University of Alberta

Iterative Receivers for Interference Limited Environments

by

Łukasz Krzymień

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Communications

Department of Electrical and Computer Engineering

©Łukasz Krzymień
Spring 2011
Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author reserves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

Examining Committee

Supervisor: Dr. Christian Schlegel, Department of Computing Science

Co-Supervisor: Dr. Yindi Jing, Department of Electrical and Computer Engineering

Examiner: Dr. Ivan Fair, Department of Electrical and Computer Engineering

Examiner: Dr. Ioanis Nikolaidis, Department of Computing Science

Examiner: Dr. Paul Ho, School of Engineering Science, Simon Fraser University

To my parents, Anna and Piotr, and my brother Kuba
- for their love, never-ending support and faith in me

Abstract

Interference dominated wireless communications systems are considered. Joint detection methods are applied to combat the negative effects of the temporal and spatial interference. However, practical joint detectors are not commonly used due to their high complexity. Therefore, there is a constant need to deliver reduced complexity solutions that approach substantial fraction of the channel capacity.

In the first part of this thesis it is shown that simple transmission technique employing repetition coding and interleaving combined with interference cancellation is an asymptotically optimal processing strategy when high interference is experienced, for instance due to the relatively high ratio of the number of signals to the number of orthogonal dimensions. Surprisingly, strong, capacity achieving codes exhibit inferior performance and are not well suited for iterative interference cancellation due to their sharp threshold characteristics. Motivated by this observation partitioned modulation is introduced and applied to a multiuser spread spectrum system, which inherently encompasses a repetition coding mechanism. The detection of the resulting signals employs a parallel interference cancellation approach, where the repetition code exchanges information iteratively with the canceller. Precise signal-to-noise ratio evolution of the proposed receiver as a function of demodulation iterations is given. It is shown, that for equal received power system, partitioned demodulator outperforms linear minimum mean squared error processor at a fraction of complexity. This receiver processing for a wide range of parameters delivers estimates that coincide with the optimal processing based on exhaustive search. For unequal received signal powers these advantages are even more visible and for a particular exponential power allocation

the proposed system reaches the capacity of the channel. The analytical investigations are verified using computer simulations.

In the second part of this dissertation, multi-user MIMO systems compliant with 3GPP LTE standard are considered. Turbo near-far resistant interference cancellation receiver is proposed. It jointly removes multi-user, multi-antenna and inter-symbol interference and outperforms traditional demodulation/decoding method adopted in the LTE standard. Semi-analytical method of predicting the performance of this joint receiver for any system setup is outlined. This approach makes it possible to tune up the performance of the system without running extensive bit-error-rate simulations.

Acknowledgements

First, I would like to express my greatest appreciation to my supervisor Prof. Christian Schlegel for providing me with many new ideas and truly broadening my horizons in the field of scientific research as well as ordinary life.

I would also like to thank my examining committee: Dr. Yindi Jing, Prof. Ivan Fair, Prof. Ioanis Nikolaidis and Prof. Paul Ho for many insightful comments and thoroughly reviewing my thesis.

To my colleagues at the HCDC laboratory, a very special thanks to Charmaine for a good time, laugh, fun and meaningful help whenever I needed it. Saeed with whom I had many discussions on various interesting topics and received practical advice on how to live life so that each day counts. Sumeeth for his kindness and Sheehan for being my beer/movies buddy. Dmitri for helping me with my research projects at the initial stage of my program and Russell for patiently polishing my thesis grammar-wise. Additionally Wesam, Shuai, Marcel, Majid and other present and former members of the lab.

To my Polish friends, especially Paweł, for a true friendship and help whenever I needed it. During my PhD journey you were my closest friend and I deeply believe this bond will last forever. Subsequently, I want to sincerely acknowledge my friend Gosia for her very patient listening during times of frustration caused by the reality of everyday life and providing me mental help to get over it. The rest of my gratitude goes to Agata and Marcin, Łukasz, Michał, Rafał and Wojtek just for being there. Finally, my special thanks I direct to my good friend Piotr "Drabol" Drabik, whose premature departure from this world made me once again realize that life is very fragile and can end unexpectedly, so living it to the fullest and having a true passion is what only matters at the end.

Passion provides balance in our lives, it can add flavor and finding it is very crucial. I found mine through salsa and other latin dances and it has been very important part of my life for the past 3.5 years. Consequently, here I want to thank my beloved Salsaddiction NYLA Dance Team directed by Daiana for providing me help with considerably improving my dance skills and making new friendships. Special thanks go to the following team members:

Ming for being there when needed, Nicole for unbelievable kindness, the greatest I have encountered in Canada so far (I wish more people were like you) and Maxwell for a good spirit. I also thank Lee, Sarah, Miranda, Katy and Ali who shared with me an enthusiasm for salsa, but due to many more similarities between us, our friendships extended far beyond just dancing. Finally, I want to thank all the Funky Buddha and On The Rocks crowd for being the reason of many extremely positive moments of my life in Edmonton.

Last, but not least, I would like to convey my sincerest gratitude to my uncle, Prof. Witold Krzymień, for his invaluable advice, as well as his encouragement to pursue the PhD program at the University of Alberta. Without him my life would probably have taken a very different path.

Contents

List of Tables

List of Figures

List of Symbols and Abbreviations

Symbols

$\mathbf{A}$	A diagonal matrix of the received signal amplitudes
$\mathbb{E}$	Expectation
$\mathbf{F}$	Fourier matrix
$\mathbf{H} = [h_{ij}]$	Complex channel gain matrix
$\mathbf{I}$	Identity matrix
K	Number of mobile users
N_r	Number of receive antennas
N_t	Number of transmit antennas
P_k	Power of k-th user (stream)
R	Channel correlation matrix $\mathbf{R} = \mathbf{H}^{\mathrm{H}}\mathbf{H}$
$\mathbf{S}$	Matrix of spreading sequences
$\mathbf{W}$	$\mathbf{W} = \mathbf{A}\mathbf{A}^{\mathrm{H}}$
α	Channel aspect ratio (load)
det	Determinant of the matrix
$\mathbf{H}^{\mathrm{H}}$	Hermitian (conjugate transpose) of $\mathbf{H}$
$\mathbf{y}$	Received signal

A

APP	A-Posteriori Probability
ARQ	Automatic Repeat-reQuest
AWGN	Additive White Gaussian Noise

B

BCJR	Bahl-Cocke-Jelinek-Raviv Algorithm
BER	Bit Error Rate
BLER	Block Error Rate
BPSK	Binary Phase-Shift Keying
BS	Base Station

C

CDF	Cumulative Distribution Function
CDMA	Code Division Multiple-Access
CP	Cyclic Prefix
CQI	Channel Quality Indicator
CRC	Cyclic Redundancy Check

D

DFT	Discrete Fourier Transform

E

EDGE	Enhanced Data for Global Evolution
eNodeB	Evolved NodeB
E-UTRA	Enhanced Universal Terrestrial Radio Access
EXIT	Extrinsic Information Transfer (charts)

F

FDD	Frequency Division Duplex
FEC	Forward Error Correction
FER	Frame Error Rate
FFT	Fast Fourier Transform

G

GMAC	Gaussian Multiple-Access Channel
GPRS	General Packet Radio Service
GSM	Global System for Mobile (Communications)

H

HARQ	Hybrid Automatic Repeat-reQuest
HSDPA	High-Speed Downlink Packet Access
HSPA+	High Speed Packet Access Evolution
HSPA	High Speed Packet Access
HSUPA	High-Speed Uplink Packet Access

I

IC	Interference Cancellation
IDFT	Inverse Discrete Fourier Transform
IEEE	Institute of Electrical and Electronics Engineers
IFFT	Inverse Fast Fourier Transform
i.i.d.	Independent and Identically-Distributed

IMT	International Mobile Telecommunications
IP	Internet Protocol
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
ISI	Inter-Symbol Interference
ISM	Industrial, Scientific and Medical (Radio Bands)

L

LAN	Local Area Network
LDPC	Low Density Parity Check (codes)
LLR	Log-Likelihood Ratio
LOS	Line-of-Sight
LTE-A	LTE-Advanced
LTE	Long Term Evolution

M

MAC	Media Access Control (Layer)
MAN	Metropolitan Area Network
MAP	Maximum A-Posteriori
MCS	Modulation and Coding Scheme
MF	Matched Filter, Correlation Receiver
MIMO	Multiple-Input and Multiple-Output
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
MS	Mobile Station

N

NLOS	Non-Line-of-Sight
NodeB	Base station in WCDMA systems

O

OFDMA	Orthogonal Frequency-Division Multiple-Access
OFDM	Orthogonal Frequency-Division Multiplexing

P

P/S	Parallel-to-Serial
PAPR	Peak-to-Average Power Ratio
PDF	Probability Density Function
PHY	Physical (Layer)
PIC	Parallel Interference Cancellation

Q

QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase-Shift Keying

R

RA	Repeat Accumulate (code)
RB	Resource Block
RC	Repetition Code
RE	Resource Element
RF	Radio Frequency
RLC	Radio Link Control (Layer)

S

S/P	Serial-to-Parallel
SAE	System Architecture Evolution
SC-FDMA	Single-Carrier Frequency-Division Multiple-Access
SCM	Spatial Channel Model
SIC	Successive Interference Cancellation
SINR	Signal to Interference-plus-Noise Ratio
SNR	Signal to Noise Ratio
SISO	Single-Input and Single-Output (channel)
SISO	Soft-Input and Soft-Output (decoder)
SMS	Short Message Service
SPA	Sum-Product Algorithm
SPC	Single Parity Check (code)
SRS	Sounding Reference Signal
SVD	Singular Value Decomposition

T

TCP	Transmission Control Protocol
TDD	Time Division Duplex
TDMA	Time-Division Multiple-Access
TEQ	Turbo Equalization
3GPP	3^{rd} Generation Partnership Project
3GPP TS	3GPP Technical Specification
3GPP2	3^{rd} Generation Partnership Project 2
TTI	Transport Transmission Interval

U

UE	User Equipment
UMTS	Universal Mobile Telecommunications System

UTRA	Universal Terrestrial Radio Access

V

VLSI	Very-Large-Scale Integration
VoIP	Voice over Internet Protocol
VT	Variance Transfer (charts)

W

W-CDMA	Wideband CDMA
WiFi	Wireless LAN Network, IEEE 802.11 standard
WiMAX	Worldwide Interoperability for Microwave Access

Z

ZF	Zero-Forcing

CHAPTER 1

Introduction

It all started with the invention of the telephone by Alexander Graham Bell in the 19th century and the first wireless long distance transmission established by Guglielmo Marconi in the early 1900s. Since then a number of important discoveries have contributed to a rapid development of communications engineering. Shortly after, the thought of interconnecting a number of terminals to form a network was conceived. This invention made communications more commonly available to the extent that it became a strategic point in the development of countries, similar to that sparked by the railway system in early 1800s. With many advances in solid state physics, mainly the invention of the transistor, the digital revolution slowly started to emerge. Its first appearance is perhaps connected to PCM digital transmission using wireline network and in principle these ideas are still used today. The pinnacle of the wireline digital boom is marked with the invention of Ethernet and the development of the US military network ARPANET, which after a number of gradual improvements transformed into the internet. It is very likely the most influential human invention of the second half of the 20th century. The internet changed life of not only individuals, but whole nations. It rapidly caused the world to shrink. Worldwide access to communications is easily available and what previous generations did not even dream of, is reality today. A plethora of new services to communicate using all human senses was and is being constantly developed up to the point, that mankind's dependence on the unlimited access to information has never been as high before. Contemporary society has for the last decade witnessed another revolution in information processing and transfer. It has been vastly influenced by the unprecedented success of the GSM cellular communications system. The number of mobile subscribers exceeded that of wireline subscribers a long time ago. Until recently those provided mostly voice services with some very limited data oriented applications, like for instance short messaging service - SMS. Now, due to the existence and tremendous success of the internet, the focus is quickly moving onto packet type connections, which can carry

any type of traffic including traditional voice communications using the VoIP technology. Due to rapid development of data oriented systems like GPRS, its improvement EDGE and subsequently UMTS, ubiquitous wireless internet access is quickly becoming a reality. Practically everyone can contact anyone, anytime from almost any place in the world and at a reasonable cost, which is mainly influenced by high government fees of the spectrum usage.

There are hundreds of new services which create an excess of new problems and challenges. They can be intuitively summarized in the following. A rapid increase in the number of mobile users, each demanding more sophisticated and data-hungry services, requires extremely efficient use of available resources like bandwidth and power. Therefore, the requirements on efficiency have never been so high before. The power/bandwidth efficiency is crucial, especially for wireless systems, and as stated by Shannon [1], together with other information theoreticians, for a given system model has a certain maximum value which, for error-free transmission, cannot be exceeded.

This thesis aims at the development and mainly improvement of methods that optimize the power/bandwidth trade-off for a number of communication situations, predominantly those in which a number of independent users communicate to a common receiver, a situation that mimics the uplink (or forward link) transmission in cellular communications systems. Multiuser detection algorithms play a central role in achieving these goals. Joint detection is a powerful method of separating signals embedded in the interference and noise of coexisting transmissions. It was first adopted and later developed in the context of user isolation in multiple-access spread spectrum CDMA systems [2, 3].

1.1 Thesis Contributions and Organization

This dissertation is organized as follows. Initially a number of basic concepts developed in digital communications and information theory, that will be used extensively throughout this thesis are introduced. Then, a short introduction to multiuser information theory is given. The simple two-user multiple-access channel is investigated and a number of important concepts for this type of channel, like the capacity region, are introduced as in [4] and [5]. These results are then generalized to a multidimensional multiple-access channel and the capacity notion captured by the capacity region is given. Subsequently, the attention is focused on basic joint detection methods including the optimal approach based on the exhaustive search algorithms. Additionally popular linear detection methods like matched filtering (sometimes deemed correlation detection), decorrelator and particularly the MMSE filter are outlined. In general, however, these methods are still of high computational complexity primarily due to the need to perform matrix inversion. Therefore subsequently a

number of popular approximations aimed at efficient matrix inversion are presented. First chapter is summarized with the introduction of interference cancellation methods and its most fundamental variants. One is the successive interference cancellation which is a direct application of the chain rule of mutual information. The other is parallel method where all signals are cancelled at once, but multiple times (iterations).

Chapter two focuses on the optimal channel coding for interference cancellation systems. As a guideline to design optimal codes for the IC systems, the results from a very recent paper [6] are utilized. The bound derived there indicates the minimum residual variance that any coding scheme is possible to deliver. It is found that simple repetition codes are the best for IC systems, when the amount of interference is high. Surprisingly, strong capacity achieving codes perform poorly in this regime. A number of popular codes are examined and a couple of numerical simulations are performed to prove the validity of the theoretical findings.

The third chapter applies the ideas of light inner coding to the case of CDMA. The new system called partitioned CDMA or partitioned spreading CDMA exhibits substantial gains over the traditional CDMA system at a minor increase in complexity. The proposed system may be viewed as a generalization of IDMA [7]. However, both solutions have completely different motivations. IDMA was initially proposed to separate users using different random interleavers. Partitioned CDMA however is based on the observation that repetition coding delivers better results for high interference scenarios. Additionally, unlike the IDMA literature this thesis delivers a very accurate SNR evolution analysis. It is shown that for certain conditions the proposed system delivers SNR values that coincide with those achieved by the optimal APP detector. The output SNR of the APP detector was found in [8, 9] via complex analysis used in statistical mechanics. Additionally, it is found that the proposed detector delivers estimates that are always better than those obtained via the traditional one-shot MMSE filtering of cubic complexity.

Chapter four presents the partitioned CDMA analysis when each user's received powers are different. Since the equal power system is interference limited the attempt to remove this limitation via unequal power allocation is made. It has been shown that for the special case of exponential receive power profile the system can serve any number of users - system is not interference limited. In other words, for certain power allocation, the capacity of GMAC can be reached. Simulations for systems with a large number of users (and consequently large interference levels) are presented and validity of the proposed solution is proven.

The fifth chapter treats signal processing for the uplink transmission in LTE standard. The iterative turbo receiver that performs processing with a mixture of serial and parallel cancellation methods is proposed, analyzed, and verified. It is shown that for the so called cell-edge situation, the proposed receiver performs 2-3 dB better than the baseline processor. Intuitive variance transfer graphs method for determining the waterfall occurrence of the

proposed system is given and simulations prove that the accuracy of the proposed graphical method is surprisingly high.

The last chapter concludes this dissertation, summarizes its main contributions and briefly sketches future research directions on the subject.

1.2 Fundamental Concepts in Digital Communications and Error Control Coding

In this section, a number of fundamental ideas and approaches to the design and analysis of the communications systems are briefly outlined. The subsequent chapters introduce signal processing methods that immensely rely on these crucial ideas, therefore their brief, but rigorous description is of high importance for a smooth development of this dissertation.

1.2.1 Reliability Measure

The most commonly used measure of reliability is the LLR. Assume that the communication system transmits binary information $x = \{\pm 1\}$. These signals enter the memoryless Gaussian channel with noise variance σ^2 and are attenuated/amplified by a factor a. The received signal is

$$y = ax + n = ax + \mathcal{N}(0, \sigma^2) \tag{1.1}$$

and the LLR of x can be defined as

$$\Lambda_x(y) = \log \frac{P(x = +1|y)}{P(x = -1|y)} \tag{1.2}$$

and is a function of observation y. The term $P(x = \pm 1|y)$ is called *a posteriori* probability. Using Bayes rule

$$P(x = +1|y) = \frac{P(y|x = +1)P(x = +1)}{P(y)} \tag{1.3}$$

Equation (1.2) can be rewritten as

$$\Lambda_x(y) = \log \frac{P(y|x = +1)P(x = +1)}{P(y|x = -1)P(x = -1)}. \tag{1.4}$$

Term $P(x = \pm 1)$ is called *a priori* probability and $P(y|x = \pm 1)$ is likelihood. Assuming that prior probabilities are equal $P(x = +1) = P(x = -1) = 1/2$ Equation (1.4) reduces to

$$\Lambda_x(y) = \log \frac{P(y|x = +1)}{P(y|x = -1)}. \tag{1.5}$$

Since the channel is Gaussian its likelihood function, sometimes named as a transition probability, can be expressed using the PDF function of the Gaussian distribution as

$$P(y|x=\pm 1) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y\mp a)^2}{2\sigma^2}} \tag{1.6}$$

and after substituting it into (1.5) the following is obtained

$$\Lambda_x(y) = \frac{2a}{\sigma^2} y \tag{1.7}$$

and it states that LLR is a simple linear function of the observation y. The concept of LLR is the foundation of modern soft information processing algorithms like turbo and LDPC codes and a plethora of their derivatives. LLR has a number of useful properties. Λ_x is conditionally Gaussian i.e.

$$\Lambda_x(y)|x \sim \mathcal{N}\left(\frac{2a^2}{\sigma^2}x; \frac{4a^2}{\sigma^2}\right). \tag{1.8}$$

Finally, *a posteriori* bit probabilities can be derived from LLR values as

$$P(x=+1|y) = \frac{1}{1+e^{-\Lambda_x(y)}} \tag{1.9}$$

$$P(x=-1|y) = \frac{e^{-\Lambda_x(y)}}{1+e^{-\Lambda_x(y)}}. \tag{1.10}$$

After this extensive introduction the idea of a soft-bit, which will be reused extensively in the following chapters, should be introduced. The soft-bit of a random variable is the expectation of a bit, assuming that it is represented in antipodal format $\{\pm 1\}$, given the channel observation y in Equation (1.1)

$$\tilde{x} = E[x|y] = P(x=+1)P(y|x=+1) + P(x=-1)P(y|x=-1) \tag{1.11}$$

which for Gaussian channel using (1.9) and after some manipulations reduces to

$$\begin{aligned} \tilde{x} &= E[x|y] = +1 \cdot P(x=+1|y) + (-1) \cdot P(x=-1|y) \\ &= \frac{1}{1+e^{-\Lambda_x(y)}} - \frac{e^{-\Lambda_x(y)}}{1+e^{-\Lambda_x(y)}} = \tanh\left(\frac{\Lambda_x(y)}{2}\right). \end{aligned} \tag{1.12}$$

Finally, another very useful property of the LLR values is discussed. Namely, that the LLR of a number of independent observations of the same experiment is equal to the sum of the LLR values obtained in each observation. Although this property is not very innovative, it has been applied extensively in decoding algorithms of modern error control codes, as we will see later.

1.2.2 Iterative Processing

The invention of turbo codes [10] brought a revolution and was one of the major advances in digital communications that influenced many domains of information transmission. The basic idea is based on iterative message exchange between number of components in the system. Generally for error control codes it is an exchange of information between component decoders, whereas for joint detection/decoding, this exchange is performed between the decoder of the applied channel code and the interference channel itself.

1.2.3 EXIT Charts

EXIT charts, introduced in [11], deliver insights into the iterative decoding (detection) process by observing the input-output characteristics of each component decoder (detector). Its motivation was to predict the behavior of the turbo codes. EXIT chart is primarily used to predict the occurrence of the turbo-cliff. The accuracy of results obtained using EXIT analysis increases by increasing the length of the considered code. For large codes this approach is very precise. For the case of short codes and many decoding iterations it loses its accuracy considerably. The input-output relationship of each component code is expressed via the correlation between input LLR values and output for a given noise level. This correlation is best captured by the mutual information measure introduced by Shannon in his seminal 1948 paper [1]. The EXIT charts method can be used to predict the performance of the turbo decoder with quantized extrinsic information as in [12].

1.2.4 Sufficient Statistics

Sufficiency is one of the most important paradigms in probability theory. Sufficient statistics preserves the statistical properties of the observation, in other words is *information lossless*. The most general mathematical definition states that a statistic $T(X)$ is sufficient for underlying parameter θ, if the conditional probability distribution of the data X, given the statistic $T(X)$, is not a function of the parameter θ i.e.

$$P(X = x|T(X) = t, \theta) = P(X = x|T(X) = t) \tag{1.13}$$

or in short

$$P(x|t, \theta) = P(x|t). \tag{1.14}$$

The Fisher-Neyman factorization theorem delivers a useful characterization of sufficient statistics. If the PDF is $f_\theta(x)$, T is sufficient for θ if and only if functions g and h can be

found such that

$$f_\theta(x) = h(x)g_\theta(T(x)), \tag{1.15}$$

meaning that the probability density function f can be factored into two, where one does not depend on θ and the other that does depend on θ only through $T(x)$. Sufficient statistics is the central motivation to the idea of describing a possibly large observation set with a small number of parameters that fully describe this set. The most simple example of sufficient statistics is by using the set of Bernoulli distributed random variables with expected value of p. The sufficient statistics for p is the mean $T(X) = \frac{X_1+\ldots+X_n}{n}$. Denoting $X_n = 0$ being a failure and $X_n = 1$ a success, $T(X)$ represents the fraction of experiments that will end with a success. This example describes determining the bias of a coin in a coin tossing experiment.

1.3 Multiple-Access Channels

The typical multiple-access situation is presented in Figure 1.1. Shortly speaking, a number of transmitters communicate to a common receiver. The receiver's task is to successfully separate these data streams. Before we formally introduce the MAC channel let us concentrate on the general idea of a random channel, which is an important concept in wireless communications and a basis for treating and approximating a number of practical channels.

1.3.1 Random Channel

Random channel is a very popular way to mathematically model the wireless channel. Its usefulness comes from the fact that it is well understood and allows for elegant analytical treatment. To describe the random channel let us use the following linear memoryless channel

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1.16}$$

where $\mathbf{x}$ is the K-dimensional data vector, $\mathbf{y}$ is the N-dimensional channel output and $\mathbf{H}$ is the N by K channel matrix. The entries of matrix $\mathbf{H}$ are random and characterized by a certain probability distribution. Usually the most convenient assumption is that the entries of channel matrix $\mathbf{H}$ are i.i.d. and are valid for a number of cases like narrowband MIMO or DS-CDMA multiple-access channels. For example $h_{i,j} \sim \mathcal{CN}(0,1)$ models the narrowband MIMO channel with K transmit and N receive antennas. The capacity of this channel was found via the study of the behavior of its eigenvalues (singular values) in [13]. Assuming that the channel entries are $h_{i,j} \sim \mathcal{B}(1,1/2)$[1] a DS-CDMA channel is modeled and the capacity

[1] Binomial distribution $\Pr(h_{i,j} = -1) = \Pr(h_{i,j} = +1) = 1/2$.

of this channel was derived in [14]. In this thesis attention is focused on these two channels. More information on the subject of random channels can be found in [15]. Subsequently, a brief introduction to multiple-access channels is given and selected relevant results from multiuser information theory are presented. There are many models of multiuser channels

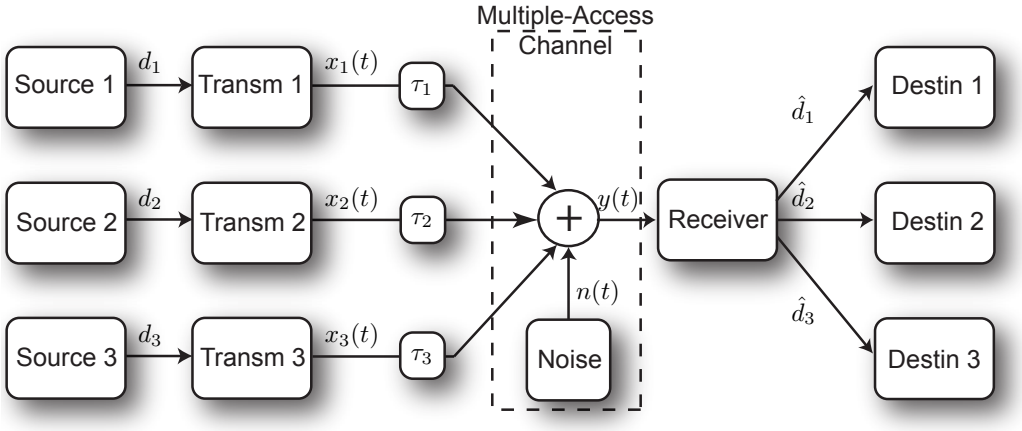

Figure 1.1: Multiple-access channel.

[4], but of most practical interest are the linear multiple-access channels and therefore, throughout the rest of this thesis, vast attention is concentrated on this particular channel. The most important model is Gaussian multiple-access channel (GMAC) model which can be described using the following equation

$$y(t) = \sum_{k=1}^{K} x_k(t - \tau_k) + n(t) \tag{1.17}$$

where $y(t)$ is the received composite signal, $x_k(t)$ is k-th user signal and $n(t)$ is additive noise, usually assumed to be zero mean Gaussian with variance σ^2 and can be easily extended to the complex case[1]. The quantity τ_k models the delay experienced by user k.

The multiple-access channel can also be modeled as a probabilistic experiment, where the channel itself is represented by a conditional probability distribution dependent on the sources as in Figure 1.2. Assuming the system consists of K users

$$\mathcal{K} = \{1, 2, \ldots, K\} \tag{1.18}$$

[1]In complex case it is circularly symmetric complex Gaussian random variable consisting of independent real and imaginary parts each with variance $\sigma^2/2$.

and that the input symbols belong to the set

$$X_{\mathcal{S}} = \{X_k, k \in \mathcal{S}\} \tag{1.19}$$

and the output alphabet be $\mathcal{Y}$, we can model multiple-access channel with the conditional probability as

$$p(y|x_{\mathcal{K}}) = p(y|x_1, x_2, \ldots, x_K) \tag{1.20}$$

and the tuple

$$(\mathcal{X}_{\mathcal{K}}, p(y|x_{\mathcal{K}}), \mathcal{Y}) \tag{1.21}$$

completely defines the channel probabilistically. An additional useful definition, that will be reused subsequently, is that users are independent, so their joint probability distribution can be expressed as

$$p(x_{\mathcal{K}}) = \prod_{k=1}^{K} p_k(x_k) \tag{1.22}$$

With this short introduction we are now ready to define the capacity region of the multiple-

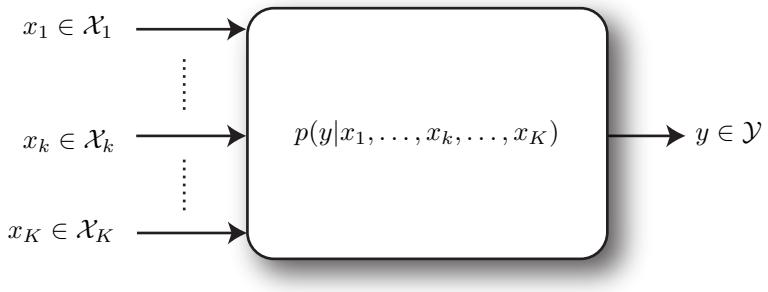

Figure 1.2: Multiple-access channel as a probabilistic experiment.

access channel[1]. This region indicates the information rates that can be achieved by every user and therefore is directly related to the overall spectral efficiency of the multiuser system. First the rate region should be introduced. Rate region is an entire set of achievable rates. Definition of rate region is as follows. With the multiple-access channel described in (1.21) and for a fixed distribution of channel inputs $\pi(x_{\mathcal{K}}) = \prod_k p_k(x_k)$ every rate vector $\mathbf{R}$ in the set

$$\mathcal{R}[\pi(x_{\mathcal{K}}), p(y|x_{\mathcal{K}})] \stackrel{\triangle}{=} \{\mathbf{R} : R(\mathcal{S}) \leq I(X_{\mathcal{S}}; Y|X_{\mathcal{K}\ \mathcal{S}}), \forall \mathcal{S} \subseteq \mathcal{K}\} \tag{1.23}$$

is achievable. To illustrate the meaning of the rate region a simple example is presented. Consider a two-user channel. For a fixed input distribution $\pi(x_1, x_2)$ the following region is

[1]This is the equivalent of channel capacity in single-user situation.

achievable

$$\mathcal{R}[p_1(x_1), p_2(x_2), p(y|x_{\mathcal{K}})] = \begin{cases} & R_1 & \leq I(X_1; Y|X_2) \\ (R_1, R_2): & R_2 & \leq I(X_2; Y|X_1) \\ & R_1 + R_2 & \leq I(X_1, X_2; Y) \end{cases} \tag{1.24}$$

and $I(A; B|C)$ represents the conditional mutual information between source A and outcome B given C. This rate region has a convenient graphical interpretation and is shown in Figure 1.3. The multiple-access channel capacity region is defined as a closure of the convex hull of union of all achievable rate regions over the entire family of product distributions on the sources, i.e.

$$\mathcal{C}[p(y|x_{\mathcal{K}})] = \text{co} \bigcup_{\pi(x_{\mathcal{K}})} \mathcal{R}[\pi(x_{\mathcal{K}}), p(y|x_{\mathcal{K}})] \tag{1.25}$$

Note the corner points (A and B) of the capacity region shown in the Figure 1.3. These points can be achieved using interference cancelation methods, which is the foundation of the proposed detector. Using the chain rule of mutual information [5] for our two–user channel we can write the mutual information as follows

$$I(X_1, X_2; Y) = I(X_1; Y) + I(X_2; Y|X_1) = R_1 + R_2 \tag{1.26}$$

which represents the point A in Figure 1.3, compare with (1.24). In the same way the equation for the case of point B can be derived. Equation (1.26) states that at the beginning the first user's data is detected ($I(X_1; Y)$) and afterwards, using the already acquired knowledge of X_1, the data of the second user is detected. Generally the capacity region can be drawn within the K-dimensional space and it is very hard to picture it graphically with a large number of users. The capacity region of real GMAC can be expressed as

$$R(\mathcal{S}) \leq \frac{1}{2} \log_2 \left(1 + \frac{P(\mathcal{S})}{\sigma^2}\right), \quad \forall \mathcal{S} \subseteq \mathcal{K} \tag{1.27}$$

and it is a generalization of Shannon's formula for the single user channel.

1.4 Optimal Multiuser Detection

Optimal multiuser detection attracted the community in the very early days of research on multiple-access channels. This problem was solved quickly, due to the fact that it was based on the most natural approach imaginable i.e. a brute force search. In this method the transmitted sequence is guessed and compared with the channel output. For a large system

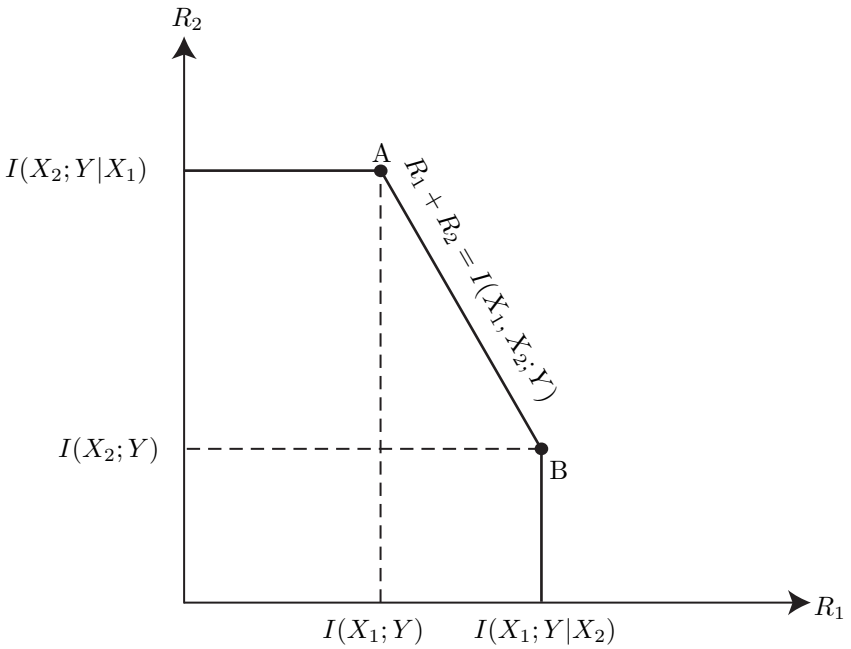

Figure 1.3: Capacity region for 2 user multiple-access channel.

the number of guesses is becoming very large therefore a search over a very large space is necessary. Such exhaustive search is very computationally expensive and this feature discards optimum detection as a method to be considered for implementation in practice. The optimum detector however, provides the best average performance in terms of sequence error rate among all possible approaches (similarly like maximum likelihood decoding on a single user channel) and is referenced today, but only to provide a benchmark for a number of reduced complexity methods that are being extensively developed.

Consider the following basic channel model

$$\mathbf{y} = \mathbf{H}\mathbf{d} + \mathbf{n} \tag{1.28}$$

where $\mathbf{H}$ is the channel matrix, $\mathbf{d}$ is the data vector and $\mathbf{n}$ is the channel noise. Assuming (1.28), the optimal multiuser, or as we may call it alternatively, joint detector can be divided into two groups. First, jointly optimal detector which minimizes the probability of error for the vector of transmitted symbols $\Pr(\hat{\mathbf{d}} \neq \mathbf{d})$. This approach does not guarantee the best possible performance for the group of users. Jointly optimal decoder tries to minimize over all possible data symbols sent the distance between received and transmitted signals i.e.

$$\hat{\mathbf{d}}_{\mathrm{ML}} = \arg \min_{\mathbf{d} \in \mathcal{D}^{\mathcal{K}_n}} ||\mathbf{y} - \mathbf{H}\mathbf{d}||_2^2 \tag{1.29}$$

This minimization can be performed using, for example, Viterbi-type algorithm, which operates on a trellis with an exponentially increasing number of states. For example, with binary data symbols its depth is 2^K, where K is the number of signals simultaneously arriving at the receiver. Equation (1.29) can be rewritten as

$$\hat{\mathbf{d}}_{\mathrm{ML}} = \arg \max_{\mathbf{d} \in \mathcal{D}^{\mathcal{K}_n}} \left(2\mathbf{d}^{\mathrm{H}} \mathbf{H}^{\mathrm{H}} \mathbf{y} - \mathbf{d}^{\mathrm{H}} \mathbf{H}^{\mathrm{H}} \mathbf{H} \mathbf{d} \right) \tag{1.30}$$

where $\mathbf{H}^{\mathrm{H}}\mathbf{y}$ is the so called matched filter output, sometimes called the de–spreader or the correlator. If the channel $\mathbf{H}$ is orthogonal implying that the correlation matrix is diagonal $\mathbf{R} = \mathbf{H}^{\mathrm{H}}\mathbf{H} = \mathbf{I}$ and information symbols $\mathbf{d}$ are uncorrelated with the constant energy then (1.30) can be rewritten as

$$\hat{\mathbf{d}}_{\mathrm{ML}} = \arg \max_{\mathbf{d}} \left(\mathbf{d}^{\mathrm{H}} \mathbf{H}^{\mathrm{H}} \mathbf{y} \right) \tag{1.31}$$

and is solved by $\hat{\mathbf{d}}_{\mathrm{ML}} = \mathrm{sgn}(\mathbf{H}^{\mathrm{H}}\mathbf{y})$. It shows that the simple matched filter (see Section 1.5) that does not perform joint detection for the case of uncorrelated spreading sequences is an optimal detector. Unfortunately such situation in multiple-access systems is hardly ever experienced.

Second method is called individually optimal detection, often abbreviated with APP (A Posteriori Probability) detection. Unlike joint optimal detector in (1.29), APP detector ensures that the given detected sequence is optimal for any given user k. Its detection rule is as follows

$$\hat{d}_k[i] = \arg \max_{\mathbf{d} \in \mathcal{D}^{\mathcal{K}_n}} \Pr(d_k[i] | \mathbf{y}) \tag{1.32}$$

and is computed via the marginalization

$$\hat{d}_k[i] = \arg \max_{\mathbf{d} \in \mathcal{D}^{\mathcal{K}_n}} \sum_{\mathbf{d}: d_k[i] = d} \Pr(\mathbf{y}|\mathbf{d}) \Pr(\mathbf{d}) \tag{1.33}$$

where index i is the index of the symbol in the frame belonging to a particular user and it is introduced since APP detection delivers optimal individual symbol estimates, not like typical detector which minimizes the error probability of the whole sequence $\mathbf{d}$. Efficient computation of the marginalization in (1.33) is performed using forward-backward method, called BCJR algorithm [16], popularized widely with the invention of turbo codes [10].

However, as mentioned, the complexity of both optimal receivers presented grows exponentially with the number of users served in the system [17, 18]. Therefore, it is prohibitively expensive to perform optimal detection even for a moderate number of users in the system. Suboptimal solutions based on optimal detection were proposed. Among them are approximate APP, sphere detectors and tree searching algorithms like LISS detector [4]. The main idea behind these approximate optimal approaches is the fact that practically

only relatively small number of candidate vectors need to be considered to generate a good estimate of the true APP. The problem is to select those most probable candidate vectors intelligently. Sphere detectors operate without the *a priori* information and the computationally expensive marginalization is only performed on much smaller number of vectors that lie inside a sphere of radius r with the center being the received point $\mathbf{y}$. The challenge is to choose the radius r such that acceptable performance is achieved. However, even those approximations of optimal methods exhibit large computational burden and are of not much practical interest.

Many attempts have been made in order to provide even less computationally expensive suboptimal solutions, which would deliver acceptable performance at much less computational expense and some of the most important of them are briefly presented below.

1.5 Linear Detection Methods and Approximations

Much attention has been devoted to linear detection using quite intuitive methods originating from linear algebraic approaches. To these belong the matched filter, sometimes called the correlator, the decorrelator and the Minimum Mean Squared Error (MMSE) detector. The channel model assumed is the same as in (1.16).

Matched filtering relies on correlating the received composite signal with known spreading codes of a given user. It is the most basic way of retrieving transmitted data from the multiple-access channel since in this way the channel operation $\mathbf{H}$ is in a sense reversed. The matched filter operation is described by the equation

$$\hat{\mathbf{d}}_{\mathrm{MF}} = \mathbf{H}^{\mathrm{H}}\mathbf{y} = \mathbf{H}^{\mathrm{H}}\mathbf{H}\mathbf{d} + \mathbf{H}^{\mathrm{H}}\mathbf{n} = \mathbf{R}\mathbf{x} + \mathbf{H}^{\mathrm{H}}\mathbf{n} \tag{1.34}$$

where $\mathbf{R} = \mathbf{H}^{\mathrm{H}}\mathbf{H}$ is the correlation matrix of the channel matrix $\mathbf{H}$. Note that operation in (1.34) is performed separately i.e. can be run on each user independently. In the case of orthogonal channel i.e. when $\mathbf{R} = \mathbf{I}$ the matched filter output is optimal (see 1.31). However, ensuring this orthogonality is in most cases impossible due to time delays and amplitude variations experienced by these sequences during transmission through the channel. It is very important to mention that the matched filter estimate is a *sufficient statistics* for the detection of transmitted symbols $\mathbf{d}$. Therefore the output of matched filter can be used in later processing as fully equivalent of the channel signal $\mathbf{y}$.

Another linear method is the decorrelator. It operates according to the following formula

$$\hat{\mathbf{d}}_{\mathrm{DEC}} = \mathbf{R}^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{y} = \mathbf{R}^{-1}\mathbf{R}\mathbf{d} + \mathbf{R}^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{n} = \mathbf{d} + \mathbf{R}^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{n} \tag{1.35}$$

Assuming that matrix $\mathbf{R}$ is invertible, decorrelator completely eliminates the multiple-access interference. Since multiple-access interference is removed, one might think that this is an optimal solution. Unfortunately it is not true, since during the elimination of interference significant enhancement of additive noise may occur and severe performance degradation is observed (term $\mathbf{R}^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{n}$ in the Equation (1.35)).

The last method called the Minimum-Mean-Square-Error (MMSE) is adopted from control and filtering theory and is basically similar to the decorrelator, however it also takes into account the channel noise, which the decorrelator ignores. This detector operates according to the minimization criterion

$$\hat{\mathbf{d}}_{\mathrm{MMSE}} = \arg \min_{\mathbf{M}} \mathbb{E}\left[||\mathbf{d} - \mathbf{M}\mathbf{y}||^2\right] \tag{1.36}$$

and matrix $\mathbf{M} = \mathbf{R} + \sigma^2\mathbf{I}$. The operation in (1.36) is equivalent to perform

$$\hat{\mathbf{d}}_{\mathrm{MMSE}} = \mathbf{M}^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{y} = \mathbf{M}^{-1}\mathbf{y}_{\mathrm{MF}}. \tag{1.37}$$

The MMSE detector takes into account channel noise as well as interference, but it ignores the data structure. As long as the data structure is not taken into account the MMSE filter provides the best performance of all linear preprocessing approaches, however, similarly to decorrelator, it has high complexity which is mostly related to the inherent need for matrix inversion. This process has a complexity growing as $O(K^3)$, where K is the number of signals present in the system.

These linear methods were extensively researched [2, 19], since they provide feasible alternative to the optimal detection. However, they have their own limitations, which appear for the so called overloaded systems[1]. In these situations linear detectors provide very poor output signal to noise ratios, which practically discard them as good detectors for highly overloaded systems.

1.5.1 Iterative Approximations

As stated the complexity of linear filtering methods lies in matrix inversion. However, many iterative filtering methods (essentially iterative matrix inverters) were proposed in order to perform the detection with the sub-cubic complexity [20, 21], with performance loss almost negligible as compared to the matrix inverters in (1.35) and (1.37). Since they operate iteratively we call them iterative or sometimes multistage processors.

[1] System is overloaded, when the number of signals present in the channel exceeds the number of available dimensions. For MIMO systems it describes the situation when the number of receive antennas N_r is smaller than that of transmit N_t. For the case of CDMA, when the number of users is larger than the spreading gain N, i.e. $K/N > 1$.

It was already observed by Gauss in 19th century that a matrix can be inverted using iterative processing methods. Technically, there are many variations of this inversion, however conceptually they have the same goal. The most popular iterative matrix inversion methods are:

- conjugate gradient,
- Gauss-Seidel,
- Jacobi,
- first and second order stationary,
- Chebyshev,
- successive relaxation method.

All of them, but the successive relaxation, are parallel cancellation methods[1]. Subsequently, the most important iterative matrix inverters are described. Assuming that the output of matched filters is available at the input of the iterative matrix inverter described by the Equation (1.34). Taking into account user amplitudes matrix $\mathbf{A}$, (1.34) can be rewritten as

$$\hat{\mathbf{d}}_{\mathrm{MF}} = \mathbf{RAd} + \mathbf{H}^{\mathrm{H}}\mathbf{n}. \tag{1.38}$$

Gauss-Seidel Method

Suppose the following linear matrix equation needs to be solved $\mathbf{Md} = \mathbf{b}$ and it is apparent that to do this the inversion of $\mathbf{M}$ is necessary. Gauss-Seidel method relies on splitting this matrix into $\mathbf{M} = \mathbf{S} - \mathbf{T}$. Now our matrix equation look as follows

$$\mathbf{Md} = \mathbf{b} \Leftrightarrow [\mathbf{S} - \mathbf{T}]\mathbf{d} = \mathbf{b} \Rightarrow \mathbf{Sd}^{(m+1)} = \mathbf{Td}^{(m)} + \mathbf{b} \tag{1.39}$$

where m represent the iteration index. Now the solution can be represented as

$$\mathbf{d}^{(m+1)} = \mathbf{S}^{-1}\left(\mathbf{Td}^{(m)} + \mathbf{b}\right) \tag{1.40}$$

Naturally, the complexity of this solution depends on the matrix $\mathbf{S}$ chosen. If it is chosen poorly, (1.40) does not show any complexity reduction as compared to the original matrix inversion. Different methods based on the Gauss-Seidel method depending on the choice of the splitting of $\mathbf{M}$ exist. Additionally, the condition required (1.40) to arrive at the desired solution is that the largest eigenvalue of $\mathbf{S}^{-1}\mathbf{T}$ must be less than one.

[1] This approach will be described in detail in the next section.

