introduced in the previous section. Consider a cluster of single-antenna sources cooperatively communicating with a destination employing a single receive antenna, which jointly result in a VMISO system, as seen in Fig 4.1. In this VMISO cluster, we assume having a total of N CS, K AS and $(N-K)$ RS. Our MSC scheme entails two phases. In *Phase-I cooperation*, the source information emanating from all K ASs is broadcast to all N CSs in a TDD manner under the assumption of having perfect synchronisation. By contrast, *Phase-II cooperation* is defined as the joint transmission of a combined signal generated by the concerted

action of all the N CSs, which will be elaborated on in more detail in the next subsection. Therefore, each CS simultaneously transmits multiple ASs' information, resulting in a high throughput. This implies that each AS is served simultaneously by multiple CSs and hence the entire set of ASs benefits from a high diversity gain.

In this section, we focus our attention on developing coding schemes for MSC in *Phase-II cooperation*, when the so-called *decode-forward* relaying technique is employed at each of the N CSs.

4.2.2.1 Superposition Coding

Following *Phase-I cooperation*, the nth of the N CSs retrieves all the K ASs' information $\hat{\mathbf{s}}_{k,n}^{(1)}, k \in [1, K]$ and the transmitted codeword is constructed as follows. Firstly, the nth CS forms K parallel codewords

$$\mathbf{c}_{n,k} = \pi_k \left[f(\hat{\mathbf{s}}_{k,n}^{(1)}) \right] \qquad k \in [1, K], \tag{4.11}$$

where π_k is referred to as the AS-specific interleaver and $f(\cdot)$ represents the outer channel coding function, which is assumed to be the same for all ASs. These AS-specific outer codewords are then punctured according to

$$\tilde{c}_{n,k}(i) = c_{n,k}[N(i-1)+n] \qquad i \in [1, N_c/N], \tag{4.12}$$

where N_c denotes the codeword lengrh. This is followed by a P/S conversion in order to create a single codeword

$$\mathbf{c}_n = [\tilde{c}_{n,1}(1), \ldots, \tilde{c}_{n,K}(1), \ldots, \tilde{c}_{n,1}(N_c/N), \ldots, \tilde{c}_{n,K}(N_c/N)]. \tag{4.13}$$

Finally, the composite codeword transmitted from the nth CS is BPSK modulated and linearly superimposed:

$$x_n^{(2)}(i) = \sum_{l=1}^{L_n} \rho_{n,l} e^{j\theta_{n,l}} x_{n,l}^{(2)}(i), \tag{4.14}$$

$$x_{n,l}^{(2)}(i) = 2\tilde{c}_n[L_n(i-1)+l] - 1, \tag{4.15}$$

where $i \in [1, N_c K/NL_n]$ and L_n is referred to as the number of *layers* contributed by the nth CS's superposition, while $\rho_{n,l}$ and $\theta_{n,l} \in [0, \pi)$ denotes the layer-specific amplitude and phase rotation respectively. In this treatise, we assume $L_n = L, \rho_{n,l} = \rho_l, \theta_{n,l} = \theta_l, \forall n \in [1, N]$.

Remarks: The rationale of allocating a different amplitude ρ_l and hence power to each of the L layers is philosophically similar to that of the multilevel coding concept of [59], where we create a number of different protection levels and detect them by gleaning extrinsic information from the previously decoded levels using multistage decoding. Imposing the associated phase rotation $\theta_{n,l}$ has two benefits, namely that of reducing the PAPR of the transmitted signal $x_n^{(2)}$ and making $x_n^{(2)}$ having L layers more distinguishable for the detector. In this section, equal amplitude allocation and uniform phase rotation are employed.

4.2.2.2 Physical-layer Algebraic Network Coding

We generalise the concept of NC as a coding function $f(\cdot)$, which jointly encodes all the incoming multiple information flows. With the aid of this generalisation, the original NC operation $\oplus$ of K linearly coded information flows $\mathbf{s}_i\mathbf{G}_i, i \in [1,K]$ becomes equivalent to encoding the vectors $\mathbf{s} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_K]$ using a nested GM, which can be written as:

$$\begin{aligned} \mathbf{c} &= \mathbf{s}_1\mathbf{G}_1 \oplus \mathbf{s}_2\mathbf{G}_2, \ldots, \oplus \mathbf{s}_K\mathbf{G}_K & (4.16)\\ &= [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_K][\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_K]^{\mathrm{T}}. & (4.17) \end{aligned}$$

We now proceed to describe the construction of codewords for our MSC taking this novel PANC principle into account. After retrieving all the K ASs' information denoted by $\mathbf{s} = [\hat{\mathbf{s}}_{1,n}^{(1)}, \ldots, \hat{\mathbf{s}}_{K,n}^{(1)}]$ and having a length of KN_i, the nth CS generates a total of κ number of versions of the differently interleaved information flow and the resultant codeword $\mathbf{c}_n$ of length $N_c = \kappa K N_i$ is given by:

$$\begin{aligned} \mathbf{c}_n &= \pi_1(\mathbf{s})\mathbf{G}_1 \oplus \pi_2(\mathbf{s})\mathbf{G}_2, \ldots, \oplus \pi_\kappa(\mathbf{s})\mathbf{G}_\kappa & (4.18)\\ &= [\pi_1(\mathbf{s}), \pi_2(\mathbf{s}), \ldots, \pi_\kappa(\mathbf{s})]\mathbf{G} & (4.19)\\ \mathbf{G} &= [\mathbf{G}_1, \mathbf{G}_2, \ldots, \mathbf{G}_\kappa]^{\mathrm{T}}, & (4.20) \end{aligned}$$

where we have $\pi_1 = \pi, \pi_i = \pi(\pi_{i-1})$ and π represents a randomly generated interleaver in this section. Although in principle an arbitrary $\mathbf{G}$ may be applicable, we adopt a simple unity-rate ACcumulate Code (ACC) [72], having a GM and Parity Check Matrix (PCM) represented as

$$\mathbf{G}_{acc} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ & 1 & & 1 \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix} \quad \mathbf{H}_{acc} = \begin{bmatrix} 1 & & & & \\ 1 & 1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & 1 & 1 \end{bmatrix}. \tag{4.21}$$

Apart from the non-systematic PANC GM of Eq. 4.21, we may also generate a systematic PANC by designing the GM as

$$\mathbf{G} = \begin{bmatrix} \mathbf{I}_{\kappa^s} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_{acc} \end{bmatrix},$$

where κ^s number of differently interleaved versions of the original information flows are created, corresponding to $[\pi_1(\mathbf{s}), \pi_2(\mathbf{s}), \ldots, \pi_{\kappa^s}(\mathbf{s})]$. Finally, the nth CS transmits a BPSK modulated punctured codeword according to

$$x_n^{(2)}(i) = 2c_n[N(i-1)+n] - 1 \qquad i \in [1, N_c/N]. \tag{4.22}$$

Remarks: The concept of NC and SPC may have some intrinsic links. In fact, the authors of [147] considered the NC concept as a SPC scheme defined over the Galois Field 2 (i.e. the operation + in SPC is replaced by $\oplus$ in terms of NC), while the authors of [148] considered the SPC

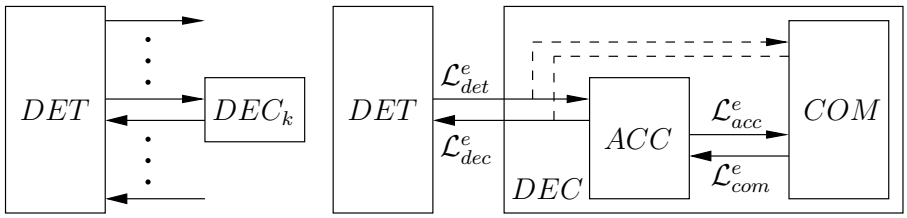


Figure 4.12: Iterative receiver architecture of both the SPC scheme employing single-stream decoding (left) and PANC scheme employing multiple-streams decoding (right), where COM represents the soft combiner of the multiple streams. Systematic and non-systematic segment are denoted as broken and solid line, respectively.

concept as a NC scheme defined over the complex field. Therefore, the PANC proposed above may be considered as a conventional NC scheme exhibiting a channel coding gain, which is a benefit of the mutual dependencies introduced by the linear module 2 addition of multiple streams. What is worth noting is that the algebraic property in our PANC emphasises the algebraic channel-coding aspects within our design.

4.2.3 Iterative Detection and Decoding

4.2.3.1 Receiver Structure

The destination receives N CSs' transmitted signals $x_n^{(2)}, n \in [1, N]$, which experienced independently faded channels h_n, yielding the received signal:

$$y_{PANC} = \sum_{n=1}^{N} h_n x_n^{(2)} + n, \tag{4.23}$$

$$y_{SPC} = \sum_{n=1}^{N} h_n \sum_{l=1}^{L} \rho_l e^{j\theta_l} x_{n,l}^{(2)} + n \tag{4.24}$$

$$= \sum_{n,l=1,1}^{N,L} \bar{h}_{n,l} x_{n,l}^{(2)} + n, \tag{4.25}$$

where $\bar{h}_{n,l} = h_n \rho_l e^{j\theta_l}$ denotes the lth layer of the nth CS's signal's equivalent channel gain and $n \sim \mathcal{CN}(0, N_0)$ is the additive circulant complex Gaussian process having a variance of $\sigma^2 = N_0/2$ per dimension.

The receiver uses iterative DET and DEC as seen in Fig 4.12. Both the SPC and PANC aided MSC may use the same data detection algorithm. However, the soft channel decoder design of SPC aided MSC depends on the choice of the specific outer channel coding function $f(\cdot)$ employed in Eq. 4.11. Hence here we discuss the soft decoding of PANC only.

4.2.3.2 Data Detection

Without loss of generality, we consider the detection of the system employing PANC, but again, the SPC detector may be identical. A host of DET schemes may be invoked, including the powerful

but complex ML detection scheme, sphere decoding [115], etc. Here we opt for employing a low-complexity soft interference cancellation scheme [140]. Aiming for the detection of $x_u^{(2)}$, Eq. 4.23 may be written as

$$y = h_u x_u^{(2)} + \xi, \tag{4.26}$$

where ξ denotes the residual interference plus noise. By approximating ξ as a joint Gaussian random vector, which can be justified by the central limit theorem, we can model the extrinsic LLR as:

$$\begin{aligned} \mathcal{L}_{det}^{e}(x_u^{(2)}) &= \log_2 \frac{p(y|x_u^{(2)}=+1)}{p(y|x_u^{(2)}=-1)} \\ &= \log_2 \frac{\exp\left[(y-\hat{\xi}-h_u)^2/2V_\xi\right]}{\exp\left[(y-\hat{\xi}+h_u)^2/2V_\xi\right]} \\ &= 2h_u(y-\hat{\xi})/V_\xi, \end{aligned} \tag{4.27}$$

where the estimated value of ξ and its variance may be expressed as

$$\hat{\xi} = \sum_{n=1}^{N} h_n \hat{x}_n^{(2)} - h_u \hat{x}_u^{(2)} \tag{4.28}$$

$$V_\xi = \sum_{n=1}^{N} v_n |h_n|^2 + \sigma^2 - v_u |h_u|^2, \tag{4.29}$$

where the soft estimate of the transmitted signal is $\hat{x}_n^{(2)} = \tanh\left[\mathcal{L}_{det}^{a}(x_n^{(2)})/2\right]$ and the 'instantaneous' variance is $v_n = 1 - \hat{x}_n^{(2)}$. It can be seen from above that only the *a priori* knowledge of $\hat{x}_n^{(2)}$ is needed in the derivation of the extrinsic information $\mathcal{L}_{det}^{e}(x_u^{(2)})$, which is gleaned from the outer DEC.

4.2.3.3 Decoding of the PANC

The soft decoding of the PANC is analogous to that of a RA code [68]. As seen in Fig 4.12, it consists of the soft ACC decoder and soft combiner (COM). The PCM of the unity-rate ACC is defined in Eq. 4.21. When considering a non-systematic PANC, the soft decoder's graphical representation is shown in Fig 4.13. After inputting the soft output information of the DET ($\mathcal{L}_{det}^{e}$) to the PANC decoder, the soft output of the ACC is forwarded to the soft combiner COM ($\mathcal{L}_{acc}^{e}$) of all κ versions of the differently interleaved information streams $[\pi_1(\mathbf{s}), \pi_2(\mathbf{s}), \ldots, \pi_\kappa(\mathbf{s})]$, which are then soft-combined and fed back to the ACC decoder ($\mathcal{L}_{com}^{e}$) for the sake of providing updated soft-information for the DET ($\mathcal{L}_{dec}^{e}$). When a systematic PANC is employed, the soft-output of the ACC decoder provided for the COM block of Fig. 4.12 corresponds to all $\kappa^{ns} = (\kappa - \kappa^s)$ versions of the differently interleaved information streams $[\pi_{\kappa^{ns}+1}(\mathbf{s}), \ldots, \pi_\kappa(\mathbf{s})]$. The rest of the soft-information related to the κ^s versions of the differently interleaved information streams $[\pi_1(\mathbf{s}), \pi_2(\mathbf{s}), \ldots, \pi_{\kappa^s}(\mathbf{s})]$ is directly fed to the soft-combiner block COM of Fig 4.12, which means that there is no ACC decoding block between the DET and COM blocks. After carrying out

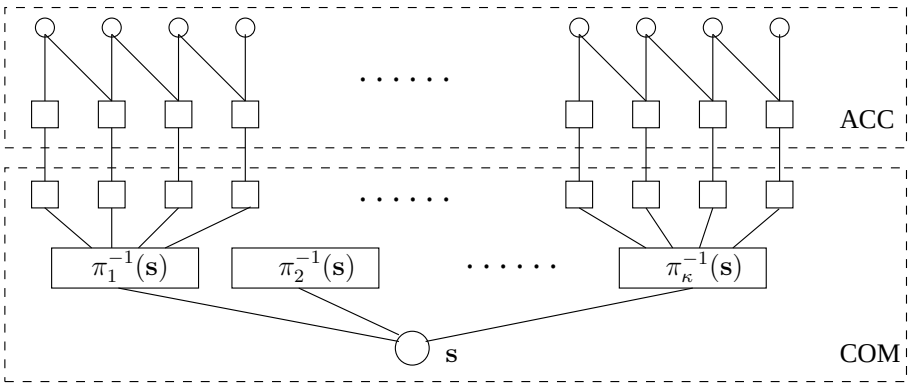


Figure 4.13: The factor graph representation of non-systematic PANC.

all the affordable iterations, the soft COM block of Fig 4.12 delivers its ultimate soft decision $\mathcal{L}_{com}^{p}$ concerning $s_i, i \in [1, K]$.

The soft-information delivered along each of the edges shown in Fig 4.13 obeys the classic sum-product algorithm [53], where variable nodes are denoted as circles and check nodes are denoted as squares. The message passed along edge j from a variable node to a check node is given by:

$$\mathcal{L}_j^e = \sum_{i=1, i\neq j}^{d_v} \mathcal{L}_i^a, \tag{4.30}$$

while that passed along edge j from a check node to a variable node is given by:

$$\mathcal{L}_j^e = 2\tanh^{-1}\left[\prod_{i=1, i\neq j}^{d_c} \tanh(\mathcal{L}_i^a/2)\right], \tag{4.31}$$

where d_c and d_v denotes the check and variable degrees, respectively, i.e. the total number of edges connected to a check node or variable node.

Remarks: The soft decoding of a PANC is similar to the conventional RA decoding [68], since both of them employ the sum-product algorithm and inherit quasi-random LDPC code-like properties [149]. The difference is that in a PANC, a full decoding iteration comprises a three-stage process, namely the soft-information exchange across the DET $\rightleftharpoons$ ACC $\rightleftharpoons$ COM decoding chain, while conventional RA decoding [68] performs a two-stage iteration denoted as ACC $\rightleftharpoons$ COM. Because of this similarity, we will employ a RA code, when the SPC scheme is used as the PANC scheme's benchmarker in our simulation. The main difference between employing a PANC and a SPC scheme from a decoding point of view is that the PANC arrangement benefits from the joint decoding of multiple information streams, while superposition coding performs single-stream channel-decoding, as seen in Fig 4.12.

		SPC				PANC	
η	N	R_1	R_2	L	ι	R_c (κ^s, κ^{ns})	ι
1/2	2	1/4	1/2	2	(1,5) x 20	1/4 (0,4)	20
1	4	1/4	1/4	4	(1,5,10) x 30	1/4 (0,4)	30
						1/4 (1,3)	30

Table 4.3: Simulation parameters in Fig 4.14 and Fig 4.15.

4.2.4 Performance Evaluation

4.2.4.1 Assumptions and Parameters

Let us now quantify the achievable performance of the proposed coding schemes. We assume error-free *Phase-I cooperation*, which is achieved with the aid of using CRC during each Phase-I transmission and by ensuring that cooperation is only activated by a perfect CRC check. The flat Rayleigh faded channels $h_n, n \in [1, N]$ between the N CSs and the destination are assumed to be independent and are perfectly known at the destination. Block fading is used, where the fades are kept constant during a SPC or PANC codeword, while changed independently between consecutive codewords.

Before comparing these two coding schemes, we firstly define our performance metric set χ, which consists of the achievable throughput η, the block error ratio P_e^{bl}, the delay τ and the complexity ι, i.e. we have $\chi = \{\eta, P_e^{bl}, \tau, \iota\}$. The system's effective throughput η may be defined as $\eta = R_c NLM$, where R_c is the channel coding rate, N is the number of CSs, L is the number of layers when the SPC scheme is employed, while we have $L = 1$ when the PANC scheme is considered. Finally, M denotes the modulation scheme used, which is BPSK for both coding schemes. For the SPC scheme, we employ a regular rate R_1 systematic RA code as the outer channel code in conjunction with a rate R_2 repetition code in oder to facilitate the multiple layers' superposition. Thus the total code-rate becomes $R_c = R_1 R_2$. On the other hand, the code-rate R_c of the PANC is defined as the number of differently interleaved versions κ, as introduced before.

Therefore, by setting the same system throughput η and the same source information segment length of $N_i = 512$ symbols, resulting in a fixed delay τ, we compare the two coding schemes in terms of their block error ratio P_e^{bl} and associated complexity ι. The complexity ι is simply quantified in terms of the number of iterations invoked. The total number of iterations of a SPC aided MSC scheme is the product of the number of DET $\rightleftharpoons$ DEC iterations and the number of iterations within the RA code, while that of a PANC aided system is deemed to be proportional to the number of iterations invoked by the three-stage DET $\rightleftharpoons$ ACC $\rightleftharpoons$ COM decoder chain. The simulation parameters used are summarised in Table 4.3.

	g_1	g_2	g_s
u.b.	$\|h_1\|^2$	$\|h_2\|^2$	$(\|h_1\|^2 + \|h_2\|^2)/2$
l.b.	$\|h_{12}\|^2 + \|h_1\|^2$	$\|h_{12}\|^2 + \|h_2\|^2$	$(\|h_1\|^2 + \|h_2\|^2)/2$

Table 4.4: Effective channel gain used in calculating the outage probability, where $|h_1|$ and $|h_2|$ denote the source to destination channel gain, while $|h_{1,2}| = |h_{2,1}|$ denotes the identical inter-source channels.

4.2.4.2 Outage Bound Analysis

We now perform an outage bound analysis as a reference for the cooperative coding schemes. Without loss of generality, we discuss the $N = 2$ MSC aided scenario. The maximum mutual information I of an $N = 2$ MSC-aided multiple access channel is equal to the minimum amongst the individual source's mutual information I_1, I_2 and the sum mutual information I_s, which is given by [34]

$$I = \min\{I_1(g_1), I_2(g_2), I_s(g_s)\}, \tag{4.32}$$

where $I_i(g_i), i \in \{1, 2, s\}$ is written in the classic form as a function of the SNR γ_s and the effective channel gain g_i, more explicitly:

$$I_i(g_i) = \log_2(1 + g_i\gamma_s) \qquad i = 1, 2 \tag{4.33}$$

$$I_s(h_s) = (1/2)\log_2(1 + g_s\gamma_s). \tag{4.34}$$

Table 4.4 collects g_i for different scenarios, where the upper bound (u.b.) corresponds to the no cooperation scenario, while the lower bound (l.b.) corresponds to the cooperation bound [34], which has a max-flow min-cut interpretation. The outage probability of a fading channel is defined as the probability of having a mutual information between the received soft value and the decided symbol, which is less than the system's target effective throughput η, formulated as:

$$p_{out} = \Pr[I < \eta]. \tag{4.35}$$

Finding the outage probability p_{out} at the system's target effective throughput η and a given SNR per-bit γ_b is equivalent to finding

$$p_{out} = \Pr[g < g_\eta], \tag{4.36}$$

where $g_\eta = (2^\eta - 1)/\eta\gamma_b$ and $g = \min\{g_1, g_2, g_s\}$. The minimum outage probability p_{out}^{min} at a given γ_b value is achieved by letting $\eta \to 0$ and it is well known that $\lim_{\eta\to 0}(2^\eta - 1)/\eta = \ln 2$.

4.2.4.3 Simulation Results

Fig 4.14 and Fig 4.15 suggest that both of our proposed coding schemes are capable of approaching the outage probability bound at their corresponding system throughput η. When $N = K = 2$ sources cooperate in a cluster as characterised in Fig 4.14, the non-systematic PANC scheme employing $\iota_{PANC} = 20$ iterations performs within a small fraction of a dB from the SPC scheme,

which requires a total of $\iota_{SPC} = 5 \times 20 = 100$ iterations, hence the former results in a significantly lower complexity. The same trend was also confirmed, when $N = K = 4$ sources cooperate in a cluster, as characterised in Fig 4.15. However, since the effective system throughput η was doubled from half to unity, both schemes exhibited a slightly higher discrepancy w.r.t. the outage probability bound. Observe in Fig 4.15 that the systematic PANC performs better than its non-systematic counterpart and its performance is close to that of the more complex SPC system. If the affordable complexity is not an issue, then the SPC system may outperform the PANC system, as seen in Fig 4.15. Since the complexity imposed determines the total power consumption, the PANC scheme may be considered as being more power-efficient.

It was found in Fig 4.15 that the non-systematic PANC is unable to fully exploit the spatial transmit diversity gain provided by $N = 4$ CSs due to its randomly designed nature. This is particularly true, when the effective throughput becomes unity, which may be referred to as a 'fully-loaded' MSC-aided scenario. The systematic PANC, on the other hand, provides an additional diversity gain, since the systematic information segment of a codeword facilitates direct communication with the soft COM of Fig 4.12 at the cost of sacrificing some of the attainable coding gain. When the effective throughput is less than unity, i.e. the system is partially-loaded, the non-systematic PANC performs sufficiently well, as seen in Fig. 4.14. From a classic coding point of view [149], the systematic information segment provides a certain amount of direct *a priori* information for the soft COM of Fig 4.12, which becomes particularly crucial, when the *a priori* information gleaned from the ACC block of Fig 4.12 is low, as in the fully-loaded scenario. This suggests that the performance of the PANC scheme proposed may be subject to a coding-gain versus diversity-gain tradeoff, which may be further investigated in our future work.

4.3 Conclusion

In this chapter, we applied the concept of non-orthogonal random waveform based multiuser communications to cooperative multiple access scenarios. Specifically, in Section 4.1 we proposed the so-called IR-STC scheme. First of all, the slot utilisation efficiency of our MSC compared to SSC enables our cooperative mechanism to achieve both a high diversity gain and a high multiplexing gain. With the aid of the matrix based representation, we compared our IR-STC to the traditional G_4 OSTBC, where the conclusion emerged that our IR-STC is capable of achieving a high throughput, while maintaining a low BER with the aid of decentralised cooperation, where decentralisation is achieved by the novel SE random interleaver generation method. All these properties render our IR-STC design eminently applicable for employment in interference-limited high-user-density *ad hoc* networks, in conventional cellular networks assisted by mobile relays complemented by fixed wireless relays and in other application-oriented scenarios.

Furthermore, we analysed the achievable performance of our IR-STC aided MSC employing various relaying techniques, where we found that when the SNR of the inter-source channel is

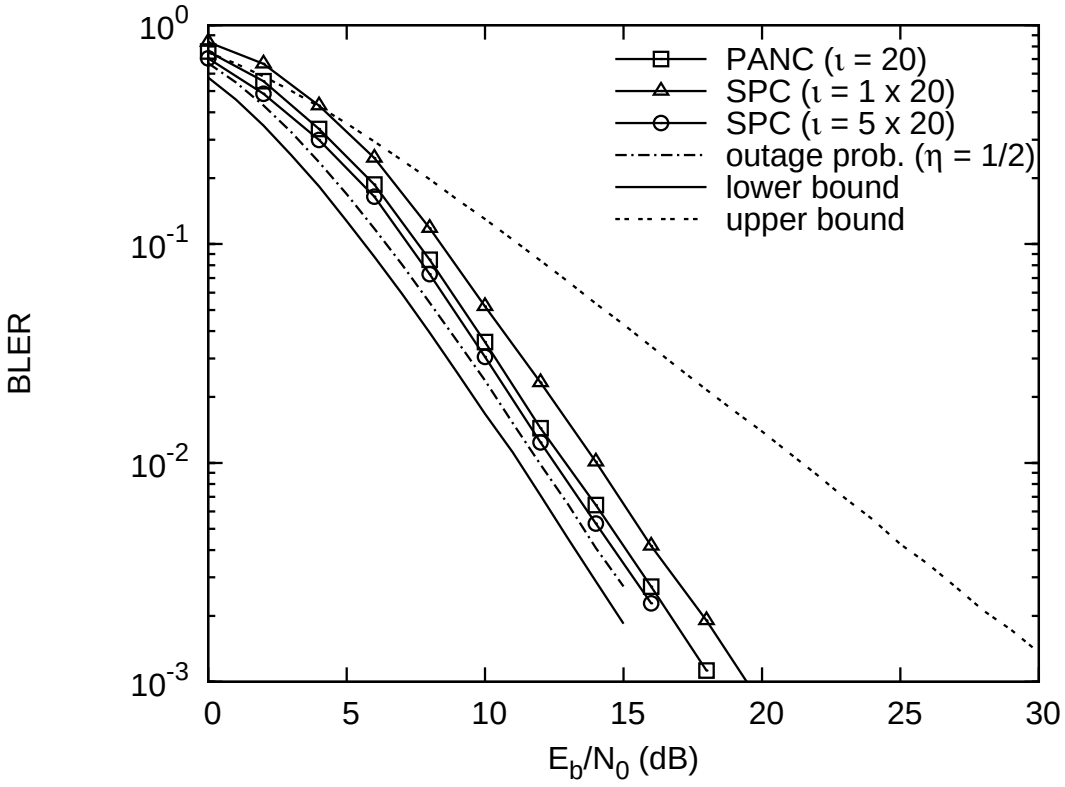


Figure 4.14: BLER of the SPC and PANC scheme aided MSC against their upper and lower bounds, where we have $K = N = 2$ and $\eta = 1/2$. The simulation parameters are listed in Table 4.3.

