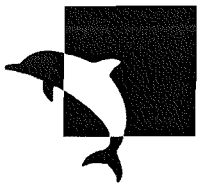


University of Southampton
Faculty of Engineering, Science and Mathematics
School of Electronics and Computer Science

Single-Carrier Time-Domain Space-Time Equalization Algorithms for the SDMA Uplink

by

Andreas Wolfgang

A thesis submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy at the University of Southampton

May 2007

SUPERVISOR:
Professor Lajos Hanzo

and

Professor Sheng Chen

This report is dedicated to:

My family

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

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

A thesis submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy

Single-Carrier Time-Domain Space-Time Equalization Algorithms for the SDMA Uplink

by Andreas Wolfgang

In this treatise I explore different detection algorithms for the employment in multiple-antenna aided wireless mobile communication systems. More explicitly, I explore the issue of detecting spatially-multiplexed information-carrying signals.

Specifically, in Chapter 2 I discuss several detection schemes, namely the linear Minimum Mean Squared Error (MMSE), the linear Minimum Bit Error Rate (MBER), the Bayesian as well as a novel Optimized Hierarchical Recursive Search Algorithm (OHRSA)-based detector, when operating in a narrowband communication environment. Furthermore, both the computational complexity and the achievable turbo-coded Bit Error Rate (BER) performance are characterized. The achievable BER performance is compared to the theoretical channel capacity limit and it was observed that the optimum Bayesian detector as well as the OHRSA-aided detector operate approximately 5 dB from the channel capacity limit.

In Chapter 3 I extend the narrowband detection strategies discussed in Chapter 2 to Space-Time Equalization (STE) algorithms, which also may benefit from Decision Feedback (DF). In addition to DF a further complexity reduction scheme is proposed for the OHRSA-aided STE, which exploits the specific characteristics of both the wideband channel and the proposed DF-STE. In comparison to the full Maximum Likelihood (ML) STE, the proposed detector is capable of achieving a complexity reduction, which is several orders of magnitude. The achievable turbo-coded BER performance is again compared to the channel capacity limit and it was observed that the distance with respect to the capacity bound is of the same order of magnitude as the distances observed for the narrowband system of Chapter 2, at the cost of an increased computational complexity.

In Chapter 4 I extend the detectors of Chapter 2 to Soft-Input Soft-Output (SISO) detection algorithms

and employ them in multiple-stage concatenated iterative systems. In order to analyze the behaviour of these multi-stage systems I introduce a new 3-Dimensional (3D) EXtrinsic Information Transfer (EXIT) chart analysis. The proposed iterative Bayesian detector as well as the novel Reduced-complexity Max-Log (RML) detector approach the channel capacity limit within the range of 1 dB, while the considered iterative MMSE detector is only capable of approaching the channel capacity limit, if the specific system considered is not rank-deficient.

Based on the derivations of Chapter 4 for narrowband MIMO systems I further evolve the SISO STEs of Chapter 3 to SISO STEs. In addition, the SISO STEs introduced may benefit from a novel soft-DF scheme. Owing to the increased search-space dimensions associated with STE the practical applicability of the Bayesian and the RML aided iterative STE is limited. Therefore a novel Markov Chain Monte Carlo (MCMC) aided STE is introduced, which is capable of exploring the increased search-space efficiently and which approaches the theoretical channel capacity limit within less than 2 dB also for high-dimensional search-spaces associated with higher-order modulation schemes.

Acknowledgment

I would like to take this opportunity to express my sincere gratitude to Professor Lajos Hanzo and Professor Sheng Chen. Their exceptional supervision and encouragement made this work possible; especially their generous investment of time and energy deserve my utmost acknowledgment.

I would like also to thank staff members and colleagues in the Communications Research Group, namely Dr. Lie Liang Yang, Dr. Soon X Ng, the group secretary Denise Harvey and many others, too numerous to mention here. My special thanks also goes to Dr. Hubert Dietl, Dr. Jos Akhtman, Ronald Tee and Seung H Won for the fruitful discussions as well as for their friendship.

Finally, my personal thanks are due to many precious friends both here as well as in Germany, my dear parents and my sister. Last but not least I want to thank my wonderful girlfriend Annika, who has supported me exceptionally well during the last three years and sacrificed much of her own career to be with me in England. I also want to thank my son Emil for making me laugh a thousand times a day.

List of Publications

[1] A. Wolfgang, S. Chen, and L. Hanzo, "Parallel interference cancellation based turbo space-time equalization in the SDMA uplink," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 609–616, 2007.

[2] J. Akhtman, A. Wolfgang, S. Chen, and L. Hanzo, "An optimized-hierarchy-aided approximate Log-MAP detector for MIMO systems." *IEEE Transactions on Wireless Communications.*, vol. 6, pp. 1900–1909, 2007.

[3] A. Wolfgang, J. Akhtman, S. Chen, and L. Hanzo, "Reduced complexity near maximum-likelihood uplink detection for decision feedback assisted space-time equalization." To appear in IEEE Transactions on Wireless Communications.

[4] A. Wolfgang, J. Akhtman, S. Chen, and L. Hanzo, "Iterative MIMO detection for rank-deficient systems," *IEEE Signal Processing Letters*, vol. 13, pp. 699–702, Nov. 2006.

[5] S. Chen, L. Hanzo, and A. Wolfgang, "Kernel-based nonlinear beamforming construction using orthogonal forward selection with Fisher ratio class separability measure.," *IEEE Signal Processing Letters*, vol. 11, pp. 478–481, 2004.

[6] S. Chen, L. Hanzo, N. Ahmad, and A. Wolfgang, "Adaptive minimum bit error rate beamforming assisted receiver for QPSK wireless communication," *Digital Signal Processing*, vol. 15, pp. 545–567, 2004.

[7] A. Wolfgang, S. Chen, and L. Hanzo, "Radial basis function network assisted space-time equalisation for dispersive fading environments," *IEE Electronics Letters*, vol. 40, no. 16, pp. 1006–1007, 2004.

[8] A. Wolfgang, N. Ahmad, S. Chen, and L. Hanzo, "Genetic algorithm assisted error probability optimisation for beamforming," *IEE Electronics Letters*, vol. 40, no. 5, pp. 320–322, 2004.

[9] S. Chen, L. Hanzo, and A. Wolfgang, "Nonlinear multiantenna detection methods," *EURASIP J. Applied Signal Processing,*, vol. 2004, no. 9, pp. 1225–1237, 2004. Special Issue on Advances in Smart Antennas 2.

[10] S. Chen, A. Wolfgang, C.J. Harris and L. Hanzo, "Adaptive nonlinear least bit error rate detection for symmetric RBF beamforming," invited submission to Special Issue of *Neural Networks*, 2007.

[11] S. Chen, A. Wolfgang, C.J. Harris and L. Hanzo, "Symmetric RBF classifier for nonlinear detection in multiple-antenna aided systems," *IEEE Transactions on Neural Networks*, accepted for publication, 2007.

[12] S. Chen, A. Wolfgang and L. Hanzo, "Constant modulus algorithm aided soft decision directed scheme for blind space-time equalisation of SIMO channels," *Signal Processing*, to appear, 2007.

[13] J. Wang, S. X. Ng, A. Wolfgang, L. L. Yang, S. Chen, and L. Hanzo, "Near-capacity three-stage MMSE turbo equalization using irregular convolutional codes," in *Proceedings of the 4th International Symposium on Turbo Codes*, April 2006, Electronic Publication

[14] S. Chen, A. Wolfgang, S. Benedetto, P. Dubamet, and L. Hanzo, "Symmetric radial basis function network equaliser," in *Proceedings of NEWCOM-ACoRN Joint Workshop*, (Vienna, Austria), 2006.

[15] A. Wolfgang, J. Akhtman, S. Chen, and L. Hanzo, "Reduced complexity single-carrier maximum-likelihood detection for decision feedback assisted space-time equalization," in *Proceedings of IEEE WCNC'06, Las Vegas, Nevada, USA.*, vol. 1, pp. 142–146, 2006.

[16] A. Wolfgang, S. Chen, and L. Hanzo, "Decision feedback aided Bayesian turbo space-time equalizer for parallel interference cancellation in SDMA systems," in *Proceedings of IEEE Vehicular Technology Conference*, Stockholm, 30 May-1 June 2005, vol. 4, pp. 2112–2116, 2005.

[17] A. Wolfgang, S. Chen, and L. Hanzo, "Radial basis function aided space-time equalization in dispersive fading uplink environments," in *Proceedings of IEEE Vehicular Technology Conference*, Stockholm, 30 May-1 June 2005, vol. 3, pp. 1552–1556, 2005.

[18] S. Chen, A. Wolfgang, and L. Hanzo, "A robust nonlinear beamforming assisted receiver for BPSK signalling," in *Proceedings of IEEE Vehicular Technology Conference*, Dallas, 26-28 of Sept. 2005, vol. 3, pp. 1921–1925, 2005.

[19] A. Wolfgang, S. Chen, and L. Hanzo, "Radial basis function network assisted wide-band beamforming," in *Proceedings of IEEE Vehicular Technology Conference*, Los Angeles, USA, 26-29 Sept., pp. 266–270, Sept. 2004.

[20] S. Chen, L. Hanzo, N. Ahmad, and A. Wolfgang, "Adaptive minimum bit error rate beamforming assisted QPSK receiver," in *Proceedings of IEEE International Conference on Communications*, Paris, May 2004, vol. 6, pp. 3389–3393, 2004.

[21] S. Chen, A. Wolfgang, and L. Hanzo, "Constant modulus algorithm aided soft decision-directed blind space-time equalization for SIMO channels," in *Proceedings of IEEE Vehicular Technology Conference*, Los Angeles, USA, 26-29 Sept., vol. 3, pp. 1718–1722, 2004.

[22] S. Chen, A. Wolfgang, Y. Shi and L. Hanzo, "Space-time decision feedback equalization using a minimum bit error rate design for single-input multi-output channels," *IET Communications*, to appear, 2007.

[23] S. Chen, A. Wolfgang, C.J. Harris and L. Hanzo, "Symmetric kernel detector for multiple-antenna aided beamforming systems," to be presented at *International Joint Conference on Neural Networks* (Orlando, Florida), Aug.12-17, 2007.

[24] M. Abuthinien, S. Chen, A. Wolfgang and L. Hanzo, "Joint maximum likelihood channel estimation and data detection for MIMO systems," to be presented at *International Conference on Communications* (Glasgow, Scotland, UK), June 24-28, 2007.

Contents

Notations

General

x	Scalar variable
$\check{x}$	Indicates a hypothetical solution to an optimization problem
$\hat{x}$	Estimated variable
$\check{\hat{x}}$	Hypothetically estimated variable
$\tilde{x}$	Sliced (hard) estimate of a variable
$\bar{x}$	Soft-estimate of a symbol
$p(\bullet)$	Probability density function of a random variable
$P(\bullet)$	Probability of a random variable

Vectors

$\boldsymbol{x}$	Column vector of scalars
x_i	i^{th} scalar element of the vector $\boldsymbol{x}$
$\mathbf{x}$	Column super-vector (vector of vectors)
$\boldsymbol{x}_i$	i^{th} vector element of the super-vector $\mathbf{x}$

Matrices

$\boldsymbol{X}$	Matrix of scalars
x_{ij}	$(i,j)^{\text{th}}$ scalar element of the matrix $\boldsymbol{X}$
$\mathbf{X}$	Super-matrix (Matrix of matrices)
$\boldsymbol{X}_{ij}$	$(i,j)^{\text{th}}$ matrix element of the super-matrix $\mathbf{X}$
$\boldsymbol{x}_i$	Vector equals the i^{th} column of the matrix $\boldsymbol{X}$
$\mathbf{x}_i$	Super-vector equals the i^{th} column of the super-matrix $\mathbf{X}$

Operators

*	Conjugate complex
^{T}	Transpose of a vector or matrix
^{H}	Hermitian (conjugate transpose) of a vector or matrix
$\otimes$	Kronecker product
$\mathtt{vec}(\bullet)$	Column wise vector operator
$\mathtt{trace}(\bullet)$	Trace of a matrix
$\mathtt{sign}(\bullet)$	Signum function
$\mathtt{diag}(\bullet)$	Diagonal elements of a matrix
$\Re(\bullet)$	Real part of a variable
$\Im(\bullet)$	Imaginary part of a variable

List of symbols

k	Discrete time index
$\mathcal{M}$	Modulation index
N_{bit}	Number of bits transmitted per channel use
R_{c}	Code-rate
R	Throughput
$\boldsymbol{I}_i$	$(i \times i)$-dimensional identity matrix
σ^2_{Tx}	Total transmit power of a mobile station
σ_{n}	Standard deviation of the Additive White Gaussian Noise (AWGN)
L	Length of the Channel Impulse Response (CIR)
l	Index associated with a particular CIR path
M	Number of transmit Antenna Elements (AEs) per Mobile Station (MS)
m	Transmit AE index
N	Number of receive AEs at the Base Station (BS)
n	Receive AE index
y	Channel output
$\boldsymbol{y}$	Channel output vector
h	Channel coefficient
$\boldsymbol{H}$	Channel matrix
f	Frequency Domain (FD) channel coefficient
$\boldsymbol{F}$	FD channel matrix
x	Transmitted symbol
$\boldsymbol{x}$	Transmitted symbol vector
$\boldsymbol{W}$	Filter weight matrix
$\boldsymbol{w}$	Filter weight vector
Q	Number of MSs (Users)

q	MS index
$\mathfrak{M}$	Total number of transmit AEs in the system ($\mathfrak{M} = Q \cdot M$)
$\mathfrak{m}$	Index associated with $\mathfrak{M}$
λ	Log Likelihood Ratio
η	AWGN
$\mathcal{Y}$	Set of legitimate channel output states
$\mathcal{X}$	Set of legitimate transmitted symbols
$\hat{\mathcal{X}}$	Set of legitimate detector output states

Chapter 1

Introduction

1.1 State-of-the-art and Trends

The increasing demand for the higher data rates to be supported by future wireless systems [1] inevitably comes at the cost of an increased bandwidth requirement, since the bandwidth is proportional to the symbol rate of the transmitted signal. In order to efficiently exploit the limited bandwidth available, advanced communication systems consider the employment of multiple transmit and receive antennas for the transmission of independent bit-streams. The employment of these Multiple-Input Multiple-Output (MIMO) [2] systems for spatial multiplexing has been motivated by the work of Foschini and Gans, who demonstrated [3] that under idealized assumptions the capacity of the wireless channel increases linearly with the minimum of the number of transmit and receive antennas. Their work constitutes a logical evolution of the channel capacity analysis provided by Shannon in [4–6]. In order to exploit this potential increase in channel capacity and the associated increase of the theoretically achievable data throughput, MIMO transceivers are expected to become an integral part of wireless communication systems, as outlined for example in [7–9]. Further work on the capacity of MIMO channels has been provided for example in [10] by Marzetta and Hochwald, in [11, 12] by Goldsmith *et. al.* as well as by Shamai and Marzetta in [13], who have contributed an extensive tutorial overview in partnership with Biglieri *et. al.* in [14].

However, even with the advent of MIMO transceivers and higher order modulation schemes, the bandwidth occupied by the transmitted signal is often significantly higher than the coherence bandwidth [15] of the wireless channel and will therefore result in multi-path propagation of the transmitted signal. The multi-path channels result in a frequency selective wireless propagation medium, which are also often referred to as wideband channels.

As an example, we will now briefly outline how the 3rd Generation Partnership Project (3GPP) and its Long-Term Evolution (LTE) [7, 16] task-force attempt to solve the problems associated with the high signal

bandwidth of future wireless mobile communication systems[1].

First of all, one has to distinguish between the bit-rate demand on the downlink also referred to as the forward link and the uplink often referred to as the reverse link. The two main differences are:

- The traffic is expected to be asymmetric, where the downlink typically has to support higher data rates than the uplink.
- The power consumption as well as the hardware cost imposes a less problematic constraint on the Base-Station (BS) than it does for the Mobile-Station (MS).

Given these characteristics, Orthogonal Frequency-Division Multiplexing (OFDM) [17] is considered to be the most potent transmission scheme in the 3GPP LTE drafts. The OFDM scheme, which belongs to the family of Multi-Carrier (MC) [18] transmission schemes is capable of supporting high data-rates, while the detection of the signal at the MS can be achieved at a relatively low computational cost. The problem of a high Peak-to-Average Power Ratio (PAPR) [17] can be solved by employing a more linear and hence less power-efficient amplifier at the BS transmitter's Radio Frequency (RF) front-end or the PAPR reduction techniques of [17, 19, 20] may be employed.

However, since mobile operators want to keep the cost and the power consumption of MSs low, OFDM is less attractive for employment in the uplink of the communication system. For this reason and owing to the lower PAPR, it was proposed in [7, 16] to consider Single-Carrier (SC) transmission for the reverse link as part of the 3GPP LTE.

The 3GPP LTE initiative was used as an example here, but the above-mentioned arguments in favour of SC transmissions in the uplink also hold for other wireless communication systems. Against this background, the performance of different SC STE approaches is discussed in this treatise. In Table 1.1 a list of selected publications discussing the LTE of the 3GPP system is provided.

1.2 Detection Techniques

In this treatise we will refer to a MIMO system, when the transmitter as well as receiver employ multiple antennas. Since we mainly consider reverse link or uplink communications, we also may use the term MS instead of the term transmitter and equivalently use the term BS rather than receiver. If not stated otherwise, we will assume that independent data streams are transmitted over the different transmit Antenna Elements (AEs) and thus the system considered is an uplink Space Division Multiplexing (SDM) [17] system. By contrast, when considering multiple MSs, which may or may not employ multiple transmit AEs, the system

[1]The exact proposal has not been finalized at the time of writing.

Table 1.1: Papers discussing the trade-off between MC systems and SC transmissions in the context of the 3GPP LTE.

[21] Zhou *et. al.*	Discusses different downlink multiple-access techniques for 3G and beyond with a special focus on the Chinese system proposals.
[22] Zhang *et. al.*	Presents an OFDM-based Time Division Duplex (TDD) system.
[23] Jamalipour *et. al.*	Compares different multiple-access architectures and comes to the conclusion that only a combination of existing multiple-access techniques is suitable for future wireless systems.
[7] Ekström *et. al.*	Introduces the evolution of the current system defined in the 3GPP proposals [16] in order to meet future user expectations. Multiple-access schemes as well as the physical layer architecture are discussed. For the uplink a SC system architecture is proposed.
[16] 3GPP	Presents different proposals for the physical layer of the 3GPP LTE. The proposals discussed are - among others - the employment of SC transmission in the reverse link as well as OFDM signals in the forward link. Other schemes discussed are for example MC - Code Division Multiple Access (CDMA). A finalized proposal is expected to appear in September 2007.

is referred to as a Space Division Multiple Access (SDMA) [17] system. Both SDM and SDMA systems aim for maximizing the throughput or multiplexing gain [24] of the system. Rather than maximizing the multiplexing gain, the other option is to maximize the diversity gain [25] of a system, which may be achieved by transmitting the same or related data over several transmit antennas, which renders communications more reliable. Naturally, there is a trade-off between the achievable multiplexing gain and diversity gain, which has been discussed for example by Tse *et. al.* in [25] or in [26].

The algorithms used for detecting signals in SDM and SDMA systems are in principle similar. The difference is that in a SDM system the transmitter might employ Space-Time Coding (STC) [27] and that the entries of the channel matrix might be constructed differently. Since the detection strategies are similar for SDM and SDMA, we do not explicitly distinguish between them when discussing the existing literature relevant for this treatise. Furthermore, since MC systems divide the wideband channel into numerous parallel narrowband channels, we assume that they both employ the same detection strategies.

1.2.1 Narrowband Single-Carrier and Multi-Carrier Detection

In Figure 1.1 the family of different MIMO detection approaches are classified. Firstly, we distinguish between linear and non-linear detectors. The most well known linear MIMO detectors are the Minimum Mean Squared Error (MMSE) detector as well as the Least Square (LS) detector, which are discussed in detail in numerous text-books, such as [17, 28]. In [29] Wang and Poor have extended the conventional MMSE detector to a Soft-Input Soft-Output (SISO) detector suitable for iterative detection. In Section 2.3.1 we derive the MMSE SDM detector for our specific system model, which is introduced in Section 2.1.1 and discuss the impact of choosing the MSE as our optimization criterion. Furthermore, both the computational complexity and the achievable Bit Error Rate (BER) performance are portrayed in Section 2.3.1.1 and Section 2.3.1.2,

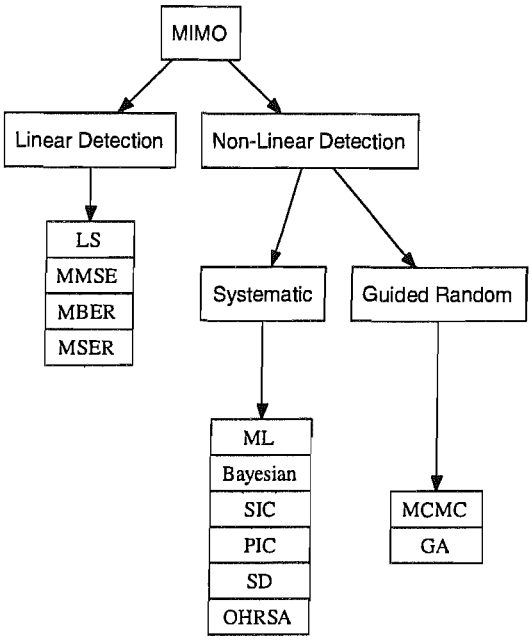


Figure 1.1: Classification of MIMO detection algorithms.

respectively. The corresponding extension to SISO MMSE detection is provided in Section 4.4.1.

