

Dissertation

# Iterative Multi-User Receivers for CDMA Systems

ausgeführt zum Zwecke der Erlangung des akademischen Grades  
eines Doktors der technischen Wissenschaften

eingereicht an der  
Technischen Universität Wien  
Fakultät für Elektrotechnik und Informationstechnik

von  
Dipl.-Ing. Joachim Wehinger

Wien, Juli 2005



TECHNISCHE  
UNIVERSITÄT  
WIEN

VIENNA  
UNIVERSITY OF  
TECHNOLOGY

Supervisor

Prof. Markus Rupp

Institut für Nachrichten- und Hochfrequenztechnik  
Technische Universität Wien, Austria

Examiner

Prof. Erik Ström

Department of Signals and Systems  
Chalmers Tekniska Högskola, Göteborg, Sweden

– Für Margarethe und Clara –



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Motivation . . . . .	1
1.2	Related Work and Contributions . . . . .	3
1.3	Thesis Outline . . . . .	5
<b>2</b>	<b>Leaving Optimality by Graceful Degradation</b>	<b>7</b>
2.1	The Communications Channel . . . . .	8
2.2	Diversity . . . . .	13
2.3	Code Division Multiple-Access . . . . .	14
2.4	Transmitter Structure . . . . .	17
2.5	Transmission Models . . . . .	18
2.5.1	Model for Detection in Flat Fading Channel . . . . .	18
2.5.2	Model for Detection in Frequency Selective Channel . . . . .	19
2.5.3	Model for Channel Estimation in Frequency Selective Channels	20
2.6	Definition of $E_b/N_0$ . . . . .	21
2.7	The Optimum Receiver . . . . .	22
2.8	Suboptimum Receivers . . . . .	23
2.8.1	Multi-User Efficiency . . . . .	25
2.8.2	Separation Theorem . . . . .	26
2.9	Iterative Multi-User Receiver . . . . .	26
2.10	Decoding . . . . .	28
<b>3</b>	<b>Iterative Multi-User Detection</b>	<b>33</b>
3.1	Overview . . . . .	33
3.2	Optimum Multi-User Detection . . . . .	34
3.3	Linear Multi-User Detection . . . . .	35
3.4	Parallel Interference Cancellation (PIC) . . . . .	36
3.4.1	Single-User Matched Filter (SUMF) . . . . .	36
3.4.2	Linear MMSE Filter (LMMSE) . . . . .	39
3.5	Successive Interference Cancellation (SIC) . . . . .	43
3.5.1	Single-User Matched Filter (SUMF) . . . . .	43
3.5.2	Linear MMSE Filter (LMMSE) . . . . .	44

3.6	Antenna Combining . . . . .	45
3.7	Convergence Analysis . . . . .	45
3.8	Simulation Results . . . . .	47
3.8.1	Feedback Information . . . . .	47
3.8.2	Inter-Symbol Interference . . . . .	49
3.8.3	Impact of Fading . . . . .	50
3.8.4	System Capacity . . . . .	53
3.8.5	Multiple Receive Antennas . . . . .	53
3.8.6	Impact of Code Constraint Length . . . . .	58
3.9	Complexity Issues . . . . .	58
3.10	Summary . . . . .	60
<b>4</b>	<b>Iterative Channel Estimation</b>	<b>63</b>
4.1	Channel Estimation Algorithms . . . . .	63
4.1.1	Initial Estimation . . . . .	64
4.1.2	Approximate Least Squares (ALS) . . . . .	65
4.1.3	Approximate Linear MMSE (ALMMSE) . . . . .	65
4.1.4	Linear MMSE (LMMSE) . . . . .	65
4.1.5	Interference Cancellation and Correlation (IC-Corr) . . . . .	67
4.2	Mean Square Error Analysis . . . . .	67
4.3	Pilot Sequence Design . . . . .	70
4.3.1	Optimal Sequences . . . . .	70
4.3.2	Pilot Sequences . . . . .	70
4.3.3	Identifiability . . . . .	73
4.3.4	Pilot Placement . . . . .	73
4.4	Simulation Results . . . . .	74
4.4.1	Feedback Information . . . . .	74
4.4.2	Inter-Symbol Interference . . . . .	75
4.4.3	Channel Estimators . . . . .	75
4.4.4	Pilot Sequences . . . . .	79
4.4.5	Soft Multi-User Detectors . . . . .	80
4.4.6	Multiple Receive Antennas . . . . .	81
4.4.7	Realistic Channels . . . . .	82
4.5	Summary . . . . .	83
<b>5</b>	<b>Iterative UMTS Receiver</b>	<b>87</b>
5.1	UMTS-FDD Standard . . . . .	87
5.1.1	Data Block Structure . . . . .	88
5.1.2	Spreading and Modulation . . . . .	88
5.1.3	Channel Coding . . . . .	89

---

5.2	Receiver Design . . . . .	90
5.3	Processing Chain for Simulation . . . . .	92
5.4	Simulation Results . . . . .	92
5.5	Summary . . . . .	94
<b>6</b>	<b>Iterative Receiver Analysis</b>	<b>99</b>
6.1	Signal Model . . . . .	100
6.2	Density Evolution . . . . .	101
6.3	Decoding . . . . .	102
6.4	Mismatched Channel Estimation . . . . .	104
6.5	Mismatched Multi-User Detection . . . . .	105
6.6	Results . . . . .	107
6.7	Summary . . . . .	111
<b>7</b>	<b>Conclusions</b>	<b>113</b>
<b>A</b>	<b>Acronyms</b>	<b>115</b>
<b>B</b>	<b>Notation</b>	<b>119</b>
<b>C</b>	<b>Variables</b>	<b>123</b>
<b>D</b>	<b>The MAP Decoding Algorithm</b>	<b>125</b>
<b>E</b>	<b>Proof of Equation (3.20)</b>	<b>129</b>
<b>F</b>	<b>Remark on Conjecture 1</b>	<b>131</b>
<b>G</b>	<b>Mathematical Tools</b>	<b>137</b>
G.1	The Matrix Inversion Lemma . . . . .	137
G.2	Numerical Integration . . . . .	137
	<b>Bibliography</b>	<b>139</b>



# Kurzfassung

Mobile Kommunikationsnetze der dritten als auch der zukünftigen Generationen sind dafür ausgelegt, den Benutzern hochratige Datendienste für Anwendungen wie Videotelefonie und Datentransfer zur Verfügung zu stellen. Mit der derzeit üblichen Rake-Empfängerstruktur wird über kurz oder lang ein Kapazitätsengpass entstehen, der nicht auf eine Begrenztheit des vorhandenen Spektrums zurückzuführen ist, sondern vielmehr auf ineffizienten Empfangsarchitekturen beruht.

Komplexe Empfangsalgorithmen können die spektrale Effizienz durch Mehrbenutzerdetektionsverfahren erheblich steigern. Die vorliegende Arbeit untersucht iterative Empfänger für kodierte CDMA Übertragung in der Aufwärtsstrecke. Der iterative Empfänger ist ein suboptimaler, aber von der Komplexität beherrschbarer Detektionsalgorithmus. Er besteht aus einem interferenzbeseitigenden Mehrbenutzerdetektor, einer Dekoderbank sowie einem Kanalschätzer. Anstatt die Entscheidungen über die tatsächlich gesendeten Symbole direkt nach der ersten Dekodierung zu fällen, werden die vorläufigen Entscheidungen rückgekoppelt, um damit die Interferenz auf dem Kanal zu schätzen und diese vor der Detektion in der nächsten Iteration zu eliminieren. Gleichsam werden die Rückkopplungssymbole zur Unterstützung der Kanalschätzung verwendet. Diese Arbeit untersucht, welche Möglichkeiten und Grenzen iterative Mehrbenutzerempfänger in realistischen Mobilfunkkanälen mit Mehrwegeausbreitung aufweisen.

Wir gehen auf die Frage ein, unter welcher Voraussetzung Schwundkanäle für die Dauer einer Blockübertragung als konstant angesehen werden können. Es stellt sich heraus, dass der Mobilfunkkanal für die Dauer eines UMTS-Slots konstant ist, wenn die Geschwindigkeit des mobilen Endgeräts 30 km/h nicht überschreitet. Weiters untersuchen wir den Einfluss der Intersymbolinterferenz auf die Bitfehlerrate. Im Falle von UMTS kann die Intersymbolinterferenz erst für Spreizfaktoren größer als 32 vernachlässigt werden. Für kürzere Spreizfaktoren kann die Bitfehlerrate nur dann minimiert werden, wenn die Intersymbolinterferenz im Detektor eliminiert wird. In Folge wird die Frage erörtert, ob *extrinsische* oder *a-posteriori* Dekoderinformation zur Interferenzbeseitigung herangezogen werden soll. In diesem Zusammenhang wird Last als das Verhältnis von Anzahl von Benutzern zu Spreizfaktor definiert. Prinzipiell zeigt die Verwendung von entscheidungsrückgekoppelten Symbolen, welche aus *a-posteriori* Information abgeleitet werden, über einen großen Lastbereich schnellere Konvergenz. Die maximale Last wird aber nur durch *extrin-*

*sische* Rückkopplungssymbole erreicht. Wir entwickeln ein lineares MMSE Filter zur Interferenzunterdrückung, welches eine erheblich höhere Last als ein signalangepasstes Einbenutzerfilter erlaubt. Parallele Interferenzbeseitigung wird mit sukzessiver Interferenzbeseitigung verglichen. Beide Strukturen ermöglichen praktisch gleich grosse Lasten, wobei sukzessive Beseitigung weniger Iterationen benötigt. Wir erweitern die Verarbeitungsstruktur auf Mehrantennenempfang und zeigen, dass die dadurch erzielte Stabilisierung der mittleren Empfangsleistung die Last erheblich steigern kann.

Im Speziellen erweitern wir den iterativen Mehrbenutzerdetektor mit einer Kanalschätzereinheit. Für die Kanalschätzung werden neben der bekannten Pilotsymbolfolge auch die Dekoderausgangswerte benutzt. Wir entwerfen Kanalschätzer, die unterschiedliches *a-priori* Wissen über die Statistik zweiter Ordnung des Rauschens, der Kanalkoeffizienten und der rückgekoppelten Datensymbole voraussetzen. Es zeigt sich, dass ein kleinste-Quadrate-Schätzer, der keine Statistik benötigt, nach sechs Iterationen nahezu die gleichen Bitfehlerraten erzielt wie die komplexeren Verfahren, welche auf linearer MMSE-Schätzung beruhen. Weiters wurde beobachtet, dass eine einfache QPSK-Pilotsymbolfolge, abhängig vom Ausbreitungskanal, Verluste bis zu 1,75 dB in  $E_b/N_0$  aufweist, gegenüber Pilotsequenzen mit perfekter Korrelationseigenschaft. Letztere zeigen besonders dann einen Vorteil, wenn der Kanal nicht in jeder Iteration geschätzt wird, sondern nur in der ersten Iteration mit Hilfe der Pilotsymbole.

Die gewonnenen Erkenntnisse über Mehrbenutzerdetektion und Kanalschätzung werden für den Entwurf eines UMTS-konformen Empfängers niedriger Komplexität verwendet. Wir zeigen, dass die iterative UMTS Empfängerstruktur unter der Annahme von Benutzersynchronität auch für eine grosse Anzahl von Benutzern eine erhebliche Senkung der Bitfehlerrate auf praktisch interessante Werte ermöglicht. In diesem Zusammenhang werden auch die praktischen Grenzen einer solchen aufgezeigt.

Weiters analysieren wir das Konvergenzverhalten von iterativen Empfängern mit integrierter Kanalschätzung auf Basis von Dichteevolution für den Fall eines Kanals mit konstanter Amplitude und einer zufällig gewählten Phase. Die Resultate für die Herleitung der Kanalschätzeigenschaften bedienen sich Analysemethoden aus dem Gebiet der statistischen Physik. Die Leistungsfähigkeit des Empfängers wird anhand seiner Mehrbenutzereffizienz beurteilt, welche eine Funktion der Symbolblocklänge, der Anzahl der Pilotsymbole, der Last, sowie des Störabstandes ist.

**Schlagwörter:** CDMA, MAP Dekodierung, Mehrbenutzerdetektion, Mehrwegeausbreitungskanäle, Interferenzbeseitigung, Intersymbolinterferenz, Kanalschätzung, Mehrfachempfangsantennen, Pilotfolgen, UMTS, Dichteevolution, asymptotische Systemanalyse.

# Abstract

Mobile communication networks of the third and future generations are designed to offer high-data rate services like video-telephony and data-transfer. The current Rake receiver architecture will create a shortage in available bandwidth offered to the users. This is not due to a shortage in spectrum but results from inefficient receiver architectures.

Spectral efficiency can be increased considerably through multi-user detection techniques in the receiver algorithms. The present thesis investigates iterative receivers for encoded CDMA transmission in the uplink. The iterative receiver is a suboptimal receiver algorithm with manageable complexity. It consists of an interference mitigating multi-user detector, a bank of single-user decoders, and a channel estimator. Instead of deciding on the transmitted symbols right after the first decoding, the receiver feeds back tentative decision symbols to mitigate multiple-access interference in the next iteration. Similarly, soft decision symbols are used to support channel estimation in every iteration. This thesis analyses the possibilities and limits of an iterative multi-user receiver in realistic mobile radio channels with multi-path propagation.

First, we ask the question under which assumption fading channels show block-fading characteristic. It turns out that, for the duration of a UMTS-slot, channels are quasi-block static if the velocity of the mobile does not exceed 30 km/h. We investigate the influence of inter-symbol interference on bit error rate. In the framework of UMTS, inter-symbol interference can only be neglected for spreading factors larger than 32. With short spreading factors, the bit error rates can only be minimised if inter-symbol interference is eliminated in the detector. Further, we analyse whether *extrinsic* or *a-posteriori* decoder information shall be used for interference mitigation. In this context we define load as the quotient between number of users to spreading factor. In general, soft decision symbols that are derived from *a-posteriori* information allow for faster convergence over a large region of the load. However, the maximal load is only achieved when soft decision symbols are used that are derived from *extrinsic* information. We develop a low-complexity linear MMSE interference suppression filter with which the load can be increased significantly over a single-user matched filter. Parallel interference cancellation is compared to successive interference cancellation. The achievable loads of the two schemes are practically the same, although successive cancellation requires fewer iterations. We

extend the multi-user receiver to multiple-antenna reception and show that the load can be increased considerably through the stabilisation of average receive power.

In particular, we extend the iterative multi-user detector with a channel estimator. Additionally to pilot symbols, also soft decision symbols are used for channel estimation. We design channel estimators that use different *a-priori* knowledge on the second order statistics of the noise, the channel coefficients, and the soft decision symbols. After six iterations the least squares estimator, that does not require any knowledge on the statistics, attains nearly the same bit error rates as the more advanced estimators based on linear MMSE estimation. We further observed, that a simple random QPSK pilot sequence, spread with the same signature sequence as the data, shows a loss of up to 1,75 dB, over a perfect correlation pilot sequence, depending on the particular multi-path channel. Perfect correlation pilot sequences are particularly advantageous when the channel is estimated only once with the pilot symbols.

The gained knowledge on multi-user detection and channel estimation is used in the design of a UMTS compliant iterative receiver with low-complexity components. With symbol synchronous transmission the receiver decreases the bit error rates significantly to practically interesting levels. Limits in practical deployment will be discussed.

Finally, we analyse the convergence behaviour of iterative receivers with integrated channel estimation by means of density evolution for a channel with constant amplitude and random phase. The results for the derivation of the properties of the channel estimator are gained by using analysis tools from statistical physics. The performance of the iterative receiver is assessed by its multi-user efficiency which is parameterised by the transmission block length, the number of pilot symbols, the load, and the signal to noise ratio.

**Keywords:** CDMA, MAP decoding, multi-user detection, multi-path channels, interference cancellation, inter-symbol interference, channel estimation, multiple receive antennas, pilot sequences, UMTS, density evolution, large system analysis.

# Acknowledgements

I think that a thesis is not written without the help of many other people. Hence, I want to express my dearest gratitude to my supervisor Prof. Markus Rupp who was a critical reader and a very constructive contributor to this work. My thankfulness is also directed to Prof. Erik Ström for pointing out many details and sharing a truly attentive view on this thesis.

The major part of this work was conducted under the guidance of Christoph Mecklenbräuker and Ralf Müller. I want to thank Christoph for his steady support, encouragement, and for being a constant source of inspiration. I am also thankful to Ralf who introduced me to the realm of iterative receivers and multi-user communications.

My sincere appreciation goes to my colleagues Thomas Zemen, Helmut Hofstetter, Maja Lončar, Jossy Sayir, and Laura Cottatellucci for their assistance and the various fruitful discussions they were involved in.

Enduring the research work is only possible with the help of some non-technicalities. With this respect I am indebted to Florian Hammer, Ed Schofield, Stavros Toumpis, Driton Statovci, and many other colleagues for their hilarious contributions and all kinds of entertainment.

My work at ftw. was generously supported by Markus Kommenda and Horst Rode who I want to thank for creating a truly exciting and friendly work environment. I have appreciated this since I have started working with ftw.

This is the right place to say “thank you” to my wife Margarethe who was so patient and supportive in all the stages of my PhD work.



# 1 Introduction

## 1.1 Motivation

With ongoing integration of electronic circuits and growing computational power, mobile digital communication systems of the second generation became a mass product in the 1990s. GSM (Global System of Mobile Communications) and IS-95 (Interim Standard 95) are representatives of this generation. Typical net data rates are in the order of 13 kb/s and with the help of advanced techniques like EDGE (Enhanced Data Rate for Global Evolution) maximum theoretical data rates of 473,6 kb/s are possible. The advent of the Internet at around the same time triggered applications like MP3 file sharing, Voice over IP (VoIP), web browsing, video telephony, and mobile data services. These demands were considered in the design of the third generation of mobile communication systems including UMTS (Universal Mobile Telecommunication System) in Europe, CDMA2000 (Code Division Multiple-Access 2.000 MHz) in the US, and TD-SCDMA (Time Duplex - Space CDMA) in China. UMTS is a commercial multi-user communication system based on CDMA allowing net-data rates up to 384 kb/s per user (Release 99). The basic limitations in the supportable data rates is imposed by multiple-access interference: every user causes noise for the reception of every other user. Since the computational complexity of systems of the third generation is several hundred times higher than those of the second generation [Sar03], system manufacturers are striving to implement the simplest solutions rather than those supporting highest system capacity. This currently casts the promise of high data rates in a multi-user scenario as an illusion rather than a reality. There are, however, sophisticated techniques that can change this situation. One of them is iterative receivers that are the topic of this thesis.

The challenge in mobile communications is the mitigation of interference and distortions that are introduced on the mobile radio link during transmission. These are path loss, shadowing, Doppler shift, inter-symbol interference due to multipath propagation, multiple-access interference, and noise. The propagation medium represents a time-variant system that influences communications. In order to process information in an optimum way the parameters describing the channel are required at the receiver. To characterise the channel, so-called sounding or training sequences,

known to the receiver, are sent together with the data. These enable estimation of the channel characteristics at moderate cost.

The system capacity in wireless communications is limited by multiple-access interference that is introduced by other users sharing the same resources. Classical receivers, based on the Rake concept, model the interference as a white Gaussian process and do not combat cross-correlation terms from other users. Rake reception is currently the most widely used technology in CDMA but far from the optimum. Interference has a structure and can be mitigated effectively in a receiver.

This thesis is devoted to a promising concept based on an iterative receiver structure. In contrast to classical approaches, where detection and decoding are performed successively in a so-called one-shot manner [Ver98], iterative receivers exchange probabilistic quantities and approach the actually transmitted information gradually. As we will see in Chapters 3–5 this technology allows tremendous gains in terms of achievable system capacity while at the same time keeping complexity low. Throughout this work we deal with the uplink, that is, the receiver in the base station, detecting individual user streams from a commonly received signal. The results could easily be particularised to the downlink that is technically less complex since all data streams face the same channel. This thesis puts special emphasis on high capacity systems that support loads larger than one, *i.e.*, where the length of the spreading sequence is smaller than the number of accommodated users.

The practical implication of the use of iterative receivers in current and evolving systems can be projected as:

1. **Higher reception quality** - data is detected with less errors.
2. **Fewer base stations** - this affects not only the deployment costs of additional base stations but also the operational costs of the sites.
3. **Higher system capacity** - the number of supported users and/or their data-rate can be increased within a cell.
4. **Less average radiation power** - mobiles can lower their transmission power since their information is decoded with less power at the base-station. This saves energy in the mobiles.

Finally, the aim of this thesis is to address the following questions:

1. How can near-optimal multi-user performance be achieved with approximate algorithms?
2. How does a multi-user detector influence the performance?

3. Are iterative receivers suitable for deployment in multi-path environments?
4. What is the impact of channel estimation on the overall receiver performance?
5. Which channel estimation algorithms are well suited to iterative structures?
6. What are the limits of iterative multi-user processing?
7. How can iterative structures be deployed in the context of UMTS?
8. Is there analytical insight into the behaviour of iterative receivers?

In this work iterative receivers are considered in connection with CDMA systems. However, most concepts can equally be applied to other multiple-access schemes like OFDM, MC-CDMA, or TDMA.

## 1.2 Related Work and Contributions

An iterative receiver exchanges information between the detector, the decoder, and the channel estimator in an iterative fashion before a final decision on the transmitted information is reached. We restrict our attention to iterative multi-user receivers for the uplink of coded CDMA. The first publications presenting an iterative CDMA receiver with hard decoder feedback information are due to Alexander *et al.* [Ale98] and Reed *et al.* [Ree98]. Moher [Moh98] considers successive interference cancellation in connection with Rake reception and perfect channel knowledge. A paper on iterative detection and decoding with parallel interference cancellation and LMMSE filtering assuming perfect channel knowledge was published by Wang and Poor [Wan99]. The mentioned papers on iterative detection and decoding do not consider the issue of channel estimation.

The contribution of this thesis lies in the design and analysis of iterative receiver algorithms that deal with inter-symbol interference as well as multi-path channel estimation. In practical deployment iterative multi-user receivers require channel estimates. El Gamal and Gereanotis [ElGa00] and Kobayashi *et al.* [Kob01] presented estimators for the particular case of a flat fading channel. We extend iterative channel estimation to deal with multi-path channels and analyse several channel estimators using soft decision symbol information. This thesis also proposes detectors that consider *inter-symbol interference* (ISI), which is important for high data rate communications. We extend the iterative receiver to multiple-antenna reception and show that increased spatial diversity allows a significantly higher number of accommodated users. We design an iterative receiver with low-complexity components that is compliant with the UMTS

standard and demonstrate its tremendous gain over a one-shot Rake reception. This thesis also presents a convergence analysis of an iterative receivers with integrated channel estimation using soft decision symbols. The analysis is based on density evolution and describes the receiver behaviour by its multi-user efficiency.

This thesis is based upon the following publications:

- J. Wehinger, R. R. Müller, M. Lončar, and C. F. Mecklenbräuker, “Performance of Iterative CDMA Receivers with Channel Estimation in Multipath Environments”, in *Proc. 36th Asilomar Conf. on Sig., Sys. and Comp.*, pp. 1439–1443, Pacific Grove (CA), USA, Nov. 2002.
- J. Wehinger, “Performance Analysis of Iterative CDMA Receivers in Flat Fading Channels”, *Proc. Canadian Inform. Theory Workshop (CWIT)*, pp. 231–234, Waterloo (ON), Canada, May 2003.
- J. Wehinger and R. R. Müller, “Analysis of Iterative CDMA Receivers in Flat Fading”, *Proc. Allerton Conf. on Comm, Control and Comp.*, pp. 1506–1515, Monticello (IL), USA, Oct. 2003.
- J. Wehinger, C. F. Mecklenbräuker, R. R. Müller, T. Zemen, and M. Lončar, “On Channel Estimators for Iterative CDMA Multiuser Receivers in Flat Fading”, *Proc. IEEE Int. Conf. on Comm. (ICC)*, pp. 2497–2501, Paris, France, May 2004.
- J. Wehinger and C. F. Mecklenbräuker, “Iterative Space-Time Multiuser Receiver with Data-Aided Channel Estimation for CDMA”, *submitted to IEEE Trans. on Sig. Proc.*, Oct. 2004, last revision Aug. 2005.
- C. F. Mecklenbräuker, J. Wehinger, T. Zemen, F. Hlawatsch, and H. Artés, Chapter on Multiuser MIMO Equalization, *Smart Antennas – State-of-the-Art*, Hindawi Publishing Company, to appear in 2005.

To our knowledge there are four PhD theses treating similar issues. Mark Reed [Ree99] considers the issue of synchronous and asynchronous CDMA reception and beamforming with perfect channel knowledge. Pei Xiao [Xia04] analyses iterative receivers with and without coding. The contribution on coded systems is on a *maximum a-posteriori* (MAP) multi-user trellis detector. The contribution by Alexander Kocian [Koc03a] is the rigorous derivation of an iterative receiver from the EM algorithm without coding. The system channel model is a flat fading channel for synchronous and asynchronous user transmission. Alexander Lampe [Lam03] considers an iterative receiver with dedicated pilot symbols and the assumption that inter-symbol interference is neglected.

## 1.3 Thesis Outline

The content of the individual chapters is briefly described in the following.

**Chapter 2** motivates the assumptions imposed on the communication system that hold throughout this thesis. It starts with discussing mobile communication channels and the assumption of block fading. Then, we will explain the vital concept of diversity and illustrate its benefits to communication systems. We discuss CDMA and introduce the mathematical transmission models for fading channels in the uplink. Further, we describe the potential of iterative communication systems to perform near-optimally while exhibiting affordable complexity. The soft-in soft-out decoding algorithm is shortly described before we review its performance on fading channels.

In **Chapter 3** we design iterative detection and decoding algorithms that process inter-symbol interference. All detectors perform interference cancellation in a parallel or in a successive manner. Interference cancellation is followed by a linear *single-user matched filter* (SUMF) or a linear *minimum mean square error* (LMMSE) filter. First, we derive the filters and discuss their properties, before we present results on the joint behaviour of detection and decoding. We show that mitigation of ISI is crucial for short spreading factors in multi-path environments. We resolve the question of which kind of decoder output information shall be used in interference cancellation - either *extrinsic* or *a-posteriori*. Further, numerical results suggest that parallel and successive cancellation allow for the same remarkable number of users in the system. However, successive cancellation requires less iterations, in particular in fading environments. We show that the number of supported users is higher when using a linear MMSE filter rather than a SUMF. An extension to multiple receive antennas shows a further increase in the number of supported users.

**Chapter 4** is devoted to iterative receivers with feedback-supported channel estimation. Several estimator types are developed for multi-path channels that assume different statistical *a-priori* knowledge. They are classified into joint estimators without interference cancelling and those with interference cancellation. Interference cancellation on a multi-path fading channel turns out to work poorly. Reliable estimates are provided by joint estimators only. Their performance will be discussed and compared against each other. Simulations with several multi-path power delay profiles indicate their feasibility. We discuss the choice of pilot sequences for initial channel estimation. The impact of processing either *extrinsic* or *a-posteriori* information in the channel estimator is also investigated.

In **Chapter 5** the insight gained about iterative detection and channel estimation is applied to UMTS receivers. We develop a standard compliant receiver that performs parallel interference cancellation and single-user matched filtering. For channel estimation we employ an approximated least-squares estimator. An alternative low-complexity Viterbi decoding strategy is used instead of the rather costly BCJR implementation. The results indicate a significant improvement in terms of achievable bit error rate when compared to standard Rake receivers.

Finally, **Chapter 6** presents an analysis of iterative receivers based on *density evolution* (DE). To manage the theoretical tractability we restrict ourselves to a single tap channel and random spreading sequences for the case when the quotient of number of users to spreading factor stays constant while both parameters tend to infinity. Under this circumstance, mathematical results, that have their origin in statistical physics, can be used to conjecture the average performance in terms of multi-user efficiency. The resulting expression for the evolution of the multi-user efficiency predicts the practical performance limits of iterative receivers with channel estimation.

**Chapter 7** concludes this thesis with a summary of the most important insight gained in this work.

## 2 Leaving Optimality by Graceful Degradation

This chapter is devoted to the development of iterative receivers as a suboptimal solution to the problem of optimum *maximum a-posteriori* (MAP) multi-user processing. We start with an introduction to the mobile wireless channel, illustrate the vital concept of diversity, and explain *code division multiple-access* (CDMA). Then, we present the transmission models and deal with channel decoding. The assumptions that hold for this work are summarised in the following - they are stated in detail in the remaining parts of this chapter.

1. Uplink of a multi-user CDMA system operating in burst mode.
2. Chip and symbol synchronous transmission model.
3. Pulse shaping filters of the chips are not considered.
4. Short random signature sequences for data spreading.
5. Perfect power control. Only small scale fading is considered.
6. Block Rayleigh fading channel with independent taps.
7. No frequency offsets due to Doppler or oscillator mismatch.
8. Perfect synchronisation at the receiver.
9. Convolutional encoding of information streams.
10. Random interleaver.

For the treatment of UMTS in Chapter 5, some of the assumptions need to be dropped due to the restrictions given by the standard.

## 2.1 The Communications Channel

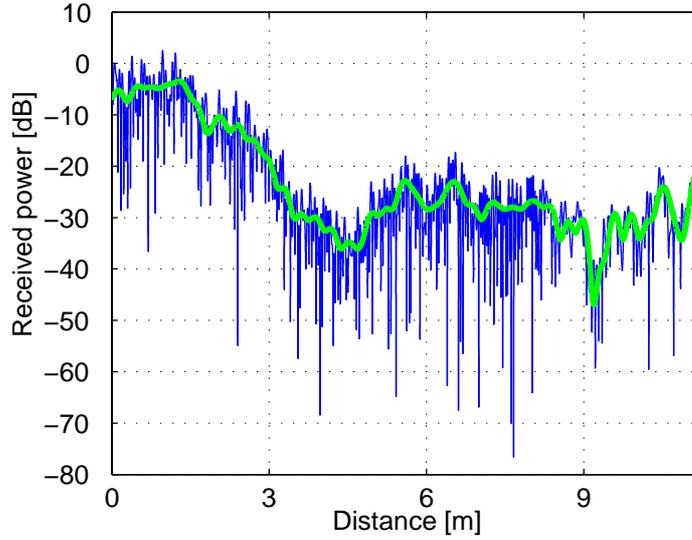
The transmission of information through electromagnetic waves is influenced by a large number of environmental effects that cause fading of the instantaneous received power. These effects can be classified into two categories: large-scale and small-scale effects. An example of fading, the variation in receive power, is shown in Fig. 2.1. The fading was generated by a geometry-based stochastic channel model [Hof04a] for an urban environment and carrier frequency  $f_C = 2$  GHz. The received power is plotted versus the distance between the transmitter and the receiver. The smoothed curve corresponds roughly to large scale fading and the residual to small scale fading. We will treat aspects of large scale effects first and then draw our attention to the small scale effects.

When waves propagate in free space their power is spread along a surface that has a distinct distance from the transmit point. The power density decreases inversely to the  $n$ -th power of the distance

$$P(d) \propto d^{-n}$$

and we call this reduction in power *path loss*, where  $n$  is the path-loss coefficient. In free space, a transmit antenna with omni-directional radiation pattern has the value  $n = 2$ . Due to reflections, refraction, and scattering in terrestrial communications, the path loss coefficient is typically between two and six [Rap96].

Path loss is one component of large-scale fading; another is shadowing. This means that the receiver whose direct *line of sight* (LOS) is obstructed receives energy by reflection and refraction, which causes significant variation of receive power over a relatively long time period. Typically, large scale effects have slowly changing characteristics and can be counteracted in communication systems by means of power control. In this thesis it is assumed that the variations caused by large scale fading are compensated perfectly. Small scale fading is much more severe. It has fast varying characteristics and cannot be counteracted with power control. Small scale fading is the result of a multitude of impinging wavefronts that interact constructively as well as destructively. This causes an interference pattern with maxima and minima that are separated by a distance in the order of  $\lambda/2$  from each other with  $\lambda$  denoting the wavelength of the carrier signal. The best known model for small-scale fading is the Rayleigh fading model. Imagine that the transmitter is located in one of the two focal points of an ellipse and the receiver is in the other focal point. Further we assume that there is no direct LOS component. All received wavefronts with the same propagation time are generated by scattering that lies on the ellipse. At the receiver they mingle together to one single concentration of effective receiver power. According to the central limit theorem the superposition of all these complex valued contributions has an amplitude that is Rayleigh-distributed,



**Figure 2.1:** A realization of received power as function of the distance between transmitter and receiver for an urban environment and carrier frequency  $f_C = 2$  GHz.

*i.e.*, has the *probability density function* (pdf)

$$f_{\text{Rayleigh}}(p) = \frac{p}{\sigma_p^2} \exp\left(-\frac{p^2}{\sigma_p^2}\right) \quad \text{for } p \geq 0.$$

Scattering contributions that can be attributed to a larger ellipse cause an energy concentration at some later point. In general the contributions decay when they have longer delays. The temporal dispersion is characterised by either the distance between the first and the last contribution, the maximal delay, or the square root of the second central moment of the power delay profile, referred to as *root mean square* (rms) delay spread [Rap96]. The emerging small scale fading channel is described by its temporal impulse response  $h(\tau)$ <sup>1</sup>. In the following we associate delay spread  $\tau_D$  to the continuous representation and maximum delay  $L$  to the discrete case. The discrete version of the channel is obtained via sampling at period  $T_C$

$$h[n] = h(nT_C)$$

where  $T_C$  denotes chip duration. The associated signal bandwidth is given through  $B = 1/T_C$ . The channel is described as a sequence of length  $L$

$$\mathbf{h} \triangleq [h[0], h[1], \dots, h[L-1]]^T.$$

<sup>1</sup>In this section time indices in parentheses  $(\cdot)$  indicate continuous time and brackets  $[\cdot]$  imply discrete time.

This allows a convenient representation of the received signal  $y[n]$  as the result of a finite convolution

$$y[n] = a[n] * h[n] \triangleq \sum_{i=0}^{L-1} a[n-i]h[i] \quad (2.1)$$

where  $*$  is the time discrete convolution and  $a[n]$  the transmit signal. Since the filtering (2.1) can be modelled by a tapped delay line, the individual weighting coefficients  $h[n]$  are called taps.

The statistical properties of the small scale fading can be expressed by its *power delay profile* (PDP) and its time-variation. The PDP reflects the strength of the arriving paths and is defined as

$$\sigma_h^2[n] = \mathbb{E} \{ |h[n]|^2 \}, n \in \{0, \dots, L-1\}.$$

Contributions arriving at later time instances have less power in average due to the larger propagation distance. A model for a PDP with exponentially decaying power is presented in the COST259 initiative [Cor01]. Fig. 2.2 illustrates its PDP for rms delay spreads of 30,9 ns and 154,7 ns. The range of the corresponding maximum delays is up to 200 ns for indoor environments and typically up to 1  $\mu$ s for outdoor environments [Hof02].

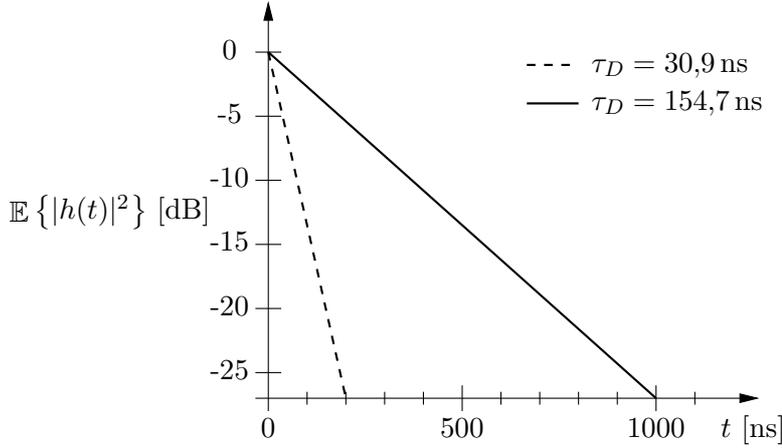
To compare the average performance results over a set of fading channel realizations for different users we make use of the following normalisation, formulated for a system with multiple receive antennas  $N_R$ :

$$\sum_{n=0}^{L-1} \sum_{r=1}^{N_R} \mathbb{E} \{ |h_r[n]|^2 \} = 1. \quad (2.2)$$

The time-variation of a channel is initiated by either a movement of the scattering, motion of the receiver, motion of the transmitter, or a combination of any of the three. When fixing the position of the transmitter and the scattering, we can view the variation of the channel as a motion through a pattern of standing waves.

The channel model used in this thesis is based on the *wide sense stationary and uncorrelated scattering* (WSSUS) assumptions that date back to Bello's work [Bel63]. This model uses  $h(\tau, t)$  for the channel to reflect its variation in a second temporal variable  $t$ . Bello's model assumes that the taps are

1. *wide-sense stationary* (WSS) with respect to  $t$
2. and that scattering contributions attributed to different propagation paths are mutually uncorrelated (*uncorrelated scattering* – US) leading to independent realizations of each tap, modelled by complex Gaussian random variables.



**Figure 2.2:** Exponentially decaying power delay profiles according to COST259 with different rms delay spreads  $\tau_D$ .

A widely used model for the resulting Doppler power spectrum from above conditions is the Jakes' model [Jak75]. Its autocorrelation for transmitting in a plane is given by

$$R_{h,h}(\tau_1, \tau_2, \Delta t) = \mathbb{E} \{h^*(\tau_1, t)h(\tau_2, t + \Delta t)\} = J_0(2\pi f_D \Delta t) \delta(\tau_1 - \tau_2)$$

with  $J_0(\cdot)$  denoting the Bessel function of first kind and zero-th order. Particularising to the interesting case  $\tau_1 = \tau_2$  we obtain the characteristic U-shaped power spectral density

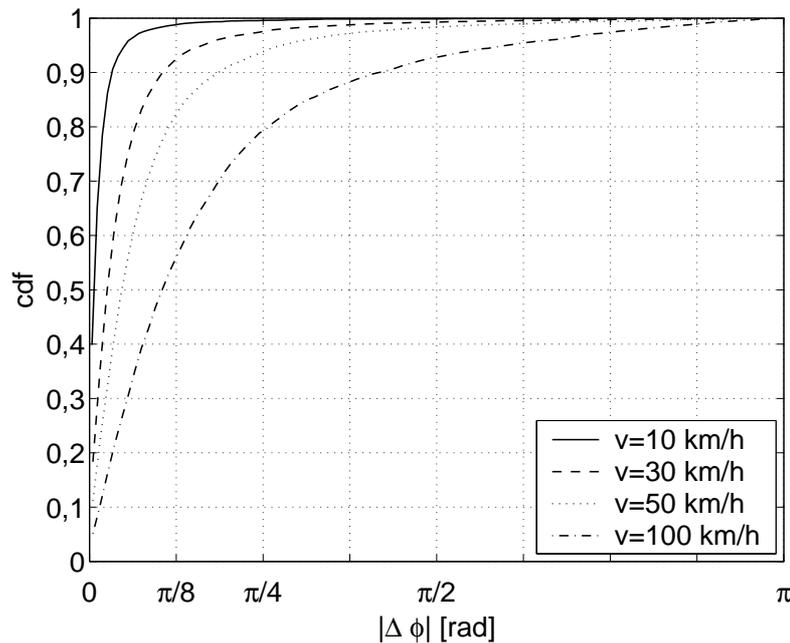
$$S_{h,h}(f) = \mathcal{F} \{R_{h,h}(\Delta t)\} = \begin{cases} \frac{1}{\pi \sqrt{f_D^2 - f^2}} & \text{for } |f| < f_D, \\ 0 & \text{elsewhere.} \end{cases}$$

The variable  $f_D$  designates the Doppler frequency given by

$$f_D = f_C \frac{v}{c_0}$$

where  $f_C$  is the carrier frequency,  $v$  the maximum relative speed between transmitter and receiver, and  $c_0$  the speed of light.

During this thesis we assume that the channel is constant for the duration of  $M$  symbols. This feature will be referred to as *block fading* characteristic. In this paragraph we investigate how well the block fading assumption is justified for the UMTS standard. UMTS operates at a chip-rate of 3,84 Mchips/s and the shortest segment of a continuous data stream is called a slot and has 2.560 chips [TS25.201]. We consider one such slot as a block for validating the block fading characteristic. One slot has a duration of  $\Delta T = 666,67 \mu\text{s}$ . The carrier frequency is 2 GHz. Assume that *quarternary phase shift keying* (QPSK) modulation is applied, then the phase shift between the beginning and the end of a slot shall not be larger than  $\pi/4$  to lead



**Figure 2.3:** Cumulative distribution functions of phase differences between start and end of a UMTS slot with duration  $666,7 \mu\text{s}$ . The carrier frequency  $f_C$  is 2 GHz and the user velocities are  $\{10, 30, 50, 100\}$  km/h.

to correct symbol decisions. We employ the simulation model presented in [Zhe03] with 50 individual scattering objects to create the time-variant fading channel. Our subject of interest is the *cumulative distribution function* (cdf) of the absolute phase difference  $|\Delta\phi| = |\phi(t + \Delta T) - \phi(t)|$  between the phase  $\phi(t)$  at the beginning and the phase  $\phi(t + \Delta T)$  at the end of the slot. The distribution of the phase difference  $\Delta\phi$  is a symmetric distribution with centre zero. Given the slot duration of  $\Delta T = 666,67 \mu\text{s}$  and a maximum speed of  $v = 100$  km/h, the phase difference is essentially limited to the range  $[-\pi, \pi]$ . The range increases for higher velocities or longer blocks. Fig. 2.3 shows four cdfs corresponding to velocities  $v = \{10, 20, 50, 100\}$  km/h. The different cumulative density curves, associated to different velocities, can be assigned outage probabilities for exceeding the maximum admissible phase difference of  $\pi/4$ . In case of a fast moving terminal with  $v = 100$  km/h this is 21 % and decreases to 7 %, 4 %, and less than 0,5 % for 50 km/h, 30 km/h, and 10 km/h, respectively. We consider outage values of 10 % a reasonable limit to justify block fading.

We like to note that a simple Rayleigh multi-path fading model is a rough approximation to reality and that state-of-the art channel models include a huge number of environmental parameters. These would comprise the propagation environment (indoor – office, home; outdoor – urban, suburban, rural, hilly terrains), number of scattering clusters, correlations among temporal as well as spatial echos, geometric

heights of transmitter and receiver etc. The modelling initiatives are driven by the European Union's projects COST259 [Cor01] and COST273 [Cor05], as well as by standardisation activities for third generation mobile communications systems in 3GPP [3GPP] and for LAN/MAN systems in IEEE802 [IEEE802]. In standardisation of UMTS, various test scenarios termed Vehicular A/B and Pedestrian A/B models [TR101 112] were introduced for verifying receiver algorithms. We use some of them in Chapter 4.

## 2.2 Diversity

The ability to exploit information on transmitted data from more than one transmission path is called diversity. Employing diversity on a wireless link increases the reception quality. As we will see shortly, the concept of fading is instrumental to mobile communications. Instead of relying on rapid fading as it is the case in a flat fading environment, *temporal* replicas can be exploited to stabilise the receive power level. The same concept is feasible for combining *spatially* separated information sources like in multiple-antenna reception. Different information sources  $x_l$  are combined via the *maximum ratio combining* (MRC) criterion, maximising the SNR of the joint information  $x$  [Bre59]:

$$x = \sum_{l=1}^L |h_l|^2 x_l / \sum_{k=1}^L |h_k|^2. \quad (2.3)$$

MRC implies that all sources are weighted by their SNR  $\gamma_l = |h_l|^2 / \sigma_v^2$  before combining. The random variable of power  $|h_l|^2$  of one single Rayleigh fading tap with variance  $\sigma_{h,l}^2$  is  $\chi^2$  ("chi-square") distributed. An MRC-combination of  $L$  equal power paths has a  $\chi^2$ -distribution with  $2L$  degrees of freedom. The corresponding pdf reads

$$f_\chi(s) = \begin{cases} \frac{s^{L-1}}{(\sigma_{h,l}^2)^L (L-1)!} e^{-\frac{s}{\sigma_{h,l}^2}} & \text{for } s \geq 0 \\ 0 & \text{else.} \end{cases}$$

Using QPSK modulation and Gray coding we obtain the *bit error rate* (BER) [Pro00]

$$\text{BER} = \left(\frac{1-\mu}{2}\right)^L \sum_{i=0}^{L-1} \binom{L-1+i}{i} \left(\frac{1+\mu}{2}\right)^i \quad (2.4)$$

when  $L$  equal power Rayleigh taps are MRC combined with

$$\mu = \sqrt{\frac{\gamma}{1+\gamma}}.$$

The variable  $\gamma$  is the SNR of a single path. A generalised expression for paths with unequal power can be found in [Alo99]. The lower bound for the BER of uncoded transmission is given for the case of an *additive white Gaussian noise* (AWGN) channel. It reads

$$\text{BER} = Q\left(\sqrt{2\gamma}\right)$$

where Marcum's  $Q(x)$  function is given through

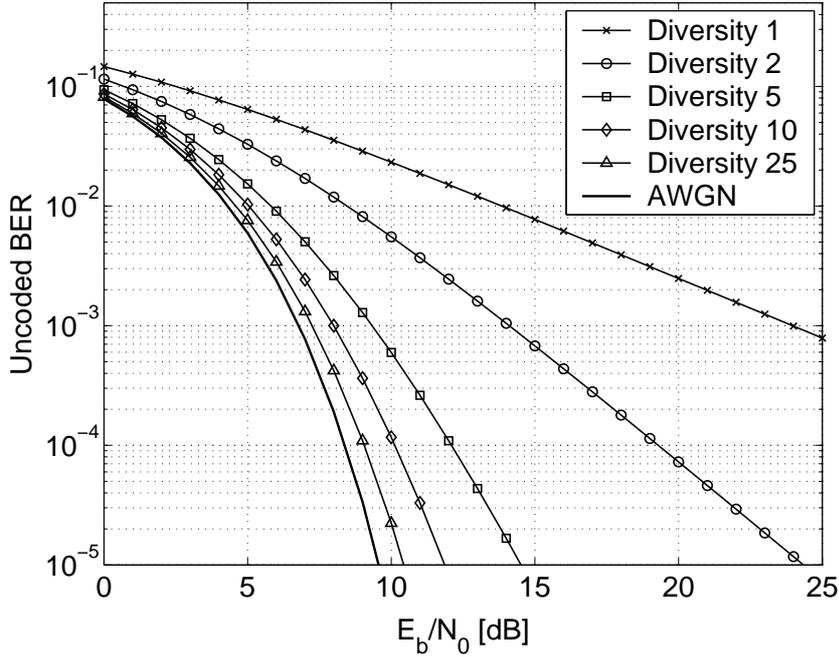
$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt. \quad (2.5)$$

Fig. 2.4 shows BER curves for different orders of diversity. Two extreme cases can be identified. The first one is when no diversity is available where  $L = 1$  and is considered as worst case. The best case is when  $L \rightarrow \infty$ . This channel becomes an AWGN channel with diversity order infinity [Jak75]. The difference in terms of BER is tremendous. The worst case achieves a BER of  $10^{-3}$  at an  $E_b/N_0$  of 24 dB while for the best case the same BER is already earned at 7 dB. The implication is that the ability to transmit at low values of  $E_b/N_0$  is made possible through exploitation of diversity. Diversity causes a steeper slope in the BER vs.  $E_b/N_0$  curve. Increasing the diversity order by one causes an improvement in BER by a factor of ten over a segment of 10 dB in  $E_b/N_0$  [Pau03]. However, this holds only for uncorrelated channels taps. In case of correlation among taps the diversity order is reduced.