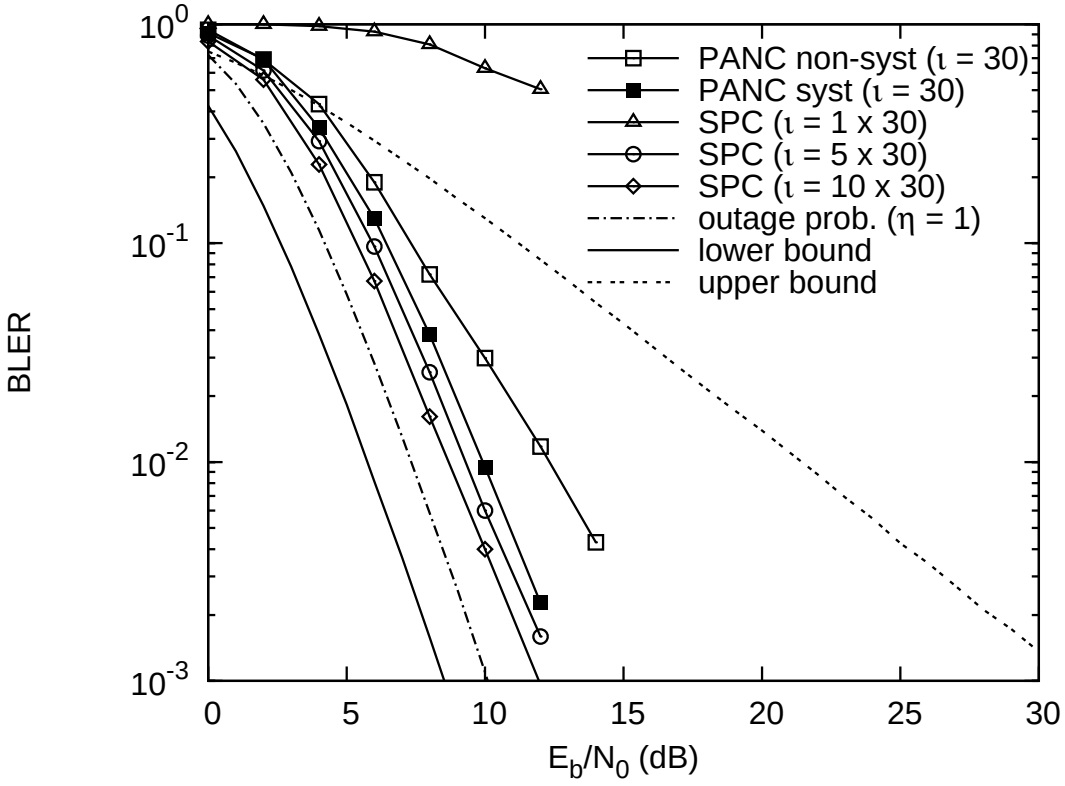


Figure 4.15: BLER of the SPC and PANC scheme aided MSC against their upper and lower bounds, where we have $K = N = 4$ and $\eta = 1$. The simulation parameters are listed in Table 4.3.

better than that of the source to destination channel, DF is the best relaying strategy in the presence of benign fading. The AF technique is only preferable at high SNRs, when Rayleigh fading is encountered. The DDF technique performs consistently worse than DF, while surprisingly, when the fading is benign, non-coherent DDF operating without the cost of estimating the inter-source channels outperforms the coherent detection aided AF technique.

In order to improve the achievable performance of the uncoded scenario, we also proposed two different coding schemes for energy efficient MSC, namely the SPC and PANC schemes of Section 4.2. The powerful channel coded SPC scheme may be considered as a direct extension of the uncoded system of Section 4.1, while the PANC scheme may be referred to as a joint network coding and channel coding scheme, which inherits the benefits of multiplexed random codes in a distributed communications scenario. The simulation results of Fig 4.14 and Fig 4.15 demonstrate that both schemes are capable of performing close to the outage probability bound. When compared to the SPC arrangement, the novel PANC scheme of Section 4.2 exhibits a lower complexity at the cost of a slight performance degradation, while maintaining the same effective throughput and delay.

4.4 Appendix - Nakagami-m Distribution and Simulator

The Nakagami-m distribution [62] is capable of modelling severe, moderate, light, or non-fading environments by adjusting its parameter, m. In fact, the sum of the mutually exclusive Hoyt and Rican models is the Nakagami-m fading distribution [62]. Consequently, Nakagami-m fading signals have a better fit with empirical measurement-based data. In this appendix, we describe the method of generating the Nakagami-m distribution with the aid of Gaussian distribution.

4.4.1 Nakagami-m Distribution

A non-negative real-valued random variable $x \in \mathcal{R}_{\geq 0}$ with its PDF:

$$p_Z(x) = \frac{2m^m}{\Gamma(m)\Omega^m} x^{2m-1} \exp\left\{-\frac{mx^2}{\Omega}\right\}, x \geq 0 \tag{4.37}$$

is referred to as *Nakagami-m distributed*, where $m > 0.5$ is the Nakagami-fading parameter, $\Omega = \mathrm{E}\left\{x^2\right\}$ is the variance of x and $\Gamma(\cdot)$ is the Gamma function. It describes the random channel amplitude distribution of a rich-scattering fading environment having a uniformly distributed phase in $[0, 2\pi)$. It is capable of modelling both severe fading ($m < 1$) and moderate fading conditions ($m > 1$) with typical m-values of:

- $m = 0$ one-side Gaussian distribution
- $m = 1$ Rayleigh distribution
- $m \to \infty$ Gaussian distribution

Moreover, for $m >> 1$, the Nakagami-m distribution is well approximated by a Rican distribution with a Rican K-factor of

$$K = \frac{\sqrt{m^m - m}}{m - \sqrt{m^m - m}}. \tag{4.38}$$

The PDF of the corresponding instantaneous power $\gamma = x^2$ is given by

$$p(\gamma) = \frac{1}{\Gamma(m)(\bar{\gamma}/m)^m}\gamma^{m-1}\exp\left\{-\frac{\gamma}{\bar{\gamma}/m}\right\}, \gamma \geq 0 \tag{4.39}$$

where we have $\bar{\gamma} = \mathrm{E}\{\gamma\} = \Omega$.

4.4.2 Analysis of Nakagami-m Distribution

For the sake of generating Nakagami-m fading using Gaussian source samples, we first underline the associated probability fundamentals. The rationale of employing Gaussian source samples is conditioned on the relationship of:

$$Z \Longleftrightarrow \Gamma \Longleftrightarrow \chi^2 \Longleftrightarrow \mathcal{N}, \tag{4.40}$$

which indicates that the *Nakagami-m* distributed variable Z , is the square root of the *Gamma* distribution, i.e. we have $Z = \sqrt{\Gamma}$. In essence, the Gamma distribution is the sum of a set of statistically independent and identically distributed (i.i.d) *central chi-square* distributions χ^2. Moreover, the χ^2 distribution can be attained by $\chi^2 = x^2$, where $x \sim \mathcal{N}(0, \sigma^2)$ is the normal (Gaussian) distribution. Note that when $x \sim \mathcal{N}(\bar{x}, \sigma^2)$ with $\bar{x} \neq 0$, the resultant χ^2 distribution is referred to as the *non-central chi-square* distribution.

Let us now commence from the normal distribution $\mathcal{N}$ and generate the Nakagami-m distribution Z according to Eq. (4.40). The PDF of χ^2 is

$$p_{\chi^2}(x) = \frac{1}{\sqrt{2\pi x\sigma^2}}\exp\left\{-\frac{x}{2\sigma^2}\right\}, x \geq 0. \tag{4.41}$$

The PDF of the χ^2_n distribution with n degrees of freedom, which is created from n Gaussian distributed variables x_i according to $\chi^2_n = \sum_{i=1}^{n} x_i^2$ with i.i.d $x_i \sim \mathcal{N}(0, \sigma^2), \forall i$ is

$$p_{\chi^2_n}(x) = \frac{1}{\Gamma(\frac{n}{2})(2\sigma^2)^{\frac{n}{2}}}x^{\frac{n}{2}-1}\exp\left\{-\frac{x}{2\sigma^2}\right\}, x \geq 0. \tag{4.42}$$

On the other hand, the PDF of the Gamma distribution with shape parameter k and scale parameter θ is

$$p_\Gamma(x) = \frac{1}{\Gamma(k)(\theta)^k}x^{k-1}\exp\left\{-\frac{x}{\theta}\right\}, x \geq 0, \tag{4.43}$$

where we have $k > 0$ and $\theta > 0$. Comparing Eq. (4.43) and Eq. (4.42), the gamma distribution is equivalent to χ^2_n if we have $k = n/2$ and $\theta = 2\sigma^2$. Furthermore, by letting $m = n/2$ and $\Omega = 2m\sigma^2$, Eq. (4.43) can be expressed as

$$p_\Gamma(x) = \frac{1}{\Gamma(m)(\Omega/m)^m}x^{m-1}\exp\left\{-\frac{x}{\Omega/m}\right\}, x \geq 0. \tag{4.44}$$

It is widely exploited that if x has a Nakagami-m distribution with parameters of m and Ω, then x^2 has a Gamma distribution with a shape parameter of m and scale parameter of Ω/m. Thus Eq. (4.44) is related to the Nakagami-m distribution

$$p_Z(x) = \frac{2}{\Gamma(m)(\Omega/m)^m} x^{2m-1} \exp\left\{-\frac{x^2}{\Omega/m}\right\}, x \geq 0. \tag{4.45}$$

It is now clear that Eq. (4.44) and Eq. (4.45) are exactly the same as Eq. (4.39) and Eq. (4.37), respectively.

From our previous discussions, we find that the Gamma distribution is equivalent to χ^2_n with $n = 2m$ degrees of freedom, i.e. we have

$$\Gamma = \sum_{i=1}^{2m} x_i^2. \tag{4.46}$$

To elaborate a little further, Eq. (4.46) can be extended as

$$\Gamma = \alpha \sum_{i=1}^{\mu} x_i^2 + \beta x_{\mu+1}^2 \tag{4.47}$$

where $\mu = \llcorner 2m \lrcorner$ is the integer part of $2m$ and $x_{\mu+1}^2$ is the fractional part, respectively. The weights α and β are given by:

$$\alpha = \frac{2\mu m + \sqrt{2\mu m\,(1+\mu-2m)}}{\mu\,(1+\mu)}, \tag{4.48}$$

$$\beta = 2m - \alpha\mu. \tag{4.49}$$

Hence, the resultant Nakagami-m distribution is:

$$Z = \sqrt{\alpha \sum_{i=1}^{\mu} x_i^2 + \beta x_{\mu+1}^2}. \tag{4.50}$$

4.4.3 Nakagami-m Simulator

Based on Eq. (4.50), which essentially describes the relation between the Nakagami-m distribution and the Gaussian distribution, we can generate Nakagami-m distributed complex random variables r using existing Gaussian random samples based on the following steps:

1. Generate (correlated) Gaussian random variables $x_i \sim \mathcal{N}(0, \sigma^2), i = 1, \ldots, \mu + 1$;
2. Generate the sum $y = \sum_{i=1}^{\mu} x_i^2$;
3. The real part (Re) of r is attained by $\mathrm{Re}\{r\} = \sqrt{\alpha y}$;
4. The imaginary part (Im) of r is attained by $\mathrm{Im}\{r\} = \sqrt{\beta x_{\mu+1}^2}$;
5. Add a random phase rotation $r = r\exp\{j\phi\}$, where ϕ is uniformly distributed in $[0, 2\pi)$.

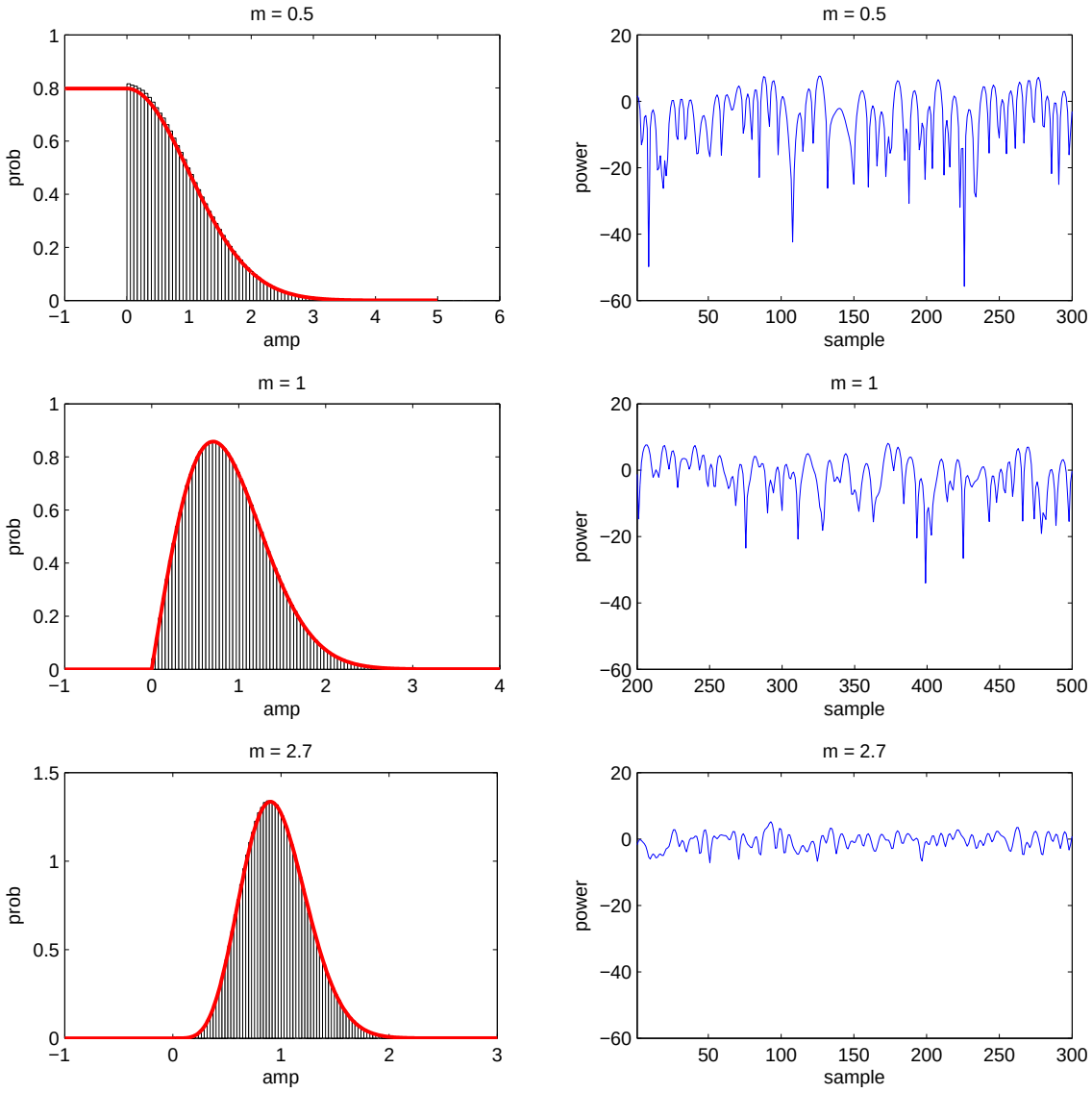


Figure 4.16: The PDF and power envelope of Nakagami-m fading, where $m = 0.5$, $m = 1$ and $m = 2.7$ from top to bottom and normalised Doppler frequency is 0.1.

The resultant amplitude of the complex random variable r is Nakagami-m distributed, i.e. we have $|r| \sim Z(m, \Omega)$, where $\Omega = 2m\sigma^2$. Some of our simulation results are shown in Fig. 4.16, characterising the PDF as well as a time-domain segment of the Nakagami-m distribution for $m = 0.5, m = 1, m = 2.7$.

Chapter 5

Superposition Coding Aided Multiplexed Hybrid ARQ

In Chapters 3 and 4, the concept of non-orthogonal random waveform based multiuser communications have been discussed, namely cellular systems and cooperative systems. In this chapter, we propose a novel superposition coding aided multiplexed HARQ scheme for the sake of improving both the link-layer and TCP layer efficiency. We present the detailed system design and discuss the transmission efficiency metrics of both layers, namely the link-layer's effective throughput and the TCP layer's mean frame arrival rate, where the latter is modelled by the classic Markov chain. Furthermore, the achievable link-layer packet error rate is quantified. It is demonstrated that our scheme substantially improves the attainable transmission efficiency of both layers and it is particularly suitable for delay-sensitive services. Hence, the proposed M-HARQ scheme constitutes a practical application of our general concepts considered in this thesis.

5.1 Introduction

For the sake of further improving the robustness against link adaptation inaccuracy due to various implementation impairments, such as channel estimation/prediction errors, feedback delay, unpredictable co-channel interference etc, HARQ schemes have been proposed [2, 6], which combine channel coding with the ARQ protocol. It has been considered as one of the key link-layer techniques in various standards, such as HSPA [2], the 3GPP LTE initiative [7] and in the WIMAX system [8]. In HARQ, the receiver asks for a packet's retransmission using the reverse-direction channel with the aid of a single-bit Negative-ACKnowledgement (NACK) flag, whenever its currently decoded packet is deemed to be erroneous based on the decision of the CRC scheme. In general, the retransmissions in ARQ-aided systems can be carried out in different manners, for example using a Type-I packet combing scheme [6] and a Type-II incremental redundancy scheme [6].

The corresponding theoretical throughput, error rate and delay analysis of the classic HARQ

scheme can be found in [150–152] and in references therein. Advanced MIMO-HARQ schemes were discussed in [153]. As a virtual MIMO interpretation, the relation between relay-aided networks and HARQ schemes was revealed in [154]. Numerous efforts were also made to specifically design channel codes for Type-II HARQ, using for example Rate Compatible Punctured Codes (RCPC) [155], Raptor codes [156] and rateless codes [157]. Moreover, instead of employing the traditional CRC, the inherent error detection capability of the channel code was exploited in the packet reliability estimation aided HARQ scheme of [158]. On the other hand, comprehensive solutions were designed for HSPA in [159] and for DL Orthogonal Frequency Code Division Multiplexing (OFCDM) in [160]. Additionally, the modelling methodology of HARQ schemes used for system-level simulations was developed in [161] and [162].

From a cross-layer point of view [163], HARQ also plays a crucial role in the overall system's transmission efficiency. Multiple retransmissions may occur at different layers of the seven-layer Open Systems Interconnection (OSI) architecture, which have different functionalities. Higher-layer retransmissions ensure reliable delivery of the packets over the entire radio access network, which may require *end-to-end* retransmission. By contrast, lower-layer retransmission aims to correct transmission errors inflicted by the physical channel based on a *hop-by-hop* per-packet retransmission. These two interacting retransmission functions jointly contribute towards the overall efficiency of the system.

The TCP supports reliable end-to-end data transmission and facilitates congestion control, where the transmission frame loss due to link errors is often assumed to be negligible in wired networks. However, directly transplanting the TCP into wireless applications suffers from link impairments, such as radio link attenuation, fading, handover, mobility and co-channel interference. For the sake of avoiding congestion due to physical retransmissions induced by channel errors, we may either conceal the effects of the wireless link from the TCP-enabled transmitter by invoking channel coding or we make it aware of the wireless channel effects. Various strategies have been proposed in the literature for combating this retransmission-induced congestion problem, which can be classified as: split-connection-based solutions [164], proxy-based solutions [165], link-layer-based solutions and end-to-end solutions [166]. For example, link-layer approaches such as HARQ attempt to conceal the channel-induced packet loss events from the TCP-enabled transmitter by reducing the effects of wireless link errors with the aid of channel coding combined with retransmissions on a prompt packet-based timescale. This solution is appealing as it does not incur the typical overhead associated with TCP-awareness and yet obeys the TCP semantics. However, this HARQ aided approach introduces extra delay due to local link layer retransmissions, which may potentially lead to a timeout and hence may trigger the slow-start phase of the TCP transmission.

Against this backdrop, in this chapter we aim at improving the overall *end-to-end* TCP transmission efficiency by reducing the link layer's *hop-by-hop* HARQ retransmission delay with the aid of our proposed superposition coding aided Multiplexed HARQ (M-HARQ) scheme, which jointly encodes the current new packet to be transmitted and any packets that are about to be retransmitted.

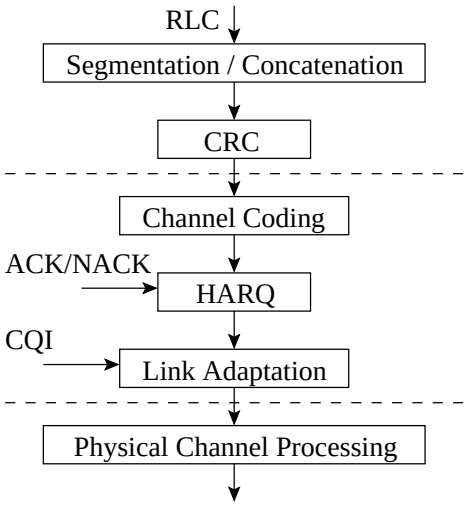


Figure 5.1: Cross-layer interactions of the HARQ scheme.

In other words, the link-layer retransmissions are embedded in the next new packet's transmission, which avoids any potential throughput reduction imposed by retransmissions although naturally, they do impose additional interference. A similar idea was proposed in [167], which requires a specifically designed channel code and its application is limited to twin-packet joint transmissions. As a benefit, our proposed scheme is capable of jointly and simultaneously transmitting multiple packets and it is equally applicable to both Type I and II HARQ techniques. In addition, our proposed scheme may even be combined with other TCP-related cross-layer designs, such as the Explicit Loss Notification (ELN) scheme of [168]. Hence, the advocated technique can be seamlessly integrated with diverse existing and future systems.

In a nutshell, the contribution of this chapter is that *we propose a novel M-HARQ scheme, which improves both the link layer as well as the end-to-end TCP layer transmission efficiency at the cost of a marginal link layer Packet-Error-Ratio (PER) performance degradation, which is imposed by the associated slight interference degradation.*

The rest of the chapter is organised as follows. In Section 5.2, we provide a general description of the HARQ scheme proposed and contrast it to the classic HARQ approach. Furthermore, the structure of our proposed M-HARQ arrangement is described, followed by the associated encoding and iterative decoding algorithms. In Section 5.3, the link layer and TCP layer transmission efficiency metrics used are introduced. In Section 5.4, both the link layer PER performance and the transmission efficiency of both the conventional HARQ and the proposed M-HARQ scheme are evaluated and discussed. Finally, we conclude our discourse in Section 5.5.

5.2 Multiplexed Hybrid ARQ

5.2.1 Conventional Approach

Fig. 5.1 provides a simplified model of the classic HARQ scheme along with its interactions between different layers. The information to be passed from the Radio Link Control (RLC) layer to the physical layer for transmission over the wireless channel is segmented into a number of packets having the appropriate size, which are then extended by the classic CRC. These packets are then subjected to channel coding, link adaptation as well as bit-to-symbol mapping based on the Channel Quality Indicator (CQI) and further processed for multicarrier transmission, beamforming etc. The HARQ scheme is only activated, when an NACK signal is fed back by the receiver to the transmitter through the feedback channel, which typically occurs owing to inaccurate link adaptation. This model suggests that the proposed bit-based M-HARQ scheme is transparent to both the actual link adaptation procedure considered as well as to the specific physical layer processing employed.

Being a physical-layer-aware ARQ scheme, HARQ combines the CRC encoding function of the link layer with channel coding in the physical layer. The family of HARQ schemes has two subclasses. In Type-I HARQ, the same coded packet is used in consecutive retransmissions, allowing the receiver having a sufficiently large memory to perform soft combining of the various replicas of the packets before decoding. Naturally, each packet is also individually decodable for a receiver without sufficient memory to decode each replica of retransmitted packets, although typically this results in a residual PER penalty or in an increased number of retransmissions. By contrast, in Type-II HARQ, only additional parity is conveyed in consecutive retransmission attempts and the receiver carries out joint decoding of all collected information as a new reduced-rate code. In order to facilitate decoding with the aid of a single decoder regardless of the code-rate, typically the family of nested codes is used, where the higher-rate codes are typically punctured versions of the lower-rate mother-code. Finally, Type-I HARQ may also be viewed as self-decodable Type-II HARQ using a repetition code.

Following the above conceptual introduction of the two basic types of HARQ schemes, let us now describe them mathematically. The information arriving from the TCP layer, which is referred to here as a *frame*, is partitioned into M *packets* of equal length N_i, $\mathbf{u}_m \in \{0,1\}^{N_i}, m \in [1,M]$. These packets are protected by the channel coding function $f_{c,v} \in \Omega = \{f_1,\ldots,f_V\}$ of rate $r_{c,v} \in \mathrm{R} = \{r_1,\ldots,r_V\}$, where Ω and R represent a set of predefined discrete rate-compatible codes and their corresponding rates. The selection of a particular code-rate is based on the CQI controlling the link adaptation procedure [5]. The maximum number of retransmissions is $L < M$, i.e. had a total of $(L+1)$ transmission attempts.

For Type-I HARQ, the same coded packet is repeated L times, i.e. we have $f^0_{c,v}(\mathbf{u}_m) = f^l_v(\mathbf{u}_m), l \in [1,L]$, where the superscript $'0'$ stands for the initial transmission. However, for Type-II HARQ, additional parity is transmitted during consecutive retransmissions without retransmitting

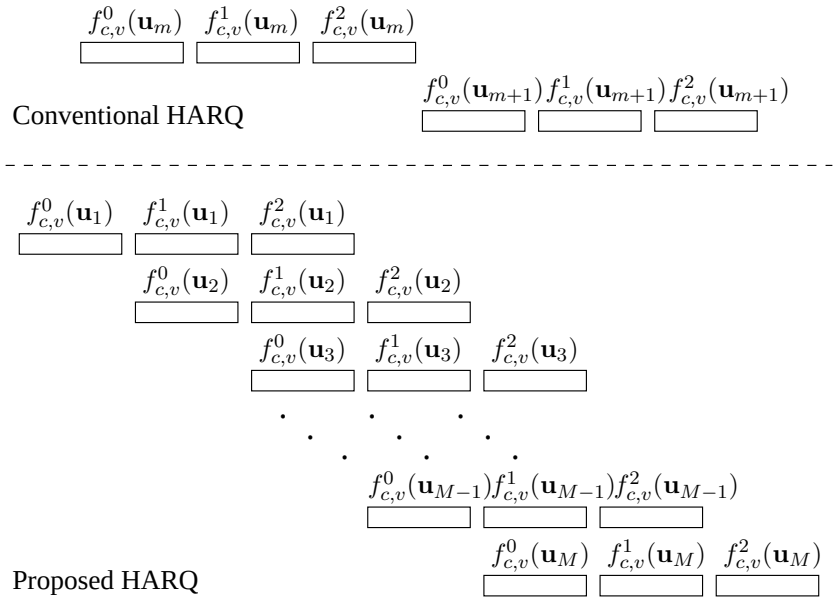


Figure 5.2: Classic HARQ and the proposed multiplexed HARQ in conjunction with $L = 2$.

the original information, i.e. we have $f^i_{c,v}(\mathbf{u}_m) \neq f^j_{c,v}(\mathbf{u}_m), i, j \in [0, L]$. These additional parities were obtained by either reinserting the previously punctured bits or by further reducing the rate of the mother code employed, thus the size of retransmitted packets $f^i_{c,v}(\mathbf{u}_m), i \in [1, L]$ may not be identical to that of the original one $f^0_{c,v}(\mathbf{u}_m)$. After successfully decoding the mth packet during the $(L+1)$st transmission attempt, the transmission of the $(m+1)$st packet is activated. The whole process is illustrated in Fig. 5.2.

5.2.2 Proposed Approach

The strategy of transmitting the next new packet only when the successful reception of the current one was confirmed is highly inefficient, which is analogous to the widely recognised drawback of conventional Stop-and-Wait ARQ [6]. However, if the receiver is capable of tolerating a modest amount of additional interference, the next new packet can be simultaneously transmitted with the retransmissions of the previous $K \in [1, L]$ erroneous packets, as seen in Fig. 5.2. In other words, M new packets are continuously transmitted, while the K erroneous packets are transmitted on a virtual channel, appropriately combined with the new packets.

5.2.2.1 Structure

In general, different packets require different number of retransmissions, depending on the instantaneous channel conditions. We consider the worst-case scenario, where each packet exploited the maximum number of retransmissions L, so that we can evaluate the maximum of the PER after L retransmissions. In the worst-case scenario considered and when employing the superposition coding scheme to be introduced shortly, the resultant interference of our M-HARQ arrangement

becomes similar to that of the ISI effects experienced for transmission over a dispersive channel in the absence of HARQ transmissions. Analogously, our scheme may be interpreted as generating Inter-Packet-Interference (IPI) and hence can be represented with the aid of a Toeplitz-matrix in the form of:

$$\mathbf{G}_{M-HARQ} = \begin{bmatrix} 1 & 1 & 1 & & & & \\ & 1 & 1 & 1 & & & \\ & & 1 & 1 & 1 & & \\ & & & 1 & 1 & 1 & \\ & & & & 1 & 1 & 1 \end{bmatrix}. \tag{5.1}$$

This band-structured matrix describes the proposed M-HARQ scheme for the specific example of $L = 2$ and $M = 5$, where a total of $M + L = 7$ packet transmissions are required. More generally, it may be inferred from Fig. 5.2 that the conventional scheme requires a total of $M_r = M(L+1)$ packet transmissions, while our scheme necessitates only $M_r = M + L$ transmissions.