The MMSE detection approach is optimum in terms of minimizing the Mean Squared Error (MSE) of a linear detector. However, a linear detector designed for achieving the lowest Bit Error Rate (BER) is optimum in terms of minimizing the BER. Hence, the set of linear detectors, which achieve the minimum BER are referred to as Minimum Bit Error Rate (MBER) detectors. They have for example been studied in [30, 31] by Chen *et al.*. A simplified MBER detector has been proposed for example by Gesbert in [32] on the basis of a closed-form expression for certain channel conditions. At the time of writing MBER detectors have mainly been proposed for Binary Phase Shift Keying (BPSK) signals [33] as well as 4-level Quadrature-Amplitude Modulation (QAM) signals [34]. Since MBER detectors are challenging to derive for higher-order QAM, in [35] Yeh and Barry have proposed an approximate Minimum Symbol Error Rate (MSER) detector for higher order QAM, while in [36] Chen *et. al.* have derived an exact MSER detector for higher order QAM. In this treatise we consider the MBER detector designed for BPSK as well as 4QAM signals, as discussed in Section 2.3.2. In Section 2.3.2.3 we propose a novel way of calculating the MBER solution for 4QAM signals, which relies on a real-valued binary representation of the complex-valued system model. The complexity of the MBER MIMO detector is discussed in Section 2.3.2.4, while the achievable performance is investigated in Section 2.3.2.5. Furthermore, we highlight in Section 4.4.4 why it may be unattractive to use iterative MBER detection. Further applications of the MBER approach include multi-user detection discussed by Li *et. al.* in [37] as well as MBER power allocation introduced by Wang and Blostein in [38].

In systems, where the number of transmit Antenna Elements (AEs) is higher than the number of receive AEs

the linear detectors of Figure 1.1 are often incapable of correctly detecting the desired signal and hence the family of non-linear detectors has to be employed. In Figure 1.1 we make a distinction between non-linear detectors, which are constructed in a systematic way and algorithms, which rely on guided random search strategies.

The classic non-linear detectors are the Maximum Likelihood (ML) [39] and the Bayesian [39] detector. The philosophical difference between these two approaches is discussed in Sections 2.3.3 and 2.3.4. It was shown for example in [39] that the Bayesian detector achieves the lowest possible BER of all non-linear detectors. In Section 2.3.3 the Bayesian detector is discussed and it is shown using a simple example, how it constructs a non-linear decision boundary. The SISO Bayesian detector is discussed in Section 4.4.2

Calculating the exact ML detector with 'brute force' imposes a high computational complexity on the receiver. Therefore, inspired by the Sphere Decoding (SD) algorithm introduced by Porst and Finke [40], Vitelko and Boutros have opened up a whole new research area with their work presented in [41], where they applied the original SD algorithm to communication systems. Many of the earlier publication on SD, such as [42] by Brunel and Boutros or [43] by Hassibi and Vikalo consider systems, where the number of transmit and receive antennas is identical.

Damen *et. al.* [44] have extended the classical SD to the so called Generalized Sphere Decoder (GSD), which is capable of operating in systems, where the number of transmit AEs is higher than the number of receiver AEs. These systems are often also referred to as rank-deficient [2] systems. Further SDs, which were designed for rank-deficient systems have been proposed by Pham *et. al.* [45], by Yang *et. al.* [46] as well as by Guo and Nilsson [47]. Note that none of the SD algorithms mentioned in this paragraph is guaranteed to find the ML solution, neither are they capable of processing soft-inputs, which is essential for efficient iterative detection.

Hochwald and ten Brink were among the first researchers to extend SD to Soft-Input Soft-Outpout (SISO) output algorithms, which resulted in the List Sphere Decoder (LSD) portrayed in [48]. Furthermore, also Vitalko *et. al.* [49] as well as Wang and Gianakis [50] have presented an iterative SD, which is capable of operating close to the channel's capacity. Pauli *et. al.* have proposed a SD [51], which is capable of detecting Differential Binary Phase Shift Keying (DPSK) modulated signals employed in a MC- Code Division Multiple Access (CDMA) system.

In contrast to the above-mentioned SD, the Optimized Hierarchy Reduced Search Algorithm (OHRSA) presented in [60, 61] and to be discussed in Section 2.3.4, has the capability of efficiently operating in so-called rank-deficient systems and is guaranteed to find the ML solution. In Section 4.4.3 the OHRSA algorithm is extended to accommodate soft information in the context of iterative receivers.

[2] A system is referred to as rank-deficient if the number of linearly independent columns of the associated system matrix is higher than the number of linearly independent rows.

Table 1.2: Selection of narrowband MIMO detection papers.

[52] Thoen *et. al.*	A LS MIMO detector is introduced and its performance is investigated when applied in a SDMA OFDM system.
[29] Wang and Poor	In this paper conventional MMSE detection is extended to SISO MMSE detection, which may be employed in iterative receivers.
[31] Chen *et. al.*	Presents a MBER beamformer designed for BPSK and 4QAM signals as well as for static channel conditions.
[32] Gesbert	The author presents a robust MBER MIMO detector, which can be constructed using a closed-form expression, provided that certain channel conditions are fulfilled.
[53] Foschini	The first low complexity V-BLAST receiver designed for MIMO based systems is proposed.
[54] Lee *et. al.*	The original V-BLAST technique is extended to a SISO algorithm and is employed in an OFDM system. For the scenarios considered the proposed V-BLAST detector achieves the ML performance.
[41] Viterbo and Boutros	First paper, which applied the SD proposed in [40] to the detection of received signals. This contribution inspired a whole new research area.
[44] Damen *et. al.*	In this paper the original SD is extended to a GSD, which is capable of operating in rank-deficient systems.
[48] Hochwald and ten Brink	The authors propose a List Sphere Decoder (LSD), which is capable of processing soft information and compare the attainable performance of their LSD to the channel capacity bound, which is also derived in the paper.
[49] Vikalo *et. al.*	A SISO SD is proposed, which is employed in an iterative system using different convolutional codes as well as Low-Density Parity-Check (LDPC) [55] codes.
[46] Yang *et. al.*	An improved generalized hard-output SD is introduced, which is designed for rank-deficient MIMO systems. The high-complexity detection process is divided into two detection stages, which with significantly reduced the complexity.
[51] Pauli *et. al.*	A SD is proposed in order to reduce the high computational complexity associated with the SISO detection of DPSK signals.
[47] Guo and Nilsson	A SD algorithm based on K-best Schnorr-Euchner (KSE) decoding is proposed, which is capable of providing both hard as well as soft outputs. Furthermore, hardware based performance results are presented.
[50] Wang and Giannakis	The original SD algorithm is extended to an exact Max-Log detector, which is employed in an iterative system.
[56] Santiago Mozos and Fernandez-Getino Garcia	Extends the SD using real-valued signals to a SD considering complex-valued signals, which is capable of detecting arbitrary modulation constellations. The performance of the proposed SD is investigated in the context of MC CDMA
[57] Zhu *et. al.*	In this paper MCMC aided MIMO detection is proposed and the performance is compared to SD algorithms. The results suggest that MCMC aided detection is capable of outperforming SD at a similar computational cost.
[58] Farhang-Boroujeny *et. al.*	MCMC based MIMO detection is discussed and the effects of different methods used for generating soft-information are presented. These methods include taking the empirical average as well as using importance sampling.
[59] Aggarwal and Wang	Presents a MCMC based detector optimized for MIMO systems employing higher-order QAM signals. In order to reduce the computational complexity of the proposed system, the received signal space is partitioned into subspaces, each of which is optimized independently.

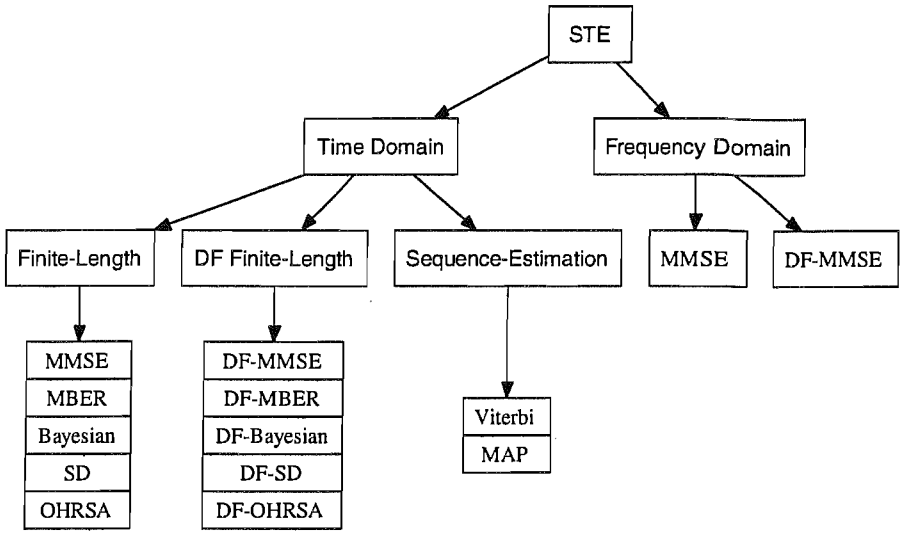


Figure 1.2: Classification of STE algorithms.

A set of detectors, which exhibits a low computational complexity is constituted by the family of interference cancellation based algorithms [62]. The Vertical Bell LAbs Layered Space Time (V-BLAST) technique introduced by Foschini [53], which is based on Serial Interference Cancellation (SIC) has inspired a spate of papers [54, 63, 64] investigating MIMO detection algorithms. More specifically, Zanella *et al.* [63] has provided a comparative study of MMSE and V-BLAST detection, while Lee *et al.* [54] as well as Kim and Kim [64] have recently extended the original V-BLAST algorithm to a more powerful SISO detector. In contrast to SIC, Parallel Interference Cancellation (PIC) has been discussed in [17, 65–68].

All SD algorithms are known to impose a low computational complexity on the receiver, if the channel Signal to Noise Ratio (SNR) is sufficiently high. By contrast, at low SNR values the complexity of SD algorithms dramatically increases, if no specific precautions are taken. Researchers have therefore considered the employment of guided random detection strategies, such as for example Markov Chain Monte Carlo (MCMC) methods [57, 58, 69] as well as Genetic Algorithms (GAs) [70]. MCMC techniques have been shown to exhibit a lower computational complexity than classic SD algorithms, especially in the low-SNR region [57].

In Table 1.2 we have provided a number of relevant references for MIMO detection, including their short description. The MMSE, LS as well as ML MIMO detectors can be found in numerous textbooks such as for example in [17, 28, 39] and they are therefore not included in Table 1.2. Non-linear detection algorithms, such as for example those introduced in [71, 72] by Chen *et al.*, which do not rely on knowing the channel at the receiver, are not included in Table 1.2, either.

1.2.2 Single-Carrier Wideband Detection

In addition to the classification of the different detection algorithms featuring in Figure 1.1, SC STEs might be further classified according to Figure 1.2. The first distinction is made between Time-Domain (TD) and Frequency-Domain (FD) STEs. While most detection algorithms listed in Figure 1.1 might be extended to finite-length STEs or Decision Feedback (DF) assisted finite-length STEs, the FD-STE has mainly been investigated in the context MMSE and LS detection.

Again, linear finite-length TD STEs can be directly constructed from narrowband detectors. The MMSE STE as well as the DF aided MMSE STE has been discussed for example in [73–75]. In this treatise, we will discuss the MMSE STE in Section 3.4.1, while its extension to DF-aided MMSE STE will be provided in Section 3.5.1. The first SISO MMSE STE has been proposed by Abe and Matsumoto in [76], while the extension from narrowband MBER detection to MBER aided STE has been published in [77] by Chen *et al.*. In Section 3.4.2 and Section 3.5.2, respectively, we will introduce a novel representation of both the MBER and the DF-aided MBER STE, while novel Bayesian as well as OHRSA aided STEs will be discussed in Sections 3.4.3, 3.4.4 as well as in Section 3.5.3 and in Section 3.5.4. A detailed complexity evaluation will be provided for each of the algorithms and their performance is studied in Section 3.5.5.

In addition to finite-length filtering-based equalizers, trellis-based sequence-estimation aided STEs may be constructed, which aim for detecting a whole sequence of transmitted data rather than performing a bit-by-bit based detection. The Maximum A Posteriori (MAP) STE has for example been discussed by Tonello in [78], while in [79] Vitelko *et al.* have introduced a hybrid detection scheme combining SD and the classic Viterbi [80, 81] algorithm.

Let us now first discuss the fundamental difference between TD and FD STEs. For channels having severe delay spread and hence requiring a high-order TD equalizer, FD equalization - as in OFDM - is computationally more advantageous than using high-order TD equalization. When comparing FD equalization aided classic serial modems and OFDM systems, it was pointed out in [82] by Falconer *et al.* that SC FD equalization has the capability of achieving a better BER performance than OFDM systems. A comparison of their achievable throughput has recently been provided by Louveaux *et al.* in [83]. In [84] a SD based cyclic-prefix aided STE was proposed, for short transmission block lengths typically on the order of 32 bits. In Table 1.3 a selection of papers discussing different STE approaches is provided.

1.3 Outline and Novelty

Against the background discussed in Section 1.2, in **Chapter 2** we first explore the performance of different narrowband MIMO detection techniques under diverse channel conditions. The differences between

Table 1.3: Selection of STE papers.

[73] Al-Dhahir and Sayed	A theoretical analysis of DF-aided MMSE STE is provided, including its diversity analysis.
[75] Tidestav *et. al.*	A MMSE DF STE is discussed in the context of a SDMA system and particular emphasis is dedicated to the near-far effects.
[85] Tüchler *et. al.*	In this paper BPSK SISO equalization is extended to higher-order modulation schemes.
[76] Abe and Matsumoto	The authors extend the MMSE SISO detector introduced in [29] to a STE employed in a SDMA system in conjunction with simple iterative channel estimation.
[77] Chen *et. al.*	A MBER STE is designed for BPSK modulated signals transmitted over a time-invariant CIR. The design is closely related to MBER beamforming and MBER single-user equalization.
[79] Vikalo *et. al.*	TD STE based on the Viterbi algorithm, which employs sphere decoding for complexity reduction.
[86] Dong and Wang	SISO STE based on the MCMC method
[87] Dietl and Utschick	The authors propose a linear TD STE based on the Wiener filter. In order to reduce the associated complexity, reduced-rank subspaces are considered for detection. Furthermore, the performance of the proposed scheme in iterative systems is investigated.
[88] Kwan and Kok	TD STE in precoded MIMO systems.
[89] Duan *et. al.*	TD wideband beamforming using Infinite Impulse Response (IIR) rather than Finite Impulse Response (FIR) filters. The paper proposed different methods for combating the stability problems of the proposed IIR filter structures.
[90] Li and Wong	SISO TD turbo equalizer based on the Kalman algorithm. The weights of the SISO equalizer are updated directly rather than by considering channel estimation and equalization as two separate blocks.
[91] Roy *et. al.*	SISO TD STE based on Kalman filtering. The STE weights are calculated directly without estimating the channel first.
[82] Falconer *et. al.*	MMSE FD-STE is compared to multi-carrier systems such as OFDM and the associated system differences are highlighted.
[83] Louveaux *et. al.*	Another paper comparing OFDM and SC FD DF-STE.
[84] Li *et. al.*	The authors propose a FD SD based STE for short block-length.
[92] Benvenuto *et. al.*	A set of different receiver architectures for DF STE are proposed. While the feed-forward filter is implemented in the FD, different feedback-filter implementations are proposed both in the frequency and the time-domain.
[93] Ng *et. al.*	Linear SISO FD STE using soft-decision feedback. Also iterative channel estimation is considered.

the various detection strategies considered, namely the MMSE, the MBER [94–96], the Bayesian as well as the novel OHRSA aided detector [97] are highlighted using a simple example. Furthermore, we derive channel capacity bounds for the system considered and benchmark the performance of the detection algorithms against this theoretical bound. In addition, the effects of using estimated rather than perfect CSI is investigated. The aim of Chapter 2 is to provide the theoretical basis of the detection strategies using the same underlying system model and to discuss the complexity associated with the different algorithms.

In **Chapter 3** we extend the narrowband detection strategies discussed in Chapter 2 to STE algorithms, which additionally may benefit from DF. Chapter 3 contains the following novel contributions

- A MBER and DF-MBER STE is proposed for both BPSK and 4QAM signals [98];
- A non-recursive Bayesian as well as DF aided Bayesian STE is designed [98–100];
- An OHRSA STE as well as a DF-aided OHRSA STE is contrived [101, 102].

Furthermore, we provide a detailed discussion on how the DF structure may be interpreted physically. Finally, the achievable performance is compared to the theoretical channel capacity bound and a detailed complexity discussion is provided.

In **Chapter 4** we extend the detectors of Chapter 2 to SISO detection algorithms and employ them in multiple-stage concatenated iterative systems. In order to analyze the behaviour of these multi-stage systems we introduce a new 3-Dimensional (3D) EXtrinsic Information Transfer (EXIT) chart analysis. We show that with the aid of a Reduced complexity Max-Log (RML) detector we are capable of operating within 2 dB of the channel's capacity bound. The main novel contributions of Chapter 4 are

- A 3D-EXIT chart, which is reduced to a 2-Dimensional (2D) EXIT chart using projection techniques [103];
- A novel reduced complexity Max-Log detector, which is based on the OHRSA [104].

It will be shown in **Chapter 5** that simply extending the SISO MIMO detection algorithm of Chapter 4 to iterative MIMO STE results in both a severe performance degradation due to error-propagation as well as a significantly increased computational cost. In order to address these difficulties we therefore introduce:

- A novel soft-decision technique in Section 5.1.3 in order to combat error-propagation;
- A MCMC aided STE is designed in Section 5.3 in order to perform non-linear detection in the increased search-space.

Our conclusions provided in **Chapter 6** establish the logical connection between the different chapters and highlight general performance as well as complexity trends of the systems considered.

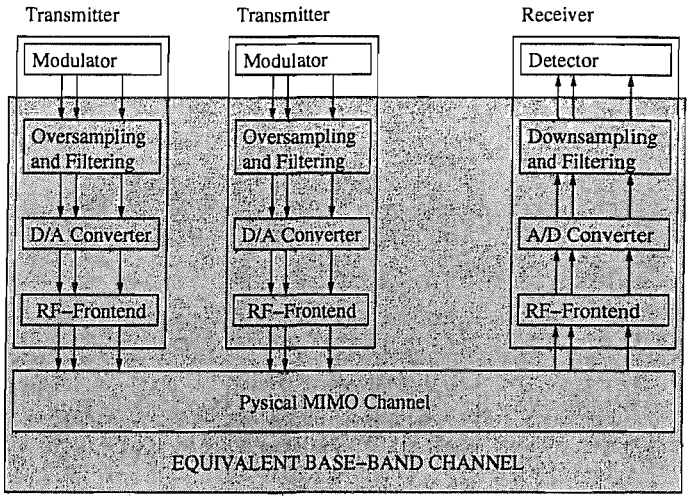


Figure 1.3: Equivalent baseband channel model.

1.4 System and Channel Model

In order to allow a simple mathematical representation of the communication systems considered in this treatise we will employ the so-called equivalent baseband channel model.

Transmitter: In Figure 1.3 we have illustrated a communication system composed of two transmitters and a single receiver, where it can be seen that the output of the MIMO transmitters, which is in general a complex-valued time-domain signal also referred to as the baseband signal, is oversampled and passed through a pulse-shaping filter. The resultant band-limited signal is then Digital-to-Analog (D/A) converted and passed on to the RF-frontend, where it is up-converted to the RF band.

Physical Channel: The RF signal is transmitted over a wireless channel having a given Channel Impulse Response (CIR). If the bandwidth of the transmitted signal is wider than the coherence bandwidth of the channel, then the signal will be subject to multi-path propagation. Multi-path channels are typically characterized by their Power Delay Profile (PDP), which describes the time-averaged CIR as illustrated in Figure 1.4. Realistic PDPs proposed in the various standard bodies are typically not symbol spaced.