## 2.3 Code Division Multiple-Access

This work deals with *code-division multiple-access* (CDMA). CDMA is the basic technology of the third and most likely also the fourth generation wireless communication system. It is a spread spectrum technology where narrow-band information with bandwidth  $B' = 1/T_S$  is spread to a wider bandwidth  $B = 1/T_C$  by the factor  $N = B/B' = T_S/T_C$ . The advantages that are brought in by spreading allow for the utilisation of the following features:

1. Flexibility in the allocation of variant data rates.
2. Flexible support of different number of users with various data rates.
3. Means to resolve multi-path components and to exploit temporal diversity.
4. Facilitates support of macro diversity for receiving information from different sources.



**Figure 2.4:** Influence of diversity on the bit error rate.

Users transmit data in the same frequency band simultaneously. Distinction between users is based on individually assigned sequences (signatures, codes) with which the transmitted information is spread in bandwidth. Users perceive other users as an increase in their noise floor. The more users are in a system, the higher is the *multiple-access interference* (MAI), *i.e.*, the distortions introduced by other users due to cross-correlations between the signatures. The basic relation in the noiseless case is given by

$$\mathbf{y} = \mathbf{S}\mathbf{b}$$

with the matrix of signature sequences [Ver98]

$$\mathbf{S} \triangleq \begin{pmatrix} s_1[0] & s_2[0] & \cdots & s_K[0] \\ s_1[1] & s_2[1] & \cdots & s_K[1] \\ \vdots & \vdots & \ddots & \vdots \\ s_1[N-1] & s_2[N-1] & \cdots & s_K[N-1] \end{pmatrix}$$

The matrix  $\mathbf{S}$  is of dimension  $N \times K$  and has *independent and identically distributed* (i.i.d.) entries from the set  $\{\pm 1 \pm j\}/\sqrt{2N}$ . The number of users is  $K$  and  $N$  denotes the spreading factor. The  $k$ -th column corresponds to the spreading sequences  $\mathbf{s}_k \in \mathbb{C}^{N \times 1}$  of user  $k$  and fulfils the energy constraint  $\|\mathbf{s}_k\|^2 = \mathbf{s}_k^H \mathbf{s}_k = 1$ . The transmitted symbols are stacked vertically in the vector  $\mathbf{b} \triangleq [b_1, b_2, \dots, b_K]^T \in \mathbb{X}^{K \times 1}$  with  $\mathbb{X}$  denoting the set of normalised QPSK symbols  $\{\pm 1 \pm j\}/\sqrt{2}$ . The symbols can

be reconstructed by convolving the corresponding matched filter  $s_k^*[N-1-n]$ ,  $n \in \{0, \dots, N-1\}$  with the receive signal  $y[n]$ ,  $n \in \{0, \dots, N-1\}$ . Practically, this operation is implemented as correlator performing chip-wise multiplication of the received signal  $y[n]$  and  $s_k^*[n]$ ,  $n \in \{0, \dots, N-1\}$  [Ver98]. For convenience the vector  $\mathbf{s}_k$  is called the single-user matched filter in this work instead of the implementation of the single-user matched filter. In matrix notation symbol detection for all users is conveniently expressed as

$$\hat{\mathbf{b}} = \underset{\mathbb{X}}{\text{quant}} \{ \mathbf{S}^T \mathbf{y} \} = \underset{\mathbb{X}}{\text{quant}} \{ \mathbf{S}^T \mathbf{S} \mathbf{b} \} = \underset{\mathbb{X}}{\text{quant}} \{ \mathbf{R} \mathbf{b} \} \quad (2.6)$$

with

$$\underset{\mathbb{X}}{\text{quant}} \{ a \} \triangleq \underset{x_i \in \mathbb{X}}{\text{argmin}} \{ |a - x_i| \}.$$

and  $\mathbf{R} \in \mathbb{C}^{K \times K}$  denoting the correlation matrix. With this detection method the transmitted information can be restored if  $\mathbf{R}$  is diagonal. Off-diagonal terms in  $\mathbf{R}$  indicate interference across users which can lead to errors. With a spreading factor of  $N$  one can select up to  $N$  orthogonal spreading sequences for which no errors are introduced in the noiseless case. In this context it is interesting to define the *system load*  $\alpha$  denoting the ratio of users  $K$  to the spreading factor  $N$

$$\alpha \triangleq \frac{K}{N}. \quad (2.7)$$

Hence, a maximal load of  $\alpha = 1$  can be supported by orthogonal spreading sequences.

However, due to multi-path propagation, the orthogonality is lost during transmission. To demonstrate this, let us define the *virtual spreading sequence* as the convolution of spreading sequence and the *true* channel

$$\check{s}_k[n] \triangleq s_k[n] * h_k[n] = \sum_{l=0}^{L-1} s_k[n-l] h_k[l], \quad n \in \{0, 1, \dots, N+L-2\}. \quad (2.8)$$

It has length  $N+L-1$ . *Virtual* underlines the idealisation of perfect channel knowledge at the receiver. Practically, the channel is estimated and will be denoted by  $\hat{h}_k[n]$ . When the estimated channel is applied in the convolution (2.8) together with the known spreading sequence  $s_k[n]$ , we obtain the *effective spreading sequence*

$$\tilde{s}_k[n] \triangleq s_k[n] * \hat{h}_k[n] = \sum_{l=0}^{L-1} s_k[n-l] \hat{h}_k[l], \quad n \in \{0, 1, \dots, N+L-2\}. \quad (2.9)$$

Concerning the influence of the channel for one single symbol transmission we notice that it is possible that

$$\check{\mathbf{s}}_k^H \check{\mathbf{s}}_l \neq 0$$

even when the associated sequences were orthogonal  $\mathbf{s}_k^H \mathbf{s}_l = 0$  for  $k \neq l$  at the transmit side.

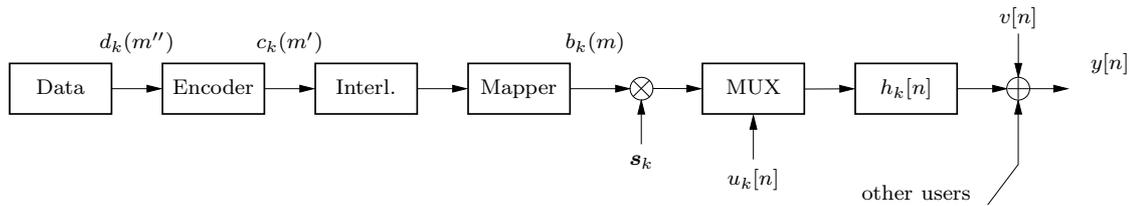
## 2.4 Transmitter Structure

This thesis treats multi-user receiver for two transmitter structures: the first is a general setup that allows for a compact mathematical representation with freely chosen components and the second is compliant with the UMTS standard. The general setup is used in all chapters except Chapter 5. In Chapter 5 the UMTS standard is revised and the adaptations to the standard are presented.

The general transmitter structure is depicted in Fig. 2.5. It shows the signal generation for user  $k$  with the propagation channel from one transmit antenna to one receive antenna at the base-station. We stick to the following convention concerning discrete time indices: parentheses  $(\cdot)$  indicate timing on a symbol level while brackets  $[\cdot]$  refer to chip instances. We consider block transmission with  $M$  symbols, where the data symbols are QPSK modulated. The first  $J$  symbols  $b_k(m)$ ,  $m \in \{0, \dots, J-1\}$  of each block are reserved for the pilot chip sequence  $u_k[n]$  with length  $JN$ . The remaining  $M-J$  symbols  $b_k(m)$ ,  $m \in \{J, \dots, M-1\}$  are data symbols. Pilot symbols are integrated in the transmitted block in order to conduct parametric channel estimation. In contrary, a non-parametric approach would not aim at directly estimating the channel but rather at using the pilots to train the equaliser. We will only consider parametric approaches. The channel can be estimated at the receiver with a known sequence of pilot symbols. The pilots are either a unique chip sequence or generated by a symbol sequence that is spread with the same signature sequence as the data.

A block is generated by encoding  $2(M-J) R_C$  bits with equally likely information symbols  $d_k(m'') \in \{0, 1\}$ . The code of user  $k$  is denoted as  $\mathcal{C}_k$  and its rate is  $R_C$ . Coding results in the code bit stream  $c_k(m') \in \{0, 1\}$  of length  $2(M-J)$  and is applied in order to introduce redundancy that allows *forward error correction* (FEC) on the receive side in case of erroneous transmissions. As we will see in Section 3.8 this becomes particularly relevant in systems that are dominated by MAI. After the encoder we pass on the code bit stream to a random interleaver. The interleaver is usually applied to alleviate the effect of burst errors that are introduced through fading on a time-variant channel or through bursty interferers. Interleaving is conducted by shuffling the encoded bits within the block prior to transmission. In case of block constant fading and encoding over a single block there is no gain to be expected due to the interleaver. Still, the interleaver is required to resolve the statistical dependencies among symbols that are introduced by the code. This will be a crucial property in the derivation of the interference suppression filters in Sections 3.4 and 3.5.

The symbols are then mapped to a QPSK symbol constellation such that  $b_{k,\text{code}}(m) \in \mathbb{X}$ . Let  $\mathbb{F}$  be the set of binary symbols  $\{0, 1\}$ , then with the mapping



**Figure 2.5:** The transmitter modules and the channel for user  $k$ .

rule

$$c \in \mathbb{F} \mapsto c' \in \{-1, +1\} : 0 \mapsto +1, 1 \mapsto -1$$

QPSK Gray labelling [Cai98] for the  $M - J$  data symbols reads

$$b_{k,\text{code}}(m) = \left( c'_k(2m) + j c'_k(2m + 1) \right) / \sqrt{2}. \quad (2.10)$$

The pilot and data symbols  $b_k(m)$  are then spread with the sequence  $\mathbf{s}_k$  before they are sent over the multi-path channel  $\mathbf{h}_k$ . This transmission scheme resembles a serially concatenated convolutional code, where the convolutional code is the outer code and the multi-path channel takes on the role of an inner convolutional code.

## 2.5 Transmission Models

In this section three transmission models for the uplink are presented. The first two models describe the transmission of one single symbol and they are used in the detector. They are devoted to the case when the channel is single path and multi-path, respectively. The third model is used for channel estimation and represents the total observation vector in matrix form. We assume that symbol transmissions are synchronous. This is motivated in [Lam02, Tse00a] where it was shown that synchronous transmissions refer to worst case scenarios in terms of mutual interference. In other words, a general asynchronous transmission is expected to have a performance not worse the one that is achieved in the synchronous case.

### 2.5.1 Model for Detection in Flat Fading Channel

For a single path channel with  $L = 1$  the transmitted signal of user  $k$  faces a stochastic, complex, multiplicative modulation due to the channel realization  $h_k$ . In the CDMA multi-user scenario the observation vector of length  $N$  that is associated with the  $m$ -th symbol transmission reads

$$\mathbf{y}(m) = \check{\mathbf{S}}(m)\mathbf{b}(m) + \mathbf{v}(m). \quad (2.11)$$

The symbols of the individual users are placed in the vector  $\mathbf{b}(m) \in \mathbb{X}^{K \times 1}$  as  $[b_1(m), b_2(m), \dots, b_K(m)]^T$ . They are modulated by the corresponding virtual spreading sequences  $\check{\mathbf{s}}_k(m)$ . The signatures are placed in the  $k$ -th column of the matrix  $\check{\mathbf{S}}(m) \in \mathbb{C}^{N \times K}$ . The symbol index  $m$  reflects that the sequences can change from one symbol transmission to the next. When they change, they are called *long* spreading sequences, as in UMTS [TS25.213]. On the other hand if  $\check{\mathbf{S}}(m)$  is constant for all symbols  $m$ , the sequences are termed *short*. The additive term  $\mathbf{v}(m) \in \mathbb{C}^{N \times 1}$  accounts for zero-mean noise with covariance matrix  $\sigma_v^2 \mathbf{I}_N$ .

### 2.5.2 Model for Detection in Frequency Selective Channel

In case of multi-path propagation, characterised through a channel of length  $L$ , energy is spread over  $N + L - 1$  chips. The observation vector  $\mathbf{y}(m)$  needs to be extended to this length in order to capture all symbol energy and thus to achieve the best receiver performance. This observation vector is distorted by inter-symbol interference attributed to previous and future symbol transmissions. We restrict ourselves to the case  $2 \leq L \leq N$  such that the number of previous as well as future symbols that affect the symbol of interest is one. The corresponding model for short sequences is formulated as

$$\mathbf{y}(m) = \sum_{q=-1}^1 \check{\mathbf{S}}_q \mathbf{b}(m - q) + \mathbf{v}(m). \quad (2.12)$$

The partial influence of the last as well as the next symbol interval stretches over  $L - 1$  chips. Terms indexed by  $q = -1$  in the sum above correspond to pre-cursor ISI caused by future symbols while terms indexed by  $q = +1$  are related to post-cursor ISI. The matrices  $\check{\mathbf{S}}_q \in \mathbb{C}^{(N+L-1) \times K}$  are attributed to previous, future, and current symbol transmissions. They are defined by

$$\check{\mathbf{S}}_q = \begin{cases} \begin{pmatrix} \check{s}_1[N] & \cdots & \check{s}_K[N] \\ \check{s}_1[N+1] & \cdots & \check{s}_K[N+1] \\ \vdots & \ddots & \vdots \\ \check{s}_1[N+L-2] & \cdots & \check{s}_K[N+L-2] \\ \mathbf{0}_N & \cdots & \mathbf{0}_N \end{pmatrix} & \text{for } q = +1 \quad (\text{post-cursor}), \\ \begin{pmatrix} \check{s}_1[0] & \cdots & \check{s}_K[0] \\ \check{s}_1[1] & \cdots & \check{s}_K[1] \\ \vdots & \ddots & \vdots \\ \check{s}_1[N+L-2] & \cdots & \check{s}_K[N+L-2] \end{pmatrix} & \text{for } q = 0, \end{cases}$$

$$\check{\mathbf{S}}_q = \begin{cases} \begin{pmatrix} \mathbf{0}_N & \cdots & \mathbf{0}_N \\ \check{s}_1[0] & \cdots & \check{s}_K[0] \\ \check{s}_1[1] & \cdots & \check{s}_K[1] \\ \vdots & \ddots & \vdots \\ \check{s}_1[L-2] & \cdots & \check{s}_K[L-2] \end{pmatrix} & \text{for } q = -1 \quad (\text{pre-cursor}). \end{cases}$$

The additive term  $\mathbf{v}(m) \in \mathbb{C}^{(N+L-1) \times 1}$  denotes again the zero-mean complex noise with covariance matrix  $\sigma_v^2 \mathbf{I}_N$ .

### 2.5.3 Model for Channel Estimation in Frequency Selective Channels

In contrast to detection where observation vectors are attributed to a symbol for time instance  $m$  we want to focus on a model that covers the samples of all symbols for a whole block. This is convenient for channel estimation since the estimate is retrieved from all the observation samples  $\mathbf{y} \in \mathbb{C}^{MN \times 1}$  together. Note that there is no explicit time dependency any longer like in (2.11) and (2.12) since all symbols are considered jointly now. The vector  $\mathbf{y}$  is also used in the derivation of the optimum receiver in Section 2.7. A linear matrix model with an isolated channel vector  $\mathbf{h}$  is given by

$$\mathbf{y} = \left( \sum_{q=0}^1 \mathcal{D}(q) \mathcal{B}(q) \right) \mathbf{h} + \mathbf{v} = \begin{bmatrix} \mathcal{A}_P \\ \mathcal{A}_D \end{bmatrix} \mathbf{h} + \mathbf{v} = \mathcal{A} \mathbf{h} + \mathbf{v}. \quad (2.13)$$

The middle expression with the matrices  $\mathcal{A}_P$  and  $\mathcal{A}_D$  stacked vertically reflects the contribution that is due to the dedicated pilot symbols  $\mathcal{A}_P$ , typically at hand for the very first channel estimation. At the receiver the true data is not known. However, in an iterative receiver we can use the soft feedback symbols to replace the true  $\mathcal{A}_D$ . In this way the quality of the channel estimates can be enhanced. The various quantities in the linear matrix model are described in the sequel:

- $\mathbf{y} \in \mathbb{C}^{MN \times 1}$  contains samples in the chip-rate. The  $L - 1$  chip samples after the block of  $MN$  samples are neglected in channel estimation.
- $\mathcal{D}(q) \in \mathbb{C}^{MN \times MKL}$  is called the “delayed chip matrix”. It is block diagonal and defined as  $\text{diag}(\underbrace{\mathbf{U}(q), \dots, \mathbf{U}(q)}_J, \underbrace{\mathbf{D}(q), \dots, \mathbf{D}(q)}_{M-J}, \dots, \mathbf{D}_1(q), \dots, \mathbf{D}_K(q)) \in \mathbb{C}^{N \times KL}$ . Let us define the following  $2N \times L$  dimensional

auxiliary Toeplitz matrix

$$\mathbf{C}_k \triangleq \begin{pmatrix} s_k[0] & 0 & 0 & 0 & \\ s_k[1] & s_k[0] & 0 & \vdots & \\ s_k[2] & s_k[1] & s_k[0] & \vdots & \\ \vdots & \vdots & \ddots & \ddots & 0 \\ s_k[N-1] & s_k[N-2] & s_k[N-3] & \ddots & s_k[0] \\ \hline 0 & s_k[N-1] & s_k[N-2] & \ddots & s_k[1] \\ \vdots & 0 & s_k[N-1] & \ddots & s_k[2] \\ \vdots & \vdots & 0 & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

where elements in  $\mathbf{C}_k$  that are not covered by the shifted spreading sequence are zero. The matrix  $\mathbf{D}_k(q)$  contains the rows of  $\mathbf{C}_k$  with number  $qN + 1$  to  $(q + 1)N$ . The construction of  $\mathbf{U}(q)$  works correspondingly where instead of the spreading sequences  $\mathbf{s}_k$  the pilot sounding sequences  $\mathbf{u}_k$  are used.

- $\mathbf{B}(q) = [\mathbf{B}(0 - q), \mathbf{B}(1 - q), \dots, \mathbf{B}(M - 1 - q)]^T$  is an  $MKL \times KL$  dimensional vertically stacked matrix consisting of the  $KL \times KL$  diagonal matrices

$$\mathbf{B}(m) = \text{diag}(\mathbf{b}(m) \otimes \mathbf{1}_L)$$

where  $\otimes$  denotes the Kronecker product. Symbol vectors  $\mathbf{b}(m)$  with  $0 \leq m < J$  contain the random pilot symbols and symbols  $J \leq m < M$  represent the QPSK data symbols. The vector  $\mathbf{b}(m = -1)$  is the zero vector  $\mathbf{0}_K$ .

- $\mathbf{h} \in \mathbb{C}^{KL \times 1}$  is a vector obtained by vertically stacking the *channel impulse responses* (CIRs)  $\mathbf{h}_k \in \mathbb{C}^{L \times 1}$  of all users' channels, *i.e.*,  $\mathbf{h} = [\mathbf{h}_1^T, \mathbf{h}_2^T, \dots, \mathbf{h}_K^T]^T$ .
- $\mathbf{v} \in \mathbb{C}^{MN \times 1}$  is modelled as zero-mean spherically invariant complex Gaussian random vector with covariance matrix  $\sigma_v^2 \mathbf{I}_{NM}$ .

## 2.6 Definition of $E_b/N_0$

For the assessment of iterative receivers we will illustrate bit error rate curves versus  $E_b/N_0$ . In this work  $E_b/N_0$  is understood as the received energy per information bit over the spectral noise density  $N_0$ . The signal to noise ratio is defined as the quotient of received signal power and noise power. It reads

$$\frac{P}{\sigma_v^2} = \frac{E_S B}{N_0 B} = G \frac{E_b}{N_0}$$

where  $E_S$  is the energy per symbol,  $B = 1/T_C$  the signal bandwidth, and  $G$  the number of bits per modulation symbol. This is easily recast to

$$\frac{E_b}{N_0} = \frac{P}{G\sigma_v^2}.$$

We keep the transmit power  $P = 1$  and the modulation is QPSK, *i.e.*,  $G = 2$ . When coding is considered, the energy per information bit becomes

$$\frac{E_b}{N_0} = \frac{1}{2\sigma_v^2 R_C}$$

where  $R_C$  denotes the code rate. In Chapters 4–6 we will make use of  $J$  pilot symbols per transmission block of length  $M$ . To stay consistent with our definition of  $E_b/N_0$  we introduce a penalisation for the energy loss in pilot symbols such that we arrive at

$$\frac{E_b}{N_0} = \frac{1}{2\sigma_v^2 R_C} \frac{M}{M - J}. \quad (2.14)$$

## 2.7 The Optimum Receiver

In this section, the optimum receiver is understood as the joint *maximum a-posteriori* (MAP) sequence detector that minimises the error probability of all user information sequences. The optimal way of estimating the information symbols  $d_k(m'')$  for all users  $k \in \{1, \dots, K\}$  and time instances  $m'' \in \{2JR_C, \dots, 2MR_C - 1\}$  from the receive vector  $\mathbf{y} \in \mathbb{C}^{MN \times 1}$  is a joint maximum *a-posteriori* approach over the transmitted code and pilot symbols  $b_k(m)$ ,  $m \in \{0, \dots, M - 1\}$ ,  $k \in \{1, \dots, K\}$  and the channel impulse responses  $\mathbf{h} \in \mathbb{C}^{KL \times 1}$ . Let us denote  $\mathbf{B} \in \mathbb{X}^{M \times K}$  the matrix with the symbols  $b_k(m)$  in the  $k$ -th column and  $m$ -th row. We consider the matrix transmission model (2.13) for the flat fading case. Then, the optimisation problem reads:

$$\begin{aligned} (\hat{\mathbf{B}}, \hat{\mathbf{h}}) &= \underset{\mathbf{B}, \mathbf{h}}{\operatorname{argmax}} f_{\mathbf{B}, \mathbf{h} | \mathbf{y}}(\mathbf{B}, \mathbf{h}, \mathbf{y}) \\ &= \underset{\mathbf{B}, \mathbf{h}}{\operatorname{argmax}} f_{\mathbf{y} | \mathbf{B}, \mathbf{h}}(\mathbf{B}, \mathbf{h}, \mathbf{y}) p_{\mathbf{B}}(\mathbf{B}) f_{\mathbf{h}}(\mathbf{h}) \end{aligned} \quad (2.15)$$

subject to

$$\begin{aligned} \mathbf{B} &\in \{\text{set of all valid code words}\}, \\ \mathbf{h} &\in \mathbb{C}^{KL \times 1}. \end{aligned}$$

The last expression in (2.15) follows from the assumption that the code symbols  $\mathbf{B}$  are independent of the channel vector  $\mathbf{h}$ . With  $f(\cdot)$  and  $p(\cdot)$  we distinguish between

the continuous *probability density function* (pdf) and the discrete *probability mass function* (pmf). With the elements of  $\mathbf{B}$  placed in  $\mathcal{B}$ , according to Paragraph 2.5.3, the conditional density distribution of the received vector given the code symbols and the channel impulse responses is the complex Gaussian multi-variate distribution

$$f_{\mathbf{y}|\mathcal{B},\mathbf{h}}(\mathcal{B}, \mathbf{h}, \mathbf{y}) = \frac{1}{(\pi\sigma_v^2)^{MN}} \exp\left(-\frac{1}{\sigma_v^2} (\mathbf{y} - \mathcal{D}\mathcal{B}\mathbf{h})^H (\mathbf{y} - \mathcal{D}\mathcal{B}\mathbf{h})\right).$$

The joint pmf of the code symbols is denoted by  $p_{\mathbf{B}}(\mathbf{B})$  and the pdf  $f_{\mathbf{h}}(\mathbf{h})$  contains the knowledge of the channel statistics.

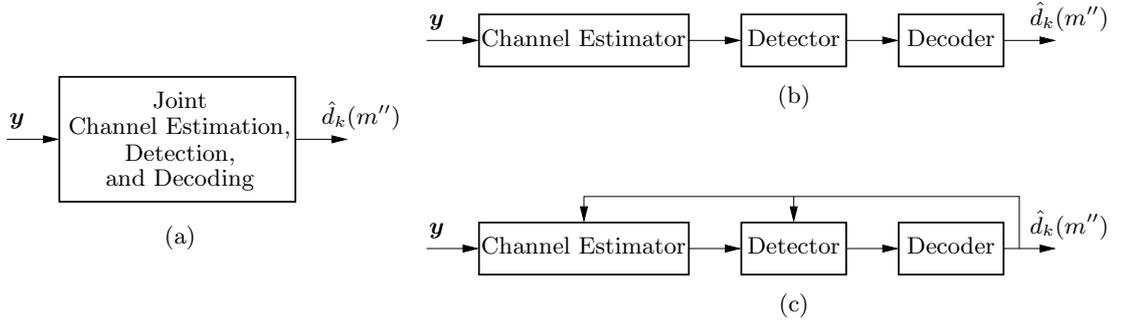
The problem formulated in (2.15) is a mixed-alphabet optimisation problem (the entries of  $\mathbf{B}$  are from a finite alphabet and those of  $\mathbf{h}$  from an infinite one) which is hard to solve. The only known exact solution to the optimisation over the code-words requires an exhaustive search over  $2^{2(M-J)R_C K}$  symbol combinations in case of QPSK modulation and  $J$  pilot symbols, thus exponential complexity in  $M - J$  and  $K$ . Its structure is the same as the travelling-salesman problem and is known as *NP-complete* (nondeterministic polynomial) problem. This means that there is no known algorithm that is able to find a solution in a duration that can be expressed as a polynomial function of time. If a solution to any of the problems in the NP-complete class can be found, then there would also exist solutions to the other problems in this class [Man89]. Although the problem formulation (2.15) looks similar as problems in complex estimation, its particular difficulty is brought in by the restriction to discrete solutions and the size of the search space.

Additionally to the NP-complete code-word problem, a maximum-likelihood estimation of the continuous channel estimate is required. Since solely the task of detection becomes computational prohibitive, the approach of joint data detection and channel estimation exceeds the computational capabilities of current processing technologies, and is hence not feasible. Therefore, we are interested in the development of low-complexity approximations to the solution of (2.15).

## 2.8 Suboptimum Receivers

Practical receiver implementations separate the individual processing tasks. Typically a single-user matched filter is followed by a decoder in a one-shot manner. This is nowadays the state of the art. Comparisons to joint optimum symbol sequence processing illustrate that separate one-shot processing causes big losses in terms of system capacity.

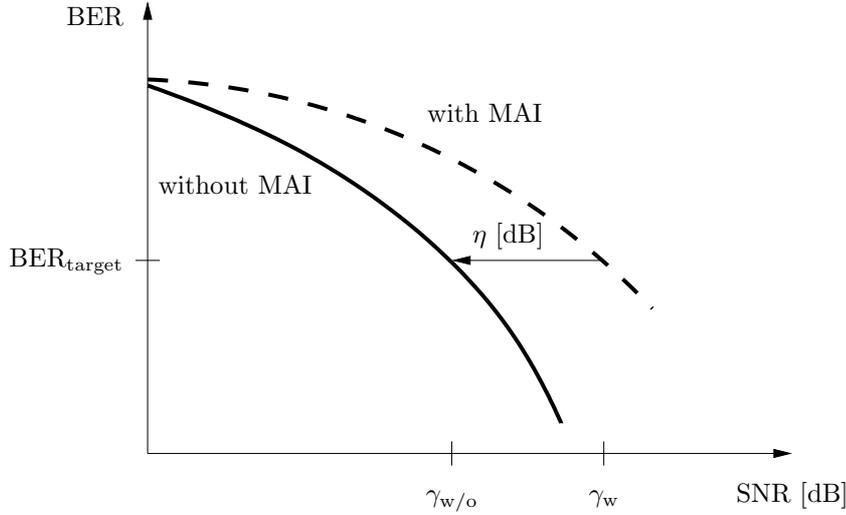
In 1993 a coding scheme was presented that allows to approach the Shannon-bound up to fractions of a dB [Ber93]. The underlying principle was a decoding system that has two independent decoders supporting each other by feeding back



**Figure 2.6:** Receiver processing methods - (a) joint processing, (b) classical one-shot processing, (c) iterative processing.

information, similar to a Turbo engine. The latter is the reason for calling these systems Turbo decoders. Few years after this important discovery, the turbo structure was applied to equalise channels with inter-symbol interference [Dou95] and to perform multi-user detection and decoding [Ale98]. Instead of implementing the joint optimum sequence receiver, the system is translated into an iterative system consisting of a multi-user detector and a bank of single-user decoders. We adapt this scheme to deal with inter-symbol interference processing and extend it to include multi-path channel estimation in the iterative loop. The discussed receiver structures are illustrated as block-diagrams in Fig. 2.6. We recap that structure (a) is the joint approach including channel estimation, detection, and decoding and was identified as too complex for practical implementation. The type (b) receiver is the classical one-shot concept that allows only a very low system load. The multi-user receiver that can support high loads at moderate complexity is based on the iterative structure (c). This iterative receiver is the subject of interest in this thesis.

Due to the separation of detection, decoding, and channel estimation the code symbols are detected individually regardless of their code constraints. This means that we perform symbol by symbol detection rather than joint sequence detection of  $M - J$  code symbols like in the optimum receiver. Chapter 3 will review symbol by symbol detection methods and detectors are developed that are suitable for iterative processing. Decoding happens individually for all users in a bank of single-user decoders that receive user code symbols containing distortions from residual multiple-access interference and remaining noise that could not be resolved by the detector in the previous processing stage. Finally, channel estimation is based on dedicated pilot symbols that are supported by soft decision data from the decoder output. In Chapter 4 we develop and investigate different channel estimation algorithms that fit into the iterative framework. Comparing the presented iterative approach with formulation (2.15) for the optimum receiver shows that a relaxation on code symbols takes place. It does not only detect symbols independently but



**Figure 2.7:** Schematic illustration of multi-user efficiency.

approaches the solution by mapping and solving the problem into the continuous domain before reaching a hard decision.

### 2.8.1 Multi-User Efficiency

In order to compare different multi-user detectors Verdú introduced the term *multi-user efficiency* (MUE) [Ver98]. We denote MUE as  $\eta$  and it is defined as the quotient

$$\eta = \frac{\gamma_{w/o}(\text{BER}_{\text{target}}, \text{Detector})}{\gamma_w(\text{BER}_{\text{target}}, \text{Detector}, \alpha)}.$$

Expression  $\gamma_{w/o}$  refers to the signal to noise ratio that is required to achieve a target bit error rate  $\text{BER}_{\text{target}}$  without multiple-access interference, *i.e.*, in the single-user case. The value  $\gamma_w$  in the denominator is the signal to noise ratio that is required to achieve a target bit error rate  $\text{BER}_{\text{target}}$  with a certain multi-user detector in case when the system load is  $\alpha$ . By definition MUE quantifies the efficiency of a multi-user detection scheme in mitigating multiple-access interference. In the worst case,  $\eta$  becomes zero, and in the best case  $\eta$  is one which refers to a total mitigation of multiple-access interference. In Fig. 2.7 the lower curve shows the BER versus SNR in case of the single-user setting while the upper dashed curve refers to the BER performance in the multiple-access scenario. The horizontal distance is the MUE expressed in dB. Its range is  $(-\infty, 0]$  dB.

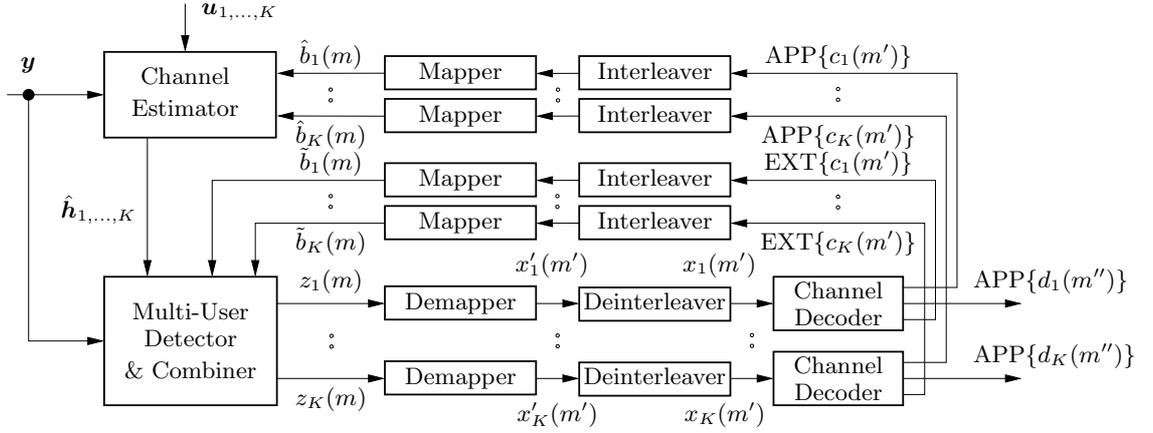


Figure 2.8: The iterative multi-user receiver.

## 2.8.2 Separation Theorem

A recent contribution [Mül04] is directed to the question on the loss in spectral efficiency that is brought in by separating detection and decoding compared to optimum joint detection and decoding. The paper reveals that for Gaussian as well as binary codes the loss on the AWGN channel is a mere function of the multi-user efficiency  $\eta$  of the detector and reads

$$C_{\text{loss}} = C_{\text{joint}} - C_{\text{separate}} = \frac{1}{2} \left( (\eta - 1) \log_2(e) - \log_2(\eta) \right). \quad (2.16)$$

The quantity  $C_{\text{joint}}$  denotes the spectral efficiency in case of joint detection and decoding and  $C_{\text{separate}}$  is the one when separate processing is applied. The loss  $C_{\text{loss}}$  is a monotonically decreasing function of  $\eta \in \{0, \dots, 1\}$ , from “ $\infty$ ” to 0. This means that the loss to optimum joint detection and decoding vanishes for  $\eta \rightarrow 1$ . We will see in Section 3.7 that for a particular set of system parameters an iterative receiver is able to achieve  $\eta = 1$  practically within a finite number of iterations.

## 2.9 Iterative Multi-User Receiver

The proposal for the iterative CDMA multi-user receiver is depicted in Fig. 2.8. It results from the arguments on the trade-off between complexity and performance that were discussed in Section 2.8. We will explain its operation principle and present an interpretation based on a graph representation.

### Operation Principle

The receiver starts with estimating the *channel impulse response* (CIRs) of all users by using the chip pilot sequences  $\mathbf{u}_k$  that cover the first  $J$  symbols in the transmitted data block of length  $M$ . The channel estimates  $\hat{\mathbf{h}}_k^{(1)} \in \mathbb{C}^{L \times 1}$  of each user  $k$  in the first iteration are then passed on to the multi-user detector to create the effective spreading sequence  $\tilde{\mathbf{s}}_k^{(1)} = s_k[n] * \hat{h}_k^{(1)}[n]$ . Note that in contrast to the virtual sequences (2.8), effective sequences  $\tilde{\mathbf{s}}_k^{(1)} \in \mathbb{C}^{N+L-1 \times 1}$  are obtained via the estimated CIR  $\hat{\mathbf{h}}_k^{(1)}$ . With these we obtain soft data estimates  $z_k^{(1)}(m), m \in \{J, \dots, M-1\}$  from the multi-user detector. The complex soft estimates are mapped to a real-valued symbol stream  $x_k^{(1)}(m'), m' \in \{2J, \dots, 2M-1\}$ , deinterleaved, and fed as  $x_k^{(1)}(m')$  to the input of the channel decoders. The decoders are so called *soft-in soft-out* (SISO) decoders that provide *a-posteriori* and *extrinsic* probabilities on code symbols,  $\text{APP}\{c_k(m')\}$  and  $\text{EXT}\{c_k(m')\}$ , respectively. The SISO decoder and its output quantities will be discussed in more detail in Section 2.10. For the moment let us consider these values to be probabilistic reliability values on code bits that are computed from the observations  $x_k^{(1)}(m')$  and the code dependencies between individual code bits. We use the probabilities  $\text{APP}\{c_k^{(1)}(m')\}$  to support the channel estimator to yield  $\hat{\mathbf{h}}_k^{(2)}$ . In this sense the data symbols are exploited as additional pilot symbols. The iterative structure will use the decoder output values also to mitigate interference. The quantities  $\text{EXT}\{c_k^{(1)}(m')\}$  are mapped to soft decision symbols  $\tilde{b}_k^{(1)}(m)$ . These are then used to generate an estimate of the interference by multiplication with the effective spreading sequence  $\tilde{\mathbf{s}}_k^{(2)}$ . The interference estimate is then subtracted from the received vector  $\mathbf{y}$  prior to detection. We then obtain symbols  $z_k^{(2)}(m)$  in the second iteration. This process is continued for several iterations. After the last iteration, the *a-posteriori* output probabilities on information symbols  $\text{APP}\{d_k(m'')\}$  delivered from the decoders are utilised to decide on the transmitted information bits along the following rule:

$$\hat{d}_k(m'') = \begin{cases} 0 & \text{for } \text{APP}\{d_k(m'')\} > 0,5, \\ 1 & \text{for } \text{APP}\{d_k(m'')\} \leq 0,5. \end{cases}$$

### Interpretation

The flow of information in iterative detection and decoding systems can be described by graphs. If a graph is free of cycles, the corresponding system is an instance of the *belief propagation* (BP) algorithm that computes correct *a-posteriori* values on information symbols [Wick03]. This applies also to iterative receivers [Wor01]. Though iterative receivers have some sort of cycles and hence violate this rule, it was observed that, when cycles are large enough, the analysis based on BP still

yields excellent results. A first systematic approach to study iterative receivers with graphs was reported in [Bou02]. This reference points out to use *extrinsic* information in the feedback-branch rather than *a-posteriori* information. It is not obvious which kind of information to use since *a-posteriori* information seems to increase the convergence speed as we will see in Paragraph 3.8.1. There, we will also resolve the issue of the usage of feedback information. A conceptual extension to the graph representation of an iterative receiver including channel estimation is given in [Cai01c, Wor01].

## 2.10 Decoding

In the following we explain decoding of information for one particular user. For notational convenience we omit the user index  $(\cdot)_k$ . At the transmitter single information bits  $d(m'')$  are mapped to code bits  $c(m')$  according to  $\mathbb{F} \mapsto \mathbb{F}^{1/R_C}$  by applying the code  $\mathcal{C}$ . In the receiver the symbols that are fed into the decoder are scaled and noisy versions of the bipolarly mapped symbols  $c(m')$ . The BPSK-mapping goes along  $c'(m') = 1 - 2c(m')$  and the received values read

$$x(m') = \mu_x c'(m') + w(m'). \quad (2.17)$$

The output of a multi-user detector is typically biased as we will see in Chapter 3. This is taken into account by  $\mu_x$ . The variable  $w(m')$  accounts for the residual noise contribution after the detector, blending system noise and multiple-access interference. It was shown in [Zha01] that for an LMMSE multi-user detector the residual noise is well modelled as Gaussian process. The detector output distribution was also modelled as Gaussian process in interference cancelling multi-user detectors in [Wan99, Kob01]. In other words, the influences to which the signal  $c'(m')$  is exposed to during transmission are summarised in the scaling factor  $\mu_x$  and the single additive noise component  $w(m')$ . The received value  $x(m')$  has the conditional distribution  $\mathcal{N}(\mu_x, \sigma_x^2)$  and  $w(m')$  is zero-mean Gaussian distributed with variance  $\sigma_x^2$ . A practical method to estimate  $\mu_x$  and  $\sigma_x^2$  from the observations  $x(m')$  is the decision directed estimator discussed in [Mec05].

The art of decoding is concerned with finding the most likely transmitted sequence of information bits  $d(m'')$ . This is made possible through mutual dependencies of code bits that are introduced in the encoder. Let  $p(\cdot)$  denote a pmf,  $\mathbf{d} \triangleq [c(2J), \dots, c(2MR_C - 1)]^T$ , and  $\mathbf{x} \triangleq [x(2J), \dots, x(2M - 1)]^T$ . The two most prominent optimisation criteria for decoding are:

- *Maximum-likelihood* sequence decoding (Viterbi [Bel57, Vit67]) for which the

optimisation criterion reads

$$\hat{\mathbf{d}} = \operatorname{argmax}_{\mathbf{d}} \{p_{\mathbf{x}|\mathbf{d}}(\mathbf{d}, \mathbf{x})\}.$$

- *Maximum a-posteriori* (MAP) symbol-by-symbol decoding (BCJR [Bah74]) that is solving

$$\hat{d}(m') = \operatorname{argmax}_{d(m')} \{p_{d(m')|\mathbf{x}}(d(m'), \mathbf{x})\}.$$

Viterbi sequence decoding is widely used in one-shot receivers and delivers hard output values on estimated information bits  $\hat{d}(m'')$ .

In iterative structures soft values are preferred since hard values cause the propagation of errors and eventually become unstable or reach an error floor. Symbol-by-symbol MAP decoding generates probabilities as output quantities that can be mapped to soft decision symbols suitable for structures with feedback. The MAP algorithm minimises the single bit error probabilities. The *soft-output Viterbi algorithm* (SOVA) [Hag89] and the max-log-MAP implementation [Koc90] have lower implementation complexity but cause higher bit error probabilities in general.

For our purpose we need a MAP decoder with the input and output measures depicted in Fig. 2.9. Since the input and all output variables are soft measures the MAP decoder is termed *soft-input soft-output* (SISO) decoder. The input are the noisy observations  $x(m')$  and the output *a-posteriori probability* (APP) values on the information bits  $d(m'')$  as well as their APP and EXT on the coded bits  $c(m')$ . In the following we stick to the definitions:

$$\text{APP}\{d(m'')\} \triangleq p_{d(m'')=0|\mathbf{x}}(d(m'') = 0, \mathbf{x}), \quad (2.18)$$

$$\text{APP}\{c(m')\} \triangleq p_{c(m')=0|\mathbf{x}}(c(m') = 0, \mathbf{x}). \quad (2.19)$$

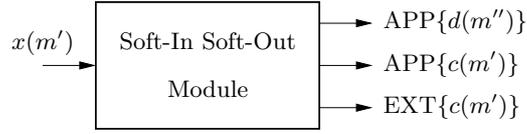
The conditional pmf for the symbol  $d(m'')$  is normalised such that

$$p_{d(m'')=0|\mathbf{x}}(d(m'') = 0, \mathbf{x}) + p_{d(m'')=1|\mathbf{x}}(d(m'') = 1, \mathbf{x}) = 1 \quad (2.20)$$

and similarly for the pmf of  $c(m')$ . The gain in information between the received values and the APP on code bits is quantified as *extrinsic* information (EXT) probability. For an AWGN channel the *extrinsic* probability is given by

$$\text{EXT}\{c(m')\} \propto \text{APP}\{c(m')\} \exp\left(\frac{(x(m') - \mu_x)^2}{2\sigma_x^2}\right) \quad (2.21)$$

where  $\propto$  denotes “proportional up to a constant factor”. The MAP algorithm and some important aspects for its practical implementation are explained in Appendix D.



**Figure 2.9:** The soft-in soft-out (SISO) decoder.

### Soft Decision Data

After decoding, the *a-posteriori* and *extrinsic* probabilities are interleaved and used to produce a conditional soft mean estimate of the transmitted symbols, given the probabilities  $\text{APP}\{c(m')\}$  or  $\text{EXT}\{c(m')\}$ , respectively. With the definitions (2.19) and (2.21) the soft QPSK mapping for the APPs and the EXTs is given by

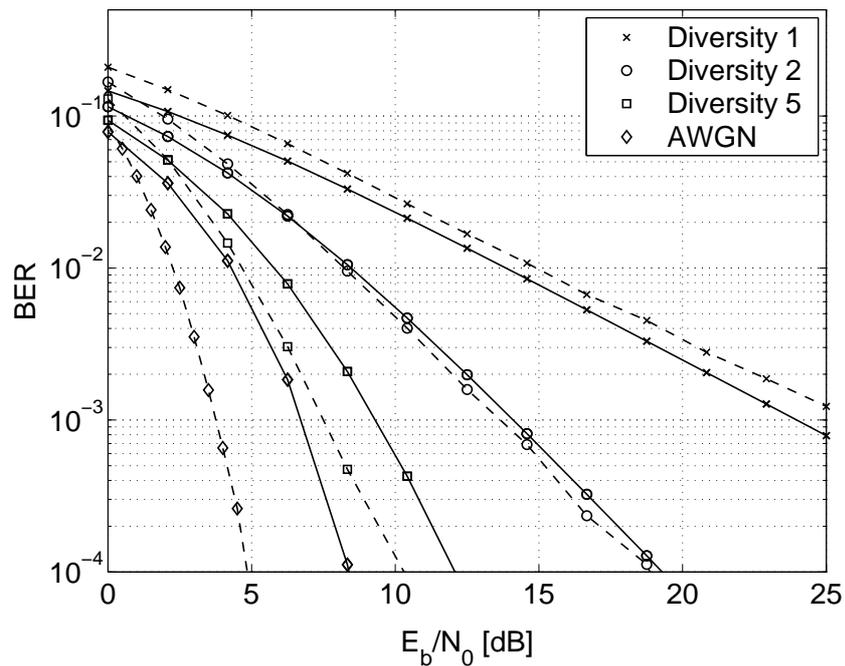
$$\begin{aligned} \hat{b}(m) &= [+1 \times \text{APP}\{c(2m)\} - 1 \times (1 - \text{APP}\{c(2m)\}) \\ &= +j(+1 \times \text{APP}\{c(2m+1)\} - 1 \times (1 - \text{APP}\{c(2m+1)\})) / \sqrt{2} \\ &= [2 \text{APP}\{c(2m)\} - 1 + j(2 \text{APP}\{c(2m+1)\} - 1)] / \sqrt{2}, \end{aligned} \quad (2.22)$$

$$\tilde{b}(m) = [2 \text{EXT}\{c(2m)\} - 1 + j(2 \text{EXT}\{c(2m+1)\} - 1)] / \sqrt{2}, \quad (2.23)$$

with  $m \in \{J, \dots, M-1\}$ .

### Impact of Fading

In this section we illuminate the impact of fading on coding. We assume that the receive amplitude is fading with  $L$  diversity branches that can all be maximum ratio combined. For such a channel there exist upper bounds on the coded bit error rate. The bounds depend on the free distance of the code. However, no exact analytical descriptions are known [Pro00]. Hence, we will study the impact of diversity on a coded system by numerical means. For this, we consider a four-state convolutional code with rate  $R_C = 1/2$  and generator polynomials  $(5, 7)_8$ . Data is transmitted in blocks with  $M = 160$  QPSK symbols that are Gray coded. The fading coefficients stay constant for the duration of a block. In Fig. 2.10 we plot the bit error rate curves for diversity orders  $L \in \{1, 2, 5\}$  for the uncoded and the coded system. Uncoded transmission is indicated by solid lines while coded transmission is marked by dashed lines. The uncoded curves were obtained through the analytical expression (2.4) and those of the coded system by Monte-Carlo simulations over 10.000 blocks. We learn that for a target bit error rate of  $10^{-3}$  the gain due to coding is strongest for the AWGN channel and that it becomes less with decreasing diversity order. For  $L = 1$  there is no gain, but rather a loss of 1 dB. For  $L = 2$  the gain is roughly 0,5 dB, and finally for the AWGN channel, *i.e.*,  $L \rightarrow \infty$ , we achieve a gain of 3,1 dB. When codes are used in an iterative multi-user receiver, they have the purpose to accelerate the



**Figure 2.10:** Impact of diversity combining on bit error rate for coded and uncoded transmission. Symbols are QPSK modulated with Gray mapping. For the coded case we apply a convolutional code with rate  $R_C = 1/2$  and generator polynomials  $(5, 7)_8$ . Dashed lines (--) refer to coded BERs and solid lines (—) to uncoded BERs.

convergence towards lower bit error rates by combating multiple-access interference. For the extreme case with  $L = 1$  the only purpose of codes is to combat multiple-access interference but there is no gain with respect to the channel over the uncoded system.



