Low complexity receiver architectures for high-speed wireless multiple-input multiple-output (MIMO) systems

Holger Claussen

A thesis submitted for the degree of Doctor of Philosophy.
The University of Edinburgh.
March 2004

Abstract

In modern wireless networks the demand for high-speed transmissions is ever increasing to provide access to data and enable new services anywhere and anytime. Mobile internet, video telephony, music and video on demand are examples for the possible applications which demand high data-rates. However, the available frequency spectrum is limited and expensive. To satisfy the demand for high data-rates, turbo encoded multiple-input multiple-output (MIMO) radio links have been recently proposed for the support of high-speed downlink packet access (HSDPA) in UMTS, where the re-use of spreading codes across the transmitter antennas results in high levels of interference. In this thesis, low complexity MIMO receiver architectures and their components are investigated to enable high-speed receivers capable of dealing with high-order modulations. For detection, multi-stage partial parallel interference cancellation (MS-PPIC) and matched filter based ordered serial interference cancellation (MF-SIC) are proposed as low complexity alternatives to the *a posteriori* probability (APP) detector and its Max-Log-APP variant. Non-linear cancellation metrics are derived for the MS-PPIC and the performance of the proposed detectors is investigated for different channel conditions. It is shown that the MS-PPIC can provide similar performance compared to the APP and, for low coding rates, superior performance compared to the Max-Log-APP, at a substantially lower computational complexity. While the MF-SIC cannot compete with the APP and MS-PPIC detectors in non-iterative receivers, it provides impressive performance when used in an iterative receiver architecture where *a priori* information from a decoder is available for ordering and interference cancellation. For decoding, a novel modification of the turbo decoder using the Max-Log-MAP algorithm is proposed, which results in a performance approaching that of a turbo decoder using the optimum Log-MAP or MAP algorithms. The approach aims to maximise the mutual information at the input of each component decoder by correcting the bias in the *a priori* information caused by the Max-Log approximation in the previous component decoder. This is performed by scaling the *a priori* information by optimised iteration-specific weight factors at each turbo iteration. Another contribution is a method for the off-line computation of the optimal weights according to the maximum mutual information criterion. Subsequently, different versions of non-iterative and iterative MIMO receiver architectures are investigated and compared in terms of performance and computational complexity for a wide range of detection algorithms using 4-QAM modulation. For iterative receivers which rely on hard cancellation, a soft-output combining scheme which maximises the mutual information at each iteration is proposed and a corresponding method for the offline computation of the optimal combining weights is presented. Finally, a novel layered encoding scheme is proposed which overcomes the problem of exponential growth in complexity of the APP detector when higher order modulations such as 16- and 64-QAM are employed. This could be achieved without any loss in performance. Layered encoding also solves the performance and convergence problems of low complexity detectors such as the proposed MS-PPIC and the MF-SIC which occur at higher order modulations. In addition, the applicability of high order modulations in MIMO systems is investigated using system-level simulations for a 2-cell indoor and a 7-cell urban scenario. The results indicate that high-order modulations could be used in a substantial area of the cell.

Declaration of originality

I hereby declare that the research recorded in this thesis and the thesis itself was composed and originated entirely by myself in the Global Wireless Systems Research Department at Bell Laboratories, Lucent Technologies, Swindon and in the Department of Electronics and Electrical Engineering at the University of Edinburgh.

Holger Claussen

Acknowledgements

I would like to thank my supervisors Dr. Reza Karimi from Bell Laboratories, Lucent Technologies and Prof. Bernard Mulgrew from the University of Edinburgh for their invaluable guidance and assistance during the course of the last tree years of research.

I would also like to thank Dr. Magnus Sandell for his help in the area of turbo-decoding, Dr. Francis Mullany for valuable discussions on receiver architectures and Dr. Stephan ten Brink for his useful input during discussions on the subject of EXIT charts. In addition, I would like to acknowledge Dr. Sivarama Venkatesan and Dr. Laurence Mailaender for their correspondence on space-time equalisation.

Finally, I would like to thank Lucent Technologies for sponsoring my work, and particularly Dr. Richard Howard, Dr. Ran Yan and Dr. Sam Samuel for giving me the opportunity to undertake my studies in cooperation with the Global Wireless Systems Research Department of Bell Laboratories, Lucent Technologies in Swindon, UK.

Contents

List of figures

List of tables

List of symbols

$\boldsymbol{A}_k$	transformation matrix corresponding to the signal, de-spread with spreading code k
$\underline{a}^{(i)}$	i^{th} column of the transformation matrix $\boldsymbol{A}$ corresponding to the contributions of the i^{th} transmitter antenna
$b_{k,m}^{(i)}(t)$	bit transmitted as m^{th} bit in the modulation scheme via the i^{th} antenna and the k^{th} spreading code at the symbol interval t
b_q	q^{th} bit in the symbol vector $\underline{x}$
b_t	bit corresponding to the binary symbol x_t of a turbo encoded data block
$\boldsymbol{C}$	spreading matrix
$\boldsymbol{C}_k$	spreading matrix for spreading code k
C	channel capacity
$\underline{c}_k$	spreading code k
$\boldsymbol{D}$	distortion matrix after equalisation
E_{b}	energy per information bit
$\underline{e}$	equalised signal vector
G	forward code polynomial for turbo code
G_{r}	backward (recursive) code polynomial for turbo code
$G_{\text{ant}}^{\text{UE}}$	antenna gain at the mobile (UE)
$G_{\text{ant}}^{\text{BS}}$	antenna gain at the base-station (BS)
${}^{(u)}G$	power gain between the u^{th} base-station and the mobile
$\boldsymbol{H}$	channel matrix
${}^{(j)}\boldsymbol{H}^{(i)}$	channel matrix corresponding to the channel between the i^{th} transmitter antenna and the j^{th} receiver antenna
$\boldsymbol{I}$	identity matrix
I_{a}	*a priori* mutual information
I_{e}	extrinsic mutual information
i	transmit antenna index $i \in \{1 \ldots N_{\text{T}}\}$
j	receive antenna index $j \in \{1 \ldots N_{\text{R}}\}$
K	number of used spreading codes
k	spreading code index $k \in \{1 \ldots K\}$
kT	single sided thermal noise power spectral density
$L_{\text{cable}}^{\text{UE}}$	cable loss at the mobile (UE)
$L_{\text{cable}}^{\text{BS}}$	cable loss at the base-station (UE)
L_{shadow}	shadow fading loss
L_{path}	path loss
l	code-state index of turbo code
M	number of bits in the modulation scheme
m	bit index in the modulation scheme $m \in \{1 \ldots M\}$
N_{T}	number of transmitter antennas
N_{R}	number of receiver antennas
N_{I}	number of iterations

N_{E}	Equaliser length (symbol epochs)
N_{V}	Transversal equaliser length (taps)
N_0	noise energy per symbol period
NF_{UE}	noise figure at the mobile
$\underline{n}$	noise vector of i.i.d. zero-mean complex Gaussian random variables (i.e. $\mathrm{E}\left\{\underline{n}\underline{n}^{\mathrm{H}}\right\} = N_0 I$)
${}^{(j)}\underline{n}$	noise vector at the j^{th} receiver antenna
n_t	noise at symbol interval t of a turbo-encoded data block (i.i.d. zero-mean complex Gaussian random variable where $\mathrm{E}\left\{n_t n_t^*\right\} = N_0$)
n	iteration index $n \in \{1 \ldots N_{\mathrm{I}}\}$
$P_{\mathrm{noise}}^{\mathrm{UE}}$	noise power at the mobile
$P_{\mathrm{interf}}^{\mathrm{UE}}$	interference power at the mobile
$P_{\mathrm{signal}}^{\mathrm{UE}}$	signal power at the mobile
P_{Tx}	power of the transmitted signal (at the base-station)
Q	spreading factor
q	bit index in the symbol vector
$\boldsymbol{R}_n$	correlation matrix of the noise vector $\underline{n}$
$\boldsymbol{R}_s$	correlation matrix of the spread symbol vector $\underline{s}$
$\boldsymbol{R}_\epsilon$	correlation matrix of the disturbance vector $\underline{\epsilon}$
R	coding rate
${}^{(j)}\underline{r}$	received signal at the j^{th} receiver antenna
r_t	received binary symbol at symbol interval t of a turbo encoded data block
S	number of states in the code trellis
$\underline{s}$	vector of spread symbols
T	number of transmitted symbols
t	symbol index, $t \in \{1 \ldots T\}$
u	base station index
$\boldsymbol{V}$	equaliser matrix
${}^{(i)}\boldsymbol{V}^{(j)}$	equaliser sub-matrix corresponding to the channel between the i^{th} transmitter antenna and the j^{th} receiver antenna
v	code memory of a turbo code
W	number of channel taps (chip-spaced)
${}^{(i)}\underline{w}^{(j)}$	transversal equaliser weights for the channel between the i^{th} transmitter antenna and the j^{th} receiver antenna
$w_a^{(n)}$	decoder scaling factor for the *a priori* information in a maximum mutual information combining (MMIC) scheme for turbo decoding
$\underline{x}$	transmitted symbol vector
$\underline{x}_k^{(i)}$	symbol vector containing symbols transmitted via the i^{th} antenna and the k^{th} spreading code
$x_k^{(i)}(t)$	symbol transmitted via the i^{th} antenna and the k^{th} spreading code at the symbol interval t
x_t	binary symbol at symbol interval t of a turbo encoded data block
$x_{t,0}$	binary symbol corresponding to the systematic bit at symbol interval t of a turbo encoded data block
$x_{t,1}$	binary symbol corresponding to the parity bit 1 bit at symbol interval t of a turbo encoded data block

$x_{t,2}$	binary symbol corresponding to the parity bit 2 bit at symbol interval t of a turbo encoded data block
$\underline{\hat{x}}_{k,L}$	reconstructed symbol vector of layer L containing symbols transmitted via the k^{th} spreading code
$\underline{y}$	detector (matched filter) output
$\underline{z}_k$	equalised signal vector, despread with spreading code k
$\underline{z}_{w,k}$	equalised signal vector, despread with spreading code k after pre-whitening (pre-whitened sufficient statistics)
$\overline{\alpha}_t(l')$	forward accumulated metric for turbo decoding at time t
$\overline{\beta}_t(l)$	backward accumulated metric for turbo decoding at time t
$\overline{\gamma}_t^{[q]}(l', l)$	logarithm of the probability of a transition from state l' to state l of the encoder trellis at time instant t
$\underline{\epsilon}$	disturbance vector of i.i.d. zero-mean complex Gaussian random variables of unit variance (i.e. $\mathrm{E}\left\{\underline{\epsilon}\underline{\epsilon}^{\mathrm{H}}\right\} = I$)
$\Lambda(.)$	LLR value of (.)
$\Lambda_c(.)$	LLR value of the channel output of (.) at the turbo decoder input
$\Lambda_a^{(n)}(.)$	LLR value of the *a priori* information on (.) at the component decoder input at iteration n
$\Lambda_e^{(n)}(.)$	LLR value of the extrinsic information on (.) at the component decoder output at iteration n
$\Lambda^{(n)}(.)$	LLR value of (.) at iteration n
$\lambda_a^{(n)}(.)$	"uncorrupted" LLR value of the *a priori* information on (.)
$\lambda_e^{(n)}(.)$	"uncorrupted" LLR value of the extrinsic information on (.)
τ	number of systematic bits (information bits) in a coded data block

Acronyms and abbreviations

3GPP	3rd generation partnership project
APP	*a posteriori* probability
AWGN	additive white gaussian noise
BER	bit error rate
BLAST	Bell Laboratories layered space-time
cdf	cumulative distribution function
CDMA	code division multiple access
DSC	decision statistics combining
EXIT	extrinsic information transfer chart
FER	frame error rate
GSM	global system for mobile
HSDPA	high-speed downlink packet access
ISI	intersymbol interference
LDPC	low density parity check
LE	layered encoding
LLR	log-likelihood ratio
MAI	multiple access interference
MAP	maximum *a posteriori*
MF	matched filter
MF-SIC	matched filter based serial interference cancellation
MIMO	multiple-input multiple-output
ML	maximum likelihood
MMIC	maximum mutual information combining
MMSE	minimum mean square error
MS-PPIC	multi-stage partial parallel interference cancellation
pdf	probability density function
PIC	parallel interference cancellation
QAM	quadrature amplitude modulation
QPSK	quadrature phase shift keying
SNR	signal to noise ratio
SIC	serial interference cancellation
SINR	signal to interference and noise ratio
UMTS	universal mobile telecommunication system
ZF	zero forcing

Chapter 1
Introduction

1.1 Motivation

New mobile applications and services such as high-speed internet, video telephony, music and video on demand provide new opportunities for the telecommunication providers to increase their revenues. To make these applications available for the customers, the data throughput of the mobile devices and base-stations needs to be increased significantly, and simultaneously the cost per transmitted bit must be reduced to make the new services attractive. Therefore, the demand for high-speed data capability is rapidly increasing from each generation of a wireless communication system to the next. However, on the way to higher data throughput, several problems arise:

- the available frequency spectrum is limited and expensive, therefore other solutions than higher bandwidth must be applied to increase the throughput;
- higher data-rates require more sophisticated receiver algorithms, which are more expensive and also consume more energy per received bit;
- the battery capacity is not increasing as fast, as the complexity of the receiver algorithms and the data-rates, which inevitably results in a shorter battery life for mobile devices.

To satisfy the demand for high data-rates, turbo encoded multiple-input multiple-output (MIMO) radio links have been recently proposed [1] for the support of high-speed downlink packet access (HSDPA) in UMTS, where the re-use of spreading codes across the transmitter antennas results in high levels of interference. MIMO radio communications offer the key to achieve high data rates by introducing space as a new dimension to the existing ones of frequency and time for data transmission [2–4]. In MIMO systems, data is transmitted simultaneously from several antennas in the same frequency spectrum. At the receiver, only the different signal paths are exploited to separate the data from the different antennas. In this way, the potential data-rate

scales linearly with the number of transmitter antennas, which allows extremely high spectral efficiencies.

While the optimum space-time algorithms for detection in terms of performance are well known, they have the disadvantage of exponential growth in computational complexity with the number of transmitter antennas, modulation order, spreading codes and transmitted symbols. As a result, the computational complexity is clearly prohibitive for typical parameter values, and would also lead to a very high energy consumption. Therefore, low complexity algorithms are required which enable both, high data-rates and long battery life for mobile devices, and also sacrifice as little performance as possible in comparison to the optimum space-time detector. However, not only the detection and decoding algorithms play a major role in achieving a high receiver performance and low complexity. Additionally, the receiver architecture must employ these algorithms as efficiently as possible to obtain the optimal results.

1.2 Thesis overview

In this thesis, low complexity MIMO receiver architectures and their components are investigated to enable low-complexity and high-performance receivers, capable of dealing with high-order modulations to meet the requirements for future generations of wireless communication systems. The thesis is organised in the following way:

- **Chapter 2:** In this chapter, space-time equalisation and several detection algorithms are discussed for MIMO radio links. It is shown that performing space-time equalisation prior to detection can significantly reduce the computational complexity, since it removes the effect of a dispersive channel and thereby restores orthogonality between the spreading codes. Then, the detection algorithm only needs to deal with the co-channel interference from other transmitter antennas in a MIMO system, and not with the interference from data transmitted via other spreading codes. MMSE space-time equalisation is described for block-based and transversal equalisation, followed by de-spreading and a pre-whitening process. Then, the performance of the equalisers is investigated for different block- and equaliser lengths, and their computational complexity is presented. For detection, the optimal *a posteriori* probability (APP) detector and its Max-Log variant are described. These algorithms are used as reference throughout the thesis. Subsequently, multistage partial parallel interference cancellation (MS-PPIC) is proposed as a

low complexity alternative to the APP detector, and non-linear cancellation metrics are derived for the MS-PPIC. In addition, an ordered serial interference cancellation scheme based on simple matched filter detection (MF-SIC) is proposed as an ultra-low complexity detector for iterative receiver architectures, where *a priori* information from a decoder is available for ordering and interference cancellation. Finally, the performance and the computational complexity of the proposed detection algorithms is compared with the APP and Max-Log-APP reference detectors.

- **Chapter 3:** Here, a novel modification of the Max-Log-MAP turbo decoder is proposed, which results in a performance approaching that of a turbo decoder using the optimum Log-MAP or MAP algorithms. The approach aims to maximise the mutual information at the input of each component decoder by correcting the bias in the *a priori* information caused by the Max-Log approximation in the previous component decoder. This is performed by scaling the *a priori* information by optimised iteration specific weight factors at each turbo iteration. A second contribution of this chapter is a method for the offline computation of the optimal weights according to the maximum mutual information criterion.

- **Chapter 4:** This chapter gives an overview of possible receiver architectures for MIMO. Employing the previously discussed components for equalisation, detection and decoding, a non-iterative and an iterative MIMO receiver architecture is proposed which provides both, high-performance and low-complexity. Different versions of the proposed receiver architectures are compared in terms of performance and computational complexity for a wide range of detection algorithms with 4-QAM modulation. For iterative receivers which rely on hard cancellation such as receivers employing the MF-SIC detector, a novel soft-output combining scheme, which maximises the mutual information at each iteration, is proposed and a corresponding method for the offline computation of the optimal combining weights is presented. The proposed weight calculation is then compared with a combining method based on decision statistics from the literature, and it is shown that the proposed method offers the same performance at a lower computational complexity, since the combining weights are calculated offline.

- **Chapter 5:** When higher order modulations such as 16- and 64-QAM are employed, several problems arise. For APP detection, the computational complexity grows exponentially with the modulation order, which results in a far too high complexity for

implementation. The proposed low complexity MS-PPIC detector offers only poor performance, and the MF-SIC detector shows convergence problems due to the poor initial matched filter estimates. In this chapter, a novel layered encoding scheme is proposed, which overcomes the problem of exponential growth in complexity of the APP detector when higher order modulations are employed. Layered encoding is essentially a special form of coded modulation using set-partitioning where bits with the same Euclidean distance in the modulation scheme are encoded in one code block at the transmitter. In this way, a high-order modulation scheme can be treated as a sum of 4-QAM modulated data blocks, which can be detected and decoded separately. It is shown that with layered encoding, the computational complexity of receivers using the APP algorithm for detection grows only linearly with the modulation order, and not exponentially as before. This can be achieved without any loss in performance. The proposed scheme also solves the performance and convergence problems of low complexity detectors such as the MS-PPIC and the MF-SIC which occur at higher order modulations and thereby enables to extend the proposed low-complexity MIMO receivers to modulations such as 16- and 64-QAM. In addition, the applicability of high-order modulations in MIMO systems is investigated using system-level simulations for a 2-cell indoor and a 7-cell urban scenario. It is shown that high-order modulations can be used in a substantial area of the cell for both investigated scenarios.

- **Chapter 6:** This chapter summarises the results of the thesis and gives an overview of the contributions to knowledge.

Chapter 2
High-performance MIMO detectors

2.1 Introduction

Turbo encoded multiple-input multiple-output (MIMO) radio links have been recently proposed for the support of high-speed downlink packet access (HSDPA) in UMTS [1]. The objective is to increase the achievable data rates for a particular user through a combination of spreading code re-use across transmit antennas and higher-order modulation schemes. The code re-use inevitably results in high levels of interference at the mobile receiver, even under non-dispersive channel conditions. In order to tackle such high interference levels, MIMO receivers based on the *a posteriori* probability (APP) detector have been considered [5] where, in order to deal with dispersive channels (while avoiding sequence estimation), the detector is preceded by space-time channel equalisation [6]. The equaliser is then followed by a de-spreading operation which allows the APP to perform joint detection of bits transmitted from multiple antennas but corresponding to a single spreading code only, thereby resulting in a significant reduction in computational complexity.

In this chapter, the receiver components for space-time equalisation and detection are presented in detail. The MIMO signal model is presented in Section 2.2. Section 2.3 concentrates on space-time equalisation, and in Section 2.4 different algorithms for detection are investigated. Multi-stage partial parallel interference cancellation (MS-PPIC) is proposed as a low complexity alternative to the APP detector for the CDMA MIMO downlink. An optimum strategy for nonlinear cancellation is derived analytically, and the performance of the MS-PPIC detector is compared with the APP and Max-Log-APP detectors. In addition to the MS-PPIC detector, matched filter based, ordered serial interference cancellation (MF-SIC) is proposed as another low complexity detector for the use in iterative receivers, where *a priori* information is available due to iterations between the detector and the decoder as presented in detail in Chapter 4. The MF-SIC detector performs hard cancellation of interference and relies on *a priori* information for bit-based cancellation and ordering to achieve a very good performance. Finally, the reductions in computational complexity are evaluated by a complexity analysis of the detection algorithms.

2.2 The MIMO signal model

Figure 2.1 illustrates the transmission and reception scheme for the MIMO link under investigation.

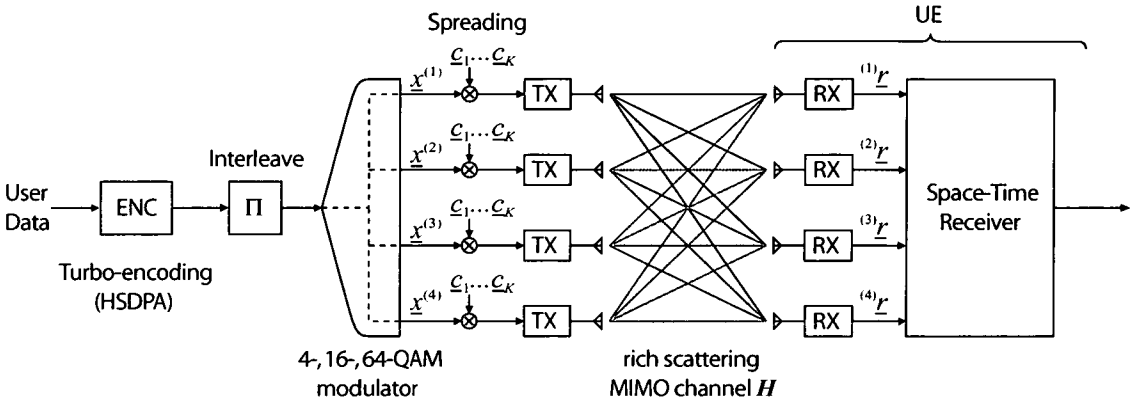


Figure 2.1: 4×4 *antenna MIMO radio link for HSDPA.*

At the transmitter, user data is encoded and interleaved. The coded data stream is de-multiplexed into N_{T} sub-streams, corresponding to the N_{T} transmitter antennas. Each sub-stream is then modulated on to TK 4-QAM symbols and subsequently spread by a factor Q via a set of K orthogonal spreading codes prior to transmission. Each transmitted spread stream then occupies T symbol intervals. Also note that the same set of K codes are re-used across all transmitter antennas. Therefore, the MIMO propagation environment, which is assumed to exhibit significant multipath, plays a major role in achieving signal separation at the receiver. The transmitted signals are received by N_{R} receiver antennas after propagation through dispersive radio channels with impulse response lengths of W chips. The received signal vector may then be written as

$$\begin{bmatrix} {}^{(1)}\underline{r} \\ \vdots \\ {}^{(N_{\mathrm{R}})}\underline{r} \end{bmatrix} = \begin{bmatrix} {}^{(1)}\boldsymbol{H}^{(1)} & \dots & {}^{(1)}\boldsymbol{H}^{(N_{\mathrm{T}})} \\ \vdots & \ddots & \vdots \\ {}^{(N_{\mathrm{R}})}\boldsymbol{H}^{(1)} & \dots & {}^{(N_{\mathrm{R}})}\boldsymbol{H}^{(N_{\mathrm{T}})} \end{bmatrix} \sum_{k=1}^{K} \boldsymbol{C}_k \begin{bmatrix} \underline{x}_k^{(1)} \\ \vdots \\ \underline{x}_k^{(N_{\mathrm{T}})} \end{bmatrix} + \begin{bmatrix} {}^{(1)}\underline{n} \\ \vdots \\ {}^{(N_{\mathrm{R}})}\underline{n} \end{bmatrix} \tag{2.1}$$

or

$$\underline{r} = \boldsymbol{H} \sum_{k=1}^{K} \boldsymbol{C}_k \underline{x}_k + \underline{n} = \boldsymbol{H}\boldsymbol{C}\underline{x} + \underline{n} = \boldsymbol{H}\underline{s} + \underline{n} \tag{2.2}$$

where ${}^{(j)}\underline{r} \in C^{(QT+W-1)\times 1}$ is the signal received at the j^{th} antenna, ${}^{(j)}\boldsymbol{H}^{(i)} \in C^{(QT+W-1)\times QT}$

is the channel matrix from the i^{th} transmitter antenna to the j^{th} receiver antenna and $\underline{x}_k^{(i)} \in C^{T\times 1}$ is the sequence of T symbols $\left[x_k^{(i)}(1)\ldots x_k^{(i)}(T)\right]^{\text{T}}$ transmitted from the i^{th} antenna via the k^{th} spreading code. Noise vector $\underline{n} \in C^{N_{\text{R}}(QT+W-1)\times 1}$ consists of i.i.d. zero-mean complex Gaussian random variables representing additive thermal noise and inter-cell interference such that $\boldsymbol{R}_n = \text{E}\left\{\underline{n}\underline{n}^{\text{H}}\right\} = N_0\boldsymbol{I}$. Finally $\boldsymbol{C}_k$ is the spreading matrix for k^{th} spreading code, $\underline{c}_k \in C^{Q\times 1}$, such that

$$\boldsymbol{C}_k = \begin{bmatrix} \underline{c}_k & \cdots & \underline{0} \\ \vdots & \ddots & \vdots \\ \underline{0} & \cdots & \underline{c}_k \end{bmatrix} \in C^{QN_{\text{T}}T\times N_{\text{T}}T}. \tag{2.3}$$

The mapping of the 4-QAM constellation is such that $x_k^{(i)}(t) = b_{k,0}^{(i)}(t) + jb_{k,1}^{(i)}(t)$ with $b_{k,m}^{(i)}(t) \in \{+1,-1\}$.

2.2.1 Interleaving and antenna multiplexing

The interleaving and the antenna multiplexing process is important for the overall system performance. Since in MIMO radio links some spatial channels are weaker than others, the corresponding bit estimates of the encoded data block have different average error probabilities dependent on their channel quality. When a number of adjacent bits in a coded data block are transmitted through weak radio channels, this can lead to decoding errors. Therefore the data blocks should be transmitted in such a way that adjacent bits in the code block are transmitted via different antennas in the MIMO system. This is the task for the interleaving and antenna multiplexing.

One option to achieve an optimum bit allocation is to design the interleaver in a way that, in combination with the antenna multiplexing, the adjacent bits in the code block are transmitted on different antennas and spread over the transmission time interval of each block to achieve maximum gain from the channel fading. Another option is to use a random interleaver for the coded data. Although this option is suboptimal, the random interleaving results in good performance and for simplicity, this solution is chosen for the investigations in the following chapters.

2.2.2 Channel modelling

For the MIMO channel simulation, tapped delay models are considered. The incoming reflections of the signal at each of the W channel taps are represented as complex channel values κ_w with $w = 1 \ldots W$. Each of the channel values consists of a real and imaginary part, modelled as random numbers whose elements are normally distributed with zero mean and unit variance. This leads to Rayleigh distributed channel values with zero mean and variance $\sigma_\kappa^2 = 2$. For the following investigations channels of two different dynamics are used: an ergodic channel and several block fading channels.

For the ergodic channel, the single tap channel values vary independently from symbol to symbol. This is the case at high mobile speeds, when the symbol interval is close to the coherence time $t_\mathrm{c} = \frac{1}{f_\mathrm{d}}$ with the Doppler shift $f_\mathrm{d} = \frac{v f_\mathrm{c}}{c} \cos(\Phi)$ where v is the mobile speed, f_c is the carrier frequency, c is the speed of light and Φ is the angle of the incoming signal with respect to the direction of travel of the mobile. While this is not the case in reality for HSDPA, this scenario can be useful for information theoretic investigations such as comparisons with the channel capacity limit. For the ergodic simulations, the 1 tap channel has a relative power gain of $p_\mathrm{erg} = 0$ dB.

For HSDPA scenarios at low mobile speeds, block fading channels are considered. In other words, for each transmitted data block, new channel realisations are generated, which stay constant during the whole coded data block. Each channel realisation consist of one (flat channel) or more taps with a specified power profile (dispersive channel), each corresponding to the incoming reflections of the signal through the rich scattering environment. While the corresponding delays can be arbitrary, oversampling factor of unity and chip-spaced taps are considered. In the following chapters, three different power profiles are used: Firstly, a flat channel with 1 tap and $p_\mathrm{1tap} = 0$ dB relative power gain, a 3 tap channel with equal power $\underline{p}_\mathrm{3tap}$ = [-4.8149; -4.8149; -4.8149] dB, and finally a 10 tap typical urban (TU) channel with the power profile $\underline{p}_\mathrm{TU}$ = [-5.2477; -8.2105; -5.0724; -11.5707; -15.915; -11.3861; -14.6251; -16.1761; -19.5195; -30.8758] dB.

The MIMO channel matrix $\boldsymbol{H}$ consists of $N_\mathrm{R} \times N_\mathrm{T}$ sub-matrices ${}^{(j)}\boldsymbol{H}^{(i)}$ which represent the separate channels between all receiver and transmitter antennas. Each of the sub-matrices has the form of a convolution matrix, corresponding to the channel vector ${}^{(j)}\underline{h}^{(i)}$, whose product with the transmitted symbol vector is equivalent to the convolution of the two vectors. The

channel vectors, containing the channel values weighted according to the power profile for each of the MIMO sub-channels, can be written as

$$^{(j)}\underline{h}^{(i)} = \sqrt{\frac{1}{2}10^{(\underline{p}/10)}} \otimes \underline{\kappa} \tag{2.4}$$

where the symbol $\otimes$ represents element-by-element multiplication, $\underline{p}$ represents the power profile and $\underline{\kappa}$ is the vector containing the channel values κ_w with $w = 1 \ldots W$.

2.3 Space-time channel equalisation

In a CDMA radio communication system, data streams are spread with orthogonal Walsh sequences and transmitted simultaneously. For typical radio channels, several copies of the transmitted signal arrive with different time delays due to reflections, dependent on the environment. Because of this phenomenon, called dispersion, the spreading codes lose orthogonality which leads to interference between the codes, and also to inter-symbol interference (ISI).

Optimum space-time detection would imply joint detection of KN_{T} transmitted symbols per symbol epoch. For 4-QAM modulation, and for dispersive channels with ISI extending over L symbols, this requires a search over a trellis containing $2^{2(L+1)KN_{\mathrm{T}}}$ states. The computational complexity is clearly prohibitive for typical parameter values. Note that, in flat fading conditions ($L = 0$) and for K orthogonal codes re-used over the transmitter antennas, the number of trellis states reduces to a more realistic value of $2^{2N_{\mathrm{T}}}$. As a result, an efficient strategy for dealing with dispersive channels is to perform detection only after a process of space-time equalisation, which effectively eliminates dispersion [7], followed by de-spreading. The equalisation process inevitably causes noise colouring, which needs to be accounted for by a pre-whitening process as shown in Figure 2.2.

The received signal over T symbol epochs is given by $\underline{r} = \boldsymbol{H}\boldsymbol{C}\underline{x} + \underline{n} = \boldsymbol{H}\underline{s} + \underline{n}$ where $\underline{s} = \boldsymbol{C}\underline{x}$ is the vector of spread symbols (see Equation 2.2). A minimum mean-square error (MMSE) equaliser represents a space-time matrix $\boldsymbol{V}$ which minimises the term $E\{\|\underline{s} - \boldsymbol{V}\underline{r}\|^2\}$. It is straightforward to show [6, 8–10] that the solution to this problem is given by

$$\boldsymbol{V} = \boldsymbol{R}_s\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{R}_s\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{R}_n\right)^{-1} \tag{2.5}$$

where $\boldsymbol{R}_s = \mathrm{E}\{\underline{s}\,\underline{s}^{\mathrm{H}}\} = 2\boldsymbol{C}\boldsymbol{C}^{\mathrm{H}}$ since $\mathrm{E}\{\underline{x}\,\underline{x}^{\mathrm{H}}\} = 2\boldsymbol{I}$ for 4-QAM. The derivation of the

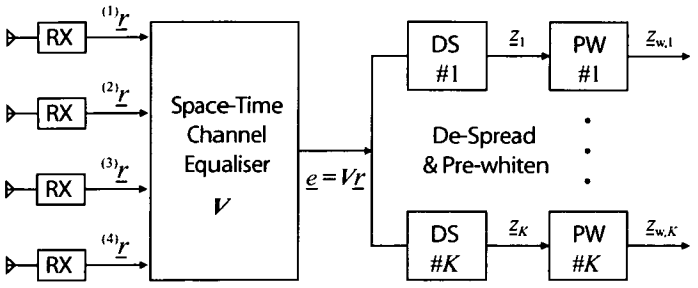


Figure 2.2: *Space-time equalisation followed by de-spreading and pre-whitening for a* 4×4 *antenna MIMO link.*

equaliser matrix $\boldsymbol{V}$ is also shown in Appendix A.1. Then, the equalisation process may be described as

$$\underline{e} = \boldsymbol{V}\underline{r} = \boldsymbol{V}\boldsymbol{H}\sum_{k=1}^{K}\boldsymbol{C}_k\underline{x}_k + \boldsymbol{V}\underline{n} \in \mathrm{C}^{QN_\mathrm{T}T} \tag{2.6}$$

and clearly results in coloured noise. To avoid excessive computational complexity, space-time equalisation is usually performed over a block of $N_\mathrm{E} < T$ symbol epochs and repeated T/N_E times to cover the entire transmission period. However, this reduction in complexity comes at the expense of degraded performance due to inaccuracies at the equaliser block edges.

The structures of the channel matrix $\boldsymbol{H}$ and the equaliser matrix $\boldsymbol{V}$ are shown in Figures 2.3 and 2.4 respectively for 4×4 antennas, TU channel, spreading factor $Q = 16$ and equaliser matrix size $N_\mathrm{E} = 1$ symbol epochs. Both matrices consist of 4×4 sub-matrices ${}^{(j)}\boldsymbol{H}^{(i)}$ and ${}^{(i)}\boldsymbol{V}^{(j)}$, each for the corresponding channel between the transmitter antenna i and receiver antenna j.

2.3.1 De-spreading and pre-whitening

The space-time equaliser nominally removes the influence of the channel matrix $\boldsymbol{H}$. As a result, assuming orthogonal spreading codes, the contribution of symbols transmitted using the k^{th} spreading code can be retrieved at the output of the equaliser via a de-spreading operation.

Even with complete access to channel state information, a space time equaliser can never fully eliminate the influence of the MIMO channel (the zero-forcing equaliser achieves this at the expense of noise enhancement). In other words, $\boldsymbol{V}\boldsymbol{H} = \boldsymbol{D} \neq \boldsymbol{I}$, where $\boldsymbol{D}$ is a non-diagonal distortion matrix. The structure of the distortion matrix resulting from the channel- and equaliser matrix under investigation is shown in Figure 2.5.

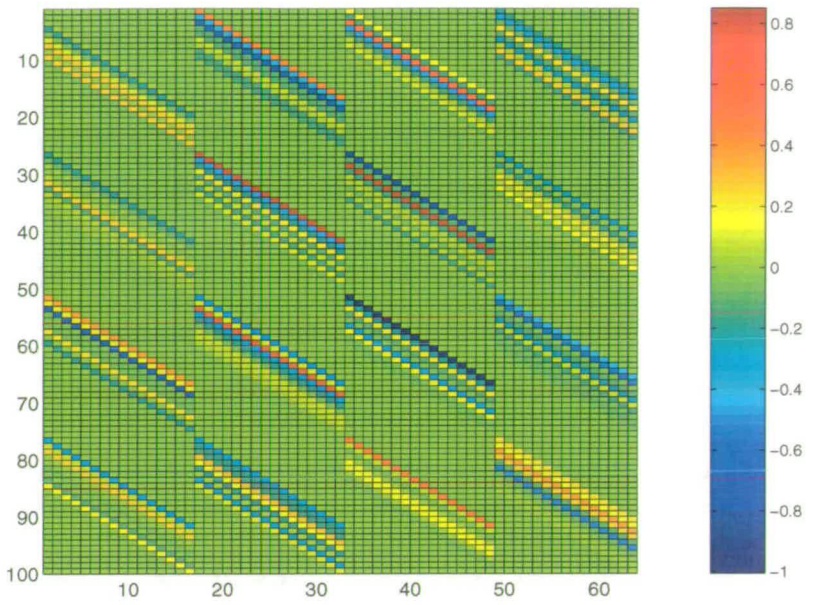


Figure 2.3: *Real part of the channel matrix* $\boldsymbol{H} \in C^{N_{\mathrm{R}}(Q+W-1)\times N_{\mathrm{T}}Q}$.

This has a number of implications with respect to the computation of pre-whitened sufficient statistics for input to the detector, as described next. The output of the equaliser may be written as

$$\underline{e} = \boldsymbol{V}\boldsymbol{H}\sum_{k=1}^{K}\boldsymbol{C}_k\underline{x}_k + \boldsymbol{V}\underline{n} = \boldsymbol{D}\sum_{k=1}^{K}\boldsymbol{C}_k\underline{x}_k + \boldsymbol{V}\underline{n} \tag{2.7}$$

and so the normalised de-spreading operation for the k^{th} spreading code may be interpreted as

$$\begin{aligned}\underline{z}_k &= \|\underline{c}_k\|^{-2}\boldsymbol{C}_k^{\mathrm{H}}\underline{e} \\ &= \|\underline{c}_k\|^{-2}\boldsymbol{C}_k^{\mathrm{H}}\boldsymbol{D}\boldsymbol{C}_k\underline{x}_k + \|\underline{c}_k\|^{-2}\boldsymbol{C}_k^{\mathrm{H}}\boldsymbol{D}\boldsymbol{C}_{\mathrm{I},k}\underline{x}_{\mathrm{I},k} + \|\underline{c}_k\|^{-2}\boldsymbol{C}_k^{\mathrm{H}}\boldsymbol{V}\underline{n} \\ &= \boldsymbol{G}_k\underline{x}_k + \boldsymbol{T}_{\mathrm{I},k}\underline{x}_{\mathrm{I},k} + \boldsymbol{T}_k\underline{n} \\ &= \boldsymbol{G}_k\underline{x}_k + \underline{v}_{\mathrm{I},k} + \underline{v}_k \in \mathrm{C}^{N_{\mathrm{T}}T}\end{aligned} \tag{2.8}$$

where $\boldsymbol{C}_{\mathrm{I},k} \in \mathrm{C}^{QN_{\mathrm{T}}T\times N_{\mathrm{T}}T(K-1)}$ and $\underline{x}_{\mathrm{I},k} \in \mathrm{C}^{N_{\mathrm{T}}T(K-1)}$ are simply equal to the spreading matrix $\boldsymbol{C}$ and symbol vector $\underline{x}$ respectively with the elements associated with the k^{th} spreading code removed. The subscript '$_{\mathrm{I}}$' represents interference. Vector $\underline{z}_k$ consists of the equalised and de-spread contributions of $N_{\mathrm{T}}T$ symbols transmitted via the k^{th} spreading code over a total of T symbol epochs.

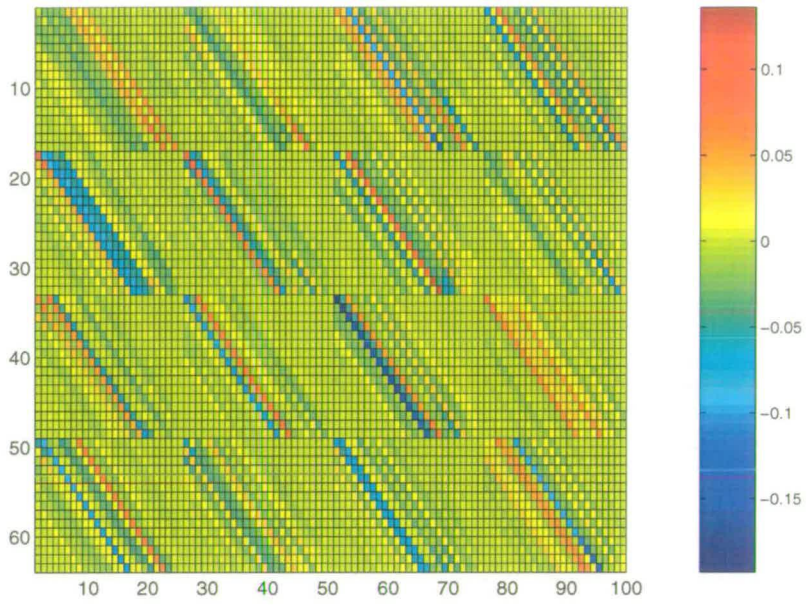


Figure 2.4: *Real part of the equaliser matrix* $\boldsymbol{V} \in C^{N_{\mathrm{T}}Q \times N_{\mathrm{R}}(Q+W-1)}$.

Considering only the N_{T} rows of Equation 2.8 corresponding to the t^{th} symbol epoch, we have for $t = (1 \ldots T)$

$$\begin{aligned}
\underline{z}_k(t) &= \boldsymbol{G}_k(t)\,\underline{x}_k + \boldsymbol{T}_{\mathrm{I},k}(t)\,\underline{x}_{\mathrm{I},k} + \boldsymbol{T}_k(t)\,\underline{n} \\
&= \boldsymbol{B}_k(t)\,\underline{x}_k(t) + \widetilde{\boldsymbol{B}}_k(t)\,\widetilde{\underline{x}}_k(t) + \boldsymbol{T}_{\mathrm{I},k}(t)\,\underline{x}_{\mathrm{I},k}(t) + \widetilde{\boldsymbol{T}}_{\mathrm{I},k}(t)\,\widetilde{\underline{x}}_{\mathrm{I},k}(t) + \boldsymbol{T}_k(t)\,\underline{n} \\
&= \boldsymbol{B}_k(t)\,\underline{x}_k(t) + \widetilde{\underline{s}}_{\mathrm{I},k}(t) + \underline{v}_{\mathrm{I},k}(t) + \widetilde{\underline{v}}_{\mathrm{I},k}(t) + \underline{v}_k(t) \\
&= \boldsymbol{B}_k(t)\,\underline{x}_k(t) + \underline{u}_k(t) \in \mathrm{C}^{N_{\mathrm{T}}}
\end{aligned} \tag{2.9}$$

where $\underline{x}_k(t) \in \mathrm{C}^{N_{\mathrm{T}}}$ is the vector of symbols transmitted during the t^{th} epoch while $\widetilde{\underline{x}}_k(t) \in \mathrm{C}^{N_{\mathrm{T}}(T-1)}$ is the vector of symbols not transmitted during the t^{th} epoch via the k^{th} spreading code. Note that while $\boldsymbol{B}_k(t)$ represents (spatial) self-interference, $\widetilde{\underline{s}}_{\mathrm{I},k}(t)$ identifies space-time interference at the de-spreader output due to symbols transmitted via the k^{th}spreading code but at other symbol epochs. The imperfect operation of the space-time equaliser also implies that in addition to coloured noise, $\underline{v}_k(t)$, a certain amount of coloured interference, $\underline{v}_{\mathrm{I},k}(t)$ and $\widetilde{\underline{v}}_{\mathrm{I},k}(t)$, (originating from other spreading codes) also "leaks" through to the de-spreader

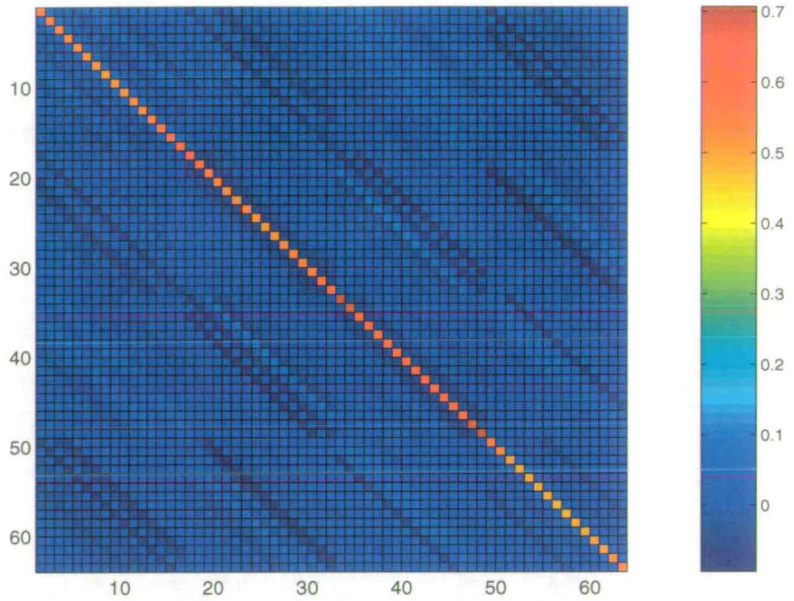


Figure 2.5: *Real part of the distortion matrix* $\boldsymbol{D} \in C^{N_{\mathrm{T}}Q \times N_{\mathrm{T}}Q}$.

output. Assuming that noise and interference are independent, one may write

$$\begin{aligned}\boldsymbol{R}_{\underline{u}_k(t)} &= \mathrm{E}\{\underline{u}_k(t)\,\underline{u}_k^{\mathrm{H}}(t)\} \\ &= 2\{\widetilde{\boldsymbol{B}}_k(t)\,\widetilde{\boldsymbol{B}}_k^{\mathrm{H}}(t) + \boldsymbol{T}_{\mathrm{I},k}(t)\,\boldsymbol{T}_{\mathrm{I},k}^{\mathrm{H}}(t) + \widetilde{\boldsymbol{T}}_{\mathrm{I},k}(t)\,\widetilde{\boldsymbol{T}}_{\mathrm{I},k}^{\mathrm{H}}(t)\} \\ &\quad + N_0\boldsymbol{T}_k(t)\,\boldsymbol{T}_k^{\mathrm{H}}(t) \end{aligned} \tag{2.10}$$

since $\mathrm{E}\{\underline{x}_{\mathrm{I},k}(t)\,\underline{x}_{\mathrm{I},k}^{\mathrm{H}}(t)\} = 2\boldsymbol{I}_{N_{\mathrm{T}}T(K-1)}$ and $\mathrm{E}\{\widetilde{\underline{x}}_k(t)\,\widetilde{\underline{x}}_k^{\mathrm{H}}(t)\} = 2\boldsymbol{I}_{N_{\mathrm{T}}(K-1)}$.

- **Pre-whitening w.r.t. Interference + Noise**

Pre-whitened sufficient statistics for the detector input can then be computed as

$$\begin{aligned}\underline{z}_{\mathrm{w},k}(t) &= \boldsymbol{R}_{\underline{u}_k(t)}^{-\frac{1}{2}}\underline{z}_k(t) \\ &= \boldsymbol{R}_{\underline{u}_k(t)}^{-\frac{1}{2}}\boldsymbol{B}_k(t)\,\underline{x}_k(t) + \boldsymbol{R}_{\underline{u}_k(t)}^{-\frac{1}{2}}\underline{u}_k(t) \\ &= \boldsymbol{R}_{\underline{u}_k(t)}^{-\frac{1}{2}}\boldsymbol{B}_k(t)\,\underline{x}_k(t) + \underline{\varepsilon}_k(t) \in \mathrm{C}^{N_{\mathrm{T}}} \end{aligned} \tag{2.11}$$

where $\mathrm{E}\{\underline{\varepsilon}_k\underline{\varepsilon}_k^{\mathrm{H}}\} = \boldsymbol{I}_{N_{\mathrm{T}}}$.

2.3.2 Transversal equalisation

The channel matrix $\boldsymbol{H}$ consists of $N_{\mathrm{R}} \times N_{\mathrm{T}}$ sub-matrices, each of the form of a convolution matrix with the coefficients of the corresponding channel from transmitter antenna i to receiver antenna j. The property that the MMSE equaliser matrix $\boldsymbol{V}$ also consists of convolution matrix type sub-matrices, which perform a filter operation in order to equalise each of the channels, can be exploited to implement the equaliser with transversal filters, as shown in Figure 2.6.

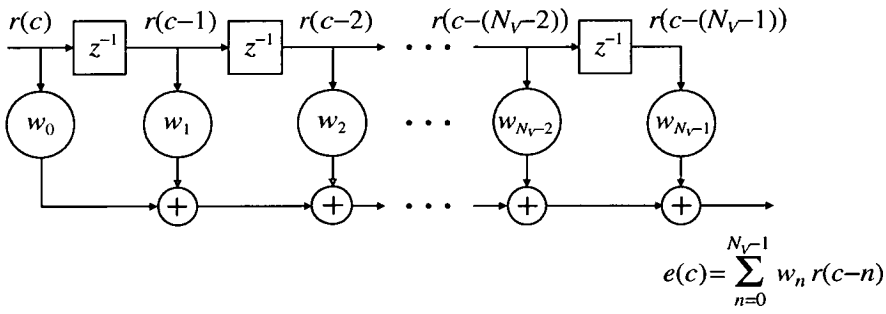


Figure 2.6: *Transversal filter.*

The weight coefficients ${}^{(i)}w^{(j)}$ for each of the channels can be derived from the block equaliser sub-matrices ${}^{(i)}\boldsymbol{V}^{(j)}$. As shown in Figure 2.7, for an equaliser with $N_{\mathrm{V}} = 16$ taps, the coefficients ${}^{(1)}w^{(1)}$ can be obtained for by selecting the $(Q + W - 1)/2$ $^{\mathrm{th}}$ column of the corresponding equaliser sub-matrix [6], where Q denotes the spreading factor and W the channel length. The example of ${}^{(1)}\boldsymbol{V}^{(1)}$ in Figure 2.7 shows that the strongest elements of ${}^{(1)}w^{(1)}$ are located in the middle. With increasing distance from the diagonal of the sub-matrix, the coefficients of ${}^{(i)}w^{(j)}$ become smaller, and approach zero for a sufficient number of equaliser taps. Using this method, the maximum number of tap coefficients obtainable is $N_{\mathrm{E}}Q$. However, since the calculation of $\boldsymbol{V}$ includes a matrix inversion, increasing N_{E} is undesirable due to the high increase in computational complexity.

For the transversal equaliser, the equalised signal from each transmitter antenna can be written as

$$\underline{e}^{(i)} = \sum_{j=1}^{N_{\mathrm{R}}} \mathrm{conv}\{{}^{(j)}\underline{r},\ {}^{(i)}\underline{w}^{(j)}\}. \tag{2.12}$$

This operation is equivalent to the block equalisation in Equation 2.7 for a block size over all T symbol epochs, assuming the number of taps of the filter are sufficiently large, that the coefficients in upper right and lower left triangle of the matrix ${}^{(i)}\boldsymbol{V}^{(j)}$ which are not covered

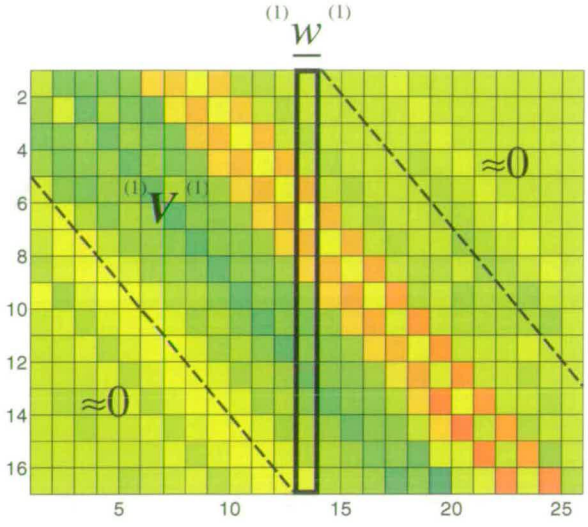


Figure 2.7: *Selection of 16 filter coefficients from the equaliser sub-matrix for a matrix size over $N_{\mathrm{E}} = 1$ symbol interval.*

by the transversal equaliser approach zero. For the calculation of the pre-whitening matrix, the equaliser matrix $\boldsymbol{V}$ is modified to match exactly the transversal filter operation. Then, the de-spreading and pre-whitening operation can be performed as for the block-based equalisation in Equations 2.8 – 2.11.

2.3.3 Approximate modelling of the equaliser output

Since the equaliser effectively eliminates the channel dispersion, the remaining ISI which leaks from each symbol to the next, is relatively small in comparison to the remaining distortion after equalisation. Therefore, the contribution from other symbols to the sufficient statistics may be neglected and the N_{T} rows of Equation 2.9 corresponding to the t^{th} symbol epoch can be written as

$$\begin{aligned} \underline{z}_k(t) &\approx \boldsymbol{B}_k(t)\,\underline{x}_k(t) + \boldsymbol{T}_{\mathrm{I},k}(t)\,\underline{x}_{\mathrm{I},k}(t) + \boldsymbol{T}_k(t)\,\underline{n} \\ &= \boldsymbol{B}_k(t)\,\underline{x}_k(t) + \underline{v}_{\mathrm{I},k}(t) + \underline{v}_k(t) \\ &= \boldsymbol{B}_k(t)\,\underline{x}_k(t) + \underline{\nu}_k(t) \end{aligned} \tag{2.13}$$

where $\underline{\nu}_k$ contains both the remaining interference from the other spreading codes $\underline{v}_{\mathrm{I},k}(t)$ and the coloured noise $\underline{v}_k(t)$. The resulting correlation of interference and noise is

$$\begin{aligned}\boldsymbol{R}_{\underline{\nu}_k(t)} &= \mathrm{E}\{\underline{\nu}_k(t)\,\underline{\nu}_k^{\mathrm{H}}(t)\} \\ &= \mathrm{E}\{\underline{v}_{\mathrm{I},k}(t)\,\underline{v}_{\mathrm{I},k}^{\mathrm{H}}(t)\} + \mathrm{E}\{\underline{v}_k(t)\,\underline{v}_k^{\mathrm{H}}(t)\} \\ &= 2\boldsymbol{T}_{\mathrm{I},k}(t)\,\boldsymbol{T}_{\mathrm{I},k}^{\mathrm{H}}(t) + N_0\boldsymbol{T}_k(t)\,\boldsymbol{T}_k^{\mathrm{H}}(t)\,. \end{aligned} \tag{2.14}$$

Then, the sufficient statistics for the detector input can be pre-whitened with respect to the interference from symbol t and noise:

$$\begin{aligned}\underline{z}_{\mathrm{w},k}(t) &= \boldsymbol{R}_{\underline{\nu}_k(t)}^{-\frac{1}{2}}\underline{z}_k(t) \\ &= \boldsymbol{R}_{\underline{\nu}_k(t)}^{-\frac{1}{2}}\boldsymbol{B}_k(t)\,\underline{x}_k(t) + \boldsymbol{R}_{\underline{\nu}_k(t)}^{-\frac{1}{2}}\underline{\nu}_k(t) \\ &= \boldsymbol{A}_k(t)\,\underline{x}_k(t) + \underline{\varepsilon}_k(t) \in \mathrm{C}^{N_{\mathrm{T}}} \end{aligned} \tag{2.15}$$

where $\mathrm{E}\{\underline{\varepsilon}_k\underline{\varepsilon}_k^{\mathrm{H}}\} = \boldsymbol{I}_{N_{\mathrm{T}}}$.

2.3.4 LLR calculation after de-spreading

The equaliser can also be employed without subsequent detection. In this case, soft outputs for decoding, in the form of log-likelihood ratios (LLRs), are computed directly after de-spreading and pre-whitening based on the remaining signal correlations $\boldsymbol{R}_{\underline{v}_k}$ under the assumption that the noise term $\underline{\varepsilon}_k(t)$ is Gaussian. Then, the LLRs can be computed as shown in Appendix A.3 for unnormalised matched filter outputs:

$$\Lambda(\underline{b}_0) = 4\,\mathrm{diag}\{\boldsymbol{A}_k(t)\}\,\mathrm{Re}\{\underline{z}_{w,k}(t)\} \tag{2.16}$$

and

$$\Lambda(\underline{b}_1) = 4\,\mathrm{diag}\{\boldsymbol{A}_k(t)\}\,\mathrm{Im}\{\underline{z}_{w,k}(t)\} \tag{2.17}$$

where $\Lambda(\underline{b}_0)$ represent the LLRs corresponding to the bits transmitted in the real part, and $\Lambda(\underline{b}_1)$ represent the LLRs corresponding to the bits transmitted in the imaginary part of the symbols $\underline{x}_k(t)$.

2.3.5 Equaliser performance comparison

The equaliser performance is investigated for both block based (Equation 2.6) and transversal (Equation 2.12) equalisation, dependent on the block size or number of taps. The bits are estimated via a sgn (.) operation after performing space-time equalisation, de-spreading and pre-whitening as shown in Figure 2.8. A MIMO link with $N_T = N_R = 4$ and 4-QAM modulation is investigated in two different channel scenarios. The data is spread with a spreading factor of $Q = 16$ and the same set of $K = 16$ orthogonal Walsh spreading codes are simultaneously transmitted from each antenna resulting in a high-interference code re-use scenario. No error correction coding is employed and the channel conditions are assumed to be known at the equaliser. Figure 2.9 shows the performance comparison for a 3-tap dispersive channel with chip-spaced equal power taps (Scenario 1). The results in Figure 2.10 consider a typically urban (TU) channel with $W = 10$ chip-spaced taps (Scenario 2). It is shown that with increasing channel length, the edge effects significantly impact the performance of the block based equalisers. These effects become smaller with increasing block size. The transversal equalisers perform significantly better, since they do not suffer from edge effects. An increased number of filter coefficients in ${}^{(i)}\underline{w}^{(j)}$ further improves the performance for transversal equalisers, especially for channels with long delay. For all following simulations, which employ channel equalisation, the 32-tap transversal equaliser is used as described above. For a real implementation, the 16-tap equaliser is more attractive due to the lower complexity and still good performance in the raw BER range of interest for coded radio systems in the range of 10^{-1} or above.