Remarks: The structure of our M-HARQ scheme may also be related to the relaying scenario, where the continuously transmitted M packets are oriented from the direct source-to-destination link while the maximum of L retransmissions of a specific packet are activated during the consecutive original packet transmissions from a set of L relay-to-destination links. Hence, the rate-loss of the consecutive retransmissions of a packet due to orthogonal time diversity achieved by the conventional HARQ scheme is mitigated by the proposed non-orthogonal spatial diversity approach facilitated by the relaying scenario considered. Relaying scenario was also referred to as a so-called 'opportunistic multipath scenario' [169], which more explicitly justifies the efficiency of our proposed M-HARQ scheme.

5.2.2.2 Encoding

Generally speaking, the joint encoding function F of the mth transmission can be represented as $F(\mathbf{u}_{a_1}, \ldots, \mathbf{u}_{a_2})$, where we have:

$$\begin{cases} (a_1, a_2) = (m, 1) & 1 \leq m \leq L, \\ (a_1, a_2) = (m, m-L) & L < m \leq M, \\ (a_1, a_2) = (M, m-L) & M < m \leq M+L. \end{cases} \tag{5.2}$$

Although in principle specifically designed coding functions may be created, we opt for the powerful superposition coding concept in this chapter:

$$F(\cdot) \quad = \quad \sum_{i=a_2}^{a_1} \rho_i e^{j\theta_i} f\left[f_{c,v}^{m-i}(\mathbf{u}_i)\right], \tag{5.3}$$

where each superimposed packet is referred to as a layer, while ρ_i and $\theta_i \in [0, \pi)$ denote the layer-specific amplitude- and phase-rotation, respectively. The benefit of choosing this particular superposition coding technique is that by opting for this simple linear operation, the specific modulation

function $f(\cdot)$ and channel coding function $f_{c,v}(\cdot)$ of the individual layers may be retained. Imposing the associated phase rotation θ_i has two benefits, namely that of reducing the Peak-to-Average Power Ratio (PAPR) of the transmitted signal and making the multiple layers more distinguishable. In this chapter, an identical amplitude allocation and uniform phase rotations are employed for the individual superimposed layers.

5.2.2.3 Decoding

Our M-HARQ scheme employs iterative Multiple Packets Detection (MPD) and DEC exchanging extrinsic information between these two receiver components, as seen in Fig. 5.3. We focus our attention on the MPD algorithm, since the choice of the DEC algorithm depends on the specific channel code employed. A host of MPD schemes may be invoked, including the powerful but high-complexity ML detection scheme, sphere decoding [115] etc. Here we opt for employing a low-complexity soft interference cancellation scheme.

The signal received after the mth packet's transmission may be represented as:

$$\mathbf{y} = \sum_{i=a_2}^{a_1} h_m^b \rho_i e^{j\theta_i} f\left[f_{c,v}^{m-i}(\mathbf{u}_i)\right] + \mathbf{n}, \tag{5.4}$$

where h_m^b is the block-fading channel's impulse response and $h_i = h_m \rho_i e^{j\theta_i}$ denotes the ith layer's equivalent channel gain, while $n \sim \mathcal{CN}(0, N_0)$ is the additive circulant complex Gaussian noise process having a variance of $\sigma^2 = N_0/2$ per dimension. When denoting the modulated packet as $\mathbf{x}_i = f\left[f_{c,v}^{m-i}(\mathbf{u}_i)\right]$, we consider the nth symbol of the mth transmission packet and aim for the detection of the jth layer's symbol $x_j = \mathbf{x}_j(n)$, then Eq. 5.4 may be written as

$$y = h_j x_j + \xi, \tag{5.5}$$

where ξ denotes the residual interference plus noise. By approximating ξ as a joint Gaussian random vector, which can be justified by the central limit theorem, we can model the extrinsic symbol probability as:

$$P^e(x_j = x) \quad \propto \quad \exp\left[-|y - \hat{\xi} - h_j x|^2 / 2V_\xi\right], \tag{5.6}$$

where $x \in A$ is the particular realization drawn from the modulation alphabet $\mathcal{A}$. The estimated value of ξ and its variance may be expressed as

$$\hat{\xi} \quad = \quad \sum_{i=a_2}^{a_1} h_i \hat{x}_i - h_j \hat{x}_j, \tag{5.7}$$

$$V_\xi \quad = \quad \sum_{i=a_2}^{a_1} v_i |h_i|^2 + \sigma^2 - v_j |h_j|^2, \tag{5.8}$$

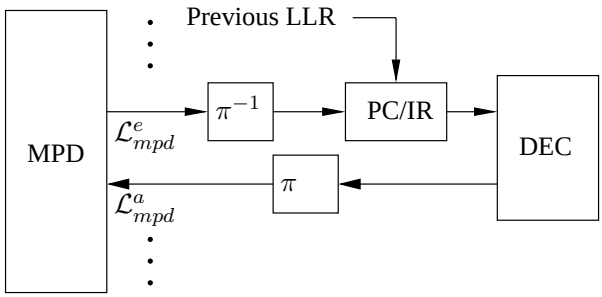


Figure 5.3: Iterative receiver architecture of the mth packet's reception.

where the soft symbol $\hat{x}_i$ and the 'instantaneous' variance v_i are given by:

$$\begin{aligned} \hat{x}_i &= \mathbb{E}[x_i] && (5.9)\\ &= \sum_{x\in\mathcal{A}} xP^a(x_j = x), && (5.10)\\ v_i &= \mathrm{Cov}[x_i, x_i] && (5.11)\\ &= \sum_{x\in\mathcal{A}} |x|^2 P^a(x_j = x) - |\hat{x}_i|^2. && (5.12) \end{aligned}$$

For the decoder of a binary code, the extrinsic non-binary symbol probability $P^e(x_j)$ may be converted to the bit-based extrinsic LLR $\mathcal{L}^e_{mpd}(d^q_j), q \in [1, Q]$, where we have $Q = \log_2 |\mathcal{A}|$ and $|\mathcal{A}|$ is the cardinality, i.e. the number of phases in the modulation alphabet $\mathcal{A}$. The extrinsic LLR of the qth bit is thus given by:

$$\mathcal{L}^e_{mpd}(d^q_j) = \log_2 \frac{\sum_{x\in\mathcal{A}^+_q} P^e(x_j = x)P^a(x_j = x)}{\sum_{x\in\mathcal{A}^-_q} P^e(x_j = x)P^a(x_j = x)}, \tag{5.13}$$

where $\mathcal{A}^+_q$ and $\mathcal{A}^-_q$ denotes the two subsets of $\mathcal{A}$ hosting symbols with their qth bit being +1 and -1, respectively. It can be seen from Eq. (5.13) that in the derivation of the extrinsic information $\mathcal{L}^e_{mpd}(d^q_j)$, only the *a priori* symbol probability $P^a(x_j = x)$ is needed, which is given by:

$$P^a(x_j = x) = \prod_{q\in[1,Q]} \frac{1}{2}\left\{1 + d^q \tanh\left[\mathcal{L}^a_{mpd}(d^q_j)/2\right]\right\}, \tag{5.14}$$

where $d^q \in \{\pm 1\}$ is the qth bit's polarity in symbol x. This corresponds to a bit-LLR to symbol-probability conversion, where the bit LLR $\mathcal{L}^a_{mpd}(d^q_j)$ is gleaned from the output of the DEC block of Fig. 5.3.

The jth layer's extrinsic LLR $\mathcal{L}^e_{mpd}(d^q_j)$ for the mth packet's transmission is then maximum-ratio-combined with the corresponding previously detected LLRs stored in the receiver's buffer, when Type-I HARQ is employed before soft decoding. When Type-II HARQ is used, the appropriately concatenated detected LLRs of all the past $K \in [1, L]$ retransmission attempts jointly constitute a codeword, which is then subjected to rate-compatible soft decoding.

Remarks: Instead of superposition coding, multiple packets may be orthogonally multiplexed within a specific transmission attempt without imposing any IPI. However, maintaining orthogonality amongst the packets requires additional DS-spreading of the original channel coded packet,

hence resulting in a rate-loss. Since orthogonal channel codes are hard to design, we may exploit the multiplexing capability inherently provided by channel codes having a rate less than unity [79] by differentiating the layers with the aid of their unique, layer-specific channel codes. Naturally, this is achieved at the cost of an increased complexity and marginal PER performance degradation.

5.3 Transmission Efficiency Metric

A general packet-based wireless network is constituted by a wired link spanning from a server to an Access Point (AP) and a wireless link from the AP to the Mobile Terminal (MT). In this section, we discuss the transmission efficiency metrics to be applied for the link layer and for the TCP layer, respectively, which are the effective throughput η and the mean frame arrival rate λ.

The following simplifying assumptions are stipulated. Firstly, when considering the TCP layer's transmission efficiency, we ignore any propagation delay over the wireless channel and the feedback delay encountered during the transmission of the link layer's error-free ACK and NACK indication. Furthermore, for transmission over the block fading channel considered in this chapter, the packet error events are independently and identically distributed (i.i.d.), resulting in a low probability of TCP timeouts due to consecutive packet loss events. Hence, we ignore the so-called slow-start TCP phase and only consider the congestion-avoidance phase of the widely used TCP Reno [170]. We also assume that the M/G/1 queue [170] has a Poissonian arrival process having an arrival rate of λ, a general i.i.d. service time duration T and a single server.

TCP Reno is one of the protocol variants that is used in order to achieve congestion avoidance. It uses a mechanism ironically referred to as *slow start* in order to increase the congestion window duration after a connection is initialised and after a timeout. When the congestion window duration exceeds a threshold T_{ss}, the algorithm enters a new state, termed as *congestion avoidance*. For every successful transmission, the congestion window is doubled for each RTT. If a frame loss occurred, Reno will halve the congestion window duration and enters a phase called *Fast Recovery*. On the other hand, if a timeout occurs, slow start procedure is activated.

5.3.1 Effective Throughput η

Let us define the normalised effective throughput η as the product of the throughput per packet η_0 and the total number of packets M divided by the total number M_r of transmissions required, yielding $\eta = \eta_0 M/M_r$, where the per-packet throughput is given by:

$$\eta_0(\gamma_b) = r \cdot b \left[1 - p_e(\gamma_b)\right], \tag{5.15}$$

where r and b are the channel coding rate and the number of bits per symbol determined by the modulation scheme employed. Furthermore, p_e denotes the link layer's PER as a function of the

τ	the average Round Trip Time (RTT)
D	propagation delay in wired link from the server to AP
T	expected duration per TCP frame
D_q	expected queueing delay
w	The congestion window size
B	Buffer-size
N	the average number of TCP frames transmitted every RTT τ
λ	mean transmission frame arrival rate

Table 5.1: Notations used in subsection 5.3.2.

SNR per-bit denoted by γ_b. This metric assumes that each packet exhausts all the L retransmissions for the sake of simplified comparisons.

5.3.2 Mean Frame Arrival Rate λ

Before continuing our discussions, the terminology used is summarised in Table 5.1. Our objective in this subsection is to quantify the TCP layer's effective throughput for both the conventional and for the proposed HARQ schemes.

The normalised effective throughput achieved at the TCP layer may be measured by the so-called mean transmission frame arrival rate λ encountered, which is determined by the average number of TCP frames N successfully transmitted within the average Round Trip Time (RTT) τ [171]. In other words, the mean transmission frame arrival rate reflects the average number of TCP frames successfully served within a unit time, i.e. normalised to the average RTT. More explicitly, we have $\lambda = N/\tau$, where $N = 3w/4$ for the Stop-and Wait ARQ regime assumed in the TCP layer [171]. Below, we will elaborate on the effects of two parameters used in the calculation of λ, namely those of the congestion window size w and of the average RTT τ.

5.3.2.1 Average Round Trip Time

The average RTT is given by

$$\tau = 2D + D_q + \mathrm{E}[T], \tag{5.16}$$

where D is the propagation delay in the wired link spanning from the server to the AP, $\mathrm{E}[T]$ is the expected transmission duration of a TCP frame and D_q is the expected queueing delay, which is given by the Pollaczek-Khinchine formula of [171]:

$$D_q = \lambda \mathrm{E}\left[T^2\right] / \left[2(1 - \lambda \mathrm{E}\left[T\right])\right]. \tag{5.17}$$

Let us now derive the first and second moments of the transmission duration T of a TCP frame, i.e. $\mathrm{E}\left[T\right]$ and $\mathrm{E}\left[T^2\right]$ in order to get the average RTT τ.

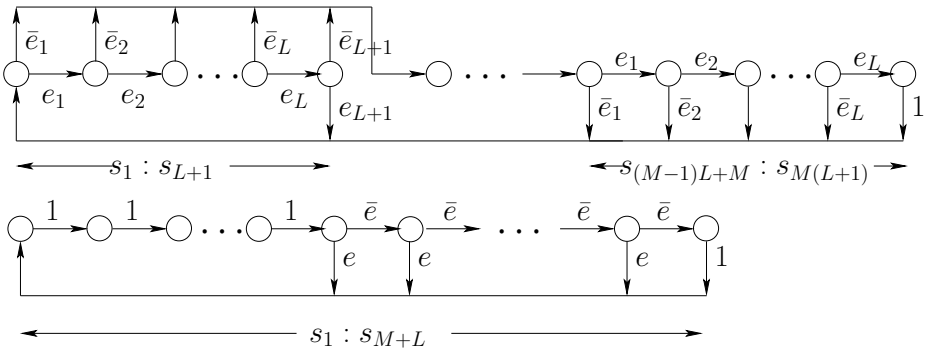


Figure 5.4: The Markov model of both the conventional scheme (top) and of the M-HARQ scheme (bottom).

A TCP frame is deemed to be successfully received only when all of its M link-layer packets are correctly received. If any packet of a TCP frame fails the CRC-test after $(L+1)$ transmissions, this frame will be dropped and the next frame queueing in the buffer of size B frames is to be sent. This procedure may be modelled with the aid of a Markov chain as follows [171–173].

Markov chain model Denote the set of states in the Markov chain by $\mathcal{S}$, where $|\mathcal{S}| = M_r$ stands for the maximum number of link layer packet transmissions within a single TCP frame. Let us consider the conventional HARQ first as seen in the upper part of Fig. 5.4. There are a total of $M_r = M(L+1)$ transmission states and each link layer packet of a TCP frame has a maximum of $(L+1)$ transmissions. Every erroneous packet transmission occurring with a probability of $e_j, j \in [1, L]$ leads from the current state represented by a circle to a state transition within the current packet. Following the final erroneous transmission of the link layer, the entire current TCP frame is dropped and the corresponding state transition leads directly to the next TCP frame represented by s_1 with a probability of e_{L+1}. By contrast, a successful transmission having a probability of $\bar{e}_j, j \in [1, L+1]$ leads to a state transition by-passing the retransmissions and initiating a new packet's transmission.

However, the proposed M-HARQ scheme exhibits different characteristics, as seen in the lower part of Fig. 5.4. There are a total of $M_r = (M+L)$ transmission states and each packet has a maximum of $(L+1)$ transmissions. The first L states are associated with a unity transmission probability, since the maximum number of $(L+1)$ transmissions of the first packet has not been exhausted. Once the $(L+1)$ packet transmissions of the first packet have been exhausted, the state transmission diagram of Fig. 5.4 traverses to the next state corresponding to the next new packet's transmission in the current TCP frame with a probability of $\bar{e}$. By contrast, if none of the $(L+1)$ transmissions were successful, then the state transition evolves to the next new frame with a probability of e.

Transition probability The transition probability $p_c(s, s')$ of the conventional HARQ scheme's Markov chain from state $s \in \mathcal{S}$ to $s' \in \mathcal{S}$ and that of the proposed M-HARQ scheme denoted by $p_p(s, s')$ are given in Table 5.2, where e_j and $\bar{e}_j = 1 - e_j$ denotes the probability of erroneous and

correct link layer packet reception during the jth transmission, respectively. Likewise, $e = \prod_{j=1}^{L+1} e_j$ and $\bar{e} = 1 - e$ denotes the probability of erroneous and correct link layer packet reception after $(L+1)$ transmissions, respectively. These transition probabilities are then incorporated in the transition probability matrix $\mathbf{P}$ of size $|\mathcal{S}| \times |\mathcal{S}|$.

State probability The state probability vector $\mathbf{p}$ hosts the probability of each state $s \in \mathcal{S}$. The initial state probability $\mathbf{p}^0$ is modelled by the entries $p^0(s=1) = 1$ and $p^0(s>1) = 0$. The transition state probability $\mathbf{p}^q$ at the qth transition phase is given by $\mathbf{p}^q = \mathbf{p}^0\hat{\mathbf{P}}$, where we have $q \in [\min(L+1, M), |\mathcal{S}|]$ and $\hat{\mathbf{P}}$ equals to the transition probability matrix $\mathbf{P}$, except that the transition to state $s=1$ corresponding to the first column of $\mathbf{P}$, namely to $\mathbf{z} = \mathbf{P}(:,1)$ is omitted from $\mathbf{P}$ to arrive at $\hat{\mathbf{P}}$. Finally, the steady state probability $\mathbf{p}^s$ is obtained by solving the linear equations $p^s(s) = \sum_{s' \in \mathcal{S}} p^s(s')p(s',s)$ subject to the constraint $\sum_{s \in \mathcal{S}} p^s(s) = 1$, where $p(s',s)$ is given in Table 5.2.

Hence, the first and second moments of T are expressed as

$$\mathrm{E}[T] = \sum_{q=a_2}^{a_1} \mathbf{p}^q \mathbf{z} t_q \tag{5.18}$$

$$\mathrm{E}\left[T^2\right] = \sum_{q=a_2}^{a_1} \mathbf{p}^q \mathbf{z} t_q^2, \tag{5.19}$$

where we have $a_1 = |\mathcal{S}|$ and $a_2 = \min(L+1, M)$. Furthermore, $t_q = qt_0$ stands for the time required for the transmission of q link-layer packets with each packet's transmission duration t_0 set to unity in this chapter.

5.3.2.2 Congestion Window Size

The congestion window size w in the TCP mechanism is employed to control the frame flow by additively increasing its window size every RTT, until either the maximum allowable buffer-size is reached or the congestion is resolved. If so, then the window size is halved. Hence the maximum congestion window size is determined by two design parameters, namely the affordable buffer-size and the acceptable Frame Error Rate (FER).

We refer to a buffer-size limited scenario, when potentially no frame error events are experienced, because the window size of w_B is increased until the buffer is filled. The subsequently arriving TCP frames are then dropped and the window size of w_B is halved in the interest of getting ready for future window expansion to prevent future frame dropping events. The window size w_B is given by $w_B = B + \tau/\mathrm{E}[T] + 1$ [170], where $1/\mathrm{E}[T]$ represents the average TCP frame transmission rate.

On the other hand, if residual frame error events persist after link layer HARQ retransmissions, the window size w_e is determined by the residual FER p_e^f experienced and w_e is halved for the same reason as argued above. The FER p_e^f corresponds to the maximum of $(L+1)$ link layer HARQ transmissions, which is given by $p_e^f = \mathbf{p}^s\mathbf{z}_e/\mathbf{p}^s\mathbf{z}$, where $\mathbf{z}_e$ equals to the vector of $\mathbf{z}$ hosting

$p_c(s,s')$	$\{s,s'\}$
e_j	$\mathcal{S}_1 = \{ij,1\}$ $i \in [1,M-1], j = L+1$
e_j	$\mathcal{S}_2 = \{(i-1)(L+1)+j, s+1\}$ $i \in [1,M], j \in [1,L]$
$\bar{e}_j$	$\mathcal{S}_3 = \{(M-1)(L+1)+j,1\}$ $j \in [1,L]$
$\bar{e}_j$	$\mathcal{S}_4 = \{(i-1)(L+1)+j, i(L+1)+1\}$ $i \in [1,M-1], j \in [1,L+1]$
1	$\mathcal{S}_5 = \{M(L+1),1\}$
0	$s,s' \in \mathcal{S} - \{\mathcal{S}_i\}_{i=1}^{5}$
$p_p(s,s')$	$\{s,s'\}$
e	$\mathcal{S}_1 = \{i,1\}$ $i \in [L+1, M+L-1]$
$\bar{e}$	$\mathcal{S}_2 = \{i, i+1\}$ $i \in [L+1, M+L-1]$
1	$\mathcal{S}_3 = \{M+L,1\}; \mathcal{S}_4 = \{i, i+1\}$ $i \in [1,L]$
0	$s,s' \in \mathcal{S} - \{\mathcal{S}_i\}_{i=1}^{4}$

Table 5.2: The transition probability of the conventional scheme $p_c(s,s')$ from state $s \in \mathcal{S}$ to $s' \in \mathcal{S}$ and of the proposed scheme $p_p(s,s')$.

the transition probabilities leading to state $s = 1$ due to failure. Hence the average number of TCP frames transmitted between two consecutive frame loss events is roughly $1/p_e^f$, which is also approximately given by $3w_e(w_e/2+1)/4$, hence we arrive at the window size of $w_e \approx \sqrt{8/3p_e^f}$.

When considering both the buffer-size limited and the FER limited behaviour, the ultimate window size w is recommended to be set to $w = \min(w_B, w_e)$:

$$w \quad = \quad \min(\sqrt{8/3p_e^f}, B + \tau/\mathrm{E}[T] + 1). \tag{5.20}$$

Based on the value of both the average RTT τ and the congestion window size w, the mean frame arrival rate may be calculated as $\lambda = 3w/4\tau$, where we obtain the value of λ by fixed point iterations.

5.4 Performance Evaluation

5.4.1 PER Investigations

Let us now evaluate the link layer PER performance of our proposed M-HARQ scheme. Fig. 5.5 shows the PER performance of the proposed arrangement against that of the conventional scheme

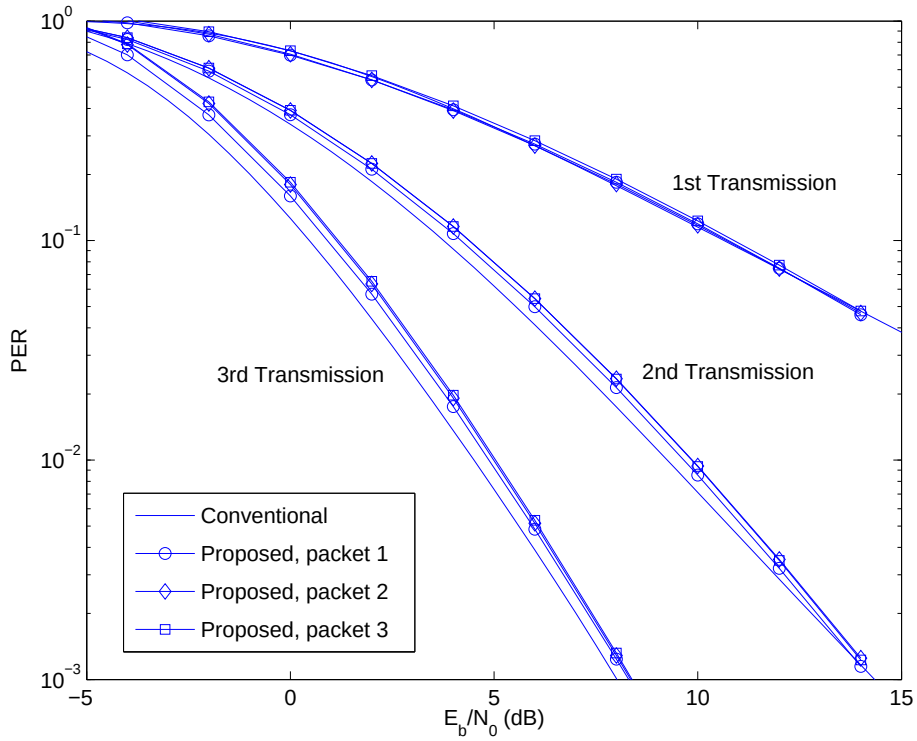


Figure 5.5: The PER performance of all $L+1=3$ transmissions for both the conventional HARQ and for the first $L+1=3$ packets of the proposed M-HARQ scheme.

for a total of $L+1=3$ transmissions employing Type-I HARQ. In practice, a total of two or three transmissions are sufficient, since the HARQ scheme acts like a 'safety net' in support of the link adaptation procedure, which is capable of preventing most of the potential packet loss events. In our simulations, each packet of length $N_i = 256$ bits is QPSK modulated and channel coded by a rate-1/3 irregular systematic RA code [68]. A Rayleigh distributed block-fading channel is used and the feedback channel conveying the NACK indicator is assumed to be error-free. Again, we consider the worst case scenario, where each of the M packets employs the maximum affordable number of $L=2$ retransmissions.

In general, according to the Toeplitz-matrix-like arrangement having L number of retransmissions, we may investigate the PER of all the $(L+1)$ transmissions for each of the first $(L+1)$ packets, since they correspond to different typical interference patterns Ω. For instance, when $L=2$ is considered, the number of layers for each of the $L+1=3$ transmissions of the $L+1=3$ first packets is given by $\Omega_{pck1}=[1,2,3]$, $\Omega_{pck2}=[2,3,3]$ and $\Omega_{pck3}=[3,3,3]$. Fig. 5.5 suggests that during the first transmission the PER performance of our proposed scheme is the same as that of the conventional scheme. By contrast, for two and three transmissions, there is an observable but marginal PER degradation for our proposed scheme compared to that of the conventional one. Apart from this slight difference, all packets experience a near-identical PER performance.

5.4.2 Efficiency Evaluation

Let us now investigate the link layer's effective throughput η for both our proposed M-HARQ scheme and for the conventional scheme. The PER p_e versus the SNR γ_b per bit of both schemes was approximated by a 6th-order polynomial fitted to the simulated curves shown in Fig. 5.5. Then, the normalised effective throughput was calculated and plotted in Fig. 5.6. Observe in Fig. 5.6 the significantly improved effective throughput of our proposed M-HARQ arrangement as compared to that of the conventional one for both $M = 4$ and $M = 12$. When the total number of transmitted packets M is significantly higher than L, the effective throughput η of our proposed scheme approaches that of the single-transmission scenario, which can be verified by comparing the results of both the $L = 1$ and $L = 2$ scenarios corresponding to $M = 4$ and $M = 12$ in Fig. 5.6, where the $L = 0$ curve is printed using the continuous line. This implies that there is only a marginal retransmission delay penalty for our proposed M-HARQ scheme for $M \to \infty$. More explicitly, the delay penalty may be related to the reduction of the effective throughput by a factor of M/M_r, which is $1/(L+1)$ for the conventional scheme and $M/(M+L) \approx 1$ for our proposed scheme for $M \to \infty$.

For example, as illustrated in Fig. 5.6, where point A related to our M-HARQ scheme is calculated as $\eta_A(p_e \approx 0) = r \cdot b[1 - p_e(\gamma_b)]M/(M+L) = 0.44$ and point B related to the conventional scheme is calculated as $\eta_B(p_e \approx 0) = r \cdot b[1 - p_e(\gamma_b)]/(L+1) = 0.22$, where we have $M = 4$, $L = 2$ and $r = 1/3$ is the channel code rate, while $b = 2$ for the QPSK scheme employed and $p_e(\gamma_b = 20) \approx 0$ for both scenarios. This is also true when considering a PER of $p_e = 10^{-1}$, which is of practical relevance, where the effective throughput of both schemes in the above example are weighted by $1 - p_e = 0.9$, which are $\eta_A(p_e = 10^{-1}) = 0.396$ and $\eta_B(p_e = 10^{-1}) = 0.198$ respectively, although the proposed M-HARQ arrangement has a slightly higher required γ_b value due to the superimposed interference, as characterised in Fig. 5.5. Thus, for $M = 4$ and a total of $L = 2$ retransmissions, we have a throughput penalty of 1/3, while a throughput penalty of 2/3 for our proposed M-HARQ scheme.