Most of the recent channel models represent the multi-path channel as a set of about six scattering clusters [105], each corresponding to one propagation path. If there exists no Line-of-Sight (LOS) paths between the transmitter and the receiver, then each multi-path component $h_i(t)$ is typically modeled as a Wide Sense Stationary (WSS) complex Gaussian process. The time-varying behaviour, i.e. the correlation between the complex-valued channel gain h_i at time t_1 and time t_2 may be written as

$$r_h(t_1, t_2) = \mathrm{E}\left[h_i(t_1)h_i^*(t_2)\right]. \tag{1.1}$$

The TD correlation of a specific path is characterized by the TD auto-correlation function, while the corre-

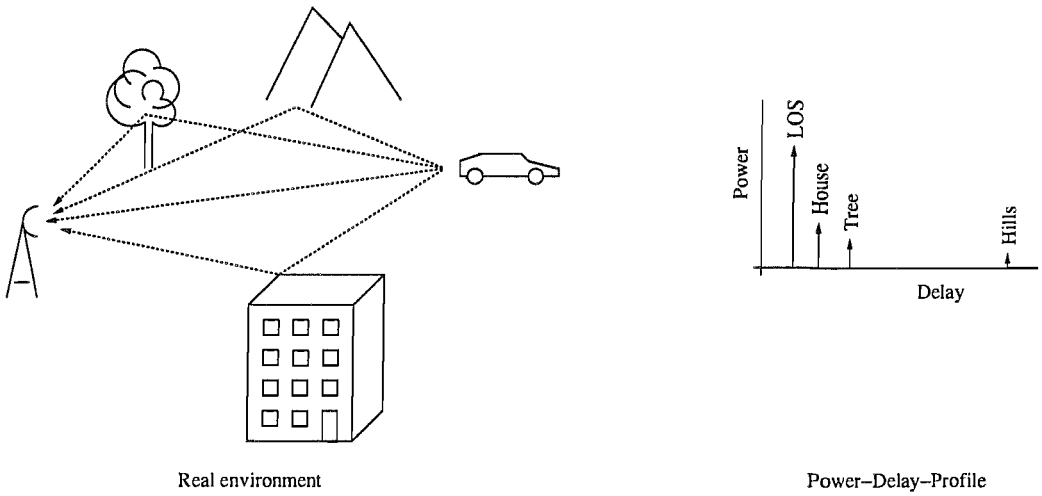


Figure 1.4: Communication environment and corresponding PDP.

sponding FD characterization is given by the Power Spectral Density (PSD), which is the Fourier transform of the TD auto-correlation function. The PSD is mainly characterized by the Doppler frequency f_D , where we have

$$\mathrm{PSD(f)} = 0 \text{ if } |\mathrm{f}| > \mathrm{f_D}. \tag{1.2}$$

The most popular PSD is the U-shaped Jakes spectrum [15] . Its auto-correlation function is described by the zero-order Bessel function as [15]

$$r_h(t_1, t_2) = J_0(2\pi f_D(t_1 - t_2)). \tag{1.3}$$

Receiver: At the receiver, the signal is down-converted again to the baseband and Analog-to-Digital (A/D) converted. The receiver then estimates the optimum timing-delay, applies the receive filter and down-samples the signal again. Throughout this treatise we assume that the baseband signal is sampled at the symbol rate, resulting in a symbol-spaced equivalent baseband CIR. In other words, the combination of transmit filter, physical channel and receive filter is sampled at symbol rate resulting in the equivalent baseband CIR.

Chapter 2

Narrowband Single-Carrier Multiple-Input Multiple-Output Detection

This chapter commences with an introduction to the basic system model and continues with an illustration of the challenges a Multiple-Input Multiple-Output (MIMO) detector faces, when operating in a single-carrier narrowband environment. In the second section of this chapter theoretical channel capacity limits are presented for the system model considered, which serve as benchmark schemes throughout this treatise. In Section 2.3 a set of different MIMO detectors, namely the Minimum Mean Squared Error (MMSE), the Minimum Bit Error Rate (MBER), the Bayesian detector as well as a reduced complexity Maximum Likelihood (ML) detector are derived. At the end of Section 2.3 the achievable performance of the various methods is evaluated under the assumption of having perfect channel knowledge at the receiver. This assumption is relaxed in Section 2.4, where a set of different MIMO channel estimators are introduced and the performance of the proposed MIMO detectors is investigated using the estimated, rather than the perfect Channel State Information (CSI).

2.1 Introduction

The aim of this overview section is to first introduce the system model and then to present our assumptions stipulated for the communication scenario considered. Based on this system model it will be illustrated with the aid of a simple example, how MIMO detectors operate.

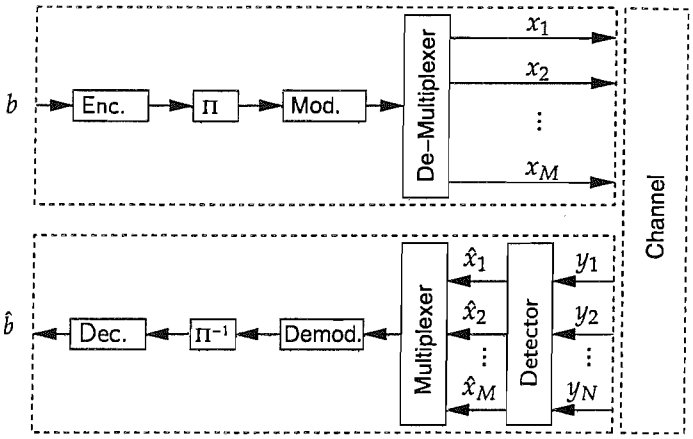


Figure 2.1: Baseband MIMO system model.

2.1.1 System Model

The system considered consists of a single Mobile-Station (MS) employing an M-element transmit antenna array and a Base-Station (BS) receiver, which has N number of Antenna Elements (AEs). The MS's transmitter channel encodes the input bit-stream at a code-rate R_c, pass the coded bits through an interleaver, modulate the signals and map them to the M different transmit AEs, as illustrated in Figure 2.1. The modulated symbols are transmitted to the BS over a narrowband fading channel characterized by the channel coefficients h_{nm}, as illustrated in Figure 2.1.

The channel coefficient h_{nm} represents the complex-valued channel coefficient between transmit AE m and the n^{th} BS receiver AE. Given the transmitted symbol $x_m(k)$, which is associated with transmit AE m, the output signal of the n^{th} AE of the BS receiver at time instant k can be written as

$$y_n(k) = \sum_{m=1}^{M} h_{nm} s_m(k) + \eta_n(k), \tag{2.1}$$

where $\eta(k)$ is the complex-valued Additive White Gaussian Noise (AWGN) having a variance of $\mathrm{E}\left[|\eta_n(k)|^2\right] = 2\sigma_n^2$.

Assuming that the MS transmits the modulated symbols to the BS over a narrowband channel at a power σ_{Tx}^2, the resultant Signal to Noise Ratio (SNR) is given as

$$\text{SNR} = \frac{\sigma_{\text{Tx}}^2 \sum_{n=1}^{N} \sum_{m=1}^{M} \mathrm{E}\left[|h_{nm}|^2\right]}{NM2\sigma_n^2}. \tag{2.2}$$

For a given modulation scheme having $\mathcal{M}$ number of modulation levels and a channel code with code-rate

R_c the associated energy per information bit to noise ratio of E_b/N_0 can be written as

$$\frac{E_b}{N_0} = \frac{\sigma_{Tx}^2 \sum_{n=1}^{N} \sum_{m=1}^{M} \mathrm{E}\left[|h_{nm}|^2\right]}{R_c \log_2(\mathcal{M}) MN 2\sigma_n^2}. \tag{2.3}$$

The number of bits per transmitted symbol is denoted as $N_{bit} = \log_2(\mathcal{M})$.

2.1.1.1 Complex-Valued Vector Model

Under the assumption of perfect synchronization, the signal transmitted by the MS's AEs and the channel's output is described by a $(N \times M)$-dimensional matrix $\boldsymbol{H}$, where the $(n,m)^{\text{th}}$ element of the matrix is given by h_{nm}. The channel's output vector $\boldsymbol{y}(k)$ can now be expressed as

$$\boldsymbol{y}(k) = \boldsymbol{H}(k)\boldsymbol{x}(k) + \boldsymbol{\eta}(k), \tag{2.4}$$

where the column vector $\boldsymbol{x}(k) = [x_1(k), \ldots, x_M(k)]^T$ contains the symbols transmitted by the MS's AEs and the associated AWGN is given by $\boldsymbol{\eta}(k) = [\eta_1(k), \ldots, \eta_N(k)]^T$.

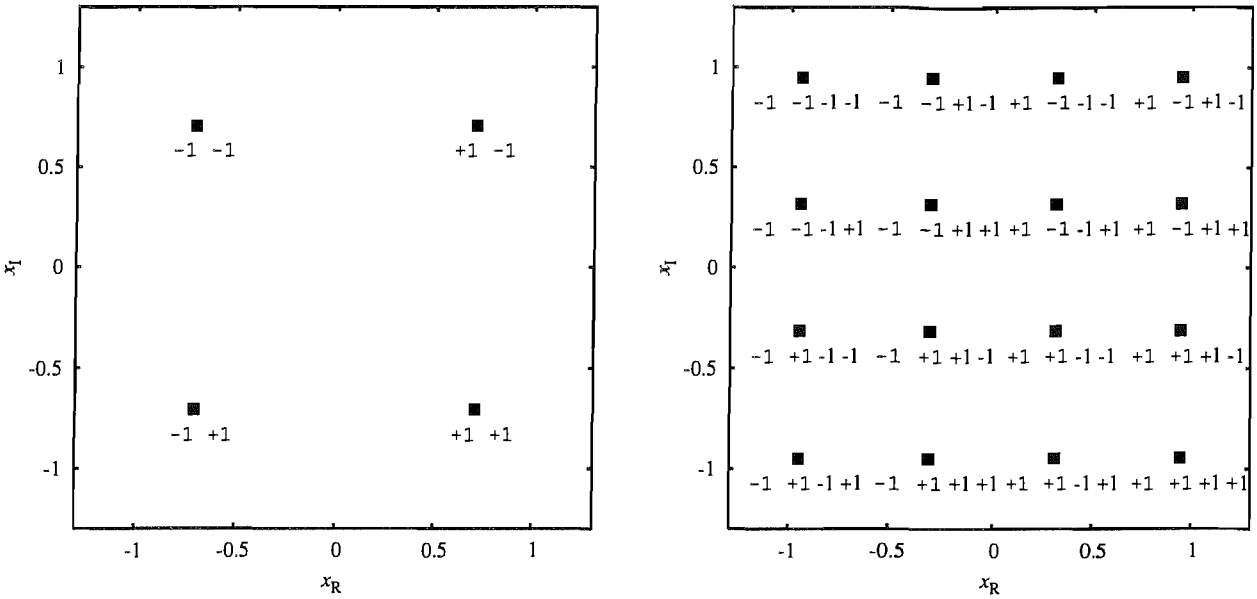
2.1.1.2 Real-Valued Binary Vector Model

For the derivation of bit-based detection algorithms and for the evaluation of the computational complexity imposed by different detection algorithms, it is often desirable to characterize the system using purely real-valued channel matrices and binary input symbols. Many detection strategies referred to as lattice[1] based detectors consider a search space of hypothetical, discrete solutions to an optimization problem. This family of detectors [42, 43, 49, 107] can be derived on the basis of a real-valued binary system representation, as outlined for example in [49]. In order to arrive at a real-valued binary system representation when encountering complex-valued CIR matrices combined with higher order modulation schemes, the system may also be represented in a binary form. Therefore the transformation matrix $\boldsymbol{Q}$ is introduced, which is defined as

$$\mathbf{Q} = \mathbf{I}_M \otimes \boldsymbol{q}, \tag{2.5}$$

where $\otimes$ denotes the Kronecker product and $\boldsymbol{I}_M$ is the $(M \times M)$-dimensional identity matrix. The choice of the row-vector $\boldsymbol{q}$ depends on the specific modulation scheme and is defined to ensure that a sequence of N_{bit} binary input symbols, which can assume a value of ± 1 is transformed into a complex-valued symbol of the specific modulation scheme used. For example, in case of BPSK we have $q = [1]$ and for 4QAM we

[1]The term lattice refers to a discrete subgroup of the Euclidean space [106]. Lattice based detectors are therefore based on a discrete model of the observation space rather than a continuous space model.



(a) 4QAM constellation.

(b) 16QAM constellation.

Figure 2.2: Phasor constellation obtained using Equation (2.6) and the binary symbols $\underline{x}$ as given in the graphs.

Table 2.1: Transformation vector for higher-order modulation schemes. The quantization vectors for higher-order QAM constellation do not use Gray mapping but rather natural enumerate.

Modulation scheme	Transformation vector
BPSK	[1]
4QAM	$1/\sqrt{2} \cdot [1 \; -j]$
16QAM	$1/\sqrt{10} \cdot [2 \; 1 \; -2j \; -1j]$

employ $q = [\frac{1}{\sqrt{2}} \; -\frac{1}{\sqrt{2}}j]$. More specifically, the modulator is represented by a matrix operation as follows

$$x = q\underline{x}, \tag{2.6}$$

where $\underline{x}$ is a vector of length $N_{\text{bit}} = \log_2(\mathcal{M})$, whose elements may assume the values of ± 1. In Table 2.1 the transformation vector values q are summarized for the modulation schemes considered in this treatise.

In Figure 2.2a the phasor constellation for a 4QAM constellation obtained with the aid of Equation (2.6) is illustrated, where the transformation vector q can be found in Table 2.1 and the binary symbol vector $\underline{x}$ is given below each phasor point. Similarly, the phasor constellation for the considered 16QAM scheme is illustrated in Figure 2.2a where it can be observed that the 16QAM constellation is a hierarchical extension of the 4QAM constellation, i.e in each quadrant of the complex plane the phasor points are a 4QAM constellation.

Equation (2.4) can now be re-written as

$$\begin{bmatrix} \Re(\boldsymbol{y}(k)) \\ \Im(\boldsymbol{y}(k)) \end{bmatrix} = \begin{bmatrix} \Re(\boldsymbol{H}(k)\boldsymbol{Q}) \\ \Im(\boldsymbol{H}(k)\boldsymbol{Q}) \end{bmatrix} \begin{bmatrix} \underline{x}_1(k) \\ \vdots \\ \underline{x}_{M\log_2(\mathcal{M})}(k) \end{bmatrix} + \underline{\eta}(k), \tag{2.7}$$

where $\Re(\cdot)$ and $\Im(\cdot)$ indicate the real and the complex part of a complex-valued variable, respectively, the index of $\underline{x}_{i+(m-1)M}$ indicates the i^{th} binary symbol of the complex-valued symbol x_m and $\underline{\eta}(k) = [(\Re(\eta(k)))^T\ (\Im(\eta(k)))^T]^T$ is the real-valued AWGN having a variance of σ_n^2. *Note that for the real-valued binary system model the antenna index m ranges from 1 to $M\log(\mathcal{M})$*

Applying the transformation described by Equation (2.7) to Equation (2.4) transforms the system having complex-valued input symbols into a bit-based real-valued system having $\log_2(\mathcal{M})$-times the number of real-valued input symbols. For BPSK modulated signals the matrix $\boldsymbol{Q}$ degenerates to the identity matrix. For the remainder of this treatise the 'underlined' variables are real-valued. If the transformation described by Equation (2.7) is applied to Equation (2.4), the following real-valued binary system is generated

$$\underline{\boldsymbol{y}}(k) = \underline{\boldsymbol{H}}(k)\underline{\boldsymbol{x}}(k) + \underline{\eta}(k). \tag{2.8}$$

2.1.2 Problem Definition

A major problem in MIMO based communication systems is the Co-Channel Interference (CCI), which is caused by several interfering transmit AEs communicating over the same channel as the desired transmit AE. Different techniques of combating the CCI have been proposed and a very effective one is constituted by the employment of multiple receive AEs, where the spatial characteristics of the channel constituted by the unique antenna-specific CIRs are exploited to distinguish between the different transmit AEs' signals. Provided that the CIRs are sufficiently accurately estimated, unique CIRs experienced by each transmitter may be viewed as a unique signature of the transmitters' signals similarly to the unique user-specific spreading sequences used in the operational Code Division Multiple Access (CDMA) systems.

Example 2.1: MIMO Detection Problem

Let us now consider a three-element transmit and a two-element receive antenna array as well as BPSK modulated signals transmitted over a narrowband channel characterized by the CIR matrix

$$\boldsymbol{H} = \begin{bmatrix} -0.50 & 0.75 & 0.55 \\ -0.25 & -0.85 & 0.70 \end{bmatrix}, \tag{2.9}$$

where the non-dispersive CIR taps of the three AEs transmitting to the first receive AE are given by the first row, while those with respect to the second one are given by the second row of the CIR matrix.

Then in the absence of noise constituted by all the legitimate combinations of the three bits of the three transmit AEs, the legitimate channel output states $\check{\mathbf{y}}(k) = [\check{y}_1(k), \check{y}_2(k)]^T$ assume

$$N_s = 2^M = 8 \tag{2.10}$$

number of different values, which are associated with the N_s number of legitimate transmitted symbol sequences. All the $N_s = 8$ legitimate transmitted symbol sequences are listed in Table 2.2 together with the associated noise-free channel output states.

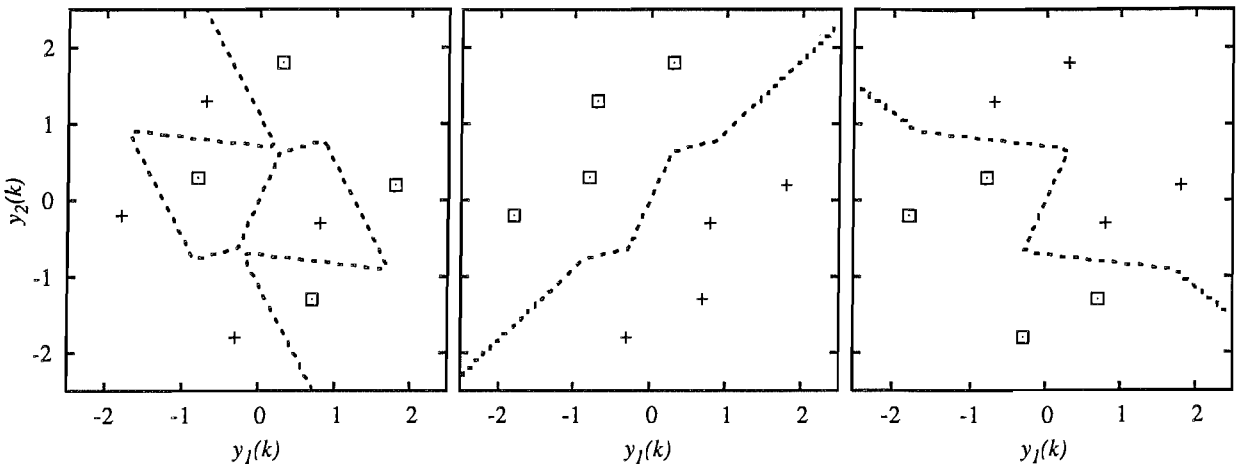
In Figure 2.3 the channel output $y_1(k)$ of array element one is plotted versus the output $y_2(k)$ of array element two. Assuming that the detector aims for detecting the signal transmitted by AE m, it has to separate the two subsets of channel output states defined as

$$\begin{aligned} \mathcal{Y}_m^+ &= \{[\check{y}_1(k), \check{y}_2(k)]^T \in \mathcal{Y} | \check{x}_m(k) = +1\} \\ \mathcal{Y}_m^- &= \{[\check{y}_1(k), \check{y}_2(k)]^T \in \mathcal{Y} | \check{x}_m(k) = -1\} \end{aligned} \tag{2.11}$$

depending on the binary value of the desired transmitter's hypothetically transmitted signal $\check{x}_m(k)$. Again, the detector which performs the processing in the spatial domain has to detect the transmitted signals by separating the two binary subsets of $\mathcal{Y}_m^+$ and $\mathcal{Y}_m^-$. The two subsets indicated by the legends $+$ and $\boxdot$ shown in Figure 2.3a, which are associated with $m = 1$ cannot be separated by a straight-line or linear decision boundary, which might be realized using a linear filter. By contrast, the so-called Bayesian decision boundary indicated by the dotted line is capable of separating the two binary subsets, hence correctly detecting the transmitted signal, provided that the noise power is sufficiently low. For the detection problem posed by Figure 2.3a and assuming equiprobable transmitted bits a linear detector would encounter an error probability of

$$P_e = \frac{\text{No. falsely classified states}}{\text{Total No. states}} = \frac{2}{8} = 0.25. \tag{2.12}$$

The separation problem associated with the detection of the $m = 2^{\text{nd}}$ and the $m = 3^{\text{rd}}$ transmit AE illustrated in Figure 2.3b and Figure 2.3c, respectively, is linearly separable but the optimal Bayesian decision boundary is again non-linear. A range of linear detectors realizing the decision boundaries considered can be found in [108]. Non-linear detection has for example been



(a) Transmit AE $m = 1$ (b) Transmit AE $m = 2$ (c) Transmit AE $m = 3$

Figure 2.3: Decision-space encountered by the detector for a narrowband MIMO system characterized by $M = 3$ transmit AEs, $N = 2$ receive AEs and the channel matrix defined in Equation (2.9).