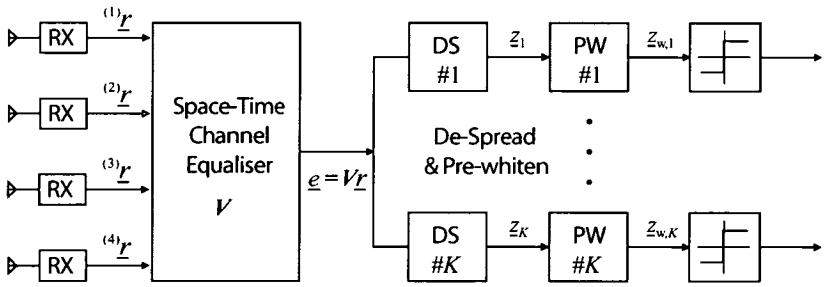


Figure 2.8: *BER calculation for space-time equalisation for a* 4×4 *antenna MIMO link.*

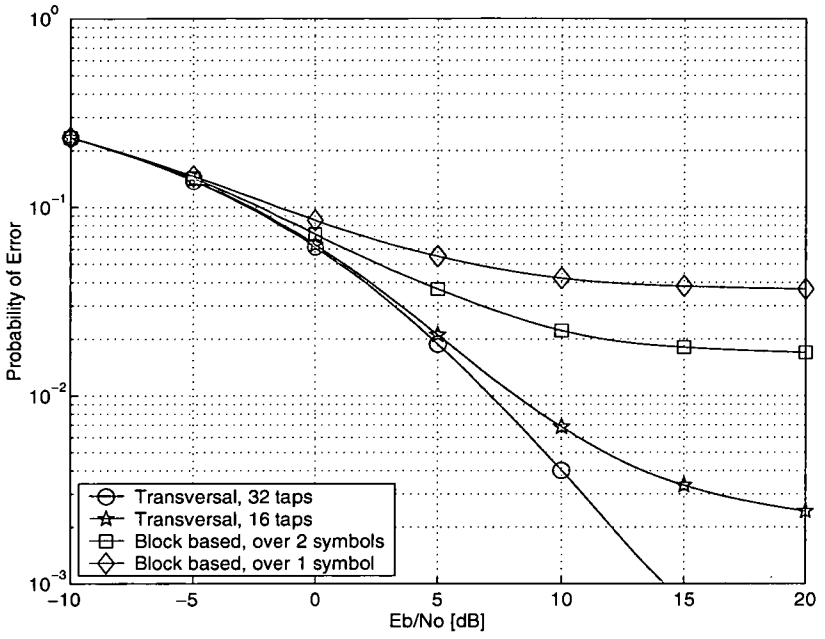


Figure 2.9: *Equaliser performance for scenario 1: 3-tap channel (chip-spaced),* 4×4 *antennas, 4-QAM modulation,* $Q = 16$*,* $K = 16$*.*

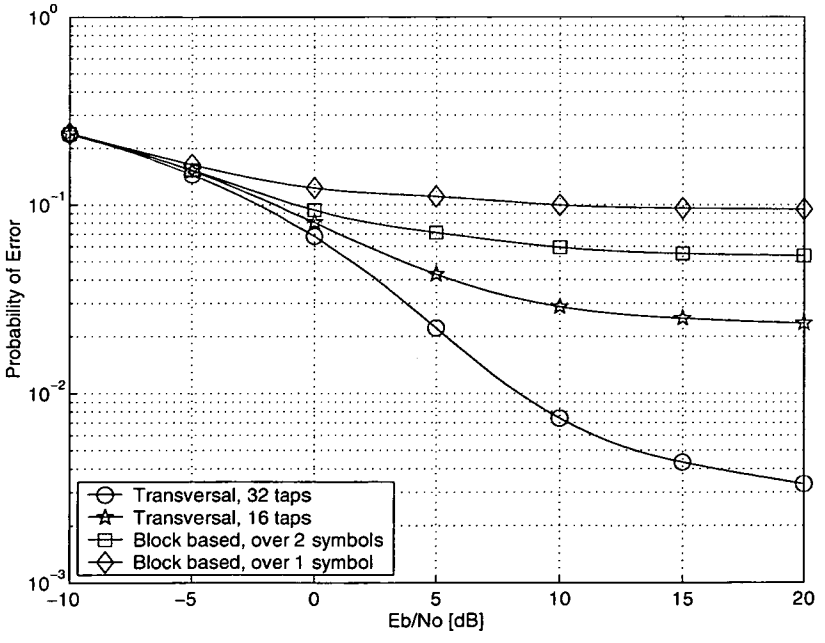


Figure 2.10: *Equaliser performance for scenario 2: TU channel (10 chip-spaced taps),* 4×4 *antennas, 4-QAM modulation,* $Q = 16$*,* $K = 16$*.*

2.3.6 Computational complexity of transversal equalisation

For the complexity calculation, the equalisation process is considered without initialisation of the equaliser weights. The transversal filter operation requires one complex multiplication and one complex addition per filter tap per equalised chip for each transmitter and receiver antenna. Therefore the required operations per symbol period can be calculated as $N_{\mathrm{CMULT}} = N_{\mathrm{T}} N_{\mathrm{R}} N_V Q$ and $N_{\mathrm{CADD}} = N_{\mathrm{T}} N_{\mathrm{R}} N_V Q$ where N_V is the number equaliser taps and Q is the spreading factor. In terms of real multiplications and real additions, the complexity is therefore $N_{\mathrm{MULT}} = 4 N_{\mathrm{T}} N_{\mathrm{R}} N_V Q$ and $N_{\mathrm{ADD}} = 4 N_{\mathrm{T}} N_{\mathrm{R}} N_V Q$. For the example of a 4×4 antenna MIMO system with $Q = 16$ a space-time receiver with 32 taps, this results in $N_{\mathrm{MULT}} = N_{\mathrm{ADD}} = 32768$. For a 16 tap equaliser the complexity reduces to $N_{\mathrm{MULT}} = N_{\mathrm{ADD}} = 16384$.

2.4 MIMO detection algorithms

Even when space-time equalisation followed by de-spreading and pre-whitening is employed as discussed previously, a certain amount of co-channel interference is still remaining. This interference can be taken into account in the LLR calculation using a detection unit after the equalisation process. In this section the optimum APP detection and its Max-Log approximation are presented as reference detectors. Then, two detectors based on interference cancellation are proposed as low complexity alternatives, followed by a complexity and performance comparison.

All detectors operate on pre-whitened sufficient statistics for each symbol interval t and spreading code k of the form (see Equation 2.15)

$$\underline{z}_w = \boldsymbol{A}\underline{x} + \underline{\epsilon} \tag{2.18}$$

where the indices for the symbol t and the code k are omitted for simplicity. $\underline{x} \in C^{N_{\mathrm{T}}}$ denotes the vector of transmitted symbols and $\boldsymbol{A} \in C^{N_{\mathrm{T}} \times N_{\mathrm{T}}}$ is the transformation matrix. It is assumed that the elements of the additive disturbance vector are i.i.d. zero-mean complex Gaussian random variables of unit variance (i.e. $\mathrm{E}\{\underline{\epsilon}\underline{\epsilon}^{\mathrm{H}}\} = \sigma_\epsilon^2 \boldsymbol{I}$).

2.4.1 Optimum MIMO detection

Considering the sufficient statistics from Equation 2.18, the likelihood function or conditional probability density of $\underline{z}_w$ may be written as

$$\begin{aligned}
\mathrm{f}\left(\underline{z}_w|\underline{x}\right) &= \prod_{i=1}^{N_\mathrm{T}} \mathrm{f}\left(\underline{z}_w^{(i)}|\underline{x}\right) \\
&= \prod_{i=1}^{N_\mathrm{T}} \frac{1}{\pi\sigma_\epsilon^2} \exp\left\{\frac{-1}{\sigma_\epsilon^2}\left|\underline{z}_w^{(i)} - \boldsymbol{A}\underline{x}^{(i)}\right|^2\right\} \\
&= \frac{1}{\left(\pi\sigma_\epsilon^2\right)^{N_\mathrm{T}}} \exp\left\{\frac{-1}{\sigma_\epsilon^2}\sum_{i=1}^{N_\mathrm{T}}\left|\underline{z}_w^{(i)} - \boldsymbol{A}\underline{x}^{(i)}\right|^2\right\} \\
&= \frac{1}{\left(\pi\sigma_\epsilon^2\right)^{N_\mathrm{T}}} \exp\left\{\frac{-1}{\sigma_\epsilon^2}\left\|\underline{z}_w - \boldsymbol{A}\underline{x}\right\|^2\right\}. \qquad (2.19)
\end{aligned}$$

With the availability of sufficient statistics $\underline{z}_w$, a detector is in a position to make a hypothesis $\underline{x}_0$ regarding the transmitted symbols. The probability that this hypothesis is correct is equal to the probability, $\mathrm{P}\{\underline{x}_0|\underline{z}_w\}$, that $\underline{x}_0$ was indeed transmitted given $\underline{z}_w$. The maximum *a posteriori* probability (MAP) detector is defined as that which minimises the probability of an incorrect hypothesis:

$$\begin{aligned}
\hat{\underline{x}}_{\mathrm{MAP}} &= \arg\max_{\underline{x}} \mathrm{P}\{\underline{x}|\underline{z}_w\} \\
&= \arg\max_{\underline{x}} \frac{\mathrm{P}\{\underline{x}, \underline{z}_w\}}{\mathrm{f}\left(\underline{z}_w\right) d\underline{z}_w} \\
&= \arg\max_{\underline{x}} \frac{\mathrm{f}\left(\underline{z}_w|\underline{x}\right) d\underline{z}_w \, \mathrm{P}\{\underline{x}\}}{\mathrm{f}\left(\underline{z}_w\right) d\underline{z}_w} \\
&= \arg\max_{\underline{x}} \left(\mathrm{f}\left(\underline{z}_w|\underline{x}\right) \, \mathrm{P}\{\underline{x}\}\right) \\
&= \arg\max_{\underline{x}} \frac{\mathrm{P}\{\underline{x}\}}{\left(\pi\sigma_\epsilon^2\right)^{N_\mathrm{T}}} \exp\left\{\frac{-1}{\sigma_\epsilon^2}\left\|\underline{z}_w - \boldsymbol{A}\underline{x}\right\|^2\right\} \\
&= \arg\max_{\underline{x}} \left\{\ln\left(\mathrm{P}\left\{\underline{x}\right\}\right) - \frac{1}{\sigma_\epsilon^2}\left\|\underline{z}_w - \boldsymbol{A}\underline{x}\right\|^2\right\} \\
&= \arg\min_{\underline{x}} \left\{\frac{1}{\sigma_\epsilon^2}\left\|\underline{z}_w - \boldsymbol{A}\underline{x}\right\|^2 - \ln\left(\mathrm{P}\left\{\underline{x}\right\}\right)\right\} \qquad (2.20)
\end{aligned}$$

where $\mathrm{P}\{\underline{x}\}$ is the *a priori* probability of $\underline{x}$ which can be supplied by a decoder in an iterative receiver architecture as shown in Chapter 4. The *a priori* information in form of LLRs can be

written as [11]

$$\begin{aligned}\Lambda_a\left(b_q\right) &= \ln\frac{\mathrm{P}\left\{b_q=+1\right\}}{\mathrm{P}\left\{b_q=-1\right\}}\\ &= \ln\frac{\mathrm{P}\left\{b_q=+1\right\}}{1-\mathrm{P}\left\{b_q=+1\right\}}\end{aligned} \tag{2.21}$$

where $\Lambda_a\left(b_q\right)$ is the LLR value corresponding to the q^{th} bit of the symbol vector $\underline{x}$. From this follows that the probabilities are

$$\begin{aligned}\mathrm{P}\left\{b_q=+1\right\} &= \left(1-\mathrm{P}\left\{b_q=+1\right\}\right)e^{\Lambda_a(b_q)}\\ &= \frac{e^{\Lambda_a(b_q)}}{e^{\Lambda_a(b_q)}+1}\\ &= \frac{e^{\Lambda_a(b_q)/2}}{e^{\Lambda_a(b_q)/2}+e^{-\Lambda_a(b_q)/2}} \qquad (2.22)\\ \mathrm{P}\left\{b_q=-1\right\} &= \frac{e^{-\Lambda_a(b_q)/2}}{e^{\Lambda_a(b_q)/2}+e^{-\Lambda_a(b_q)/2}}. \qquad (2.23)\end{aligned}$$

Hence, one can rewrite the probabilities for an arbitrary bit $b_q \in \{+1,-1\}$ and further simplify the equation using logarithmic probabilities:

$$\mathrm{P}\left\{b_q\right\} = \frac{e^{b_q\Lambda_a(b_q)/2}}{e^{\Lambda_a(b_q)/2}+e^{-\Lambda_a(b_q)/2}} \tag{2.24}$$

$$\ln\mathrm{P}\left\{b_q\right\} = \frac{1}{2}b_q\Lambda_a\left(b_q\right). \tag{2.25}$$

The logarithmic probability of the symbol vector $\underline{x}$ can then be calculated as sum of the probabilities $\ln\mathrm{P}\left\{b_q\right\}$ corresponding to the symbol vector:

$$\ln\mathrm{P}\{\underline{x}\} = \frac{1}{2}\underline{b}^{\mathrm{T}}\underline{\Lambda}_a\left(\underline{b}\right) \tag{2.26}$$

where $\underline{\Lambda}_a\left(\underline{b}\right)$ is the *a priori* LLR values of the bit-vector $\underline{b}$. In the absence of such *a priori* information (i.e. $\mathrm{P}\{\underline{x}\} = 0$), the MAP detector degenerates into the maximum likelihood (ML) detector.

Soft outputs for the q^{th} bit of the symbol vector $\underline{x}$ may be derived in the form of log-likelihood

ratios (LLR) at the output of the MAP detector as

$$\begin{aligned}
\Lambda(b_q) &= \ln \frac{\mathrm{P}\{b_q = +1|\underline{z}_w\}}{\mathrm{P}\{b_q = -1|\underline{z}_w\}} \\
&= \ln \frac{\sum_{\underline{x}|b_q=+1} \mathrm{P}\{\underline{x}|\underline{z}_w\}}{\sum_{\underline{x}|b_q=-1} \mathrm{P}\{\underline{x}|\underline{z}_w\}} \\
&= \ln \frac{\sum_{\underline{x}|b_q=+1} \mathrm{f}(\underline{z}_w|\underline{x})\,\mathrm{P}\{\underline{x}\}}{\sum_{\underline{x}|b_q=-1} \mathrm{f}(\underline{z}_w|\underline{x})\,\mathrm{P}\{\underline{x}\}} \\
&= \ln \frac{\sum_{\underline{x}|b_q=+1} \pi^{-N_\mathrm{T}} \exp\left\{\frac{-1}{\sigma_\epsilon^2}\|\underline{z}_w - \boldsymbol{A}\underline{x}\|^2\right\}\mathrm{P}\{\underline{x}\}}{\sum_{\underline{x}|b_q=-1} \pi^{-N_\mathrm{T}} \exp\left\{\frac{-1}{\sigma_\epsilon^2}\|\underline{z}_w - \boldsymbol{A}\underline{x}\|^2\right\}\mathrm{P}\{\underline{x}\}} \\
&= \ln \frac{\sum_{\underline{x}|b_q=+1} \exp\left\{\frac{-1}{\sigma_\epsilon^2}\|\underline{z}_w - \boldsymbol{A}\underline{x}\|^2 + \ln \mathrm{P}\{\underline{x}\}\right\}}{\sum_{\underline{x}|b_q=-1} \exp\left\{\frac{-1}{\sigma_\epsilon^2}\|\underline{z}_w - \boldsymbol{A}\underline{x}\|^2 + \ln \mathrm{P}\{\underline{x}\}\right\}}.
\end{aligned} \tag{2.27}$$

Assuming equi-probable symbols, the LLR computation can be simplified to

$$\Lambda(b_q) = \ln \frac{\sum_{\underline{x}|b_q=+1} \exp\left\{\frac{-1}{\sigma_\epsilon^2}\|\underline{z}_w - \boldsymbol{A}\underline{x}\|^2\right\}}{\sum_{\underline{x}|b_q=-1} \exp\left\{\frac{-1}{\sigma_\epsilon^2}\|\underline{z}_w - \boldsymbol{A}\underline{x}\|^2\right\}}. \tag{2.28}$$

Equation 2.28 represents what is commonly known as the *a posteriori* probability (APP) detector. Comparison of Equations 2.20 and 2.27 indicate that the signs of the above LLR values are equivalent to minimum probability of error (MAP) bit estimates.

The expression for the LLR is not computationally friendly and involves divisions, logarithms and exponentials. The sum of the exponential terms can be avoided using the Jacobian algorithm [12]: $\ln\left(e^{\delta_1} + e^{\delta_2}\right) = \max(\delta_1, \delta_2) + \ln\left(1 + e^{-|\delta_1 - \delta_2|}\right)$. To reduce the computational complexity, the correction function $\ln\left(1 + e^{-|\delta_1 - \delta_2|}\right)$ can be implemented as a lookup table [13].

The computation of the LLR can be further simplified by exploiting the Max-Log approximation $\ln\left(e^{\delta_1} + e^{\delta_2} + \ldots + e^{\delta_n}\right) \approx \max\left(e^{\delta_1}, e^{\delta_2}, \ldots, e^{\delta_n}\right)$ where the correction function is omitted [5, 13]. Then, the LLR calculation of the Max-Log-APP detector may be written as

$$\begin{aligned}
\Lambda(b_q) &\approx \max_{\underline{x}|b_q=+1} \left\{ -\|\underline{z}_w - \boldsymbol{A}\underline{x}\|^2 + \ln \mathrm{P}\{\underline{x}\} \right\} - \max_{\underline{x}|b_q=-1} \left\{ -\|\underline{z}_w - \boldsymbol{A}\underline{x}\|^2 + \ln \mathrm{P}\{\underline{x}\} \right\} \\
&= \min_{\underline{x}|b_q=-1} \left\{ \|\underline{z}_w - \boldsymbol{A}\underline{x}\|^2 - \ln \mathrm{P}\{\underline{x}\} \right\} - \min_{\underline{x}|b_q=+1} \left\{ \|\underline{z}_w - \boldsymbol{A}\underline{x}\|^2 - \ln \mathrm{P}\{\underline{x}\} \right\}.
\end{aligned} \tag{2.29}$$

Again, assuming equi-probable symbols, the equation simplifies to

$$\Lambda(b_q) \approx \min_{\underline{x}|b_q=-1} \{ \|\underline{z}_w - \boldsymbol{A}\underline{x}\|^2 \} - \min_{\underline{x}|b_q=+1} \{ \|\underline{z}_w - \boldsymbol{A}\underline{x}\|^2 \}. \tag{2.30}$$

Note that, when using the the Max-Log approximation the LLR computation is suboptimal. While the approximation error is insignificant when operating at high ranges of E_b/N_0, at the low E_b/N_0 ranges of interest this cannot be ignored.

2.4.2 Multi-stage partial parallel interference cancellation

Various forms of parallel interference cancellation have been considered in the past in the context of multi-user detection for the CDMA uplink [14, 15]. Here, the multi-stage partial parallel interference canceller (MS-PPIC) is considered as an alternative to APP detection in the context of the CDMA MIMO downlink. Unlike APP, which is a single-shot joint detection process, the MS-PPIC involves multiple stages of "non-linear" cancellation, where at each stage the contributions due to interfering antennas are removed from the sufficient statistics at the detector input, thereby enhancing the detection process. Antenna interference contributions at the n^{th} stage of cancellation are constructed from "soft symbols" derived in the previous $(n-1)^{\text{th}}$ stage, as well as from those derived most recently in the current n^{th} stage. Log-likelihood ratios are finally computed after the last stage where, subsequent to multiple stages of cancellation, additive Gaussian noise is the only remaining source of disturbance.

Performing matched filtering on the sufficient statistics of Equation 2.18 and normalising, we have

$$\underline{y} = \Delta^{-1}\boldsymbol{A}^{\mathrm{H}}\underline{z}_w = \Delta^{-1}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\underline{x} + \Delta^{-1}\boldsymbol{A}^{\mathrm{H}}\underline{\epsilon} = \Delta^{-1}\boldsymbol{R}\underline{x} + \underline{\eta} \tag{2.31}$$

where $\boldsymbol{R} = \boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$, $\Delta = \mathrm{diag}\{\boldsymbol{R}\}$ and $\mathrm{E}\{\underline{\eta}\underline{\eta}^{\mathrm{H}}\} = \sigma_\epsilon^2\Delta^{-1}\boldsymbol{R}\Delta^{-\mathrm{H}}$. One may re-write this in the form

$$\underline{y} = \underline{x} + \Delta^{-1}(\boldsymbol{R} - \Delta)\underline{x} + \underline{\eta} = \underline{x} + \Delta^{-1}\boldsymbol{R}'\underline{x} + \underline{\eta} = \underline{x} + \boldsymbol{S}\underline{x} + \underline{\eta} \tag{2.32}$$

where, given that $\boldsymbol{R}'$ and so $\boldsymbol{S}$ have zero diagonal elements, it is clear that the term $\boldsymbol{S}\underline{x}$ represents interference contributions which need to be cancelled. The sufficient statistics of 2.32 are applied to the MS-PPIC and may be viewed as the 0^{th} stage output, $\underline{y}[0]$, of the detector. The cancellation architecture is illustrated in Figure 2.11.

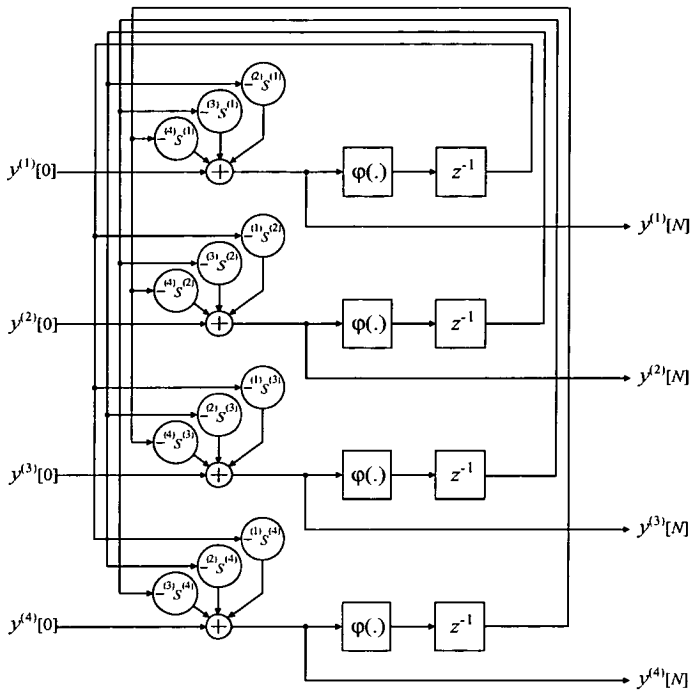


Figure 2.11: *MS-PPIC Architecture.*

Consequently, writing $\underline{y} = \underline{y}[0]$, N stages of parallel cancellation may be described as

for $n = 1 \ldots N_{\mathrm{I}}$ stages

$$\underline{\xi} = \underline{y}[n-1] \in C^{N_{\mathrm{T}} \times 1} \tag{2.33}$$

for $i = 1 \ldots N_{\mathrm{T}}$ antennas

$$y^{(i)}[n] = y^{(i)}[0] - \underline{s}^{(i)} \left\{ \tanh\left\{ \boldsymbol{G}\, \mathrm{Re}\left(\underline{\xi}\right)\right\} + j \tanh\left\{ \boldsymbol{G}\, \mathrm{Im}\left(\underline{\xi}\right)\right\}\right\} \tag{2.34}$$

$$\xi^{(i)} = y^{(i)}[n] \tag{2.35}$$

end

end

where $y^{(i)}[n]$ is the i^{th} element of $\underline{y}[n]$ and n is the stage index. $\underline{s}^{(i)}$ and ${}^{(j)}s^{(i)}$ represent the i^{th} row and $(n,n)^{\mathrm{th}}$ element of $\boldsymbol{S}$ respectively. Essentially, at each stage the contributions due to other antennas are removed from the elements $\underline{y}[0]$. The contributions at the n^{th} stage are constructed from "soft symbols" derived in the previous $(n-1)^{\mathrm{th}}$ stage as well as those derived most recently in the current stage.

Note that $\boldsymbol{G}$ is a diagonal matrix of reliability weights which define the "softness" of the can-

cellation process which can be optimised as shown in Appendix A.2 resulting in

$$\boldsymbol{G} = 2\left\{\text{diag}\left\{2\boldsymbol{S}\boldsymbol{S}^{\text{H}}\right\} + \Delta^{-1}\right\}^{-1}. \tag{2.36}$$

Log-likelihood ratios are computed after the last stage where, due to multiple stages of cancellation, $y^{(i)}[N] \approx x^{(i)} + \eta^{(i)}$. Then, the LLRs can be computed as shown in Appendix A.3 for normalised matched filter outputs:

$$\Lambda(\underline{b}_0) = \frac{4}{\sigma_\epsilon^2}\left[\text{diag}\left(\boldsymbol{R}^{-1}\right)\right]^{-1}\text{Re}\left\{\underline{y}[N]\right\} \tag{2.37}$$

and

$$\Lambda(\underline{b}_1) = \frac{4}{\sigma_\epsilon^2}\left[\text{diag}\left(\boldsymbol{R}^{-1}\right)\right]^{-1}\text{Im}\left\{\underline{y}[N]\right\} \tag{2.38}$$

since $\text{diag}\left(\sigma_\epsilon^2\Delta^{-1}\boldsymbol{R}\Delta^{-\text{H}}\right) = \sigma_\epsilon^2\text{diag}\left(\Delta^{-1}\right) = \sigma_\epsilon^2\text{diag}\left(\boldsymbol{R}^{-1}\right)$. $\Lambda(\underline{b}_0)$ represent the LLRs corresponding to the bits transmitted in the real part, and $\Lambda(\underline{b}_1)$ represent the LLRs corresponding to the bits transmitted in the imaginary part of the symbol vector $\underline{x}$. Note that, when space-time equalisation followed by de-spreading and pre-whitening is applied, $\sigma_\epsilon^2 = 1$.

2.4.3 Serial interference cancellation

Serial interference cancellation (SIC) schemes have been considered for many years in the context of multi-user detection for the CDMA uplink [15–17]. These schemes combat interference by successively detecting and cancelling the influence of data streams from the received signal. The best results are achieved, when the more reliable data streams are detected and cancelled first. In the context of MIMO receivers, the original BLAST detector [2] is essentially a SIC architecture incorporating ordering and detection based on the minimum mean-squared error (MMSE) criterion. Furthermore, significant performance improvements have been demonstrated through iterations between the BLAST detector and a convolutional decoder [18]. Here, ordered matched filter based serial interference cancellation is proposed as detection scheme for MIMO receivers with very low complexity. When no *a priori* information is available, the MF-SIC detector performance cannot compete with the APP and MS-PPIC algorithms above, due to the simple matched filter based detection and hard cancellation which can lead to error propagation. Therefore, the MF-SIC is intended as an ultra low complexity detector for iterative receiver architectures, where it can achieve impressive performance.

When no *a priori* information is available at bit-level, the MF-SIC operates at a symbol level. The first step is to determine the most reliable symbol according to a reliability criterion, at each symbol interval t. Ideally, the symbol with the lowest error probability is selected [17]. Lacking such information, the symbol $x^{(i)}(t) \quad n = 1 \ldots N_{\mathrm{T}}$ with the highest signature energy, $\left\|\underline{a}^{(i)}(t)\right\|^2$ (or least mean-square estimation error), is selected. The next step is to estimate the selected symbol (soft-output derived via matched filter detection), make a hard decision on the estimate, and finally reconstruct and cancel its contribution from the received signal. Performing matched filtering on the sufficient statistics from Equation 2.18 for the symbol of interest, we have

$$y^{(i)} = \underline{a}^{(i)\mathrm{H}} \underline{z}_w = \underline{a}^{(i)\mathrm{H}} \underline{a}^{(i)} x^{(i)} + v^{(i)} + \eta^{(i)} \tag{2.39}$$

where $\underline{a}^{(i)}$ is the i^{th} column of matrix $\boldsymbol{A}$ representing the signature of the symbol transmitted from the i^{th} antenna of interest. The interfering contributions from other antennas are denoted by $v^{(i)}$, and $\mathrm{E}\{\eta^{(i)}\eta^{(i)^*}\} = \sigma^2_{\eta^{(i)}}$. Then, the contributions of the detected symbol are cancelled from the received signal (hard-cancellation):

$$\underline{z}_w := \underline{z}_w - \underline{a}^{(i)} \left\{ \mathrm{sgn}\left[\mathrm{Re}\left(y^{(i)}\right)\right] + j\, \mathrm{sgn}\left[\mathrm{Im}\left(y^{(i)}\right)\right] \right\}. \tag{2.40}$$

The process is then repeated for the next most reliable symbol. If the decision on the selected symbol is correct, then its interference towards other symbols can be completely suppressed. However, a wrong decision doubles the level of interference caused by the erroneously detected symbol. Consequently, the reliability criterion used for the ordering of symbols is of critical importance in any form of successive cancellation based on hard-cancellation. The symbol-based serial detection and cancellation process is depicted in Figure 2.12.

When the MF-SIC detector has access to reliability information at a bit level, in the form of log-likelihood ratios (LLRs) $\Lambda\left(b_m^{(i)}\right)$, ordering and cancellation can be performed at a bit level (rather than symbol level) based on the LLRs. In other words, the bit $b_m^{(i)}$ with the largest LLR value $\left|\Lambda\left(b_m^{(i)}\right)\right|$ (or minimum estimation error probability), can be selected as most reliable. Since bit estimates corresponding to a particular symbol can have different reliabilities, the use of LLR values represents an optimum ordering policy. First, matched filtering is performed on the sufficient statistics from Equation 2.18 to obtain the estimate for the bit of interest:

$$y_m^{(i)} = \begin{cases} \mathrm{Re}\left\{\underline{a}^{(i)\mathrm{H}} \underline{z}_w\right\} & \text{if } (m = 0) \\ \mathrm{Im}\left\{\underline{a}^{(i)\mathrm{H}} \underline{z}_w\right\} & \text{if } (m = 1) \end{cases} \tag{2.41}$$

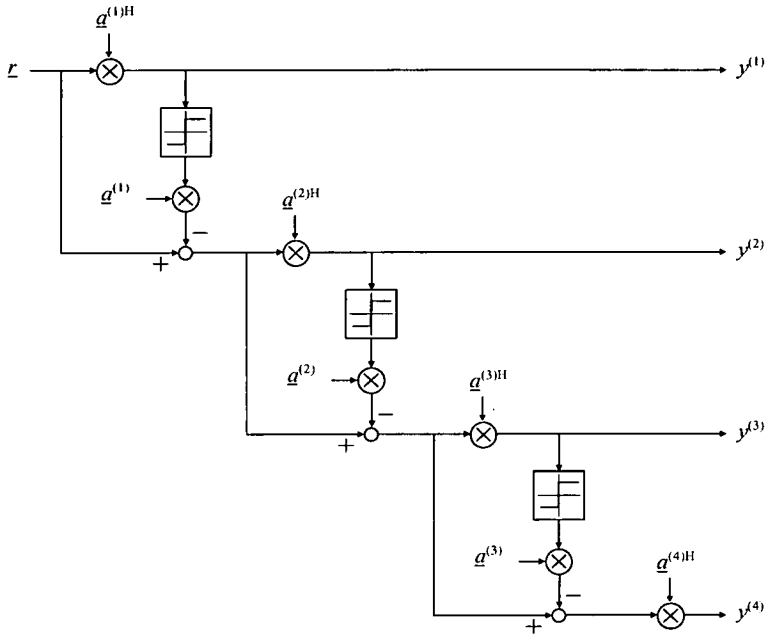


Figure 2.12: *Symbol-based MF-SIC overview.*

where $m = 0$ or 1 depending on whether the bit of interest forms the real or imaginary part of the 4-QAM symbol. Then the contribution from the detected bit is cancelled from the received signal. The cancellation process at each symbol interval is based on the more reliable hard bit estimates derived from the LLR values (for example from a soft-output decoder in an iterative receiver architecture:

$$\underline{z}_w := \underline{z}_w - j^m \underline{a}^{(i)} \mathrm{sgn}\left[\Lambda\left(b_m^{(i)}\right)\right]. \tag{2.42}$$

The process is again repeated for the next most reliable bit until all bits are detected.

Finally, LLR values can be calculated from the MF-SIC output as shown in Appendix A.3 for unnormalised matched filter outputs as

$$\Lambda\left(b_0^{(i)}\right) = \frac{4\left\|\underline{a}^{(i)}\right\|^2}{\sigma^2_{\eta^{(i)}}} \mathrm{Re}\left\{y^{(i)}\right\} \tag{2.43}$$

and

$$\Lambda\left(b_1^{(i)}\right) = \frac{4\left\|\underline{a}^{(i)}\right\|^2}{\sigma^2_{\eta^{(i)}}} \mathrm{Im}\left\{y^{(i)}\right\}. \tag{2.44}$$

For the LLR calculation, the interfering contributions $v^{(i)}$ from other antennas are ignored.

2.4.4 Detector performance comparison

In this section, the performance of the presented detectors is investigated for a 4×4 antenna MIMO link ($N_{\mathrm{T}} = N_{\mathrm{R}} = 4$) with 4-QAM modulation in two different scenarios. The MS-PPIC detector operates with 6 iterations. No error correction coding is employed and the channel conditions are assumed to be known at the detector.

Scenario 1: ergodic channel

A 1-tap ergodic MIMO channel is considered. No CDMA ($K = 1$, $Q = 1$) and no space-time equalisation is employed, therefore no pre-whitening is required. Figure 2.13 shows the BER results for the ergodic channel. As expected, the APP detector performs best. Note that at low E_{b}/N_0 the Max-Log-APP degrades slightly in comparison to the APP. Also note that at low E_{b}/N_0, the performance of the MS-PPIC approaches that of the APP and in fact becomes superior to that of the Max-Log-APP. The BER results show clearly that the low complexity MF-SIC algorithm is not suitable without *a priori* information.

Scenario 2: TU channel

A typical urban (TU) scenario with 10 chip-spaced taps is considered. The data is spread with a spreading factor of $Q = 16$, and the same set of $K = 16$ orthogonal Walsh spreading codes are simultaneously transmitted from each antenna resulting in a high-interference code re-use scenario. Here, the interference cancelling detectors are simulated both, with and without the space-time equaliser. Figure 2.14 shows the performance results for the TU channel. With space-time equalisation, the optimum APP algorithm performs best, followed by the Max-Log-APP. Again, the MS-PPIC approaches the APP and outperforms the Max-Log-APP at low E_{b}/N_0. In addition to the results with equalisation, the performance of the interference cancelling detectors is also shown without equalisation. Here, the detectors have to cancel the interference from all N_{T} antennas as well as from all K spreading codes. It is shown that, due to the higher interference, the performance of the MS-PPIC drops without equalisation, but also approaches the APP performance at low E_{b}/N_0. The MF-SIC based detectors with and without equalisation offer again only poor results, since no *a priori* information is available.

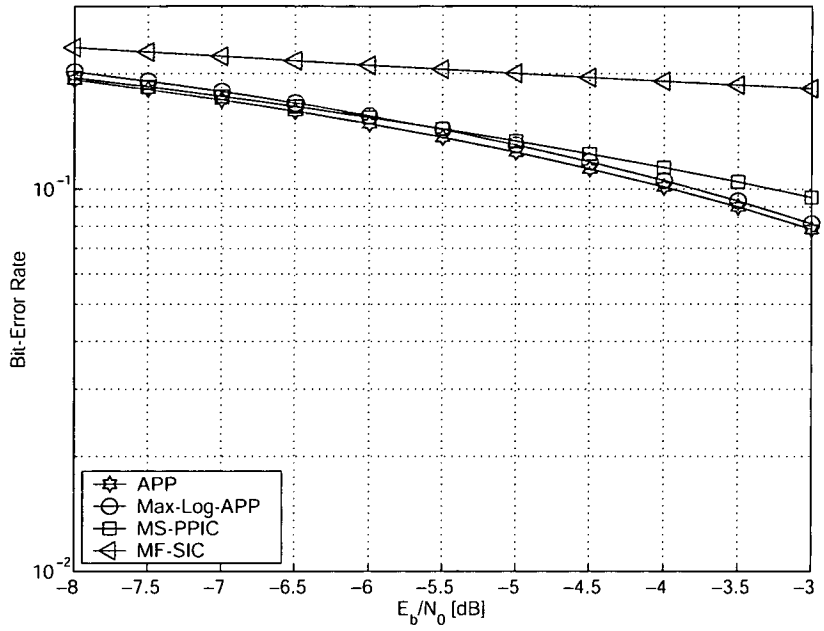


Figure 2.13: *Performance comparison for Scenario 1: ergodic channel,* 4×4 *antennas, 4-QAM modulation,* $Q = 1$*,* $K = 1$*, without equalisation.*

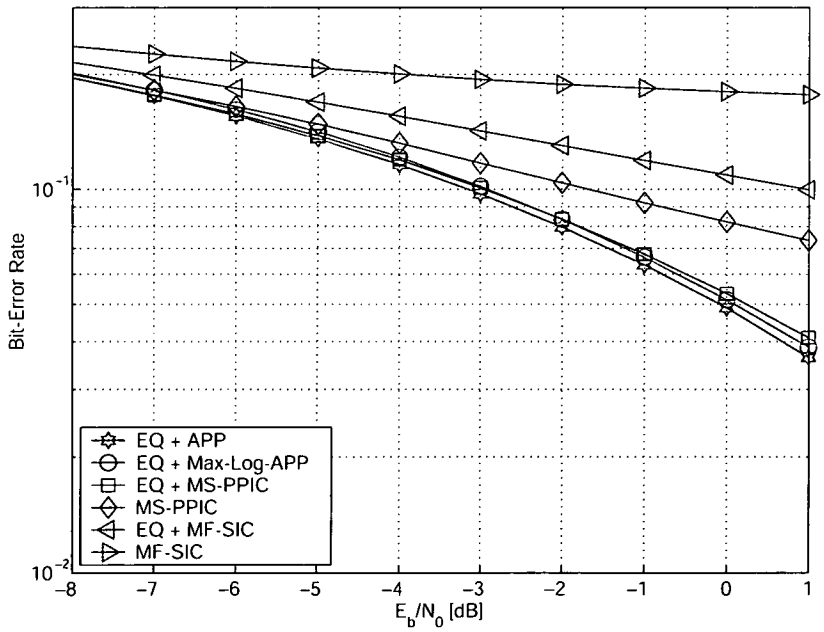


Figure 2.14: *Performance comparison for Scenario 2: TU channel,* 4×4 *antennas, 4-QAM modulation,* $Q = 16$*,* $K = 16$*, with equalisation.*

2.4.5 Detector complexity comparison

The computational complexity of the above detection algorithms can be compared in terms of the required number of real multiplications n_{MULT}, real additions n_{ADD}, table look-ups n_{TAB} and compare operations n_{CMP}.

For the initialisation of the APP and Max-Log-APP detectors, the required number of operations are $n_{\mathrm{ADD}} \approx 2^{2N_{\mathrm{T}}} 4N_{\mathrm{T}}^2$, where $2^{2N_{\mathrm{T}}}$ is the number of hypotheses over which the detector searches for 4-QAM modulation. This initialisation must be performed once for each spreading code during each channel estimation interval. At low mobile speeds, the latter is typically equal to a coding block interval. Following initialisation, the required number of operations per symbol period and per spreading code are $n_{\mathrm{MULT}} \approx 2^{2N_{\mathrm{T}}} 2N_{\mathrm{T}}$, $n_{\mathrm{ADD}} \approx 2^{2N_{\mathrm{T}}} 6N_{\mathrm{T}}$, $n_{\mathrm{CMP}} \approx 2^{2N_{\mathrm{T}}} 2N_{\mathrm{T}}$ and $n_{\mathrm{TAB}} \approx 2^{2N_{\mathrm{T}}} 2N_{\mathrm{T}}$ for the APP. The Jacobian algorithm [13] is used for efficient computation of Equation 2.28 where the correction function is implemented via a look-up table. The corresponding values for the Max-Log-APP are $n_{\mathrm{MULT}} \approx 2^{2N_{\mathrm{T}}} 2N_{\mathrm{T}}$, $n_{\mathrm{ADD}} \approx 2^{2N_{\mathrm{T}}} 4N_{\mathrm{T}}$ and $n_{\mathrm{CMP}} \approx 2^{2N_{\mathrm{T}}} 2N_{\mathrm{T}}$.

For the initialisation of the proposed MS-PPIC detector, the required number of operations are $n_{\mathrm{MULT}} \approx 4N_{\mathrm{T}}^3 + 6N_{\mathrm{T}}^2$, $n_{\mathrm{ADD}} \approx 4N_{\mathrm{T}}^3 + 2N_{\mathrm{T}}^2$ and $n_{\mathrm{DIV}} \approx 2N_{\mathrm{T}}$, where n_{DIV} is the number of real divisions. Following initialisation, the required number of operations per symbol period and per spreading code are $n_{\mathrm{MULT}} \approx N\left(4N_{\mathrm{T}}^2 + 4N_{\mathrm{T}}\right) + 4N_{\mathrm{T}}^2$, $n_{\mathrm{ADD}} \approx N\left(4N_{\mathrm{T}}^2 + 4N_{\mathrm{T}}\right) + 4N_{\mathrm{T}}^2$ and $n_{\mathrm{TAB}} \approx 4N_{\mathrm{T}}N$. The $\tanh()$ function used for partial cancellation is implemented via a look-up table.

The lowest computational complexity is required for the proposed MF-SIC detector. When bit-level detection and cancellation is assumed, the detection and cancellation order is updated for each symbol epoch and spreading code with $n_{\mathrm{CMP}} \approx 2N_{\mathrm{T}}^2$. Then, the required number of operations for matched filtering and interference cancellation are $n_{\mathrm{MULT}} \approx 8N_{\mathrm{T}}^2$ and $n_{\mathrm{ADD}} \approx 8N_{\mathrm{T}}^2$. Symbol-based operation of the MF-SIC requires only half the amount of operations.

Table 2.1 and Table 2.2 show a complexity comparison for the case of $N_{\mathrm{T}} = N_{\mathrm{R}} = 4$ and 4-QAM modulation. It is assumed that the MS-PPIC detector operates with $N_{\mathrm{I}} = 6$ stages. As can be seen, the computational complexity of the proposed MS-PPIC detector is significantly lower than those of the APP and Max-Log-APP detectors. The MF-SIC detector requires the lowest complexity. Of course, the real complexity depends on the implementation of the algorithms.

operations	Initialisation			
	APP	Max-Log APP	MS-PPIC	MF-SIC
n_{MULT}	0	0	353	0
n_{ADD}	16384	16384	288	0
n_{TAB}	0	0	0	0
n_{CMP}	0	0	0	0
n_{DIV}	0	0	8	0

Table 2.1: *Detector complexity comparison: Initialisation.*

operations	Normal Operation			
	APP	Max-Log APP	MS-PPIC	MF-SIC
n_{MULT}	2048	2048	544	128
n_{ADD}	6144	4096	544	128
n_{TAB}	2048	0	96	0
n_{CMP}	2048	2048	0	32
n_{DIV}	0	0	0	0

Table 2.2: *Detector complexity comparison: Normal operation.*

2.5 Summary and conclusions

In this chapter, receiver components such as space-time equalisation and different detection schemes were discussed for MIMO radio communications.

It was shown that an efficient strategy for dealing with dispersive channels is to perform detection only after a process of space-time equalisation, which effectively eliminates dispersion, followed by de-spreading. The equalisation process inevitably causes noise colouring, which needs to be accounted for in the detection process. The main advantage of space-time equalisation is that the detection process can deal with the pre-whitened signal for each spreading code separately. This reduces the computational complexity of the detector significantly, and thereby makes it possible to use the APP algorithm for detection in MIMO systems using CDMA.

For MIMO detection, multistage partial parallel interference cancellation (MS-PPIC) is proposed as a low complexity alternative to the APP detector based on space-time equalisation, de-spreading and pre-whitening. It was shown that the proposed MS-PPIC detector exhibits a performance close to the optimum APP detector, and furthermore, is superior to Max-Log-MAP detection at low E_b/N_0. In addition, it is shown that the MS-PPIC achieves this at a mere 25% of the Max-Log-APP complexity for a 4×4 antenna scenario. More importantly, the MS-PPIC is highly scalable in that its complexity grows only linearly with the number of transmitter antennas and not exponentially as the APP and Max-Log-APP detectors.

In addition, an ordered serial interference cancellation scheme based on matched filtering (MF-SIC) is presented. The complexity of this detector is only 5% of the Max-Log-APP complexity and is also lower than that of the proposed MS-PPIC. The MF-SIC detector is mainly aimed for the use in iterative receiver architectures, where the low complexity of the components is very important. It was shown that, when no *a priori* information is available for detection and cancellation, the MF-SIC offers only poor performance. However in the following chapters it will be demonstrated that in a iterative receiver architecture, the detector is able to achieve impressive performance.

Chapter 3
Improved turbo decoding via maximisation of mutual information transfer

3.1 Introduction

Since the discovery of turbo codes in 1993 [19] there has been renewed interest in the field of coding theory, with the aim of approaching the Shannon limit. Furthermore, with the proliferation of wireless mobile devices in recent years, the availability of low-cost and low-power decoder chips is of paramount importance. To this end, several techniques for reducing the complexity of the optimum MAP decoding algorithm [11] have been proposed. Examples include the Log-MAP and Max-Log-MAP algorithms [13, 20]. It has also been shown that the performance of the Max-Log-MAP algorithm can be improved by employing a fixed scaling factor for the exchanged *a priori* information in the decoding process [21].

In this chapter, a novel modification of the Max-Log-MAP algorithm is proposed for use in a turbo decoding process. The approach aims to maximise the mutual information at the input of each component decoder by correcting the bias in the *a priori* information caused by the Max-Log approximation in the previous component decoder [22]. This is performed by scaling the *a priori* information by optimised, iteration specific weight factors at each turbo iteration. A second contribution of this chapter is a method for the off-line computation of the optimal weights according to the maximum mutual information criterion. These are developed in Section 3.4. Sections 3.2 and 3.3 provide the necessary background. Finally, Section 3.5 presents a performance comparison and Section 3.6 provides a computational complexity comparison. It is shown that the performance of a turbo decoder using the modified Max-Log-MAP algorithm can be improved to approach that of a turbo decoder using the optimum Log-MAP or MAP algorithms. This is achieved at the expense of only two additional multiplications per systematic bit per turbo iteration. Furthermore, the insensitivity of the Max-Log-MAP algorithm to an arbitrary scaling of its input log-likelihood ratios is maintained. Extrinsic information transfer (EXIT) charts [5] are used to analyse the convergence behaviour of the turbo decoder.

3.2 Turbo decoding overview

Consider the received signal, $r_t = x_t + n_t$, at the output of an AWGN channel at time instant t, where $x_t \in \{+1, -1\}$ is the transmitted binary symbol (corresponding to the encoded bit $b_t \in \{1, 0\}$ and n_t is zero-mean Gaussian noise of variance $\mathrm{E}\{n_t^2\} = N_0$. Then the log-likelihood ratio (LLR) of the transmitted symbol is defined as

$$\Lambda(x_t) = \log\frac{\mathrm{P}\{x_t = +1\}}{\mathrm{P}\{x_t = -1\}} = \frac{2}{N_0} r_t \tag{3.1}$$

where $\mathrm{P}\{A\}$ represents the probability of event A. Let us also consider a parallel concatenated turbo encoding process of rate $1/3$ at the transmitter. This consists of two $1/2$ rate recursive systematic convolutional (RSC) encoders separated by an interleaving process, resulting in transmitted systematic symbol $x_{t,0}$ and parity symbols $x_{t,1}$ and $x_{t,2}$ as described in [13]. The corresponding likelihood variables at the output of the channel (input of the decoder) may then be expressed as $\Lambda_\mathrm{c}(x_{t,0})$, $\Lambda_\mathrm{c}(x_{t,1})$ and $\Lambda_\mathrm{c}(x_{t,2})$.

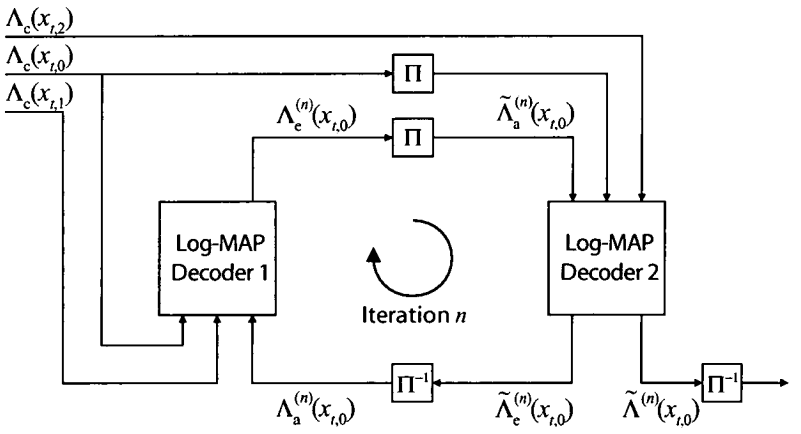


Figure 3.1: *Turbo decoding for parallel concatenated codes.*

Figure 3.1 depicts the turbo decoding procedure whereby decoding is performed in an iterative manner via two soft-output component decoders, separated by an interleaver, with the objective of improving the estimates of $x_{t,0}$ from iteration n to iteration $n+1$. The first decoder generates extrinsic information $\Lambda_\mathrm{e}^{(n)}(x_{t,0})$ on the systematic bits, which then serves as *a priori* information $\tilde{\Lambda}_\mathrm{a}^{(n)}(x_{t,0})$ for the second decoding process. The symbol ' ~ ' denotes interleaved quantities.

The maximum a posteriori probability (MAP) algorithm is the optimum strategy for the decoding of RSC codes, as it results in a minimum probability of bit error. However, due to its high computational complexity, the MAP algorithm is usually implemented in the logarithmic domain in the form of the Log-MAP or Max-Log-MAP algorithms. While the former is mathematically equivalent to the MAP algorithm, the latter involves an approximation which results in even lower complexity, albeit at the expense of some degradation in performance [13, 20]. For purposes of brevity, the expressions presented in this chapter are written for the first component decoder, with obvious extensions to the second decoder.

3.2.1 The Log-MAP algorithm

The Log-MAP algorithm [13, 23] is the log-domain implementation of the MAP algorithm and operates directly on LLRs. Given the LLRs for the systematic and parity bits as well as *a priori* LLRs for the systematic bits, the Log-MAP algorithm computes new LLRs for the systematic bits as described below

$$\begin{aligned}\Lambda\left(x_{t,0}\right) &= \log\frac{\sum_{l=0}^{S-1}\exp\left\{\bar{\alpha}_{t-1}\left(l'\right)+\bar{\gamma}_t^{[1]}\left(l',l\right)+\bar{\beta}_t\left(l\right)\right\}}{\sum_{l=0}^{S-1}\exp\left\{\bar{\alpha}_{t-1}\left(l'\right)+\bar{\gamma}_t^{[0]}\left(l',l\right)+\bar{\beta}_t\left(l\right)\right\}} \\ &= \Lambda_{\mathrm{a}}\left(x_{t,0}\right)+\Lambda_{\mathrm{c}}\left(x_{t,0}\right)+\Lambda_{\mathrm{e}}\left(x_{t,0}\right) \end{aligned} \tag{3.2}$$

where $\bar{\gamma}_t^{[q]}\left(l',l\right)$ is the logarithm of the probability of a transition from state l' to state l of the encoder trellis at time instant t, given that the systematic bit takes on value $q \in \{1,0\}$ and S is the total number of states in the trellis. Note that the new information at the decoder output regarding the systematic bits is encapsulated in the extrinsic information term $\Lambda_{\mathrm{e}}\left(x_{t,0}\right)$. Coefficients $\bar{\alpha}_t\left(l'\right)$ and $\bar{\beta}_t\left(l\right)$ are forward- and backward-accumulated metrics at time t. For a data block of τ systematic bits $\left(x_{1,0}\ldots x_{\tau,0}\right)$ and the corresponding parity bits $\left(x_{1,1}\ldots x_{\tau,1}\right)$, these coefficients are calculated as follows:

- **Forward Recursion** – Initialise $\bar{\alpha}_0\left(l\right)$, $l=0,1\ldots S-1$ such that $\bar{\alpha}_0\left(0\right)=0$ and $\bar{\alpha}_0\left(l\right)=-\infty$ for $l\neq 0$. Then

$$\bar{\gamma}_t^{[q]}\left(l',l\right)=\frac{1}{2}\left\{\left\{\Lambda_{\mathrm{a}}\left(x_{t,0}\right)+\Lambda_{\mathrm{c}}\left(x_{t,0}\right)\right\}x_{t,0}^{[q]}+\Lambda_{\mathrm{c}}\left(x_{t,1}\right)x_{t,1}^{[q]}\right\} \tag{3.3}$$

$$\text{and} \qquad \bar{\alpha}_t(l) = \log \sum_{l'=0}^{S-1} \sum_{q=0,1} \exp\left\{\bar{\alpha}_{t-1}(l') + \bar{\gamma}_t^{[q]}(l', l)\right\}. \tag{3.4}$$

- **Backward Recursion** – Initialise $\bar{\beta}_\tau(l)$, $l = 0, 1 \ldots S-1$ such that $\bar{\beta}_\tau(0) = 0$ and $\bar{\beta}_\tau(l) = -\infty$ for $l \neq 0$. Then

$$\bar{\beta}_t(l) = \log \sum_{l'=0}^{S-1} \sum_{q=0,1} \exp\left\{\bar{\beta}_{t+1}(l') + \bar{\gamma}_{t+1}^{[q]}(l, l')\right\}. \tag{3.5}$$

Equation 3.2 can be readily implemented via the Jacobian equality $\log\left(e^{\delta_1} + e^{\delta_2}\right) = \max(\delta_1 + \delta_2) + \log\left(1 + e^{-|\delta_1 - \delta_2|}\right)$ and using a look-up table to evaluate the correction function $\log\left(1 + e^{-|\delta_1 - \delta_2|}\right)$.

3.2.2 The Max-Log-MAP algorithm

The complexity of the Log-MAP algorithm can be further reduced by using the Max-Log approximation $\log\left(e^{\delta_1} + e^{\delta_2}\right) \approx \max(\delta_1 + \delta_2)$ for evaluating Equation 3.2. Clearly, this results in biased soft outputs and degrades the performance of the decoder. Nevertheless, the Max-Log-MAP algorithm is often the preferred choice for implementing a MAP decoder since it has the added advantage that its operation is insensitive to a scaling of the input LLRs. Using the Max-Log-MAP algorithm, the LLRs for the systematic bits can be calculated as

$$\begin{aligned}\Lambda(x_{t,0}) = &\max_l \left[\bar{\alpha}_{t-1}(l') + \bar{\gamma}_t^{[1]}(l', l) + \bar{\beta}_t(l)\right] \\ &- \max_l \left[\bar{\alpha}_{t-1}(l') + \bar{\gamma}_t^{[0]}(l', l) + \bar{\beta}_t(l)\right].\end{aligned} \tag{3.6}$$

When the inputs of the decoder are scaled by a factor ξ, this results in scaled values of $\bar{\alpha}_t(l)$, $\bar{\beta}_t(l)$ and $\bar{\gamma}_t^{[q]}(l', l)$ by the same factor. When the Log-MAP algorithm is used, the LLR calculation is clearly affected, however it can be shown that when the Max-Log-MAP algorithm is employed, this simply results in a scaled decoder output:

$$\begin{aligned}\xi\Lambda(x_{t,0}) = &\max_l \left[\xi\bar{\alpha}_{t-1}(l') + \xi\bar{\gamma}_t^{[1]}(l', l) + \xi\bar{\beta}_t(l)\right] \\ &- \max_l \left[\xi\bar{\alpha}_{t-1}(l') + \xi\bar{\gamma}_t^{[0]}(l', l) + \xi\bar{\beta}_t(l)\right].\end{aligned} \tag{3.7}$$

This implies that knowledge or estimation of the channel noise variance N_0 is not required to scale the decoder inputs to correct LLR values when using the Max-Log-MAP algorithm for decoding.

3.3 Extrinsic information transfer (EXIT) charts

The performance and convergence behaviour of turbo codes can be analysed using extrinsic information transfer (EXIT) charts, as proposed in [24]. The idea is to visualise the evolution of the mutual information exchanged between the component decoders from iteration to iteration. EXIT charts operate under the following assumptions: a) The *a priori* information is fairly uncorrelated from channel observations. This is valid for large interleaver sizes. b) The extrinsic information $\Lambda_{\mathrm{e}}(x_{t,0})$ has a Gaussian-like distribution, as shown in [25] for the MAP decoder.

An EXIT chart consists of a pair of curves which represent the mutual information transfer functions of the component decoders in the turbo process. Each curve is essentially a plot of *a priori* mutual information I_{a} against extrinsic mutual information I_{e} for the component decoder of interest. Here, the mutual information is a measure of the degree of dependency between the log-likelihood variables $\Lambda_{\mathrm{a}}(x_{t,0})$ or $\Lambda_{\mathrm{e}}(x_{t,0})$, and the corresponding transmitted bits $x_{t,0}$. The mutual information takes on values between 0 for no knowledge, and 1 for perfect knowledge of the transmitted bits, dependent on the reliability of the likelihood variables. The terms I_{a} and I_{e} are related to the probability density functions (pdfs) of $\Lambda_{\mathrm{a}}(x_{t,0})$ and $\Lambda_{\mathrm{e}}(x_{t,0})$, the signal-to-noise ratio E_{b}/N_0 and the RSC encoder polynomials. If the component decoders are identical, the two curves are naturally mirror images. The required pdfs can be estimated by generating histograms $p(\Lambda_{\mathrm{a}})$ and $p(\Lambda_{\mathrm{e}})$ of $\Lambda_{\mathrm{a}}(x_{t,0})$ and $\Lambda_{\mathrm{e}}(x_{t,0})$ respectively for a particular value of E_{b}/N_0 where E_{b} denotes the energy per information bit. This can be achieved by applying *a priori* information modelled as $\Lambda_{\mathrm{a}}(x_{t,0}) = \mu_{\mathrm{a}} x_{t,0} + n_{\mathrm{a},t}$, $t = 1 \ldots \tau$ to the input of a component decoder and observing the output $\Lambda_{\mathrm{e}}(x_{t,0})$ for a coded data block corresponding to τ information bits. The random variable $n_{\mathrm{a},t}$ is zero-mean Gaussian with variance $\mathrm{E}\{n_{\mathrm{a},t}^2\} = \sigma_{\mathrm{a}}^2$ such that $\sigma_{\mathrm{a}}^2 = 2\mu_{\mathrm{a}}$. The latter is a requirement for $\Lambda_{\mathrm{a}}(x_{t,0})$ to be a LLR. The mutual information I_{a} may then be computed as

$$I_{\mathrm{a}} = \sum_{q=-1,1} \frac{1}{2} \int_{-\infty}^{+\infty} p(\Lambda_{\mathrm{a}} | x_{t,0} = q) \log_2 \frac{2p(\Lambda_{\mathrm{a}} | x_{t,0} = q)}{p_{\mathrm{a}}} d\Lambda_{\mathrm{a}} \tag{3.8}$$

where $p_{\mathrm{a}} = p(\Lambda_{\mathrm{a}} | x_{t,0} = -1) + p(\Lambda_{\mathrm{a}} | x_{t,0} = +1)$. Similarly, I_{e} can be computed as

$$I_{\mathrm{e}} = \sum_{q=-1,1} \frac{1}{2} \int_{-\infty}^{+\infty} p\left(\Lambda_{\mathrm{e}} | x_{t,0} = q\right) \log_2 \frac{2p\left(\Lambda_{\mathrm{e}} | x_{t,0} = q\right)}{p_{\mathrm{e}}} d\Lambda_{\mathrm{e}} \tag{3.9}$$

where $p_{\mathrm{e}} = p\left(\Lambda_{\mathrm{e}} | x_{t,0} = -1\right) + p\left(\Lambda_{\mathrm{e}} | x_{t,0} = +1\right)$. The resulting pair $(I_{\mathrm{a}}, I_{\mathrm{e}})$ defines one point on the transfer function curve. To improve the reliability of the results, the mutual information can be calculated by averaging the statistics over several coded blocks. Different points (for the same E_{b}/N_0) can be obtained by varying the value of σ_{a}^2.

Having derived the transfer functions, we may now observe the trajectory of mutual information at various iterations of an actual turbo decoding process. At each iteration, mutual information is again computed as in 3.8 and 3.9, however the *a priori* LLR, $\Lambda_{\mathrm{a}}\left(x_{t,0}\right)$, at the input of the component decoder is no longer a modelled random variable but corresponds to the actual extrinsic LLR generated by the previous component decoding operation.

Figures 3.2 and 3.3 illustrate EXIT charts with trajectories of mutual information for the Log-MAP and Max-Log-MAP algorithms respectively. The "snapshot" trajectories correspond to turbo decoding iterations for a specific coded data block. The 1/2 rate (punctured) turbo encoder consists of two component RSC encoders, each operating at 1/2 rate with a memory of $v = 4$ (16 states) and octal generator polynomials $(G_{\mathrm{r}}, G) = (1 + D + D^4, 1 + D + D^2 + D^3 + D^4)$, where G_{r} denotes the recursive feedback polynomial. Note that while the mutual information trajectory for the Log-MAP algorithm in Figure 3.2 fits the predicted transfer function, the trajectory in Figure 3.3 clearly indicates the impact of numerical errors resulting from the Max-Log approximation: the trajectory stalls after only the first iteration and the turbo decoder is unable to converge at the simulated E_{b}/N_0 of 1 dB.