# 3 Iterative Multi-User Detection

This chapter is devoted to the development and analysis of detectors for the iterative multi-user receiver. Particular emphasis is directed to the processing of inter-symbol interference which plays an important role in medium data rate communications. Throughout, we assume to have perfect channel knowledge available - the aspect of channel estimation will be illuminated in Chapter 4. We start off with a review of optimum and non-iterative multi-user detection techniques for single symbol processing. Note that the optimum receiver described in Section 2.7 searches for a codeword of length  $M$ . In the iterative receiver, the multi-user detector does not have any knowledge of the code constraints and the single-user decoders do not know the waveform of the signals. Hence, the multi-user detector performs single-symbol processing. We derive detectors based on interference cancellation and discuss an analysis method for AWGN channels based on density evolution. This analysis will be extended to include channel estimation in Chapter 6. The subsequent sections in the present chapter will investigate the impact of transmission and receiver features on the achievable bit error rates and system capacities. The studied features comprise inter-symbol interference processing, the use of feedback information, detector filters, interference cancellation scheduling, user power distributions, multiple receive antennas, and code constraint lengths.

## 3.1 Overview

Until around 1984 it was believed that *multiple-access interference* (MAI) in a multi-user communication system is best treated as white noise. If this was true the *single-user matched filter* (SUMF) would be that receiver filter that maximises the output signal to noise ratio. The seminal work of Verdú in 1984 [Ver86] brought forth that multiple-access interference is structured and is hence not white, but coloured. Verdú derived the optimum detector which shows to be NP-complete, meaning that the computational effort cannot be given as a polynomial expression of the number of users. This was the starting point for the area of *multi-user detection* (MUD) that deals with suboptimum receiver structures having finite computational complexity. Excellent overviews on MUD techniques are given in [Mos96, Ver98, Ras03a].

Multi-user detection is obsolete if spreading sequences were perfectly orthogonal

to each other. Due to multi-path propagation on the wireless link and due to the use of scrambling codes in systems like UMTS, orthogonality cannot be preserved. Furthermore, the number of orthogonal sequences is strongly limited by the length of the spreading sequences. For high capacity systems, that operate in the overloaded mode, there are more users present than orthogonal sequences are available. We show that the use of non-orthogonal signalling and coding allows for very low bit error rates when iterative receivers are employed.

## 3.2 Optimum Multi-User Detection

Let us consider the multi-user system model defined in (2.11) for the particular case of an AWGN channel and all users having power  $\|\mathbf{s}_k\|^2 = \mathbf{s}_k^H \mathbf{s}_k = 1$ . Omitting the symbol index  $m$  we can conceptually write for the chip level

$$\mathbf{y} = \mathbf{S}\mathbf{b} + \mathbf{v}$$

with  $\mathbf{S} \in \mathbb{C}^{N \times K}$  the spreading matrix and  $\mathbf{v} \in \mathbb{C}^{N \times 1}$  the additive white Gaussian noise with covariance matrix  $\sigma_v^2 \mathbf{I}_N$ . The statistics on the symbol vector  $\mathbf{b} \in \mathbb{X}^{K \times 1}$  contained in  $\mathbf{y}$  do not change if the observed vector  $\mathbf{y}$  is passed through a single-user matched filter. Then, we obtain a model that is mapped into the symbol space as

$$\mathbf{r} = \mathbf{S}^H \mathbf{y} = \mathbf{R}\mathbf{b} + \mathbf{z} \quad (3.1)$$

with  $\mathbf{R} \triangleq \mathbf{S}^H \mathbf{S} \in \mathbb{C}^{K \times K}$  denoting the cross-correlation matrix and  $\mathbf{z}$  is a  $K$ -dimensional, zero-mean complex Gaussian random vector with covariance matrix  $\sigma_v^2 \mathbf{R}$ . Hence, the output  $\mathbf{r}$  forms a sufficient statistic for the estimation of the symbols  $\mathbf{b}$  of all users. The conditional pdf of  $\mathbf{r}$  given the symbol vector  $\mathbf{b}$  is a Gaussian multi-variate distribution expressed as [Ver86]

$$f_{\mathbf{r}|\mathbf{b}}(\mathbf{r}, \mathbf{b}) = \frac{1}{(\pi\sigma_v^2)^K \det \mathbf{R}} \exp \left( -\frac{(\mathbf{r} - \mathbf{R}\mathbf{b})^H \mathbf{R}^{-1} (\mathbf{r} - \mathbf{R}\mathbf{b})}{\sigma_v^2} \right). \quad (3.2)$$

The optimum symbol detector, in the sense the bit error probability is minimised, is a *maximum a-posteriori* (MAP) detector. It can be formulated for either each user *individually*

$$\hat{b}_k = \operatorname{argmax}_{\tilde{b}_k \in \mathbb{X}} f_{b_k|\mathbf{r}}(\tilde{b}_k, \mathbf{r}),$$

or for all users *jointly*

$$\hat{\mathbf{b}} = \operatorname{argmax}_{\tilde{\mathbf{b}} \in \mathbb{X}^{K \times 1}} f_{\mathbf{b}|\mathbf{r}}(\tilde{\mathbf{b}}, \mathbf{r}).$$

The bit error probability of the individually optimum detector is always lower or equal the one of the jointly optimum detector. The possible symbols  $b_k$  can take

on values from the set  $\mathbb{X}$ . The symbol vector of all users is  $\tilde{\mathbf{b}} \in \mathbb{X}^{K \times 1}$ . Under the assumption that all symbol vectors appear equally likely, *i.e.*,  $p(\tilde{\mathbf{b}}) = 4^{-K}$ , the MAP detector reduces to a *maximum likelihood* (ML) detector via Bayes' law. The individual ML estimate is

$$\hat{b}_k = \operatorname{argmax}_{b_k \in \mathbb{X}} \left\{ \sum_{\tilde{\mathbf{b}} \in \mathbb{X}^{K \times 1}: \tilde{b}_k = b_k} f_{r|\mathbf{b}}(\tilde{\mathbf{b}}, \mathbf{r}) \right\} \quad (3.3)$$

and the joint ML estimate of the data vector reads

$$\hat{\mathbf{b}} = \operatorname{argmax}_{\tilde{\mathbf{b}} \in \mathbb{X}^{K \times 1}} f_{r|\mathbf{b}}(\tilde{\mathbf{b}}, \mathbf{r}) = \operatorname{argmin}_{\tilde{\mathbf{b}} \in \mathbb{X}^{K \times 1}} \left\{ (\mathbf{r} - \mathbf{R}\tilde{\mathbf{b}})^{\mathrm{H}} \mathbf{R}^{-1} (\mathbf{r} - \mathbf{R}\tilde{\mathbf{b}}) \right\}.$$

Latter is equal to the minimisation of the quadratic form

$$\hat{\mathbf{b}} = \operatorname{argmin}_{\tilde{\mathbf{b}} \in \mathbb{X}^{K \times 1}} \left\{ \tilde{\mathbf{b}}^{\mathrm{H}} \mathbf{R} \tilde{\mathbf{b}} - 2\Re \left\{ \tilde{\mathbf{b}}^{\mathrm{H}} \mathbf{r} \right\} \right\}. \quad (3.4)$$

The individual ML detector (3.3) involves the evaluation of  $4^K$  conditional likelihood functions (3.2) per user. The joint ML solution requires an exhaustive search over  $4^K$  symbol hypotheses. Therefore, the complexity of both detectors is exponential in the number of users  $\mathcal{O}(4^K)$ . This is the reason why they have limited practical importance, *i.e.*, only in systems with few users.

### 3.3 Linear Multi-User Detection

This section reviews the most prominent linear detectors for an AWGN multiple-access channel. The canonical model presented in (3.1) is used. The filter is applied onto the chip-matched receive vector and assigned to a decision region like

$$\hat{\mathbf{z}} = \operatorname{quant}_{\mathbb{X}^K} \{ \mathbf{L}^{\mathrm{H}} \mathbf{r} \}.$$

All succeeding linear operations work with the sufficient statistics of  $\mathbf{r}$ . In this sense the *single-user matched filter* (SUMF) detector becomes

$$\mathbf{L} = \mathbf{I}_K \quad (3.5)$$

that treats all user interference as white noise. A linear filter that exploits the cross-correlation is the *zero-forcer* (ZF) also known as *decorrelating* detector. It simply neglects the noise and minimises the multiple-access interference to zero. We obtain

$$\mathbf{L} = \mathbf{R}^{-1}.$$

The decorrelator has the disadvantage that it amplifies noise and in low SNR regions it performs even worse than the SUMF. A compromise in treating noise and MAI is achieved by the *LMMSE* detector that has its origin in signal estimation. It reads

$$\mathbf{L} = (\mathbf{R} + \sigma_v^2 \mathbf{I}_K)^{-1} \quad (3.6)$$

and performs for low SNRs like the matched filter solution, while for high SNRs the LMMSE solution converges to the decorrelator performance. The performance of linear detectors is deteriorated in situations with unequal receive power levels caused by, *e.g.*, the near-far effect [Ver98]. It was reported in [Tse99] that the equal power distribution maximises the spectral efficiency for the SUMF, the decorrelator, and the LMMSE detector. For the decorrelator this result was already reported in [Lup89]. This is the main reason why commercial communication systems exercise power control [Sch04a]. However, we show in the subsequent sections that unequal power distributions are actually of benefit for interference-cancellation based detection schemes.

### 3.4 Parallel Interference Cancellation (PIC)

We map symbols obtained in previous iterations to soft information  $\tilde{b}_k(m)$  in order to get an estimate of the total *multiple-access interference* (MAI) symbolised by  $\tilde{\mathbf{y}}(m)$ . This estimate is used to lower the effective interference and is subtracted from the receive vector  $\mathbf{y}(m)$  prior to symbol detection. In mathematical notation interference cancellation for a general frequency selective channel (2.12) with short spreading sequences, *i.e.*,  $\mathbf{s}_k(m) = \mathbf{s}_k$ , is expressed as

$$\tilde{\mathbf{y}}_k^{(i)}(m) = \mathbf{y}(m) - \sum_{q=-1}^1 \tilde{\mathbf{S}}_q \tilde{\mathbf{b}}^{(i-1)}(m-q) + \tilde{\mathbf{s}}_k \tilde{b}_k^{(i-1)}(m). \quad (3.7)$$

The index  $(\cdot)^{(i)}$  designates iterations  $i > 0$ . After interference cancellation we apply a linear filter  $\mathbf{f}_k^{(i)}(m)$  to the vector  $\tilde{\mathbf{y}}_k^{(i)}(m)$  to obtain an estimate of the transmitted symbol in the form  $\hat{z}_k^{(i)}(m) = \left(\mathbf{f}_k^{(i)}(m)\right)^H \tilde{\mathbf{y}}_k^{(i)}(m)$ . In the next two paragraphs we develop linear filters that are suitable for this task.

#### 3.4.1 Single-User Matched Filter (SUMF)

The simplest linear detector for retrieving  $k$ -th user's information is the single-user matched filter whose elements are given by

$$f_k[n] \triangleq \check{s}_k[n] = \sum_{l=0}^{L-1} s_k[n-l] h_k[l], \quad n \in \{0, 1, \dots, N+L-2\}. \quad (3.8)$$

The filter reads as vector  $\mathbf{f}_k = [f_k[0], f_k[1], \dots, f_k[N + L - 2]]^T$ . For a short code CDMA system and a block fading environment the time index  $m$  and iteration index  $i$  can be omitted, since the filter stays constant for all symbols and all iterations.

### Matched and Mismatched Detectors

Precise *channel state information* (CSI) is practically never available at the receiver. It is rather obtained by estimation and associated with an error. Instead of the true channel  $\mathbf{h}$ , its estimate  $\hat{\mathbf{h}}$  is used in the detector. Such a detector is called a *mismatched detector* in contrast to a *matched detector*, that has perfect CSI available. The mismatched detector requires the error covariance matrix of the channel taps. The error variance is only available, if the power delay profile and the statistical dependencies of the channel taps are known to the receiver. Since this is a strong requirement, we will use an approximation by neglecting the term including the error covariance matrix. Under this approximation, we can use the expressions for the matched detectors and replace the virtual by the effective spreading sequences. Subsequently, we restrict ourselves to the treatment of the detector filters in the matched case.

### Filter Output Distribution

The filter output  $z_k^{(i)}(m) = \mathbf{f}_k^H(m) \tilde{\mathbf{y}}_k^{(i)}(m)$  can be modelled by

$$z_k^{(i)}(m) = \mu_k^{(i)}(m) b_k(m) + \nu_k^{(i)}(m) \quad (3.9)$$

as suggested in [Wan99]. The residual noise and multiple-access interference is jointly modelled as spherically invariant complex Gaussian process with  $\nu_k^{(i)}(m) \sim \mathcal{CN}\left(0, (\sigma_k^2(m))^{(i)}\right)$ . The resulting *signal to interference and noise ratio* (SINR) reads

$$\gamma_k^{(i)}(m) = \frac{|\mu_k^{(i)}(m)|^2}{(\sigma_k^2(m))^{(i)}}. \quad (3.10)$$

The upper index  $(\cdot)^{(i)}$  underlines that the statistical output values depend on the iteration  $i$ . In the subsequent treatment of this section we will omit the iteration index for notational convenience.

### SINR of PIC-SUMF Detector in Flat Fading

First, let us consider the particular case of a frequency flat channel. The output of the matched filter is

$$z_k(m) = |h_k|^2 b_k(m) + \check{\mathbf{s}}_k^H \left( \check{\mathbf{S}} \mathbf{b}(m) - \check{\mathbf{s}}_k b_k(m) - \check{\mathbf{S}} \tilde{\mathbf{b}}(m) - \check{\mathbf{s}}_k \tilde{b}_k(m) + \mathbf{v}(m) \right). \quad (3.11)$$

The conditional mean of the output symbol  $z_k(m)$  (3.11) evaluates to

$$\mu_k(m) = \mathbb{E}_{\mathbf{v}} \{b_k^*(m)z_k(m)\} = |h_k|^2$$

and the variance of the multiple access interference and noise reads

$$\begin{aligned} \sigma_k^2(m) &= \mathbb{E}_{\mathbf{v}} \{\nu_k(m)\nu_k^*(m)\} \\ &= \mathbf{s}_k^H \left( \check{\mathbf{S}}\mathbf{V}(m)\check{\mathbf{S}}^H - \left(1 - |\tilde{b}_k(m)|^2\right) \check{\mathbf{s}}_k\check{\mathbf{s}}_k^H + \sigma_v^2\mathbf{I}_N \right) \mathbf{s}_k. \end{aligned}$$

The covariance matrix of the soft decision symbol errors, or equivalently the residual power matrix,  $\mathbf{V}(m) \in \mathbb{C}^{K \times K}$  is defined as

$$\mathbf{V}(m) \triangleq \text{diag} \left( 1 - |\tilde{b}_1(m)|^2, 1 - |\tilde{b}_2(m)|^2, \dots, 1 - |\tilde{b}_K(m)|^2 \right) \quad (3.12)$$

and the corresponding SINR is given through

$$\gamma_k(m) = \frac{|h_k|^4}{\check{\mathbf{s}}_k^H \left( \check{\mathbf{S}}\mathbf{V}(m)\check{\mathbf{S}}^H - \left(1 - |\tilde{b}_k(m)|^2\right) \check{\mathbf{s}}_k\check{\mathbf{s}}_k^H + \sigma_v^2\mathbf{I}_N \right) \check{\mathbf{s}}_k}. \quad (3.13)$$

### SINR of PIC-SUMF Detector in Frequency Selective Fading

The extension to the multi-path case is straightforward. The desired signal power is determined through the square of

$$\mu_k(m) = \|\mathbf{h}_k\|^2$$

and the variance is increased by the contributions of pre- and post-cursor ISI terms along

$$\sigma_k^2(m) = \check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q\mathbf{V}(m-q)\check{\mathbf{S}}_q^H - \left(1 - |\tilde{b}_k(m)|^2\right) \check{\mathbf{s}}_k\check{\mathbf{s}}_k^H + \sigma_v^2\mathbf{I}_{N+L-1} \right) \check{\mathbf{s}}_k.$$

We highlight two particular cases for the SINR:

- $\tilde{\mathbf{B}}(m) = \text{diag}(\mathbf{0}_K) \forall m$ : The feedback symbols are all unknown. This is the case in the first iteration. The symbol variance matrices become diagonal matrices  $\mathbf{V}(m-q) = \mathbf{I}_K$  and the SUMF output SINR becomes that of a conventional SUMF detector (3.5) without feedback

$$\gamma_k(m) = \frac{\|\mathbf{h}_k\|^4}{\check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q\check{\mathbf{S}}_q^H - \check{\mathbf{s}}_k\check{\mathbf{s}}_k^H + \sigma_v^2\mathbf{I}_{N+L-1} \right) \check{\mathbf{s}}_k}.$$

- $\tilde{\mathbf{B}}(m) = \text{diag}(\mathbf{b}(m)) \forall m$ : This happens when all feedback symbols correspond to the actually transmitted symbols. The SINR becomes that of a single-user system and it reads

$$\gamma_k(m) = \frac{\|\mathbf{h}_k\|^2}{\sigma_v^2}.$$

### 3.4.2 Linear MMSE Filter (LMMSE)

The second filter which is suitable after parallel interference cancellation, is the linear minimum mean square error filter. We discuss first the flat fading case and explore then the properties in the frequency selective case.

#### PIC-LMMSE Detector for Flat Fading

The linear minimum mean square error (LMMSE) filter for the interference cancelled observation vector  $\tilde{\mathbf{y}}_k(m)$  (3.7) and the flat fading CDMA system model presented in Paragraph 2.5.1 is the solution to the Wiener-Hopf equations [Hay91]

$$\mathbf{f}_k^H(m) = \underbrace{\mathbb{E}_{b,v} \{b_k(m)\tilde{\mathbf{y}}_k^H(m)\}}_1 \left( \underbrace{\mathbb{E}_{b,v} \{\tilde{\mathbf{y}}_k(m)\tilde{\mathbf{y}}_k^H(m)\}}_2 \right)^{-1}. \quad (3.14)$$

In the derivation of the LMMSE filter, expectation over  $b$  is understood as conditional expectation, given the *extrinsic* decoder output probability  $\text{EXT}\{c(m')\}$  defined in (2.21). Let us elaborate on the two expectation terms separately.

1. The first term in (3.14) equates to

$$\mathbb{E}_{b,v} \{b_k(m)\tilde{\mathbf{y}}_k^H(m)\} = \check{\mathbf{s}}_k^H.$$

2. The second term expands to

$$\begin{aligned} \mathbb{E}_{b,v} \{\tilde{\mathbf{y}}_k(m)\tilde{\mathbf{y}}_k^H(m)\} &= \check{\mathbf{S}} \mathbb{E}_{b,v} \left\{ \left( \mathbf{b}(m) - \tilde{\mathbf{b}}(m) + \tilde{b}_k(m)\mathbf{e}_k \right) \times \right. \\ &\quad \left. \left( \mathbf{b}(m) - \tilde{\mathbf{b}}(m) + \tilde{b}_k(m)\mathbf{e}_k \right)^H \right\} \check{\mathbf{S}}^H + \sigma_v^2 \mathbf{I}_N \\ &= \check{\mathbf{S}}\mathbf{V}(m)\check{\mathbf{S}}^H + \left| \tilde{b}_k(m) \right|^2 \check{\mathbf{s}}_k\check{\mathbf{s}}_k^H + \sigma_v^2 \mathbf{I}_N \end{aligned}$$

where  $\mathbf{e}_k \triangleq [0, \dots, 0, 1, 0, \dots, 0]^T$  denotes the standard basis unit vector with a one at the  $k$ -th element and zero elsewhere.

Let us define the Hermitian matrix  $\mathbf{A} \in \mathbb{C}^{N \times N}$  as

$$\mathbf{A} \triangleq \check{\mathbf{S}}\mathbf{V}(m)\check{\mathbf{S}}^H + \left| \tilde{b}_k(m) \right|^2 \check{\mathbf{s}}_k \check{\mathbf{s}}_k^H + \sigma_v^2 \mathbf{I}_N.$$

The matrix  $\mathbf{V}(m)$  (3.12) is a covariance matrix that is always positive semi-definite [Hor99]. Multiplying the matrix with  $\check{\mathbf{S}}$  from the left and  $\check{\mathbf{S}}^H$  from the right leaves the product positive semi-definite. Hence, it has non-negative eigenvalues. Since the second expression has non-negative eigenvalues, and the last matrix  $N$  non-zero eigenvalues, the sum has only positive eigenvalues. Hence, the matrix is invertible. The LMMSE filter yields

$$\mathbf{f}_k^H(m) = \check{\mathbf{s}}_k^H \mathbf{A}^{-1} = \check{\mathbf{s}}_k^H \left( \check{\mathbf{S}}\mathbf{V}(m)\check{\mathbf{S}}^H + \left| \tilde{b}_k(m) \right|^2 \check{\mathbf{s}}_k \check{\mathbf{s}}_k^H + \sigma_v^2 \mathbf{I}_N \right)^{-1}. \quad (3.15)$$

The filter output distribution conditioned on  $b_k(m)$  is modelled as spherically invariant complex Gaussian process like in (3.9). The conditional mean of the filter output is given by

$$\begin{aligned} \mu_k(m) &= \mathbb{E}_{b,v} \{ b_k^*(m) z_k(m) \} \\ &= \mathbb{E}_{b,v} \{ b_k^* \mathbf{f}_k^H(m) \tilde{\mathbf{y}}_k(m) \} \\ &= \mathbb{E}_{b,v} \left\{ b_k^*(m) \mathbf{f}_k^H(m) \left( \check{\mathbf{S}}\mathbf{B}(m) - \check{\mathbf{S}}\tilde{\mathbf{B}}(m) + \check{\mathbf{s}}_k \tilde{b}_k(m) + \mathbf{v}(m) \right) \right\} \\ &= \mathbf{f}_k^H(m) \check{\mathbf{s}}_k \\ &= \check{\mathbf{s}}_k^H \left( \check{\mathbf{S}}\mathbf{V}(m)\check{\mathbf{S}}^H + \left| \tilde{b}_k(m) \right|^2 \check{\mathbf{s}}_k \check{\mathbf{s}}_k^H + \sigma_v^2 \mathbf{I}_N \right)^{-1} \check{\mathbf{s}}_k \end{aligned} \quad (3.16)$$

and its conditional variance computes as

$$\begin{aligned} \sigma_k^2(m) &= \mathbb{E}_{b,v} \left\{ \left| z_k(m) - \mathbb{E}_{b,v} \{ z_k(m) \} \right|^2 \right\} \\ &= \mathbb{E}_{b,v} \left\{ z_k(m) z_k^*(m) - 2\Re \{ z_k(m) b_k^*(m) \mu_k^*(m) \} + b_k(m) \mu_k(m) \mu_k^*(m) b_k^*(m) \right\} \\ &= \mathbb{E}_{b,v} \left\{ z_k(m) z_k^*(m) \right\} - \mu_k^2(m). \end{aligned}$$

The variance evaluates to

$$\begin{aligned} \sigma_k^2(m) &= \mathbf{f}_k^H(m) \left( \check{\mathbf{S}}\mathbf{V}(m)\check{\mathbf{S}}^H + \left| \tilde{b}_k(m) \right|^2 \check{\mathbf{s}}_k \check{\mathbf{s}}_k^H + \sigma_v^2 \mathbf{I}_N \right) \mathbf{f}_k(m) - \mu_k^2(m) \\ &= \check{\mathbf{s}}_k^H \left( \check{\mathbf{S}}\mathbf{V}(m)\check{\mathbf{S}}^H + \left| \tilde{b}_k(m) \right|^2 \check{\mathbf{s}}_k \check{\mathbf{s}}_k^H + \sigma_v^2 \mathbf{I}_N \right)^{-1} \end{aligned}$$

$$\begin{aligned}
& \times \left( \check{\mathbf{S}}\mathbf{V}(m)\check{\mathbf{S}}^H + |\tilde{b}_k(m)|^2 \check{\mathbf{s}}_k\check{\mathbf{s}}_k^H + \sigma_v^2\mathbf{I}_N \right) \mathbf{f}_k(m) - \mu_k^2(m) \\
& = \check{\mathbf{s}}_k^H \mathbf{f}_k(m) - \mu_k^2(m) \\
& = \mu_k(m)(1 - \mu_k(m)). \tag{3.17}
\end{aligned}$$

### PIC-LMMSE Detector for Frequency Selective Fading

In the multi-path case, the LMMSE filter is computed as in (3.14). The computation assumes that adjacent symbols of a particular user are statistically independent, *i.e.*,  $\mathbb{E}\{b_k(m)b_k^*(n)\} = 0, \forall n \neq m$ . This is fulfilled when a random interleaver with infinite length is used. The LMMSE filter that deals with ISI reads

$$\mathbf{f}_k^H(m) = \check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q \mathbf{V}(m-q) \check{\mathbf{S}}_q^H + |\tilde{b}_k(m)|^2 \check{\mathbf{s}}_k\check{\mathbf{s}}_k^H + \sigma_v^2\mathbf{I}_{N+L-1} \right)^{-1}. \tag{3.18}$$

With the definition of the SINR in (3.10) and the expressions (3.16) for the mean and the variance (3.17), we obtain  $\gamma_k(m) = \mu_k(m)/(1 - \mu_k(m))$  and the following special cases emerge:

- $\tilde{\mathbf{B}}(m) = \text{diag}(\mathbf{0}_K) \forall m$ : In this case the symbol variance matrices become diagonal matrices  $\mathbf{V}(m) = \mathbf{I}_K$  and the LMMSE detector becomes the conventional LMMSE multi-user detector, similar to (3.6) for the AWGN channel,

$$\mathbf{f}_k^H(m) = \check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q \check{\mathbf{S}}_q^H + \sigma_v^2\mathbf{I}_{N+L-1} \right)^{-1}.$$

The corresponding SINR reads

$$\gamma_k(m) = \frac{\check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q \check{\mathbf{S}}_q^H + \sigma_v^2\mathbf{I}_{N+L-1} \right)^{-1} \check{\mathbf{s}}_k}{1 - \check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q \check{\mathbf{S}}_q^H + \sigma_v^2\mathbf{I}_{N+L-1} \right)^{-1} \check{\mathbf{s}}_k}.$$

- $\tilde{\mathbf{B}}(m) = \mathbf{B}(m) \forall m$ : When the feedback symbols are completely known, the mean  $\mu_k(m)$  becomes

$$\mu_k(m) = \check{\mathbf{s}}_k^H (\check{\mathbf{s}}_k\check{\mathbf{s}}_k^H + \sigma_v^2)^{-1} \check{\mathbf{s}}_k = \frac{\|\mathbf{h}_k\|^2}{\|\mathbf{h}_k\|^2 + \sigma_v^2}$$

and the corresponding SINR is given through

$$\gamma_k(m) = \frac{\|\mathbf{h}_k\|^2}{\sigma_v^2}.$$

### Low Complexity Implementation of the PIC-LMMSE Detector

In general the conditional mean  $\mu_k(m)$  of the output of the LMMSE filter (3.16) is not equal to one. This implies that the output symbols (3.9) are biased. Dividing the filter by  $\mu_k(m)$  yields the *unbiased LMMSE filter*

$$\mathbf{f}'_k{}^H(m) = \frac{\check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q \mathbf{V}(m-q) \check{\mathbf{S}}_q^H + |\tilde{b}_k(m)|^2 \check{\mathbf{s}}_k \check{\mathbf{s}}_k^H + \sigma_v^2 \mathbf{I}_{N+L-1} \right)^{-1}}{\check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q \mathbf{V}(m-q) \check{\mathbf{S}}_q^H + |\tilde{b}_k(m)|^2 \check{\mathbf{s}}_k \check{\mathbf{s}}_k^H + \sigma_v^2 \mathbf{I}_{N+L-1} \right)^{-1} \check{\mathbf{s}}_k}. \quad (3.19)$$

In Appendix E it is shown that, by applying the matrix inversion lemma (G.2), expression (3.19) can be translated into the simple expression

$$\mathbf{f}'_k{}^H(m) = \frac{\check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q \mathbf{V}(m-q) \check{\mathbf{S}}_q^H + \sigma_v^2 \mathbf{I}_{N+L-1} \right)^{-1}}{\check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q \mathbf{V}(m-q) \check{\mathbf{S}}_q^H + \sigma_v^2 \mathbf{I}_{N+L-1} \right)^{-1} \check{\mathbf{s}}_k}. \quad (3.20)$$

Comparing (3.18) and (3.20) shows that making the filter unbiased, causes a reduction in complexity. The inverse of (3.18) does not need to be computed for every user any longer but can be pre-computed for all users at a particular iteration. The filter output SINR does not change due to this scaling. The expression for the unbiased LMMSE filter was first used in [Mül02] for an AWGN channel and for multi-path channels causing inter-symbol interference in [Weh02].

A further reduction in complexity can be achieved when the covariance matrix of the soft symbol decision  $\mathbf{V}(m)$  is computed as empirical average over all symbols instead of each symbol individually:

$$[\mathbf{V}]_{k,k} = 1 - \frac{1}{M-J} \sum_{m=J}^{M-1} |\tilde{b}_k(m)|^2. \quad (3.21)$$

This idea was proposed in [Cai01a]. When the filter uses the average covariance matrix (3.21), the resulting filter is termed *unconditional*, otherwise, when the covariance matrix is computed for every symbol (3.12), it is called *conditional*. In the *conditional* case, the computation of the filter involves a matrix inversion for each symbol instance as it can be seen in (3.20). When the *unconditional* filter is used, an inverse of a matrix needs to be computed once for every user and every iteration

only. The *unconditional* version of (3.20) is expressed as

$$\mathbf{f}''_k^H = \frac{\check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q \mathbf{V} \check{\mathbf{S}}_q^H + \sigma_v^2 \mathbf{I}_{N+L-1} \right)^{-1}}{\check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q \mathbf{V} \check{\mathbf{S}}_q^H + \sigma_v^2 \mathbf{I}_{N+L-1} \right)^{-1} \check{\mathbf{s}}_k}. \quad (3.22)$$

### 3.5 Successive Interference Cancellation (SIC)

In parallel interference cancellation, that we have considered in the previous section, there is no ordering of the users. For a system where users are received with large differences in their power due to fading, parallel interference cancellation might not be the best strategy. In this case it may be better to detect those users first that are the strongest. Let us exemplify this for a two user scenario with one single iteration. If first the stronger user is detected, it would result in a good detection probability since the weak user does not cause much interference. If the correctly detected signal is subtracted from the received signal, the user with low power can also be detected reliably. The worst case scenario occurs when both have the same power.

The strategy described above is called *successive interference cancellation* (SIC). Without loss of generality, users are ordered according to their instantaneous gain  $\|\mathbf{h}_k\|^2$  such that the strongest user has index one, the second strongest index two and so forth. Then, SIC in the frequency selective case is formulated for  $k \in \{1, 2, \dots, K\}$  as

$$\tilde{\mathbf{y}}_k^{(i)}(m) = \mathbf{y}(m) - \sum_{q=-1}^1 \sum_{l=1}^{k-1} \check{\mathbf{s}}_{q,l}(m) \tilde{b}_l^{(i)}(m-q) - \sum_{q=-1}^1 \sum_{p=k+1}^K \check{\mathbf{s}}_{q,p}(m) \tilde{b}_p^{(i-1)}(m-q). \quad (3.23)$$

The first sum accounts for all  $k-1$  users that have already been detected in the current iteration  $i$  whereas the second sum accounts for those that still need to be detected. This is reflected in the upper index at the soft decision symbol  $\tilde{b}_l^{(i)}$ . Under the assumption that previously detected symbols are correct, it was shown in [Var97] that SIC can attain the capacity bound.

#### 3.5.1 Single-User Matched Filter (SUMF)

The single-user matched filter with perfect channel impulse response has length  $N+L-1$  and is defined as

$$f_k[n] \triangleq \check{s}_k[n] = \sum_{l=0}^{L-1} s_k[n-l] h_k[l].$$

The corresponding detector output SINR at iteration  $i$  is given through

$$\gamma_k^{(i)}(m) = \frac{\|\mathbf{h}_k\|^4}{\check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q \mathbf{V}_k^{(i)}(m-q) \check{\mathbf{S}}_q^H + |\tilde{b}_k(m)|^2 \check{\mathbf{s}}_k \check{\mathbf{s}}_k^H + \sigma_v^2 \mathbf{I}_{N+L-1} \right) \check{\mathbf{s}}_k - \|\mathbf{h}_k\|^4}$$

where the difference to the SINR of the PIC-SUMF (3.13) lies in the computation of the symbol covariance matrix  $\mathbf{V}_k^{(i)}(m)$ . Latter is now dependent on the scheduled user in the following way

$$\left[ \mathbf{V}_k^{(i)}(m) \right]_{l,l} = \begin{cases} 1 - |\tilde{b}_l^{(i)}(m)|^2 & \text{for } l \leq k, \\ 1 - |\tilde{b}_l^{(i-1)}(m)|^2 & \text{for } l > k. \end{cases} \quad (3.24)$$

### 3.5.2 Linear MMSE Filter (LMMSE)

Similarly to the LMMSE for the PIC detector, an unbiased and unconditional low-complexity filter can be developed for the SIC detector. The difference lies again in the definition of the symbol variance matrix which is caused by the successive interference cancellation. The unconditional symbol covariance matrix is computed along

$$\left[ \mathbf{V}_k^{(i)} \right]_{l,l} = \begin{cases} 1 - \frac{1}{M-J} \sum_{m=J}^{M-1} |\tilde{b}_l^{(i)}(m)|^2 & \text{for } l \leq k, \\ 1 - \frac{1}{M-J} \sum_{m=J}^{M-1} |\tilde{b}_l^{(i-1)}(m)|^2 & \text{for } l > k. \end{cases}$$

With this new definition of the symbol variance matrix we obtain the following expression for the (unbiased) unconditional filter in the frequency selective case:

$$\mathbf{f}''_k^H = \frac{\check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q \mathbf{V}_k^{(i)} \check{\mathbf{S}}_q^H + \sigma_v^2 \mathbf{I}_{N+L-1} \right)^{-1}}{\check{\mathbf{s}}_k^H \left( \sum_{q=-1}^1 \check{\mathbf{S}}_q \mathbf{V}_k^{(i)} \check{\mathbf{S}}_q^H + \sigma_v^2 \mathbf{I}_{N+L-1} \right)^{-1} \check{\mathbf{s}}_k}. \quad (3.25)$$

This filter needs to be computed for every user individually. Hence, the computational increase in complexity grows by a factor of  $K$  over the unbiased unconditional PIC-LMMSE filter (3.22).

### 3.6 Antenna Combining

In case of multiple receive antennas, the symbols  $z_{k,r}(m)$ ,  $m \in \{J, \dots, M-1\}$  that result from detection at the different antenna elements  $r$  are maximum ratio combined along (2.3) for a particular user  $k$  according to their SINRs  $\gamma_{k,r}$ . We have investigated two methods to obtain an estimate of the SINR. The first is obtained through the post-detection SINR  $\hat{\gamma}_{k,r}^{(1)} = \hat{\mu}_{k,r}^2 / \hat{\sigma}_{k,r}^2$  of the complex symbols after detection. We obtained this estimate through ‘‘Estimator 2’’ described in [Bea00]. The second method is related to the estimated channel  $\hat{\gamma}_{k,r}^{(2)} = \|\hat{\mathbf{h}}_{k,r}\|^2 / \sigma^2$  where  $\sigma^2$  is the variance of the noise and the multiple-access interference which is assumed approximately constant at all antennas. Hence, the latter combining only requires the values  $\hat{\mathbf{h}}_{k,r}$ . We have observed that the combining along the estimated channel coefficients leads to the best results. The combined output symbol for user  $k$  at antenna element  $r$  becomes

$$z_k(m) = \frac{\sum_{r=1}^{N_R} \|\hat{\mathbf{h}}_{k,r}\|^2 z_{k,r}(m)}{\sum_{r=1}^{N_R} \|\hat{\mathbf{h}}_{k,r}\|^2 \hat{\mu}_{k,r}}. \quad (3.26)$$

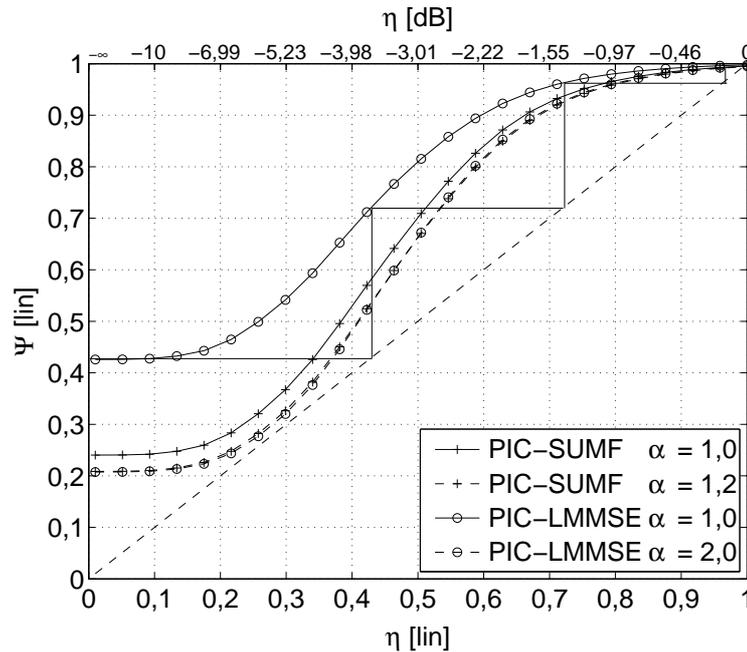
The scaling of the symbols streams  $z_{k,r}(m)$  by a factor of  $1/\hat{\mu}_{k,r}$  is required, since the mean of the symbols is different at the  $N_R$  antenna elements in general. Hence, the symbols are scaled by their estimated mean value

$$\hat{\mu}_{k,r} = \frac{1}{M-J} \sum_{m=J}^{M-1} |z_{k,r}(m)|$$

prior to combining.

### 3.7 Convergence Analysis

Insight into the convergence of iterative detection and decoding is required for learning their fundamental performance limits. An early method based on variance evolution is presented in [Ale98]. Recently, EXIT charts have become popular for analysing iterative systems [Bri01, Li03]. Particular importance in this work is devoted to the density evolution presented in [Bou02], where the dynamic behaviour of iterative detection and decoding is studied by means of a one-dimensional measure, the multi-user-efficiency  $\eta$  (see Paragraph 2.8.1 for the definition). In Chapter 6 we explain the details of this method and will extend it to include channel estimation. In the following, we discuss how the results can be interpreted and restrict ourselves



**Figure 3.1:** Evolution of the multi-user-efficiency  $\Psi(\eta)$  for a PIC-SUMF and a PIC-LMMSE detector on an AWGN channel with  $E_b/N_0 = 5$  dB, equal power users, and convolutional code  $(5, 7)_8$ .

to the case of iterative detection and decoding on an AWGN channel where all users have the same power. An exemplified evolution of the multi-user efficiency for a PIC-SUMF and a PIC-LMMSE without channel estimation is depicted in Fig. 3.1 for different loads  $\alpha$ . The  $E_b/N_0$  was chosen 5 dB. Let us consider the PIC-LMMSE curve for load  $\alpha = 1$ . Starting point is  $\eta = 0$  on the horizontal axis and on the vertical axis the multi-user efficiency  $\Psi$  after the next iteration is illustrated. The value  $\Psi$  becomes the starting point in the next iteration. Hence, it is projected onto the diagonal  $\Psi = \eta$  and a new multi-user efficiency value is computed for the new  $\eta$  value. In this way, a staircase is built up. In the illustrated case, the staircase leads to the right upper corner that indicates that all multiple-access interference has been mitigated. The shape depends on  $E_b/N_0$ , the system load  $\alpha$ , the detector type, and the received user power distribution. The number of vertical steps is the number of iterations needed to attain a particular multi-user efficiency. For the PIC-LMMSE filter the multi-user efficiency for a maximum load of roughly 2,0 is illustrated. For  $\eta = 0,3$  the gap between the multi-user efficiency curve and the diagonal  $\Psi = \eta$  is very narrow and it requires many iterations to pass through. But the analysis tells us that for the given parameters all multiple-access interference can be mitigated. If the curve intersected the diagonal  $\Psi = \eta$ , the convergence would come to an end prematurely. The second group of curves belongs to the PIC-SUMF detector, showing

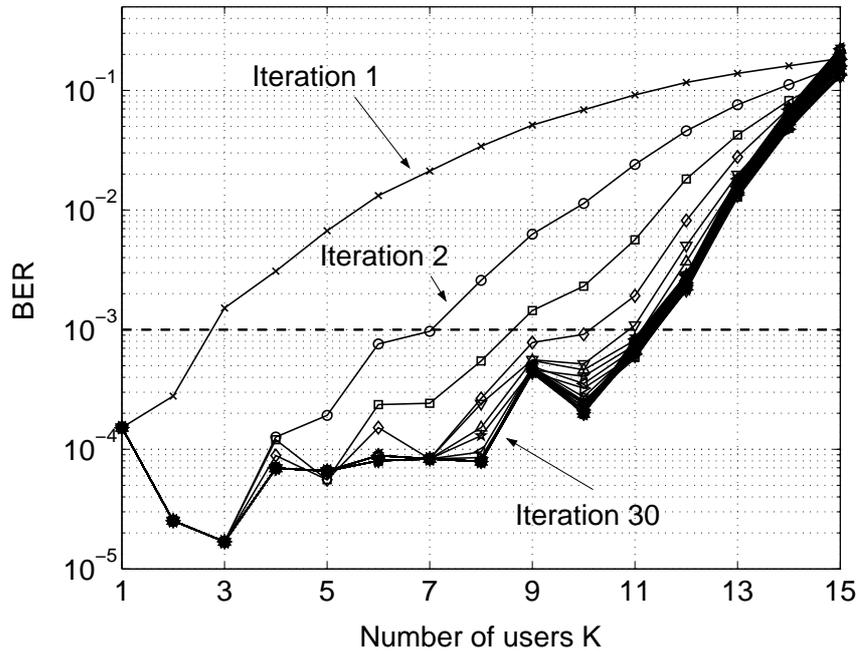
loads  $\alpha = 1,0$  and  $\alpha = 1,2$ . Compared to the PIC-LMMSE case the curve is closer to the diagonal  $\Psi = \eta$ . This means that for the given parameters, total MAI mitigation is still possible but at the expense of more iterations. The second curve shows a load of  $\alpha = 1,2$ . This is the maximal load that can be supported by a PIC-SUMF filter. Note that this case is practically identical to the PIC-LMMSE scenario for  $\alpha = 2,0$ . This means that the symbol SINR at the output of the two detectors is the same. The analysis shows us that a higher system capacity can be achieved by employing an LMMSE filter rather than a SUMF after parallel interference cancellation.

## 3.8 Simulation Results

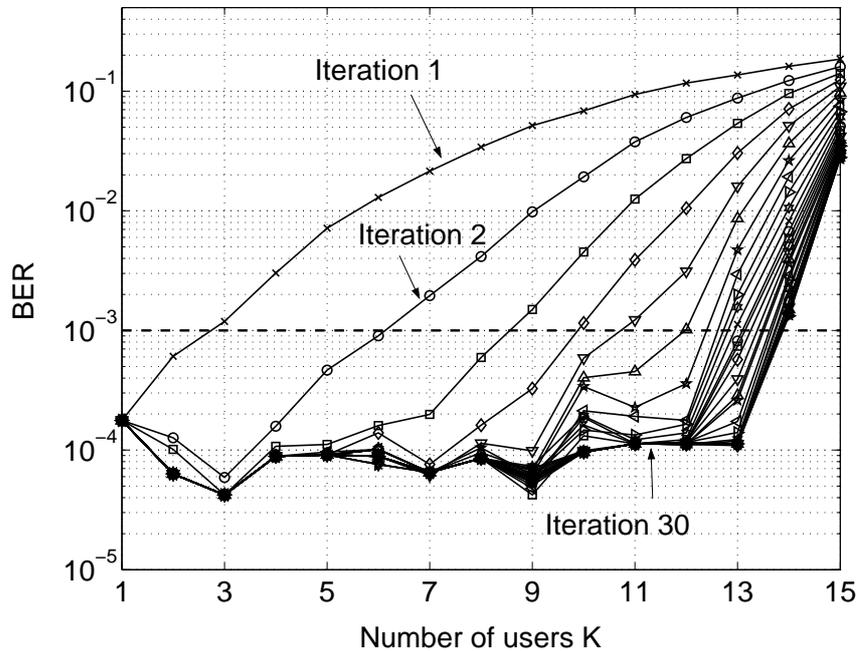
The assessment of the interference cancelling multi-user detection schemes is carried out via simulations. We focus on combinations of the parallel (3.7) and successive (3.23) interference cancellers with either a single-user matched filter or the unconditional unbiased LMMSE filters (3.22) and (3.25) for PIC and SIC, respectively. The biased LMMSE filter (3.18) and its conditional unbiased variant (3.20) are not considered due to prohibitive complexity. We consider receivers that have perfect channel knowledge available and therefore the schemes are matched, *i.e.*,  $\Sigma = \mathbf{0}_{KL}$  (see Paragraph 3.4.1 for the definition of matched detectors). Throughout the simulations we assume the following. The block length was chosen  $M = 160$  QPSK symbols and no pilots are interlaced as preamble. The number of 160 QPSK symbols is equivalent to 320 BPSK symbols in UMTS slot-processing with spreading factor  $N = 8$ , neglecting the control symbols [TS25.101]. The encoders are all non-recursive and non-systematic convolutional codes with code rate  $R_C = 1/2$  and, if not explicitly stated differently, four states. We define the *single-user bound* (SUB) as the BER for a single-user that has perfect channel knowledge in the receiver.

### 3.8.1 Feedback Information

In previously published papers on iterative detector and decoding structures there was quite some debate on which kind of decoder output information shall be used for interference cancellation. While early papers in the field like [Ale98, Moh98, ElGa00] use *a-posteriori* (APP) information, other contributions have shown that it is better to use *extrinsic* (EXT) information only. The earliest experimental observation was published in [Mar01], where the authors observe a bias in the residual interference due to the statistical dependencies of the feedback symbols and the observation vector. They state that with *extrinsic* information higher capacities are achievable. This was later shown analytically in [Bou02]. To resolve the reported differences we carry out an experiment that gives rise to an interesting conclusion. We compare



**Figure 3.2:** Bit error rate versus number of users illustrated for 30 iterations for a PIC-LMMSE detector with APP soft decision feedback on an AWGN channel with  $E_b/N_0 = 5$  dB,  $N = 8$ .



**Figure 3.3:** Bit error rate versus number of users illustrated for 30 iterations for a PIC-LMMSE detector with EXT soft decision feedback on an AWGN channel with  $E_b/N_0 = 5$  dB,  $N = 8$ .

the capacity of a PIC-LMMSE detector with APP feedback (2.22) and another one with EXT feedback (2.23). The considered channel is AWGN with  $E_b/N_0 = 5$  dB, a spreading factor of  $N = 8$ , and the number of iterations is limited to 30. The BER curves are drawn versus the number of users and they are illustrated in Fig. 3.2 for the APP feedback and in Fig. 3.3 for the EXT symbols. Taking the BER of  $10^{-3}$  as a criterion we observe that for  $K \leq 11$  the APP feedback causes a faster or equal convergence speed than the EXT feedback. While with  $N = 11$  users, a BER of  $10^{-3}$  is obtained after five iterations, it requires one more iteration for the EXT feedback. However, we observe that with EXT feedback a total of 14 users can be supported. This happens at the expense of more iterations. While the APP feedback leads to a clear stop in convergence, the EXT feedback is able to further exploit the information leading to higher capacities with more iterations. The minimum BER for eleven users after 30 iterations is  $6 \times 10^{-3}$  in the APP case and it is  $10^{-4}$  for the EXT case.