Table 2.2: Legitimate transmitted symbol sequences and the corresponding channel output states in the absence of noise for a narrowband MIMO system characterized by $M = 3$ transmit AEs, $N = 2$ receive AEs and the channel matrix defined in Equation (2.9).

$\check{x}_1$	$\check{x}_2$	$\check{x}_3$	$\check{y}_1(k)$	$\check{y}_2(k)$
+1	+1	+1	+0.80	−0.35
+1	+1	−1	−0.30	−1.75
+1	−1	+1	−0.70	+1.25
+1	−1	−1	−1.80	−0.15
−1	+1	+1	+1.80	0.15
−1	+1	−1	+0.70	−1.25
−1	−1	+1	+0.30	−1.75
−1	−1	−1	−0.80	0.35

discussed in [71].

All the algorithms, which are presented in this chapter constitute manifestations of different linear and non-linear detectors, all of which aim for separating the two binary subsets according to a given cost-function. Their exact behaviour will be analyzed using the same system model as considered in Example 2.1 and in addition by considering Rayleigh fading channels.

2.2 Benchmarking - Channel Capacity and Information Rates

The capacity of narrowband single-user MIMO channels was quantified in [3]. When comparing different antenna-array configurations, the total transmitted power of a MS is kept constant, regardless of the number M of transmit antennas, as defined in Equation (2.2).

In this section we consider the specific case of a MS having no knowledge of the CIR, which therefore uniformly distributes the transmit power over all transmit AEs. The information rate $I(\boldsymbol{x},\boldsymbol{y})$ supported by a channel is obtained through the mutual information between a transmitted signal vector $\boldsymbol{x}$ and the corresponding received signal vector $\boldsymbol{y}$, for a given CIR realization $\boldsymbol{H}(k)$ under the assumption of having a noise vector $\boldsymbol{\eta}$, which is uncorrelated both with the CIR as well as with the transmitted symbols [109] . It is defined as

$$\begin{aligned} I(\boldsymbol{x},\boldsymbol{y}) &= H(\boldsymbol{y}) - H(\boldsymbol{y}|\boldsymbol{x}) \\ &= H(\boldsymbol{y}) - H(\boldsymbol{\eta}), \end{aligned} \tag{2.13}$$

where H represents the entropy of a random variable defined as $H(\cdot) = -\log_2(p(\cdot))$, with $p(\cdot)$ denoting the Probability Density Function (PDF).

Given the Equation (2.13), the capacity of a complex-valued non-dispersive MIMO channel contaminated by AWGN is given as [3]

$$C(k) = \mathrm{E}\left[\max I(\boldsymbol{x},\boldsymbol{y})\right], \tag{2.14}$$

2.2.1 Gaussian Input Signal

Although in practice the $N_{\text{bit}} = \log_2 \mathcal{M}$-bit signal to be transmitted has $\mathcal{M}$ discrete values, it is computationally appealing to quantify the MIMO channel's capacity under idealized conditions. More specifically, it can be shown [4–6] that if the values assumed by the input signal x_m are Gaussian distributed, the mutual information $I(\boldsymbol{x},\boldsymbol{y})$ quantified in Equation (2.13) is maximized. The associated channel capacity for a given MIMO CIR realization $\boldsymbol{H}(k)$ can then be written as [110]

$$C(k) = \mathrm{E}\left[\log_2\left(\det\left(\boldsymbol{I}_N + \mathrm{SNR}\,\boldsymbol{H}(k)\boldsymbol{H}(k)^H\right)\right)\right], \tag{2.15}$$

where the expectation value is taken with respect to different realizations of the noise, which is taken into account through the SNR defined in Equation (2.2). The channel capacity C of a Rayleigh fading channel, rather than that associated with a single time-invariant CIR realization may be obtained on the basis of Equation (2.15) by taking the expectation value with respect to the CIR matrix $\boldsymbol{H}(k)$, yielding

$$C = \mathrm{E}\left[C(k)\right], \tag{2.16}$$

which may be evaluated using Monte Carlo simulation.

2.2.2 Discrete Input Signal

In contrast to the previous section, if the input signal of the channel is a discrete valued signal, which can assume $\mathcal{M}$ different signal values depending on the specific modulation scheme used, no closed form solution exists for the frequency flat MIMO channel [111].

In order to obtain the capacity of a narrowband MIMO channel for non-Gaussian input signals, we briefly return to the definition of the transmission rate supported by the channel in Equation (2.13). The entropy $H(\eta)$ of the noise required in Equation (2.13) can readily be calculated as [110]

$$\begin{aligned} H(\eta) &= \log_2\left(\det\left(\pi e 2\sigma_n^2 \boldsymbol{I}_N\right)\right) \\ &= \log_2\left(N\pi e 2\sigma_n^2\right). \end{aligned} \tag{2.17}$$

The entropy of the received signal vector can be calculated as [110]

$$H(\boldsymbol{y}) = -\mathrm{E}\left[\log_2\left(\frac{1}{\mathcal{M}^M(2\pi\sigma_n^2)^N}\sum_{\check{\boldsymbol{x}}^{(i)}\in\mathcal{X}}\exp\left(-\frac{1}{2\sigma_n^2}||\boldsymbol{y}-\boldsymbol{H}(k)\check{\boldsymbol{x}}^{(i)}||^2\right)\right)\right], \tag{2.18}$$

where $\mathcal{X}$ is the set of all hypothetically transmitted symbol vectors $\check{\boldsymbol{x}}^{(i)}$ and $1 \leq i \leq \mathcal{M}^M$. The expectation value in Equation (2.18) is taken with respect to different MIMO CIR realizations and to the noise. For a moderate number of transmit AEs Equation (2.18) can be evaluated using Monte-Carlo simulations. For higher-order modulation schemes and for a high number of transmit AEs, approximate formulae might be applied [112].

For the remainder of this treatise we stipulate the following assumptions. The **'channel capacity'** C quantifies the maximally achievable throughput of the MIMO channel assuming a Gaussian distributed input signal. By contrast, the term **'achievable information rate'** or **'achievable throughput'** I is used when we refer to the theoretically possible maximal throughput, which can be achieved for a given MIMO channel under the constraint of the specific $\mathcal{M}$-array modulation scheme used.

2.2.3 Results

Let us first assume encountering an ergodic Rayleigh-distributed channel, where the time average over the channel is equal to the corresponding ensemble average. In other words, the random fluctuation of the channel gain can be averaged out over time. Furthermore, the signal transmitted by the different transmit AEs and received by the different AEs is assumed to be independently faded. For this specific scenario error-free transmission is theoretically possible, if the overall transmission rate R of all transmitters, also referred to as the throughput is lower than or equal to the channel capacity C or the achievable throughput

I, respectively.

Figure 2.4 shows both the capacity and the achievable information rate for a system using M number of transmit AEs and having a CIR matrix having independently faded Rayleigh components. The receiver was assumed to employ N number of receive AEs.

When comparing the channel capacity of Equations (2.15) and (2.16) for different array configurations in Figure 2.4a, it can be observed how the channel capacity increases with the number of receive as well as transmit AEs.

The achievable information rate for BPSK modulated signals, when considering the same system configuration is illustrated in Figure 2.4b. It can be seen that for $M = 2$ and $M = 4$ the achievable information rate approaches $I_{\text{BPSK}}(M = 2) = 2\ \frac{\text{bps}}{\text{Hz}}$ and $I_{\text{BPSK}}(M = 4) = 4\ \frac{\text{bps}}{\text{Hz}}$, respectively, when assuming perfect Nyquist filtration associated with an excess bandwidth of zero which corresponds to the maximum possible Shannonian signaling rate of 2 Baud/Hz

The achievable information rate for 4QAM modulated signals considering the same system configuration is illustrated in Figure 2.4c, where it can be observed that for $M = 2$ and $M = 4$ the achievable information rate approaches $I_{\text{BPSK}}(M = 2) = 4\ \frac{\text{bps}}{\text{Hz}}$ and $I_{\text{BPSK}}(M = 4) = 8\ \frac{\text{bps}}{\text{Hz}}$, respectively.

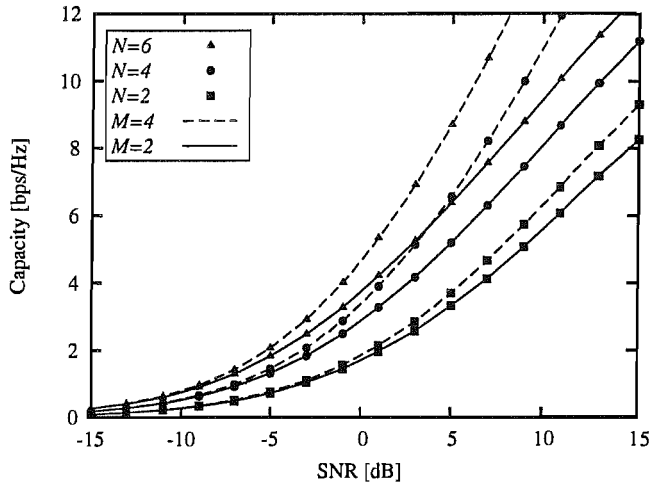
Example 2.2: Channel Capacity

From the graphs plotted in Figure 2.4 the channel capacity might be obtained as follows. For example, a $(N \times M) = (4 \times 2)$-dimensional MIMO system transmitting 4QAM signals and using a half-rate channel code $R_c = 1/2$ results in a transmission rate $R = R_c \cdot \log_2(M) \cdot M$=2. Then the channel capacity bound evaluated from Equation (2.15) becomes $\text{SNR}_{C(4,2)} = -2.5$ dB and the SNR bound associated with the information rate of Equation (2.13) is given as $\text{SNR}_{I_{4\text{QAM}}(4,2)} = -2.1$ dB. According to the findings of Shannon [110], error-free transmission for the considered system becomes possible for $\text{SNR} \geq \text{SNR}_{I_{4\text{QAM}}(4,2)} = -2.1$ dB.

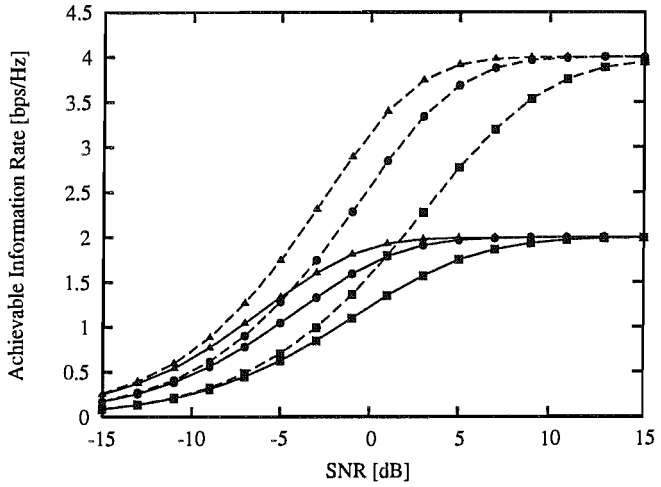
The ergodic channel capacity is only suitable as a theoretical performance bound. In realistic propagation environments the ergodicity of the channel is often not satisfied. Therefore a performance measure called the outage capacity, which is more practically motivated, has been introduced [3]. The general definition of the outage probability for a single-user system is given as

$$P_{\text{out}} = P(C < R) = \int_{-\infty}^{R} p_C\, \partial C, \tag{2.19}$$

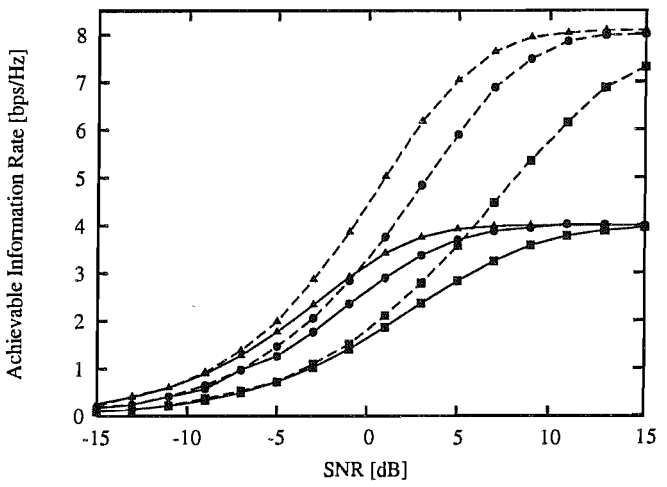
where p_C is the PDF of the channel capacity C. P_{out} is the probability with which the capacity of the channel C is smaller than a given target transmission rate R. In these cases when outage occurs, error-free transmission is not possible. Assuming the channel is constant over one transmitted codeword, the outage



(a) Gaussian distributed transmitted signals.

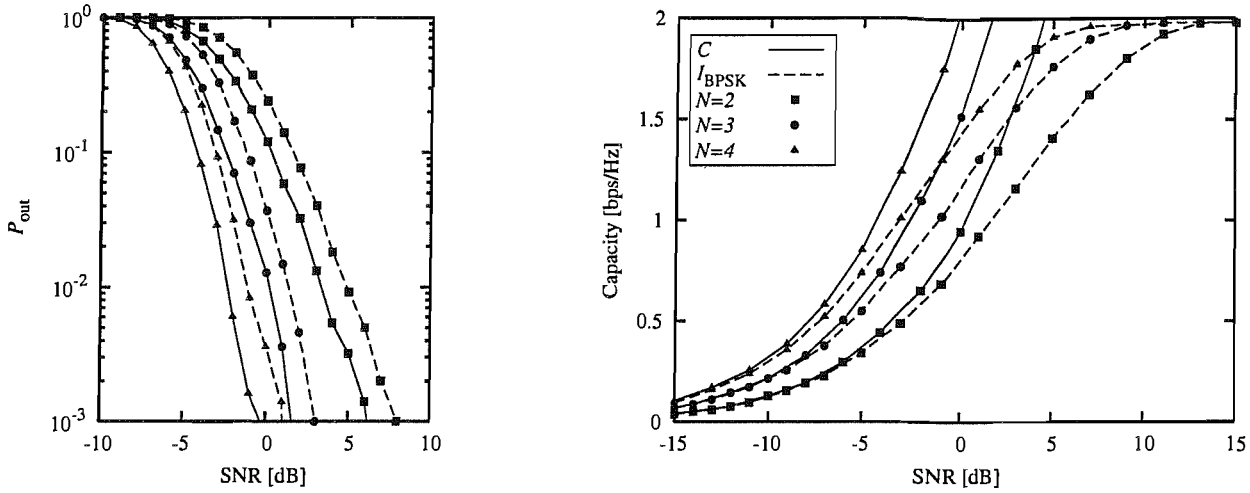


(b) BPSK modulated signals.



(c) 4QAM modulated signals.

Figure 2.4: Channel capacity and achievable information rates for different MIMO system configurations evaluated from Equations (2.15) and (2.13), respectively.



(a) Outage probability for a transmission rate $R = 1$ bps/Hz .

(b) Outage capacity associated with an outage probability $P_{out} = 0.1$.

Figure 2.5: Outage probability and outage capacity evaluated as outlined in Section 2.2.3, for a system using $M = 2$ transmit AEs and a variable number N of receive AEs.

probability is a lower bound for the codeword error rate.

Let us now focus our attention on evaluating the associated outage probability, which was defined in Equation (2.19). The outage probability recorded at a given SNR is obtained by invoking a sufficiently high number of channel realizations $\boldsymbol{H}(k)$ and by determining the channel capacity or the achievable information rate for each of these realizations according to Equation (2.15) and Equation (2.13), respectively. The outage probability is then given as

$$P_{\text{out}} = \frac{\text{No. channel realizations obeying } C < R}{\text{Total No. channel realizations}}. \tag{2.20}$$

The outage probability is always defined for a given target transmission throughput R.

Example 2.3: Outage Capacity

If we consider for example a system supporting $M = 4$ transmit AEs, 4QAM signals and a channel code having code-rate $R_c = 1/2$, the target throughput used when evaluating the outage probability is given as $R = 4 \cdot \log_2(4) \cdot 1/2 = 4$ bps/Hz. The outage probability recorded for this scenario is defined as the probability that the capacity of the channel is lower than 4 bps/Hz. Assume for example that the channel capacity has been recorded for 10 000 CIR realizations at a given SNR and 300 out of the 10 000 CIR realizations have a channel capacity below 4 bps/Hz. Then the outage probability is given as $P_{\text{out}} = 300/10\,000 = 3\%$.

In Figure 2.5 both the outage probability and the outage capacity are plotted for a system employing $M = 2$ transmit AEs and a variable number N of receive AEs. The outage probability recorded for a transmission rate of $R = 1$ bps/Hz is shown in Figure 2.5a, where it can be observed that as expected the outage probabil-

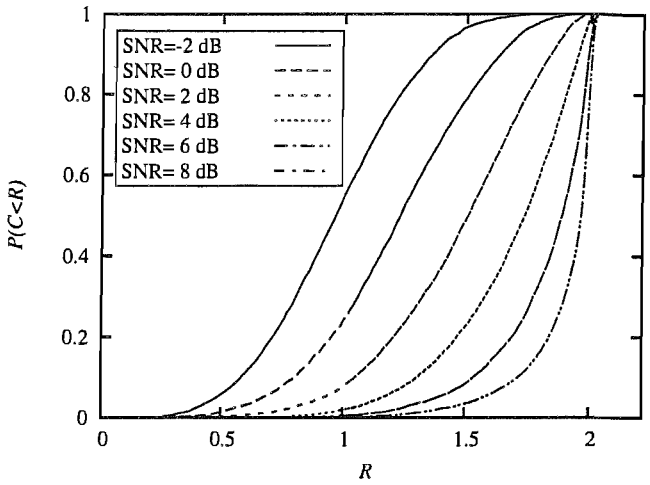


Figure 2.6: CDF for BPSK modulated input signals, for $M = 2$ transmit AEs and $N = 2$ receive AEs at different SNR.

ity decreases upon increasing the number of receive AEs. For a higher number of receive AEs we observed that the increased spatial diversity of the channel was capable of reducing the probability of outages. Furthermore, observe in Figure 2.5a that there is a significant gap between the outage probability recorded for the Gaussian and the BPSK modulated inputs, as represented by the continuous and dashed lines.

In order to characterize the outage probability, the transmission rate was fixed and the probability of an outage event was plotted as a function of SNR. In certain circumstances, the metric of interest might not be the outage probability but rather the outage capacity. The outage capacity is always defined for a given target outage probability P_{out}. If we consider the same scenario as before but we now fix the outage probability for example to $P_{\text{out}} = 3\%$, the associated outage capacity is now defined as the throughput supported by the channel, while ensuring the target outage probability is not exceeded.

In order to arrive at the outage capacity curves, shown in Figure 2.5b we evaluate the capacity C and the information rate I according to Equations (2.15) and (2.13), respectively, for a sufficiently high number of channel realizations. We then gradually increase the transmission rate, until the corresponding outage probability becomes $P_{\text{out}} = 0.1$. The specific transmission rate resulting in a given outage probability P_{out} is denoted as the outage capacity.

When comparing the outage results of Figure 2.5 to the ergodic capacity plotted in Figure 2.4 for the same setting, it can be observed that the SNR gain achieved by having a higher number of receive AEs is higher when the transmitted signal experiences a non-ergodic channel where outage events occur.

The third statistically pertinent way of illustrating the outage capacity is in the form of its Cumulative Distribution Function (CDF) shown in Figure 2.6 for different SNRs. Note that Figure 2.5a, Figure 2.5b and Figure 2.6 all represent the same channel and user scenario but for each figure a different variable was fixed, namely the transmission rate R, the outage probability P_{out} and the SNR, respectively.

In the remainder of this Chapter, we aim for introducing different MIMO detection algorithms with the ultimate goal of approaching the channel capacity limits introduced.