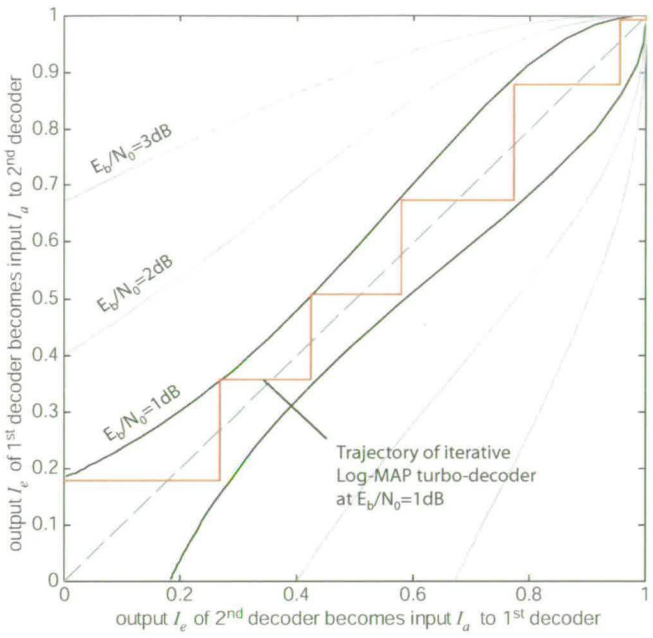


Figure 3.2: *EXIT chart with snapshot trajectory for the* $1/2$ *rate (punctured) turbo decoder using the Log-MAP algorithm (16 states and* 10^5*-bit interleaver).*

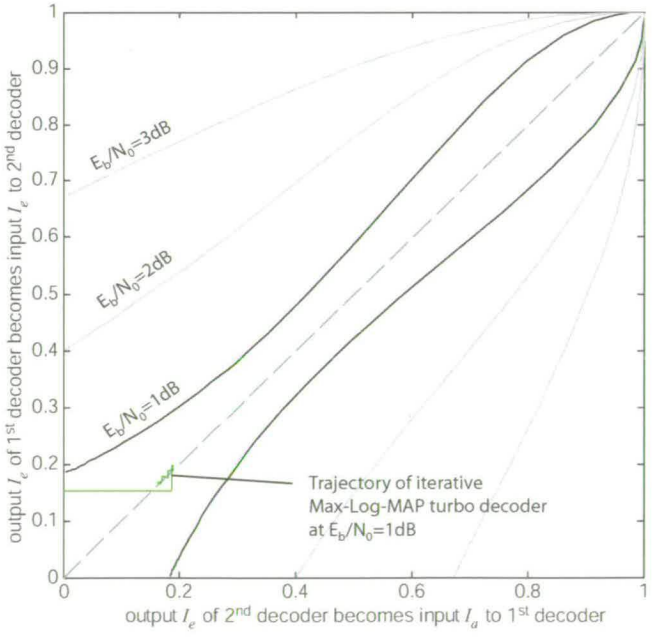


Figure 3.3: *EXIT chart with snapshot trajectory for the* $1/2$ *rate (punctured) turbo decoder using the Max-Log-MAP algorithm (16 states and* 10^5*-bit interleaver).*

3.4 The maximum mutual information principle

The poor convergence of the turbo decoder using the Max-Log-MAP algorithm is due to the accumulating bias in the extrinsic information caused by the max() operations. Since extrinsic information is used as *a priori* information, $\Lambda_{\mathrm{a}}(x_{t,0})$, for the next component decoding operation, and is combined with channel observations $\Lambda_{\mathrm{c}}(x_{t,0})$, as shown in [20], this bias leads to sub-optimal combining proportions in the decoder. To correct for this phenomenon, the logarithmic transition probabilities at the n^{th} iteration may be modified as follows:

$$\bar{\gamma}_t^{[q]}(l',l) = \frac{1}{2}\left\{\left\{w_{\mathrm{a}}^{(n)}\Lambda_{\mathrm{a}}^{(n)}(x_{t,0}) + \Lambda_{\mathrm{c}}^{(n)}(x_{t,0})\right\}x_{t,0}^{[q]} + \Lambda_{\mathrm{c}}^{(n)}(x_{t,1})\,x_{t,1}^{[q]}\right\}. \tag{3.10}$$

In other words, the bias of the *a priori* information can be corrected by scaling it by a factor $w_{\mathrm{a}}^{(n)}$ at the n^{th} iteration, as depicted in Figure 3.4.

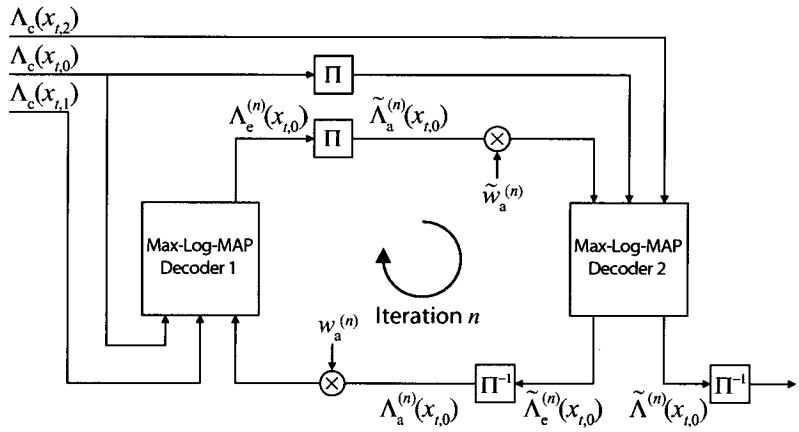


Figure 3.4: *Turbo decoding with weighing of a priori information.*

This correction procedure for the Max-Log-MAP algorithm is far less complex than the correction function employed in the Log-MAP algorithm. Furthermore, and perhaps more importantly from a practical point of view, the corrected Max-Log-MAP algorithm remains insensitive to an arbitrary scaling of the LLR values at its input, thereby eliminating the need to estimate the noise variance at the channel output. From observations of the EXIT charts in the previous section, it is evident that rapid convergence of the turbo process relies on the effective exchange of mutual information between the component decoders. It may be inferred that the optimum value for the weight factor $w_{\mathrm{a}}^{(n)}$ is that which maximises the mutual information between the term $\zeta_t^{(n)} = w_{\mathrm{a}}^{(n)}\Lambda_{\mathrm{a}}^{(n)}(x_{t,0}) + \Lambda_{\mathrm{c}}^{(n)}(x_{t,0})$ and the vector of "uncorrupted" LLRs $\underline{\lambda}_t^{(n)}$ for each

component decoder and at each iteration n. Using vector notation, $\zeta_t^{(n)}$ may be modelled as

$$\begin{aligned}\zeta_t^{(n)} &= \begin{bmatrix} w_{\mathrm{a}}^{(n)} & 1 \end{bmatrix} \begin{bmatrix} \Lambda_{\mathrm{a}}^{(n)}(x_{t,0}) \\ \Lambda_{\mathrm{c}}^{(n)}(x_{t,0}) \end{bmatrix} = \left(\underline{w}^{(n)}\right)^{\mathrm{T}} \underline{\Lambda}_t^{(n)} \\ &= \left(\underline{w}^{(n)}\right)^{\mathrm{T}} \left(\underline{\lambda}_t^{(n)} + \underline{\epsilon}_t^{(n)}\right) = \left(\underline{w}^{(n)}\right)^{\mathrm{T}} \underline{\lambda}_t^{(n)} + v_t^{(n)} \end{aligned} \tag{3.11}$$

where $\underline{\epsilon}_t^{(n)}$ represents the contributions of channel noise plus the numerical approximation error inherent in the Max-Log-MAP algorithm. Given variances

$$s_\zeta^{(n)} = \mathrm{E}\left\{\left(\underline{w}^{(n)}\right)^{\mathrm{T}} \left(\underline{\lambda}_t^{(n)} + \underline{\epsilon}_t^{(n)}\right)\left(\underline{\lambda}_t^{(n)} + \underline{\epsilon}_t^{(n)}\right)^{\mathrm{T}} \underline{w}^{(n)}\right\} = \left(\underline{w}^{(n)}\right)^{\mathrm{T}} \boldsymbol{R}_{\lambda+\epsilon}^{(n)} \underline{w}^{(n)} \tag{3.12}$$

$$s_v^{(n)} = \mathrm{E}\left\{\left(\underline{w}^{(n)}\right)^{\mathrm{T}} \underline{\epsilon}_t^{(n)} \left(\underline{\epsilon}_t^{(n)}\right)^{\mathrm{T}} \underline{w}^{(n)}\right\} = \left(\underline{w}^{(n)}\right)^{\mathrm{T}} \boldsymbol{R}_\epsilon^{(n)} \underline{w}^{(n)} \tag{3.13}$$

and modelling $v_t^{(n)}$ as a Gaussian random variable, the differential and conditional entropies of $\zeta_t^{(n)}$ are

$$h\left(\zeta_t^{(n)}\right) = \frac{1}{2}\left[1 + \ln\left(2\pi s_\zeta^{(n)}\right)\right] \tag{3.14}$$

$$h\left(\zeta_t^{(n)} \mid \underline{\lambda}_t^{(n)}\right) = \frac{1}{2}\left[1 + \ln\left(2\pi s_v^{(n)}\right)\right]. \tag{3.15}$$

By definition [26], the mutual information can be written as

$$I\left(\zeta_t^{(n)} ; \underline{\lambda}_t^{(n)}\right) = h\left(\zeta_t^{(n)}\right) - h\left(\zeta_t^{(n)} \mid \underline{\lambda}_t^{(n)}\right) = \frac{1}{2}\log\frac{s_\zeta^{(n)}}{s_v^{(n)}} \tag{3.16}$$

and the optimum weight factors can then be derived as

$$\underline{w}_{\mathrm{OPT}}^{(n)} = \arg\max_{\underline{w}^{(n)}} \frac{s_\zeta^{(n)}}{s_v^{(n)}} = \arg\max_{\underline{w}^{(n)}} \frac{\left(\underline{w}^{(n)}\right)^{\mathrm{T}} \boldsymbol{R}_{\lambda+\epsilon}^{(n)} \underline{w}^{(n)}}{\left(\underline{w}^{(n)}\right)^{\mathrm{T}} \boldsymbol{R}_\epsilon^{(n)} \underline{w}^{(n)}}. \tag{3.17}$$

Setting $\underline{z} = \left(\boldsymbol{R}_\epsilon^{(n)}\right)^{\mathrm{T}/2} \underline{w}^{(n)}$, we arrive at the Rayleigh-quotient problem [27]

$$\underline{z}_{\mathrm{OPT}}^{(n)} = \arg\max_{\underline{w}^{(n)}} \frac{\underline{z}^{\mathrm{T}} \left(\boldsymbol{R}_\epsilon^{(n)}\right)^{-1/2} \boldsymbol{R}_{\lambda+\epsilon}^{(n)} \left(\boldsymbol{R}_\epsilon^{(n)}\right)^{-\mathrm{T}/2} \underline{z}}{\underline{z}^{\mathrm{T}} \underline{z}} \tag{3.18}$$

with solutions

$$\underline{z}_{\mathrm{OPT}}^{(n)} = k\,\mathrm{eig}_{\max}\left\{\left(\boldsymbol{R}_{\epsilon}^{(n)}\right)^{-1/2}\boldsymbol{R}_{\lambda+\epsilon}^{(n)}\left(\boldsymbol{R}_{\epsilon}^{(n)}\right)^{-\mathrm{T}/2}\right\} \tag{3.19}$$

$$\underline{w}_{\mathrm{OPT}}^{(n)} = k\left(\boldsymbol{R}_{\epsilon}^{(n)}\right)^{-\mathrm{T}/2}\mathrm{eig}_{\max}\left\{\left(\boldsymbol{R}_{\epsilon}^{(n)}\right)^{-1/2}\boldsymbol{R}_{\lambda+\epsilon}^{(n)}\left(\boldsymbol{R}_{\epsilon}^{(n)}\right)^{-\mathrm{T}/2}\right\} \tag{3.20}$$

where $\mathrm{eig}_{\max}\{\boldsymbol{A}\}$ is the eigenvector of $\boldsymbol{A}$ corresponding to its largest eigenvalue. The scalar k is chosen such that the second element of $\underline{w}_{\mathrm{OPT}}^{(n)}$, i.e. the weight factor of $\Lambda_{\mathrm{c}}^{(n)}(x_{t,0})$, equals unity. Inspection of 3.11 to 3.19 reveals that the optimum weights are functions of the iteration index, the error correcting capabilities of the component decoders (i.e. encoder polynomials) and the signal to noise ratio. The optimum weights $\underline{w}_{\mathrm{OPT}}^{(n)}$ can be computed or "trained" off-line based on time-averaged estimates of correlation matrices $\boldsymbol{R}_{\lambda+\epsilon}^{(n)}$ and $\boldsymbol{R}_{\epsilon}^{(n)}$ derived over a sufficiently long data block corresponding to τ encoded information bits. Specifically

$$\boldsymbol{R}_{\lambda+\epsilon}^{(n)} = \mathrm{E}\left\{\underline{\Lambda}_t^{(n)}\left(\underline{\Lambda}_t^{(n)}\right)^{\mathrm{T}}\right\} = \lim_{\tau\to\infty}\frac{1}{\tau}\sum_{t=1}^{\tau}\underline{\Lambda}_t^{(n)}\left(\underline{\Lambda}_t^{(n)}\right)^{\mathrm{T}}. \tag{3.21}$$

Furthermore, a vector $\underline{\lambda}_t^{(n)}$ of "uncorrupted" LLRs can be defined using the fixed amplitudes of $\mathrm{E}\{\Lambda_{\mathrm{a}}^{(n)}(x_{t,0})\cdot x_{t,0}\}$ for the *a priori* information, and $\mathrm{E}\{\Lambda_{\mathrm{c}}^{(n)}(x_{t,0})\cdot x_{t,0}\}$ for the channel information. Note that the transmitted symbols $x_{t,0}$ are available for offline optimisation.

$$\underline{\lambda}_t^{(n)} = \begin{bmatrix} \lambda_{\mathrm{a}}^{(n)}(x_{t,0}) \\ \lambda_{\mathrm{c}}^{(n)}(x_{t,0}) \end{bmatrix} = \begin{bmatrix} \mathrm{E}\{\Lambda_{\mathrm{a}}^{(n)}(x_{t,0})\cdot x_{t,0}\}x_{t,0} \\ \mathrm{E}\{\Lambda_{\mathrm{c}}^{(n)}(x_{t,0})\cdot x_{t,0}\}x_{t,0} \end{bmatrix} \tag{3.22}$$

so that

$$\boldsymbol{R}_{\lambda}^{(n)} = \mathrm{E}\left\{\underline{\lambda}_t^{(n)}\left(\underline{\lambda}_t^{(n)}\right)^{\mathrm{T}}\right\} = \lim_{\tau\to\infty}\begin{bmatrix} \left(\phi^{(n)}\right)^2 & \phi^{(n)}\theta^{(n)} \\ \theta^{(n)}\phi^{(n)} & \left(\theta^{(n)}\right)^2 \end{bmatrix} \tag{3.23}$$

where $\phi^{(n)} = \frac{1}{\tau}\sum_{t=1}^{\tau}\Lambda_{\mathrm{a}}^{(n)}(x_{t,0})\cdot x_{t,0}$ and $\theta^{(n)} = \frac{1}{\tau}\sum_{t=1}^{\tau}\Lambda_{\mathrm{c}}^{(n)}(x_{t,0})\cdot x_{t,0}$. Finally, assuming that vectors $\underline{\epsilon}_t^{(n)}$ and $\underline{\lambda}_t^{(n)}$ are uncorrelated, one may derive $\boldsymbol{R}_{\epsilon}^{(n)}$ as $\boldsymbol{R}_{\lambda+\epsilon}^{(n)} - \boldsymbol{R}_{\lambda}^{(n)}$. The above training procedure should be performed under E_{b}/N_0 conditions that are typical at the bit-error rate range of interest.

3.5 Decoder performance comparison

To investigate the decoder performance for differend code polynomials, two turbo encoders are considered at the input of an AWGN channel. The first $1/2$ rate (punctured) turbo encoder consists of two $1/2$ rate component RSC codes of memory 4, octal polynomials $(G_\mathrm{r}, G) = \left(1 + D + D^4, 1 + D + D^2 + D^3 + D^4\right)$ and an interleaver size of 10^5 bits. The second $1/2$ rate (punctured) turbo encoder is that specified for UMTS [8] and consists of two $1/2$ rate component RSC codes of memory 3, octal polynomials $(G_\mathrm{r}, G) = \left(1 + D^2 + D^3, 1 + D + D^3\right)$ and an interleaver size of 5114 bits. This is the maximum block size specified for high speed downlink packet access (HSDPA) in UMTS. Table 3.1 shows the optimum weight factors derived off-line for each iteration of the two turbo decoders at E_b/N_0 of 1.0 and 0.7 dB respectively. The impact of the combining scheme of Equation 3.10 on the mutual information trajectory of the first turbo decoder is indicated in Figure 3.5. In comparison to the original trajectory of Figure 3.3, turbo decoding with the improved Max-Log-MAP algorithm does not stall and is able to converge almost as well as with the Log-MAP algorithm for the same input data as before. However, there is still some degradation with respect to the full Log-MAP decoder, since more iterations are required to achieve the same level of performance. These results are achieved at the expense of only two additional multiplications per iteration per systematic bit. Figure 3.6 shows the BER performance of the first turbo decoder after 6 iterations. The results show that the proposed MMIC scheme significantly improves the performance of the Max-Log-MAP turbo decoder. Figure 3.7 shows the BER results for the UMTS turbo decoder after 6 iterations. Again, the performance of the turbo decoder using the Max-Log-MAP algorithm and MMIC approaches that of the turbo decoder using the optimum Log-MAP algorithm. The performance difference can be reduced down to only 0.05 dB at a BER of 10^{-4}.

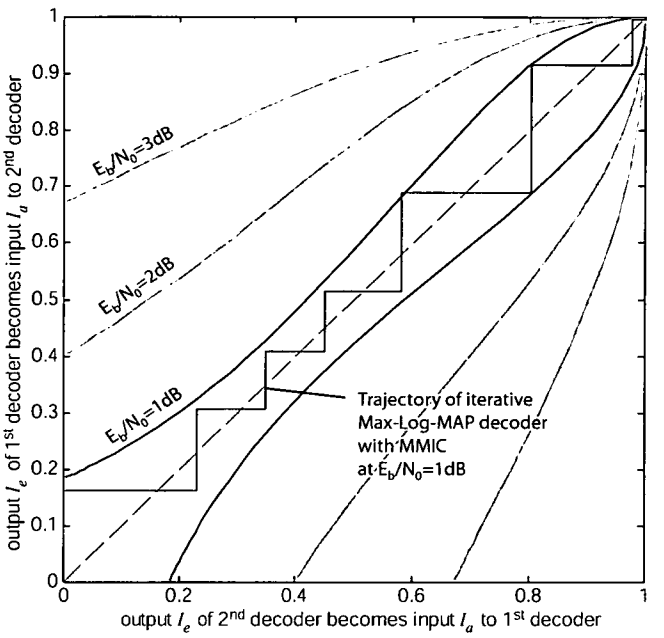


Figure 3.5: *EXIT chart with snapshot trajectory for the* $1/2$ *rate (punctured) turbo decoder using the proposed Max-Log-MAP algorithm with MMIC (16 states and* 10^5*-bit interleaver).*

iteration n	Decoder 1, 1.0 dB		Decoder 2 (UMTS), 0.7 dB	
	$w_a^{(n)}$	$\tilde{w}_a^{(n)}$	$w_a^{(n)}$	$\tilde{w}_a^{(n)}$
1	0*	0.505	0*	0.517
2	0.566	0.602	0.581	0.617
3	0.629	0.656	0.640	0.668
4	0.682	0.712	0.683	0.713
5	0.754	0.814	0.732	0.769
6	0.892	1.020	0.792	0.837

* no a priori knowledge in iteration 1 for the first component decoder.

Table 3.1: *Optimised weight factors.*

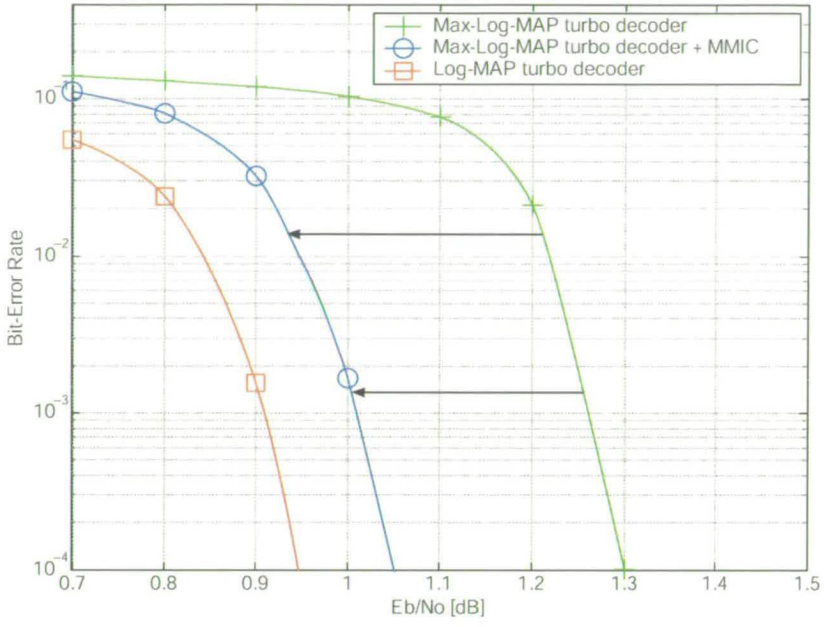


Figure 3.6: *Performance for the first* $1/2$ *rate (punctured) turbo decoder (16 states,* 10^5*-bit interleaver).*

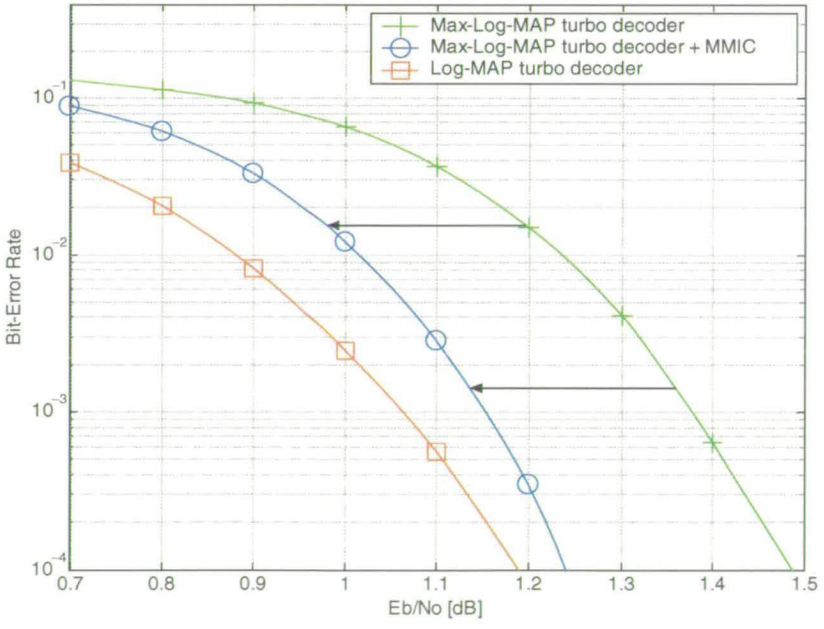


Figure 3.7: *Performance for the* $1/2$ *rate (punctured) UMTS turbo decoder (8 states, 5114-bit interleaver).*

3.6 Decoder complexity comparison

The computational complexity of the decoder algorithms and the proposed MMIC technique can be compared in terms of the required number of real multiplications n_{MULT}, real additions n_{ADD}, table look-ups n_{TAB} and compare operations n_{CMP}. A complexity comparison of the standard decoding algorithms is given in [13]. The complexity is given per information bit for N_{I} iterations of the rate $1/3$ turbo decoder consisting of two rate $1/2$ convolutional component decoders of memory order v.

For the Log-MAP algorithm the number of required operations are: $n_{\mathrm{MULT}} = 2N_{\mathrm{I}}\,(2 \cdot 2^{v})$, $n_{\mathrm{ADD}} = 2N_{\mathrm{I}}\,(12 \cdot 2^{v} + 6)$, $n_{\mathrm{CMP}} = 2N_{\mathrm{I}}\,(4 \cdot 2^{v} - 2)$ and $n_{\mathrm{TAB}} = 2N_{\mathrm{I}}\,(4 \cdot 2^{v} - 2)$.

For the Max-Log-MAP algorithm, the number of required operations are $n_{\mathrm{MULT}} = 2N_{\mathrm{I}}\,(2 \cdot 2^{v})$, $n_{\mathrm{ADD}} = 2N_{\mathrm{I}}\,(8 \cdot 2^{v} + 8)$ and $n_{\mathrm{CMP}} = 2N_{\mathrm{I}}\,(4 \cdot 2^{v} - 2)$. The proposed MMIC technique increases the number of multiplications by only 1 per iteration and component decoder to $n_{\mathrm{MULT}} = 2N_{\mathrm{I}}\,(2 \cdot 2^{v} + 1)$.

Table 3.2 and Table 3.3 show the turbo decoder complexities for the two decoders under investigation for $N_{\mathrm{I}} = 6$ decoding iterations. It is shown that the memory order v has a significant impact on the complexity. The proposed MMIC technique adds only little complexity.

operations	Log-MAP	Max-Log-MAP	Max-Log-MAP + MMIC
n_{MULT}	384	384	396
n_{ADD}	2376	1632	1632
n_{CMP}	744	744	744
n_{TAB}	744	0	0

Table 3.2: *Turbo decoder complexity per information bit with $v = 4$ and $N_{\mathrm{I}} = 6$ decoding iterations.*

operations	Log-MAP	Max-Log-MAP	Max-Log-MAP + MMIC
n_{MULT}	192	192	204
n_{ADD}	1224	864	864
n_{CMP}	360	360	360
n_{TAB}	360	0	0

Table 3.3: *UMTS Turbo decoder complexity per information bit with $v = 3$ and $N_{\mathrm{I}} = 6$ decoding iterations.*

To compare the complexity results with the results for equalisation and detection in the previous chapter, the given values need to be converted to operations per symbol period. For this, the values are multiplied by the factor $\beta = R \cdot N_{\mathrm{T}} \cdot M$, where R is the coding rate, N_{T} is the number of transmitter antennas and M is the modulation order.

3.7 Summary and conclusions

A maximum mutual information combining (MMIC) scheme was proposed as a means to improve the performance of turbo decoders whose component decoders use the Max-Log-MAP algorithm. The convergence behaviour of such turbo decoders was investigated by using extrinsic information transfer (EXIT) charts. The proposed combining scheme is achieved by iteration-specific scaling of the a priori information at the input of each component decoder in order to maximise the transfer of mutual information to the next component decoder, as suggested by the EXIT charts. The scaling corrects the accumulated bias introduced by the Max-Log approximation. A method for off-line computation of the optimum weight values was also described. It was shown that the proposed combining scheme significantly improves the performance of a turbo decoder using the Max-Log-MAP algorithm to within 0.05 dB of a turbo decoder using the optimum Log-MAP algorithm. The improved decoder retains the low complexity and insensitivity to input scaling which are inherent advantages of the Max-Log-MAP algorithm.

Chapter 4
MIMO receiver architectures

4.1 Introduction

In the previous chapters the receiver components such as the equaliser, detector and decoder were discussed separately. Now, all these components are combined for space-time reception in MIMO radio communication systems. In general, the receiver architectures can be classified in two different types: non-iterative and iterative receivers.

Currently, nearly all existing receivers for mobile communications fall in the category of non-iterative receivers like for GSM, UMTS, IS-95 and CDMA2000. Receivers which do not involve iterations between the various components, such as for instance the detector and the decoder, can be classified as non-iterative receivers. However, this does not necessarily imply that the components themselves are non-iterative in nature. The possibilities of non-iterative receivers are illustrated in Figure 4.1.

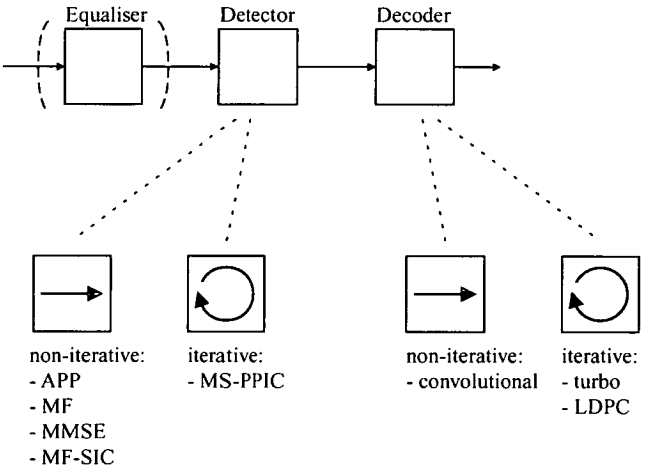


Figure 4.1: *Architecture overview for non-iterative receivers.*

The receiver consists of two essential components, the detector and the decoder. Optionally, the detector can be preceeded by a space-time equaliser, which eliminates the dispersion and

restores orthogonality between the spreading codes. This allows separate detection and decoding of data transmitted via each spreading code and therefore reduces the complexity of the detection algorithm significantly. As indicated above, the detector can use non-iterative techniques such as the *a posteriori* probability (APP) algorithm, simple matched filtering (MF), minimum mean-square error (MMSE) detection or successive interference cancellation. Iterative detectors include, for example, the multi-stage partial parallel interference canceller (MS-PPIC) which feeds back tentative estimates in order to improve the detector soft-output from iteration to iteration as presented in Chapter 2. For decoding, the same flexibility exists as for detection. Popular methods are convolutional codes as used in GSM, the iterative turbo codes which are employed in UMTS, or low-density parity check codes which allow decoders with very low complexity and therefore are very attractive for the downlink.

The receiver performance can further be improved by using the turbo principle. In the context of turbo codes it has been shown that iterations between two low complexity component decoders can outperform a single complex decoder and achieve performance close to the capacity limit [19]. The same principle can be used in iterative receivers as shown in Figure 4.2.

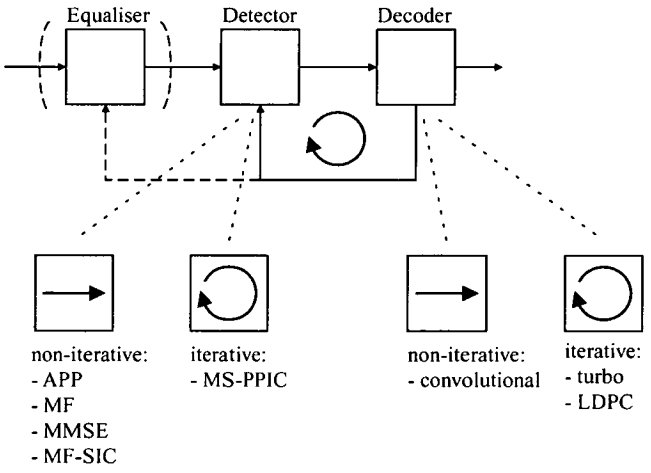


Figure 4.2: *Architecture overview for iterative receivers.*

For detectors whose operation depends on the quality of bit estimates in order to cancel interference, or can accept *a priori* information to improve the soft outputs, the feedback of information on the transmitted bits can improve the receiver performance significantly. Examples for interference cancelling detectors are MS-PPIC and MF-SIC, but also the APP and Max-Log-APP algorithms can benefit from *a priori* information. These more reliable soft outputs can be obtained from the decoding stage, where errors in the bit estimates are corrected. For this, the

decoder must supply soft outputs in form of LLRs not only on the information bits, but on all transmitted bits.

In this chapter, two architectures for non-iterative and iterative receivers are proposed and presented in detail. In the Sections 4.2 and 4.3, non-iterative receivers are investigated and their performance is compared under various channel conditions for 4-QAM modulation. The following Sections 4.4 and 4.5 concentrate on iterative receiver architectures. Here, a novel soft-output combining scheme is proposed for receivers based on hard-cancellation such as the MF-SIC, which maximises the mutual information transfer from iteration to iteration and improves the receiver convergence significantly. The performance of the iterative receivers is investigated under different channel conditions in Section 4.6. Finally, the complexity of the non-iterative and iterative receivers is investigated in Section 4.7.

4.2 Proposed non-iterative receiver architecture

In this section, a non-iterative receiver architecture is proposed for MIMO CDMA based on the components presented in the previous chapters. In Chapter 2, it was shown that space-time equalisation can significantly reduce the detector complexity. In dispersive channels, the equaliser removes the interference between the spreading codes so that the detector only needs to deal with the remaining interference from the different antennas in a MIMO system, and not also with the interference of data transmitted via other spreading codes. Different detector options were also discussed in Chapter 2, and turbo decoding was presented in Chapter 3.

Figure 4.3 depicts the proposed receiver architecture. The space-time equaliser mitigates the impact of the dispersive MIMO channel $\boldsymbol{H}$, and nominally restores the signal conditions to those present at the transmitter. Signal contributions from the various spreading codes are then retrieved via K de-spreading operations. The equalisation and de-spreading processes inevitably result in different spatio-temporal colouring of the noise vector $\underline{n}$ along each of the K branches. This is accounted for via pre-whitening of the sufficient statistics prior to the detection process, which leads again to spatial interference between the antennas, for each spreading code.

The pre-whitened sufficient statistics at the input of the k^{th} detector over the t^{th} symbol interval

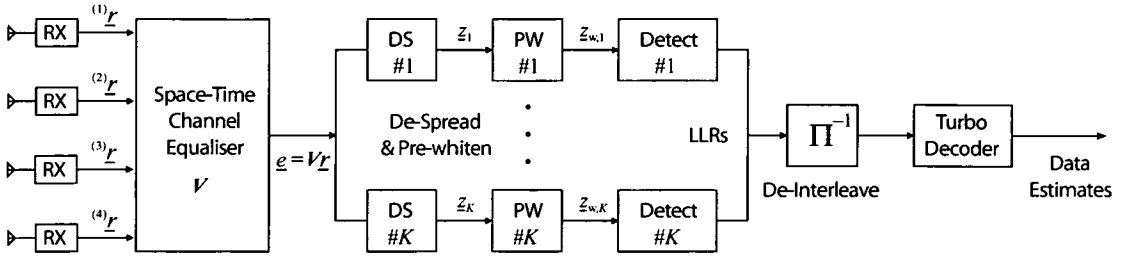


Figure 4.3: *Non-iterative MIMO receiver chain.*

may be written as (see Equation 2.15)

$$\underline{z}_{\mathrm{w},k}(t) = \boldsymbol{A}_k(t)\,\underline{x}_k(t) + \underline{\varepsilon}_k(t) \tag{4.1}$$

for $k = 1 \ldots K$ and $t = 1 \ldots N$. Vector $\underline{x}_k(t) \in C^{N_\mathrm{T} \times 1}$ consists of symbols transmitted via the k^{th} code during the t^{th} symbol epoch. $\boldsymbol{A}_k(t) \in C^{N_\mathrm{T} \times N_\mathrm{T}}$ denotes the corresponding linear transformation (a function of the channel matrix $\boldsymbol{H}$, equaliser matrix $\boldsymbol{V}$ and spreading matrix $\boldsymbol{C}_k$). Vector $\underline{\varepsilon}_k(t)$ contains contributions due to residual spatio-temporal self- and cross-code interference, intercell interference and thermal noise. As a result of pre-whitening, the elements of $\underline{\varepsilon}_k(t)$ are uncorrelated, i.e. $\mathrm{E}\{\underline{\varepsilon}_k(t)\,\underline{\varepsilon}_k^{\mathrm{H}}(t)\} = \boldsymbol{I}$. Note that for the following investigation full knowledge of the channel state information, as well as the noise power spectral density, N_0, is assumed. The sets of pre-whitened sufficient statistics $\underline{z}_{\mathrm{w},k}(t)$ $k = 1 \ldots K$ with $t = 1 \ldots N$ are then applied to the detector resulting in soft outputs in form of LLRs. The LLRs are then applied to a de-interleaver followed by a turbo decoder for error correction. The performance of the proposed receiver architecture is presented next.

4.3 Performance of non-iterative receivers

In this section, different versions of the proposed non-iterative receiver architecture are investigated using detection algorithms such as APP, Max-Log-APP and MS-PPIC. A MIMO system with $N_\mathrm{T} = N_\mathrm{R} = 4$ antennas and 4-QAM modulation is considered. The performance of the proposed non-iterative receiver architecture is investigated in different versions for three different channel scenarios. For decoding, an 8-state $1/3$ rate turbo decoder according to the HSDPA specification is used with 8 decoding iterations and a coding-block size of 1000 information bits. The results for $1/2$ rate are achieved via puncturing. The MS-PPIC detector operates with

6 iterations.

Version 1: Space-time equalisation and detection

Version 1 of the proposed receiver architecture is shown in Figure 4.3. As detection algorithms, APP, Max-Log-APP and MS-PPIC are employed.

Version 2: Detection only

The proposed architecture is also compared with a receiver architecture without an equaliser, as shown in Figure 4.4, which deals with interference from all N_{T} antennas and K spreading codes. Here, only the MS-PPIC algorithm is employed for detection, since the number of possible states for the APP and Max-Log-APP algorithm is extremely high, resulting in a prohibitive computational complexity.

Version 3: Space-time equalisation only

As Version 3, space-time equalisation followed by de-spreading and prewhitening only is considered in a receiver as shown in Figure 4.5.

Furthermore, the performance of the above receiver versions are investigated in the following 3 radio scenarios.

Scenario 1: Ergodic Channel

Here, a 1-tap ergodic MIMO channel is investigated. A non CDMA transmission ($K = 1, Q = 1$) is considered. To compare the detection algorithms without the influence of the equaliser, and since the ergodic channel is non-dispersive, all receivers are simulated in Version 2 for this scenario. The receiver performance is also compared with the MIMO channel capacity limit which can be calculated as shown in Appendix A.4. Figures 4.6 and 4.7 show the BER and FER performance results for Scenario 1 with the coding rates 1/3 and 1/2. As expected, the optimum APP algorithm performs best for both rates, at 2.7 dB from the capacity limit of -5.7 dB for rate 1/3, and at 3.2 dB from the capacity limit of -4.8 dB for rate 1/2. It is also shown that, when the Max-Log approximation is employed for the APP algorithm, the soft-output quality of the detector degrades, which leads to a decreased performance after decoding. It is evident, that at low E_{b}/N_0 (rate 1/3) the impact of the approximation is higher than at high values of E_{b}/N_0. The results also show that the MS-PPIC algorithm offers similar performance as the APP based algorithms. For low ranges of E_{b}/N_0 with coding rate 1/3, the MS-PPIC is even able to outperform the Max-Log-APP algorithm and performs close to the optimum APP algorithm.

Scenario 2: Flat fading channel

In this scenario, a 1-tap flat fading channel is considered. Data is spread with a spreading factor of $Q = 16$ and the same set of $K = 16$ orthogonal Walsh spreading codes are simultaneously transmitted from each antenna resulting in a high-interference code re-use scenario. Here, additionally to the receivers of Version 1, the receivers based on interference cancellation are also simulated without equalisation as receiver Version 2. Finally, to investigate the performance with space-time equalisation only, the receiver Version 3 is simulated as well. The simulation results for Scenario 2 are shown in Figure 4.8 for coding rate 1/3 and in Figure 4.9 for coding rate 1/2. For coding rate 1/3 the APP based receiver performs best, followed by the two MS-PPIC based receivers with and without equalisation, which slightly outperform the Max-Log-APP based receiver. When LLRs are calculated at the output of the equaliser after de-spreading and pre-whitening, without any further detection algorithm (Version 3), the performance drops by 1 dB at the FER of interest at 10^{-1}. As observed in the ergodic simulations, for rate 1/2 coding the Max-Log-APP receiver performs very close to the APP based receiver. This results from the operation at higher E_b/N_0, where the soft-output quality is less important since fewer errors need to be corrected by the decoder. Here, the two MS-PPIC based algorithms cannot match the APP performance and perform 0.7 dB worse. When operating at higher E_b/N_0, the MS-PPIC degrades slightly in comparison to the APP algorithm as shown previously in the detector comparison in Chapter 2. When only de-spreading is employed, the performance drops by 2 dB in comparison to the APP based receiver (Figure 4.9). This demonstrates that for flat fading channels, the detection after equalisation is essential in order to achieve good performance. The performance difference between receivers with and without detector increases at higher coding rates such as 1/2.

Scenario 3: TU channel

Here, a typical urban (TU) scenario with 10 chip-spaced taps is considered. Again, the data is spread with a spreading factor of $Q = 16$ and the same set of $K = 16$ orthogonal Walsh spreading codes are simultaneously transmitted from each antenna. As in Scenario 2, the receivers are simulated in all 3 versions. The performance of the receivers for Scenario 3 is depicted in Figure 4.10 for rate 1/3 coding and in Figure 4.11 for rate 1/2 coding. In comparison to the flat fading results (Scenario 2), the performance improves for the TU channel, since the diversity gain from the multiple taps is higher than the impact of the resulting interference caused by dispersion and lost orthogonality of the received signatures. For coding rate 1/3, all receivers perform well. The APP and the MS-PPIC receiver with equalisation of-

fer the best performance. Even with de-spreading only, the performance loss is only 0.3 dB. For coding rate 1/2, APP, Max-Log-APP and MS-PPIC based receiver offer nearly equivalent performance. This confirms that the MS-PPIC can be used as a low complexity alternative to the APP based algorithms without sacrificing performance. When equalisation followed by de-spreading and pre-whitening is performed without detection (Version 3), the performance drops by 0.6 dB. When equalisation is not performed, the receiver has to deal not only with inter-antenna interference but also with inter-code interference. Therefore, the MS-PPIC based receiver of Version 2 offers the worst performance, 1.5 dB poorer than the APP results. This shows the importance of an equaliser for channels with high dispersion.

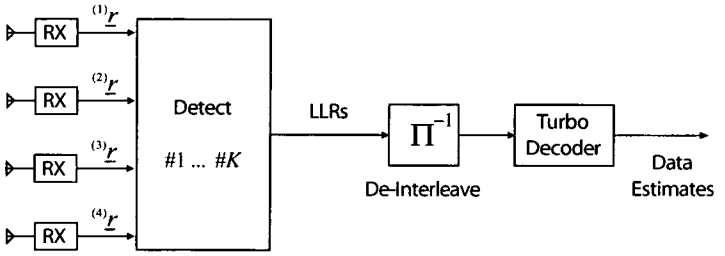


Figure 4.4: *Architecture overview of Version 2 of the non-iterative receivers.*

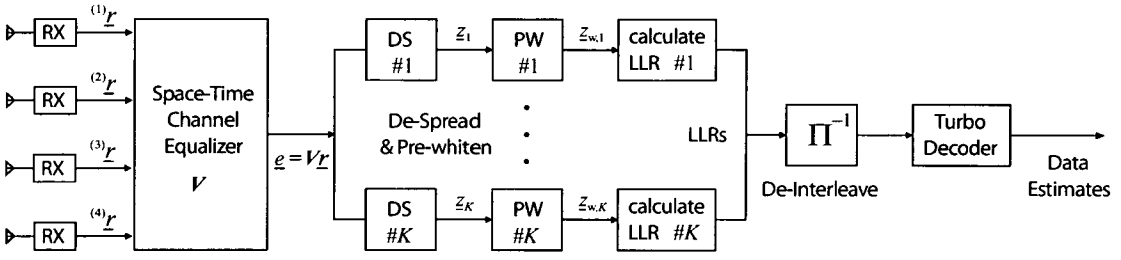


Figure 4.5: *Architecture overview of Version 3 of the non-iterative receivers.*

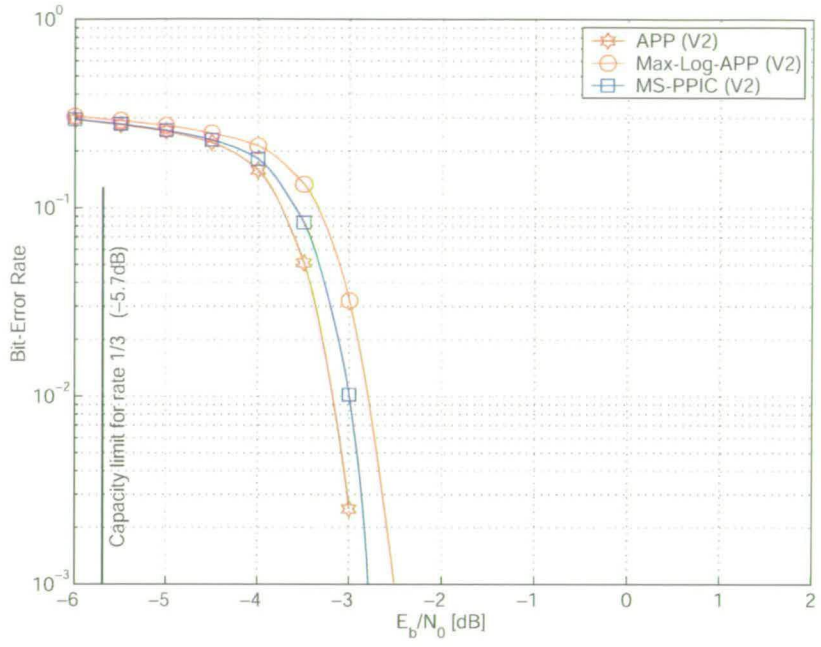


Figure 4.6: *Performance comparison for Scenario 1: ergodic channel, 4x4 antennas, 4-QAM modulation, rate 1/3 turbo coding, no equaliser, Q=1, K=1.*

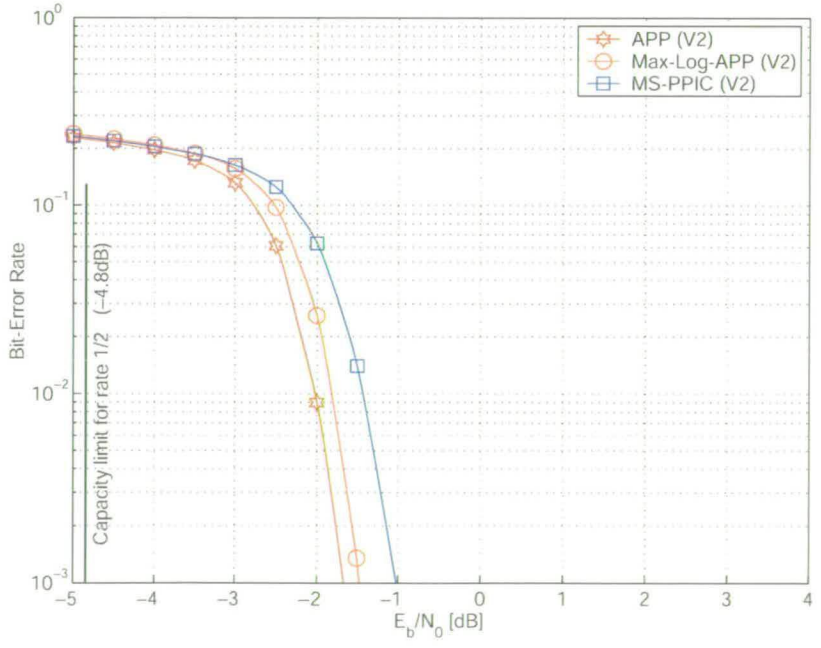


Figure 4.7: *Performance comparison for Scenario 1: ergodic channel, 4x4 antennas, 4-QAM modulation, rate 1/2 turbo coding, no equaliser, Q=1, K=1.*

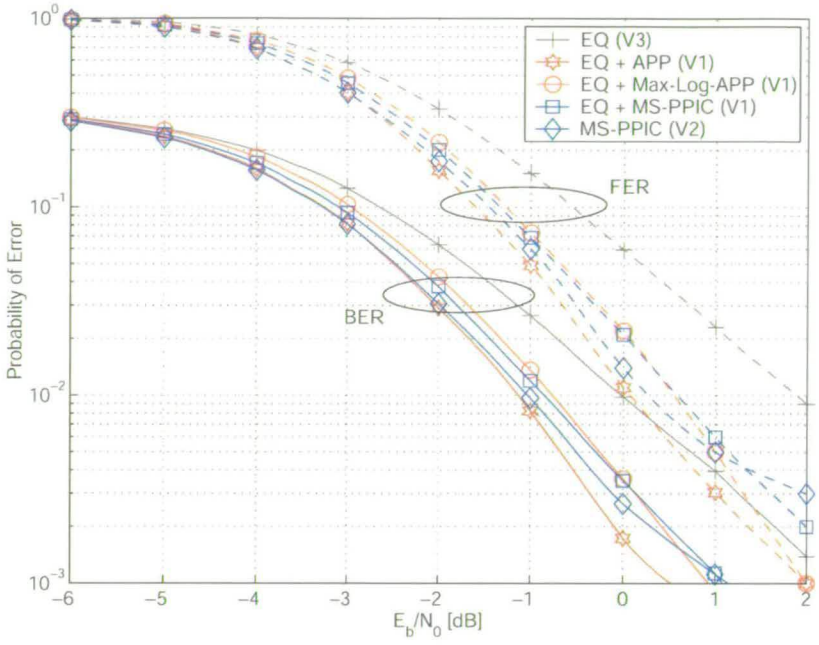


Figure 4.8: *Performance comparison for Scenario 2: 1 tap channel, 4x4 antennas, 4-QAM modulation, rate 1/3 turbo coding, Q=16, K=16.*

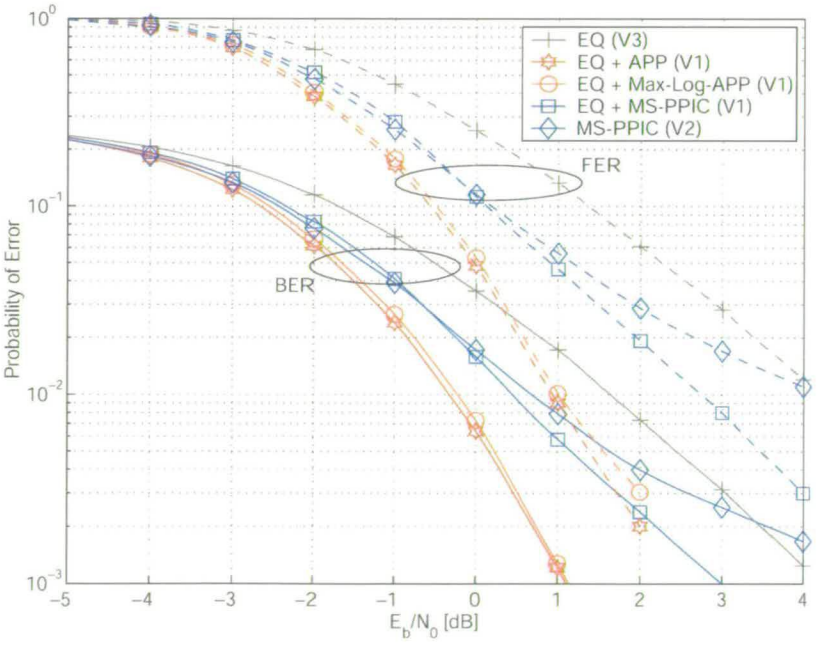


Figure 4.9: *Performance comparison for Scenario 2: 1 tap channel, 4x4 antennas, 4-QAM modulation, rate 1/2 turbo coding, Q=16, K=16.*

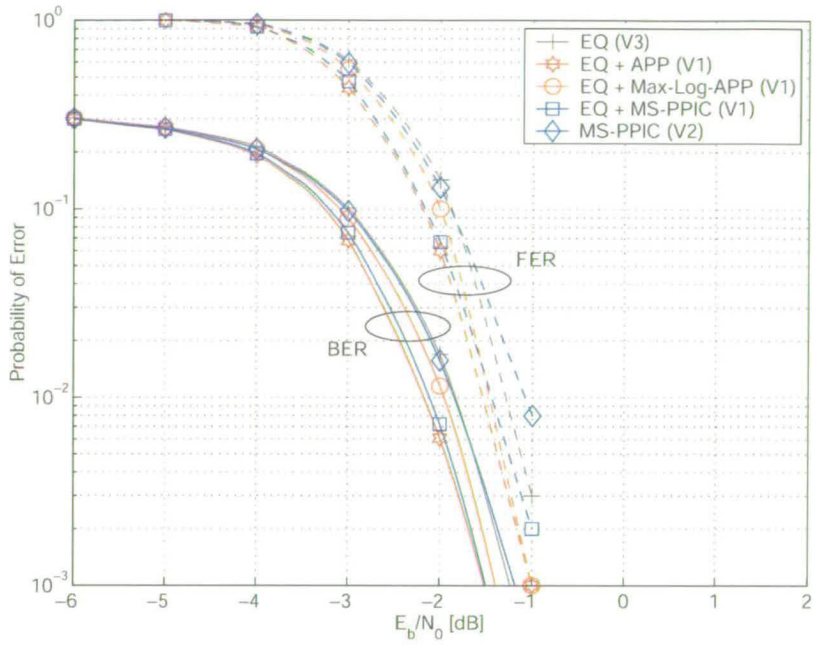


Figure 4.10: *Performance comparison for Scenario 3: TU channel, 4x4 antennas, 4-QAM modulation, rate 1/3 turbo coding, Q=16, K=16.*

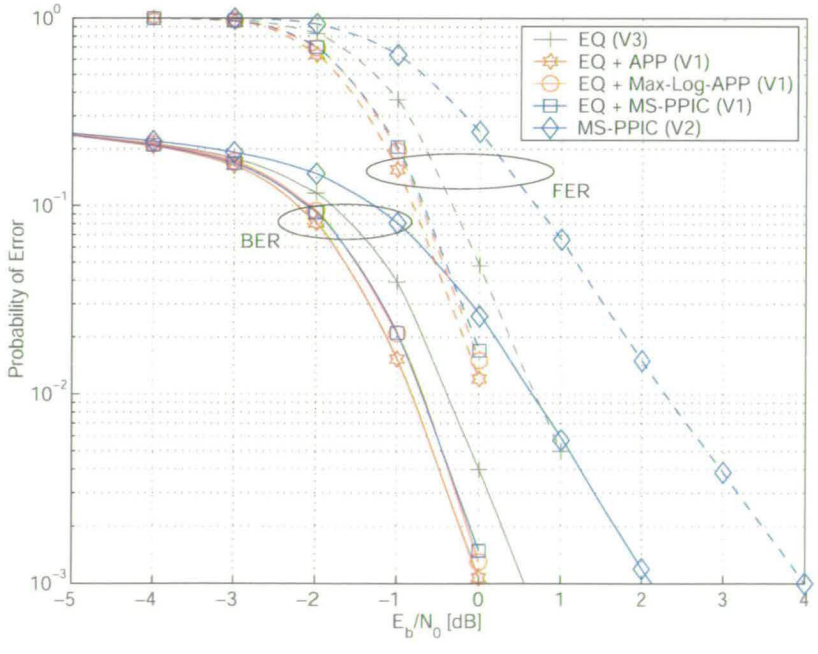


Figure 4.11: *Performance comparison for Scenario 3: TU channel, 4x4 antennas, 4-QAM modulation, rate 1/2 turbo coding, Q=16, K=16.*

4.4 Proposed iterative receiver architecture

The receiver performance can be further improved when reliable *a priori* information is available at the detection stage. In the architecture presented in the previous section, this *a priori* information can be provided by the decoder. In this way, the receiver can be modified to a iterative architecture as shown in Figure 4.12.

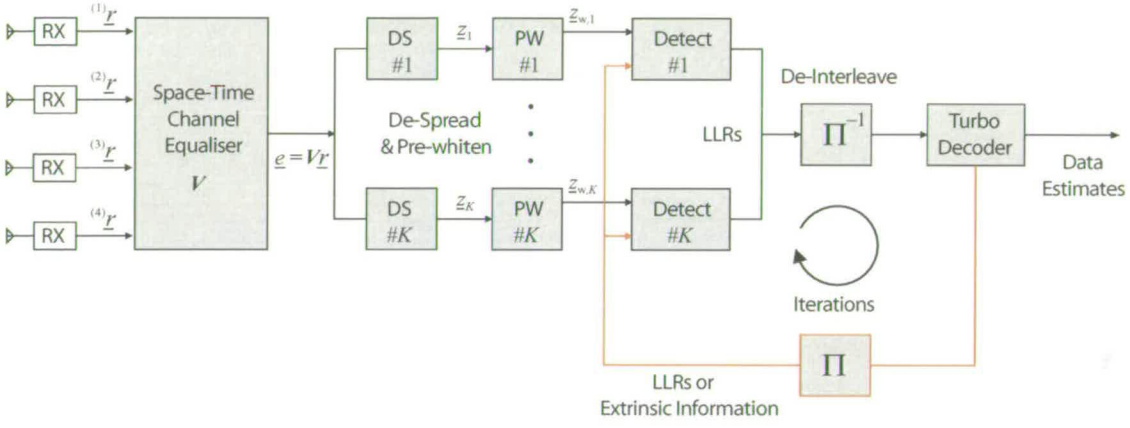


Figure 4.12: *Iterative MIMO receiver chain.*

The decoder produces soft-outputs in form of LLRs for all transmitted bits. This information is then interleaved and fed back to the detection stage. This feedback can also be used to improve the channel estimates or the equaliser parameters but, since in this investigation perfect channel knowledge is assumed, this is not considered here. When reliable *a priori* information is available, the detector produces more reliable soft outputs which result again in less errors after decoding. This process can be repeated for several iterations in order to improve the soft output quality from iteration to iteration and to converge to reliable estimates.

In the context of iterative receivers, often the question arises as to whether extrinsic information or LLRs need to be fed back as *a priori* information to the detector. This depends on the detector type and how the *a priori* information is used in the detection process.

The most common technique is to feed back extrinsic information, in other words only the additional information which was gained by the decoding process, for the use in an APP or Max-Log-APP detector [28]. Here the *a priori* information is used to calculate the logarithmic probabilities $\ln \mathrm{P}(\underline{x})$ from Equation 2.27 for the APP and Equation 2.29 for the Max-Log-APP. This results in a more reliable detector output, which again improves the reliability of the

decoder output. This process is repeated for several iterations until the receiver converges.

When MS-PPIC detection is employed, the *a priori* information is used in a different way, for partial interference cancellation. In the first iteration, the MS-PPIC detector operates as usual using multiple stages to converge to a reliable first estimate. The detector output is then de-interleaved and decoded to correct errors in the detector estimates. This represents the first iteration. In all following iterations, instead of the detector output, the more reliable LLRs obtained from the soft-output decoder are used for one cancellation stage per iteration. Unlike for the APP detector, the MS-PPIC requires LLRs and not extrinsic information, since the cancellation should be based on the most reliable bit estimates and not on only the gained information of the decoding process.

When the receiver is based on the MF-SIC detector, the *a priori* information is used for two purposes: Firstly, for interference cancellation and secondly, for ordering of the cancellation. As for MS-PPIC detection, LLRs are used as *a priori* information, since the ordering and cancellation should be based on the most reliable bit estimates. The order metric is of critical importance for high performance in any form of successive cancellation process, as described in Chapter 2. In the first iteration the detector operates at symbol level. The cancellation is based on the detector estimates, ordered in decreasing order of symbol energy (see Chapter 2). After detection, the bit estimates are de-interleaved and decoded. This completes the first iteration. In all following iterations of the receiver, the MF-SIC has access to reliability information at a bit level, in the form of LLRs, generated by the soft-output decoder in the previous iteration. As a result, at each symbol interval t, ordering can be performed at a bit level (rather than symbol level) based on the LLRs, which represents an optimum ordering policy. The bit $b_{k,m}^{(i)}(t)$ with the largest LLR value $\left|\Lambda^{(n-1)}\left(b_{k,m}^{(i)}(t)\right)\right|$ (or minimum estimation error probability) based on the most recent decoder output from the $(n-1)^{\text{th}}$ iteration, can be selected as most reliable. The cancellation process at the t^{th} symbol interval is based on the more reliable hard bit estimates derived from the LLR values. The process is again repeated for the next most reliable bit. After the MF-SIC detection of a complete code-block, the soft-outputs are again multiplexed into a single stream for de-interleaving and decoding. The performance of the MF-SIC (and hence the receiver) improves at each iteration as the quality of the decoder output improves. This is depicted in Figure 4.13 for a TU channel scenario with 10 chip-spaced taps, $N_{\text{T}} = N_{\text{T}} = 4$ antennas, spreading factor $Q = 16$ and $K = 16$ spreading codes reused across the transmitter antennas. The data is 4-QAM modulated and a rate 1/2 (punctured) turbo

code with the polynomials $(G_r, G) = \left(1 + D + D^2, 1 + D^2\right)$ and Log-Map decoding with 8 iterations is employed. It is shown that the MF-SIC based receiver starts with poor performance in iteration 1, but its performance improves significantly in the following iterations.

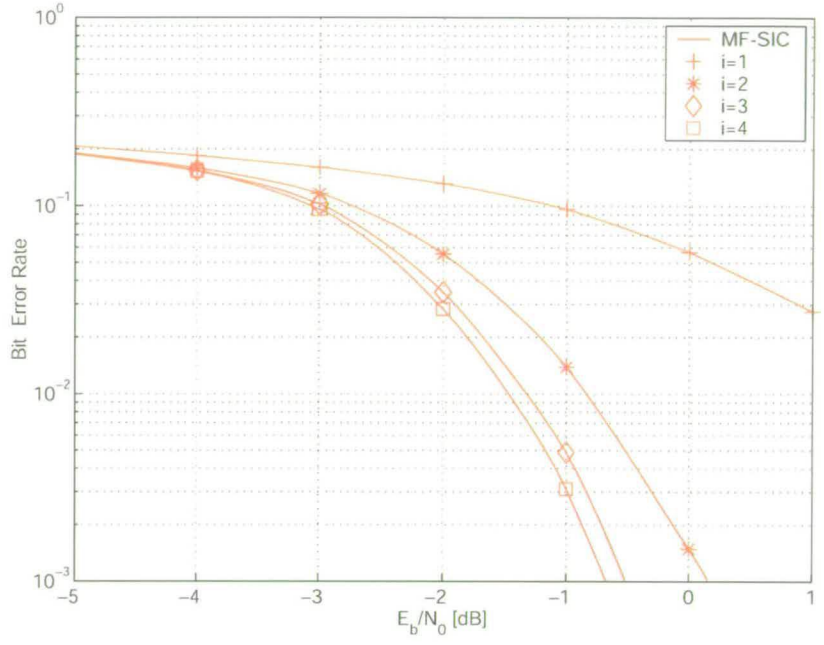


Figure 4.13: *Performance improvement from iteration to iteration for an MF-SIC based receiver (TU channel, 4x4 antennas, 4-QAM modulation, rate 1/2 turbo coding, Q=16, K=16).*

4.5 Improved iterative receiver using soft-output combining

When iterative receivers are based on hard-cancellation of interference, such as is the case when MF-SIC detection is employed, erroneously detected bits can lead to error propagation due to incorrectly cancelled bits. The different detection and cancellation order in each iteration also leads to different error propagation and thereby to a certain degree of independence in the bit estimates from iteration to iteration. Therefore, the performance can be improved by combining the soft-outputs of the detector from the current iteration with the values of the previous iteration as proposed in [29] for a MF-SIC based receiver and [30, 31] for a receiver based on parallel interference cancellation. The MS-PPIC cannot benefit from this scheme, since the optimum partial cancellation is employed which already minimises these error propagation effects. The modification of the receiver is shown in Figure 4.14.