Jacobi Method

This solution is a particular example of Gauss-Seidel method, when the matrix $\mathbf{H}$ is chosen to be the identity matrix $\mathbf{I}$. And this recursive process can be started with the initial vector being the correlator estimate i.e. $\mathbf{d}^{(0)} = \mathbf{y}$. Using the Jacobi method and with the right choice of matrix $\mathbf{M}$, every linear detector as decorrelator and MMSE can be implemented. The method converges to the desired solution of decorrelator only when the spectral radius (maximum eigenvalue) of the correlation matrix $\mathbf{R}$ is $\rho(\mathbf{R}) < 2$ and for random channels it occurs only for loads $\alpha < (\sqrt{2}-1)^2 \approx 17\%$ [20], which are relatively small when comparing to current standards. In case of MMSE realization using Jacobi, much higher loads, up to 100%, can be accommodated. The most important aspect of Jacobi iteration is that this method and its variations lead to parallel interference cancellation methods discussed later.

Stationary Iterative Methods

Generalizing the Jacobi method and overcoming its limitations about the size of spectral radius by introducing the *cancellation factor* $\tau^{(m)}$, we arrive at first-order stationary solution [4]

$$\mathbf{A}\mathbf{d}^{(m+1)} = \mathbf{A}\mathbf{d}^{(m)} - \tau^{(m)} \left(\mathbf{M}\mathbf{A}\mathbf{d}^{(m)} - \mathbf{y} \right). \tag{1.41}$$

It is seen that new estimate $\mathbf{d}^{(m+1)}$ in Equation (1.41) depends only on the previous estimate i.e. $\mathbf{d}^{(m)}$. If we allow the dependence on n previous estimates the method is called n-th order method. The *cancellation factor* τ is the equivalent of the learning factor well known from the adaptive filter theory (gradient based adaptive filters and LMS algorithm). If it is constant through all the iterations the method is called *stationary*. Setting $\mathbf{M} = \mathbf{R}$ and $\mathbf{M} = \mathbf{R} + \sigma^2\mathbf{A}^{-2}$ in (1.41) decorrelator and MMSE filter can be implemented, respectively. The value $\tau^{(m)}$ can control the spectral radius of $\mathbf{H}^{-1}\mathbf{T}$. Stationary iterative methods can converge to decorrelator or MMSE filter for system loads $\alpha < 1$ and the fastest convergence is assured when $\tau = \tau_{\text{opt}} = 2/(\lambda_{\min} + \lambda_{\max})$. As before, λ represent the eigenvalues of matrix $\mathbf{M}$. Stationary iterative methods are very attractive from the complexity and implementation point of view, however they fail, just like the approximated methods, in situations when the number of available dimensions decreases, i.e. when system loading exceeds 100%.

1.6 Coded Transmission

The improvement can be obtained when using error control codes (ECC) that would help to improve successively the reliability information passed to those above mentioned linear filters. Coded multiuser communications can be viewed as a serial concatenation of an error control code with multiple-access channel just like the serial turbo coding schemes. An error control code plays the role of outer code whereas the multiple–access channel the role of inner code. With such interpretation the soft information exchange principle can be applied. It is shown in [22] that such configuration can achieve the capacity of the MAC. However, this solution may be very complex, since many filtering/decoding runs are needed. This approach is named later full decoding and represents a performance reference for the evaluation of the proposed PS-CDMA detector and its various possible configurations.

The question arises what codes are best suited for the concatenation with multiple–access channel. Contrary to what intuition might say the use of strong error control codes, like turbo or LDPC, in the iterative multiuser detection loop is not the best possible approach as reported extensively in [23, 24] and simpler codes like convolutional or even repetition are more preferred. Those ideas will be thoroughly explained in Chapter 2.

1.7 Iterative Detection and Interference Cancellation

The invention of turbo codes and reinvention of Gallager's LDPC codes brought attention to iterative detection of signals in multiple-access channels. In [25, 26] it is demonstrated that iterative detection of CDMA signals can provide substantial gains over traditional one shot detection and still be considerably less complex than optimum ML detection.

Every iterative scheme contains a cancellation step, where the reconstructed signals are subtracted from the composite channel signal. The cancellation step is a direct realization of the conditional mutual information in (1.24). Estimates of the data used to produce the reconstructed signal can be taken from the error control code, performing MAP detection on coded symbols.

1.7.1 Successive (Serial) Cancellation

This idea of performing this detection is a direct application of chain rule of mutual information [5]

$$I(X_1, X_2, \ldots, X_K; Y) = \sum_{k=1}^{K} I(X_k; Y | X_{k-1}, X_{k-2}, \ldots, X_1) \tag{1.42}$$

and is shown, for example in [27], to achieve the capacity of GMAC. This cancellation approach relies on ordering (according to some specific criteria, usually with respect to the received powers) users and processing them according to this order. Its principle is simple: subtract off the strongest remaining signal successively user by user. This process is shown in Figure 1.4. At the beginning, the first user (usually the strongest) is detected/decoded treating all $K-1$ other users as the noise. Then the decoded data is subtracted from the composite channel signal and the second user is processed, but now there are only $K-2$ noise sources. The process is repeated successively until no users in the system are left and the cancellation process results in additive channel noise independent from users' data. This method was proven to achieve the vertices of capacity region mentioned during the description of Figure 1.3 and Equation (1.26). Of course, in the description of detection/decoding process it is assumed that users transmit with rates that allow the perfect decoding of their data streams. Otherwise the system performance is strictly away from the capacity limits. The power ordering (or rate in the case of equal powers) is very reasonable, since it is most probable that the strong user is estimated correctly and its errors will not propagate in the cancellation chain to the other users. This is the first problem

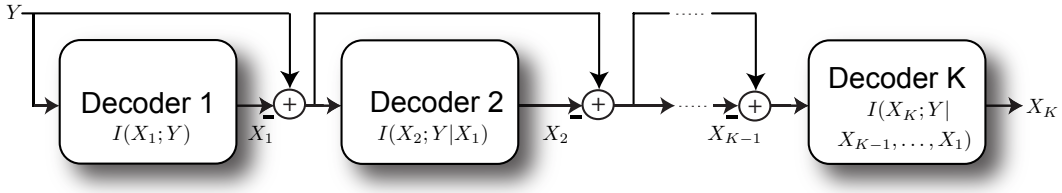

Figure 1.4: Successive (serial) interference cancellation receiver.

arising in this method - the error propagation. An error in decoding k–th data stream will propagate further down to all subsequent users $k-1,\ldots,1$. Another evident drawback is that a successive cancellation scheme exhibits long processing delay, which is proportional to the number of users present in the system. To overcome this effect parallel cancellation was introduced.

1.7.2 Parallel Cancellation

In parallel cancellation shown in Figure 1.5 no ordering is needed by definition. It relies on subtracting off all of the users' signals from all the others. All users are detected/decoded at once, however many times[1]. In many cases the number of these iterations is much less than the number of users, which provides a clear advantage over the serial scheme described previously. The proposed partitioned detector operates on the principle of parallel

[1]To assure satisfactory performance.

cancellation and its core characteristics will be explored in subsequent chapters. In [22] it was concluded that parallel cancellation methods can achieve the capacity of the channel, as long as proper power assignment is experienced.

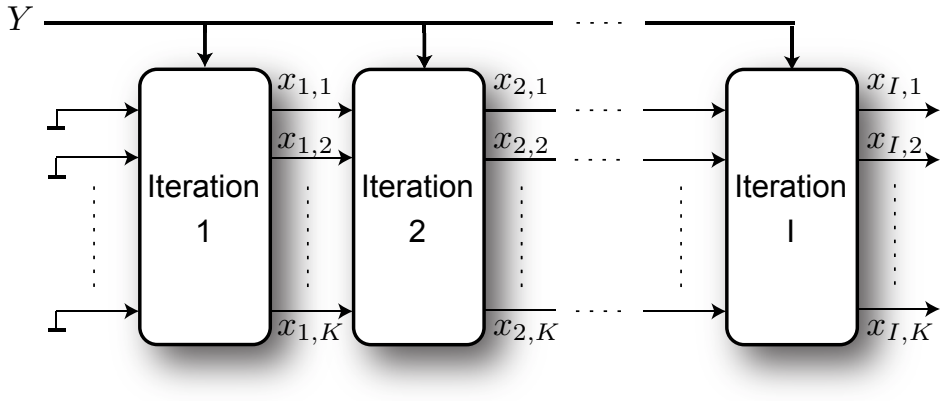

Figure 1.5: Parallel interference cancellation receiver.

CHAPTER 2

Optimal Coding for Interference Cancellation Systems

As stated in the previous chapter, achieving the limits of the multiple-access channel requires the use of the channel codes. This situation is analogous to point-to-point communications, where channel coding considerably increases the transmission efficiency. Invoking error control coding in the equalization process is referred to as Turbo Equalization. In this method the equalizer exchanges information with the error control code and both entities are separated by an interleaver, preventing build-up of correlation during iterations. Since the central idea of this thesis is interference cancellation, in this chapter, channel coding for such systems will be discussed. As it will become evident the multiple-access situation has quite different demands in terms of coding schemes as compared to a single-access model impaired only with additive noise. Very often the ideas from the latter cannot be fully transplanted to the former. This behavior can be captured and expressed analytically in the form of a bound, representing the optimal performance of any coding scheme in IC situations. Below, such a bound for the optimal coding scheme is presented and a number of traditional coding schemes examined and verified from the point of their suitability for IC purposes. Additionally computer simulations are performed to prove the validity of the proposed approach.

The main characteristics that determines whether a given coding scheme is adequate for interference cancellation systems is that such a scheme should deliver good performance not only in the close-to-capacity SNR region, but also in very low SNR regions, where errors inevitably occur. The main reasoning behind this statement is that if the codeword estimates are poor error propagation occurs, very often preventing the system from reaching acceptable performance. Besides that, it is known that the main issue with the majority of channel codes is that for low signal-to-noise ratios they actually deliver very poor codeword (as well as information contained in these codewords) estimates, often poorer than straight-

forward uncoded transmission. This effect, surprisingly, is more pronounced for strong capacity achieving codes, see for example [24]. The bound mentioned above formalizes this phenomenon and therefore gives a guideline on how to design the optimal coding method for the interference cancellation systems.

2.1 System Model

Starting with the system model, Figure 2.1 presents the generic interference-dominated system. The channel considered here is generic and in fact can be any linear channel, like CDMA, ISI, MIMO, OFDM, etc. The model operates as follows. First, the users' or streams' signals $\mathbf{b}$ are independently encoded using specific channel codes. The choice of the channel code can be arbitrary as long as soft decoding of such a code is possible. The K coded streams $\mathbf{c}_k$ are subsequently independently interleaved, modulated, and passed to the linear channel[1] in the form of $\mathbf{x}_k$. The main attention of this chapter is focused on channel codes allowing the best performance at the receiver side. The output of the channel can be expressed using a linear model as

$$\mathbf{y} = \mathbf{HAx} + \mathbf{n} \tag{2.1}$$

where $\mathbf{H}$ is the channel matrix, $\mathbf{A}$ is the matrix of user amplitudes, and $\mathbf{n}$ is the AWGN noise with variance σ_n^2. The need for interleaving after encoding is solely related to the nature of

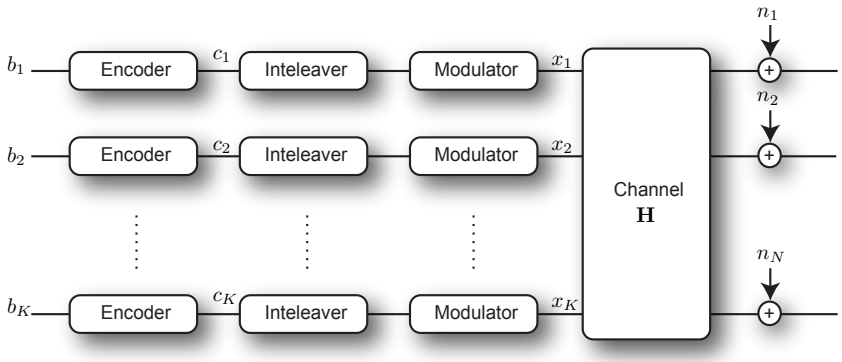

Figure 2.1: Transmitter and linear channel model with coding.

the detection process performed at the receiver, where information is iteratively exchanged between the channel code and the linear channel itself, via the interference cancellation operation. In this way the turbo principle [10, 28, 29] is realized and the well-known turbo cliff behavior expected. Figure 2.2 presents a generic receiver for the considered transmission

[1] We concentrate on the Gaussian Multiple-Access Channel (GMAC), unless otherwise stated.

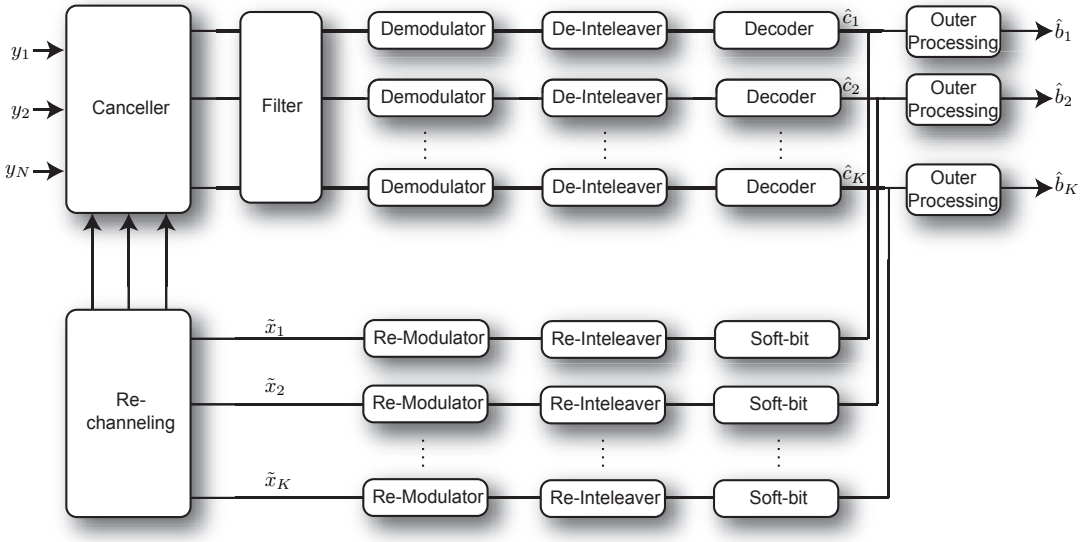

Figure 2.2: Generic receiver with interference cancellation.

system. The receiver's dimensionality is N, and, generally receiver dimensionality is different from that of the transmitter, i.e. $N \neq K$. In practice the number of receive dimensions can be associated with the number of receive antennas or spreading gain for the case of CDMA. Perfect channel knowledge at the receiver is assumed. The detection/decoding process is as follows. The first operation of the receiver is filtering in order to retrieve initial estimates of the codewords. For simplicity we concentrate now on simple matched filtering to the channel[1] that results in

$$\hat{\mathbf{x}} = \mathbf{H}^{\mathrm{H}}\mathbf{y} \tag{2.2}$$

where $\mathbf{H}^{\mathrm{H}}$ is the Hermitian (conjugate-transpose) of the matrix $\mathbf{H}$. Symbols $\hat{\mathbf{x}}$ are soft demodulated (this procedure is discussed extensively in Chapter 1) producing the LLRs of the coded bits, de-interleaved and passed to the inner error control code decoders. Based on the code constraints, the decoders deliver (hopefully more reliable) estimates of the transmitted signals, which are subsequently re-interleaved, re-modulated and then used in the interference cancellation block. The error control code decoders output extrinsic LLRs, which are obtained as

$$L_e = L_{\mathrm{APP}} - L_a \tag{2.3}$$

[1]Naturally, more sophisticated pre-processors are allowed, like the decorrelator or the MMSE filter.

where L_{APP} is the complete A Posteriori Probability (APP) information from the decoder and L_a is the a priori information. This subtraction is performed in order not to violate the extrinsic information exchange principle [10] i.e. to avoid a negative correlation build-up during the iterative cancellation process. In our case the channel LLRs are obtained from soft demodulators. The process continues for a number of iterations, say I, and finally outer decoding of the complete information is performed to obtain the estimates of the transmitted bits, formally described as

$$\hat{b} = f(L^I_{\mathrm{APP}}) \tag{2.4}$$

where $f(x)$ can be a sign function or any other outer processing function.

2.2 Analysis of IC Systems

As in any communications system, depending on the level of interference and the level of noise affecting the transmission, convergence to error-free performance may or may not be achievable. For the case of iterative processors, this means that successful detection is possible only if each entity participating in the information exchange has an output SNR larger than that on the input, i.e. each block monotonically improves the reliability of the processed information. This statement brings us to the main idea of the graphical representation of the convergence of iterative systems captured by the variance transfer charts.

2.2.1 Variance Transfer Charts

The variance transfer chart [23] is a graphical representation of the behavior of the iterative system and conceptually is very similar to the EXIT [11] charts as shown in [30]. The notion of variance transfer charts can be applied to any iterative system, however they have been found particularly useful in the analysis of the interference systems, like multi-user CDMA or MIMO. The primary reason for this is the fact that the interference process can very often be modelled as a Gaussian process with given, usually zero, mean and a certain variance. The VT method aims at tracking the variance changes during the iterative process and as such is practically analogous to the density evolution approach described in the previous chapter in the context of LDPC codes.

Practically, the variance transfer curve is a characterization of the behavior of a certain block in the communications system structure, like a filter or a channel decoder. The latter is of main interest here. The curve visualizes the relationship between input signals and the resulting output signals. Those input/output quantities can be the SNR, or in the case

of fixed signal power, the variance, usually describing the power of noise plus any form of interference. However, as is customary for soft processing algorithms, the input and output signals are LLR values. From chapter 1 we know that the expression for the LLR in the case of the Gaussian channel is just a scaled version of the received signal.

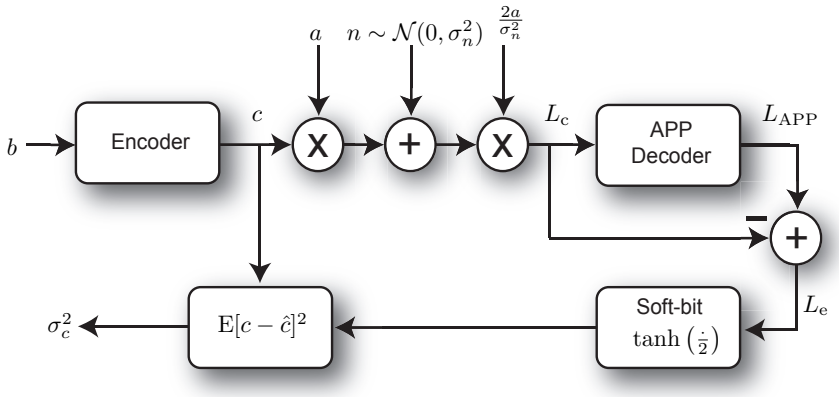

Figure 2.3: Variance transfer curve evaluation procedure.

The procedure of obtaining the variance transfer curve is presented in Figure 2.3 and is as follows. The information bits are represented as b. They are encoded into c codeword, attenuated by a factor a and transmitted through the AWGN channel with noise power σ_n^2. Before feeding the resulting signals to the decoder, they are converted into LLR values using (1.7). The decoder uses these LLRs to produce output LLRs, hopefully with improved reliability. From these extrinsic LLRs, the soft-bit values are extracted. They describe the influence of the decoder errors on the system performance in the subsequent cancellation stages. The soft-bits (or soft-symbols) introduced in (1.12) measure the power of the error introduced by the decoder. Value $\sigma_c^2 = 1$ implies that the decoder output is independent of the transmitted coded data bits c. On the other hand, $\sigma_c^2 = 0$ describes the opposite situation - perfect decoding, i.e. the output of the decoder is exactly the same as the input to the modulator. Every other situation resides between those two extremes. The soft-bit variance actually describes the amount of interference present in the system. The purpose of the iterative detection/decoding is to reduce the soft-bit variance iteratively from the initial value of full uncertainty of the transmitted information, denoted by $\sigma_c^2 = 1$ to full knowledge of them, $\sigma_c^2 = 0$. The VT curve is simply a graphical representation of the following relationship

$$\sigma_c^2 = f\left(\frac{\sigma^2}{a^2}\right) \tag{2.5}$$

and a^2 is the power of the desired signal and σ_n^2 is the power of the AWGN distortion. For the vast majority of known channel codes, the function $f(\cdot)$ has to be obtained numerically using the procedure depicted in Figure 2.3. However, for simple repetition codes, as well as LDPC codes, an analytical evaluation is possible using the DE approach.

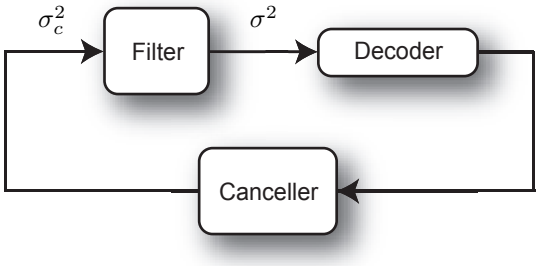

Figure 2.4: Simplified diagram of the iterative information exchange.

The simplified information exchange describing this process is shown in Figure 2.4. As we can see the filter accepts signals from the canceller and these signals contain residual interference power described by the soft-symbol variance $\sigma_{c,i}^2$. The response to this input is a new signal, that is assumed to be Gaussian[1] with zero mean and variance σ^2. These messages are converted into LLR values and passed to the error control decoder. This is exactly the same situation as in the case of single user Gaussian channel. The decoder returns soft codewords estimates which are passed to the canceller and after cancellation results in new signal with residual interference power $\sigma_{c,i+1}^2$. This process continues for a number of iterations until both variances[2] stabilize. This stabilization point may be equivalent to achieving error free performance or in the opposite situation when the iterative receiver gets stuck strictly away from errorless performance. When plotting VT curves it is usually assumed that the decoder and the canceller are one entity accepting σ^2 at the input and delivering σ_c^2 at the output.

Figure 2.5 shows an example of a variance transfer graph related to no particular filter or inner channel coding scheme. This graph has the following meaning. The decoding process starts with the filtering operation. At this moment the soft-symbol variance is maximum, $\sigma_{c,0}^2 = 1$ and the filter returns a signal embedded in noise plus interference of power $\sigma_0^2 = 0.85$. This signal is passed to the channel decoder which returns soft symbol estimates that are fed to the cancellation block resulting in soft-symbol variance $\sigma_{c,1}^2 = 0.25$. Again the filter is invoked and the process continues for a number of iterations until the final

[1]For large systems, $K \longrightarrow \infty$, it is exactly Gaussian, via the Central Limit Theorem arguments.
[2]Stabilization of soft-symbol variance implies stabilization of the σ^2 and vice versa.

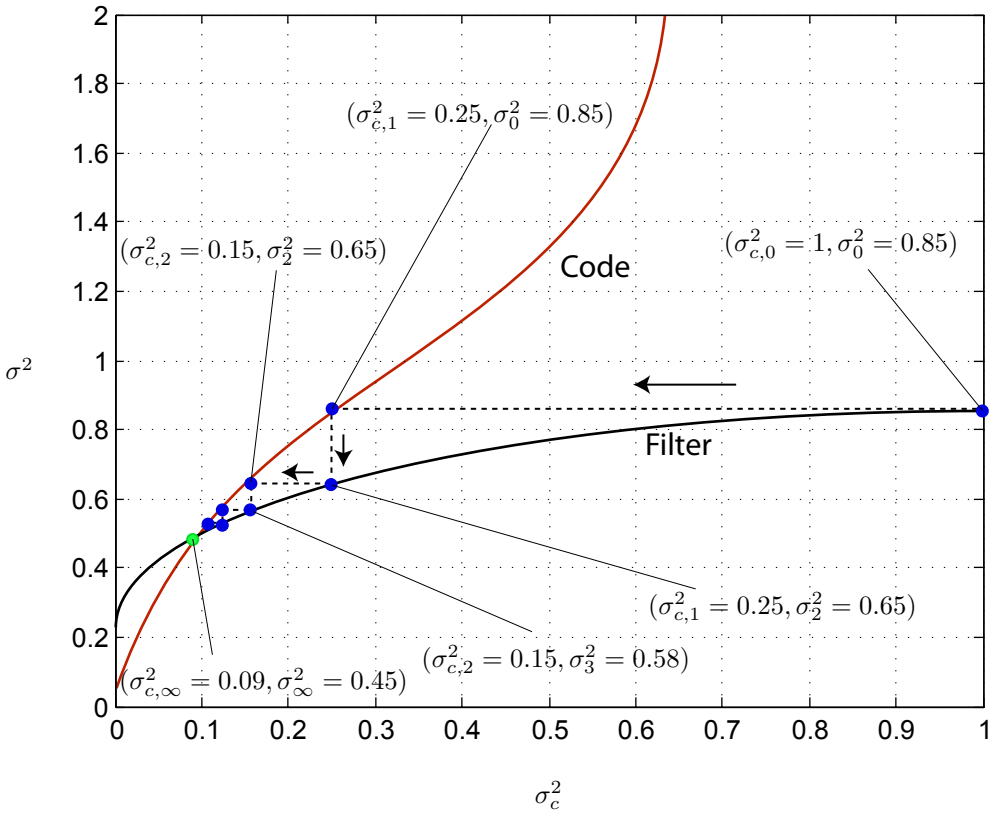

Figure 2.5: A generic example of a set of VT curves.

stable point is reached, for our case this point is $(\sigma^2_{c,\infty} = 0.09, \sigma^2_\infty = 0.45)$. Therefore perfect decoding is not achieved, since the soft-symbol variance is greater than zero, $\sigma^2_{c,\infty} > 0$. Perfect decoding after a number of iterations, is possible when both variance transfer curves, do not intersect. This condition is analogous to one determining convergence of concatenated systems using EXIT analysis. Naturally, both curves always intersect at a point, where the noise power is $\sigma^2_n > 0$.

2.3 Optimal Coding for IC

As with any complete theory it is always good to indicate what performance can be expected out of the system. Especially in information and communications theory a sort of best

performance is very crucial in assessing the efficiency of a particular solution. Since the early 1950s, such a guideline was the Shannon limit [1] which relates transmission power (or equivalently SNR) to a maximum transmission rate where error-free communication is possible.

The situation is not any different for the case of iterative interference cancellation systems. We would like to find which codes output the best codeword estimates for a certain SNR. The measure of this quality is the already introduced soft-bit (or symbol) variance σ_c^2. We are interested in codes producing codeword estimates with the smallest soft-bit variance possible for a given SNR value.

There is a very vital connection between the perfect codes for cancellation and the Shannon bound. The latter precisely describes the at which SNR level errorless codeword estimation is possible i.e. $\sigma_c^2 = 0$. However, for iterative cancellation systems, not only the final value of σ_c^2 is important, but even more important are its intermediate values. In this sense the bound derived subsequently could be called an extension of the Shannon bound. The derivation of the bound on optimal coding scheme presented below is taken from a very recent paper by Schlegel and Burnashev [6].

Using the Gaussian approximation it can be assumed that signals available after the filtering operation are scaled and noisy versions of the transmitted codewords, i.e.

$$y_k = a_k c_k + n_k \tag{2.6}$$

where due to the Gaussian assumption, the noise $n_k \sim \mathcal{N}(0, \sigma_n^2)$. Using the signal $\mathbf{y} = [y_1, \ldots, y_K]$, the estimate of the codeword $\mathbf{c}$

$$\hat{\mathbf{c}}(\mathbf{y}) \tag{2.7}$$

is produced by the inner decoder. For clarity of derivation we can assume, without loss of generality, that the attenuation constant a_k is unity and can be omitted in the subsequent derivations. We are interested in the decoder minimizing the error between the actual transmitted codeword $\mathbf{c}$ and its decoded estimate $\hat{\mathbf{c}}$

$$\min_{\mathcal{C}} \mathbb{E}||\mathbf{c} - \hat{\mathbf{c}}||^2 \tag{2.8}$$

for any value of the signal-to-noise[1] ratio ρ. Three random variables are involved here. The encoder outputs C, the information sequence U and the set of observed channel symbols Y. Using Shannon theory, it can be inferred that the uncertainty about the information symbol

[1] plus interference

u_k is given by the difference between rate R and the channel capacity C

$$H(U|\mathbf{y}) \geq \frac{R - C(\rho)}{R} \tag{2.9}$$

where $\rho = 1/\sigma^2$ is the normalized SNR w.r.t to unit signal power. Subsequently, for any reasonable, uniquely decodable code the following holds

$$H(C|\mathbf{y}) \geq H(U|\mathbf{y}) \tag{2.10}$$

meaning that the uncertainty about the transmitted codeword is at least as large as that of the information symbols. This fact comes from the assumption that the code itself is linear. The vast majority of known channel codes are linear, so here we consider solely such classes of codes. Therefore the uncertainty about the coded bits can be reproduced in the same fashion as for information bits by combining (2.9) and (2.10) which results in

$$H(C|\mathbf{y}) \geq \frac{R - C(\rho)}{R}. \tag{2.11}$$

Now, armed with this mathematical apparatus, we can ask the fundamental question. Given $H(C|\mathbf{y})$ what is the mean-squared error $\mathrm{E}||\mathbf{c} - \hat{\mathbf{c}}||^2$ of the best estimate $\hat{\mathbf{c}}(\mathbf{y})$ that a soft codeword estimator can produce from the received sequence $\mathbf{y}$. This question is answered by the following

Proposition 1. *For a given $\sigma_c^2 = E||\mathbf{c} - \hat{\mathbf{c}}||^2$, the maximum value of uncertainty $H(C|\mathbf{y})$ is achieved when $Pr(C = 1|\mathbf{y})$ takes on only two values c_0 and $1 - c_0$ and is given by*

$$\max_{(C,\mathbf{y})} H(C|\mathbf{y}) = h\left(\frac{1 - \sqrt{1 - \sigma_c^2}}{2}\right) \tag{2.12}$$

where $h(\cdot)$ is the binary entropy function.

Proof: see [6].

Extracting σ_c^2 from (2.12) we obtain the minimum possible variance of the best codeword estimator as

$$\sigma_{c,\min}^2 = 4h^{-1}(H(C|\mathbf{y}))\left(1 - h^{-1}(H(C|\mathbf{y}))\right). \tag{2.13}$$

Combining (2.11) and (2.13) we arrive at

Theorem 1. *The estimation error of the output symbols of an optimal binary error control symbol estimator is lower bounded by*

$$\sigma_c^2 \geq 4h^{-1}\left(\frac{R - C(\rho)}{R}\right)\left(1 - h^{-1}\left(\frac{R - C(\rho)}{R}\right)\right). \tag{2.14}$$

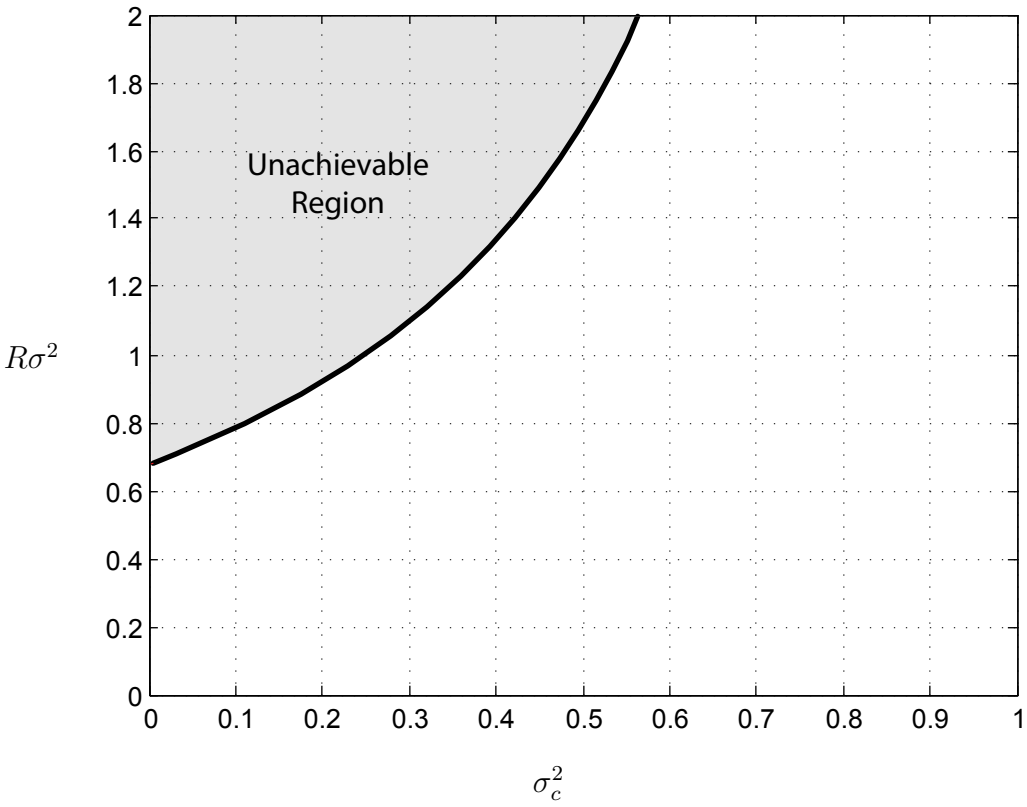

Figure 2.6: Graphical representation of the bound in (2.14).

Figure 2.6 shows the bound on optimal codes as states in (2.14). It has one particularly interesting point. Namely the point where $\sigma_c^2 = 0$. It's corresponding Y-axis point equals

$$\frac{1}{2R\sigma^2(\sigma_c^2 = 0)} = \ln 2 = -1.59\text{dB} \tag{2.15}$$

which is the famous Shannon bound. Therefore it is appropriate to name the derived bound on optimal cancellation codes an extended Shannon bound, since it is extended to error rates (or soft-symbol variances) strictly away from zero.

2.4 Performance Evaluation

In this section the performance with respect to the bound described previously for certain popular coding schemes and their suitability to IC systems are presented. For better clarity,

the performance of each code is briefly preceded with the explanation of the particular channel coding scheme.

2.4.1 Repetition Codes

Repetition coding is the simplest coding scheme imaginable and a very basic method of encoding signals. It is described here due to its optimality in the high noise region as proved in [6]. A repetition code of rate $R = 1/M$ is a block code and therefore can be described by the generator matrix

$$\mathbf{G}_{(M,1)} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & \dots & 1 \end{bmatrix} \tag{2.16}$$

consisting of M ones and a parity-check matrix

$$\mathbf{H}_{(M,1)} = \begin{bmatrix} \mathbf{1}_{M-1} & | & \mathbf{I}_{M-1} \end{bmatrix} \tag{2.17}$$

where $\mathbf{1}_m$ is a column vector of m ones and $\mathbf{I}_m$ is $m \times m$ identity matrix. Soft decoding of this code is used extensively throughout this thesis as follows. After collecting M outputs of a channel the following sequence of the LLR values is obtained

$$\{\Lambda_x(y_1), \Lambda_x(y_1), \dots, \Lambda_x(y_M)\} \tag{2.18}$$

the LLR of the transmitted information bit x is simply

$$\Lambda_x(\mathbf{y}) = \sum_{i=1}^{M} \Lambda_x(y_i). \tag{2.19}$$

This operation equals the traditional "majority vote decoding" approach. The formula (2.19) delivers full APP information about the information bit. For the purposes of interference cancellation, the estimates of the coded bits are needed. In these circumstances equation (2.19) must be modified. The LLR of the m-th coded bit can be obtained as

$$\Lambda_{x,m}(\mathbf{y}) = \sum_{i \neq m}^{M} \Lambda_x(y_i) \tag{2.20}$$

where the extrinsic information exchange rule is clearly adopted, since symbol m does not depend on itself - the summation excludes this particular term. We will return to this idea in Chapter 3, during the analysis of partitioned modulation on the example of CDMA.

The BER curve for the repetition code coincides with that of an uncoded system, which for

the Gaussian channel model is expressed using the Q[1] error function, and the signal-to-noise ratio is measured per bit. As shown the main advantage of repetition codes is their decoding simplicity, which boils down to a simple addition of the reliability information. Variance transfer curves of those codes are obtained using

$$\sigma_c^2(x) = \mathbb{E}\left[1 - \tanh(x + \sqrt{x}\xi)\right]^2 \tag{2.21}$$

for a chosen code rate $R = 1/M$ and $x = \frac{(M-1)P}{\sigma_n^2}$. The expectation is taken over ξ, which is a standard Gaussian random variable $\mathcal{N}(0, 1)$. Closed form expression for formula (2.21) does not exist, but its accurate approximations were presented in [31] and will be discussed in more detail in Chapter 3. Variance transfer curves for a number of coding rates are presented in Figure 2.7. As evidenced, repetition codes approach the bound relatively closely in the high noise region, which conforms with the observations made previously, stating that in the limit of high noise, repetition codes reach the bound. However, the gap increases visibly in the high SNR regime.

2.4.2 Hamming Codes

Hamming codes are one the earliest invented error correction/detecting codes and were introduced by Richard Hamming in [32]. They are block codes and are capable of correcting single errors since its parity-check matrix columns are linearly independent. The total length of the code is $n = 2^r - 1$ when encoding $k = 2^r - r - 1$ of information bits. Hamming codes are capable of correcting single errors exactly in accordance with the Hamming bound and therefore they are a rare example of perfect codes. It is known that there are no other perfect codes capable of correcting t errors, $t > 1$, except for the Golay codes: a binary $(23, 12, 7)$ that can correct triple errors and a ternary $(11, 6, 5)$ capable of correcting double errors. For more information on Golay codes see [33].

The parity-check matrix of a Hamming code has a very characteristic structure. Namely for an (n, k) code the columns are binary representations of the numbers $1, \ldots, n$. For example for the simplest $(7, 4)$ code we have

$$\mathbf{H}_{(7,4)} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}. \tag{2.22}$$

Hamming codes can be decoded using SISO algorithms. The results below are obtained using a simple one-way APP algorithm introduced in [34]. Figure 2.8 presents variance transfer of

[1] $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} dt$.

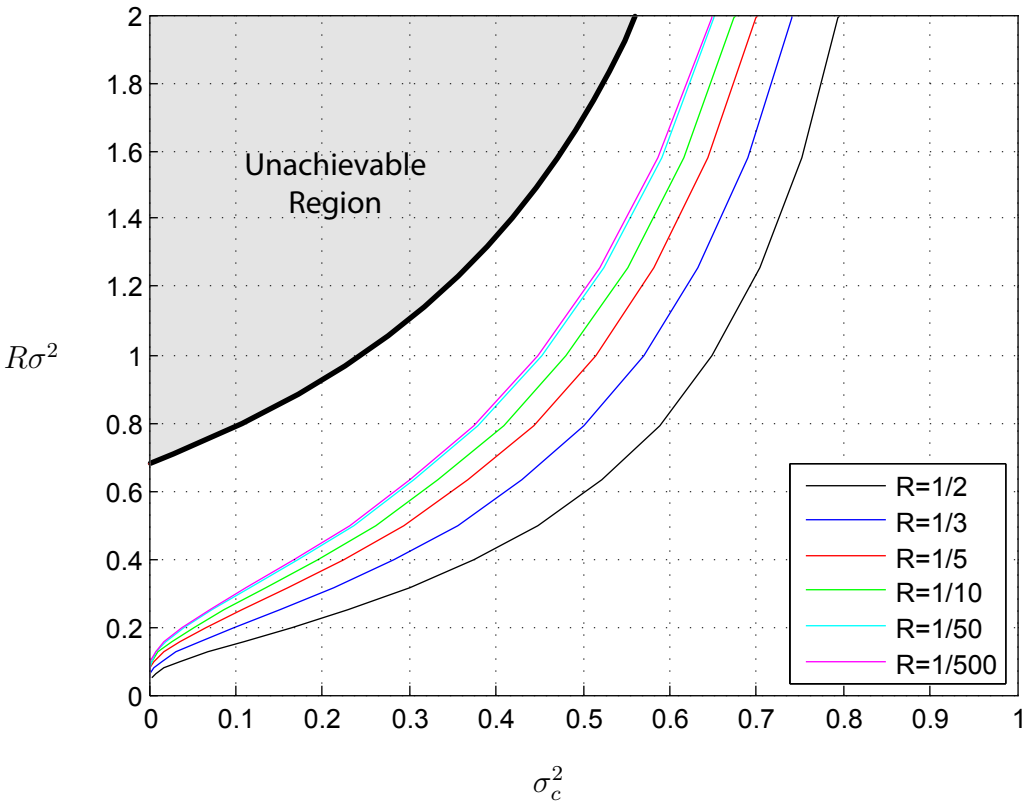

Figure 2.7: VT curves of repetition codes.

different rates and lengths. As clearly seen Hamming codes overall performance is quite far from the bound. Additionally, these codes have a very limited choice of achievable rates, i.e. they cannot be flexibly set. As we see from the graph and will observe later, Hamming codes have a very remarkable property. By increasing the rate they come closer to the bound, opposite as with repetition codes and other codes examined subsequently.

2.4.3 Convolutional Codes

Convolutional codes were introduced by Peter Elias in [35]. The canonical version of a simple convolutional encoder is presented in Figure 2.9. It consists of as shift register with delay elements. It has one input and two outputs, implying rate of 1/2. The connection from the shift register cells to the output is usually given using the octal notation. Since the first output is a function of all delay elements as well as the input, the connection can

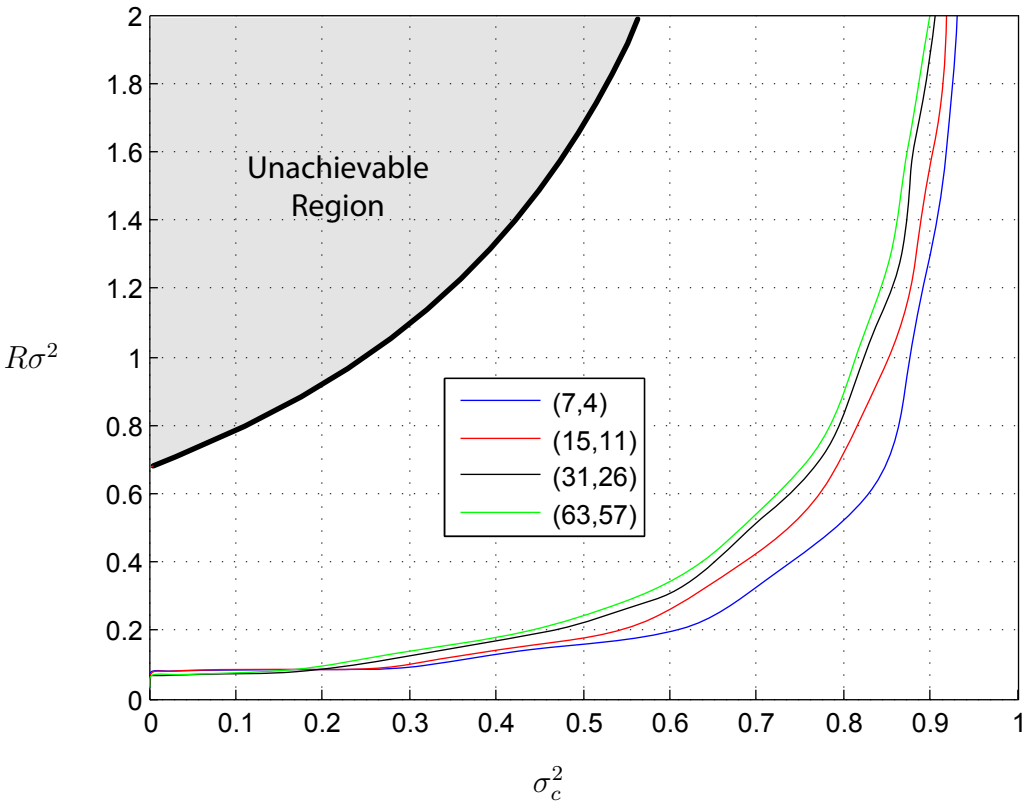

Figure 2.8: VT curves of Hamming codes of various rates.

be represented in binary form as [1 1 1 1], which translates into octal form of 17. Second output has a binary description [1 0 0 1], which is 12 in octal. Therefore this particular code is described as [17,12] convolutional code and its generator matrix is

$$\mathbf{G}(D) = [1 + D + D^2 + D^3 \quad 1 + D^2] \tag{2.23}$$

A convolutional code has a number of characteristic parameters that arise from the structure of its generator matrix. One is memory, which equals the number of delay elements in the encoder and in our case equals 3. Another parameter is the constraint length of the convolutional code, which is simply the memory plus the current input. Our example code has a constraint length of 4.

Convolutional codes are in fact the first class of codes that facilitated soft-decoding via the BCJR algorithm [16] or SOVA introduced in [36]. Convolutional codes were primarily of

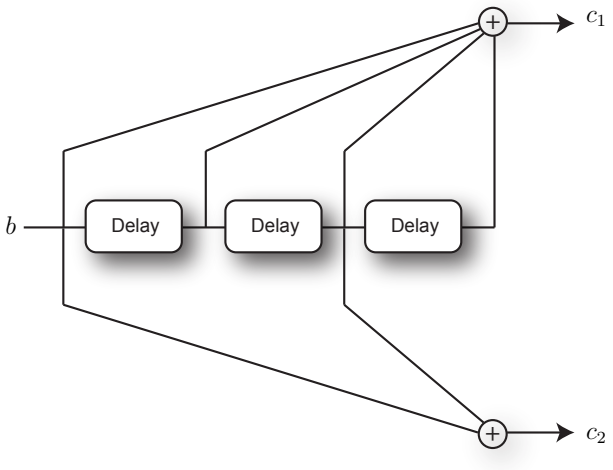

Figure 2.9: Encoder for a canonical version of rate 1/2 convolutional code with generator [17,12], i.e. $[1 + D + D^2 + D^3, 1 + D^2]$.

interest in the pre-turbo era as they could approach capacity limits within 2-3 dB, however at enormous expense in the decoding complexity [37]. This complexity grows exponentially with the memory of the code, since the number of possible states of the decoder grow according to exponential law. For example, for memory 3 convolutional code the decoding trellis of a binary code has $2^3 = 8$ states, but for memory 10 it is already exceeding 1000 states. This is the main drawback of convolutional codes, which was partly overcome with the invention of the turbo codes and iterative message processing. For more detailed overview of the convolutional codes, please refer to [37, 38, 39].