Fig. 5.7 compares the mean successful frame arrival rate λ recorded for a Poissonian source frame generation process for both the conventional HARQ scheme and for our proposed arrangement at the TCP layer. Two different phenomena may be observed in Fig. 5.7, namely the 'buffer-size limited' and the 'FER limited' situations seen at the right and left of the figure. More explicitly, at low E_b/N_0 values, λ is limited by the high FER imposed by the channel, which activates more retransmissions. On the other hand, λ is limited by the finite buffer-size of B TCP frames, when the FER is low, such as $p_e \leq 0.02$ for our proposed scheme and $p_e \leq 0.003$ for the conventional one. This is clearly seen in Fig. 5.7 upon increasing the buffer-size from $B = 1$ to 20 TCP frames. It can also be seen in Fig. 5.7 that our proposed scheme substantially improves the mean successful TCP frame arrival rate at low γ_b values, before the conventional scheme reaches its buffer-limited maximum. Although both schemes reach the same maximum value of λ bounded by the buffer-size

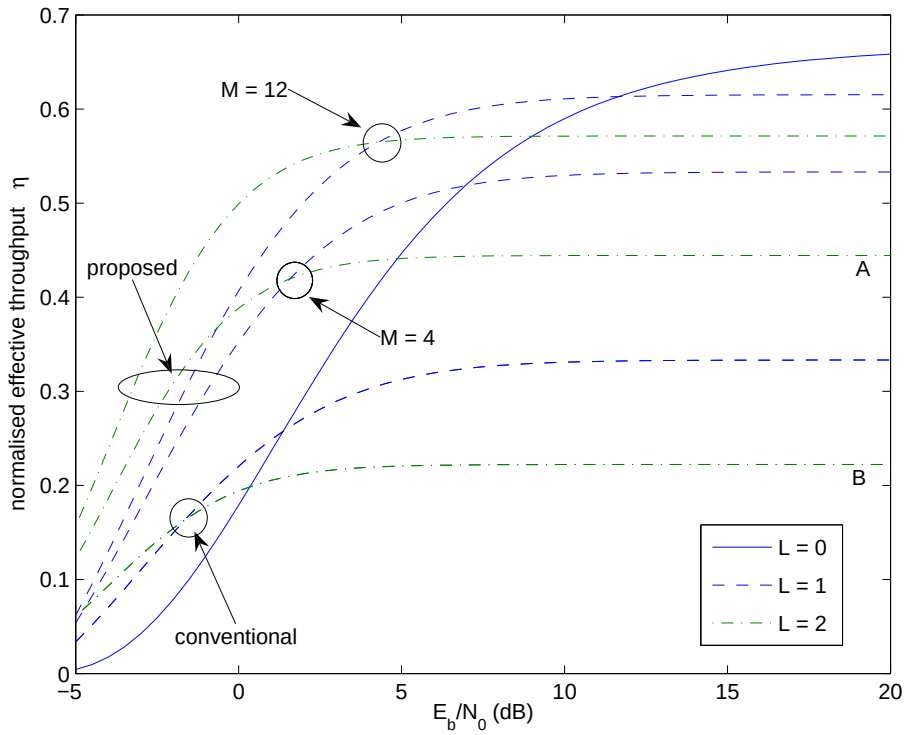


Figure 5.6: The effective throughput of the first and second retransmissions for both the conventional HARQ and the proposed M-HARQ scheme.

B, the proposed M-HARQ arrangement requires lower E_b/N_0 values than the conventional HARQ, when comparing points A and C in Fig. 5.7. This clearly implies that our scheme is more tolerant to frame error events and hence has a higher end-to-end throughput as a benefit of employing the proposed M-HARQ scheme.

5.4.3 Discussion

Let us now discuss both the limitations and the beneficial applications of the M-HARQ arrangement. Our proposed scheme is based on the superposition coding approach and hence the resultant composite packet of multiple superimposed layers becomes effectively 'interference-limited'. Therefore, the per-layer throughput rate should not be excessive in order to ensure that the decoded PER remains low and approaches the single-layer best-case performance, as illustrated in Fig. 5.5. More explicitly, this requirement discourages the employment of high-throughput, but interference-sensitive high-PER, high-order modulation schemes, although sophisticated MPD algorithms may be employed to relax this requirement, provided that the complexity imposed remains affordable.

Furthermore, relatively low-rate channel codes are preferred for the sake of supporting the low PER transmission of multiple superimposed layers at a near-single-layer PER performance. Since the number of retransmissions L is typically low in practice, so is the number of superimposed layers. Hence for example different channel codes and/or interleavers may be used to separate the layers using the principles of channel code aided [37] or interleave division multiplexing [31].

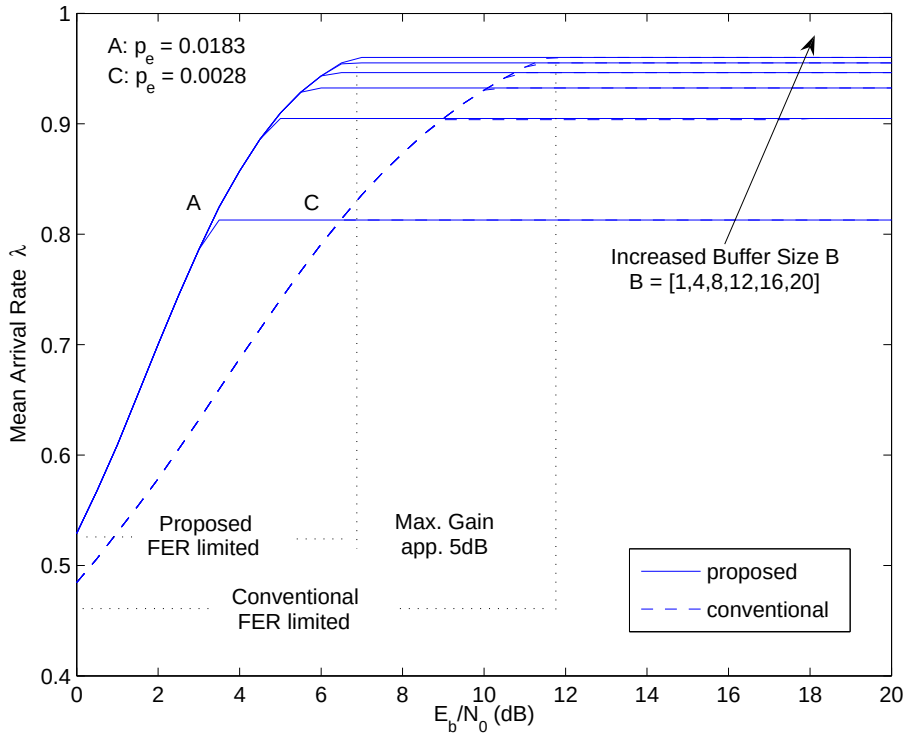


Figure 5.7: The mean arrival rate for both the conventional HARQ and for the proposed M-HARQ scheme.

Alternatively, orthogonal spreading sequences can be employed for separating the layers at the cost of reducing their effective throughput proportionately to the spreading factor. Furthermore, in our investigations, only a Type-I HARQ scheme was employed based on the argument that Type-II HARQ provides a limited extra gain over Type-I HARQ for low-order modulation and low-rate channel coding [159], although the employment of Type-II HARQ is also straightforward.

Remarks: According to Fig. 5.6 and Fig. 5.7, our scheme achieves a significantly higher effective link layer throughput η and TCP layer mean frame arrival rate λ than the conventional scheme, which makes our scheme particularly suitable for delay-sensitive low-rate applications, while providing cell-edge users with an improved end-to-end throughput and transmission integrity.

5.5 Conclusion

In this chapter, as a practical application of the concept of non-orthogonal random waveform based multiuser communications, a novel superposition-aided multiplexed HARQ scheme was proposed, which jointly encodes the current new packet to be transmitted and any packets that are about to be retransmitted. The proposed scheme is capable of substantially improving the link layer's effective throughput for all transmitted packets at a marginal PER performance degradation. As a result, this improved link layer transmission efficiency also contributed towards an improved overall end-to-end TCP transmission efficiency.

Furthermore, our superposition coding aided arrangement may be readily integrated with existing systems without substantially modifying the current design. Although only Type-I HARQ was employed, the extension to Type-II HARQ is also straightforward. However, the proposed scheme discourages the employment of high-order modulation schemes and favours relatively low-rate channel codes. Hence, the proposed scheme is particularly suitable for delay-sensitive, low-rate services and for providing cell-edge users with an improved end-to-end throughput and/or transmission integrity.

5.6 Appendix - Pollaczek-Khintchine Formula

Let us assume that the packets arrive at an AP according to a Poissonian packet arrival process with a rate of λ. The service time T of packets are i.i.d. random variables having a common PDF and first as well as second moments of $\mathbb{E}[T]$ and $\mathbb{E}T^2]$. There is a single server and sufficient storage for packets to be queued, when finding the server busy. Each packet waits, until it can be processed. The server can handle only one customer at a time. This particular queueing model is referred to as the M/G/1 queue in the form of Kendall notation, which is used to describe and classify the queueing model, where 'M' represents the Poissonian packet arrival process, 'G' represents the i.i.d. service time distribution and '1' stands for one service channel (server) [171].

We define the *service load* as $\rho = \lambda\mathbb{E}[T]$ and assume that we have $\rho < 1$. Alternatively, we define the *average service rate* as $\lambda_s = 1/\mathbb{E}[T]$ and thus, the service load is given by $\rho = \lambda/\lambda_s$. For a stable system the average packet arrival rate λ should be lower than the average service rate λ_s, namely $\rho < 1$. Hence, the service load can be interpreted as the long-term fraction of time that the server is busy.

The Pollaczek-Khintchine formula states that the long-term *average delay* D_q per packet spent in the queue is given by [171]:

$$D_q = \lambda\mathbb{E}[T^2]/[2(1-\rho)]. \tag{5.21}$$

This formula quantifies D_q and more importantly, it provides further useful insights. Denoting the squared normalised variance of the service time by $c_s^2 = \sigma^2(T)/(\mathbb{E}[T])^2$, we can formulate a more insightful expression for the Pollaczek-Khintchine formula as:

$$D_q = (1+c_s^2)\rho\mathbb{E}[T]/[2(1-\rho)]. \tag{5.22}$$

The above formula shows that the average delay per packet is proportional to the factor $(1+c_s^2)/2$, when the average packet arrival rate λ and the mean service time $\mathbb{E}[T]$ are kept fixed.

Differentiation of D_q as a function of ρ shows that the slope of increase of D_q as a function of ρ is proportional to $(1-\rho)^{-2}$. Hence, a small increase in the average packet arrival rate at a load of $\rho = 0.9$ causes an increase in the average delay, which is about 25 times higher than it would be for a load of $\rho = 0.5$. This demonstrates the danger of designing a stochastic system with too

high a utilisation level, since a small increase in the service load will in general cause a dramatic degradation in the system’s performance.

Chapter 6

Harmony Search Aided Iterative Joint Channel Estimation, Multiuser Detection and Channel Decoding for DS-CDMA

All the detection algorithms employed in our previous chapters are low-complexity IC schemes. However, when a large search space is encountered for example in case of multiple antennas, high order modulation schemes or heavily-loaded multiuser systems, more sophisticated, yet low-complexity detection algorithm will be required.

In this chapter, a novel MUD scheme is proposed for DS-CDMA systems employing the so-called HS algorithm, which is a novel meta-heuristic optimisation method. We specifically design the HS aided MUD for the communications problem considered and apply it in an iterative joint Channel Estimation (CE), MUD and channel decoding framework. The simulation results demonstrate that a near-single-user performance can be achieved by the proposed algorithm while avoiding the excessive-complexity full-search-based optimum detector even in the context of DS-CDMA systems supporting more users than the number of chips per sequence. Moreover, the HS algorithm can be efficiently incorporated in an EM based CE framework. As an example, our candidate system using $N_c = 7$-chip random sequences is capable of supporting six times more users than the number of chips per spreading code at a BER of 10^{-5}, while approaching the single-user performance within about 1dB.

6.1 Introduction

Coded MIMO multiuser systems are typical scenarios in the context of next generation wireless communications employed for the sake of achieving a high spectral efficiency, which led to high-

dimensional adaptive modulation and coding in the HSPA system [2]. Near-single-user detection in such a high-dimensional and highly-correlated space is a non-trivial task. Inspired by the turbo principle [64], optimum MAP based Joint Detection and Decoding (JDD) scheme can be achieved by reduced-complexity Iterative Detection and Decoding (IDD) upon exchanging extrinsic information between the receiver components without a significant performance compromise [116]. The associated decoupling of the excessive-complexity joint MAP receiver into components is facilitated by the inclusion of an interleaver employed between the detector and decoder, which randomises and decorrelates the soft information exchanged between them during the iterations. Furthermore, the interleavers are expected to have a high depth [64] and be data-specific [31]. Early work on IDD can be found for example in [74, 75, 116, 117]. Similar to the analysis of LDPC codes [17] and serial concatenated codes [18], the convergence of IDD schemes is typically investigated by semi-analytical methods, such as the DE charts [117, 174] and EXIT charts [36, 112].

In the context of the MUD scheme proposed in this chapter, the optimum Bayesian MAP detector has an excessive computational complexity. One may either use the classic MMSE detector [75] or develop reduced complexity detection algorithms, such as the well-known Sphere Decoding (SD) aided detector [175] and the Markov Chain Monte Carlo (MCMC) aided detector [176, 177]. Alternatively, stochastic global optimisation techniques may be pursued in order to reduce the complexity, while still capturing the MAP solution with a high probability using for example the class of Evolutionary Algorithms (EAs) [28], such as the well-known Genetic Algorithms (GA) [178]. Another class of algorithms, which also captured substantial research attention is constituted by the Swarm Intelligence (SI) algorithms, such as Particle Swarm Optimisation (PSO) [179] and Ant Colony Optimisation (ACO) [180, 181]. Apart from these algorithms, imitating the improvisation process of musicians, a new naturally-inspired meta-heuristic optimisation method was proposed recently, leading to the so-called HS algorithm [47].

In addition to detection and decoding, the receiver also has to acquire sufficiently accurate CE [178, 182]. Conventional pilot-aided CE suffers from an extra overhead in terms of an increased bandwidth and power. Furthermore, stand-alone ML suffers from the lack of extrinsic information, which would be readily available from the outer channel decoder. Hence, the data-aided CE may be incorporated in the aforementioned IDD framework, so that the estimated channel parameters can be iteratively improved by the error correction decoder. In particular, the well-established EM algorithm [183] provides a general framework for solving such problems. In this chapter, the key idea of our HS aided EM algorithm is that instead of using the true *a posteriori* distribution of the transmitted data, given the observation and the current channel estimate, we use the *a posteriori* distribution based on the data estimates provided by the HS algorithm.

The novel contribution of this chapter is that *we design a HS based global optimisation method for iterative joint CE, MUD and channel decoding. The simulation results demonstrate that the single-user performance can be approached by our specifically designed HS algorithm without the excessive complexity of the MAP based MUD even in the challenging scenario of supporting more*

users than the number of chips per sequence. Furthermore, the HS-aided MUD is amalgamated with the EM based CE framework. Finally, the proposed algorithm is sufficiently general to be applicable to the linear Gaussian vector model employed for MIMO detection.

The remainder of this chapter is organised as follows. In Section 6.2, we highlight the general approach of EA and review the original HS algorithm. In Section 6.3 we specifically design the HS algorithm for our uncoded DS-CDMA MUD problem and develop it in the context of iterative joint CE, MUD and channel decoding in Section 6.4. We then characterise its performance in Section 6.5 and conclude in Section 6.6.

6.2 Preliminaries of Harmony Search Algorithm

6.2.1 Problem Formulation and the Routine of EAs

Consider the optimisation of $f(\mathbf{x})$ subject to $x_k \in \mathcal{A}_k, k \in [1, K]$, where $f(\cdot)$ is the *Fitness Function* (FF), $\mathbf{x} = [x_1, \ldots, x_K]^{\mathrm{T}}$ is the set of *candidates* containing K legitimate variables chosen from a discrete or continuous alphabet $\mathcal{A}_k, k \in [1, K]$, where the superscript $(\cdot)^{\mathrm{T}}$ denotes transpose. In our MUD problem, K is the number of users and $\mathbf{x}$ hosts one of the 2^K possible legitimate K-user BPSK modulated candidate vector.

The general routine of employing EAs may be summarised as follows. A population of K-user candidates is randomly initialised, where each one represents a potential K-user MUD solution. The 'quality' of each K-user candidate is evaluated using the FF. Then a selection process is applied during each iteration in order to form a new population, where the selection process tends to favour high-quality K-user candidate. Two tasks are distinguished in this selection process. The *exploitation* step provides the algorithm with the capability of improving and combining the traits of the current K-user candidate, while the *exploration* step prevents the candidates from premature convergence to a potentially local rather than global optimum, without exploring sufficiently diverse regions of the entire search area, i.e. different K-user vectors. These two steps jointly maintain a sufficiently diverse K-user candidate population to avoid getting trapped in local optima. This procedure is repeated, until the affordable complexity-budget is exhausted.

6.2.2 The Harmony Search Algorithm

We now briefly review the HS algorithm introduced in [47] and relate it to our K-user MUD context in Section 6.3. When a musician improvises, the aesthetic quantification (FF) results from a set of pitches produced by the music instruments (variables) involved. The musician seeks to produce aesthetically pleasing harmony (the optimum K-user vector) as determined by his/her aesthetic perception (FF) inferred from rehearsals (iterations).

There is a range of parameters associated with the HS algorithm. The harmony memory size M

specifies the number of initial K-user harmony candidates stored in the Harmony Memory Matrix (HMM), which represents the initial population size. The harmony memory activation probability P_{ma} specifies the typically less than unity probability of a new K-user candidate being selected from the HMM rather than randomly, where the latter has a probability of $(1 - P_{ma})$. The pitch adjustment probability P_{pa} specifies the chance of changing the pitch, i.e. that of opting for a newly generated K-user candidate from the HMM, rather than using variants of the previous one from the HMM. Finally, Q represents the total number of iterations or improvisations carried out throughout the HS algorithm. More specifically, the HS algorithm can be summarised as follows:

6.2.2.1 Step 1 - Initialisation

A set of M initial K-user harmony candidates is generated and stored in the HMM. More explicitly, to initialise the HMM $\mathbf{X}^0 = [\mathbf{x}_1^0, \dots \mathbf{x}_M^0]$, each variable $x_{m,k}^0, m \in [1, M], k \in [1, K]$ is generated either randomly from a uniform distribution based on its legitimate alphabet $\mathcal{A}_k$ or from a mutated version of a rough initial value, such as the MMSE estimate, where $x_{m,k}^0$ denotes the kth variable of the mth K-user harmony candidate vector $\mathbf{x}_m^0$ in the HMM.

6.2.2.2 Step 2 - Improvisation

At iteration $q \in [1, Q]$, the new K-user harmony candidate $\mathbf{x}^{new}$ is generated using an appropriate combination of the following HS operations: *memory activation*, *pitch adjustment* and *random selection*. More particularly, the new K-user harmony candidate of the qth iteration is either selected from the HMM $\mathbf{X}^{q-1}$ of the previous iteration with a probability of P_{ma} or chosen randomly from the legitimate alphabet with a probability of $(1 - P_{ma})$.

Once a K-user candidate was selected from $\mathbf{X}^{q-1}$, then according to [47] a further pitch adjustment characterised by a step of Δ may be applied with a pitch adjustment probability of P_{pa}, where the specific value of each variable $x_k^{new}, k \in [1, K]$ of the new K-user harmony candidate is tuned to match the neighbouring values in its legitimate candidate solution alphabet $\mathcal{A}_k$. We emphasise in advance that this is a challenging task, when incorporating the HS algorithm into the MUD problem having a binary alphabet, since the original fine-tuning technique of [47] would entail binary toggling of the related bit and hence lead to a dramatically different K-user counterpart. Our solution to this problem will be proposed in the next section.

Again, the probabilities P_{ma} and P_{pa} control the exploration and exploitation phase of the optimisation algorithm, respectively and the corresponding pseudo-code of this procedure is seen in Table 6.1.

Table 6.1: Pseudo-code of the improvisation step

for $k = 1, \ldots, K$ **do**
  if $\mathcal{U}(0,1) \leq P_{ma}$ **do**
    $x_k^{new} = x_{m_i,k}^{q-1}, m_i \sim \mathcal{U}[1, M]$
    if $\mathcal{U}(0,1) \leq P_{pa}$ **do**
      $x_k^{new} = x_k^{new} \pm \Delta \in \mathcal{A}_k$
    end if
  else $x_k^{new} \in \mathcal{A}_k$
  end if
end for

6.2.2.3 Step 3 - Updating

The new harmony candidate $\mathbf{x}^{new}$ generated replaces the worst harmony of the HMM $\mathbf{X}^{q-1}$, provided that its score measured in terms of the FF is better than that of the worst harmony in $\mathbf{X}^{q-1}$. Otherwise, no changes are made in the HMM, namely we have $\mathbf{X}^q = \mathbf{X}^{q-1}$.

6.3 Harmony Search Aided Multiuser Detection of Uncoded DS-CDMA

6.3.1 System Model and the Fitness Function

Let us now first apply the HS algorithm in our MUD problem of an uncoded DS-CDMA system. Consider an uncoded BPSK modulated K-user DS-CDMA system employing user-specific N_c-chip spreading sequences. The canonical linear Gaussian system model of the nth symbol can be written as:

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{n}, \tag{6.1}$$

where $\mathbf{y} \in \mathcal{R}^{N_c \times 1}$, $\mathbf{x} \in \mathcal{R}^{K \times 1}$ and $\mathbf{C} \in \mathcal{R}^{N_c \times K}$ are the N_c-chip received sample vector where each chip-value is constituted by the superposition of the K users' chips, the K-user transmitted symbol vector and the DS-CDMA spreading matrix, respectively as visualised in Fig 9.9 of [28]. Furthermore, $\mathbf{n} \in \mathcal{R}^{N_c \times 1}, n_i \sim \mathcal{N}(0, \sigma^2)$ is the AWGN vector.

Given the knowledge of the spreading matrix $\mathbf{C}$ at the receiver, the objective of the receiver is to maximise the joint APP $P(\mathbf{x}|\mathbf{y})$, which is given by:

$$\begin{aligned} P(\mathbf{x}|\mathbf{y}) &= p(\mathbf{y}|\mathbf{x})P(\mathbf{x})/p(\mathbf{y}) \\ &\propto p(\mathbf{y}|\mathbf{x}). \end{aligned} \tag{6.2}$$

This is proportional to the likelihood function $p(\mathbf{y}|\mathbf{x})$, where we ignored the irrelevant part of $p(\mathbf{y})$ and assumed having an identical *a priori* probability $P(\mathbf{x})$ for each legitimate transmitted symbol

as far as the uncoded system is considered. The likelihood function $p(\mathbf{y}|\mathbf{x})$ may then be expressed as:

$$p(\mathbf{y}|\mathbf{x}) \propto \exp\left\{-||\mathbf{y}-\mathbf{C}\mathbf{x}||^2/2\sigma^2\right\}. \tag{6.3}$$

In this case, the MAP criterion reduces to the ML criterion. Hence, the FF equals to the likelihood function and we aim at finding the solution:

$$\mathbf{x}^* = arg\min_{\mathbf{x}\in\mathcal{A}} ||\mathbf{y}-\mathbf{C}\mathbf{x}||^2, \tag{6.4}$$

where the cardinality of the set is given by $\mathcal{A} = \prod_{k=1}^{K}\mathcal{A}_k$, which is 2^K for the BPSK scheme employed. Hence the complexity imposed by the ML detector is exponentially increased with K. Therefore, we resort to an alternative near-ML optimisation method, namely to the HS aided MUD, which evaluates Eq. (6.4) only for a fraction of the 2^K candidates and yet captures the ML solution with a high probability.

6.3.2 Harmony Search Aided Multiuser Detection

6.3.2.1 Naive Transplanting

We initialise our specific HS algorithm by randomly generating M number of K-user candidate harmony vectors $\mathbf{X}^0 = [\mathbf{x}_1^0, \ldots \mathbf{x}_M^0]$ from the set of 2^K legitimate solutions. At each iteration q, a new K-user harmony vector $\mathbf{x}^{new}$ is generated either from the HMM $\mathbf{X}^{q-1}$ with a probability of P_{ma} according to $x_k^{new} = x_{m_i,k}^{q-1}, m_i \sim \mathcal{U}[1,M], k \in [1,K]$, which implies inheriting the kth bit of one of the M candidate vectors in the $(q-1)$th iteration, $k \in [1,K]$ or randomly from the alphabet $x_k^{new} \in \mathcal{A}_k = \{\pm 1\}$ with a probability of $(1-P_{ma})$. The newly generated K-user harmony vector replaces the harmony vector of the worst FF in $\mathbf{X}^{q-1}$, provided it is fitter than the worst one, otherwise it will be dropped. Then we continue by setting $\mathbf{X}^q = \mathbf{X}^{q-1}$ for the next iteration.

However, the above direct employment of the HS aided MUD results in a poor performance, since in the particular MUD problem only binary data of $\{\pm 1\}$ are considered, which is incompatible with the pitch adjustment step of the original HS proposal. Thus, instead of 'tuning' the kth binary value of x_k^{new} to its opposite based on the predefined pitch adjustment probability of P_{pa} in an binary manner, we propose to take the soft information associated with x_k^{new} into consideration for determining P_{pa}. As a result, the specific value of x_k^{new} will be toggled to the opposite binary value during each pitch adjustment step based on the probability containing its soft information.

6.3.2.2 Pitch Adjustment

More explicitly, after randomly selecting a K-user base harmony vector $\mathbf{x}^b$ at iteration q from the HMM $\mathbf{X}^{q-1}$ with a probability of P_{ma}, the pitch adjustment is carried out by generating the kth variable x_k^{new} based on the marginal APP $P(x_k^b|\mathbf{y},\mathbf{x}_{-k}^b)$ of the base harmony vector, where $\mathbf{x}_{-k}^b$ denotes the K-user base vector with the kth element x_k^b excluded. This marginal APP represents

the soft information of x_k^{new} and acts as the replacement of P_{pa} in the original HS proposal [47]. An additional benefit of this is that the marginal APP is automatically updated after each iteration, rather than being used unaltered throughout the MUD process, as it originally set before the optimisation. The Log-Likelihood Ratio (LLR) of this marginal APP may be conveniently evaluated in the following way:

$$\begin{aligned}\mathcal{L}_{pa} &= \ln \frac{P(x_k^b = +1|\mathbf{y}, \mathbf{x}_{-k}^b)}{P(x_k^b = -1|\mathbf{y}, \mathbf{x}_{-k}^b)} \\ &\propto \ln \frac{p(\mathbf{y}|x_k^b = +1, \mathbf{x}_{-k}^b)}{p(\mathbf{y}|x_k^b = -1, \mathbf{x}_{-k}^b)} \\ &= (||\mathbf{y} - \mathbf{C}\mathbf{x}_{k-}^b||^2 - ||\mathbf{y} - \mathbf{C}\mathbf{x}_{k+}^b||^2)/2\sigma^2, \end{aligned} \tag{6.5}$$

where the subscript $+/-$ indicates the sign of the kth element x_k^b of the K-user base vector. Note that in Eq. (6.5) we dropped the identical *a priori* probability of $P(x_k^b)$. Then we have:

$$P_{pa}(x_k^{new} = +1) = \frac{1}{1 + e^{-\mathcal{L}_{pa}}}. \tag{6.6}$$

In this case, the P_{ma}-based *memory activation* and the P_{pa}-based *pitch adjustment* merge into a single joint step and the exploration of a sufficiently diverse range of solutions is ensured by using both a *random selection* of the K-user harmony candidate vector with a probability of $(1 - P_{ma})$ and by a randomly selected K-user base vector from the HMM at each iteration. The remainder of the algorithm follows the same procedure as discussed in Section 6.2 and the pseudo-code of our HS-aided MUD algorithm may be seen in Table 6.2.