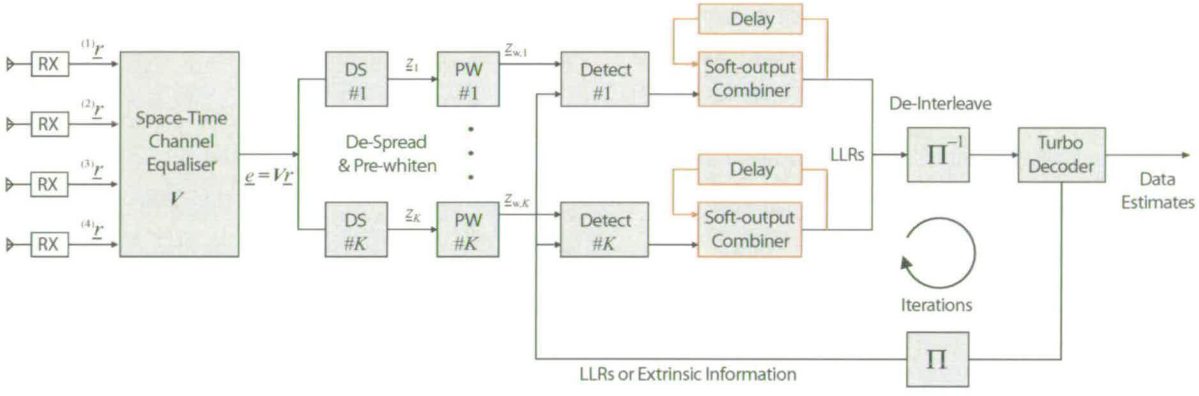


Figure 4.14: *Iterative MIMO receiver chain with soft-output combining.*

In the second and all following iterations, the LLRs at the detector output are combined with the corresponding values from the previous iteration. In this operation, the combining ratio is of paramount importance for the receiver performance. If the ratio between current and previous LLRs is too low, the combining method shows only little effect. On the other hand, a too high combining ratio results in slow convergence of the iterative receiver. In addition, the calculation of the combining factors should be of low computational complexity. In this section, a novel method for the off-line calculation of the combining factors is proposed which maximises the mutual information between the LLRs after combining and the corresponding bits. This has the advantage that in normal operation it is not necessary to calculate and update the combining factors. The proposed maximum mutual information combining (MMIC) can be

compared with decision statistics combining (DSC), as suggested in the literature [30, 31] for parallel interference cancellation. In the latter case, the combining factors are calculated online based on the variances of the LLRs of the current and previous iteration.

The combining operation may be written as

$$\Lambda^{(n)}\left(b_{k,m}^{(i)}(t)\right) := w_1^{(n)}\Lambda^{(n)}\left(b_{k,m}^{(i)}(t)\right) + w_2^{(n)}\Lambda^{(n-1)}\left(b_{k,m}^{(i)}(t)\right) \tag{4.2}$$

where $\Lambda^{(n)}\left(b_{k,m}^{(i)}(t)\right)$ represents the LLR at the output of the detector or at the n^{th} iteration corresponding to the bit $b_{k,m}^{(i)}(t) \in \{1, -1\}$ transmitted via the k^{th} spreading code from the i^{th} antenna and as the m^{th} bit of the modulation scheme. The proportions of the weight factors $w_1^{(n)}$ and $w_2^{(n)}$ have a significant influence on the performance and speed of convergence of the iterative receiver.

4.5.1 Proposed maximum mutual information combining (MMIC)

The weight factors $w_1^{(n)}$ and $w_2^{(n)}$ for each iteration n can be optimised by off-line maximisation of the mutual information between the combined term $\Lambda^{(n)}\left(b_{k,m}^{(i)}(t)\right)$ and the corresponding bits $b_{k,m}^{(i)}(t)$ in each combining step. Using vector notation

$$\begin{aligned}\Lambda^{(n)}\left(b_{k,m}^{(i)}(t)\right) &= \begin{bmatrix} w_1^{(n)} & w_2^{(n)} \end{bmatrix} \begin{bmatrix} \Lambda^{(n)}\left(b_{k,m}^{(i)}(t)\right) \\ \Lambda^{(n-1)}\left(b_{k,m}^{(i)}(t)\right) \end{bmatrix} = \underline{w}^{(n)\mathrm{T}}\underline{\Lambda}^{(n)}(t) \\ &= \underline{w}^{(n)\mathrm{T}}\left(\underline{\lambda}^{(n)}(t) + \underline{\epsilon}^{(n)}(t)\right) = \underline{w}^{(n)\mathrm{T}}\underline{\lambda}^{(n)}(t) + v^{(n)}(t)\end{aligned} \tag{4.3}$$

where $\underline{\epsilon}^{(n)}(t)$ represents the contributions of channel noise plus the numerical approximation error at the detector output, and $\underline{\lambda}^{(n)}(t)$ is the vector of "uncorrupted" LLRs. Now we end up with the same combining problem solved in Chapter 3 in the context of improved Max-Log-MAP turbo decoding with the solution

$$\underline{w}_{\text{OPT}}^{(n)} = k\left(\boldsymbol{R}_{\epsilon}^{(n)}\right)^{-\mathrm{T}/2}\text{eig}_{\max}\left\{\left(\boldsymbol{R}_{\epsilon}^{(n)}\right)^{-1/2}\boldsymbol{R}_{\lambda+\epsilon}^{(n)}\left(\boldsymbol{R}_{\epsilon}^{(n)}\right)^{-\mathrm{T}/2}\right\} \tag{4.4}$$

where $\text{eig}_{\max}\{\boldsymbol{A}\}$ is the eigenvector of $\boldsymbol{A}$ corresponding to its largest eigenvalue. The scalar k is chosen such that $w_1^{(n)} + w_2^{(n)} = 1$. Here, the optimum weights are functions of the iteration index, the performance of the detector and the decoder (i.e. detector type, encoder polynomials) and the signal to noise ratio. The optimum weights $\underline{w}_{\text{OPT}}^{(n)}$ can be computed or "trained" off-

line based on time-averaged estimates of correlation matrices $\boldsymbol{R}_{\lambda+\epsilon}^{(n)}$ and $\boldsymbol{R}_{\epsilon}^{(n)}$ derived over a sufficiently long data block corresponding to τ encoded information bits. Specifically

$$\boldsymbol{R}_{\lambda+\epsilon}^{(n)} = \mathrm{E}\left\{\underline{\Lambda}^{(n)}(t)\,\underline{\Lambda}^{(n)\mathrm{T}}(t)\right\} = \lim_{\tau\to\infty}\frac{1}{\tau}\sum_{t=1}^{\tau}\underline{\Lambda}^{(n)}(t)\,\underline{\Lambda}^{(n)\mathrm{T}}(t)\,. \tag{4.5}$$

Furthermore, the vector $\underline{\lambda}^{(n)}(t)$ of "uncorrupted" LLRs may be modelled as as scaled versions of the transmitted bits $b_{k,m}^{(i)}(t) \in \{1,-1\}$ using the expected absolute value of the corresponding LLRs as amplitude:

$$\underline{\lambda}^{(n)}(t) = \begin{bmatrix} \mathrm{E}\{\Lambda^{(n)}\left(b_{k,m}^{(i)}(t)\right)\cdot b_{k,m}^{(i)}(t)\}b_{k,m}^{(i)}(t) \\ \mathrm{E}\{\Lambda^{(n-1)}\left(b_{k,m}^{(i)}(t)\right)\cdot b_{k,m}^{(i)}(t)\}b_{k,m}^{(i)}(t) \end{bmatrix} \tag{4.6}$$

so that

$$\boldsymbol{R}_{\lambda}^{(n)} = \mathrm{E}\left\{\underline{\lambda}^{(n)}(t)\,\underline{\lambda}^{(n)\mathrm{T}}(t)\right\} = \lim_{\tau\to\infty}\begin{bmatrix} \left(\phi^{(n)}\right)^2 & \phi^{(n)}\theta^{(n)} \\ \theta^{(n)}\phi^{(n)} & \left(\theta^{(n)}\right)^2 \end{bmatrix} \tag{4.7}$$

where $\phi^{(n)} = \frac{1}{\tau}\sum_{t=1}^{\tau}\Lambda^{(n)}\left(b_{k,m}^{(i)}(t)\right)\cdot b_{k,m}^{(i)}(t)$ and $\theta^{(n)} = \frac{1}{\tau}\sum_{t=1}^{\tau}\Lambda^{(n-1)}\left(b_{k,m}^{(i)}(t)\right)\cdot b_{k,m}^{(i)}(t)$. Finally, assuming that vectors $\underline{\epsilon}^{(n)}(t)$ and $\underline{\lambda}^{(n)}(t)$ are uncorrelated, one may derive $\boldsymbol{R}_{\epsilon}^{(n)}$ as $\boldsymbol{R}_{\lambda+\epsilon}^{(n)} - \boldsymbol{R}_{\lambda}^{(n)}$. The above training procedure should be performed under E_b/N_0 conditions that are typical at the bit-error rate range of interest.

4.5.2 Decision statistics combining (DSC)

Alternatively, the combining weights can be calculated online based on the variances of the LLRs of the current and previous iteration as proposed in the literature [30, 31] in context of parallel interference cancellation. Here it is assumed that in the first iterations the LLRs at the input of the combiner are weakly correlated and the weights are calculated in a way similar to receive diversity maximum ratio combining. Then, the weights for the n^{th} iteration can be calculated as

$$w_1^{(n)} = \frac{\left(\sigma^{(n-1)}\right)^2}{\left(\sigma^{(n-1)}\right)^2 + \left(\sigma^{(n)}\right)^2} \tag{4.8}$$

$$w_2^{(n)} = \frac{\left(\sigma^{(n)}\right)^2}{\left(\sigma^{(n)}\right)^2 + \left(\sigma^{(n-1)}\right)^2} \tag{4.9}$$

where $\left(\sigma^{(n)}\right)^2$ and $\left(\sigma^{(n-1)}\right)^2$ are the estimated variances of the LLRs at the detector output of the current and previous iteration.

4.5.3 Performance comparison (MMIC vs. DSC)

In this subsection, the performance of soft-output combining based on the proposed maximum mutual information criterion (MMIC) is compared with decision statistics combining (DSC) for the iterative receiver of Figure 4.14, based on MF-SIC detection. A total number of 4 receiver iterations are performed between the detector and the decoder. A system with $N_{\mathrm{T}} = N_{\mathrm{T}} = 4$ is considered for two different channel scenarios. Instead of the HSDPA turbo decoder, a low complexity 4-state $1/3$ rate turbo decoder punctured to rate $1/2$ based on the Log-MAP algorithm is used with 8 decoding iterations and a coding-block size of 1000 information bits.

Scenario 1: Ergodic Channel

A 1-tap ergodic MIMO channel is considered for Scenario 1. No CDMA ($K = 1, Q = 1$) and no space-time equalisation is employed. Therefore, no pre-whitening is required. Figure 4.15 presents the simulation results for Scenario 1. The results show that under this conditions MMIC performs best, 0.3 dB better than DSC. It is also shown that the MMIC technique can improve the performance in comparison to a receiver without combining by 1 dB at a FER of 10^{-1}.

Scenario 2: TU channel

For the second scenario, a typical urban (TU) channel with 10 chip-spaced taps is considered. Here the data is spread with a spreading factor of $Q = 16$, the same set of $K = 16$ orthogonal Walsh spreading codes is re-used across the transmitter antennas and space-time equalisation is employed. For this scenario, the performance comparison is shown in Figure 4.16. Here, the performance differences between the three receivers is far less pronounced. MMIC improves the performance by only 0.1 dB in comparison to DSC and a receiver without combining.

MMIC will be used as the default combining method in all following iterative receivers based on MF-SIC detection.

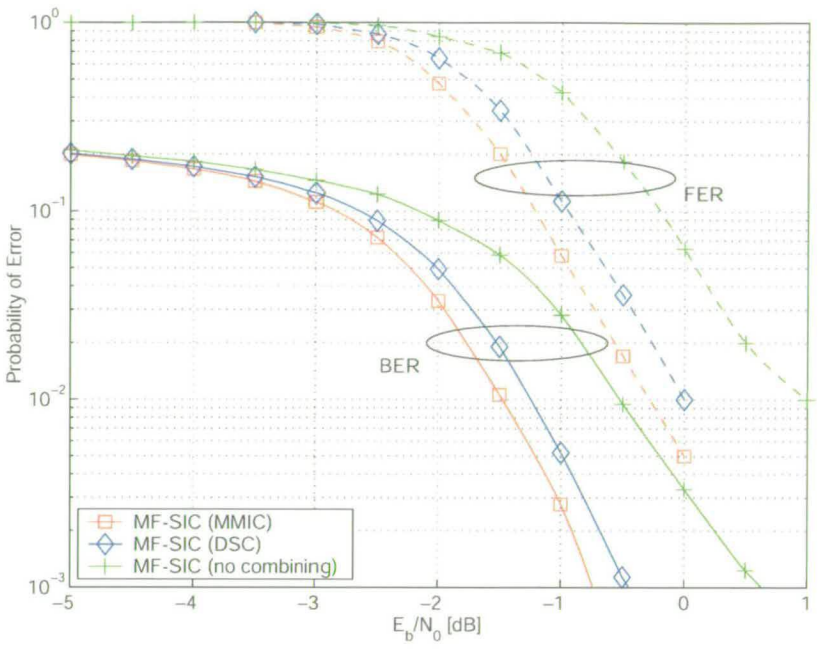


Figure 4.15: *Performance comparison of iterative MF-SIC receivers with MMIC and DSC for Scenario 1 with ergodic channel and rate 1/2 turbo coding.*

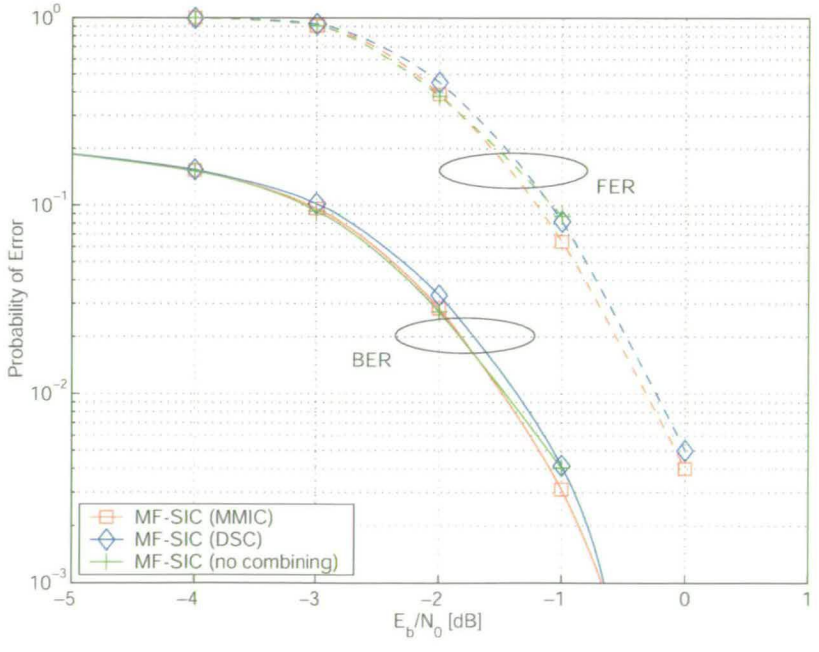


Figure 4.16: *Performance comparison of iterative MF-SIC receivers with MMIC and DSC for Scenario 2 with TU channel and rate 1/2 turbo coding.*

4.6 Performance of iterative receivers for 4-QAM

In this section, the performance of the proposed iterative receiver architectures is investigated using Max-Log-APP, MS-PPIC and MF-SIC as detection algorithms. A total number of 4 receiver iterations are performed between the detector and the decoder. The MF-SIC based receivers employ the proposed soft-output combining technique based on the maximum mutual information criterion. Instead of the HSDPA turbo decoder, a low complexity 4-state $1/3$ rate turbo decoder based on the Log-MAP algorithm is used with 8 decoding iterations and a code-block size of 1000 information bits. The results for $1/2$ rate are achieved via puncturing. The MS-PPIC detector operates with 6 iterations. Here, the iterative receiver architectures are considered in 4 versions.

Version 1: Space-time equalisation and detection

Version 1 is the proposed iterative receiver architecture including space-time equalisation as shown in Figure 4.12. For detection, Max-Log-APP and MS-PPIC algorithms are employed.

Version 2: Space-time equalisation and detection with MMIC

When the iterative receiver is based on space-time equalisation and MF-SIC detection, maximum mutual information combining is employed in a receiver architecture as shown in Figure 4.14.

Version 3: Detection only

The proposed architectures are also compared with a receiver architecture without equaliser as shown in Figure 4.17, which deals with interference from all N_{T} antennas and K spreading codes. For dispersive channels, only the MS-PPIC algorithm is employed for detection, since the number of possible states for the Max-Log-APP algorithm is extremely high, resulting in a prohibitive computational complexity.

Version 4: Detection only with MMIC

For iterative receivers without space-time equalisation, based on MF-SIC detection with maximum mutual information combining, an architecture as shown in Figure 4.18 is employed.

The performance of the iterative receivers is investigated in 3 different channel scenarios. In all three cases, the channel conditions are assumed to be known at the receiver.

Scenario 1: Ergodic Channel

Here, a 1-tap ergodic MIMO channel is considered without CDMA ($K = 1$, $Q = 1$). To compare the detection algorithms without the influence of the equaliser, and since the ergodic channel is non-dispersive, the receivers based on Max-Log-MAP and MS-PPIC detection are simulated in Version 3 and the receiver based on MF-SIC detection is simulated in Version 4 in this scenario. Here, the receiver performance is also compared with the MIMO channel capacity limit which can be calculated as shown in Appendix A.4. Figures 4.19 and 4.20 show the BER performance for this scenario. For rate 1/3 coding, the iterative receiver based on the Max-Log-MAP detector can gain most from the *a priori* information and shows the best performance. It is demonstrated that, when reliable *a priori* information is available the iterative MF-SIC receiver performs very well and offers the same performance as the MS-PPIC receiver. The receiver based on the Max-Log-APP algorithm performs best, only 1.6 dB from the channel capacity limit of -5.7 dB. The simulations with coding rate 1/2, presented in Figure 4.20, show the same trend as for non-iterative receivers in Figure 4.7: the receivers based on interference cancellation such as MS-PPIC and MF-SIC lose performance in comparion to the Max-Log-APP detector due to the operation at higher E_b/N_0. The Max-Log-APP based receiver performs best, at only 1.9 dB from the channel capacity limit of -4.8 dB. In iterative receivers, small performance differences of the detectors become larger from iteration to iteration. This results in the more pronounced difference in BER between the Max-Log-APP and MS-PPIC for iterative receivers in comparison to the corresponding results for non-iterative receivers. It is also shown that the iterative MF-SIC based receiver suffers more from the less reliable *a priori* information and offers the worst performance. In this scenario, the iterative receivers can outperform the non-iterative receivers by approximately 1 dB. Only the MF-SIC based receiver performance benefits significantly from the iterations and becomes competitive. This questions if the cost of the higher computational complexity, due to the 4 iterations, is reasonable for the resulting gain.

Scenario 2: Flat fading channel

A 1-tap flat fading channel is considered. As above for the non-iterative receiver simulations, a system with $N_T = N_R = 4$, $Q = 16$ and $K = 16$ is investigated. The receivers based on Max-Log-MAP and MS-PPIC detection are simulated in Version 1 and the receivers based on MF-SIC detection in Version 2. Additionally, the receivers which employ interference cancelling detectors are simulated without equalisation as receiver Version 3 with MS-PPIC detection and as Version 4 for MF-SIC detection. The performance results for this flat fading scenario are

shown in Figure 4.21 and Figure 4.22 for coding rate 1/3 and coding rate 1/2 respectively. For coding rate 1/3, the Max-Log-APP receiver performs best and outperforms the MS-PPIC and MF-SIC based receivers slightly by 0.1 to 0.5 dB at the target FER of 10^{-1}. In this scenario, the MS-PPIC and MF-SIC receivers without channel equalisation offer nearly equivalent performance to their counterparts with equaliser. For coding rate 1/2, the performance difference between Max-Log-APP and the interference cancelling receivers becomes more evident. The iterative Max-Log-APP receiver outperforms the MS-PPIC receiver by 0.9 dB and the MF-SIC receiver by 1.2 dB. Here, the receivers without an equaliser show a degraded performance in comparison to their counterparts with an equaliser: The MS-PPIC degrades by 0.7 dB and the MF-SIC by 2.2 dB at the FER of 10^{-1} when no equalisation is employed. In comparion to the non-iterative results in Figures 4.8 and 4.9 the iterative receivers can improve the performance by approximately 1 dB in flat fading.

Scenario 3: TU channel

Here, a typical urban (TU) scenario with 10 chip-spaced taps is considered. Again, a system with $N_{\mathrm{T}} = N_{\mathrm{R}} = 4$, $Q = 16$ and $K = 16$ is investigated. The receivers are simulated in the same versions as in Scenario 2. The performance comparisons for Scenario 3 are shown in Figure 4.23 for coding rate 1/3 and in Figure 4.24 for coding rate 1/2. For coding rate 1/3, all receivers offer nearly the same performance. The Max-Log-APP performs best followed by the MF-SIC and MS-PPIC at only 0.1 to 0.3 dB higher E_{b}/N_0 at the target FER of 10^{-1}. Even the receivers without equalisation perform equally well. For coding rate 1/2, this is not the case anymore. Without equalisation, the MS-PPIC performs 1.1 dB worse and the MF-SIC performs even 2.3 dB worse. This underlines the importance of equalisation, when higher coding rates are employed like rate 1/2. Here, the MS-PPIC and MF-SIC perform very well. The Max-Log-APP based receiver outperforms them by only 0.4 dB at the FER of 10^{-1}. This confirms that low complexity detectors based on interference cancellation can be used as a Max-Log-APP replacement in iterative receiver architectures without losing much performance in TU channel conditions.

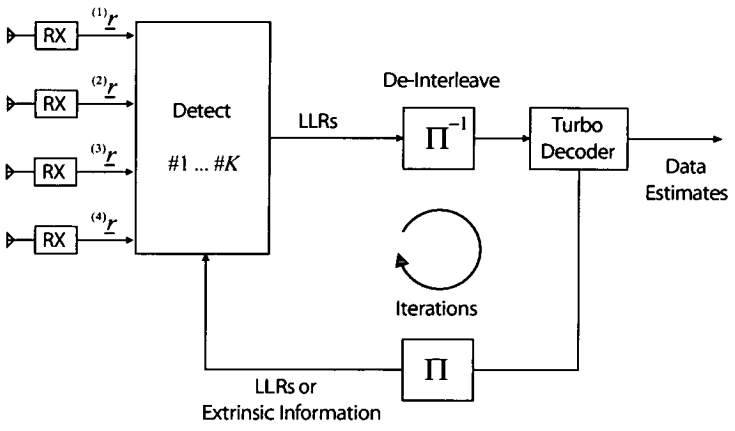


Figure 4.17: *Architecture overview of Version 3 of the iterative receivers.*

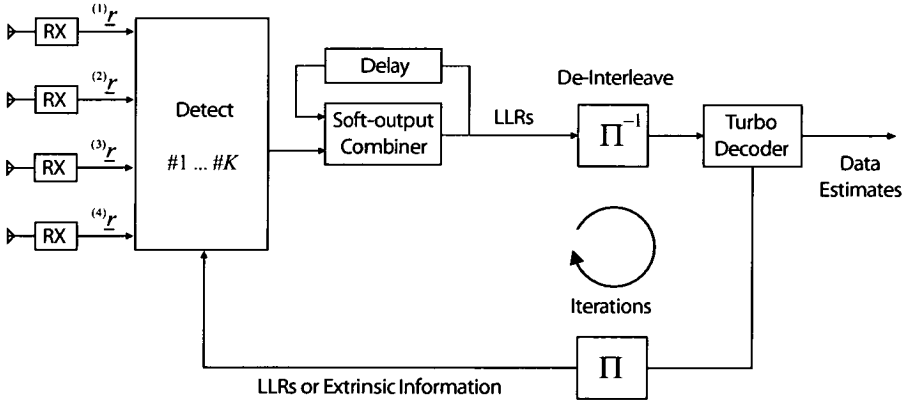


Figure 4.18: *Architecture overview of Version 4 of the iterative receivers.*

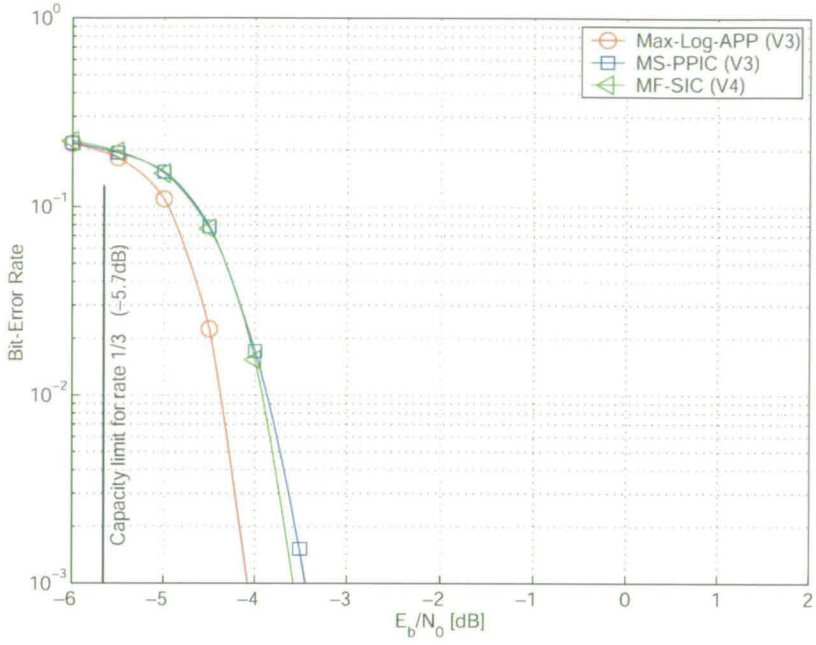


Figure 4.19: *Performance comparison: ergodic channel, 4x4 antennas, 4-QAM modulation, 4 iterations, rate 1/3 turbo coding, no equaliser, no CDMA.*

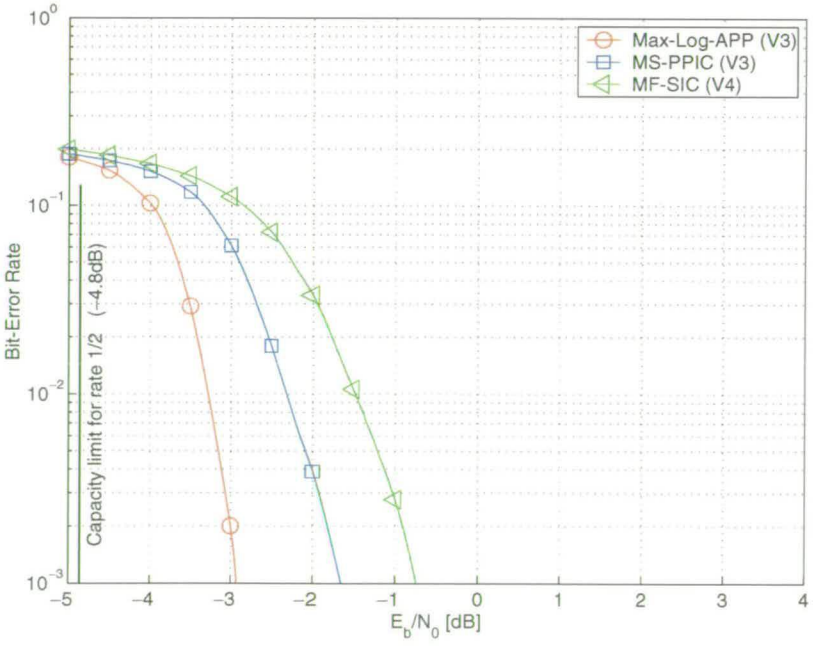


Figure 4.20: *Performance comparison: ergodic channel, 4x4 antennas, 4-QAM modulation, 4 iterations, rate 1/2 turbo coding, no equaliser, no CDMA.*

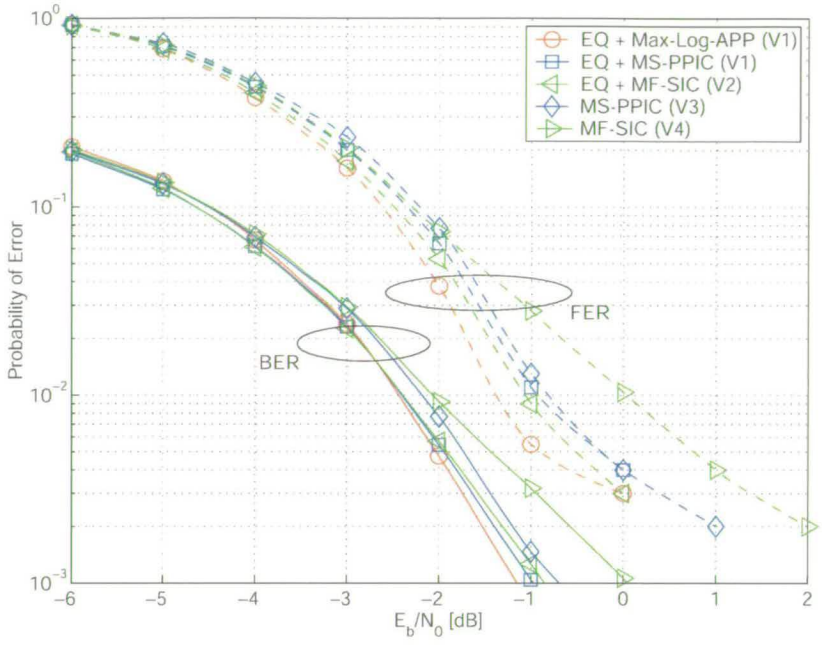


Figure 4.21: *Performance comparison: flat fading channel, 4x4 antennas, 4-QAM modulation, 4 iterations, rate 1/3 turbo coding.*

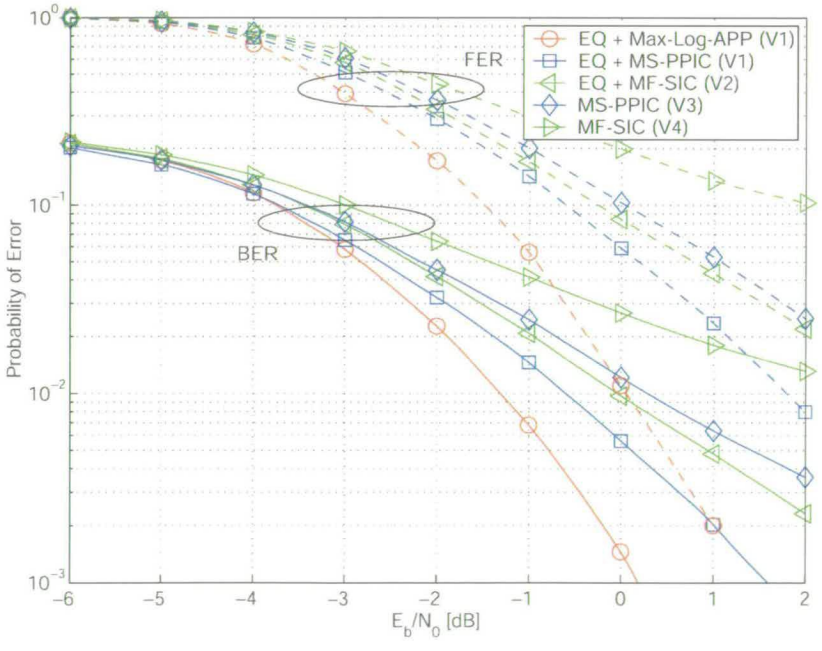


Figure 4.22: *Performance comparison: flat fading channel, 4x4 antennas, 4-QAM modulation, 4 iterations, rate 1/2 turbo coding.*

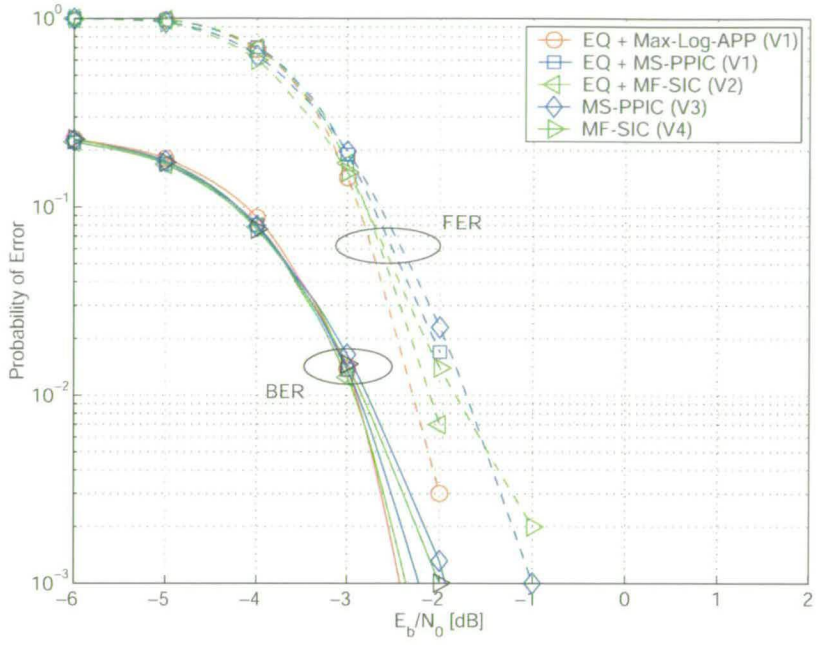


Figure 4.23: *Performance comparison: TU channel, 4x4 antennas, 4-QAM modulation, 4 iterations, rate 1/3 turbo coding.*

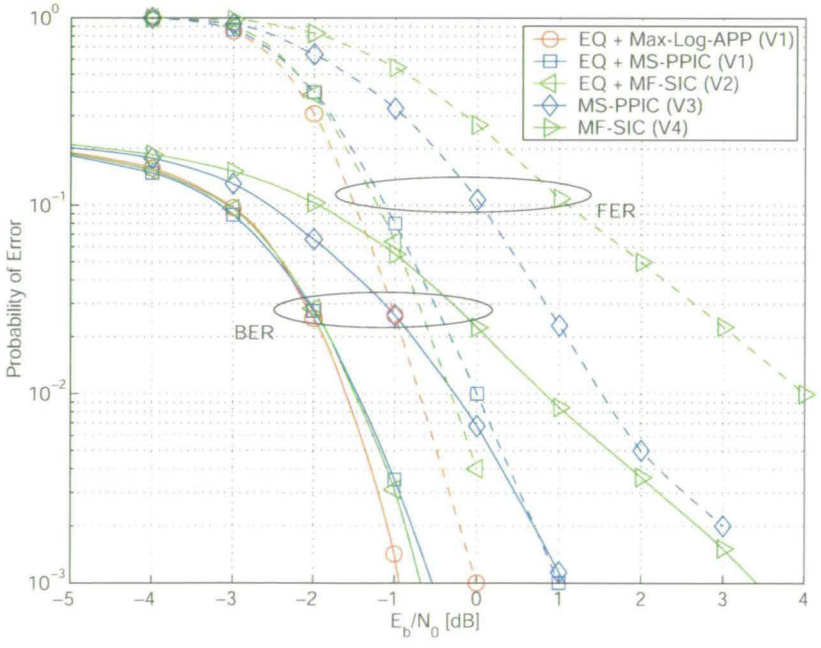


Figure 4.24: *Performance comparison: TU channel, 4x4 antennas, 4-QAM modulation, 4 iterations, rate 1/2 turbo coding.*

4.7 Receiver complexity comparison

A complexity comparison of the discussed non-iterative and iterative receivers is shown in Table 4.1 and Table 4.2 respectively. The results shown are first order estimates and are given in terms of operations per symbol period (multiplications, additions, ...) for normal operation. The complexity required for initialisation is neglected here. The table also allows a complexity comparison of the components and therefore gives an indication in which part of the receiver the greatest complexity can be found.

Non-iterative receivers

It is shown in Table 4.1 that the non-iterative receiver based on the optimum APP algorithm requires the most operations, and the detector is clearly the most complex receiver component. Using the Max-Log-APP algorithm instead reduces the detector complexity to only 70% of the APP detector. The best tradeoff between computational complexity and performance is offered by the MS-PPIC based receiver. The MS-PPIC detector requires only 10% of the APP complexity and offers nearly the same performance. This reduces the overall receiver complexity by approximately 50% and leaves the turbo decoder as component with the highest complexity. If even lower complexity is required, the most sensible step would therefore be to reduce the decoding complexity, for example by reducing the number of decoding iterations. It is also confirmed that the use of an equaliser is not only advantageous in terms of performance, but also reduces the required number of operations significantly.

Iterative receivers

The complexity comparison for iterative receivers shown in Table 4.2 demonstrates that the receivers based on the proposed interference cancelling detectors MS-PPIC and MF-SIC offer a significant complexity advantage. The receiver complexity with the proposed algorithms could be reduced by approximately 50% in comparison to the Max-Log-APP based iterative receiver. It is also shown that for the iterative receivers under investigation, the complexity for space-time equalisation is low in comparison to the detection and decoding components since equalisation is performed only once and is not included in the iterative loop. As for the non-iterative receivers, the solutions without space-time equalisation cannot compete with the equaliser based receivers neither in terms of complexity nor in performance.

In comparison with the non-iterative receivers, the total complexity of the iterative receivers is, despite the low complexity decoder, significantly higher. In most applications, the 1 dB perfor-

mance gain does not justify the more than two times higher complexity. Note that the iterative receivers under investigation are dimensioned for high performance. The use of a Max-Log-MAP decoder with only 4 instead of 8 decoding iterations would reduce the receiver-complexity by nearly 50% and could be a very competitive solution to the non-iterative receivers in terms of performance and complexity.

Components	**Non-iterative receiver complexity [operations]**				
	EQ (V3)	APP (V1)	Max-Log APP (V1)	MS-PPIC (V1)	MS-PPIC no EQ (V2)
Equaliser	65536	65536	65536	65536	0
Detector	0	196608	131072	18944	233984
Decoder	120832	120832	120832	120832	120832
$\sum$	186368	382976	317440	205312	354816

Table 4.1: *Non-iterative receiver complexity comparison (Normal operation).*

Components	**Iterative receiver complexity [operations]**				
	Max-Log APP (V1)	MS-PPIC (V1)	MS-PPIC no EQ (V3)	MF-SIC (V2)	MF-SIC no EQ (V4)
Equaliser	1×65536	1×65536	0	1×65536	0
Detector	4×131072	4×18944	4×233984	4×4608	4×73728
Decoder	4×92160	4×92160	4×92160	4×92160	4×92160
$\sum$	958464	509952	1304576	452608	663552

Table 4.2: *Iterative receiver complexity comparison (Normal operation).*

4.8 Summary and conclusions

In this chapter, an overview of non-iterative and iterative receiver architectures was presented. Two receiver architectures based on space-time equalisation, de-spreading and pre-whitening followed by detection and decoding were proposed and their performance and complexity was compared using different detection algorithms.

The simulation results for non-iterative receivers confirm that the MS-PPIC can be used as a low-complexity alternative for the APP based algorithms. In the simulations with coding rate 1/3 for dispersive channels, the MS-PPIC even outperforms the Max-Log-MAP algorithm in the proposed receiver architecture. In addition, the use of the proposed MS-PPIC algorithm for detection reduces the receiver complexity by approximately 50% in comparison to the APP based receiver. It was also shown that the use of a detection stage after de-spreading and pre-whitening is important to achieve good performance, especially for flat fading or high code rates (e.g. rate 1/2). Finally, the results for the TU channel have demonstrated that the use of a space-time equaliser prior to detection improves the performance by eliminating the interference between the spreading codes. This also leads to a significantly reduced computational complexity of the detection algorithm.

To further improve the receiver performance, the proposed architecture can be modified to an iterative receiver architecture, where reliable soft-outputs from the decoder are fed back to improve the detector outputs using *a priori* information. In addition to the detectors investigated for non-iterative receivers, simple, matched filter based successive interference cancellation becomes interesting, where despite the use of a simplified detection scheme, a high level of receiver performance is achieved via iterations with the decoder. In fact, it is shown that, via a combination of bit-level ordering, detection and cancellation, and the use of a novel soft-output combining technique which maximises the mutual information transfer in each iteration, the low-complexity iterative MF-SIC receiver can compete with the Max-Log-APP and MS-PPIC based iterative receivers. While the Max-Log-APP algorithm performs best, the interference cancelling detectors (MS-PPIC and MF-SIC) can offer similar performance and reduce the computational complexity of the iterative receiver by approximately 50%. It was demonstrated that with the proposed iterative receiver architecture, the receiver performance can be improved by approximately 1 dB in comparison to the original non-iterative receivers, but this requires more than twice the computational complexity. Therefore, it is questionable if the additional complexity for the iterations is justified with respect to the performance gain.

Chapter 5
Layered encoding for high-order modulations

5.1 Introduction

When higher data rates are required, one solution in MIMO systems is to increase the number of antennas at the transmitter and the receiver. However, this is not always preferred due to size restrictions at the user equipment or the resulting higher manufacturing costs. Alternatively, the modulation scheme can be changed from 4-QAM to high-order modulations such as 16- or 64-QAM to increase the data rate in adequate channel conditions by the factor 2 or 3 respectively. This has several implications for the receivers under investigation:

- The optimum APP detector becomes too complex for a 4×4 antenna system, since the complexity grows exponentially with the modulation order and the number of transmitting antennas.
- The MS-PPIC detector offers only poor performance in comparison to the optimum APP detector.
- The iterative MF-SIC based receiver cannot converge due to the poor initial performance of the MF detection stage.

Recent research has mainly concentrated on reduction of the search space for the APP detector with promising results. One example is the "spherical APP" detector which searches for candidates in a certain sphere radius around the soft symbol output [28]. This method leads to reliable estimates especially at high E_b/N_0, but in the range of interest at low E_b/N_0 the spherical APP does not produce the same soft-output quality as the full APP. In addition, the search algorithm is complicated to implement in hardware. In this chapter, all three problems are addressed by introducing a novel layered encoding scheme, whereby bits which map to different Euclidean distances on the modulation constellation are encoded separately. This essentially represents a form of multilevel coded modulation utilising set partitioning [23, 32, 33].

The concept of layered encoding is presented in detail in Section 5.2. To show the impact of layered encoding on the receiver performance two cases are investigated: Firstly, in Section 5.3 the proposed layered encoding scheme is simulated in a non-CDMA scenario for an ergodic MIMO channel and different numbers of antennas at the transmitter and receiver. The simulations are performed without equalisation since the ergodic channels are non-dispersive. Secondly, in Section 5.4 the performance is investigated in a CDMA typical urban (TU) scenario for the receivers proposed in the previous chapter. It will be shown that the proposed layered encoding technique significantly reduces the computational complexity for receivers using high-order modulation without loss in performance in comparion to the original full APP receiver. Finally in Section 5.5, the system throughput and the applicability of MIMO systems using high-order modulation schemes is investigated using system-level simulations. It will be demonstrated that even 64-QAM modulation can be used in a significant area of the cell.

5.2 The concept of layered encoding

The high levels of interference caused by code reuse across the transmitting antennas of a MIMO link can lead to poor performance when using high-order constellations such as 16- or 64-QAM. This is particularly the case for receivers based on low complexity detection algorithms like the MS-PPIC and the iterative MF-SIC. The quality of the soft outputs at the detector or decoder output can be so poor that the iterative processes cannot converge adequately. APP based algorithms still offer good performance but at the expense of a prohibitive computational complexity. The problems described above can be resolved through a judicious encoding process as described next.

For a Gray-mapped 16-QAM constellation, each symbol $x_k^{(i)}(t)$ transmitted by the i^{th} transmitter antenna via the k^{th} spreading code at time t is given by

$$x_k^{(i)}(t) = 2\left\{-b_{k,0}^{(i)}(t) - jb_{k,1}^{(i)}(t)\right\} + \left\{-b_{k,0}^{(i)}(t)\, b_{k,2}^{(i)}(t) - jb_{k,1}^{(i)}(t)\, b_{k,3}^{(i)}(t)\right\} \tag{5.1}$$

as a function of encoded bits $b_{k,0}^{(i)}(t), b_{k,1}^{(i)}(t), b_{k,2}^{(i)}(t), b_{k,3}^{(i)}(t) \in \{-1, +1\}$. The corresponding constellation is illustrated in Figure 5.1. As can be seen, for such high-order constellations, the Euclidean distance is not the same for all modulated bits. This implies that the modulation scheme affords different levels of protection to different bits. For the Gray mapped 16-QAM constellation of Figure 5.3, it is clear that $b_{k,0}^{(i)}(t)$ and $b_{k,1}^{(i)}(t)$ are equally better protected than

$b_{k,2}^{(i)}(t)$ and $b_{k,3}^{(i)}(t)$.

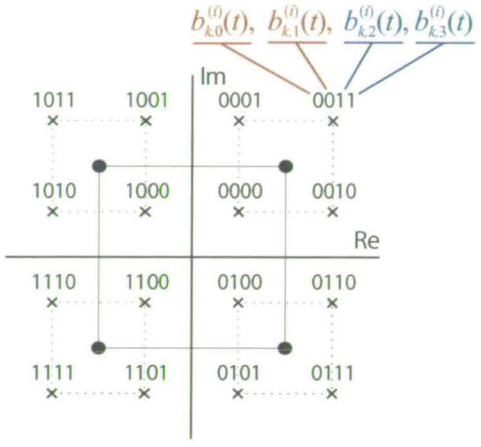


Figure 5.1: *16-QAM constellation with Gray mapping.*

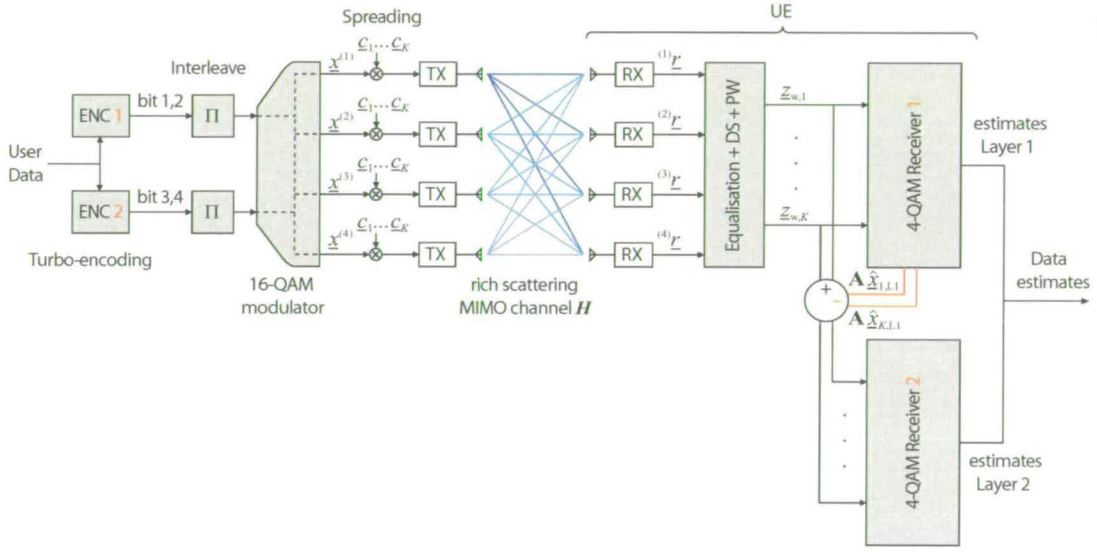


Figure 5.2: *MIMO link with layered encoding for 16-QAM.*

The above feature may be exploited by encoding the well-protected and less-protected bits separately at the transmitter to be able to detect and decode them sequentially. The proposed modified receiver architecture is shown in Figure 5.2. User data is split into two classes of bits, denoted as layers, and is encoded and interleaved. The encoded bits of Layer 1 correspond to $b_{k,0}^{(i)}(t)$ and $b_{k,1}^{(i)}(t)$, while the encoded bits of Layer 2 correspond to $b_{k,2}^{(i)}(t)$ and $b_{k,3}^{(i)}(t)$. The bits are then mapped on to 16-QAM symbols according to Equation 5.1.

At the receiver, the sufficient statistics of the of the received signal after space-time equalisation,

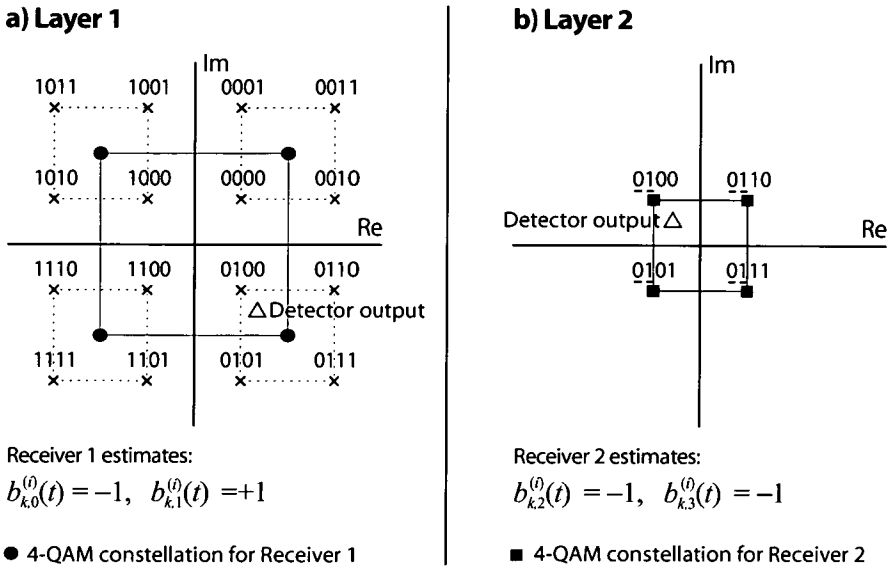


Figure 5.3: *16-QAM modulation as an aggregate of 2 interdependent 4-QAM modulations.*

de-spreading and pre-whitening can be written as (see Equation 2.15)

$$\underline{z}_{w,k}(t) = \boldsymbol{A}_k(t)\,\underline{x}_k(t) + \underline{\epsilon}(t) = \boldsymbol{A}_k(t)\,\underline{x}_{k,\mathrm{L1}}(t) + \underline{\epsilon}(t) + \underline{\delta}(t). \tag{5.2}$$

Then, the well-protected bits $b_{k,0}^{(i)}(t)$ and $b_{k,1}^{(i)}(t)$ corresponding to the symbols $\underline{x}_{k,\mathrm{L1}}(t)$ of Layer 1 are detected and decoded first. Due to the greater Euclidean distance associated with these bits, the decoding process is able correct any errors reliably. Subsequently, the contribution of the estimated bits $b_{k,0}^{(i)}(t)$ and $b_{k,1}^{(i)}(t)$ from all transmitter antennas is cancelled from the sufficient statistics for each spreading code k:

$$\underline{z}_{w,k}(t) := \underline{z}_{w,k}(t) - \boldsymbol{A}_k(t)\,\hat{\underline{x}}_{k,\mathrm{L1}}(t) \tag{5.3}$$

where $\boldsymbol{A}_k\hat{\underline{x}}_{k,\mathrm{L1}}$ is the reconstructed first layer contribution transmitted via the k^{th} spreading code. The reconstructed symbol vector for the first layer can be written as

$$\hat{\underline{x}}_{k,\mathrm{L1}}(t) = \begin{bmatrix} 2\left\{-b_{k,0}^{(1)}(t) - jb_{k,1}^{(1)}(t)\right\} \\ \vdots \\ 2\left\{-b_{k,0}^{(N_\mathrm{T})}(t) - jb_{k,1}^{(N_\mathrm{T})}(t)\right\} \end{bmatrix}. \tag{5.4}$$

This significantly reduces the interference for the remaining less-protected bits $b_{k,2}^{(i)}(t)$ and

$b_{k,3}^{(i)}(t)$, which are only then detected and decoded. The detection and cancellation process is illustrated in Figure 5.3 for a 16-QAM constellation. In the example shown, the bits corresponding to Layer 1 are estimated as $b_{k,0}^{(i)}(t) = -1$ and $b_{k,1}^{(i)}(t) = 1$ by Receiver 1 (a) and the contributions are cancelled. Subsequently, the remaining bits corresponding to Layer 2 are estimated as $b_{k,2}^{(i)}(t) = -1$ and $b_{k,3}^{(i)}(t) = -1$ by Receiver 2 (b). For 64-QAM, the procedure is identical, except that three classes are considered, according to the three levels of protection provided by the modulation scheme.

To further increase the performance of the proposed layered encoding scheme, the coding rate of each layer can be adapted to the detection algorithm and the channel conditions via puncturing or repetition. In this way the cancellation process of each layer can be optimised and thereby improve the bit-error rate of the receiver without changing the overall code rate for the transmitted data block. An optimal rate allocation achieves the best tradeoff between protecting the subsequent layers and minimising the error propagation from previous layers.

The optimal rate allocation protects the subsequent layers well enough to prevent poor results and also keeps the error propagation between the layers low.

5.3 Performance of layered encoding for non-CDMA ergodic radio links

In this section, the proposed layered encoding scheme is investigated for an ergodic channel scenario, with 16-QAM modulation and different antenna settings starting at 1×1 up to 4×4. No CDMA ($K = 1, Q = 1$) and also no equalisation is employed. The optimum APP algorithm is used. A 8-state $1/3$ rate parallel concatenated turbo encoder is applied according to the HSDPA specifications. Rate 1/2 encoding is achieved via puncturing. Decoding is performed via 8 turbo iterations. The coding block-size is 1000 information bits.

The complexity for the APP based receivers under investigation depends on the number of signal constellation points which need to be searched by the receivers. The full APP receiver searches over all 2^{MN_T} where M is the number of bits per symbol ($M = 4$ for 16-QAM) and N_T is the number of transmitter antennas. With layered encoding the number of states per detection stage reduces to 2^{2N_T} (each layer is 4-QAM modulated with $M = 2$). The total computational complexity is proportional to $(M/2) \times 2^{2N_T}$ where $M/2$ is the number of layers

(2 for 16-QAM). In this way, the computational complexity increases only linearly with the modulation order and not exponentially.

Figure 5.4 shows the BER performance for the 1×1 antenna case with equal coding rates for each layer. Here, the full APP detector shows the best performance, approximately 1.4 dB better than the APP with layered encoding. The reason for the performance loss lies in the different Euclidian distances in the modulation of the separately and equivalently encoded layers: the first layer is too well protected which leaves insufficient protection for the second layer. This leads to a poor average BER performance. Figure 5.5 shows the performance comparison using layer-specific coding rates. For Layer 1, a rate 1/1.7 coding is employed and for Layer 2 a coding rate of 1/2.3 resulting in and average coding rate of 1/1.955 ($\approx 1/2$). It can be seen that now the BER performance of the two layers is better balanced and even results in an improvement of 0.4 dB in comparison to the full APP. In terms of computational complexity, the full APP receiver for the single antenna case is twice as complex as the APP with layered encoding (APP: 16 search states, LE+APP: 2×4 search states).

Figure 5.6 shows the performance for a 2×2 antenna MIMO system. In this specific example, no layer-specific coding rates are required since the performance of the individual layers is already well balanced. However under other channel conditions, variable coding rates for the layers can improve the performance, as shown later. In terms of computational complexity, the full APP receiver for 2×2 antennas is 8 times as complex as the APP with layered encoding (APP: 256 search states, LE+APP: 2×16 search states)

The results for 4×4 antennas are shown in Figure 5.7. Again, the receiver with layered encoding performs close to the full APP detector with no need for layer-specific coding rates for the simulated channel conditions. In terms of computational complexity, the full APP receiver for 4×4 antennas is 128 times as complex as the APP with layered encoding (APP: 65536 search states, LE+APP: 2×256 search states)

The simulations in the non-CDMA ergodic scenario have shown that with the proposed layered encoding method high-order modulations can be treated as interdependent 4-QAM constellations which can be detected and decoded separately. This leads to a significant reduction in computational complexity. It was shown that for a 4×4 antenna MIMO link, the complexity can be reduced to only 0.8% of the original APP complexity without a significant loss in performance.