The early works dealing with iterative detection and decoding were investigating systems with  $\alpha < 1$ . From the view-point of the number of required iterations, it is well justified to utilise APP to achieve a particular BER. However, we have seen that this is not true for all loads. In particular, high loads can be only accommodated with EXT feedback.

For completeness, we mention that in [Mar01] a weighting of the feedback symbols was proposed that leads to a trade-off between the bias and the increased symbol variance, that is caused by the usage of EXT. This weighting is not considered in the present work. In the remaining parts of this thesis we deal with large loads  $\alpha$  and thus we will employ soft decisions that are retrieved from *extrinsic* probabilities via (2.23).

### 3.8.2 Inter-Symbol Interference

The multi-path wireless channel with  $L$  taps spreads the energy of a symbol  $b_k(m)$  over  $N + L - 1$  chip intervals, causing *inter-symbol interference* (ISI). To maximise the SNR of the signal at the output of the detector, all temporal signal terms need to be collected. This requires more processing in the computation of the detector filters and in detection itself. The question is, whether ISI processing, *i.e.*, gathering spread energy and mitigating MAI that is introduced by ISI, needs to be processed exactly along (3.7) and (3.23) for parallel and successive interference cancellation, respectively, or whether an approximation is sufficient? We consider an approximated processing that neglects energy contributions that are not covered by the  $N$  chips attributed to the  $m$ -th symbol. In this case, the observation vector  $\mathbf{y}(m)$  has length  $N$ . In case of parallel interference cancellation (3.7) the approximated

processing becomes

$$\tilde{\mathbf{y}}_k^{(i)}(m) = \mathbf{y}(m) - \check{\mathbf{S}}_a \tilde{\mathbf{b}}^{(i-1)}(m) + \check{\mathbf{s}}_{k,a} \tilde{b}_k^{(i-1)}(m).$$

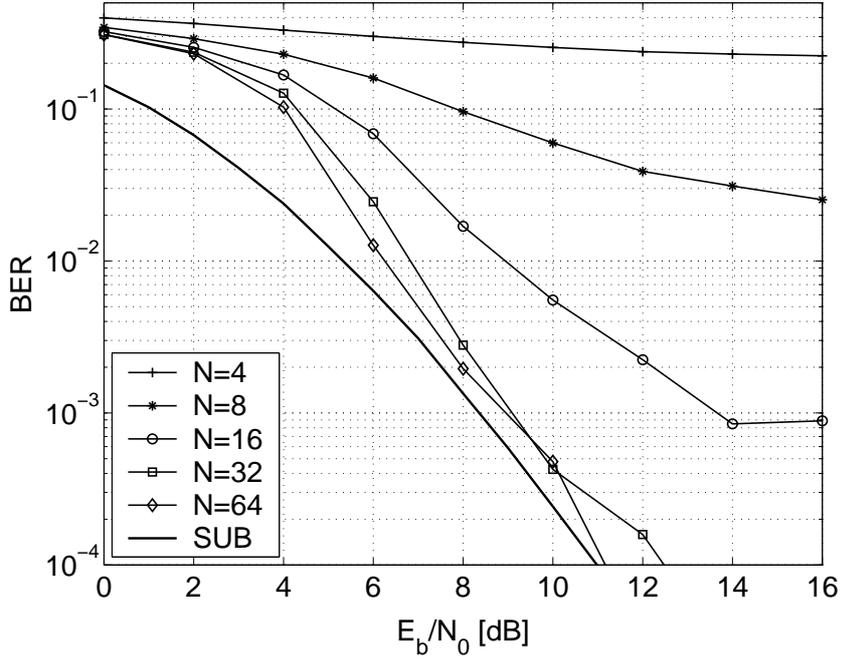
with the matrix

$$\check{\mathbf{S}}_a \triangleq \begin{pmatrix} \tilde{s}_1[0] & \tilde{s}_2[0] & \cdots & \tilde{s}_K[0] \\ \tilde{s}_1[1] & \tilde{s}_2[1] & \cdots & \tilde{s}_K[1] \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{s}_1[N-1] & \tilde{s}_2[N-1] & \cdots & \tilde{s}_K[N-1] \end{pmatrix}$$

consisting of the first  $N$  chips of the effective signature sequences of each user that are placed along the columns. The matrix  $\check{\mathbf{S}}_a$  has dimension  $N \times K$ . The single-user matched filter  $\mathbf{f}_k$  of user  $k$  for the approximated processing corresponds to the  $k$ -th column  $\check{\mathbf{s}}_{k,a}$  of the matrix  $\check{\mathbf{S}}_a$ . For spreading sequences with large spreading factor and short channels, *i.e.*,  $L \ll N$ , neglecting ISI is a justified assumption. However, it is unclear when a channel becomes short enough. To answer this question, we study a system with  $L = 5$  equally spaced i.i.d. Rayleigh taps. The taps are separated by a UMTS chip period, *i.e.*, by 260 ns [TS25.101]. Hence, the total channel length is 1.040 ns. This choice of channel is motivated by the measurement campaign described in [Hof02] where dispersions of up to 1 ns have been observed. For the present investigation, the spreading factor  $N$  is changed while the system load is kept fixed at  $\alpha = K/N = 1$ . The receiver has perfect channel knowledge and a PIC-SUMF is employed. Fig. 3.4 shows the attained BER after six iterations for different spreading factors  $N$ . We use a four-state convolutional code with rate  $R_C = 1/2$  and generator polynomials  $(5, 7)_8$ . The figure shows a dramatic loss for the spreading factors  $N = \{4, 8, 16\}$  compared to  $N = 32$ . The bit error rates for  $N = 32$  and  $N = 64$  are already very similar. We conclude that in channels of length  $L = 5$  with spreading factor  $N < 32$ , allowing for medium data rate communications, the consideration of ISI creates a huge gain in performance. In hilly terrains, channels might have even a larger spread such that proper ISI processing will also become vital for low-data rate communications.

### 3.8.3 Impact of Fading

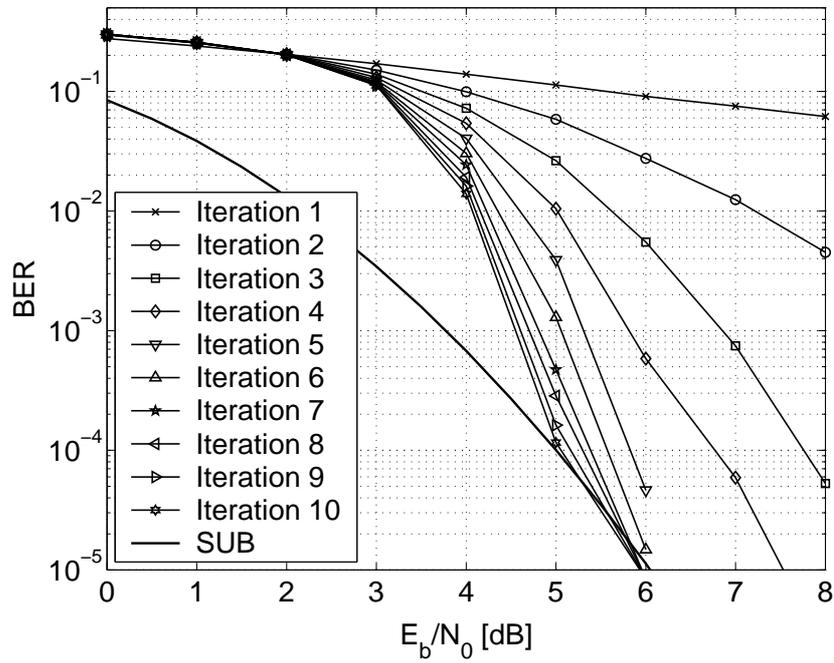
From the principle of successive interference cancellation it is intuitive that an equal power situation among the users is the worst case. Small scale fading causes a “natural” instantaneous power distribution that helps to accelerate achieving low BERs in an iterative receiver. It was shown in [Cai04] that particular power distributions increase the capacity when PIC-LMMSE filtering is employed. In the following we consider a system with load  $\alpha = 1.5$  with unconditional PIC- (3.22) as well as unconditional SIC-LMMSE (3.25) detection for spreading factor  $N = 8$ . First, let us



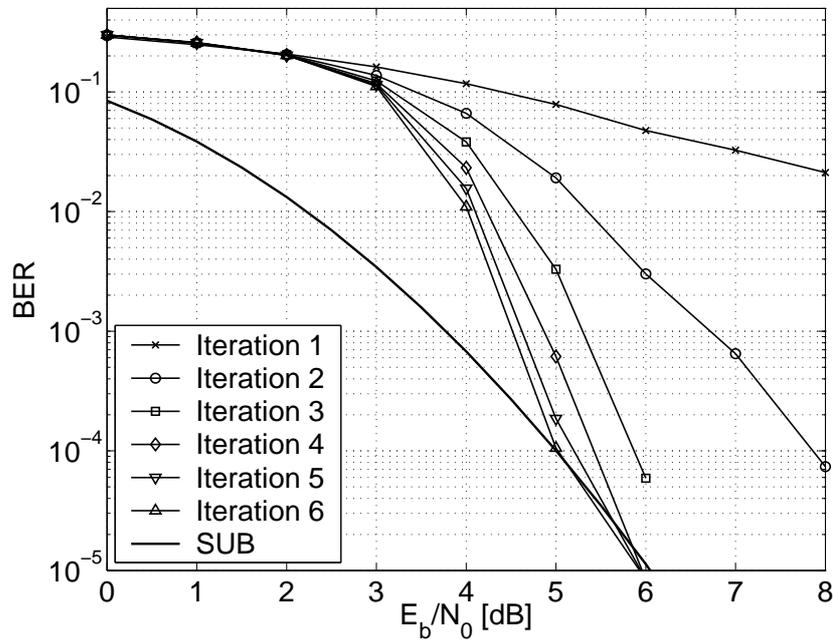
**Figure 3.4:** Influence of neglecting inter-symbol interference for a Rayleigh fading channel with  $L = 5$  i.i.d. taps and PIC-SUMF detection. The curves are parameterised by the spreading factor  $N = \{4, 8, 16, 32, 64\}$  for a load  $\alpha = 1$ .

focus on the AWGN case. The BER curves for the PIC are illustrated in Fig. 3.5 and the SIC is depicted in Fig. 3.6. The individual BER curves correspond to different iterations and the monotonic decrease indicates convergence towards the single-user bound. In this particular setting the PIC as well as the SIC attain the single-user bound at an  $E_b/N_0$  of 5 dB. The single-user bound is the bit error rate curve that is obtained when one single user is active employing a PIC-LMMSE detector and a symbol-by-symbol MAP-decoder. For mobile applications the bit error rate regime of interest is below  $10^{-3}$ . A BER of  $10^{-3}$  is attained at  $E_b/N_0$  of 4,5 dB. The PIC and SIC show different convergence behaviour. While the PIC requires ten iterations to achieve the SUB, the SIC achieves this already after six iterations. The corresponding BER curves attain lower values as their PIC counterparts. This is due to the scheduling structure of the SIC that takes into account the already correctly decoded bits and hence achieves an improved performance.

The second scenario reflects an i.i.d. Rayleigh channel with  $L = 5$  taps. The corresponding bit error rates are depicted in Figs. 3.7 and 3.8 for PIC and SIC, respectively. The first observation is that due to fading the single-user bound is flatter than its AWGN pendant due to limited diversity. Interestingly, the PIC attains the SUB already after five iterations. In case of the SIC just two iterations are needed to attain the SUB at an  $E_b/N_0$  of 8 dB where the BER reaches  $10^{-3}$ .



**Figure 3.5:** PIC-LMMSE detector with  $\alpha = K/N = 12/8 = 1,5$  on an AWGN channel.



**Figure 3.6:** SIC-LMMSE detector with  $\alpha = K/N = 12/8 = 1,5$  on an AWGN channel.

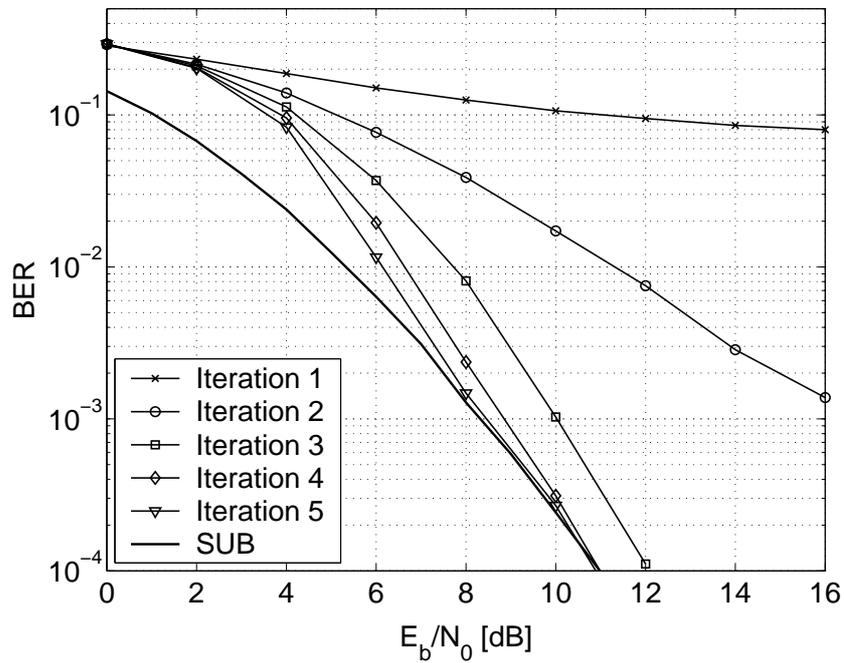
The plots also reveal an interesting detail. Those curves that correspond to the first iteration illustrate the bit error rate performance of non-iterative systems based on LMMSE detection (3.6). The achievable bit error rate is relatively poor and cannot be used in a practical setup. A classical Rake structure then has an even worse performance than the linear MUD schemes. What can also be learnt is that a SIC scheme can considerably improve with just one iteration. At the expense of a higher  $E_b/N_0$ , feasible BERs can be attained.

### 3.8.4 System Capacity

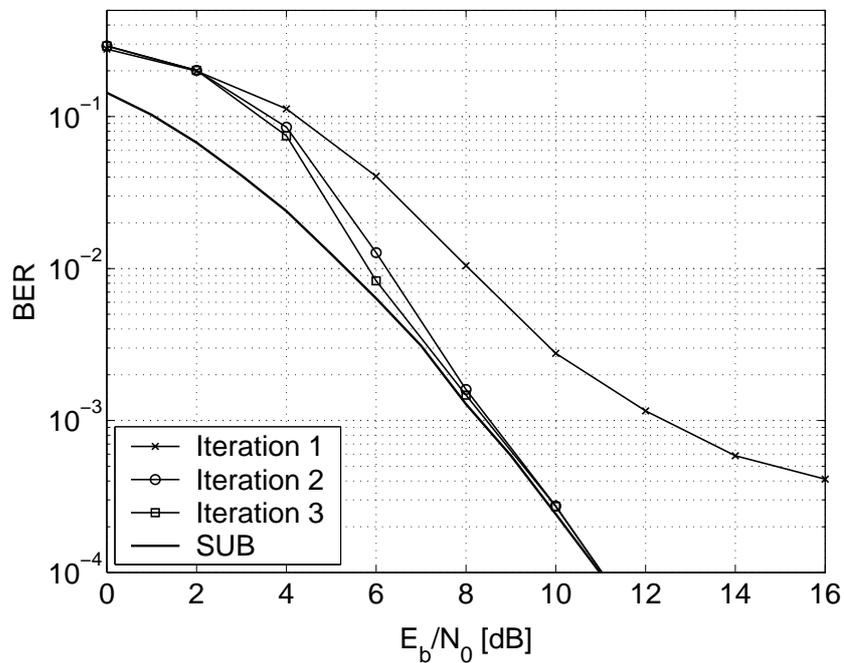
This section compares the effectivity of the interference cancellation methods in terms of number of accommodated users. Latter measure is also referred to as system capacity throughout this work. We consider a five tap i.i.d. Rayleigh fading channel with  $E_b/N_0 = 9$  dB and let the detectors have perfect CSI available. Let us start with the results for the PIC with a SUMF or a LMMSE filter. The BER curves are portrayed in Fig. 3.9 and Fig. 3.10, respectively. Taking the BER of  $10^{-3}$  as assessment criterion we find that the PIC-SUMF supports up to nine users while the PIC-LMMSE can support up to 17 users. This is a significant increase in capacity. Similar results are obtained for the SIC with SUMF and LMMSE. Fig. 3.11 shows that the SIC-SUMF can accommodate ten users while for the SIC-LMMSE, depicted in Fig. 3.12, this number increases to 17. From the figures it is obvious that the successive IC strategy converges with less iterations. The bit error rate curves below  $10^{-3}$  are jagged due to a limited number of simulations.

### 3.8.5 Multiple Receive Antennas

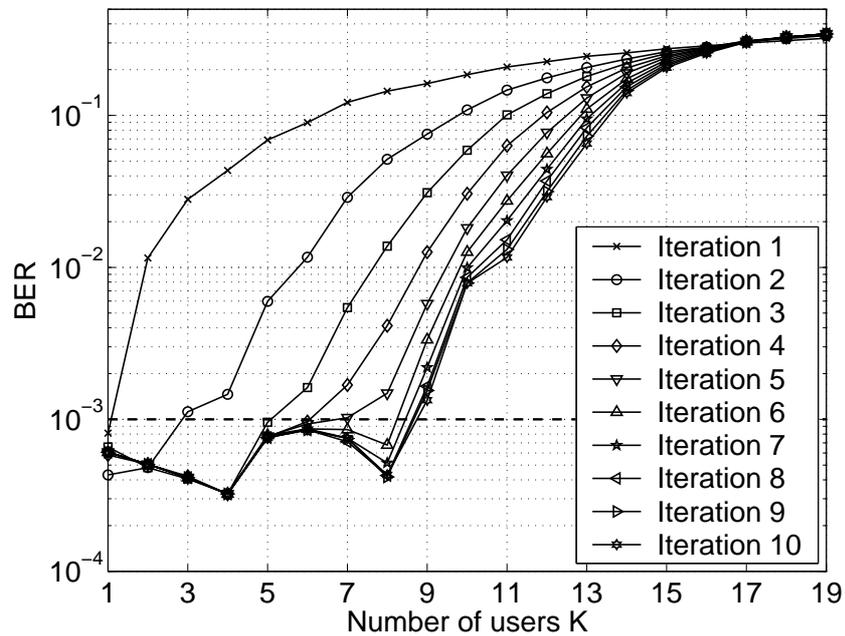
The benefit from installing two receive antennas at the receiver is portrayed in Fig. 3.13 and Fig. 3.14 for the PIC-LMMSE detector and the SIC-LMMSE detector. We assume an  $L = 5$  i.i.d. Rayleigh fading channel to both receive elements with the energy restriction formulated in (2.2). Antenna combining is achieved via MRC given in (3.26). Again, for the target BER of  $10^{-3}$  the number of users can be elevated from 17 to 29 for the PIC-LMMSE and from 17 to 30 for the SIC-LMMSE. This corresponds to a capacity increase of more than 70%. The extension into the spatial domain increases the effective diversity order which in turn leads first to an improved link reliability. Second, the assumption of independent multi-path realization imposes a spatial signature on the transmitted signal such that the degree of freedom in the spreading sequences is increased by the factor of two. This reduces the effective system load  $\alpha$  by a factor of two.



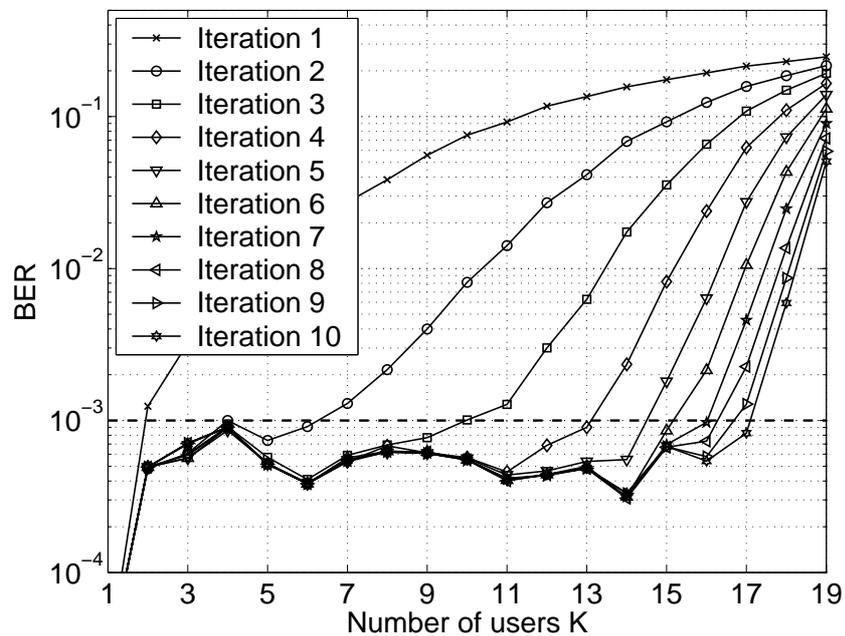
**Figure 3.7:** PIC-LMMSE detector with  $\alpha = K/N = 12/8 = 1,5$  on an  $L = 5$  tap Rayleigh fading channel.



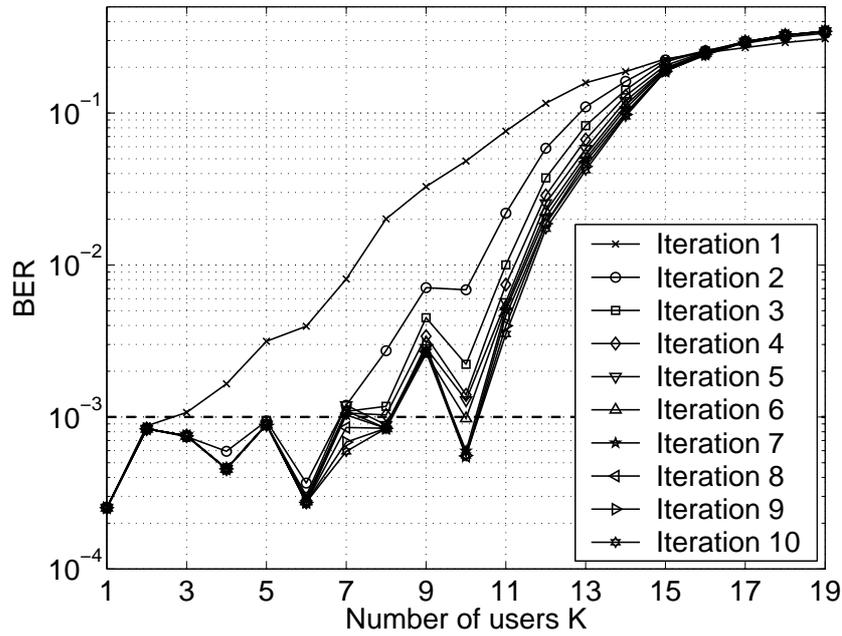
**Figure 3.8:** SIC-LMMSE detector with  $\alpha = K/N = 12/8 = 1,5$  on an  $L = 5$  tap Rayleigh fading channel.



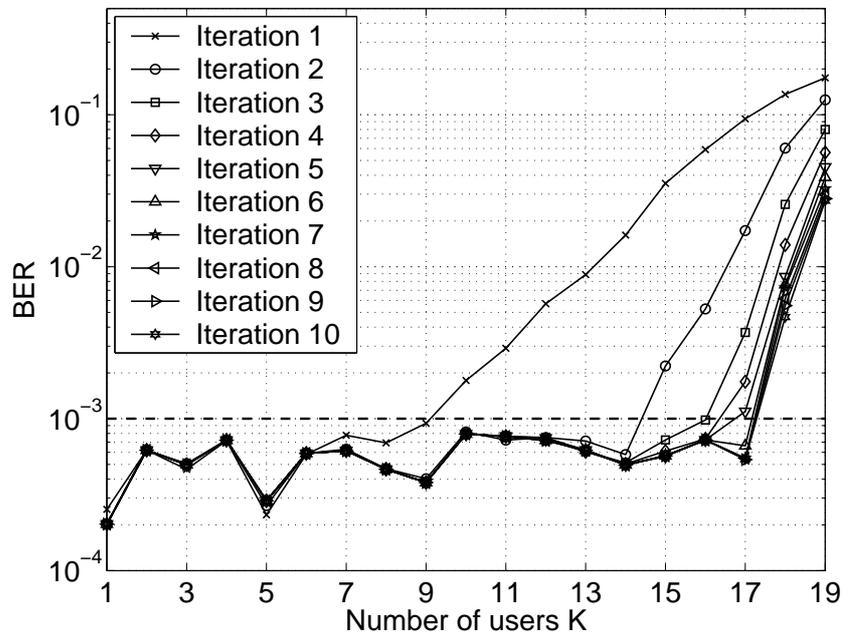
**Figure 3.9:** BER versus number of users of the PIC-SUMF detector on an  $L = 5$  tap Rayleigh fading channel with  $E_b/N_0 = 9$  dB and spreading factor  $N = 8$ .



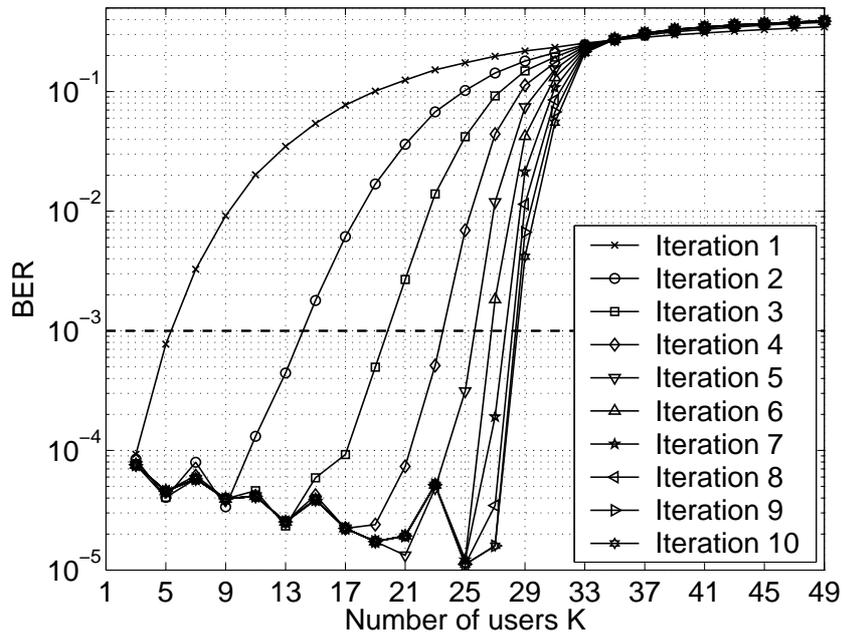
**Figure 3.10:** BER versus number of users of the PIC-LMMSE detector on an  $L = 5$  tap Rayleigh fading channel with  $E_b/N_0 = 9$  dB and spreading factor  $N = 8$ .



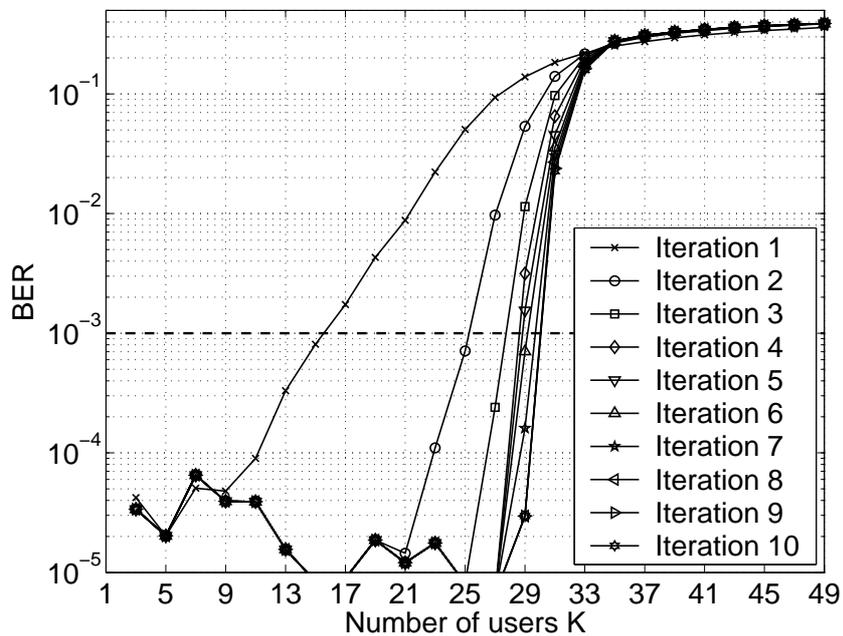
**Figure 3.11:** BER versus number of users of the SIC-SUMF detector on an  $L = 5$  tap Rayleigh fading channel with  $E_b/N_0 = 9$  dB and spreading factor  $N = 8$ .



**Figure 3.12:** BER versus number of users of the SIC-LMMSE detector on an  $L = 5$  tap Rayleigh fading channel with  $E_b/N_0 = 9$  dB and spreading factor  $N = 8$ .



**Figure 3.13:** BER versus number of users of the PIC-LMMSE detector on an  $L = 5$  tap Rayleigh fading channel with  $E_b/N_0 = 9$  dB, spreading factor  $N = 8$ , and two receive antennas.



**Figure 3.14:** BER versus number of users of the SIC-LMMSE detector on an  $L = 5$  tap Rayleigh fading channel with  $E_b/N_0 = 9$  dB, spreading factor  $N = 8$ , and two receive antennas.

**Table 3.1:** Coding gains (CG) of convolutional codes with rate  $R_C = 1/2$ .

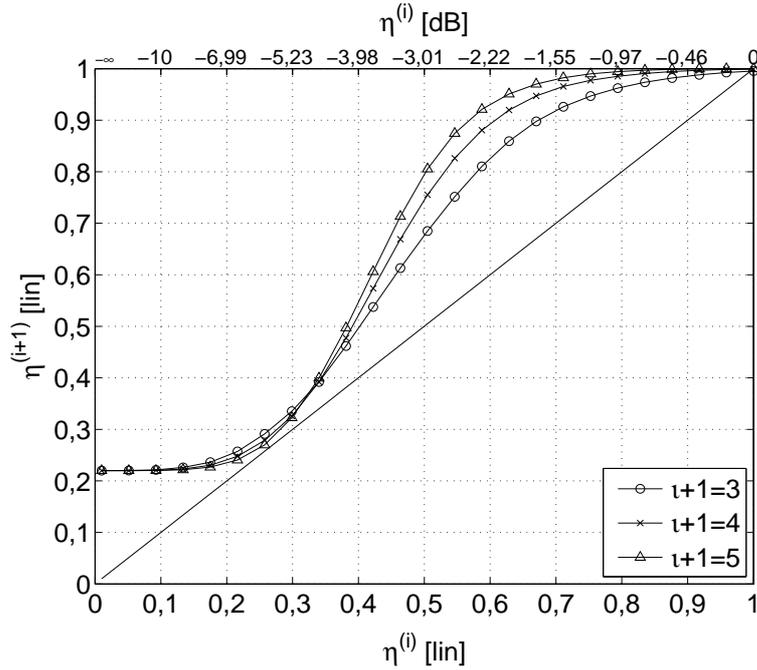
CONSTRAINT LENGTH $\iota + 1$	POLYNOMIALS	CG AT BER OF $10^{-3}$
3	$(5, 7)_8$	3,10 dB
4	$(15, 17)_8$	3,39 dB
5	$(23, 25)_8$	3,63 dB

### 3.8.6 Impact of Code Constraint Length

Due to the involvement of coding in iterative detection and decoding, the cancellation process is accelerated thanks to the self-correcting capabilities of the code. In Fig. 3.15 the evolution of the multi-user efficiency is illustrated for convolutional codes with rate  $R_C = 1/2$  and different constraint lengths  $\iota + 1$ , with  $\iota$  denoting the memory length. They are obtained through the analysis based on density evolution that we introduced in Section 3.7. The codes have constraint lengths  $\{3, 4, 5\}$  and their properties are listed in Tab. 3.1. The plots are generated for a system load of  $\alpha = 1,9$ . Despite the difference in coding gain, the plot shows that the maximum supportable load is nearly the same for all variants. However, the gap between the multi-user efficiency curve and the  $\Psi = \eta$  line below  $\eta < 0,3$  is the critical point in the convergence since it determines whether a receiver converges at all. From that region we learn that codes with larger constraint length have a closer distance to the diagonal line and experience a stopping earlier than codes with shorter constraint length. The different slopes of the multi-user efficiency curves for the region  $\eta > 0,3$  suggest that a different number of iterations is required to reach convergence. However, they do not determine whether or not they converge at all. Codes with short constraint length will require few more iterations but eventually they will also converge. The same behaviour was observed in an experimental setup in [Küh03] for an iterative receiver in OFDM-CDMA. The rationale behind this lies in the fact that for low SNRs the BER curve is rather flat and that it suddenly becomes steep as soon as a threshold value is reached. This behaviour gets more pronounced the more coding gain a particular code offers. In case of codes with longer constraint lengths, this threshold is shifted towards higher SINRs and can imply that the threshold SINR is never exceeded. Then the iterative process stops prematurely. For choosing among the investigated codes, it is feasible to select the one with constraint length three since decoding becomes less complex and highest loads can be supported.

## 3.9 Complexity Issues

From the results on parallel and successive interference cancellation it becomes clear that both techniques practically attain the same limits in bit error rate and/or num-



**Figure 3.15:** Multi-user efficiency evolution of the PIC-LMMSE detector on an AWGN channel with  $E_b/N_0 = 5$  dB with load  $\alpha = 1,9$ . The codes are all convolutional codes with different constraint lengths  $\iota + 1$ .

ber of accommodated users. However, SIC shows a faster progression and requires fewer iterations. The question, whether successive cancellation can be realized, is important for practical systems. Sequential user processing causes delays. A new user cannot be processed before the results from the previous are accessible. For a system, that accommodates many users, a real-time implementation problem arises. The PIC can parallelise the user processing but requires more iterations. A trade-off between processing power and admissible processing-delay has to be found.

The single-user matched filter (3.8), which is the same for PIC and SIC, requires the spreading sequence  $\mathbf{s}_k$  and the channel estimate  $\hat{\mathbf{h}}_k$ . The LMMSE-filter requires additional knowledge of the variance of the noise as well as the variance of the feedback symbols. The calculation of the biased LMMSE filter (3.18) requires the inverse of an  $(N + L - 1) \times (N + L - 1)$  matrix that needs to be calculated for all users and all symbols per iteration. An unbiased filter is obtained by applying the matrix inversion lemma. It turns out that the computation of the inverse can then be unified for all users (3.20) which saves a factor of  $K$  in complexity. The computation can be further simplified when we average over the variances of the soft feedback symbols along (3.21). In this way the inversion needs to be computed only once per iteration in case of PIC scheduling (3.22). The saving in complexity over the initial filter (3.20) is then  $K \times (M - J)$ . In case of successive interference

cancellation the inverse needs to be computed for every user separately per iteration (3.25). Its complexity is increased by a factor of  $K$  over the PIC implementation (3.20).

### 3.10 Summary

We have presented and discussed interference cancellation techniques with linear filters for iterative receivers. Cancellation can be classified into parallel and successive methods according to the deployed scheduling procedure. The cancellation step is followed by a linear detector filter that is realized as either single-user matched filter or linear MMSE filter. Instead of implementing the exact PIC-LMMSE filter (3.20) we employ an unconditional and unbiased variant (3.22) that saves roughly  $K \times (M - J)$  in complexity. A similar strategy is envisaged for the SIC where the complexity is decreased by a factor of  $M - J$  over the exact implementation.

**Feedback Symbols:** Whether soft decision data shall be derived from *extrinsic* or *a-posteriori* probabilities depends on the objectives. If the prime goal is the optimization of system capacity, *i.e.*, number of users, without considering the number of iterations, *extrinsic* probabilities are the preferred choice. If convergence speed is the criterion, *a-posteriori* probabilities shall be used for a particular load region. For loads smaller than one, *a-posteriori* probabilities tend to attain the same performance as *extrinsic* probabilities. However, for high load systems, *a-posteriori* probabilities show a stop in the convergence cycle due to residual interference.

**Detector Filter:** Receiver analysis based on density evolution as well as experiments made obvious that an LMMSE filter after interference cancellation allows more than 50% increase in load compared to the SUMF filter.

**Scheduling of Interference Cancellation:** The number of users that can be accommodated with successive interference cancellation is very little higher than for parallel interference cancellation. Practically, the two accommodate the same number of users. However, the total number of required iterations to achieve the maximum is for successive interference cancellation lower than for parallel interference cancellation. This comes to an expense in processing delay time and processing power for the LMMSE filter.

**Impact of Fading:** Fading causes a significant acceleration in convergence. This holds true for parallel as well as successive interference cancellation. The number of iterations to reach convergence is reduced from ten to five for the PIC-LMMSE detector when comparing the AWGN channel and the i.i.d. Rayleigh

case with  $L = 5$  taps. Respectively, for the SIC-LMMSE the number reduces from six to two.

**Code Constraint Length:** From the three investigated convolutional codes with rate  $R_C = 1/2$  and constraint lengths  $\iota + 1 \in \{3, 4, 5\}$ , the code with constraint length three supports the highest system capacity.

**Multiple Receive Antennas:** An increase in receive diversity enhances link reliability and offers additional degrees of freedom that lower the effective system load.

**Improvement in Bit Error Rate:** Iterative processing allows coded bit error rates below  $10^{-3}$  in an  $E_b/N_0$  region where linear one-shot multi-user detection performs two to three decades worse in terms of bit error rate.



# 4 Iterative Channel Estimation

In the previous chapter iterative multi-user detectors were developed with knowing the channel impulse response of all users or at least with an estimate thereof. In a practical situation the channel information is obtained through so called pilot symbols and it is typically associated with errors. This chapter investigates techniques that refine the quality of the channel estimates by using soft decision symbols gained from the decoder output symbols as additional pilots. This was illustrated for a single-tap channel in [ElGa00] and [Kob01]. The current work extends to multi-path scenarios and develops various estimation algorithms that assume different amount of statistical *a-priori* knowledge on the noise, the channel taps, and the feedback symbols. Their performance is assessed in terms of bit error rate (BER) and normalized mean square error (NMSE) for several channels with different power delay profiles. Further, two schemes for dedicated pilot sequences are investigated and compared against each other in terms of bit error rate and complexity. The iterative receiver is then extended to multiple receive antennas, and system capacities of the different schemes are evaluated.

## 4.1 Channel Estimation Algorithms

For the derivation of the estimators we will use the linear matrix signal model

$$\mathbf{y} = \begin{bmatrix} \mathcal{A}_P \\ \mathcal{A}_D \end{bmatrix} \mathbf{h} + \mathbf{v} = \mathcal{A}\mathbf{h} + \mathbf{v} \quad (4.1)$$

that we have introduced in Paragraph 2.5.3. The matrices  $\mathcal{A}_P$  and  $\mathcal{A}_D$  account for the contribution due to the dedicated pilots and the soft decision data, respectively. The channel is assumed to be block fading over  $M$  transmitted QPSK symbols. The channel impulse response  $\mathbf{h} \in \mathbb{C}^{KL \times 1}$  is estimated from the received vector  $\mathbf{y} \in \mathbb{C}^{MN \times 1}$  in the two following modes:

1. Pilot estimation: For initial channel estimation in the first iteration.
2. Pilot and soft decision symbol based estimation: For channel estimation in iteration two and subsequent iterations.

When the symbols are known to the receiver, the associated estimation is also called *data-aided* estimation like in mode one. When detected symbols are used if they were the true symbols, this is known as decision-directed estimation as in mode two [Mey97]. In the following, we will first present linear estimators for the case when only dedicated pilots are employed and will then develop estimators that employ feedback symbols. For the subsequent derivations we assume to have precise second order statistics available for the noise and the channel taps.

### 4.1.1 Initial Estimation

The first channel estimate is gained with the help of the dedicated pilot symbols. Design rules for pilot sequences with different constellation sizes will be discussed in Section 4.3. We use the linear matrix model (4.1) and consider that part of the observation only that corresponds to the pilot contribution  $\mathbf{A}_P$ . Instead of taking the total receive vector  $\mathbf{y}$  we use the first  $JN$  chips and define

$$\mathbf{y}_P \triangleq [y[0], \dots, y[JN - 1]]^T \in \mathbb{C}^{JN \times 1}.$$

This vector is the basis for estimating the unknown channel coefficients with the matrix  $\mathbf{A}_P$  that is known to the receiver. Assuming  $\mathbf{A}_P$  has full column rank, the *least squares* (LS) estimator reads

$$\hat{\mathbf{h}}_{\text{LS},P} = (\mathbf{A}_P^H \mathbf{A}_P)^{-1} \mathbf{A}_P^H \mathbf{y}_P \quad (4.2)$$

which is also the *maximum-likelihood* (ML) estimator due to the linear Gaussian signal model [Kay93]. It does not require knowledge about the statistics of the fading nor of the additive white Gaussian noise. The formula illustrates that the least squares estimator performs matched filtering and applies a decorrelator afterwards.

The linear *minimum mean square* (LMMSE) estimator is given by

$$\hat{\mathbf{h}}_{\text{LMMSE},P} = (\mathbf{A}_P^H \mathbf{C}_{v,P}^{-1} \mathbf{A}_P + \mathbf{C}_h^{-1})^{-1} \mathbf{A}_P^H \mathbf{C}_{v,P}^{-1} \mathbf{y}_P \quad (4.3)$$

with the covariance matrix of the noise part for the pilot preamble

$$\mathbf{C}_{v,P} = \mathbb{E}_{\mathbf{v}} \{ \mathbf{v}_P \mathbf{v}_P^H \} = \sigma_v^2 \mathbf{I}_{JN} \in \mathbb{R}^{JN \times JN} \quad (4.4)$$

and that of the channel

$$\mathbf{C}_h = \mathbb{E}_{\mathbf{h}} \{ \mathbf{h} \mathbf{h}^H \} = \text{diag} ([\mathbf{p}_1^T, \mathbf{p}_2^T, \dots, \mathbf{p}_K^T]) \in \mathbb{R}^{KL \times KL}.$$

The vectors  $\mathbf{p}_k$  for all users  $k$  that are vertically stacked and diagonalized are defined as

$$\mathbf{p}_k^T = [\sigma_{h,k}^2[0], \sigma_{h,k}^2[1], \dots, \sigma_{h,k}^2[L - 1]]^T \in \mathbb{R}^{L \times 1}.$$

### 4.1.2 Approximate Least Squares (ALS)

A least squares estimator for  $\mathbf{h}$  that operates on the total receive vector  $\mathbf{y}$  requires to know  $\mathcal{A}_P$  as well as  $\mathcal{A}_D$ . Since the latter matrix is not known we replace the symbols  $b_k(m)$  by their soft data estimates  $\tilde{b}_k(m)$  from (2.22) and hence get an approximation  $\tilde{\mathcal{A}}_D$ . The approximation is reflected in the name of the estimator - *approximate least squares* (ALS). One could also call it *tentative least squares* estimator. Using the definition

$$\tilde{\mathcal{A}} \triangleq \begin{bmatrix} \mathcal{A}_P \\ \tilde{\mathcal{A}}_D \end{bmatrix} \in \mathbb{C}^{MN \times KL}$$

in (4.1) and assuming that  $\tilde{\mathcal{A}}$  has full column rank, we obtain the ALS estimate

$$\hat{\mathbf{h}}_{\text{ALS}} = \left( \tilde{\mathcal{A}}^H \tilde{\mathcal{A}} \right)^{-1} \tilde{\mathcal{A}}^H \mathbf{y}. \quad (4.5)$$

### 4.1.3 Approximate Linear MMSE (ALMMSE)

We assume i.i.d. and zero-mean Rayleigh fading taps and a deterministic matrix  $\mathbf{A}$ . Latter assumption is an approximation since the soft feedback elements are stochastic. If the transmitted symbols were correctly at hand we would obtain the true linear MMSE estimator. The ALMMSE solution is

$$\hat{\mathbf{h}}_{\text{ALMMSE}} = \left( \tilde{\mathcal{A}}^H \mathbf{C}_v^{-1} \tilde{\mathcal{A}} + \mathbf{C}_h^{-1} \right)^{-1} \tilde{\mathcal{A}}^H \mathbf{C}_v^{-1} \mathbf{y}. \quad (4.6)$$

### 4.1.4 Linear MMSE (LMMSE)

A more advanced method than the ALS and ALMMSE estimator takes into account the variance of the soft decision symbols [Zem03a, Weh04, Lon04]. This is the objective of the linear estimator that is obtained via the Wiener-Hopf equations

$$\hat{\mathbf{h}}_{\text{LMMSE}} = \mathbf{C}_{yh}^H \mathbf{C}_{yy}^{-1} \mathbf{y}. \quad (4.7)$$

For the computation of the covariance matrices we assume that the receive vector  $\mathbf{y} \in \mathbb{C}^{MN \times 1}$ , the channel vector  $\mathbf{h} \in \mathbb{C}^{KL \times 1}$ , and the noise vector  $\mathbf{v} \in \mathbb{C}^{MN \times 1}$  are zero-mean. Furthermore,  $\mathbf{h}$ ,  $\mathbf{v}$ , and the data matrix  $\mathcal{B}$  are assumed statistically independent. The expectation over a data symbol  $\mathbb{E}_b\{b_k(m)\}$  is understood as *conditional expectation*, given the *a-posteriori* decoder output probability mass function from the previous iteration. Latter is expressed through the single variable APP $\{c(m')\}$  given in (2.19). The corresponding mapping was carried out in (2.22) and we obtain  $\tilde{b}_k(m) = \mathbb{E}_b\{b_k(m)\}$ . The covariance matrices of the receive vector and the channel

vector reads

$$\begin{aligned}
\mathbf{C}_{yh} &= \mathbb{E}_{\mathcal{B}, \mathbf{h}, \mathbf{v}} \{ \mathbf{y} \mathbf{h}^H \} \\
&= \mathbb{E}_{\mathcal{B}, \mathbf{h}, \mathbf{v}} \left\{ \sum_{i=0}^1 \mathcal{D}(i) \mathcal{B}(i) \mathbf{h} \mathbf{h}^H + \mathbf{v} \mathbf{h}^H \right\} \\
&= \sum_{i=0}^1 \mathcal{D}(i) \tilde{\mathcal{B}}(i) \mathbf{C}_h.
\end{aligned}$$

The auto-covariance matrix of the receive vector  $\mathbf{y}$  is computed as

$$\begin{aligned}
\mathbf{C}_{yy} &= \mathbb{E}_{\mathcal{B}, \mathbf{h}, \mathbf{v}} \{ \mathbf{y} \mathbf{y}^H \} \\
&= \mathbb{E}_{\mathcal{B}, \mathbf{h}, \mathbf{v}} \left\{ \sum_{i=0}^1 \sum_{j=0}^1 \mathcal{D}(i) \mathcal{B}(i) \mathbf{h} \mathbf{h}^H \mathcal{B}^H(j) \mathcal{D}^H(j) + \mathbf{v} \mathbf{v}^H \right\} \\
&= \sum_{i=0}^1 \sum_{j=0}^1 \mathcal{D}(i) \mathbb{E}_{\mathcal{B}} \{ \mathcal{B}(i) \mathbf{C}_h \mathcal{B}^H(j) \} \mathcal{D}^H(j) + \sigma_v^2 \mathbf{I}_{MN}. \tag{4.8}
\end{aligned}$$

In order to come up with the solution of (4.8) we need to compute the expectation  $\mathbb{E}_{\mathcal{B}} \{ \mathcal{B}(i) \mathbf{C}_h \mathcal{B}^H(j) \}$ . If we assume independence among individual code symbols we arrive at

$$\mathbb{E}_b \{ b_p(m) b_q(n) \} = \begin{cases} \tilde{b}_p(m) \tilde{b}_q(n) & \text{for } p \neq q, m \neq n \\ 1 & \text{for } p = q, m = n. \end{cases}$$

$p, q \in \{1, \dots, K\}$  denote the user indices and  $m, n \in \{0, \dots, M-1\}$  stand for the symbol indices. The product  $\mathbb{E}_{\mathcal{B}} \{ \mathcal{B}(i) \mathbf{C}_h \mathcal{B}^H(j) \}$  is expressed as

$$\mathbb{E}_{\mathcal{B}} \{ \mathcal{B}(i) \mathbf{C}_h \mathcal{B}^H(j) \} = \mathbb{E}_{\mathcal{B}} \{ \mathcal{B}(i) \} \mathbf{C}_h \mathbb{E}_{\mathcal{B}} \{ \mathcal{B}^H(j) \} + \mathbf{\Lambda}_{i,j} = \tilde{\mathcal{B}}(i) \mathbf{C}_h \tilde{\mathcal{B}}^H(j) + \mathbf{\Lambda}_{i,j}$$

where the matrix  $\mathbf{\Lambda}_{i,j} \in \mathbb{R}^{KLM \times KLM}$  denotes a diagonal matrix. Three cases can be distinguished:

$i = j$ :  $\mathbf{\Lambda}_{i,j}$  has entries on the main diagonal.