2.3 Multiple-Input Multiple-Output Detection Algorithms

In this section different MIMO detectors, namely the MMSE, the MBER, the Bayesian and a reduced complexity ML MIMO detector will be presented. All the algorithms presented in this section are derived on the basis of both the complex-valued and the real-valued binary system model described in Section 2.1.1 and their characteristic behaviour is analyzed using the simple Example 2.1. At the end of this section simulation results are presented for the different algorithms both for block-fading channels, where the channel's envelope is kept constant for a specific number of transmitted bits, as well as for uncorrelated fading channels. The Frame Error Rate (FER) results obtained for the block-fading channel are benchmarked against the outage capacity results of Figure 2.5, while the BER results generated for the uncorrelated fading channel are benchmarked against the capacity limits.

2.3.1 Minimum Mean Squared Error Detection

The design of angularly selective linear filters based on the MMSE criterion has its origin in the design of classic frequency-selective filters, for example channel equalizers. In the context of MIMO detection, the maximization of the SNR or the Signal to Interference plus Noise Ratio (SINR) has been carried out by minimizing the Mean Squared Error (MSE) between the received and the transmitted reference sequence at the output of the linear MIMO detection filter. Owing to its convenient optimization characteristics, i.e due to having a single minimum, the MMSE design approach has been popular in digital receiver design [28]. In the subsequent discussions, the time index k is dropped for notational convenience.

The output of a linear MIMO detector can be written as [113]

$$\hat{\boldsymbol{x}} = \boldsymbol{W}^H \boldsymbol{y}, \tag{2.21}$$

where $\boldsymbol{W}$ is the $(N \times M)$-dimensional complex-valued MIMO receiver's weight matrix. The different columns of $\boldsymbol{W}$ denoted as $\boldsymbol{w}_m$ are associated with the different transmitters' symbols and $\hat{\boldsymbol{x}}$ is the $(M \times 1)$-dimensional filter output vector generated for the sake of detecting the associated transmitted signal vector $\boldsymbol{x}$. The corresponding real-valued filter output is given as

$$\underline{\hat{\boldsymbol{x}}} = \underline{\boldsymbol{W}}^T \underline{\boldsymbol{y}}, \tag{2.22}$$

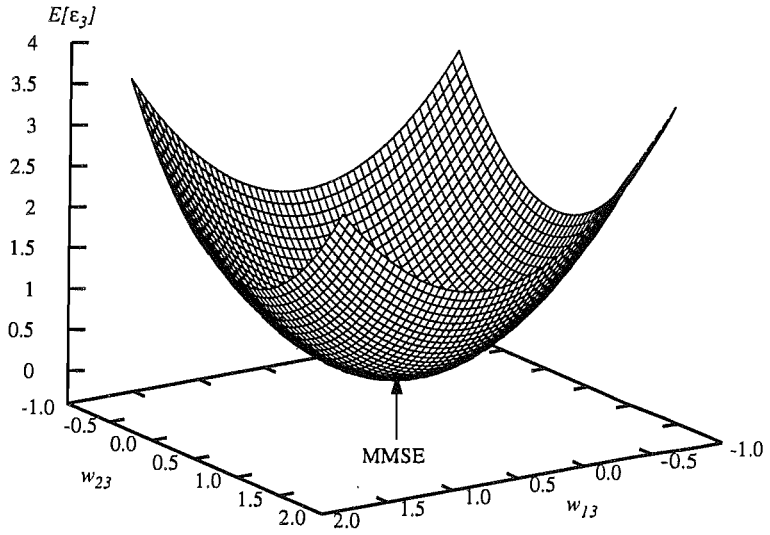


Figure 2.7: MSE as a function of the array weights associated with the third transmit AE for the example system of Section 2.1.2 at SNR=17 dB.

where $\underline{\underline{W}}$ is the $(2N \times N_{\text{bit}}M)$-dimensional real-valued weight matrix. Again, the different columns of $\underline{\underline{W}}$ denoted as $\underline{w}_m$ are associated with the different transmitters' binary symbols, $\underline{\hat{x}}$ is the $(N_{\text{bit}}M \times 1)$ dimensional filter output vector generated for the detection of the associated binary transmitted signal vector $\underline{x}$.

The squared error at the filter output can then be written as

$$\begin{aligned} \epsilon^2 &= ||\boldsymbol{x} - \boldsymbol{W}^H\boldsymbol{y}||^2 = ||\boldsymbol{x} - \hat{\boldsymbol{x}}||^2 && (2.23) \\ &= ||\underline{x} - \underline{\underline{W}}^H\underline{y}||^2 = ||\underline{x} - \underline{\hat{x}}||^2. && (2.24) \end{aligned}$$

The MMSE optimization problem is now defined as

$$\boldsymbol{W} = \arg \min_{\boldsymbol{W} \in \mathbb{C}^{N \times M}} \mathrm{E}\left[||\boldsymbol{x} - \boldsymbol{W}^H\boldsymbol{y}||^2\right] \quad (2.25)$$

or

$$\underline{\underline{W}} = \arg \min_{\underline{\underline{W}} \in \mathbb{R}^{2N \times \log_2(M)M}} \mathrm{E}\left[||\underline{x} - \underline{\underline{W}}^T\underline{y}||^2\right], \quad (2.26)$$

respectively.

The MSE can now be expressed as

$$\begin{aligned} \epsilon^2 &= \mathtt{trace}\left((\boldsymbol{x} - \boldsymbol{W}^H\boldsymbol{y})(\boldsymbol{x} - \boldsymbol{W}^H\boldsymbol{y})^H\right) \\ &= \mathtt{trace}\left((\underline{x} - \underline{\underline{W}}^T\underline{y})(\underline{x} - \underline{\underline{W}}^T\underline{y})^T\right). && (2.27) \end{aligned}$$

Figure 2.7 shows the MSE surface for user three of Example 2.1 introduced in Section 2.1.2 at SNR=17 dB

as a function of the MIMO array weight w_{13} and weight w_{23}. We can see the typical quadratic shape of the cost-function, which may be minimized by taking the derivative of Equation (2.27) and setting it to zero. The derivative of Equation (2.27) can be expressed as

$$\nabla\epsilon^2 = \frac{\partial\epsilon^2}{\partial \boldsymbol{W}} = \mathrm{E}\left[-2(\boldsymbol{W}^H\boldsymbol{y} - \boldsymbol{x})\boldsymbol{y}^H\right]. \tag{2.28}$$

Setting Equation (2.28) equals zero yields

$$\begin{aligned} \mathrm{E}\left[-2(\boldsymbol{W}^H\boldsymbol{y} - \boldsymbol{x})\boldsymbol{y}^H\right] &= 0 \\ -\boldsymbol{W}^H\mathrm{E}\left[\boldsymbol{y}\boldsymbol{y}^H\right] + \mathrm{E}\left[\boldsymbol{x}\boldsymbol{y}^H\right] &= 0 \\ \boldsymbol{W} &= \mathrm{E}\left[\boldsymbol{y}\boldsymbol{y}^H\right]^{-1}\mathrm{E}\left[\boldsymbol{y}\boldsymbol{x}^H\right] \\ \boldsymbol{W} &= \boldsymbol{R}_{yy}^{-1}\boldsymbol{R}_{yx}, \end{aligned} \tag{2.29}$$

where the cross-correlation matrix $\boldsymbol{R}_{yx}$ is given as the channel matrix $\boldsymbol{H}$ since we can write

$$\begin{aligned} \boldsymbol{R}_{yx} &= \mathrm{E}\left[\boldsymbol{y}\boldsymbol{x}^H\right] \\ &= \mathrm{E}\left[\boldsymbol{H}\boldsymbol{x}\boldsymbol{x}^H\right] + \mathrm{E}\left[\boldsymbol{\eta}\boldsymbol{x}^H\right] \\ &= \mathrm{E}\left[\boldsymbol{H}\right]\mathrm{E}\left[\boldsymbol{x}\boldsymbol{x}^H\right] + 0 \\ &= \boldsymbol{H} \end{aligned} \tag{2.30}$$

if the symbols are uncorrelated from the noise as well as the channel matrix and $\mathrm{E}\left[\boldsymbol{x}\boldsymbol{x}^H\right] = \boldsymbol{I}_M$. The correlation matrix $\boldsymbol{R}_{yy}$ can be written as

$$\begin{aligned} \boldsymbol{R}_{yy} &= \mathrm{E}\left[\boldsymbol{y}\boldsymbol{y}^H\right] \\ &= \mathrm{E}\left[(\boldsymbol{H}\boldsymbol{x} + \boldsymbol{\eta})(\boldsymbol{H}\boldsymbol{x} + \boldsymbol{\eta})^H\right] \\ &= \boldsymbol{H}\boldsymbol{H}^H + 2\sigma_{\mathrm{n}}^2\boldsymbol{I}_N, \end{aligned} \tag{2.31}$$

where it again was assumed that $\mathrm{E}\left[\boldsymbol{x}\boldsymbol{x}^H\right] = \boldsymbol{I}_M$.

The corresponding weight vector for the corresponding real-valued binary system is given as

$$\underline{\boldsymbol{W}} = \left(\underline{\boldsymbol{H}}\underline{\boldsymbol{H}}^T + \sigma_{\mathrm{n}}^2\boldsymbol{I}_{2N}\right)^{-1}\underline{\boldsymbol{R}}_{yx}. \tag{2.32}$$

Note, that the noise variance considered in the context of the real-valued binary system is only σ_{n}^2 rather than $2\sigma_{\mathrm{n}}^2$ as considered in the context of the complex-valued MMSE detector.

For BPSK modulated transmitters, the detected symbol vector is given as[2]

$$\tilde{\boldsymbol{x}} = \mathtt{sign}\left(\Re\left(\boldsymbol{W}^H\boldsymbol{y}\right)\right) = \underline{\tilde{\boldsymbol{x}}} = \mathtt{sign}\left(\underline{\boldsymbol{W}}^T\underline{\boldsymbol{y}}\right) \tag{2.33}$$

and for 4QAM modulated sources the corresponding detected symbol is given by

$$\tilde{\boldsymbol{x}} = \frac{1}{\sqrt{2}}\mathtt{sign}\left(\Re\left(\boldsymbol{W}^H\boldsymbol{y}\right)\right) + \frac{\mathrm{j}}{\sqrt{2}}\mathtt{sign}\left(\Im\left(\boldsymbol{W}^H\boldsymbol{y}\right)\right), \tag{2.34}$$

which is identical to

$$\underline{\tilde{\boldsymbol{x}}} = \mathtt{sign}\left(\underline{\boldsymbol{W}}^T\underline{\boldsymbol{y}}\right). \tag{2.35}$$

The filter weights given in Equation (2.29) and in Equation (2.32) are optimal in the MMSE sense and this solution is often also referred to as the Wiener filter [28]. Under the previously stipulated assumptions no other linear filter is capable of achieving a higher SNR at the receiver's output. The implementation of the Wiener solution is in practice relatively simple, since the MSE is a quadratic function associated with a single minimum, which corresponds to the MMSE solution. For the adaptive realization of the algorithm numerous examples can be found in the literature [28]. The most common ones are the Least Mean Square (LMS) and the Recursive Least Squares (RLS) algorithm, which update the MIMO array's weights according to a certain step-size and forgetting factor, respectively.

For BPSK and 4QAM signals, the BER at the output of the MMSE detector may be calculated directly as a function of the MSE, if the conditional PDF of the filter output can be assumed to be Gaussian distributed. If this approximation holds, minimizing the MSE error is equivalent to minimizing the BER and thus optimality is achieved in terms of both MSE and BER.

2.3.1.1 Computational Complexity of the Real-Valued MMSE Detector

For the evaluation of the computational complexity associated with the MMSE detector, the real-valued system model is considered. This will facilitate its fair comparison to other algorithms, since in the case of real-valued operations the complexity imposed by a multiplication and an addition might be considered equivalent.

The complexity evaluation of any linear detector may be divided into two parts, namely the calculation of the MIMO array's weight vector and the detection of the symbol itself. Upon introducing the variables $A = N_{\text{bit}}M$ as well as $B = 2N$, and assuming that the weight vector is calculated using direct matrix

[2]Note, that the optimal MMSE receiver designed for BPSK modulated signals would only minimize the MSE between the transmitted signal and the *real* part of the received signal. However, there exists no closed form solution for this problem and we will therefore employ the complex-value MMSE detector also for BPSK modulated signals. The performance of the real-valued and complex-valued solutions was studied for example in [114].

Table 2.3: The normalized computational complexity (real-valued additions and multiplications) of the MMSE detector when considering the detection of a single transmitted bit.

Weight vector calculation of Equation (2.29)	$16N^2 - \frac{2N(2N+MN_{\text{bit}}-1)}{MN_{\text{bit}}} + \frac{1}{MN_{\text{bit}}} O\left(N^3\right)$
Detection according to Equation (2.35)	$4N-1$

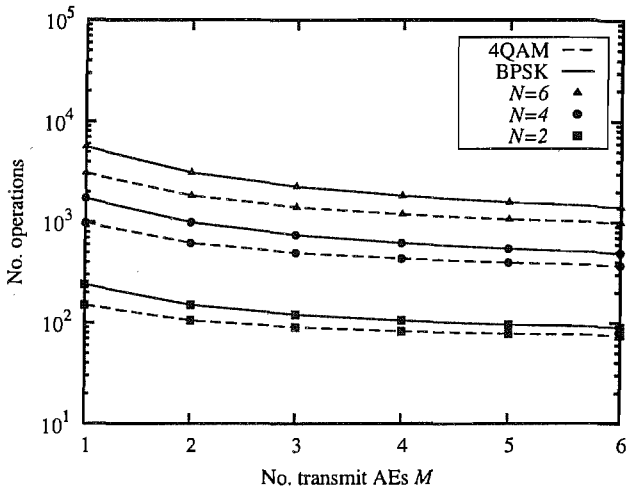


Figure 2.8: The normalized computational complexity required for evaluating the MMSE array weights as a function of the number of transmit AEs for different modulation schemes and a different number of MIMO array weights, when considering the detection of a single transmitted bit calculated according to Table 2.3. The proportionality factor of the matrix inversion was chosen to be 3.

inversion rather than an iterative algorithm and assuming that the channel matrix is known to the receiver, the associated complexity is given as follows:

- The calculation of the $(B \times A)$-dimensional correlation matrix of Equation (2.29) invokes $(2A-1)B^2$ additions and multiplications for the channel correlation matrix plus B further additions are required to take into account the AWGN.
- The matrix inversion of Equation (2.32) might be carried out at a computational cost, which is proportional to $O(B^3)$.
- The matrix multiplication required for the calculation of the weight matrix in Equation (2.32) can be performed at a cost of $(2B-1)BA$ number of operations.

The total number of operations constituted by the number of multiplications plus the number of additions can be obtained by substituting A as well as B. The complexity imposed by the detection of a transmitted bit involves B multiplications and $(B-1)$ additions. The total complexity is summarized in Table 2.3.

In Figure 2.8 the normalized computational complexity defined as the number of real-valued additions plus multiplications required for the evaluation of the MMSE weight vector plotted from Table 2.3 as a function of the number of transmit AEs M for BPSK modulated as well as 4QAM signals, when considering $N = 2, 4$ and 6 receive AEs.

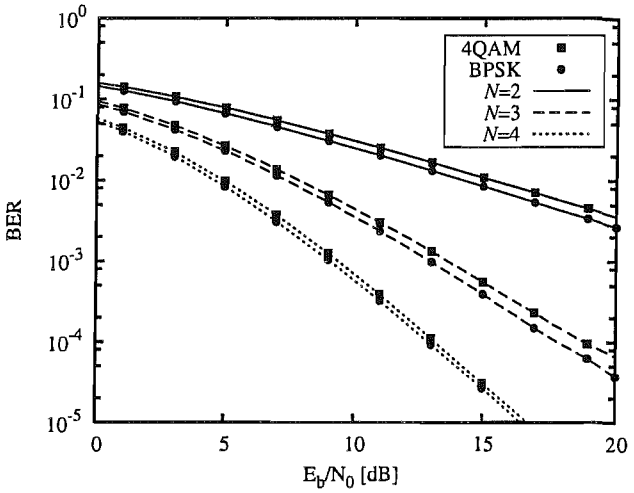


Figure 2.9: BER versus E_b/N_0 performance for a MIMO system employing $M = 2$ transmit AEs and $N = 2, 3$ as well as 4 receive AEs. BPSK and 4QAM modulated signals as well as having perfect channel knowledge were considered. The channel was assumed to be uncorrelated, independent Rayleigh fading with $\mathrm{E}\left[|h_{nm}|^2\right] = 1$.

2.3.1.2 Performance of the MMSE Detector

The MIMO system considered employs $M = 2$ transmit AEs and different number of receive AEs N. The channel was assumed to be an independent, uncorrelated Rayleigh fading channel between all antenna links with $\mathrm{E}\left[|h_{nm}|^2\right] = 1$. Perfect channel knowledge is assumed by the receiver. Figure 2.9 shows the BER versus E_b/N_0 performance for the uncoded system. It can be observed that as the number of receive AEs is increased, the increased diversity gain results in an improved BER performance. Furthermore, it can be observed, that the BER versus E_b/N_0 performance of the BPSK modulated and the 4QAM signal is very similar.

2.3.2 Minimum Bit Error Rate Detection

As it was pointed out in Section 2.3.1, the MMSE receiver design is optimum in terms of the achievable BER only if the Gaussian assumption on the detector's output holds. This however is not necessarily guaranteed in a communication system, as it was shown in [31] in the context of channel equalization and in [33] for conventional narrowband beamforming. In practice the so-called near-far effect results in a non-Gaussian distribution and this observation motivated the employment of other optimization criteria, such as the direct optimization of the BER.

More explicitly, it is a more judicious option to minimize directly the BER rather than minimizing the MSE. Minimizing the BER directly will often result in a lower BER than the MMSE approach and will become equivalent to the MMSE approach, if the filter's output is Gaussian distributed. The MBER solution may be

defined as

$$\hat{\boldsymbol{x}} = \boldsymbol{W}^H \boldsymbol{y}, \tag{2.36}$$

where

$$\boldsymbol{W} = \arg \min_{\boldsymbol{W} \in \mathbb{C}^{N \times M}} P_e(\boldsymbol{W}) \tag{2.37}$$

or equivalently

$$\underline{\hat{\boldsymbol{x}}} = \underline{\boldsymbol{W}}^T \underline{\boldsymbol{y}}, \tag{2.38}$$

where

$$\underline{\boldsymbol{W}} = \arg \min_{\underline{\boldsymbol{W}} \in \mathbb{R}^{2N \times N_{\text{bit}} M}} P_e(\underline{\boldsymbol{W}}). \tag{2.39}$$

The first step in the derivation of the filter weights, which minimize the Bit Error Probability (BEP) is to find an expression for the BEP at the output of a linear receiver, which is typically achieved by integrating the conditional PDF of the filter output over its dynamic range. Assuming a MIMO system, which employs M jointly detected transmit AEs, the transmitted symbol vector may assume $N_s = \mathcal{M}^M$ different values, where $\mathcal{M}$ is the number of modulated phasor points [115]. This N_s number of possible transmitted symbol vectors $\check{\boldsymbol{x}}$ form the set

$$\mathcal{X} = \{\check{\boldsymbol{x}}^{(i)}, 1 \leq i \leq N_s\}. \tag{2.40}$$

The set of possible channel output vectors associated with the set $\mathcal{X}$ may be defined as

$$\mathcal{Y} = \{\check{\boldsymbol{y}}^{(i)} = \boldsymbol{H}\check{\boldsymbol{x}}^{(i)}, 1 \leq i \leq N_s\}. \tag{2.41}$$

2.3.2.1 MBER Detection of BPSK Modulated Signals

As already illustrated in the example of Section 2.1.2, this set can be partitioned into two subsets depending on the value of $\check{x}_m$, yielding

$$\mathcal{Y}_m^{\pm} = \{\check{\boldsymbol{y}}^{(i)} \in \mathcal{Y} \mid \check{x}_m^{(i)} = \pm 1\}. \tag{2.42}$$

In a similar manner, the noise-free part of the detector's output signal is limited to values of the set

$$\hat{\mathcal{X}}_m = \{\check{\hat{x}}_m^{(i)} = \boldsymbol{w}_m^H \check{\boldsymbol{y}}^{(i)}, 1 \leq i \leq N_s\}, \tag{2.43}$$

where $\boldsymbol{w}_m$ is the m^{th} column of the MIMO array's weight matrix $\boldsymbol{W}$. Again, this set can be partitioned into two subsets defined as

$$\hat{\mathcal{X}}_m^{\pm} = \{\check{\hat{x}}_m^{(i)} \in \hat{\mathcal{X}}_m \mid \check{x}_m^{(i)} = \pm 1\}. \tag{2.44}$$

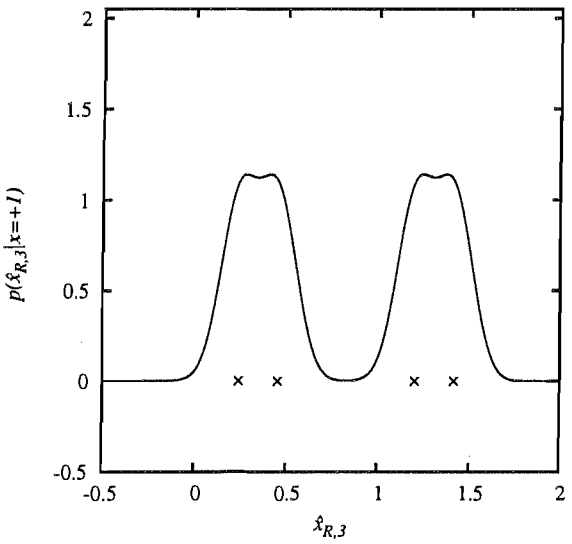


Figure 2.10: Conditional PDF of the MMSE detector's filter output evaluated from Equation (2.45) for the example used in Section 2.1.2 associated with transmit AE $m = 3$ at SNR=17 dB.