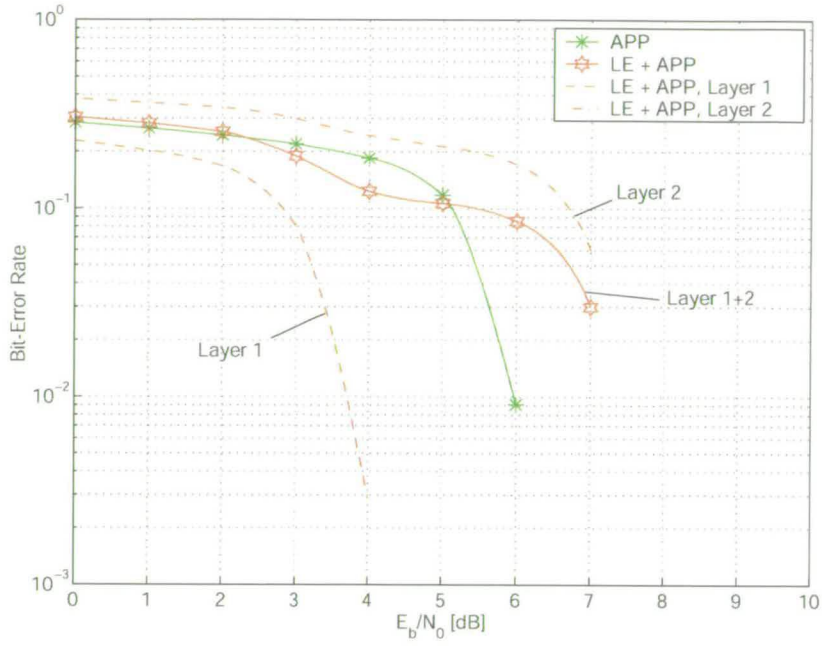


Figure 5.4: *BER comparison for 16-QAM, 1×1 antennas and ergodic channel with fixed rate 1/2 turbo coding.*

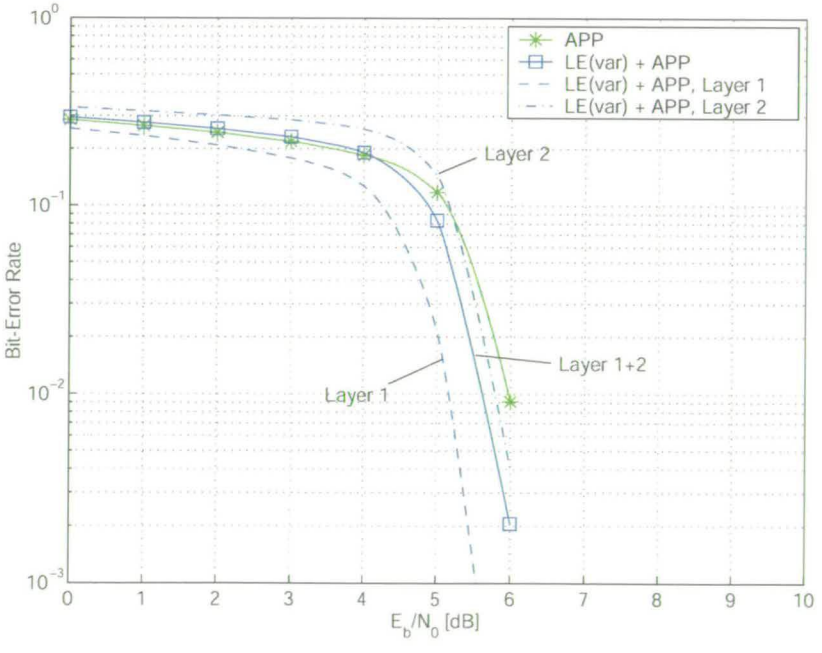


Figure 5.5: *BER comparison for 16-QAM, 1×1 antennas and ergodic channel with variable rate 1/2 turbo coding.*

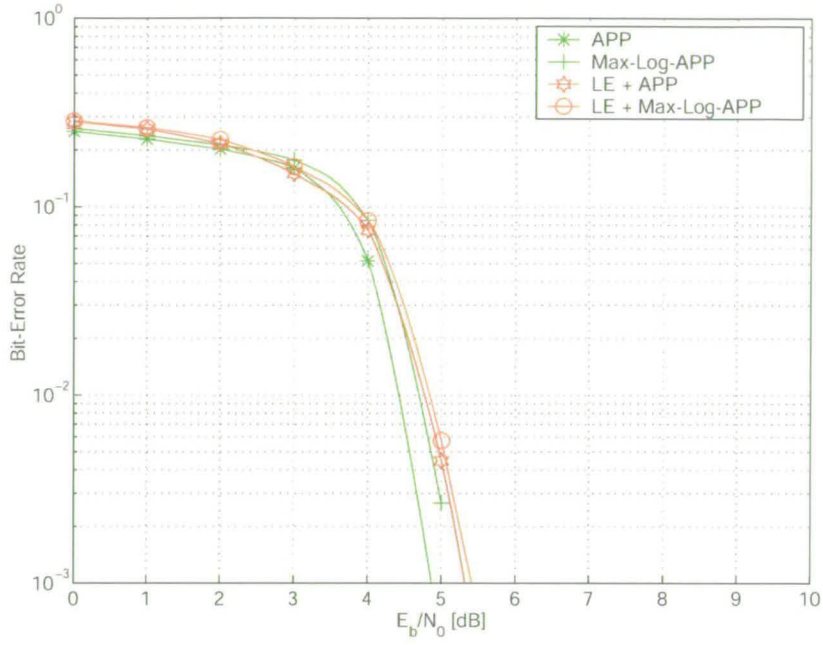


Figure 5.6: *BER comparison for 16-QAM, 2×2 antennas and ergodic channel with fixed rate 1/2 turbo coding.*

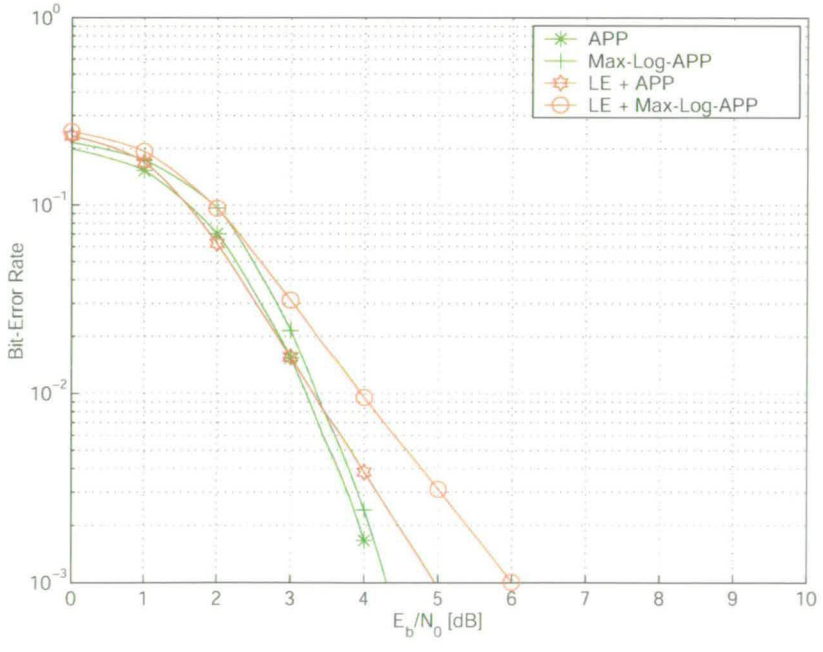


Figure 5.7: *BER comparison for 16-QAM, 4×4 antennas and ergodic channel with fixed rate 1/2 turbo coding.*

5.4 Performance of layered encoding with CDMA and TU channel

In this section, the performance of layered encoding is investigated for a CDMA scenario with TU channel conditions for 16- and 64-QAM modulation. The receiver architecture is the same as presented in the previous chapters, employing space-time equalisation, de-spreading and pre-whitening, detection and turbo decoding. Detection of each of the 4-QAM layers, is performed via either simple de-spreading or more sophisticated algorithms such as APP, Max-Log-APP and MS-PPIC. For 16-QAM, the results are compared to the full APP detector performance which requires a search over 65536 symbol vector candidates for each spreading code and symbol period. This comparison is not possible for 64-QAM, since the required search over 16.7×10^6 possible candidates is far too complex to simulate. A system with $N_{\mathrm{T}} = N_{\mathrm{R}} = 4$, $Q = 16$ and $K = 16$ is considered. A block-fading channel with 10 chip-spaced taps and a TU power profile is examined and the channel conditions are assumed to be known at the receiver. For decoding, a 8-state $1/3$ rate turbo decoder according to the HSDPA specification is used with 8 decoding iterations and a coding-block size of 1000 information bits. The results for $1/2$ rate are achieved via puncturing. The MS-PPIC detector operates with 6 iterations.

5.4.1 Simulation results for 16-QAM

Figures 5.8 and 5.9 show the performance of the reference receiver using the full APP and Max-Log-APP algorithm. Due to the high computational complexity, the simulations took over 4 weeks on a 800 MHz Pentium PC. As might be expected, errors due to the Max-Log approximation are more evident when operating at low SNRs. This can be seen by comparing the results for coding rates of $1/3$ and $1/2$. With rate $1/3$ coding (operating at low SNR) the difference between the APP and the Max-Log-APP is 0.5 dB at a BER of 10^{-2} whereas for 1/2 rate coding (operating at higher SNR), the difference diminishes.

Figure 5.10 shows the performance of the receivers with layered encoding with rate 1/3 coding. It can be seen, that the low complexity receivers with layered encoding are able to achieve the same BER and FER performance like the full APP reference shown in Figure 5.8. Figure 5.11 shows the impact of layer-specific coding rates ($R_{\mathrm{L1}} = 1/2.6$, $R_{\mathrm{L2}} = 1/3.4 \Rightarrow R_{\mathrm{AV}} = 1/2.946$). The performance improves significantly and can now outperform the reference APP receiver by approximately 1 dB. Note that, the MS-PPIC based receiver is able to perform as

well as the APP receiver with layered encoding and thereby outperform the Max-Log-APP.

Figures 5.12 and 5.13 show the layered encoding results for rate 1/2 coding with and without variable coding rates for the layers. Compared to the reference receiver in Figure 5.9, the performance with layered encoding is better by 0.5 - 1 dB, even without layer-specific coding rates. When variable coding rates are used as shown in Figure 5.13, the performance improves further by 0.5 dB ($R_{\mathrm{L1}} = 1/1.7$, $R_{\mathrm{L2}} = 1/2.3 \Rightarrow R_{\mathrm{AV}} = 1/1.995$). For the coding rate 1/2, the receivers based on APP, Max-Log-APP and MS-PPIC detection, all with layered encoding, offer the same performance.

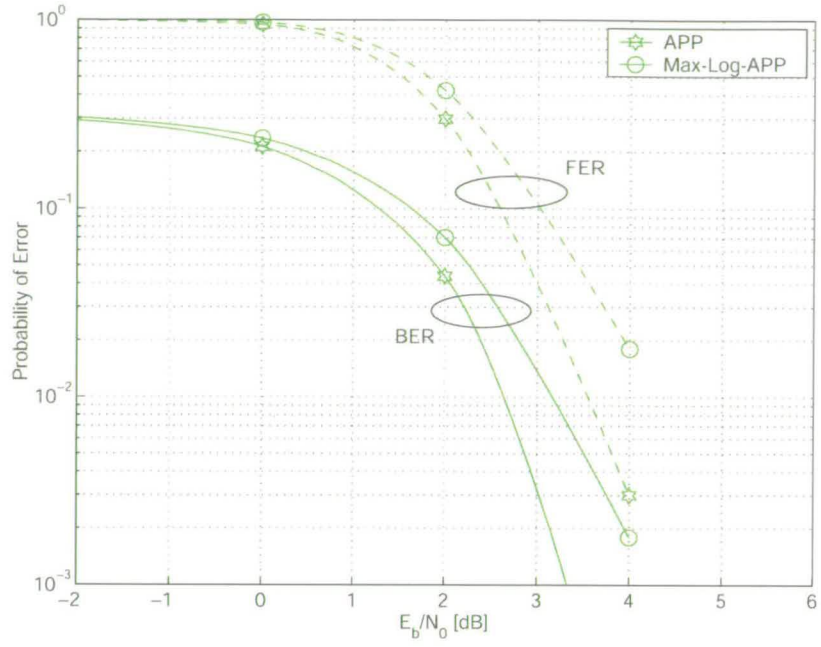


Figure 5.8: *APP Performance without layered encoding for 16-QAM, 4 × 4 antennas and TU channel with rate 1/3 turbo coding.*

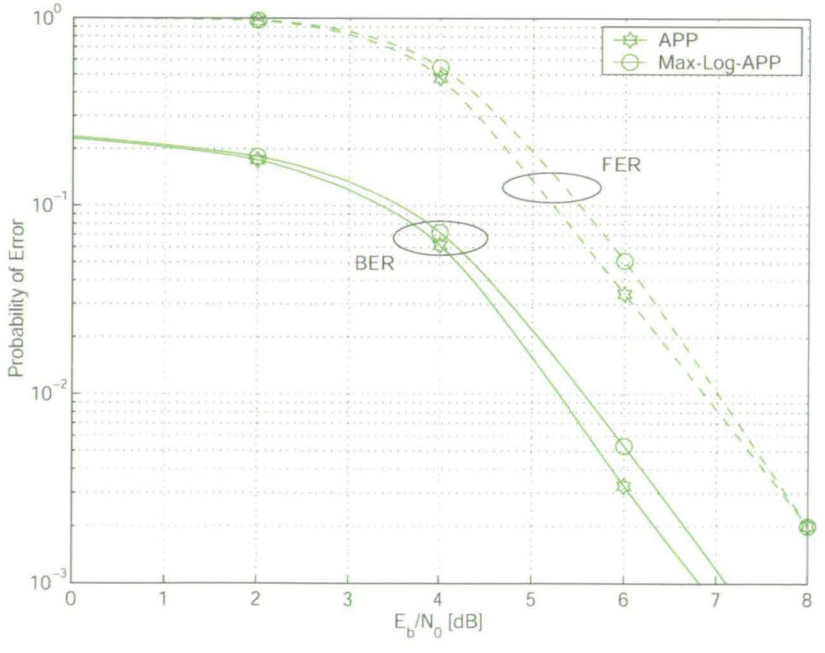


Figure 5.9: *APP Performance without layered encoding for 16-QAM, 4 × 4 antennas and TU channel with rate 1/2 turbo coding.*

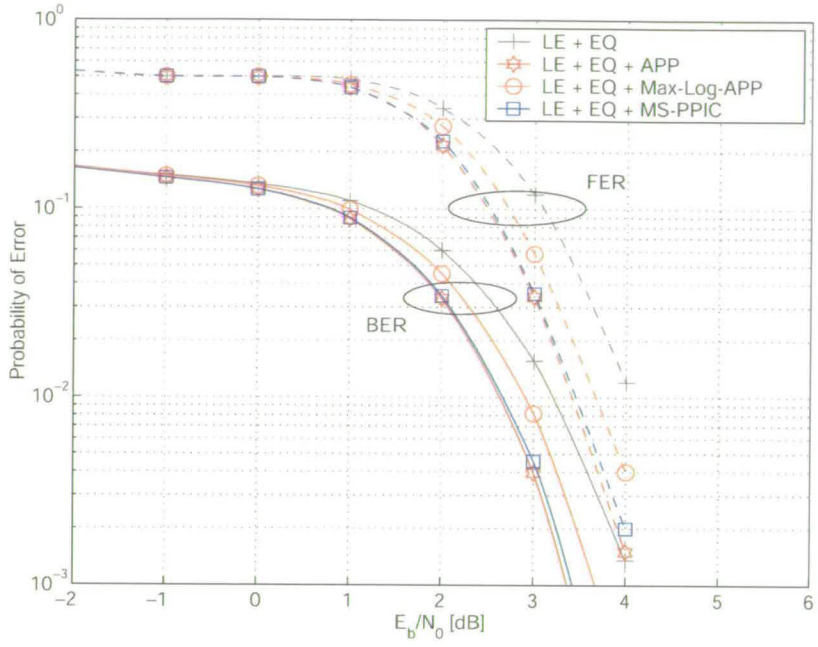


Figure 5.10: *Performance comparison for 16-QAM, 4×4 antennas and TU channel with fixed rate 1/3 turbo coding.*

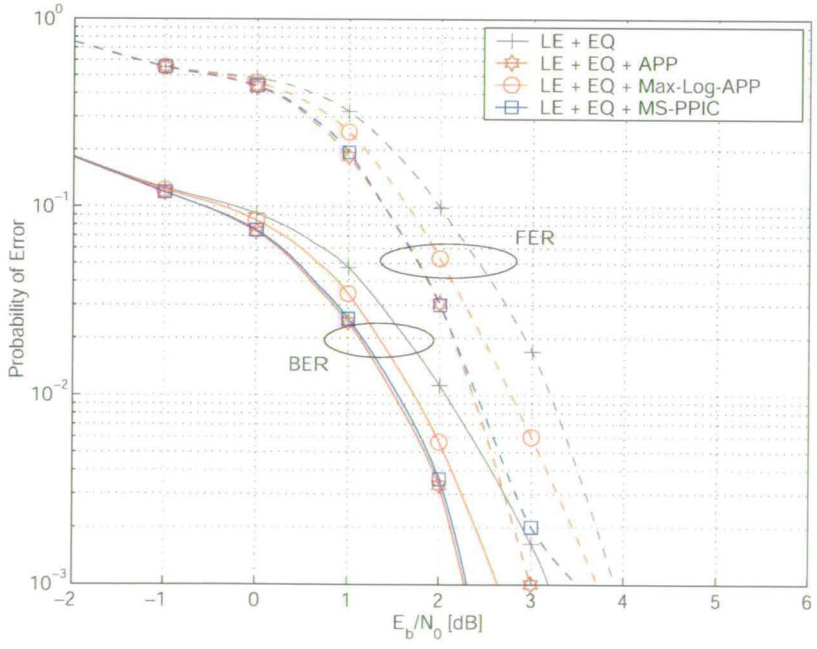


Figure 5.11: *Performance comparison for 16-QAM, 4×4 antennas and TU channel with variable rate 1/3 turbo coding.*

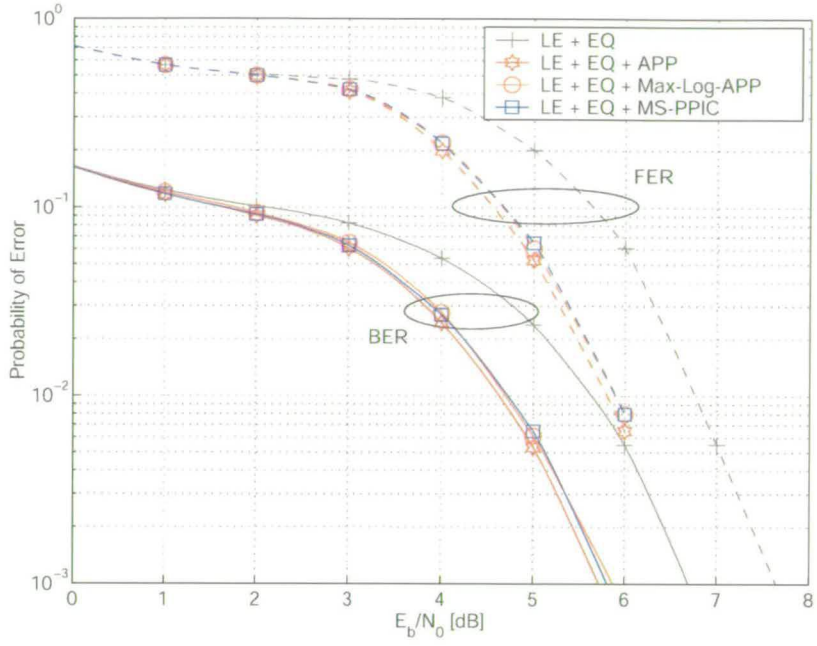


Figure 5.12: *Performance comparison for 16-QAM, 4×4 antennas and TU channel with fixed rate 1/2 turbo coding.*

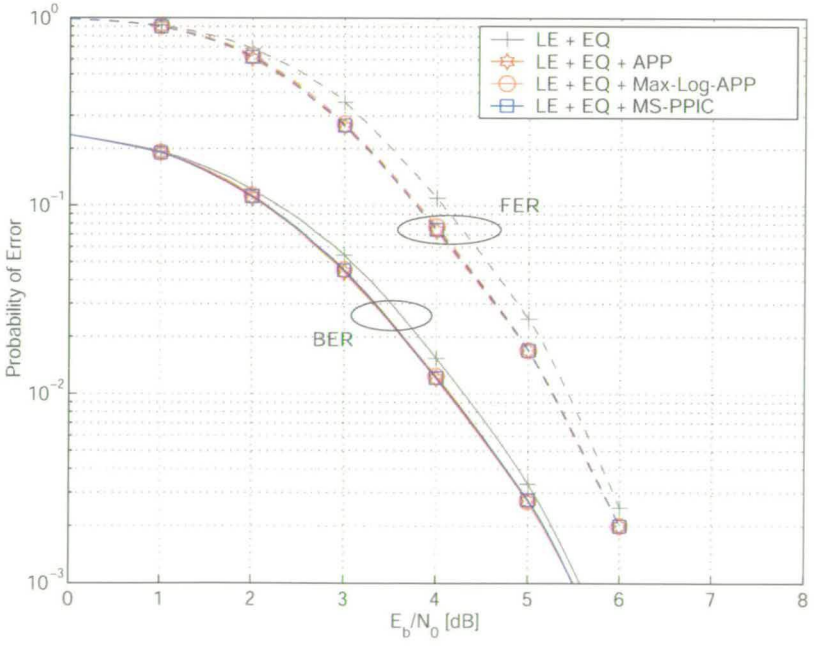


Figure 5.13: *Performance comparison for 16-QAM, 4×4 antennas and TU channel with variable rate 1/2 turbo coding.*

5.4.2 Simulation results for 64-QAM

Figures 5.14 and 5.15 show the BER and FER performance of the receivers with layered encoding for 64-QAM modulation, rate 1/3 coding, without and with layer-specific coding rates respectively. Here, the reference APP receiver becomes with 16.7×10^6 possible states far too complex to simulate. The results show that the proposed layered encoding method is capable of coping with high-order modulation schemes without problems and all receivers offer similar good results. With variable layer rates ($R_{\mathrm{L1}} = 1/2.1$, $R_{\mathrm{L2}} = 1/2.9$, $R_{\mathrm{L3}} = 1/4.0 \Rightarrow R_{\mathrm{AV}} = 1/2.801$), the performance can be improved further by 3.5 dB at a FER of 10^{-1}.

The same results for coding rate 1/2 are presented in Figures 5.16 and 5.17. Again, the different detection algorithms offer essentially the same performance. With variable layer rate ($R_{\mathrm{L1}} = 1/1.6$, $R_{\mathrm{L2}} = 1/2.1$, $R_{\mathrm{L3}} = 1/2.3 \Rightarrow R_{\mathrm{AV}} = 1/1.953$), the performance can be improved by 2 dB at the optimised FER range of interest at 10^{-1}.

For 64-QAM, the complexity of the full APP detector can be reduced by the factor 21.8×10^3 by using the proposed layered encoding scheme.

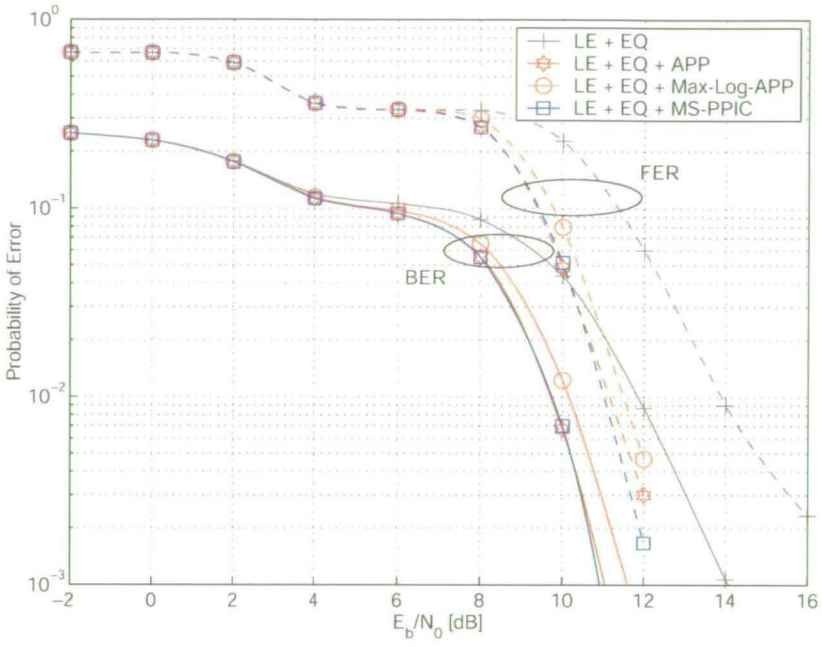


Figure 5.14: *Performance comparison for 64-QAM, 4×4 antennas and TU channel with fixed rate 1/3 turbo coding.*

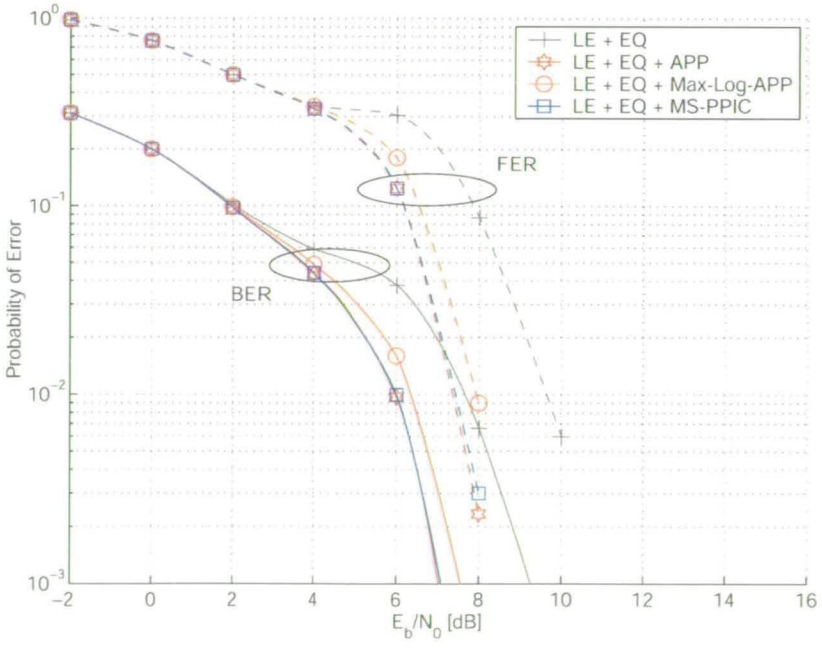


Figure 5.15: *Performance comparison for 64-QAM, 4×4 antennas and TU channel with variable rate 1/3 turbo coding.*

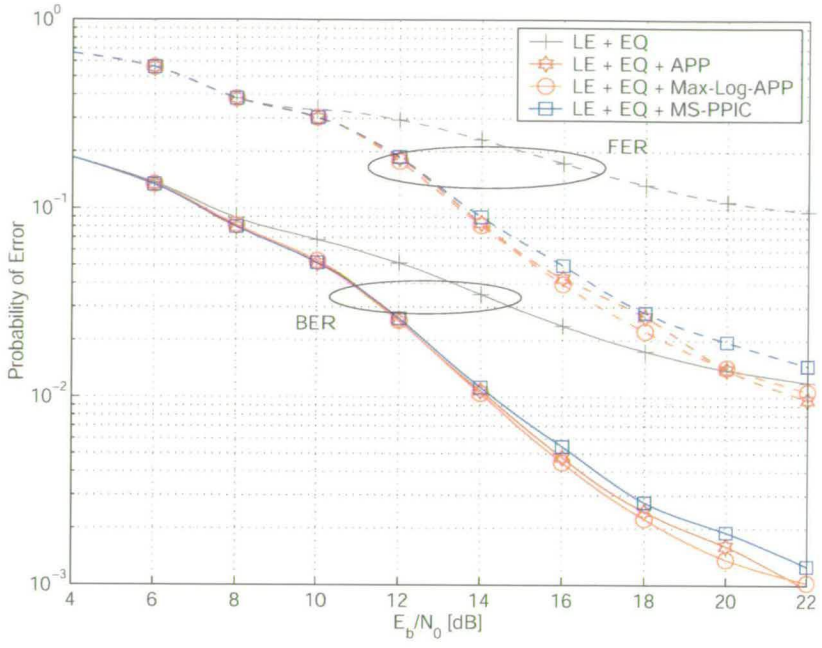


Figure 5.16: *Performance comparison for 64-QAM, 4×4 antennas and TU channel with fixed rate 1/2 turbo coding.*

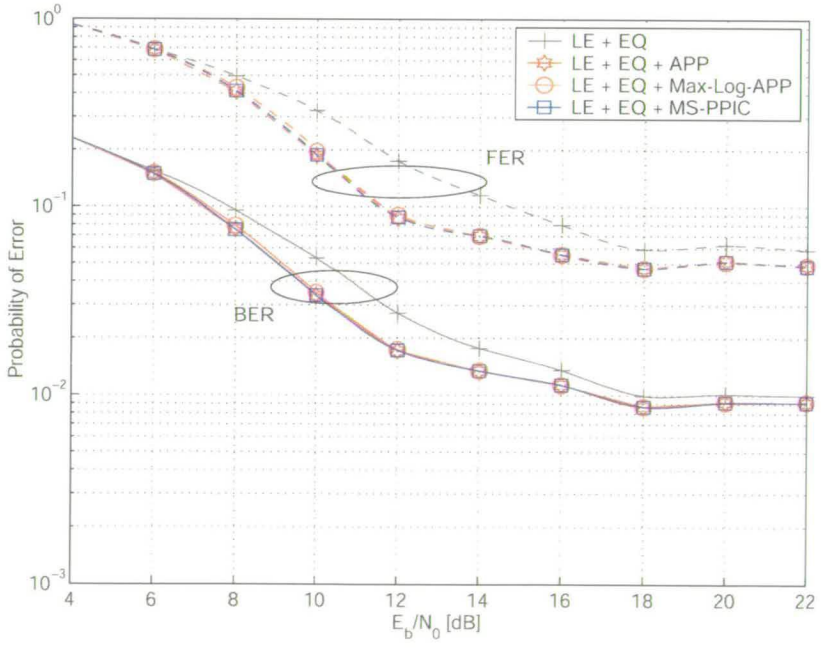


Figure 5.17: *Performance comparison for 64-QAM, 4×4 antennas and TU channel with variable rate 1/2 turbo coding.*

5.5 System-level considerations

In the previous sections it has been shown that high-order modulation schemes for MIMO radio communications are feasible and the problems regarding computational complexity and performance have been addressed. However, the question remains as to whether these high-order modulation schemes can be used in a significant area of the cell in real systems, or if they can only operate within a limited area surrounding the base-station. In this section, the applicability of high-order modulation schemes is assessed using system-level simulations for both a two cell indoor scenario and a 7-cell urban scenario for the HSDPA downlink. In addition, the obtainable system throughput with the different modulation schemes and coding rates is investigated.

5.5.1 System throughput

Figures 5.18 and 5.19 show the obtainable system throughput with 4-, 16- and 64-QAM modulation and coding rates 1/2 and 1/3 for TU channels (as used in Chapter 4) where the MS-PPIC algorithm with space-time equalisation is employed. For 16- and 64-QAM, the proposed layered encoding scheme was used. Since the FER performances of the MS-PPIC, APP and Max-Log-APP algorithms are practically identical for the TU channel, the throughput values can be generalised to all three algorithms. The results show that, for rate 1/3 coding, throughputs of up to 30 MBit/s can be achieved for a 4×4 antenna MIMO system with 64-QAM modulation. For coding rate 1/2, throughputs of up to 45 MBit/s are possible. Note that throughput is given in terms of information bits per second. In combination with the FER results in the previous sections, it is shown that even for high FER of 10-20%, a throughput close to the maximum is achieved for each modulation scheme. This also justifies the target FER of 10^{-1} which was assumed in the previous chapters, and shows that in order to achieve the maximum throughput, minimising the FER is not necessarily the best strategy. It is often more advantageous to increase the coding rate or the modulation order as soon as possible. For example, in Figure 5.19 at a E_b/N_0 of 3 dB, using 4-QAM modulation results in a throughput of 15 MBit/s with virtually no frame-errors as shown in Figure 4.11. Using 16-QAM modulation instead, leads to a higher FER of 15% as shown in Figure 5.15, however the throughput increases significantly to 22 MBit/s.

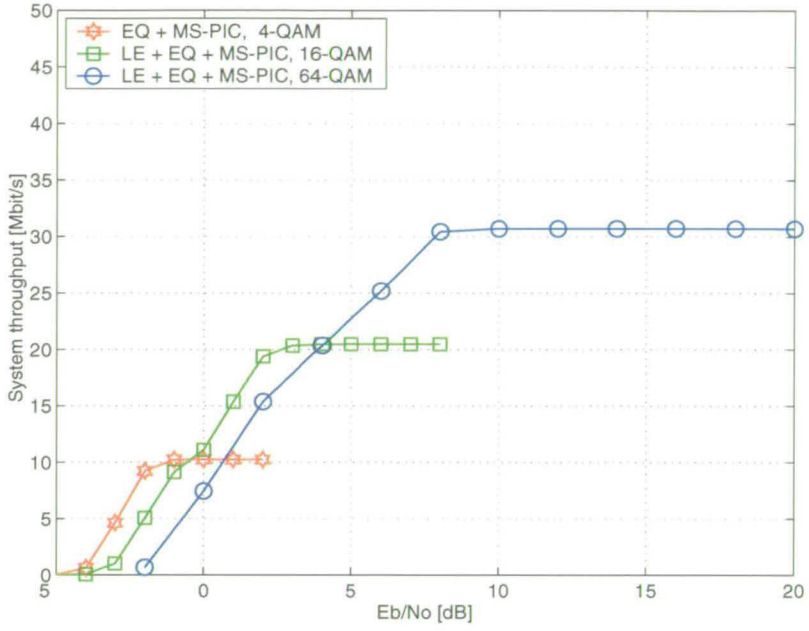


Figure 5.18: *System throughput for a 4×4 antenna MIMO link, TU channel, 4-QAM, 16-QAM and 64-QAM modulation and 1/3 rate coding.*

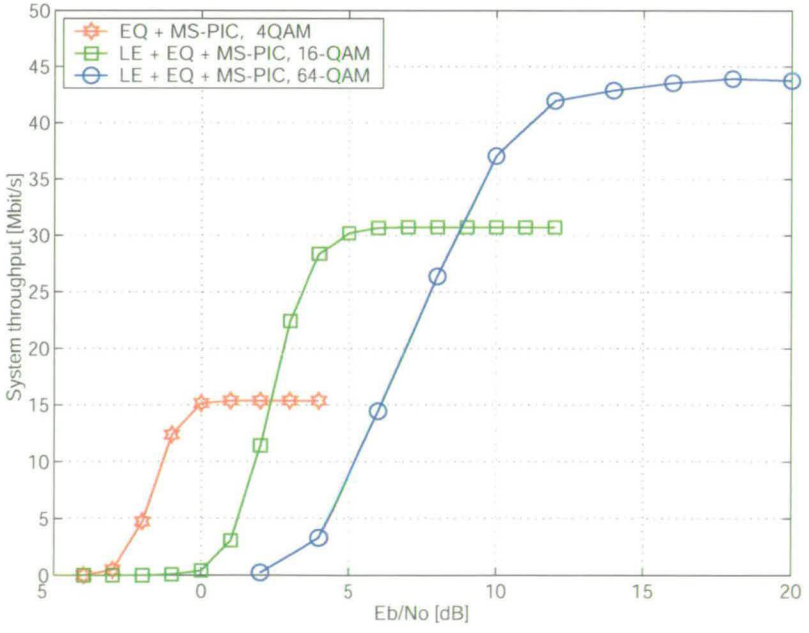


Figure 5.19: *System throughput for a 4×4 antenna MIMO link, TU channel, 4-QAM, 16-QAM and 64-QAM modulation and 1/2 rate coding.*

5.5.2 System level simulations

System level simulations can be used to gain insight and assess the viability of the higher order modulation schemes for different scenarios.

The system level simulator is used to generate records of

- intra-cell powers
- intra-cell interference powers
- Signal-to-interference+noise ratios (SINRs)

for the downlink of a HSDPA-type system as a function of the location of the serviced mobile within its cell. This is done by dropping the mobile to random positions in the cell as shown in Figure 5.20 for each simulation trial. Then the link budget from each base-station to the mobile is calculated. It is assumed that the mobile receives data from the base-station with the best channel condition with other base-stations acting as interference. The resulting signal and interference powers are then used to calculate the probability density function (pdf) and the corresponding cumulative distribution function (cdf) for the signal-to-interference+noise ratio (SINR) at the mobile for the scenario under investigation. Then, for each target SINR value the coverage in the cell can be determined, based on the cdf. Coverage is defined as the proportion of the cell where the SINR at the UE is higher than the target SINR value.

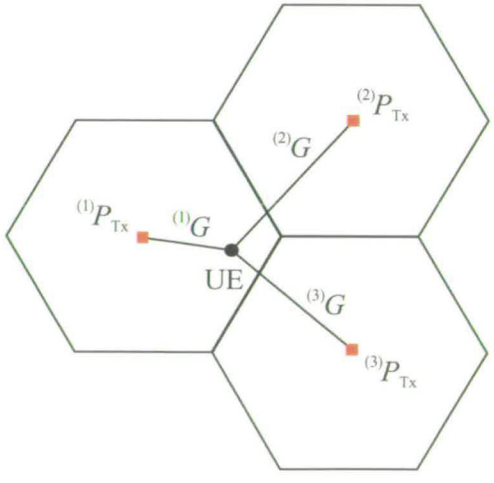


Figure 5.20: *SINR calculation for system-level simulations.*

Based on the received signal power at the user equipment (UE) $P_{\mathrm{signal}}^{\mathrm{UE}}$, the received interference

power $P_{\text{interf}}^{\text{UE}}$ and the noise-power $P_{\text{noise}}^{\text{UE}}$, the signal-to-interference+noise ratio (SINR) can be calculated as follows:

$$\begin{aligned}\text{SINR}_{\text{DL}} &= \frac{P_{\text{signal}}^{\text{UE}}}{P_{\text{noise}}^{\text{UE}} + P_{\text{interf}}^{\text{UE}}} \\ &= \frac{{}^{(1)}P_{\text{Rx}}^{\text{UE}}}{P_{\text{noise}}^{\text{UE}} + \sum_{u\neq 1} {}^{(u)}P_{\text{Rx}}^{\text{UE}}} \\ &= \frac{{}^{(1)}P_{\text{Tx}} \times {}^{(u)}G}{P_{\text{noise}}^{\text{UE}} + \sum_{u\neq 1} {}^{(u)}P_{\text{Tx}} \times {}^{(u)}G}.\end{aligned} \tag{5.5}$$

$P_{\text{signal}}^{\text{UE}}$ denotes the received power at the user equipment (UE) from the u^{th} base-station, ${}^{(u)}P_{\text{Tx}}$ represents the transmitted power of the u^{th} base-station and ${}^{(u)}G$ denotes the power gain between the u^{th} base-station and the UE. The noise-power at the receiver can be calculated as

$$P_{\text{noise}}^{(\text{UE})}\,[\text{dB}] = 10\log_{10}\left(kT \times \text{NF}_{\text{UE}} \times B\right) \tag{5.6}$$

where $kT = 1.3804 \times 10^{-23} \times 290$ W/Hz is the single-sided thermal-noise power spectral density. The UE receiver noise figure is assumed to be $\text{NF}_{\text{UE}} = 10$ dB and the UE noise-equivalent bandwidth is $B = 3.84 \times 10^6$ Hz. The power gain between the u^{th} base-station (BS) and the UE is denoted by

$$\begin{aligned}{}^{(u)}G\,[dB] = &\left(G_{\text{ant}}^{\text{BS}}\,[\text{dB}] - L_{\text{cable}}^{\text{BS}}\,[\text{dB}]\right) + \left(G_{\text{ant}}^{\text{UE}}\,[\text{dB}] - L_{\text{cable}}^{\text{UE}}\,[\text{dB}]\right) \\ &- L_{\text{shadow}}\,[\text{dB}] - L_{\text{path}}\,[\text{dB}]\end{aligned} \tag{5.7}$$

where $\left(G_{\text{ant}}^{\text{BS}}\,[\text{dB}] - L_{\text{cable}}^{\text{BS}}\,[\text{dB}]\right)$ is the antenna gain at the base-station including cable loss and $\left(G_{\text{ant}}^{\text{UE}}\,[\text{dB}] - L_{\text{cable}}^{\text{UE}}\,[\text{dB}]\right)$ is the UE antenna gain including cable loss. The shadowing attenuation is denoted by $L_{\text{shadow}}\,[\text{dB}]$ and $L_{\text{path}}\,[\text{dB}]$ is the path-loss.

In combination with the link-level results from the previous chapters one may calculate the percentage of the cell area in which a particular modulation scheme and coding rate can be supported. For this, the interference is modelled as gaussian noise of equivalent power, which is a worst-case assumption. Then, assuming a frame-error rate of 10%, the corresponding E_{b}/N_0 values from the link-level simulations are converted to equivalent SINR values:

$$\text{SINR}_{\text{DL}} = \frac{M \times R \times K}{Q}\frac{E_{\text{b}}}{N_0} \tag{5.8}$$

for a 2^M-QAM modulated radio-link where Q is the spreading factor, R is the coding rate and K is the number of used spreading codes.

To assess the viability of higher order modulations in realistic conditions, a 2-cell indoor scenario and a 7-cell urban scenario are investigated. At the mobile, the proposed space-time receiver employing space-time equalisation followed by detection based on the MS-PPIC algorithm and turbo decoding is considered. For 16- and 64-QAM modulations, the proposed layered encoding scheme is employed. The following results can be considered as worst-case scenarios, since all base-stations always transmit with full power (no power control), the system is fully loaded and a conservative antenna gain at the base-station is assumed.

Scenario 1: 2-cell indoor scenario

For Scenario 1, two cells with a cell radius of 100 m and transmitter powers of ${}^{(u)}P_{\mathrm{Tx}} = 0.25\ W$ per base-station are considered. For the investigation, a transmitter gain (antenna gain − cable loss) of $G_{\mathrm{ant}}^{\mathrm{BS}} - L_{\mathrm{cable}}^{\mathrm{BS}} = 5$ dB is assumed. At the receiver, the gain is $G_{\mathrm{ant}}^{\mathrm{UE}} - L_{\mathrm{cable}}^{\mathrm{UE}} = 0$ dB. The shadowing attenuation L_{shadow} is log-normal distributed with zero mean and variance σ^2, where $\sigma = 12$ dB with a correlation of 50% across the base-stations. The path loss is $L_{\mathrm{path}} = 37 + 30 \log_{10}(d)$ where d is the distance in metres.

The system-level results for Scenario 1 are shown in Figure 5.21 for coding rate $1/3$, and in Figure 5.22 for coding rate $1/2$. For coding rate $1/3$, 4-QAM modulation is supported in the whole cell. 16-QAM is supported in 86% (up to 93 m from the base-station) and 64-QAM is supported in 65% (up to 81 m from the base-station) of the cell. For coding rate $1/2$, 4-QAM modulation is supported in 97% (98 m), 16-QAM is supported in 75% (87 m) and 64-QAM is supported in 40% (63 m) of the cell area. For both coding rates it can be seen that even the higher order modulation schemes are usable in a substantial area of the cell for the indoor environment.

Scenario 2: 7-cell urban scenario

In this scenario, 7 cells with a cell radius of 1500 m and transmitter powers of ${}^{(u)}P_{\mathrm{Tx}} = 20\ W$ per base-station are considered. The transmitter gain (antenna gain − cable loss) is $G_{\mathrm{ant}}^{\mathrm{BS}} - L_{\mathrm{cable}}^{\mathrm{BS}} = 11$ dB. At the receiver, the gain is $G_{\mathrm{ant}}^{\mathrm{UE}} - L_{\mathrm{cable}}^{\mathrm{UE}} = 0$ dB. As before, the shadowing attenuation L_{shadow} is log-normal distributed with zero mean and variance σ^2, where $\sigma = 8$ dB with a correlation of 50% across the base-stations. Here, the path-loss is assumed to be $L_{\mathrm{path}} = 29 + 35 \log_{10}(d)$

The simulation results for Scenario 2 are shown in Figure 5.23 for coding rate $1/3$, and in Figure 5.24 for coding rate $1/2$. Since the simulation parameters are chosen conservatively, and no power-control is used at the base-station, 4-QAM modulation is supported in only 95% (1462 m) of the cell for coding rate $1/3$. 16-QAM is supported in 55% (1112 m) and 64-QAM is supported in 29% (808) of the cell. When coding rate $1/2$ is employed, 4-QAM modulation is supported in 80% (1341 m), 16-QAM is supported in 40% (949 m), and 64-QAM is supported in 12% (520 m) of the cell. As in Scenario 1, the system-level simulations show that high-order modulation schemes such as 16- and 64-QAM can be used in a significant area of the cell.

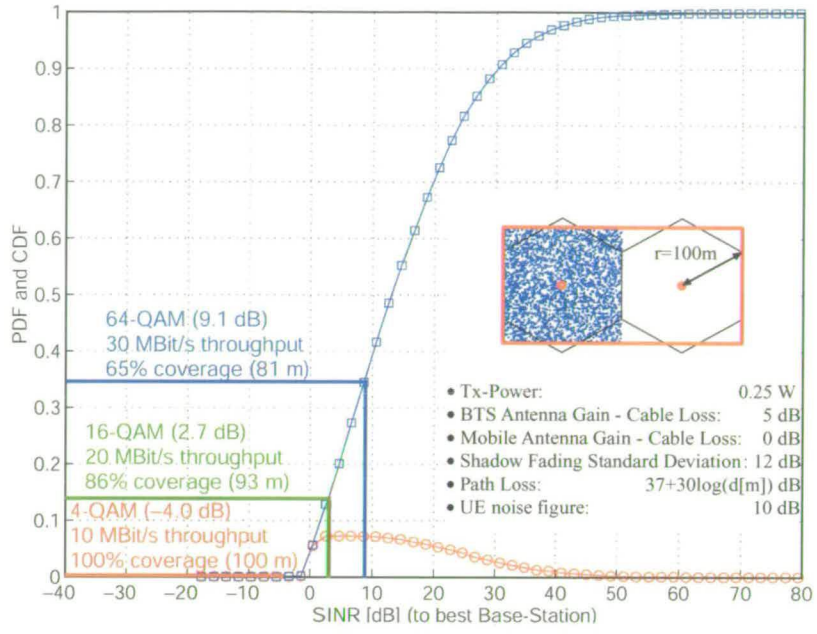


Figure 5.21: *pdf, cdf and the resulting system coverage for a 2 cell indoor scenario 4-QAM, 16-QAM and 64QAM modulation for a* 4×4 *antenna MIMO link with rate 1/3 coding.*

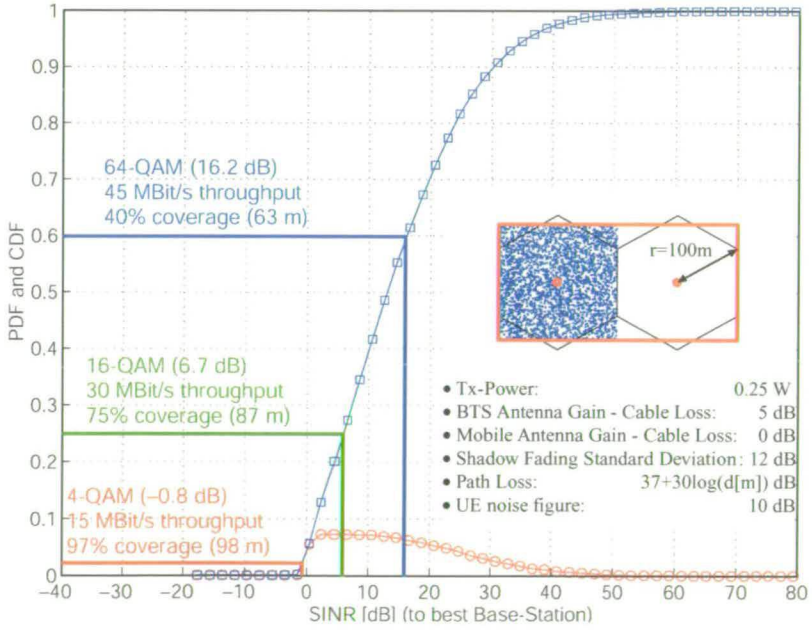


Figure 5.22: *pdf, cdf and the resulting system coverage for a 2 cell indoor scenario 4-QAM, 16-QAM and 64QAM modulation for a* 4×4 *antenna MIMO link with rate 1/2 coding.*

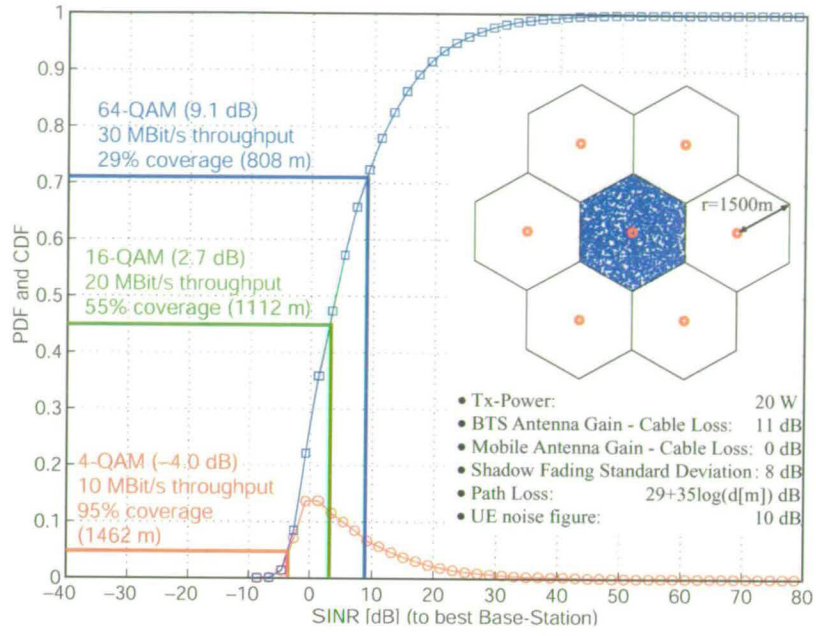


Figure 5.23: *pdf, cdf and the resulting system throughput for a 7 cell urban scenario 4-QAM, 16-QAM and 64QAM modulation for a* 4×4 *antenna MIMO link with rate 1/3 coding.*

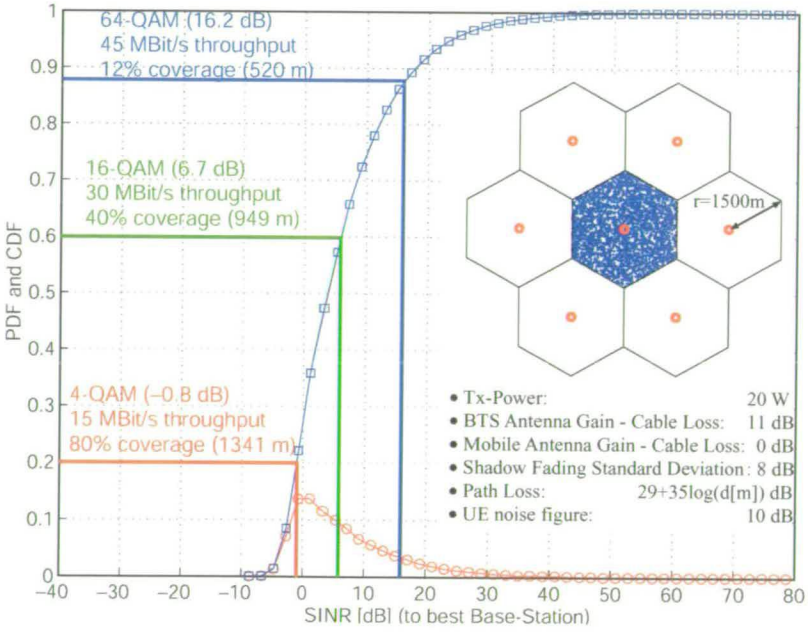


Figure 5.24: *pdf, cdf and the resulting system throughput for a 7 cell urban scenario 4-QAM, 16-QAM and 64QAM modulation for a* 4×4 *antenna MIMO link with rate 1/2 coding.*

5.6 Summary and conclusions

In this chapter, a novel layered encoding scheme was proposed which enables high-performance MIMO receivers for high-order modulations at a very low computational complexity. Through a judicious encoding process, layered encoding makes it possible to treat a 2^{2M}-QAM modulation as a set of M interdependent 4-QAM modulations. The proposed scheme solves the problems with the exponential growth in complexity of the APP detector and also the performance issues with low complexity detection algorithms such as the MS-PPIC.

It was shown that, in a non-CDMA ergodic scenario with different numbers of transmitting and receiving antennas, an APP based receiver with layered encoding can achieve the same performance as the full APP receiver without layered encoding. With layer-specific coding rates, the proposed scheme can even outperform the full APP at a significantly lower computational complexity.

The simulation results with CDMA and TU channel profiles, including space-time equalisation, de-spreading and pre-whitening followed by detection and turbo decoding, show that the receivers with the proposed layered encoding scheme offer the same performance as the full APP based receiver. With variable coding rates for the layers, the proposed receivers even outperform the APP reference by 1 dB for 16-QAM modulation at only 0.8% of the original APP complexity. The layered encoding scheme is highly scalable and shows similar results for 64-QAM. Here, variable coding rates for the layers can further improve the FER performance by 2 - 3.5 dB. A comparison with the full APP receiver is not possible for 64-QAM due to the 21.8×10^3 times higher computational complexity compared to the layered encoding scheme.

Finally, the system throughput and the applicability of high-order modulations in MIMO systems was investigated using system-level simulations for a 2-cell indoor scenario and a 7-cell urban scenario. The results show, that for a 4×4 antenna MIMO system with 64-QAM modulation and rate 1/3 coding, information troughputs of up to 30 MBit/s (information bits) can be achieved. For coding rate 1/2, information troughputs of up to 45 MBit/s are possible. In addition, the system-level simulations indicate that higher order modulation schemes such as 16-QAM and even 64-QAM can be used for MIMO radio communications in a substantial area of the cell.

Chapter 6
Summary and conclusions

6.1 Results of the thesis

In this chapter, a brief summary of the contents and the results of the thesis is provided. As outlined in Chapter 1, the thesis focuses on low complexity MIMO receiver architectures and its components in the context of HSDPA for the high-speed downlink of UMTS. The thesis is organised in 4 main chapters. In Chapters 2 and 3, different components for detection and decoding were investigated. In Chapter 4, different MIMO receiver architectures were proposed using the components from the previous chapters for 4-QAM modulation. Finally, in Chapter 5 the complexity and performance problems for higher order modulations such as 16- and 64-QAM were solved. The detailed results of the chapters are described next.

In **Chapter 2**, the MIMO signal model was introduced and details about the interleaving and antenna multiplexing as well as the channel simulations were given. It was shown that space-time equalisation can significantly reduce the detector complexity in dispersive channels. Block based and transversal MMSE equalisation was investigated for different block or filter lengths. It was demonstrated that the transversal equaliser always outperforms its block-based counterpart, since it avoids inaccuracies at the block edges. Then, the optimum APP detection algorithm and its Max-Log variant were presented. These detectors were used as reference throughout the thesis. The optimum detector has the disadvantage that its computational complexity grows exponentially with the number of transmitter antennas as well as with the modulation order. Multi-stage partial parallel interference cancellation (MS-PPIC) was proposed as a low complexity alternative to the APP detector and the coefficients for non-linear cancellation were derived. It was shown that the MS-PPIC detector offers performance close to the APP detector, and furthermore outperforms the Max-Log-APP detector at low E_b/N_0. This could be achieved at a mere 25% of the Max-Log-APP complexity for a 4×4 antenna MIMO scenario. More importantly, the MS-PPIC is highly scalable, since its complexity grows only linearly with the number of antennas and not exponentially as for the APP and Max-Log-APP detectors. In addition, simple, matched filter based ordered serial interference cancellation was proposed as

an ultra-low complexity detector for receivers where *a priori* information is available from a decoder. Here, bit-based ordering and hard-cancellation is performed which results in only 5% of the Max-Log-APP complexity for a 4×4 antenna MIMO scenario. It was shown that, when no *a priori* information is available, the detector can not compete with the APP and MS-PPIC. However, the results in Chapter 4 and show that, the MF-SIC offers impressive performance, when used in an iterative receiver architecture.

In **Chapter 3** turbo decoding with the optimum Log-MAP and the Max-Log-MAP algorithms were presented and their performance and convergence was investigated using simulations and extrinsic information transfer (EXIT) charts. While the optimum Log-Map decoder offers the best performance, the Max-Log-MAP decoder has the advantage of lower computational complexity, and more importantly, it is insensitive to input scaling. Therefore, when the Max-Log-MAP decoder is used, it is not necessary to estimate the variance of the noise to convert the detector outputs into log-likelihood ratios (LLRs). To overcome the performance hit, maximum mutual information combining (MMIC) was proposed as a novel modification of the Max-Log-MAP turbo decoder. This was achieved by introducing iteration specific weight factors to correct the increasing bias in the extrinsic information resulting from the Max-Log approximation. Then, the extrinsic information, which becomes *a priori* information for the next component decoder is combined with the channel values in a way that maximises the mutual information between the combined term and the transmitted bits. A second contribution of this chapter was a method for offline calculation of the optimum weight factors according to the maximum mutual information criterion. It was shown that the proposed combining scheme improves the performance of a turbo decoder using Max-Log-MAP algorithm to within 0.05 dB of a turbo decoder using the optimum Log-MAP algorithm. The improved decoder retains the low complexity and insensitivity to input scaling which are inherent advantages of the Max-Log-MAP algorithm.

In **Chapter 4**, an overview of non-iterative and iterative receiver architectures based on the components discussed in the previous chapters was presented. Both a non-iterative and an iterative receiver architecture, based on space-time equalisation, de-spreading and pre-whitening followed by detection and decoding were proposed. Then, the proposed architectures and alternative architecture versions were compared in terms of performance and complexity using different detection algorithms. The simulation results have shown that the proposed architectures offer both the best performance and also a very low computational complexity. For the

non-iterative receiver, the proposed architecture with MS-PPIC detection offers clearly the best compromise between performance and complexity. While offering virtually the same performance as with the optimum APP detection algorithm, the use of the MS-PPIC detector reduces the total receiver complexity by 50% compared to a APP based receiver. The same complexity reduction was observed for the proposed iterative receiver architecture, where the detector can benefit from *a priori* information supplied by the decoder, when MS-PPIC detection was employed. While the Max-Log-APP based iterative receiver offers the best performance, simple, matched filter based successive interference cancellation was investigated, where despite the use of a simplified detection scheme, a high level of receiver performance is achieved via iterations with the decoder. In fact, it was shown that, via a combination of bit-level ordering, detection and cancellation, and the use of a novel soft-output combining technique which maximises the mutual information transfer in each iteration, the low-complexity iterative MF-SIC receiver can compete with the Max-Log-APP and MS-PPIC based iterative receivers. Finally, it was demonstrated that when using an iterative receiver architecture, the receiver performance can be improved by approximately 1 dB in comparison to the non-iterative receivers under investigation. However, the iterative architecture requires more than twice the computational complexity. Therefore, it is questionable if the additional complexity for the iterations is justified with respect to the performance gain.

In **Chapter 5**, the problem of the exponential growth in computational complexity with the modulation order of the APP detector was solved by introducing a novel layered encoding scheme. In addition, layered encoding also solves the performance issues with low complexity detection algorithms like the MS-PPIC when high-order modulations are employed. It was shown that in a non-CDMA ergodic scenario and different numbers of transmitting and receiving antennas, an APP based receiver with layered encoding can achieve the same performance as the full APP receiver without layered encoding. With layer-specific coding rates, the proposed scheme can even outperform the full APP at a significantly lower computational complexity. It was also shown that for a CDMA scenario with a TU channel, a receiver with layered encoding employing space-time equalisation followed by detection and decoding offers virtually the same performance as a receiver of equivalent architecture employing the full APP algorithm for detection. With variable coding rates for the layers, the proposed receivers even outperform the APP reference by 1 dB for 16-QAM modulation at only 0.8% of the original APP complexity. The layered encoding scheme is highly scalable and shows similarly good results for 64-QAM. Here, variable coding rates for the layers can further improve the FER

performance by 2 - 3.5 dB. For 64-QAM, the proposed scheme reduces the APP detector complexity by a factor of 21.8×10^3. In addition, the applicability of higher order modulation schemes was investigated with system-level simulations for a 2-cell indoor scenario and a 7-cell urban scenario. The simulations indicate that high-order modulations such as 16-QAM and 64-QAM can be used in a substantial area of the cell for both scenarios. With coding rate $1/2$ and 64-QAM modulation, the system is able to achieve a throughput of up to 45 MBit/s (information bits) for a 4×4 antenna MIMO link.

6.2 Contributions to knowledge

This section highlights the contributions to knowledge of this thesis.

- The coefficients for non-linear cancellation of the MS-PPIC detector were derived for the proposed application.
- Bit-based serial interference cancellation was proposed as a ultra-low complexity detector for iterative MIMO receivers.
- The Max-Log-MAP turbo decoder was improved by maximising the mutual information transfer from iteration to iteration.
- A method for offline computation of the optimal combining weights was presented for the improved Max-Log-MAP turbo decoder.
- A non-iterative receiver architecture was proposed which allows low-complexity and high-performance MIMO receivers.
- An iterative receiver architecture was proposed which can lead to further improvements in terms of performance.
- A soft-output combining technique was proposed which can mitigate errors resulting from incorrect cancellation in iterative receivers, when hard interference cancellation is employed at the detector.
- An extensive comparison of different non-iterative and iterative receiver architectures with different detection algorithms was presented.

- It was shown that the proposed iterative receiver architecture makes the use of ultra-low complexity interference cancelling detectors based on simple matched filtering possible and impressive performance results were presented.

- The problem of the exponential growth in computational complexity of the APP detector, at higher order modulations, was solved without any loss in performance by introducing a novel layered encoding scheme. This scheme also solves the performance and convergence problems of low complexity interference cancelling detectors such as the MS-PPIC and the MF-SIC.

- It was shown with system level simulations that high-order modulation such as 16- and 64-QAM can be used in substantial areas of the cell in a MIMO radio communication system. For this 2-cell indoor with a cell radius of 100m and a 7-cell urban scenario with a cell radius of 1500m were investigated.