6.4 Harmony Search Aided Iterative Receiver

6.4.1 System Model and Optimum Soft Multiuser Detection

We now proceed to extend our HS-aided MUD in the context of the entire iterative receiver, leading to the soft HS-aided MUD. In addition to Eq. (6.1), the system model becomes:

$$\mathbf{Y} = \mathbf{CHX} + \mathbf{N}, \tag{6.7}$$

where $\mathbf{Y} \in \mathcal{C}^{N_c \times N}$, $\mathbf{X} \in \mathcal{R}^{K \times N}$ and $\mathbf{N} \in \mathcal{C}^{N_c \times N}$ denotes the matrix of received signal samples, transmitted symbols and noise samples of a given transmission frame having a block length of N bits. Furthermore, $\mathbf{H} = \text{diag}\,[h_1, \ldots, h_k]$ contains the block-invariant complex channel to be estimated. Furthermore, R and N_i represent the channel code rate and information frame length, hence we have $N = N_i/R$.

Turbo-style iterative CE, MUD and channel decoding is employed as seen in Fig. 6.1, where each CE update will be based on the soft output of the IDD scheme and the new CE results will again be used by the IDD. We will first design the Soft-HS (SHS) aided MUD algorithm assuming perfect channel knowledge and then demonstrate that our SHS-aided MUD algorithm can be efficiently applied in the EM based channel estimation framework.

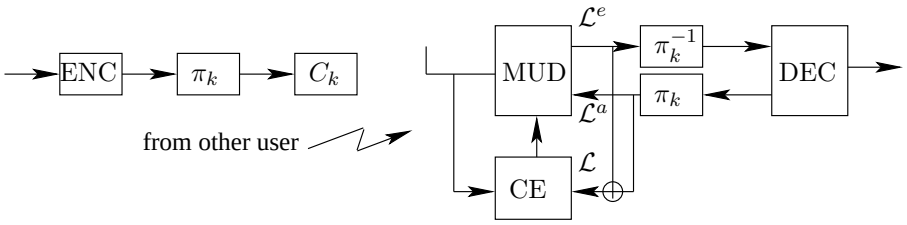


Figure 6.1: Iterative channel estimation, multiuser detection and channel decoding.

Assuming that a long interleaver is employed between the MUD and channel decoder, which decorrelates the soft information exchanged during the consecutive iterations, we focus on the nth symbol interval $\mathbf{y}_n = \mathbf{C}\mathbf{H}\mathbf{x}_n + \mathbf{n}$ and drop the symbol index n. In the BPSK modulated system, the MUD delivers soft information in terms of *extrinsic* LLRs denoted by $\mathcal{L}^e(\mathbf{x})$ to the outer channel decoder based on the observation of the input *a priori* LLRs denoted by $\mathcal{L}^a(\mathbf{x})$. Since all users' information is independent of each other, the *extrinsic* LLR of the kth user is given by:

$$\mathcal{L}^e(x_k) = \mathcal{L}(x_k) - \mathcal{L}^a(x_k), \tag{6.8}$$

where the *a posteriori* LLR $\mathcal{L}(x_k)$ is given by:

$$\begin{aligned} \mathcal{L}(x_k) &= \ln \frac{P\left[x_k = +1|\mathbf{y}, \mathcal{L}^a(\mathbf{x})\right]}{P\left[x_k = -1|\mathbf{y}, \mathcal{L}^a(\mathbf{x})\right]} \\ &= \ln \frac{\sum_{\forall \mathbf{x}_{-k}} P\left[x_k = +1, \mathbf{x}_{-k}|\mathbf{y}, \mathcal{L}^a(\mathbf{x})\right]}{\sum_{\forall \mathbf{x}_{-k}} P\left[x_k = -1, \mathbf{x}_{-k}|\mathbf{y}, \mathcal{L}^a(\mathbf{x})\right]}, \end{aligned} \tag{6.9}$$

noting that the optimum Bayesian approach aided summation leads to the consideration of all possible 2^{K-1} such vectors. Obviously, this potentially prohibitive complexity should be avoided. In the following, we will show that the SHS algorithm is capable of providing a low-complexity alternative.

6.4.2 Soft Harmony Search Aided Multiuser Detection

The idea of the SHS algorithm is that we gather a sufficient number of K-user candidate vectors after Q iterations in the final HMM, each contributing significantly (in enchanting harmony) to the evaluation of the overall summation, where the significance is quantified by the FF. This in principle approximates the summation of all legitimate K-user vectors in terms of the Bayesian optimum of Eq. (6.9). Compared to the hard HS-aided MUD of Section 6.3, the following three aspects have to be modified in order to formulate the SHS algorithm.

6.4.2.1 Fitness Function

The FF to be evaluated in the SHS is defined as the joint APP of the K-user transmitted vector $\mathbf{x}$ based on the observation $\mathbf{y}$ and the *a priori* LLRs $\mathcal{L}^a(\mathbf{x})$ provided by the channel decoder, which

may be expressed as:

$$\begin{aligned} f(\mathbf{x}) &= P\left[\mathbf{x}|\mathbf{y}, \mathcal{L}^a(\mathbf{x})\right] \\ &\propto p(\mathbf{y}|\mathbf{x}) P\left[\mathbf{x}|\mathcal{L}^a(\mathbf{x})\right] \\ &\propto \exp\left\{-||\mathbf{y}-\mathbf{CHx}||^2/2\sigma^2\right\} \prod_{\forall k} P\left[x_k|\mathcal{L}^a(x_k)\right] \\ &= -||\mathbf{y}-\mathbf{CHx}||^2/2\sigma^2 + \sum_{\forall k} \ln P^a(x_k), \end{aligned} \tag{6.10}$$

where in the third line, we exploited the fact that each user's bits are independent of each other, leading to $P(\mathbf{x}|\mathcal{L}^a(\mathbf{x})) = \prod_{\forall k} P\left[x_k|\mathcal{L}^a(x_k)\right]$, while in the last line, we exploited the monotonic nature of the FF, leading to a numerically efficient log-domain expression. Moreover, $P^a(x_k)$ of Eq. (6.10) denoting the *a priori* probability of each user x_k may be expressed as:

$$P^a(x_k = +1) = \frac{1}{1 + e^{-\mathcal{L}^a(x_k)}}. \tag{6.11}$$

6.4.2.2 Pitch Adjustment

The relation between the LLR $\mathcal{L}_{pa}$ of the marginal APP and the *a priori* information $\mathcal{L}^a(x_k)$ is revealed upon expanding Eq. (6.5) by taking into account the *a priori* information $\mathcal{L}^a(x_k)$:

$$\begin{aligned} \mathcal{L}_{pa} &= \ln \frac{P\left[x_k^b = +1|\mathbf{y}, \mathbf{x}_{-k}^b, \mathcal{L}^a(x_k)\right]}{P\left[x_k^b = -1|\mathbf{y}, \mathbf{x}_{-k}^b, \mathcal{L}^a(x_k)\right]} \\ &= \ln \frac{p(\mathbf{y}|x_k^b = +1, \mathbf{x}_{-k}^b)}{p(\mathbf{y}|x_k^b = -1, \mathbf{x}_{-k}^b)} + \mathcal{L}^a(x_k). \end{aligned} \tag{6.12}$$

This value is then used in Eq. (6.6).

6.4.2.3 Soft Output

After Q iterations, a list of K-user candidate vectors with reasonably good fitness was generated. Then Eq. (6.9) is evaluated based on the K-user candidate vectors in the final HMM $\mathbf{X}^Q$, which can be expressed as:

$$\begin{aligned} \mathcal{L}(x_k) &= \ln \frac{\sum_{\mathbf{x}_{-k} \in \mathbf{X}^Q} P\left[x_k = +1, \mathbf{x}_{-k}|\mathbf{y}, \mathcal{L}^a(\mathbf{x})\right]}{\sum_{\mathbf{x}_{-k} \in \mathbf{X}^Q} P\left[x_k = -1, \mathbf{x}_{-k}|\mathbf{y}, \mathcal{L}^a(\mathbf{x})\right]} \\ &= \ln \sum_{\mathbf{x}_{-k} \in \mathbf{X}^Q} \exp\left\{f(x_k = +1, \mathbf{x}_{-k})\right\} - \\ &\quad \ln \sum_{\mathbf{x}_{-k} \in \mathbf{X}^Q} \exp\left\{f(x_k = -1, \mathbf{x}_{-k})\right\} \qquad (6.13) \\ &\approx \max_{\mathbf{x}_{-k} \in \mathbf{X}^Q} f(x_k = +1, \mathbf{x}_{-k}) - \\ &\quad \max_{\mathbf{x}_{-k} \in \mathbf{X}^Q} f(x_k = -1, \mathbf{x}_{-k}), \qquad (6.14) \end{aligned}$$

where the above approximation exploited the Jacobian logarithm $\ln(e^a + e^b) = \max(a, b) + \ln(1 + e^{-|a-b|})$. Eq. (6.14) represents the LLRs of the soft APP output of considering all the

K-user candidate vectors in the final HMM. When considering the *extrinsic* LLR $\mathcal{L}^e(x_k)$, the corresponding *extrinsic* FF value $f^e(\mathbf{x})$ is substituted in Eq. (6.14), which may be given by $f^e(\mathbf{x}) = f(\mathbf{x}) - \ln P^a(x_k)$.

We may seek a further reduced-complexity sub-optimum output by considering only the best harmony vector $\mathbf{x}^* = \max_{\mathbf{x}\in\mathbf{X}^Q} f(\mathbf{x})$ of the Qth iteration. In this case, Eq. (6.9) becomes:

$$\mathcal{L}(x_k) = \ln \frac{P\left[x_k = +1, \mathbf{x}^*_{-k}|\mathbf{y}, \mathcal{L}^a(\mathbf{x})\right]}{P\left[x_k = -1, \mathbf{x}^*_{-k}|\mathbf{y}, \mathcal{L}^a(\mathbf{x})\right]}, \tag{6.15}$$

and we will use Eq. (6.15) as the soft APP output of our SHS-aided MUD.

6.4.3 Soft Harmony Search Aided Channel Estimation

Let us now demonstrate that our SHS-aided MUD algorithm can be efficiently applied in the EM based CE framework, leading to the new joint iterative CE, MUD and channel decoding. In this case, the entire observation frame of received samples $\mathbf{Y}$ is considered rather than just a single symbol interval $\mathbf{y}_n, n \in [1, N]$. Let us now introduce the notation $\mathbf{h} = [h_1, \ldots, h_k]^{\mathrm{T}}$ and recall from Eq. (6.7) that the ML estimate of the channel conditioned on the observation of $\mathbf{Y}$ is given by:

$$\mathbf{h}^* = arg \max_{\mathbf{h}} p(\mathbf{Y}|\mathbf{h}). \tag{6.16}$$

However, the explicit expression of the likelihood function $p(\mathbf{Y}|\mathbf{h})$ is unknown, since the transmitted data $\mathbf{X}$ is unknown. We thus resort to the so-called EM algorithm [183], which iteratively finds the ML solution for $\mathbf{h}$.

Since we do not know the transmitted data $\mathbf{X}$, we cannot evaluate the exact value of the likelihood function $p(\mathbf{Y}|\mathbf{h})$. However, given the channel-contaminated received data $\mathbf{Y}$ that we do know, we can find *a posteriori* estimates of the probabilities for the transmitted data $\mathbf{X}$, given the previously estimated value of $\mathbf{h}^{q-1}$. Then, for each detected set of $\mathbf{X}$, we can thus calculate an expected value of the likelihood function with the aid of the channel-contaminated received data $\mathbf{Y}$ and the previous channel estimate. In the language of EM, we may refer to $\mathbf{Y}$ and $\mathbf{Z} = \{\mathbf{Y}, \mathbf{X}\}$ as the *incomplete data* and the *complete data*, respectively, while $\mathbf{X}$ is referred to as the *missing data*.

6.4.3.1 The Expectation Step

More formally, we commence from the following expectation calculation:

$$\begin{aligned} \Omega(\mathbf{h}, \mathbf{h}^{q-1}) &= \mathbb{E}\left[\ln p(\mathbf{X}, \mathbf{Y}|\mathbf{h})|\mathbf{Y}, \mathbf{h}^{q-1}\right] \\ &= \sum_{\forall \mathbf{X}} P(\mathbf{X}|\mathbf{Y}, \mathbf{h}^{q-1}) \ln p(\mathbf{X}, \mathbf{Y}|\mathbf{h}), \end{aligned} \tag{6.17}$$

where the log-domain likelihood function of the complete data $p(\mathbf{X}, \mathbf{Y}|\mathbf{h})$ averaged over all possible 2^{KN} number of transmitted data sets $\mathbf{X}$, which is written as $\ln p(\tilde{\mathbf{X}}, \mathbf{Y}|\mathbf{h})$ can be further reformulated

as:

$$\begin{aligned}\ln p(\tilde{\mathbf{X}},\mathbf{Y}|\mathbf{h}) &\propto \ln p(\mathbf{Y}|\tilde{\mathbf{X}},\mathbf{h})P(\tilde{\mathbf{X}}|\mathbf{h}) \\ &\propto -\sum_{n=1}^{N}||\mathbf{y}_n - \mathbf{C}\boldsymbol{\mathcal{O}}_n\mathbf{h}||^2/2\sigma^2 \\ &\propto \left(2\mathrm{Re}\left\{\mathbf{r}^{\mathrm{H}}\mathbf{h}\right\} - \mathbf{h}^{\mathrm{H}}\mathbf{R}\mathbf{h}\right)/2\sigma^2, \end{aligned} \tag{6.18}$$

where $\mathbf{y}_n$ and $\boldsymbol{\mathcal{O}}_n = \mathrm{diag}[\tilde{x}_{1,n},\ldots,\tilde{x}_{K,n}]$ denote the received signal vector and the expected value of the transmitted vector during the nth symbol interval, respectively. We also removed the contribution of $P(\tilde{\mathbf{X}}|\mathbf{h})$ in the second line based on the fact that the estimated transmitted symbol $\tilde{\mathbf{X}}$ is independent of the channel $\mathbf{h}$. Furthermore, $\mathbf{r} \in \mathcal{C}^{K\times 1}$ and $\mathbf{R} \in \mathcal{C}^{K\times K}$ can be expressed as [184]:

$$\mathbf{r} = \sum_{n=1}^{N}\boldsymbol{\mathcal{O}}_n\mathbf{C}^{\mathrm{T}}\mathbf{y}_n, \tag{6.19}$$

$$\mathbf{R} = \sum_{n=1}^{N}\boldsymbol{\mathcal{O}}_n\mathbf{C}^{\mathrm{T}}\mathbf{C}\boldsymbol{\mathcal{O}}_n, \tag{6.20}$$

where the entries $\tilde{x}_{k,n}$ of $\boldsymbol{\mathcal{O}}_n$ are given by the soft output values of the SHS MUD:

$$\begin{aligned}\tilde{x}_{k,n} &= \sum_{\forall\mathbf{X}} x_{k,n}P(\mathbf{X}|\mathbf{Y},\mathbf{h}^{q-1}) \\ &= \sum_{\forall\mathbf{x}_n} x_{k,n}P(\mathbf{x}_n|\mathbf{y}_n,\mathbf{h}^{q-1})\Sigma_c \\ &\propto \sum_{\forall\mathbf{x}_{-k,n}} P(x_{k,n}=+1,\mathbf{x}_{-k,n}|\mathbf{y}_n,\mathbf{h}^{q-1}) \\ &\quad -P(x_{k,n}=-1,\mathbf{x}_{-k,n}|\mathbf{y}_n,\mathbf{h}^{q-1}). \end{aligned} \tag{6.21}$$

Noting that $\Sigma_c = \sum_{\forall\mathbf{X}_{-n}} P(\mathbf{X}_{-n}|\mathbf{Y}_{-n},\mathbf{h}^{q-1})$ is a constant and hence was cancelled out, where $\mathbf{X}_{-n}$ and $\mathbf{Y}_{-n}$ represents the transmitted data frame and received signal frame with the nth symbol interval excluded, respectively.

The key idea of our HS aided EM algorithm is that instead of using the true *a posteriori* distribution of the transmitted data $\mathbf{X}$, given both the observation constituted by $\mathbf{Y}$ and the parameter estimate constituted by the channel $\mathbf{h}$, we use the approximated *a posteriori* distribution based on the K-user harmony vectors in the final HMM, which are believed to contribute significantly to the true *a posteriori* distribution. We may now utilise the set of significantly contributing K-user vectors stored in the final HMM instead of summing them over all possible 2^{K-1} number of values $\mathbf{x}_{-k,n}$. Hence, Eq. (6.14) or Eq. (6.15) may now be directly applied for the calculation of the soft value $\tilde{x}_{k,n}$, hence we arrive at $\tilde{x}_{k,n} = \tanh[\mathcal{L}(x_{k,n})]$.

6.4.3.2 The Maximisation Step

We may now maximise the expected value of Eq. (6.17) in the sense of the MMSE, which is explicitly given by [184]:

$$\mathbf{h}^q = arg\max_{\mathbf{h}}\Omega(\mathbf{h},\mathbf{h}^{q-1}) = \mathbf{R}^{-1}\mathbf{r}. \tag{6.22}$$

However, the EM based blind-type CE suffers from the classic phase ambiguity of π in the BPSK modulated system considered. Hence, a few initial training symbols are required to remove the phase ambiguity of π before activating the blind EM algorithm, leading to a semi-blind CE. Since the EM algorithm is a data-aided scheme, the length of the initial training sequence is typically short compared to a pilot-aided CE. In this case, we may use the rough initial CEs based on the training symbols to initialise the receiver, while the subsequent channel updates of Eq. (6.22) can be generated by combining the data-aided and pilot-aided channel estimates. More explicitly, we have:

$$\mathbf{h}^q = \left[\mathbf{R} + \mathbf{R}^p\right]^{-1} \left(\mathbf{r} + \mathbf{r}^p\right), \tag{6.23}$$

where the calculation of $\mathbf{R}^p$ and $\mathbf{r}^p$ is based on the true value of the pilot symbols.

6.5 Performance Evaluation

6.5.1 Complexity

Let us now discuss the complexity of our HS-aided MUD in terms of the required FF evaluations. In the uncoded DS-CDMA case, the ML detector requires $\mathcal{Q}_{ML} = 2^K$ evaluations, while the HM algorithm requires a total of

$$\mathcal{Q}_{HS} = Q \times 2 \times K \times P_{ma} + Q + M \tag{6.24}$$

evaluations of the FF of Eq. (6.4). In detail, it includes M FF evaluations of the initial HMM, an evaluation at the end of each of the Q improvisations and $2 \times K$ evaluations, when generating the marginal APP for the pitch adjustment of each of the Q improvisations.

On the other hand, when the SHS aided MUD is considered, generating the *a posteriori* LLRs given by Eq. (6.9) for the K users requires $\mathcal{Q}_{MAP} = K \times 2^K$ evaluations of Eq. (6.10), while the SHS algorithm requires a total of

$$\mathcal{Q}_{SHS} \quad = \quad \mathcal{Q}_{HS} + 2 \times K \tag{6.25}$$

evaluations of the FF of Eq. (6.10), where in addition to the $\mathcal{Q}_{HS}$ FF evaluations according to Eq. (6.10), the soft output generation requires a further $2 \times K$ evaluations of Eq. (6.15).

6.5.2 Uncoded DS-CDMA

6.5.2.1 Effects of the Number of Improvisations

Fig. 6.2 and Fig. 6.3 show the achievable performance of the HS-aided MUD for a $K = 15$-user uncoded BPSK modulated fully-loaded DS-CDMA system using both $N_c = 15$-chip m-sequences

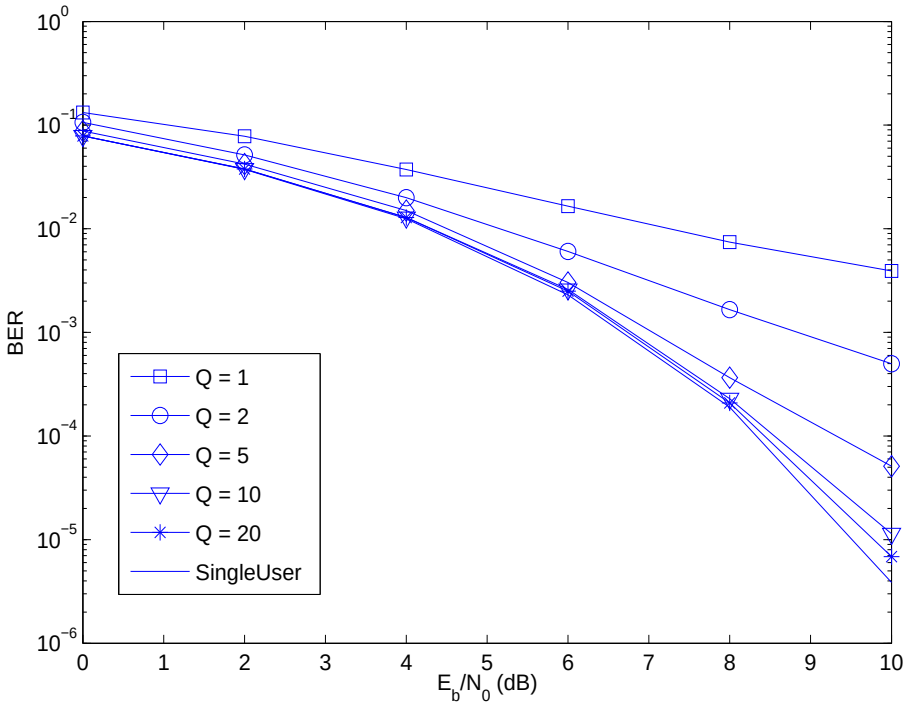


Figure 6.2: BER performance of HS-aided MUD of a $K = 15$-user uncoded BPSK modulated fully-loaded DS-CDMA system using $N = 15$-chip m-sequences, when transmitting over an AWGN channel. Furthermore, we have $M = 10$, $P_{ma} = 1$. The number of FF evaluation was $\mathcal{Q}_{HS} = 630$ for $Q = 20$, while $\mathcal{Q}_{ML} = 32768$.

and random sequences, respectively, when transmitting over an AWGN channel. The information frame length was $N_i = 512$. In our simulations, we varied the value of Q and M and set $P_{ma} = 1$.

Fig. 6.2 demonstrates that a near-single-user performance can be achieved by the HS MUD at a significantly reduced complexity compared to the ML detector. Quantitatively speaking, according to Eq. (6.24) for $Q = 20$ and $M = 10$ the FF is evaluated $\mathcal{Q}_{HS} = 630$ times using the HS-aided MUD, while the ML requires $\mathcal{Q}_{ML} = 2^{15} = 32768$ evaluations. On the other hand, when employing random sequences, the performance of the HS aided MUD failed to reach near-single-user performance even for $M = 20$ and $Q = 50$, as shown in Fig. 6.3. This was the consequence of the potentially high cross-correlation of the randomly chosen sequences, which imposed substantial interference. However, we will demonstrate that this detrimental effect can be compensated, when a channel-coded iterative receiver is employed, where the soft decoding is capable of removing the interference imposed.

6.5.2.2 Effects of Other Parameters

Fig 6.4 shows the effects of varying the HMM size M and the harmony memory activation probability P_{ma}, while fixing $Q = 20$. The other simulation parameters remain the same. The HMM size M represents the initial population size and predetermines the solution-space exploration capability of the EA algorithm, where a large M ensures a sufficiently diverse population. As demonstrated in

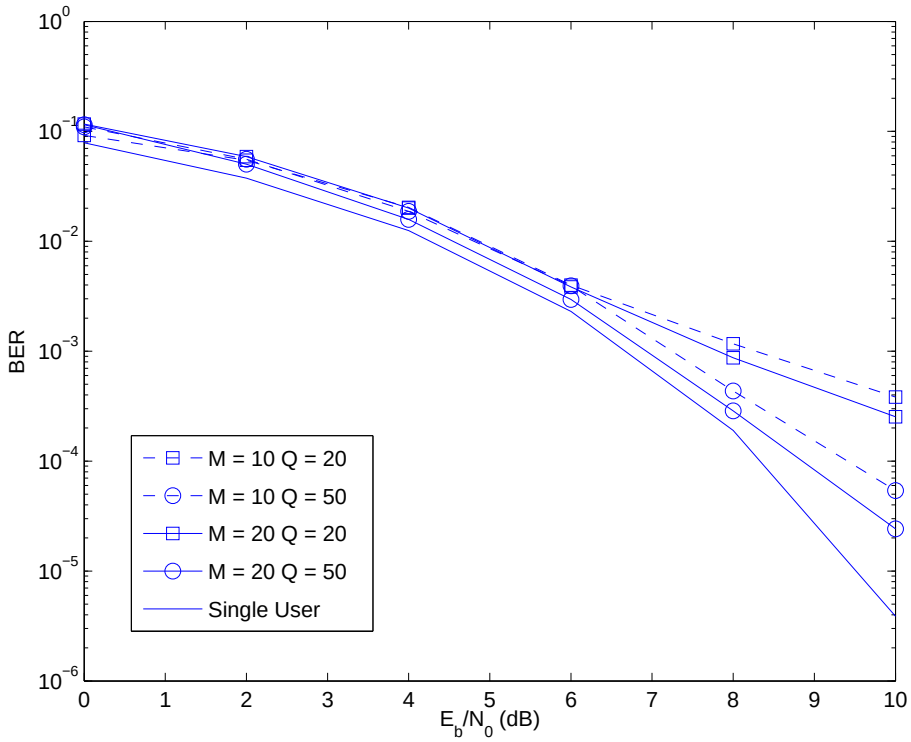


Figure 6.3: BER performance of HS-aided MUD of a $K = 15$-user uncoded BPSK modulated fully-loaded DS-CDMA system using $N = 15$-chip random sequences, when transmitting over an AWGN channel. Furthermore, we have $P_{ma} = 1$. For the best BER performance simulated, the number of FF evaluation was $\mathcal{Q}_{HS} = 1570$ for $Q = 50$ and $M = 20$, while $\mathcal{Q}_{ML} = 32768$.

Fig. 6.4, the achievable performance is improved substantially upon increasing $M = 2, 5, 10, 20$, yielding small incremental evaluations of the FF of Eq. (6.4) i.e. $\mathcal{Q}_{HS} = 622, 625, 630, 640$ according to Eq. (6.24).

Furthermore, the probability of $P_{ma} = 1$ was favoured. The reason for this choice is that the marginal APP based pitch adjustment of Eq. (6.6) is capable of providing sufficient decision reliability, hence the magnitude of the LLRs is improved during the successive iterations or improvisations. This suggests that the selection of a new K-user harmony vector should be biased towards the HMM rather than towards a random choice from the legitimate alphabet. Hence, artificially enforcing a random selection with a probability of $(1 - P_{ma})$ is not recommended and hence the improvisations based on this selection may be wasted. Importantly, ignoring the associated random selection owing to $P_{ma} = 1$ does not limit the exploration capability of the HS-aided MUD, because if we have a sufficiently high M, randomly selected base vectors have already been generated for inclusion in the HMM before fine-tuning the pitch adjustment.