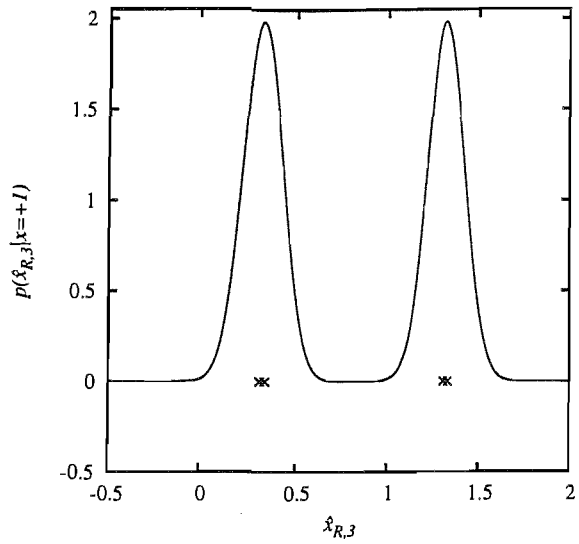


Figure 2.11: Conditional PDF of the MBER detector's filter output evaluated from Equation (2.45) for the example used in Section 2.1.2 associated with transmit AE $m = 3$ at SNR=17 dB.

Let us now define $\hat{x}_{\mathrm{R},m}(k) = \Re\left(\hat{x}_m(k)\right)$ and $\check{x}_{\mathrm{R},m}^{(i)} = \Re\left(\check{x}_m^{(i)}\right)$. Then the conditional PDF of $\hat{x}_{\mathrm{R},m}(k)$ derived for BPSK modulated sources can be expressed as the superposition of $N_{\mathrm{sb}} = \frac{N_{\mathrm{s}}}{2}$ Gaussian PDFs positioned at the legitimate noiseless output states and averaged over the N_{sb} equiprobable states:

$$p(\hat{x}_{\mathrm{R},m}|x_m = +1) = \frac{1}{N_{\mathrm{sb}}\sqrt{2\pi\sigma_{\mathrm{n}}^2 \boldsymbol{w}_m^H \boldsymbol{w}_m}} \sum_{\check{x}_m^{(i)} \in \check{\mathcal{X}}_m^+} \exp\left(-\frac{(\hat{x}_{\mathrm{R},m} - \check{x}_{\mathrm{R},m}^{(i)})^2}{2\sigma_{\mathrm{n}}^2 \boldsymbol{w}_m^H \boldsymbol{w}_m}\right), \tag{2.45}$$

where $N_{\mathrm{sb}} = \frac{N_{\mathrm{s}}}{2}$ is the number of legitimate noise-less filter output states in the set $\check{\mathcal{X}}_m^+$. The variance of the noise after filtering with the MBER detector's weights is given by $2\sigma_{\mathrm{n}}^2 \boldsymbol{w}_m^H \boldsymbol{w}_m$ since the noise is filtered by the same detector's weights as the received signal.

In Figure 2.11 the conditional PDF of the channel output associated with the $m = 3^{\mathrm{rd}}$ transmit AE is plotted for the scenario already considered in Example 2.1 of Section 2.1.2 at SNR=17 dB. The PDFs associated with both the MMSE and the MBER detectors' weights were plotted using Equation (2.45). It can be seen that the filter output is not Gaussian distributed regardless whether the MMSE solution or the MBER solution is considered. It can also be observed that the minimum distance between the possible filter output states indicated as small crosses and the decision boundary found at $\hat{x}_{\mathrm{R},3} = 0$ is slightly larger for the MBER solution. This will result in a lower BER associated with the MBER-aided MIMO detector.

The BER can now be obtained by integrating the conditional PDF $p(\hat{x}_{\mathrm{R},m}|x_m = +1)$ over the negative

half-plane, namely over the range, where erroneous decision are carried out, as seen below:

$$P_e(\boldsymbol{w}_m) = \int_{-\infty}^{0} p(\hat{x}_{R,m}|\check{x}_m = +1)\,\partial\hat{x}_{R,m} \tag{2.46}$$

$$= \sum_{\check{\boldsymbol{x}}_m^{(i)} \in \hat{\mathcal{X}}_m^+} \int_{-\infty}^{0} \frac{1}{N_{sb}\sqrt{2\pi\sigma_n^2 \boldsymbol{w}_m^H \boldsymbol{w}_m}} \exp\left(-\frac{(\hat{x}_{R,m} - \check{x}_{R,m}^{(i)})^2}{2\sigma_n^2 \boldsymbol{w}_m^H \boldsymbol{w}_m}\right) \partial\hat{x}_{R,m}, \tag{2.47}$$

yielding

$$P_e(\boldsymbol{w}_m) = \frac{1}{N_{sb}} \sum_{\check{\boldsymbol{x}}_m^{(i)} \in \hat{\mathcal{X}}_m^+} \mathcal{Q}\left(g_i(\boldsymbol{w}_m)\right), \tag{2.48}$$

where we have

$$g_i(\boldsymbol{w}_m) = \frac{\mathtt{sign}\left(\check{x}_{R,m}^{(i)}\right)\check{x}_{R,m}^{(i)}}{\sigma_n\sqrt{\boldsymbol{w}_m^H \boldsymbol{w}_m}} \tag{2.49}$$

and the Gaussian Q-function given as [28]

$$\mathcal{Q}(z) = \int_z^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-y^2}{2}\right) \partial y. \tag{2.50}$$

Equation (2.48) may be employed to calculate the BER performance of any linear detector characterized by the weight vector $\boldsymbol{w}_m$, enabling us to avoid a Monte Carlo simulation for Gaussian channels.

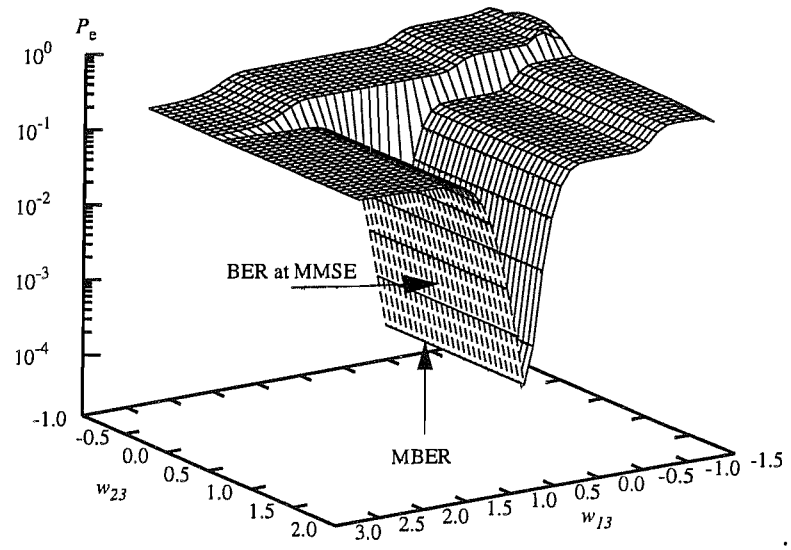
In Figure 2.12 the BER surface is plotted as a function of the weight coefficients for the system already considered in Example 2.1 of Section 2.1.2 at SNR=10 dB. It can be seen in Figure 2.12a, which shows the BER surface associated with the $m = 3^{rd}$ transmit AE, that in contrast to the MSE surface of Figure 2.7 plotted for the same scenario, the MBER surface has a very distinct minimum, which is restricted to a small array weight region. Hence, a slight variation of the weights will cause large BER variations, when operating close to the MBER solution. The different location of the MMSE and the MBER solution is indicated by the arrows. The corresponding BERs are distinctly different. The specific shape of the BER surface makes it difficult to find the optimum solution to Equation (2.37) and Equation (2.39), respectively. As already mentioned above, the solution to this problem can be found by taking the derivative of Equation (2.48) and setting it to zero. The derivative of Equation (2.48) is given as (Appendix A)

$$\nabla P_e(\boldsymbol{w}_m) = C \sum_{\check{\boldsymbol{x}}_m^{(i)} \in \hat{\mathcal{X}}_m^+} \exp\left(-\frac{\left(\check{x}_{R,m}^{(i)}\right)^2}{2\sigma_n^2 \boldsymbol{w}_m^H \boldsymbol{w}_m}\right) \mathtt{sign}\left(\check{x}_{R,m}^{(i)}\right) \left(\frac{\check{\boldsymbol{x}}_{R,m}^{(i)} \boldsymbol{w}_m}{\boldsymbol{w}_m^H \boldsymbol{w}_m} - \check{y}_m^{(i)}\right), \tag{2.51}$$

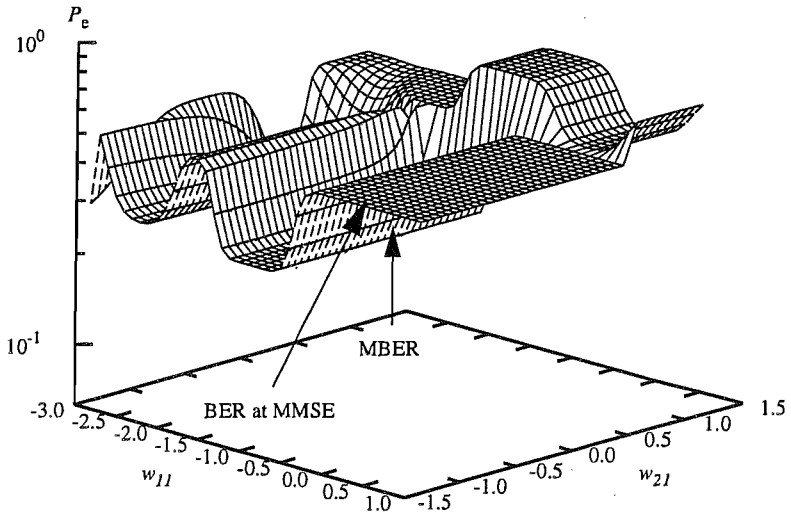
where we have

$$C = \frac{1}{2N_{sb}\sqrt{2\pi\sigma_n^2 \boldsymbol{w}_m^H \boldsymbol{w}_m}}. \tag{2.52}$$

The weight values where we have $P_e(\boldsymbol{w}_m) = 0$ might be found by using a simplified conjugate gradient algorithm, as given in Appendix B.



(a) BER surface for $m = 3$



(b) BER surface for $m = 1$

Figure 2.12: BER as a function of the array weights for the (2×3)-dimensional MIMO example used in Section 2.1.2 at SNR=10 dB.

In Figure 2.12b the BER surface is shown as a function of the weights associated with the $m = 1^{st}$, transmit AE, which corresponds to the linearly non-separable detection problem outlined in Figure 2.3a. It can be seen that the MBER surface now becomes irregular, exhibiting numerous local extrema even for this low-dimensional detection problem. Finding the minimum of this surface is not a trivial task and hence the search algorithm has to be carefully initialized. Employing a conjugate gradient algorithm has in some cases shown to be suboptimal, since it does not always converge to the MBER solution, it may even converge to a weight vector associated with a BER performance worse than that of the MMSE algorithm. In scenarios where convergence of the algorithm is difficult to achieve, it is more advisable to consider the employment of the steepest decent algorithm, which is more likely to find the minimum of the BER surface at the cost of an increased number of iterations, i.e. increased complexity. In order to circumvent convergence problems other optimization algorithms [94, 95] such as for example Genetic Algorithms (GAs) [116] have also been proposed, which are more robust in the presence of local minima occurring in the BER cost-function.

2.3.2.2 MBER Detection of 4QAM Signals

Before commencing with the derivation of the 4QAM weights, we reformulate the output of a linear filter given by Equation (2.21) under the absence of AWGN as

$$\hat{\boldsymbol{x}} = \boldsymbol{W}^H \boldsymbol{H} \boldsymbol{x} \tag{2.53}$$

$$= \boldsymbol{C} \boldsymbol{x}. \tag{2.54}$$

We refer to the $(i,j)^{th}$ element of the combined impulse response matrix $\boldsymbol{C}$ as c_{ij}. The output of the MBER detector associated with the desired transmit antenna m can now be expressed by

$$\hat{x}_m = c_{mm} x_m + \sum_{i \neq m} c_{mi} x_i \,, \tag{2.55}$$

where the first term is the desired signal and the sum represents the residual interference term. Under the assumption that c_{mm} is real-valued and positive, the optimal decision rule for 4QAM is given by

$$\tilde{x}_m = \frac{1}{\sqrt{2}} \mathtt{sign}\left(\Re\left(\hat{x}_m\right)\right) + \frac{\mathrm{j}}{\sqrt{2}} \mathtt{sign}\left(\Im\left(\hat{x}_m\right)\right). \tag{2.56}$$

It is, however, not guaranteed that c_{mm} is real-valued and positive, and therefore a complex-valued coefficient c_{mm} might rotate the transmitted symbol constellation according to Equation (2.55). In order to remove this phase rotation, the detector has to rotate its weight vector according to

$$\boldsymbol{w}_{\mathrm{rot},m} = \frac{c_{mm}^*}{|c_{mm}|} \boldsymbol{w}_m. \tag{2.57}$$

Without this rotation the MBER algorithm still might produce a perfect 4QAM constellation at the MBER filter's output, but may introduce a phase shift similar to that often imposed by the Constant Modulus Algorithm (CMA) [115, 117]. Note that for the MMSE algorithm this weight rotation is not required since for the MMSE weight vector we have $c_{mm} \approx 1$.

Having introduced the weight rotation of Equation (2.57), the derivation of the 4QAM MBER solution closely follows the derivation of the BPSK MBER weights. Depending on the hypothetically transmitted symbol $\check{x}_m$, the set of possible channel output states $\mathcal{Y}$ for a given weight vector $\boldsymbol{w}_m$, introduced in Equation (2.41) can be partitioned into four subsets given as

$$\mathcal{Y}_m^{\pm,\pm} = \{\check{\boldsymbol{y}}^{(i)} \in \mathcal{Y} \mid \check{x}_m = \pm 1 \pm \mathrm{j}\}. \tag{2.58}$$

In the same way, the set of noise-free detector outputs $\hat{\mathcal{X}}$ generated for a given weight vector $\boldsymbol{w}_m$, which was defined in Equation (2.43) can be partitioned into four subsets according to

$$\hat{\mathcal{X}}_m^{\pm,\pm} = \{\check{x}_m^{(i)} \in \hat{\mathcal{X}}_m \mid \check{x}_m = \pm 1 \pm \mathrm{j}\}. \tag{2.59}$$

The conditional PDF of $\hat{x}_m$ derived for 4QAM sources can now be expressed as

$$p(\hat{x}_m | x_m = +1+\mathrm{j}) = \frac{1}{N_{\mathrm{sb}}\sqrt{2\pi\sigma_{\mathrm{n}}^2 \boldsymbol{w}_m^H \boldsymbol{w}_m}} \sum_{\check{x}_m^{(i)} \in \hat{\mathcal{X}}_m^{+,+}} \exp\left(-\frac{||\hat{x}_m - \check{x}_m^{(i)}||^2}{2\sigma_{\mathrm{n}}^2 \boldsymbol{w}_m^H \boldsymbol{w}_m}\right), \tag{2.60}$$

where $N_{\mathrm{sb}} = \frac{N_{\mathrm{s}}}{4}$ is the number of legitimate noiseless filter output states in the set $\mathcal{Y}_m^{+,+}$. Here we have chosen the symbol $x_m = +1+\mathrm{j}$ for the derivation of the BER expression. Choosing any other symbol for our derivation, would yield the same result.

The resultant BER can now be calculated as

$$P_{\mathrm{e}} = \frac{1}{2}\left(P_{\mathrm{e,R}} + P_{\mathrm{e,I}}\right), \tag{2.61}$$

where following the derivation of the BPSK weights, $P_{\mathrm{e,R}}$ and $P_{\mathrm{e,I}}$ can be expressed as

$$P_{\mathrm{e,R}}(\boldsymbol{w}_m) = \frac{1}{N_{\mathrm{sb}}} \sum_{\check{x}_m^{(i)} \in \hat{\mathcal{X}}_m^{+,+}} \mathcal{Q}\left(g_{\mathrm{R},i}(\boldsymbol{w}_m)\right) \tag{2.62}$$

and

$$P_{\mathrm{e,I}}(\boldsymbol{w}_m) = \frac{1}{N_{\mathrm{sb}}} \sum_{\check{x}_m^{(i)} \in \hat{\mathcal{X}}_m^{+,+}} \mathcal{Q}\left(g_{\mathrm{I},i}(\boldsymbol{w}_m)\right), \tag{2.63}$$

where we have

$$g_{\mathrm{R},i}(\boldsymbol{w}_m) = \frac{\mathrm{sign}\left(\Re\left(\check{x}_m^{(i)}\right)\right)\check{x}_{\mathrm{R},m}^{(i)}}{\sigma_{\mathrm{n}}\sqrt{\boldsymbol{w}_m^H\boldsymbol{w}_m}} \tag{2.64}$$

and

$$g_{\mathrm{I},i}(\boldsymbol{w}_m) = \frac{\mathrm{sign}\left(\Im\left(\check{x}_m^{(i)}\right)\right)\check{x}_{\mathrm{I},m}^{(i)}}{\sigma_{\mathrm{n}}\sqrt{\boldsymbol{w}_m^H\boldsymbol{w}_m}}, \tag{2.65}$$

while $\check{x}_{\mathrm{I},m}^{(i)} = \Im\left(\check{x}_m^{(i)}\right)$ and $\check{x}_{\mathrm{R},m}^{(i)} = \Re\left(\check{x}_m^{(i)}\right)$. The gradient of the BER given in Equation (2.61) can now be calculated as

$$\nabla P_{\mathrm{e}}(\boldsymbol{w}_m) = \frac{1}{2}\left(\nabla P_{\mathrm{e,I}}(\boldsymbol{w}_m) + \nabla P_{\mathrm{e,R}}(\boldsymbol{w}_m)\right), \tag{2.66}$$

where we have

$$\nabla P_{\mathrm{e,R}}(\boldsymbol{w}_m) = C \sum_{\check{\boldsymbol{x}}_m^{(i)} \in \check{\mathcal{X}}_m^{+,+}} \exp\left(-\frac{\left(\check{x}_{\mathrm{R},m}^{(i)}\right)^2}{2\sigma_{\mathrm{n}}^2\boldsymbol{w}_m^H\boldsymbol{w}_m}\right) \mathrm{sign}\left(\Re\left(\check{x}_m^{(i)}\right)\right)\left(\frac{\check{x}_{\mathrm{R},m}^{(i)}\boldsymbol{w}_m}{\boldsymbol{w}_m^H\boldsymbol{w}_m} - \check{y}_m^{(i)}\right) \tag{2.67}$$

and

$$\nabla P_{\mathrm{e,I}}(\boldsymbol{w}_m) = C \sum_{\check{\boldsymbol{x}}_m^{(i)} \in \check{\mathcal{X}}_m^{+,+}} \exp\left(-\frac{\left(\check{x}_{\mathrm{I},m}^{(i)}\right)^2}{2\sigma_{\mathrm{n}}^2\boldsymbol{w}_m^H\boldsymbol{w}_m}\right) \mathrm{sign}\left(\Im\left(\check{x}_m^{(i)}\right)\right)\left(\frac{\check{x}_{\mathrm{I},m}^{(i)}\boldsymbol{w}_m}{\boldsymbol{w}_m^H\boldsymbol{w}_m} + \mathrm{j}\check{y}_m^{(i)}\right), \tag{2.68}$$

while

$$C = \frac{1}{2N_{\mathrm{sb}}\sqrt{2\pi\sigma_{\mathrm{n}}^2\boldsymbol{w}_m^H\boldsymbol{w}_m}}. \tag{2.69}$$

The MBER weights, again, can be found using the simplified conjugate gradient algorithm, where the weights have to be rotated after each iteration using Equation (2.57).