Appendix A
Derivations

A.1 Derivation of the MMSE equaliser matrix

Given the received signal $\underline{r} = \boldsymbol{H}\underline{s} + \underline{n}$ where $\underline{s} = \sum_{k=1}^{K} \boldsymbol{C}_k \underline{x}_k$, the equalised signal can then be written as

$$\underline{e} = \boldsymbol{V}\underline{r} \tag{A.1}$$

where the equaliser matrix V is calculated using the MMSE criterion $\min_{\underline{s}} \mathrm{E}\,|\underline{s} - \underline{e}|^2$. Denoting the argument as

$$\xi = \mathrm{E}\,|\underline{s} - \underline{e}|^2 = \underbrace{\mathrm{E}\left(\underline{e}^{\mathrm{H}}\underline{e}\right)}_{(1)} + \mathrm{E}\left(\underline{s}^{\mathrm{H}}\underline{s}\right) - \underbrace{\mathrm{E}\left(\underline{e}^{\mathrm{H}}\underline{s}\right)}_{(2)} - \underbrace{\mathrm{E}\left(\underline{s}^{\mathrm{H}}\underline{e}\right)}_{(3)} \tag{A.2}$$

with

$$(1) \qquad \mathrm{E}\left(\underline{e}^{\mathrm{H}}\underline{e}\right) = \mathrm{E}\left[\left(\boldsymbol{V}\underline{r}\right)^{\mathrm{H}}\left(\boldsymbol{V}\underline{r}\right)\right] = \mathrm{E}\left[\underline{r}^{\mathrm{H}}\boldsymbol{V}^{\mathrm{H}}\boldsymbol{V}\underline{r}\right] = \mathrm{Tr}\left(\boldsymbol{V}\overline{\underline{r}\underline{r}^{\mathrm{H}}}\boldsymbol{V}^{\mathrm{H}}\right) \tag{A.3}$$

$$\frac{\partial}{\partial \boldsymbol{V}}\mathrm{E}\left(\underline{e}^{\mathrm{H}}\underline{e}\right) = 2\boldsymbol{V}\overline{\underline{r}\underline{r}^{\mathrm{H}}} \tag{A.4}$$

$$(2) \qquad \mathrm{E}\left(\underline{e}^{\mathrm{H}}\underline{s}\right) = \mathrm{E}\left[\left(\boldsymbol{V}\underline{r}\right)^{\mathrm{H}}\underline{s}\right] = \mathrm{E}\left[\underline{r}^{\mathrm{H}}\boldsymbol{V}^{\mathrm{H}}\underline{s}\right] = \mathrm{Tr}\left(\boldsymbol{V}\overline{\underline{s}\underline{r}^{\mathrm{H}}}\right) \tag{A.5}$$

$$\frac{\partial}{\partial \boldsymbol{V}}\mathrm{E}\left(\underline{e}^{\mathrm{H}}\underline{s}\right) = 2\overline{\underline{s}\underline{r}^{\mathrm{H}}} \tag{A.6}$$

$$(3) \qquad \mathrm{E}\left(\underline{s}^{\mathrm{H}}\underline{e}\right) = \mathrm{E}\left[\underline{s}^{\mathrm{H}}\boldsymbol{V}\underline{r}\right] = \mathrm{Tr}\left(\boldsymbol{V}\overline{\underline{s}\underline{r}^{\mathrm{H}}}\right) \tag{A.7}$$

$$\frac{\partial}{\partial \boldsymbol{V}}\mathrm{E}\left(\underline{s}^{\mathrm{H}}\underline{e}\right) = 0 \tag{A.8}$$

and

$$\overline{\underline{r}\underline{r}^{\mathrm{H}}} = \overline{\left(\boldsymbol{H}\underline{s} + \underline{n}\right)\left(\underline{s}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}} + \underline{n}^{\boldsymbol{H}}\right)} = \boldsymbol{H}\overline{\underline{s}\underline{s}^{\mathrm{H}}}\boldsymbol{H}^{\mathrm{H}} + \overline{\underline{n}\underline{n}^{\mathrm{H}}} = \boldsymbol{H}\boldsymbol{R}_s\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{R}_n \tag{A.9}$$

$$\overline{\underline{s}\underline{r}^{\mathrm{H}}} = \overline{\underline{s}\left(\underline{s}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}} + \underline{n}^{\boldsymbol{H}}\right)} = \overline{\underline{s}\underline{s}^{\mathrm{H}}}\boldsymbol{H}^{\mathrm{H}} = \boldsymbol{R}_s\boldsymbol{H}^{\mathrm{H}}. \tag{A.10}$$

The solution of the MMSE metric $\min_{\underline{s}} \mathrm{E}\,|\underline{s} - \underline{e}|^2 = \min_{\underline{s}} \xi$ is found for $\frac{\partial \xi}{\partial \boldsymbol{V}} = 0$:

$$\frac{\partial \xi}{\partial \boldsymbol{V}} = 2\boldsymbol{V}\left(\boldsymbol{H}\boldsymbol{R}_s\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{R}_n\right) - 2\left(\boldsymbol{R}_s\boldsymbol{H}^{\mathrm{H}}\right) = 0. \tag{A.11}$$

Therefore, the MMSE equaliser matrix V can be written as

$$\boldsymbol{V} = \boldsymbol{R}_s \boldsymbol{H}^{\mathrm{H}} \left(\boldsymbol{H} \boldsymbol{R}_s \boldsymbol{H}^{\mathrm{H}} + \boldsymbol{R}_n \right)^{-1}. \tag{A.12}$$

A.2 Non-linear soft cancellation

From (2.32), denoting the i^{th} element of $\underline{y}$ as $\underline{y}^{(i)}$ and the i^{th} row of $\boldsymbol{S}$ as $\underline{s}^{(i)\mathrm{T}}$, we have

$$y^{(i)}[0] = y^{(i)} = x^{(i)} + \underline{s}^{(i)\mathrm{T}} \underline{x} + \eta^{(i)} = x^{(i)} + v^{(i)}[0] \tag{A.13}$$

and it immediately follows that cancellation at the n^{th} stage of the detector should be of the form

$$\begin{aligned} y^{(i)}[n] &= y^{(i)}[0] - \underline{s}^{(i)\mathrm{T}} \varphi \{ \underline{\tilde{x}}[n-1] \} \\ &= x^{(i)} + v^{(i)}[n] \\ &= x^{(i)} + \underline{s}^{(i)\mathrm{T}} \left(\underline{x} - \varphi \{ \underline{\tilde{x}}[n-1] \} \right) + \eta^{(i)} \end{aligned} \tag{A.14}$$

where $\varphi \{ \underline{\tilde{x}}[n-1] \}$ is in general a non-linear function of tentative estimates, $\underline{\tilde{x}}[n-1]$, derived in the previous stage. One could ignore the non-linearity and simply use the tentative estimates $\underline{\tilde{x}}[n-1]$ directly in a linear cancellation process. It has been shown that (under certain constraints on the eigenvalues of $\boldsymbol{S}$) the resulting linear MS-PIC converges to the MMSE joint-detector as the number of stages approaches infinity [34]. At the other extreme, one could choose the function φ to be a mapping to the 4-QAM alphabet (i.e. a threshold operation). Such hard cancellation would perform well if and only if there was a high level of confidence regarding the reliability of tentative estimates $\underline{\tilde{x}}[n-1]$. In order to deal with cases where the tentative estimates are unreliable, one may instead use soft symbols. A soft symbol can be written as the expected value, $\mathrm{E}\{\underline{\tilde{x}}[n-1]\}$, of the tentative estimates in the cancellation process. Since $y^{(i)}[n-1] = x^{(i)} + v^{(i)}[n-1]$, then $\tilde{x}^{(i)}[n-1] = y^{(i)}[n-1]$ and assuming that the

noise plus interference term $v^{(i)}$ has a Gaussian distribution, it can readily be shown that

$$\begin{aligned}\varphi\left\{\tilde{x}^{(i)}[n-1]\right\} &\equiv \mathrm{E}\left\{\mathrm{Re}\left(\tilde{x}^{(i)}[n-1]\right)\right\} + j\,\mathrm{E}\left\{\mathrm{Im}\left(\tilde{x}^{(i)}[n-1]\right)\right\} \\ &\equiv \mathrm{E}\left\{\tilde{b}_0^{(i)}[n-1]\right\} + j\,\mathrm{E}\left\{\tilde{b}_1^{(i)}[n-1]\right\} \\ &\equiv \tanh\left\{\frac{1}{2}\Lambda\left(b_0^{(i)}[n-1]\right)\right\} + j\tanh\left\{\frac{1}{2}\Lambda\left(b_1^{(i)}[n-1]\right)\right\} \\ &\equiv \tanh\left\{g^{(i)}[n]\,\mathrm{Re}\left(y^{(i)}[n-1]\right)\right\} \\ &\quad + j\tanh\left\{g^{(i)}[n]\,\mathrm{Im}\left(y^{(i)}[n-1]\right)\right\} \end{aligned} \tag{A.15}$$

where λ is the log-likelihood ratio and

$$g^{(i)}[n] = \frac{2}{\mathrm{E}\left\{\left|v^{(i)}[n-1]\right|^2\right\}} \tag{A.16}$$

can be viewed as an antenna-dependent "softness" factor for the n^{th} stage. Factor $g^{(i)}[n]$ can be readily computed for the first stage:

$$g^{(i)}[1] = \frac{2}{\mathrm{E}\left\{\left|\underline{s}^{(i)\mathrm{T}}\underline{x} + \eta^{(i)}\right|^2\right\}} = \frac{2}{2\underline{s}^{(i)\mathrm{T}}\underline{s}^{(i)*} + \sigma_\epsilon^2 \boldsymbol{R}_{i,i}^{-1}} \tag{A.17}$$

where $\boldsymbol{R}_{i,i}$ the i^{th} diagonal element of $\boldsymbol{R}$. The computation of $g^{(i)}[n]$ is more involved for subsequent stages. Consequently, $g^{(i)}[1]$ is used for all stages $n = 1 \ldots N$. Though sub-optimal, this strategy should not significantly degrade performance in the SNR range of interest. Note that, when space-time equalisation followed by de-spreading and pre-whitening is applied, $\sigma_\epsilon^2 = 1$.

A.3 LLR computation from matched filter outputs

Considering a single component model, the received signal can be written as

$$\underline{r} = \underline{a}x + \underline{v} \tag{A.18}$$

where x is the transmitted symbol, $\underline{a}$ is the signature and $\underline{v}$ is the noise vector, where $\mathrm{E}\{\underline{v}\underline{v}^{\mathrm{H}}\} = \sigma^2\boldsymbol{I}$ and v_i are independent, zero-mean, complex Gaussian random variables. Note that $\underline{v}$ here also represents interference from other co-channel transmitted symbols.

The normalised matched filter output is given by

$$y = \|\underline{a}\|^{-2} \underline{a}^{\mathrm{H}} \underline{r} = x + \|\underline{a}\|^{-2} \underline{a}^{\mathrm{H}} \underline{v} = x + \eta \tag{A.19}$$

where

$$\mathrm{E}\{|\eta|^2\} = \sigma_\eta^2 = \mathrm{E}\{\|\underline{a}\|^{-2} \underline{a}^{\mathrm{H}} \underline{v} \, \|\underline{a}\|^{-2} \underline{v}^{\mathrm{H}} \underline{a}\} = \|\underline{a}\|^{-2} \sigma^2. \tag{A.20}$$

Then, the likelihood function which is the probability density function of the matched filter output y conditioned on the transmitted symbol x can be written as

$$\mathrm{f}(y|x) = \frac{1}{\pi\sigma_\eta^2} \exp\left\{\frac{-1}{\sigma_\eta^2} |\eta|^2\right\}. \tag{A.21}$$

The LLRs of the bits corresponding to the symbol x can be calculated in the same way as for the APP detector described above, by considering only the possible states of the modulation:

$$\begin{aligned}
\Lambda(b_m) &= \ln \frac{\mathrm{P}\{b_m = +1|y\}}{\mathrm{P}\{b_m = -1|y\}} = \ln \frac{\sum_{x|b_m=+1} \mathrm{P}\{x|y\}}{\sum_{x|b_m=-1} \mathrm{P}(x|y)} = \ln \frac{\sum_{x|b_m=+1} \mathrm{f}(y|x)\,\mathrm{P}\{x\}}{\sum_{x|b_m=-1} \mathrm{f}(y|x)\,\mathrm{P}\{x\}} \\
&= \ln \frac{\sum_{x|b_m=+1} \frac{1}{\pi\sigma_\eta^2} \exp\left\{\frac{-1}{\sigma_\eta^2} |\eta|^2\right\} \mathrm{P}\{x\}}{\sum_{x|b_m=-1} \frac{1}{\pi\sigma_\eta^2} \exp\left\{\frac{-1}{\sigma_\eta^2} |\eta|^2\right\} \mathrm{P}\{x\}} \\
&= \ln \frac{\sum_{x|b_m=+1} \exp\left\{\frac{-1}{\sigma_\eta^2} |y-x|^2 + \ln \mathrm{P}\{x\}\right\}}{\sum_{x|b_m=-1} \exp\left\{\frac{-1}{\sigma_\eta^2} |y-x|^2 + \ln \mathrm{P}\{x\}\right\}}.
\end{aligned} \tag{A.22}$$

Using the Max-Log approximation, the LLRs can be written as

$$\begin{aligned}
\Lambda(b_m) &\approx \max_{\underline{x}|b_i=+1} \left\{\frac{-1}{\sigma_\eta^2} |y-x|^2 + \ln \mathrm{P}\{x\}\right\} - \max_{\underline{x}|b_i=-1} \left\{\frac{-1}{\sigma_\eta^2} |y-x|^2 + \ln \mathrm{P}\{x\}\right\} \\
&= \min_{\underline{x}|b_i=-1} \left\{\frac{1}{\sigma_\eta^2} |y-x|^2 - \ln \mathrm{P}\{x\}\right\} - \min_{\underline{x}|b_i=+1} \left\{\frac{1}{\sigma_\eta^2} |y-x|^2 - \ln \mathrm{P}\{x\}\right\}
\end{aligned} \tag{A.23}$$

Now, assuming equi-probable symbols, the LLRs are

$$\Lambda(b_m) \approx \min_{\underline{x}|b_i=+1} \left\{\frac{1}{\sigma_\eta^2} |y-x|^2\right\} - \min_{\underline{x}|b_i=-1} \left\{\frac{1}{\sigma_\eta^2} |y-x|^2\right\}. \tag{A.24}$$

When 4-QAM modulation is employed as shown in Figure A.1, the LLR computation can be simplified further. The following example shows the calculation of the LLR for the bit b_1 which is represented by the real part of the symbol x. The LLR for bit b_1 can be calculated in the same way:

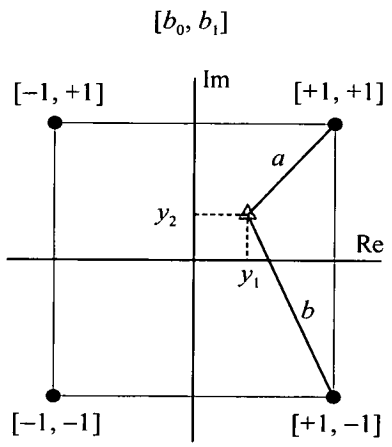


Figure A.1: *LLR calculation for a 4-QAM symbol.*

$$\begin{aligned}
\Lambda(b_1) &\approx \min_{\underline{x}|b_i=+1}\left\{\frac{1}{\sigma_\eta^2}|y-x|^2\right\} - \min_{\underline{x}|b_i=-1}\left\{\frac{1}{\sigma_\eta^2}|y-x|^2\right\} = b^2 - a^2 \\
&= \frac{1}{\sigma_\eta^2}\left\{\left[(1+y_2)^2 + (1-y_1)^2\right] - \left[(1-y_2)^2 + (1-y_1)^2\right]\right\} \\
&= \frac{1}{\sigma_\eta^2}\left\{(1+y_2)^2 - (1-y_2)^2\right\} \\
&= \frac{1}{\sigma_\eta^2}\left\{1 + y_2^2 + 2y_2 - 1 - y_2^2 + 2y_2\right\} \\
&= \frac{4}{\sigma_\eta^2}\mathrm{Re}(y)\,. \qquad \text{(A.25)}
\end{aligned}$$

Now considering the more general case, where the matched filtering is not normalised. Then the matched filter output can be written as

$$\mathring{y} = \underline{a}^{\mathrm{H}}\underline{r} = \underline{a}^{\mathrm{H}}\underline{a}\,x + \underline{a}^{\mathrm{H}}\underline{v} = \|\underline{a}\|^2 x + \mathring{\eta} = \frac{1}{\|\underline{a}\|^{-2}} y \qquad \text{(A.26)}$$

where

$$\mathrm{E}\{|\mathring{\eta}|^2\} = \sigma_{\mathring{\eta}}^2 = \mathrm{E}\{\underline{a}^{\mathrm{H}}\underline{v}\underline{v}^{\mathrm{H}}\underline{a}\} = \|\underline{a}\|^2\sigma^2 = \frac{1}{\|\underline{a}\|^{-4}}\sigma_\eta^2. \qquad \text{(A.27)}$$

Substituting $y = \|\underline{a}\|^{-2}\mathring{y}$ and the variance $\sigma_\eta^2 = \|\underline{a}\|^{-4}\sigma_{\mathring{\eta}}^2$ in Equation A.25, the LLRs for the

unnormalised matched filter are

$$\begin{aligned}\Lambda(b_1) &\approx \frac{4}{\sigma_\eta^2}\mathrm{Re}(y) \\ &= \frac{4}{\|\underline{a}\|^{-4}\sigma_{\mathring{\eta}}^2}\mathrm{Re}\left(\|\underline{a}\|^{-2}\mathring{y}\right) \\ &= \frac{4\|\underline{a}\|^2}{\sigma_{\mathring{\eta}}^2}\mathrm{Re}(\mathring{y}).\end{aligned} \tag{A.28}$$

It is shown that, for 4-QAM, the matched filter operation generates soft outputs which can be converted to LLRs by a simple scaling operation. Note that for Max-Log-MAP decoding no correction of the soft-outputs is required, since the decoder is insensitive to input scaling.

A.4 Calculation of the MIMO capacity limit for ergodic channels

The capacity limit for an ergodic MIMO radio link can be calculated as follows. Firstly, the capacity C for the MIMO system under investigation is determined:

$$C = RN_{\mathrm{T}}M \qquad \text{[bits / channel use]} \tag{A.29}$$

where R is the coding rate, N_{T} is the number of transmitter antennas and M is the number of bits per symbol (2 for 4-QAM). Based on the capacity, the Shannon limit in terms of E_{s}/N_0 can be calculated using the equation

$$C = \mathrm{E}\log_2 \det\left\{\boldsymbol{I}_{N_{\mathrm{R}}} + \frac{E_{\mathrm{s}}/N_0}{N_{\mathrm{T}}}\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}\right\} \tag{A.30}$$

where E_{s} is the energy per symbol interval, and therefore $E_{\mathrm{s}}/N_{\mathrm{T}}$ is the signal energy per transmitted QAM constellation symbol. Then, by definition the resulting capacity limit can be converted from E_{s}/N_0 to E_{b}/N_0:

$$\left.\frac{E_{\mathrm{b}}}{N_0}\right|_{\mathrm{dB}} = \left.\frac{E_{\mathrm{s}}}{N_0}\right|_{\mathrm{dB}} + 10\log_{10}\frac{1}{RN_{\mathrm{T}}M} \tag{A.31}$$

where E_{b}/N_0 is the transmitted energy per information bit.

Example: $R = 1/2$, $N_\mathrm{T} = 4$, $M = 2$

This scenario results in a channel capacity of $C = 4$ bits per channel use. Equation A.30 is solved numerically, resulting in a Shannon capacity limit of $E_\mathrm{s}/N_0 = 1.19$ dB. Converted to E_b/N_0, this corresponds to a value of -4.83 dB.

Appendix B
Awards, patent applications and publications

B.1 Awards

The following awards were received by the authors Holger Claussen, Reza Karimi and Bernard Mulgrew:

- *Excellent Paper Award* for the paper "Improved Max-Log-MAP Turbo Decoding using Maximum Mutual Information Combining" at the 14^{th} IEEE International Symposium on Indoor and Mobile Radio Communications PIMRC 2003, Beijing, China.
- *Best Paper Award* for the paper "Layered Encoding for iterative 16- and 64-QAM MIMO Receivers" at the 5^{th} European Personal Mobile Communications Conference EPMCC 2003, Glasgow, UK.

B.2 Patent applications

Patent applications were filed by Lucent Technologies under the titles:

- Improved Max-Log-MAP turbo decoding using maximum mutual information combining.
- Layered encoding for APP and neural network based low complexity MIMO receivers for high-order modulations (16-QAM, 64-QAM, 256-QAM).
- A radio telecommunication system operative by iterative determination of soft estimates and a corresponding method.
- A radio telecommunications receiver operative to receive digital data symbols or bits of iterative determination of soft estimates and corresponding method.

B.3 Publications

During the course of the thesis, the following papers were published (copies of the papers are attached at the end of the thesis):

- H. R. Karimi, H. Claussen, "Impact of Equalizer Output Modelling Errors on MIMO Receivers based on APP and PIC detection", *Globecom 2003, San Francisco, USA.*
- H. Claussen, H. R. Karimi, B. Mulgrew, "High Performance MIMO Receivers based on Multi-Stage Partial Parallel Interference Cancellation", *Vehicular Technology Conference VTC 2003 fall, Orlando, USA.*
- H. Claussen, H. R. Karimi, B. Mulgrew, "Improved Max-Log-MAP Turbo Decoding using Maximum Mutual Information Combining", *Personal Indoor Mobile Radio Communications Conference PIMRC 2003, Beijing, China.*
- H. Claussen, H. R. Karimi, B. Mulgrew, "Layered Encoding for 16- and 64-QAM Iterative MIMO Receivers", *5th European Personal Mobile Communications Conference EPMCC 2003, Glasgow, UK.*
- H. Claussen, H. R. Karimi, B. Mulgrew, "A Low Complexity Iterative Receiver based on Successive Cancellation for MIMO", *Conference on Personal Wireless Communications PWC 2002, Singapore.*
- H. Claussen, B. Mulgrew, H. R. Karimi, "Performance Optimization of Successive Interference Cancellation Detectors", *World Wireless Congress 2002, San Francisco, USA.*

B.4 Papers in review

The following papers are submitted for publication and are currently in review.

- H. Claussen, H. R. Karimi, B. Mulgrew, "Improved Max-Log-MAP Turbo-decoding via Maximization of Mutual Information Transfer", *submitted to EURASIP Journal on Applied Signal Processing, Special Issue on Turbo Processing.*
- H. Claussen, H. R. Karimi, B. Mulgrew, "Layered Encoding for Low Complexity Detection of High-Order Modulations in MIMO Channels", *submitted to IEEE International*

Symposium on Spread Spectrum Techniques and Applications ISSSTA 2004, Sydney, Australia.

References

[1] 3GPP, *Technical Specification Group Radio Access Network, High Speed Downlink Packet Access: Physical Layer Aspects (Rel.5)*, 2002.

[2] G. J. Foschini, "Layered space-time archtiecture for wireless communications in a fading environment when using multiple antennas," *Bell Labs Technical Journal*, vol. 6, pp. 41–59, 1996.

[3] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," in *Wireless Personal Communications*, pp. 311–335, 1998.

[4] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, November 1999.

[5] D. Garret, L. Davis, G. Woodward, and M. Sandell, "VLSI implementation of a maximum likelihood a posteriori probability detector for BLAST to support HSDPA for 3GPP," tech. rep., Bell Labs, Lucent Technologies, August 2001.

[6] L. Mailaender, "Linear MIMO equalization for CDMA downlink signals," tech. rep., Bell Labs, Lucent Technologies, September 2002.

[7] H. Claussen, H. R. Karimi, and B. Mulgrew, "High-performance MIMO receivers based on multi-stage partial parallel interference cancellation," in *Vehicular Technology Conference, Orlando, USA*, October 2003.

[8] A. Klein, G. K. Kaleh, and P. W. Baier, "Zero forcing and minimum mean-square-error equalization for multiuser detection in code-division multiple-access channels," *IEEE Transactions on Vehicular Technology*, vol. 45, pp. 276 – 287, May 1996.

[9] A. Klein, "Data detection algorithms specially designed for the downlink of CDMA mobile radio systems," in *Vehicular Technology Conference*, vol. 1, pp. 203–207, May 1997.

[10] H. R. Karimi and H. Claussen, "Impact of equalizer output modelling errors on MIMO receivers based on APP and PIC detectors," in *Globecom, San Francisco, USA*, October 2003.

[11] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Transactions on Information Theory*, vol. 42, pp. 429–445, March 1996.

[12] J. A. Erfanian, S. Pasupathy, and G. Gulot, "Reduced complexity symbol detectors with parallel structure for ISI channels," *IEEE Transactions on Communications*, vol. 42, pp. 1661–1671, February 1994.

[13] B. Vucetic and J. Juan, *Turbo Codes*. Kluwer Academic Publishers, 2000.

[14] S. Moshavi, "Multi-user detection for DS-CDMA communications," *IEEE Communications Magazine*, vol. 34, pp. 124–136, October 1996.

[15] S. Verdu, *Multiuser Detection.* Cambridge University Press, 1998.

[16] P. S. Guinand, R. W. Kerr, and M. Moher, "Serial interference cancellation for highly correlated users," in *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, pp. 133–136, August 1999.

[17] H. Claussen, B. Mulgrew, and H. R. Karimi, "Performance optimization of successive interference cancellation detectors," in *World Wireless Congress WWC, San Francisco, USA*, pp. 797–802, May 2002.

[18] X. Li, H. Huang, G. J. Foschini, and R. A. Valenzuela, "Effects of iterative detection and decoding on the performance of BLAST," in *Globecom 2000, San Francisco, USA*, vol. 2, pp. 1061–1066, November 2000.

[19] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo-codes," in *IEEE International Conference on Communications, Geneva*, vol. 2, pp. 1064–1070, May 1993.

[20] P. Robertson, E. Villebrun, and P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain," in *IEEE International Conference on Communication*, vol. 2, pp. 1009–1013, June 1995.

[21] J. Vogt and A. Finger, "Improving the Max-Log-MAP turbo-decoder," *Electronic Letters*, vol. 36, pp. 1937–1939, November 2000.

[22] H. Claussen, H. R. Karimi, and B. Mulgrew, "Improved Max-Log-MAP turbo-decoding using maximum mutual information combining," in *Personal Indoor Mobile Radio Communications Conference PIMRC 2003, Beijing, China*, September 2003. The Authors received the Excellent Paper Award.

[23] R. H. Morelos-Zaragoza, *The Art of Error Correcting Coding.* John Wiley & Sons, Ltd, 2002.

[24] S. ten Brink, "Convergence behaviour of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, pp. 1727–1737, October 2001.

[25] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," in *IEEE International Symposium on Information Theory*, vol. 47, pp. 599–618, February 2001.

[26] S. Haykin, *Neural Networks.* Prentice Hall, 2 ed., 1999.

[27] G. Strang, *Linear algebra and its applications.* Harcourt Brace Jovanovich, Publishers, 1986.

[28] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple antenna channel," *IEEE Transactions on Communications*, vol. 51, pp. 389–399, March 2003.

[29] H. Claussen, H. R. Karimi, and B. Mulgrew, "Layered encoding for 16- and 64-QAM iterative MIMO receivers," in *5th European Personal Mobile Communications Conference EPMCC 2003, Glasgow, UK*, April 2003. The Authors received the Best Paper Award.

[30] S. Marinkovic, B. Vucetic, and A. Ushirokawa, "Space-time iterative and multi-stage receiver structures for CDMA mobile communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 1594–1604, August 2001.

[31] B. Vucetic and J. Juan, *Space-Time Coding*. John Wiley & Sons, Ltd, 2003.

[32] U. Wachsmann, R. F. H. Fischer, and J. B. Huber, "Multilevel codes: Theoretical concepts and practical design rules," *IEEE Transactions on Information Theory*, vol. 45, pp. 1361–1391, July 1999.

[33] R. H. Morelos-Zaragoza, M. P. C. Fossorier, S. Lin, and H. Imai, "Multilevel coded modulation for unequal error protection and multistage decoding - Part I: Symmetric constellations," *IEEE Transactions on Communications*, vol. 48, pp. 204–213, February 2000.

[34] L. K. Rasmussen, D. Guo, T. J. Lim, and Y. Ma, "Aspects on linear parallel interference cancellation in CDMA," in *IEEE International Symposium on Information Theory*, p. 37, August 1998.

Published Papers

PERFORMANCE OPTIMIZATION OF SUCCESSIVE CANCELLATION DETECTORS

Holger Claussen [1], Bernard Mulgrew [1], H. Reza Karimi [2]

[1] Signals & Systems Group, University of Edinburgh, {Holger.Claussen, B.Mulgrew}@ee.ed.ac.uk
[2] Bell Labs Research, Lucent Technologies, HKarimi@lucent.com

Abstract *- Successive interference cancellation (SIC) is a well known technique to improve the bit-error rate of multiuser detectors. Since an incorrectly detected and cancelled symbol doubles the interference caused by the corresponding user, the performance gain is highly dependent on the cancellation order of the users. A popular approach is to order the detection and cancellation based on the received energy of each user, or more sophisticatedly, the average error probability. However, this is not necessarily optimal. In this paper, new ordering methods, based on the instantaneous error probabilities within each symbol period are proposed, to improve the performance of successive cancellation detectors.*

1. Introduction

Successive cancellation can significantly improve the performance of a receiver. Once a user symbol is detected, its contribution to the received signal is subtracted, thereby resulting in reduced multiple access interference (MAI) experienced by the remaining users. If the user symbol was detected correctly, its contribution could be cancelled completely, assuming the channel knowledge for the reconstruction of the signal was perfect. For incorrectly detected symbols, the MAI caused by the cancelled user is doubled. Successive cancellation can be applied to different types of detectors, such as matched filter (MF) or Rake detector, decorrelating or zero-forcing (ZF) detector, and the minimum mean-square error (MMSE) detector. To achieve the best results, the user with the lowest error probability must be detected and cancelled first. A popular approach is to order the detection according to the received energy. Although it improves the bit-error rate, it is not optimum, because it does not take the crosscorrelations between the spreading codes into account. Alternative ordering methods have been presented which correct the detected energies using the crosscorrelations between the users [1][3][4], based on the least mean-square error or the average error probability [5]. These methods improve the performance, but are still not optimal, because they neglect the fact that for most detectors the error probability for each user is not only dependent on the crosscorrelations, but also on the transmitted symbols of the interfering users. This problem does not exist for the decorrelating detector, because it eliminates the MAI completely. For this detector, the optimal order metric was proposed in [2].

In this paper, two novel ordering metrics, for the MF detector, and the MMSE joint detector are proposed. Both are based on the instantaneous error probabilities within each symbol period, dependent on the crosscorrelations between the spreading codes and the transmitted symbols. Since the instantaneous error probability for the decorrelating detector is independent of the transmitted symbols, this detector can not be improved by this scheme. For the sake of completeness, the optimal order metric for the decorrelating detector is presented as well.

2. System model

The transmission model for spreading, power control and channel simulation is shown in figure 1.

For each user k, the transmission data is BPSK modulated to symbols $\underline{b}^{(k)}$, and spread with a unique orthogonal variable spreading factor (OVSF) code $\underline{c}^{(k)}$ of spreading factor Q.

To evaluate the performance of the different order metrics, perfect power control within each timeslot is applied, so that

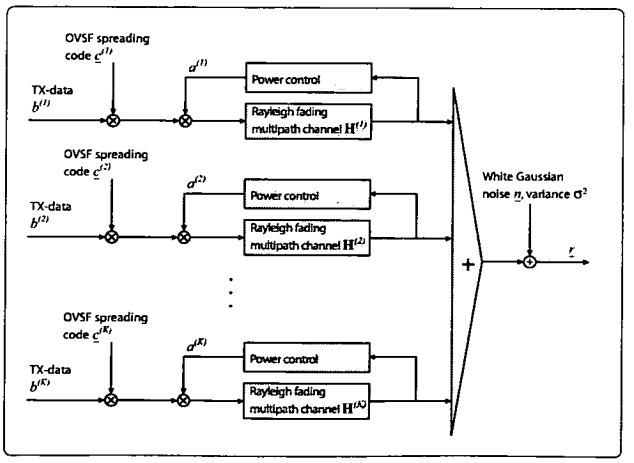


Figure 1: Overview of transmission model for K users.

all users are received with equivalent energy. This prevents trivial choices for the ordering. Therefore, the energy loss of each user due to the multipath is calculated separately without noise, and the spreading code $\underline{c}^{(k)}$ of each user k is adjusted by the correction amplitudes $a^{(k)}$ to result in the same received energy. The amplitudes $a^{(k)}$ for the power correction are updated for each timeslot, depending on the changing fading values for each path.

Each signal is transmitted through a Rayleigh fading dispersive multipath channel. The channel is simulated by a tapped delay model with independent fading values for each delay. After the channel simulation, the signals of all users are received at one antenna, and white Gaussian noise $\underline{n}$ with the variance σ^2 is added. This represents the uplink in a cellular system with synchronous received users.

An overview of the successive cancellation detector is shown in figure 2. The received signal $\underline{r}$, is fed into a detector which detects all transmitted symbols $x^{(1)}$... $x^{(K)}$ within a symbol period. Then, the most reliable symbol estimate is determined according to an order metric. Ideally, the user symbol with lowest error probability would be selected. The next step is to make a hard decision on the selected symbol, and to reconstruct its interference. Finally, the interference is subtracted from the received signal $\underline{r}$. If the decision on the selected symbol is correct, its MAI is cancelled completely, however a wrong decision doubles the interference caused by the detected symbol. Therefore, the order metric is of crucial importance for the performance of successive cancellation detectors. This detection and cancellation process is continued until all user symbols are detected.

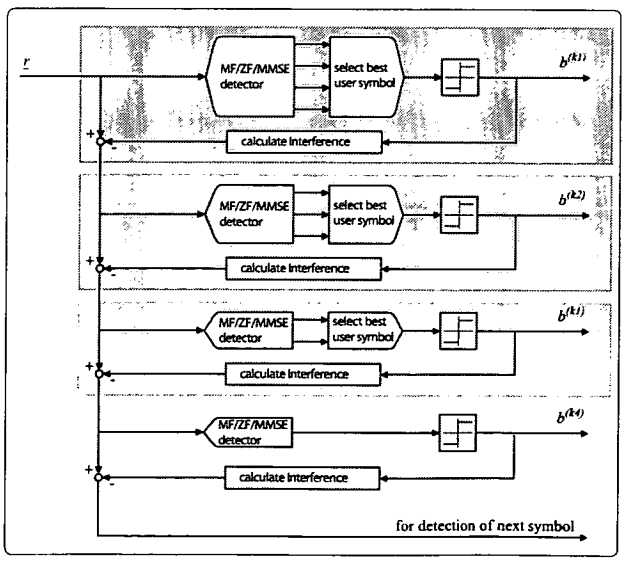


Figure 2: Successive interference cancellation overview (for 4 users).

The received signal vector $\underline{r}$ for the uplink can be described mathematically as follows:

$$\underline{r} = \mathbf{A}\underline{b} + \underline{n} \tag{1}$$

$$\text{with:} \quad \mathbf{A} = \left[\mathbf{H}^{(1)}\underline{c}^{(1)}a^{(1)},\ldots,\mathbf{H}^{(K)}\underline{c}^{(K)}a^{(K)}\right] \tag{2}$$

where:

$\underline{b}$	$\in \{-1,1\}^{K\times 1}$	transmitted symbols (BPSK)
$\mathbf{A}$	$\in C^{(Q+W-1)\times K}$	channel-code matrix
$\underline{n}$	$\in C^{(Q+W-1)\times 1}$	noise vector, $\mathrm{E}\{\underline{n}\underline{n}^H\} = \sigma^2\mathbf{I}$
$\underline{r}$	$\in C^{(Q+W-1)\times 1}$	received signal
$\mathbf{H}^{(k)}$	$\in C^{(Q+W-1)\times Q}$	channel matrix for user k
$\underline{c}^{(k)}$	$\in R^{Q\times 1}$	spreading code of user k

K is the number of users, Q is the spreading factor, W is the number of chip-spaced channel taps, and $a^{(k)}$ is calculated such that:

$$\left[\mathbf{A}^H\mathbf{A}\right]_{k,k} = 1 \quad \text{for each user } k \tag{3}$$

The subscripts k,k denote the element at the position (k,k) of the matrix $[\mathbf{A}^H\mathbf{A}]$.

For MF detection, the detector output vector $\underline{y}$ is:

$$\underline{y}_{MF} = \mathrm{Re}\left\{\mathbf{A}^H\underline{r}\right\} \tag{4}$$

$$\underline{b}_{MF} = \mathrm{sign}\left(\underline{y}_{MF}\right) \tag{5}$$

The sign of the detector output corresponds to the estimates of the transmitted symbols (equation 5).

For ZF detection, the detector output vector $\underline{y}$ is [6][7]:

$$\underline{y}_{ZF} = \left[\mathrm{Re}\left\{\mathbf{A}^H\mathbf{A}\right\}\right]^{-1}\mathrm{Re}\left\{\mathbf{A}^H\underline{r}\right\} \tag{6}$$

$$\mathbf{A} = \mathbf{A}_R + j\mathbf{A}_I$$

$$\mathbf{A}^H\underline{r} = \left(\mathbf{A}_R^T - j\mathbf{A}_I^T\right)\left(\underline{r}_R + j\underline{r}_I\right)$$

$$\mathrm{Re}\left\{\mathbf{A}^H\underline{r}\right\} = \mathbf{A}_R^T\underline{r}_R + \mathbf{A}_I^T\underline{r}_I$$

$$\underline{y}_{ZF} = \mathbf{R}_R^{-1}\mathbf{A}_R^T\underline{r}_R + \mathbf{R}_R^{-1}\mathbf{A}_I^T\underline{r}_I \tag{7}$$

$$\text{with:} \quad \mathbf{R} = \mathbf{A}^H\mathbf{A} \tag{8}$$

$$\underline{b}_{ZF} = \mathrm{sign}\left(\underline{y}_{ZF}\right) \tag{9}$$

Equation 7 represents the filter implementation of the ZF detector. This form is necessary when only the user of interest is detected and cancelled. The subscripts R and I denote the real, or imaginary part of a matrix or vector.

For MMSE detection, the detector output vector $\underline{y}$ is [6][7]:

$$\underline{y}_{MMSE} = \left[\mathrm{Re}\left\{ \mathbf{A}^H \mathbf{A} + \frac{\sigma^2}{2}\mathbf{I} \right\} \right]^{-1} \mathrm{Re}\left\{ \mathbf{A}^H \underline{r} \right\} \quad (10)$$

$$\underline{y}_{MMSE} = \mathbf{M}_R^{-1}\mathbf{A}_R^T \underline{r}_R + \mathbf{M}_R^{-1}\mathbf{A}_I^T \underline{r}_I \quad (11)$$

$$\text{with:} \quad \mathbf{M} = \mathbf{A}^H \mathbf{A} + \frac{\sigma^2}{2}\mathbf{I} \quad (12)$$

$$\underline{b}_{MMSE} = \mathrm{sign}\left(\underline{y}_{MMSE} \right) \quad (13)$$

The filter implementation of the MMSE detector is represented by equation 11.

3. Derivation of the order metrics

Since every wrong decision for detection and cancellation doubles the interference caused by the corresponding user, it is important to detect the most reliable user first. When the symbol of the user of interest is detected correctly, its signal can be reconstructed and cancelled, which leads to lower MAI for all remaining users. Therefore, the optimum order of detection and cancellation is that which minimizes the bit error rate of the user of interest in each cancellation step. As each code interferes with the others differently, the order must be recalculated after every cancellation step.

The probability of error is different for each type of detector. Hence, for the most common types, the MF detector, the ZF detector, and the MMSE detector, the optimal or near optimal ordering metrics are derived. This allows the optimization of successive cancellation detectors, based on these basic types. Since the error probabilities are dependent on the Q-function, the characteristics which are important for the ordering are reviewed.

3.1 The Q-Function

The error probabilities of the investigated detectors are based on the complementary Gaussian cumulative distribution function:

$$Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt \quad (14)$$

The important characteristics of the Q-function for the order of the error probabilities are:

- $Q(x)$ is monotonically decreasing
- $Q(x)$ is convex on the interval $(0,+\infty)$

3.2 MF detector

The average error probability of the MF detector for user 1 with one interfering user is:

$$P_{MF}^{(1)} = \frac{1}{2} Q\left(\frac{\rho_{11} - \mathrm{Re}\{\rho_{12}\}}{\sqrt{\frac{1}{2}\rho_{11}\sigma^2}} \right) + \frac{1}{2} Q\left(\frac{\rho_{11} + \mathrm{Re}\{\rho_{12}\}}{\sqrt{\frac{1}{2}\rho_{11}\sigma^2}} \right) \quad (15)$$

The first term describes the case when the signal of user 1 is attenuated by the interfering user, which increases the error probability for user 1. The second term represents the case, where the signal of user 1 is amplified by the interfering signal, which reduces the error probability for user 1. Both terms are weighted equally, because on average both effects occur with equal probability. As the Q-function is monotonically decreasing and convex in the interval $(0,+\infty)$, the first term is dominating for increasing real crosscorrelations, $\mathrm{Re}\{\rho_{12}\} = [\mathbf{R}_R]_{12}$ between the two users.

Figure 3 shows the conditional distributions of the MF detector output for user 1 with one interfering user. Dependent on the transmitted symbols, the distributions are shifted from their original positions by $+\delta$ or $-\delta$.

For the case of K-1 interfering users, the average error probability can be averaged over all possible interfering symbols:

$$P_{MF}^{(k)} = \frac{1}{2^{K-1}} \sum_{e_1} \cdots \sum_{e_{j\neq k}} \cdots \sum_{e_K} Q\left(\frac{[\mathbf{R}_R]_{k,k} + \sum_{j\neq k} [\mathbf{R}_R]_{k,j} e^{(j)}}{\sqrt{\frac{1}{2}\sigma^2 [\mathbf{R}_R]_{k,k}}} \right) \quad (16)$$

Where $e^{(j)}$ denotes the two states $\{-1,+1\}$ where either the user signal is attenuated or amplified by the interfering user. $[\mathbf{R}_R^{-1}]_{kk}$ denotes the element at the position (k,k) of the inverted real crosscorrelation matrix $\mathbf{R}_R^{-1}$ of the spreading codes.

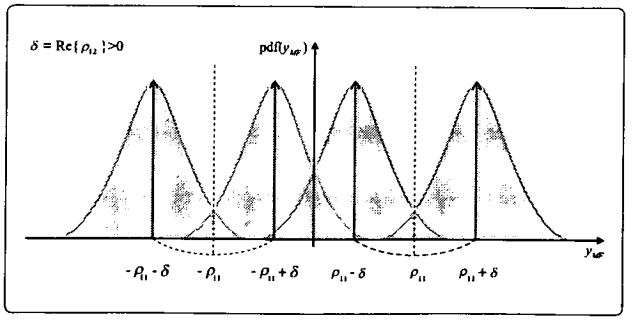


Figure 3: Conditional distributions of matched filter output for user 1 with one interfering user.

For successive cancellation, the instantaneous error probability is more significant than the average error probability. Since the transmitted symbols $b^{(k)}$ can be estimated for each user, the equation can be reduced from an exponential number of terms (2 to the power of K-1) to one term for the probability that user k is in error, given that all decisions about all users are correct.

$$P_{MF}^{(k)} = \mathrm{Q}\left(\frac{\left[\mathbf{R}_R\right]_{k,k} + b_{MF}^{(k)} \sum_{j \neq k} \left[\mathbf{R}_R\right]_{k,j} b_{MF}^{(j)}}{\sqrt{\frac{1}{2}\sigma^2 \left[\mathbf{R}_R\right]_{k,k}}} \right) \tag{17}$$

The optimal order for detection and successive cancellation requires that the user with the lowest error probability be detected first. The index $k_{opt,MF}$ of this user is given by:

$$k_{opt,MF} = \arg\max_k \left(\frac{\left[\mathbf{R}_R\right]_{k,k} + b_{MF}^{(k)} \sum_{j \neq k} \left[\mathbf{R}_R\right]_{k,j} b_{MF}^{(j)}}{\sqrt{\left[\mathbf{R}_R\right]_{k,k}}} \right) \tag{18}$$

The drawback of this order metric is that it is very sensitive to errors in $b_{MF}^{(k)}$. Since the interference is a sum of K-1 user signals, errors in $b_{MF}^{(j)}$ do not have such a large impact on the resulting order values. The order metric can be improved by using soft estimates, and made more robust without the influence of $b_{MF}^{(k)}$, assuming the worst case for $b_{MF}^{(k)}$:

$$k_{opt,MF} = \arg\max_k \left(\frac{\left[\mathbf{R}_R\right]_{k,k} - \left| \sum_{j \neq k} \left[\mathbf{R}_R\right]_{k,j} \tanh\left(\alpha y_{MF}^{(j)}\right) \right|}{\sqrt{\left[\mathbf{R}_R\right]_{k,k}}} \right) \tag{19}$$

with the scaling factor α=2 for the soft values. After each cancellation step the order metrics must be re-calculated, as each spreading code can be affected differently by the previously cancelled user.

3.3 ZF detector

The average error probability of the ZF detector for user 1 with K-1 interfering users is:

$$P_{ZF}^{(k)} = \mathrm{Q}\left(\frac{1}{\sqrt{\frac{1}{2}\sigma^2 \left[\mathbf{R}_R^{-1}\right]_{k,k}}} \right) \tag{20}$$

Because for the decorrelating detector, the error probability does not depend on the transmitted symbols, the optimal order mainly depends on the received amplitudes and the crosscorrelations which are constantly changing due to the Rayleigh fading multipath channel. Therefore, the order metric of the decorrelating detector can not be improved by taking the estimated symbols into account.

As for the MF detector, the index for the most reliable user $k_{opt,ZF}$ is given by:

$$k_{opt,ZF} = \arg\min_k \left(\left[\mathbf{R}_R^{-1}\right]_{k,k} \right) \tag{21}$$

Since the crosscorrelation matrix changes after each cancellation, its inverse must be updated in every cancellation step.

3.4 MMSE detector

The average error probability of the minimum mean-square error detector (MMSE) for user 1 with K-1 interfering users is:

$$P_{MMSE}^{(k)} = \frac{1}{2^{K-1}} \sum_{e_1} \cdots \sum_{e_{k-1}} \sum_{e_{k+1}} \cdots \sum_{e_K} \mathrm{Q}\left(\frac{\left[\mathbf{M}_R^{-1}\mathbf{R}_R\right]_{k,k} + \sum_{j \neq k} \left[\mathbf{M}_R^{-1}\mathbf{R}_R\right]_{k,j} e^{(j)}}{\sqrt{\frac{1}{2}\sigma^2 \left[\mathbf{M}_R^{-1}\mathbf{R}_R\mathbf{M}_R^{-1}\right]_{k,k}}} \right) \tag{22}$$

where $\mathbf{M}_R$ is the real part of the modified crosscorrelation matrix with increased diagonal elements of the MMSE detector. As in the case for the MF detector, the probability of an error can be reduced from the exponential number of terms to one term, when the transmitted symbols can be estimated. Therefore, the estimated instantaneous error probability can be written as:

$$P_{MMSE}^{(k)} = \mathrm{Q}\left(\frac{\left[\mathbf{M}_R^{-1}\mathbf{R}_R\right]_{k,k} + b_{MMSE}^{(k)} \sum_{j \neq k} \left[\mathbf{M}_R^{-1}\mathbf{R}_R\right]_{k,j} b_{MMSE}^{(j)}}{\sqrt{\frac{1}{2}\sigma^2 \left[\mathbf{M}_R^{-1}\mathbf{R}_R\mathbf{M}_R^{-1}\right]_{k,k}}} \right) \tag{23}$$

Again, the index for the most reliable user $k_{opt,MMSE}$ is given by:

$$k_{opt,MMSE} = \arg\max_k \left(\frac{\left[\mathbf{M}_R^{-1}\mathbf{R}_R\right]_{k,k} + b_{MMSE}^{(k)} \sum_{j \neq k} \left[\mathbf{M}_R^{-1}\mathbf{R}_R\right]_{k,j} b_{MMSE}^{(j)}}{\sqrt{\left[\mathbf{M}_R^{-1}\mathbf{R}_R\mathbf{M}_R^{-1}\right]_{k,k}}} \right) \tag{24}$$

As for the matched filter, the order metric can be improved by using soft estimates and made more robust against errors of $b_{MMSE}^{(k)}$, assuming the worst case of $b_{MMSE}^{(k)}$:

$$k_{opt,MMSE} = \arg\max_{k} \left(\frac{\left[\mathbf{M}_R^{-1}\mathbf{R}_R\right]_{k,k} - \left|\sum_{j\neq k}\left[\mathbf{M}_R^{-1}\mathbf{R}_R\right]_{k,j} \tanh\left(\alpha\, y_{MMSE}^{(j)}\right)\right|}{\sqrt{\left[\mathbf{M}_R^{-1}\mathbf{R}_R\mathbf{M}_R^{-1}\right]_{k,k}}} \right) \tag{25}$$

with the scaling factor α=2 for the soft values. The crosscorrelation matrixes and the order calculation for the next user must be updated after each cancellation step.

4. Performance comparison

To compare the ordering metrics, the MF detector, the ZF detector, and the MMSE detector are simulated with and without successive cancellation. For successive cancellation, order metrics based on the received energy (maxE), the least mean-sqare error (LMSE), and the proposed ordering based on the instantaneous error probability (IEP) are investigated.

For the simulations, the case of a transmission with 16 users, a spreading factor of 16 with OVSF codes, and a Rayleigh fading multipath channel with average chip spaced tap powers of [0 dB, -3 dB, -6 dB, -9 dB] at a mobile speed of 3 km/h is examined. The data is transmitted in time slots of 2560 chips at a chip rate of 3.84 MHz (according to the UMTS standard). Perfect power control is assumed, so that the order metrics which are based only on the received energy can not derive a definite order. Order metrics which are based on the least mean-square error take the crosscorrelations of the received signatures into account, and are able to adjust the order dependent on the changing channel conditions. It will be shown, that even when all users are received with equivalent energy, the detection order has a high influence on the bit-error rate of the receiver, and that the new proposed ordering metrics perform much better than the ones which are based on the least mean-square error or the highest received signal energy.

4.1 MF detector

Figure 4 shows the bit-error rate (BER) simulation results for the MF detector. Both successive cancelling versions of the detector offer better bit-error rates than the standard version, but their performance is greatly dependent on the applied ordering metric. Since the received power of all users is equivalent, detectors based on the received power

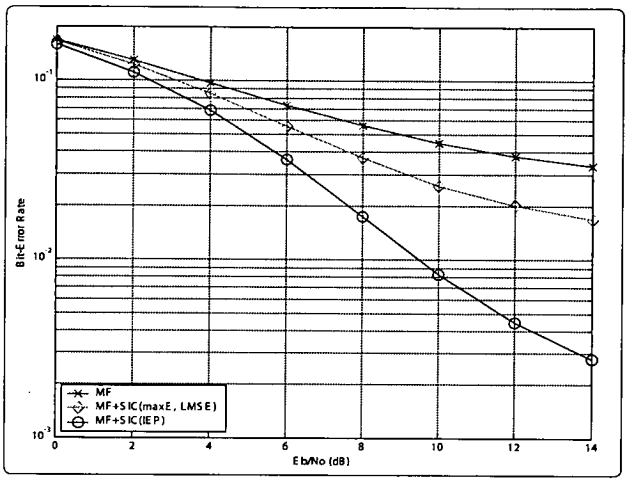


Figure 4: Comparison of MF detectors with successive cancellation.

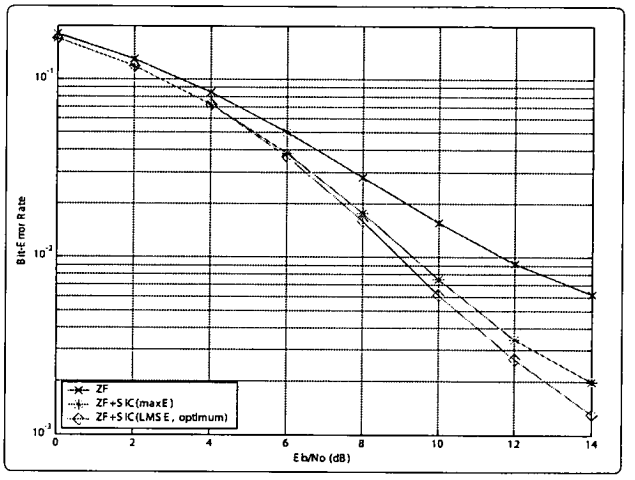


Figure 5: Comparison of ZF detectors with successive cancellation.

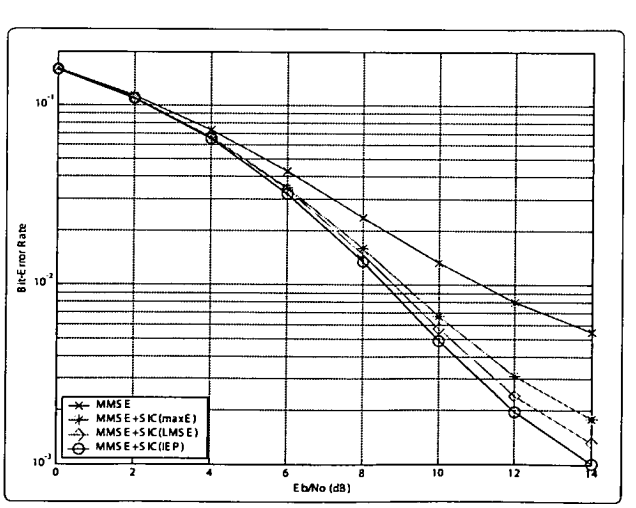


Figure 6: Comparison of MMSE detectors with successive cancellation.

either can not make a definite ordering decision, or will apply an order which disregards the transmitted symbols. For the MF detector, ordering according to the least mean-square error is equivalent to the ordering according to the highest received energy. Using the estimated transmitted symbols, to calculate the ordering metric based on the instantaneous error probability, improves the detector performance significantly. It can even outperform the standard MMSE detector, and in the BER range of interest between 10^{-1} and $10^{-1.8}$ offers a performance comparable to the MMSE SIC detector. This can be achieved with much lower computational complexity than for the ZF of MMSE detector, because no matrix inversions are required.

4.2 ZF detector

The simulation results for the ZF detector are shown in Figure 5. Since this detector completely suppresses the MAI, the error probability for the ZF detector is independent of the transmitted symbols of all users. Therefore, it cannot be improved by taking these estimates into account and the optimal order metric for the decorrelating detector is equivalent to ordering according to the least mean-square error. The price of complete MAI suppression is noise enhancement, which degrades the performance. The simulation results show, that the successive cancellation versions of the ZF detector perform much better than the joint detection version. In the BER range of interest, the MF detector with the proposed ordering metric outperforms the ZF+SIC detector at lower complexity.

4.3 MMSE detector

Figure 6 shows the simulation results for the MMSE detector. As for all other detectors, successive cancellation versions improves the bit-error rate in comparison to the joint detector significantly. Since the MMSE detector does not suppress the MAI completely, the performance degradation due to noise enhancement is lower than for the ZF detector. For higher E_b/N_0, the MMSE detector with the new ordering metric, based on the instantaneous error probability offers the best results. However, in the BER range of interest, the MF detector with the proposed ordering metric offers the same performance at lower complexity.

5. Conclusions

In this paper, two novel order metrics, for the MF detector and the MMSE detector were proposed. The metrics are based on the minimum instantaneous error probability within one symbol period, and take the estimated symbols for each user into account, which are available at each stage of the successive cancellation. Additionally the optimum ordering metric for the decorrelating detector was presented. As shown, successive cancellation, based on the proposed ordering metrics significantly improves the bit-error rate of the presented detectors. For the BER range of interest between 10^{-1} and $10^{-1.8}$ the MF SIC detector with the proposed order metric outperforms the ZF and MMSE joint detectors and is able to achieve the same performance as the MMSE SIC detector. These large improvements are especially interesting, because of the low complexity of this detector. For higher E_b/N_0, the MMSE detector with the proposed ordering metric offers the best results. Its performance was also improved, by optimizing the detection and cancellation order.

References

[1] S. Verdu, "Multiuser Detection", Cambridge University Press, 1998.

[2] G. D. Golden, C. J. Foschini, R. A. Valenzuela and P. W. Wolniansky, "Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture", Electronic Letters 1999 Vol. 35 No.1, pp.14-16.

[3] M. Debbah, B. Muquet, M. Courville, M. Muck, S. Simoens and P. Loubaton, "A MMSE Successive Interference Cancellation Scheme for a New Adjustable Hybrid Spread OFDM system", VTC2000 pp. 745-749.

[4] B. Hassibi, "A fast square-root implementation for BLAST", Signals, Systems and Computers, 2000. Conference Record of the Thirty-Fourth Asilomar Conference on , Volume: 2 , 2000 pp. 1255-1259 vol.2

[5] H.V. Poor and S. Verdu, "Probability of Error in MMSE Multiuser Detection" IEEE Trans. on Information Theory, May 1997, pp. 858-871.

[6] A. Klein, G. K. Kaleh, P. W. Baier, "Zero Forcing and Minimum Mean-Square-Error Equalization for Multiuser Detection in Code-Division Multiple-Access Channels", IEEE Trans. VT, Vol 45, No. 2, pp. 276-287, May 1996

[7] H. R. Karimi, "Multi-Rate Multi-User Joint-Detection for the Asynchonous UTRA-FDD WCDMA Uplink", Bell Labs Technical Memorandum, ITD-99-38342A, 10041690-991027-01TM, September 1999.

A Low Complexity Iterative Receiver based on Successive Cancellation for MIMO

Holger Claussen [1], Hamid Reza Karimi [2], Bernard Mulgrew [1]

[1] *Signals & Systems Group, University of Edinburgh, UK*
[2] *Bell Labs Research, Lucent Technologies, Swindon, UK*

Abstract: Turbo-encoded multiple-input multiple-output (MIMO) systems have recently been proposed for the support of high-speed downlink packet access (HSDPA) in UMTS, where the re-use of spreading codes across the transmitter antennas results in high levels of interference. The state of the art receiver for this system incorporates a channel equalizer, followed by an *a posteriori* probability (APP) detector and a turbo decoder. However, the complexity of APP detection can become prohibitive since it grows exponentially with the number of transmitter antennas and the modulation order. In this paper, a MIMO receiver is proposed which replaces the optimum but complex APP detector by successive interference cancellation (SIC) incorporating sub-optimal matched filter detection. Using convolutional encoding at the transmitter, the receiver performance is sustained via iterations between the simplified detector and the convolutional decoder. In combination with a proposed novel soft-output combining scheme, it is shown that the new receiver can outperform the APP-based receiver at a much lower complexity and with no need for channel equalization.

Key words: Iterative detection, MIMO, successive cancellation, serial interference cancellation, SIC, soft output combining, order metric, ordering

1. INTRODUCTION

Turbo-encoded multiple-input multiple-output (MIMO) systems have recently been proposed for the support of high-speed downlink packet access (HSDPA) in UMTS [1]. The concept here is to increase the achievable data rates for a particular user through a combination of code re-use across transmit antennas and higher-order modulation schemes. The code re-use inevitably results in high levels of interference at the mobile receiver, even under non-dispersive channel conditions. In order to tackle such high interference levels, receivers based on the optimal *a posteriori* probability (APP) detector [2] followed by turbo decoding have been proposed [3][4]. To cope with dispersive channels and in order to avoid sequence estimation, it is necessary to use an APP detector preceded by a matrix channel equaliser.

Essentially, the APP detector operates by computing soft-outputs for the transmitted bits which most closely match the received signal in an Euclidian sense. The

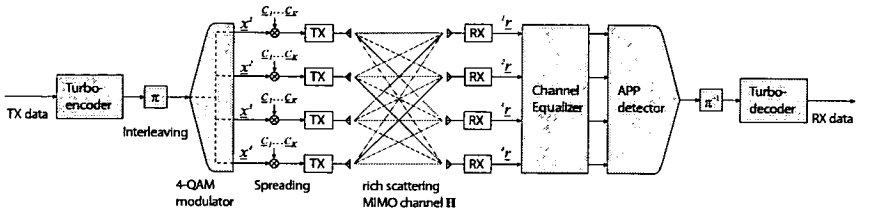


Figure 1. Equalized APP based receiver for a 4x4 MIMO system (Reference)

computational complexity of the APP detector is an exponential function of the total number of bits transmitted during a symbol epoch, which is equal to the product of the number of transmit antennas and the number of bits per symbol. Consequently, the complexity of the APP detector can become prohibitive for increasing numbers of transmit antennas and (perhaps more importantly) modulation orders. This inflexibility of the optimal APP detector has resulted in renewed interest in the use of sub-optimal but less complex MIMO detectors.

Successive interference cancellation (SIC) schemes have been considered for many years in the context of multi-user detection for the CDMA uplink [5][6][7]. These schemes combat interference by successively detecting and cancelling the influence of data streams from the received signal. The more reliable data streams are detected and cancelled first. In the context of MIMO receivers, the original BLAST detector [8] is essentially a SIC architecture incorporating ordering and detection based on the minimum mean-squared error (MMSE) criterion. Furthermore, significant performance improvements have been demonstrated through iterations between the BLAST detector and a convolutional decoder [9].

In this paper, a bit-based SIC scheme incorporating simple matched filters (MF) as the basic detection unit is considered as a receiver for a convolutionally-encoded MIMO link. The MF-SIC detector performs iterations with a convolutional decoder in conjunction with a novel soft-output combining technique. Convolutional coding is used, since it provides better convergence than turbo coding in iterative schemes. The combining acts to suppress instabilities caused by erroneously detected and cancelled bits. The resulting receiver architecture is highly scalable in terms of dealing with growing numbers of transmit antennas and high-order modulation schemes.