As stated before, for the purposes of interference cancellation, the decoder has to estimate the most likely transmitted codeword. For the case of convolutional codes, it means that the decoding complexity will be somewhat increased and this increase depends on the number of code bits per information bit, i.e. code rate. The BCJR algorithm needs to compute its internal variables (α, β, γ). The most computationally expensive of them is the calculation of γ, which is the state transition probability directly dependent on the channel information. When estimating LLRs of all the coded bits, the algorithm has to calculate set of γ variables only once and save them in different positions in a trellis according to the parity bits. Remaining variables (α, β) however need to be computed for each stream of parity bits.

In Figure 2.10 variance transfer curves for convolutional codes of different rates are presented. All codes are of length 720 information bits. As evidenced, decreasing the rate of the code does not help much in terms of decreasing the soft-bit variance for a given SNR

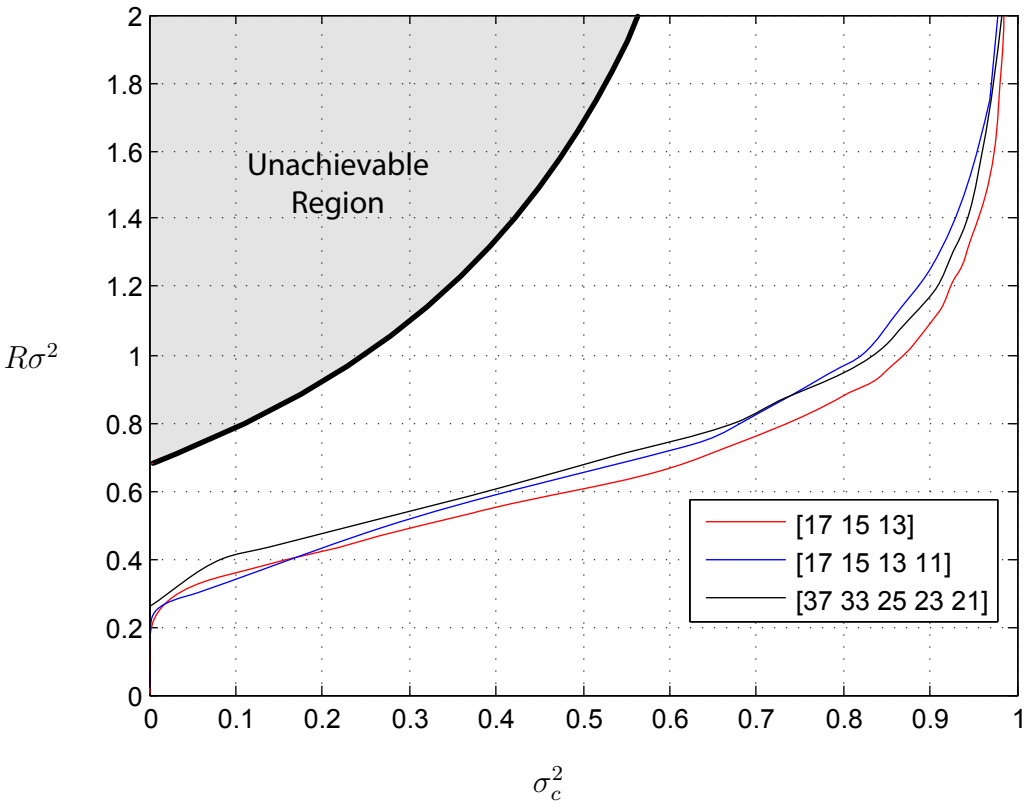

Figure 2.10: VT curves of convolutional codes of low rates.

value - it barely has any impact on the performance. On the other hand, decoding complexity of the low rate convolutional codes increases with lowering rate due to arguments discussed previously. As pessimistic as it may seem, lowering the rate does not help significantly in making convolutional codes good for IC purposes. We gain almost nothing in terms of lowering output variance of the estimated codeword by adding extra complexity to the inner coding. Additionally, lowering the rate in order to avoid mimicking a repetition code requires the use of longer generating polynomials which implies increasing the depth of the trellis. This causes the increase of the number of states of each constituent code. Summarizing, simple convolutional codes are almost as good as it gets for this family of codes and adding extra complexity is counterproductive.

2.4.4 Parallel Turbo Codes

Parallel turbo codes were also examined due to their excellent performance on the AWGN channel. Turbo code is a concatenation of simple convolutional codes separated by an interleaver. Similar concatenations were proposed before the invention of the turbo codes, however no interleaving operation was used then. A Turbo encoder, in order to guarantee an expected performance during the decoding process, has to meet certain criteria. First of them is that each component convolutional encoder has to be recursive systematic. Secondly, at least one of the encoders needs to be terminated to the initial zero state. This requirement is a consequence of using the BCJR decoding algorithm, which performs backward recursion assuming the final state of the trellis to be known. The structure of the famous Berrou turbo encoder is presented in Figure 2.11. It consists of two identical recursive systematic codes with an interleaver in between. The output of the encoder is the input information sequence (systematic output) and two parity streams, one from each RSC encoder.

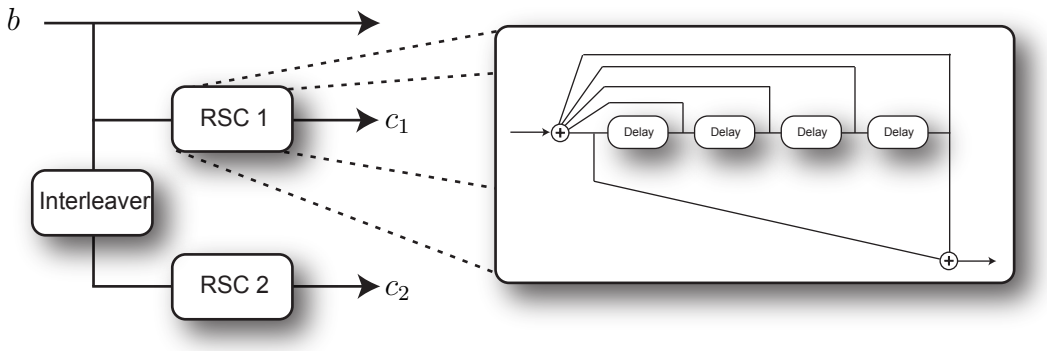

Figure 2.11: Original Berrou code with two identical component RSC codes [1, 21/37], i.e. $[1, \frac{1+D^4}{1+D+D^2+D^3+D^4}]$.

As mentioned, concatenating the number of component codes is not a new idea [40]. What is revolutionary and crucial to the fantastic performance of turbo codes is the decoding strategy. It is performed iteratively and consists of soft-decoding of each convolutional code. Component decoders exchange extrinsic information. Therefore extrinsic information is the new inference about the transmitted codeword, possible to obtain due to the code constraints, delivered by a single MAP decoding run. For more information on turbo codes see for example [41, 39].

To illustrate suitability for interference cancellation two famous codes were chosen, namely the Berrou code [10] and the 3GPP turbo code, being the choice for modern wireless broadband standards like HSPA and LTE [42]. The first code consists of two identical constituent codes of polynomials [37,21] and its decoding trellis has 16 states. The latter code has lower decoding complexity since it has only 8 states and generating polynomials [17,15]. Variance

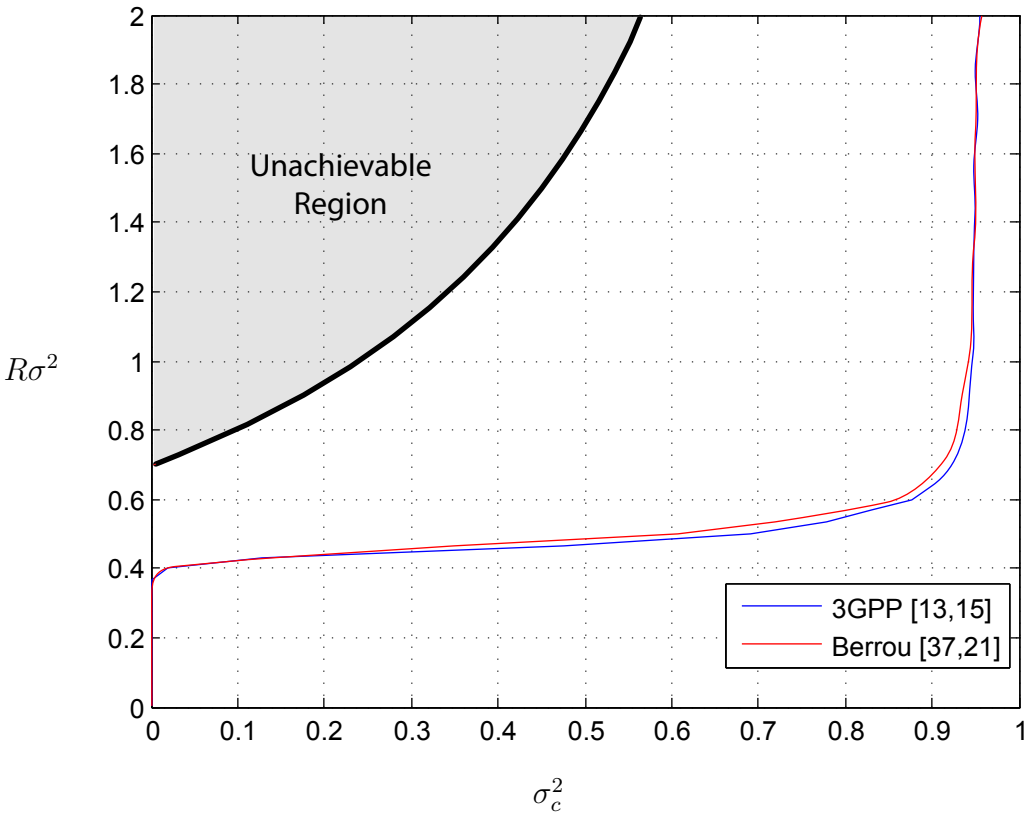

Figure 2.12: VT curves of popular parallel turbo codes of rate $R = 1/3$.

transfer curves of these codes are presented in Figure 2.12. It is well seen that turbo codes are not well suited for iterative cancellation systems. Of course only a fraction of possible codes is presented here, however there is strong evidence that any turbo code will perform equally poorly with respect to the optimal cancellation code. Note that a very pronounced sharp threshold (cliff) characteristics is present.

2.4.5 Regular LDPC Codes

The idea of LDPC codes was introduced by Gallager in his PhD thesis [43] and because of their decoding complexity largely forgotten until the discovery of turbo codes and major advances in the VLSI circuits domain. Today they have become a very intensively researched subject since they can perform close to the Shannon limit and their decoding complexity is in many cases lower than that of the turbo codes. Also the analysis of LDPC codes is relatively simple and delivers very precise results using the DE method.

Since an LDPC code is a linear block code it can be described uniquely using the parity-check matrix $\mathbf{H}$. Such a matrix has a very regular structure and for large codes, which is very often the case, is sparse (this fact is the main reason for calling these codes low density) - it contains a small number of non-zero entries. Gallager proposed constructing LDPC codes by randomly placing 1's and 0's in a parity-check matrix, subject to constraints that each row of $\mathbf{H}$ had the same number d_v of 1's and each column has the same number d_c of 1's. This condition determines whether the code is regular. If it is not fulfilled for each column/row pair, an irregular code is created. Gallager showed that the minimum distance of a "typical" regular LDPC code increases linearly with code size N as long as $d_v \geq 3$. Below the parity-check matrix of a code that has $d_v = 3$ and $d_c = 6$ that defined LDPC code of length $N = 18$ is presented[1]. From this parity-check matrix it is straightforward to derive the rate of the code. The number of coded symbols equals to the number of columns of $\mathbf{H}$ and is $N = 18$. The number of information bits is the number of columns less the number of rows, in this case $K = 9$, which gives the code rate $R = 9/18 = 1/2$. Using the (d_v, d_c) notation, rate is

$$R = \frac{d_c - d_v}{d_c} \tag{2.24}$$

implying that $d_c > d_v$.

$$\mathbf{H}_{(3,6,18)} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}. \tag{2.25}$$

An LDPC code can be viewed as a concatenation of a short repetition code and single

[1] This particular matrix is not actually "low density", but it serves mainly for illustration purpose.

parity-check code with a large interleaver in between, which from these short component codes create a large code, which as per Shannon theory, is necessary to approach the capacity limits. An LDPC code can be described graphically using a bi-partite Tanner graph [44]. Such a method of describing these codes provides useful insights into the decoding process and, generally, the behavior of these codes. The Tanner graph of our $(3, 6, 18)$ regular LDPC code is presented in Figure 2.13. Decoding of LDPC codes is performed by an iterative

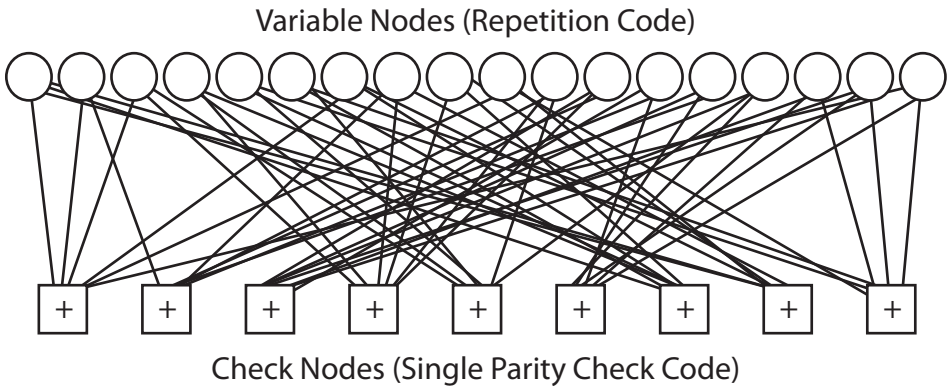

Figure 2.13: Tanner graph of the considered (3,6,18) regular LDPC code.

information exchange between variable and check nodes along the edges defined by the interleaver. This approach is referred to as message passing and this decoding methodology is a variation of the SPA algorithm. The convergence behavior of LDPC codes can be analyzed using the DE tool.

Density Evolution

Density Evolution is one of the most successful analytical tools used in predicting the performance of iterative systems. Originally, it was developed for LDPC codes [45], but it can be used in the analysis of any codes which are composed of a number of component codes separated by an interleaver. The DE idea is as follows. Since a classical LDPC code can be viewed as a concatenation of repetition code and a single parity-check code with an (preferably large) interleaver in between, the convergence behavior can be predicted via the input-output characteristics of each of these component codes. The DE relies on the *average values* of the LLRs exchanged between those two component codes. Although only average messages are used in this analysis, this approach describes the behavior of the decoder very accurately.

Since in this thesis regular LDPC codes are considered, where the degrees of variable and check nodes are the same for each type of node, DE for a such a family of codes is briefly outlined. For every LDPC code a variable node is an RC decoding node. Its input is the LLR from the channel $m_v^{(0)}$ and LLR values from the $d_v - 1$ check nodes $m_c^{(i-1)}$ connected to it[1]. Therefore its operation can be written as

$$m_v^{(i)} = m_v^{(0)} + (d_v - 1)m_c^{(i-1)}. \tag{2.26}$$

The variable node performs just the summation of incoming messages. On the other hand, the average output LLRs of the check nodes have a slightly more complicated expression

$$m_c^{(i-1)} = \Phi^{-1}\left(1 - [1 - \Phi(m_v^{(i-1)})]^{d_c - 1}\right) \tag{2.27}$$

where

$$\Phi(m_c) = 1 - \mathrm{E}\left[\tanh\left(\frac{U}{2}\right)\right] = 1 - \frac{1}{\sqrt{4\pi m_c}}\int_{\Re}\tanh\left(\frac{u}{2}\right)e^{-\frac{1}{4m_c}(u-m_c)^2}du. \tag{2.28}$$

Evaluation of this integral is computationally expensive therefore the following practical approximation was developed in [45]

$$\Phi(m) = \begin{cases} e^{-0.4527m^{0.86}+0.0218}, & m < 19.89 \\ \sqrt{\frac{\pi}{m}}e^{-\frac{m}{4}}(1 - \frac{11}{7m}), & m \geq 19.89. \end{cases} \tag{2.29}$$

where $m > 0$. The DE algorithm converges to error-free performance if $m_v^{(i)} \to \infty$. Otherwise the decoder delivers error rates strictly away from zero.

Suitability for IC

Variance transfer curves for the popular family of regular LDPC codes with variable node degree $d_v = 3$ are presented in Figure 2.14. A wide range of codes with different rates are investigated. The results suggest that those powerful codes are a poor choice for the cancellation systems. They exhibit a relatively sharp cliff behavior, but for low SNR value (large values of $R\sigma^2$) they do not provide any error correction capability. This results inevitably in severe error propagation phenomena, which effectively prevents the iterative system from converging even for moderate system loads.

[1] These are mean values.

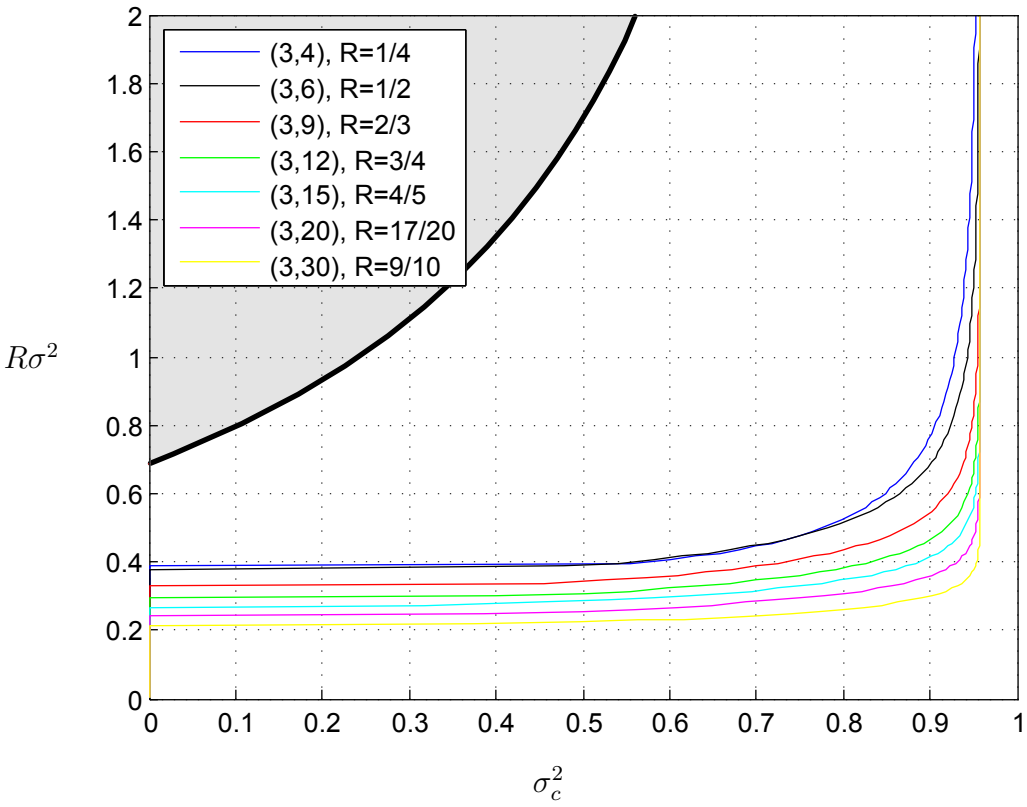

Figure 2.14: VT curves of regular $(3, x)$ LDPC codes.

2.4.6 VT Curve of the Correlation Detector

A number of transfer characteristics for modern channel coding schemes were presented above. What is left is the variance transfer function for the filter (detector) part, since the code VT interacts with the detector during the iterative interference cancellation process. Unlike for the majority of channel codes, the characteristics of the detector can be derived analytically. Such derivation for the simplest detector, the matched filter, is given in the following.

Recalling Equation (2.1) for an equal power channel, $\mathbf{A} = \mathbf{I}$, we have

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}. \tag{2.30}$$

The desired stream k can be separated out as

$$\mathbf{y} = \mathbf{h}_k x_k + \mathbf{H}_k \mathbf{x}_k + \mathbf{n}. \tag{2.31}$$

$\mathbf{H}_k$ denotes matrix $\mathbf{H}$ with the k-th column removed. Similarly $\mathbf{x}_k$ is a vector of modulation symbols $\mathbf{x}$ with the k-th element removed. After interference cancellation with signal estimates $\tilde{\mathbf{x}}_k$ we have the following signal for stream k

$$\mathbf{y}_k = \mathbf{y} - \sum_{\substack{i=1 \\ i \neq k}}^{N_t} \mathbf{h}_i \tilde{x}_i = \mathbf{h}_k x_k + \sum_{\substack{i=1 \\ i \neq k}}^{N_t} \mathbf{h}_i (x_i - \tilde{x}_i) + \mathbf{n}. \tag{2.32}$$

This signal is passed to the matched filter resulting in estimates of k-th stream data

$$\hat{x}_{\mathrm{MF},k} = {\mathbf{h}_k}^{\mathrm{H}} \mathbf{y}_k \tag{2.33}$$

which after a number of trivial operations becomes

$$\hat{x}_{\mathrm{MF,k}} = \underbrace{x_k}_{\text{signal}} + \underbrace{\sum_{\substack{i=1 \\ i \neq k}}^{N_t} {\mathbf{h}_k}^{\mathrm{H}} \mathbf{h}_i (x_i - \tilde{x}_i)}_{\text{interference}} + \underbrace{{\mathbf{h}_k}^{\mathrm{H}} \mathbf{n}}_{\text{noise}} \tag{2.34}$$

where signal, interference from other streams affecting stream k and noise components are clearly separated. We are primarily interested in the power of these three terms. Since all three terms in (2.34) are statistically independent we can look at them separately. Therefore the signal power is

$$P_S = \mathbb{E}[x_k {x_k}^{\mathrm{H}}] = P_k, \tag{2.35}$$

interference power is

$$P_I = \mathrm{E}\left[\left(\sum_{\substack{i=1 \\ i \neq k}}^{N_t} {\mathbf{h}_k}^{\mathrm{H}} \mathbf{h}_i (x_i - \tilde{x}_i) \right) \left(\sum_{\substack{i=1 \\ i \neq k}}^{N_t} {\mathbf{h}_k}^{\mathrm{H}} \mathbf{h}_i (x_i - \tilde{x}_i) \right)^{\mathrm{H}} \right] \tag{2.36}$$

which after some algebraic operations, and the fact that the expectation of cross-terms is zero, becomes

$$\mathrm{E}\left[\sum_{\substack{i=1 \\ i \neq k}}^{N_t} {\mathbf{h}_k}^{\mathrm{H}} \mathbf{h}_i (x_i - \tilde{x}_i)(x_i - \tilde{x}_i)^{\mathrm{H}} {\mathbf{h}_i}^{\mathrm{H}} \mathbf{h}_k \right]. \tag{2.37}$$

The sum can be carried in front of the expectation

$$\sum_{\substack{i=1\\ i\neq k}}^{N_t} \mathbb{E}\left[\mathbf{h}_k{}^{\mathrm{H}}\mathbf{h}_i(x_i-\tilde{x}_i)(x_i-\tilde{x}_i)^{\mathrm{H}}\mathbf{h}_i{}^{\mathrm{H}}\mathbf{h}_k\right]. \tag{2.38}$$

Since modulation symbols x are statistically independent of the channel $\mathbf{H}$, (2.38) can be rewritten as

$$\sum_{\substack{i=1\\ i\neq k}}^{N_t} \mathbb{E}\left[(x_i-\tilde{x}_i)(x_i-\tilde{x}_i)^{\mathrm{H}}\right]\mathbb{E}\left[\mathbf{h}_k{}^{\mathrm{H}}\mathbf{h}_i\mathbf{h}_i{}^{\mathrm{H}}\mathbf{h}_k\right]. \tag{2.39}$$

The quantity $\mathbb{E}\left[(x_i-\tilde{x}_i)(x_i-\tilde{x}_i)^{\mathrm{H}}\right]$ is the soft-symbol variance $\sigma^2_{c,i}$. Subsequently, if the random channel (see 1.3.1) is column normalized to unity the expectation over channel coefficients becomes

$$\mathbb{E}\left[\mathbf{h}_k{}^{\mathrm{H}}\mathbf{h}_i\mathbf{h}_i{}^{\mathrm{H}}\mathbf{h}_k\right] = \frac{1}{N_r}. \tag{2.40}$$

Combining (2.39) with (2.40) we obtain the final expression for the interference power as

$$P_I = \frac{1}{N_r}\sum_{\substack{i=1\\ i\neq k}}^{N_t}\sigma^2_{c,i}. \tag{2.41}$$

which is a linear function of individual soft-symbol variances.

The power of the noise component is

$$\begin{aligned} P_N &= \mathbb{E}[(\mathbf{h}_k{}^{\mathrm{H}}\mathbf{n})(\mathbf{h}_k{}^{\mathrm{H}}\mathbf{n})^*] \\ &= \mathbb{E}[(h^*_{k,1}n_1+\ldots+h^*_{k,N_r}n_{N_r})(h^*_{k,1}n_1+\ldots+h^*_{k,N_r}n_{N_r})^*] \\ &= \mathbb{E}[h^*_{k,1}n_1h_{k,1}n^*_1+\ldots+h^*_{k,N_r}n_{N_r}h_{k,N_r}n^*_{N_r}] \\ &= \mathbb{E}[h^*_{k,1}h_{k,1}n_1n^*_1+\ldots+h^*_{k,N_r}h_{k,N_r}n_{N_r}n^*_{N_r}] \end{aligned} \tag{2.42}$$

where cross-terms disappear due to the white noise assumption and due to independence of noise and channel coefficients we obtain

$$P_N = \sum_{i=1}^{N_r}\underbrace{\mathbb{E}[h^*_{k,i}h_{k,i}]}_{=1/N_r}\underbrace{\mathbb{E}[n_in^*_i]}_{=\sigma^2_n} = \sigma^2_n \tag{2.43}$$

Summarizing, the noise plus interference powers experienced by stream k at the output of

the matched filter detector in iteration j can be expressed as

$$\sigma_{k,j}^2 = \sigma_n^2 + \frac{1}{N_r} \sum_{\substack{i=1 \\ i \neq k}}^{N_t} \sigma_{c,i,j}^2. \tag{2.44}$$

Assuming that each stream has a different power P_i (2.44) becomes

$$\sigma_{k,j}^2 = \sigma_n^2 + \frac{1}{N_r} \sum_{\substack{i=1 \\ i \neq k}}^{N_t} P_i \sigma_{c,i,j}^2 \tag{2.45}$$

and values of $\sigma_{c,i,j}^2$ represent the quality of the codeword estimates delivered by the inner error control code for stream i in iteration j.

The variance transfer of the matched filter receiver with interference cancellation can be plotted. Since the curve is just a straight line it is enough to calculate two points and connect them. First point would be for the case where the soft-symbol variance is one. Then (2.45) becomes

$$\sigma_{k,j}^2 = \sigma_n^2 + \frac{1}{N_r} \sum_{\substack{i=1 \\ i \neq k}}^{N_t} P_i. \tag{2.46}$$

The second point is when the soft-symbol variance is zero and then (2.45) turns into

$$\sigma_{k,j}^2 = \sigma_n^2. \tag{2.47}$$

Figure 2.15 presents a number of matched filter detector variance transfer curves for various noise levels and antenna configurations with the assumption of equal, unit power streams, i.e. $P_k = 1$. The dependence between the amount of interference in the system, described by the slope of the curve, and the amount of noise, captured by the shift in the Y-axis direction, is very clearly visible. Usually the slope of these curves is called the system load, and in the large-system limit[1], when the array size grows ($N_r, N_t \to \infty$), it equals

$$\alpha = \frac{N_t}{N_r}. \tag{2.48}$$

In highly loaded systems the interference component dominates, the performance of lightly loaded systems is governed by the amount of additive noise. Armed with these tools it is now straightforward to predict the behavior of any MIMO system by combining the detector variance transfer curve as in Figure 2.15 with that of the inner code.

[1]For finite size systems it is $\alpha = \frac{N_t - 1}{N_r}$.

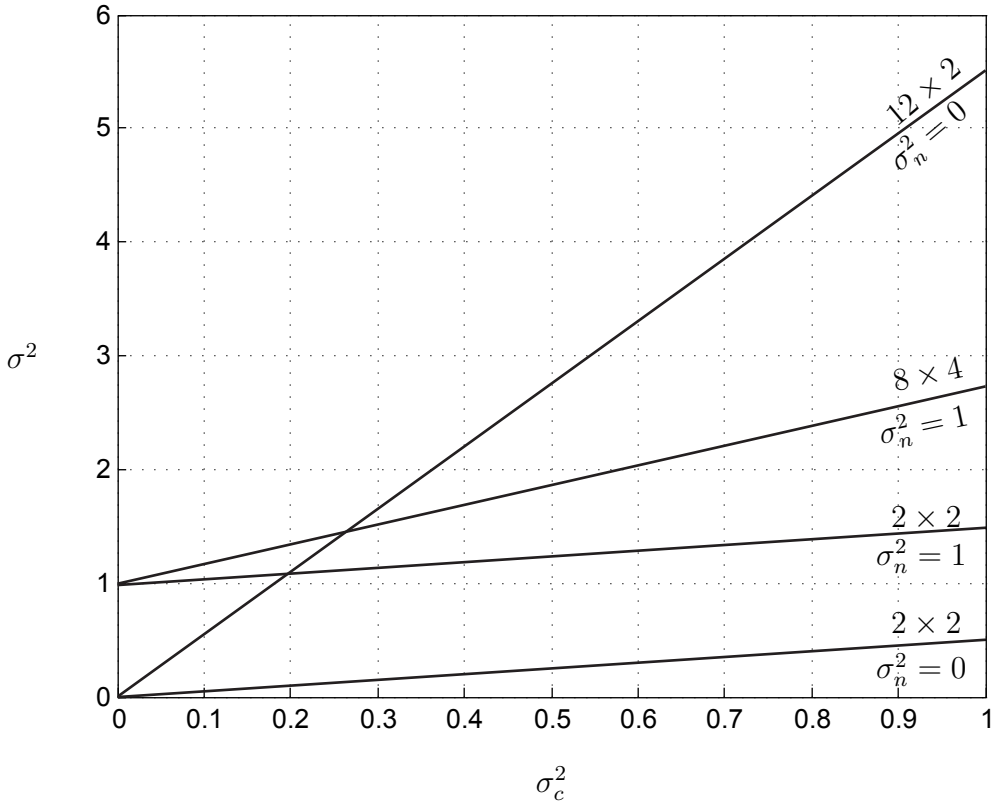

Figure 2.15: Variance transfer curves for matched filter detector and a number of MIMO $N_t \times N_r$ array configurations.

2.5 Numerical Experiments

The purpose of this section is to show that variance transfer charts indeed predict the behavior of iterative cancellation systems with the help of a series of computer simulation experiments. For this purpose a Rayleigh fading MIMO flat channel with N_t transmit and N_r received antennas is chosen. This channel model is an adequate representative of a general linear channel and attracted theoretical [13, 46] as well as practical interest. The channel matrix consists of N_t columns describing gains between given transmit antenna and all N_r received antennas

$$\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{N_t}]. \tag{2.49}$$

Each entry $h_{i,j}$ is complex Gaussian with unit variance, that is $h_{i,j} \sim \mathcal{CN}(0,1)$. In the simulations the following channel normalization is applied

$$\mathbf{h}_i{}^{\mathrm{H}}\mathbf{h}_i = 1 \tag{2.50}$$

meaning that columns are normalized to unity. This assures that transmitted power is the same as the received power for each stream. In other words the overall channel gain is equal to the number of streams[1].

The transmitter and receiver schemes are identical to those presented in Figure 2.1 and Figure 2.2, respectively. Such a system deals with 3 types of signals. One is the target signal, second the multi-antenna (multi-stream) interference, the third is additive noise. The signal to noise ratio of such a system, often acronymed by SINR, can be expressed as

$$\gamma = \frac{P}{\sigma_I^2 + \sigma_N^2}. \tag{2.51}$$

The evolution of this parameter during the iterative process is important and can be visualized with VT curves. The amount of interference in the MIMO system considered depends on the antenna array configuration, i.e. the number of transmit, receive antennas or more precisely the ratio of these i.e. N_t/N_r. This quantity is called the channel load α.

In order to compare different systems fairly the maximum spectral efficiency must be kept constant which translates into

$$\frac{N_t}{N_r} R_i R_o = \alpha R_i R_o = \text{const.} \tag{2.52}$$

where R_i and R_o denote inner and outer code rate, respectively. Subsequently, attention is drawn to three inner coding schemes: repetition, convolutional and turbo codes. We will show via simulations and VT analysis in what regions which coding scheme performs best. Moreover, the total transmit power is kept constant

$$P_t N_t = \text{const.} \tag{2.53}$$

to allow for fair comparisons. P_t denotes transmitted of each antenna.

2.5.1 System Parameters Setup

The question arises which inner channel code in which situation to use in order to achieve the highest spectral efficiencies. The answer can be delivered via careful analysis of the

[1] Each stream has a unity gain on average.

variance transfer charts. For the purpose of comparison we use the inner channel code like turbo, convolutional and repetition of various rates. Note that the Y-axis of the plots presented subsequently are not rate normalized anymore, because the interference canceller and matched filter does not depend on the rate of the applied coding scheme, since both of these entities operate on symbols, not the information bits.

Repetition Code

We can identify what spectral efficiencies are achievable for this coding scheme. Also, it is interesting to ask what maximum spectral efficiency can be attained for the interference cancellation receiver in the case when the SNR is high, in fact is going to infinity. To do that the non-normalized variance transfer curves of the code together with those of the matched filter receiver are necessary. They are presented in Figure 2.16. For clarity, a separate Figure 2.17 is given for low rate codes ($R = 1/15, \ldots, R = 1/50$). For example the maximum spectral efficiency for the case of a half-rate repetition code is

$$C_{\max} = \frac{N_t}{N_r} R = 2.1 \times \frac{1}{2} = 1.05 \quad \text{[bits/dimension]}. \tag{2.54}$$

where 2.1 describes maximum system load α for which rate 1/2 repetition code can converge - see rate 1/2 curve in Figure 2.17. When decreasing the rate, the spectral efficiency actually increases, since this decrease is compensated by a larger number of users (streams) that the system can accommodate. The values of these spectral efficiencies are gathered in Table 2.1. Summarizing, for the case of repetition codes and equal received powers, the maximum

Table 2.1: Maximum achievable spectral efficiencies for repetition codes.

Rate R	$\alpha_{\max}$	$\alpha_{\max} R$
1/2	2.1	1.05
1/3	4.3	4.3/3=1.43
1/4	6	1.5
1/5	8	8/5=1.6
1/6	10	5/3=1.66
1/7	12	12/7=1.71
⋮	⋮	⋮
1/15	30	2.0
⋮	⋮	⋮
1/50	102	102/50=2.01

spectral efficiency of roughly 2 bits/dimension can be attained in the high SNR limit. As it will be seen in the next chapters this limitation can be surpassed via unequal power allocations.

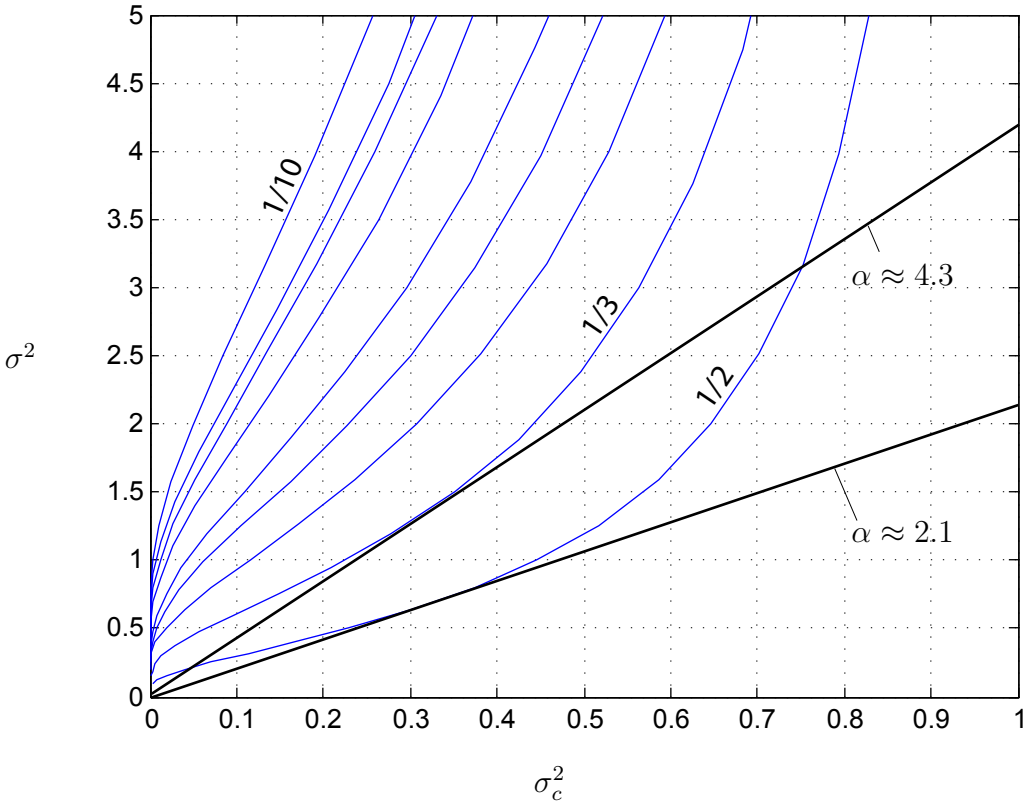

Figure 2.16: VT curves of repetition codes of various rates $R = 1/2, \ldots, 1/10$.

Convolutional Code

Table 2.2 shows the maximum spectral efficiencies of the considered convolutional codes and are obtained via the examination of Figure 2.18. These efficiencies are visibly lower than those of the repetition codes. Currently, there is no indication that other convolutional codes could have the potential of outperforming the repetition codes.

Turbo Code

Figure 2.19 presents the maximum spectral efficiency that strong turbo coding schemes can achieve. As it can be clearly seen both codes considered perform almost identically in this

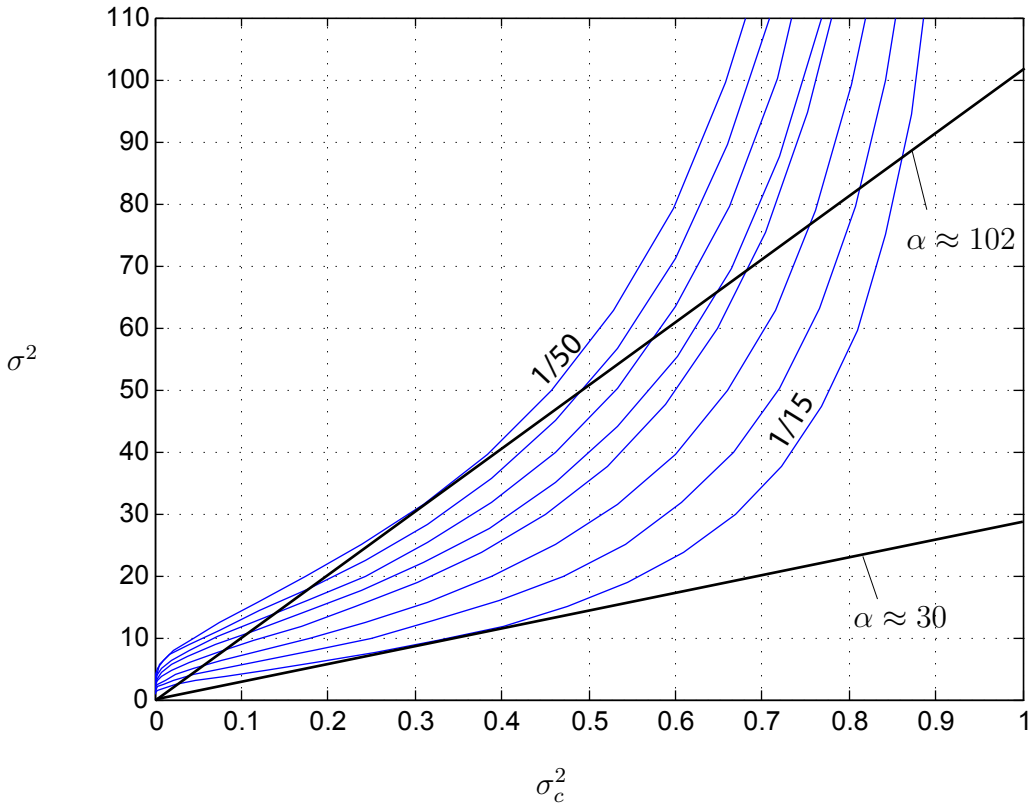

Figure 2.17: Non-normalized VT curves of repetition codes of various rates $R = 1/15, \ldots, 1/50$.

regard. The maximum spectral efficiency they can achieve is

$$C_{\max} = \frac{1}{3} \times 2 = \frac{2}{3} \quad \text{[bits/dimension]}. \tag{2.55}$$

Similarly, as in the case of convolutional codes, its doubtful that any other turbo scheme could perform significantly better.

2.5.2 Simulation Results

Subsequently, a couple of simulation results will be presented and explained. As it was already established, different coding schemes are optimal in different situations. Repetition codes are best for high spectral efficiencies and large SNR values. Turbo codes are the

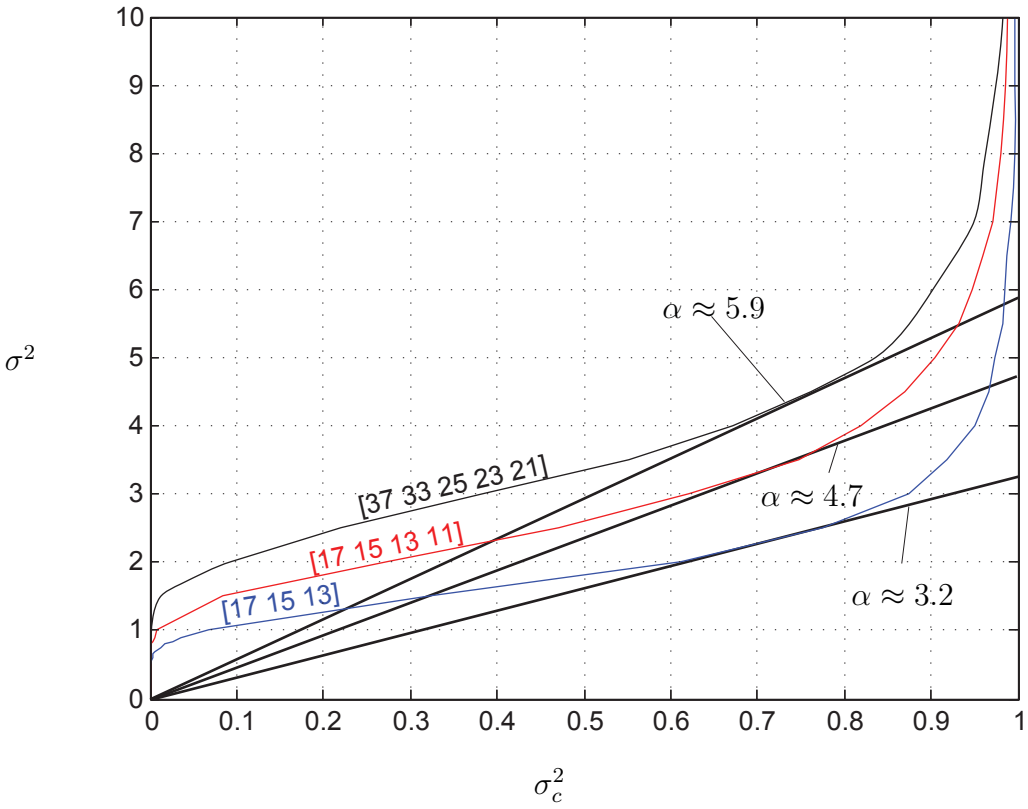

Figure 2.18: Non-normalized VT curves of convolutional codes of various rates $R = 1/3, 1/4, 1/5$.

winning choice for low spectral efficiencies and high noise levels and convolutional codes are placed between the two. Therefore, it is proposed to examine those three coding schemes in three different scenarios of low, moderate, and high spectral efficiency. For the case of an inner repetition code there is an outer LDPC code of high rate applied to allow the system to achieve error-free performance for reasonable SNR values. In subsequent discussions I is the number of turbo iterations, i.e. extrinsic information exchanges between the channel and the code.

Table 2.2: Maximum achievable spectral efficiencies for the considered convolutional codes.

Rate R	$\alpha_{\max}$	$\alpha_{\max}R$
1/3	3.2	1.06
1/4	4.7	1.17
1/5	5.9	1.18

Very Low Spectral Efficiencies, $\alpha R < 0.5$

For this setup it is certain that interference cancellation with inner turbo codes will converge for the lowest SNR values. To show this the following simulations were performed:

1. Inner turbo 3GPP code of length 720 information bits of rate 1/3, number of transmit antennas $N_t = 2$, number of receive antennas $N_r = 2$. Modulation: QPSK. The spectral efficiency is 2/3=0.66[1] bits/complex dimension

2. Inner convolutional code of length 720 information bits of rate 1/3 and generator polynomial [17 15 13], number of transmit antennas $N_t = 2$, number of receive antennas $N_r = 2$. Modulation: QPSK. The spectral efficiency is 2/3=0.66 bits/complex dimension

3. Inner repetition code of length 720 information bits of rate 1/3, outer (3,30) LDPC code of rate 0.9, number of transmit antennas $N_t = 2$, number of receive antennas $N_r = 2$. Modulation: QPSK. The spectral efficiency is 0.6 bits/complex dimension

The performance for the case of very low spectral efficiencies is presented in Figure 2.20. As expected the system employing inner turbo code provides the best performance beating both remaining systems by 2 and 3 dB. These results are not surprising and are in good agreement with the theoretical predictions.