6.5.3 Coded DS-CDMA

Let us now investigate the SHS-aided MUD in the context of a channel-coded DS-CDMA system using an iterative receiver. Each user employed a random spreading sequence of length $N_c = 7$ and an outer repetition code of rate of $R = 1/3$, where the information frame length was $N_i = 512$

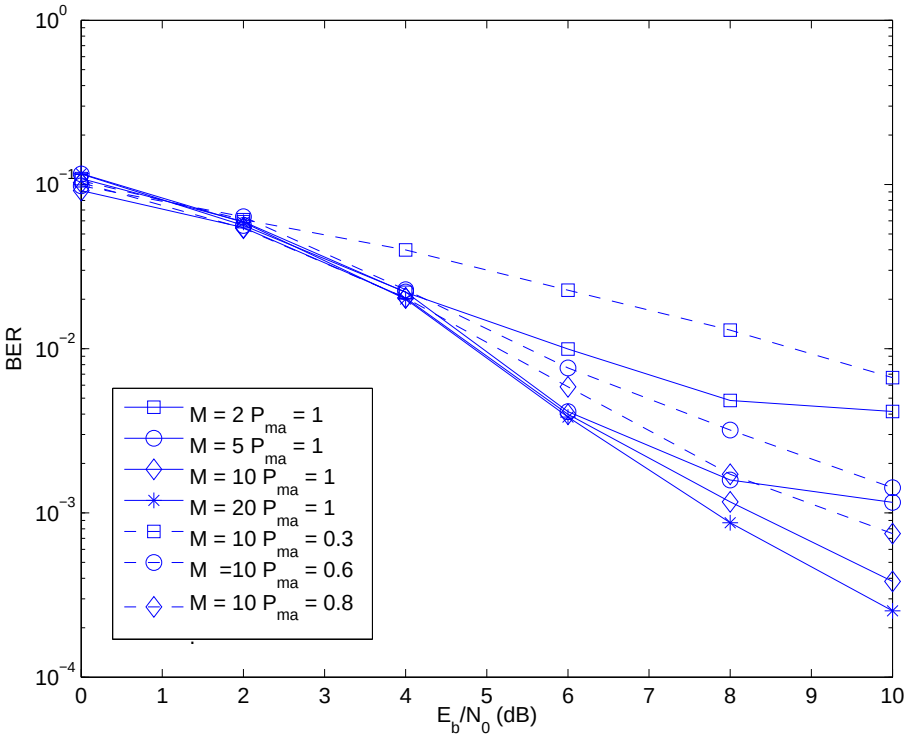


Figure 6.4: The effect of the harmony memory size M and the harmony memory considering probability P_{ma} on the BER performance of the HS-aided MUD of a $K = 15$-user uncoded BPSK modulated fully-loaded DS-CDMA system using $N = 15$-chip random sequences, when transmitting over an AWGN channel. Furthermore, we have $Q = 20$. According to Eq. (6.24) we had $\mathcal{Q}_{HS} = 622, 625, 630, 640$ for $M = 2, 5, 10, 20$, respectively, while $\mathcal{Q}_{ML} = 32768$.

and the number of iterations between the SHS-aided MUD and the soft decoder was set to $I = 10$. The resultant channel coded transmission frame length was $N = N_i/R = 1536$ bits. We set the HS parameters as $M = 10$ and $P_{ma} = 1$. Furthermore, we define the so-called normalised system-load β as the ratio of the number of users supported to the spreading sequence length employed. Hence we have $\beta = K/N_c$. An AWGN channel was assumed and additionally we modelled the false-locking imposed by blind EM-based CE by a uniformly distributed channel phase noise, namely when we had $h_k = e^{j\theta_k}, \theta_k \in [-\pi, \pi)$, where the phase noise remained constant over a transmission block of $N = 1536$ bits. Apart from employing the EM based CE algorithm using the above-mentioned random phase-noise based false-locking model, we will also use an idealised benchmarker, where we assume perfect knowledge of the channel's phase noise.

6.5.3.1 Effects of the Number of Improvisations

Fig 6.5 shows the effects of varying the number of improvisations from $Q = 2, 5, 8, 10, 20$, when the normalised system load was $\beta = 4$ and we had $\mathcal{Q}_{SHS} = 180, 351, 522, 636, 1206$, while $\mathcal{Q}_{MAP} = 28 \times 2^{28}$. Compared to Fig 6.3 recorded for uncoded DS-CDMA, the effect of interference due to the high cross-correlation of the random sequences employed is compensated by the rate $R = 1/3$ outer repetition code and iterative receiver. The outer repetition code employed is known as being capable of providing the best multiple access capability [35], since it gener-

ates the highest extrinsic information in interference-limited, i.e. low Signal-to-Interference-Noise Ratio (SINR) scenarios compared to other channel codes [35]. Fig 6.5 demonstrated furthermore that even in this overloaded scenario, the SHS-aided MUD remained capable of attaining a performance, which was within $E_b/N_0 = 0.5$dB from the single-user performance measured at the BER of $P_e = 10^{-5}$.

6.5.3.2 Effects of System Load

Fig 6.6 shows the effects of varying the normalised system load β, when we have $Q = 20$ improvisations. As shown in this figure, having an almost unprecedented system load as high as $\beta = 6$ is possible for the SHS aided MUD. More specifically, when the normalised system load is $\beta = 6$, and $\mathcal{Q}_{SHS} = 1794$ FF evaluations are used instead of $\mathcal{Q}_{MAP} = 42 \times 2^{42}$ the performance is only about $E_b/N_0 = 1$dB away from the single user performance measured at the BER of $P_e \approx 2 \times 10^{-5}$. We note furthermore that the normalised system load of $\beta = 6$ is supported using the same parameter, as employed when we have a normalised system load of $\beta = 3$. In fact, none of the HS parameters were changed during these simulations. This implies that within certain limits the SHS-aided MUD is capable of maintaining a near-single-user performance at a fixed complexity region for a high normalised system load β.

However, although it was not explicitly shown here, we found that further increasing the number M or Q of improvisations does not necessarily improve the achievable performance. This implies that the SHS-aided MUD may suffer from an irreducible BER owing to getting trapped in a local optimum [12], when the normalised system load is high. An intuitive argument for this is that the APP-based pitch adjustment may fail to function reliably since there are cases, where the probability of a specific bit is near-unity despite being erroneous, which prevents the MUD from exploring other regions of the search space. This problem becomes more prominent, when the search space is large.

6.5.3.3 SHS Assisted EM Based CE

Fig 6.7 shows the achievable performance of our SHS assisted EM based CE for a DS-CDMA system having a normalised system load of $\beta = 4$, when the channel's phase noise was unknown to the receiver. The number of iterations between the CE and IDD was set to $I_{EM} = 5$. Fig 6.7 demonstrates that employing training sequences alone, i.e. without a data-aided mode, failed to generate sufficiently accurate phase estimates. On the other hand, the semi-blind data-aided EM approach, which jointly considered the pilot- and data-aided channel estimates become capable of acquiring accurate phase noise estimates. Furthermore, the longer the training sequences employed, the faster the convergence to the best possible performance associated with perfect knowledge of the channel's phase noise, when modelling the effects of false-locking. Furthermore, $T = 16$ pilots are required in the data-aided EM based CE to achieving convergence, while employing $T = 64$

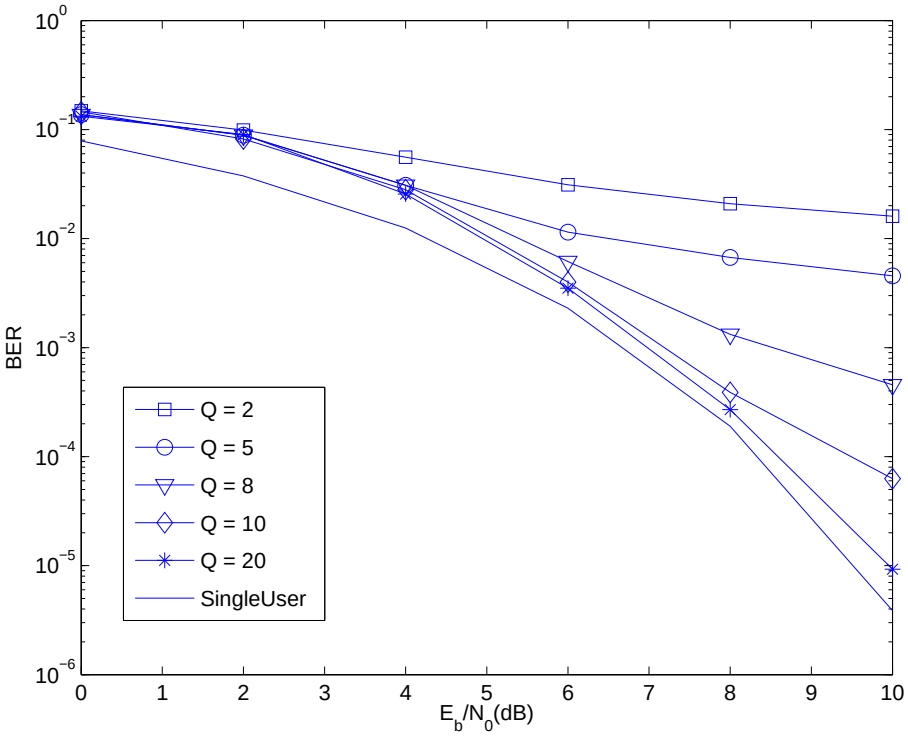


Figure 6.5: BER performance of SHS-aided MUD of a $K = 28$-user $R = 1/3$-repetition coded BPSK modulated DS-CDMA system using $N_c = 7$-chip random sequences and $I = 10$ iterations. Furthermore, we have $M = 10$, $P_{ma} = 1$. According to Eq. (6.25), we have $\mathcal{Q}_{SHS} = 180, 351, 522, 636, 1206$ for $Q = 2, 5, 8, 10, 20$, respectively, while $\mathcal{Q}_{MAP} = 28 \times 2^{28}$.

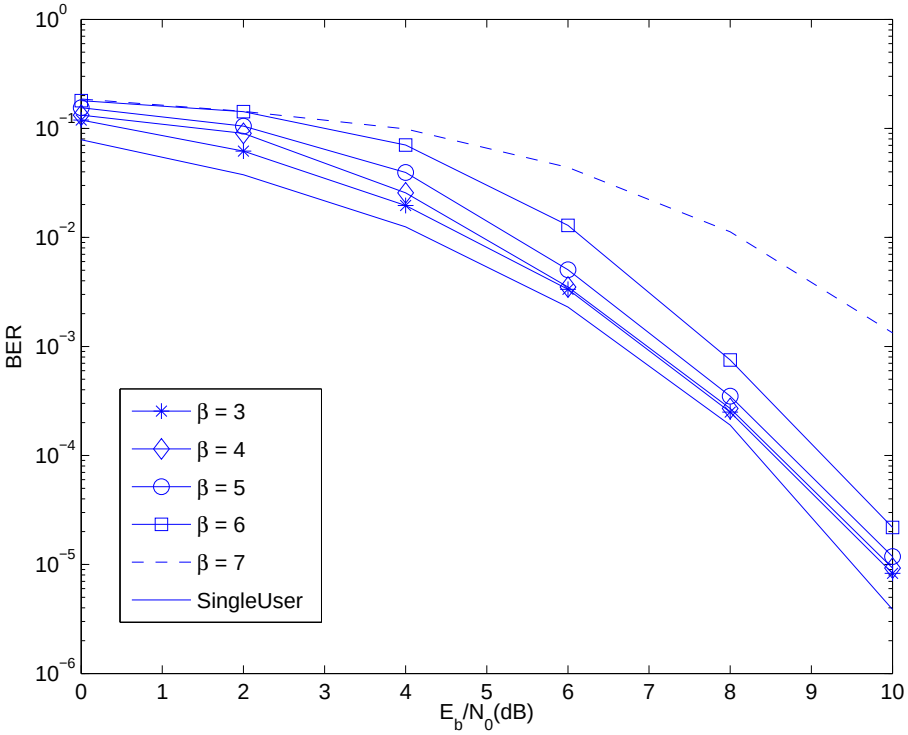


Figure 6.6: BER performance of SHS-aided MUD of a $R = 1/3$-repetition coded BPSK modulated DS-CDMA system using $N_c = 7$-chip random sequences for supporting the system load of $\beta = 3, 4, 5, 6, 7$. Furthermore, we have $Q = 20$, $M = 10$, $P_{ma} = 1$. According to Eq. (6.25), we have $\mathcal{Q}_{SHS} = 1794$ for $\beta = 6$, while $\mathcal{Q}_{MAP} = 42 \times 2^{42}$.

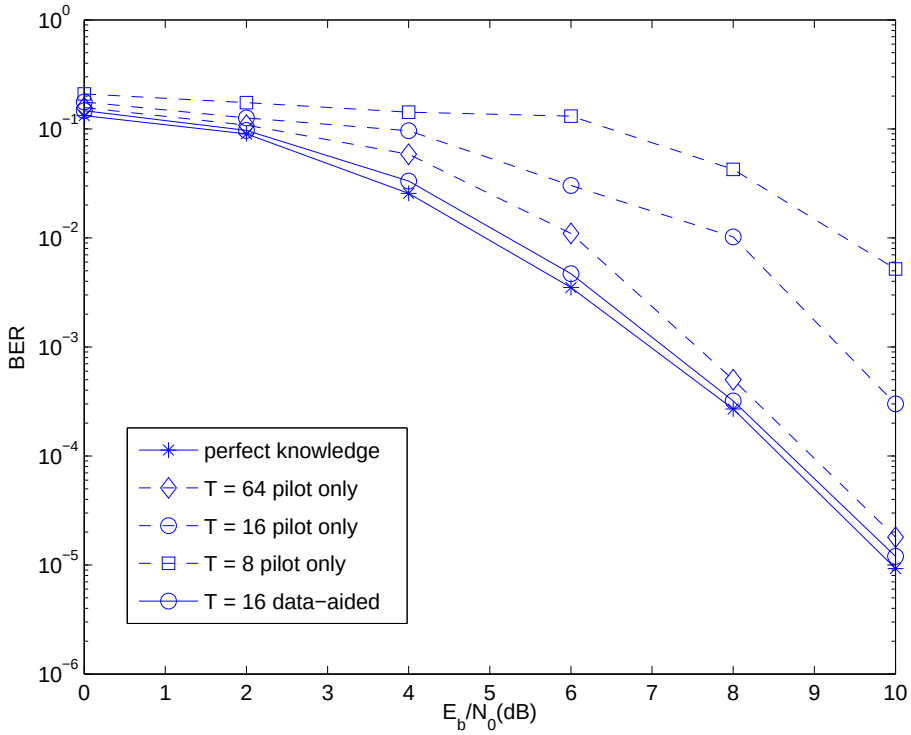


Figure 6.7: BER performance of SHS-assisted EM based CE algorithm of a $R = 1/3$-repetition coded BPSK modulated DS-CDMA system using $N_c = 7$-chip random sequences for supporting a system load of $\beta = 4$. Furthermore, we have $Q = 20$, $M = 10$, $P_{ma} = 1$. According to Eq. (6.25), we have $\mathcal{Q}_{SHS} = 1206$, while $\mathcal{Q}_{MAP} = 28 \times 2^{28}$.

pilot-aided CE can not lead to convergence.

6.5.4 Discussion

Based on the above investigations, we may conclude that the proposed HS-aided algorithm is capable of achieving a near-single-user performance without the excessive complexity of the optimum detector and that within limits, the complexity of the algorithm is reasonably independent of β. In comparison to other EAs, which require the tuning of a range of parameters, as long as the initial HMM size of M is sufficiently large, the HS-aided algorithm requires the tuning of a single parameter, namely of the number of improvisations Q. This property is acquired as a benefit of the marginal APP-based pitch adjustment step, which equips it with the capability of attaining convergence from a Bayesian inference point of view. At the same time, the entire search space is visited by the HS MUD with the aid of the randomly generated base vector from the set of M rather diverse candidate vectors at each improvisation.

The generation of the pitch adjustment probability based on the marginal APP may be compared to the so-called Metropolis-Hastings (MH) method [12], which is also known as the MCMC technique [12]. In contrast to the conventional MH method, which generates successive sample vectors that are correlated, our proposed HS algorithm does not form a correlated Markov chain, in other words, the resultant harmony vectors are independent. This is because in the HS algorithm

each new harmony vector may be accepted or rejected based on the value of the FF. As a result, after each improvisation, the newly generated HMM will have an identical, or a higher score than the previous HMM. Together with a sufficiently high M value, the random-walk behaviour of the conventional MH method may be avoided. On the other hand, our HS algorithm is also different from the traditional importance sampling or rejection sampling technique [12], which requires a so-called *proposal distribution* and is only effective in low-dimensional problems.

Unfortunately, the HS aided algorithm employing the marginal APP P_{pa} suffers from two problems similar to those of the conventional Monte Carlo algorithm, which are the potential occurrence of local optima and that the exact proof of convergence is not available at the time writing.

6.6 Conclusion

Stimulated by the random guided HS algorithm of applied mathematics, in this chapter, we proposed a novel HS-aided MUD for DS-CDMA systems and developed it into a joint iterative CE, MUD and channel decoding framework. We used the APP based soft information as a variant of the pitch adjustment probability in the original HS-aided algorithm. As a result, the harmony memory activation step and the pitch adjustment step were jointly considered, while the random selection step was removed, since the APP based pitch adjustment probability provides sufficiently accurate guideline during the optimisation process. Hence, by adjusting a single parameter, namely the improvisation number, the HS-aided MUD was shown to be capable of approaching the single-user performance even for DS-CDMA systems supporting an extremely high normalised system load of $\beta = 6$ at a low complexity. As a further benefit, the soft HS-aided MUD can be efficiently combined with the EM based CE framework, when communicating over a block-invariant channel, where the EM algorithm is recognised as optimum from a Bayesian inference point of view. Our future work will involve comparions to other optimisation algorithms.

6.7 Appendix - Example of Harmony Search Algorithm

We now provide a simple example to augment the explanation of our HS-aided MUD for uncoded DS-CDMA. Consider a $K = 7$ user BPSK modulated DS-CDMA system employing $N_c = 7$-chip random spreading codes, when transmitting over the AWGN channel at $E_b/N_0 = 6dB$, and using $M = 5$, $P_{ma} = 1$.

Assume that we have the transmitted vector of

$$x = [+1, +1, -1, +1, +1, +1, -1]^{\mathrm{T}}$$

for the $K = 7$ users and the spreading matrix of

$$\mathbf{C} = \begin{bmatrix} +1 & +1 & +1 & +1 & +1 & +1 & -1 \\ -1 & +1 & +1 & -1 & -1 & -1 & -1 \\ +1 & +1 & +1 & -1 & +1 & -1 & -1 \\ +1 & -1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & +1 & +1 & +1 & -1 \\ +1 & -1 & -1 & +1 & -1 & +1 & +1 \\ +1 & +1 & +1 & -1 & +1 & +1 & +1 \end{bmatrix},$$

where each column of the spreading matrix represents each user's spreading code. The received vector at the output of the AWGN channel was:

$$\mathbf{y} = \begin{bmatrix} +4.0686 \\ -3.1287 \\ +2.5033 \\ -2.0255 \\ +1.9452 \\ -0.6461 \\ +1.1707 \end{bmatrix}.$$

According to Eq. (6.4) the minimum value of the FF may be calculated as

$$f(\mathbf{x})_{min} = ||\mathbf{y} - \mathbf{C}\mathbf{x}|| = 2.8368.$$

Let us now track the operation of the HS process during the first improvisation step by step as follows:

6.7.1 Generation of the Initial HMM

We randomly generate the initial HMM $\mathbf{X}^0$ having $M = 5$ independent initial harmony candidate vectors:

$$\mathbf{X}^0 = \begin{bmatrix} -1 & -1 & -1 & +1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 & -1 & +1 & -1 \\ -1 & -1 & -1 & +1 & +1 & +1 & -1 \\ -1 & +1 & -1 & -1 & -1 & +1 & +1 \\ +1 & -1 & +1 & +1 & +1 & -1 & -1 \end{bmatrix}.$$

Their corresponding FF values are calculated from Eq. (6.4) as:

$$f(\mathbf{X}^0) = \begin{bmatrix} f(\mathbf{x}_1^0) = 12.5052 \\ f(\mathbf{x}_2^0) = 5.7130 \\ f(\mathbf{x}_3^0) = 9.2966 \\ f(\mathbf{x}_4^0) = 10.4767 \\ f(\mathbf{x}_5^0) = 4.5612 \end{bmatrix}.$$

6.7.2 Generation of New Harmony Candidate

Base Vector When generating a new harmony candidate, the HS-aided MUD firstly selects a base vector from the above initial HMM, since $P_{ma} = 1$ was set. The base harmony vector $\mathbf{x}^b$ is selected as $x_k^b = x_{m_i,k}^0, m_i \sim \mathcal{U}[1, M], k \in [1, K]$, where $\mathcal{U}$ represents the random uniform distribution. In this example, we have generated the index randomly as $m_i = [1, 3, 3, 2, 2, 2, 1]$, the base harmony vector becomes:

$$\begin{aligned} \mathbf{x}^b &= [x_{1,1}^0, x_{3,2}^0, x_{3,3}^0, x_{2,4}^0, x_{2,5}^0, x_{2,6}^0, x_{1,7}^0] \\ &= [-1, -1, -1, +1, -1, +1, -1]. \end{aligned}$$

Pitch Adjustment The algorithm then generates the pitch adjustment probability P_{pa} according to Eq (6.5) and Eq (6.6), yielding $P_{pa}(x_1^{new} = +1) = 0.9975$, which implies that the first bit of the base harmony vector should be "+1" with a probability of 0.9975, given the values of other bits in the base harmony vector remain unchanged. In this example, we generated a uniformly distributed random value of $\mathcal{U}(0, 1) \leq P_{pa}(x_1^{new} = +1)$, thus the updated base harmony vector becomes $\mathbf{x}^b = [+1, -1, -1, +1, -1, +1, -1]$.

This process continues, until all K bits of the base vector become updated with the aid of Eq (6.5) and Eq (6.6) in a similar manner. The pitch adjustment probabilities of all K bits calculated are:

$$P_{pa} = [0.9975, 0.9975, 0.9975, 0.9975, 0.9975, 0.0025].$$

and their corresponding base harmony vectors updated during each pitch adjustment are:

$$\mathbf{X}^b = \begin{bmatrix} +1 & -1 & -1 & +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & +1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 & +1 & +1 & -1 \\ +1 & +1 & +1 & +1 & +1 & +1 & -1 \\ +1 & +1 & +1 & +1 & +1 & +1 & -1 \end{bmatrix}.$$

As a result, the last row of the matrix $\mathbf{X}^b$ is the newly generated harmony vector, namely:

$$\mathbf{x}^{new} = [+1, +1, +1, +1, +1, +1, -1].$$

6.7.3 Generation of New HMM

Since the FF value of the new harmony vector becomes $f(\mathbf{x}^{new}) = 4.3227$, which is less than the worst one in the initial HMM $f(\mathbf{x}_1^0) = 12.5052$, this new harmony vector will replace the lowest fitness harmony vector, i.e. $\mathbf{x}_1^0$ for the generation of the new HMM, where we have $\mathbf{x}_1^1 = \mathbf{x}^{new}$.

Although we have not as yet found the optimum K-bit vector associated with $f(\mathbf{x})_{min}$, the newly generated HMM excludes the low-fitness and hence unlikely candidate $\mathbf{x}_1^0$. The above procedure continues until the affordable complexity is exhausted. Observe that at this stage, there is only a single bit difference between the new harmony vector $\mathbf{x}^{new}$ and the true transmitted vector $\mathbf{x}$, the correct K-bit vector is likely to be found with a high probability during the forthcoming HS improvisations.

Table 6.2: Pseudo-code of the HS-aided MUD

/* Initial HMM Generation */
Initialise $\mathbf{X}^0$ Define $f(\cdot) = ||\mathbf{y} - \mathbf{C}\mathbf{x}||$ Compute $f(\mathbf{x}_m^0), m \in [1, M]$
/* Improvisation Loop */
for $q = 1, \ldots, Q$ **do**
  /* New Harmony Candidate Generation */
  if $\mathcal{U}(0,1) \leq P_{ma}$ **do**
    /* Base Vector Generation */
    for $k = 1, \ldots, K$ **do**
      $x_k^b = x_{m_i,k}^{q-1}, m_i \sim \mathcal{U}[1, M]$
    end for
    /* Pitch Adjustment */
    for $k = 1, \ldots, K$ **do**
      Compute pitch adjustment probability based on Eq (6.6)
      if $\mathcal{U}(0,1) \leq P_{pa}(x_k^{new} = +1)$ **do** $x_k^{new} = +1$
      else $x_k^{new} = -1$
      end if
      Set $\mathbf{x}^b = [x_1^{new}, \ldots, x_k^{new}, x_{k+1}^b, \ldots, x_K^b]$
    end for
    Set $\mathbf{x}^{new} = \mathbf{x}^b$
  else
    /* Random Selection */
    for $k = 1, \ldots, K$ **do**
      $x_k^{new} \in \{+1, -1\}$
    end for
  end if
  /* New HMM Generation */
  Find $\epsilon = arg\max f(\mathbf{x}_\epsilon^0), \epsilon \in [1, M]$
  if $f(\mathbf{x}^{new}) < f(\mathbf{x}_\epsilon^{q-1})$ **and** $f(\mathbf{x}^{new}) \neq f(\mathbf{x}_m^{q-1}), \forall m$ **do**
    $\mathbf{x}_\epsilon^q = \mathbf{x}^{new}, \mathbf{x}_m^q = \mathbf{x}_m^{q-1}, \forall m \neq \epsilon$
  else $\mathbf{X}^q = \mathbf{X}^{q-1}$
  end if
end for

Chapter 7

Conclusion

In this concluding chapter, the main findings of our investigations presented in this treatise are summarised in Section 7.1. Additionally, a range of ideas concerning our future research is presented in Section 7.2.

7.1 Summary of Findings

7.1.1 Chapter One

In this introduction chapter, we outlined the objective of this thesis, namely the aim of providing an unified treatment of near-capacity non-orthogonal random waveform based multiuser communications. The specific techniques proposed in this thesis were discussed in more detail in Chapter 2, where we considered the SPC concept aiming for achieving a high data rate, the random coding principle aiming for maintaining a low error ratio, both of which benefit from the iterative receiver architecture presented in Section 1.1. The brief history of multiuser communications was reviewed in Section 1.2. These considerations assisted us in formulating the motivation of employing non-orthogonal random waveform based multiuser communications.

7.1.2 Chapter Two

We elaborated further in Chapter 2, where we focused our attention on three key topics, namely, on the attainable maximal mutual information of multiuser communications in Section 2.2, on an ideal as well as a practical random coding scheme in Section 2.3 and on an advanced iterative receiver architecture in Section 2.4.

We constructed a generalised linear system model, which is capable of describing a wide range of systems, including multiuser systems, MIMO systems and circulant systems. The fundamental information theoretic analysis of Section 2.2.1.3 suggested that non-orthogonal multiuser commu-

nications is superior in comparison to the traditional orthogonal approach, as illustrated in Fig 2.4 and Fig 2.5. The achievable capacity improvement can be attained by invoking SIC as well as SPC. For a real system having a discrete channel input in Section 2.2.2, our discussion on the constellation-constrained capacity resulted in the derivation of EXIT curve. Furthermore, we revealed that the conventional multilayer mapping suffered from two undesired properties, namely, from having a reduced cardinality and a non-equiprobable signalling constellation, as shown in Fig 2.6. For the sake of eliminating the capacity penalty due to the above-mentioned drawbacks as well as for retaining the benefit of having a Gaussian-like channel input constellation for the multilayer mapping, simple layer-specific phase rotations and power allocation can be applied. The resultant system subsumes the classic multilevel coding and bit-interleaved coded modulation architectures, as discussed in Section 2.2.2.4.

The artificially imposed phase rotation and the layer-specific power allocation underlined the importance of combining the concept of channel division with that of code division. On the other hand, the code division principle is also crucial. The random coding principle had been widely used, both explicitly and implicitly in various designs. With the aid of a comprehensive introduction to factor graph and the sum-product algorithm of Section 2.3.1, we discussed two practical quasi-random coding approaches, namely, LDPC codes and interleaved codes. It may be concluded that interleaved codewords are essentially random codewords. In the context of non-orthogonal multiuser communications, differently interleaved codewords render the receiver's factor graph fully connected and prevent the graph from being decomposed into local subgraphs, as demonstrated in Fig 2.16(a) and Fig 2.16(b).

Concerning the receiver architecture, the decomposition of a MAP based optimum receiver leads to the so-called iterative data detection and channel decoding principle, which was justified by the positioning the interleavers between the channel decoder and the detector. Two primary benefits of the interleavers were also emphasised in Section 2.4.2. Firstly, the interleavers made the consecutive coded bits uncorrelated. Secondly, the interleaved codewords may be considered as codewords of a quasi-random coding scheme. In particular, we focused our attention on the data detection algorithm, where both the optimum Bayesian detector and a low-complexity IC aided detector algorithm were derived as detailed in Section 2.4.2.1 and Section 2.4.2.2, respectively.