The proposed MF-SIC receiver is compared with the APP-based receiver considered for an equivalent turbo-encoded MIMO link [4] and is shown to achieve superior performance at a much lower complexity. The performance loss due to the use of a sub-optimal detector is regained via iterations with the decoder, enabled by the novel soft-output combining technique.

2. SIGNAL MODEL

Figures (1) and (2) illustrate the transmission scheme for the MIMO system under investigation. At the transmitter, user data is convolutional or turbo encoded and interleaved. The coded data stream is de-multiplexed into N_T sub-streams, corresponding to the N_T transmit antennas. Each sub-stream is then modulated on to NK 4-QAM symbols and subsequently spread by a factor Q via a set of K orthogonal spreading codes prior to transmission. Each transmitted spread stream then occupies N symbol

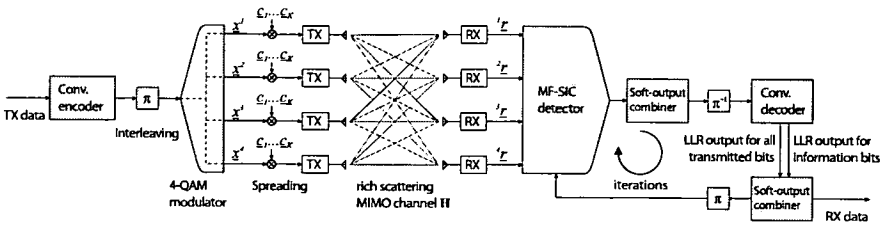


Figure 2. Iterative MF-SIC based receiver for a 4x4 MIMO system

intervals. Also note that the same set of K codes are re-used across all transmit antennas. Therefore, the MIMO propagation environment, which is assumed to exhibit significant multipath, plays a major role in achieving signal separation at the receiver. The transmitted signals are received by N_R receive antennas after propagation through dispersive radio channels with impulse response lengths of W chips. The received signal vector can then be modelled as follows:

$$\begin{bmatrix} {}^{1}\underline{r} \\ \vdots \\ {}^{N_R}\underline{r} \end{bmatrix} = \begin{bmatrix} {}^{1}\boldsymbol{H}^{1} & \cdots & {}^{1}\boldsymbol{H}^{N_T} \\ \vdots & \ddots & \vdots \\ {}^{N_R}\boldsymbol{H}^{1} & \cdots & {}^{N_R}\boldsymbol{H}^{N_T} \end{bmatrix} \sum_{k=1}^{K} \begin{bmatrix} \boldsymbol{C}'_k & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \boldsymbol{C}'_k \end{bmatrix} \begin{bmatrix} \underline{x}_k^{1} \\ \vdots \\ \underline{x}_k^{N_T} \end{bmatrix} + \begin{bmatrix} {}^{1}\underline{v} \\ \vdots \\ {}^{N_R}\underline{v} \end{bmatrix} \tag{1}$$

$$\underline{r} = \boldsymbol{H} \sum_{k=1}^{K} \boldsymbol{C}_k \underline{x}_k + \underline{v} \tag{2}$$

where: ${}^{m}\underline{r} \in C^{(QN+W-1)\times 1}$ Signal at Rx antenna m.

${}^{m}\underline{v} \in C^{(QN+W-1)\times 1}$ Noise + inter-cell interference at Rx antenna m.

$\underline{x}_k^n \in C^{N\times 1}$ Symbol sequence $[x_k^n(1) \ldots x_k^n(N)]^{\mathrm{T}}$ at Tx antenna n spread via k^{th} spreading code.

$\boldsymbol{C}'_k \in C^{QN\times N}$ Spreading matrix for k^{th} spreading code $\underline{c}_k \in C^{Q\times 1}$.

$$\boldsymbol{C}'_k = \underbrace{\begin{bmatrix} \underline{c}_k & \cdots & \underline{0} \\ \vdots & \ddots & \vdots \\ \underline{0} & \cdots & \underline{c}_k \end{bmatrix}}_{N \text{ Times}}$$

${}^{m}\boldsymbol{H}^{n} \in C^{(QN+W-1)\times QN}$ Channel matrix from Tx antenna n to Rx antenna m.

and $\underline{v}$ is a vector of iid complex Gaussian variables, $\boldsymbol{R}_v = \mathrm{E}\{\underline{v}\underline{v}^{\mathrm{H}}\} = N_0\boldsymbol{I}$. The 4-QAM modulation mapping is such that $x_k^n(t) = b_{k,0}^n(t) + jb_{k,1}^n(t)$ with $b_{k,i}^n(t) \in \{+1, -1\}$.

3. APP RECEIVER

As indicated in Figure (1), in this receiver the signal vector $\underline{r}$ is applied to an *a posteriori* probability (APP) detector following a process of channel equalization. The soft outputs from the APP detector are then applied to a turbo decoder which generates reliable estimates of the transmitted bits.

A full space-time APP detector implies joint detection of KN_T transmitted symbols per symbol epoch. For 4-QAM modulation, and for dispersive channels with ISI extending over L symbols, this requires a search over a trellis containing $2^{2(L+1)KN_T}$ states.

The computational complexity is clearly inhibitive for typical parameter values. Note that, in flat fading conditions (L=0) and for K orthogonal codes re-used over the transmit antennas, the number of trellis states reduces to a more realistic value of 2^{2N_T}. As a result, a strategy for dealing with dispersive channels is to apply APP detection after a process of space-time equalization which effectively eliminates dispersion. The equalization process inevitably causes noise colouring, which needs to be accounted for in APP detection.

Space-Time Equalization

The received signal over N symbol epochs may be written as

$$\underline{r} = \boldsymbol{H}\sum_{k=1}^{K}\boldsymbol{C}_k\,\underline{x}_k + \underline{v} = \boldsymbol{HC}\underline{x} + \underline{v} \tag{3}$$

where $\underline{s}=\boldsymbol{C}\underline{x}$ is the vector of spread symbols. A minimum mean-square error (MMSE) equalizer represents a space-time matrix $\boldsymbol{V}$ which minimizes the term $\mathrm{E}\{\|\underline{s}-\boldsymbol{V}\underline{r}\|^2\}$. It is easy to show [10], that the solution to this problem is given by

$$\boldsymbol{V} = \boldsymbol{R}_s\boldsymbol{H}^{\mathrm{H}}\left(\boldsymbol{H}\boldsymbol{R}_s\boldsymbol{H}^{\mathrm{H}} + \boldsymbol{R}_v\right)^{-1} \tag{4}$$

where $\boldsymbol{R}_s=\mathrm{E}\{\underline{s}\underline{s}^{\mathrm{H}}\}= 2\boldsymbol{C}\boldsymbol{C}^{\mathrm{H}}$ since $\mathrm{E}\{\underline{x}\underline{x}^{\mathrm{H}}\}=2\boldsymbol{I}$. The equalization process $\boldsymbol{V}\underline{r}$ essentially eliminates the effects of the channel matrix $\boldsymbol{H}$. As a result, assuming orthogonal spreading codes, the contribution of symbols transmitted using the k^{th} spreading code can be retrieved at the output of the equalizer via the de-spreading operation

$$\underline{z}_k = \boldsymbol{C}_k^{\mathrm{H}}\boldsymbol{V}\underline{r} = \boldsymbol{T}_k\underline{r} \approx \|\underline{c}_k\|^2\,\underline{x}_k + \boldsymbol{T}_k\underline{v} \quad \in \quad \mathrm{C}^{N_T N\times 1} \qquad \text{and} \qquad k=1\ldots K \tag{5}$$

where $\underline{v}_k=\boldsymbol{T}_k\underline{v}$ is coloured noise. Due to excessive computational complexity, space-time equalization is usually performed over a block of $N_E<N$ symbols and repeated N/N_E times to cover the overall transmission period.

APP Detection

Vector $\underline{z}_k$ consists of the equalized and de-spread contributions of $N_T N$ symbols transmitted via the k^{th} spreading code over a total of N symbol epochs. Considering only the N_T rows of Eq. (5) corresponding to the t^{th} symbol epoch, we have

$$\underline{z}_k(t) = \|\underline{c}_k\|^2\,\underline{x}_k(t) + \boldsymbol{T}_k(t)\underline{v} \in \mathrm{C}^{N_T\times 1} \qquad \text{and} \qquad k=1\ldots K \tag{6}$$

The APP algorithm can then be applied to derive log-likelihood ratios for each (equally probable) transmitted bit, $b_{k,i}^{n}(t)$ n=1…N_T $i\in\{0,1\}$, in the form of soft outputs

$$\begin{aligned} y\left(b_{k,i}^{n}(t)\right) \;=\; & \min_{\underline{x}_k(t)\,|\,b_{k,i}^{n}(t)=-1}\left\{\left\|\boldsymbol{R}_{u_k(t)}^{\frac{-1}{2}}\left(\underline{z}_k(t)-\|\underline{c}_k\|^2\,\underline{x}_k(t)\right)\right\|^2\right\} \\ & - \min_{\underline{x}_k(t)\,|\,b_{k,i}^{n}(t)=+1}\left\{\left\|\boldsymbol{R}_{u_k(t)}^{\frac{-1}{2}}\left(\underline{z}_k(t)-\|\underline{c}_k\|^2\,\underline{x}_k(t)\right)\right\|^2\right\} \end{aligned} \tag{7}$$

where $\underline{u}_k(t)=\boldsymbol{T}_k(t)\underline{v}$ is coloured noise and $\boldsymbol{R}_{u_k(t)}=\mathrm{E}\{\underline{u}_k(t)\underline{u}_k(t)^{\mathrm{H}}\}=N_o\,\boldsymbol{T}_k(t)\boldsymbol{T}_k(t)^{\mathrm{H}}$. The soft outputs are then applied to a Turbo decoder whose constituent decoders operate based on the max-log MAP algorithm.

4. ITERATIVE RECEIVER

At the receiver of Figure (2), the signal vector $\underline{r}$ is fed into a successive interference canceller incorporating matched filter detection (MF-SIC). The received signal of Eq. (1) observed over the t^{th} symbol epoch may be written as

$$\underline{r}(t) = \sum_{n=1}^{N_T} \sum_{k=1}^{K} \underline{a}_k^n(t) x_k^n(t) + \mathrm{ISI} + \underline{v}(t) \in C^{N_R(Q+W-1)\times 1} \tag{8}$$

where $x_k^n(t)$ is a transmitted symbol at the t^{th} symbol epoch and $\underline{a}_k^n(t)$ is its code-channel signature at the receiver. The output of the MF-SIC is then de-interleaved and applied to a convolutional decoder. This represents the first iteration of the receiver. Soft outputs from the decoder are then re-interleaved and applied to the MF-SIC for further iterations.

Iteration 1:

Here, the MF-SIC operates at a symbol level. The first step is to determine, at each symbol interval t, the most reliable symbol according to a reliability criterion. Ideally, the symbol with the lowest error probability is selected [6]. Lacking such information, the symbol $x_k^n(t)$ k=1...K n=1...N_T with the highest signature energy, $|\underline{a}_k^n(t)|^2$ (or least mean-square estimation error), is selected. The next step is to estimate the selected symbol (soft-output derived via matched filter detection), make a hard decision on the estimate, reconstruct and cancel its contribution from the received signal:

$$y_k^n(t) = \underline{a}_k^n(t)^{\mathrm{H}} \underline{r}(t) \tag{9}$$

$$\underline{r}(t) = \underline{r}(t) - \underline{a}_k^n(t)\left\{\mathrm{sgn}\left\{\mathrm{Re}[y_k^n(t)]\right\} + j\,\mathrm{sgn}\left\{\mathrm{Im}[y_k^n(t)]\right\}\right\} \tag{10}$$

The process is then repeated for the next most reliable symbol. If the decision on the selected symbol is correct, then its interference towards other symbols can be completely suppressed. However, a wrong decision doubles the level of interference caused by the erroneously detected symbol. Consequently, the reliability criterion used for the ordering of symbols is of critical importance in any form of successive cancellation.

After the MF-SIC detection of a complete code-block, the corresponding soft-outputs, $y_{k,0}^n(t)$=Re$[y_k^n(t)]$ and $y_{k,1}^n(t)$=Im$[y_k^n(t)]$, are multiplexed into a single stream for de-interleaving and convolutional decoding (max-log MAP algorithm). The decoder output is fed into the soft-output combiner and an interleaver prior to re-application to the MF-SIC for subsequent iterations.

Iteration 2 and beyond:

In the second iteration of the receiver, the MF-SIC has access to reliability information at a bit level, in the form of log-likelihood ratios, $\Lambda(b_{k,i}^n(t))$, generated by the soft-output decoder in the previous iteration. As a result, at each symbol interval t, ordering can be performed at a bit level (rather than symbol level) based on the log-likelihood ratios (LLRs). In other words, the bit $b_{k,i}^n(t)$ with the largest LLR value $|\Lambda(b_{k,i}^n(t))|$ (or minimum estimation error probability), can be selected as most reliable. Since bit estimates corresponding to a particular symbol can have different reliabilities, the use of LLR values represents an optimum ordering policy. The cancellation

process at the t^{th} symbol interval is based on the more reliable hard bit estimates derived from the LLR values:

$$y_{k,i}^{n}(t) = \frac{1}{2j^{i}}\left\{\underline{a}_{k}^{n}(t)^{H}\,\underline{r}(t) + (-1)^{i}\,\underline{r}(t)^{H}\,\underline{a}_{k}^{n}(t)\right\} \tag{11}$$

$$\underline{r}(t) = \underline{r}(t) - j^{i}\,\underline{a}_{k}^{n}(t)\,\mathrm{sgn}\left\{\Lambda\left(\hat{b}_{k,i}^{n}(t)\right)\right\} \tag{12}$$

where i=0 or 1 depending on whether the bit of interest forms the real or imaginary part of the 4-QAM symbol. The process is again repeated for the next most reliable bit. After the MF-SIC detection of a complete code-block, the soft-outputs y, are again multiplexed into a single stream for de-interleaving and decoding (max-log MAP algorithm). The performance of the MF-SIC (and hence the receiver) should improve at each iteration as the quality of the decoder output improves.

Soft-Output Combining

In the proposed iterative receiver, mutual information is exchanged between the MF-SIC detector and the convolutional decoder. Therefore, at each iteration, soft estimates (in the form of LLR values) at the output of the decoder are fed back to the detector for purposes of interference cancellation. Consequently, new and hopefully more reliable soft-output values are made available at the output of the decoder after each iteration. However, in some cases, the interference cancellation process can lead to poorer soft-outputs for certain bits. This can result in error propagation and therefore unstable bit-error rate performance in subsequent iterations.

Such instabilities can be avoided by combining the soft-output values computed in the current iteration with those computed in the previous iteration(s). The combining weight factors have a significant influence on the stability and the speed of convergence of the iterative receiver. Using this combining process, reliability information already gained for a certain transmitted bit is not lost in the next iteration.

While soft-output combining can be performed either at the output of the detector or that of the decoder, simulations indicate that a combination of both is most effective. If q indicates the iteration index, then soft-output combining may be described as

$$y_{k,i}^{n}(t)\,[q] = \alpha\; y_{k,i}^{n}(t)\,[q] \;+\; (1-\alpha)\; y_{k,i}^{n}(t)\,[q-1] \tag{13}$$

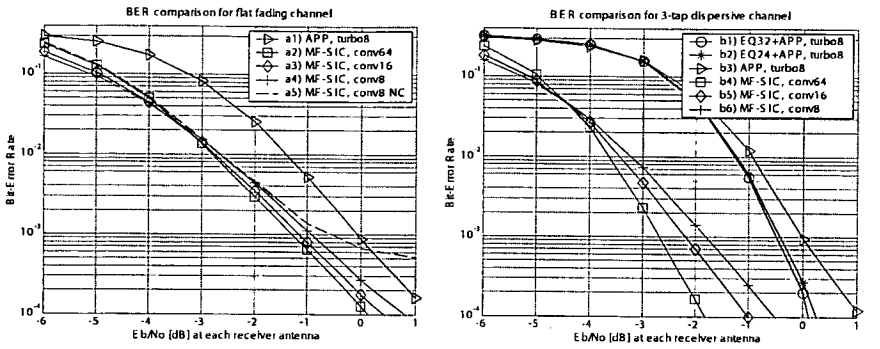
$$\lambda_{k,i}^{n}(t)\,[q] = \beta\; \Lambda_{k,i}^{n}(t)\,[q] \;+\; (1-\beta)\; \Lambda_{k,i}^{n}(t)\,[q-1] \tag{14}$$

Good performance results were found to be achieved via combining factors of α=0.9 and β =0.75. Soft-output combining is an essential element of the proposed iterative receiver.

5. PERFORMANCE AND COMPLEXITY COMPARISON

The performance of the APP-based receiver for a Turbo-encoded MIMO link [4] is considered as reference for comparison with that of the proposed MF-SIC based receiver for an equivalent convolutionally-encoded MIMO link.

A system with N_T=N_R=4, Q=16 and K=16, similar to the HSDPA specifications is considered. In addition to a flat Rayleigh fading channel, a dispersive channel with 3 equal-power, chip-spaced taps is also considered. The assumed mobile speed is 3



(a) (b)

Figure 3. Performance comparison

km/h and the receiver has perfect knowledge of the average channel conditions during each transmitted data block. For the turbo-encoded MIMO link, a 8-state rate 1/3 turbo encoder is used in accordance with the HSDPA specifications, resulting in a block size with up to 5114 information bits. A total of 6 iterations of the turbo decoder are performed in the receiver. For the convolutionally-encoded MIMO link, rate 1/3, 8- 16- and 64-state convolutional encoders are considered to allow a comprehensive comparison in terms of performance and complexity. A total of 4 iterations between the MF-SIC detector and the convolutional decoder are performed. Soft-output combiners with the coefficients α=0.9 after detection and β=0.75 after decoding were used.

The BER performance comparison is presented in Figure (3) for flat (a) and dispersive (b) channels, while Figure (4) illustrates the corresponding computational complexities in terms of the number of real multiplications.

For flat fading channels, the proposed iterative MF-SIC based receiver outperforms the APP-based receiver by approximately 1 dB, dependent on the memory size of the convolutional code, consistently at a lower total computational complexity (a2-4). Simulation result (a5) clearly demonstrates the degradation in performance when soft-output combining is not used.

For the dispersive channel, the MF-SIC detectors offer again significant improvements in BER. In fact, the performance improvement over the equalized APP reference is even higher than for flat fading (result b4-6). Even with a simple 8-state convolutional decoder (b6), the proposed receiver offers improved BER results at approximately only 20% of the APP-based receiver complexity. The small performance differences for the equalized APP detector, between using equalizer block sizes of 32 or 24 chips (b2, b1), shows that the edge effects are negligible for this channel.

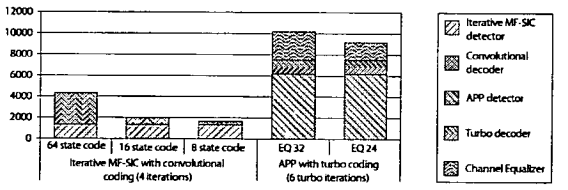


Figure 4. Complexity comparison: Multiplications per information bit

6. CONCLUSIONS

In this paper, a low complexity iterative receiver for a convolutionally-encoded MIMO system is proposed and compared with an APP-based receiver for an equivalent turbo-encoded MIMO system. The proposed iterative receiver utilizes a successive interference cancellation architecture based on simple matched filters (MF-SIC). Despite the use of a simplified detection scheme, a high level of receiver performance is achieved via iterations with the decoder. In fact, it is shown that, via a combination of bit-based ordering/detection/cancellation and the use of a novel soft-output combining technique at the decoder output, the proposed low-complexity iterative receiver outperforms the APP-based receiver. Depending on the complexity of the convolutional decoder, the BER performance can be improved by up to 2 dB, always at a significantly lower computational complexity. In contrast to the APP detector where the computational complexity grows exponentially with the number of antennas and the order of the modulation scheme, the complexity of the proposed iterative solution only grows linearly. This makes the proposed solution highly scalable and even more attractive for 16- and 64-QAM. Furthermore, unlike the APP-based receiver, the proposed solution does not require a matrix-channel equalizer to cope with dispersive propagation environments, making it attractive from an implementation standpoint. Finally, the iterative concept can be exploited to improve the channel estimation, which can further improve the performance in comparison to a non-iterative APP detector.

REFERENCES

[1] *3GPP TSG RAN WG1*, "PARC with APP Decoding for HSDPA", TSG-R1(02)0549, April 2002, Paris, France.

[2] *Benedetto, S.; Divsalar, D.; Montorsi, G.; Pollara, F.*, "A Soft Input Soft-Output APP Module for Iterative Decoding of Concatenated Codes", IEEE Communictions Letters, Volume1, Issue 1, pp. 22 -24, January 1997.

[3] *3GPP TSG RAN WG1*, "Further link level results for HSDPA using multiple antennas", TSG-R1#17(00)1386, November 2000, Stockholm, Sweden.

[4] *3GPP TSG RAN WG1*, "Link Level Results for HSDPA using Multiple Antennas in Correlated and Measured Channels", TSG-1#19(01)0302. February 2001, Las Vegas, U.S.A.

[5] *Verdu S.*, "Multiuser Detection", Cambridge University Press, 1998.

[6] *Claussen, H.; Mulgrew, B; Karimi, H.R.*, "Performance Optimization of Successive Cancellation Detectors", World Wireless Congress, pp.797-802, May 2002, San Francicso, U.S.A.

[7] *Guinand, P.S.; Kerr, R.W.; Moher, M.*, "Serial Interference Cancellation for Highly Correlated Users", IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, pp. 133-136, August 1999, Victoria, Canada.

[8] *Foschini G.J.*, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment when using Multielement Antennas", Bell Labs Technical Journal, pp. 41-59, Autumn 1996.

[9] *Li, X.; Huang, H.; Foschini, G.J.; Valenzuela, R.A.*, "Effects of Iterative Detection and Decoding on the Performance of BLAST", GLOBECOM '00. Volume 2, pp. 1061-1066, November 2000, San Francisco, U.S.A.

[10] *Klein, A.; Kaleh, G.K.; Baier, P.W.*, "Zero Forcing and Minimum Mean-Square-Error Equalization for Multiuser Detection in Code-Division Multiple-Access Channels", IEEE Transactions on Vehicular Technology, Volume 45, No. 2, May 1996.

Layered Encoding for 16- and 64-QAM Iterative MIMO Receivers

Holger Claussen[1], Hamid Reza Karimi[2], Bernard Mulgrew[1]

[1] Signals & Systems Group, University of Edinburgh, UK
[2] Bell Labs Research, Lucent Technologies, Swindon, UK

Abstract *- In a previous work, the authors demonstrated that, for MIMO links using 4-QAM, receivers which involve iterations between a successive interference cancellation (SIC) detector and a convolutional decoder are able to outperform receivers based on the a posteriori probability (APP) detector and a turbo decoder. This is achieved at significantly lower complexity and without the need for matrix channel equalization. However, for higher-order modulation schemes such as 16- and 64-QAM, the iterative receiver is unable to converge appropriately, while the APP-based receiver performs adequately but at the expense of prohibitive computational complexity. In this paper, a novel layered encoding scheme is proposed at the transmitter which allows the iterative receiver to operate with higher-order modulations. Here, iterations are performed between the detector and a turbo-decoder. In addition, soft output combining, inherent to the iterative receiver architecture, is optimized in this paper so as to maximize the mutual information transfer at each iteration. It is shown that, in conjunction with the layered encoding scheme, the iterative receiver again outperforms the APP-based receiver, at a lower computational complexity.*

I. Introduction

The reuse of spreading codes across transmitter antennas in a multiple-input multiple-output (MIMO) link has been proposed for the support of high-speed downlink packet access (HSDPA) in UMTS [1]. While code reuse can allow increased data rates, it also results in high levels of interference which can be particularly detrimental when using high-order modulation schemes.

In order to tackle such high interference levels in a turbo-encoded MIMO link, receivers based on an *a posteriori* probability (APP) detector followed by a turbo decoder [2][3] have been proposed. To cope with dispersive channels and in order to avoid sequence estimation, the APP detector is also preceded by a matrix channel equalizer. The APP detector, which is described in more detail in [4][5], operates by computing soft-outputs for estimates of transmitted bits which most closely match the received signal in an Euclidian sense. The computational complexity of the APP detector is an exponential function of the total number of bits transmitted during a symbol epoch, which is equal to the product of the number of transmit antennas and the number of bits per symbol. Consequently, the complexity of the APP detector can become prohibitive for increasing numbers of transmit antennas and (perhaps more importantly) modulation orders. This inflexibility of the APP detector has resulted in renewed interest in the use of sub-optimal but less complex MIMO detectors.

In a previous work, it was shown by the authors that receivers which involve iterations between a decoder and a successive interference cancellation (SIC) detector based on simple matched filters (MF) are able to provide impressive performance at low computational complexity for a convolutionally encoded MIMO link using 4-QAM [5]. The performance of the iterative receiver is in fact superior to that of the APP-based receiver for the equivalent turbo-encoded MIMO link, and is achieved at a lower computational complexity. For high-order modulation schemes such as 16- and 64-QAM, however, the high levels of interference imply that the originally proposed iterative receiver is unable to converge due to the poor initial estimate of the matched filter detector.

In this paper, the above problem is solved via a novel layered encoding scheme, whereby bits which map to different Euclidian distances on the modulation constellation are encoded separately. This essentially represents a form of multilevel coded modulation utilising set partitioning [6][7]. The original 4-QAM iterative receiver of [5] is modified to employ turbo-decoding and to take advantage of the layered encoding at the transmitter and as a result converges appropriately in the presence of 16- and 64-QAM, again providing superior performance compared to the APP-based receiver. Soft output combining, which is an essential element of the iterative receiver, has been optimized in this paper to maximize the mutual information transfer at each iteration. Without loss of generality, and in order to allow a direct comparison with the APP-based receiver, Gray mapping is always considered.

Section II describes the signal model for the MIMO link and the proposed layered encoding scheme at the transmitter. Section III concentrates on the receiver processing and the soft output combining. Finally, the simulation results and performance comparison with the APP-based receiver are presented in section IV.

II. Layered Encoding Scheme and Signal Model

The high levels of interference caused by code reuse across the transmitting antennas of a MIMO link can result in poor receiver performance when using high order constellations such as 16- and 64-QAM. This is particularly the case for receivers whose operation relies on iterations between a low-complexity MIMO detector and a decoder. The quality of the soft outputs at the detector

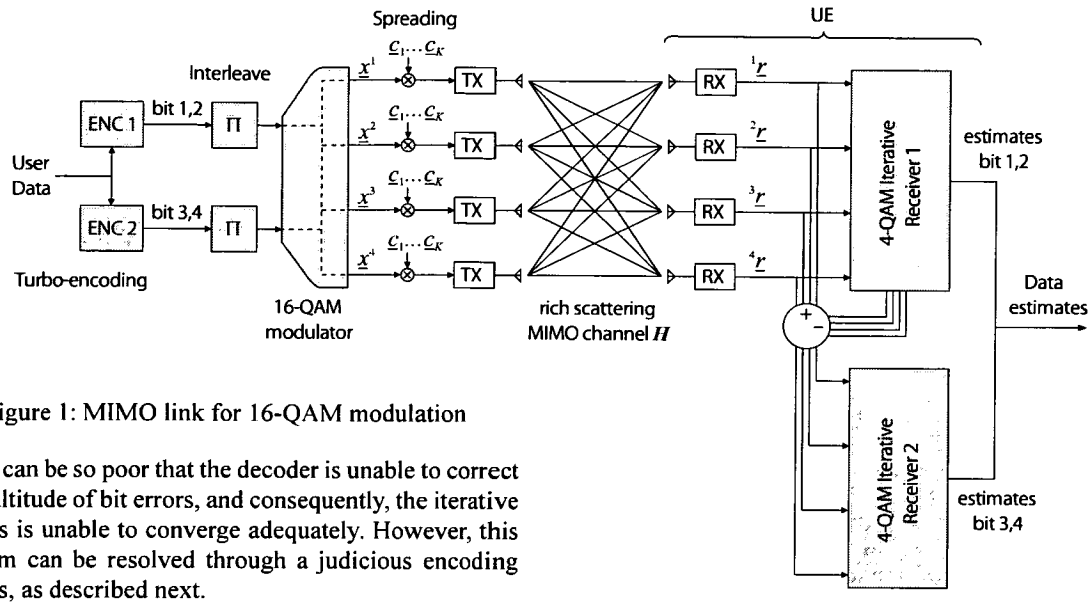


Figure 1: MIMO link for 16-QAM modulation

output can be so poor that the decoder is unable to correct the multitude of bit errors, and consequently, the iterative process is unable to converge adequately. However, this problem can be resolved through a judicious encoding process, as described next.

1. Signal Model

Figure (1) illustrates the system overview for the MIMO link under investigation for the case of 16-QAM modulation. At the transmitter, user data is encoded and interleaved. The coded data stream is de-multiplexed into N_T sub-streams, corresponding to the N_T transmit antennas. Each sub-stream is then modulated on to NK 16-QAM / 64-QAM symbols and subsequently spread by a factor Q via a set of K orthogonal spreading codes prior to transmission. Each transmitted spread stream then occupies N symbol intervals. Also note that the same set of K codes are re-used across all transmit antennas. Therefore, the MIMO propagation environment, which is assumed to exhibit significant multipath, plays a major role in achieving signal separation at the receiver. The transmitted signals are received by N_R receive antennas after propagation through dispersive radio channels with impulse response lengths of W chips. The received signal vector observed over the t^{th} symbol interval may then be written as

$$\underline{r}(t) = \sum_{n=1}^{N_T} \sum_{k=1}^{K} \underline{a}_k^n(t) x_k^n(t) + \text{ISI} + \underline{v}(t) \in \mathrm{C}^{N_R(Q+W-1)\times 1} \qquad (1)$$

where $x_k^n(t) \in \mathrm{C}$ is a symbol transmitted at the t^{th} symbol epoch, from antenna n, via code k and $\underline{a}_k^n(t)$ is its corresponding code-channel signature at the receiver. $\underline{v}(t)$ is a vector of iid complex Gaussian random variables representing thermal noise and inter-cell interference, with $\boldsymbol{R}_v = \mathrm{E}\{\underline{v}(t)\underline{v}^{\mathrm{H}}(t)\} = N_0 \boldsymbol{I}$.

2. Layered encoding

For a Gray-mapped 16-QAM constellation, each symbol $x_k^n(t)$ is given by

$$x_k^n(t) = 2\left\{-b_{k,0}^n(t) - jb_{k,1}^n(t)\right\} + \left\{-b_{k,0}^n(t)b_{k,2}^n(t) - jb_{k,1}^n(t)b_{k,3}^n(t)\right\} \qquad (2)$$

as a function of encoded bits $b_{k,0}^n$, $b_{k,1}^n$, $b_{k,2}^n$, $b_{k,3}^n \in \{-1, +1\}$. The corresponding constellation is illustrated in Figure (2a). As can be seen, for such high-order constellations, the Euclidean distance is not the same for all modulated bits. This implies that the modulation scheme affords different levels of protection to different bits. For the Gray mapped 16-QAM constellation of Figure (2), it is clear that $b_{k,0}^n$ and $b_{k,1}^n$ are equally better protected than $b_{k,2}^n$ and $b_{k,3}^n$.

The above feature may be exploited in the context of an iterative receiver, whereby the well-protected bits $b_{k,0}^n(t)$ and $b_{k,1}^n(t)$ are iteratively detected and decoded first. Due to the greater Euclidean distance associated with these bits, the iterative process is able converge reliably. The contribution of the estimated bits is subsequently cancelled from the received signal. This significantly reduces the interference for the remaining less-protected bits $b_{k,2}^n(t)$ and $b_{k,3}^n(t)$, which are only then iteratively detected and decoded.

In order for the well-protected and less-protected bits to be detected and decoded separately, it is required that they are also encoded separately at the transmitter. This is indicated in Figure (1), where the user data is split into two classes and encoded/interleaved independently. The encoded bits of class-1 correspond to $b_{k,0}^n(t)$ and $b_{k,1}^n(t)$, while the encoded bits of class-2 correspond to $b_{k,2}^n(t)$ and $b_{k,3}^n(t)$. The bits are then mapped on to 16-QAM symbols according to Equation (2). For 64-QAM, the procedure is identical, except that three classes are considered, according to the three levels of protection provided by the modulation scheme.

III. Layered Iterative Receiver

The iterative receiver of Figure (3) was originally proposed [5] in the context of 4-QAM. However, the layered encoding scheme in conjunction with the 16-QAM transmitter described in the previous section allows the receiver to treat the transmitted symbols as the aggregate of two inter-dependent 4-QAM constellations. Bits $b_{k,0}^n$ and $b_{k,1}^n$ contribute to the first 4-QAM constellation, while bits $b_{k,2}^n$ and $b_{k,3}^n$ contribute to the second

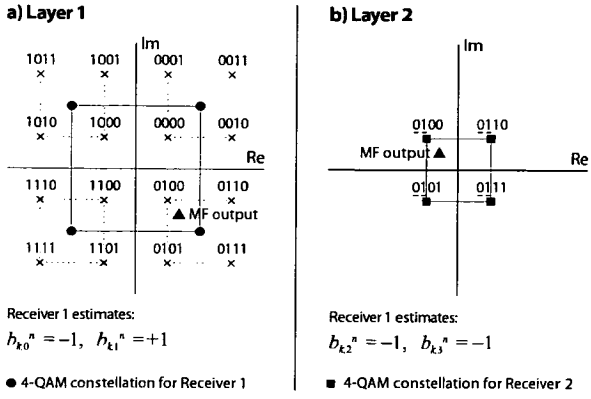


Figure 2: 16-QAM modulation as an aggregate of 2 interdependent 4-QAM modulations

constellation (with the latter mapping depending on the values of $\{b_{k,0}^n, b_{k,1}^n\}$ for an overall Gray mapping). The 4-QAM receiver first derives estimates of $\{b_{k,0}^n, b_{k,1}^n\}$ via a certain number of iterations, cancels their contribution from the received signal, and then derives estimates of $\{b_{k,2}^n, b_{k,3}^n\}$ via a second set of iterations. Figure (2) illustrate the effect of the cancellation process. It is clearly seen that once the contributions of $b_{k,0}^n$ and $b_{k,1}^n$ are subtracted from the 16-QAM constellation, the modulation is reduced to 4-QAM.

While the layered receiver process has been described for 16-QAM, it can be readily extended to 64-QAM, whereby the receiver treats the transmitted symbols as the aggregate of three inter-dependent 4-QAM constellations corresponding to three classes of reliability.

The operation of the iterative 4-QAM receiver is described next.

1. 4-QAM Iterative MF-SIC Receiver

As shown in Figure (3), the signal vector $\underline{r}$ is fed into a successive interference canceller incorporating matched filter detection (MF-SIC). The output of the MF-SIC is then de-interleaved and applied to a turbo-decoder. This represents the first iteration of the receiver. Soft outputs from the decoder are then re-interleaved and applied to the MF-SIC for further iterations.

Iteration 1:

Here, the MF-SIC operates at a 4-QAM symbol level, where a symbol corresponds either to class-1 bits $\{b_{k,0}^n(t), b_{k,1}^n(t)\}$ or class-2 bits $\{b_{k,2}^n(t), b_{k,3}^n(t)\}$. The first step is to determine, at each symbol interval t, the most reliable symbol according to some reliability criterion. Ideally, the symbol with the lowest error probability is selected. Lacking such information, the symbol $x_k^n(t)$ $k=1...K$ $n=1...N_T$ with the highest signature energy, $|\underline{a}_k^n(t)|^2$ (or least mean-square estimation error), is selected. The next step is to estimate the selected symbol (soft-output derived via matched filter detection), make a hard decision on the estimate, reconstruct and cancel its contribution from the received signal, as described next.

$$y_k^n(t) = \underline{a}_k^n(t)^H \underline{r}(t) \tag{3}$$

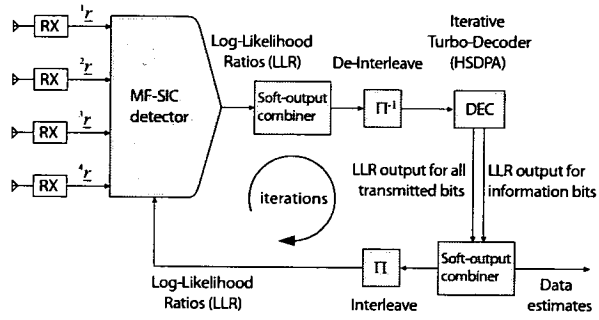


Figure 3: Iterative receiver based on Successive Interference Cancellation

For cancelling contributions of $b_{k,0}^n(t)$ or $b_{k,1}^n(t)$ (4)

$$\underline{r}(t) := \underline{r}(t) - 2\underline{a}_k^n(t)\left\{-\mathrm{sgn}\left\{\mathrm{Re}[y_k^n(t)]\right\} - j\,\mathrm{sgn}\left\{\mathrm{Im}[y_k^n(t)]\right\}\right\}$$

For cancelling contributions of $b_{k,2}^n(t)$ or $b_{k,3}^n(t)$

$$\underline{r}(t) := \underline{r}(t) - \underline{a}_k^n(t)\left\{-\hat{b}_{k,0}^n(t)\,\mathrm{sgn}\left\{\mathrm{Re}[y_k^n(t)]\right\} - j\hat{b}_{k,1}^n(t)\,\mathrm{sgn}\left\{\mathrm{Im}[y_k^n(t)]\right\}\right\} \tag{5}$$

where $\hat{b}$ is an estimate of b. The process is then repeated for the next most reliable symbol. If the decision on the selected symbol is correct, then its interference towards other symbols can be completely suppressed. However, a wrong decision doubles the level of interference caused by the erroneously detected symbol. Consequently, the reliability criterion used for the ordering of symbols is of critical importance in any form of successive cancellation. After the MF-SIC detection of a complete code-block, the corresponding soft-outputs,

$$\left\{y_{k,0}^n(t) = -\mathrm{Re}[y_k^n(t)],\ y_{k,1}^n(t) = -\mathrm{Im}[y_k^n(t)]\right\} \tag{6}$$

or (7)

$$\left\{y_{k,2}^n(t) = -\mathrm{Re}[y_k^n(t)]/\hat{b}_{k,0}^n(t),\ y_{k,1}^n(t) = -\mathrm{Im}[y_k^n(t)]/\hat{b}_{k,1}^n(t)\right\}$$

are multiplexed into a single stream for de-interleaving and turbo-decoding. The turbo decoding is performed for only 2 iterations (Max-Log MAP algorithm), to keep the computational complexity low. The decoder output is fed into the soft-output combiner and an interleaver prior to re-application to the MF-SIC for subsequent iterations.

Iteration 2 and beyond:

In the second iteration of the receiver, the MF-SIC has access to reliability information at a bit level, in the form of log-likelihood ratios, $\Lambda(\hat{b}_{k,i}^n(t))$, generated by the soft-output decoder in the previous iteration. As a result, at each symbol interval t, ordering can be performed at a bit level (rather than symbol level) based on the log-likelihood ratios (LLRs). In other words, the bit $\hat{b}_{k,i}^n(t)$ with the largest LLR value $|\Lambda(\hat{b}_{k,i}^n(t))|$ (or minimum estimation error probability), can be selected as most reliable. Since bit estimates corresponding to a particular symbol can have different reliabilities, the use of LLR values represents an optimum ordering policy. The cancellation process at the i^{th} symbol interval is based on the more reliable hard bit estimates derived from the LLR values:

$$y_{k,i}^n(t) = \frac{1}{2j^i}\left\{\underline{a}_k^n(t)^H \underline{r}(t) + (-1)^i \underline{r}(t)^H \underline{a}_k^n(t)\right\} \tag{8}$$

For cancelling contributions of $b_{k,0}^n(t)$ and $b_{k,1}^n(t)$

$$\underline{r}(t) := \underline{r}(t) - 2j^i \underline{a}_k^n(t)\left[-\mathrm{sgn}\left\{\Lambda\left(\hat{b}_{k,i}^n(t)\right)\right\}\right] \quad (9)$$

For cancelling contributions of $b_{k,2}^n(t)$ and $b_{k,3}^n(t)$

$$\underline{r}(t) := \underline{r}(t) - j^{(i-2)} \underline{a}_k^n(t)\left[-\hat{b}_{k,(i-2)}^n(t)\mathrm{sgn}\left\{\Lambda\left(\hat{b}_{k,i}^n(t)\right)\right\}\right] \quad (10)$$

The process is again repeated for the next most reliable bit. After the MF-SIC detection of a complete code-block, the soft-outputs y are again multiplexed into a single stream for de-interleaving and only 2 iterations of the turbo-decoder (Max-Log MAP algorithm). The performance of the MF-SIC (and hence the receiver) should improve at each iteration as the quality of the decoder output improves. Only in the final iteration, six iterations of the turbo-decoder are performed to achieve the best frame-error rate (FER) results.

2. Optimized Soft-Output Combining

As shown in Figure (3), the detector and decoder soft-outputs are combined with the corresponding values from the previous iteration [5]. This avoids instabilities due to error propagation caused by incorrectly detected and cancelled bits. Additional details are provided in [9] in the context of turbo decoding. The combining operation may be written as

$$\Lambda^q(\hat{b}_{k,i}^n(t)) := w_1^q \cdot \Lambda^q(\hat{b}_{k,i}^n(t)) + w_2^q \cdot \Lambda^{q-1}(\hat{b}_{k,i}^n(t)) \quad (11)$$

where $\Lambda^q(\hat{b}_{k,i}^n(t))$ represents soft values at the output of the detector or decoder at the q^{th} iteration. The proportion of the weight factors w_1^q and w_2^q have significant influence on the performance and speed of convergence of the iterative receiver. The weight factors w_1^q and w_2^q for each iteration q can be optimized by off-line maximiziation of the mutual information of the combined term $\Lambda^q(\hat{b}_{k,i}^n(t))$ in each combining step. Using vector notation

$$\begin{aligned}\Lambda^q(\hat{b}_{k,i}^n(t)) &:= \begin{bmatrix} w_1^q & w_2^q \end{bmatrix}\begin{bmatrix} \Lambda^q(\hat{b}_{k,i}^n(t)) \\ \Lambda^{q-1}(\hat{b}_{k,i}^n(t)) \end{bmatrix} \\ &= \underline{w}^{q\,\mathrm{T}}\left(\underline{\lambda}^q(t) + \underline{\varepsilon}^q(t)\right) = \underline{w}^{q\,\mathrm{T}} \cdot \underline{\lambda}^q(t) + v^q(t)\end{aligned} \quad (12)$$

where $\underline{\varepsilon}^q(t)$ represents the contributions of channel noise and the detection errors due to interference, and $\underline{\lambda}^q(t)$ is the vector of uncorrupted LLRs.

$$\begin{aligned}[\sigma_\Lambda^q]^2 &= \mathrm{E}\left\{\underline{w}^{q\,\mathrm{T}}(\underline{\lambda}^q(t) + \underline{\varepsilon}^q(t))(\underline{\lambda}^q(t) + \underline{\varepsilon}^q(t))^{\mathrm{T}}\underline{w}^q\right\} \\ &= \underline{w}^{q\,\mathrm{T}} \boldsymbol{R}_{\lambda+\varepsilon}^q \underline{w}^q \end{aligned} \quad (13)$$

$$[\sigma_v^q]^2 = \mathrm{E}\left\{\underline{w}^{q\,\mathrm{T}}\underline{\varepsilon}^q(t)\underline{\varepsilon}^q(t)^{\mathrm{T}}\underline{w}^q\right\} = \underline{w}^{q\,\mathrm{T}}\boldsymbol{R}_\varepsilon^q\underline{w}^q \quad (14)$$

Modelling $v^q(t)$ as a Gaussian random variable, the mutual information $I(\Lambda^q(b_{k,i}^n(t)); \underline{\lambda}^q(t))$ can be written as [8]

$$I\left(\Lambda^q(\hat{b}_{k,i}^n(t)); \underline{\lambda}^q(t)\right) = \frac{1}{2}\log\left(\frac{[\sigma_\Lambda^q]^2}{[\sigma_v^q]^2}\right) \quad (15)$$

The optimum combining weight factors must maximize the mutual information and can be derived as

$$\underline{w}_{opt}^q = \arg\max_{\underline{w}^q}\left(\frac{[\sigma_\Lambda^q]^2}{[\sigma_v^q]^2}\right) = \arg\max_{\underline{w}^q}\left(\frac{\underline{w}^{q\,\mathrm{T}}\boldsymbol{R}_{(\lambda+\varepsilon)}^q\underline{w}^q}{\underline{w}^{q\,\mathrm{T}}\boldsymbol{R}_\varepsilon^q\underline{w}^q}\right) \quad (16)$$

As shown in previous work [9], the solution to the weight optimization problem is given by

$$\underline{w}_{opt}^q = \left[\boldsymbol{R}_\varepsilon^q\right]^{-\mathrm{T}/2}\mathrm{eig}_{\max}\left(\left[\boldsymbol{R}_\varepsilon^q\right]^{-\mathrm{I}/2}\boldsymbol{R}_{\lambda+\varepsilon}^q\left[\boldsymbol{R}_\varepsilon^q\right]^{-\mathrm{T}/2}\right) \quad (17)$$

$$\text{with } \boldsymbol{R}_\varepsilon^q = \boldsymbol{R}_{\lambda+\varepsilon}^q - \boldsymbol{R}_\lambda^q \text{ and } \boldsymbol{R}_\lambda^q = \mathrm{E}\left\{\underline{\lambda}^q(t)\underline{\lambda}^q(t)^{\mathrm{T}}\right\} \quad (18)$$

$$\underline{\lambda}^q(t) = \begin{bmatrix} E\left\{\Lambda^q(\hat{b}_{k,i}^n(t))\cdot b_{k,i}^n(t)\right\}\cdot b_{k,i}^n(t) \\ E\left\{\Lambda^{q-1}(\hat{b}_{k,i}^n(t))\cdot b_{k,i}^n(t)\right\}\cdot b_{k,i}^n(t) \end{bmatrix} \quad (19)$$

The maximum eigenvector specifies the solution to within a scaling factor. However, since the Max-Log-MAP algorithm used for decoding is insensitive to input scaling, the scaling factor does not influence the receiver performance.

IV. Simulation Results

The performance of the APP-based receiver is considered as reference for comparison with that of the proposed iterative receiver for equivalent turbo-encoded MIMO links.

A system with N_T=N_R=4, Q=16 and K=16 is considered. In addition to a flat Rayleigh fading channel, a dispersive channel with 3 equal-power, chip-spaced taps is examined. The assumed mobile speed is 3 km/h and the receiver has perfect knowledge of the average channel conditions during each transmitted data block. A 8-state rate 1/3 turbo encoder is used in accordance with the HSDPA specifications, resulting in a block size of up to 5114 information bits. A total of 6 turbo-decoder iterations are performed in the receiver. For the proposed iterative receiver, a total of 6 iterations between the MF-SIC detector and the turbo-decoder are performed. The optimized weight factors w_1^q for the two soft-output combiners are shown in Table (1), where w_2^q=1−w_1^q.

Figure (4) shows the BER and FER performance comparison for 16- and 64-QAM for flat fading. The solid lines represent the bit-error rates and the dashed lines the corresponding frame-error rates. Here, the proposed iterative receiver outperforms the APP-based reference by approximately 1 dB in the FER range of interest at around 10^{-1}. The iterative MF-SIC receiver exhibits an error floor, both for 16- and 64-QAM. However, the FER results show that the error floor has no significant effect on the FER rates of interest. For 64-QAM modulation, the performance of the APP receiver is not shown since its computational complexity becomes clearly prohibitive in this scenario.

For dispersive channels, the iterative receiver outperforms the equalized APP receiver by over 5 dB, at a significantly lower computational complexity as shown in Figure (5). Again, the APP based receiver for 64-QAM is too complex to simulate. In addition, the error floor is much lower than for flat fading.

Figure (6) shows the throughput provided by the iterative receiver for a HSDPA scenario and flat fading. This demonstrates that the effect of the error floor on the system throughput, experienced by the proposed receiver is negligible.

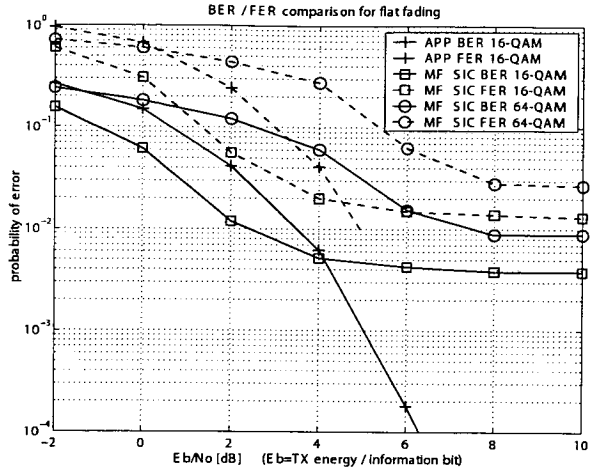


Figure 4: BER / FER comparison for 16- and 64-QAM modulation and flat fading

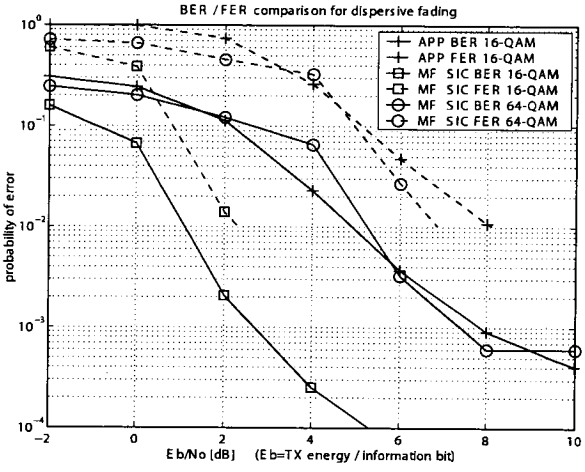


Figure 5: BER / FER comparison for 16- and 64-QAM modulation and dispersive fading

V. Conclusions

In this paper, a novel layered encoding scheme was proposed which enables the use of low-complexity iterative receivers (based on matched filters and successive interference cancellation) in MIMO links with high-order modulation schemes such as 16-QAM and 64-QAM. By taking advantage of the layered encoding scheme, the iterative receiver which was originally proposed by the authors in the context of 4-QAM, was extended here for higher-order constellations. The receiver also relies on an optimized soft-output combining scheme which maximizes the mutual information transfer in each iteration. Based on FER results for 16-QAM in dispersive channels, it was shown that proposed layered iterative receiver outperforms the reference APP-based receiver by over 5 dB, at significantly lower computational complexity. Also for 64-QAM, where the complexity of the APP-based receiver becomes clearly prohibitive, the proposed layered iterative receiver still exhibits impressive BER and FER performance.

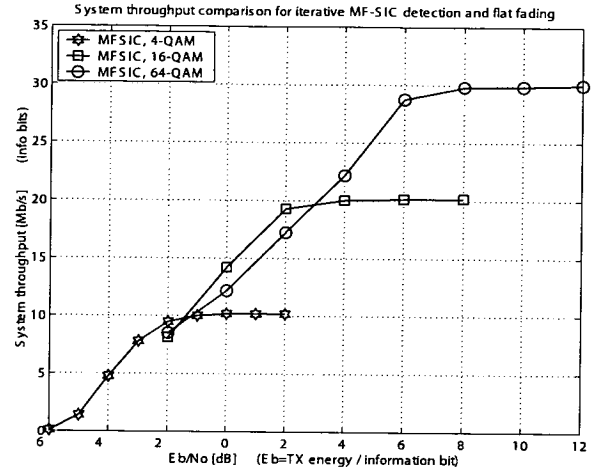


Figure 6: System troughput comparison

Table 1: Optimized soft-output combining factors

iteration q	1	2	3	4	5	6
w_1^q (combiner 1)	1.00	0.93	0.63	0.54	0.46	0.48
w_1^q (combiner 2)	1.00	0.82	0.59	0.45	0.34	0.72

References

[1] 3GPP TSG RAN WG1, "PARC with APP Decoding for HSDPA", TSG-R1(02)0549, April 2002, Paris, France.

[2] 3GPP TSG RAN WG1, "Further link level results for HSDPA using multiple antennas", TSG-R1#17(00)1386, November 2000, Stockholm, Sweden.

[3] 3GPP TSG RAN WG1, "Link Level Results for HSDPA using Multiple Antennas in Correlated and Measured Channels", TSG-1#19(01)0302. February 2001, Las Vegas, U.S.A.

[4] S.Benedetto, D. Divsalar, G. Montorsi, F. Pollara, "A Soft Input Soft-Output APP Module for Iterative Decoding of Concentrated Codes", IEEE Communications Letters, Volume 1, Issue 1, pp. 22 -24, January 1997.

[5] H. Claussen, H. R. Karimi, B. Mulgrew, "A Low Complexity Iterative Receiver based on Successive Cancellation for MIMO", Conference on Personal Wireless Communications PWC'2002, pp. 105-112, October 2002, Singapore.

[6] U. Wachsmann, R. F. H. Fischer, J. B. Huber, "Multilevel Codes: Theoretical Concepts and Practical Design Rules", IEEE Transactions on Information Theory, Volume 45, No. 5, pp. 1361-1391, July 1999.

[7] R. H. Morelos-Zaragoza, M. P. C. Fossorier, S. Lin, H. Imai, "Multilevel Coded Modulation for Unequal Error Protection and Multistage Decoding - Part I: Symmetric Constellations", IEEE Transactions on Communications, Volume 48, No. 2, pp. 204-213, February 2000.

[8] S. Haykin, "Neural Networks", Prentice Hall, 2nd edition, 1998.

[9] H. Claussen, H. R. Karimi, B. Mulgrew, "Improved Max-Log-MAP Turbo-decoding using Maximum Mutual Information Combining", submitted to IEEE International Conference on Communications, May 2003, Anchorage, U.S.A.

Improved Max-Log MAP Turbo Decoding using Maximum Mutual Information Combining

Holger Claussen[1], Hamid Reza Karimi[2], Bernard Mulgrew[1]

[1] Signals & Systems Group, University of Edinburgh, UK
[2] Bell Labs Research, Lucent Technologies, Swindon, UK

Abstract—**The demand for low-cost and low-power decoder chips has resulted in renewed interest in low-complexity decoding algorithms. In this paper a novel modification of the Max-Log-MAP algorithm is proposed for use in a turbo decoding process. This is achieved by scaling the *a priori* information by correction weights at each iteration, in order to maximize the exchange of mutual information between the component decoders. It is shown that the proposed technique results in a performance which approaches that of a turbo decoder using the optimum MAP algorithm, while maintaining the advantages of low complexity and insensitivity to input scaling inherent in the Max-Log-MAP algorithm. A second contribution of this paper is a method for off-line computation of the optimum weight values. The convergence behaviour of the proposed decoder is analysed via extrinsic information transfer (EXIT) charts.**

1. INTRODUCTION

Since the discovery of turbo codes in 1993 [1] there has been renewed interest in the field of coding theory, with the aim of approaching the Shannon limit. Furthermore, with the proliferation of wireless mobile devices in recent years, the availability of low-cost and low-power decoder chips is of paramount importance. To this end, several techniques for reducing the complexity of the optimum MAP decoding algorithm [2] have been proposed. Examples include the Log-MAP and Max-Log-MAP algorithms [3][4].

In this paper a novel modification of the Max-Log-MAP algorithm is proposed for use in a turbo decoding process. The approach aims to maximize the mutual information at the input of each component decoder by correcting the bias in the *a priori* information caused by the Max-Log approximation in the previous component decoder. This is performed by scaling the *a priori* information by optimised weight factors at each turbo iteration. A second contribution of this paper is a method for the off-line computation of the optimal weights according to the maximum mutual information criterion. These are developed in Section 4. Sections 2 and 3 provide the necessary background, and Section 5 presents a performance comparison. It is shown that the performance of a turbo decoder using the modified Max-Log-MAP algorithm can be improved to approach that of a turbo decoder using the optimum Log-MAP or MAP algorithms. This is achieved at the expense of only two additional multiplications per systematic bit per turbo iteration. Furthermore, the insensitivity of the Max-Log-MAP algorithm to an arbitrary scaling of its input log-likelihood ratios is maintained. Extrinsic information transfer (EXIT) charts [5] are used to analyse the convergence behaviour of the turbo decoder.

2. TURBO DECODING

Consider the received signal, $r_t = x_t + n_t$, at the output of an AWGN channel at time instant t, where $x_t \in \{+1,-1\}$ is the transmitted binary symbol (corresponding to the encoded bit $b_t \in \{1,0\}$) and n_t is zero-mean Gaussian noise of variance $\mathrm{E}\{n_t^2\} = N_o$. Then the log-likelihood ratio (LLR) of the transmitted symbol is defined as

$$L(x_t) = \log \frac{\mathrm{P}\{x_t = +1\}}{\mathrm{P}\{x_t = -1\}} = \frac{2}{N_o} r_t \qquad (1)$$

where $\mathrm{P}\{A\}$ represents the probability of event A. Let us also consider, without loss of generality, a parallel concatenated turbo encoding process of rate 1/3 at the transmitter. This consists of two 1/2 rate recursive systematic convolutional (RSC) encoders separated by an interleaving process, resulting in transmitted systematic symbol $x_{t,0}$ and parity symbols $x_{t,1}$ and $x_{t,2}$. The corresponding signals at the output of the channel (input of the decoder) may then be expressed as $L_c(x_{t,0})$, $L_c(x_{t,1})$ and $L_c(x_{t,2})$.

Figure 1 depicts the turbo decoding procedure whereby decoding is performed in an iterative manner via two soft-output component decoders, separated by an interleaver, with the objective of improving the estimates of $x_{t,0}$ from iteration i to iteration $i+1$. The first decoder generates extrinsic information $L_e^{(i)}(x_{t,0})$ on the systematic bits, which then serves as *a priori* information $\tilde{L}_a^{(i)}(x_{t,0})$ for the second decoding process. The symbol '~' denotes interleaved quantities.

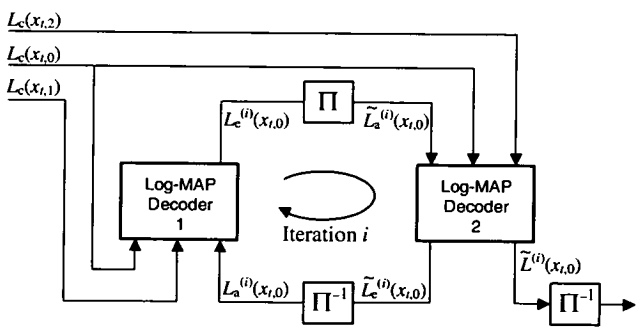


Figure 1 – Turbo decoding for parallel concatenated codes.

The maximum *a posteriori* probability (MAP) algorithm is the optimum strategy for the decoding of RSC codes, as it results in a minimum probability of bit error. However, due to its high computational complexity, the MAP algorithm is usually implemented in the logarithmic domain in the form of the Log-MAP or Max-Log-MAP algorithms. While the former is mathematically equivalent to the MAP algorithm, the latter involves an approximation which results in even lower complexity, albeit at the expense of some degradation in performance [3][4]. For purposes of brevity, the expressions presented in this paper are written for the first component decoder, with obvious extensions to the second decoder.