[1] It is half, i.e. 0.33 with real modulation.

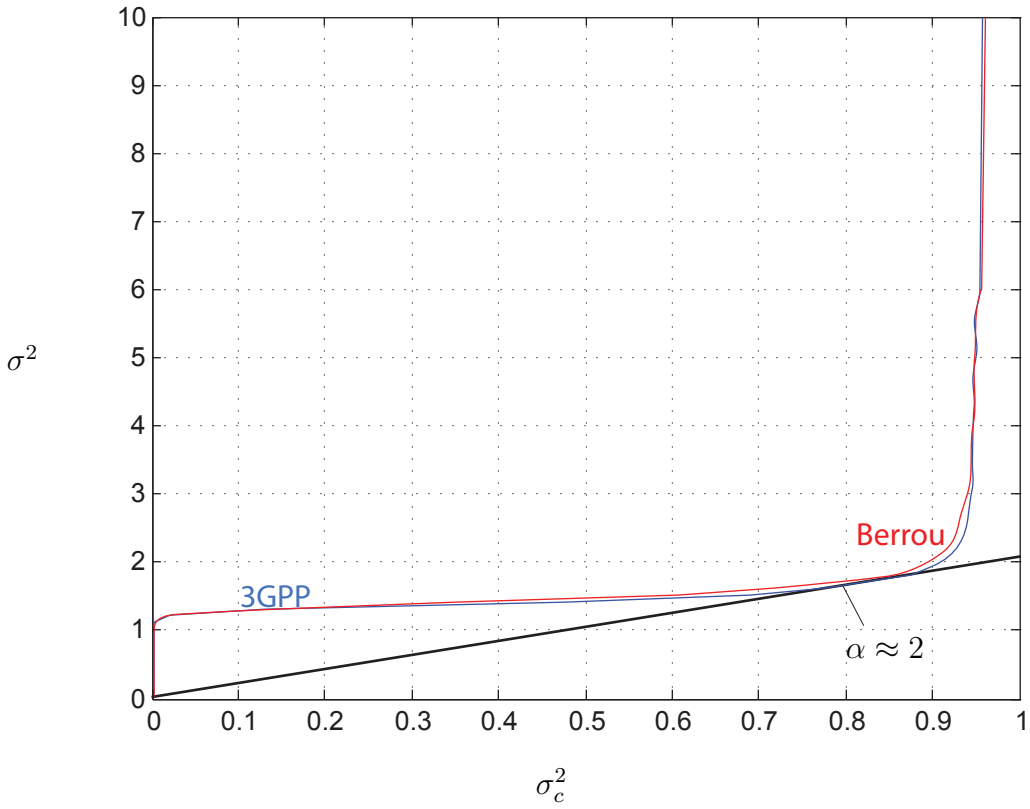

Figure 2.19: Non-normalized rate VT curves of popular turbo codes of rate $R = 1/3$.

Low Spectral Efficiencies, $\alpha R < 1.0$

As in the previous example for this setup it also is certain that all coding schemes will converge, but again for different SNR values. To show this the following simulations were performed:

1. Inner turbo 3GPP code of length 720 information bits of rate 1/3, number of transmit antennas $N_t = 4$, number of receive antennas $N_r = 2$. Modulation: QPSK. The spectral efficiency is 4/3=1.33 bits/complex dimension

2. Inner convolutional code of length 720 information bits of rate 1/3, number of transmit antennas $N_t = 4$, number of receive antennas $N_r = 2$. Modulation: QPSK. The spectral efficiency is 4/3=1.33 bits/complex dimension

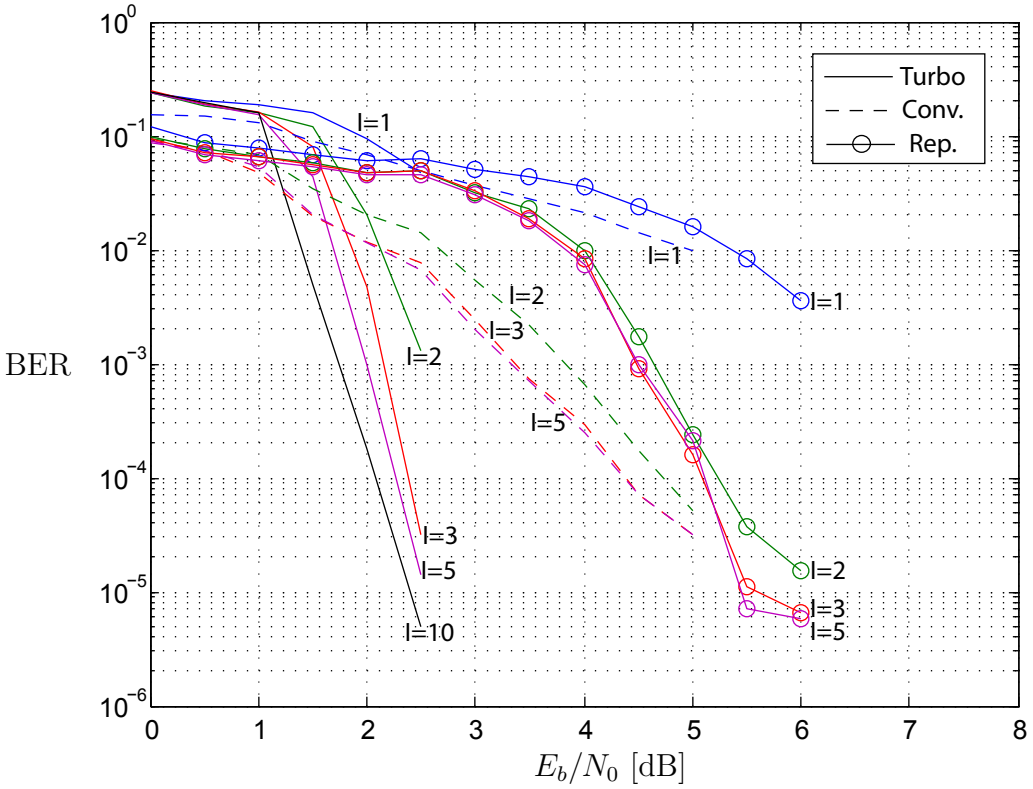

Figure 2.20: Bit error rates for very low spectral efficiencies $\alpha R \approx 1/3$.

3. Inner repetition code of length 720 information bits of rate 1/3, outer (3,30) LDPC code of rate 0.9, number of transmit antennas $N_t = 4$, number of receive antennas $N_r = 2$. Modulation: QPSK. The spectral efficiency is $0.9 \cdot 4/3 = 1.2$ bits/complex dimension

Figure 2.21 presents the performance for the case of very low spectral efficiencies $\alpha R < 1.0$. As expected by the analysis of variance transfer charts for this particular situation the system employing inner turbo code provides the worst performance beaten by both remaining coding systems. The gap in favor of the best scheme employing convolutional code is about 4 dB. Again these results are in good agreement with the theoretical predictions.

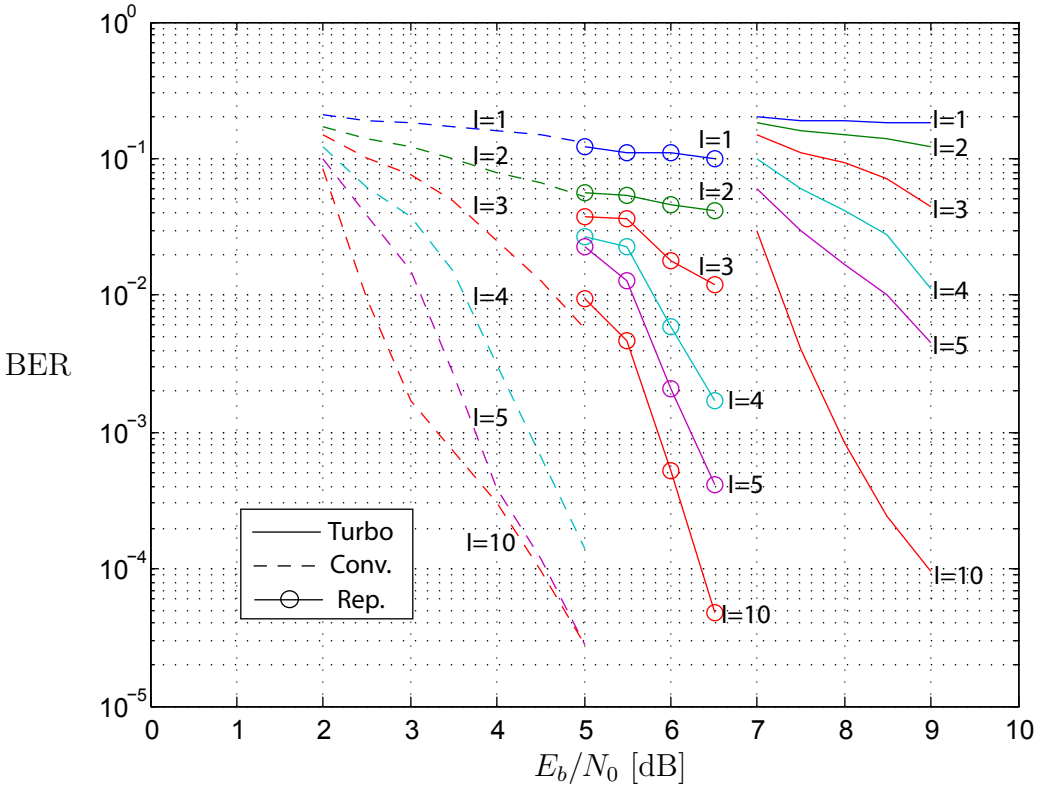

Figure 2.21: Bit error rates for low spectral efficiencies $\alpha R \approx 2/3$.

Medium Spectral Efficiencies, $\alpha R \approx 1.0$

For this setup it is certain that interference cancellation with inner turbo codes is expected not to converge at all. The convolutional code will converge first and after the repetition code. To show this the following simulations were performed:

1. Inner turbo 3GPP code of length 720 information bits of rate 1/3, number of transmit antennas $N_t = 6$, number of receive antennas $N_r = 2$. Modulation: QPSK. The spectral efficiency is 2.0 bits/ complex dimension

2. Inner convolutional code of length 720 information bits of rate 1/4, number of transmit antennas $N_t = 8$, number of receive antennas $N_r = 2$. Modulation: QPSK. The spectral efficiency is 2.0 bits/complex dimension

3. Inner repetition code of length 720 information bits of rate 1/6, outer (3,9) LDPC

code of rate 2/3, number of transmit antennas $N_t = 18$, number of receive antennas $N_r = 2$. Modulation: QPSK. The spectral efficiency is 2.0 bits/complex dimension

The number of transmit antennas for the case of repetition code may appear very large, much larger than current and even future MIMO systems could support. However, instead of streams from a single user, the system could accept and successively resolve number of streams from multiple users. For the case considered, where $N_t = 18$, it could be for example 9 users each equipped with two transmit antennas. Of course, the transmission of all the data streams would occur at the same time and utilizing the same bandwidth. As evidenced

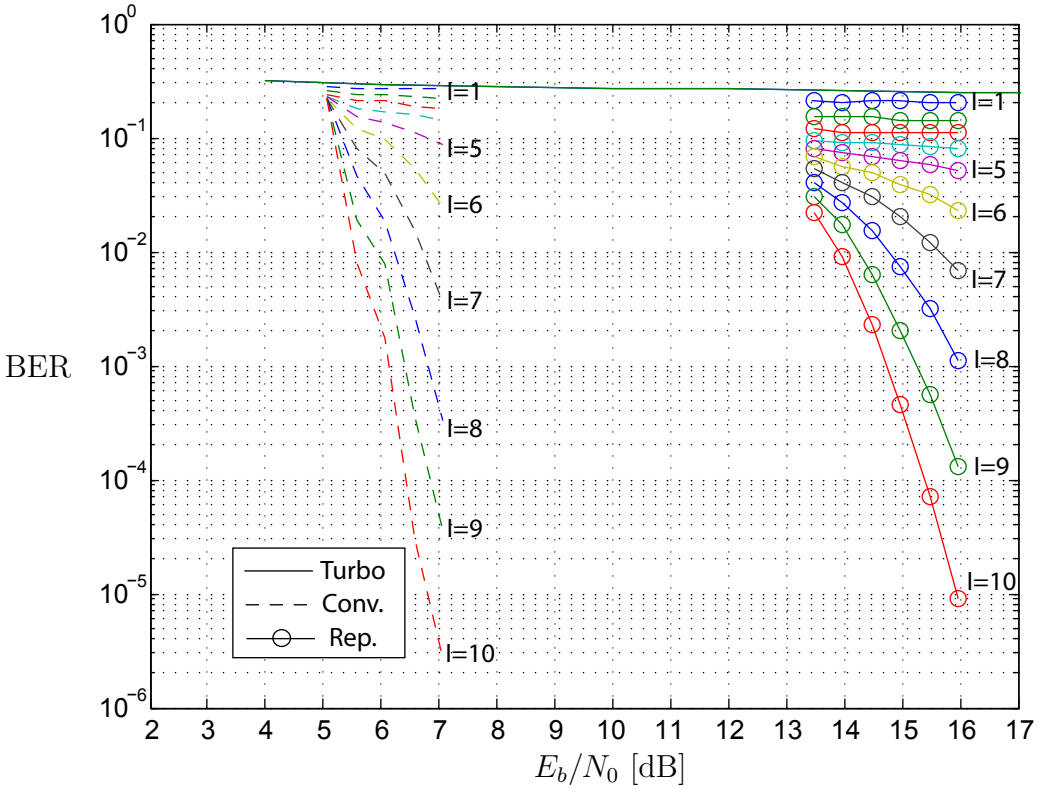

Figure 2.22: Bit error rates for medium spectral efficiencies $\alpha R \approx 1$.

in Figure 2.22 it is seen that strong inner turbo code does not converge at all, no matter how many turbo iterations are performed - the performance in the 10th iteration is shown with green solid curve. The experiment for the system with inner turbo code was performed until very high value of SNR $E_b/N_0 = 30$ dB and still the convergence has not been achieved. The clear winner for this system setup is the convolutional code beating the repetition code system by ca 9 dB.

High Spectral Efficiencies, $\alpha R \approx 1.1$

For this setup as previously, only the system with inner convolutional and repetition coding is expected to converge. The turbo coding system will not converge no matter how high of an SNR is available, therefore it is neglected in the numerical experiments. To illustrate this the following simulation was performed:

1. Inner convolutional code of length 720 information bits of rate 1/5, number of transmit antennas $N_t = 11$, number of receive antennas $N_r = 2$. Modulation: QPSK. The spectral efficiency is 2.2 bits/complex dimension

2. Inner repetition code of length 720 information bits of rate 1/5, outer (3,30) LDPC code of rate 0.9, number of transmit antennas $N_t = 12$, number of receive antennas $N_r = 2$. Modulation: QPSK. The spectral efficiency is 2.16 bits/complex dimension

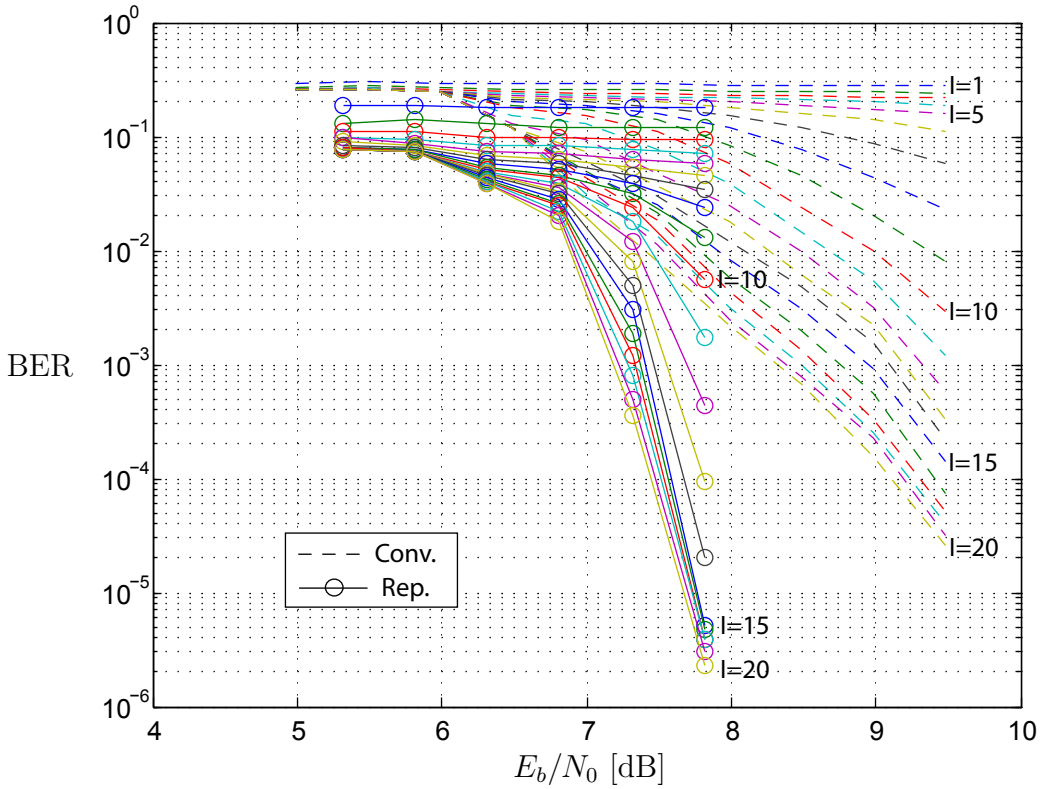

Figure 2.23: Bit error rates for spectral efficiencies $\alpha R \approx 1.1$.

Results for high spectral efficiencies are presented in Figure 2.23. We can see that in this case the best performer is the system employing inner repetition code and it has an advantage of ca 2 dB over the system with the inner convolutional code.

Highest Spectral Efficiencies, $\alpha R = 1.2$

For this setup only the system with inner repetition code is expected to converge. The turbo coding system will not converge no matter how high SNR is available, therefore it is neglected in the numerical experiments. To illustrate this the following simulation was performed:

1. Inner convolutional code of length 720 information bits of rate 1/5, number of transmit antennas $N_t = 12$, number of receive antennas $N_r = 2$. Modulation: QPSK. The spectral efficiency is 2.4 bits/complex dimension
2. Inner repetition code of length 720 information bits of rate 1/6, outer (3,30) LDPC code of rate 0.9, number of transmit antennas $N_t = 16$, number of receive antennas $N_r = 2$. Modulation: QPSK. The spectral efficiency is 2.4 bits/complex dimension

Finally, the spectral efficiency is boosted even higher. For $\alpha R = 1.2$ only the system based on inner repetition code can converge to error-free performance as evidenced in Figure 2.24. The setup using a convolutional code does not converge for examined signal-to-noise ratios. Its BER curve is very shallow and possibly error-free behavior can be expected but for very large values of SNR. Therefore for this scenario the repetition code is a clear winner.

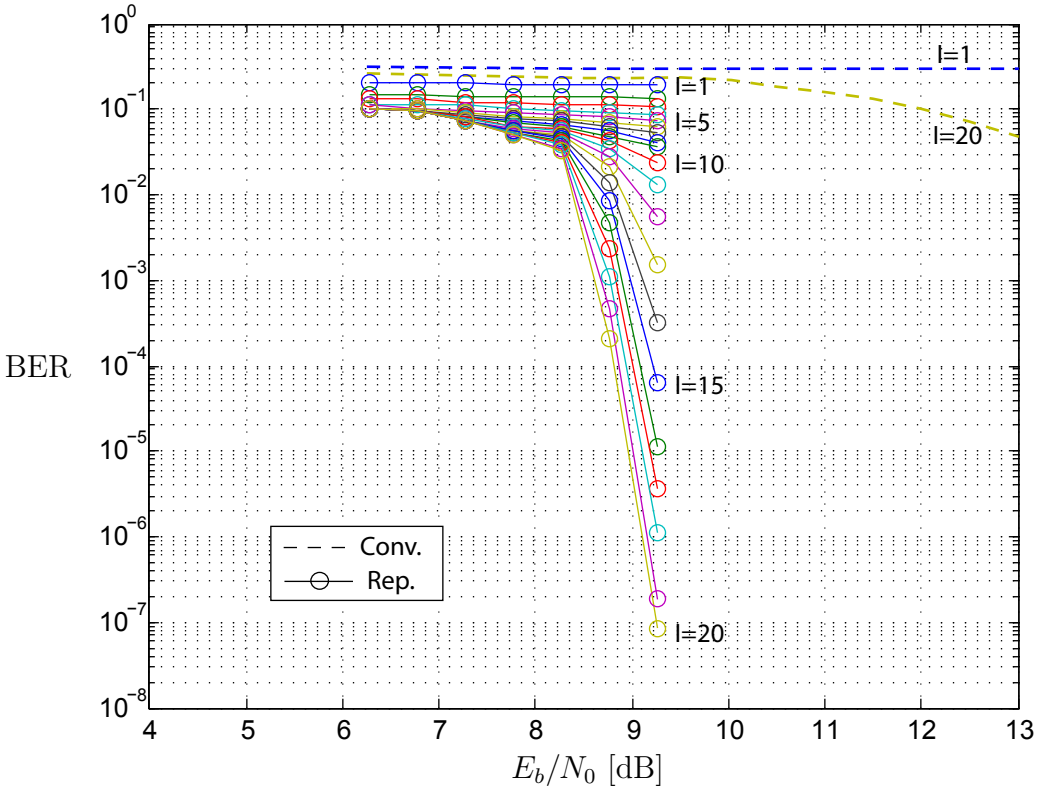

Figure 2.24: Bit error rates for spectral efficiencies $\alpha R \approx 1.2$.

2.6 Chapter Summary

This chapter contains an extensive introduction to the methods of describing the behavior of iterative interference cancellation systems. The concept of variance transfer charts is explained and a number of popular coding schemes are examined for their suitability in IC systems.

Although we do not know how to find the optimal coding schemes and moreover we do not know if such practical codes even exist, the desired properties of such a code can be easily identified by careful examination of the variance transfer charts. Such a code must successfully combine two contradicting characteristics. Namely, it should act as a repetition code in a very high noise region and as strong, capacity-achieving code in the moderate to low-noise regime.

The positive message is that such an optimal code can be approximated with a serial con-

catenation of two codes with the before-mentioned characteristics. The outer code should be a strong capacity-achieving code and the inner, the one facing the channel, a repetition code. Such concatenation however would likely not be decoded as customary in well known turbo fashion. The new processing schedule can be easily outlined using the variance transfer charts. This brings us to the key idea of this thesis - two-stage processing. It states that iterative message exchange between the canceller and the inner repetition code should be performed initially to remove the interference and then the outer code decoder just for removing the effects of additive noise should be invoked. Such an approach is especially suitable for high interference situations, for example in the case of MIMO communications it happens when the number of active data streams exceeds the number of receive antennas. A comparison between two-stage processing and traditional turbo equalization, where a feedback connection between the outer code and the canceller exist is the subject of the next chapter.

CHAPTER 3

Partitioned Spreading Code Division Multiple-Access

The main message of the previous chapter is that repetition coding is the best choice for an error control code facing the channel in iterative interference cancellation systems as long as interference is high. As it is known, CDMA modulation has an inherent built-in error control mechanism, i.e. spreading mimics the repetition coding. Utilizing and developing further this particular observation the PS-CDMA transmission is proposed [47, 48]. This chapter, besides the explanation of the system model and demodulation process, provides a precise analysis and performance evaluation of the proposed transmission methodology. A careful reader may immediately observe that PS-CDMA system is very similar to IDMA proposed in [7, 49]. However, PS-CDMA can be viewed as a generalization of IDMA and is, as presented in the previous chapter, motivated using precise arguments flowing from Chapter 2. IDMA on the other hand was motivated by the idea to distinguish user in the CDMA system not by signature sequence but by the interleavers. Additionally, in this thesis a precise analytical treatment of PS-CDMA is outlined, which was unknown at the time that IDMA was introduced.

The inner repetition code delivers symbol estimates with minimal residual variance for a range of SINR values, which means it effectively combats the interference coming from other users/streams. However, for high noise levels it performs poorly as shown in the previous chapter. Therefore an outer error control mechanism is required. Below attention is focused on regular LDPC codes, since they are easy to analyze, have excellent error correction capabilities and are very flexible in terms of adjusting the code rate.

The proposed two-stage processor relies on iterative information exchange between the inner repetition code and the interference canceller. It presents a turbo system. However, the inner repetition code is very easy to decode therefore implying small complexity requirements as opposed to traditionally viewed turbo equalization where stronger codes, like convolutional

codes, are used. Additionally, the way the repetition coding is utilized in the CDMA system (with normalizing the power to unity) incurs no rate loss as compared to the traditional CDMA with correlation or MMSE detection followed by the decoding of the applied channel code. In other words partitioned modulation is the concatenation of a rate $1/M$ repetition code with a CDMA channel of reduced dimensionality of N/M with interleaver in between.

3.1 Transmission Model

In this chapter chip-asynchronous CDMA, where a number of concurrent users transmit information over a GMAC channel, is considered. No fading or multipath is experienced and packet arrivals are synchronous, implying the level of interference remains constant in time. Chip asynchronousity models chip sampling process imperfections. For simplicity BPSK modulation is used, but the analysis can be extended to higher order formats (see Chapter 5). The system model is presented in Figure 3.1. Each user's information vector

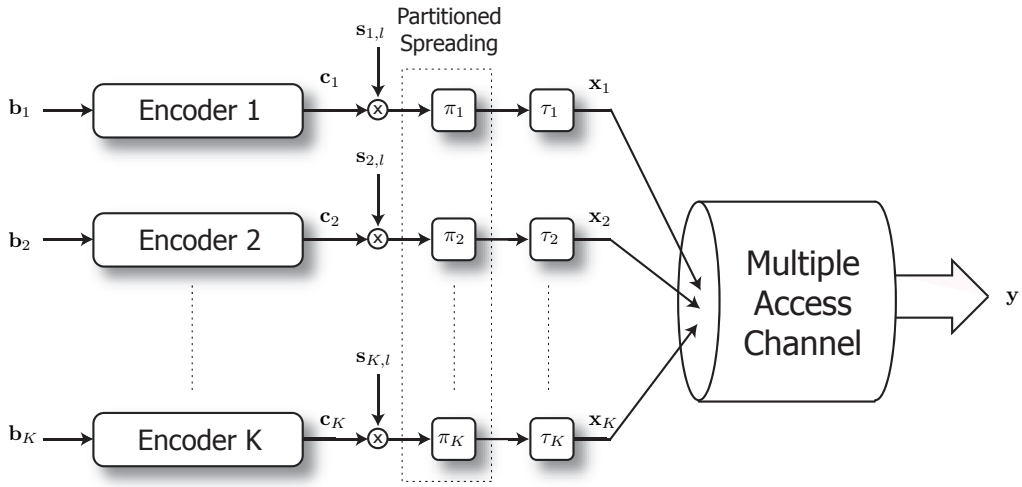

Figure 3.1: Partitioned Spreading CDMA transmitter.

$\mathbf{b}_k$ is initially encoded by the outer error control code producing the output vector $\mathbf{c}_k$. The users are distinguished by the unique pseudo–random spreading sequences $\mathbf{s}_1, \ldots, \mathbf{s}_K$, with chips chosen randomly and uniformly from the set $\{-1/\sqrt{N}, 1/\sqrt{N}\}$, where N is the bit interval in terms of chip times, usually referred to as spreading gain. Partitioned spreading relies on dividing length LN spreading sequences into a number of sections of equal size, say LM, and interleaving them using the permutation function $\pi_k(j)$, $0 \leq j < LM$, where

L is the length of the entire user data block in bits. Formally these operations are described as follows and visualized in Figure 3.2. Let T_s be the symbol (bit) duration, $T_c = T_s/N$ the duration of a chip, and $p(t)$ the unit–energy chip waveform. The k-th user signal $x_k(t)$ is given by

$$x_k(t) = \sqrt{P_k} \sum_{l=0}^{L-1} c_{k,l} \sum_{m=0}^{M-1} {s'}_{k,l,m} \left(t - lT_s - \tau_k - \pi_k(lM+m)\frac{T_s}{M} \right) \tag{3.1}$$

where

$${s'}_{k,l,m}(t) = \sum_{n=0}^{N/M-1} s_{k,n+lN+m\frac{N}{M}} p\left(t - nT_c\right) \tag{3.2}$$

represents the m-th partition for the coded bit $c_{k,l}$. P_k is the power of user k, and τ_k is the time delay of user k, which models the asynchronous situation. $p(t)$ is the impulse response of the chip pulse shaping filter. After spreading, the signals are transmitted over an AWGN channel and the resulting received signal is

$$y(t) = \sum_{k=1}^{K} x_k(t) + n(t) \tag{3.3}$$

where the noise power per dimension equals $\mathbf{E}[n^2(t)] = N_0/2 = \sigma^2$. Note that while (3.3) is a real system model, all results translate to complex equivalent systems without loss of generality.

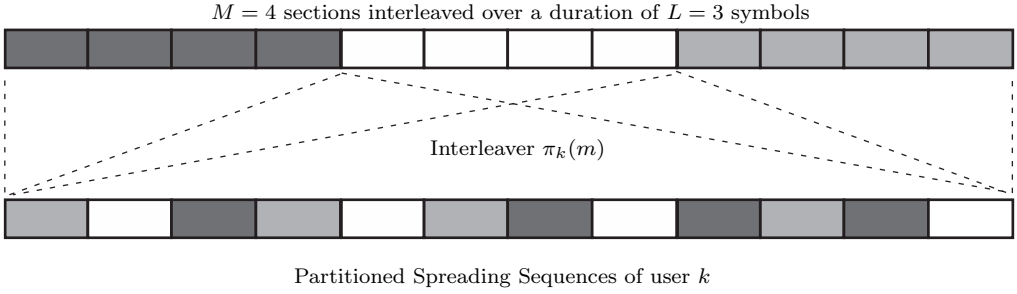

Figure 3.2: PS-CDMA modulation.

At the receiver, signals pass through a bank of filters $\sqrt{M}{s'}_{k,l,m}(-t)$ matched to the spread partitions. This detector operation is illustrated in Figure 3.3. The filter outputs of these

filters can be expressed as

$$z_{k,l,m} = \sqrt{\frac{P_k}{M}} c_{k,l} + I_{k,l} + n_{k,l} \tag{3.4}$$

where the noise samples $n_{k,l}$ have power σ_n^2 and the interference $I_{k,l}$ has power

$$\frac{\kappa}{N} \sum_{k' \neq k}^{K} P_{k'}. \tag{3.5}$$

The factor κ depends on the specific chip waveform used and is calculated for various pulses later in this chapter and in [50]. The purpose of interleaving with $\pi_k(\cdot)$ is to decrease the correlation between $z_{k,l,m}$, $m = 0, \ldots, M-1$, which deliver M (largely) independent estimates of the coded bits $c_{k,l}$. The noise and interference in (3.4) are well approximated

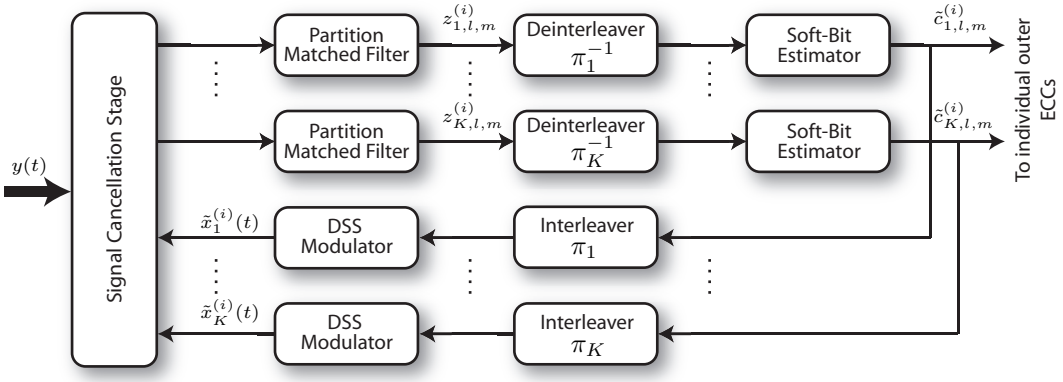

Figure 3.3: Structure of the partitioned CDMA iterative detector.

by a Gaussian random variable [22] and therefore its distribution is given by

$$f(z_{k,l,m}|c_{k,l} = \pm 1) \sim \mathcal{N}\left(\pm\sqrt{\frac{P_k}{M}}, \sigma_{0,k}^2\right) \tag{3.6}$$

where

$$\sigma_{0,k}^2 = \sigma_n^2 + \frac{\kappa}{N} \sum_{k' \neq k}^{K} P_{k'} \tag{3.7}$$

is the interference and noise power of the partition $z_{k,l,m}$. The cumulative a posteriori

probability taking into account signals from all M partitions is given by

$$\Pr\left(z_{k,l,1},\ldots,z_{k,l,M}|c_{k,l}=\pm 1\right)\sim\exp\left\{-\frac{1}{2\sigma_{0,k}^2}\sum_{m=0}^{M-1}\left(z_{k,l,m}\mp\sqrt{\frac{P_k}{M}}\right)^2\right\}. \tag{3.8}$$

Assuming $\Pr(c_{k,l}=\pm 1)=1/2$ the log-likelihood ratio (LLR) of $c_{k,l}$ at each iteration i is

$$\ln\left(\frac{\Pr(c_{k,l}=1|z_{k,l,0}^{(i)},\ldots,z_{k,l,M-1}^{(i)})}{\Pr(c_{k,l}=-1|z_{k,l,0}^{(i)},\ldots,z_{k,l,M-1}^{(i)})}\right)=\sum_{m=0}^{M-1}\frac{2}{\sigma_{i,k}^2}\sqrt{\frac{P_k}{M}}z_{k,l,m}^{(i)}=\sum_{m=0}^{M-1}\lambda_{k,l,m}^{(i)} \tag{3.9}$$

which is the summation of M LLR values of each partition

$$\lambda_{k,l,m}^{(i)}=\frac{2}{\sigma_{i,k}^2}\sqrt{\frac{P_k}{M}}z_{k,l,m}^{(i)}. \tag{3.10}$$

Equation (3.10) assumes that these LLRs are Gaussian (see the derivation in (1.1) and thereafter). Thus using (1.12) and not violating the extrinsic information principle we obtain a soft estimate of the coded partition at iteration $\tilde{c}_{k,l,m}^{(i)}$ as

$$\tilde{c}_{k,l,m}=\tanh\left(\frac{1}{2}\sum_{\substack{m'=0\\m'\neq m}}^{M-1}\lambda_{k,l,m'}^{(i)}\right)=\tanh\left(\frac{1}{\sigma_{i,k}^2}\sqrt{\frac{P_k}{M}}\sum_{\substack{m'=0\\m'\neq m}}^{M-1}z_{k,l,m'}\right). \tag{3.11}$$

The interference variance of the k-th user becomes independent of k as K becomes large, i.e., $\sigma_{0,k}^2\to\sigma_0^2=\sigma_n^2+\frac{\kappa}{N}\sum_{k=1}^{K}P_k$. Equation (3.10) describes the decoding of the inner repetition code while excluding self-message so not to violate the extrinsic information exchange rule. Second, equation (3.11) refers to obtaining soft-bit estimates from the LLR value obtained from the Gaussian channel, explained in (1.11) and (1.12). These soft-bits are then used to form a canceled signal for each user k as

$$\begin{aligned}y_k^{(i)}(t)&=y(t)-\sum_{k'\neq k}^{K}\tilde{x}_{k'}(t) \qquad (3.12)\\&=x_k(t)+\sum_{k'\neq k}^{K}\sqrt{\frac{P_{k'}}{M}}\sum_{l=0}^{L-1}\sum_{m=0}^{M-1}(c_{k',l}-\tilde{c}_{k',l,m}^{(i)})\cdot {s'}_{k',m}\left(t-\tau_{k'}-\pi_{k'}(lM+m)\frac{T_s}{M}\right)+n(t)\\&=x_k(t)+\sum_{k'\neq k}^{K}\sqrt{\frac{P_{k'}}{M}}\sum_{m=0}^{M-1}\delta_{k',l,m}^{(i)}\cdot {s'}_{k',m}\left(t-\tau_{k'}-\pi_{k'}(lM+m)\frac{T_s}{M}\right)+n(t)\end{aligned}$$

These canceled signals are used in iteration $i+1$ via repeated matched filtering to produce $z_{k,l,m}^{(i+1)}$. The quantity $\delta_{k',l,m}^{(i)}$ is the soft–bit error at partition m in the i–th iteration step and

its power is the soft-bit variance introduced in Chapter 2.

3.2 Analysis of the Detection Dynamics

The main analysis approach uses variance evolution. It can be shown that the noise and interference values in the iterative process follow a Gaussian distribution [22], and $\Pr(z_{k,l,m}(i)|c_{k,l} = \pm 1) \sim \mathcal{N}\left(\pm\sqrt{\frac{P_k}{M}}, \sigma_i^2\right)$. Together with equations (3.10) and (3.11), soft bit estimates can be expressed as

$$\tilde{c}_{k,l,m}(i) = \tanh\left(\frac{1}{\sigma_i^2}\sqrt{\frac{P_k}{M}}\sum_{m\neq m'}^{M} z_{k,l,m'}\right) \tag{3.13}$$

$$= \tanh\left(\frac{(M-1)P_k}{M\sigma_i^2} + \xi\sqrt{\frac{(M-1)P_k}{M\sigma_i^2}}\right) \tag{3.14}$$

where $\xi \sim \mathcal{N}(0,1)$. The factor $(M-1)/M$ is a consequence of excluding the self-message in (3.10). These messages will be fed back subsequently to the canceller.

Define the following quantity which constitutes the expectation of the soft–bit error $\delta_{k,l,m}$ as

$$g(s) = \mathbb{E}\left[\left(1 - \tanh\left(s + \xi\sqrt{s}\right)\right)^2\right] \tag{3.15}$$

Although no trivial closed form expression exists for $g(s)$, tight upper bounds are known [31], such as

$$\hat{g}(s) \leq \begin{cases} \frac{1}{1+s}, & s < 1 \\ \pi Q(\sqrt{s}), & s \geq 1 \end{cases} \tag{3.16}$$

and $Q(\cdot)$ is the Gaussian error function already defined in Chapter 2. Now, following (3.12), the evolution of the interference and noise power can be written as

$$\sigma_{i+1}^2 = \frac{\kappa}{N}\sum_{k=1}^{K} P_k g\left(\frac{(M-1)P_k}{M\sigma_i^2}\right) + \sigma_n^2. \tag{3.17}$$

For an equal power scenario, $P_k = P$, and for N, M and number of detection iterations, I, all tending to infinity (3.17) simplifies to

$$\sigma_\infty^2 = \kappa\alpha P g\left(\frac{P}{\sigma_\infty^2}\right) + \sigma_n^2 \tag{3.18}$$

where α is the CDMA channel load

$$\alpha = \frac{K}{N} \tag{3.19}$$

and is analogous to the channel load in (2.48) expressed for the case of the multi-stream MIMO transmission.

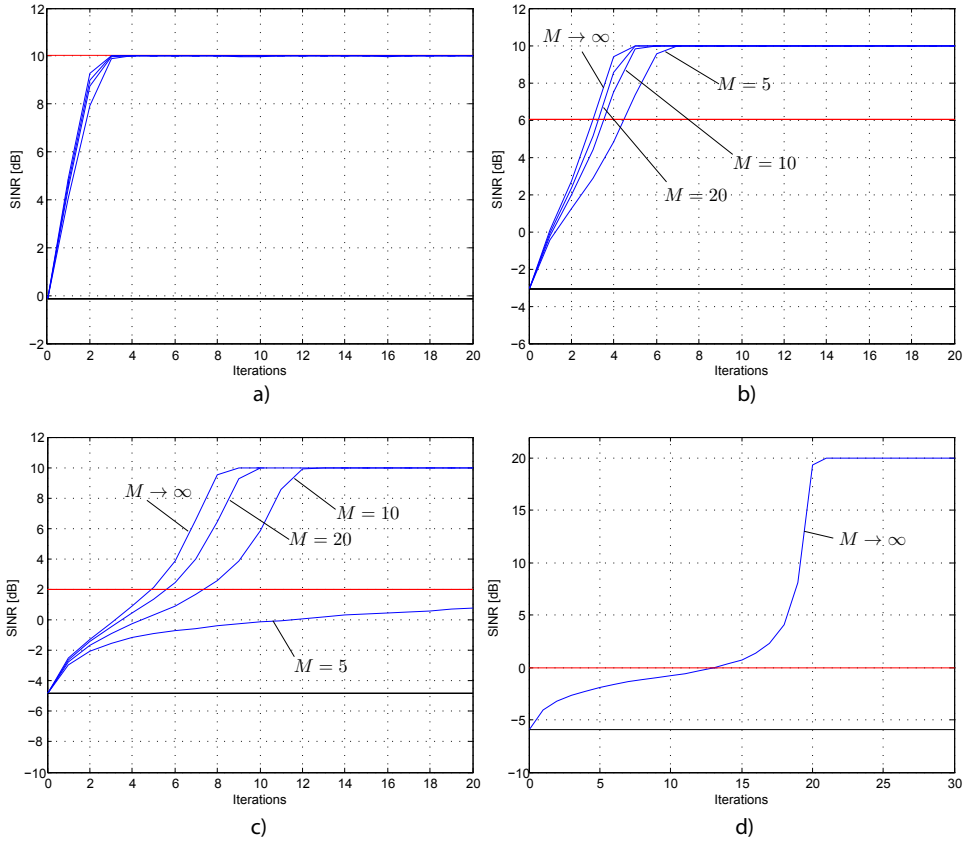

Figure 3.4: SNR evolution of different CDMA detectors for different channel loads a) $\alpha = 0.5$ b) $\alpha = 1.0$ c) $\alpha = 1.5$ when $E_b/N_0 = 10$ dB. d) $\alpha = 2.0$ and $E_b/N_0 = 20$ dB, $M \to \infty$. Spreading gain $N = 30$.

Figure 3.4 presents the SINR evolution of popular CDMA detectors for a number of system loads. For comparisons a performance matched filter (black curve) as well as MMSE receiver (red curve) is shown. We can see that partitioned demodulation with matched filtering outperforms both traditionally adopted detectors in every considered situation. Since the partitioned demodulator is based on the matched filter the initial SNR delivered by both is the same. As evidenced, the performance curves are affected by the repetition factor M. Its impact is more pronounced for higher loaded systems. Additionally, the partitioned system achieves interference-free performance in almost all the cases, except the situation when $\alpha = 150\%$ and M is small.

It is very interesting to ask what maximum system load can be supported by the partitioned demodulation. For limiting arguments, $\sigma^2 \to 0$ as well as increasing the repetition factor such that $M \to \infty$, the maximum supportable load α for synchronous partitioned CDMA demodulator is described by the following

$$\sigma_\infty^2 = \alpha_{\max} P g\left(\frac{P}{\sigma_\infty^2}\right) \tag{3.20}$$

which implies that

$$\alpha_{\max} = \lim_{\sigma_\infty^2 \to 0} \frac{\sigma_\infty^2}{Pg\left(\frac{P}{\sigma_\infty^2}\right)} \approx 2.09. \tag{3.21}$$

It is assumed here that the system does not get stuck before on some $\sigma_\infty^2 \neq 0$. In other words maximum load is such that it prevents variance transfer curves of the repetition code and the matched-filter canceller to overlap except at the final point ($\sigma_{c,\infty}^2 = 0, \sigma_\infty^2 = 0$). This condition can be grasped graphically with the variance transfer charts as presented in Figure 3.5. As customary for a matched filter canceller the load is equivalent to the slope of the cancellation curve. Two sets of VT curves correspond to exact evaluation of the soft-bit variance function as in (3.15) and to the approximation which is the upper bound. As is indicated the maximum load is ≈ 2.1 and ≈ 1.9 corresponding to the exact evaluation and the approximation, respectively. Naturally, in practical situations these maximum load figures will be slightly different (smaller) due to the number of factors, like noise level and finite repetition factor M. Nevertheless, this result indicates a great potential of the partitioned modulation not only in terms of efficiency but also implementation feasibility [51].