7.1.3 Chapter Three

This chapter presented a specific instantiation of non-orthogonal random waveform based multiuser communications in cellular applications, where interleaved quasi-random codes were employed as a means of differentiating the users. More explicitly, the IDMA system was introduced, which may be considered as a code-spread and chip-interleaved DS-CDMA system as seen in Fig 3.1. The properties of IDMA, which are different from traditional DS-CDMA include the concept of near-capacity random coding, an increased time diversity, flexible joint coding and spreading sequence

design and the employment of a powerful chip-level iterative receiver, as detailed in Section 3.1.1.

We also proposed a generalised system model, which we referred to a MC-IDM-IDMA and analysed its convergence behaviour with the aid of EXIT charts. We firstly highlighted in Section 3.2.3.1 that the employment of two-dimensional EXIT charts in the analysis of MUD is appropriate, when we study PIC and an equal received power system communicating over an AWGN channel or over an uncorrelated fading channel. By observing the EXIT function T_{mud} seen in Fig 3.4, we were able to distinguish two different MUD behaviours, namely the right-hand-side of the EXIT-chart was the noise-limited region, while its left-hand-side was the interference-limited region. Furthermore, as illustrated in Fig 3.5, the terminology of error-correction capability and multiple access capability were introduced for characterising the outer channel code's EXIT function T_{dec}.

The EXIT chart analysis facilitated our investigations of three design aspects, namely the coding versus spreading tradeoff, the multiplexing versus diversity tradeoff and the complexity versus performance tradeoff. The associated coding versus DS-spreading tradeoffs were discussed and our findings were compared to the characteristics of traditional DS-CDMA in Section 3.2.4.1. It was found that in contrast to the separate DS-spreading and channel coding design of traditional DS-CDMA systems, the DS-spreading operation should be replaced by a repetition code in the context of IDMA type systems so as to be jointly designed with the channel code in order to improve the achievable bandwidth efficiency. A general design rule was also provided for invoking low-rate codes in our system. More specifically, when designing MC-IDM and IDMA systems for operating in various channel conditions at a given E_b/N_0 constraint and at an affordable decoding complexity, it is beneficial to employ a channel code having the highest possible rate, which is 'just' capable of meeting the specific target BER requirements and then dedicating the rest of the affordable bandwidth expansion to DS-spreading, i.e. to DS repetition codes. Furthermore, in Section 3.2.4.2, the differences between MC-IDM and MC-IDMA were investigated and the multiuser diversity benefits of MC-IDMA were revealed. The proposed MC-IDM-IDMA system is thus flexible in terms of either assigning the total system throughput to a single user or providing a multiple access capability for several users, as demonstrated in Fig 3.17. Finally, the associated complexity versus performance tradeoffs were discussed by investigating two MUDs' EXIT curves and a reduced-complexity hybrid MUD concept was provided in the context of Fig 3.22.

While the family of IDMA-type systems employ explicit interleavers, we proposed the so-called IR-CDMA concept, where different users are distinguished by their quasi-random codes having implicit interleavers, such as RA codes. We specially designed novel user-specific SE interleaver aided RA codes for the sake of reducing the memory storage requirements in the context of multiuser transceivers with no performance penalty, as seen in Fig 3.21. Quantatitively, our proposed SE interleaver, which exhibits similar correlation properties to those of random interleavers as illustrated in Fig 3.20, does not require the storage of all the K users' interleavers. Only m bits have to be stored, where m in our system is related to the length $Q = 2^m$ of the source information packets. It was also demonstrated that low-rate RA code aided IR-CDMA systems using SE in-

terleavers are capable of approaching the Gaussian channel's capacity with the aid of the proposed low-complexity unequal power allocation scheme discussed in Section 3.3.3.3. More explicitly, the discrepancy between the true capacity and the achievable capacity of our IR-CDMA arrangement is a function of both the block length as well as of the number of users. As a result, the IR-CDMA system becomes more suitable for employment in moderate multiuser sum-rate scenarios.

7.1.4 Chapter Four

In Chapter 4, we introduced non-orthogonal random waveform based multiuser communications framework in the context of cooperative systems, where we specifically designed the so-called IR-STC and outlined its benefits. More specifically, they are capable of maintaining a high throughput with the aid of SPC, while maintaining a low bit-error-ratio with the aid of the iterative detection.

After considering the family of both multiplexing-oriented and diversity-oriented configurations in Section 4.1.3.2, we compared the performance of our IR-STC scheme to the traditional G_4 OSTBC arrangement in Section 4.1.4.1. It was concluded that the maximum number of layers supported by our IR-STC scheme was $L = 7$, which is equivalent to a G_4 OSTBC scheme using a large and hence error-sensitive 128-QAM constellation. Moreover, our scheme required a lower power than the lower-throughput 4 bit/symbol G_4 OSTBC aided 16-QAM scheme, as observed at $BER \leq 10^{-5}$, which can be seen in Fig. 4.5. Furthermore, the achievable gain of IR-STC was summarised in Table 7.1.

In addition, we analysed in Section 4.1.4.2 the achievable performance of our IR-STC aided MSC, when communicating over Nakagami faded inter-source channels employing various relaying techniques, including AF, DF, SDF and DDF. We concluded that when the SNR of the inter-source channel is better than that of the source to destination channel, DF is the best relaying strategy in the presence of benign fading, provided that the Nakagami fading parameter m value larger than 1. The AF technique is only preferable at high SNRs, when severe inter-source channel fading is encountered. DDF performs consistently worse than DF due to the doubled noise variance of non-coherent detection. Surprisingly, when the fading is benign, non-coherent DDF outperforms the coherent detected AF technique without the cost of estimating all inter-source channels.

In contrast to the above-mentioned uncoded system, we also designed a novel energy efficient coding scheme in Section 4.2, namely the PANC scheme, which was inspired by the network coding philosophy. It may be viewed as a technique of conveying a linear combination of multiple sources' information, rather than using conventional relaying for delivering these information flows individually with the aid of classic resource allocation, such as time-multiplexing or code-multiplexing. We also designed a code-multiplexing oriented SPC scheme, which was employed as a benchmarker. In Section 4.2.2, we demonstrated that, intrinsicly, the PANC scheme may viewed as a SPC scheme defined over the Galois Field (GF) 2, while the SPC may be considered as a NC scheme defined over the complex-valued field. The simulation results of Fig 4.14 and Fig 4.15

Effective Throughput	G_4	IR-STC	Δ_1	Δ_2
$\eta_{IR-STC} = 2$	$\mathcal{M} = 16$	$L = 4$	6	-
$\eta_{IR-STC} = 2.5$	$\mathcal{M} = 32$	$L = 5$	5	-
$\eta_{IR-STC} = 3$	$\mathcal{M} = 64$	$L = 6$	4	9
$\eta_{IR-STC} = 3.5$	$\mathcal{M} = 128$	$L = 7$	2	7
$\eta_{IR-STC} = 4$	$\mathcal{M} = 256$	$L = 8$	-1	4

Table 7.1: Summary of the power gain Δ in dB of IR-STC compared to G_4 OSTBC in MSC, where Δ_1, Δ_2 denotes the power gain corresponding to 16-QAM and 64-QAM in G_4 OSTBC, respectively.

demonstrate that both schemes are capable of performing close to the outage probability bound. When compared to the SPC arrangement, the PANC scheme exhibits a lower complexity, requiring about one-fifth of the total number of iterations, when compared to the SPC scheme. This was achieved at the cost of a slight performance degradation, i.e. within 1dB measured at BLER of 10^{-3}, while maintaining the same effective throughput and delay.

7.1.5 Chapter Five

This chapter considered a realistic scenario for wireless TCP packet based transmission employing a HARQ scheme. HARQ schemes were widely used for reducing the channel induced packet loss events. Our proposed SPC aided multiplexed HARQ scheme was capable of reducing the link-layer HARQ retransmission delay, so that the TCP layer's end-to-end transmission efficiency was substantially improved.

More specifically, as discussed in Section 5.2.2, our scheme jointly encodes the current new packet to be transmitted and any packets that are about to be retransmitted, as illustrated in Fig 5.2. We invoked our non-orthogonal random waveform based multiuser communications principle, where the SPC scheme may also be considered as an instantiation of network coding in the complex-valued field. Two efficiency metrics were introduced, namely the link-layer effective throughput and the TCP layer mean frame arrival rate, where the latter was analysed in more detail. In particular, we assumed that the M/G/1 queue has a Poissonian arrival process associated with an arrival rate of λ, a general i.i.d. service time duration T and a single server, where the propagation delay, feedback delay and the slow start phase were ignored. The analysis of Section 5.3.2 was based on a Markov chain model, as seen in Fig 5.4. Our simulation results the achievable performance in Fig 5.5, Fig 5.6 and Fig 5.7. Quantatively, as seen in Fig 5.7, at the same mean frame arrival rate, the highest E_b/N_0 gain may be observed when the TCP buffer size was $B = 20$, which was approximately 5dB. Hence, it may be concluded that our proposed scheme is capable of substantially improving the TCP layer's end-to-end transmission efficiency for all transmitted packets at a marginal link-layer packet error ratio performance degradation.

The benefits and limitations of our proposed scheme may be summarised as follows. The scheme is more effective in the context of low-rate channel coded systems and discourages the employment of high-order modulation schemes. On the other hand, it achieves a significantly higher TCP layer throughput and may be readily integrated with existing systems without substantially modifying the current design. These properties make the proposed scheme particularly suitable for delay-sensitive low-rate applications.

7.1.6 Chapter Six

This chapter presented a novel HS aided multiuser detection algorithm, which is particularly suitable for detection in MIMO systems employing high order modulation scheme and in high user-load DS-CDMA systems. We commenced our algorithm development by introducing the HS algorithm in Section 6.2, where the general philosophy of EAs was summarised. Similarly to other EAs, the original HS algorithm has a range of parameters and may be partitioned into three steps, namely the initialisation step, improvisation step and the updating step. The improvisation step was summarised in Table 6.1.

This algorithm was then incorporated into a multiuser detector designed for DS-CDMA systems, where in contrast to the naive transplanting based technique of Section 6.3.2.1, we specifically designed the marginal APP based pitch adjustment step in Section 6.3.2.2, which effectively combines the memory consideration activation and pitch adjustment functions. We found that this novel modification efficiently guided the MUD's search procedure and made the original random selection redundant. Hence, compared to conventional EAs using a range of parameters based on trial-and-error testing, our proposed algorithm requires only two tuning parameters, namely the number of improvisations and the initial number of harmony candidates. The detailed pseudo-code of our algorithm was given in Table 6.2. As a further step, we developed our algorithm in the context of the entire iterative receiver, where the soft HS algorithm was proposed and integrated with the classic EM based channel estimation algorithm. We proposed three different types of soft outputs for our algorithm, namely the full soft-output generation technique of Eq (6.13), the approximate formula of Eq (6.14) and the expression of Eq (6.15).

Finally, various performance results were provided in Section 6.5. For an uncoded system, our HS algorithm was capable of approaching the single-user performance without the employment of full-search based optimum detector, as shown in Fig 6.2 and Fig 6.3. Moreover, the effect of the various parameters was illustrated in Fig 6.4. On the other hand, for a coded system, the HS-aided MUD was shown in Fig 6.6 to be capable of approaching the single-user performance even for DS-CDMA systems supporting an extremely high normalised system load of $\beta = 6$ at a low complexity. In this case, according to Eq. (6.25), the number of fitness function evaluations for the HS-aided MUD was $\mathcal{Q}_{SHS} = 1794$ for $K = 42$ users, while that for the MAP MUD was $\mathcal{Q}_{MAP} = 42 \times 2^{42}$. As a further benefit, the soft HS-aided MUD can be efficiently combined with

the EM based CE framework, when communicating over a block-invariant channel, as shown in Fig 6.7.

7.2 Future Work

7.2.1 Issue of Power Allocation

The choice of the appropriate power allocation was shown to be important for IDMA systems [85], where the reported maximum bandwidth efficiency was as high as 8 bits/symbol. An inherent assumption was the availability of perfect knowledge of all mobile stations' channel, both at the receiver as well as at the mobile transmitter. In order to eliminate the idealised assumptions of [85], the following research ideas maybe pursued:

1. Centralised and decentralised power allocation schemes may be designed based on game-theory. The decentralised power-allocation may be considered to be an instantiation of non-cooperative game aiming at Nash Equilibrium, while the centralised power-allocation may be viewed as cooperative game aiming for finding the Pareto Optimum (PO).

2. The investigation of non-cooperative game in the context of iterative receivers is novel, since all receiver algorithms considered in the open literature are simple linear receivers under the so-called Large System Assumption (LSA) [185]. Importantly, the concept of utility is exposed throughout the discussion of game-theory aided transceiver design as a result of its economic nature, while this terminology may be potentially related to the overall efficiency definition.

3. Concerning the power allocation scheme, apart from employing a HS-based new optimisation algorithm, future research should be focused on the robustness of power allocation, where a range of realistic impariments, such as the step-size, the delay, etc has to be investigated. This is expected to lead to practical power allocation schemes or discourage the employment of power allocation all together.

7.2.2 Cross Layer Design

The benefits of cross-layer designs are multi-field. As an extension of our cross-layer oriented SPC aided multiplexed HARQ scheme, we have the following further research ideas:

1. First and foremost, we will focus our attention on diverse modulation schemes. As an interesting practical aspect, asynchronous HARQ schemes may be designed. The concept of our multiplexed HARQ scheme may be compared to that of the corresponding relay aided system owing to their strong similarities [186, 187].

2. The cross-layer design of an adaptive modulation is also promising. Intuitively, delay-constrained services require a low latency as well as high-order modulation schemes [188, 189]. The lower the tolerable delay, the higher number of bits per symbol to be used. However, it is well-known that high-order modulation scheme tend to have a high BER and low power efficiency. Hence, keep the delay requirement constant, reducing the latency of TCP and HARQ may ultimately relax the choice of modulation scheme, i.e. a low-order modulation scheme may be selected. This leads to a joint design of many important aspects of the physical layer transceiver.

Glossary

2G	Second-Generation
3GPP	Third-Generation Partnership Project's
ACC	ACcumulate Code
ACO	Ant Colony Optimisation
AF	Amplify-Forward
AP	Access Point
APP	A Posteriori Probability
AS	Active Sources
AWGN	Additive White Gaussian Noise
BER	Bit-Error-Ratio
BIAWGN	Binary Input Additive White Gaussian Noise
BICM	Bit Interleaved Coded Modulation
BLAST	Bell-Labs Layered Space Time Architecture
CDec	Channel Decoder
CE	Channel Estimation
CIR	Channel Impulse Response
CMA	Cooperative Multiple Access
CP	Cyclic Prefix
CQI	Channel Quality Indicator
CS	Cooperating Sources
DBPSK	differentially encoded BPSK
DDF	Differential-Decode-Forward
DEC	DECoders
Des	De-spreader

DET	DETector
DF	Decode-Forward
DFT	Discrete Fourier Transform
DL	DownLink
DPC	Dirty Paper Coding
DPDCH	Dedicated Physical Data CHannels
DSP	Digital Signal Processing
EAs	Evolutionary Algorithms
ELN	Explicit Loss Notification
EM	Expectation Maximisation
EVD	EigenValue Decomposition
EXIT	EXtrinsic Information Transfer
FD	Frequency-Domain
FDCTF	Frequency Domain Channel Transfer Function
FDE	Frequency Domain Equalisation
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FER	Frame Error Rate
FF	*Fitness Function*
FH	Frequency Hopping
FIR	Finite Impulse Response
GA	Genetic Algorithms
GCDMA	Generalised Code Division Multiple Access
GF	Galois Field
GMAC	Gaussian Multiple Access Channel
GP	Generator Polynomials
HARQ	Hybrid Automatic Repeat reQuest
HS	Harmony Search
HSPA	High Speed Packet Access
IC	Interference Cancellation
IDD	Iterative Detection and Decoding
IDM	Interleave Division Multiplexing
IDMA	Interleave Division Multiple Access

IIR	Infinite Impulse Response
IPI	Inter-Packet-Interference
IR-CDMA	Interleaved Random Code Division Multiple Access
IR-STC	Interleaved Random Space Time Code
ISI	Inter Symbol Interference
JDD	Joint Detection and Decoding
LDC	Linear Dispersion Codes
LDPC	Low Density Parity Check
LLR	Log-Likelihood Ratio
LSA	Large System Assumption
LTE	Long Term Evolution
M-HARQ	Multiplexed HARQ
MAP	Maximum A posterior Probability
MC-DS-CDMA	Multi-Carrier Direct Sequence Code Division Multiple Access
MC-IDM	Multi-Carrier Interleave Division Multiplexing
MC-IDM-IDMA	Multi-Carrier Interleave Division Multiplexing aided Interleave Division Multiple Access
MC-IDMA	Multi-Carrier Interleave Division Multiple Access
MC-ML-IDMA	Multi-Carrier Multilayer IDMA
MCMC	Markov Chain Monte Carlo
MF	Matched Filter
MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood
MLC	MultiLevel Coding
MMSE	Minimum Mean Square Error
MPD	Multiple Packets Detection
MS	Mobile Station
MSC	Multi-Source Cooperation
MT	Mobile Terminal
MU	Multi-User
MUD	Multi-User Detector
MUI	Multi-User Interference
NACK	Negative-ACKnowledgement
NC	Network Coding

OFCDM	Orthogonal Frequency Code Division Multiplexing
OFDM	Orthogonal Frequency Division Multiplexing
OSI	Open Systems Interconnection
OSTBC	Orthogonal Space Time Block Code
OVSF	Orthogonal Variable Spreading Factor
P/S	Parallel-to-Serial
PANC	Physical-layer Algebraic Network Coding
PAPR	Peak-to-Average Power Ratio
PCM	Parity Check Matrix
PDF	Probability Density Function
PER	Packet-Error-Ratio
PIC	Parallel Interference Cancellation
PMF	Probability Mass Function
PO	Pareto Optimum
PSO	Particle Swarm Optimisation
RA	Repeat Accumulate
RCPC	Rate Compatible Punctured Codes
RLC	Radio Link Control
RS	Relaying Sources
RTT	Round Trip Time
SC-FDMA	Single Carrier Frequency Division Multiple Access
SCC	Serially Concatenated Code
SD	Sphere Decoding
SDF	Soft-Decode-Forward
SDMA	Spatial Division Multiple Access
SE	Structured Embedded
SF	Spreading Factor
SHS	Soft-HS
SI	Swarm Intelligence
SIC	Successive Interference Cancellation
SISO	Single Input Single Output
SNR	Signal-to-Noise-Ratio
SPC	SuperPosition Coding
SSC	Single Source Cooperation

STC	Space-Time Codes
SU	Single-User
TCM	Trellis Coded Modulation
TCMA	Trellis Code Multiple Access
TCP	Transport Control Protocol
TD	Time-Domain
TD-SCDMA	Time-Division Synchronised CDMA
TDD	Time Division Duplex
TDL	Tapped Delay Line
TDMA	Time Division Multiple Access
UL	UpLink
UWB	Ultra WideBand
VA	Viterbi Algorithm
VMIMO	Virtual MIMO
VMISO	Virtual Multiple Input Single Output
VT	Variance Transfer
WIMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
ZP	Zero-Padding

Bibliography

[1] R. Steele and L. Hanzo, *Mobile Radio Communications*. IEEE Press - John Wiley, 2nd ed., 1999.

[2] L. Hanzo, J. Blogh, and S. Ni, *3G, HSDPA, HSUPA and Intelligent FDD versus TDD Networking: Smart Antennas and Adaptive Modulation*. IEEE Press - John Wiley, 2008.

[3] L. Hanzo, O. Alamri, M. El-Hajjar, and N. Wu, *Advanced Space-Time Coding: Near-Capacity Sphere Packing Modulation, Cooperative Communication and Multi-functional MIMO*. IEEE Press - John Wiley, April 2009.

[4] L. Hanzo, M. Munster, B. J. Choi, and T. Keller, *OFDM and MC-CDMA for Broadcasting Multi-User Communications, WLANs and Broadcasting*. Wiley-IEEE Press, 2003.

[5] L. Hanzo, C. H. Wong, and M. S. Yee, *Adaptive Wireless Transceivers: Turbo-Coded, Space-Time Coded TDMA, CDMA and OFDM Systems*. Wiley-IEEE Press, 2002.

[6] S. Lin and D. J.Costello, *Error Control Coding: Fundamentals and Applications*. New York: Prentice-Hall, Inc., 2nd ed., 2005.

[7] H. Ekstrom, A. Furuskar, J. Karlsson, M. Meyer, S. Parkvall, J. Torsner, and M. Wahlqvist, "Technical solutions for the 3G long-term evolution," *IEEE Communications Magazine*, vol. 44, pp. 38– 45, Mar. 2006.

[8] A. Ghosh, D. R. Wolter, J. G. Andrews, and R. Chen, "Broadband wireless access with WiMax/802.16: current performance benchmarks and future potential," *IEEE Communications Magazine*, vol. 43, pp. 129–136, Feb. 2005.

[9] D. Soldani and S. Dixit, "Wireless relays for broadband access," *IEEE Communications Magazine*, vol. 46, pp. 58–66, Mar. 2008.

[10] L. Hanzo, S. X. Ng, T. Keller, and W. T. Webb, *Quadrature Amplitude Modulation: From Basics to Adaptive Trellis-Coded, Turbo-Equalised and Space-Time Coded OFDM, CDMA and MC-CDMA Systems*. Wiley-IEEE Press, 2004.

[11] A. R. Calderbank and L. H. Ozarow, "Non-equiprobable signalling on the Gaussian channel," *IEEE Transactions on Information Theory*, vol. 36, pp. 126–740, Jan. 1990.

[12] J. C. MacKay, *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003.

[13] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Transactions on Information Theory*, vol. 48, pp. 1804–1824, June 2002.

[14] M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, pp. 439 – 441, May 1983.

[15] P. Sweeney, *Error Control Coding: From Theory to Practice*. John Wiley, 2002.

[16] R. G. Gallager, "A simple derivation of the coding and some applications," *IEEE Transactions on Information Theory*, vol. 11, pp. 3–8, Jan. 1965.

[17] R. G. Gallager, *Low-Density Parity-Check Codes*. MIT Press, Cambridge, 1963.

[18] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: performance analysis, design, and iterative decoding," *IEEE Transactions on Information Theory*, vol. 44, pp. 909–926, May 1998.

[19] J. Hagenauer, "The turbo principle in mobile communications," in *Proc. of the 2004 Nordic Radio Symposium*, (Oulu, Finland), Aug.16-18, 2004.

[20] N. C. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[21] E. Zehavi, "8-PSK trellis codes for a Rayleigh channel," *IEEE Transactions on Communications*, vol. 40, pp. 873–884, May 1992.

[22] A. J. Goldsmith, *Wireless Communication*. Cambridge University Press, 2005.

[23] T. S. Rappaport, *Wireless Communications - Principles and Practice*. IEEE Press, 1996.

[24] P. Wang, J. Xiao, and P. Li, "Comparison of orthogonal and non-orthogonal approaches to future wireless cellular systems," *IEEE Vehicular Technology Magazine*, vol. 1, pp. 4–11, Sept. 2006.

[25] L. L. Yang and L. Hanzo, "Multicarrier DS-CDMA: A multiple access scheme for ubiquitous broadband wireless communications," *IEEE Communications Magazine*, vol. 41, pp. 116–124, Oct. 2003.

[26] L. L. Yang and L. Hanzo, "Performance of generalized multicarrier DS-CDMA over Nakagami-m fading channels," *IEEE Transactions on Communications*, vol. 50, pp. 956–966, June 2002.

[27] L. L. Yang and L. Hanzo, "Software-defined-radio-assisted adaptive broadband frequency hopping multicarrier DS-CDMA," *IEEE Communications Magazine*, vol. 40, pp. 174–183, Mar. 2002.

[28] L. Hanzo, L. L. Yang, E. L. Kuan, and K. Yen, *Single- and Multi-Carrier DS-CDMA: Multi-User Detection, Space-Time Spreading, Synchronisation, Networking and Standards*. Wiley-IEEE Press, 2003.

[29] F. Brannstrom, T. M. Aulin, and L. K. Rasmussen, "Iterative detectors for trellis-code multiple-access," *IEEE Transactions on Communications*, vol. 50, pp. 1478– 1485, Sept. 2002.

[30] F. Brannstrom, T. M. Aulin, and L. K. Rasmussen, "Iterative multi-user detection of trellis code multiple access using a posteriori probabilities," in *Proc. of IEEE ICC '01*, vol. 1, (Helsinki, Finland), pp. 11–15, June 11–14, 2001.

[31] P. Li, L. H. Liu, K. Y. Wu, and W. K. Leung, "Interleave-division multiple-access," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 938–947, Apr. 2006.

[32] H. Schoeneich and P. A. Hoeher, "Adaptive interleave-division multiple access-A potential air interference for 4G bearer services and wireless LANs," in *Proc. WOCN 04*, (Muscat, Oman), pp. 179 – 182, June7–9, 2004.

[33] G. Y. Liu, J. H. Zhang, P. Zhang, Y. Wang, X. T. Liu, and S. Li, "Evolution map from TD-SCDMA to FuTURE B3G TDD," *IEEE Communications Magazine*, vol. 44, pp. 54–61, Mar. 2006.

[34] T. Cover and J. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[35] R. Zhang and L. Hanzo, "Three design aspects of multicarrier interleave division multiple access," *IEEE Transactions on Vehicular Technology*, vol. 57, pp. 3607–3617, Nov. 2008.

[36] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, pp. 1727–1737, Oct. 2001.

[37] R. Zhang, L. Xu, S. Chen, and L. Hanzo, "Repeat accumulate code division multiple access and its hybrid detection," in *Proc. of IEEE ICC '08*, (Beijing, China), pp. 4790–4794, May19-23, 2008.

[38] R. Zhang and L. Hanzo, "Interleave division multiplexing aided space-time coding for high-throughput uplink cooperative communications," in *Proc. of IEEE WCNC 08'*, (Las Vegas, USA), pp. 465–469, Mar./Apr. 2008.

[39] R. Zhang and L. Hanzo, "Space-time coding for high-throughput interleave division multiplexing aided multi-source cooperation," *IEE Electronics Letters*, vol. 28, pp. 367–368, Feb. 2008.

[40] R. Zhang and L. Hanzo, "High-throughput non-orthogonal interleaved random space-time coding for multi-source cooperation," in *Proc. of IEEE Globecom '08*, (New Orleans, USA), pp. 1–5, Nov./Dec. 2008.

[41] R. Zhang and L. Hanzo, "Interleaved random space-time coding for multisource cooperations," *IEEE Transactions on Vehicular Technology*, vol. 58, pp. 2120 – 2125, May 2009.

[42] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 4, pp. 1204–1216, July 2000.

[43] R. Zhang and L. Hanzo, "Physical-layer algebraic network coding and superposition coding for the multi-source cooperation aided uplink," in *Proc. of IEEE VTC '09 Spring*, to be published.

[44] R. Zhang and L. Hanzo, "Coding schemes for energy efficient multi-source cooperation aided uplink transmission," *IEEE Signal Processing Letters*, vol. 16, pp. 438 – 441, May 2009.

[45] R. Zhang and L. Hanzo, "Superposition-coding aided multiplexed hybrid ARQ scheme for improved link-layer transmission efficiency," in *Proc. of IEEE ICC '09*, to be published.

[46] R. Zhang and L. Hanzo, "Superposition-coding aided multiplexed hybrid ARQ scheme for improved end-to-end transmission efficiency," *IEEE Transactions on Vehicular Technology*, to be published.

[47] Z. W. Geem, J. H. Kim, and G. V. Loganathan, "A new heuristic optimization algorithm: Harmony Search," *Simulation*, vol. 76, pp. 60–68, 2001.

[48] R. Zhang and L. Hanzo, "EXIT chart based joint code-rate and spreading-factor optimisation of single-carrier interleave division multiple access," in *Proc. of IEEE WCNC 07'*, (HongKong, China), pp. 735–739, Mar.11-15, 2007.

[49] R. Zhang and L. Hanzo, "Harmony search aided iterative multiuser detection and channel decoding for DS-CDMA," in *Proc. of IEEE Globecom '09*, submitted.