2.3.2.3 MBER Detection for the Real-Valued Binary System Model

In contrast to the complex-valued representation of the MBER algorithm found in the literature, a real-valued representation is independent of the modulation scheme and resembles the MBER solution derived for BPSK modulated signals. Let us denote

$$\underline{\check{\mathcal{X}}} = \{\underline{\check{\boldsymbol{x}}}^{(i)},\ 1 \le i \le N_{\mathrm{s}}\}, \tag{2.70}$$

where $\underline{\check{\boldsymbol{x}}}$ is a hypothetically transmitted binary symbol associated with the real-valued binary system described by Equation (2.8). Associated with the set $\check{\mathcal{X}}$, the set of legitimate real-valued channel output vectors may be defined as

$$\underline{\mathcal{Y}} = \{\underline{\check{\boldsymbol{y}}}^{(i)} = \underline{\boldsymbol{H}}\underline{\check{\boldsymbol{x}}}^{(i)},\ 1 \le i \le N_{\mathrm{s}}\}, \tag{2.71}$$

which can be partitioned according to

$$\underline{\mathcal{Y}}_m^{\pm} = \{\underline{\check{y}}^{(i)} \in \underline{\mathcal{Y}} \mid \check{x}_m^{(i)} = \pm 1\}. \tag{2.72}$$

In a similar manner, the noise-free part of the linear MIMO detector's output signal only may assume discrete values of the set

$$\underline{\hat{\mathcal{X}}}_m = \{\check{x}_m^{(i)} = \underline{w}_m^T \underline{\check{y}}^{(i)},\ 1 \leq i \leq N_s\}, \tag{2.73}$$

where $\underline{w}_m$ is the m^{th} column of the weight matrix $\underline{\underline{W}}$. Note, that for the real-valued binary system model the index m ranges from 1 to MN_{bit}. This set again can be partitioned into two subsets defined as

$$\underline{\hat{\mathcal{X}}}_m^{\pm} = \{\check{x}_m^{(i)} \in \underline{\hat{\mathcal{X}}}_m \mid \check{x}_m^{(i)} = \pm 1\}. \tag{2.74}$$

Following the derivation of the complex-valued MBER detector derived for BPSK modulation, the BEP can be shown to be

$$P_e(\underline{w}_m) = \frac{1}{N_{sb}} \sum_{\check{x}_m^{(i)} \in \underline{\hat{\mathcal{X}}}_m^+} Q\left(g_i(\underline{w}_m)\right) \tag{2.75}$$

in conjunction with

$$g_i(\underline{w}_m) = \frac{\texttt{sign}\left(\check{x}_m^{(i)}\right)\check{x}_m^{(i)}}{\sigma_n \sqrt{\underline{w}_m^H \underline{w}_m}}. \tag{2.76}$$

The derivative of Equation (2.75) is given as

$$\nabla P_e(\underline{w}_m) = C \sum_{\check{x}_m^{(i)} \in \underline{\hat{\mathcal{X}}}_m^+} \exp\left(-\frac{\left(\check{x}_m^{(i)}\right)^2}{2\sigma_n^2 \underline{w}_m^T \underline{w}_m}\right) \texttt{sign}\left(\check{x}_m^{(i)}\right)\left(\frac{\check{x}_m^{(i)} \underline{w}_m}{\underline{w}_m^T \underline{w}_m} - \underline{\check{y}}_m^{(i)}\right), \tag{2.77}$$

where we have

$$C = \frac{1}{2N_{sb}\sqrt{2\pi\sigma_n^2 \underline{w}_m^T \underline{w}_m}}. \tag{2.78}$$

Assuming a normalized weight vector, where $\underline{w}_m^T \underline{w}_m = 1$, this further simplifies to

$$\nabla P_e(\underline{w}_m) = C \sum_{\check{x}_m^{(i)} \in \underline{\hat{\mathcal{X}}}_m^+} \exp\left(-\frac{\left(\check{x}_m^{(i)}\right)^2}{2\sigma_n^2}\right)\left(\check{x}_m^{(i)} \underline{w}_m - \underline{\check{y}}_m^{(i)}\right)$$

where

$$C = \frac{1}{2N_{sb}\sqrt{2\pi\sigma_n^2}}. \tag{2.79}$$

If the MBER solution is expressed with the aid of the binary real-valued system representation of this section, the phase rotation of Equation (2.57) required for the complex-valued 4QAM MBER solution becomes

Table 2.4: The normalized computational complexity (real-valued additions and multiplications) of the MBER detector when considering the detection of a single transmitted bit.

Weight vector calculation	$N_{\text{grad}}N_{\text{sb}}(2+16N)+$ $N_{\text{s}}\left(\frac{2N(2MN_{\text{bit}}-1)}{MN_{\text{bit}}}\right)$
Detection	$4N-1$

unnecessary.

2.3.2.4 Computational Complexity of the Real-Valued MBER Detector

The computational cost associated with the MBER algorithm is evaluated on the basis of the real-valued binary system model and under the assumption that the channel matrix is known to the receiver. The associated computational cost can again be divided into two parts similarly to the complexity evaluation of the MMSE algorithm. The computational complexity required for the calculation of the MBER weights is difficult to quantify and depends strongly on the initial weights chosen for the conjugate gradient algorithms as well as the shape of the BER surface.

Let us first quantify the computational complexity imposed by the evaluation of the BEP gradient of Equation (2.79) for a normalized weight vector. The evaluation of the expression in the sum of Equation (2.79) requires $3+3\cdot 2\cdot N = 3+6N$ number of real-valued multiplications and additions as well as one evaluation of the exp function. Each term in the sum of Equation (2.79) has to be evaluated N_{sb} times and summed up, which requires further $N_{\text{sb}}-1$ additions. The normalization of the weight vector of Equation (2.79) contributes $2\cdot N+(2\cdot N-1)+2\cdot N = 6N-1$ real-valued operations. The computational complexity difference between the detection of BPSK and 4QAM signals is included in the factor N_{sb}. Additionally the receiver has to evaluate the legitimate channel output states, which requires $N_{\text{s}}(2N(2MN_{\text{bit}}-1))$ real-valued additions and multiplications. The evaluation of a hypothetical detector output state associated with a given channel output state requires $(4N-1)$ real-valued additions and multiplications.

When considering the complexity of the conjugate gradient algorithm presented in Appendix B, it can be seen that the computational complexity associated with the evaluation of the gradient constitutes the main contribution. We therefore approximate the complexity required for the calculation of the weight vector associated with a single transmitted bit as $N_{\text{grad}}C^{\nabla P_{\text{e}}}$, where N_{grad} is the number of iterations used by the conjugate gradient algorithm and

$$C^{\nabla P_{\text{e}}} \approx N_{\text{sb}}(3+6N+6N-1+4N-1+1) = N_{\text{sb}}(16N+2) \tag{2.80}$$

is the computational complexity required for the calculation of the BEP gradient of Equation (2.79) for a given weight vector. The total normalized complexity required for carrying out the detection of a single bit

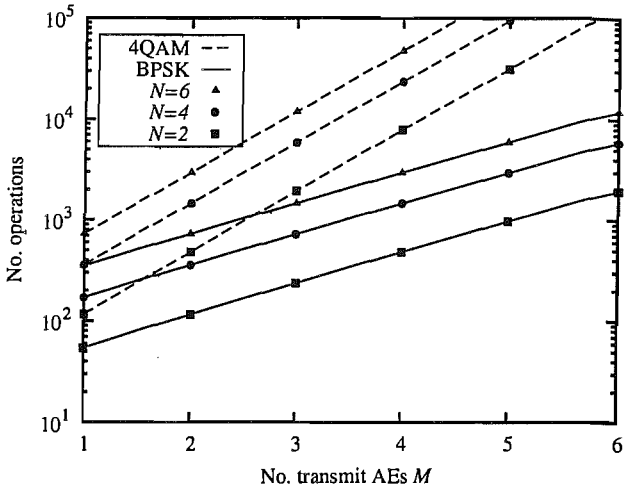


Figure 2.13: The normalized computational complexity required for evaluating the MBER array weights as a function of the number of transmit AEs for different modulation schemes and a different number of MIMO array weights, when considering the detection of a single transmitted bit calculated according to Table 2.4. The number N_{grad} of conjugate gradient algorithm iterations was set to 1.

is summarized in Table 2.4, where the evaluation of the exp function has been neglected.

In Figure 2.13 the normalized computational complexity defined as the number of real-valued additions plus multiplications required for the evaluation of the MBER weight vector - as summarized in Table 2.4 - is plotted as a function of the number of transmit AEs M for BPSK modulated as well as 4QAM signals when considering $N = 2, 4$ and 6 receive AEs. In contrast to the computational complexity imposed by the MMSE detector, which was illustrated in Figure 2.8, it can be observed that the complexity associated with the calculation of the MBER weights increases exponentially with the number of transmit AEs M and the bits per transmitted symbol $N_{\text{bit}} = \log_2(\mathcal{M})$. Note that the number of iterations required by the simplified conjugate gradient algorithm for the evaluation of the MBER weights might be very high and thus further increases the associated complexity. Once the weight vector has been evaluated, the detection of the symbol itself can be achieved at a computational cost, which is identical to that imposed by the MMSE detector, which is linear in M, as it can be seen from Tables 2.4 and 2.3, respectively.

2.3.2.5 Performance of the MBER Detector

The single-user MIMO system considered is identical to the scenario used in Section 2.3.1.2, which is characterized by $M = 2$ transmit AEs at the MS and a BS employing $N = 2, 3$ and 4 receive antennas. The channel was assumed to be Rayleigh faded with $\mathrm{E}\left[|h_{nm}|^2\right] = 1$. Figure 2.14 shows the BER versus E_b/N_0 performance of the uncoded system. For BPSK modulated signals, the MBER detector clearly outperforms the MMSE detector of Figure 2.9. This is mainly due-to the fact that the MBER detector only optimizes the real-part of the filter output, whereas the MMSE algorithm minimizes the MSE composed of both the real and the imaginary part of the filter output. It was shown in [114] that the real-valued

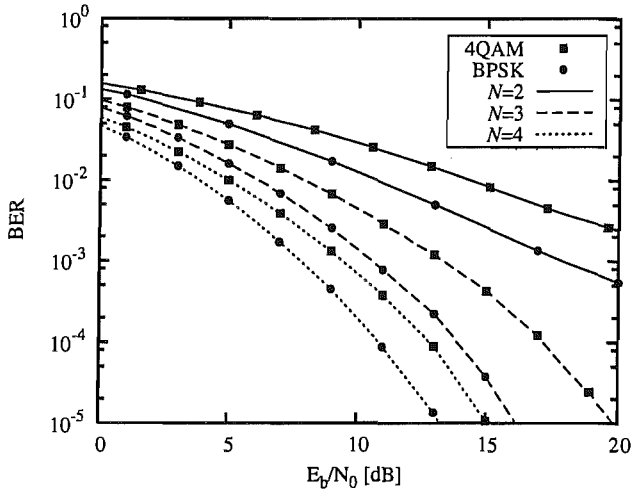


Figure 2.14: BER versus E_b/N_0 performance for a MIMO system employing $M = 2$ transmit AEs and $N = 2, 3$ as well as 4 receive AEs. BPSK and 4QAM modulated signals as well as having perfect channel knowledge were considered.

MMSE detector performs similarly to the MBER detector in the context of BPSK modulated signals. If 4QAM signals are considered, the MMSE and the MBER algorithm perform similarly for $N = 2$ receive antennas, since for a low number of receive AEs the linearly non-separable channel-output constellation points degrade the BER. If the number of receive AEs is increased to $N = 3$ and $N = 4$, respectively, the MBER detector outperforms the MMSE algorithm in the high-SNR region. This performance difference indicates that the MMSE algorithm is incapable of optimally separating two linearly separable data sets owing to its MSE-based optimization function, whereas the MBER algorithm achieves this goal. We note, that both the MBER and the MMSE algorithm have to perform the same classification task by constructing a linear decision boundary defined by the filter weights. In general the MBER algorithm is more likely to outperform the MMSE algorithm in the high-SNR region, since the PDF is then mainly determined by the non-Gaussian interference, rather than by the Gaussian noise.

2.3.3 Bayesian Detection

Both the presented MMSE and MBER algorithms presented in Section 2.3.1 and Section 2.3.2, respectively, constitute linear receivers. The optimal receiver designed for a given channel model however can only be realized by a non-linear structure designed for implementing the decision boundaries indicated in Section 2.1.2.

For the derivation of the Bayesian detector let us assume that the transmitted signal is BPSK modulated. The probability of making an erroneous decision P_e is then defined as

$$P_e = P(\hat{x}_m = +1|x_m = -1)P(x_m = -1) + P(\hat{x}_m = -1|x_m = +1)P(x_m = +1). \quad (2.81)$$

It was proved for example in [39] that Equation (2.81) is minimized when the detector decides for the hypothesis $\check{x}_m = \pm 1$ if

$$p(\boldsymbol{y}|x_m = \pm 1) > p(\boldsymbol{y}|x_m = \mp 1) \tag{2.82}$$

assuming that all hypothesis are equally likely. This implies that we decide for the hypothesis of $\check{x}_m$, which maximizes the conditional likelihood. We may now write the decision rule as

$$\hat{x}_m = \arg \max_{x_m \in \{+1,-1\}} p(\boldsymbol{y}|x_m). \tag{2.83}$$

In decision theory most of the time the conditional probability density $p(\boldsymbol{y}|x_m = \pm 1)$ is known or can be estimated. It becomes clear when considering the Bayes theorem, which enables us to express the *a posteriori* probability as

$$P(x_m = \pm 1|\boldsymbol{y}) = \frac{p(\boldsymbol{y}|x_m = \pm 1)P(x_m = \pm 1)}{p(\boldsymbol{y})} \tag{2.84}$$

that maximizing the conditional likelihood, is identical to maximizing the *a posteriori* probability if the hypothesis are equally likely. The Bayes' theorem assists in transforming the conditional probability density function $p(\boldsymbol{y}|x_m = \pm 1)$ in order to obtain the conditional *a posteriori* probability $P(x_m = \pm 1|\boldsymbol{y})$.

The probabilities given in Equation (2.84) can now be directly used to perform a Maximum A Posteriori (MAP) decision based on the soft output of the Bayesian detector. The corresponding decision rule can be written as

$$\hat{x}_m = \begin{cases} +1 & \text{if } P(x_m = +1)p(\boldsymbol{y}|x_m = +1) - P(x_m = -1)p(\boldsymbol{y}|x_m = -1) > 0; \\ -1 & \text{if } P(x_m = +1)p(\boldsymbol{y}|x_m = +1) - P(x_m = -1)p(\boldsymbol{y}|x_m = -1) < 0. \end{cases} \tag{2.85}$$

It is noted that the MBER detector introduced in Section 2.3.2 may be viewed as the linear detector achieving the lowest possible miss-classification error, whereas the Bayesian detector characterized by Equation (2.85) realizes the optimum non-linear detector attaining the lowest possible miss-classification error.

2.3.3.1 Bayesian Detection of BPSK Modulated Signals

Equation (2.85) can be directly used to design the Bayesian detector for BPSK modulated signals. All that needs to be done to construct a Bayesian detector, is to find an expression for the two conditional probabilities $p(\boldsymbol{y}|x_m = +1)$ and $p(\boldsymbol{y}|x_m = -1)$. This can be done in analogy to the derivation of the MBER detector presented in Section 2.3.2. Again, the partitioning of the legitimate noise-free channel output states into the two subsets $\mathcal{Y}^{\pm}$ is considered, which has been defined in Equation (2.42). Let us now

rewrite the decision function of Equation (2.85) as

$$\hat{x}_m = \mathtt{sign}\left(f_{\mathrm{B},m}(\boldsymbol{y})\right) = \begin{cases} +1 & \text{if } f_{\mathrm{B},m}(\boldsymbol{y}) \geq 0 \\ -1 & \text{if } f_{\mathrm{B},m}(\boldsymbol{y}) < 0 \end{cases}, \tag{2.86}$$

where the optimal Bayesian decision function [118] $f_{\mathrm{B},m}(\cdot)$ based on the difference of the associated conditional PDFs is given as

$$\begin{aligned} f_{\mathrm{B},m}(\boldsymbol{y}) &= p(\boldsymbol{y}|x_m = +1) - p(\boldsymbol{y}|x_m = -1) \\ &= \sum_{\check{\boldsymbol{x}}_m^{(i,+)} \in \mathcal{X}_m^+} p\left(\boldsymbol{y}|\check{\boldsymbol{x}}_m^{(i,+)}\right) - \sum_{\check{\boldsymbol{x}}_m^{(i,-)} \in \mathcal{X}_m^-} p\left(\boldsymbol{y}|\check{\boldsymbol{x}}_m^{(i,-)}\right) \\ &= \sum_{\check{\boldsymbol{y}}_m^{(i,+)} \in \mathcal{Y}_m^+} \frac{1}{(2\pi\sigma_{\mathrm{n}}^2)^N} \exp\left(-\frac{||\boldsymbol{y} - \check{\boldsymbol{y}}_m^{(i,+)}||^2}{2\sigma_{\mathrm{n}}^2}\right) \\ &\quad - \sum_{\check{\boldsymbol{y}}_m^{(i,-)} \in \mathcal{Y}_m^-} \frac{1}{(2\pi\sigma_{\mathrm{n}}^2)^N} \exp\left(-\frac{||\boldsymbol{y} - \check{\boldsymbol{y}}_m^{(i,-)}||^2}{2\sigma_{\mathrm{n}}^2}\right), \end{aligned} \tag{2.87}$$

where we have $\check{\boldsymbol{y}}_m^{(i,\pm)} \in \mathcal{Y}_m^{\pm}$. It is shown in Figure 2.15 how the Bayesian detector formulates an optimum non-linear decision boundary. The surface illustrates the values of the Bayesian decision function as a function of the channel-output for the specific channel setup used in Example 2.1 of Section 2.1.2 when considering the detection of the signal associated with the $m = 3^{\mathrm{rd}}$ transmit AE. Figure 2.15 corresponds to Figure 2.3c of Example 2.1 of Section 2.1.2 of this chapter. The non-linear decision boundary is placed exactly where the decision function of Equation (2.87) is equal to zero, which is illustrated at the bottom of the graph as a dotted line.

2.3.3.2 Bayesian Detection for 4QAM Signals

In the context of 4QAM a symbol x_m is associated with the two binary bits $\underline{x}_{2m-1}$ and $\underline{x}_{2m}$. The Bayesian detector makes a decision for each of this sets separately according to

$$\hat{\underline{x}}_{2m-1+i} = \begin{cases} +1 & \text{if } f_{\mathrm{B},m}^{(i)}(\boldsymbol{y}) \geq 0 \\ -1 & \text{if } f_{\mathrm{B},m}^{(i)}(\boldsymbol{y}) < 0 \end{cases} \quad \text{for } i = 0, 1\,, \tag{2.88}$$

where we have

$$\begin{aligned} f_{\mathrm{B},m}^{(i)}(\boldsymbol{y}) &= p(\boldsymbol{y} \mid \underline{x}_{2m-1+i} = +1) + p(\boldsymbol{y} \mid \underline{x}_{2m-1+i} = -1) \\ &= \sum_{\check{\boldsymbol{x}}_m^{(i,+)} \in \mathcal{X}_m^{+,i}} p\left(\boldsymbol{y}|\check{\boldsymbol{x}}_m^{(i,+)}\right) - \sum_{\check{\boldsymbol{x}}_m^{(i,-)} \in \mathcal{X}_m^{-,i}} p\left(\boldsymbol{y}|\check{\boldsymbol{x}}_m^{(i,-)}\right) \end{aligned} \tag{2.89}$$

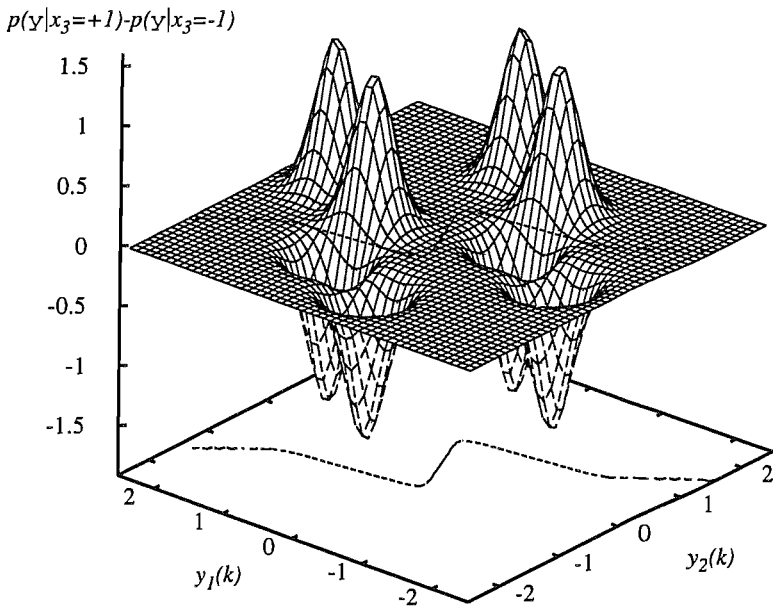


Figure 2.15: Difference of the conditional probabilities as a function of the channel output associated with the $m = 3^{\text{rd}}$ transmit AE for the example used in Section 2.1.2 at SNR=10 dB.

with

$$\begin{aligned} \mathcal{X}_m^{\pm,0} &= \mathcal{X}_m^{\pm,+} \bigcup \mathcal{X}_m^{\pm,-} \\ \mathcal{X}_m^{\pm,1} &= \mathcal{X}_m^{+,\pm} \bigcup \mathcal{X}_m^{-,\pm}, \end{aligned}$$

where $\mathcal{X}_m^{\pm,\pm}$ have been defined in the context of Equation (2.58).