3.2.1 Optimality of Partitioned Modulation

Expressing the partitioned detector's output SINR after an infinite number of iterations and in the limiting case $M \to \infty$ as $\gamma_\infty = P/\sigma_\infty^2$ it can be derived that

$$\gamma_\infty = \frac{P}{\sigma_n^2 + \kappa\alpha P \mathbf{E}\left[1 - \tanh\left(\gamma_\infty + \sqrt{\gamma_\infty}\xi\right)\right]^2} \tag{3.22}$$

where equation (3.18) is used. In [8] it was shown that the output signal-to-noise and interference ratio of the individually optimal (APP) detector for the original (unpartitioned) CDMA system satisfies (3.22) for $\kappa\alpha < \alpha_s \approx 1.49$. This shows that partitioned CDMA with iterative processing is an asymptotically optimal demodulation method for equal–power CDMA, as long as $\kappa\alpha < 1.49$. The analysis carried out in [8] for dense case (without interleaving) breaks for values of loads higher than 1.49 and it is unknown whether above this threshold the APP estimates are still delivered. Due to that, the PS-CDMA method

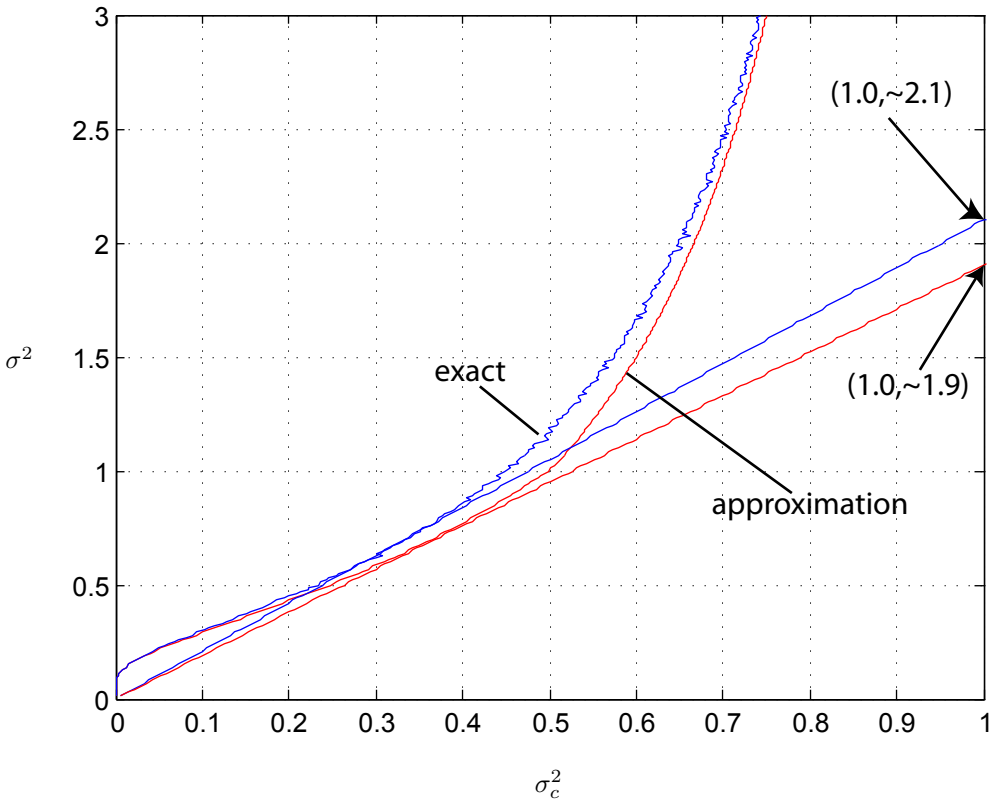

Figure 3.5: Maximum system load α achievable for PS CDMA demodulator in equal power scenario.

is an example of an optimal detection method with bounded (non-exponential) complexity. Moreover, following [19, equation (6)] the noise plus interference power at the output of the minimum mean–squared error (MMSE) filter again applied to the equal power CDMA satisfies

$$\sigma^2_{\text{MMSE}} = \frac{\kappa\alpha P}{1+\frac{P}{\sigma^2_{\text{MMSE}}}} + \sigma_n^2 \geq \kappa\alpha P g\left(\frac{P}{\sigma^2_{\text{MMSE}}}\right) + \sigma_n^2 \tag{3.23}$$

which leads to $\sigma^2_\infty \leq \sigma^2_{\text{MMSE}}$ using (3.16). Therefore it is implied that partitioned CDMA always performs better than MMSE–based CDMA detection, and, where the two are near equal, MMSE filtering is near optimal by the argument above.

3.2.2 Simulations

Until this point performance evaluation of the partitioned demodulator relied on solving the iterative equation (3.17). In this subsection the validity of this approach will be examined using numerical simulation experiments. This comparison is mainly based on variance evolution approach. It will be clearly seen that simulations are in perfect agreement with the analysis given previously.

Two systems were simulated, with $K = 45$ and $K = 60$ users. Remaining parameters are shared, namely spreading gain $N = 30$, repetition factor $M = 10$ and channel $E_b/N_0 = 10$ dB. The frame length is set to $L = 120$ implying the interleaver length of $LM = 1200$. SNR evolution is presented in Figure 3.6. It is clearly observed that the analysis captures the

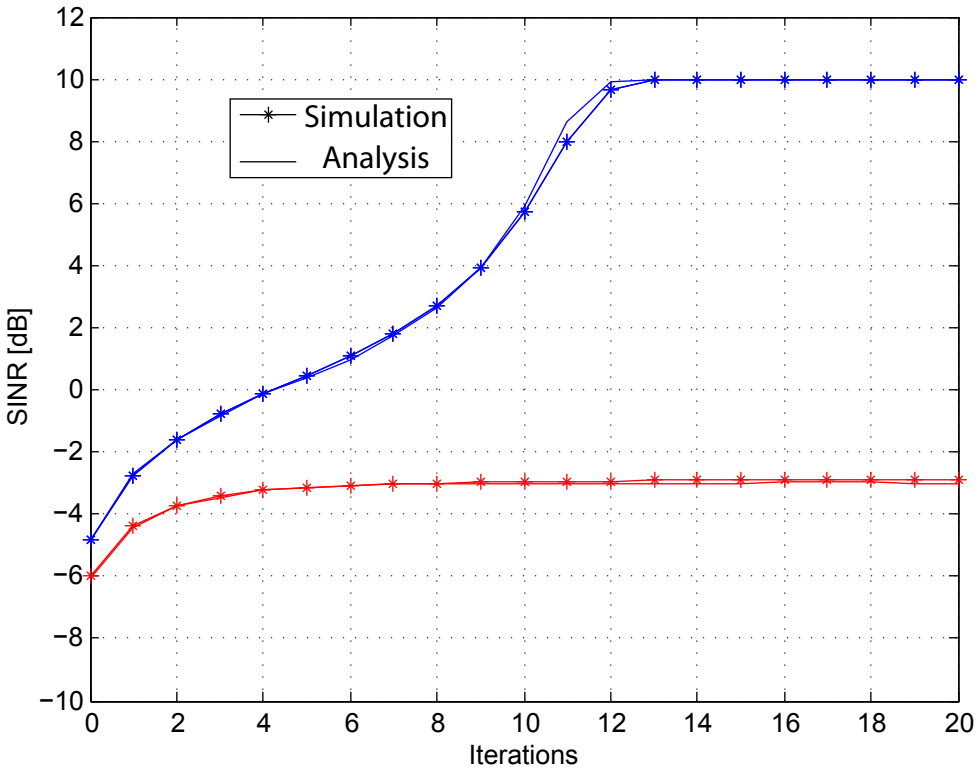

Figure 3.6: Comparison between analytical evaluation of variance evolution and that obtained from the numerical simulation.

behavior of the partitioned demodulator very precisely and virtually there is no difference

between the analysis and the simulation. With this observation we can safely assume that SNR evolution for different systems presented previously in Figure 3.4 is the same as the simulation would deliver. Summarizing, it is enough to evaluate the system performance analytically without the need to run expensive and time consuming numerical experiments. This observation is true however as long as interleaver size, which is the product of frame length L and number of partitions M, is large. With the decrease of frame length the gap between the analysis and simulation might start to show up, since in the analysis it is assumed that the LLR values in the repetition code decoding block are not correlated. When the correlation starts to build-up the analysis is no longer valid. The effect of correlations will be more pronounced for high loaded systems and it will be discussed in the subsequent paragraph.

3.3 Practical Considerations

3.3.1 Finite Interleaving

One of the most pronounced characteristics of the partitioned transmission is interleaving. The analysis presented in the previous section bears the crucial assumption that interleaving is large enough to assure independence between partition estimates $z_{k,l,m}$. Variance transfer (see Equation 3.17) and all its derivatives are valid if interleavers are large. From practical point of view large interleavers are unfeasible, especially in wireless environments, where propagation conditions change very rapidly and large frame sizes cannot be easily accommodated.

Correlation is the phenomena that destroys partitioned system performance as compared to the theoretical predictions. The shorter interleaving the more severely correlation affects the system. Figure 3.7 shows the divergence from the theory due to use of finite interleaver. As it can be observed short interleaving makes system performance saturate away from final performance predicted by the theory. However, for larger interleaver (length 2000) system results in behavior very closely to theory. It must be stressed here that gap seen in the Figure 3.7 translates into an SNR loss of several decibels, which is very considerable, looking at today's standards[1]. In Figure 3.8 the minimum required interleaver size depending on the system load α for PS-CDMA detector to follow (3.17) is presented. It is seen that for underloaded system ($\alpha < 1$) interleaver size practically does not matter. Its role is becoming more important as channel load increases. As it can be seen above certain threshold load (about 1.75) interleaver size does not matter again. This is due to the fact that the system does not converge for a given load ($\alpha > 1.75$) even for the infinite case. As already stated,

[1] In Figure 3.7, final variance of the finite system is about 4 times higher than that of the perfect

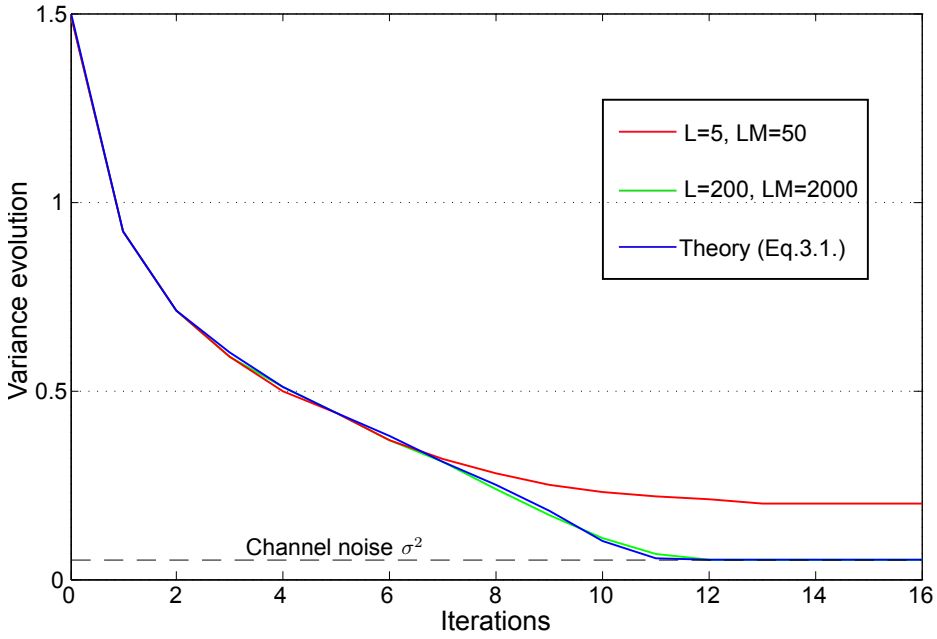

Figure 3.7: Comparison of theoretical predictions and simulated performance of PS-CDMA for the following parameters: number of users $K = 30$, partition multiplicity $M = 10$, spreading gain $N = 20$, $E_b/N_0 = 10$ dB. Interleaver sizes are 50 and 2000.

performance degradation due to short interleavers is observed. To understand the cause of this effect let us concentrate on equation (3.14) which evaluates the soft–bits. This equation is only valid if estimates $z^i_{k,l,m'}$ are uncorrelated, which is true for large interleavers. With interleavers size reduction this correlation starts to play a significant role. With the correlation between messages corresponding log–likelihood–values

$$\lambda^i_{k,l,m'} = \frac{2}{\sigma_i^2} z_{k,l,m'} \tag{3.24}$$

coming from the repetition code are correlated and their sum

$$\sum_{m' \neq m}^{M} z^i_{k,l,m'} \tag{3.25}$$

theoretical variance obtained via (3.17). It translates into SNR loss of 6 dB.

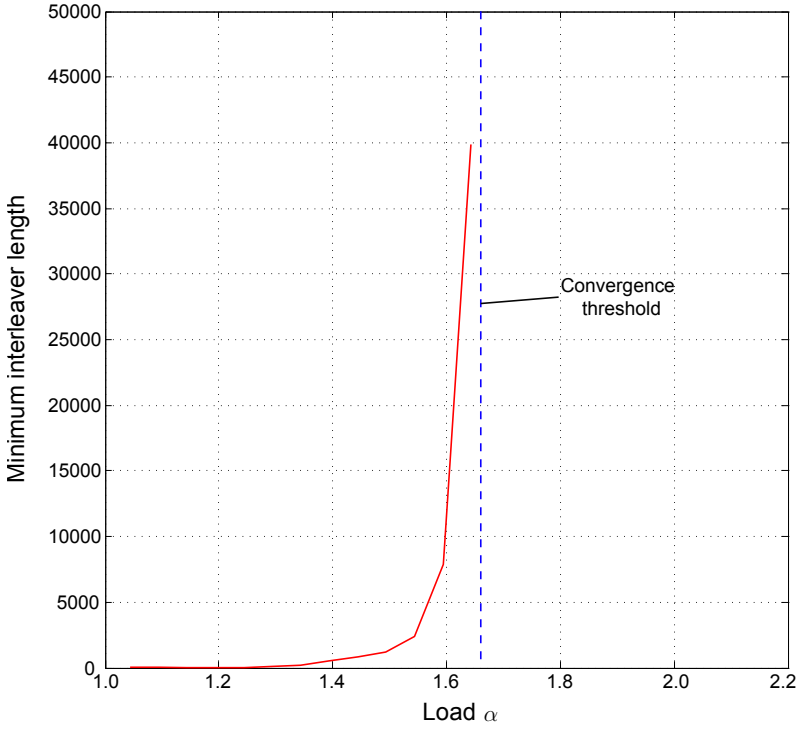

Figure 3.8: Approximate minimum interleaver length required to ensure that the simulation matches the theoretical predictions. Partition multiplicity $M = 10$, spreading gain $N = 20$, $E_b/N_0 = 10$ dB. Number of detection iterations performed $I = 30$.

generally can be no longer substituted with the random variable[1]

$$\mathcal{N}(\mu^i, \omega^i) = \mathcal{N}\left(\sum_{m' \neq m}^{M} s^i_{m'}; \sum_{m' \neq m}^{M} v^i_{m'}\right) \tag{3.26}$$

where $s^i_{m'} = \mathbb{E}[z^i_{k,l,m'}]$ is the mean of a single partition and $v^i_{m'} = \mathbb{E}[z^i_{k,l,m'} - s^i_{m'}]^2$ is its variance. This fact translates into signal-to-noise degradation in Equation (3.17) i.e. the increase of power (variance) of the noise that partitions are immersed in. Also the nature of this variance is changed due to the following observation. The distribution of LLRs in the first detection iteration is Gaussian, however as iterations progress the deviation from Gaussian occurs and is more severe with increasing system load, decreasing the interleaver size or both. Below a number of figures depicting the distribution of $z_{k,l,m}$ estimates for various

[1] Like it is assumed in (3.14).

scenarios are presented. For clarity of presentation, it is assumed that all-one codewords are transmitted.

Figures 3.9, 3.10 and 3.11 present the distributions of the local field

$$\sum_{m' \neq m}^{M} z_{k,l,m'} \tag{3.27}$$

for iterations, $I = 1, 3, 5, 7, 10, 20$, for various block sizes L and therefore interleaver lengths. As shown the distribution in the finite interleaving case (red curves) has tails that are more extended and the main lobe is shrunk as compared to the Gaussian case (black curves) and this is more pronounced for higher iterations. Both distributions, true Gaussian and the "new distribution" presented in each plot have the same means and variances. Note, that all the figures are in the logarithmic scale in order to show the tail behavior better. Since the shrinking is much more visible on the linear scale than on the logarithmic scale, Figure 3.13 is also plotted. Such similar distributions of the local field are observed in fluid dynamics theory and treated solely via simulation experiments. See for example [52, Figure 11] and [53]. Figure 3.12 shows how the distribution in iteration $I = 20$ is affected by increasing the block size L. As it is clearly seen the distribution approaches Gaussian as long as frame length is increased. Since local field diverges from Gaussian for $i > 0$, evolution of other parameters besides the mean and variance must be tracked in order to retrieve the actual distribution and predict the actual performance of the finite system. These are likely higher moments like skewness and kurtosis, which are the simplest measures of non-gaussianity. Such tracking makes not only the problem of finite interleaving complex, but also quite separated from the plethora of solutions and concepts developed in the communications and information theory, which heavily rely on the Gaussian assumption.

3.3.2 Complexity

The main advantage of the proposed partitioned modulation besides good performance is complexity. Since there are many definitions of complexity, here the focus on the complexity related to the number of basic operations that hardware has to execute to facilitate the detection. Looking at the algorithm, it is noticeable that the major complexity burden comes from evaluating the non-linear hyperbolic tangent function together with the division by the noise variance, which is estimated at the receiver constantly, see (3.11). However, this non-linear operation can be evaluated in hardware quite effectively using the approximation as a lookup table, which is proven to work fast and efficient[1]. All the other PS-CDMA receiver

[1] As shown in [51] via fixed point simulations the imperfect LLR estimation, due to non perfect division and $\tanh(\cdot)$ evaluation, has very little impact on the overall performance of the PS-CDMA detector. What matters is to perform cancellation and matched filtering precisely.

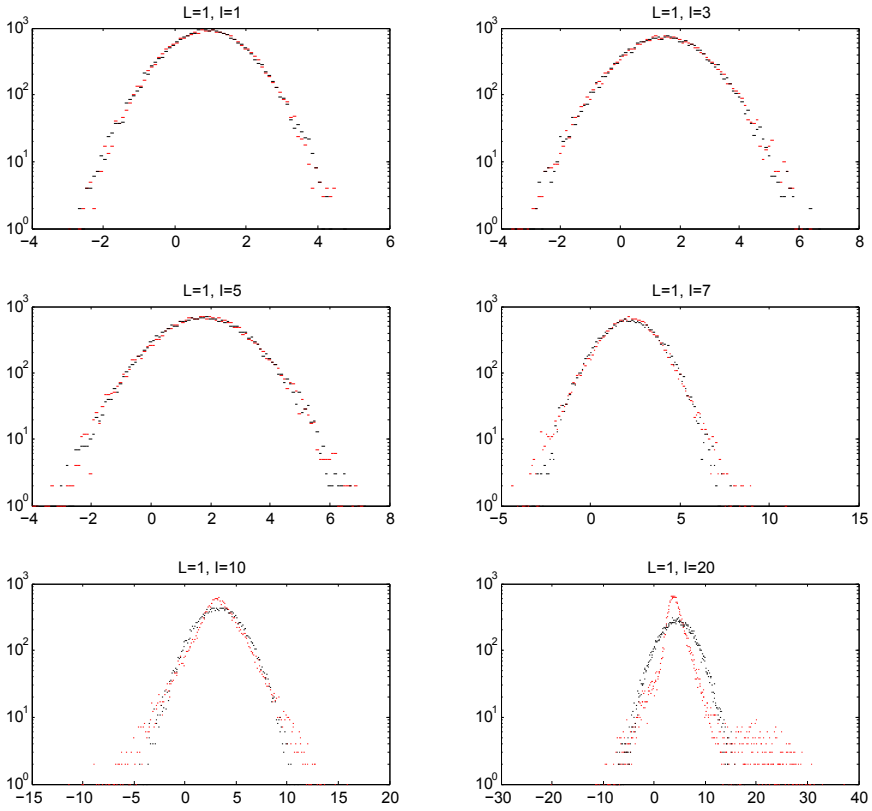

Figure 3.9: Logarithmic-scale distributions of the local field in iteration I. Partition multiplicity $M = 10$, spreading gain $N = 100$, number of users $K = 150$, $E_b/N_0 = 10$ dB. Frame length $L = 1$. Interleaver length $LM = 10$.

operations is the message exchange, very similar to those existing in LDPC decoders. After careful examination of the algorithm, the following complexity estimate can be given for the proposed detector[1]

Theorem 2. *The complexity of partitioned spreading CDMA system with K users, M partitions and processing gain N in terms of the number of multiplication/division arithmetic operations per iteration is*

$$O\left(K(4M+N)\right). \tag{3.28}$$

Therefore, the complexity of detecting bits of K users in I iterations shapes as $O\left(IK(4M+N)\right)$.

[1]Computational load associated with interleaving is omitted, since it is negligible compared to the remaining operations.

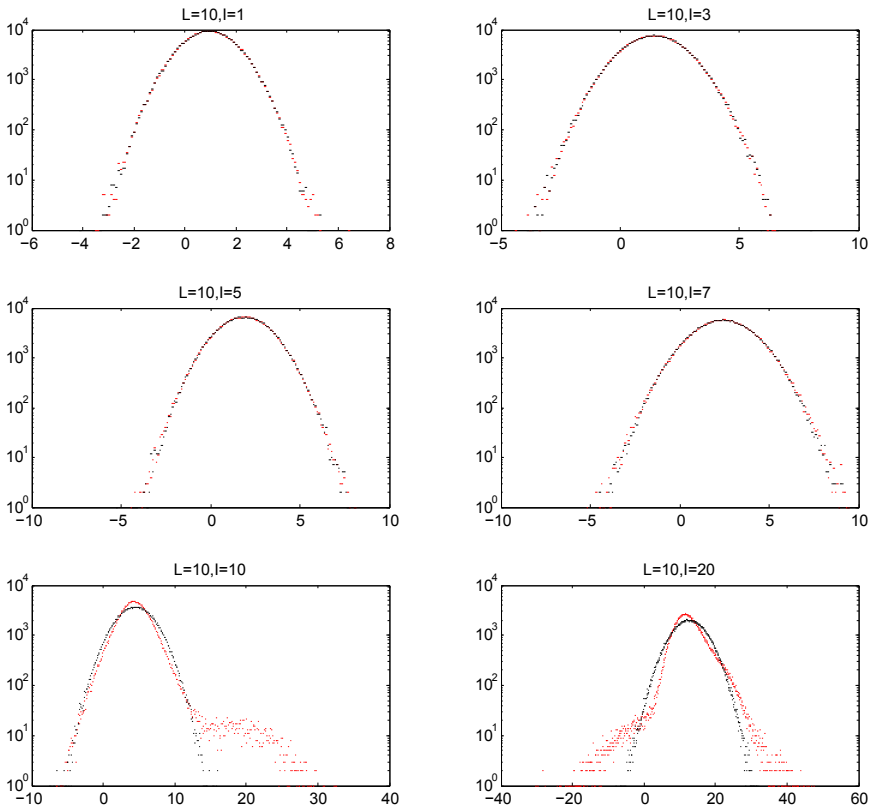

Figure 3.10: Logarithmic-scale distributions of the local field in iteration I. Partition multiplicity $M = 10$, spreading gain $N = 100$, number of users $K = 150$, $E_b/N_0 = 10$ dB. Frame length $L = 10$. Interleaver length $LM = 100$.

It is a linear growth with the number of users K in the system, whereas linear MMSE filtering associated with matched filtering followed by matrix inversion is cubic $O(KN + K^3)$. The optimal, maximum likelihood detection complexity is exponential $O(KN + a^K)$ and it also depends on the constellation size. The total amount of computations per bit per iteration is presented in Appendix A.

3.4 Coded PS-CDMA

Up to this point the attention was focused on the uncoded transmission, i.e. the case that no outer error control coding mechanisms were present and only inner repetition coding

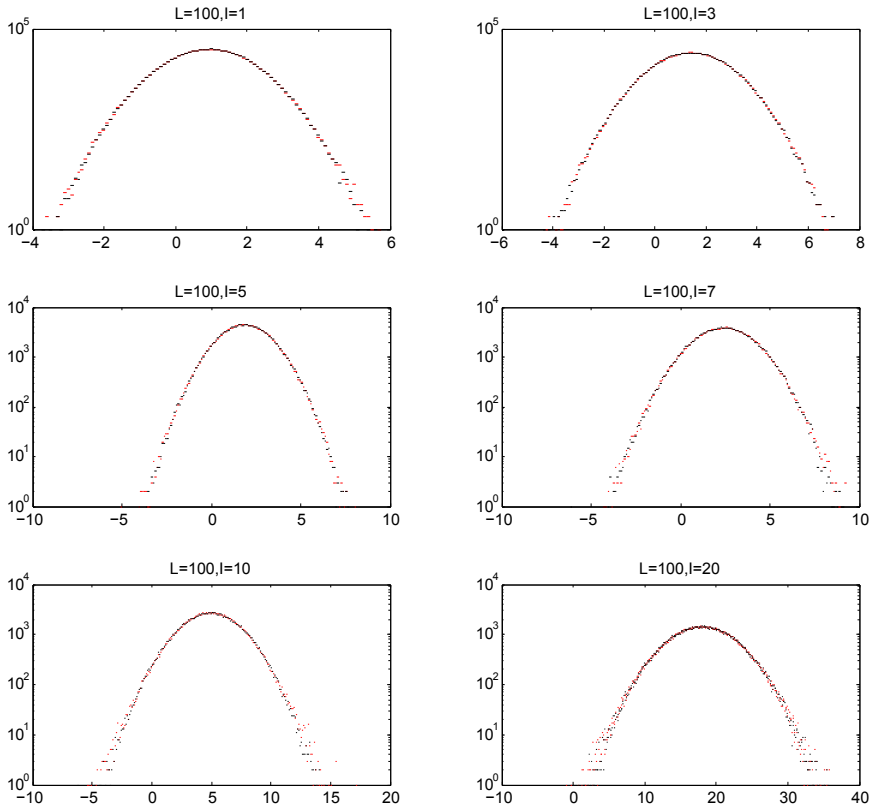

Figure 3.11: Logarithmic-scale distributions of the local field in iteration I. Partition multiplicity $M = 10$, spreading gain $N = 100$, number of users $K = 150$, $E_b/N_0 = 10$ dB. Frame length $L = 100$. Interleaver length $LM = 1000$.

was applied. In this section the situation when outer coding is utilized will be investigated. A number of processing schedules will be examined and compared in terms of attainable performance. Since it was already established that the inner repetition code is optimal for cancellation systems and powerful outer error control codes are best for fighting-off the additive noise such codes will be used subsequently and the attention is particularly focused on regular LDPC codes. They have good error correcting capabilities and very elegant analytical treatment allowing for more insights into the iterative detection/decoding process.

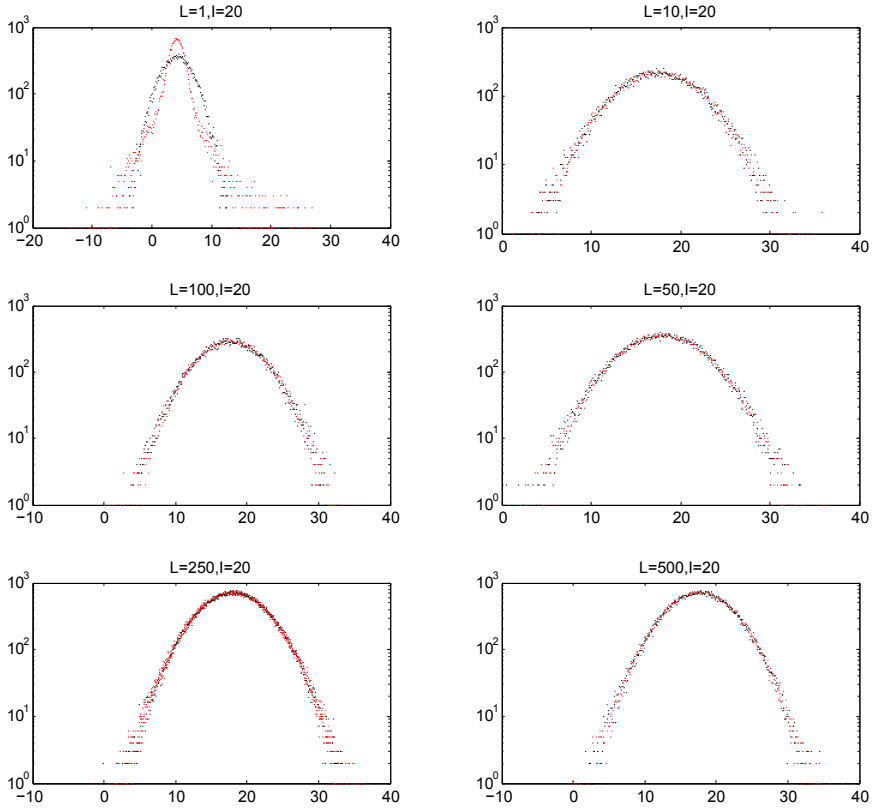

Figure 3.12: Logarithmic-scale distributions of the local field in iteration $I = 20$ for various frame lengths $L = 1, \ldots, 500$. Partition multiplicity $M = 10$, spreading gain $N = 100$, number of users $K = 150$, $E_b/N_0 = 10$ dB.

3.4.1 Processing Schedules

Since in coded PS-CDMA system three entities are interacting together, a number of processing schedules can be outlined. In the following, the attention is concentrated on two particular schemes. The first is the two-stage schedule. In this scheme the signal processing is already presented in Figure 3.3 and is as follows. First the iterations between the channel (MF, MMSE filter) and the inner repetition code are performed. It is just an operation of partitioned demodulator as introduced at the beginning of this chapter. A number of partitioned iterations are performed and if a given stopping criterion is satisfied (for example predefined maximum number of iterations reached or stabilization of the output variance) the signals are passed in the form of LLRs to the outer channel code decoder. This procedure

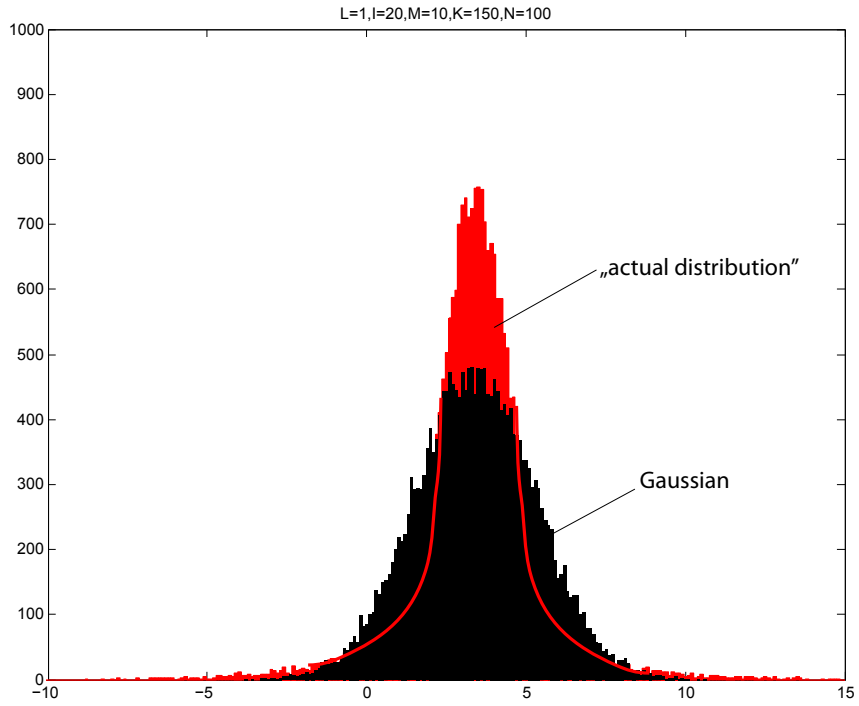

Figure 3.13: Linear-scale distribution of the local field in iteration $I = 20$. Partition multiplicity $M = 10$, spreading gain $N = 100$, number of users $K = 150$, $E_b/N_0 = 10$ dB. Frame length $L = 1$. Interleaver length $LM = 10$.

can be described formally as

1. For $i = 1, \ldots, I_{\mathrm{PS}}$ perform customary PS detector iterations according to

$$\sigma_{i+1}^2 = \sigma_n^2 + \alpha P g\left(\frac{(M-1)P}{M\sigma_i^2}\right), \tag{3.29}$$

   starting with $\sigma_0^2 = \sigma_n^2 + \alpha P$.

2. Return resulting (full APP[1]) LLRs

$$LLR_{k,l} = \frac{2\sqrt{P}}{\sigma_I^2} \sum_{m=1}^{M} z_{k,l,m}^I \tag{3.30}$$

   to the outer error control decoder.

[1]Factor $\frac{M-1}{M}$ disappears, since no message exclusion is performed.

Table 3.1: Table of thresholds [in dB] of LDPC/BPSK for a given rates R.

Rate R	d_v	d_c	τ_{LDPC}**[dB]**	τ_{BPSK}**[dB]**	$P/\sigma_\infty^2 = 2R\tau_{\mathbf{LDPC}}$	P/σ_∞^2 **[dB]**
0.25	3	4	1.05	-0.79	0.636	-1.965
0.40	3	5	0.9	-0.23	0.984	-0.070
0.50	3	6	1.15	0.19	1.303	1.149
0.66	3	9	1.9	1.02	2.044	3.1048
0.75	3	12	2.2	1.63	2.489	3.9602
0.80	3	15	2.55	2.04	2.878	4.5909
0.85	3	20	3.05	2.55	3.431	5.3542
0.90	3	30	3.75	3.2	4.268	6.3022

In order to find the maximum channel loading for the two-stage detector it is enough to find such maximum α that allows the partitioned demodulator to return LLRs with signal-to-noise-ratio equal to the outer LDPC code convergence threshold. Such thresholds can be evaluated analytically using the density evolution outlined before. The code thresholds for considered (3,x) family of regular LDPC codes is presented in Table 3.1. Also capacity for BPSK modulation on AWGN channel is indicated. An interesting observation can be made is that high rate LDPC codes from this family approach capacity limits much closer than low rate ensembles of codes and the gap is less than 0.5 dB. Of course these results are achieved for very long codes and for finite configurations numerical simulations need to be performed. Nevertheless, the limiting performance of such coding scheme using the threshold in the table can be indicated.

The second schedule is a generalization of the two-stage processing. It is deemed full-schedule and relies on returning the LLRs from the outer code decoder to the interference canceller after their conversion into soft-symbols using (1.12). Therefore its analysis is exactly the same as with the exception of the first iteration of the partitioned detector. Instead of using equation

$$\sigma_{i+1}^2 = \frac{1}{N}\sum_{k=1}^{K} P_k g\left(\frac{(M-1)P_k}{M\sigma_i^2}\right) + \sigma_n^2 \tag{3.31}$$

the following is applied

$$\sigma_{i+1}^2 = \frac{1}{N}\sum_{k=1}^{K} P_k g\left(\frac{m_v}{2}\right) + \sigma_n^2 \tag{3.32}$$

where m_v is the LLR received from the variable node of the LDPC code and is simply expressed as

$$m_v = d_v m_c \tag{3.33}$$

where m_c are final messages obtained from the check nodes. Then the iterations proceed as in (3.17) and LLRs from the detector are passed again to the outer LDPC code. The process

continues for a number of iterations until some stopping criterion is satisfied. As clearly seen, the full processing schedule is far more complex than the two-stage methodology. It invokes an outer channel code decoder multiple times. However, as it is the case, such joint processing results in a better performance. Figure 3.14 presents the maximum achievable

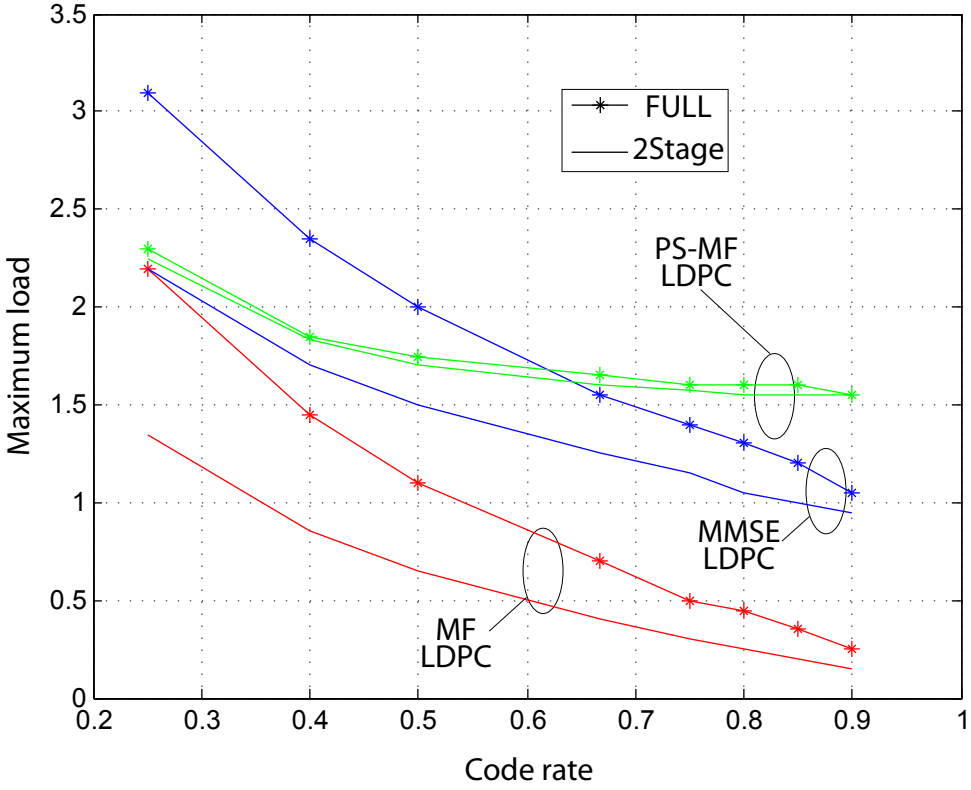

Figure 3.14: Maximum supportable system loads with partitioned CDMA, plotted as a function of rate R for regular (3,x) LDPC codes and $E_b/N_0 = 10$ dB. $M = 20$.

system loads for a number of detectors. A number of interesting observations can be derived from this plot. First, the partitioned CDMA detector with simple matched filters for the two-stage configuration always perform better than the much more complex MMSE filter requiring matrix inversion and it coincides with observations inferred from the analysis of Equation (3.23). However, the MMSE filter however results in better performance when full schedule is employed. However, in this method such MMSE filtering is computationally very expensive, since separate MMSE inversion per each user per each modulation symbol has to be performed. Such MMSE filtering, where a priori information is accepted for the case

of CDMA channels was proposed in [26]. However, for high rate outer coding schemes the difference in favor of the partitioned detector can be more than 50% and simple partitioned demodulation outperforms per-user MMSE filtering. Secondly, the advantage of the full decoding schedule in case of PS detector is negligible. Both schedules perform virtually the same. Note that the simple matched filter (traditionally used in today's CDMA systems) with coding performs really poor with performance noticeably away from other detectors considered.

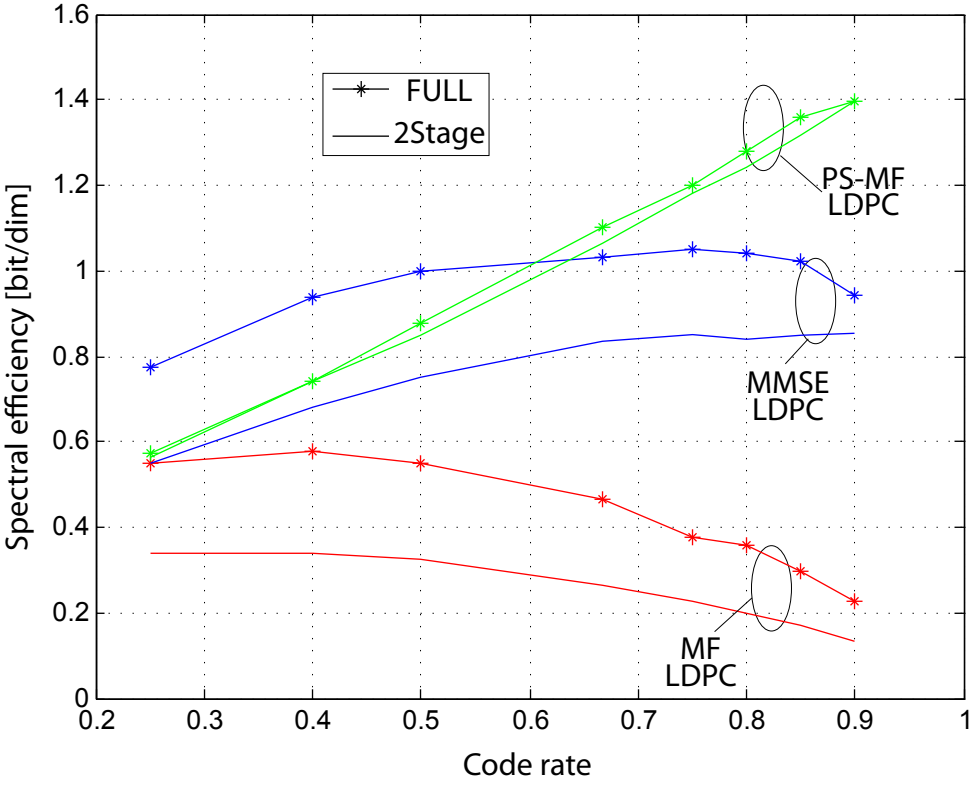

Figure 3.15: Spectral efficiency with partitioned CDMA, plotted as a function of rate R for regular (3,x) LDPC codes and $E_b/N_0 = 10$ dB. $M = 20$.

Another important conclusion can be drawn. Namely, although it may seem that larger loads can be supported when lower rate codes are used and naturally it would imply that use of such low rate codes is the best, but the reality is actually opposite. The most important performance measure of any communication system is the spectral efficiency. For the case

considered here the spectral efficiency is

$$C = \alpha R \qquad \text{[bits/dimension]} \qquad (3.34)$$

which is a product of achievable load (related to number of users/streams that the system can support) and information rate R of each stream, which is exactly the outer code rate. This is the straightforward derivative of the equation (2.52). Bearing that in mind, we plot Figure 3.15, which delivers quite the opposite message from the one by Figure 3.14. In fact it is most beneficial for partitioned system to use a high rate powerful outer code, since this maximizes the system's spectral efficiency. This message brings the attention to the previous chapter, when it was suggested that repetition code should be combined with powerful outer code to facilitate optimal performance.

3.5 Notes on Asynchronicity

To model asynchronous sampling of a user's chip waveforms the constant κ was previously used. Its value depends on one parameter, namely the chip waveform. Assuming that the desired user is sampled perfectly, asynchronous arrivals of remaining users will cause the interference that can be simply described using the κ constant. This value is less than unity and exactly equal to unity in the case of perfect synchronism between all the users. A brief explanation of calculating the value of this parameter is presented subsequently.

Assume that chips of all users are modulated using the same waveform $p(t)$. The receiver correlates the interferers signals shifted in time by τ[1] with this waveform. This process has the following mathematical description

$$\int_{-\infty}^{\infty} p(t)p(t-\tau)dt = c_{p,p}(\tau). \qquad (3.35)$$

Equation (3.35) is the definition of the autocorrelation function of the pulse $p(t)$ evaluated in point τ. Therefore finding the power of the interference boils down to sampling the squared autocorrelation function of the pulse shaping filter at specific points for a number of possible time shifts and subsequently averaging over these shifts. It is as follows

$$P_I = \frac{1}{T_c}\int_0^{T_c} c_{p,p}(\tau)^2 d\tau. \qquad (3.36)$$

This equation describes the energy from one pulse of the interfering user. In order to take

[1] The value of $\tau = 0$ or $\tau | T_c$ represents the chip synchronous situation.

into account the other $2L$ pulses[1] of the interfering users[2] the following straightforward expression has to be evaluated

$$P = \lim_{L\to\infty} \sum_{l=-L}^{L} \frac{1}{T_c} \int_0^{T_c} c_{p,p}(\tau - lT_c)^2 d\tau. \tag{3.37}$$

The results of the evaluation of (3.37) are gathered in Table 3.2 for a number of popular shaping pulses. It is clearly visible that the interference coming from other users is dimin-

Table 3.2: Table of coefficients κ for selected pulse shaping filters

Filter	Roll-off β	κ	κ [dB]
Sinc	0	1	0
Root RC	0.1	0.975	-0.11
	0.2	0.949	-0.23
	0.3	0.925	-0.34
	0.4	0.899	-0.46
	0.5	0.874	-0.58
	0.6	0.849	-0.71
	0.7	0.824	-0.84
	0.8	0.799	-0.97
	0.9	0.774	-1.11
Rect	-	2/3	-1.76

ished due to the lack of synchronicity and this reduction ranges from 1.75 dB for rectangular pulses to no loss (0 dB) when using sinc-type chip waveforms.

Summarizing, it is beneficial to have asynchronous arrivals of user signals since the interference power is reduced as compared to the synchronous situation. Naturally, it is assumed here that the desired user is sampled perfectly, which is possible due to sophisticated synchronization algorithms, like the one in [54].

3.6 Chapter Summary

This chapter describes the application of partitioned modulation for CDMA signalling. It is shown that this method significantly outperforms traditional signal processing and in a vast majority of cases is even better than complex MMSE filtering, while still having lower computational demand. The case when finite interleavers are employed was briefly outlined and precise complexity measures of the proposed algorithm were given. Additionally, the

[1] L preceding and L succeeding. It is assumed that non-causal shaping filters can be used.

[2] For the continuous transmission there is an infinite number of these pulses, i.e. $L \to \infty$, however their power vanishes in the limit and therefore their sum is finite.

concatenation of partitioned modulation with the outer error control code was investigated on the example of strong regular LDPC codes. Contrary to intuition, the use of the outer error correction mechanism does not give a substantial advantage and in many cases the system performance is virtually identical to the two–stage case where the outer error control code does not feed back its output messages to the interference canceller. This argument favors the use of two–stage processing rather than complex turbo processing, since it is often counterproductive.

Besides many attractive features, partitioned modulation has its own drawbacks. Namely this modulation is interference limited, meaning that beyond certain channel load threshold, the partitioned detector cannot deliver sufficient SNR, allowing for outer error control code mechanism to converge. The good news however is that this obstacle can be overcome with careful received power assignments. The case of removing this limitation by varying the users' received powers is the focus of the next chapter.