[50] R. Zhang and L. Hanzo, "Harmony search aided iterative channel estimation, multiuser detection and channel decoding for DS-CDMA," *IEEE Transactions on Wireless Communications*, submitted.

[51] R. Zhang and L. Hanzo, "Iterative multiuser detection and channel decoding for DS-CDMA using harmony search," *IEEE Signal Processing Letter*, submitted.

[52] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379/623–423/656, 1948.

[53] F. Kschischang, B. Frey, and H. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on Information Theory*, vol. 47, pp. 498–519, Feb. 2001.

[54] L. Hanzo, P. Cherriman, and J. Streit, *Video Compression and Communications: From Basics to H.261, H.263, H.264, MPEG4 for DVB and HSDPA-Style Adaptive Turbo-Transceivers*. IEEE Press - John Wiley, 2007.

[55] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Transactions on Information Theory*, vol. 28, pp. 55–67, Jan. 1982.

[56] M. H. DeGroot and M. J. Schervish, *Probability and Statistics*. Addison Wesley, 2001.

[57] E. G. Larsson and B. R. Vojcic, "Cooperative transmit diversity based on superposition modulation," *IEEE Communications Letters*, vol. 9, pp. 778–780, Sept. 2005.

[58] X. Ma and P. Li, "Coded modulation using superimposed binary codes," *IEEE Transactions on Information Theory*, vol. 50, pp. 3331–3343, Dec. 2004.

[59] H. Imai and S. Hirakawa, "A new multilevel coding method using error-correcting codes," *IEEE Transactions on Information Theory*, vol. 23, pp. 371–377, May 1977.

[60] M. Isaka and H. Imai, "On the iterative decoding of multilevel codes," *IEEE Journal on Selected Areas in Communication*, vol. 19, pp. 935–943, May 2001.

[61] G. Caire, G. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Transactions on Information Theory*, vol. 44, pp. 927–946, May 1998.

[62] J.G.Proakis, *Digital Communications*. New York: McGrawHill, 2001.

[63] B. Sklar, "A primer on turbo code concepts," *IEEE Communications Magazine*, vol. 35, pp. 94–102, Dec. 1997.

[64] C. Berrou and A. Glavieux, "Near optimum error-correcting coding and decoding: turbo codes," *IEEE Transactions on Communications*, vol. 44, pp. 1261–1271, Oct. 1996.

[65] B. Schlegel and L. C. Perez, *Trellis and Turbo Coding*. IEEE Press - John Wiley, 2004.

[66] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Transactions on Information Theory*, vol. 27, p. 533 ?547, Sept. 1981.

[67] N. Wiberg, *Codes and Decoding on General Graphs*. PhD thesis, Linköping University, Linköping, Sweden, 1996.

[68] J. Jin, A. Khandekar, and R. J. McEliece, "Irregular repeat-accumulate codes," in *Proc. of the 2nd International Conference on Turbo Codes*, (Munich, Germany), pp. 125–127, Sept.4 – 7, 2000.

[69] A. Tarable, G. Montorsi, and S. Benedetto, "Analysis and design of interleavers for iterative multiuser receivers in coded CDMA systems," *IEEE Transactions on Information Theory*, vol. 51, pp. 1650–1666, May 2005.

[70] H. V. Poor, "Turbo multiuser detection: A primer," *Jounal on Communications. Networks*, vol. 3, pp. 196–201, Sept. 2001.

[71] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Transactions on Information Theory*, vol. 13, pp. 260–269, Apr. 1967.

[72] L. Hanzo, T. H. Liew, and B. L. Yeap, *Turbo Coding, Turbo Equalisation and Space-Time Coding for Transmission over Fading Channels*. Wiley-IEEE Press, 2002.

[73] M. Moher, "An iterative multiuser decoder for near-capacity communications," *IEEE Transactions on Communications*, vol. 46, pp. 870–880, July 1998.

[74] M. C. Reed, C. B. Schlegel, P. D. Alexander, and J. A. Asenstorfer, "Iterative multiuser detection for CDMA with FEC: Near-single-user performance," *IEEE Transactions on Communications*, vol. 46, pp. 1693–1699, Dec. 1998.

[75] X. D. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Transactions on Communications*, vol. 47, pp. 1046–1061, July 1999.

[76] P. D. Alexander, A. J. Grant, and M. C. Reed, "Iterative detection in code-division multiple access with error control coding," *European Transactions Telecommunications*, vol. ETT-9, pp. 419–425, 1998.

[77] J. Y. Hui, "Throughput analysis for code division multiple accessing of the spread spectrum channel," *IEEE Journal on Selected Areas in Communications*, vol. SAC-2, pp. 482–486, 1984.

[78] A. J. Viterbi, "Spread spectrum communications–myths and realities," *IEEE Communications Magazine*, vol. 17, pp. 11–18, May 1979.

[79] A. J. Viterbi, "Very low rate convolutional codes for maximum theoretical performance of spread spectrum multiple-access channels," *IEEE Journal on Selected Areas in Communications*, vol. 8, pp. 641–649, Aug. 1990.

[80] S. Verdu and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Transactions on Information Theory*, vol. 45, pp. 622–640, Mar. 1999.

[81] P. Frenger, P. Orten, and T. Ottosson, "Code-spread CDMA using maximum free distance low-rate convolutional codes," *IEEE Transactions on Wireless Communications*, vol. 48, pp. 135 – 144, Jan. 2000.

[82] D. Gargy and F. Adachi, "Comprehensive evaluation of chip interleaving effect on turbo-coded DS-SS in a Rayleigh fading channel with antenna diversity reception," *Wireless Communications and Mobile Computing*, vol. 6, pp. 49–60, 2006.

[83] R. H. Mahadevappa and J. G. Proakis, "Mitigating multiple access interference and inter-symbol interference in uncoded CDMA systems with chip-level interleaving," *IEEE Transactions on Wireless Communications*, vol. 1, pp. 781–792, Oct. 2002.

[84] X. Gui and T. S. Ng, "A novel chip-interleaving DS SS system," *IEEE Transactions on Vehicular Technology*, vol. 49, pp. 21–27, Jan. 2000.

[85] L. H. Liu, J. Tong, and P. Li, "Analysis and optimization of CDMA systems with chip-level interleavers," *IEEE Journal on Selected Areas in Communication*, vol. 24, pp. 141 – 150, Jan. 2006.

[86] P. A. Hoeher and H. Schoeneich, "Interleave-division multiple access from a multiuser point of view," in *Proc. 5th Int. Symposium on Turbo Codes and Related Topics in connection with the 6th Int. ITG-Conference on Source and Channel Coding*, (Munich, Germany), pp. 140 – 144, Apr.3–7, 2006.

[87] K. Li, X. D. Wang, and P. Li, "Analysis and optimization of interleave-division multiple-access communication systems," in *Proc. of IEEE ICASSP '05*, (Philadelphia, USA), pp. iii/917– iii/920, Mar.18–23, 2005.

[88] P. Li, W. K. Leung, and K. Y. Wu, "Low-rate turbo-Hadamard codes," *IEEE Transactions on Information Theory*, vol. 49, pp. 3213–3224, Dec. 2003.

[89] P. Li, X. Huang, and N. Phamdo, "Zigzag codes and concatenated zigzag codes," *IEEE Transactions on Information Theory*, vol. 47, pp. 800–807, Feb. 2001.

[90] P. Wang, P. Li, and L. H. Liu, "Power allocation for multiple access systems with practical coding and iterative multiuser detection," in *Proc. of IEEE ICC'06*, (Istanbul, Turkey), pp. 260–269, June11–15, 2006.

[91] P. A. Hoeher and H. Schoeneich, "Semi-blind pilot-layer aided channel estimation with emphasis on interleave-division multiple access systems," in *Proc. of IEEE GLOBECOM '05*, (St. Louis, Mo, USA), pp. 3513–3517, Nov-Dec 2005.

[92] K. Kusume and G. Bauch, "CDMA and IDMA: Iterative multiuser detections for near-far asynchronous communications," in *IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications, 2005*, (Berlin, Germany), pp. 426– 431, Sept.11–14, 2005.

[93] M. M. Butt, J. C. Fricke, and P. A. Hoeher, "Reliability-based packet combining with application to interleave-division multiple access," in *Proc. 5th Int. Symposium on Turbo Codes*

and Related Topics in connection with the 6th Int. ITG-Conference on Source and Channel Coding, (Munich, Germany), pp. 140 – 144, Apr.3–7, 2006.

[94] Q. Huang, S. Chan, K. T. Ko, P. Li, and P. Wang, "A QoS architecture for IDMA-based multi-service wireless networks," in *Proc. of IEEE ICC'07*, (Glasgow, U.K.), pp. 5070–5075, June24-28, 2007.

[95] L. H. Liu and P. Li, "A comparative study on low-cost multiuser detectors," in *Proc. of IEEE ICC'06*, (Istanbul, Turkey), pp. 4947 – 4952, June11–15, 2006.

[96] I. M. Mahafeno, C. Langlais, and C. Jégo, "Reduced complexity iterative multi-user detector for IDMA (Interleave-Division Multiple Access) system," in *Proc. of IEEE GLOBECOM'06*, (USA), pp. 140–144, 2006.

[97] K. Kusume, G. Dietl, W. Utschick, and G. Bauch, "Performance of interleave division multiple access based on minimum mean square error detection," in *Proc. of IEEE ICC'07*, (Glasgow, U.K.), pp. 2961–2966, June24-28, 2007.

[98] K. Li, X. D. Wang, and P. Li, "Analysis and optimization of interleave-division multiple-access communication systems," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 1973–1983, May 2007.

[99] Q. H. Guo and P. Li, "LMMSE turbo equalization based on factor graphs," *IEEE Transactions on Selected Areas in Communications*, vol. 26, pp. 311–319, Feb. 2008.

[100] O. Nagy, M. C. Reed, and Z. N. Shi, "Optimal detection of IDMA signals," in *Proceedings of WCNC '07*, (Hong Kong, China), pp. 1236–1240, Mar.11-15, 2007.

[101] Z. N. Shi, C. B. Schlegel, R. Kempter, and M. C. Reed, "On the performance of partitioned-spreading CDMA," in *40th Conference on Information Sciences and Systems*, (Princeton, NJ), pp. 5508–5513, Mar.22-24, 2006.

[102] K. Li, X. D. Wang, G. S. Yue, and P. Li, "A low-rate code-spread and chip-interleaved time-hopping UWB system," *IEEE Journal on Selected Areas in Communication*, vol. 24, pp. 864 – 870, Apr. 2006.

[103] P. Weitkemper and K. D. Kammeyer, "Power optimization of IDMA systems with different target BER constraints," in *Proceedings of VTC '07 Spring*, (Dublin, Ireland), pp. 2812–2816, Apr.22-25, 2007.

[104] S. K. Lau and W. Y. Yue, "A linear zero-one formulation of optimal power allocation of IDMA systems," in *Proceedings of Wireless Communications, Networking and Mobile Computing (WiCom 2007)*, (New York, USA), pp. 637–640, Sept.21-25, 2007.

[105] S. Zhou, Y. Li, M. Zhao, X. Xu, J. Wang, and Y. Yao, "Novel techniques to improve downlink multiple access capacity for Beyond 3G," *IEEE Communications Magazine*, vol. 43, pp. 61 – 69, jun 2005.

[106] K. Kusume and G. Bauch, "Some aspects of interleave division multiple access in ad hoc networks," in *Proc. 5th Int. Symposium on Turbo Codes and Related Topics in connection with the 6th Int. ITG-Conference on Source and Channel Coding*, (Munich, Germany), pp. 135 – 139, Apr.3–7, 2006.

[107] I. Mahafeno, C. Langlais, and C. Jego, "OFDM-IDMA versus IDMA with ISI cancellation for quasi-static Rayleigh fading multipath channels.," in *Proc. of 4th Int. Symposium on Turbo Codes and Related Topics in connection with the 6th Int. ITG-Conference on Source and Channel Coding*, (Munich, Germany), pp. 140–144, Sept.3-7, 2006.

[108] R. Zhang and L. Hanzo, "Iteratively detected multicarrier interleave division multiple access," in *Proc. of IEEE MICROCOLL 07'*, (Budapest, Hungary), pp. 381–385, May14-16, 2007.

[109] P. Li, Q. H. Guo, and J. Tong, "The OFDM-IDMA approach to wireless communication systems," *IEEE Transactions on Wireless Communications*, vol. 14, pp. 18–24, June 2007.

[110] Q. H. Guo, X. J. Yuan, and P. Li, "Single- and multi- carrier IDMA schemes with cyclic prefixing and zero padding techniques," *European Transactions on Telecommunications*, vol. 14, pp. 18–24, May 2008.

[111] P. Hoeher and W. Xu, "Multi-layer interleave-division multiple access for 3GPP long term evolution," in *Proceddings of ICC '07*, (Glasgow, U.K.), pp. 5508–5513, June24-28, 2007.

[112] K. Li and X. D. Wang, "EXIT chart analysis of turbo multiuser detection," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 300–311, Jan. 2005.

[113] G. S. Yue and X. D. Wang, "Coding-spreading tradeoff in LDPC-coded CDMA with turbo multiuser detection," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 1734–1745, Sept. 2004.

[114] H. Wei and L. Hanzo, "Coding versus spreading gain optimization of a nonbinary BCH coded CDMA system," *IEE Electronics Letters*, vol. 41, pp. 816–817, July 2005.

[115] L. Hanzo and T. Keller, *OFDM and MC-CDMA: A Primer*. Wiley-IEEE Press, 2006.

[116] J. Boutros and G. Caire, "Iterative multiuser decoding: unified framework and asymptotic performance analysis," *IEEE Transactions on Information Theory*, vol. 48, pp. 1772–1793, 2002.

[117] Z. N. Shi and C. B. Schlegel, "Iterative multiuser detection and error control code decoding in random CDMA," *IEEE Transactions on Signal Processing*, vol. 54, pp. 1886– 1895, May 2006.

[118] V. V. Veeravalli and A. Mantravadi, "The coding-spreading tradeoff in CDMA systems," *IEEE Transactions on Communications*, vol. 20, pp. 396–408, Feb. 2002.

[119] Z. Q. Tang and W. E. Ryan, "Achievable information rates and the coding-spreading tradeoff in finite-sized synchronous CDMA systems," *IEEE Transactions on Wireless Communications*, vol. 53, pp. 1432–1437, Sept. 2005.

[120] S. Shamai and S. Verdu, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Transactions on Information Theory*, vol. 47, pp. 1302–1327, May 2001.

[121] L. Xu, R. Zhang, S. Chen, and L. Hanzo, "EXIT-chart aided hybrid multiuser detector design for frequency-domain-spread chip-interleaved MC-CDMA," in *Proc. of IEEE VTC '08 Spring*, (Marina Bay, Singapore), pp. 1816 – 1820, May11-14, 2008.

[122] P. Li, W. K. Leung, and K. Y. Wu, "Low-rate turbo-Hadamard codes," *IEEE Transactions on Information Theory*, vol. 49, pp. 3213–3224, Dec. 2003.

[123] N. Varnica, A. Kavcic, X. Ma, and P. Li, "Density evolution and LDPC code optimization for interleaver division multiple accesss," in *Proc. of IEEE GMC 04'*, (Shanghai, China), pp. 381–385, May11-13, 2004.

[124] H. Jin, *Analysis and Design of Turbo-like Codes*. Ph.D. Thesis, California Institute of Technology, Pasadena, 2001.

[125] I. Pupeza, A. Kavcic, and P. Li, "Efficient generation of interleavers for IDMA," in *Proc. of IEEE ICC 06'*, (Istanbul, Turkey), pp. 196–201, June11-15 2006.

[126] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communication*, vol. 16, pp. 1451–1458, Oct. 1998.

[127] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–355, Mar. 1998.

[128] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. Part I. System description," *IEEE Transactions on Communications*, vol. 51, pp. 1927–1938, Nov. 2003.

[129] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. Part II. Implementation aspects and performance analysis," *IEEE Transactions on Communications*, vol. 51, pp. 1939–1948, Nov. 2003.

[130] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. Part I and II," *IEEE Transactions on Communications*, vol. 51, pp. 1927–1948, Nov. 2003.

[131] R. Pabst, B. H. Walke, D. C. Schultz, P. Herhold, H. Yanikomeroglu, S. Mukherjee, H. Viswanathan, M. Lott, W. Zirwas, M. Dohler, H. Aghvami, D. D. Falconer, and G. P. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," *IEEE Communications Magazine*, vol. 42, pp. 80–89, Sept. 2004.

[132] A. Ozgur, O. Leveque, and N. C. Tse, "Hierarchical cooperation achieves optimal capacity scaling in ad hoc networks," *IEEE Transactions on Information Theory*, vol. 53, pp. 3549–3572, Oct. 2007.

[133] J. N. Laneman, N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[134] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Transactions on Information Theory*, vol. 51, pp. 3037–3063, Sept. 2005.

[135] K. Azarian, H. E. Gamal, and P. Schniter, "On the achievable diversity multiplexing tradeoff in half-duplex cooperative channels," *IEEE Transactions on Information Theory*, vol. 51, pp. 4152–4172, Dec. 2005.

[136] O. Shalvi, "Multiple source cooperation diversity," *IEEE Communications Letters*, vol. 8, pp. 712–714, Dec. 2004.

[137] A. Ribeiro, R. Q. Wang, and G. B. Giannakis, "Multi-source cooperation with full-diversity spectral-efficiency and controllable-complexity," *IEEE Journal on Selected Areas in Communications*, vol. 25, pp. 415–425, Feb. 2007.

[138] N. C. Tse, P. Viswanath, and L. Z. Zheng, "Diversity-multiplexing tradeoff in multiple-access channels," *IEEE Transactions on Information Theory*, vol. 50, pp. 1859–1874, Sept. 2004.

[139] M. Sellathurai and S. Haykin, "Turbo-BLAST for wireless communications: theory and experiments," *IEEE Transactions on Signal Processing*, vol. 50, pp. 2538–2546, Oct. 2002.

[140] K. Y. Wu and P. Li, "Multilayer turbo space-time codes," *IEEE Transactions on Communications*, vol. 9, pp. 55–57, Jan. 2005.

[141] C. Li, K. Li, X. D. Wang, and P. Li, "An interleave-division-multiplexing MISO system with partial CSI at transmitter," *IEEE Transactions on Vehicular Technology*, vol. 56, pp. 1197–1208, May 2007.

[142] B. S. Mergen and A. Scaglione, "Randomized space-time coding for distributed cooperative communication," *IEEE Transactions on Signal Processing*, vol. 55, pp. 5003–5017, Oct. 2007.

[143] Y. D. Chen, S. Kishore, and J. Li, "Wireless diversity through network coding," in *Proc. of IEEE WCNC '06*, (Las Vegas, USA), pp. 1681–1686, Apr.3-6, 2006.

[144] C. Hausl and J. Hagenauer, "Iterative network and channel decoding for the two-way relay channel," in *Proc. of IEEE ICC '06*, (Istanbul, Turkey), pp. 1568–1573, June11-15, 2006.

[145] C. Hausl and P. Dupraz, "Joint network-channel coding for the multiple-access relay channel," in *Proc. of IEEE Sensor and Ad Hoc Communications and Networks (SECON) '06*, (Hyatt Regency, USA), pp. 817–822, Sept.25-28, 2006.

[146] N. Fawaz, D. Gesbert, and M. Debbah, "When network coding and dirty paper coding meet in a cooperative ad hoc network," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 1862–1867, May 2008.

[147] L. Xiao, T. Fuja, J. Kliewer, and D. Costello, "A network coding approach to cooperative diversity," *IEEE Transactions on Information Theory*, vol. 53, pp. 3714–3722, Oct. 2007.

[148] T. R. Wang and G. B. Giannakis, "Complex field network coding for multiuser cooperative communications," *IEEE Journal on Selected Areas in Communications*, vol. 26, pp. 561–571, Apr. 2008.

[149] S. ten Brink and G. Kramer, "Design of repeat-accumulate codes for iterative detection and decoding," *IEEE Transactions on Signal Processing*, vol. 51, pp. 2764–2772, Nov. 2003.

[150] G. Caire and D. Tuninetti, "The throughput of hybrid-ARQ protocols for the Gaussian collision channel," *IEEE Transactions on Information Theory*, vol. 47, pp. 1971–1988, July 2001.

[151] S. Sesia, G. Caire, and G. Vivier, "Incremental redundancy hybrid ARQ schemes based on low-density parity-check codes," *IEEE Transactions on Communications*, vol. 52, pp. 1311–1321, Aug. 2004.

[152] C. F. Leanderson and G. Caire, "The performance of incremental redundancy schemes based on convolutional codes in the block-fading Gaussian collision channel," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 843–854, May 2004.

[153] H. E. Gamal, G. Caire, and M. O. Damen, "The MIMO ARQ channel: diversity multiplexing delay tradeoff," *IEEE Transactions on Information Theory*, vol. 52, pp. 3601–3621, Aug. 2006.

[154] B. Zhao and M. C. Valenti, "Practical relay networks: a generalization of hybrid-ARQ," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 7–18, Jan. 2005.

[155] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Transactions on Communications*, vol. 36, pp. 389–400, Apr. 1988.

[156] A. Shokrollahi, "Raptor codes," *IEEE Transactions on Information Theory*, vol. 52, pp. 2551–2567, June 2006.

[157] N. Bonello, R. Zhang, S. Chen, and L. Hanzo, "Reconfigurable rateless codes," submitted to IEEE Transactions on Communications.

[158] A. Roongta, J. W. Moon, and J. M. Shea, "Reliability-based hybrid ARQ as an adaptive response to jamming," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 1045–1055, May 2005.

[159] P. Frenger, S. Parkvall, and E. Dahlman, "Performance comparison of HARQ with chase combining and incremental redundancy for HSDPA," in *Proc. of IEEE VTC '01 Fall*, (Atlantic City, New Jersey, USA), pp. 1829–1833, Oct.7-11, 2001.

[160] N. Miki, H. Atarashi, K. Higuchi, S. Abeta, and M. Sawahashi, "Experimental evaluation on effect of hybrid ARQ with packet combining in forward link for VSF-OFCDM broadband wireless access," in *Proc. of IEEE PIMRC '03*, (Beijing, China), pp. 360–365, Sept.7-10, 2003.

[161] F. Frederiksen and T. E. Kolding, "Performance and modeling of WCDMA/HSDPA transmission/H-ARQ schemes," in *Proc. of IEEE VTC '02 Fall*, (Vancouver, BC, Canada), pp. 472–476, Sept.24-28, 2002.

[162] J. F. Cheng, "Coding performance of hybrid ARQ schemes," *IEEE Transactions on Communications*, vol. 54, pp. 1017– 1029, June 2006.

[163] S. Shakkottai, T. S. Rappaport, and P. C. Karlsson, "Cross-layer design for wireless networks," *IEEE Communications Magazine*, vol. 41, pp. 74– 80, Oct. 2003.

[164] A. Bakre and B. R. Badrinath, "I-TCP: indirect TCP for mobile hosts," in *Proc. of IEEE 15th International Conference on Distributed Computing Systems*, (Vancouver, Canada), pp. 136–143, May28-31 1995.

[165] H. Balakrishnan, S. Seshan, and R. Katz, "Improving reliable transport and handoff performance in cellular wireless networks," *ACM Wireless Networks*, vol. 4, pp. 129–136, Dec. 1995.

[166] R. Caceres and L. Iftode, "Improving the performance of reliable transport protocols in mobile computing environments," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 850–857, June 1995.

[167] C. Hausl and A. Chindapol, "Hybrid ARQ with cross-packet channel coding," *IEEE Communications Letters*, vol. 11, pp. 434–436, May 2007.

[168] H. Balakrishnan, V. N. Padmanabhan, S. Seshan, and R. H. Katz, "A comparison of mechanisms for improving TCP performance overwireless links," *IEEE/ACM Transactions Networking*, vol. 5, pp. 756–769, Dec. 1997.

[169] A. Ribeiro, X. D. Cai, and G. B. Giannakis, "Opportunistic multipath for bandwidth-efficient cooperative multiple access," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 2321–2327, Sept. 2006.

[170] W. R. Stevens, *TCP/IP Illustrated, Volume I: The Protocols*. MA: Addison-Wesley, 1994.

[171] L. Kleinrock, *Queueing Systems, Volume I and II*. New York: Wiley, 1976.

[172] Q. Q. Huang, S. Chan, P. Li, and M. Zukerman, "Improving wireless TCP throughput by a novel TCM-based hybrid ARQ," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 2476–2485, 2007.

[173] A. Bakre and B. R. Badrinath, "Improving TCP over wireless through adaptive link layer setting," in *Proc. of IEEE GLOBECOM '01*, (San Antonio, USA), pp. 1766–17703, Nov.25-29 2001.

[174] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Transactions on Information Theory*, vol. 47, pp. 619–637, Feb. 2001.

[175] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Transactions on Communications*, vol. 51, pp. 389–399, Mar. 2003.

[176] R. Chen, J. S. Liu, and X. D. Wang, "Convergence analysis and comparisons of Markov Chain Monte Carlo algorithms in digital communications," *IEEE Transactions on Signal Processing*, vol. 50, pp. 255–270, Feb. 2002.

[177] B. Farhang-Boroujeny, H. D. Zhu, and Z. N. Shi, "Markov Chain Monte Carlo algorithms for CDMA and MIMO communication systems," *IEEE Transactions on Signal Processing*, vol. 54, pp. 1896– 1909, May 2006.

[178] K. Yen and L. Hanzo, "Genetic algorithm assisted joint multiuser symbol detection and fading channel estimation for synchronous CDMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 985–998, June 2001.

[179] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. of the IEEE Int. Conf. on Neural Networks*, (Piscataway, NJ), pp. 1942–1948, 1995.

[180] M. Dorigo and L. Gambardella, "Ant colony system: A cooperative learning approach to the traveling salesman problem," *IEEE Transactions on Evolutionary Computation*, vol. 1, pp. 53–66, Apr. 1997.

[181] C. Xu, B. Hu, L. L. Yang, and L. Hanzo, "Ant-colony-based multiuser detection for multi-functional antenna array assisted MC DS-CDMA systems," *IEEE Transactions on Vehicular Technology*, vol. 57, pp. 658–663, Jan. 2008.

[182] M. Jiang, J. Akhtman, and L. Hanzo, "Iterative joint channel estimation and multi-user detection for multiple-antenna aided OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 2904–2914, Aug. 2007.

[183] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society, Series B*, vol. 39, pp. 1–38, Jan. 1977.

[184] M. Kobayashi, J. Boutros, and G. Caire, "Successive interference cancellation with SISO decoding and EM channel estimation," *IEEE Journal On Selected Areas In Communications*, vol. 19, pp. 1450–1460, Aug. 2001.

[185] S. Buzzi and H. V. Poor, "Joint receiver and transmitter optimization for energy-efficient CDMA communications," *IEEE Journal on Selected Areas in Communications*, vol. 26, pp. 459–472, Sept. 2008.

[186] B. Zhao and M. C. Valenti, "Practical relay networks: a generalization of hybrid-ARQ," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 7–18, Jan. 2005.

[187] I. Stanojev, O. Simeone, Y. Bar-Ness, and C. You, "Performance of multi-relay collaborative hybrid-ARQ protocols over fading channels," *IEEE Communications Letters*, vol. 10, pp. 522–524, July 2006.

[188] F. Meshkati, A. J. Goldsmith, H. V. Poor, and S. C. Schwartz, "A game-theoretic approach to energy-efficient modulation in CDMA networks with delay constraints," *IEEE Journal on Selected Areas in Communications*, vol. 25, pp. 1069–1078, Aug. 2007.

[189] X. Wang, Q. Liu, and G. B. Giannakis, "Analyzing and optimizing adaptive modulation-coding jointly with ARQ for QoS-guaranteed traffic," *IEEE Transactions on Vehicular Technology*, vol. 56, pp. 710–719, Mar. 2007.

Subject Index

T

U

V

W

Z

Author Index

I

J

K

L

M

T

U

V

W

X

Y

Z