$i > j$ :  $\mathbf{\Lambda}_{i,j}$  has off-diagonal elements in the upper triangle and the diagonal starts in the  $\{(i-j)KL + 1\}$ -th column.

$i < j$ :  $\mathbf{\Lambda}_{i,j}$  has off-diagonal elements in the lower triangle and begins in the  $\{(j-i)KL + 1\}$ -th row.

The individual entries on the diagonals are the symbol variances weighted by the variance of the corresponding fading tap:  $\sigma_{h,k}^2[l]\sigma_{b,k}^2(m) \approx \sigma_{h,k}^2[l](1-|\tilde{b}_k(m)|^2)$ . There are  $(M - |i - j|)KL$  diagonal entries and in vector form they read

$$[\mathbf{p}_1^T \sigma_{b,1}^2(1), \dots, \mathbf{p}_K^T \sigma_{b,K}^2(1), \mathbf{p}_1^T \sigma_{b,1}^2(2), \dots, \mathbf{p}_K^T \sigma_{b,K}^2(M - |i - j|)]^T \in \mathbb{R}^{(M-|i-j|KL) \times 1}.$$

Note that, for perfectly known  $\tilde{b}_k(m)$ , the elements in  $\mathbf{\Lambda}_{i,j}$  become one and the estimator reduces to the ALMMSE estimator given by (4.6).

### 4.1.5 Interference Cancellation and Correlation (IC-Corr)

We discuss an interference canceler with post-correlation for channel estimation. This approach is similar to the interference canceling multi-user detectors with single-user matched filtering presented in Sections 3.4 and 3.5. For iterations  $i > 1$ , multiple-access interference is canceled from the receive vector  $\mathbf{y}$  for the identification of the  $l$ -th path of user  $k$  along

$$\tilde{\mathbf{y}}_{k,l}^{(i)} = \mathbf{y} - \left( \sum_{q=0}^1 \mathcal{D}(q) \tilde{\mathcal{B}}^{(i-1)}(q) \right) \hat{\mathbf{h}}_{\sim k,l}^{(i-1)} = \mathbf{y} - \tilde{\mathcal{A}} \hat{\mathbf{h}}_{\sim k,l}^{(i-1)}.$$

The subindex  $(\cdot)_{\sim k,l}$  reflects that entry  $kL + l - 1$  of the vector  $\hat{\mathbf{h}}$  is set to zero. The definition of the matrices follows Paragraph 2.5.3. A new estimate of channel tap  $l \in \{0, \dots, L - 1\}$  is obtained by correlating the interference mitigated observation vector  $\tilde{\mathbf{y}}_{k,l}$  with a delayed version of a stacked chip sequence consisting of a zero pre-cursor, the known pilot sequence  $\mathbf{u}_k$  and the soft decision data:

$$\left( \hat{h}_k[l] \right)_{\text{IC-Corr}}^{(i)} = \frac{N}{MN - l} [\mathbf{0}_l, \mathbf{u}_k^H, \mathbf{a}_{k,l}^H] \tilde{\mathbf{y}}_{k,l}^{(i-1)}.$$

The vector  $\mathbf{u}_k \in \mathbb{C}^{JN \times 1}$  denotes the pilot sequence of user  $k$ . Let  $\mathbf{a}_k$  be the chip sequence

$$\mathbf{a}_k \triangleq \left[ \tilde{b}_k^{(i-1)}(J), \dots, \tilde{b}_k^{(i-1)}(M - 1) \right]^T \otimes \mathbf{s}_k \in \mathbb{C}^{(M-J)N \times 1}$$

then

$$\mathbf{a}_{k,l} = \left[ a_k[0], a_k[1], \dots, a_k[(M - J)N - 1 - l] \right]^T \in \mathbb{C}^{(M-J)N-l \times 1}.$$

## 4.2 Mean Square Error Analysis

The quality of the estimated tap  $i$  can be assessed by its *mean square error* (MSE) given through

$$\text{MSE}_i \triangleq [\boldsymbol{\Sigma}]_{i,i} \quad (4.9)$$

with

$$\boldsymbol{\Sigma} \triangleq \mathbb{E}_v \left\{ \left( \hat{\mathbf{h}} - \mathbf{h} \right) \left( \hat{\mathbf{h}} - \mathbf{h} \right)^H \right\}$$

being the channel estimation error covariance matrix. For the *least squares* estimator (4.2) with known pilot symbols the error covariance matrix is given by

$$\boldsymbol{\Sigma}_{\text{LS}} = \sigma_v^2 \left( \mathcal{A}_{\text{P}}^H \mathcal{A}_{\text{P}} \right)^{-1}. \quad (4.10)$$

The corresponding matrix in the case of LMMSE estimation (4.3) reads

$$\boldsymbol{\Sigma}_{\text{LMMSE}} = \left( \mathbf{C}_h^{-1} + \frac{1}{\sigma_v^2} \mathcal{A}_{\text{P}}^H \mathcal{A}_{\text{P}} \right)^{-1}. \quad (4.11)$$

During the iterative refinement process we track the mean square error (4.9) of all taps of all users in one single quantity that we call *normalized mean square error* (NMSE). We define it as

$$\text{NMSE} \triangleq \frac{1}{KL} \text{tr} \{ \boldsymbol{\Sigma} \}. \quad (4.12)$$

On average, taps with larger power can be expected to be estimated with more precision than taps with low power. In the NMSE measure this is taken into account. The strong taps predominantly influence the bit error rate. The convergence behaviour of an iterative receiver strongly depends on the quality of the channel estimates that are required in the interference canceler. Interference mitigation is influenced by the phase as well as by the amplitude and can have a detrimental influence on detection in case of strong deviations.

The lower limit for the error covariance matrix  $\boldsymbol{\Sigma}$  is given by the Cramér-Rao lower bound (CRLB) that holds for unbiased estimators. In Paragraph 4.4.3 we make use of the CRLB for two cases: the first case arises when we consider the model with  $J$  pilot symbols only (CRLB1) – it is valid for the first iteration, *i.e.*, when only the pilot symbols are involved. The second is based on the assumption that we know the  $J$  pilot symbols and the  $M - J$  data symbols  $b_k(m)$  perfectly (CRLB2). Then, the matrix  $\mathcal{A}_{\text{P}}$  is replaced by the complete pilot and data matrix  $\mathcal{A}$  (4.1) in (4.11). Our MSE analysis does not include the data-aided cases since we do not know how to evaluate the value of the soft symbols after the  $i$ -th iteration for a general multi-path fading channel. Hence, only the two above mentioned extreme cases are considered. With known pilot sequences the least squares estimator is unbiased and the CRLB bounds its MSE. For the LMMSE estimator, the MSE can be bounded by the Bayesian CRLB [Tre68]. For the individual channel tap  $l$  the Bayesian CRLB of user  $k$  is

$$\text{MSE}_{\hat{h}_k[n]} \geq \left[ \mathbf{K}^{-1} \right]_{(k-1)L+n, (k-1)L+n} \quad (4.13)$$

where  $\mathbf{K}$  is the Bayesian information matrix. The Bayesian information matrix is composed of the Fisher information matrix and the prior information matrix<sup>1</sup>

$$\mathbf{K} = \mathbf{J} + \mathbf{J}_p$$

with

$$\mathbf{J} = -\mathbb{E}_{\mathbf{h}, \mathbf{y}} \left\{ \frac{\partial}{\partial \mathbf{h}^*} \left( \frac{\partial}{\partial \mathbf{h}^*} \right)^T \log f_{\mathbf{y}|\mathbf{h}}(\mathbf{h}, \mathbf{y}) \right\}, \quad (4.14)$$

$$\mathbf{J}_p = -\mathbb{E}_{\mathbf{h}} \left\{ \frac{\partial}{\partial \mathbf{h}^*} \left( \frac{\partial}{\partial \mathbf{h}^*} \right)^T \log f_{\mathbf{h}}(\mathbf{h}) \right\} \quad (4.15)$$

with  $\log(\cdot)$  denoting the natural logarithm. The formalism of the complex derivative follows the definitions in [Kay93]. The observation vector  $\mathbf{y}$  is conditionally distributed as  $\mathcal{CN}(\mathbf{A}\mathbf{h}, \mathbf{C}_v)$  and its conditional likelihood function is the complex Gaussian multi-variate distribution

$$f_{\mathbf{y}|\mathbf{h}}(\mathbf{h}, \mathbf{y}) = \frac{1}{\pi^{KL}(\sigma_v^2)^{MN}} \exp \left( -\frac{1}{\sigma_v^2} (\mathbf{y} - \mathbf{A}\mathbf{h})^H (\mathbf{y} - \mathbf{A}\mathbf{h}) \right).$$

The channel vector  $\mathbf{h}$  has distribution  $\mathcal{CN}(\mathbf{0}, \mathbf{C}_h)$  and its likelihood function reads

$$f_{\mathbf{h}}(\mathbf{h}) = \frac{1}{\pi^{KL} \det(\mathbf{C}_h)} \exp(-\mathbf{h}^H \mathbf{C}_h^{-1} \mathbf{h}).$$

We now evaluate the information matrices (4.14), (4.15). They involve the second derivative of the log-likelihood function. Alternatively, we can utilize the following property [Kay93]: An efficient estimator, *i.e.*, one that attains the CRLB, exists iff  $\frac{\partial}{\partial \mathbf{h}^*} \log f_{\mathbf{y}|\mathbf{h}}(\mathbf{h}, \mathbf{y})$  can be written as

$$\frac{\partial}{\partial \mathbf{h}^*} \log f_{\mathbf{y}|\mathbf{h}}(\mathbf{h}, \mathbf{y}) = \mathbf{J}(\mathbf{g}(\mathbf{y}) - \mathbf{h}). \quad (4.16)$$

In this case  $\mathbf{J}$  is the Fisher information matrix and  $\mathbf{g}(\mathbf{y})$  is the expression for the estimator attaining the bound. The first derivative of the log-likelihood function (4.14) is

$$\frac{\partial}{\partial \mathbf{h}^*} \log f_{\mathbf{y}|\mathbf{h}}(\mathbf{h}, \mathbf{y}) = \frac{1}{\sigma_v^2} \mathbf{A}^H \mathbf{A} ((\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{y} - \mathbf{h}).$$

A comparison with (4.16) identifies the Fisher information matrix (4.14) as

$$\mathbf{J} = \frac{1}{\sigma_v^2} \mathbf{A}^H \mathbf{A}$$

---

<sup>1</sup>If the prior information matrix  $\mathbf{J}_p$  is chosen to be the zero matrix then this reduces to the classical CRLB.

and  $\mathbf{g}(\mathbf{y}) = (\mathcal{A}^H \mathcal{A})^{-1} \mathcal{A}^H \mathbf{y}$ . This estimator corresponds to the modified least squares estimator and equals the ALMMSE in (4.6) for deterministic  $\mathbf{h}$ , *i.e.*,  $\mathbf{C}_h = \mathbf{0}$ . We are left with the computation of the second derivate (4.15) of the prior. After some algebra we obtain

$$\mathbf{J}_p = -\mathbb{E}_{\mathbf{h}} \left\{ \frac{\partial}{\partial \mathbf{h}^*} \left( \frac{\partial}{\partial \mathbf{h}^*} \right)^T \log f_{\mathbf{h}}(\mathbf{h}) \right\} = \mathbf{C}_h^{-1}.$$

### 4.3 Pilot Sequence Design

The design of pilot sounding sequences  $\mathbf{u}_k$  is of utmost importance since the quality of the channel estimates severely depends on their correlation properties. First, to resolve multi-path components, one is interested in sequences that exhibit good auto-correlation properties to keep distortions low that result from multi-path components. Second, good cross-correlation properties are desired to keep distortion low that arise due to the other active users.

#### 4.3.1 Optimal Sequences

Optimal channel sounding sequences minimize the normalized mean square error (4.12) under an energy constraint. It was shown in [Cro91] that optimal sequences must satisfy

$$\mathcal{A}_P^H \mathcal{A}_P \propto \mathbf{I}_{JN}. \quad (4.17)$$

The estimation error is inversely proportional to the total pilot symbol energy  $\epsilon$ . When imposing a trace constraint  $\epsilon = \text{tr}\{\mathcal{A}_P^H \mathcal{A}_P\}$  with  $\epsilon$  being the total pilot energy, the minimization of the NMSE (4.12) can be understood as a constraint optimization problem which can be solved by Langrangian multipliers [Cai01] or by matrix inequalities [Cro91]. Pilot sequences with a particular length  $JN$  and  $R$  constellation points are chosen such that this error is minimized. Sequences that minimize the NMSE have been identified in numerous works. To mention a few the reader is referred to modified m-sequences [Cro91] and *perfect root of unity sequences* (PRUS) [Ng98, Bal00, Rup00].

#### 4.3.2 Pilot Sequences

In this paragraph we present one representative of suboptimal and one from optimal training sequences.

### Random Symbol Sequence with Spreading (RAND)

A straightforward approach for finding an identification sequence is to apply the short spreading sequences  $\mathbf{s}_k$ , with which symbol spreading is performed, also on a sequence

$$\mathbf{b}_{k,\text{pilots}} \triangleq [b_k(0), \dots, b_k(J-1)]^T \in \mathbb{X}^{J \times 1}$$

of random QPSK pilot symbols. The resulting chip sequence  $\mathbf{u}_k \triangleq [u_k[0], \dots, u_k[JN-1]]^T \in (\mathbb{X}/\sqrt{N})^{JN \times 1}$  of user  $k$  can be written as

$$\mathbf{u}_k = \mathbf{b}_{k,\text{pilots}} \otimes \mathbf{s}_k.$$

The individual chip values  $u_k[n]$  lie on the four possible points of the unit circle given through the set  $\mathbb{X}/\sqrt{N}$ . This sequence does not minimize the NMSE (4.12) under a transmit power constraint in general but it has only  $R = 4$  roots.

### Perfect Root of Unity Sequences (PRUS)

Periodic sequences for optimal identification  $\mathbf{x} \triangleq [x[0], \dots, x[JN-1]]^T \in \mathcal{X}$  of length  $JN$  exhibit the perfect autocorrelation property

$$\delta[l] = \frac{1}{JN} \sum_{n=0}^{JN-1} x[n+l]x^*[n] = \begin{cases} 1 & \text{for } l = 0, \\ 0 & \text{else.} \end{cases} \quad (4.18)$$

One set of sequences fulfilling this requirement under a constant modulus constraint are *perfect root of unity sequences* (PRUS). These are sequences whose elements are complex roots of unity in the form  $\exp(j2\pi x)$  with  $x$  being a rational number. Perfect refers to the fact that the autocorrelation is zero except for lag  $l = 0$ . It was shown in [Mow95] that for a constellation with  $R$  roots (this is the alphabet size) and length  $JN$  a lower bound on the number  $|\mathcal{X}|$  of available sequences is given by

$$|\mathcal{X}| \geq \begin{cases} p! q^p \phi^p(q) R^p & \text{if } R_{\min} \text{ divides } R, \\ 0 & \text{else,} \end{cases} \quad (4.19)$$

where

$$R_{\min} = \begin{cases} 2pq & \text{for } q \text{ even and } p \text{ odd,} \\ pq & \text{else.} \end{cases} \quad (4.20)$$

$R_{\min}$  denotes the minimum alphabet size. It was conjectured in [Mow95a] that for a given length, PRUS exist if and only if the alphabet size is an integer multiple of  $R_{\min}$ . The factor  $p$  should be chosen such that  $p^2$  is the greatest perfect square factor in  $JN$  and  $q$  is an integer [Ng98]. Both numbers are chosen such that (4.20) and

$$JN = p^2 q$$

are met. The equations imply that for a particular PRUS sequence length  $JN$  a minimum number of roots  $R_{\min}$  is required. In (4.19) the expression  $\phi(q)$  denotes the Euler Totient function [Abr70]. It is defined as the number of integers in the set  $\{1, 2, \dots, q-1\}$  that are relatively prime to  $q$ , *i.e.*, have no common factors.

A design of PRUS based on consecutive root of unity phase difference sequences is given in [Ng98]. With the definition of phase differences

$$d[l] \triangleq \frac{x[n+l]}{x[n]} \quad (4.21)$$

we can see that the individual terms in the correlation (4.18)

$$x[n+l]x^*[n] = \frac{x[n+l]}{x[n]} = d[l]$$

correspond to phase differences. Hence, for a sequence with perfect autocorrelation properties the sum of differences for different time lags  $l$  diminishes. A solution for the consecutive root of unity phase difference sequences  $d[l]$  fulfilling requirement (4.18) is given by

$$d[pk+s] = A[s]E_r(k), \quad s \in \{0, 1, \dots, p-1\}$$

where

$$E_r[k] = \exp\left(j\frac{2\pi}{pq}kr\right), \quad k \in \{0, 1, \dots, pq-1\}.$$

The variable  $r$  is an integer factor that is relatively prime to  $JN$ .  $A[s]$  represents an arbitrary complex constant with unit magnitude. Setting  $x[0] = 1$  we can determine the consecutive symbols from (4.21). The number of shifted sequences corresponds to the sequence length  $JN$ . In case of channel identification  $JN$  must be larger than  $KL$ , the number of paths that need to be estimated. This in turn imposes a lower bound on the number of required pilot symbols  $J$

$$JN \geq KL. \quad (4.22)$$

Due to property (4.18) one obtains sounding sequences of length  $JN$  for different users by shifted versions of periodically extended perfect root of unity sequences. A possible assignment of identification sequences based on a PRUS to the individual users is

$$u_k[n] = x[(n + (k-1)L) \bmod JN].$$

### 4.3.3 Identifiability

To obtain a meaningful estimate of the  $KL$  unknown channel coefficients, the matrix  $\mathcal{A}^H \mathcal{A}$  needs to be invertible, *i.e.*,  $\mathcal{A} \in \mathbb{C}^{MN \times KL}$  needs to have full column rank  $KL$ . Matrices  $\mathcal{A}$ , that include soft decision data, almost surely show full rank. However, pilot matrices  $\mathcal{A}_P \in \mathbb{C}^{JN \times KL}$ , that are used for initial channel estimation, can be a rank deficient. This applies also when a big number of parameters needs to be identified as it happens in overloaded systems, *i.e.*, where  $K > N$ , or in long channels where  $L$  is large.

It was mentioned in [Mar99], in context of a MIMO-BLAST system on a flat fading channel, that the number of training symbols needs to be at least as great as the number of transmit antennas. In other words, the number of training symbols  $JN$  in our case must be at least  $KL$ . This result corresponds to condition (4.22).

The issue of the optimum amount of pilot symbols dedicated to MIMO channel sounding was addressed in [Has03] by maximizing mutual information. However, we do not make explicit use of these results. We simply employ the shortest number of pilots fulfilling (4.22).

### 4.3.4 Pilot Placement

Throughout this work we consider block fading, *i.e.*, the channel is supposed to be constant during the transmission of a block of  $M$  symbols. For the transmit block structure and channel estimation we make the following assumptions:

1. Symbols are placed in a preamble of length  $J$ .
2. We assume a guard interval between blocks. This guard interval has length  $L-1$  and assures that the beginning of each transmission block is not distorted by previous transmissions. Hence, inter-block interference does not occur.
3. The fraction of energy, that extends the length of a block in mode (i) pilot estimation and (ii) pilot and soft decision supported estimation, is negligible:  $L-1 \ll JN < MN$ .

We will now justify the placement scheme. The optimum placement of pilot symbols within a transmission block is studied in [Don02] for a decision-feedback receiver structure that is able to exploit also the data symbols which are known perfectly. When we further restrict that data and pilot symbols have the same modulus, the pilots can be placed anywhere in the block as long as (4.22) is met. However, since the correctness of feedback data symbols in our iterative receiver depends on the initial channel estimate, we need to provide a good channel estimate that is based only on the pilots, *i.e.*, those ones that keep the NMSE low. For the case when only

pilots are at hand and no data available, a clustered placement of pilots is optimal since no guard intervals need to be introduced.

## 4.4 Simulation Results

The achievable bit error rates as well as system capacities of the iterative receivers with integrated channel estimation are studied in the present section by means of Monte-Carlo simulations. We will consider a system where all users have a non-systematic, non-recursive convolutional code  $\mathcal{C}_k$  with rate  $R_C = 1/2$  and generator polynomials  $(5, 7)_8$ . We assume channels with  $L$  consecutive taps that fulfill the power constraint formulated in (2.2). The particular power delay profile in force is given in the description of the considered setup. For the case of ALMMSE and LMMSE estimation we assume that the noise variance  $\sigma_v^2$  and the channel covariance matrix  $\mathbf{C}_h$  are perfectly known. We choose  $M = 160$  QPSK symbols per block and use spreading factor  $N = 8$ . We motivate this by UMTS slot processing [TS25.213] where for a data spreading factor  $N = 8$  we would also obtain 320 bits for a slot length of 2.560 chips (here, we neglect the ten bits on the control channel). To take into account the loss of energy due to the  $J = 10$  pilots we stick to definition (2.14) for  $E_b/N_0$  in the following. In the sequel we will study the influence of system parameters, channel estimators, and detectors on the overall performance. For all BER vs.  $E_b/N_0$  curves we consider overloaded systems with load  $\alpha = K/N = 1.5$ . In the following, the *single-user bound* (SUB) is understood as the bit error rate that is achieved when only one single user is active having perfect channel knowledge available and employing unbiased, unconditional PIC-LMMSE detection and MAP decoding. The SUB is thus the lower bound on the BER for the multi-user case. A receiver can attain the SUB if all multiple-access interference is mitigated successfully. To facilitate the comparison to a situation with channel estimation, the single-user bound is evaluated for the  $E_b/N_0$  in accordance with definition (2.14).

### 4.4.1 Feedback Information

Using soft decision data in the system with perfect channel information was discussed in Paragraph 3.8.1. Naturally, the question whether to apply *extrinsic* or *a-posteriori* mapped symbols arises also for channel estimation. We have observed that a mapping from *a-posteriori* information results in an accelerated convergence towards the hard values in the discrete symbol set  $\mathbb{X}$  [Lon04] for frequency selective fading. Detrimental correlation effects as they were observed for *a-posteriori* mapped symbols in case of interference cancellation in the detector cannot be noticed. We will therefore only use *a-posteriori* symbols.

### 4.4.2 Inter-Symbol Interference

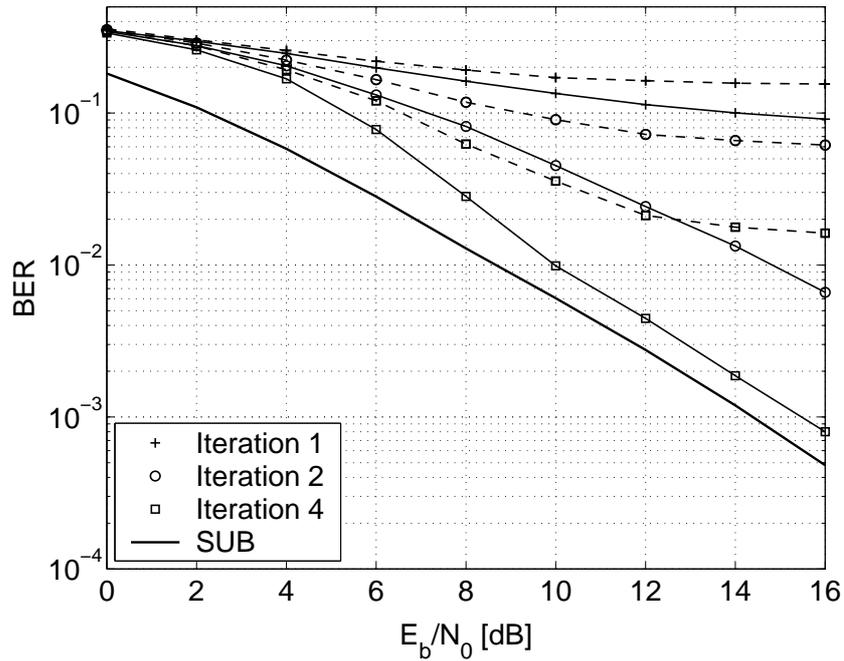
Some papers assume that if the channel is short compared to the chip sequence length  $L \ll N$ , inter-symbol interference can be neglected. In this case only the first  $N$  chips of the effective sequence  $\tilde{\mathbf{s}}_k$  are used for detection. The impact of neglecting the remaining chips was investigated in Paragraph 3.8.2 for the case when perfect channel knowledge is available at the receiver. Now, we will investigate how this looks like for different iterations when the channel is estimated. In Fig. 4.1 we illustrate the difference of a system neglecting ISI to one that considers ISI. We assume an i.i.d. Rayleigh channel of length  $L = 2$  and the spreading sequence length  $N$  is eight. We employ the random pilot scheme described in Paragraph 4.3.2 with  $J = 10$ . The number of users is  $K = 12$  and the channel is estimated via the ALS algorithm. The multi-user detector is the unconditional, unbiased PIC-LMMSE. The plot reveals that, in case of neglecting ISI in this short channel, the loss in terms of bit error rate is significant. There is no possibility to attain a bit error rate of  $10^{-3}$  by neglecting ISI. The curves show an error floor after all iterations. However, when ISI is treated properly a steady improvement of the BER after each iteration can be observed that allows a BER of  $10^{-3}$  at an  $E_b/N_0$  of 15,5 dB. For a larger channel this effect becomes more pronounced and points out the significance of considering ISI in medium data rate communication systems like UMTS.

### 4.4.3 Channel Estimators

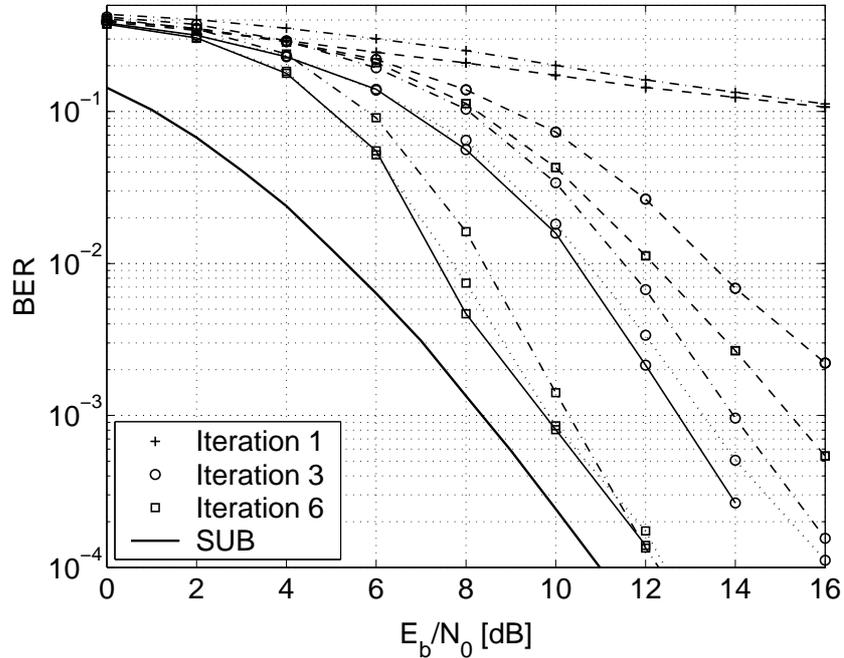
The channel estimators that were developed in Section 4.1 can be classified into two groups: the first group comprises those estimators that carry out joint channel estimation and the second group is based on interference cancellation. We comment separately on the findings for the two groups in the sequel.

#### Joint Estimators

The ALS, ALMMSE, and LMMSE estimators require different statistical knowledge on the random parameters  $\mathbf{v}$ ,  $\mathbf{h}$ , and  $\tilde{b}(m)$ . The simplest scheme is the ALS estimator (4.5) that does not use any assumption on either of the three random variables. The ALMMSE (4.6) assumes to know the variance of the noise as well as the covariance matrix of the channel vector. The ALS and the ALMMSE estimator assume that the soft data estimates obtained from the feedback are deterministic. However, due to the probabilistic nature of the decoder output this is not the case. One way to incorporate all statistical information is to include also the variance of the individual symbols. This leads to the LMMSE estimator formulated in (4.7) which can be considered the optimum linear estimator.



**Figure 4.1:** Benefits of ISI processing in i.i.d. Rayleigh channel with  $L = 2$ . Solid lines (—) refer to exact ISI processing while dashed lines (---) indicate that ISI was neglected. ALS channel estimator with random pilot symbols.  $N = 8$ ,  $K = 12$ ,  $J = 10$ .



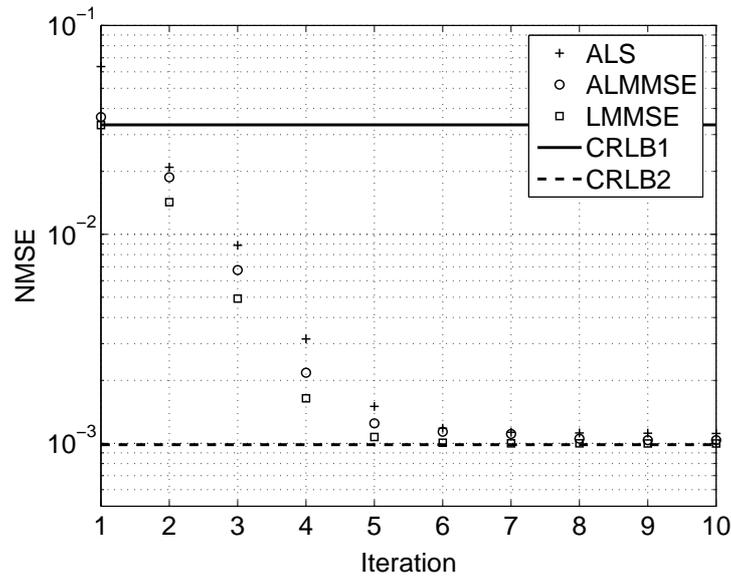
**Figure 4.2:** Comparison of BER with channel estimators PILOTS only (---), ALS (—·), ALMMSE (···), and LMMSE (—). The pilot symbols are RAND and the parameters  $N = 8$ ,  $K = 12$ , i.i.d. Rayleigh channel with  $L = 5$ ,  $M = 160$ ,  $J = 10$ .

Fig. 4.2 shows bit error rate curves plotted versus  $E_b/N_0$  corresponding to iterations one, three, and six for an i.i.d. multi-path Rayleigh fading channel with  $L = 5$  taps and different channel estimators. We do not illustrate results for iterations higher than six since there is no further improvement. We use the random pilot sequence (RAND) which was described in Paragraph 4.3.2. The plot shows two groups of curves. The first group refers to the case when only the dedicated pilots are used for channel estimation. The corresponding BER curves are plotted as dashed lines (PILOTS). The second group is related to the case when also soft decision symbols are used in the feedback and is illustrated by dash-dotted (ALS), dotted (ALMMSE), and solid (LMMSE) lines.

The BER curve referring to the first iteration was generated with the pilots only approach in connection with the ALS or the ALMMSE channel estimator. They represent the performance of a non-iterative, linear multi-user receiver. If we use the initial ALMMSE channel estimate in the further iterations we achieve BERs marked by dashed curves after the third and the sixth iteration. Although there is no iterative refinement of channel estimates involved during iterations, the receiver is able to mitigate multiple-access interference due to the code, even with a rather poor channel estimate. A BER of  $10^{-3}$  is obtained after the sixth iteration at an  $E_b/N_0$  of 15,1 dB.

All other estimators presented in the plots use, additionally to the pilots, also the data symbols for channel estimation. After the third iteration, the ALS, the ALMMSE, and the LMMSE scheme reach a BER of  $10^{-3}$  at roughly 12,8 – 14 dB. After the sixth iteration the required  $E_b/N_0$  lies between 9,5 and 10,5 dB. This corresponds to an improvement over the pilot estimation of approximately 4 – 6 dB after the third iteration and 5 – 6 dB after the sixth iteration. The performance of the receiver with ALS estimation is 0,5 – 1 dB worse than the one with the ALMMSE estimator. The LMMSE estimator achieves an additional gain of up to 0,5 dB. For a target bit error rate of  $10^{-3}$ , the gap to the single-user bound is 2 dB for the ALS estimator and 1,5 dB for the LMMSE estimator. For the ALMMSE as well as the LMMSE estimator we have assumed that we know the noise variance and the channel covariance matrix perfectly.

A plot of the evolution of the normalized mean square error (4.12) for different iterations is given in Fig. 4.3. The system load is  $\alpha = 1,5$  and  $E_b/N_0 = 10$  dB. The NMSE is shown for the ALS, the ALMMSE, and the LMMSE estimator. For the Cramér-Rao lower bound (4.11) we assume that in the first iteration only the pilots can be exploited (CRLB1) and for all other iterations the complete data (CRLB2). Since the bounds need to be evaluated for a particular system model, we randomly choose one particular realization of pilot/data symbols and also one particular realization of the matrix  $\mathbf{S}$  of signature sequences. The NMSE decreases most rapidly for the LMMSE, followed by the ALMMSE, and then the ALS.

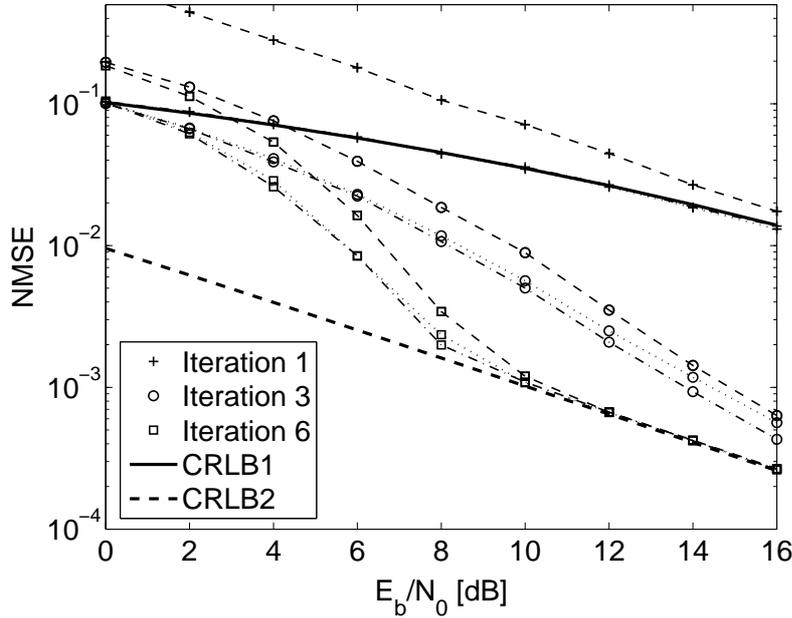


**Figure 4.3:** Evolution of the NMSE for RAND pilots and different channel estimators,  $N = 8$ ,  $K = 12$ , i.i.d. Rayleigh channel with  $L = 5$ ,  $E_b/N_0 = 10$  dB,  $M = 160$ ,  $J = 10$ . CRLB1 designates the Cramér-Rao bound for the  $J$  pilots only and CRLB2 for  $M$  symbols.

The NMSE as a function of  $E_b/N_0$  is depicted in Fig. 4.4. Results are illustrated for the ALS, the ALMMSE, and the LMMSE estimator. Further, the Cramér-Rao lower bounds for  $J$  pilot symbols (CRLB1) and for all  $M$  symbols (CRLB2) are shown. To improve readability the curves are plotted for iterations one, three, and six only. For the first iteration the ALMMSE estimator is the same as the LMMSE and it represents the efficient estimator attaining the CRLB1 with the  $J$  pilot symbols. After six iterations the CRLB2 for  $M$  symbols is practically reached by all estimators at 10 dB.

### IC-Corr Estimator

On a first glance the idea of interference cancellation for channel estimation is appealing, since subsequent processing involves only a correlation. And, in fact, the concept works well as long as there are few channel coefficients,  $L \leq 3$ , that do not undergo fading, *i.e.*, have constant energy. For practical cases where fading is involved this scheme does not show satisfying results. The reason is that the estimate of multiple-access interference is influenced by the estimated channel coefficients of all other users. Small errors propagate very quickly and cause the system to stop its convergence at unacceptable high bit error rates.



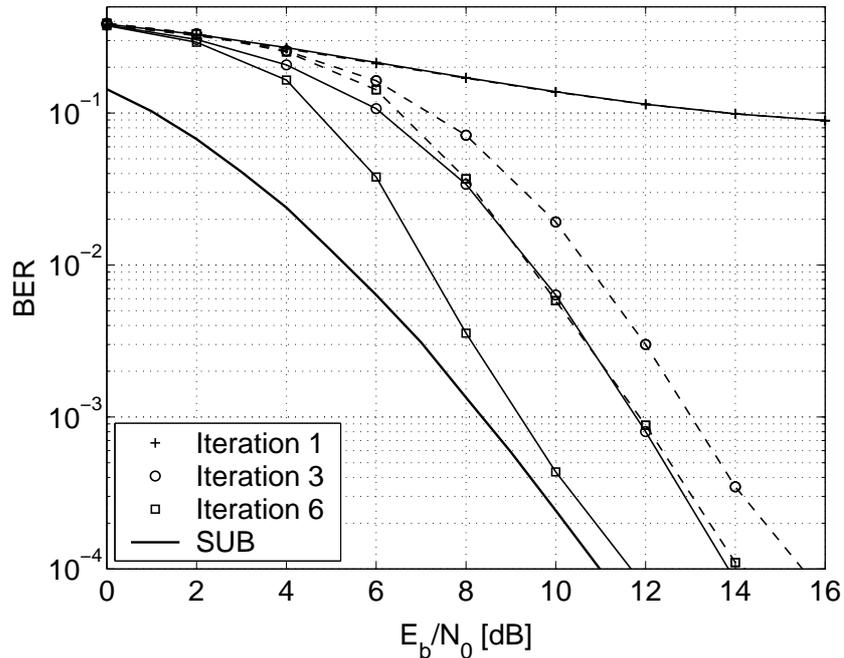
**Figure 4.4:** NMSE for RAND pilots and with channel estimators: ALS (---), ALMMSE (···), and LMMSE (-·-). CRLB1 designates the Cramér-Rao bound for the  $J$  pilots only and CRLB2 for  $M$  symbols. The parameters are  $N = 8$ ,  $K = 12$ , i.i.d. Rayleigh channel with  $L = 5$ ,  $M = 160$ ,  $J = 10$ .

#### 4.4.4 Pilot Sequences

In the following we assess the influence of the random pilot sequence (RAND) and the one of the perfect root of unity sequence (PRUS) on the iterative receiver. The first  $J$  symbols, *i.e.*, the first  $JN$  chips of each block represent the pilot sequence  $\mathbf{u}_k$ . The pilot pattern RAND is based on a random sequence of training symbols that are spread with  $\mathbf{s}_k$  like the data. This is desirable for a practical implementation since all symbols are spread with the same signature. The PRUS scheme uses a chip sounding sequence that is a perfect root of unity sequence.

In the following each user has a channel with  $L = 5$  i.i.d. Rayleigh taps. This requires  $KL = 60$  taps to be estimated and we choose  $J = 10$ , *i.e.*,  $JN = 80$  chips. The corresponding PRUS sequence has  $R = 20$  roots, the parameters used for the construction method presented in Paragraph 4.3.2 are  $p = 2$ ,  $q = 20$ , and  $r = 3$ . To estimate the channel we employ the ALS algorithm.

The case of the RAND pilot scheme was already discussed in Paragraph 4.4.3 and is depicted in Fig. 4.2. We use the same setup to study the behaviour of the multi-user receiver with perfect root of unity sequences. The corresponding results are depicted in Fig. 4.5. Perfect root of unity sequences are, in the mean square error



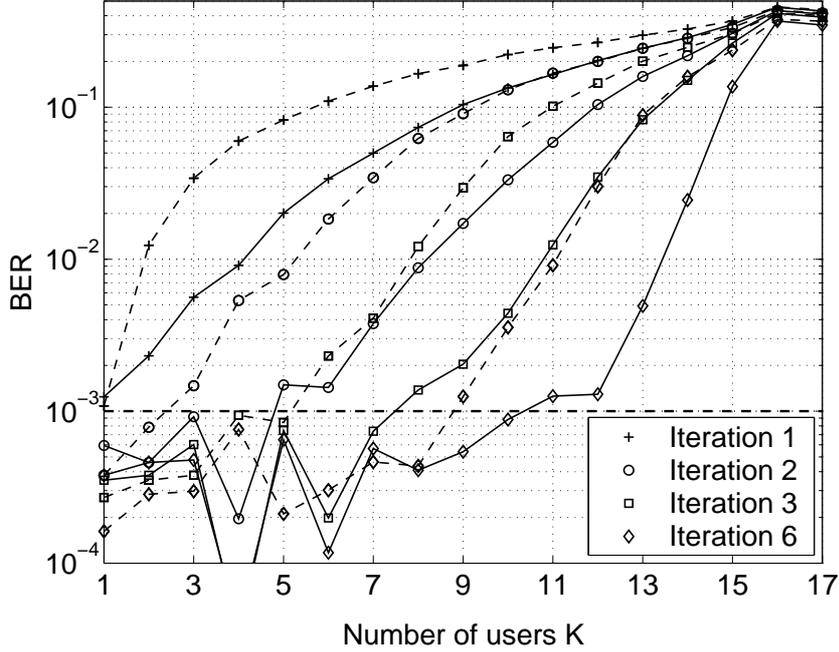
**Figure 4.5:** Comparison of BER with channel estimators PILOTS only (---), ALS  $\approx$  ALMMSE  $\approx$  LMMSE (—). The pilot symbols are PRUS and  $N = 8$ ,  $K = 12$ , i.i.d. Rayleigh channel with  $L = 5$ ,  $M = 160$ ,  $J = 10$ .

sense, the optimal sequences for channel estimation. When only  $J$  pilots are used for channel identification and aiming at a target BER of  $10^{-3}$ , we observe a gain of 4 dB over the RAND scheme, that is depicted in Fig. 4.2, after the third iteration and 3,3 dB after the sixth iteration. However, this gain diminishes when turning to the approach where soft-decision symbols are used. The results for the ALS, ALMMSE, and the LMMSE were practically the same. After the third iteration there is a gain of 0,8 – 2 dB and after the sixth iteration a gain of 0,3 – 0,6 can be observed. The gap to the single-user bound at a BER of  $10^{-3}$  is roughly 1 dB.

We conclude that for the considered load  $\alpha = 1,5$  and a target bit error rate of  $10^{-3}$  the PRUS has a gain of 3,3 dB over RAND in case of channel estimation with pilots only, and shows also an improvement of 1 dB in case when soft decision symbols are used to support channel estimation.

#### 4.4.5 Soft Multi-User Detectors

In Fig. 4.6 the bit error rate is plotted versus the number of users  $K$  for the two detectors PIC-SUMF (3.8) and unconditional, unbiased PIC-LMMSE (3.22) after the first, the second, the third, and the sixth iteration. We set  $E_b/N_0$  to 10 dB and employ the random pilot scheme of Paragraph 4.3.2. For channel estimation we use

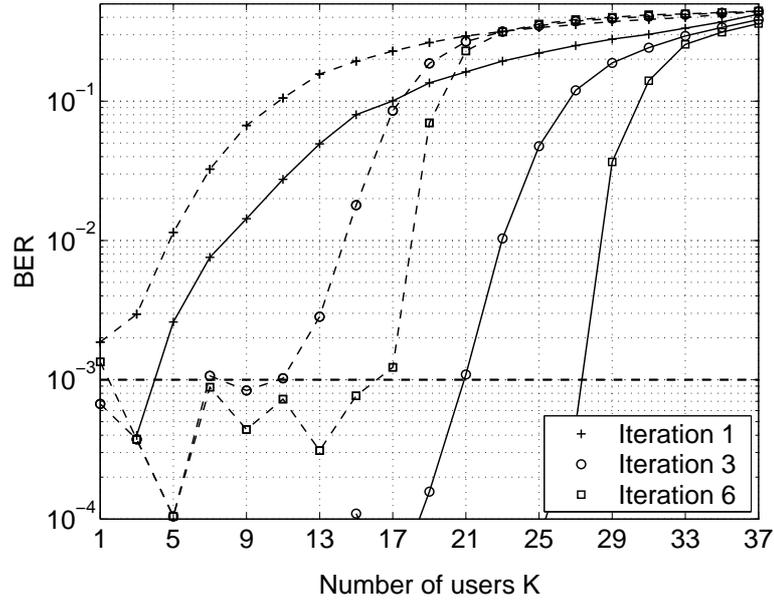


**Figure 4.6:** Impact of detector on number of accommodated users. SUMF (--) and LMMSE (-) with ALS channel estimation and RAND pilots. The system parameters are  $N = 8$ , i.i.d. Rayleigh channel with  $L = 5$ ,  $E_b/N_0 = 10$  dB,  $M = 160$ ,  $J = 10$ .

the ALS algorithm (4.5). The curves reflect that with six iterations and a target BER of  $10^{-3}$  the PIC-LMMSE detector supports twelve users while the PIC-SUMF detector accommodates only nine users. The jaggedness of the BER curves in the bit error rate regime below  $10^{-3}$  is attributed to an insufficient number of Monte-Carlo simulations that were producing only  $K \times 2 \times (M - J) \times 250 = K \times 75,000$  values due to prohibitive simulation time.

#### 4.4.6 Multiple Receive Antennas

In Fig. 4.7 we outline the performance of the PIC-LMMSE detector and the ALS channel estimator when  $N_R = 1$  and  $N_R = 2$  receive antennas are installed when using the random pilot sequence (RAND). We can clearly perceive a diversity gain due to an enhanced stabilization of the receive power. The plot reveals that the number of users to be served can be drastically increased when two receive antennas are used. For two receive antennas roughly 27 users can be accommodated with a BER of  $10^{-3}$ . This is made possible by the increased diversity that is due to spatial antenna combining as well as the spatial signature imposed onto the spreading sequences.