CHAPTER 4

Unequal Power Allocations for PS-CDMA

Until now all the derivations were performed with the assumption that received power levels were the same for all users/streams, which characterizes traditional power controlled system. However, as motivated by the arguments in [22] and by the fact that equal power allocations represent the worst case scenario for iterative interference cancellation receivers in this chapter the attention will be focused on relaxing this assumption and allowing received powers to differ. For the case of partitioned spreading modulation for CDMA the first observation that unequal power allocations can result in much better performance of the iterative receiver was made in [25]. It was shown there that allowing J groups of different power users will increase the maximum supportable load J-fold as compared to just only one group ($J = 1$), which of course denotes the equal power system. This increase is not obtained for free. It is possible because of the evident increase of the overall power present in the system. In other words, allowing unequal power allocations, power and spectral efficiency can be effectively traded and for certain cases, as will be seen later, the system is no longer interference limited. These observations coincide with the ones flowing from chapter 2, i.e. that systems with inner repetition codes can achieve very high spectral efficiencies and this as per information theory arguments, can only be possible in high signal-to-noise ratios regime. Figure 4.1 presents the maximum achievable loads for 2-stage schedule with 1,2 and 3 different power groups. The results in this figure are semi-analytical, and were obtained via numerical optimization. Nevertheless, the tendency and general message is well illustrated. It can be clearly observed that the maximum system load increases linearly with the increasing number of groups as

$$\alpha_{\max} = J\alpha_0 \tag{4.1}$$

where α_0 is the maximum load achieved for the 2-stage system in the equal power situation.

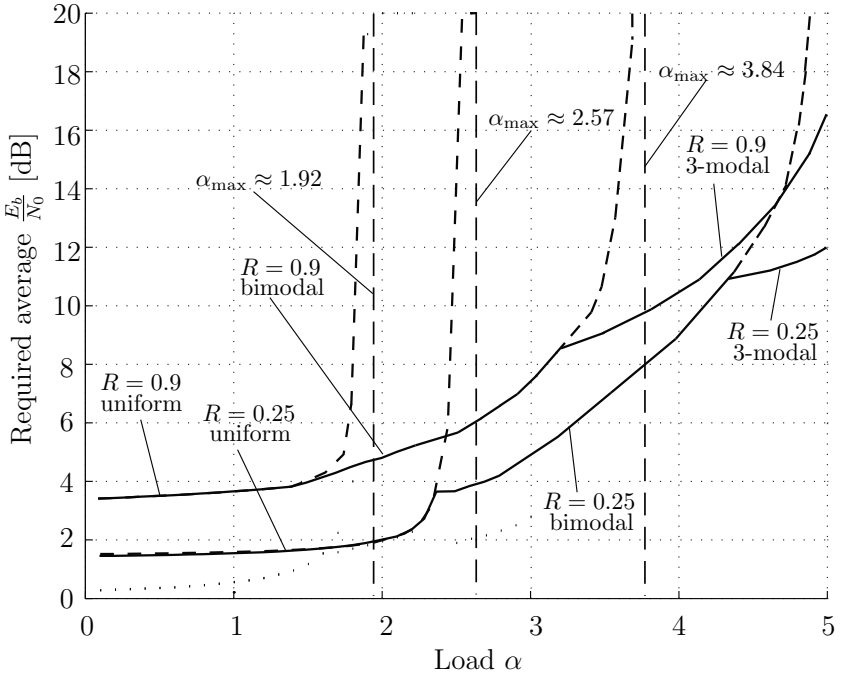

Figure 4.1: Average signal-to-noise ratio required to achieve given load α using two-stage schedule with unequal user power levels.

Moreover, in [55] and to some extent in [50] it was shown that the proposed two–stage partitioned detector can achieve the capacity of the linear multiple-access channel within 1 bit. However it is fulfilled when certain power profile is assured at the receiver. The derivation of this profile proceeds as follows.

Let us modify the variance transfer equation in (3.17) and instead of discrete analysis use continuous approach initially suggested in [56] where sums are replaced with integrals. Consider equation (3.17) with K users with powers $P_1 \leq P_2 \leq \ldots \leq P_K$. Denoting $P_k = P(k)$ and assuming $K, M \to \infty$, we may use the continuous approximation

$$\sigma_{i+1}^2 = \frac{1}{N} \int_0^K P(x) g\left(\frac{P(x)}{\sigma_i^2}\right) dx + \sigma_n^2 \quad i = 1, 2, \ldots \Leftrightarrow \tag{4.2}$$

$$\sigma_{i+1}^2 = \int_0^\alpha T(u) g\left(\frac{T(u)}{\sigma_i^2}\right) du + \sigma_n^2 \quad i = 1, 2, \ldots \tag{4.3}$$

where $T(u) = P(uN)$, $u \in [0, \alpha]$, is the power distribution and as defined a non–decreasing function.

In the case of two–stage decoding the system converges to error–free performance only if

the signal-to-noise ratio $T(0)/\sigma_\infty^2$ of the weakest user at the output of the detector is higher than the code's convergence threshold μ. Denote fixed point σ_∞^2 of the iterative equation in (4.3) as $\sigma_\infty^2 = v$. Then (4.3) becomes

$$v = \int_0^\alpha T(u) g\left(\frac{T(u)}{v}\right) du + \sigma_n^2. \tag{4.4}$$

Dividing by both sides by v we obtain

$$1 = \int_0^\alpha \frac{T(u)}{v} g\left(\frac{T(u)}{v}\right) du + \frac{\sigma_n^2}{v}. \tag{4.5}$$

The convergence condition using (4.3) and (4.5) can be rewritten as

$$\int_0^\alpha \frac{T(u)}{v} g\left(\frac{T(u)}{v}\right) du + \frac{\sigma_n^2}{v} < 1 \tag{4.6}$$

for all values of v

$$\frac{T(0)}{\mu} \le v \le \int_0^\alpha T(u) du + \sigma_n^2.$$

For the case of full decoding the soft–bit estimates (3.13) are produced by the error control code instead of simply combining partitions. Denoting the soft symbol variance corresponding to (3.15) by $g_{\text{ecc}}(\cdot)$ we can express the convergence condition for full decoding as

$$\int_0^\alpha \frac{T(u)}{v} g_{\text{ecc}}\left(\frac{T(u)}{v}\right) du + \frac{\sigma_n^2}{v} < 1 \tag{4.7}$$

for all

$$\frac{T(0)}{\mu} \le v \le \int_0^\alpha T(u) du + \sigma_n^2 \ .$$

Equations (4.6) and (4.7) determine the maximal loads α for a given average power to noise ratio and user power profile.

4.1 Achievable Loads with Two–Stage Decoding

In Figure 3.14 we argued that the performance of the two–stage decoding of partitioned CDMA can be virtually identical to the full decoding. In this section we present an analysis further supporting the potential of two–stage decoding.

Lemma 1. *Any system load α can be achieved with two–stage decoding and an appropriate*

choice of a power distribution $T(u)$ i.e. two-stage demodulation is not interference limited.

Proof. See [50] and B.1. □

Figure 4.2 presents the performance of the partitioned detector in an unequal power scenario and illustrates that power and spectral efficiency can be traded effectively with partitioned CDMA. The parameters used are $K = 80$, $N = 20$, $M = 20$, with rate $R = 0.9$ regular (3,30) LDPC outer code of length 2100 coded bits. Number of detector iterations is $I = 60$ and decoder was set to 30. The exponential power with $a = 1.5$ was used. The figure labels the average E_b/N_0, the BER performance is that of the weakest user. The received power distribution range was ca 25 dB following $T(u)$ in Lemma 1. It can be seen that the theory matches the simulation very precisely.

Partitioned CDMA with two-stage demodulation also has a very appealing property of dynamic user assignment. It says that the constants a (and σ^2) describing the required power distribution do not depend on the system loads α. Thus, for two different system loads $\alpha < \alpha'$, the corresponding power distributions $T(u)$ and $T'(u)$ coincide for $u < \alpha$, i.e.,

$$T'(u) = T(u) = e^{au}, \qquad u \in [0, \alpha] \ . \tag{4.8}$$

The implication of this results is that, in an operational system, new users can always be added without disturbing the existing population. Lemma 1 proves that all users can still be decoded, and (4.8) proves that previously active users do not need to adjust their power.

4.2 Approaching Channel Capacity with Two–Stage Decoding

Below are the steps showing that two-stage demodulation of partitioned CDMA can achieve the capacity of the GMAC channel. For details see Appendix B.

After having shown that arbitrary loads can be achieved with exponential power distributions $T(u) = e^{au}$ if $a \geq 2 \ln 2 \stackrel{\text{def}}{=} a_0$, we consider now partitioned signaling using the limiting distribution $T(u) = e^{a_0 u}$. We will assume that the individual error correcting codes (ECC) are capacity-achieving on the BIAWGN channel. Let us define

$$f(t) \stackrel{\text{def}}{=} \frac{1}{a_0} \int_t^{te^{a_0 \alpha}} g(x)dx + t\sigma_n^2 - 1 \ , \tag{4.9}$$

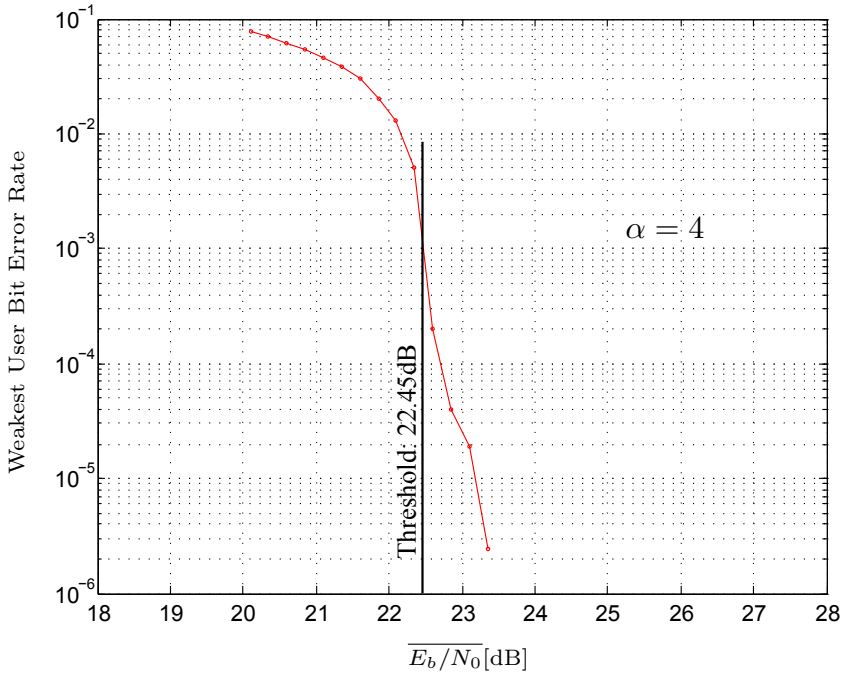

Figure 4.2: Bit error rate performance of the weakest user for a high-rate LDPC coded partitioned CDMA system at the high load of $\alpha = 4$.

and rewrite $f(t)$ as

$$\begin{aligned} f(t) &= \frac{1}{a_0}\int_t^{te^{a_0\alpha}} g(x)dx + t\sigma_n^2 - \frac{1}{a_0}\int_0^{\infty} g(x)dx \\ &= \frac{1}{a_0}\left(-\int_0^t g(x)dx - \int_{te^{a_0\alpha}}^{\infty} g(x)dx\right) + t\sigma_n^2 \\ &\leq -\frac{1}{a_0}\int_0^t g(x)dx + t\sigma_n^2 \stackrel{\text{def}}{=} \tilde{f}(t)\ . \end{aligned} \tag{4.10}$$

The following lemma assures a system with $a = a_0$ converges to a residual variance $\sigma_\infty^2 < 1/a_0$.

Lemma 2. *For $\sigma_n^2 = 1/a_0 - \epsilon$ where $\epsilon \geq 0.3$, the system converges to the residual noise variance $\sigma_\infty^2 \leq 1/a_0$.*

Proof. See [55]. □

The next lemma gives a lower bound on the spectral efficiency of the system.

Lemma 3. *Consider the power distribution* $T(u) = e^{a_0 u}$, $u \in [0, \alpha]$ *and* $\sigma_n^2 = 1/a_0 - \epsilon$. *Assume capacity approaching codes for each user. The parameter* ϵ *is chosen according to Lemma 2, such that* $\sigma_\infty^2 < 1/a_0$. *The spectral efficiency of the system is lower bounded by*

$$C_{\mathit{eff}} \geq \alpha - 0.1985 \tag{4.11}$$

for any $\alpha > 0$.

Proof. See [55] and B.2. □

The next lemma upper bounds the average signal-to-noise ratio of the system.

Lemma 4. *Consider the power distribution* $T(u) = e^{a_0 u}$, $u \in [0, \alpha]$ *and* $\sigma_n^2 = 1/a_0 - 0.3 = 0.4213$. *Assume capacity-achieving error control codes for every user.* E_b/N_0 *denote the average signal-to-noise of the system. The AWGN channel capacity which can be achieved for this* E_b/N_0 *satisfies*

$$C_{AWGN} \leq \alpha + 0.8 \tag{4.12}$$

for any $\alpha \geq 1$.

Proof. See [55] and B.3. □

Combining Lemmas 3 and 4 we obtain main result:

Theorem 3. *The spectral efficiency per dimension of the partitioned transmission under two–stage decoding satisfies*

$$C_{AWGN} - C_{\mathit{eff}} < 1 \qquad \textit{[bits/dimension]} \tag{4.13}$$

for any system load $\alpha \geq 1$.

Those findings are illustrated in Figure 4.3, where it can be seen that spectral efficiency achieved by the partitioned detector using two-stage demodulation is following the channel capacity for a broad range of signal-to-noise ratios. For low SNR values there exists a visible gap to AWGN capacity. This is due to fact that in evaluating the capacity of two–stage processor, a high rate outer LDPC code was used. This code has a theoretical convergence point at ca 3.75 dB. As we can see this is roughly the same as the cut-off SNR in Figure 4.3. Below this value, only a fraction of users (the stronger ones) converge to error-free performance, others inevitably deliver erroneous estimates of transmitted bits.

Figure 4.4 presents the achievable rates for each user in the population. We can see that the weakest user achieves the rate $R \approx 0.9$ whereas other users have higher rate, quickly

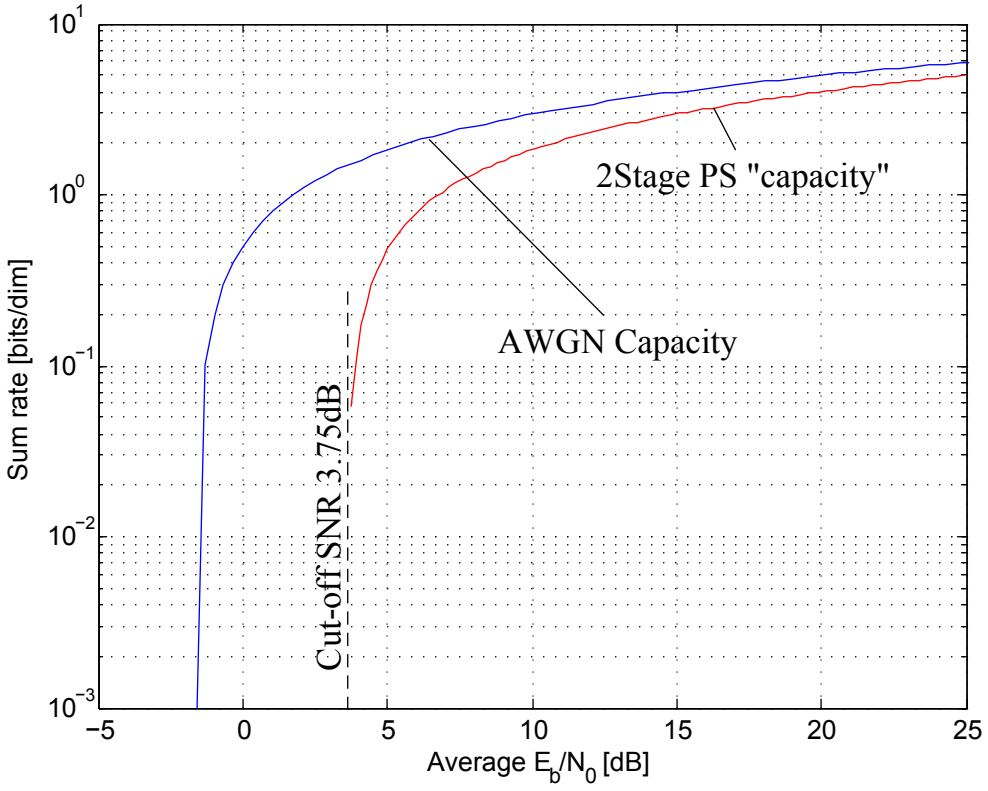

Figure 4.3: Spectral efficiency achieved by PS-CDMA two–stage detector using optimal power allocation and outer (3,30) LDPC code of rate $R = 0.9$.

achieving the maximum offered by the binary modulation employed. It is also important to note that modern binary error control coding systems of high rate, like high rate LDPC codes, operate very close to the capacity limits. This gap increases for low rate scenarios. This statement simply implies that the high rate codes are not a limiting factor in keeping the resulting performance of PS detector to be away from the capacity. It is sitting inherently in the detection process.

Figure 4.5, on the other hand, illustrates how the maximum achievable system load α varies with the choice of power exponent a. A clear asymptote at $a_0 = 1.3863$ is observed. Additionally the power dynamic range between strongest and weakest user in the pool as a function of system load is indicated (in red). Those power ranges shown are usual in wireless scenarios, but here they give the system advantage over, for example, linear receivers, which

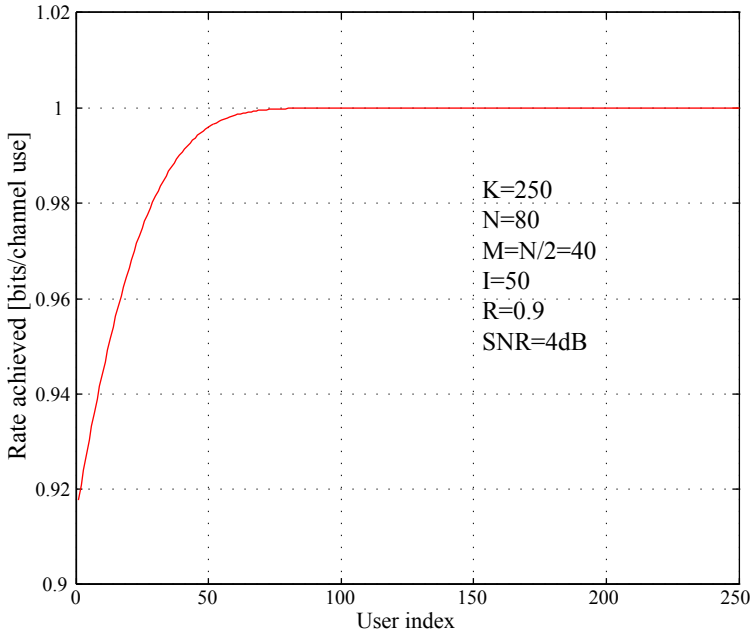

Figure 4.4: Example of rate distribution for unequal power scenario. $E_b/N_0 = 4$ dB for the weakest user is slightly above the code threshold i.e. 3.75 dB.

in such situation experiences the so called *near–far* problem.

4.3 Power Control Elimination

The need to assure a certain power distribution at the receiver imposes some form of power control, previously known from CDMA receivers working on the principle of simple correlation (matched filtering). In fact power control is a complex mechanism, which requires exchange of many control messages, which generally significantly contribute to the drop of information rate achieved as compared to the ideal, capacity achieving system.

As seen previously, users can be added to the system, without disturbing the already existing population. Whenever a user wants to join the pool of users and transmit to the base station it must after sending its signal with power level P_1 receive the connection confirmation from the BS. If it does not receive this message then it changes its power level and tries to get connected again. The process continues until the connection is finally established.

Figure 4.6 shows the performance of the system with the distribution as a function of the transmitter–referred signal–to–noise ratio[1]. The relationship between transmitted and

[1]As transmitted SNR we express the value $\bar{P}_{\mathrm{tr}}/2R\sigma^2$ in order to compare systems fairly based on the transmitted power.

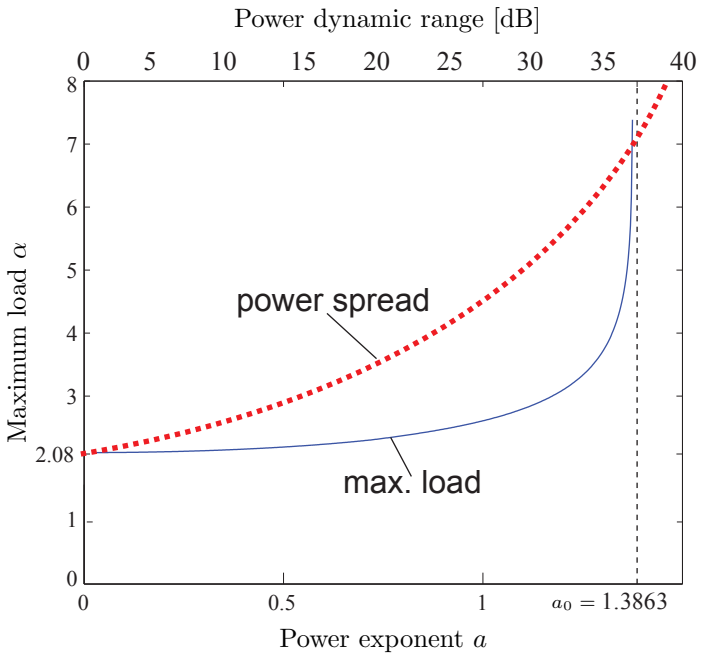

Figure 4.5: Maximum achievable loads α together with required power spread as a function of power exponent a. Power dynamic range for a given load is presented.

received power is assumed to follow the relationship

$$P_r = \frac{P_t}{(1+d)^{\beta}} \tag{4.14}$$

where d is distance between transmitter and receiver and β is path loss factor which is at least 2, for the free space case, or higher for other cases. We further assume that transmitters are uniformly spatially distributed. The performance behavior of equal power (traditional power controlled) CDMA using partitioned spreading is shown for comparison. In both cases it was assured that each user converges and the convergence threshold corresponds to the BPSK capacity. In the figure we used $\mu = 2.04$ dB which corresponds to the rate $R = 0.8$. It is clearly seen that the system with random received powers is behaving almost identically to the equal-power system for loads up to an $\alpha \approx 1.75$. However, past this value their performances diverge. An equal power system saturates around $\alpha \approx 2.08$, whereas uncontrolled power system can further increase the load until the saturation point at about $\alpha \approx 3.05$. It correspond to an almost 50% increase in the system capacity in favor of uncontrolled powers.

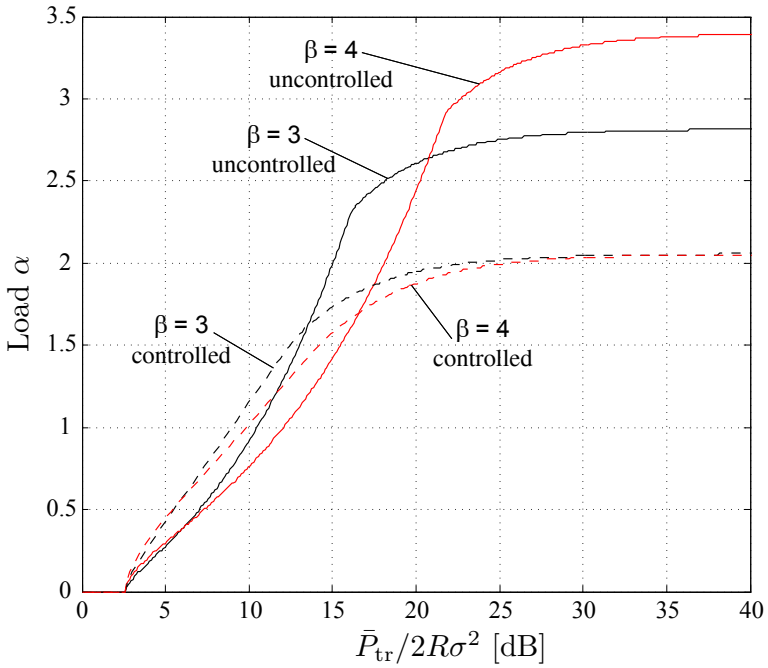

Figure 4.6: Achievable loads as a function of average transmitter–referred signal-to-noise ratio. Path loss factors $\beta = 3$ and $\beta = 4$.

4.4 Distribution of PS–CDMA Outputs

It is important to describe the statistics of partition estimates delivered by the partitioned CDMA front-end in unequal power scenarios[1]. Partition estimates are obtained via simple matched filtering of the detector input $\mathbf{y} = [y_1, y_2, \ldots, y_N]$, where N is the spreading gain. This composite signal is

$$y_n = \sum_{k=1}^{K} s_{k,n} d_k \sqrt{P_k} + b_n \tag{4.15}$$

where $b_n \sim \mathcal{N}(0, \sigma_n^2)$ is the channel noise. In subsequent derivations we assume, without loss of generality, that for every user $d_k = 1$. The filtered signal $z_{k,l,m}$, corresponding to one partition can be expressed as a sum of desired signal, interference and transformed channel noise

$$z_{k,l,m} = S_{k,l,m} + I_{k,l} + N_{k,l} \tag{4.16}$$

[1]For equal power scenarios it is known that asymptotic distribution tends to Gaussian, via Central Limit Theorem arguments, see [4].

Assume that we want to recover one partition. Then the matched filter estimate (normalized by $\sqrt{M}$) looks like

$$\begin{aligned} z_{k,1} &= \sum_{n=1}^{N/M} \sqrt{M} s_{k,n} y_n = \sum_{n=1}^{N/M} \sqrt{M} s_{k,n} d_k \sqrt{P_k} s_{k,n} + \\ &+ \sum_{n=1}^{N/M} \sqrt{M} s_{k,n} \sum_{k' \neq k}^{K} d_{k'} \sqrt{P_{k'}} s_{k',n} + \sum_{n=1}^{N/M} \sqrt{M} s_{k,n} b_n \end{aligned} \tag{4.17}$$

where for the simplicity we omit bit index l. From this we can see that each partition estimate $z_{k,m}$, $k = 1, \ldots, K$ and $m = 1, \ldots, M$ is a random variable, a summation of $K+1$ terms (random variables), where one is desired signal, $K-1$ is interference and one is noise. As the number of terms (users K) increases we can apply the Central Limit Theorem [57]. The number of partitions per symbol does not matter here, the only important detail is to allow a large number of users K. It appears to be true, since in regular CDMA ($M = 1$) we get Gaussian outputs from a matched filter by summation of terms belonging to other users. Let us look at the interference contained in each partition $z_{k,m}$. It is

$$I_{k,1} = \sum_{n=1}^{N/M} \sqrt{M} s_{k,n} \sum_{k' \neq k}^{K} d_{k'} \sqrt{P_{k'}} s_{k',n} = \sum_{k' \neq k}^{K} d_{k'} \sqrt{P_{k'}} \sqrt{M} \sum_{n=1}^{N/M} s_{k,n} s_{k',n} = \sum_{k' \neq k}^{K} \omega_{k'} \tag{4.18}$$

where $\omega_{k'}$ is a random variable that has mean and variance as (the expectations are taken over spreading sequences)

$$\begin{aligned} \mathrm{E}[\omega_{k'}] &= 0 \\ \mathrm{E}[\omega_{k'}^2] &= \frac{1}{N} P_{k'} = \sigma_{k'}^2 \end{aligned} \tag{4.19}$$

Now it is seen that interference is just a summation of $K-1$ random variables with bounded variances as in (4.19). If some necessary conditions are met [58] (*Lindeberg condition*) the summation of large number of independent random variables with bounded variance according to Central Limit Theorem [57, 58] tends to a Gaussian distribution with zero mean and variance

$$\theta_k^2 = \sum_{k' \neq k}^{K} \sigma_{k'}^2 = \frac{1}{N} \sum_{k' \neq k}^{K} P_{k'} \tag{4.20}$$

Also noise and the signal itself have bounded variances ($\frac{P_{k'}}{M}$ and σ^2 respectively), therefore we can conclude that the distribution of partition estimates $z_{k,l,m}$ is

$$p(z_{k,l,m}|d_k = 1) = \mathcal{N}\left(\sqrt{\frac{P_k}{M}}, \theta_k^2 + \sigma^2\right) \tag{4.21}$$

Having this established, we can now use the LLR derivation as well as soft-bit extraction as described in section 1.2.1. This derivation will be correct and analogous to the single user Gaussian channel, when the Lindeberg condition [58] is satisfied. This condition states that the individual powers $\frac{P_k}{M}$ are small compared to the their sum $\sum_{k'=1}^{K} \frac{P'_k}{M}$ i.e. contribute negligibly to the sum. Formally it is expressed as follows. Assume that the powers for each user k are ordered such that $P_1 < P_2 \ldots < P_K$. Then for any $k < K$ the ratio

$$\frac{P_k}{\sum_{k'=1}^{K} P_{k'}} \tag{4.22}$$

approaches zero. In other words, each user has a negligible impact on the system as a whole or that the individual powers P_k are small compared to their sum $\sum_{k=1}^{K} P_k$. Moreover since the powers are ordered it can be written that

$$\frac{P_k}{\sum_{k'=1}^{K} P_{k'}} \leq \frac{P_K}{\sum_{k'=1}^{K} P_{k'}}. \tag{4.23}$$

Therefore it is sufficient to show that the (4.22) vanishes for the case of the strongest user

$$\zeta(a, \alpha, N) = \frac{P_K}{\sum_{k=1}^{K} P_k} = \frac{P_0 e^{a\alpha}}{N \int_0^{\alpha} P_0 e^{ax} dx} = \frac{a e^{a\alpha}}{N(e^{a\alpha} - 1)} \tag{4.24}$$

Now it is only enough to show that $\zeta(a, \alpha, N) \to 0$ for any choice of parameters a, α, N. Even for moderate values of α and a we have:

$$\zeta(a, \alpha, N) = O(N^{-1}) \tag{4.25}$$

Interestingly, (4.25) is independent of the number of users in the system.

4.5 Chapter Summary

This chapter demonstrated that varying the received powers can be beneficial to the performance of the partitioned detector. It is of no surprise, since equal power allocations are the worst case for the iterative interference cancellation demodulators. Unequal power assignments can virtually remove the interference limited nature of the system. Optimal

power distribution is derived and has an exponential nature with control parameter a. It is shown that assigning the user pool according to this distribution results in system's spectral efficiency to be only constant away from the channel capacity. The most visible drawback of the partitioned two–stage demodulator is that in order to achieve high spectral efficiencies and deliver interference-free estimates, large number of iterations has to be performed. However, as shown in the previous chapter, each such iteration is not computationally expensive. But even if the received power assignment is random and typical uplink cellular environment is considered in many situations the partitioned detector exhibits performance advantages over the power controlled system.

CHAPTER 5

Signal Processing for Uplink LTE

Previous chapters discussed the application of interference cancellation for MIMO and CDMA transmission systems. The key signal processing parameters like transmission format, coding scheme, interleaving, modulation etc. were freely chosen to maximize the system's overall spectral efficiency. In this chapter parameters like channel coding and modulation, are chosen according to the LTE standard [42, 59], which is presently the object of intensive industrial as well as academic research worldwide.

To lay down the ground for specific receiver design solutions, a brief introduction to the most relevant aspects related to the receiver processing for the LTE standard is subsequently provided.

5.1 High-Level Overview

The development of the next generation cellular technology started with the introduction of broadband transmission techniques such as WiMAX which offer very high data rates. The UMTS technology upgrade has been dubbed LTE - Long Term Evolution. The aim is to enable much higher throughput values to be achieved along with considerably lower packet latency (a growing requirement for many mobile services), and enable cellular communications services to move forward to meet the needs of cellular systems technology in 2017 and beyond.

HSPA, a combination of HSDPA and HSUPA, and HSPA+ are now being deployed, the 3G LTE development is being named 3.99G as it is not meeting formally the requirements of the IMT-4G standard, although in reality there are many similarities with the cellular

technologies being touted as 4G. However, regardless of the terminology, it is certain that LTE will offer significant improvements in performance over the existing 3G standards.

Many operators have not yet upgraded their basic 3G networks, and LTE is seen as the next logical step for them as they will leapfrog from basic 3G straight to LTE. Such an approach will avoid providing several stages of upgrade. The use of LTE will also provide the data capabilities that will be required for many years and until the full launch of the full 4G standard, known as LTE Advanced [60].

The Third Generation Partnership Project, 3GPP, that oversaw the development of the UMTS 3G system started a campaign on the evolution of the 3G technology with a workshop that was held in late 2004. LTE core specifications were then included in Release 8, dated March 2009. It set down a number of high level requirements for 3G LTE such as:

- Increased service provisioning - more services at lower cost with better user experience
- Flexibility of use of existing and new frequency bands
- Simplified architecture
- Reasonable terminal power consumption

5.1.1 Requirements

The following key LTE specifications give an overall view of the performance that the system offers. It meets the requirements of industry for high data download speeds as well as reduced latency - a factor important for many applications from VoIP to gaming and interactive use of data. It also provides significant improvements in the use of the available spectrum. In terms of actual figures, targets for LTE included download rates of 100Mbps, and upload rates of 50Mbps for every 20MHz of spectrum. In addition to this, LTE is required to support at least 200 active users in every 5MHz cell. (i.e. 200 active phone calls). Targets have also been set for the latency in IP packet delivery. With the growing use of service including VoIP, gaming and many other applications, where latency is of primary concern a requirement of sub-10ms latency for small IP packets has been set. The detailed requirements for LTE are summarized in Table 1. For comparison the table also illustrates the performance of HSPA.

There are many aspects regarding the full characterization of the uplink transmission in the LTE like:

- Transmission organization and parameters
- Transmitter and receiver signal processing

Metric	LTE Requirement	HSPA+
Peak data rate (Single antenna config)	DL: 100Mbps UL: 50Mbps (20MHz spectrum)	DL: 42Mbps UL: 11Mbps
Mobility	Up to 350kmph, optimized for low speeds 0-15kmph	
Spectrum flexibility	1.4, 3.0, 5.0, 10.0, 15.0, 20.0 MHz	5.0, 10.0 MHz
Duplex	FDD and TDD	FDD only
Multi antenna configurations	DL: 4x4, 4x2, 2x2, 1x2, 1x1 (BSxMS) UL: 1x2, 1x1 (MSxBS)	DL: 2x2, 1x1 (BSxMS) UL: 1x2, 1x1 (MSxBS)
Control plane latency	<100ms	
User plane latency	<5ms	50ms
User capacity	200 users per cell (5 MHz spectrum)	
Coverage size	5-100km with slight degradation after 30km	

Figure 5.1: LTE high-level requirements.

- Uplink reference signals - the channel estimation procedures
- Physical channel structure - multiplexing of various control and data channels
- Capacity, coverage and link budget
- Random access and synchronization
- Power control

Here the attention is mainly concentrated on the first two aspects, transmission format and eNodeB signal processing since these experienced most modifications as compared to the previous standards and are main reason for the dramatic improvement of the end-user experience.

5.1.2 Evolution

Although there are major advances between LTE and its 3G predecessors, LTE is nevertheless looked upon as an evolution of the UMTS/3GPP 3G standards. However, despite

employing a different radio interface, using OFDMA/SC-FDMA instead of CDMA, there are many similarities with the earlier forms of 3G architecture and there is a much scope for re-use of existing architectures.

LTE can be seen as a further evolution of functionality, increased speeds and generally, improved performance. In addition to this, LTE is an all IP based network, supporting both IPv4 and IPv6. There is also no basic provision for voice, since this can be carried using VoIP packets.

5.1.3 Mobility and Latency

3GPP puts a strong emphasis on LTE's mobility aspects. Even though the system is designed to provide its peak performance in slow environment (up to 15 km/h) it is also expected to maintain connection and support voice and real time services for very high speeds up to 350 km/h or even 500 km/h.

5.1.4 Coverage

Apart form increasing data rates for LTE, 3GPP specifies the service area which considerably exceeds those existing in today's 3G standard like UMTS. The highest throughput and mobility will be supported in cell ranges up to 5 km. The maximum range where acceptable data rate degradation and overall quality of service reduction is allowed is 100 km.

5.1.5 Core Technology

LTE introduces a number of new technologies comparing to the previous cellular systems. They enable LTE to be able to operate more efficiently with respect to the use of the available bandwidth and power.

- OFDM (Orthogonal Frequency Division Multiplexing): this technology has been incorporated into LTE because it enables high data bandwidths to be transmitted efficiently.
- MIMO (Multiple-Input Multiple-Output): One of the main problems that previous telecommunications systems have encountered is that of multiple signals arising from the many reflections that are encountered. By using MIMO, these additional signal paths can be used to advantage and increase the throughput.

- SAE (System Architecture Evolution): With the very high data rate and low latency requirements for 3G LTE, it is necessary to evolve the system architecture to enable the improved performance to be achieved. One change is that the number of functions previously handled by the core network has been transferred out to the periphery. Essentially this provides a much "flatter" form of network architecture. In this way latency times can be reduced and data can be routed more directly to the destination.

5.2 Physical Layer Overview

5.2.1 OFDM

OFDM modulation has been used in practical systems for many years. Most important of them are the WiFi networks deployed worldwide. However, the use of OFDM in cellular systems was largely opposed in the previous standards like 3G, in favor of WCDMA technology. This will change with the introduction of LTE. It is the first cellular communications standard based on principles of OFDM, both on the uplink as well as the downlink. The major advantage of OFDM is its robustness to the effects of ISI, which are typically dominant in broadband communication systems. The equalization process necessary to recover the transmitted data streams is straightforward and usually boils down to a single-tap equalization. Mainly due to this particular property, OFDM is very attractive in conjunction with multiple antenna arrays, which are the second main breakthrough in modern wireless communications. The access schemes differ between the uplink and downlink: OFDMA is used in the downlink while SC-FDMA in the uplink. SC-FDMA is used in view of the fact that its peak to average power ratio is small and the more constant power enables high efficiency of the RF power amplifier in the mobile handsets - an important factor for battery-powered equipment.

5.2.2 MIMO

Until recently, communications consisted of a single transmit and a single receive antenna and with advances in communications theory as well as VLSI the theoretical limits of such single antenna configurations were approached very closely. That means that without increasing the bandwidth or/and transmit power, higher data rates could not be achieved. In order to further increase data rates a fresh look at the channel itself was necessary. It was delivered with the invention of MIMO technology, which alters the channel between a transmitter-receiver pair. When using MIMO, it is necessary to use multiple antennas to enable different propagation paths to be distinguished. It was shown theoretically as well

as practically that the use of multiple antennas both at the transmitter and at the receiver can significantly increase the data rates without increasing the bandwidth and the overall system transmit power. This increase scales linearly with the number of antennas employed. Shortly after laying down the theoretical foundations for MIMO channel analysis, practical experimental solutions were presented. Although those were in laboratory conditions, the large data rate increases predicted by the theoretical investigations were observed. In LTE 2-by-2, 4-by-2, or 4-by-4 antenna arrays will be used, at least in the first phase of deployment of the standard. While it is relatively easy to add further antennas to a base station, the same is not true for mobile handsets, where the dimensions of the user equipment limit the number of antennas, which should be placed at least a half wavelength apart in order to reduce the correlation between signals.

5.2.3 Dynamic Spectrum Allocation and Duplexing

Bandwidth can be allocated dynamically on both the downlink and the uplink depending on the service requirements. Six configurations are defined ranging from 1.4MHz up to 20MHz in both FDD and TDD mode. TDD has a strong advantage in the case where an asymmetry between the uplink and downlink data speeds exists.

5.2.4 Modulation Formats

LTE employs three quadrature modulation formats: QPSK, 16QAM and 64QAM. The choice of a particular scheme depends on prevailing conditions expressed in terms of SNR and the CQI. Only when there is a sufficient SNR, a higher order modulation formats is used, as discussed later.

5.2.5 Downlink

As mentioned above, the data rates anticipated on the downlink can reach 100Mbps in the first release of the standard. The use of MIMO configurations will further increase those numbers to roughly 300Mbps. To achieve such peak data rates clever and efficient designs of radio interfaces are necessary. On the downlink, OFDMA is employed, which is a multi-user version of OFDM, where certain users can use only a subset of the available subcarriers, the remaining subcarriers being used by other users.

5.2.6 Uplink

The uplink radio interface is based on SC-FDMA, sometimes called DFT-spread OFDM (DFT-S-OFDM). It is a modification of OFDM that aims at better power efficiency. It combines a low PAPR of a single carrier system with the multipath interference resistance and subcarrier frequency allocation of the OFDM systems. Another important advantage of SC-FDMA over DS-CDMA used in HSPA is that it achieves intra-cell orthogonality between users even on highly dispersive channels.

In SC-FDMA, multiple access is facilitated by inserting silent Fourier-coefficients (zero padding) on the transmitter side before the iDFT (inner, OFDM specific DFT), and removing them on the receiver side before the iDFT (outer SC-FDMA specific DFT). Different users are assigned to different Fourier-coefficients (sub-carriers). The difference between the size of both DFTs ($M \leq N$) indicates the maximum number of orthogonal users that can be served within a single radio link and allows for very flexible bandwidth usage depending on the instantaneous demands of the particular user. The size of the outer DFT used in SC-FDMA is restricted such that it can be constructed from multiplies of 2,3 and 5, i.e.

$$M = 2^j 3^k 5^l \tag{5.1}$$

where j, k, l are integers, enabling efficient, relatively low-complexity mixed-radix FFT implementations.

5.2.7 Modulation/Coding Schemes and Channel Reporting

In LTE, the base station can report the channel quality information to assist the mobile in choosing an appropriate MCS for transmission via the use of CQI control messages. The criterion is block error rate BLER $\leq 10^{-1}$. Table 5.1 presents the available choices as a function of the CQI index [61].

The adjustment to a specific MCS scheme, or equivalently CQI index, is performed via the rate matching operation, which relies on puncturing or repeating bits coming from the channel encoder [42].

5.2.8 Transmission Format

Although LTE supports both TDD as well as FDD, here stress will be put on symmetric alignment used in FDD, since this mode is more intuitive and therefore easier to follow. For more detailed information on TDD mode please refer to [61, 62]. Since SC-FDMA is based

Table 5.1: Available modulation and coding schemes in LTE.

CQI index	Modulation	Approx. code rate	Info bits per symbol
0	No transmission	-	-
1	QPSK	0.076	0.1523
2	QPSK	0.12	0.2344
3	QPSK	0.19	0.3770
4	QPSK	0.30	0.6016
5	QPSK	0.44	0.8770
6	QPSK	0.59	1.1758
7	16QAM	0.37	1.4766
8	16QAM	0.48	1.9141
9	16QAM	0.60	2.4063
10	64QAM	0.45	2.7305
11	64QAM	0.55	3.3223
12	64QAM	0.65	3.9023
13	64QAM	0.75	4.5234
14	64QAM	0.85	5.1152
15	64QAM	0.93	5.5547

on the same fundamental processing as OFDM the details presented below are the same for the uplink as well as the downlink transmission.

FDD Frame Structure

The transmission format hierarchy of LTE is depicted in Figure 5.2. In LTE a maximum of 1200 modulation symbols per OFDM time slot are allowed, assuming that the FFT size is maximum, i.e. $N = 2048$. The duration of a subframe (1ms) specifies the TTI. As seen in the Figure 5.2, an SC-FDMA symbol has a duration of 67μs[1], which is the inverse of the carrier spacing (15 kHz). Therefore there are a maximum of 16800 useful modulation symbols within one subframe of duration 1ms. These numbers can be extended to every possible configuration adopted in LTE and are presented in Table 5.2.

Table 5.2: LTE bandwidth/resource configuration

Channel bandwidth [MHz]	1.4	3	5	10	15	20
No. of resource blocks	6	15	25	50	75	100
No. of occupied subcarriers	72	180	300	600	900	1200
FFT size	128	256	512	1024	1536	2048
Sample rate [MHz]	1.92	3.84	7.68	15.36	23.04	30.72
Samples per slot	960	1920	3840	7680	11520	15360

[1]Assuming short (normal) cyclic prefix of duration $4.7\mu s$.

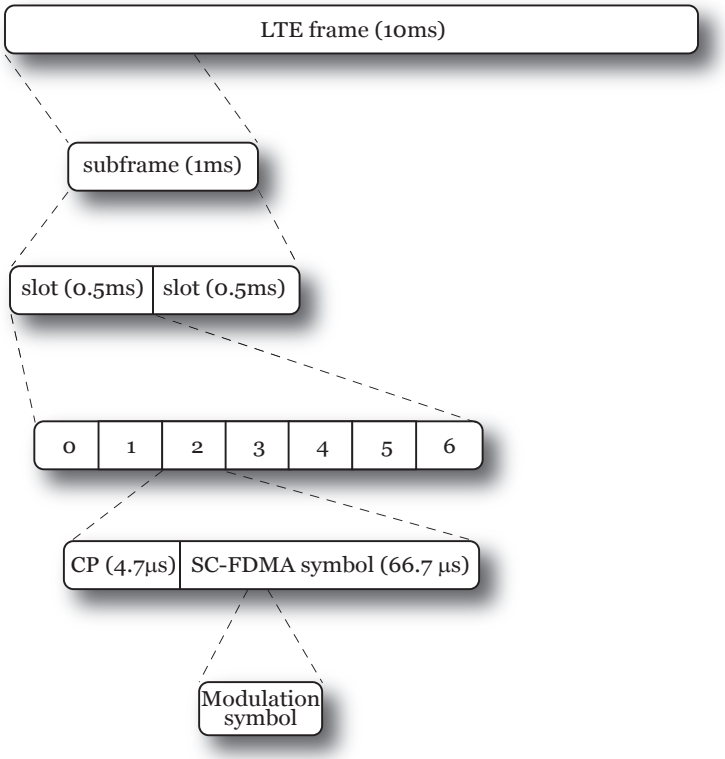

Figure 5.2: LTE uplink frame structure for short cyclic prefix length.