2.1. Log-MAP Algorithm

The Log-MAP algorithm is the log-domain implementation of the MAP algorithm and operates directly on LLRs. Given the LLRs for the systematic and parity bits as well as *a priori* LLRs for the systematic bits, the Log-MAP algorithm computes new LLRs for the systematic bits as described below

$$L(x_{t,0}) = \log \frac{\sum_{l=0}^{M-1} \exp\{\bar{\alpha}_{t-1}(l') + \bar{\gamma}_t^{[1]}(l',l) + \bar{\beta}_t(l)\}}{\sum_{l=0}^{M-1} \exp\{\bar{\alpha}_{t-1}(l') + \bar{\gamma}_t^{[0]}(l',l) + \bar{\beta}_t(l)\}} \quad (2)$$

$$= L_a(x_{t,0}) + L_c(x_{t,0}) + L_e(x_{t,0}) \quad (3)$$

where $\bar{\gamma}_t^{[q]}(l',l)$ is the logarithm of the probability of a transition from state l' to state l of the encoder trellis at time instant t, given that the systematic bit takes on value $q \in \{1,0\}$ and M is the total number of states in the trellis. Note that the new information at the decoder output regarding the systematic bits is encapsulated in the extrinsic information term $L_e(x_{t,0})$. Coefficients $\bar{\alpha}_t(l')$ and $\bar{\beta}_t(l)$ are forward- and backward-accumulated metrics at time t. For a data block of τ systematic bits $(x_{1,0} \cdots x_{\tau,0})$ and the corresponding parity bits $(x_{1,1} \cdots x_{\tau,1})$, these coefficients are calculated as follows:

Forward Recursion– Initialise $\bar{\alpha}_0(l)$, $l = 0,1,\cdots,M-1$ such that $\bar{\alpha}_0(0) = 0$ and $\bar{\alpha}_0(l) = -\infty$ for $l \neq 0$. Then

$$\bar{\gamma}_t^{[q]}(l',l) = \frac{1}{2}\left\{\left\{L_a(x_{t,0}) + L_c(x_{t,0})\right\}x_{t,0}^{[q]} + L_c(x_{t,1})x_{t,1}^{[q]}\right\} \quad (4)$$

and $$\bar{\alpha}_t(l) = \log \sum_{l'=0}^{M-1} \sum_{q=0,1} \exp\{\bar{\alpha}_{t-1}(l') + \bar{\gamma}_t^{[q]}(l',l)\} \quad (5)$$

Backward Recursion – Initialise $\bar{\beta}_\tau(l)$, $l = 0,1,\cdots,M-1$ such that $\bar{\beta}_\tau(0) = 0$ and $\bar{\beta}_\tau(l) = -\infty$ for $l \neq 0$. Then

$$\bar{\beta}_t(l) = \log \sum_{l'=0}^{M-1} \sum_{q=0,1} \exp\{\bar{\beta}_{t+1}(l') + \bar{\gamma}_{t+1}^{[q]}(l,l')\} \quad (6)$$

where $x_{t,n}^{[0]} = -1$ and $x_{t,n}^{[1]} = +1$.

Equation (2) can be readily implemented via the Jacobian equality $\log(e^{\delta_1} + e^{\delta_2}) = \max(\delta_1,\delta_2) + \log(1 + e^{-|\delta_2 - \delta_1|})$ and using a look up table to evaluate the correction function $\log(1 + e^{-|\delta_2 - \delta_1|})$.

2.2. Max-Log MAP Algorithm

The complexity of the Log-MAP algorithm can be further reduced by using the Max-Log approximation $\log(e^{\delta_1} + e^{\delta_2}) \approx \max(\delta_1,\delta_2)$ for evaluating (2). Clearly, this results in biased soft outputs and degrades the performance of the decoder. Nevertheless, the Max-Log-MAP algorithm is often the preferred choice for implementing a MAP decoder since it has the added advantage that its operation is insensitive to a scaling of the input LLRs. This implies that knowledge or estimation of the channel noise variance N_o is not required.

3. EXIT CHARTS

The performance and convergence behaviour of turbo codes can be analysed using extrinsic information transfer (EXIT) charts, as proposed in [5]. The idea is to visualize the evolution of the mutual information exchanged between the component decoders from iteration to iteration. EXIT charts operate under the following assumptions: a) The *a priori* information is fairly uncorrelated from channel observations. This is valid for large interleaver sizes. b) The extrinsic information $L_e(x_{t,0})$ has a Gaussian-like distribution, as shown in [6] for the MAP decoder.

An EXIT chart consists of a pair of curves which represent the mutual information transfer functions of the component decoders in the turbo process. Each curve is essentially a plot of *a priori* mutual information I_a against extrinsic mutual information I_e for the component decoder of interest. The terms I_a and I_e are related to the probability density functions (pdfs) of $L_a(x_{t,0})$ and $L_e(x_{t,0})$, the signal-to-noise ratio E_b/N_o and the RSC encoder polynomials. If the component decoders are identical, the two curves are naturally mirror images. The required pdfs can be estimated by generating histograms $p(L_a)$ and $p(L_e)$ of $L_a(x_{t,0})$ and $L_e(x_{t,0})$ respectively for a particular value of E_b/N_o where E_b denotes the energy per information bit. This can be achieved by applying *a priori* information modelled as $L_a(x_{t,0}) = \mu_a x_{t,0} + n_{a,t}$ $t = 1 \cdots \tau$ to the input of a component decoder and observing the output $L_e(x_{t,0})$ for a coded data block corresponding to τ information bits. The random variable $n_{a,t}$ is zero-mean Gaussian with variance $E\{n_{a,t}^2\} = \sigma_a^2$ such that $\sigma_a^2 = 2\mu_a$. The latter is a requirement for $L_a(x_{t,0})$ to be a LLR. The mutual information I_a may then be computed as

$$I_a = \sum_{q=-1,1} \frac{1}{2} \int_{-\infty}^{+\infty} p(L_a | x_{t,0} = q) \log_2 \frac{2p(L_a | x_{t,0} = q)}{p_a} dL_a \quad (7)$$

where $p_a = p(L_a \mid x_{t,0} = -1) + p(L_a \mid x_{t,0} = +1)$. Similarly, I_e can be computed as

$$I_e = \sum_{q=-1,1} \frac{1}{2} \int_{-\infty}^{+\infty} p(L_e | x_{t,0} = q) \log_2 \frac{2p(L_e | x_{t,0} = q)}{p_e} dL_e \quad (8)$$

where $p_e = p(L_e \mid x_{t,0} = -1) + p(L_e \mid x_{t,0} = +1)$. The resulting pair (I_a, I_e) defines one point on the transfer function curve.

Different points (for the same E_b/N_o) can be obtained by varying the value of σ_a^2.

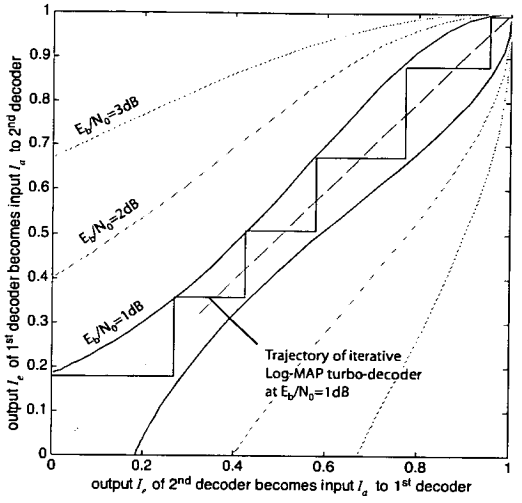


Figure 2 – EXIT chart for Log-MAP turbo decoder.

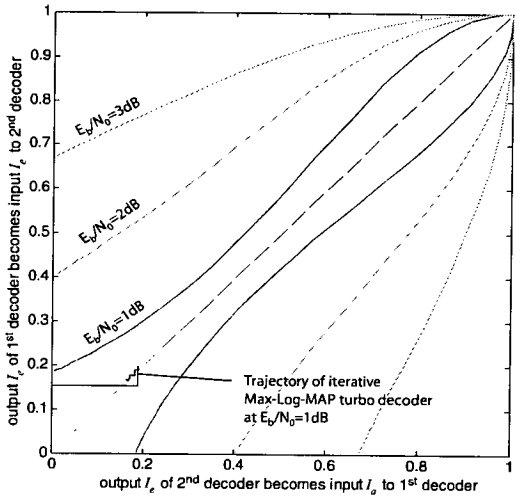


Figure 3 – EXIT chart for Max-Log MAP turbo decoder.

Having derived the transfer functions, we may now observe the trajectory of mutual information at various iterations of an actual turbo decoding process. At each iteration, mutual information is again computed as in (7) and (8), however the *a priori* LLR, $L_a(x_{t,0})$, at the input of the component decoder is no longer a modelled random variable but corresponds to the actual extrinsic LLR generated by the previous component decoding operation.

Figures 2 and 3 illustrate EXIT charts with trajectories of mutual information for the Log-MAP and Max-Log-MAP algorithms respectively. The "snapshot" trajectories correspond to turbo decoding iterations for a specific coded data block. The 1/2 rate (punctured) turbo encoder consists of two component RSC encoders, each operating at 1/2 rate with a memory of 4 and octal generator polynomials (G_r, G) = (23,37), where G_r denotes the recursive feedback polynomial. Note that while the mutual information trajectory for the Log-MAP algorithm in Figure 2 fits the predicted transfer function, the trajectory in Figure 3 clearly indicates the impact of numerical errors resulting from the Max-Log approximation: the trajectory stalls after only the first iteration and the turbo decoder is unable to converge at the simulated E_b/N_o of 1dB.

4. Maximum Mutual Information Combining (MMIC)

The poor convergence of the turbo decoder using the Max-Log-MAP algorithm is due to the accumulating bias in the extrinsic information caused by the max() operations. Since extrinsic information is used as *a priori* information, $L_a(x_{t,0})$, for the next component decoding operation, and is combined with channel observations $L_c(x_{t,0})$, as shown in (4), this bias leads to sub-optimal combining proportions in the decoder. To correct for this phenomenon, the logarithmic transition probabilities at the i^{th} iteration may be modified as follows

$$\bar{\gamma}_t^{[q]}(l',l) = \frac{1}{2}\left\{\left\{w_a^{(i)} L_a^{(i)}(x_{t,0}) + L_c^{(i)}(x_{t,0})\right\} x_{t,0}^{[q]} + L_c^{(i)}(x_{t,1}) x_{t,1}^{[q]}\right\} \quad (9)$$

In other words, the bias of the *a priori* information can be corrected by scaling it by a factor $w_a^{(i)}$ at the i^{th} iteration, as depicted in Figure 4. This correction procedure for the Max-Log-MAP algorithm is far less complex than the correction function employed in the Log-MAP algorithm. Furthermore, and perhaps more importantly from a practical point of view, the corrected Max-Log-MAP algorithm remains insensitive to an arbitrary scaling of the LLR values at its input, thereby eliminating the need to estimate the noise variance at the channel output. From observations of the EXIT charts in the previous section, it is evident that rapid convergence of the turbo process relies on the effective exchange of mutual information between the component decoders. Consequently, it may be inferred that the optimum value for the weight factor $w_a^{(i)}$ is that which maximizes the mutual information of the term $\zeta_t^{(i)} = w_a^{(i)} L_a^{(i)}(x_{t,0}) + L_c^{(i)}(x_{t,0})$ for each component decoder and at each iteration i. Using vector notation, $\zeta_t^{(i)}$ may be modelled as

$$\begin{aligned}\zeta_t^{(i)} &= \begin{bmatrix} w_a^{(i)} & 1 \end{bmatrix} \begin{bmatrix} L_a^{(i)}(x_{t,0}) \\ L_c^{(i)}(x_{t,0}) \end{bmatrix} \\ &= (\underline{w}^{(i)})^T \underline{L}_t^{(i)} \\ &= (\underline{w}^{(i)})^T (\underline{\lambda}_t^{(i)} + \underline{\varepsilon}_t^{(i)}) = (\underline{w}^{(i)})^T \underline{\lambda}_t^{(i)} + v_t^{(i)} \end{aligned} \quad (10)$$

where $\underline{\varepsilon}_t^{(i)}$ represents the contributions of channel noise plus the numerical approximation error inherent in the Max-Log-MAP algorithm, and $\underline{\lambda}_t^{(i)}$ is the vector of uncorrupted LLRs.

Given variances

$$s_\zeta^{(i)} = \mathrm{E}\{(\underline{w}^{(i)})^{\mathrm{T}}(\underline{\lambda}_t^{(i)} + \underline{\varepsilon}_t^{(i)})(\underline{\lambda}_t^{(i)} + \underline{\varepsilon}_t^{(i)})^{\mathrm{T}}\underline{w}^{(i)}\}$$
$$= (\underline{w}^{(i)})^{\mathrm{T}}\boldsymbol{R}_{\lambda+\varepsilon}^{(i)}\underline{w}^{(i)} \quad (11)$$

$$s_v^{(i)} = \mathrm{E}\{(\underline{w}^{(i)})^{\mathrm{T}}\underline{\varepsilon}_t^{(i)}(\underline{\varepsilon}_t^{(i)})^{\mathrm{T}}\underline{w}^{(i)}\}$$
$$= (\underline{w}^{(i)})^{\mathrm{T}}\boldsymbol{R}_{\varepsilon}^{(i)}\underline{w}^{(i)} \quad (12)$$

and modelling $v_t^{(i)}$ as a Gaussian random variable, the differential and conditional entropies of $\zeta_t^{(i)}$ are

$$h(\zeta_t^{(i)}) = \tfrac{1}{2}\log\{2\pi e\, s_\zeta^{(i)}\} \quad (13)$$

$$h(\zeta_t^{(i)} \mid \underline{\lambda}_t^{(i)}) = \tfrac{1}{2}\log\{2\pi e\, s_v^{(i)}\} \quad (14)$$

By definition [7], the mutual information can be written as

$$I(\zeta_t^{(i)}; \underline{\lambda}_t^{(i)}) = h(\zeta_t^{(i)}) - h(\zeta_t^{(i)} \mid \underline{\lambda}_t^{(i)}) = \tfrac{1}{2}\log\frac{s_\zeta^{(i)}}{s_v^{(i)}} \quad (15)$$

and the optimum weight factors can then be derived as

$$\underline{w}_{\mathrm{OPT}}^{(i)} = \arg\max_{\underline{w}^{(i)}}\frac{s_\zeta^{(i)}}{s_v^{(i)}} = \arg\max_{\underline{w}^{(i)}}\frac{(\underline{w}^{(i)})^{\mathrm{T}}\boldsymbol{R}_{\lambda+\varepsilon}^{(i)}\underline{w}^{(i)}}{(\underline{w}^{(i)})^{\mathrm{T}}\boldsymbol{R}_{\varepsilon}^{(i)}\underline{w}^{(i)}} \quad (16)$$

Setting $\underline{z} = (\boldsymbol{R}_{\varepsilon}^{(i)})^{\mathrm{T}/2}\underline{w}^{(i)}$, we arrive at the quotient problem

$$\underline{z}_{\mathrm{OPT}}^{(i)} = \arg\max_{\underline{z}}\frac{\underline{z}^{\mathrm{T}}(\boldsymbol{R}_{\varepsilon}^{(i)})^{-1/2}\boldsymbol{R}_{\lambda+\varepsilon}^{(i)}(\boldsymbol{R}_{\varepsilon}^{(i)})^{-\mathrm{T}/2}\underline{z}}{\underline{z}^{\mathrm{T}}\underline{z}} \quad (17)$$

with solutions

$$\underline{z}_{\mathrm{OPT}}^{(i)} = k\,\mathrm{eig}_{\max}\left\{(\boldsymbol{R}_{\varepsilon}^{(i)})^{-1/2}\boldsymbol{R}_{\lambda+\varepsilon}^{(i)}(\boldsymbol{R}_{\varepsilon}^{(i)})^{-\mathrm{T}/2}\right\} \quad (18)$$

$$\underline{w}_{\mathrm{OPT}}^{(i)} = k\,(\boldsymbol{R}_{\varepsilon}^{(i)})^{-\mathrm{T}/2}\mathrm{eig}_{\max}\left\{(\boldsymbol{R}_{\varepsilon}^{(i)})^{-1/2}\boldsymbol{R}_{\lambda+\varepsilon}^{(i)}(\boldsymbol{R}_{\varepsilon}^{(i)})^{-\mathrm{T}/2}\right\} \quad (19)$$

where $\mathrm{eig}_{\max}(\boldsymbol{A})$ is the eigenvector of $\boldsymbol{A}$ corresponding to its largest eigenvalue. The scalar k is chosen such that the second element of $\underline{w}_{\mathrm{OPT}}^{(i)}$, i.e. the weight factor of $L_{\mathrm{c}}^{(i)}(x_{t,0})$, equals unity. Inspection of (10) to (19) reveals that the optimum weights are functions of the iteration index, the error correcting capabilities of the component decoders (i.e. encoder polynomials) and the signal to noise ratio. The optimum weights $\underline{w}_{\mathrm{OPT}}^{(i)}$ can be computed or "trained" off-line based on time-averaged estimates of correlation matrices $\boldsymbol{R}_{\lambda+\varepsilon}^{(i)}$ and $\boldsymbol{R}_{\varepsilon}^{(i)}$ derived over a sufficiently long data block corresponding to τ encoded information bits. Specifically

$$\boldsymbol{R}_{\lambda+\varepsilon}^{(i)} = \mathrm{E}\{\underline{L}_t^{(i)}(\underline{L}_t^{(i)})^{\mathrm{T}}\} = \lim_{\tau\to\infty}\sum_{t=1}^{\tau}\underline{L}_t^{(i)}(\underline{L}_t^{(i)})^{\mathrm{T}} \quad (20)$$

Furthermore, the vector $\underline{\lambda}_t^{(i)}$ of "uncorrupted" LLRs may be written as

$$\underline{\lambda}_t^{(i)} = \begin{bmatrix}\lambda_{\mathrm{a}}^{(i)}(x_{t,0})\\ \lambda_{\mathrm{c}}^{(i)}(x_{t,0})\end{bmatrix} = \begin{bmatrix}\mathrm{E}\{L_{\mathrm{a}}^{(i)}(x_{t,0})\,.\,x_{t,0}\}\,x_{t,0}\\ \mathrm{E}\{L_{\mathrm{c}}^{(i)}(x_{t,0})\,.\,x_{t,0}\}\,x_{t,0}\end{bmatrix} \quad (21)$$

so that

$$\boldsymbol{R}_{\lambda}^{(i)} = \mathrm{E}\{\underline{\lambda}_t^{(i)}(\underline{\lambda}_t^{(i)})^{\mathrm{T}}\} = \lim_{\tau\to\infty}\begin{bmatrix}(\varphi^{(i)})^2 & \varphi^{(i)}\theta^{(i)}\\ \varphi^{(i)}\theta^{(i)} & (\theta^{(i)})^2\end{bmatrix} \quad (22)$$

where $\varphi^{(i)} = \frac{1}{\tau}\sum_{t=1}^{\tau}L_{\mathrm{a}}^{(i)}(x_{t,0})\,.\,x_{t,0}$ and $\theta^{(i)} = \frac{1}{\tau}\sum_{t=1}^{\tau}L_{\mathrm{c}}^{(i)}(x_{t,0})\,.\,x_{t,0}$.

Finally, assuming that vectors $\underline{\varepsilon}_t^{(i)}$ and $\underline{\lambda}_t^{(i)}$ are uncorrelated, one may derive $\boldsymbol{R}_{\varepsilon}^{(i)}$ as $\boldsymbol{R}_{\lambda+\varepsilon}^{(i)} - \boldsymbol{R}_{\lambda}^{(i)}$. The above training procedure should be performed under E_b/N_o conditions that are typical at the bit-error rate range of interest.

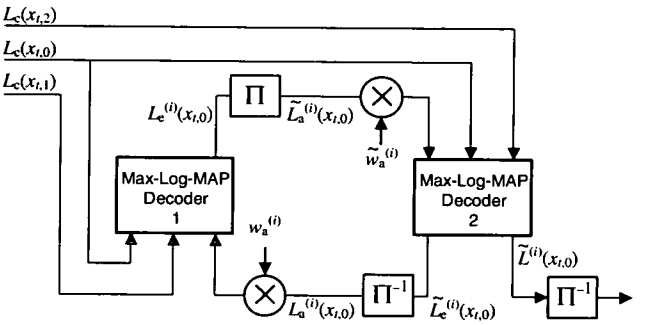


Figure 4 – Turbo decoding with weighting of *a priori* information.

5. SIMULATION RESULTS

Two different turbo encoders are considered at the input of an AWGN channel. The first 1/2 rate (punctured) turbo encoder consists of two 1/2 rate component RSC codes of memory 4, octal polynomials $(G_r, G) = (23,37)$ and an interleaver size of 10^5 bits. The second 1/2 rate (punctured) turbo encoder is that specified for UMTS [8] and consists of two 1/2 rate component RSC codes of memory 3, octal polynomials $(G_r, G) = (11,13)$ and an interleaver size of 5114 bits. This is the maximum block size specified for high speed downlink packet access (HSDPA) in UMTS. Table 1 shows the optimum weight factors derived off-line for each iteration of the two turbo decoders at E_b/N_o of 1.0 and 0.7 dB respectively. The impact of the combining scheme of (9) on the mutual information trajectory of the first turbo decoder is indicated in Figure 5. In comparison to the original trajectory of Figure 3, turbo decoding with the improved Max-Log-MAP algorithm does not stall and is able to converge almost as well as with the Log-MAP algorithm. This is achieved at the expense of only two additional multiplications per iteration per systematic bit. Figure 6 shows the BER performance of the first turbo decoder after 6 iterations.

Table 1 – Optimized weight factors.

	$(G_r, G) = (23,37)$ E_b/N_o=1.0dB		$(G_r, G) = (11,13)$ E_b/N_o=0.7dB	
Iteration i	$w_{\mathrm{a}}^{(i)}$	$\tilde{w}_{\mathrm{a}}^{(i)}$	$w_{\mathrm{a}}^{(i)}$	$\tilde{w}_{\mathrm{a}}^{(i)}$
1	0*	0.505	0*	0.517
2	0.566	0.602	0.581	0.617
3	0.629	0.656	0.640	0.668
4	0.682	0.712	0.683	0.713
5	0.754	0.814	0.732	0.769
6	0.892	1.020	0.792	0.837

* no *a priori* knowledge in iteration 1 for first component decoder.

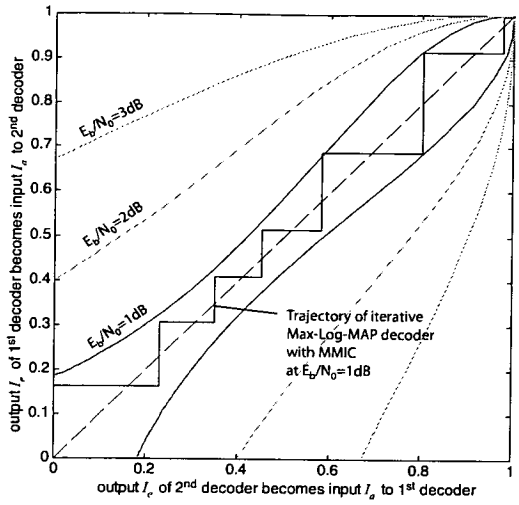


Figure 5 – EXIT chart for Max-Log MAP turbo decoder with MMIC.

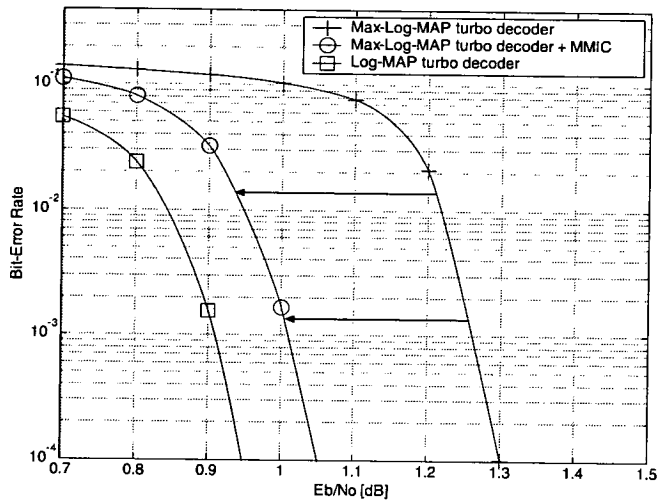


Figure 6 – Performance of first turbo decoder (16 states, 10^5-bit interleaver).

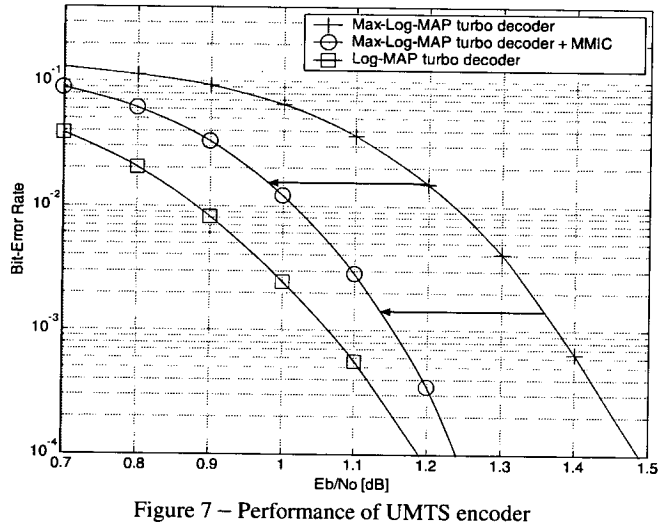


Figure 7 – Performance of UMTS encoder (8 states, 5114-bit interleaver).

The results show that the proposed MMIC scheme significantly improves the performance of the turbo decoder. Figure 7 shows the BER results for the UMTS turbo decoder after 6 iterations. Again, the performance of the turbo decoder using the Max-Log-MAP algorithm and MMIC approaches that of the turbo decoder using the optimum Log-MAP algorithm. The performance difference can be reduced down to only 0.05 dB at a BER of 10^{-4}.

6. CONCLUSIONS

A maximum mutual information combining (MMIC) scheme was proposed as a means to improve the performance of turbo decoders whose component decoders use the Max-Log-MAP algorithm. The convergence behaviour of such turbo decoders was investigated by using extrinsic information transfer (EXIT) charts. The combining scheme is achieved by iteration-specific scaling of the *a priori* information at the input of each component decoder in order to maximize the transfer of mutual information to the next component decoder, as suggested by the EXIT charts. The scaling corrects the accumulated bias introduced by the Max-Log approximation. A method for off-line computation of the optimum weight values was also described. It was shown that the proposed combining scheme significantly improves the performance of a turbo decoder using the Max-Log-MAP algorithm to within 0.05dB of a turbo decoder using the optimum Log-MAP algorithm. The improved decoder retains the low complexity and insensitivity to input scaling which are inherent advantages of the Max-Log-MAP algorithm.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. Stephan ten Brink and Dr. Magnus Sandell for their valuable input on the subjects of EXIT charts and MAP decoding.

REFERENCES

[1] C. Berrou, A. Glavieux, P. Thitimajshima, "Near Shannon limit error correcting coding and decoding: turbo codes", IEEE International Conference on Communications, pp. 1064-1070, May 1993.

[2] J. Hagenauer, E. Offer, L. Papke, "Iterative decoding of binary block and convolutional codes", IEEE Transactions on Information Theory, Vol. 42, No. 2, pp. 429-445, March 1996.

[3] B. Vucetic, J.Yuan, "Turbo codes", Kluwer Academic Publishers, 2000.

[4] P. Robertson, E. Villebrun, P. Hoeher, "A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain", IEEE International Conference on Communications, pp. 1009-1013, June 1995.

[5] S. ten Brink, "Convergence behaviour of iteratively decoded parallel concatenated codes", IEEE Transactions on Communications, Vol. 49, No. 10, pp. 1727-1737, October 2001.

[6] T. Richardson, R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding", IEEE Transactions on Information Theory, Vol. 47, No. 2, pp. 599–618, February 2001.

[7] S. Haykin, "Neural Networks", Prentice-Hall, 2nd edition, 1999.

[8] 3GPP, "TS 25.212 V3.11.0 (2002-09)", Technical Specification Group Radio Access Network, Multiplexing and Channel Coding (FDD), Release 1999.

High-Performance MIMO Receivers based on Multi-Stage Partial Parallel Interference Cancellation

Holger Claussen[1], Hamid Reza Karimi[2], Bernard Mulgrew[1]

[1] Signals & Systems Group, University of Edinburgh, UK
[2] Bell Labs Research, Lucent Technologies, Swindon, UK

Abstract – **Turbo-encoded multiple-input multiple-output (MIMO) radio links have been recently proposed for the support of high-speed downlink packet access (HSDPA) in UMTS, where the re-use of spreading codes across the transmitter antennas results in high levels of interference. The state-of-the-art receiver chain for such a link incorporates space-time channel equalization, de-spreading, pre-whitening and finally *a posteriori* probability (APP) detection. In this paper, a multi-stage partial parallel interference canceller (MS-PPIC) is considered as a low complexity alternative to the APP detector and its Max-Log variant. Non-linear cancellation metrics are derived for the MS-PPIC and its performance is compared with the APP detector for flat and dispersive channels. It is shown that the MS-PPIC can provide similar performance compared to APP and, for low coding rates, superior performance compared to Max-Log-APP, at a substantially lower computational complexity.**

1. INTRODUCTION

Turbo-encoded multiple-input multiple-output (MIMO) radio links have been recently proposed for the support of high-speed downlink packet access (HSDPA) in UMTS [1]. The objective is to increase the achievable data rates for a particular user through a combination of spreading code re-use across transmit antennas and higher-order modulation schemes. The code re-use inevitably results in high levels of interference at the mobile receiver, even under non-dispersive channel conditions. In order to tackle such high interference levels, MIMO receivers based on the *a posteriori* probability (APP) detector have been considered by the authors [2] where in order to deal with dispersive channels (while avoiding sequence estimation) the detector is preceded by space-time channel equalization [3]. The equalizer is then followed by a de-spreading operation which allows the APP to perform joint detection of bits transmitted from multiple antennas but corresponding to a single spreading code only, thereby resulting in a significant reduction in computational complexity.

In this paper, multi-stage partial parallel interference cancellation (MS-PPIC) is proposed as a low complexity alternative to the APP detector in the above receiver. A strategy for non-linear cancellation is derived analytically and the resulting performance of the MS-PPIC based receiver is compared with the APP and Max-Log-APP based receivers. In addition, the reductions in computational complexity are quantified by a complexity analysis of the detection algorithms.

2. SIGNAL MODEL

Figure 1 illustrates the transmission and reception scheme for the MIMO link under investigation.

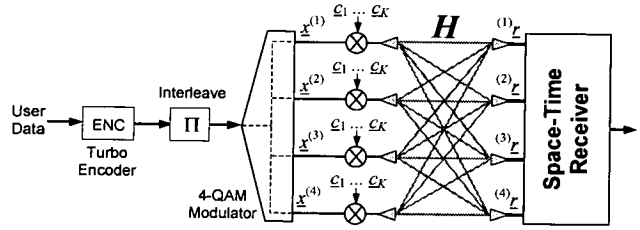


Figure 1 – 4×4 MIMO transmission and reception.

At the transmitter, user data is encoded and interleaved. The coded data stream is de-multiplexed into N_T sub-streams, corresponding to the N_T transmit antennas. Each sub-stream is then modulated on to NK 4-QAM symbols and subsequently spread by a factor Q via a set of K orthogonal spreading codes prior to transmission. Each transmitted spread stream then occupies N symbol intervals. Also note that the same set of K codes are re-used across all transmit antennas. Therefore, the MIMO propagation environment, which is assumed to exhibit significant multi-path, plays a major role in achieving signal separation at the receiver. The transmitted signals are received by N_R receive antennas after propagation through dispersive radio channels with impulse response lengths of W chips. The received signal vector may then be written as

$$\begin{bmatrix} {}^{(1)}\underline{r} \\ \vdots \\ {}^{(N_R)}\underline{r} \end{bmatrix} = \begin{bmatrix} {}^{(1)}\boldsymbol{H}^{(1)} & \cdots & {}^{(1)}\boldsymbol{H}^{(N_T)} \\ \vdots & \ddots & \vdots \\ {}^{(N_R)}\boldsymbol{H}^{(1)} & \cdots & {}^{(N_R)}\boldsymbol{H}^{(N_T)} \end{bmatrix} \sum_{k=1}^{K} \boldsymbol{C}_k \begin{bmatrix} \underline{x}_k^{(1)} \\ \vdots \\ \underline{x}_k^{(N_T)} \end{bmatrix} + \begin{bmatrix} {}^{(1)}\underline{n} \\ \vdots \\ {}^{(N_R)}\underline{n} \end{bmatrix} \quad (1)$$

or

$$\underline{r} = \boldsymbol{H} \sum_{k=1}^{K} \boldsymbol{C}_k \underline{x}_k + \underline{n} \quad (2)$$

where ${}^{(m)}\underline{r} \in C^{(QN+W-1)\times 1}$ is the signal received at the m^{th} antenna, ${}^{(\bar{m})}\boldsymbol{H}^{(i)} \in C^{(QN+W-1)\times QN}$ is the channel matrix from the i^{th} transmit antenna to the m^{th} receive antenna and $\underline{x}_k^{(i)} \in C^{N\times 1}$ is the sequence of N symbols $[x_k^{(i)}(1)\cdots x_k^{(i)}(N)]^T$ transmitted from the i^{th} antenna via the k^{th} spreading code. Noise vector $\underline{n} \in C^{N_R(QN+W-1)\times 1}$ consists of i.i.d. zero-mean complex Gaussian random variables representing additive thermal noise and inter-cell interference such that

$\boldsymbol{R}_n = \mathrm{E}\{\underline{n}\underline{n}^{\mathrm{H}}\} = N_0\boldsymbol{I}$. Finally $\boldsymbol{C}_k$ is the spreading matrix for k^{th} spreading code, $\underline{c}_k \in C^{Q\times 1}$, such that

$$\boldsymbol{C}_k = \underbrace{\begin{bmatrix} \underline{c}_k & \cdots & \underline{0} \\ \vdots & \ddots & \vdots \\ \underline{0} & \cdots & \underline{c}_k \end{bmatrix}}_{N_{\mathrm{T}}N \text{ Times}} \in C^{QN_{\mathrm{T}}N\times N_{\mathrm{T}}N} \tag{3}$$

The mapping of the 4-QAM constellation is such that $x_k^{(i)}(t) = b_{k,0}^{(i)}(t) + \mathrm{j}b_{k,1}^{(i)}(t)$ with $b_{k,q}^{(i)}(t) \in \{+1,-1\}$.

3. RECEIVER ARCHITECTURE

The optimum strategy for space-time detection at the receiver requires APP joint detection of KN_{T} symbols transmitted over N_{T} antennas and K codes at each symbol epoch. For 4-QAM modulation, and for dispersive channels with ISI extending over L symbols, this requires a search over a trellis containing $2^{2(L+1)KN_{\mathrm{T}}}$ states. The involved complexity is clearly prohibitive for typical parameter values. However, note that in flat fading conditions (L=0) and with K orthogonal codes re-used over the transmit antennas, the number of states reduces to a more realistic value of $2^{2N_{\mathrm{T}}}$. The above argument suggests that an efficient strategy for dealing with dispersive channels is to perform detection after a process of space-time equalization [3]. Figure 2 depicts the resulting receiver architecture [2]. The space-time equalizer mitigates the impact of the dispersive MIMO channel, $\boldsymbol{H}$, and nominally restores the signal conditions to those present at the transmitter. Signal contributions from the various spreading codes are then extracted via K de-spreading operations. The equalization and de-spreading processes inevitably result in different spatio-temporal colouring of the noise vector $\underline{n}$ along each of the K branches. This is accounted for via pre-whitening of the sufficient statistics prior to the detection process.

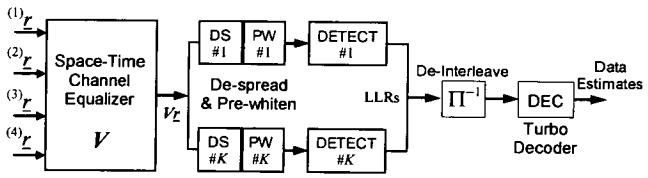


Figure 2 – MIMO receiver chain.

The pre-whitened sufficient statistics at the input of the k^{th} detector over the t^{th} symbol interval may be written as

$$\underline{z}_{\mathrm{w},k}(t) = \boldsymbol{A}_k(t)\underline{x}_k(t) + \underline{\varepsilon}_k(t) \tag{4}$$

for $k = 1\ldots K$ and $t = 1\ldots N$. Vector $\underline{x}_k(t) \in C^{N_{\mathrm{T}}\times 1}$ consists of symbols transmitted via the k^{th} code during the t^{th} symbol epoch and $\boldsymbol{A}_k(t) \in C^{N_{\mathrm{T}}\times N_{\mathrm{T}}}$ is the corresponding linear transformation (a function of the channel matrix $\boldsymbol{H}$, equalizer matrix $\boldsymbol{V}$ and spreading matrix $\boldsymbol{C}_k$). Vector $\underline{\varepsilon}_k(t)$ contains contributions due to residual spatio-temporal self- and cross-code interference, inter-cell interference and thermal noise.

As a result of pre-whitening, the elements of $\underline{\varepsilon}_k(t)$ are uncorrelated, i.e. $\mathrm{E}\{\underline{\varepsilon}_k(t)\underline{\varepsilon}_k^{\mathrm{H}}(t)\} = \boldsymbol{I}$. Note that full knowledge of the channel state information, as well as the noise power spectral density, N_0, is assumed. A detailed model of the equalizer output, de-spreading and pre-whitening is presented in [4].

The set of pre-whitened sufficient statistics, $\underline{z}_{\mathrm{w},k}(t)$ $\{k = 1\ldots K,$ $t = 1\ldots N\}$ are then applied to the detectors and the resulting soft outputs are de-interleaved and decoded. Two candidates for the detection process are described next. For clarity, the time index t, and spreading code index k, will be omitted.

3.1. A Posteriori Probability (APP) Detector

Given that the elements of $\underline{\varepsilon}$ in the expression $\underline{z}_{\mathrm{w}} = \boldsymbol{A}\underline{x} + \underline{\varepsilon}$ are i.i.d. zero-mean complex Gaussian random variables, it can be shown that soft information for the i^{th} bit of the symbol vector $\underline{x}$ may be derived in the form of log-likelihood ratios (LLR) as follows

$$\begin{aligned}\lambda(b_i) &= \ln\frac{\mathrm{P}\{b_i = +1 \mid \underline{z}_w\}}{\mathrm{P}\{b_i = -1 \mid \underline{z}_w\}} = \ln\frac{\sum_{\underline{x}|b_i=+1}\mathrm{f}(\underline{z}_w \mid \underline{x})\mathrm{P}\{\underline{x}\}}{\sum_{\underline{x}|b_i=-1}\mathrm{f}(\underline{z}_w \mid \underline{x})\mathrm{P}\{\underline{x}\}} \\ &= \ln\frac{\sum_{\underline{x}|b_i=+1}\exp\left\{-\left\|\underline{z}_w - \boldsymbol{A}\underline{x}\right\|^2 + \ln\mathrm{P}\{\underline{x}\}\right\}}{\sum_{\underline{x}|b_i=-1}\exp\left\{-\left\|\underline{z}_w - \boldsymbol{A}\underline{x}\right\|^2 + \ln\mathrm{P}\{\underline{x}\}\right\}}\end{aligned} \tag{5}$$

where f denotes probability density and $\mathrm{P}\{\underline{x}\}$ is the *a priori* probability of $\underline{x}$. Equation 5 represents what is commonly known as the maximum *a posteriori* probability (MAP) detector [5] or simply the *a posteriori* probability (APP) detector.

Clearly, the APP detector is not computationally friendly as it involves logarithms and exponentials. The computation of the LLRs can therefore be simplified by exploiting the well-known Max-Log approximation which states that $\ln(e^{\delta_1} + e^{\delta_2} + \cdots + e^{\delta_n}) \approx \max(\delta_1, \delta_2, \cdots, \delta_n)$. The Max-Log-APP detector may then be written as:

$$\begin{aligned}\lambda(b_i) \approx &\min_{\underline{x}|b_i=-1}\left\{\left\|\underline{z}_w - \boldsymbol{A}\underline{x}\right\|^2 - \ln\mathrm{P}\{x\}\right\} \\ &- \min_{\underline{x}|b_i=+1}\left\{\left\|\underline{z}_w - \boldsymbol{A}\underline{x}\right\|^2 - \ln\mathrm{P}\{x\}\right\}\end{aligned} \tag{6}$$

Note that the above reduction in computational complexity is naturally accompanied by some degradation in performance.

3.2. Multi-Stage Partial Parallel Interference Canceller (MS-PPIC)

Various forms of parallel interference cancellation have been considered in the past in the context of multi-user detection for the CDMA uplink [6]. The multi-stage partial parallel interference canceller (MS-PPIC) is considered in this paper

as an alternative to APP detection in the context of the CDMA MIMO downlink. Unlike APP, which is a single-shot joint detection process, the MS-PPIC involves multiple stages of "non-linear" cancellation, where at each stage the contributions due to interfering antennas are removed from the sufficient statistics at the detector input, thereby enhancing the detection process. Antenna interference contributions at the m^{th} stage of cancellation are constructed from "soft symbols" derived in the previous $(m-1)^{th}$ stage, as well as those derived most recently in the current m^{th} stage. Log-likelihood ratios are finally computed after the final stage where, subsequent to multiple stages of cancellation, additive Gaussian noise is the only remaining source of disturbance.

Performing matched filtering on the pre-whitened sufficient statistics of 4 and normalizing, we have

$$\underline{y} = \boldsymbol{\Delta}^{-1}\boldsymbol{A}^{\mathrm{H}}\underline{z}_{\mathrm{w}} = \boldsymbol{\Delta}^{-1}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\underline{x} + \boldsymbol{\Delta}^{-1}\boldsymbol{A}^{\mathrm{H}}\underline{\varepsilon} = \boldsymbol{\Delta}^{-1}\boldsymbol{R}\underline{x} + \underline{\eta} \qquad (7)$$

where $\boldsymbol{R}=\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}$, $\boldsymbol{\Delta}=\mathrm{diag}\{\boldsymbol{R}\}$ and $E\{\underline{\eta}\underline{\eta}^{\mathrm{H}}\} = \boldsymbol{\Delta}^{-1}\boldsymbol{R}\boldsymbol{\Delta}^{-\mathrm{H}}$. One may re-write this in the form

$$\begin{aligned}\underline{y} &= \underline{x} + \boldsymbol{\Delta}^{-1}(\boldsymbol{R}-\boldsymbol{\Delta})\underline{x} + \underline{\eta} = \underline{x} + \boldsymbol{\Delta}^{-1}\boldsymbol{R}'\underline{x} + \underline{\eta} \\ &= \underline{x} + \boldsymbol{S}\underline{x} + \underline{\eta} \end{aligned} \qquad (8)$$

where, given that $\boldsymbol{R}'$ and $\boldsymbol{S}$ both have zero diagonals, the term $\boldsymbol{S}\underline{x}$ clearly represents interference contributions which need to be cancelled. The sufficient statistics of 8 are applied to the MS-PPIC and may be viewed as the 0^{th} stage output, $\underline{y}[0]$, of the detector. The cancellation architecture is illustrated in Figure 3.

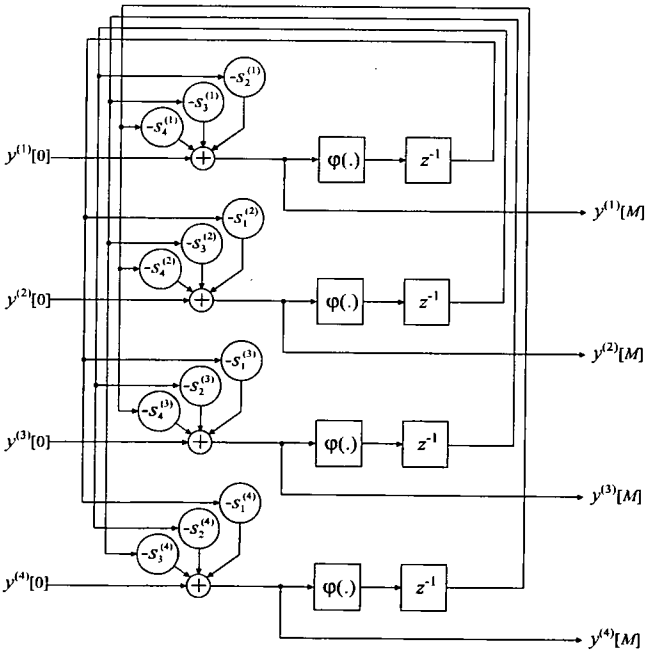


Figure 3 – MS-PPIC Architecture.

Denoting the i^{th} element of $\underline{y}$ as $y^{(i)}$, the i^{th} row of $\boldsymbol{S}$ as $\underline{s}^{(i)\mathrm{T}}$ and using '[m]' to identify the m^{th} stage, M stages of parallel cancellation may be described (see Appendix) as

for $m = 1 \cdots M$ (stages)
  $\underline{\xi} = \underline{y}[m-1] \in \mathrm{C}^{N_{\mathrm{T}}\times 1}$
  for $i = 1 \cdots N_{\mathrm{T}}$ (antennas)
    $y^{(i)}[m] = y^{(i)}[0]$
      $- s^{(i)\mathrm{T}}\left\{\tanh\{\boldsymbol{G}\,\mathrm{Re}(\underline{\xi})\} + \mathrm{j}\tanh\{\boldsymbol{G}\,\mathrm{Im}(\underline{\xi})\}\right\}$
    $\xi^{(i)} = y^{(i)}[m]$
  end
end

where $$\boldsymbol{G} = 2\left\{\mathrm{diag}\{2\boldsymbol{S}\boldsymbol{S}^{\mathrm{H}}\} + \boldsymbol{\Delta}^{-1}\right\}^{-1} \qquad (9)$$

is a diagonal matrix of reliability weights which define the "softness" of the cancellation process. Log-likelihood ratios are computed after the last stage where, due to multiple stages of cancellation, $y^{(i)}[M] \approx x^{(i)} + \eta^{(i)}$ and so

$$\lambda(b_0^{(i)}) = \frac{4\,\mathrm{Re}\{y^{(i)}[M]\}}{R_{i,i}^{-1}} \quad \lambda(b_1^{(i)}) = \frac{4\,\mathrm{Im}\{y^{(i)}[M]\}}{R_{i,i}^{-1}} \qquad (10)$$

where $R_{i,i}$ is the i^{th} diagonal element of $\boldsymbol{R}$.

4. COMPLEXITY COMPARISON

The computational complexity of the above detection algorithms can be compared in terms of the required number of real multiplications n_{MULT}, real additions n_{ADD}, table look-ups n_{TAB} and compare operations n_{CMP}.

For the initialisation of the APP and Max-Log-APP detectors, $n_{\mathrm{ADD}} \approx 2^{2N_{\mathrm{T}}}4N_{\mathrm{T}}^2$, where $2^{2N_{\mathrm{T}}}$ is the number of hypotheses over which the detector searches for 4-QAM modulation. This initialisation must be performed once for each spreading code during each channel estimation interval. At low mobile speeds, the latter is typically equal to a coding block interval. Following initialisation, the required number of operations per symbol period and per spreading code are $n_{\mathrm{MULT}} \approx 2^{2N_{\mathrm{T}}}2N_{\mathrm{T}}$, $n_{\mathrm{ADD}} \approx 2^{2N_{\mathrm{T}}}6N_{\mathrm{T}}$, $n_{\mathrm{CMP}} \approx 2^{2N_{\mathrm{T}}}2N_{\mathrm{T}}$ and $n_{\mathrm{TAB}} \approx 2^{2N_{\mathrm{T}}}2N_{\mathrm{T}}$ for the APP. The Jacobian algorithm [7] is used for efficient computation of (5) where the correction function is implemented via a look-up table. The corresponding values for the Max-Log-APP are $n_{\mathrm{MULT}} \approx 2^{2N_{\mathrm{T}}}2N_{\mathrm{T}}$, $n_{\mathrm{ADD}} \approx 2^{2N_{\mathrm{T}}}4N_{\mathrm{T}}$ and $n_{\mathrm{CMP}} \approx 2^{2N_{\mathrm{T}}}2N_{\mathrm{T}}$. For the initialisation of the proposed MS-PPIC detector $n_{\mathrm{MULT}} \approx 4N_{\mathrm{T}}^3 + 6N_{\mathrm{T}}^2$, $n_{\mathrm{ADD}} \approx 4N_{\mathrm{T}}^3 + 2N_{\mathrm{T}}^2$ and $n_{\mathrm{DIV}} \approx 2N_{\mathrm{T}}$, where n_{DIV} is the number of real divitions. Following initialisation, the required number of operations per symbol period and per spreading code are $n_{\mathrm{MULT}} \approx M(4N_{\mathrm{T}}^2 + 4N_{\mathrm{T}}) + 4N_{\mathrm{T}}^2$, $n_{\mathrm{ADD}} \approx M(4N_{\mathrm{T}}^2 + 4N_{\mathrm{T}}) + 4N_{\mathrm{T}}^2$ and $n_{\mathrm{TAB}} \approx 4N_{\mathrm{T}}M$ for the MS-PPIC. The tanh() function used for partial cancellation is implemented via a look-up table.

Table 1 shows a complexity comparison for the case of $N_T = N_R = 4$ and 4-QAM modulation. It is assumed that the MS-PPIC detector operates with $M = 6$ stages. As can be seen, the computational complexity of the proposed MS-PPIC detector is significantly lower than those of the APP and Max-Log-APP detectors.

Table 1 – Complexity comparison.

	Initalisation			Normal operation		
	APP	Max-Log APP	MS-PPIC	APP	Max-Log APP	MS-PPIC
n_{MULT}	0	0	353	2048	2048	544
n_{ADD}	16384	16384	288	6144	4096	544
n_{TAB}	0	0	0	2048	0	96
n_{CMP}	0	0	0	2048	2048	0

5. SIMULATION RESULTS

A 4×4 turbo-encoded MIMO link with parameters somewhat similar to the HSDPA specifications is considered. Data is 4-QAM modulated and spread by a factor of 16. The same set of 16 orthogonal Walsh spreading codes are simultaneously transmitted from each antenna. Coding of rate 1/3 is achieved via parallel concatenated turbo encoding (8-state component encoders) as in the UMTS specifications. A block of 1000 information bits then results in a coded block size of 3012 bits (transmitted over $N = 24$ symbols epochs). Coding of rate 1/2 is achieved via puncturing, resulting in a coded block size of 2006 bits. A total of 8 turbo decoder (Max-Log-MAP algorithm) iterations are performed at the receiver. Random space-time channel interleaving is also employed. In addition to a flat Rayleigh fading channel, a dispersive channel of 10 chip-spaced Rayleigh fading taps with a typical urban (TU) power profile is also considered. The channels are uncorrelated across the receiver antennas and the receiver is assumed to have perfect knowledge of the channel state information. Both ergodic (independent channel realisations per symbol epoch) as well as the more realistic block-fading (independent channel realisations per N symbol epochs) channel models are considered. Chip-rate transversal MMSE equalization with a span of 32 chips is applied in the block-fading scenarios.

Results for the ergodic model are compared with the MIMO capacity limit (in bits per channel use) as defined by $C = \mathrm{E}\{\log_2 \det\{\boldsymbol{I} + (E_s/N_0)\,\boldsymbol{H}\boldsymbol{H}^H\}\}$ where E_s is the symbol energy. For the simulation parameters considered, this implies a Shannon limit of $E_{b\text{-INFO}}/N_0 = -5.7$ dB for rate 1/3 and $E_{b\text{-INFO}}/N_0 = -4.8$ dB for rate 1/2 coding.

Figure 4 shows the performance results for ergodic flat fading without channel equalization. As expected, the optimal APP detector always offers the best performance, at only 2.7 dB away from the MIMO channel capacity limit. Interestingly, the results also indicate that the MS-PPIC outperforms the more practical Max-Log version of the APP detector at a coding rate of 1/3. However, at the higher coding rate of 1/2, the performance order is reversed.

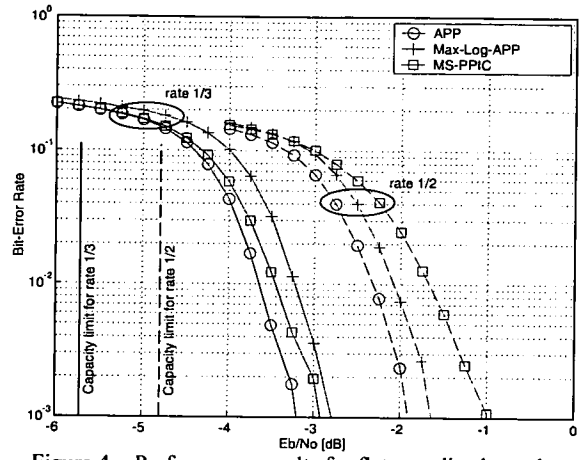


Figure 4 – Performance results for flat ergodic channels.

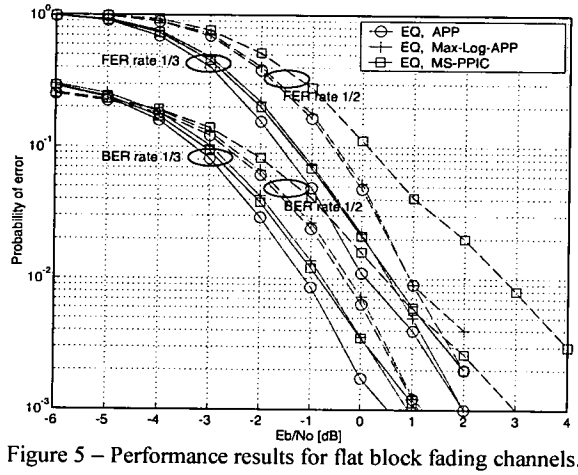


Figure 5 – Performance results for flat block fading channels.

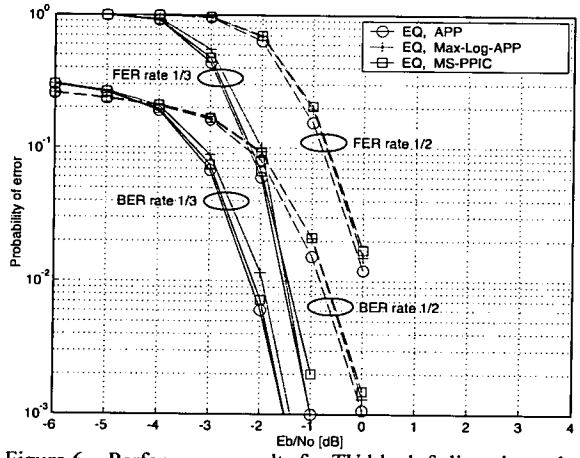


Figure 6 – Performance results for TU block fading channels.

The performance results for block fading are shown in Figures 5 and 6. For flat fading with 1/3 rate coding, the MS-PPIC receiver outperforms the Max-Log-APP. At rate 1/2, the Max-Log-APP performs slightly better than the proposed MS-PPIC receiver. In dispersive TU channels, the three receivers have

essentially the same performance for both 1/3 and 1/2 coding rates as indicated in Figure 6.

6. CONCLUSIONS

In this paper, multi-stage partial parallel interference cancellation (MS-PPIC) is proposed as an alternative to APP detection in a MIMO receiver architecture employing space-time channel equalization, de-spreading and pre-whitening. It is shown that the proposed MS-PPIC based receiver exhibits performance which is similar to that of the optimum APP based receiver (2.7dB away from the Shannon limit), and furthermore, is superior to receivers based on Max-Log-APP detection at low coding rates. In addition, it is shown that the MS-PPIC achieves this at a mere 25% of the Max-Log-APP computational complexity. More importantly, the proposed MS-PPIC is highly scalable in that its complexity grows only linearly with the number of transmitter antennas and not exponentially as in the APP detector.

APPENDIX – NON-LINEAR CANCELLATION

From (8), denoting the i^{th} element of $\underline{y}$ as $\underline{y}^{(i)}$ and the i^{th} row of $\boldsymbol{S}$ as $\underline{s}^{(i)\mathrm{T}}$, we have

$$y^{(i)}[0] = y^{(i)} = x^{(i)} + \underline{s}^{(i)\mathrm{T}}\,\underline{x} + \eta^{(i)} = x^{(i)} + v^{(i)}[0] \quad \text{(A1)}$$

and it immediately follows that cancellation at the m^{th} stage of the detector should be of the form

$$\begin{aligned} y^{(i)}[m] &= y^{(i)}[0] - \underline{s}^{(i)\mathrm{T}}\varphi\{\hat{\underline{x}}[m-1]\} \\ &= x^{(i)} + \underline{s}^{(i)\mathrm{T}}(\underline{x} - \varphi\{\hat{\underline{x}}[m-1]\}) + \eta^{(i)} \\ &= x^{(i)} + v^{(i)}[m] \end{aligned} \quad \text{(A2)}$$

where $\varphi\{\hat{\underline{x}}[m-1]\}$ is in general a non-linear function of tentative estimates, $\hat{\underline{x}}[m-1]$, derived in the previous stage. One could ignore the non-linearity and simply use the tentative estimates $\hat{\underline{x}}[m-1]$ directly in a linear cancellation process. It has been shown that (under certain constraints on the eigenvalues of $\boldsymbol{S}$) the resulting linear MS-PIC converges to the MMSE joint-detector as the number of stages approaches infinity [8]. At the other extreme, one could choose the function φ to be a mapping to the 4-QAM alphabet (i.e. a threshold operation). Such hard cancellation would perform well if and only if there was a high level of confidence regarding the reliability of tentative estimates $\hat{\underline{x}}[m-1]$. In order to deal with cases where the tentative estimates are unreliable, one may instead use the expected value, $\mathrm{E}\{\hat{\underline{x}}[m-1]\}$, of the tentative estimates in the cancellation process. Since $y^{(i)}[m-1] = x^{(i)} + v^{(i)}[m-1]$, then $\hat{x}^{(i)}[m-1] = y^{(i)}[m-1]$ and assuming that the noise plus interference term $v^{(i)}$ has a Gaussian distribution, it can readily be shown that

$$\begin{aligned} \varphi\left\{\hat{\underline{x}}^{(i)}[m-1]\right\} &\equiv \mathrm{E}\left\{\mathrm{Re}(\hat{\underline{x}}^{(i)}[m-1])\right\} + \mathrm{jE}\left\{\mathrm{Im}(\hat{\underline{x}}^{(i)}[m-1])\right\} \\ &\equiv \mathrm{E}\left\{\hat{\underline{b}}_0^{(i)}[m-1]\right\} + \mathrm{jE}\left\{\hat{\underline{b}}_1^{(i)}[m-1]\right\} \\ &\equiv \tanh\left\{\tfrac{1}{2}\lambda(\hat{\underline{b}}_0^{(i)}[m-1])\right\} + \mathrm{j}\tanh\left\{\tfrac{1}{2}\lambda(\hat{\underline{b}}_1^{(i)}[m-1])\right\} \\ &\equiv \tanh\left\{g^{(i)}[m]\ \mathrm{Re}(\underline{y}^{(i)}[m-1])\right\} \\ &\quad + \mathrm{j}\tanh\left\{g^{(i)}[m]\ \mathrm{Im}(\underline{y}^{(i)}[m-1])\right\} \end{aligned} \quad \text{(A3)}$$

where λ is the log-likelihood ratio and

$$g^{(i)}[m] = \frac{2}{\mathrm{E}\left\{\left|v^{(i)}[m-1]\right|^2\right\}} \quad \text{(A4)}$$

can be viewed as an antenna-dependent "softness" factor for the m^{th} stage. Factor $g^{(i)}[m]$ can be readily computed for the first stage:

$$g^{(i)}[1] = \frac{2}{\mathrm{E}\left\{\left|\underline{s}^{(i)\mathrm{T}}\underline{x} + \eta^{(i)}\right|^2\right\}} = \frac{2}{2\underline{s}^{(i)\mathrm{T}}\underline{s}^{(i)*} + R_{i,i}^{-1}} \quad \text{(A5)}$$

where $R_{i,i}$ the i^{th} diagonal element of $\boldsymbol{R}$. The computation of $g^{(i)}[m]$ is more involved for subsequent stages. Consequently, $g^{(i)}[1]$ is used for all stages $m = 1 \ldots M$. Though sub-optimal, this strategy should not significantly degrade performance in the SNR range of interest.

REFERENCES

[1] 3GPP, "TR 25.858 V5.0.0 (2002-03)", Technical Specification Group Radio Access Network, High Speed Downlink Packet Access: Physical Layer Aspects (Rel. 5).

[2] H.Claussen, H.R.Karimi, B.Mulgrew, "A Low Complexity Iterative Receiver based on Successive Cancellation for MIMO", Conference on Personal Wireless Communications PWC'2002, pp. 105-112, Oct. 2002, Singapore.

[3] G.Woodward, L.Mailaender, S.Venkatesan, "Chip Equalisation for Third Generation MIMO High Data Rate Applications", Proceedings of the Sensor Array and Multichannel Signal Processing Workshop, pp. 590-593, Aug. 2002.

[4] H.R.Karimi, H.Claussen, "Impact of Equalizer Output Modelling Errors on MIMO Receivers based on APP and PIC Detection", To be published Globecom 2003, Dec. 2003.

[5] B.M.Hochwald, S. ten Brink "Achieving Near-Capacity on a Multiple-Antenna Channel", IEEE Transactions on Communications, Vol. 51, No. 3, pp. 389-399, Mar. 2003.

[6] S. Moshavi, "Multi-User Detection for DS-CDMA Communications", IEEE Communications Magazine, pp. 124-136, Oct. 1996.

[7] B.Vucetic, J.Yuan, "Turbo codes", Kluwer Academic Publishers, 2000.

[8] L.K.Rasmussen, D.Guo, T.J.Lim, Y.Ma, "Aspects on Linear Parallel Interference Cancellation in CDMA", IEEE International Symposium on Information Theory", page 37, Aug. 1998.