**Figure 4.7:** Impact of multiple receive antennas,  $N_R = 1$  (—) and  $N_R = 2$  (---), on number of accommodated users with ALS channel estimation and RAND pilots. The system parameters are  $N = 8$ , i.i.d. Rayleigh channel with  $L = 5$ ,  $E_b/N_0 = 10$  dB,  $M = 160$ ,  $J = 24$ .

#### 4.4.7 Realistic Channels

The final investigation in this chapter is devoted to the assessment of the PIC-LMMSE receiver with ALS channel estimation in three multi-path scenarios. The first power delay profile is based on the exponentially decaying COST259 model with five taps [Cor01]. The second is the 3GPP test channel Pedestrian-B, specified in [TR101 112]. And the third, FTW, is obtained from a channel sounding measurement campaign conducted in a suburb of Vienna [Hof02]. In the simulations the PDPs fulfill the energy normalization constraint (2.2). They are listed with their temporal delay as well as their relative power with respect to the first tap in Table 4.1. Since oversampling is not considered in this thesis the taps  $\tau[l]$  were assigned to the discrete time value  $n$  that minimizes

$$n = \underset{k}{\operatorname{argmin}} \left\{ \left| \tau[l] - \left( k + \frac{1}{2} \right) T_C \right| \right\}.$$

The COST259 channel with  $L = 5$  is the shortest of the considered channels. The receiver shows that it can reach a BER of  $10^{-3}$  at 11,8 dB compared to the single-user bound at 11,1 dB. This is depicted in Fig. 4.8 for  $J = 16$  RAND pilots.

The Pedestrian-B channel with  $L = 15$  has the largest temporal width. Fig. 4.9 shows the BER curves for a system with  $J = 24$  and RAND pilots. It can be seen

**Table 4.1:** Power delay profiles (PDPs) of the simulated multi-path scenarios with relative power to the first tap.

	TAP 1	TAP 2	TAP 3
COST259	0 dB / 1 ns	-4,3 dB / 245 ns	-8,7 dB / 489 ns
PEDESTRIAN-B	0 dB / 1 ns	-0,9 dB / 200 ns	-4,9 dB / 800 ns
FTW	0 dB / 1 ns	-6,0 dB / 361 ns	-3,6 dB / 1.131 ns
	TAP 4	TAP 5	TAP 6
COST259	-13,0 dB / 733 ns	-17,4 dB / 977 ns	-
PEDESTRIAN-B	-8,0 dB / 1.200 ns	-7,8 dB / 2.300 ns	-23,9 dB / 3.700 ns
FTW	-9,9 dB / 1.991 ns	-7,8 dB / 2.701 ns	-14,1 dB / 3.181 ns

that for the Pedestrian-B channel there is a gap of 3,1 dB to the single-user bound at a target bit error rate of  $10^{-3}$ .

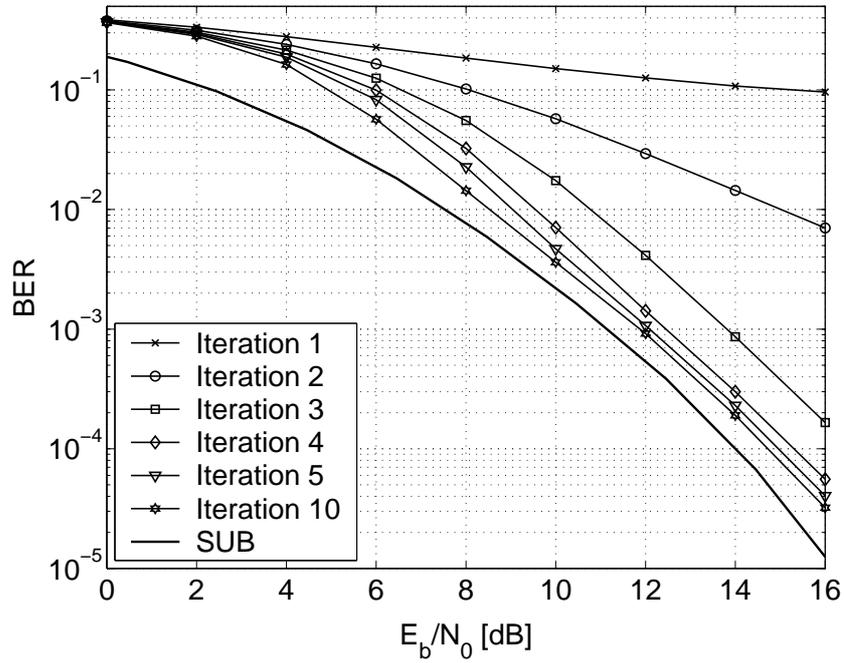
The third channel, FTW, has length  $L = 13$  and it is retrieved from measurements. Fig. 4.10 shows the result for the RAND pilots sequence. In this case the single-user bound is approached very closely.

Depending on the power delay profile, the gap to the single-user bound can make up as much as 3 dB like in the case of the Pedestrian-B channel for a target bit error rate of  $10^{-3}$ . When utilizing PRUS sequences instead of RAND sequences for channel sounding, this gap can be reduced. In case of the Pedestrian-B channel and a target bit error rate of  $10^{-3}$  an improvement of 1,75 dB could be observed. For the COST259 channel the corresponding difference is 0,35 dB and 0,5 dB for the FTW-channel.

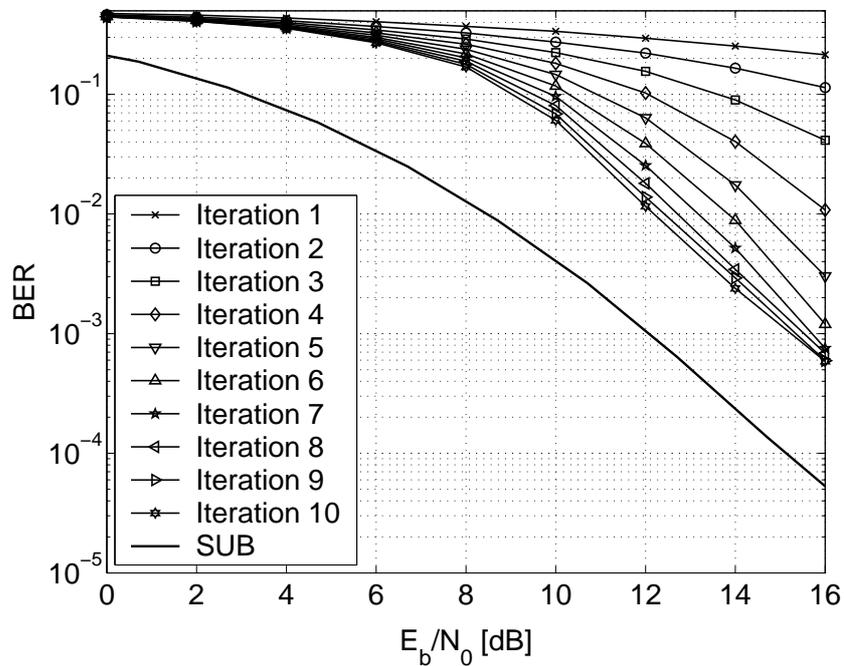
## 4.5 Summary

In this chapter we have developed and analyzed channel estimators that are suitable for integration in iterative receivers. The multi-path channel estimators utilize soft symbol estimates to refine the channel estimates within the iterative process. We have assessed multi-user receivers in terms of bit error rate and mean square error of the channel estimates. The setups included different channel estimators in combination with various detectors, pilot sequences, different number of receive antennas, and realistic multi-path channels. The channel estimators are the approximate LS (ALS) estimator, the approximate LMMSE (ALMMSE) estimator, that uses the covariance matrices of the noise and the one of the channel, and the LMMSE that requires additionally the variance of the soft symbol estimates. We have also investigated an interference canceling and correlation based type estimator.

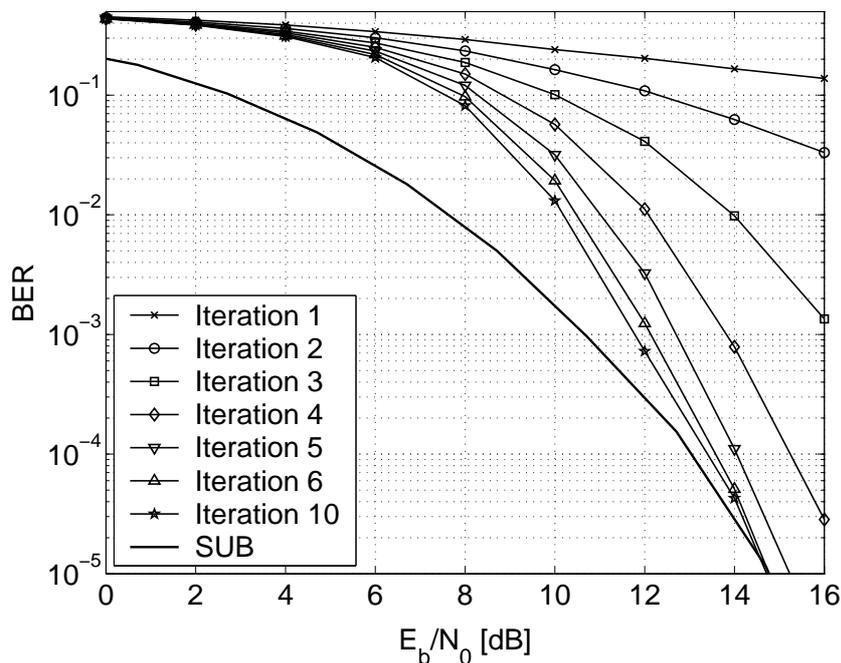
**Feedback Information** Multi-user receivers, that employ *a-posteriori* mapped sym-



**Figure 4.8:** PIC-LMMSE receiver with ALS channel estimation in a COST259 channel with  $L = 5$  paths and RAND pilots.  $K = 12$ ,  $N = 8$ ,  $M = 160$ ,  $J = 16$ .



**Figure 4.9:** PIC-LMMSE receiver with ALS channel estimation in a Pedestrian-B channel with  $L = 15$  paths and RAND pilots.  $K = 12$ ,  $N = 8$ ,  $M = 160$ ,  $J = 24$ .



**Figure 4.10:** PIC-LMMSE receiver with ALS channel estimation in an FTW channel with  $L = 13$  paths and RAND pilots.  $K = 12$ ,  $N = 8$ ,  $M = 160$ ,  $J = 24$ .

bols in the channel estimators, outperform receivers that utilize soft decision symbols gained from *extrinsic* information. *A-posteriori* symbols lead to faster convergence speed in terms of the bit error rate.

**Inter-Symbol Interference:** For an i.i.d. Rayleigh channel with  $L = 2$  taps and spreading factor  $N = 8$ , neglecting ISI degrades performance severely and leads to unacceptable high error floors. With exact ISI processing, as it was utilized in this work, ISI in practical channels with long delays can be processed effectively.

**Channel Estimators:** The ALS estimator does not require statistical *a-priori* information and shows a loss of only 0,5 dB to the more complex LMMSE estimator after six iterations when random pilots are employed. In case of PRUS, the difference between ALS and LMMSE diminishes. The IC-Corr channel estimation approach is not useful since error propagation is too strong in a fading environment with more than three taps.

**Soft-Decision Symbols Supported Estimation:** Using soft decision symbols for channel estimation reduces the  $E_b/N_0$  by 6 dB to reach a bit error rate of  $10^{-3}$  when compared to pilot only based channel identification.

**Pilot Sequences:** In an iterative receiver, where channel estimation is based only on perfect root of unity sequences, bit error rates of  $10^{-3}$  can be reached roughly 4 dB earlier than by random pilot sequences. If soft decision symbols are used as additional pilots, the gap between RAND and PRUS reduces to less than 1,75 dB depending on the channel.

**Detector:** After six iterations and a target BER of  $10^{-3}$  the unbiased, unconditional PIC-LMMSE detector can support roughly 33% more users than the PIC-SUMF detector at  $E_b/N_0 = 10$  dB.

**Receive Diversity:** The number of accommodated users can be nearly doubled by installing two receive antennas instead of a single receive antenna. This is made possible through the additional degrees of freedom offered by spatial diversity.

**Number of Iterations:** Practical convergence of the iterative receiver with channel estimation supported by soft decision symbols is typically observed after six iterations.

# 5 Iterative UMTS Receiver

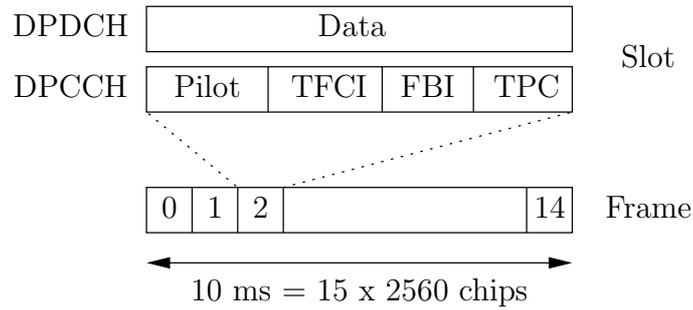
In the previous chapters the performance of several multi-user detectors and channel estimators was assessed in order to identify the minimum required complexity that leads to useful bit error rates for a particular system load and multi-path scenario. In the present chapter, we apply the insight gained in the general setup, to design an iterative receiver that is compliant with the UMTS-FDD standard. The objectives are maximization of the system capacity while at the same time keeping implementation complexity moderately low.

The transmission model in the general setup presented in Section 2.4 and that of the UMTS standard differ in several aspects. The general setup separates the pilot and data symbols by transmitting them consecutively. In UMTS, signalization of data and control information happens simultaneously on the in- and the quadrature phase. While the general setup uses short sequences, the UMTS setup applies short channelization codes separately for the control and the data information and subsequently applies a long scrambling sequence. The general transmitter structure applies a four-state convolutional code with rate  $R_C = 1/2$ . UMTS uses 64-state convolutional codes with rates  $R_C = 1/2$  and  $R_C = 1/3$  or an eight-state Turbo code with  $R_C = 1/3$ .

## 5.1 UMTS-FDD Standard

In UMTS-FDD, signaling between the base-station and the mobile takes place in two frequency bands, each having a bandwidth of 5 MHz. These bands are called paired bands - the one in the upper frequency band is used for downlink communications and the one in the lower frequency band is employed for uplink communications. This physical separation is reflected in the abbreviation FDD - frequency division duplex. Data is spread at a chip-rate of 3,84 Mchips/s.

In the subsequent paragraphs we briefly review the following parts of the UMTS-FDD standard: segmentation [TS25.211], modulation [TS25.213], spreading [TS25.213], and coding [TS25.212].



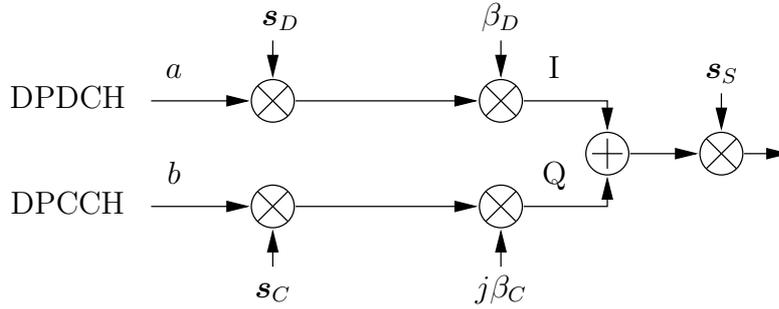
**Figure 5.1:** Structure of slots and frames in UMTS.

### 5.1.1 Data Block Structure

Dedicated channels are assigned to single users. They are segmented into frames of length 10 ms corresponding to 38.400 chips. Every frame consists of 15 slots with 2.560 chips. The segmentation in UMTS is depicted in Fig. 5.1. The dedicated channels are called *dedicated physical data channel* (DPDCH) and *dedicated physical control channel* (DPCCH). The user data is transmitted on the DPDCH while the control information is conveyed via the DPCCH. Control information contains pilot symbols for channel estimation (Pilot), the transport format combination indicator (TFCI), feedback information (FBI), and transmission power control (TPC) commands. For the purposes of this thesis, only the pilot symbols are important and we will not make use of the remaining control information.

### 5.1.2 Spreading and Modulation

Fig. 5.2 illustrates how data spreading and data modulation is exercised in the framework of UMTS. Data symbols  $a$  are spread with a channelization code  $\mathbf{s}_D$  having length  $N_D = 2^{\{2, \dots, 8\}}$ . Control symbols are spread with a different channelization code  $\mathbf{s}_C$  with constant length  $N_C = 256$ . In case of data spreading with factor four, each user can have up to six dedicated physical data channels. The employed codes are from the family of Walsh-Hadamard codes and are selected such that they exhibit mutual orthogonality regardless of the spreading factor. Note that channelization codes separate the individual channels of one particular user, but they are not responsible for separation towards other users. Separation to other users is established after the chipwise combination of the signals from the DPDCH and the DPCCH. The control information is mapped on the quadrature phase and the data on the in-phase after they are amplified with the factors  $\beta_D$  and  $\beta_C$ , respectively. The amplification factors control the relation of the power devoted to the control and the data channel. The combined chip stream is further scrambled by the sequence  $\mathbf{s}_S$  of length 38.400. This is a truncated Gold chip sequence uniquely assigned to



**Figure 5.2:** Modulation and spreading in UMTS.

each user by the base-station. After chipwise multiplication of the channelized chip stream with the scrambling code, the data is transmitted over the channel. Due to the huge number of available scrambling codes, there is no limit imposed on the number of users in the uplink. Channelization codes are *short* codes in the sense that they are reused for the spreading of every symbol. Scrambling sequences on the other hand are *long* - there is no repetitive pattern on the symbol level that can be exploited. At the receive antenna, chip sequences attributed to a particular data or control symbol are distorted by interference of other users. The cross-correlation between symbols of different users varies from symbol to symbol due to the scrambling sequences.

Channelization and addition of in-phase and quadrature components on the chip level of user  $k$  can be expressed as

$$x_k[n] = \beta_{D,k} a_k(\lceil n/N_{D,k} \rceil) s_{D,k}[n \bmod N_{D,k}] \\ + j\beta_{C,k} b_k(\lceil n/N_C \rceil) s_{C,k}[n \bmod N_C]$$

with  $\lceil x \rceil$  denoting to closest integer value to  $x$  that is equal or larger than  $x$ . The channelized chip stream is further scrambled with the user unique scrambling sequence  $s_{S,k}[n]$

$$x'_k[n] = x_k[n] s_{S,k}[n \bmod 38.400]. \quad (5.1)$$

### 5.1.3 Channel Coding

In UMTS, channel coding is applied over a block whose length is specified by the *transmission time interval* (TTI). The minimal value of the TTI is 10 ms and its maximum value is 80 ms. They correspond to 15 and 120 slots, respectively. In case of slots having different channel realizations we expect an additional diversity gain (block interleaving) that lowers the bit error rates.

The UMTS-standard contains three codes: two non-recursive convolutional codes with rate 1/2 and rate 1/3 and one Turbo code with recursive generator polynomials

**Table 5.1:** Channel codes in UMTS: CC – convolutional code, TC – Turbo code.

	CC 1/2	CC 1/3	TC
RATE $R_C$	1/2	1/3	1/3
MEMORY LENGTH $\iota$	8	8	3
GENERATORS	$[561, 753]_8$	$[557, 663, 711]_8$	$\left[1, \frac{15}{13}, \left(\frac{15}{13}\right)'\right]_8$
REQUIRED $E_b/N_0$ FOR BER= $10^{-3}$	2,4 dB	1,95 dB	0,95 dB
MAX. INPUT BLOCK SIZE	504	504	5.114

and rate 1/3. The codes and their properties are listed in Table 5.1. The length of the information sequence, that is encoded to deliver a codeword, is limited by the standard to 504 bits for the convolutional codes and 5.114 bits for the Turbo code. In Table 5.1 we obtained the values of  $E_b/N_0$ , that are required to achieve a bit error rate of  $10^{-3}$ , by simulations on an AWGN channel. For the two convolutional codes, we employed a Viterbi decoder and for the Turbo code we used the max-log-MAP approximation [Bau98] of the BCJR algorithm [Bah74]. The input blocks in the simulation had a length of 504 for the convolutional code and 1.596 for the Turbo code. For the Turbo code, the required  $E_b/N_0$  depends on the size of the code block. The value decreases for longer block lengths. Each code block is followed by terminating tail bits. The details of those are given in [TS25.213].

## 5.2 Receiver Design

The following passage describes the received signal and discusses the choice of the receiver elements, based on the observations of the preceding chapters.

### Transmission Model

The general setup used in the previous chapters makes use of short random sequences that spreads the subsequently transmitted pilot and data symbols. This prerequisite allowed us to find an elegant expression of the received signal as given in (2.12) and (2.13). There are two major differences to UMTS: (1) UMTS uses long sequences and (2) the pilot information is spread with a different channelization code than the data and both are modulated orthogonally to each other. Further, a long scrambling sequence is applied. For the sake of simplicity we give the expression for one receive sample rather than for the whole observation vector  $\mathbf{y}$ . The received chip stream

$$y[n] = \sum_{k=1}^K \sqrt{P_k} \sum_{q=0}^{L-1} x'_k[n-q] h_k[q] + v[n] \quad (5.2)$$

is a superposition of  $K$  user contributions (5.1) and zero-mean additive white Gaussian noise  $v[n]$  with variance  $\sigma_v^2$ . Every user transmits via a channel  $h_k[n], n \in \{0, \dots, L-1\}$  and its transmission power is  $P_k$ .

### Inter-Symbol Interference Processing

Paragraphs 3.8.2 and 4.4.2 study the impact on the bit error rates when inter-symbol interference is neglected. It turns out that even for short channels the loss can be dramatic. In cases where the spreading factor  $N_D < 32$ , inter-symbol interference needs to be considered in order to approach the single-user bound. Hence, our iterative UMTS receiver shall incorporate this feature.

### Detector

In Paragraph 3.8.4 it was shown that parallel interference cancellation allows for the same number of accommodated users as successive interference cancellation. This comes at an expense of additional iterations in the case of parallel cancellation. For a fading environment, SIC requires two iterations to come close to the single-user bound while PIC processing requires four to five. Hence, from the viewpoint of number of iterations, SIC seems more appealing. However, SIC is associated with a long latency time. For the present receiver design it is assumed that latency time is a more critical point than distributed parallel processing power. Hence, we decided for PIC scheduling.

The unconditional, unbiased LMMSE-filter given by (3.22) allows for an efficient implementation since the involved matrix inverse can be used for all users and all symbols in the same iteration step. This holds only, when the spreading sequences are short. Due to the long scrambling in UMTS, the sequences change from symbol to symbol and calculation of the filter is required for every symbol instance. This is an additional complexity that scales linearly with the length of a block. The increased number of updates renders LMMSE as too demanding and we decided to apply a linear single-user matched filter instead.

### Decoding

The 64-state convolutional codes with rates  $1/2$  and  $1/3$  are too complex to decode with the optimum symbol by symbol MAP decoder presented in Section 2.10. Instead, we apply a hard-output Viterbi decoder. The estimated hard information bits are re-encoded and fed back to the interference canceler as well as to the channel estimator. Improved performance is achieved when the hard feedback information from the Viterbi decoder is weighted as it was done in [Mar01, Nor02]. However, the latter is not applied in the present work. In case of the Turbo code, there is an

additional inner decoding loop that requires a MAP decoder. The number of these inner iterations was set to eight. Due to the short memory length of the Turbo code, MAP decoding does not impose a computational constraint.

### Channel Estimator

The results in Chapter 4 make evident that the approximate least squares algorithm is most practical. In contrast to the approximate linear MMSE and the exact linear MMSE scheme, it does not require any second-order statistical information on neither the noise nor the channel coefficients. In UMTS, the pilot data is no longer contained in a preamble, but resides in the quadrature path and is sent in parallel to the data which is in the in-phase. We adapt the ALS estimator correspondingly.

## 5.3 Processing Chain for Simulation

In the general setup that was used in the previous chapters, the codeword length was exactly the size of a transmission block minus the pilot sequence length. One codeword effectively undergoes one particular fading realization. The individual blocks face independent block fading realization. Now, in UMTS, code symbols are split among several slots that undergo different fading realization. In general, this causes a so called block interleaving gain. To simulate block interleaving, we need to adapt the simulation chain. The modified processing chain for UMTS is depicted in Fig. 5.3. Data  $\mathbf{d}_k$  is generated for  $K$  users in accordance with the spreading factor  $N_D$  and the number of slots  $NoS$  which is itself determined through the *transmission time interval* (TTI). For simplicity we assume that all users have the same spreading factor  $N_D$ . We create multi-path channels  $\mathbf{h}_k(s)$ , that are independent from slot to slot, and convolve them with the chip sequences  $x'_k[n]$  given in (5.1). The contributions of all users and the noise are summed up as formulated in (5.2). After the observation vectors  $\mathbf{y}(s)$  of all slots have been generated, they are passed on to the receiver. The processing is carried out over  $NoI$  iterations. After all symbols have been detected for a particular user, they are passed on to the decoder. The decoder outputs values  $\hat{\mathbf{d}}_k$  that are re-encoded to form estimated code symbols  $\hat{\mathbf{c}}_k$ . These are used for interference cancellation and channel estimation during the next iteration. In each iteration the bit error rate is evaluated.

## 5.4 Simulation Results

In the general transmission setup, with pilots in the preamble, the loss in information energy due to the overhead was considered via (2.14). Similarly, we introduce a

```

NoF := number of simulation frames
NoI := number of iterations
NoS := number of slots per TTI

for f = 1:NoF
  - Generate data  $\mathbf{d}_k$ 
  - Encode data  $\mathbf{d}_k \mapsto \mathbf{c}_k$ 
  - Split  $\mathbf{c}_k$  to NoS  $\mathbf{c}_k(s)$ 
  for s = 1:NoS
    - Generate multi-path channels  $\mathbf{h}_k(s)$ 
    - Generate  $\mathbf{y}(s)$  from all  $\mathbf{c}_k(s)$  and  $\mathbf{h}_k(s)$ 
  end s
  for i = 1:NoI
    for s = 1:NoS
      - Estimate channels  $\hat{\mathbf{h}}_k(s)$ 
      - Perform IC and linear filtering to obtain  $\hat{\mathbf{c}}_k(s)$ 
    end s
    - Assemble  $\hat{\mathbf{c}}_k$  from NoS  $\hat{\mathbf{c}}_k(s)$ 
    - Decode  $\hat{\mathbf{c}}_k$  to obtain  $\hat{\mathbf{d}}_k$ 
    - BER evaluation of  $\hat{\mathbf{d}}_k$ 
    - Re-encode  $\hat{\mathbf{d}}_k \mapsto \hat{\mathbf{c}}_k$ 
  end i
end f

```

**Figure 5.3:** Structure of the UMTS simulator.

correction factor for UMTS modulation. Given the  $E_b/N_0$  and system noise variance  $\sigma_v^2$ , the transmit powers of the individual users become

$$P_k = 2\sigma_v^2 R_C \frac{E_b}{N_0} \frac{N_C/N_D}{\beta_C^2 E_C + \beta_D^2 E_D N_C/N_D}.$$

In the following we define the overhead OH as the fraction of the total transmit energy that is spent on control information. The overhead is given by

$$\text{OH} = \frac{\beta_C^2 E_C}{\beta_C^2 E_C + \beta_D^2 E_D N_C/N_D}$$

with  $E_C = |\mathbf{s}_C|^2 = 1$  and  $E_D = |\mathbf{s}_D|^2 = 1$  being the energy of the spreading sequence used in the control and the data path, respectively. The values  $\sqrt{P_k}\beta_C$  and  $\sqrt{P_k}\beta_D$  denote the amplitude factors of the control and the data path. The scrambling sequence  $s_S[n]$  has amplitude one and does hence not influence the normalization.

We investigate the influence of coding for the simulation setup summarized in Table 5.2. The number of pilot symbols per slot is eight and the overhead energy makes up 11,1% for the rate 1/2 code and 15,8% for the simulations with rate 1/3 codes. The multi-path channel employs the exponentially decaying power delay profile of the COST259 channel covering five relevant taps.

Fig. 5.4 and Fig. 5.5 show the bit error rate curves versus  $E_b/N_0$  for the convolutional codes with rate  $R_C = 1/2$  and rate  $R_C = 1/3$ , respectively. For the code with rate  $R_C = 1/2$ , we have observed that the improvement stops after the fifth iteration. Comparing the BERs with the single user bound makes clear that the loss is less than 1 dB in the relevant regime where bit error rates are lower than  $10^{-3}$ . When studying the result for the convolutional code with rate  $R_C = 1/3$ , we notice that the main difference lies in the accelerated convergence. It is also clear from intuition - a code with a lower rate carries more redundancy that can be exploited with the cost of a lower net data rate. For the considered setup, we save one iteration compared to the scheme with rate  $R_C = 1/2$ . Both decoders attain the bit error rate of  $10^{-3}$  at roughly 11 dB.

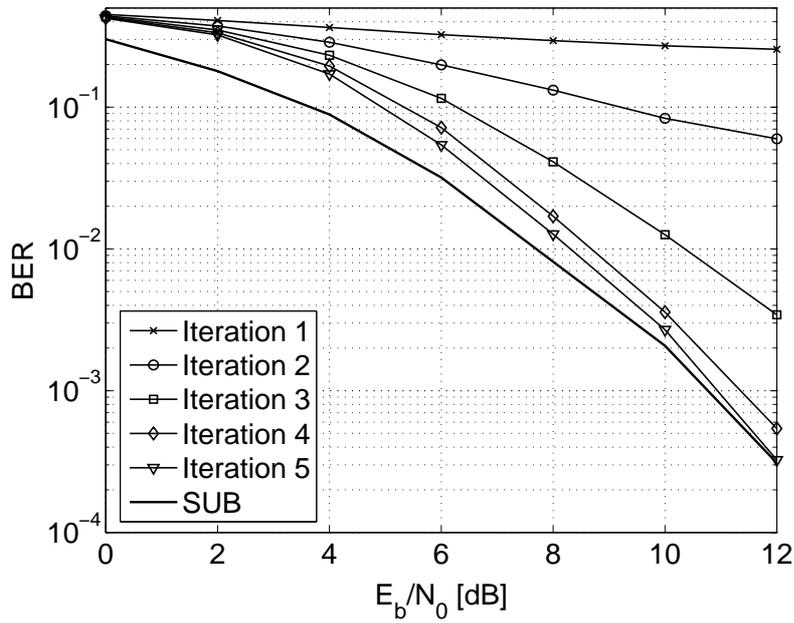
For the Turbo code we obtain the result depicted in Fig. 5.6. In each iteration of the iterative receiver, the Turbo decoder performs eight internal iterations. As long as the first iteration does not cause bit error rates that are smaller than 0,14, the Turbo decoder is not able to converge. However, as soon as this threshold is reached, a water-fall effect takes place that decreases the bit error rate drastically. A target bit error rate of  $10^{-3}$  is reached within four iterations at an  $E_b/N_0$  of 9 dB. While for the convolutional codes, the improvement stops after five iterations, the Turbo code leads to a further decrease in bit error. After ten and 15 iterations, bit error rates of  $2,2 \times 10^{-3}$  and  $10^{-4}$  are reached. The gap to the single-user bound is roughly 2 dB. The effectivity of the multi-user detector to remove multiple-access interference is poor at very low levels of  $E_b/N_0$  and it cannot produce sufficiently reliable uncoded symbols that cause the decoder to converge. Despite the increased complexity, the potential gain over the rate 1/3 convolutional code makes up roughly 6 dB for a target bit error rate of  $10^{-3}$ .

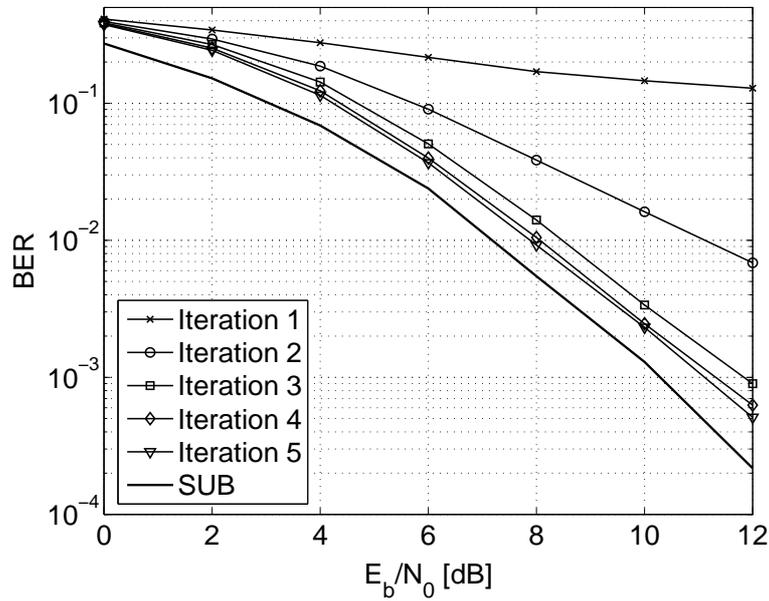
## 5.5 Summary

We have designed an iterative receiver for the UMTS-FDD standard and have demonstrated its functionality in the COST259 multi-path channel with  $L = 5$  taps. The core elements were chosen to exhibit low-complexity such as the parallel interference canceling detector with the single-user matched filter, the hard Viterbi output decoder, and the approximate least squares channel estimator. An extension to multiple receive antennas was tested and it essentially showed similar behaviour

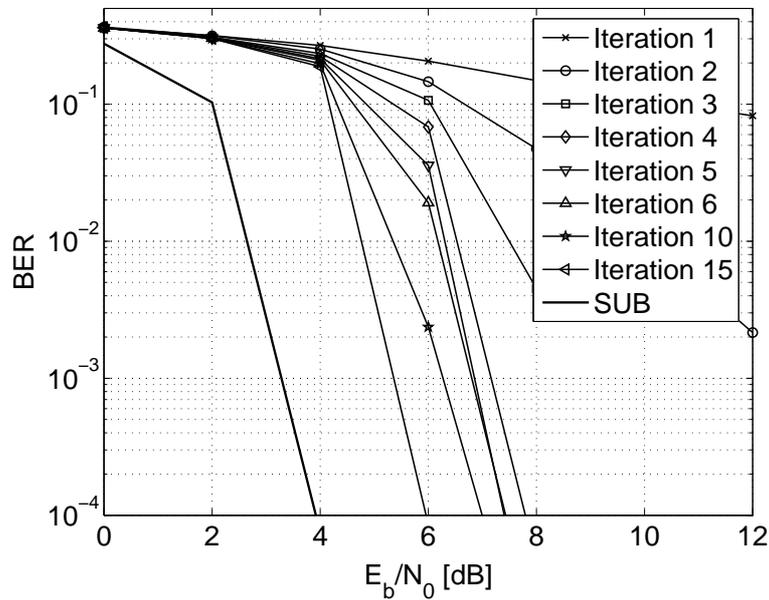
**Table 5.2:** Parameters for UMTS Uplink Simulation.

PARAMETER	VALUE
Number of users $K$	12
TTI	10 ms
$N_{\text{DCH}}, N_{\text{CCH}}$	8, 256
$\beta_c/\beta_d$	$\sqrt{2}$
OH $R_C = 1/2, R_C = 1/3$	11,1 %, 15,8 %
Number of pilot symbols per slot	8
Number of COST259 taps $L$	5
Pulse shaping	off

**Figure 5.4:** Average bit error rates for  $K = 12$  users having spreading factor  $N_{\text{DCH}} = 8$  on a COST259 channel with  $L = 5$  taps. All users employ a convolutional code with rate  $R_C = 1/2$  (CC 1/2).



**Figure 5.5:** Average bit error rates for  $K = 12$  users having spreading factor  $N_{\text{DCH}} = 8$  on a COST259 channel with  $L = 5$  taps. All users employ a convolutional code with rate  $R_C = 1/3$  (CC 1/3).



**Figure 5.6:** Average bit error rates for  $K = 12$  users having spreading factor  $N_{\text{DCH}} = 8$  on a COST259 channel with  $L = 5$  taps. All users employ a Turbo code with rate  $R_C = 1/3$  (TC).

as the more general setup presented in Paragraph 4.4.6.

The critical issue that will be faced in practice is that users transmit asynchronously in the uplink. In case of large TTI values, the codeword can comprise up to 120 slots. Then, only fractional cancellation of user contributions (MAI) can be envisaged which results in a deteriorated performance. However, in the *high speed downlink packet access* (HSDPA) mode of UMTS [TR25.858] all users are synchronous with spreading factor 16 and the TTI is fixed to three slots. Therefore, the iterative receiver can be fully employed. Since one user can be assigned up to 15 parallel data channels, decoding of all individual streams is required *a-priori*, and this supports the iterative receiver structure in a natural fashion.



## 6 Iterative Receiver Analysis

In the previous chapters we assess the convergence behaviour of iterative receivers including channel estimation by numerical means. The associated simulations require a large number of simulation frames to deliver meaningful results. This is time costly and, thus, it is beneficial to gain insight from an analytical point of view. This is a difficult problem, caused by the following two factors: First, there is no tractable exact formulation for the relation of bit error rate and signal to noise ratio for a given channel code. Second, the involved soft decision symbols that are the substantial part of the iterative processing cause a mismatched situation in the channel estimator as well as in the multi-user detector. Results for the output signal to noise and interference ratio in mismatched situations were published only recently [Cai04]. We will extend on these results and present a semi-analytical approach to predict the convergence behaviour of the iterative receiver by resorting to results from statistical physics.

A first analysis on iterative receivers with channel estimation was reported by Alexander and Grant in [Ale00b] where a certain amount of the power is devoted to pilot symbols. It uses an approximation for the decoder behaviour, and the exchanged measures are symbol output variances. Lampe presented an analysis [Lam03] that makes use of results from random matrix theory for channel estimates that are based on dedicated pilots but not on soft decision symbols. Recently, a technique called density evolution was successfully applied to accurately describe the behaviour of an iterative receiver with symbols transmitted on an AWGN channel [Bou02]. We extend that analysis to include channel estimation for a single channel tap with random phase and constant amplitude.

We include the soft decision symbols gained from the output of the decoder to enhance channel estimation as it was done in Chapter 4. This results in a mismatch of the symbol power distribution that is due to the fact that the soft decision symbols differ from the true data symbols. The mismatched behaviour can be described by deploying results from statistical physics. We conduct a semi-analytical investigation, where the bit error rate of the channel code is derived from simulations and the infinite number of users is replaced by a finite number of users. As a result the analysis delivers the achievable multi-user efficiency for a given symbol block length and amount of energy spent on pilot symbols.

## 6.1 Signal Model

We consider the uplink of a chip-synchronous CDMA system with  $K$  users as it is depicted in Fig. 2.5. Every user is assigned one single fading tap designated by  $h_k$ . The assumption is again that the channel is block fading meaning that it remains constant during the transmission of  $M$  symbols. The block of  $M$  symbols  $b_k(m)$ ,  $m \in \{0, 1, \dots, M-1\}$  of each user are QPSK modulated. All of them are spread by random signature sequences  $\mathbf{s}_k(m) \triangleq [s_k[0], \dots, s_k[N-1]]$ ,  $k \in \{1, \dots, K\}$  with spreading factor  $N$ . For the purpose of the analysis, the sequences are long, meaning that they are different for every symbol interval. This ensures that all chips are i.i.d. which is a requirement to apply results from random matrix theory and the replica method. In contrast to our previous considerations, where the chips were randomly chosen from the discrete set  $\mathbb{X}/\sqrt{N}$ , the chips are now i.i.d., zero-mean, complex Gaussian distributed

$$\mathcal{CN}(0, 1/N) \triangleq \mathcal{N}(0, 1/(2N)) + j\mathcal{N}(0, 1/(2N)).$$

Each block of symbols contains a preamble with  $J$  randomly chosen pilot symbols from the set  $\mathbb{X}$ . This choice corresponds to the random pilot symbol sequence presented in Paragraph 4.3.2. The remaining  $M - J$  symbols are obtained by mapping the encoded data stream  $c_k(m')$  onto a QPSK constellation with Gray labeling after random interleaving. The encoded data stream  $c_k(m')$  is obtained from encoding the raw data stream  $d_k(m'')$  with a code  $\mathcal{C}_k$ .

Restricting the matrix model (2.13) to the flat fading case the received vector  $\mathbf{y} \in \mathbb{C}^{MN \times 1}$  is written as

$$\mathbf{y} = \mathbf{D}\mathbf{B}\mathbf{h} + \mathbf{v} \quad (6.1)$$

where

- $\mathbf{D} \in \mathbb{C}^{MN \times KM}$  is a block diagonal matrix  $\text{diag}[\mathbf{S}(0), \mathbf{S}(1), \dots, \mathbf{S}(M-1)]$ .  $\mathbf{S}(m) \in \mathbb{C}^{N \times K}$  denotes the spreading matrix at discrete time  $m$ . It is composed as  $[\mathbf{s}_1(m), \dots, \mathbf{s}_K(m)]$  with the signature sequence  $\mathbf{s}_k(m) \in \mathbb{C}^{N \times 1}$  of user  $k$  and symbol  $m$  placed in the  $k$ -th column.
- The matrix  $\mathbf{B} \triangleq [\mathbf{B}_P^T, \mathbf{B}_D^T]^T$  has dimension  $KM \times K$  and is defined by the vertically stacked matrices  $[\mathbf{B}(0), \mathbf{B}(1), \dots, \mathbf{B}(M-1)]^T$ . The first  $J$  matrices  $\mathbf{B}(m)$  make up the pilot part  $\mathbf{B}_P \in \{\mathbb{X}, 0\}^{JK \times K}$ , and the remaining  $M - J$  matrices denote the part due to the data  $\mathbf{B}_D \in \{\mathbb{X}, 0\}^{(M-J)K \times K}$ .  $\mathbf{B}(m)$  is a diagonal matrix of dimension  $K \times K$ . The index  $m \in \{0, \dots, J-1\}$  accounts for the pilot symbols and  $m \in \{J, \dots, M-1\}$  for the code symbols.

- The vector  $\mathbf{h} \in \mathbb{C}^{K \times 1}$  denotes the  $K$  channel coefficients of all the users. The coefficients have a random phase and unit power in each realization, *i.e.*,  $|h_k|^2 = 1$ .
- The noise vector  $\mathbf{v} \in \mathbb{C}^{MN \times 1}$  has zero-mean, spherically invariant complex Gaussian distributed samples  $\mathcal{CN}(0, \sigma_v^2)$ .

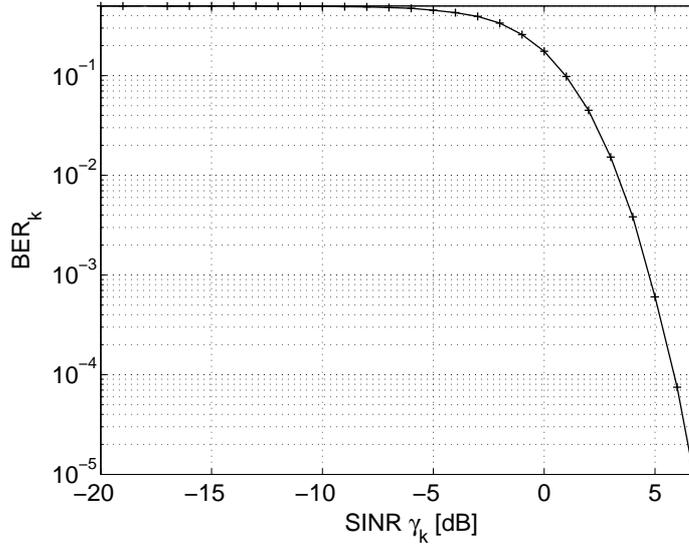
For the system  $E_b/N_0$  we use definition (2.14) that takes into account the loss due to pilot symbols.

## 6.2 Density Evolution

This analysis describes the convergence behaviour of the iterative receiver including channel estimation. The involved soft decision symbols and the channel estimates are represented through densities. Hence, this form of analysis is called density evolution. To explain the idea we present the “big picture” conveyed through Fig. 6.1. The analysis does not use any observed measures except the relation between bit error rate and the signal to noise ratio for the considered channel code  $\mathcal{C}_k$ . We study the large system case, meaning that the number of users  $K$  and the spreading factor  $N$  tend to infinity at a fixed ratio  $K, N \rightarrow \infty, K/N = \alpha$ . This has the advantage that the average receiver performance, expressed in terms of its multi-user efficiency, can be described analytically and does not depend on the particular realization of signature sequences any longer.

Our aim is to describe the multi-user efficiency  $\Psi$  of the multi-user detector, defined in Paragraph 2.8.1, after one iteration, given an assumed multi-user efficiency  $\eta$ . In the large system case the multi-user efficiency  $\eta$  is the same for all users since they all experience the same effective interference [Tse99]. The effective SINR at the input of the decoder for user  $k$  is given by  $\gamma_k = \eta|h_k|^2/\sigma_v^2$ . It is associated with the parameter  $\mu_k$  that fully describes the distribution of all log-likelihood values at the output of the decoder for the assumption that all code bits are +1. This condition does not restrict the validity of the results but helps us to find a simple description of the decoder output. The decoder output density is conveyed to the channel estimator that outputs the individual distributions of all channel estimates. Latter are characterized by the mean of the estimate  $\bar{h}_k$  and its mean square error  $\xi_k^2$ . Thus, we obtain again a density. The densities of the log-likelihood ratios and the channel estimates are passed on to the multi-user detector and the new multi-user efficiency  $\Psi$  is evaluated. This procedure is repeated for several values of  $\eta \in [0, 1]$  such that we obtain the relation  $\Psi(\eta)$  from which the convergence of the receiver can be predicted.