Mapping to Resource Blocks

LTE transmission is organized into resource blocks. One RB consists of $N_{\text{sc}}^{\text{RB}} = 12$ consecutive subcarriers, spaced 15 kHz apart and 6 or 7, depending on the length of cyclic prefix, OFDM symbols N_{symb}. The total number of available subcarriers depends on the overall transmission bandwidth of the system, which varies from 1.4 up to 20MHz. Depending on this bandwidth a different number, N_{RB}, of 180kHz resource blocks can be used (6 up to 100). The smallest unit defined in LTE is the resource element, which consists of one subcarrier of duration of one OFDM symbol, i.e. $71.4\mu s$. This implies that each RE can be viewed as the equivalent of a modulation symbol in single-carrier transmission. Resource configurations are presented in Table 5.2. Using this information, Table 5.1, and Figure 5.3, we can calculate the maximum data rate for a single-layer LTE configuration as

$$\begin{aligned} R &= N_{\text{RB,max}} N_{\text{sc}}^{\text{RB}} N_{\text{symb}} \text{CQI}(15) \frac{\text{number of 0.5ms slots}}{\text{second}} \\ &= 100 \cdot 12 \cdot 7 \cdot 5.5547 \cdot 2000 = 93.3\text{Mbit/s} \end{aligned} \tag{5.2}$$

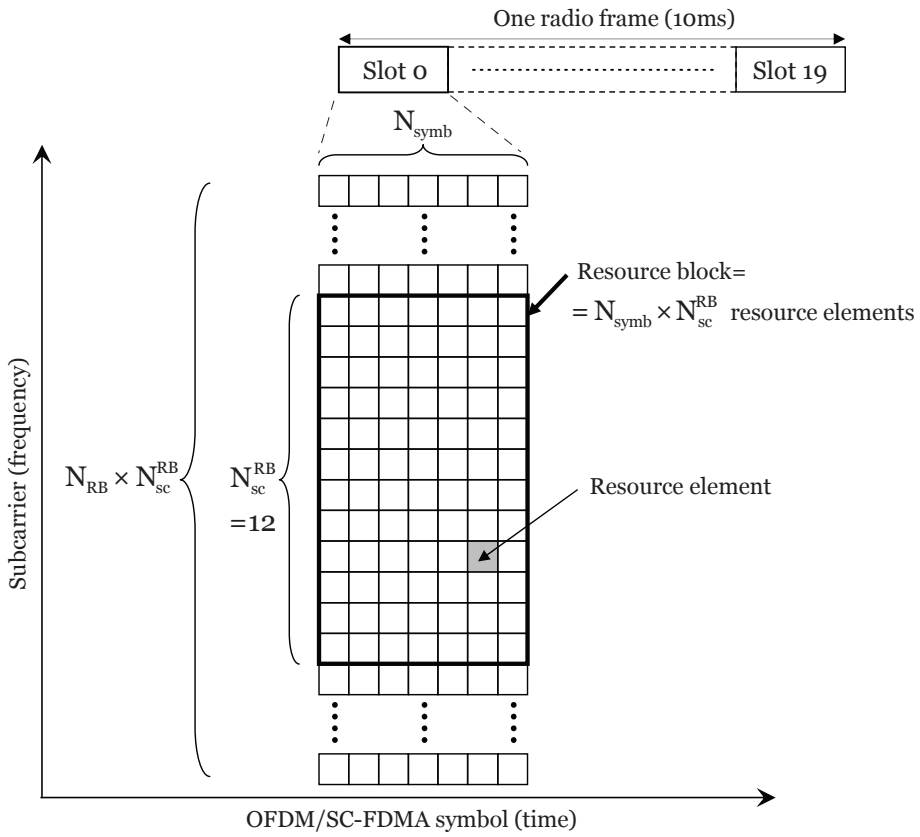

Figure 5.3: Resource block allocation in LTE.

With a dual layer configuration, like MIMO 2-by-2, that rate doubles to 186.6Mbit/s.

Example of Resource Block Allocation for 5MHz System Bandwidth

Usually N is larger than the maximum number of occupied subcarriers, providing for efficient oversampling and "sinc" pulse-shaping, i.e. not all available bandwidth is occupied by resource blocks as depicted in Figure 5.4. In this 5MHz bandwidth configuration an FFT of size 512 (300 occupied subcarriers) is used as presented in Table 5.2.

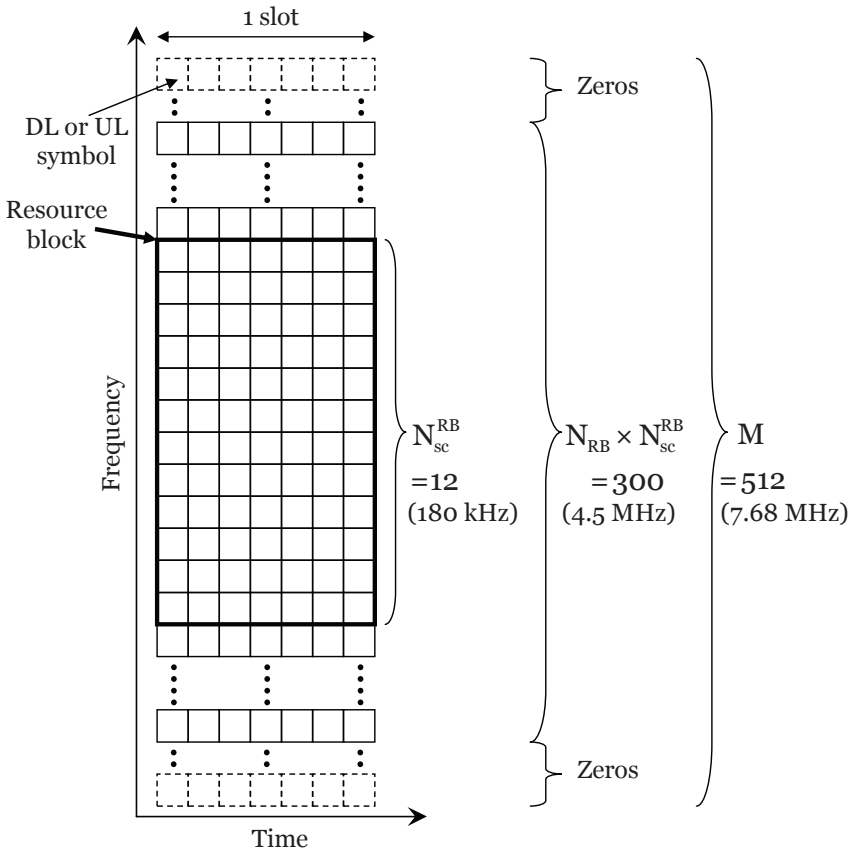

Figure 5.4: Example of resource block allocation for 5MHz bandwidth.

Resource Block Alignment

Figure 5.5 presents the relationship between resource blocks, OFDM symbols and subcarriers in time.

5.3 LTE Channel Model

Since the generation of SC-FDMA signals is almost identical to that performed in the case of OFDM, the attention is first concentrated on how the OFDM, with CP, transforms the multipath channel into a set of orthogonal flat channels. Applying subsequently outer DFT, the proper SC-FDMA stream is obtained.

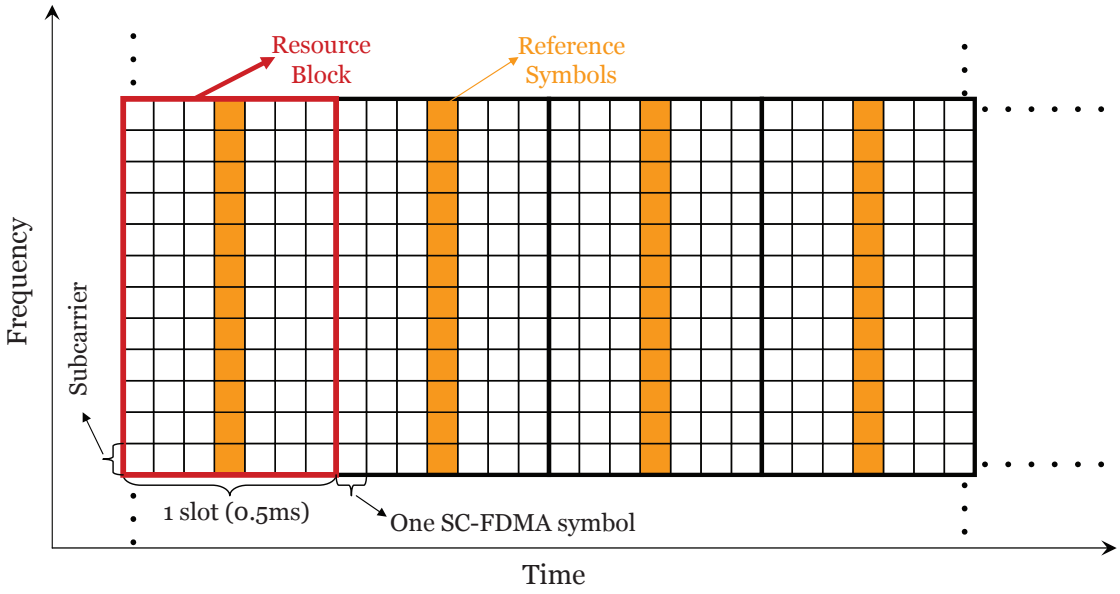

Figure 5.5: Allocation of resource blocks in time.

5.3.1 OFDM with Cyclic Prefix Channel

As already established, OFDM effectively mitigates multipath interference via the use of a particularly simple single-tap equalization in the frequency domain. Below the derivation of this idea is explained.

Assuming that the channel impulse response of length L is

$$\mathbf{h} = [h_0, h_1, \ldots, h_{L-1}] \tag{5.3}$$

and data of length N that has to be transmitted is

$$\mathbf{d} = [d_0, d_1, \ldots, d_{N-1}]^{\mathrm{T}}. \tag{5.4}$$

First, inverse DFT of size N on data vector is performed resulting in

$$\mathbf{x} = \mathbf{F}^{\mathrm{H}}\mathbf{d} = [x_0, x_1, \ldots, x_{N-1}]^{\mathrm{T}}. \tag{5.5}$$

where $\mathbf{F}$ is the N size DFT matrix[1] with entries $\mathbf{F}_{m,n} = e^{-j2\pi mn/N}$. Then cyclic prefix of

[1] It is a unitary matrix, i.e. $\mathbf{F}^{-1} = \mathbf{F}^{\mathrm{H}}$.

minimum length is added[1]

$$\tilde{\mathbf{x}} = [x_{N-(L-1)}, x_{N-(L-2)}, \ldots, x_{N-2}, x_{N-1}, x_0, x_1, \ldots, x_{N-1}]^{\mathrm{T}}. \tag{5.6}$$

Signal $\tilde{\mathbf{x}}$ is transmitted over a multipath channel which acts as a linear filter, i.e. linear convolution is performed. Therefore the output of the channel is

$$y_m = \sum_{l=0}^{L-1} h_l x_{m-l} + n_m \tag{5.7}$$

which in matrix form can be expressed as

$$\begin{bmatrix} y_{-(L-1)} \\ y_{-(L-2)} \\ \vdots \\ y_0 \\ y_1 \\ \vdots \\ y_{N-2} \\ y_{N-1} \end{bmatrix} = \begin{bmatrix} h_0 & 0 & \ldots & \ldots & \ldots & \ldots & 0 \\ h_1 & h_0 & \ldots & \ldots & \ldots & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & \ldots & h_{L-1} & \ldots & h_1 & h_0 \end{bmatrix} \begin{bmatrix} x_{N-(L-1)} \\ \vdots \\ x_{N-1} \\ x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} + \begin{bmatrix} n_{-(L-1)} \\ n_{-(L-2)} \\ \vdots \\ n_0 \\ n_1 \\ \vdots \\ n_{N-2} \\ n_{N-1} \end{bmatrix}. \tag{5.8}$$

After removing the cyclic prefix from (5.8), which is equivalent to removing first $\mu = L - 1$ rows of the channel matrix, we obtain

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{N-2} \\ y_{N-1} \end{bmatrix} = \begin{bmatrix} h_{L-1} & \ldots & h_1 & h_0 & 0 & \ldots & 0 \\ 0 & h_{L-1} & \ldots & h_1 & h_0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & \ldots & h_{L-1} & \ldots & h_1 & h_0 \end{bmatrix} \begin{bmatrix} x_{N-(L-1)} \\ \vdots \\ x_{N-1} \\ x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} + \begin{bmatrix} n_0 \\ n_1 \\ \vdots \\ n_{N-2} \\ n_{N-1} \end{bmatrix}. \tag{5.9}$$

[1]The minimum length μ of cyclic prefix is $L - 1$ for channels of length L. In practice however longer CP is used to always assure that $\mu > L - 1$. Increasing the length of CP results in increasing the rate loss of OFDM systems as compared to single carrier systems.

However, equation (5.9) is equivalent to

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{N-2} \\ y_{N-1} \end{bmatrix} = \begin{bmatrix} h_0 & 0 & \dots & h_{L-1} & \dots & h_1 \\ h_1 & h_0 & 0 & \dots & h_{L-1} & \dots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & h_{L-1} & \dots & h_1 & h_0 \end{bmatrix} \begin{bmatrix} x_0 \\ x_1 \\ \vdots \\ x_{N-1} \end{bmatrix} + \begin{bmatrix} n_0 \\ n_1 \\ \vdots \\ n_{N-2} \\ n_{N-1} \end{bmatrix} \tag{5.10}$$

where the channel matrix $\tilde{\mathbf{H}}$ is a circulant matrix. Writing (5.10) in a matrix form we obtain

$$\mathbf{y} = \tilde{\mathbf{H}}\mathbf{x} + \mathbf{n} = \tilde{\mathbf{H}}\mathbf{F}\dot{+}\mathbf{n}. \tag{5.11}$$

This signal is subsequently used as input to the DFT block

$$\mathbf{F}\mathbf{y} = \mathbf{F}\tilde{\mathbf{H}}\mathbf{F}^{\mathrm{H}}\mathbf{d} + \mathbf{F}\mathbf{n}. \tag{5.12}$$

Using the fact that the Fourier matrix diagonalizes any circulant matrix [63], we can write (5.12) as

$$\mathbf{Y} = \mathbf{H}\mathbf{d} + \mathbf{F}\mathbf{n} \tag{5.13}$$

and matrix $\mathbf{H}$ is diagonal with the DFT of the first column of $\tilde{\mathbf{H}}$ on its diagonal. $\mathbf{Y}$ is a vector and it is the Fourier transform of time vector $\mathbf{y}$. Gaussian noise $\mathbf{n}$ is transformed from time to the frequency domain, but this transformation[1] does not change its statistical properties. In this way multipath channel was transformed into a flat channel due to the use of DFT/IDFT and the application of the cyclic prefix.

OFDM with Multiple Antennas

Previously, the OFDM channel for SISO antenna configuration was derived. Here the guidelines on how to treat multipath MIMO channel by the means of OFDM with CP are given. Recalling the OFDM channel for single antenna systems, i.e. equation (5.12) the MIMO channel with N_t transmit and N_r receive antennas can be expressed as

$$\begin{bmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_{N_r} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{1,1} & \mathbf{H}_{1,2} & \dots & \mathbf{H}_{1,N_t} \\ \mathbf{H}_{2,1} & \mathbf{H}_{2,2} & \dots & \mathbf{H}_{2,N_t} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{H}_{N_r,1} & \mathbf{H}_{N_r,2} & \dots & \mathbf{H}_{N_r,N_t} \end{bmatrix} \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_{N_t} \end{bmatrix} + \begin{bmatrix} \mathbf{N}_1 \\ \mathbf{N}_2 \\ \vdots \\ \mathbf{N}_{N_r} \end{bmatrix} \tag{5.14}$$

[1] DFT is orthonormal transformation and Gaussian signals are invariant to the use of DFT/IDFT.

where each entry $\mathbf{H}_{n_r,n_t}$ is diagonal matrix of frequency gains between receive antenna n_r and transmit antenna n_t and is exactly like matrix $\mathbf{H}$ in equation (5.13). It can be seen that OFDM is particularly simple and suitable for the case of multiple antennas. Such a baseband channel model as in (5.14) applies to the downlink LTE transmission where OFDMA is used. However, for the case of uplink, where SC-FDMA is employed the situation is slightly different.

5.3.2 SC-FDMA Signal Generation

In order to generate SC-FDMA signals, the modulation symbols are spread using a size M DFT. Then zeros are inserted so the DFT outputs match the size of N subcarrier OFDM modulator. The positions of the zeros determine to which subcarriers the DFT-precoded data are mapped. Usually N is larger than the number of subcarriers used, M, thus allowing for efficient oversampling and pulse shaping [61]. For single antenna channels the SC-FDMA can be described similarly like it was in the case of OFDM in (5.12). Received signal after N size IDFT is

$$\mathbf{Y} = \mathbf{HZF}_{\mathrm{M}}\mathbf{d} + \mathbf{F}_{\mathrm{N}}\mathbf{n}. \tag{5.15}$$

Matrix $\mathbf{F}_{\mathrm{M}}$ is a Fourier matrix of size M while $\mathbf{Z}$ is the $N \times M$ zero padding matrix adding silent Fourier coefficients to the output of size M DFT. So $M + Z = N$. Two zero insertion mechanism presented in Figure 5.6 are foreseen for SC-FDMA in LTE:

- Localized mapping. In this scheme M adjacent subcarriers are used and zeros are inserted in one or two consecutive blocks.
- Distributed mapping. Here the subcarrier mapping allocates M equally spaced subcarriers, (every K-th subcarrier) and $K - 1$ zeros are inserted between the M DFT outputs. Also additional zeros are appended to either size of DFT output before the N-IDFT, because $MK < N$.

This subcarrier matching can be used to make users orthogonal by assigning each of them a different subset of all available subcarriers. In practice localized mapping is preferred, simplifying the transmission scheme where the same RB structure, as in downlink, can be utilized. Distributed allocation is only used for SRS enabling eNodeB to perform uplink frequency-selective scheduling [61].

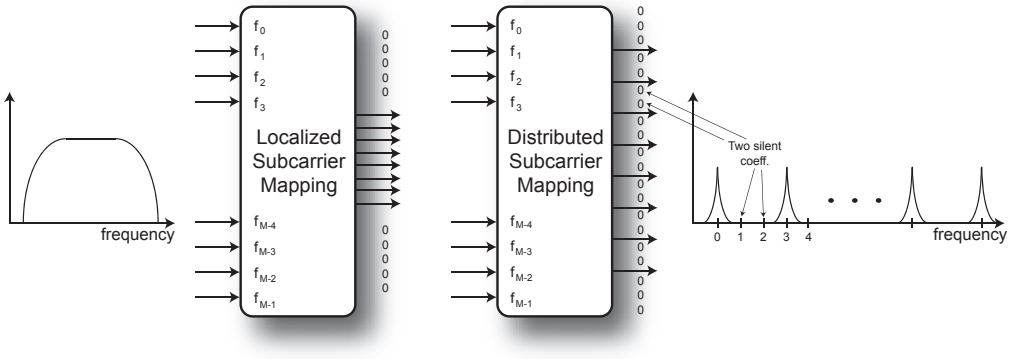

Figure 5.6: Subcarrier mapping in SC-FDMA as in [61], $K = 3$.

5.4 Signal Processing for UL-LTE

In this section the transmit signal processing chain and various receiver processing schemes are outlined. We start with the SC-FDMA signal generation for LTE. Then the channel model is briefly outlined. Subsequently, two receiver schemes are discussed and compared. First is the baseline MMSE frequency-domain filtering followed by decoding of the channel code. Second is group SIC/PIC MMSE turbo cancellation. It is shown by using the semi-analytical method based on variance transfer charts and numerical simulations, that the latter outperforms the baseline processor when the users' signals arrive at the base station with different power levels. In other words, in this development we concentrate on improving uplink transmission quality for users located on the cell edges. Such a situation is presented in Figure 5.7. Two users are accessing the base station. Their signals arriving at the BS are of different strength due to their different geographical and electrical locations. The problem that arises here, namely that weak user cannot be decoded due to the presence of strong closely-located user arises in modern cellular systems and is very vital in high performances systems, like LTE. The nature of this problem is very similar to *near-far* problem present in networks based on CDMA.

5.4.1 Transmitter

The frequency domain generation of the SC-FDMA signal is shown in Figure 5.8. Note that the figure depicts a single-stream two-transmit antenna scenario. Naturally multiple independent stream transmission (multi codeword) is also possible, however current LTE

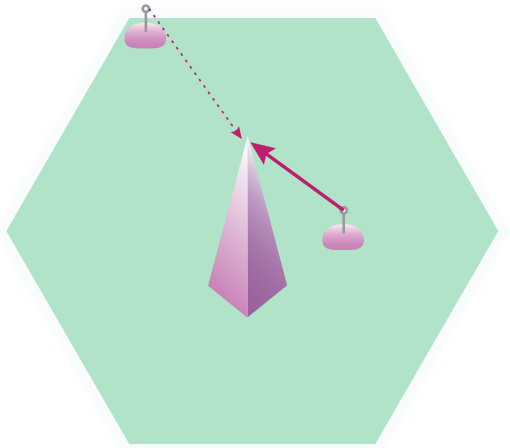

Figure 5.7: Cell-edge problem. Two-user case.

releases implement transmission of up to two codewords only. Zero padding is also called subcarrier mapping and it ascertains that the size of the block after the M-point DFT is in alignment with the inner iFFT of size N as discussed before.

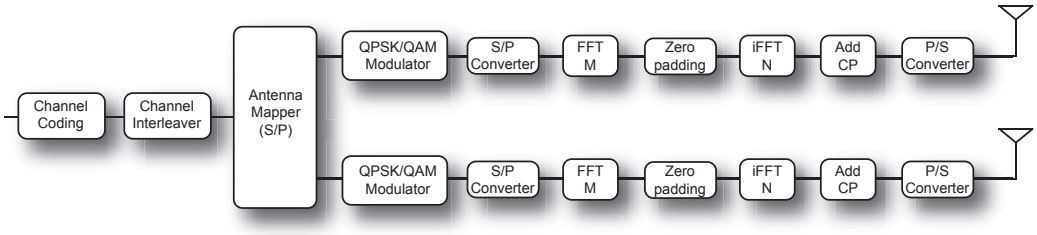

Figure 5.8: SC-FDMA dual-stream MIMO transmitter.

5.4.2 Channel Model

LTE uplink communication utilizes SC-FDMA, where a block of M time-domain symbols is transformed via an M-point DFT into the frequency domain and transmitted as M out of N OFDM tones [64]. The resulting channel can be modeled as follows

$$\mathbf{Y} = \mathbf{HAX} + \mathbf{N} = \mathbf{HAFx} + \mathbf{N}, \tag{5.16}$$

where $\mathbf{H}$ is a matrix composed of diagonal matrices in a form solely dependent on the MIMO configuration. For instance for a 2×2 MIMO case

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_{1,1} & \mathbf{H}_{1,2} \\ \mathbf{H}_{2,1} & \mathbf{H}_{2,2} \end{bmatrix}, \tag{5.17}$$

where each matrix $\mathbf{H}_{r,t}$ is a diagonal matrix of size M frequency gains between the r-th receive and t-th transmit antenna. The matrix $\mathbf{A}$ is a diagonal matrix of stream (user) amplitudes

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix}, \tag{5.18}$$

implying that $\mathbf{W} = \mathbf{A}\mathbf{A}^{\mathrm{H}}$ is the power matrix. $\mathbf{F}$ is composed of N_t size-M Fourier transform matrices

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_{\mathrm{M}} & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_{\mathrm{M}} \end{bmatrix}. \tag{5.19}$$

$\mathbf{x}$ is the vector of transmitted QAM modulation symbols

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \tag{5.20}$$

and the noise $\mathbf{N}$ is complex Gaussian with variance σ_n^2.

Below, we concentrate on a specific cellular situation where two MIMO users with different powers are communicating to the base station. They use the same resources at the same time. The strong user is close to the base station and the weak one is further off, possibly at the cell edge. Usually, in such situations the weak user experiences very low QoS and in many cases it cannot transmit. The proposed approach aims at separating both users, so both of them can operate simultaneously.

5.4.3 Receiver

The need for sophisticated receiver algorithms is dictated by the characteristics of the SC-FDMA modulation itself. Since the signals are transmitted in the time domain, unlike with downlink transmission where OFDMA is used, they experience ISI. Furthermore, for the case when spatial multiplexing is employed, multi-antenna interference starts to play a dominant role. Therefore, the receiver has to remove both types of interference in the presence of additive noise.

5.4.4 Baseline Linear MMSE MIMO Equalization

Baseline uplink signal processing is fundamentally the same as on the downlink and is based on OFDM modulation. A key advantage of adopting OFDM for LTE is that it transforms the channel from a finite-state time-domain channel into a frequency-domain multiplicative channel, greatly simplifying basic signal processing functions. The generic MMSE equalizer for the channel model (5.16) is given by

$$\begin{aligned} \tilde{\mathbf{X}}_{\mathrm{MMSE}} &= \mathbf{A}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}}\mathbf{M}^{-1}\mathbf{Y} = \\ &= \mathbf{A}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}}\left(\mathbf{H}\mathbf{W}\mathbf{H}^{\mathrm{H}} + \sigma^2\mathbf{I}\right)^{-1}\mathbf{Y}. \end{aligned} \tag{5.21}$$

which using the matrix inversion lemma [4] can be transformed into

$$\begin{aligned} \tilde{\mathbf{X}}_{\mathrm{MMSE}} &= \mathbf{A}^{-1}\mathbf{M}^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{Y} \\ &= \mathbf{A}^{-1}\left(\mathbf{H}^{\mathrm{H}}\mathbf{H} + \sigma^2\mathbf{W}^{-1}\right)^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{Y}. \end{aligned} \tag{5.22}$$

Inversion of the matrix $\mathbf{M}$ is simple, since as shown previously it consists of N_t blocks of diagonal matrices and can be accomplished by M separate inversions of $N_t \times N_t$ size matrices. Subsequently, to obtain modulation symbol estimates in case of SC-FDMA, the resulting signal in (5.21) or (5.22) is passed to the inverse DFT block

$$\tilde{\mathbf{x}}_{\mathrm{MMSE}} = \mathbf{F}^{\mathrm{H}}\tilde{\mathbf{X}}_{\mathrm{MMSE}}. \tag{5.23}$$

Finally, these signals are passed to soft QAM demodulators followed by the individual error control decoders. In case of LTE these would be turbo codes.

5.4.5 Group Turbo MMSE MIMO Equalization

The key idea of the modification of the baseline receiver relies on the use of the following splitting of the received signal depending on the received powers [65]. Rewriting (5.16) in the presence of two users as

$$\mathbf{Y} = \mathbf{H}_1\mathbf{F}\mathbf{A}_1\mathbf{x}_1 + \mathbf{H}_2\mathbf{F}\mathbf{A}_2\mathbf{x}_2 + \mathbf{N} \tag{5.24}$$

it can be seen that signals belonging to each user (group of N_t streams) are separated and we applied the trivial equality $\mathbf{AF} = \mathbf{FA}$, which is true for diagonal matrix $\mathbf{A}$. $\mathbf{H}_i$ is the i-th column of the MIMO matrix[1], see for example (5.17). The proposed group iterative equalizer is shown in Figure 5.9. The receiver operates as follows. First the received signal

[1] In the case of multiple antenna users with N_t antennas, it is a group of N_t columns.

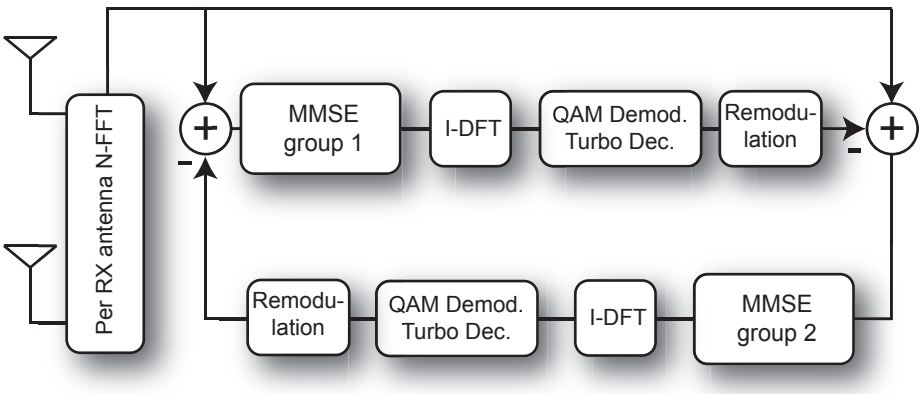

Figure 5.9: Structure of the proposed receiver for 2 power groups.

$\mathbf{Y}$ is passed to an MMSE filter tuned to the strong user. The resulting signal is[1]

$$\tilde{\mathbf{x}}_1 = \mathbf{M}_1\mathbf{Y} \tag{5.25}$$

where

$$\mathbf{M}_1 = {\mathbf{A}_1}^{\mathrm{H}}\mathbf{F}^{\mathrm{H}}{\mathbf{H}_1}^{\mathrm{H}}\left(\mathbf{H}_1\mathbf{W}_1{\mathbf{H}_1}^{\mathrm{H}} + \mathbf{H}_2\mathbf{W}_2{\mathbf{H}_2}^{\mathrm{H}} + \sigma^2\mathbf{I}\right)^{-1} \tag{5.26}$$

is the scaled[2] covariance matrix of $\mathbf{Y}$ which is the input signal to the filter at the initial iteration and a trivial relationship $\mathbf{F}\mathbf{F}^{\mathrm{H}} = \mathbf{F}^{\mathrm{H}}\mathbf{F} = \mathbf{I}$ was applied. It can now clearly be seen that the proposed receiver performs inversion in the frequency domain.

The estimate $\tilde{\mathbf{x}}_1$ is then passed to the soft-demodulator and subsequently to the turbo code decoder, which produces estimates of all (systematic and parity) bits $\hat{\mathbf{x}}_1$. These estimates can be erroneous and the influence of these errors can be captured using the soft-symbol variance parameter

$$\sigma_{s,1}^2 = \mathrm{E}\left[\mathbf{x}_1 - \hat{\mathbf{x}}_1\right]^2 \tag{5.27}$$

which is related to the power of the residual interference that the weak user will experience from imperfect cancellation of the strong user. The estimates $\hat{\mathbf{x}}_1$ are QAM remodulated to obtain $\mathbf{Y}_1$ and cancelled from the composite signal $\mathbf{Y}$ to produce the interference-reduced signal for the weak group detector/decoder

$$\mathbf{Y}_2 = \mathbf{Y} - \mathbf{Y}_1 \tag{5.28}$$

[1] Without loss of generality, it is assumed that the first user is the stronger one.
[2] by ${\mathbf{A}_1}^{\mathrm{H}}{\mathbf{F}_1}^{\mathrm{H}}{\mathbf{H}_1}^{\mathrm{H}}$.

which can be further developed into

$$\mathbf{Y}_2 = \mathbf{H}_1\mathbf{F}\mathbf{A}_1(\mathbf{x}_1 - \hat{\mathbf{x}}_1) + \mathbf{H}_2\mathbf{F}\mathbf{A}_2\mathbf{x}_2 + \mathbf{N} \tag{5.29}$$

where interstream interference[1] and intersymbol interference is present. The covariance matrix of (5.29), used in the subsequent MMSE filtering, is

$$\mathrm{E}\left[\mathbf{Y}_2\mathbf{Y}_2^H\right] = \mathbf{H}_1\mathbf{W}_1\sigma_{s,1}^2{\mathbf{H}_1}^{\mathrm{H}} + \mathbf{H}_2\mathbf{W}_2{\mathbf{H}_2}^{\mathrm{H}} + \sigma^2\mathbf{I}. \tag{5.30}$$

Similarly, the process of MMSE filtering is repeated for the weak user to obtain

$$\tilde{\mathbf{x}}_2 = \mathbf{M}_2\mathbf{Y}_2 \tag{5.31}$$

where

$$\mathbf{M}_2 = {\mathbf{A}_2}^{\mathrm{H}}\mathbf{F}^{\mathrm{H}}{\mathbf{H}_2}^{\mathrm{H}}\left(\mathbf{H}_1\mathbf{W}_1\sigma_{s,1}^2{\mathbf{H}_1}^{\mathrm{H}} + \mathbf{H}_2\mathbf{W}_2{\mathbf{H}_2}^{\mathrm{H}} + \sigma^2\mathbf{I}\right)^{-1}. \tag{5.32}$$

The cancellation and demodulation/decoding is performed and this completes a first global turbo iteration. The number of such global iterations can be arbitrary, however avoiding high complexity it should be limited to a moderate number of $2-3$ in order to facilitate practical implementation.

Before the analysis and simulation results are presented bit error rates of different modulation schemes using the 3GPP turbo code used subsequently in the LTE-MIMO setups are given. These results, depicted in Figure 5.10 act as a benchmark in evaluating the performance of the baseline as well as the proposed detector. It can be clearly seen that for chosen block lengths 700, 1400 and 2100, the decoder achieves nearly error free performance ($P_b = 10^{-5}$) for $E_b/N_0 = 2, 4.75, 9.25$ dB for QPSK, 16QAM, and 64QAM modulations, respectively.

Soft Modulation/Demodulation

In all the simulations we assume that the soft symbols $\hat{\mathbf{x}}$ are generated by first generating LLR values for each component bit of the constellation, then (after decoding) rebuilding the soft symbols using the decoded soft bits $\tilde{b}_i$. We assume that all quadrature constellations are bit-mapped in the binary superposition form, where in each dimension, a PAM symbol is created as

$$x = \sum_{i=0}^{B-1} b_i 2^i \tag{5.33}$$

and B is the order of the PAM constellation.

[1] In case of single transmit antenna, $N_t = 1$, no interstream interference occurs.

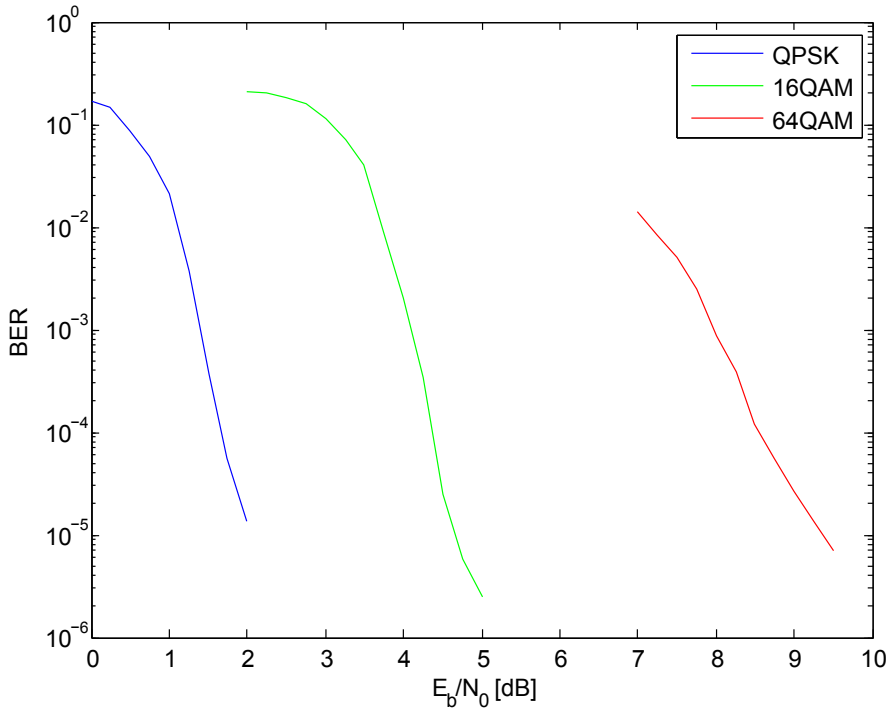

Figure 5.10: Bit-error-rates for turbo coded [13,15] data modulated using LTE standard modulations (QPSK, 16QAM and 64QAM) on AWGN channel. Code length is 700, 1400 and 2100 information bits, respectively.

5.4.6 Analysis

It is desirable to analytically predict the behavior of the proposed iterative structure. Here once again the attention is focused on the variance transfer charts that can predict the turbo-cliff occurrence of any iterative systems as described extensively in Chapter 2. With this tool the error free performance of the detector can be predicted as a function of the SNR. The only important knowledge required to sketch those graphs is the SNR transfer of each of the blocks involved in the processing chain. To comply with the LTE standards [42, 59], the choice of channel coding and modulation is fixed, therefore we do not consider other alternatives. For the case considered, two transfer functions have to be evaluated. One for the MMSE detector and one for the demodulation/decoding block as shown in Figure 5.11. The MMSE filter for user k accepts as its input the cancelled signal with residual interference power of group k' $\sigma^2_{s,k'}$ and outputs signal plus disturbance of power[1] $\sigma^2_{\text{norm},k}$ which is subsequently passed to the demodulator/decoder and re-modulator to produce soft-symbol

[1]VT analysis assumes that interference is modeled as Gaussian noise $\mathcal{CN}(0, \sigma^2_{\text{norm},k})$.

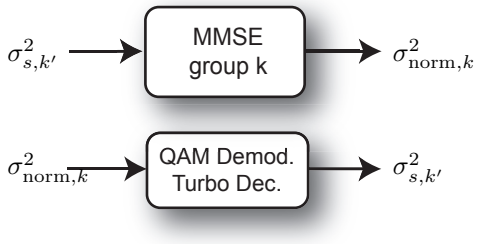

Figure 5.11: SNR transfer procedure.

values that will be used in subsequent MMSE filtering of group k'. The variance σ_k^2 can be calculated using equations (5.31) and (5.32). Usually the signal power is alternated through the MMSE filtering and in general $P_k \neq 1$. Therefore the normalization w.r.t to unit signal power is necessary and new normalized noise plus interference power is obtained as

$$\sigma^2_{\text{norm},k} = \frac{\sigma_k^2}{P_k}. \tag{5.34}$$

Having the VT plots one can precisely predict the number of turbo equalization iterations needed for the system to achieve error-free performance for both types of traffic. This is particulary interesting for mixed traffic, where one user affects others depending on parameters like modulation format, coding applied. Although the analysis given below is carried only for a two user case, it is straightforward to extend to other cases. The system converges when two VT curves of the code and detector do not cross for all users.

Subsequently, the attention is focused on two different antenna configuration systems:

- Two single antenna users communicating to BS equipped with 2 receive antennas,
- Two double antenna users communicating to BS equipped with 4 receive antennas.

The following parameters are common for both cases analyzed. A 6-path channel model with impulse response is assumed

$$\mathbb{E}[|\mathbf{h}|^2] = [0.6429 \;\; 0.2553 \;\; 0.0671 \;\; 0.0222 \;\; 0.0084 \;\; 0.0042]. \tag{5.35}$$

These values are averages over a number of runs and were obtained from an SCM channel simulator [66] for the urban macro scenario. Each of the channel impulse response taps has distribution $h_i \sim \mathcal{CN}(0, |h_i|^2)$, which implies Rayleigh fading. It is also assumed that

the entire block of M modulation symbols experiences the same fading. Since M is usually smaller than the number of symbols in a data frame therefore the whole turbo frame experiences a number of fading channels, which are assumed to be independent. At the receiver, complex Gaussian noise $\mathcal{CN}(0, \sigma^2)$ is added to every receive antenna. The average signal-to-noise ratio per bit for both systems with the assumption that both terminals have the same number of antennas is obtained as

$$\overline{E_b/N_0} = \frac{1}{2}\left(\frac{E_{s,1}}{R_{t,1}R_{m,1}\sigma^2} + \frac{E_{s,2}}{R_{t,2}R_{m,2}\sigma^2}\right) \tag{5.36}$$

and used therein. $R_{t,k}$ and, $R_{m,k}$ are the turbo code rate (here $R_{t,i} = 1/3$) and the modulation rate for user k. In the simulations, localized subcarrier mapping was used [67]. Both figures below present system behavior with the parameters listed in Table 5.3.

2 by 2 MIMO System

Figure 5.12 presents the detection/decoding process for an 2×2 MIMO system with 10 dB power difference between single antenna users for different SNR levels. The convergence trajectories are outlined. Dashed lines represent variance transfer of the QAM demodulator and turbo decoder, while solid curves describe the behavior of the group MMSE detector. To visualize the correctness of semi-analysis based on VT charts subsequently we draw the attention on a number of SNR levels that are close to the convergence point. First, let us concentrate on subfigure a) for average E_b/N_0 of 6.5 dB. The detection/decoding process starts with $\sigma_s^2 = 1$. First piece of the trajectory, horizontal line, denoted by (1), represents MMSE filtering of the strong group (equation (5.25)). Second, the vertical line (2), describes the decoding of the signal from the strong group MMSE filter. Third line (3) denotes the MMSE filtering of the weak group and (4) shows the decoding characteristics of the weak user. This completes the first turbo iteration, therefore one such iteration is described by four lines (1-4). Process continues (lines 5-7) and gets stuck. Line (8), which points in backward direction, is responsible for the lack of convergence at this SNR level. It indicates that the performance do "back-off" from the error-free situation. Precisely, due to this behavior, it can be observed that after a number of turbo iterations the red trajectory curve does get stuck in an infinite loop, denoted by the red dashed line, and the system will never achieve $\sigma_s^2 = 0$ for both users. After increasing the SNR to 7 dB we obtain system behavior presented in subfigure b). In this case both users are closer to error-free performance and as seen in subfigure c) for 7.5 dB detection/decoding trajectory arrives at $\sigma_s^2 = 0$ for both users implying that almost error-free behavior is observed.

It is however difficult to exactly point out the moment when the system achieves almost error-free performance using the variance transfer analysis. The main reason is that we do

not expect for our system to exhibit perfect cliff behavior due the application of finite length codes. Therefore, for our particular scenario we can expect vanishing error-rates for SNR values between 7-7.5 dB or even 8 dB[1] as presented later in the simulations section.

[1] As in the case of EXIT charts, VT charts do not provide any information about the actual error rates past the threshold. They just assume that these rates are small.

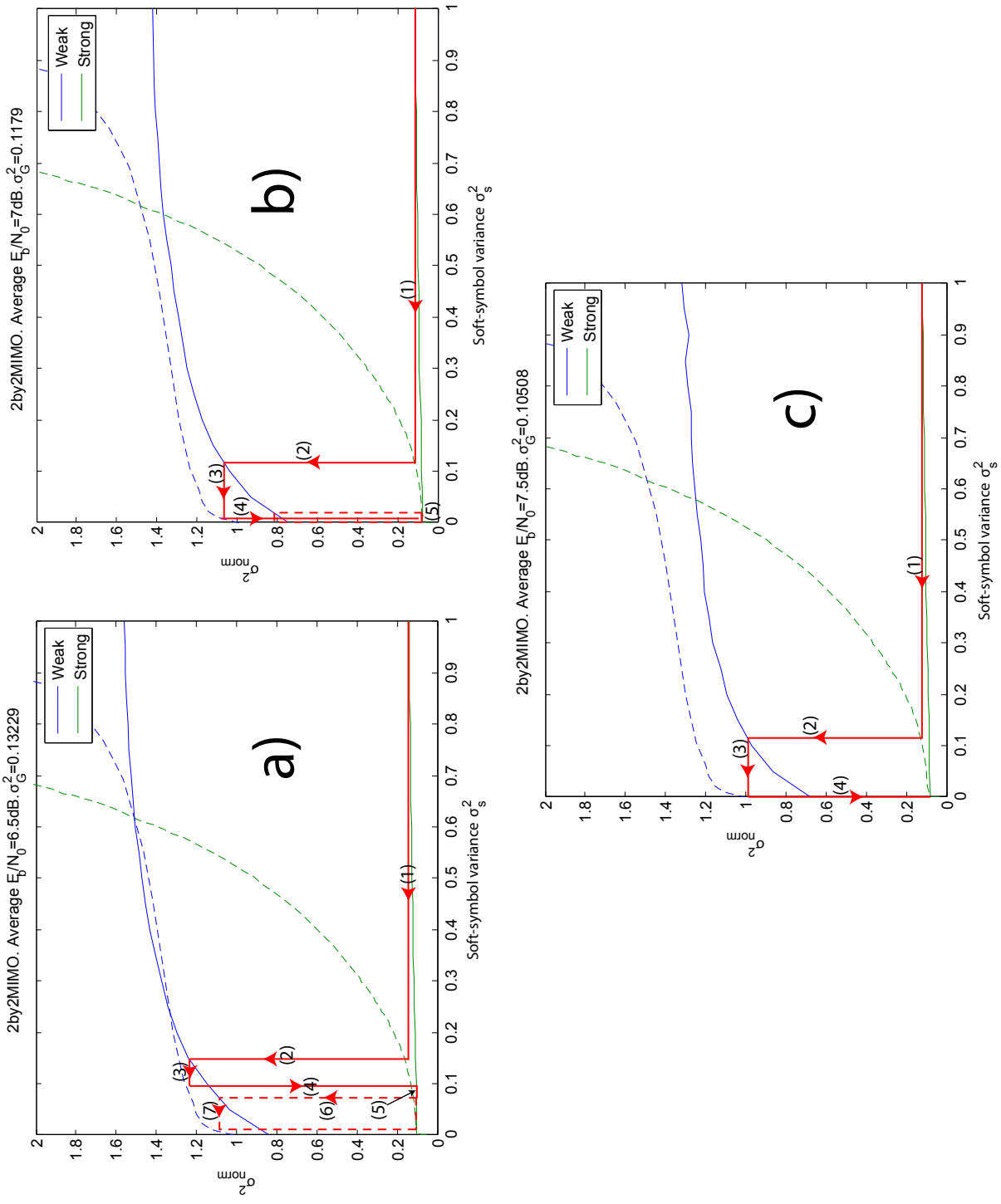

Figure 5.12: Variance transfer charts for 2×2 MIMO system at various values of average E_b/N_0

4 by 4 MIMO System

Figure 5.13 presents the detection/decoding process for an 4×4 MIMO system with 10 dB power difference between two users each equipped with two transmit antennas. Results are very similar to those for 2×2 setup. The only difference is that error-free behavior occurs for slightly larger value of SNR. Figure 5.13 c) indicates that system should give vanishing error rates when average E_b/N_0 is 8 dB. In this case everyone converges, since the detection/decoding trajectory arrives at $\sigma_s^2 = 0$.

The extension to larger number of power groups is conceptually straightforward, however a graphical representation of such a situation would require use of multidimensional[1] VT charts making them hard to read and eventually not very clear and useful. Nevertheless, the outlined method is well suited for situations when number of power groups equals two.

[1]Number of dimensions would equal to number of power groups in our case.

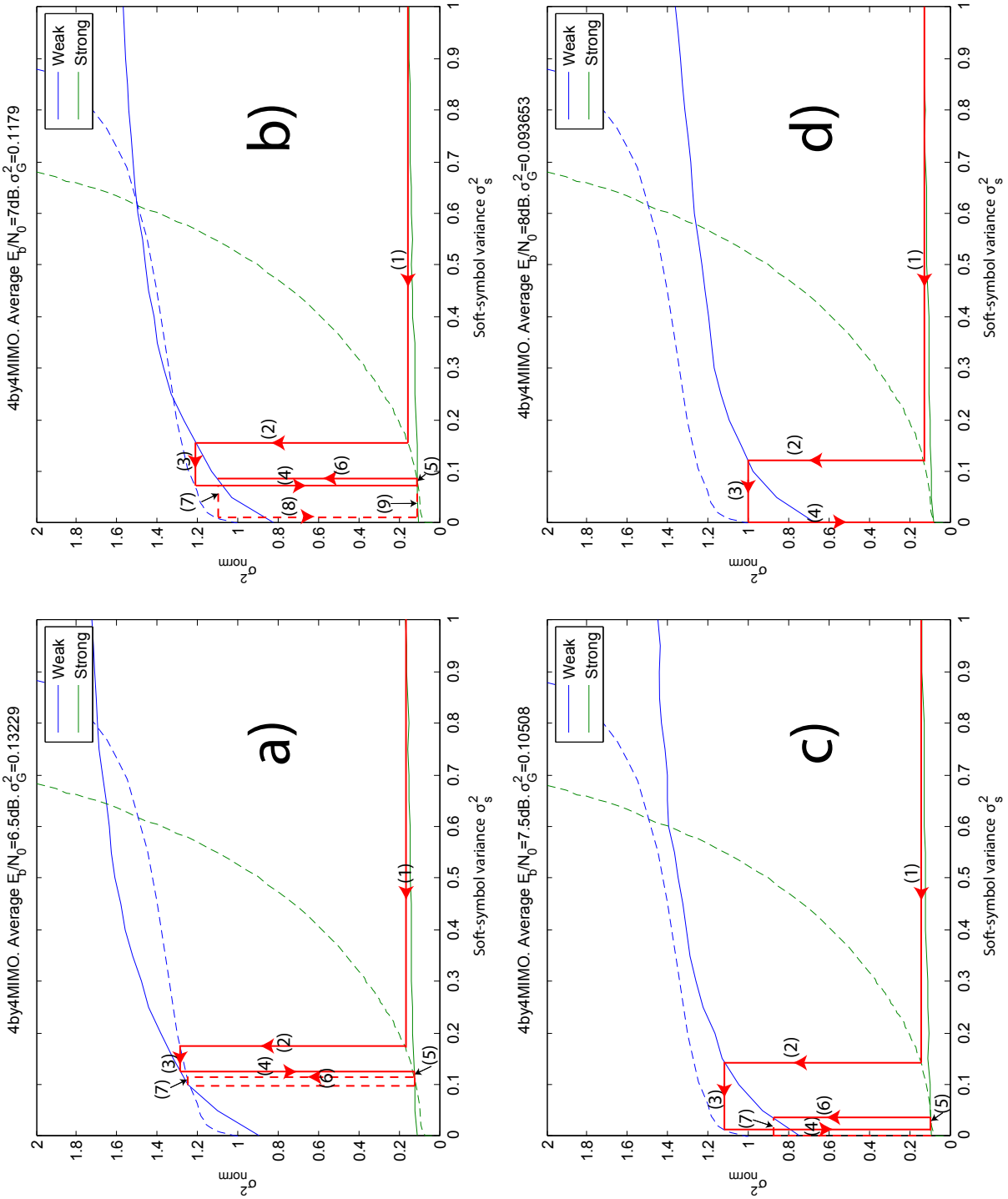

Figure 5.13: Variance transfer charts for 4×4 MIMO system at various values of average E_b/N_0.

5.4.7 Numerical Results

To prove the validity of the above proposed semi-analytical approach numerical experiments were performed. All the important simulation parameters are given in Table 5.3.

Table 5.3: MIMO-SC-FDMA LTE simulation parameters

Parameter	Value
# of users K	2
User power separation	10dB
# of transmit antennas per user N_t	1 or 2
# of receive antennas at the base station N_r	2 or 4
Turbo frame length for weak user L	700
Turbo frame length for strong user L	2100
Modulation of the weak user	QPSK
Modulation of the strong user	64QAM
Turbo encoder	$[13, 15]$
Internal turbo decoding iterations	5
Turbo equalization iterations	up to 5
DFT size M	36
Bit errors counted per user per stream	100

2 by 2 MIMO System

Simulation results for 2×2 MIMO system are presented in Figure 5.14. It can be observed that the proposed receiver provides a coding gain of about 2 dB over the baseline receiver for bit error rate of $P_b = 10^{-5}$ and this advantage is achieved after 3 turbo iterations. Performing more iterations does not provide significant improvement, it just adds to the receiver complexity. We can see that error-free performance is achieved for about SNR of 8 dB. The turbo cliff starts to occur at about 7-7.5 dB as predicted by the variance transfer charts in Figure 5.12.

Additionally, the distance to interference-free performance can be indicated which actually tells us how the proposed system combats the effects of multiuser interference. By comparing the BER curves in Figure 5.14 with those in Figure 5.10 it is seen that the weak user is still about 2.5 dB away from its AWGN performance. On the other hand this distance for the strong user is less than 0.5 dB. It practically means that by applying the proposed decoding process the strong user experiences an almost interference-free channel.

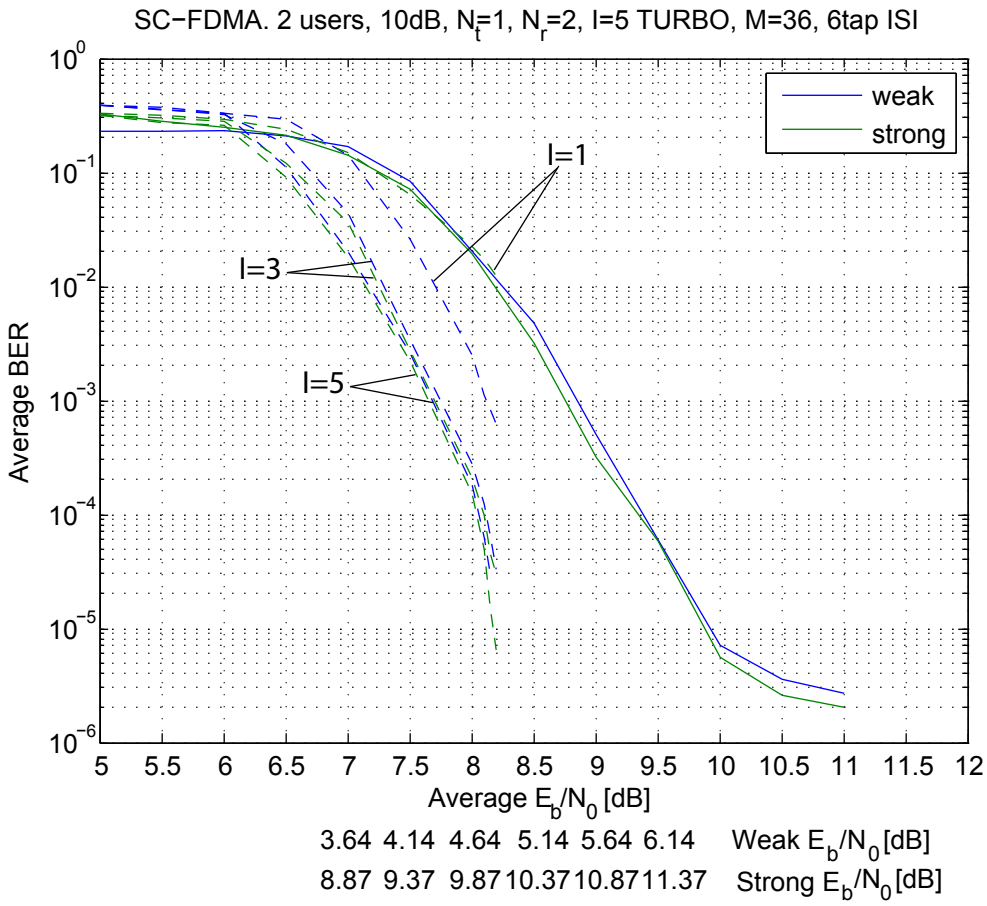

Figure 5.14: Bit error rates for 2 user 1×2 system with 10 dB power difference between users.

4 by 4 MIMO System

Similarly to the case of 2×2 system, for 4×4 system, it can be seen that the proposed receiver outperforms the baseline processor and this improvement is even larger. From the results presented in Figure 5.15 we notice that 3 global turbo iterations are necessary for the system to achieve almost error-free performance, which agrees with the predictions obtained via the variance transfer analysis in Figure 5.13. The maximum gain is ca 2.5 dB (at $P_b = 10^{-5}$) after three global turbo iterations and no visible gains were observed by performing more iterations. The performance of the strong user in the initial iteration is the same as for the baseline algorithm, since no cancellation benefits (and therefore SNR advantage) are exploited. However, the weak user experiences about 1.5 dB gain in the first iteration. The channel code thresholds for the chosen codeword lengths are 2 dB and

9.25 dB for QPSK and 64QAM modulation, respectively. This means that at least for the strong user almost interference-free performance is achieved. The weak user, on the other hand, is ca 2.5 dB away from its AWGN performance, similar to the 2×2 case. In this particular scenario, this statement implies that the weak user could be transmitting with a slightly higher rate. This may be achieved by flexible control of the code rate via puncturing and repetition, which is performed by the rate matching operation [42].

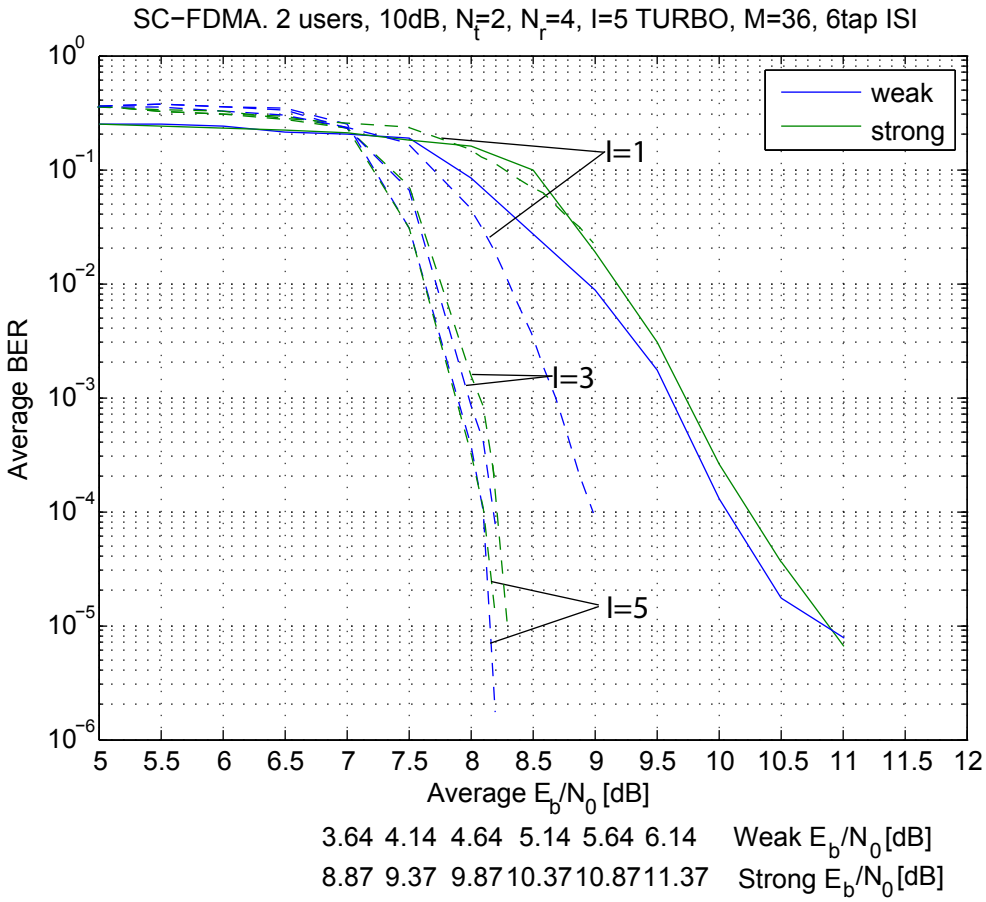

Figure 5.15: Bit error rates for 2 user 2×4 system with 10 dB power difference between users.

5.5 Chapter Summary

In this chapter the most advanced wireless communications system up to date, LTE, was presented. The radio interface as well as main characteristics and requirements for the system were briefly indicated. After a general introduction, receiver signal processing for the uplink LTE was described in detail together with its mathematical treatment. Two receivers were compared. First, the traditional two-stage receiver adopted in first release of LTE (LTE release 8) consisting of MMSE frequency domain filtering followed by the decoding of the channel code. This receiver is treated as a baseline. Using mathematical notation and applying interference cancellation methods an improved version of such a receiver is proposed. The comparisons between both systems were made on a specific scenario often present in the uplink transmission and named a cell-edge problem. In this situation different power and rate users access a common receiver. An intuitive method of predicting the convergence behavior of the proposed turbo equalization processor based on variance transfer charts is presented. The method predicts behavior of the turbo receiver accurately, as evidenced using series of computer simulations. It can be seen that an advantage of almost 2.5 dB over the baseline receiver can be experienced at a cost of reasonable complexity increase.

CHAPTER 6

Conclusions and Future Work

This thesis concentrates on methods of detecting signals in the presence of strong interference. Such a situation arises commonly in multi-user communications as well as in MIMO setups. This thesis explores methods of designing and analyzing such systems. The most dominant is the one based on variance transfer graphs. This method is based on an implicit assumption that interference is Gaussian distributed. This is certainly the case when the number of interferes is large. But as shown in Chapter 5, even for small populations VT graphs satisfactorily predict the cliff behavior of the interference cancellation receivers.

6.1 Summary of Contributions

This thesis contains a number of contributions to the field of receiver design for multi-user communications and they are summarized below.

6.1.1 Optimal Coding for IC Systems

The focus is put on answering the crucial question of what kind of channel coding is the best for interference cancellation systems. Basing on analytical results presented in [6] a number of coding schemes applied to multi-user communication systems are examined. To these belong: repetition, convolutional, Hamming, turbo and LDPC. It is shown, using variance transfer charts as well as simulation experiments, that in the high interference regime, simple coding schemes particularly repetition codes outperform the state-of-the-art capacity achieving codes like turbo and LDPC.

6.1.2 Partitioned CDMA

Using the fact that repetition coding is asymptotically optimal in the high interference regime partitioned modulation for CDMA is proposed. A precise SINR evolution is given. It is shown that this analysis is very accurate and results are virtually identical to those obtained by computer simulations. Using the outlined SINR evolution it is inferred that for certain conditions the partitioned detector of polynomial complexity outputs estimates identical to those delivered by optimal (APP) receiver of exponential complexity. Another very important observation is made. Namely, that the performance of the proposed iterative receiver is at least as good as those of widely celebrated linear MMSE filter and in many cases better.

Additionally, a concatenation of the proposed partitioned front-end with powerful error control codes (in our case regular LDPC codes) is investigated. The conclusion is that very complex per-user MMSE system with error control code input surpasses the partitioned processor only if rate of the applied code is low. For high rate codes ($R > 1/2$) the proposed structure performs significantly better resulting in considerably higher spectral efficiencies.

6.1.3 Unequal Power Allocations

It is shown that adding another degree of freedom to the partitioned CDMA system in the form of varying received signal powers results in even better performance of the partitioned iterative detector. SINR evolution equation is modified to account for power differences and using a continuous analysis it is demonstrated that arbitrary system loads can be accommodated. It is possible as long as received powers are allocated according to particular (derived) exponential distribution. Subsequently, it is shown that a capacity of a linear multiuser channel is achieved within 1 bit per dimension as long as power exponent is $a_0 = 2 \log 2$. Again these theoretical findings are proven using computer simulations. Subsequently, it is demonstrated that random received powers, mimicking non-power controlled CDMA system, outperforms the one with equal power distribution (power controlled scenario) for a number of practical scenarios.

6.1.4 Signal Processing for Uplink LTE

In order to improve the baseline system performance in a particular very often practically occurring situation, named the cell-edge problem, group-based turbo equalization scheme based on frequency domain MMSE filtering is proposed. This structure delivers almost 2 dB (2×2 MIMO) and 3 dB (4×4 MIMO) advantage over baseline processor at the expense

of a minor complexity increase. The advantage of the group-turbo equalizer increases by expanding the system size.

6.2 Future Work

Naturally, each of the points presented before can be extended. For chapter 2 all the codes analyzed are quite far from the optimal performance. Therefore, more code families can be examined like Hadamard, fountain (with its special cases LT [68] and raptor codes [69]) or repeat accumulate codes [70] to name a few. Also the modifications of codes already examined via the operations of puncturing and repetition can be a promising addition. This would fit into the design of cancellation systems for 3GPP standard, particularly LTE, where mandatory rate matching operation is mainly based on puncturing or repetition of the code bits.

For chapter 5, more analytical results are needed. Two main subjects of future research can be identified. First is the expression for the receiver output SNR for a given predefined power difference between users. The value of this output SNR can be directly translated into CQI index. On this basis, throughput advantage of the proposed receiver vs. baseline receiver can be quantified. Second is the analytical evaluation of the VT curve for the MMSE filter, so sketching VT graphs is easy and fast.

Also the reduction of the complexity during the MMSE matrix inversion in the LTE system should be addressed. It should be performed via the use of iterative filters [20] briefly introduced in Chapter 1, since these methods are very attractive alternative to direct methods. This extension seems to be particularly beneficial for large antenna arrays at the base station preferably larger than 8. Such sizes will be inevitably seen in future releases of LTE/LTE-A standard.

References

[1] C. E. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423/623–656, July/October 1948.

[2] S. Verdu and S. Shamai, "Spectral Efficiency of CDMA with Random Spreading," *IEEE Transactions on Information Theory*, vol. 45, no. 2, pp. 622–640, March 1999.

[3] M. Varanasi and A. Aazhang, "Multistage Detection in Asynchronous Code-Division Multiple-Access Communications," *IEEE Transactions on Communications*, vol. 38, no. 4, pp. 509–519, April 1990.

[4] C. Schlegel and A. Grant, *Coordinated Multiuser Communications*, Springer, 2005.

[5] T. Cover and J. Thomas, *Elements of Information Theory*, Wiley, 2006.

[6] C. Schlegel and M. Burnashev, "Optimal error control coding for iterative cancellation systems," in *Proceeedings of the IEEE International Symposium on Information Theory*, June 2010, pp. 1993–1997.

[7] L. Ping, L. Liu, K. Wu, and W. K. Leung, "Interleave-Division Multiple-Access," *IEEE Transactions on Wireless Communications*, vol. 5, no. 4, April 2006.

[8] T. Tanaka, "A Statistical Mechanics Approach to Large-System Analysis of CDMA Multiuser Detectors," *IEEE Transactions on Information Theory*, vol. 48, no. 11, pp. 2888–2910, November 2002.

[9] A. Montanari and D. Tse, "Analysis of Belief propagation for Non-Linear Problems: The Example of CDMA (or: How to Prove Tanaka's Formula)," in *Proceedings of the IEEE Information Theory Workshop*, 2006.

[10] C. Berrou and A. Glavieux, "Near Optimum Error Correcting Coding and Decoding: Turbo-Codes," *IEEE Transactions on Communications*, vol. 44, no. 10, pp. 1261–1271, October 1996.

[11] S. ten Brink, "Convergence Behavior of Iteratively Decoded Parallel Concatenated Codes," *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727–1737, October 2001.

[12] Ł. Krzymień and C. Schlegel, "Turbo Decoding with One Bit Extrinsic Quantization," *IEEE Communications Letters*, vol. 9, no. 8, pp. 732–734, August 2005.

[13] E. Telatar, "Capacity of Multi-antenna Gaussian Channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, November 1999.

[14] S. Verdu, "Capacity Region of Gaussian CDMA Channels: the Symbol Synchronous Case," in *Proceedings of the 24th Allerton Conference*, October 1986, pp. 1025–1034.

[15] A. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications*, NOW Publishers Inc., 2004.

[16] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate," *IEEE Transactions on Information Theory*, vol. 20, no. 3, pp. 284–287, March 1974.

[17] W. van Etten, "Maximum Likelihood Receiver for Multiple Channel Transmission Systems," *IEEE Transactions on Communications*, vol. 24, no. 2, pp. 276–283, February 1976.

[18] S. Verdu, "Minimum Probability of Error for Asynchronous Gaussian Multiple Access Channels," *IEEE Transactions on Information Theory*, vol. 32, no. 1, pp. 85–96, January 1986.

[19] D. Tse and S. Hanly, "Linear Multiuser Receivers: Effective Interference, Effective Bandwidth and User Capacity," *IEEE Transactions on Information Theory*, pp. 641–657, March 1999.

[20] A. Grant and C. Schlegel, "Convergence of Linear Interference Cancellation Multiuser Receivers," *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1824–1834, October 2001.

[21] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, 1996.

[22] C. Schlegel, Z. Shi, and M. Burnashev, "Optimal Power/Rate Allocation and Code Selection for Iterative Joint Detection of Coded Random CDMA," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 4286–4294, September 2006.

[23] Z. Shi and C. Schlegel, "Joint Iterative Decoding of Serially Concatenated Error Control Coded CDMA," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 2, pp. 1646–1653, August 2001.

[24] Z. Shi and C. Schlegel, "Iterative Multi-User Detection and Error Control Code Decoding in Random CDMA," *IEEE Transactions on Signal Processing*, vol. 54, no. 5, pp. 1886–1895, May 2006.

[25] Ł. Krzymień, D. Truhachev, and C. Schlegel, "Coded CDMA with Partitioned Spreading," in *Proceedings of the 44th Allerton Conference*, 2006.

[26] X. Wang and V. H. Poor, "Iterative (Turbo) Soft Interference Cancellation and Decoding for Coded CDMA," *IEEE Transactions on Communications*, vol. 47, no. 7, pp. 1046–1061, July 1999.

[27] M. Varanasi and T. Guess, "Achieving Vertices of the Capacity Region of the Synchronous Gaussian Correlated-Waveform Multiple-Access Channel with Decision-Feedback Receivers," in *Proceeedings of the IEEE International Symposium on Information Theory*, June 1997.

[28] D. MacKay and R. M. Neal, "Near Shannon Limit Performance of Low Density Parity Check Codes," *IEE Electronics Letters*, vol. 32, pp. 1645, August 2008.

[29] M. Tüchler, A. Singer, and R. Köetter, "Minimum Mean Squared Error Equalization using A Priori Information," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 673–683, March 2002.

[30] D. P. Shepherd, M. Ryan, M. C. Reed, and Z. Shi, "An Analytical Comparison of EXIT and Variance Transfer (VT) Tools for Iterative Decoder Analysis," in *Proceedings of the Asilomar Conference on Signals, Systems and Computers*, October 2005, pp. 956–960.

[31] M. V. Burnashev, C. Schlegel, W. A. Krzymień, and Z. Shi, "Characteristics Analysis of Successive Interference Cancellation Methods," *Problems of Information Trnasmission*, vol. 40, no. 4, pp. 297–317, April 2004.

[32] R. W. E. Hamming, "Error Detecting and Error Correcting Codes," *The Bell System Technical Journal*, vol. 29, pp. 147–160, 1950.

[33] M. J. E. Golay, "Notes on Digital Coding," *Proceedings of the IEEE*, vol. 37, pp. 657, 1949.

[34] T. Johansson and K. Sh. Zigangirov, "A Simple One-Sweep Algorithm for Optimal APP Symbol Decoding of Linear Codes," *IEEE Transactions on Information Theory*, November 1998.

[35] P. Elias, "Error-Free Coding," *IRE Trans. on Inform. Theory*, vol. IT-4, pp. 29–37, September 1954.

[36] J. Hagenauer and P. Höeher, "A Viterbi Algorithm with Soft-Decision Outputs and its Applications," in *Proceedings of the IEEE Global Telecommunication Conference*, November 1989, pp. 1680–1686.

[37] R. Johannesson and K. Sh. Zigangirov, *Fundamentals of Convolutional Coding*, IEEE Press, 1998.

[38] J. G. Proakis, *Digital Communications, 4th Edition*, McGraw Hill, 2000.

[39] B. Vucetic and Jinhong Y., *Turbo Codes: Principles and Applications*, Kluwer Academis Publishers, 2000.

[40] G. D. Forney, *Concatenated Codes*, MIT Press, 1966.

[41] C. Schlegel and L. Perez, *Trellis and Turbo Coding*, Wiley, 2004.

[42] 3GPP, "TS 36.212 V9.2.0 Multiplexing and Channel Coding (Release 9)," June 2010, http://www.3gpp.org/ftp/Specs/html-info/36212.htm.

[43] R. G. Gallager, *Low-Density Parity-Check Codes*, Ph.D. thesis, Massachussets Institute of Technology, 1963.

[44] R. M. Tanner, "A Recursive Approach to Low Complexity Codes," *IEEE Transactions on Information Theory*, vol. 27, no. 5, pp. 533–547, September 1981.

[45] S. Chung, T. J. Richardson, and Urbanke R. L., "Analysis of Sum-Product Decoding of Low-Density Parity-Check Codes Using a Gaussian Approximation," *IEEE Transactions on Information Theory*, vol. 47, pp. 657–670, 2001.

[46] E. Biglieri and G. Taricco, *Transmission and Reception with Multiple Antennas: Theoretical Foundations*, NOW Publishers Inc., 2004.

[47] C. Schlegel, M. Burnashev, Truhachev D., and Ł. Krzymień, "Iterative Multiuser Detection of Random CDMA Using Partitioned Spreading," in *Proceedings of the International Symposium on Turbo Coding and Applications*, April 2006.

[48] C. Schlegel, "CDMA with Partitioned Spreading," *IEEE Communications Letters*, vol. 11, no. 12, December 2007.

[49] L. Ping, L. Liu, and W. K. Leung, "A Simple Approach to Near-Optimal Multiuser Detection: Interleave-Division Multiple Access," in *Proceedings of the IEEE Wireless Communications and Networking Conference*, March 2003.

[50] Ł. Krzymień, D. Truhachev, C. Schlegel, and M. V. Burnashev, "Two–Stage Detection of Partitioned Random CDMA," *European Transactions on Telecommunications*, vol. 19, no. 5, pp. 499–509, August 2008.

[51] R. Dodd, "Hardware Implementation of Partitioned CDMA Multiuser Detector," M.S. thesis, University of Alberta, Edmonton, Canada, 2009.

[52] Z. She, E. Jackson, and S.A. Orszag, "Scale-Dependent Intermittency and Coherence in Turbulence," *Journal of Scientific Computing*, vol. 3, pp. 407–434, April 1988.

[53] K. Yamamoto and T. Kambe, "Gaussian and near-exponential probability distributions of turbulence obtained from a numerical simulation," *Fluid Dynamics Research Journal*, pp. 65–72, August 1991.

[54] S. Nagaraj, S. Khan, C. Schlegel, and M.V. Burnashev, "Differential Preamble Detection in Packet-based Wireless Networks," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 599–607, February 2009.

[55] D. Truhachev, C. Schlegel, and Ł. Krzymień, "A Two-Stage Capacity Achieving Demodulation/Decoding Method for Random Matrix Channels," *IEEE Transactions on Information Theory*, vol. 55, no. 1, pp. 136–146, January 2009.

[56] C. Schlegel, M. Burnashev, D. Truhachev, and Ł. Krzymień, "Optimized Power Distributions for Partitioned Signaling on Random Linear Matrix Channels," *Information Theory and Applications Workshop*, January 2007.

[57] H. Cramer, *Mathematical Methods of Statistics*, Princeton University Press, 1946.

[58] W. Feller, *An Introduction to Probability Theory and Its Applications, vol. 2*, Wiley, 1966.

[59] 3GPP, "TS 36.211 V9.1.0 Physical Channels and Modulation (Release 9)," March 2010, http://www.3gpp.org/ftp/Specs/html-info/36211.htm.

[60] S. Parkvall and et al., "LTE-Advanced - Evolving LTE towards IMT-Advanced," in *VTC Fall*, 2008.

[61] S. Sesia, I. Toufik, and M. Baker, *LTE, The UMTS Long Term Evolution - From Theory to Practice*, Wiley, 2009.

[62] E. Dahlman, S. Parkvall, J. Sköld, and P. Beming, *3G Evolution. HSPA and LTE for Mobile Broadband*, Elsevier, 2008.

[63] R. M. Gray, *Toeplitz and Circulant Matrices: A Review*, NOW Publishers Inc., 2006.

[64] H. G. Myung, "Single Carrier FDMA for Uplink Wireless Transmission," *IEEE Vehicular Technology Magazine*, September 2006.

[65] Ł. Krzymień and C. Schlegel, "Near-far Resistant MIMO Iterative Receiver for Uplink LTE," in *Proceeedings of the IEEE International Conference on Communications*, 2011.

[66] J. Salo, G. Del Galdo, J. Salmi, P. Kysti, M. Milojevic, D. Laselva, and C. Schneider, "MATLAB implementation of the 3GPP Spatial Channel Model (3GPP TR 25.996)," On-line, January 2005, http://www.tkk.fi/Units/Radio/scm/.

[67] H. Myung, "Technical Overview of 3GPP LTE," May 2008.

[68] M. Luby, "LT Codes," in *Foundations of Computer Science, 2002. Proceedings. The 43rd Annual IEEE Symposium on*, 2002, pp. 271 – 280.

[69] A. Shokrollahi, "Raptor Codes," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2551 –2567, June 2006.

[70] D. Divsalar, H. Jin, and R. J. McEliece, "Coding Theorems for Turbo-Like Codes," in *Proceedings of the 36th Allerton Conference*, 1998, pp. 201–210.

[71] M. S. Alencar, "A Comparison of Bounds on the Capacity of a Binary Channel," in *Proceedings of the IEEE Global Telecommunication Conference*, November 1996, pp. 1273–1275.

APPENDIX A

Detailed complexity analysis of the PS-CDMA receiver in terms of elementary operations is presented in the table below

Table A.1: PS-CDMA computational complexity per bit per iteration

Task	Computation	Operation
Matched Filtering	$KM(N/M-1)$	additions
	KN	multiplications
Soft bit evaluation	$K(2M-1)$	additions
	KM	multiplications
	$2KM$	divisions
	KM	$\tanh(\cdot)$ operations
Re-modulation	KM	multiplications
Cancellation	$2KN$	subtractions/additions
Overall	$K(M+3N-1)$	additions/subtractions
	$K(N+2M)$	multiplications
	$2KM$	divisions
	KM	$\tanh(\cdot)$ operations

APPENDIX B

B.1 Proof of Lemma 1

Any channel aspect ratio α can be achieved with two-stage decoding and an appropriate choice of a power distribution $T(u)$.

Proof. Consider the power distribution $T(u) = e^{au}$, $u \in [0, \alpha]$, and where a is a parameter. Let us consider different signal-to-noise ratios by varying the channel noise variance σ^2. Now, using $t = 1/v$, the right hand side of (4.6) can be rewritten as

$$\begin{aligned}\int_0^{\alpha} \frac{T(u)}{v} g\left(\frac{T(u)}{v}\right) du + \frac{\sigma^2}{v} &= \int_0^{\alpha} te^{au} g\left(te^{au}\right) du + t\sigma^2 \\ &= \frac{1}{a} \int_t^{te^{a\alpha}} g(x)dx + t\sigma^2 \ , \end{aligned} \tag{B.1}$$

where $x = te^{au}$, and we used $T(u) = e^{au}$. Further using the transformation

$$\begin{aligned}\int te^{au} g\left(te^{au}\right) du &= \frac{1}{a} \int g\left(te^{au}\right) d\left(te^{au}\right) \\ &= \frac{1}{a} \int g(x)dx \ , \end{aligned} \tag{B.2}$$

where $u \in [0, \alpha]$ translates into $x \in [t, te^{a\alpha}]$. Hence the convergence condition based on (4.6)

is equivalent to

$$\frac{1}{a}\int_{t}^{te^{a\alpha}} g(x)dx + t\sigma^2 < 1 \tag{B.3}$$

for $t \in \left(1/\left(\frac{e^{a\alpha}-1}{a}+\sigma^2\right), \mu\right]$ (from (4.3)). In Appendix B.4 it is shown that

$$\int_{0}^{\infty} g(x)dx = 2\ln 2 \approx 1.3863 \ , \tag{B.4}$$

and applying (B.4) in (B.3) and choosing a and σ^2 such that

$$\frac{1}{a}2\ln 2 + \mu\sigma^2 < 1 \ , \tag{B.5}$$

it is guaranteed that

$$\frac{1}{a}\int_{t}^{te^{a\alpha}} g(x)dx + t\sigma^2 < 1 \tag{B.6}$$

for $t \in (0, \mu]$ and any α. □

B.2 Proof of Lemma 3

Consider the power distribution $T(u) = e^{a_0 u}$, $u \in [0, \alpha]$ and $\sigma^2 = \sigma'^2$. Assume capacity approaching codes for each mode. Then, for any $\alpha > 0$, the spectral efficiency of the system is lower bounded by

$$C_{\text{eff}} > \alpha - 0.1985. \tag{B.7}$$

Proof. The spectral efficiency is calculated as

$$C_{\text{eff}} = \int_{0}^{\alpha} R(u)du \ , \tag{B.8}$$

where $R(u)$ are the achievable rates of the modes calculated from the capacity formula of the BIAWGN channel with signal-to-noise ratio

$$\frac{T(u)}{\sigma_\infty^2} \geq a_0 T(u) = a_0 e^{a_0 u} \ , \tag{B.9}$$

as $C_{BIAWGN}\left(T(0)/\sigma_\infty^2\right)$. For $\sigma^2 = \sigma'^2$ the system converges to $\sigma_\infty^2 = 1/a_0$, so we have

$T(u)/\sigma_\infty^2 = a_0 e^{a_0 u}$, and

$$\begin{aligned} C_{\mathit{eff}} &= \int_0^\alpha C_{BIAWGN}\left(a_0 e^{a_0 u}\right) du \\ &= \alpha - \int_0^\alpha (1 - C_{BIAWGN}\left(a_0 e^{a_0 u}\right)) du \\ &> \alpha - \int_0^\infty (1 - C_{BIAWGN}\left(a_0 e^{a_0 u}\right)) du \ . \end{aligned} \tag{B.10}$$

where the value

$$\int_0^\infty (1 - C_{BIAWGN}\left(a_0 e^{a_0 u}\right)) du = 0.1985 \tag{B.11}$$

is found by numerical integration. □

In order to avoid the infinite limit the following decomposition was used, which employs the lower bound $\underline{C}_{BIAWGN}$ on C_{BIAWGN} derived in Appendix B.5

$$\begin{aligned} &\int_0^\infty (1 - C_{BIAWGN}\left(a_0 e^{a_0 u}\right) du \\ &\quad \le \int_0^{10} (1 - C_{BIAWGN}\left(a_0 e^{a_0 u}\right) du + \int_{10}^\infty (1 - \underline{C}_{BIAWGN}\left(a_0 e^{a_0 u}\right) du \\ &\quad < \int_0^{10} (1 - C_{BIAWGN}\left(a_0 e^{a_0 u}\right) du + \frac{2\pi^{\frac{3}{2}}}{\ln 2(\pi^2 - 8)} \int_{10}^\infty e^{-\frac{1}{2} a_0 e^{a_0 u}} du \\ &\quad < \int_0^{10} (1 - C_{BIAWGN}\left(a_0 e^{a_0 u}\right) du + \frac{2\sqrt{2}\pi^2}{\ln 2(\pi^2 - 8)} Q(10) \approx 0.1985. \end{aligned}$$

B.3 Proof of Lemma 4

Consider the power distribution $T(u) = e^{a_0 u}$, $u \in [0, \alpha]$ and $\sigma^2 = \sigma'^2$. Assume capacity-achieving ECCs and let E_b/N_0 denote the average SNR of the system, where $N_0 = 2\sigma^2$, as is customary. Then for $\alpha > 1$ the AWGN channel capacity which can be achieved for this SNR satisfies

$$C_{AWGN} \le \alpha + 0.79 \ . \tag{B.12}$$

Proof. The average SNR for the system with $T(u) = e^{a_0 u}$ power distribution is expressed

as

$$\frac{E_b}{N_0} = \frac{1}{\alpha}\int_0^{\alpha} \frac{T(u)}{2R(u)\sigma^2}du \ , \tag{B.13}$$

where $R(u)$ is the rates of a BIAWGN channel facing noise with variance σ_∞^2. Since $\sigma^2 = \sigma'^2$, $\sigma_\infty^2 = 1/a_0$. Consider some fixed $\alpha' < \alpha$. We write

$$\begin{aligned}
\frac{E_b}{N_0} &= \frac{1}{\alpha}\int_0^{\alpha} \frac{T(u)}{2R(u)\sigma^2}du \\
&= \frac{1}{\alpha}\left(\int_0^{\alpha'} \frac{T(u)}{2R(u)\sigma^2}du + \int_{\alpha'}^{\alpha} \frac{T(u)}{2R(u)\sigma^2}du\right) \\
&\leq \frac{1}{\alpha}\left(\int_0^{\alpha'} \frac{T(u)}{2R(u)\sigma^2}du + \frac{1}{2R(\alpha')\sigma^2}\int_{\alpha'}^{\alpha} T(u)du\right) \\
&= \frac{1}{\alpha}\left(\int_0^{\alpha'} \frac{T(u)}{2R(u)\sigma^2}du + \frac{e^{a_0\alpha} - e^{a_0\alpha'}}{2a_0R(\alpha')\sigma^2}\right) \\
&= \frac{2^{2\alpha} + c_2}{2c_1\alpha} \ ,
\end{aligned} \tag{B.14}$$

where

$$c_1 \stackrel{\text{def}}{=} a_0R(\alpha')\sigma^2 > 0 \ , \tag{B.15}$$

$$c_2 \stackrel{\text{def}}{=} a_0R(\alpha')\sigma^2\int_0^{\alpha'} \frac{T(u)}{R(u)\sigma^2}du - e^{a_0\alpha'} \ . \tag{B.16}$$

Now define δ as the smallest positive number such that for any $\alpha \geq \alpha'$

$$\frac{2^{2\alpha} + c_2}{2c_1\alpha} \leq \frac{2^{2(\alpha+\delta)} - 1}{2(\alpha+\delta)} \qquad \Leftrightarrow \tag{B.17}$$

$$\frac{2^{2\alpha} + c_2}{c_1} + \frac{2^{2\alpha} + c_2}{\alpha c_1}\delta + 1 \leq 2^{2(\alpha+\delta)} \qquad \Leftrightarrow$$

$$\frac{1}{c_1} + \frac{c_2}{c_1 2^{2\alpha}} + \frac{\delta}{\alpha c_1} + \frac{c_2\delta}{\alpha c_1 2^{2\alpha}} + \frac{1}{2^{2\alpha}} \leq 2^{2\delta} \tag{B.18}$$

For $c_2 \in (-1, 0)$ the left hand side rapidly decreases for $\alpha > 0$ and therefore if δ satisfies

$$\frac{1}{c_1} + \frac{c_2}{c_1 2^{2\alpha'}} + \frac{\delta}{\alpha' c_1} + \frac{c_2\delta}{\alpha' c_1 2^{2\alpha'}} + \frac{1}{2^{2\alpha'}} = 2^{2\delta} \tag{B.19}$$

then (B.18) will be satisfied for all $\alpha \geq \alpha'$.

For specific numerical choices of α' we obtain different bounds:

$$\alpha' = 1 \Rightarrow \delta = 0.79 \tag{B.20}$$

$$\alpha' = 2 \Rightarrow \delta = 0.5629 \tag{B.21}$$

$$\alpha' \to \infty \Rightarrow \delta = -\frac{1}{2}\log_2(a_0\sigma^2) = 0.3796 \tag{B.22}$$

□

B.4 Integral of the Soft Bit Variance Function $g(x)$

Proposition 2.

$$\int_0^\infty g(x)dx = 2\ln 2 \tag{B.23}$$

Proof. We use the representation of the soft-bit function $g(x)$ stated in Proposition 2 from [31]. The signal-to-noise ratio parameter x used in this paper is $x = b^2$, where b is the argument of the soft-bit function in [31].

$$\begin{aligned}\int_0^\infty g(x)dx &= \int_0^\infty g(b^2)db^2 = 2\int_0^\infty bg(b^2)db \\ &= \int_0^\infty 8be^{-b^2/2}\cdot\sum_{n=0}^{\infty}(-1)^n e^{b^2(2n+1)^2/2}Q[b(2n+1)]db\ ,\end{aligned} \tag{B.24}$$

where

$$Q(y) = \frac{1}{\sqrt{2\pi}}\int_y^\infty e^{-u^2/2}du\ . \tag{B.25}$$

We start with work on the Q-function making the variable exchange $v = u - b(2n+1)$

$$\begin{aligned}Q(b(2n+1)) &= \frac{1}{\sqrt{2\pi}}\int_{b(2n+1)}^{\infty} e^{-u^2/2}du \\ &= \frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-(v+b(2n+1))^2/2}dv = \frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-v^2/2}e^{-bv(2n+1)}e^{-b^2(2n+1)^2/2}dv.\end{aligned} \tag{B.26}$$

Substituting (B.26) into (B.24) leads to

$$\begin{aligned}
&\int_0^\infty g(x)dx \\
&\quad= \int_0^\infty 8be^{-b^2/2}\sum_{n=0}^{\infty}(-1)^n e^{b^2(2n+1)^2/2}\cdot\frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-v^2/2}e^{-bv(2n+1)}e^{-b^2(2n+1)^2/2}dvdb \\
&\quad= \int_0^\infty 8be^{-b^2/2}\sum_{n=0}^{\infty}(-1)^n \cdot\frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-v^2/2}e^{-bv(2n+1)}dvdb \\
&\quad= \frac{8}{\sqrt{2\pi}}\sum_{n=0}^{\infty}(-1)^n \cdot\int_0^\infty\int_0^\infty be^{-b^2/2}e^{-v^2/2}e^{-bv(2n+1)}dvdb \\
&\quad= \frac{8}{\sqrt{2\pi}}\sum_{n=0}^{\infty}(-1)^n \frac{\sqrt{2\pi}}{2}\frac{1}{2(n+1)} = 2\sum_{n=0}^{\infty}(-1)^n\frac{1}{n+1} = 2\ln 2
\end{aligned} \tag{B.27}$$

□

B.5 Lower Bound on the Capacity of AWGN Channel with Binary Input

Consider a user with power P using BPSK modulation over AWGN channel with variance σ^2. The maximum achievable rate of this mode is determined by the capacity of the BIAWGN channel with two input levels $-\sqrt{P}, \sqrt{P}$. Below a simple convenient lower bound on the achievable rate as a function of the input SNR is derived.

Let us denote the signal to noise ratio in the channel by x, $x = P/\sigma^2$ and the corresponding BIAWGN channel capacity by $C_{BIAWGN}(x)$. We prove that

Proposition 3. *For* $x \geq a_0 = 2\ln 2$

$$C_{BIAWGN}(x) \geq 1 - \frac{2\pi^{\frac{3}{2}}}{\ln 2(\pi^2-8)}e^{-\frac{1}{2}x} \stackrel{def}{=} \underline{C}_{BIAWGN}(x) \tag{B.28}$$

Proof. In [71] (Equation (18)) gives the following lower bound on the BIAWGN channel capacity:

$$C_{BIAWGN}(x) \geq 1 - \log_2(e)\frac{2\pi}{\pi^2-8}\left[\frac{\pi}{\sqrt{2}}\mathrm{erfc}\left(\frac{2}{\pi}\sqrt{x}\right)e^{\frac{8-\pi^2}{2\pi^2}x} - 2\mathrm{erfc}\left(\frac{\sqrt{x}}{\sqrt{2}}\right)\right]. \tag{B.29}$$

For $y > 1/\sqrt{2}$ the following inequality

$$\text{erfc}(y) = 2Q(\sqrt{2}y) \leq 2\frac{e^{-(\sqrt{2}y)^2/2}}{\sqrt{2\pi}} < e^{-y^2}\sqrt{\frac{2}{\pi}} \tag{B.30}$$

holds. Now we observe that $2\sqrt{x}/\pi > 1/\sqrt{2}$ for $x \geq a_0$ and we can lower bound the right hand side of (B.29) as

$$\begin{aligned} & 1 - \log_2(e)\frac{2\pi}{\pi^2 - 8}\left[\frac{\pi}{\sqrt{2}}\text{erfc}\left(\frac{2}{\pi}\sqrt{x}\right)e^{\frac{8-\pi^2}{2\pi^2}x}\right] \\ & \geq 1 - \log_2(e)\frac{2\pi}{\pi^2 - 8}\frac{\pi}{\sqrt{2}}\frac{\sqrt{2}}{\sqrt{\pi}}e^{-\frac{4}{\pi^2}x}e^{\frac{8-\pi^2}{2\pi^2}x} \\ & = 1 - \frac{2\pi^{\frac{3}{2}}}{\ln 2(\pi^2 - 8)}e^{-\frac{1}{2}x} \ . \end{aligned} \tag{B.31}$$

□