**Figure 6.2:** Bit error rate for code bits based on *extrinsic* log-likelihood ratios. The code  $\mathcal{C}$  is a four-state convolutional code with rate  $1/2$  and polynomials  $(5, 7)_8$ .

efficiency  $\eta$ , the system noise variance  $\sigma_v^2$ , and the channel realization  $h_k$ . In the analysis we model the input to the decoder as the output of an AWGN channel

$$z_k(m') = b_k(m') + \rho_k(m')$$

with  $\rho_k(m') \sim \mathcal{N}(0, 1/\gamma_k)$  where

$$\gamma_k \triangleq \eta |h_k|^2 / \sigma_v^2 \quad (6.3)$$

denotes the *effective* signal to interference and noise ratio of user  $k$ . This quantity can range between zero and  $|h_k|^2 / \sigma_v^2$  and directly reflects the capabilities of the multi-user detector to mitigate multiple-access interference via its multi-user efficiency  $\eta$ . The decoder output bit error rate of the code bits is a function of the particular code  $\mathcal{C}_k$  and the effective output SINR  $\gamma_k$  of the multi-user detector. The input-output relation  $\text{BER} = e_s(\mathcal{C}, \gamma_k)$  of the code has to be derived by simulations since no exact analytical expressions exist. For the convolutional code with generator polynomials  $(5, 7)_8$  the corresponding bit error rate curve based on the *extrinsic* output code bits is depicted in Fig. 6.2. When the BER is evaluated along  $\hat{c}'(m') = \text{sign}(\mathcal{L}(m'))$ , it can be expressed in terms of the SINR  $\mu_k^2 / (2\mu_k) = \mu_k / 2$  of the LLR output distribution (6.2) along

$$\text{BER}_k = Q\left(\sqrt{\frac{\mu_k}{2}}\right),$$

such that we can solve for the parameter  $\mu_k$  as

$$\mu_k = 2 \left(Q^{-1}(\text{BER}_k)\right)^2. \quad (6.4)$$

## 6.4 Mismatched Channel Estimation

The receiver feeds back *extrinsic* information retrieved from code bits to the channel estimator in each iteration. This information is used to improve the channel estimate by means of transferring reliability information into soft decision symbols that act as additional pilot symbols. The log-likelihood ratio values  $\mathcal{L}(m')$ , which range from “ $-\infty$ ” to “ $+\infty$ ”, are mapped to soft decision symbols  $\tilde{c}(m')$  in the confined interval  $[-1, +1]$ , depending on their reliability. The mapping function  $f(\mathcal{L})$ , that minimizes the mean square error  $\mathbb{E}_{\mathcal{L}} \{(f(\mathcal{L}(m')) - c'(m'))^2\}$ , is  $\tanh(\mathcal{L}(m')/2)$  [Sgr01]. The QPSK code symbols are finally obtained from two consecutive LLR values along

$$\tilde{b}_k(m) = \left( \tanh \left( \frac{\mathcal{L}'_k(2m)}{2} \right) + j \tanh \left( \frac{\mathcal{L}'_k(2m+1)}{2} \right) \right) / \sqrt{2}. \quad (6.5)$$

For this analysis we employ the approximate LMMSE (ALMMSE) channel estimator (4.6) that we have derived in Paragraph 4.1.3. The ALMMSE estimator is obtained when we assume that the channel vector  $\mathbf{h}$  and the received signal  $\mathbf{y}$  are zero-mean and that the deterministic code symbol matrix  $\mathbf{B}$  in (6.1) is replaced with the mixture of known pilot symbols and mapped soft code symbols  $\tilde{\mathbf{B}} \triangleq [\mathbf{B}_P \tilde{\mathbf{B}}_D]^T$ . For the flat Rayleigh fading channel with all taps  $h_k, k \in \{1, \dots, K\}$  having variance one, the ALMMSE estimator (4.6) reads

$$\hat{\mathbf{h}} = \left( \tilde{\mathbf{B}}^H \mathcal{D}^H \mathcal{D} \tilde{\mathbf{B}} + \sigma_v^2 \mathbf{I}_K \right)^{-1} \tilde{\mathbf{B}}^H \mathcal{D}^H \mathbf{y} = \left( \tilde{\mathbf{A}}^H \tilde{\mathbf{A}} + \sigma_v^2 \mathbf{I}_K \right)^{-1} \tilde{\mathbf{A}}^H \mathbf{y}. \quad (6.6)$$

The channel estimator (6.6) is mismatched since it assumes that the soft decision symbols  $\tilde{b}(m)$  are the true data symbols  $b(m)$ . It is only matched if the  $M - J$  data symbols were perfectly known.

The statistical description of the soft decision data, characterized through  $\mu_k$  (6.4), prohibits us from finding an exact expression for  $\hat{h}_k, k \in \{1, \dots, K\}$ . Rather than looking at a particular outcome of the channel estimator, we now aim at finding an expression that describes the distribution of the estimate  $\hat{h}_k$ . If we had soft decision symbols with constant modulus we could use results from random matrix. It was shown in [Eva00] how under these circumstances a distribution of  $\hat{h}_k$  can be found. However, in our analysis, the soft decision symbols do not have constant modulus and thus the true output SINR would be too optimistic. Fortunately, a recently developed generalization [Mül04a] of Tanaka’s replica method [Tan02a] to analyze large CDMA systems provides the tool to tackle SINR prediction in mismatched channel estimators. We have found the following conjecture.

**Conjecture 1** *The SINR  $\beta_k$  of the mismatched channel estimator (6.6) for user  $k$  converges, as the number of users  $K$  and the spreading factor  $N$  grow large with*

their ratio  $\alpha = K/N$  fixed, to

$$\beta_k = \lim_{k \rightarrow \infty} \frac{\tilde{E}_k^2}{\tilde{F}_k} \quad (6.7)$$

where  $\tilde{E}_k$  and  $\tilde{F}_k$  are the unique solutions to the systems of fixed point equations (F.7) and (F.8) that are given in Appendix F.

A rigorous proof of Conjecture 1 is yet unknown. However, Monte Carlo simulations with real systems indicate an excellent match between the average observed and predicted signal to noise ratio. This evidence is presented in Appendix F.

With Conjecture 1 and results from Guo and Verdú [Guo02], the conditional mean becomes asymptotically

$$\bar{h}_k = \frac{\tilde{E}_k}{1 + \tilde{E}_k} h_k \quad (6.8)$$

and the *mean square error* (MSE), *i.e.*, the variance of the estimated fading tap  $\hat{h}_k$ , is expressed as

$$\xi_k^2 = \frac{1 + \tilde{F}_k}{(1 + \tilde{E}_k)^2}. \quad (6.9)$$

In this case the distribution of the channel estimate  $\hat{h}_k$ , conditioned on its true value  $h_k$ , is given by the complex Gaussian distribution

$$f_{\hat{h}_k}(\hat{h}_k|h_k) = \frac{(1 + \tilde{E}_k)^2}{\pi|h_k|^2(1 + \tilde{F}_k)} \exp\left(-\frac{|(1 + \tilde{E}_k)\hat{h}_k - \tilde{E}_k h_k|^2}{(1 + \tilde{F}_k)|h_k|^2}\right). \quad (6.10)$$

The conditional mean  $\bar{h}_k$  and its variance  $\xi_k^2$ , together with the distribution of  $h_k$  (which is known by assumption in the analysis), fully characterize the distribution of the estimate  $\hat{h}_k$ . This distribution is needed to evaluate the performance of the multi-user detector suffering from imperfect channel knowledge.

## 6.5 Mismatched Multi-User Detection

For soft data detection we deploy parallel interference cancellation followed by linear MMSE filtering that was introduced in Paragraph 3.4.2. In a real receiver, the input measures to the detector are the received vector  $\mathbf{y}$ , the soft decision symbols  $\tilde{\mathbf{B}}$  based on *extrinsic* information, the channel estimate  $\hat{\mathbf{h}}$ , and the spreading matrix  $\mathcal{D}$ . In the analysis we need the parameters describing the statistics of the log-likelihood feedback values, *i.e.*,  $\mu_k$  (6.4), the mean square error of the channel estimate  $\xi_k^2$  (6.9), the mean of the channel estimate  $\bar{h}_k$  (6.8), the  $E_b/N_0$ , and the system load  $\alpha$ .

In the large system limit, the multi-user efficiency  $\Psi$  of the PIC-LMMSE detector is determined by the Tse-Hanley equation [Tse99] provided that the detector is perfectly matched, *i.e.*,  $\hat{h}_k = h_k$ . Let  $U$  denote the residual interference power and  $K, N \rightarrow \infty$  with  $K/N = \alpha$ . Then, the multi-user efficiency  $\kappa$  of the IC-LMMSE multi-user detector is given by

$$\kappa = \lim_{K \rightarrow \infty} \left( \sigma_v^2 + \frac{\alpha}{K} \sum_{k=1}^K \int \frac{u}{1 + u\kappa} dH_k(u) \right)^{-1} \quad (6.11)$$

with the *cumulative distribution function* (cdf) of the residual interference

$$H_k(u) = f \left( \left| b_k - \tilde{b}_k \right|^2 \left| \hat{h}_k \right|^2 < u \right). \quad (6.12)$$

If the residual interfering power of the users is not completely known, *i.e.*,  $\hat{h}_k \neq h_k$ , expression (6.11) describes the *nominal* multi-user efficiency  $\kappa$  [Cai04]. The *true* multi-user efficiency  $\Psi$  is computed as

$$\Psi = \kappa \frac{\lim_{K \rightarrow \infty} \left( 1 + \frac{\alpha}{K} \sum_{k=1}^K \int \frac{u}{(1 + u\kappa)^2} dH_k(u) \right)}{\lim_{K \rightarrow \infty} \left( 1 + \frac{\alpha}{K} \sum_{k=1}^K \int \int \frac{p}{(1 + u\kappa)^2} dH_k(u) dH_k(p|u) \right)} \quad (6.13)$$

with  $dH_k(p|u)$  being the *true* conditional cdf of the residual interference power of all users

$$H_k(p|u) = f \left( \left| b_k h_k - \tilde{b}_k \hat{h}_k \right|^2 < p \mid u = h_k \right). \quad (6.14)$$

The nominal and true distributions of the interference in (6.12) and (6.14) are computed with the Gaussian approximation for the soft symbols  $\tilde{b}_k$  and the conditional Gaussian distribution of the channel estimates with mean  $\bar{h}_k$  and variance  $\xi_k^2$ .

This work extends the analysis presented in [Bou02] by the channel estimator and by taking into account the amplitude and phase mismatch expressed by a distribution of the channel estimate in the multi-user detector. The variances of the individual symbols depend on the soft decision symbols and on the statistics of the channel estimate  $\hat{h}_k$ . The *true symbol variance* for QPSK is computed as

$$\zeta_k(m) = \sum_{b_k \in \mathbb{X}} \left| b_k h_k - \tilde{b}_k(m) \hat{h}_k \right|^2 p(b_k)$$

and becomes

$$\zeta_k(m) = |h_k|^2 + \left| \tilde{b}_k(m) \right|^2 \left| \hat{h}_k \right|^2 - 2 \left| \tilde{b}_k(m) \right|^2 \Re \left\{ h_k^* \hat{h}_k \right\}$$

when we assume that the probability for transmitting a symbol  $b_k = (a + jb)/\sqrt{2}$  with  $a, b \in \{+1, -1\}$  is given through

$$p\left(\tilde{b}_{k,\text{real}}(m) = a\right) \times p\left(\tilde{b}_{k,\text{imag}}(m) = b\right).$$

The *nominal symbol variance* is similarly derived according to (6.12) and results in

$$\zeta_k(m) = \left(1 - \left|\tilde{b}_k(m)\right|^2\right) \left|\hat{h}_k\right|^2.$$

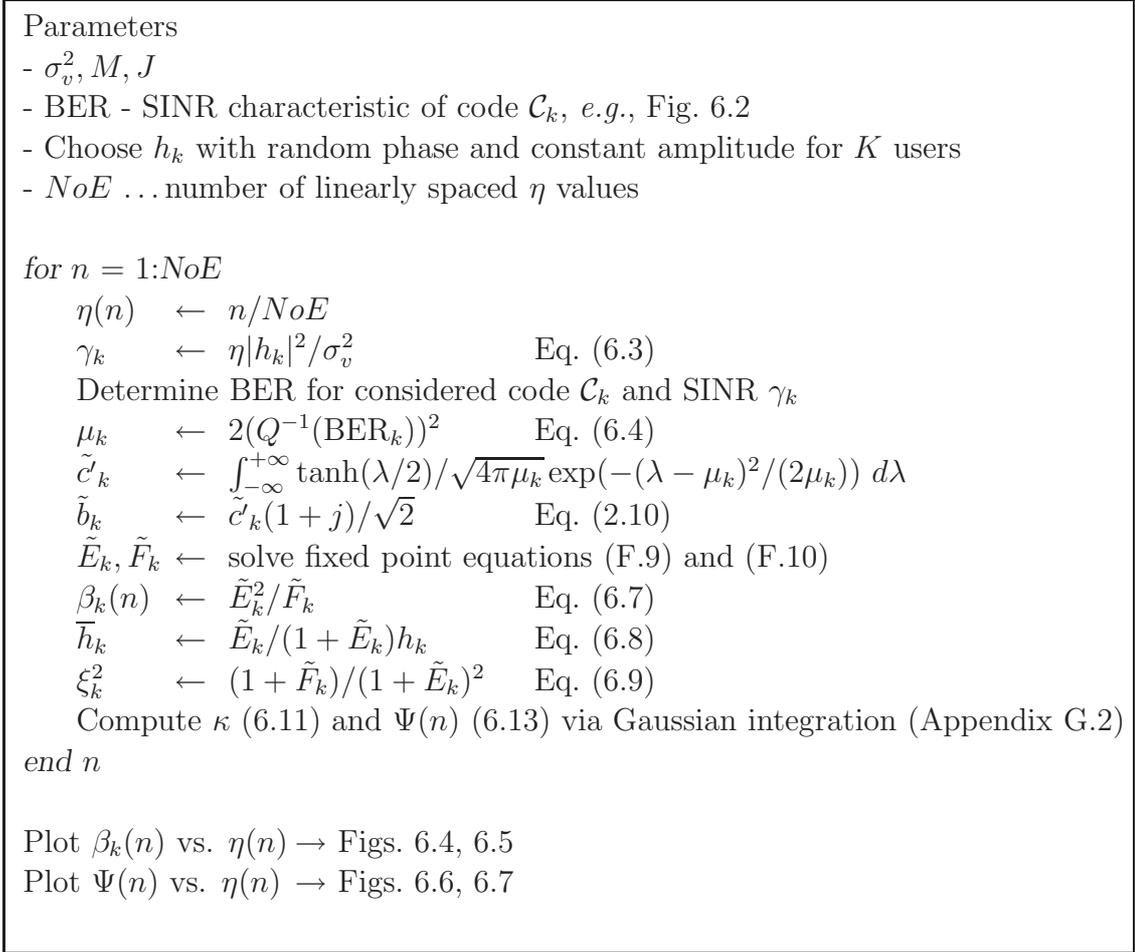
We want to point out that  $\tilde{b}_k(m)$  and  $\hat{h}_k$  are assumed to be independent and that we are aware of the fact that this is not completely true. When viewing channel estimation as a factor graph,  $\hat{h}_k$  consists of  $J$  pilot symbol contributions and  $M - J$  contributions of  $\tilde{b}_k(m)$ . The individual influence of one single symbol onto the fading tap of the same user tends to zero when the block length goes to infinity. But, for finite block length, the two variables are not fully independent.

The pdf of the code symbols  $\tilde{b}_k(m)$  is known and completely characterized by  $\mu_k$ . We can compute (6.11) and (6.13) by numerical integration as described in Appendix G.2. The individual soft feedback symbols  $\tilde{b}_k(m)$  are obtained via (6.5). The channel estimates are conditionally distributed as  $f_{\hat{h}_k}(\hat{h}_k|h_k) \sim \mathcal{CN}(\bar{h}_k, \xi_k^2)$  as shown in (6.10).

## 6.6 Results

The required steps for conducting the previously described analysis are outlined in Fig. 6.3. The large system behaviour for a continuous user distribution is well approximated by  $K = 25$  users. When applying this procedure, we produce the following results.

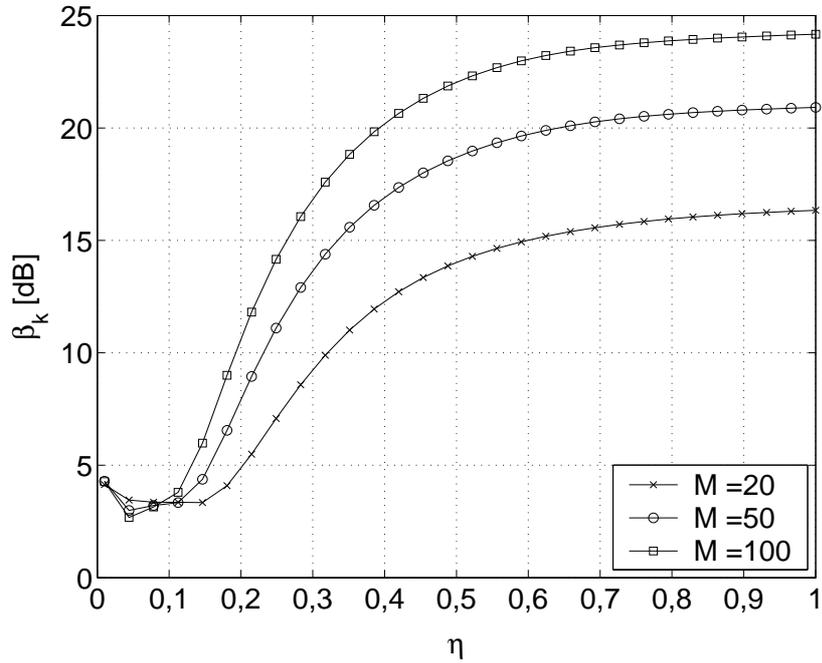
First, we illustrate the SINR  $\beta_k$  (6.7) of the channel estimate versus the multi-user efficiency  $\eta$  as an intermediate result before we assess the overall performance of the iterative receiver. We consider the channel  $|h_k|^2 = 1$  of a particular user  $k$ . In the considered setting the system load is  $\alpha = 1.5$  and  $E_b/N_0 = 5$  dB. Fig. 6.4 shows the SINR values corresponding to the block lengths  $M \in \{20, 50, 100\}$  and the number of pilot symbols  $J = 2$ . The SINR curves in Fig. 6.5 refer to the same block lengths but with  $J = 10$  pilot symbols. Due to the energy normalization (2.14) the curves of  $\beta_k$  do not originate at the same value. For  $\eta = 0$ , channel estimation relies only on the  $J$  dedicated pilot symbols. When the multi-user efficiency  $\eta$  attains larger values, the SINR  $\beta_k$  increases except for a remarkable phenomenon at low multi-user efficiencies. For small values of  $\eta$  the SINR slightly decreases. This is due to the fact that soft decision symbols are used in the channel estimator. The ALMMSE is



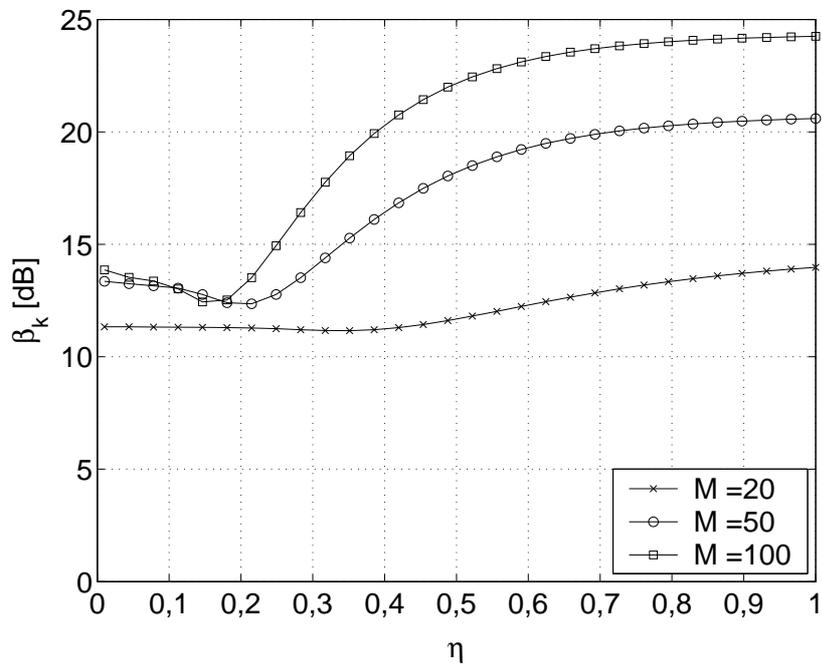
**Figure 6.3:** Outline of the iterative receiver analysis.

not aware that soft decision symbols with small amplitudes do not indicate a small amplitude of the channel taps. Hence, the estimator underestimates the amplitude of the channel. For large values of  $\eta$  the soft decision symbols become “harder” and the SINR flattens out. The maximum SINR  $\beta_k$  is achieved for  $\eta = 1$  and it implies that all feedback symbols  $\tilde{b}_k(m)$  have converged to their true values  $b_k(m)$ .

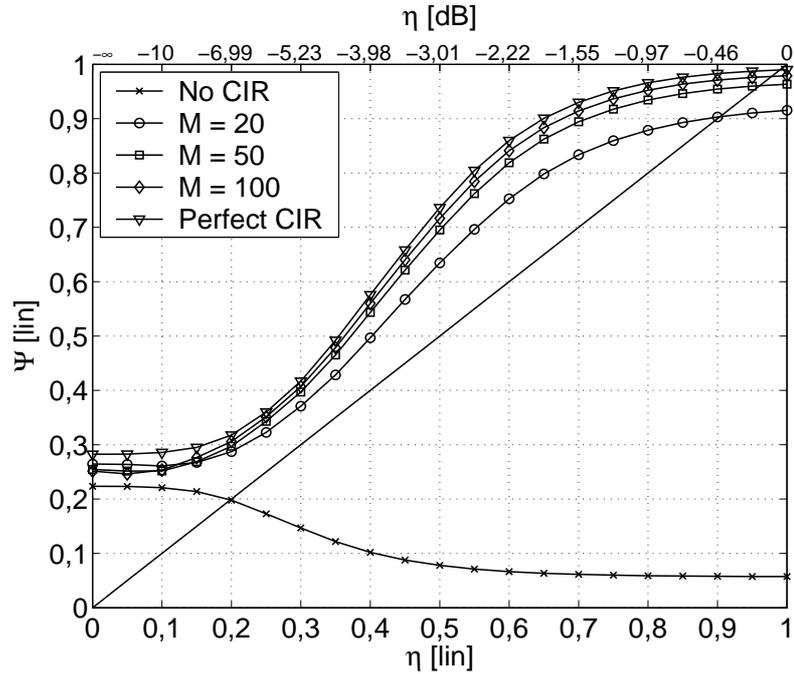
To assess the performance of the whole receiver we illustrate the evolution  $\Psi(\eta)$  of the multi-user efficiency  $\eta$ . That is, we plot the relation between the assumed multi-user efficiency  $\eta$  that is the input measure to the analysis versus the resulting multi-user efficiency  $\Psi$  after one iteration. A particular value of  $\eta$  characterizes the densities of the soft decision code symbols and the channel estimates. Evaluating (6.13) for several values of  $\eta$  results in an input-output relation  $\Psi(\eta)$  from which the behaviour of the iterative receiver can be predicted. We plot the relation  $\Psi(\eta)$  versus  $\eta$  for an  $E_b/N_0$  of 5 dB, a load  $\alpha = 1,5$ , and a rate  $R_C = 1/2$  convolutional



**Figure 6.4:** Output SINR  $\beta_k$  of the ALMMSE channel estimator for  $E_b/N_0 = 5$  dB, number of pilots  $J = 2$ , and load  $\alpha = 1,5$ .



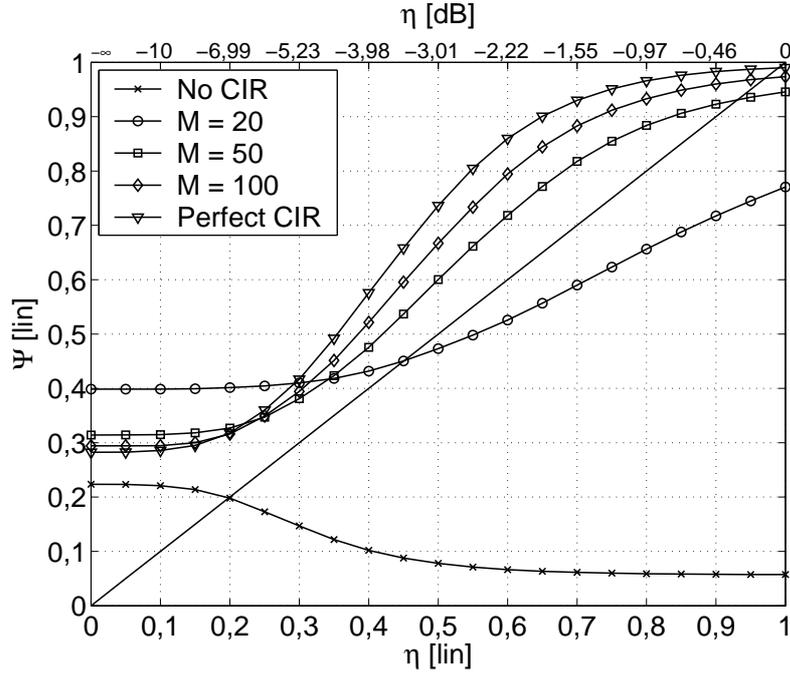
**Figure 6.5:** Output SINR  $\beta_k$  of the ALMMSE channel estimator for  $E_b/N_0 = 5$  dB, number of pilots  $J = 10$ , and load  $\alpha = 1,5$ .



**Figure 6.6:** Evolution of the multi-user efficiency  $\Psi(\eta)$  for  $E_b/N_0 = 5$  dB, number of pilots  $J = 2$ , and load  $\alpha = 1,5$ .

code with generator polynomials  $(5, 7)_8$ . The evolution curves  $\Psi(\eta)$  for the multi-user efficiency are given in Figs. 6.6 and 6.7 for block length  $M \in \{20, 50, 100\}$  and  $J = 2$  and  $J = 10$ , respectively. Like in Section 3.7 we plot the diagonal  $\Psi(\eta) = \eta$ . The crossing of a particular curve with the diagonal marks the final point in the iterative process and denotes a fixed-point. When a staircase is plotted, the number of steps to a fixed-point tells us the number of iterations required to attain where convergence comes to an end. The corresponding value  $\eta$  in dB can be determined from the upper horizontal axis and designates the distance to the single-user receiver performance.

In the following, we discuss the impact of approximate LMMSE channel estimation on the iterative receiver performance. The lowest curve describes the performance when neither the mean value  $\bar{h}_k$  nor the mean square error  $\xi_k^2$  of the distribution of the channel estimate  $\hat{h}_k$  is available in the receiver (“No CIR”) except its statistics  $\sigma_{k,h}^2 = 1$ . In this case we set  $\bar{h}_k = 1$  and  $\xi_k^2 = 1$ . The resulting fixed point indicates a gap of roughly 7 dB to the single-user bound. The most upper curve in the diagrams characterizes the performance for perfectly known channel coefficients at the receiver (“Perfect CIR”) with  $\bar{h}_k = h_k$  and  $\xi_k^2 = 0$ . At  $E_b/N_0 = 5$  dB the receiver is able to attain the single-user bound. In the plots we can further recognize three more curves representing the evolution of the multi-user efficiency  $\Psi(\eta)$  for soft decision directed



**Figure 6.7:** Evolution of the multi-user efficiency  $\Psi(\eta)$  for  $E_b/N_0 = 5$  dB,  $J = 10$ , and load  $\alpha = 1,5$ .

channel estimation described in Section 6.4. The block length  $M$  corresponds to 20, 50, and 100 symbols, respectively. Fig. 6.6 shows the case for  $J = 2$ . The fixed point for  $M = 20$  is associated with a loss of 0,46 dB to the single-user performance and those of  $M = 50$  and  $M = 100$  with 0,17 dB and 0,1 dB, respectively. The results for  $J = 10$  show an interesting detail. Due to the normalization (2.14) the noise level is increased compared to the case  $J = 2$ . For the case  $M = 20$  the SNR becomes rather poor such that the convergence stops 3,46 dB from the single-user bound. For the given  $E_b/N_0$  an increase of number of pilot symbols does not improve the performance. For blocks of length  $M = 50$  and  $M = 100$  the loss due to normalization is not that strong any more but there is still an increase of the system SNR compared to the  $J = 2$  case. Hence, for the given load and  $E_b/N_0$ , an increase in  $J$  is not beneficial.

## 6.7 Summary

In this chapter we applied density evolution to the analysis of an iterative CDMA receiver consisting of a parallel interference canceling LMMSE multi-user detector, a bank of single-user MAP decoders, and an approximate LMMSE (ALMMSE) channel estimator for a single path channel with constant amplitude and random

phase. The quantities that are delivered through the iterative process are described by parameters that describe their Gaussian distributions. This holds for the soft decision symbols and the average channel estimate. The channel estimator makes use of soft decision symbols that are derived from the output quantities of a MAP decoder. In the large system case we derive analytical expressions for the mean channel estimates  $\bar{h}_k$  and their error-variances  $\xi_k^2$  when soft decision symbols are used in the estimator. These values are passed on to the multi-user detector and they are jointly used with the soft decision symbols that are derived from the *extrinsic* information. The performance is evaluated in terms of the multi-user efficiency  $\Psi$  and we show how the performance depends on the symbol block length  $M$ , the number of dedicated pilot symbols  $J$ , the system  $E_b/N_0$ , and the load  $\alpha = K/N$ .

## 7 Conclusions

This thesis investigates iterative CDMA receivers with coding for realistic channel environments in uplink communications. We derive iterative receivers with channel estimation as suboptimum implementation of the optimum receiver that exhibits prohibitive computational complexity. An iterative receiver consists of three decoupled modules - the multi-user detector, a bank of single-user decoders, and a channel estimator. In this way the complexity is reduced to a manageable measure compared to the defying effort required by optimum joint multi-user decoding and channel estimation. The emphasis of this work is placed on inter-symbol interference processing in the detectors and multi-path channel estimation that is integrated in the iterative processing loop. We investigate various types of detectors with different realizations of interference cancellation and the subsequent linear filter and assess their support for system capacity, convergence properties, and bit error rate. Then, we develop channel estimation algorithms that use different degrees of statistical *a-priori* knowledge. An interesting aspect is also the choice of pilot sequences for initial channel estimation in the first iteration when no soft decision data is available. The estimation schemes are compared against each other with several multi-path channels. Further, we design an iterative receiver for the UMTS-FDD standard using the results on multi-user detection and channel estimation obtained from the general setup considered previously. The iterative receiver requires some modifications due to the different modulation format and the type of convolutional codes used. We demonstrate that the general concepts can be applied successfully in UMTS under the assumption of frame-synchronous transmission. Finally, we present an approach to study the iterative receiver in a semi-analytical way for the particular case of a single path channel. The most important insight gained from this work shall be summarized in the following paragraph.

**Feedback Information:** For interference cancellation in the multi-user detector *extrinsic* information leads to highest achievable loads. For a particular range of the system load *a-posteriori* feedback converges faster but does not achieve better results than when *extrinsic* information is used with more iterations. When soft decision symbols are used in channel estimation symbols derived from *a-posteriori* information lead to better channel estimates than their *extrinsic* counterparts.

**Inter-Symbol Interference:** For realistic channels, where energy can be dispersed over an observation interval of  $1\ \mu\text{s}$ , inter-symbol-interference processing is required for spreading factors  $N < 32$  to maximize the system capacity.

**Interference Cancellation:** Successive interference cancellation requires two to three iterations less to achieve the best possible performance compared to parallel interference cancellation. However, the maximum load was observed to be only slightly higher for successive IC than for the parallel IC.

**Soft-Decision Symbols Supported Channel Estimators:** Iterative structure promote the usage of channel estimators that consider soft decision symbols. We show that with the most complex scheme, that considers the variances of soft decision symbols, the lowest mean square error can be obtained. However, the relatively small gains in the mean square error do not accelerate the convergence speed of the receiver. It turns out that a simple approximated least-squares scheme shows a loss of up to 1,75 dB over the LMMSE estimator at a BER of  $10^{-3}$  in the Pedestrian-B multipath environment.

**Channel Sounding Sequences:** Optimum sequences for initial channel estimation minimize the mean square error of the channel estimates. However, the final bit error rate achieved with a setup based on perfect root of unity sequences can be replaced by the less complex random symbol sequence.

**System Capacity:** An iterative receiver with integrated ALS channel estimation and PIC-LMMSE detection allows system loads of 1,5 in various multi-path channels. This implies a significant increase in system capacity when compared to one-short Rake receivers.

**Multiple Antenna Reception:** An increase in the number of receive antennas allows roughly the same increase in the number of accommodated users.

**Convolutional Codes:** Convolutional codes with larger constraint lengths lead to a premature end in the iterative convergence process.

**UMTS Receiver:** Under the assumption of symbol synchronous transmission the UMTS compliant iterative receiver shows large performance gains compared to one-shot Rake receivers.

**Analysis:** We have developed an analysis tool based on density evolution to predict the convergence behaviour of an iterative receiver with integrated channel estimation for the case of a constant amplitude channel with random phase. The best performance is always achieved with larger block lengths  $M$ . The number of pilot symbols  $J$  can be kept at a minimum to ensure identifiability.

# A Acronyms

<b>ALS</b>	Approximate Least Squares
<b>ALMMSE</b>	Approximate Linear Minimum Mean Square
<b>AWGN</b>	Additive White Gaussian Noise
<b>BER</b>	Bit Error Rate
<b>BP</b>	Belief Propagation
<b>BPSK</b>	Binary Phase Shift Keying
<b>CC</b>	Convolutional Code
<b>cdf</b>	Cumulative Distribution Function
<b>CG</b>	Coding Gain
<b>CIR</b>	Channel Impulse Response
<b>COST</b>	COoperation europeenne dans le domaine de la recherche Scientifique et Technique
<b>CSI</b>	Channel State Information
<b>CDMA</b>	Code Division Multiple-Access
<b>DE</b>	Density Evolution
<b>DS</b>	Direct Sequence
<b>EDGE</b>	Enhanced Data Rate for Global Evolution
<b>FBI</b>	Feedback Information
<b>FEC</b>	Forward Error Correction
<b>GSM</b>	Global System of Mobile Communications
<b>HSDPA</b>	High Speed Downlink Packet Access

<b>IC</b>	Interference Cancellation
<b>IEEE</b>	Institute of Electrical and Electronics Engineers
<b>i.i.d.</b>	Independent and Identically Distributed
<b>ISI</b>	Inter-Symbol Interference
<b>IS-95</b>	Interim Standard 95
<b>LAN</b>	Local Area Network
<b>LMMSE</b>	Linear Minimum Mean Square Error
<b>LOS</b>	Line of Sight
<b>MAI</b>	Multiple-Access Interference
<b>MAN</b>	Metropolitan Area Network
<b>MAP</b>	Maximum A-Posteriori
<b>MC-CDMA</b>	Multi-Carrier CDMA
<b>MIMO</b>	Multiple-Input Multiple-Output
<b>ML</b>	Maximum Likelihood
<b>MMSE</b>	Minimum Mean Square Error
<b>MRC</b>	Maximum Ratio Combining
<b>MSE</b>	Mean Square Error
<b>MUE</b>	Multi-User Efficiency
<b>NMSE</b>	Normalized Mean Square Error
<b>OFDM</b>	Orthogonal Frequency Division Multiplexing
<b>pdf</b>	Probability Density Function
<b>pmf</b>	Probability Mass Function
<b>PIC</b>	Parallel Interference Canceler
<b>QPSK</b>	Quarternary Phase Shift Keying

---

<b>PRUS</b>	Perfect Root of Unity Sequence
<b>RAND</b>	Random Symbol Pilot Sequence
<b>rms</b>	Root Mean Square
<b>SIC</b>	Successive Interference Canceler
<b>SINR</b>	Signal to Interference and Noise Ratio
<b>SNR</b>	Signal to Noise Ratio
<b>SOVA</b>	Soft Output Viterbi Algorithm
<b>SUMF</b>	Single-User Matched Filter
<b>TC</b>	Turbo Code
<b>TDMA</b>	Time Division Multiple-Access
<b>TFCI</b>	Transmit Format Combination Indicator
<b>TPC</b>	Transmit Power Control
<b>TTI</b>	Transmission Time Interval
<b>UMTS</b>	Universal Mobile Telecommunication System
<b>UTRA</b>	UMTS Terrestrial Radio Access
<b>WLAN</b>	Wireless LAN
<b>WSSUS</b>	Wide Sense Stationary and Uncorrelated Scattering
<b>ZF</b>	Zero Forcer
<b>3GPP</b>	Third Generation Partnership Program



## B Notation

$a, b, c, \dots$	Scalars
$\mathbf{a}, \mathbf{b}, \mathbf{c}, \dots$	Vectors
$\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$	Matrices
$[\mathbf{A}]_{k,l}$	Element in the $k$ -th row and $l$ -th column of matrix $\mathbf{A}$
$\mathbf{0}_N$	All-0 vector of dimension $N \times 1$
$\mathbf{1}_N$	All-1 vector of dimension $N \times 1$
$\mathbf{I}_N$	Identity matrix of dimension $N \times N$
$\mathbf{e}_k$	The standard basis unit vector for the $k$ -th dimension
$(\cdot)^*$	Conjugate operator
$(\cdot)^T$	Transpose operator
$(\cdot)^H$	Hermitian operator
$(\cdot)^{-1}$	Inverse operator
$(\cdot)^{(i)}$	Indication for $i$ -th iteration
$j$	$\sqrt{-1}$
$\exp(\cdot)$	Exponential function
$\log(\cdot)$	Natural logarithmic function
$\log_2(\cdot)$	Logarithm with base 2
$\underset{\mathbf{x}}{\operatorname{argmin}} f(\mathbf{x})$	Argument $\mathbf{x}$ for which the scalar function $f(\mathbf{x})$ is minimized.

---

$\operatorname{argmax}_{\mathbf{x}} f(\mathbf{x})$	Argument $\mathbf{x}$ for which the scalar function $f(\mathbf{x})$ is maximized.
$\operatorname{quant}_{\mathbb{X}}\{x\}$	Assignment of that value in $\mathbb{X}$ that has the minimum Euclidian distance to the observation $x$ .
$ x $	Absolute value of $x$
$\ \mathbf{a}\ $	$l_2$ norm of $\mathbf{a}$
$\operatorname{diag}(\mathbf{a})$	Place elements of the vector $\mathbf{a}$ on the main diagonal of a matrix that has zeros elsewhere (becomes a matrix).
$\operatorname{diag}(\mathbf{A})$	Diagonal of $\mathbf{A}$ (becomes a vector)
$\operatorname{tr}(\mathbf{A})$	Trace of $\mathbf{A}$
$\mathcal{F}(\cdot)$	Fourier transform
$\Re\{\cdot\}$	Real part
$\Im\{\cdot\}$	Imaginary part
$f(\cdot)$	Probability density function (pdf)
$p(\cdot)$	Probability mass function (pmf)
$\mathbb{E}\{\cdot\}$	Expectation operator
$\triangleq$	Defined as
$\propto$	Proportional
$\sim$	Distributed as
$*$	Convolution
$\otimes$	Kronecker product
$\mathbb{C}^{m \times n}$	Set of complex matrices with dimension $m \times n$
$\mathbb{F}^{m \times n}$	Set of matrices with dimension $m \times n$ and elements $\{0, 1\}$

---

$\mathbb{R}^{m \times n}$	Set of real matrices with dimension $m \times n$
$\mathbb{X}^{m \times n}$	Set of matrices with dimension $m \times n$ and elements $\{\pm 1 \pm j\}/\sqrt{2}$ .
$\mathcal{N}(\mu, \sigma^2)$	Real Gaussian distribution with mean $\mu$ and variance $\sigma^2$ .
$\mathcal{CN}(\mu_R + j\mu_I, \sigma^2)$	Complex Gaussian distribution for $Z = X + jY$ with $X$ and $Y$ independent and distributed as $X \sim \mathcal{N}(\mu_R, \sigma^2/2)$ and $Y \sim \mathcal{N}(\mu_I, \sigma^2/2)$ , respectively.
$\delta(k - l)$	Kronecker Delta function: $\delta(k - l) = \begin{cases} 0 & : k \neq l \\ 1 & : k = l \end{cases}$
$\phi(\cdot)$	Euler Totian function
$\nabla_{\mathbf{x}} f(\mathbf{x})$	Gradient of scalar function $f(\mathbf{x})$ with respect to $\mathbf{x}$ .



## C Variables

$B$	Bandwidth
$G$	Number of bits per transmit symbol
$J$	Number of pilot symbols
$K$	Number of users
$L$	Length of propagation channel
$M$	Transmission block length
$N$	Spreading sequence length
$N_R$	Number of receive antennas
$P$	Power
$R_C$	Code rate
$T_C$	Chip duration
$T_S$	Symbol duration
$i$	Iteration index
$k$	User index
$l$	Path index
$m$	Symbol index
$n$	Chip index
$r$	Antenna index
$b(m)$	Transmit symbol

---

$c(m')$	Coded symbol
$d(m'')$	Information bit
$h[n]$	Channel impulse response
$\check{s}[n]$	Effective spreading sequence with <i>perfect</i> CSI
$\tilde{s}[n]$	Virtual spreading sequence with <i>estimated</i> CSI
$s[n]$	Spreading sequence
$u[n]$	Pilot sequence
$v(m)$	Noise contribution where $v(m) \sim \mathcal{CN}(0, \sigma_v^2)$
$x(m')$	Demapped detector output symbols
$z(m')$	Detector output symbols
$\alpha$	System load
$\eta$	True multi-user efficiency
$\gamma$	Signal to noise ratio (SNR) or Signal to noise and interference ratio (SINR)
$\sigma_h^2$	Channel tap variance
$\sigma_x^2$	Observation noise variance
$\sigma_v^2$	Additive noise variance
$\kappa$	Nominal multi-user efficiency
$\iota$	Code memory length
$\tau_D$	Channel delay spread
$\Psi$	True multi-user efficiency as output of analysis
$\Sigma$	Covariance matrix of the channel estimation error

## D The MAP Decoding Algorithm

Convolutional codes can be represented by a trellis that shows all valid transitions from a particular state  $s(m'')$  at time  $m''$  to the state  $s(m'' + 1)$ . An example of a trellis for a four-state convolutional code with rate  $R_C = 1/2$  and memory length  $\iota$  is depicted in Fig. D.1. There are  $2^\iota$  possible states denoted by  $\{s_0, s_1, s_2, s_3\}$ . For time  $m'' = 0$  we assume that the state of the code is  $S(m'' = 0) = s_0$  in the trellis. From this initial state the trellis is built up according to the  $M'' = 2(M - J)R_C$  information bits and the generator polynomials of the convolutional code. At the end the trellis is forced to its initial state by appending  $\iota$  terminating zero input symbols. From state each state  $s_i, i \in \{0, \dots, 3\}$  at time  $m''$  there are two outgoing edges:  $e_0$  is associated with the information bit  $d = 0$  and the output code word  $\mathbf{c} = [c_0 \ c_1]$ . Similarly, this holds for the information bit  $d = 1$  that is associated with edge  $e_1$ . The conditional probability for a transition from state  $S_i$  to state  $S_j$ , assuming that the states are connected, calculates as

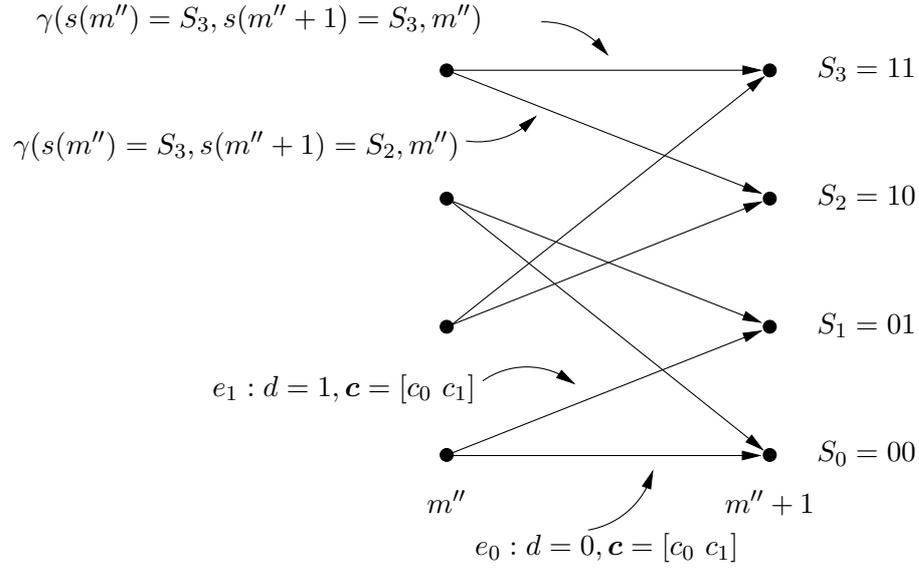
$$\gamma(S_i, S_j, m'') = \frac{1}{\sqrt{2\pi\sigma_x^2}} \prod_{u=1}^{1/R_C} \exp\left(-\frac{(x(u + m''/R_C) - \mu_x c'_{i,j}(u))^2}{2\sigma_x^2}\right)$$

where  $\mu_x$  and  $\sigma_x^2$  correspond to the parameters of the AWGN model (2.17). The value  $c'_{i,j}(u)$  is the BPSK mapped  $u$ -th code bit that is associated with the transition from state  $S_i$  to state  $S_j$ . The probability for being in a particular state  $S_i$  is computed by means of the forward propagation probabilities  $\alpha$  and the backward propagation probabilities  $\beta$ . They are expressed as

$$\begin{aligned} \alpha(S_i, m'' + 1) &= \sum_{j=0}^{2^\iota - 1} \alpha(S_j, m'') \gamma(S_j, S_i, m''), \\ \beta(S_i, m'') &= \sum_{j=0}^{2^\iota - 1} \beta(S_j, m'' + 1) \gamma(S_i, S_j, m''). \end{aligned}$$

The initial values of  $\alpha(S_i, 0)$  and  $\beta(S_i, M'' + \iota)$  are

$$\alpha(S_i, 0) = \begin{cases} 1 & i = 0, \\ 0 & i \neq 0, \end{cases} \quad \beta(S_i, M'' + \iota) = \begin{cases} 1 & i = 0, \\ 0 & i \neq 0. \end{cases}$$



**Figure D.1:** Example of a trellis for a four-state convolutional code with rate  $R_C = 1/2$ .

The forward and backward state probabilities  $\alpha(\cdot)$  and  $\beta(\cdot)$  are computed as products of probabilities. These quantities take on very small values after few stages and can cause errors due to overflows that are typically caused by the finite number representation capabilities of the hardware. The problem is removed when the following normalisations are applied at every stage

$$\sum_{i=0}^{2^e-1} \alpha(S_i, m'') = 1 \quad \text{and} \quad \sum_{i=0}^{2^e-1} \beta(S_i, m'') = 1 .$$

With  $a \in \mathbb{F}$  the APP on the information bits are computed from the state probabilities

$$\text{APP} \{d(m'') = a | \mathbf{x}\} = \frac{\sum_{d=a} \alpha(S_i, m'') \beta(S_j, m'' + 1)}{\sum_{d=0, d=1} \alpha(S_i, m'') \beta(S_j, m'' + 1)} , \quad (\text{D.1})$$

where the sum over  $d = a$  concerns all transitions from state  $S_i$  to state  $S_j$  where the information symbol  $d$  has the value  $a$ . The APP on the code bits  $c_l$ ,  $l \in \{0, 1\}$ , are computed from the transition probabilities

$$\text{APP} \{c_l(m'') = a | \mathbf{x}\} = \frac{\sum_{c_l=a} \alpha(S_i, m'') \gamma(S_i, S_j, m'') \beta(S_j, m'' + 1)}{\sum_{c_l=0, c_l=1} \alpha(S_i, m'') \gamma(S_i, S_j, m'') \beta(S_j, m'' + 1)} . \quad (\text{D.2})$$

Similar as in (D.1) the sum over  $c_l = a$  includes all transitions for which the code bit  $c_l$  is  $a$ . For the implementation the following details turn out to be important. The relation between the APP and EXT values is given by

$$\text{EXT} \{c_l(m'') = a|\mathbf{x}\} \propto \text{APP} \{c_l(m'') = a|\mathbf{x}\} \times \exp\left(\frac{(x(l + m'/R_C) - (1 - 2a)\mu_x c'(l))^2}{2\sigma_x^2}\right).$$

This relation allows to compute the EXT values from the APP and the observations. For APP values that are very close to 0 or 1 the corresponding conversion into likelihoods casts a numerical problem. Expressions like “ $\infty - \infty$ ” occur and provoke instabilities. The MAP algorithm offers an intrinsic way to compute the EXT values by reshaping expression (D.2) to

$$\text{EXT} \{c_l(m'') = a|\mathbf{x}\} = \frac{\sum_{c_l=a} \alpha(S_i, m'') \gamma_{\sim l}(S_i, S_j, m'') \beta(S_j, m'' + 1)}{\sum_{c_l=0, c_l=1} \alpha(S_i, m'') \gamma_{\sim l}(S_i, S_j, m'') \beta(S_j, m'' + 1)}$$

with

$$\gamma_{\sim l}(S_i, S_j, m'') = \frac{1}{\sqrt{2\pi\sigma_x^2}} \prod_{u=1, u \neq l}^{1/R_C} \exp\left(-\frac{(x(u + m''/R_C) - \mu_x c'_{i,j}(u))^2}{2\sigma_x^2}\right).$$



## E Proof of Equation (3.20)

The variables used in the matrix inversion lemma (G.2) are identified with those in (3.19) as

$$\begin{aligned}\mathbf{A} &\triangleq \sum_{q=-1}^1 \check{\mathbf{S}}_q \mathbf{V}(m-q) \check{\mathbf{S}}_q^H + \sigma_v^2 \mathbf{I}_{N+L-1}, \\ \mathbf{b} &\triangleq \check{\mathbf{s}}_k \tilde{b}_k(m).\end{aligned}$$

Hence, (3.19) can be recasted to

$$\begin{aligned}& \frac{\check{\mathbf{s}}_k^H \left( \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{b} (1 + \mathbf{b}^H \mathbf{A}^{-1} \mathbf{b})^{-1} \mathbf{b}^H \mathbf{A}^{-1} \right)}{\check{\mathbf{s}}_k^H \left( \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{b} (1 + \mathbf{b}^H \mathbf{A}^{-1} \mathbf{b})^{-1} \mathbf{b}^H \mathbf{A}^{-1} \right) \check{\mathbf{s}}_k} \\ = & \frac{\check{\mathbf{s}}_k^H \left( \mathbf{A}^{-1} + \mathbf{A}^{-1} \mathbf{b}^H \mathbf{A}^{-1} \mathbf{b} - \mathbf{A}^{-1} \mathbf{b} \mathbf{b}^H \mathbf{A}^{-1} \right)}{\check{\mathbf{s}}_k^H \left( \mathbf{A}^{-1} + \mathbf{A}^{-1} \mathbf{b}^H \mathbf{A}^{-1} \mathbf{b} - \mathbf{A}^{-1} \mathbf{b} \mathbf{b}^H \mathbf{A}^{-1} \right) \check{\mathbf{s}}_k} \\ = & \frac{\check{\mathbf{s}}_k^H \mathbf{A}^{-1} + \check{\mathbf{s}}_k^H \mathbf{A}^{-1} \mathbf{b}^H \mathbf{A}^{-1} \mathbf{b} - \check{\mathbf{s}}_k^H \mathbf{A}^{-1} \mathbf{b} \mathbf{b}^H \mathbf{A}^{-1}}{\left( \check{\mathbf{s}}_k^H \mathbf{A}^{-1} + \check{\mathbf{s}}_k^H \mathbf{A}^{-1} \mathbf{b}^H \mathbf{A}^{-1} \mathbf{b} - \check{\mathbf{s}}_k^H \mathbf{A}^{-1} \mathbf{b} \mathbf{b}^H \mathbf{A}^{-1} \right) \check{\mathbf{s}}_k} \\ = & \frac{\check{\mathbf{s}}_k^H \mathbf{A}^{-1} + \left| \tilde{b}_k(m) \right|^2 \check{\mathbf{s}}_k^H \mathbf{A}^{-1} \check{\mathbf{s}}_k \check{\mathbf{s}}_k^H \mathbf{A}^{-1} - \left| \tilde{b}_k(m) \right|^2 \check{\mathbf{s}}_k^H \mathbf{A}^{-1} \check{\mathbf{s}}_k \check{\mathbf{s}}_k^H \mathbf{A}^{-1}}{\left( \check{\mathbf{s}}_k^H \mathbf{A}^{-1} + \left| \tilde{b}_k(m) \right|^2 \check{\mathbf{s}}_k^H \mathbf{A}^{-1} \underbrace{\check{\mathbf{s}}_k \check{\mathbf{s}}_k^H}_{\text{scalar}} - \left| \tilde{b}_k(m) \right|^2 \underbrace{\check{\mathbf{s}}_k \check{\mathbf{s}}_k^H}_{\text{scalar}} \right) \check{\mathbf{s}}_k} \\ = & \frac{\check{\mathbf{s}}_k^H \mathbf{A}^{-1}}{\check{\mathbf{s}}_k^H \mathbf{A}^{-1} \check{\mathbf{s}}_k}\end{aligned}$$

and relation (3.20) is established.



# F Remark on Conjecture 1

In [Nis01, Tan02a] the replica method was used to analyse the output SINR of a maximum *a-posteriori* multi-user detector. The replica method is a tool from statistical physics and it describes the macroscopic behaviour of a compound of microscopic particles. In case of multi-user detection the microscopic particles correspond to the individual chips of all users and the SINR would result from two macroscopic measures. The replica method is particularly useful for mismatched situations, *i.e.*, when the noise variance  $\sigma_v^2$  or the power distribution of the users is not perfectly known. It is able to accurately describe the SINR in such a mismatched situation. A matched LMMSE multi-user detector described through the replica method is described by the same equations as those obtained through random matrix theory [Tan02a].

We apply the replica method to mismatched channel estimation in the case when all users have constant power but a random phase. The problem of linear channel estimation is structurally the same as that of linear multi-user detection. A linear multi-user detector is a two stage device that first applies a filter on the observation vector and then uses the output to decide on the transmitted symbol. In the following we review the estimation problem for the situation of a single tap channel and formulate a conjecture for the output SINR based on results obtained through the replica method. Its validity is supported through experimental results.

## Computing the SINR of the Mismatched Channel Estimator

The channel estimates obtained through approximated LMMSE estimation are

$$\tilde{\mathbf{h}} = \left( \tilde{\mathbf{A}}^H \tilde{\mathbf{A}} + \sigma_v^2 \mathbf{I}_K \right)^{-1} \tilde{\mathbf{A}}^H \mathbf{y}.$$

We want to find the associated output distribution of  $\hat{h}_k \sim N(\bar{h}_k, \xi_k^2)$ . The mean and the error variance are given through (6.8) and (6.9). They are parameterised by the macroscopic measures  $\tilde{E}_k$  and  $\tilde{F}_k$  whose quotient (6.7) is the SINR  $\beta_k$  [Nis01] of the channel estimate, assuming that the power of all taps is one. The physical meaning of  $\tilde{E}_k$  and  $\tilde{F}_k$  is that they describe a set of fixed-points that minimise the free energy of the multi-user detector (channel estimator) [Tan02a]. For the computation of the parameters we follow [Mül03] which is a generalisation of [Tan02a] allowing unequal chip variances. The variable  $w_k[m]$  denotes the variance of the chip in the  $k$ -th

column and  $m$ -th row of  $\tilde{\mathcal{A}} \in \mathbb{C}^{MN \times K}$ . The parameters  $\tilde{E}_k$  and  $\tilde{F}_k$  write

$$\begin{aligned}\tilde{E}_k &\triangleq \frac{1}{MN} \sum_{m=1}^{MN} E[m] w_k[m] = \frac{1}{M} \sum_{m=1}^M E[mN] w_k[mN], \\ \tilde{F}_k &\triangleq \frac{1}{MN} \sum_{m=1}^{MN} F[m] w_k[m] = \frac{1}{M} \sum_{m=1}^M F[mN] w_k[mN]\end{aligned}$$

where the last equation follows from the fact that the chip variances for user  $k$  associated with symbol  $m$  are all the same. To simplify notation we use  $E[mN] \rightarrow E(m)$  and  $F[mN] \rightarrow F(m)$  and reflect the dependence on the symbol level by the bracket  $(\cdot)$ . The variables  $E(m)$  and  $F(m)$  are computed as

$$\begin{aligned}E(m) &= \frac{1}{\sigma_v^2 + \alpha(p(m) - q(m))}, \\ F(m) &= \frac{\sigma_v^2 + \alpha(p_0(m) - 2r(m) + q(m))}{(\sigma_v^2 + \alpha(p(m) - q(m))^2)}.\end{aligned}$$

The set of parameters  $\{p(m), q(m), p_0(m), r(m)\}$  are directly related to the variances of the related symbol variances. They are in general different for different users and chips due to the soft decision data. To bring into play the soft decision data we make a modification in the set of equations (186), (187), (202)–(205) presented in [Mül03]. The variances are replaced by the associated variance of the soft decision value in the following way:

**Assumption 1**

$$\tilde{E}_k = \frac{1}{M} \sum_{m=1}^M E(m) \Re \left\{ b_n(m) \tilde{b}_n^*(m) \right\} \quad (\text{F.1})$$

$$\tilde{F}_k = \frac{1}{M} \sum_{m=1}^M F(m) \Re \left\{ b_k(m) \tilde{b}_k(m) \right\} \quad (\text{F.2})$$

$$r(m) = \frac{1}{K} \sum_{n=1}^K \frac{\tilde{E}_n}{1 + \tilde{E}_n} \Re \left\{ b_n(m) \tilde{b}_n^*(m) \right\} \quad (\text{F.3})$$

$$q(m) = \frac{1}{K} \sum_{n=1}^K \frac{\tilde{E}_n^2 + \tilde{F}_n}{(1 + \tilde{E}_n)^2} \left| \tilde{b}_n(m) \right|^2 \quad (\text{F.4})$$

$$p(m) = \frac{1}{K} \sum_{n=1}^K \frac{\tilde{E}_n^2 + \tilde{E}_n + \tilde{F}_n + 1}{(1 + \tilde{E}_n)^2} \left| \tilde{b}_n(m) \right|^2 \quad (\text{F.5})$$

$$p_0(m) = \frac{1}{K} \sum_{n=1}^K \left| b_n(m) \right|^2 \quad (\text{F.6})$$

With the assumptions (F.3–F.6) the parameters  $\tilde{E}_k$  and  $\tilde{F}_k$  are determined as

$$\tilde{E}_k = \sum_{m=1}^M \frac{\Re \left\{ \tilde{b}_k(m) b_k^*(m) \right\}}{\sigma_v^2 + \frac{\alpha}{K} \sum_{n=1}^K \frac{|\tilde{b}_n(m)|^2}{1 + \tilde{E}_n}} \quad (\text{F.7})$$

$$\begin{aligned} \tilde{F}_k &= \sum_{m=1}^M \left| \tilde{b}_k(i) \right|^2 \times \\ & \frac{\sigma_v^2 + \frac{\alpha}{K} \sum_{n=1}^K \frac{(1 + \tilde{E}_k)^2 - 2\Re \left\{ \tilde{b}_n(m) b_n^*(m) \right\} (\tilde{E}_k + \tilde{E}_k^2) + |\tilde{b}_n(m)|^2 (\tilde{E}_n^2 + \tilde{F}_n)}{(1 + \tilde{E}_n^2)}}{\left( \sigma_v^2 + \frac{\alpha}{K} \sum_{n=1}^K \frac{|\tilde{b}_n(m)|^2}{1 + \tilde{E}_n} \right)^2}. \end{aligned} \quad (\text{F.8})$$

For large block length  $(M - J) \gg 1$ , the summation over the soft decision symbols is equivalent to averaging over their distribution  $L_k(x, y)$ . Hence, (F.7) and (F.8) can be rewritten as

$$\tilde{E}_k = \frac{J}{\sigma_v^2 + \frac{\alpha}{K} \sum_{n=1}^K \frac{1}{1 + \tilde{E}_n}} + \int_{\mathbb{R}^2} \frac{(M - J)}{\sigma_v^2 + \frac{\alpha}{K} \sum_{n=1}^K \frac{\text{thc}(x, y)}{1 + \tilde{E}_n}} \text{ths}(x, y) dL_k(x, y) \quad (\text{F.9})$$

$$\begin{aligned} \tilde{F}_k &= J \frac{\sigma_v^2 + \frac{\alpha}{K} \sum_{n=1}^K \frac{1 + \tilde{F}_n}{(1 + \tilde{E}_n)^2}}{\left( \sigma_v^2 + \frac{\alpha}{K} \sum_{n=1}^K \frac{1}{1 + \tilde{E}_n} \right)^2} + (M - J) \times \\ & \int_{\mathbb{R}^2} \frac{\sigma_v^2 + \frac{\alpha}{K} \sum_{n=1}^K \left( 1 - \frac{2\tilde{E}_n \text{ths}(x, y)}{1 + \tilde{E}_n} + \frac{(\tilde{E}_n^2 + \tilde{F}_n) \text{thc}(x, y)}{(1 + \tilde{E}_n)^2} \right)}{\left( \sigma_v^2 + \frac{\alpha}{K} \sum_{n=1}^K \frac{\text{thc}(x, y)}{1 + \tilde{E}_n} \right)^2} \text{thc}(x, y) dL_k(x, y) \end{aligned} \quad (\text{F.10})$$

with the definitions

$$\text{ths}(x, y) \triangleq \frac{\tanh\left(\frac{x}{2}\right) + \tanh\left(\frac{y}{2}\right)}{2}$$

and

$$\text{thc}(x, y) \triangleq \frac{\tanh^2\left(\frac{x}{2}\right) + \tanh^2\left(\frac{y}{2}\right)}{2}.$$

The LLR values are approximated through a two-dimensional Gaussian distribution [Chu01]

$$\frac{\partial^2}{\partial x \partial y} L_k(x, y) = \frac{1}{4\pi\mu_k} \exp\left(-\frac{(x - \mu_k)^2 + (y - \mu_k)^2}{4\mu_k}\right).$$

The SINRs  $\beta_k$  can be calculated analytically given the mean  $\mu_k$  of the LLR values. The fixed-point equations for  $\tilde{E}_k$  (F.9) and  $\tilde{F}_k$  (F.10) are solved via Monte-Carlo integration.

The above choice of  $\{p(m), q(m), p_0(m), r(m)\}$  is an assumption and we do not know a rigorous proof that this choice leads to a valid energy function that describes the channel estimation problem. This is an open issue. However, we will show comparisons with experimentally evaluated SINRs that strongly support this choice.

### Validation

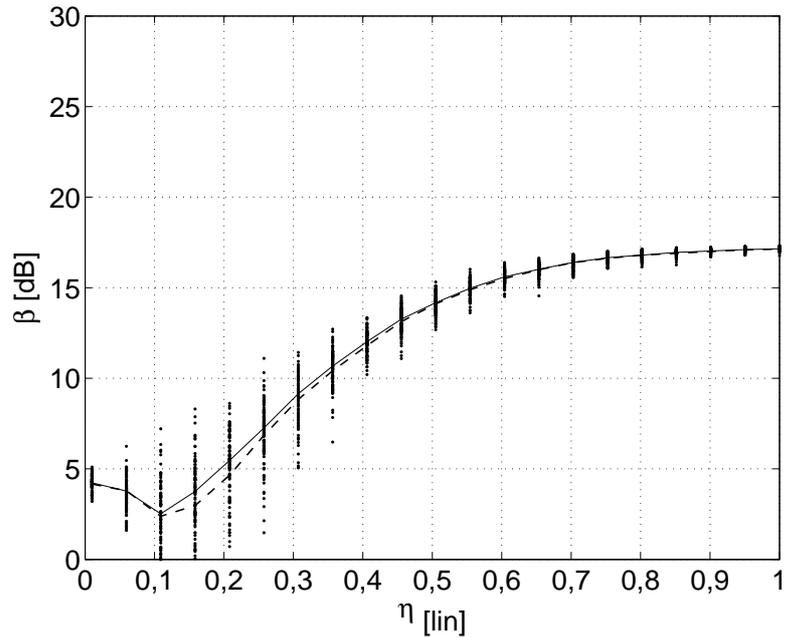
To verify the validity of our conjecture we determine the SINRs of an LMMSE estimator for a finite dimensional system that is based on random signature realizations. We consider the pairs  $\{M = 20, J = 2\}$  and  $\{M = 100, J = 10\}$  for the block length and the number of pilot symbols. The system load is  $\alpha = K/N = 48/32 = 1.5$  and the elements of the signature sequences are from the discrete set  $\mathbb{X}/\sqrt{2N}$ . The pilot symbols are chosen randomly from the set  $\mathbb{X}$ . For the data we generate  $2(M - J)$  random log-likelihood ratio values according to the statistics (6.2) given through  $\eta$  and map them to the complex soft decision symbols via (6.5). These are substituted in  $\tilde{\mathbf{B}}_D$  and we obtain the ALMMSE channel estimator (4.6) of user  $k$  as

$$\mathbf{f}_k^H = \tilde{\mathbf{a}}_k^H \left( \tilde{\mathbf{A}} \tilde{\mathbf{A}}^H + \sigma_v^2 \mathbf{I}_{MN} \right)^{-1}$$

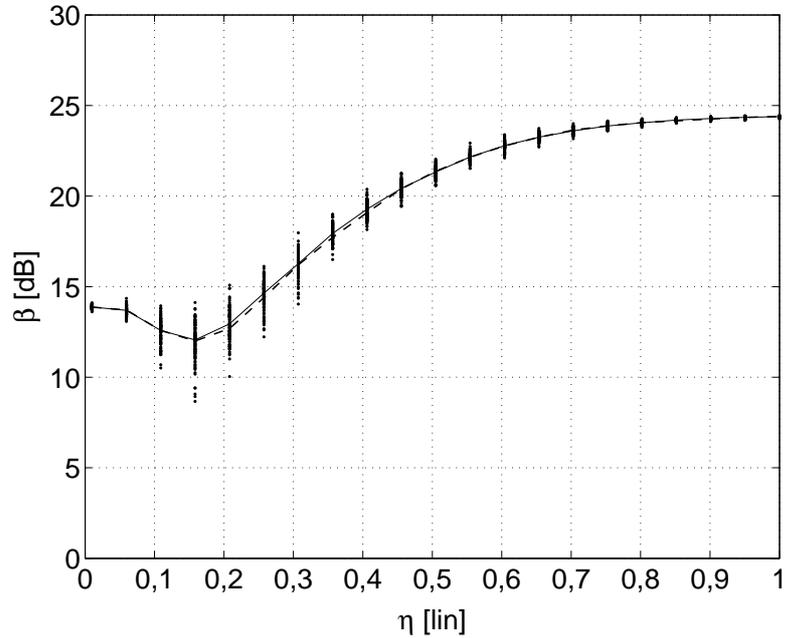
with  $\tilde{\mathbf{a}}_k \in \mathbb{C}^{MN \times 1}$  denoting the  $k$ -th column of  $\tilde{\mathbf{A}} \in \mathbb{C}^{MN \times K}$ . The corresponding output SINR of the LMMSE estimator for user  $k$  is given by

$$\text{SINR}_k = \frac{(\mathbf{f}_k^H \mathbf{a}_k)^2}{\mathbf{f}_k^H \mathbf{f}_k \sigma_v^2 + \sum_{i=1, i \neq k}^K (\mathbf{f}_i^H \mathbf{a}_i)^2}. \quad (\text{F.11})$$

Figs. F.1 and F.2 show 100 SINRs (F.11) of one particular user and the associated mean SINR for  $\{M = 20, J = 2\}$  and  $\{M = 100, J = 10\}$ , respectively. We also plot the analytical result conveyed through (6.7). The plots indicate that the analytical description through the results from the replica method show strong agreement with the experiments. The dynamic range of the SINRs becomes smaller with a growing spreading factor  $N$  [Tse99, Tse00].



**Figure F.1:** Comparison of analytical SINR (‘—’) results through the replica method and experimentally evaluated SINRs and their mean (‘-’) for a system with  $M = 20$ ,  $J = 2$ ,  $\alpha = K/N = 32/48 = 1,5$  and  $E_b/N_0 = 5$  dB.



**Figure F.2:** Comparison of analytical SINR (‘—’) results through the replica method and experimentally evaluated SINRs and their mean (‘-’) for a system with  $M = 100$ ,  $J = 10$ ,  $\alpha = K/N = 32/48 = 1,5$  and  $E_b/N_0 = 5$  dB.



# G Mathematical Tools

## G.1 The Matrix Inversion Lemma

**Lemma 1** Let  $\mathbf{A} \in \mathbb{C}^{m \times m}$  be a non-singular matrix,  $\mathbf{B}$  and  $\mathbf{C}$  be complex-valued matrices of dimension  $m \times n$  and  $n \times n$ , respectively. If  $(\mathbf{A} + \mathbf{B}\mathbf{C}\mathbf{B}^H)$  is invertible, its inverse can be rewritten as

$$(\mathbf{A} + \mathbf{B}\mathbf{C}\mathbf{B}^H)^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C} + \mathbf{B}^H\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{A}^{-1}. \quad (\text{G.1})$$

This relation is called the matrix inversion lemma, also known as Woodbury identity [Hay91]. For the special case of vectors it can be simplified to

$$\begin{aligned} (\mathbf{A} + \mathbf{b}\mathbf{b}^H)^{-1} &= \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{b}(1 + \mathbf{b}^H\mathbf{A}^{-1}\mathbf{b})^{-1}\mathbf{b}^H\mathbf{A}^{-1} \\ &= \mathbf{A}^{-1} - \frac{\mathbf{A}^{-1}\mathbf{b}\mathbf{b}^H\mathbf{A}^{-1}}{1 + \mathbf{b}^H\mathbf{A}^{-1}\mathbf{b}}. \end{aligned} \quad (\text{G.2})$$

## G.2 Numerical Integration

Whenever integrals cannot be given in closed, analytic form, numerical methods become inevitable. In this case the integration interval  $[-h, +h]$  is discretised into a finite number of  $n$  sections. The contribution of one section can be approximated by the piece-wise approximation of the integrand function  $f(x)$ . Typical approximation functions are the linear function or polynomial functions of low degree. When it comes to the evaluation of  $R$ -dimensional expressions, even numerical computation can become prohibitive since in those cases the number of operations scales with  $\mathcal{O}(n^R)$ . For these cases one can apply methods dating back to Gauss, Lagrange, and Laguerre [Abr70]. Considering a finite interval, Gaussian integration proves to be particularly useful. The rule of Gauss states that the integral over  $f(x)$  on the range  $[-h, +h]$  can be expressed as [Eng87]

$$\int_{-h}^{+h} f(x) dx = Q_n(-h, +h) + E_n(-h, +h) = \sum_{i=1}^n w_i f(x_i) + \mathcal{O}(h^{2n+1}).$$

The variable  $n$  expresses the number of points of support,  $Q_n(-h, +h)$  denotes the Gaussian quadrature part, and  $E_n(-h, +h)$  is the error part. The term  $\mathcal{O}(h^{2n+1})$

denotes the local order of the error for  $n$  points of support on the interval  $[-h, +h]$ . The integrand function  $f(x)$  is evaluated and weighted at  $n$  discrete points  $x_i$ . The values of the weights  $w_i$  as well as the points of support  $x_i$  are obtained from a transformation from values  $x'_i, w'_i$  that are given for the range  $[-1, +1]$  according to

$$\begin{aligned}x_i &= hx'_i, \\w_i &= hw'_i.\end{aligned}$$

Look-up tables parameterised by  $n$  are given in [Abr70]. The error term looks like

$$E_n(-h, +h) = \frac{2^{2n+1}(n!)^4}{(2n+1)((2n)!)^3} h^{2n+1} f^{2n}(\zeta) \quad \text{with } \zeta \in [-h, +h].$$

The degree  $n$  determines the accuracy of the approximation and is given in look-up tables, *e.g.*, [Abr70].

For the particular case of a Gaussian distribution  $f(\lambda) \sim \mathcal{N}(\mu, 2\mu)$  the integration limit  $h$  can be calculated solving the inequality

$$\frac{p(\mu + h)}{p(\mu)} = \frac{p(\mu - h)}{p(\mu)} \leq \epsilon \quad \text{with } \epsilon \in [0, 1]. \quad (\text{G.3})$$

Since, from the centre  $\mu$ , the Gaussian distribution is a monotonically decreasing function to both sides, we consider all contributions where the amplitude is larger than  $\epsilon$  times the maximum  $p(\mu)$ . The measure  $\epsilon$  can be interpreted as accuracy factor. When shifting the distribution by  $-\mu$ , the value  $h$  satisfying (G.3) is readily found by

$$h \geq \sqrt{-4 \log \epsilon}.$$

# Bibliography

- [3GPP] Members of 3GPP. The 3rd Generation Partnership Project (3GPP), Oct. 2004. <http://www.3gpp.org>.
- [Abr70] M. Abramowitz and I. A. Stegun, Editors. *Handbook of Mathematical Functions*. Dover, 1970.
- [Ale00b] P. D. Alexander and A. J. Grant. Iterative Channel and Information Sequence Estimation in CDMA. In *Proceedings IEEE Symposium on Spread Spectrum Technology Applications (ISSTA)*, pp. 593–597, Parsippany (NJ), USA, Sept. 2000.
- [Ale98] P. D. Alexander, A. J. Grant and M. C. Reed. Iterative Detection in Code-Division Multiple-Access with Error Control Coding. *European Transactions on Telecommunications*, Vol. 9, No. 5, pp. 419–426, Sept.–Oct. 1998.
- [Alo99] M.-S. Alouini and A. J. Goldsmith. A Unified Approach for Calculating Error Rates of Linearly Modulated Signals in Generalized Fading Channels. *IEEE Transactions on Communications*, Vol. 47, No. 9, pp. 1324–1334, Sept. 1999.
- [Bah74] L. R. Bahl, J. Cocke, F. Jelinek and J. Raviv. Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate. *IEEE Transactions on Information Theory*, pp. 284–287, Mar. 1974.
- [Bal00] J. Balakrishnan, M. Rupp and H. Visvanathan. Optimal Channel Training for Multiple Antenna Systems. In *Proceedings IEEE Multiaccess, Mobility, and Teletraffic for Wireless Communications*, Duck Key (FL), USA, Dec. 2000.
- [Bau98] G. Bauch and V. Franz. A Comparison of Soft-In/Soft-Out Algorithms for Turbo-Detection. In *Proceedings International Conference on Telecommunications (ICT)*, pp. 259–262, Porto Carras, Greece, June 1998.
- [Bea00] N. C. Beaulieu, A. S. Toms and D. R. Pauluzzi. Comparison of Four SNR Estimators for QPSK Modulations. *IEEE Transactions on Communications*, Vol. 4, No. 2, pp. 43–45, Feb. 2000.

- [Bel57] R. Bellman. *Dynamic Programming*. Princeton University Press, Boston, 1957.
- [Bel63] P. A. Bello. Characterization of Randomly Time-Variant Linear Channels. *IEEE Transaction Journal on Communication Systems*, Vol. CS-11, pp. 360–393, Dec. 1963.
- [Ber93] C. Berrou, A. Glavieux and P. Thitimajshima. Near Shannon Limit Error-Correcting Coding and Decoding: Turbo-Codes. In *Proceedings IEEE International Conference on Communications (ICC)*, pp. 1064–1070, Geneva, Switzerland, May 1993.
- [Bou02] J. Boutros and G. Caire. Iterative Multiuser Joint Decoding: Unified Framework and Asymptotic Analysis. *IEEE Transactions on Information Theory*, Vol. 48, No. 7, pp. 1772–1793, July 2002.
- [Bre59] D. G. Brennan. Linear Diversity Combining Techniques. *Proceedings of the IRE*, Vol. 47, pp. 1075–1102, June 1959.
- [Bri01] S. ten Brink. Convergence Behavior of Iteratively Decoded Parallel Concatenated Codes. *IEEE Transactions on Communications*, Vol. 49, No. 10, pp. 1727–1737, Oct. 2001.
- [Cai01] G. Caire and U. Mitra. Structured Multiuser Channel Estimation for Block-Synchronous DS/CDMA. *IEEE Transactions on Communications*, Vol. 49, No. 9, pp. 1605–1617, Sept. 2001.
- [Cai01a] G. Caire and R. R. Müller. The Optimal Received Power Distribution for IC-based Iterative Multiuser Joint Decoders. In *Proceedings Allerton Conference on Communications, Control and Computing*, Monticello (IL), USA, Oct. 2001.
- [Cai04] G. Caire, R. R. Müller and T. Tanaka. Iterative Multiuser Joint Decoding: Optimal Power Allocation and Low-Complexity Implementation. *IEEE Transactions on Information Theory*, Vol. 50, No. 9, pp. 1950–1973, Sept. 2004.
- [Cai98] G. Caire, G. Taricco and E. Biglieri. Bit-Interleaved Coded Modulation. *IEEE Transactions on Information Theory*, Vol. 44, No. 3, pp. 927–946, May 1998.
- [Cai01c] G. Caire, A. Tulino and E. Biglieri. Iterative Multiuser Joint Detection and Parameter Estimation: A Factor-Graph Approach. In *Proceedings IEEE Inform. Theory Workshop (ITW)*, Cairns, Australia, Sept. 2001.

- [Chu01] S. Y. Chung, T. J. Richardson and R. L. Urbanke. Analysis of Sum-Product Decoding of Low-Density Parity-Check Codes Using a Gaussian Approximation. *IEEE Transactions on Information Theory*, Vol. 2, No. 47, pp. 657–670, Feb. 2001.
- [Cor01] L. M. Correia, Editor. *Wireless Flexible Personalised Communications*. John Wiley & Sons, Ltd., 2001.
- [Cor05] L. M. Correia, Editor. *Towards Mobile Broadband Multimedia Networks*. Elsevier, to be published.
- [Cro91] S. N. Crozier, D. D. Falconer and S. A. Mahmoud. Least Sum of Squared Errors (LSSE) Channel Estimation. *Proc. Institute of Electronics Engineers*, Vol. 138, No. pt. F, pp. 371–378, Aug. 1991.
- [Don02] M. Dong and L. Tang. Optimal Design and Placement of Pilot Symbols for Channel Estimation. *IEEE Transactions on Signal Processing*, Vol. 50, No. 12, pp. 3055–3069, Dec. 2002.
- [Dou95] C. Douillard, M. Jezequel, C. Berrou, A. Picart, P. Didier and A. Glavieux. Iterative Correction of Intersymbol Interference: Turbo Equalization. *European Transactions on Telecommunications*, pp. 507–511, Sept.–Oct. 1995.
- [ElGa00] H. El-Gamal and E. Geraniotis. Iterative Multiuser Detection for Coded CDMA Signals in AWGN and Fading Channels. *IEEE Journal on Selected Areas in Communications*, Vol. 18, No. 1, pp. 30–41, Jan. 2000.
- [Eng87] G. Engeln-Müllges and F. Reutter. *Formelsammlung zur Numerischen Mathematik mit C-Programmen*. Bibliographisches Institut Wissenschaftsverlag, Mannheim-Wien-Zürich, 1987.
- [Eva00] J. S. Evans and D. N. C. Tse. Large System Performance of Linear Multiuser Receivers in Multipath Fading Channels. *IEEE Transactions on Information Theory*, Vol. 46, No. 6, pp. 2059–2078, Sept. 2000.
- [Guo02] D. Guo and S. Verdú. Multiuser Detection and Statistical Mechanics. In V. Bhargava, H. V. Poor, V. Tarokh and S. Yoon, Editors, *Communications, Information and Network Security*, pp. 229–277. Kluwer Academic Publisher, 2002.
- [Hag89] J. Hagenauer and P. Hoeher. A Viterbi Algorithm with Soft-Decision Outputs and its Applications. In *Proceedings IEEE Global Telecommunications Conference (GLOBECOM)*, Vol. 3, pp. 1680–1686, Dallas (TX), USA, Nov. 1989.

- [Has03] B. Hassibi and B. M. Hochwald. How Much Training is Needed in Multiple-Antenna Wireless Links? *IEEE Transactions on Information Theory*, Vol. 49, No. 4, pp. 951–963, Apr. 2003.
- [Hay91] S. Haykin. *Adaptive Filter Theory*. Prentice Hall, 2nd edition, 1991.
- [Hof02] H. Hofstetter. Webpage of the MIMO Channel Sounding Campaign by Forschungszentrum Telekommunikation Wien (ftw.), Nov. 2002. <http://www.ftw.at/measurements>.
- [Hof04a] H. Hofstetter and G. Steinböck. A Geometry based Stochastic Channel Model for MIMO, TD(04)060. In *Proceedings COST273 meeting*, Athens, Greece, Jan. 2004.
- [Hor99] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, 1999.
- [IEEE802] IEEE. The IEEE 802 LAN/MAN Standards Committee, Oct. 2004. <http://ieee802.org>.
- [Jak75] W. C. Jakes. *Microwave Mobile Communications*. John Wiley & Sons, Ltd., Feb. 1975.
- [Kay93] S. M. Kay. *Fundamentals of Statistical Signal Processing I - Estimation Theory*. Prentice Hall, 1993.
- [Kob01] M. Kobayashi, J. Boutros and G. Caire. Successive Interference Cancellation With SISO Decoding and EM Channel Estimation. *IEEE Transactions on Communications*, Vol. 19, No. 8, pp. 1450–1460, Aug. 2001.
- [Koc90] W. Koch and A. Baier. Optimum and Sub-Optimum Detection of Coded Data Disturbed by Time-Varying Intersymbol Interference. In *Proceedings IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 1679–1684, San Diego (CA), USA, Dec. 1990.
- [Koc03a] A. Kocian. *EM-Based Joint Data Detection and Channel Estimation of DS/CDMA Signals*. PhD thesis, Aalborg Universitet, Denmark, Oct. 2003.
- [Küh03] V. Kühn. Iterative Interference Cancellation and Channel Estimation for OFDM-CDMA. In *Proceedings IEEE International Conference on Communications (ICC)*, Anchorage (AK), USA, May 2003.
- [Lam02] A. Lampe. Iterative Multiuser Detection With Integrated Channel Estimation for Coded DS-CDMA. *IEEE Transactions on Communications*, Vol. 50, No. 8, pp. 1217–1223, Aug. 2002.

- [Lam03] A. Lampe. *Multiuser Detection and Channel Estimation for DS-CDMA Systems*. PhD thesis, University of Erlangen-Nürnberg, Germany, 2003.
- [Li03] K. Li and X. Wang. EXIT Chart Analysis of Turbo Multiuser Detection. In *Proceedings Allerton Conference on Communications, Control and Computing*, pp. 1675–1677, Monticello (IL), USA, Oct. 2003.
- [Lon04] M. Lončar, R. R. Müller, J. Wehinger, C. Mecklenbräuker and T. Abe. Iterative Channel Estimation and Data Detection in Frequency Selective Fading MIMO Channels. *European Transactions on Telecommunications*, Vol. 15, No. 5, pp. 459–470, Sept.–Oct. 2004.
- [Lup89] R. Lupas and S. Verdú. Linear Multiuser Detectors for Synchronous Code-Division Multiple-Access Channels. *IEEE Transactions on Information Theory*, Vol. 35, No. 1, pp. 123–136, Jan. 1989.
- [Man89] U. Manber. *Introduction to Algorithms*. Addison Wesley, 1989.
- [Mar01] S. Marinkovic, B. Vucetic and J. Evans. Improved Iterative Parallel Interference Cancellation. In *Proceedings IEEE Int. Symp. on Inform. Theory (ISIT)*, p. 34, Cairns, Australia, June 2001.
- [Mar99] T. Marzetta. BLAST Training: Estimating Channel Characteristics for High-Capacity Space-Time Wireless. In *Proceedings Allerton Conference on Communications, Control and Computing*, Monticello (IL), USA, Sept. 1999.
- [Mec05] C. F. Mecklenbräuker and S. Paul. On Estimating the Signal to Noise Ratio from BPSK Signals. In *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing. (ICASSP)*, pp. 65–68, Philadelphia (PA), USA, Mar. 2005.
- [Mey97] H. Meyr, M. Moeneclaey and S. A. Fechtel. *Digital Communication Receivers*. Telecommunications and Signal Processing. John Wiley & Sons, Ltd., 1997.
- [Moh98] M. Moher. An Iterative Multiuser Decoder for Near-Capacity Communications. *IEEE Transactions on Communications*, Vol. 46, No. 7, pp. 870–880, July 1998.
- [Mos96] S. Moshavi. Multi-User Detection for DS-CDMA Communications. *IEEE Communications Magazine*, pp. 124–136, Oct. 1996.
- [Mow95] W. H. Mow. *Sequence Design for Spread Spectrum*. The Chinese University Press, Hong Kong, China, 1995.

- [Mow95a] W. H. Mow. *A Study of Correlation of Sequences*. PhD thesis, The Chinese University of Hong Kong, China, May 1995.
- [Mül02] R. R. Müller and G. Caire. Efficient Implementation of Iterative Multiuser Decoding. In *Proceedings IEEE International Symposium on Information Theory (ISIT)*, Lausanne, Switzerland, July 2002.
- [Mül03] R. R. Müller. Application of Large Random Matrices in Communications Engineering. In *Proceedings Int. Conf. on Advances in the Internet, Processing, Systems and Interdisciplinary Research (IPSI)*, Sveti Stefan, Montenegro, Oct. 2003.
- [Mül04] R. R. Müller and W. Gerstacker. On the Capacity Loss Due to Separation of Detection and Decoding. *IEEE Transactions on Information Theory*, Vol. 50, No. 8, pp. 1769–1777, Aug. 2004.
- [Mül04a] R. R. Müller and A. M. Tulino. Minimum Bit Error Probability of Large Randomly Spread MC-CDMA Systems in Multipath Rayleigh Fading. In *ISSTA*, Sydney, Australia, Sept. 2004.
- [Ng98] J. C. L. Ng, K. B. Letaief and R. D. Murch. Complex Optimal Sequences with Constant Magnitude for Fast Channel Estimation Initialization. *IEEE Transactions on Communications*, Vol. 46, No. 3, pp. 305–308, Mar. 1998.
- [Nis01] H. Nishimori. *Statistical Physics of Spin Glasses and Information Processing*. Oxford Science Publications, 2001.
- [Nor02] A. Nordio, M. Hernandez and G. Caire. Design and Performance of a Low-Complexity Iterative Multiuser Joint Decoder Based on Viterbi Decoding and Parallel Interference Cancellation. In *Proceedings IEEE International Conference on Communications (ICC)*, New York (NYDC), USA, Apr. 2002.
- [Pau03] A. Paulraj, R. Nabar and D. Gore. *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2003.
- [Pro00] J. G. Proakis. *Digital Communications*. McGraw Hill, 4th edition, Aug. 2000.
- [Rap96] T. Rappaport. *Wireless Communications, Principles and Practice*. Prentice Hall, 1st edition, Jan. 1996.
- [Ras03a] L. Rasmussen, A. Grant and C. Schlegel. Leading Multiuser Detection Towards Adulthood. In *Australien Communication Theory Workshop*, Adelaide, Australia, Feb. 2003.

- [Ree99] M. C. Reed. *Iterative Receiver Techniques for Coded Multiple Access Communications Systems*. PhD thesis, The University of South Australia, Sydney, Australia, Oct. 1999.
- [Ree98] M. C. Reed, C. B. Schlegel, P. D. Alexander and J. A. Asenstorfer. Iterative Multiuser Detection for CDMA with FEC: Near-Single-User Performance. *IEEE Transactions on Communications*, Vol. 46, No. 12, pp. 1693–1699, Dec. 1998.
- [Ric01a] T. J. Richardson, A. Shokrollahi and R. L. Urbanke. Design of Capacity-Approaching Irregular Low-Density Parity-Check Codes. *IEEE Transactions on Information Theory*, Vol. 47, No. 2, pp. 619–637, Feb. 2001.
- [Rup00] M. Rupp. Fast Implementation of the LMS Algorithm. In *Proceedings European Signal Processing Conference (EUSIPCO)*, Tampere, Finland, Sept. 2000.
- [Sar03] K. Sarrigeorgidis and J. M. Rabaey. Massively Parallel Wireless Reconfigurable Processor Architecture and Programming. In *Proceedings Reconfigurable Architectures Workshop (RAW)*, Nice, France, Apr. 2003.
- [Sch04a] C. Schlegel. Iterative Joint Detection Using Recursive Signal Cancellation. In *Proceedings International Conference on 3G Mobile Communications Technology*, London, UK, Sept. 2004.
- [Sgr01] C. Sgraja, W. G. Teich, A. Engelhart and J. Lindner. Multiuser/Multisubchannel Detection Based on Recurrent Neural Network Structures for Linear Modulation Schemes with General Complex-Values Symbol Alphabet. In *Proceedings COST262 Workshop Multiuser Detection in Spread Spectrum Communications*, pp. 45–52, Schloss Reisingen, Ulm, Germany, Jan. 2001.
- [Tan02a] 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, Nov. 2002.
- [TR101 112] Members of ETSI. Selection procedures for the choice of radio transmission technologies of the Universal Mobile Telecommunications System (UMTS) (UMTS 30.03 Version 3.2.0). Standardization Document, ETSI, Apr. 1998.
- [TR25.858] Members of 3GPP. Technical Specification Group Radio Access Network; Physical Layer Aspects of UTRA High Speed Downlink Packet Access,

- (3GTS 25.858 Version 5.0.0 Release 5). Standardization Document, 3GPP, Mar. 2002.
- [Tre68] H. L. V. Trees. *Detection, Estimation, and Modulation Theory*. John Wiley & Sons, Ltd., 1968.
- [TS25.101] Members of 3GPP. Technical Specification Group Radio Access Network; User Equipment(UE) radio transmission and reception (FDD) (3GTS 25.101 Version 6.1.0). Standardization Document, 3GPP, June 2003.
- [TS25.201] Members of 3GPP. Technical Specification Group Radio Access Network; Physical layer - General description (FDD) (3GTS 25.201 Version 5.2.0). Standardization Document, 3GPP, Sept. 2002.
- [TS25.211] Members of 3GPP. Technical Specification Group Radio Access Network; Physical channel and mapping of transport channels onto physical channels (FDD) (3GTS 25.211 Version 5.3.0). Standardization Document, 3GPP, Dec. 2002.
- [TS25.212] Members of 3GPP. Technical Specification Group Radio Access Network; Multiplexing and channel coding (FDD) (3GTS 25.212 Version 5.3.0). Standardization Document, 3GPP, Dec. 2002.
- [TS25.213] Members of 3GPP. Technical Specification Group Radio Access Network; Spreading and modulation (FDD) (3GTS 25.213 Version 4.3.0). Standardization Document, 3GPP, June 2002.
- [Tse99] D. N. C. Tse and S. V. Hanly. Linear Multiuser Receivers: Effective Interference, Effective Bandwidth and User Capacity. *IEEE Transactions on Information Theory*, Vol. 45, No. 2, pp. 641–657, Mar. 1999.
- [Tse00] D. N. C. Tse and O. Zeitouni. Linear Multiuser Receivers in Random Environments. *IEEE Transactions on Information Theory*, Vol. 46, No. 1, pp. 171–188, Jan. 2000.
- [Tse00a] K. Tse and D. N. C. Tse. Effective Interference and Effective Bandwidth of Linear Multiuser Receivers in Asynchronous CDMA Systems. *IEEE Transactions on Information Theory*, Vol. 46, No. 4, pp. 1426–1447, July 2000.
- [Var97] M. K. Varanasi and T. Guess. Optimum Decision Feedback Multiuser Equalization with Successive Decoding Achieves the Total Capacity of the Gaussian Multiple-Access Channel. In *Proceedings 31st Asilomar Conf. on Sig., Sys. and Comp.*, Pacific Grove (CA), USA, Nov. 1997.

- [Ver86] S. Verdú. Minimum Probability of Error for Asynchronous Gaussian Multiple-Access Channels. *IEEE Transactions on Information Theory*, Vol. 32, No. 1, pp. 85–96, Jan. 1986.
- [Ver98] S. Verdú. *Multiuser Detection*. Cambridge University Press, 1st edition, Sept. 1998.
- [Vit67] A. J. Viterbi. Error Bounds for Convolutional Codes and an Asymptotically Optimal Decoding Algorithm. *IEEE Transactions on Information Theory*, Vol. IT-13, pp. 260–269, 1967.
- [Wan99] X. Wang and H. V. Poor. Iterative (Turbo) Soft Interference Cancellation and Decoding for Coded CDMA. *IEEE Transactions on Communications*, Vol. 47, No. 7, pp. 1046–1061, July 1999.
- [Weh04] J. Wehinger, C. F. Mecklenbräuker, R. R. Müller, T. Zemen and M. Lončar. On Channel Estimators for Iterative CDMA Multiuser Receivers in Flat Fading. In *Proceedings IEEE International Conference on Communications (ICC)*, pp. 2497–2501, Paris, France, June 2004.
- [Weh02] J. Wehinger, R. R. Müller, M. Lončar and C. F. Mecklenbräuker. Performance of Iterative CDMA Receivers with Channel Estimation in Multipath Environments. In *Proceedings 36th Asilomar Conf. on Sig., Sys. and Comp.*, pp. 1439–1443, Pacific Grove (CA), USA, Nov. 2002.
- [Wick03] S. B. Wicker and S. Kim. *Fundamentals of Codes, Graphs, and Iterative Decoding*. KAP, 2003.
- [Wor01] A. P. Worthen and W. E. Stark. Unified Design of Iterative Receivers Using Factor Graphs. *IEEE Transactions on Information Theory*, Vol. 47, No. 2, pp. 843–849, Feb. 2001.
- [Xia04] P. Xiao. *Iterative Detection, Decoding and Channel Parameter Estimation for Orthogonally Modulated DS-CDMA Systems*. PhD thesis, Chalmers Tekniska Högskola, Göteborg, Sweden, Jan. 2004.
- [Zem03a] T. Zemen, M. Lončar, J. Wehinger, C. F. Mecklenbräuker and R. R. Müller. Improved Channel Estimation for Iterative Receivers. In *Proceedings IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 257–261, San Francisco (CA), USA, Nov. 2003.
- [Zha01] J. Zhang, E. K. P. Chong and D. N. C. Tse. Output MAI Distributions of Linear MMSE Multiuser Receivers in DS-CDMA Systems. *IEEE Transactions on Information Theory*, Vol. 47, No. 3, pp. 1128–1144, Mar. 2001.

- [Zhe03] Y. R. Zheng and C. Xiao. Simulation Models With Correct Statistical Properties for Rayleigh Fading Channels. *IEEE Transactions on Communications*, Vol. 51, No. 6, pp. 920–928, June 2003.

# Curriculum Vitae

Joachim Wehinger, citizen of Austria  
born on November 8th, 1975, in Hohenems, Austria

married to Margarethe Wehinger, née Lechner, since 2003  
daughter Clara Elisabeth, born 2005

## Education

- |           |  |
|-----------|--|
| 2001-2005 | <i>Technische Universität Wien</i><br>PhD studies in electrical engineering                                |
| 1996-2001 | <i>Technische Universität Wien</i><br>Diploma studies in electrical engineering                            |
| 1999-2001 | <i>Chalmers Tekniska Högskola, Göteborg, Sweden</i><br>Master of science studies in digital communications |
| 1990-1995 | <i>Höhere Technische Lehranstalt Rankweil</i>  |
| 1986-1990 | <i>Bundesgymnasium Dornbirn</i>  |

## Employment/Internships

- |                      |  |
|----------------------|--|
| since Sep. 2001      | <i>Forschungszentrum Telekommunikation Wien (ftw.)</i><br>Research in MIMO systems and multi-user receivers                              |
| June 2000 - May 2001 | <i>Austrian Aerospace, Wien</i><br>Diploma thesis - Analysis of GPS signals for atmospheric sounding and design of occultation receivers |
| July 1999            | <i>HighQLaser, Hohenems</i><br>Control electronics testing of a cooling device   |
| July-Aug. 1998       | <i>Leica Geosystems AG, Heerbrugg, Switzerland</i><br>Database development for customer services   |
| July-Aug. 1997       | <i>Leica Geosystems AG, Heerbrugg, Switzerland</i><br>Control electronics programming for theodolites                                    |
| July-Aug. 1996       | <i>Leica Geosystems AG, Heerbrugg, Switzerland</i><br>Database development for repair requests   |

## National Service

Oct. 1995 - May 1996 Jägerregiment 9, Bludesch

## Hobbies

Playing the oboe, chamber music, mountaineering, travelling, sailing