2.3.3.3 Bayesian Detection for the Real-Valued Binary System Model

In analogy to the MBER detector also the Bayesian detector can readily be derived for the real-valued system model presented in Equation (2.7). Given the set of legitimate real-valued channel output states $\underline{\mathcal{Y}}_m^{\pm}$ defined in Equation (2.72), then the Bayesian decision function $f_{\text{B},m}(\cdot)$ may be expressed as

$$\begin{aligned} f_{\text{B},\underline{m}}(\underline{y}) &= p(\underline{y} \mid \underline{x}_m = +1) - p(\underline{y} \mid \underline{x}_m = -1) \\ &= \sum_{\underline{\check{x}}_m^{(i,+)} \in \underline{\mathcal{X}}_m^+} p\left(\underline{y} | \underline{\check{x}}_m^{(i,+)}\right) - \sum_{\underline{\check{x}}_m^{(i,-)} \in \underline{\mathcal{X}}_m^-} p\left(\underline{y} | \underline{\check{x}}_m^{(i,-)}\right) \\ &= \sum_{\underline{\check{y}}_m^{(i,+)} \in \underline{\mathcal{Y}}_m^+} \frac{1}{(2\pi\sigma_{\text{n}}^2)^N} \exp\left(-\frac{||\underline{y} - \underline{\check{y}}_m^{(i,+)}||^2}{2\sigma_{\text{n}}^2}\right) \\ &\quad - \sum_{\underline{\check{y}}_m^{(i,-)} \in \underline{\mathcal{Y}}_m^-} \frac{1}{(2\pi\sigma_{\text{n}}^2)^N} \exp\left(-\frac{||\underline{y} - \underline{\check{y}}_m^{(i,-)}||^2}{2\sigma_{\text{n}}^2}\right). \end{aligned} \tag{2.90}$$

Table 2.5: The normalized computational complexity (real-valued additions and multiplications) of the Bayesian detector when considering the detection of a single transmitted bit.

Calculation of the legitimate channel output states	$\frac{N_s(2N(2MN_{bit}-1))}{MN_{bit}}$
Detection of a single bit	$N_s 6N - 1$ plus N_s exp function evaluations

Again, the notation used for the real-valued binary interpretation of the Bayesian detector is independent of the modulation scheme and only real-valued operations have to be performed at the receiver. Note, that the factor $\frac{1}{(2\pi\sigma_n^2)^N}$ does not change compared to that of the complex-valued Bayesian detector even so the noise is now real-valued or one-dimensional. The exponent N of the factor remains unchanged, since the square root factor which would be introduced due to the 'real-valued' noise is canceled out by the fact that the channel-output states have $2N$ real-valued dimensions.

2.3.3.4 Interpretation of the Bayesian Detector as Radial-Basis-Function Network

The Bayesian detector can be realized using a so-called Radial Basis Function Network (RBFN) [119, 120] positioning a Gaussian kernel function of the form seen in Equation (2.92) at the legitimate channel output states. The response of such a RBFN, if BPSK modulation is considered, is given as

$$f_{\text{RBF}}(\boldsymbol{y}) = \sum_{i=1}^{N_c} w_i \phi(\boldsymbol{y}, \mathbf{c}^{(i)}) \tag{2.91}$$

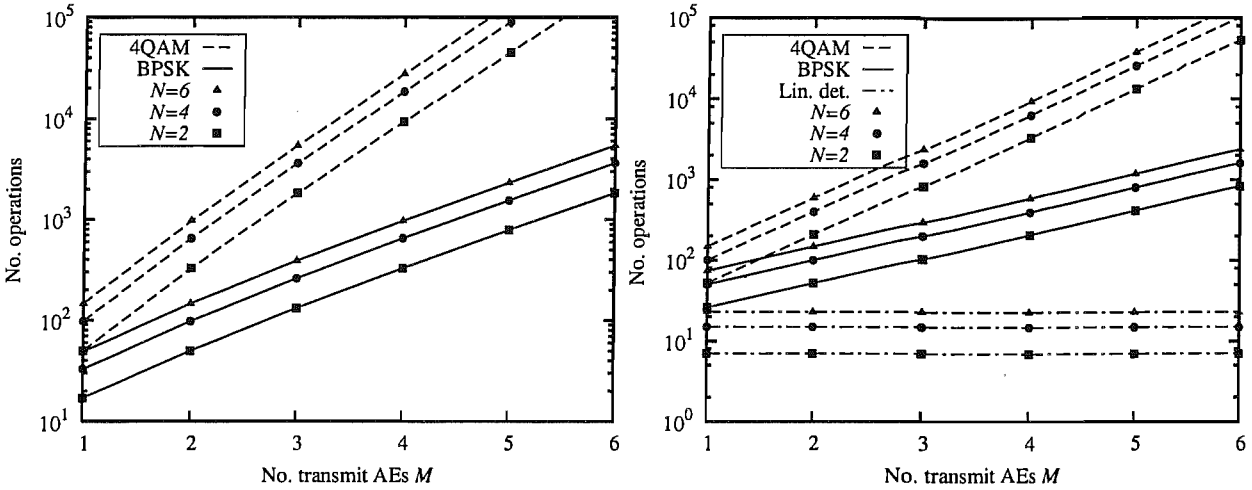
with the Gaussian kernel function given by

$$\phi(\boldsymbol{y}(k), \mathbf{c}^{(i)}) = \exp\left(-\frac{||\boldsymbol{y} - \mathbf{c}^{(i)}||^2}{\rho}\right), \tag{2.92}$$

where the RBF centers $\mathbf{c}^{(i)}$ are set to the legitimate noise-free channel output states determined by the CIR, the radius ρ is chosen to be $2\sigma_n^2$, the RBF weights w_i are set to $+1$, if we have $\mathbf{c}^{(i)} \in \mathcal{Y}_m^+$ and to -1 if $\mathbf{c}^{(i)} \in \mathcal{Y}_m^-$, and the number of RBFN centers N_c is set to the number of noise-free channel output states N_s.

2.3.3.5 Computational Complexity of the Real-Valued Bayesian Detector

Similar to both the linear MMSE and the MBER detectors discussed in Section 2.3.1 and Section 2.3.2, respectively, the computational complexity imposed by the Bayesian MIMO detector can be separated into two parts, namely the calculation of the legitimate channel output states and the detection of the transmitted bits.



(a) Number of real-valued additions plus multiplications required for the calculation of the legitimate channel output states plus the the real-valued operations required for the detection of a single bit.

(b) Number of real-valued additions plus multiplications required for the detection of a single bit once the legitimate channel output states have been determined. The complexity considered for the linear detectors only includes the detection and not the computation of the weight vector according to Table 2.3.

Figure 2.16: The normalized computational complexity imposed by the Bayesian detector for the detection of a single transmitted bit as a function of the number of transmit AEs, evaluated for different modulation schemes and a different number of MIMO array elements. The values were calculated according to Table 2.5.

According to Equation (2.71) the computation of the possible channel output states requires $N_s(2N(2MN_{bit} - 1))$ number of real-valued additions plus multiplications as derived in Section 2.3.2.4. For binary signals, the detection of a single bit requires $N_s 6N - 1$ real-valued additions plus multiplications as well as N_s evaluations of the exp function. The total complexity is summarized in Table 2.5.

In Figure 2.16a the normalized computational complexity defined as the number of real-valued additions plus multiplications required for the calculation of the legitimate noiseless channel output states is illustrated as a function of the number of transmit AEs M for BPSK modulated as well as 4QAM signals, when considering $N = 2, 4$ and 6 receive AEs. Similar to the computational complexity associated with the evaluation of the BER surface gradient illustrated in Figure 2.13, the complexity associated with the evaluation of the legitimate channel output states is exponential in the number of transmit AEs M, as well as in the number of bits per symbol N_{bit}.

However, once the MBER detector has evaluated the associated weight vector, it can carry out the detection at a linearly increasing cost as it can be seen from Table 2.4. By contrast, once the legitimate channel output states have been evaluated the complexity imposed by the Bayesian detector when detecting a single bit increases exponentially with the number of transmit AEs M, as well as with the number of bits per symbol N_{bit}. This exponential increase of computational cost required for the detection of a single bit is illustrated in Figure 2.16 as a function of the number of transmit AEs M for BPSK modulated as well as 4QAM signals, when considering $N = 2, 4$ and 6 receive AEs. The computational cost associated with the linear

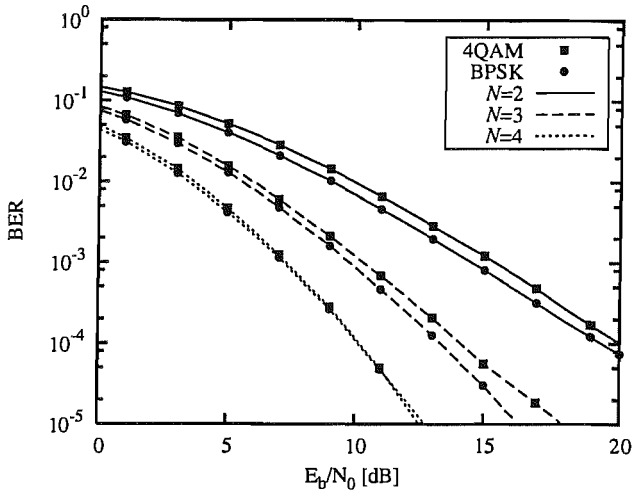


Figure 2.17: BER versus E_b/N_0 performance for a MIMO system employing $M = 2$ transmit AEs and $N = 2, 3$ as well as 4 receive AEs. BPSK and 4QAM modulated signals as well as having perfect channel knowledge were considered.

MMSE and MBER detector denoted by the label 'Lin. det.' are shown as a reference.

2.3.3.6 Performance of the Bayesian Detector

The system considered is identical to the system used in Section 2.3.1.2, which is characterized by $M = 2$ transmit AEs at the MS and a BS employing $N = 2$ receive antennas while considering independent uncorrelated Rayleigh fading channels with $\mathrm{E}\left[|h_{nm}|^2\right] = 1$. It can be observed in Figure 2.17 that the receiver exhibits a significantly better performance, than the two linear detectors, namely the MMSE and the MBER detector. This performance gain however is achieved at a significantly increased computational cost imposed by the Bayesian detection algorithm. In the next subsection we therefore introduce a low-complexity algorithm, which approaches the performance of the Bayesian detector.

2.3.4 Reduced Complexity Maximum Likelihood Detection

Let us define the ML solution for the transmitted symbol vector, which is given as

$$\hat{\boldsymbol{x}} = \arg\max_{\check{\boldsymbol{x}} \in \mathcal{X}} p(\boldsymbol{y}|\check{\boldsymbol{x}}), \tag{2.93}$$

where $\mathcal{X}$ is the set of potentially transmitted symbol vectors $\check{\boldsymbol{x}}$. Equation (2.93) may be re-written as

$$\hat{\boldsymbol{x}} = \arg\min_{\check{\boldsymbol{x}} \in \mathcal{X}} \{||\boldsymbol{y} - \boldsymbol{H}\check{\boldsymbol{x}}||^2\}. \tag{2.94}$$

When comparing the optimization performed by the Bayesian detector given in Equation (2.83) and Equation (2.93), which defines the ML solution for the reduced complexity ML detector it can be observed that

there is a fundamental philosophical difference between the two. While the Bayesian detector attempts to maximize the conditional probability for each symbol separately, the reduced complexity ML detector aims for maximizing the conditional probability of a complete symbol vector.

Let us assume for the derivation of the algorithm that the CIR matrix $\boldsymbol{H}$ as well as the noise are real-valued and that the transmitted signal is Binary Phase Shift Keying (BPSK) modulated. It was shown in [60] that the solution to the problem defined by Equation (2.94) is identical to solving

$$\hat{\boldsymbol{x}} = \arg\min_{\check{\boldsymbol{x}} \in \mathcal{X}} ||\boldsymbol{U}(\check{\boldsymbol{x}} - \hat{\boldsymbol{x}}_{\mathrm{MMSE}})||^2, \tag{2.95}$$

where the upper triangular matrix U is defined by

$$\boldsymbol{U}^H\boldsymbol{U} = \boldsymbol{H}^H\boldsymbol{H} + \sigma_{\mathrm{n}}^2\mathbf{I}_M, \tag{2.96}$$

while

$$\hat{\boldsymbol{x}}_{\mathrm{MMSE}} = (\boldsymbol{H}^H\boldsymbol{H} + \sigma_{\mathrm{n}}^2\boldsymbol{I}_M)^{-1}\boldsymbol{H}^H\boldsymbol{y}. \tag{2.97}$$

The MMSE solution presented in Equation (2.97) appears different from the one introduced in Section 2.3.1 but in fact both yield the same weight vector. The two different MMSE solutions are also referred to as the left- and the right-handed MMSE solution [17].

Let us now first define $N_{\mathrm{b}} = M\log_2(\mathcal{M}) = MN_{\mathrm{bit}}$ as the number of symbols in $\boldsymbol{x}$ considered by the detector. Exploiting the fact that the matrix $\boldsymbol{U}$ has an upper triangular structure, it can be shown that the objective function used for the detection of the transmitted symbol vector $\check{\boldsymbol{x}}$ may be written as [61]

$$\begin{aligned} J(\check{\boldsymbol{x}}) &= ||\boldsymbol{U}(\check{\boldsymbol{x}} - \hat{\boldsymbol{x}}_{\mathrm{MMSE}})||^2 && (2.98)\\ &= (\check{\boldsymbol{x}} - \hat{\boldsymbol{x}}_{\mathrm{MMSE}})^H\mathrm{U}^H\boldsymbol{U}(\check{\boldsymbol{x}} - \hat{\boldsymbol{x}}_{\mathrm{MMSE}}) && (2.99)\\ &= \sum_{i=1}^{N_{\mathrm{b}}} |\sum_{j=i}^{N_{\mathrm{b}}} u_{ij}(\check{x}_j - \hat{x}_{\mathrm{MMSE},j})|^2 && (2.100)\\ &= \sum_{i=1}^{N_{\mathrm{b}}} \phi_i(\check{\boldsymbol{x}}_i), && (2.101) \end{aligned}$$

where u_{ij} is the $(i,j)^{\mathrm{th}}$ element of $\boldsymbol{U}$, $\check{\boldsymbol{x}}_i = [\check{x}_i, \ldots, \check{x}_{N_{\mathrm{b}}}]^T$ and $\phi_i(\check{\boldsymbol{x}}_i)$ may be expressed as [61]

$$\begin{aligned} \phi_i(\check{\boldsymbol{x}}_i) &= |u_{ii}(\check{x}_i - \hat{x}_{\mathrm{MMSE},i}) \\ &\quad + \underbrace{\sum_{j=i+1}^{N_{\mathrm{b}}} u_{ij}(\check{x}_j - \hat{x}_{\mathrm{MMSE},j})}_{a_i}|^2, \end{aligned} \tag{2.102}$$

and where the second term a_i of Equation (2.102) is independent of the specific symbol's value of $\check{x}_i$. The

cost-function given in Equation (2.101) may now be re-written in a recursive manner as

$$J_i(\check{\boldsymbol{x}}_i) = J_{i+1}(\check{\boldsymbol{x}}_{i+1}) + \phi_i(\check{\boldsymbol{x}}_i), \quad i = N_b - 1, \ldots 1, \tag{2.103}$$

where we have $J_{N_b}(\check{\boldsymbol{x}}_{N_b}) = |u_{N_b N_b}(\check{\boldsymbol{x}}_{N_b} - \hat{x}_{\mathrm{MMSE},N_b})|^2$. The cost-function has the essential property that [61]

$$J(\check{\boldsymbol{x}}) = J_1(\check{\boldsymbol{x}}_1) > J_2(\check{\boldsymbol{x}}_2) > \cdots > J_{N_b}(\check{\boldsymbol{x}}_{N_b}) = J_{N_b}(\check{\boldsymbol{x}}_{N_b}) > 0. \tag{2.104}$$

Due to the fact that the cost-function $J_i(\check{\boldsymbol{x}}_i)$ only depends on $\{\check{\boldsymbol{x}}_i, \ldots, \check{\boldsymbol{x}}_{N_b}\}$, we introduce the notation

$$J_i(\check{\boldsymbol{x}}_i) = J_i([\check{\boldsymbol{x}}_i, \ldots, \check{\boldsymbol{x}}_{N_b}]^T) = J_i([\check{\boldsymbol{x}}_i\, \check{\boldsymbol{x}}_{i+1}]^T). \tag{2.105}$$

This property facilitates the creation of a low-complexity search technique, which is outlined in detail in [61] and is also referred to as the Optimized Hierarchical Recursive Search Algorithm (OHRSA).

To ensure that the algorithm operates efficiently, it is advisable to reorder the columns of the CIR matrix first in increasing order according to the norm of the columns. This will result in a 'best-first' detection strategy, as outlined in [17]. The columns of the CIR matrix $\boldsymbol{H}$ are therefore re-arranged such that the reordered CIR matrix $\boldsymbol{H}^{(o)}$ satisfies

$$||(\boldsymbol{H}^{(o)})_1||^2 \leq ||(\boldsymbol{H}^{(o)})_2||^2 \leq \cdots \leq ||(\boldsymbol{H}^{(o)})_{N_b}||^2, \tag{2.106}$$

where $||(\boldsymbol{H}^{(o)})_i||^2$ indicates the energy or squared norm of the i^{th} column of $\boldsymbol{H}^{(o)}$.

In the following subsections we will first present the conventional search tree based optimization algorithm, which allows to find the ML solution in a recursive manner. This recursive search algorithm has been proposed in [60]. We then propose an alternative, novel search method referred to as the 'Min-Path' search strategy, which is computationally more efficient.

Example 2.4: Conventional Search-tree Based Optimization

Consider a simple real-valued candidate system employing $M = 3$ transmit AEs and $N = 2$ receive AEs, resulting in a $(N \times M) = (2 \times 3)$-dimensional channel matrix. The N-element channel output $\boldsymbol{y}$ is given as the product of the CIR matrix $\boldsymbol{H}$ and the transmitted symbol vector $\boldsymbol{x}$ plus the AWGN. The characteristic quantities of such a system are given for example as

$$\boldsymbol{y} = \begin{bmatrix} 0.29 \\ -0.55 \end{bmatrix}, \quad \boldsymbol{H} = \begin{bmatrix} 0.50 & 0.75 & 0.55 \\ 0.25 & 0.95 & 0.70 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{x} = \begin{bmatrix} +1 \\ +1 \\ +1 \end{bmatrix}, \tag{2.107}$$

which were recorded at $N_0 = 1.50$.

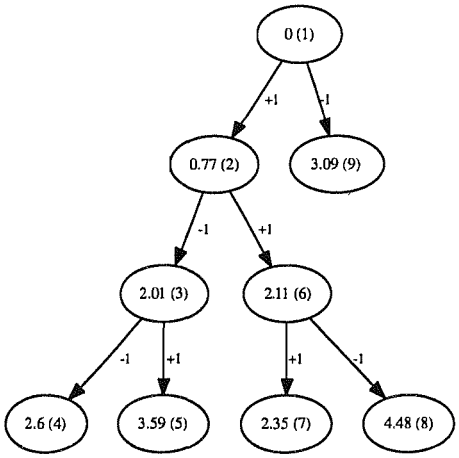


Figure 2.18: Example of a search tree formed by the OHRSA based detector in a scenario employing BPSK modulation, $N = 2$ receive AEs and $M = 2$ transmit AEs at $N_0 = 1.5$. The exact received signal vector $\boldsymbol{y}$, the channel matrix $\boldsymbol{H}$ and the transmitted sequence $\boldsymbol{x}$ are given in Equation (2.107). The values in the ellipses indicate the cost-function values computed from Equation (2.103), while the node index is given in brackets.

As mentioned earlier, the convergence of the algorithm can be improved by reordering the columns of the CIR matrix $\boldsymbol{H}$, so that the columns of the reordered CIR matrix have a monotonously increasing ordered norm [17, 61]. Commencing the algorithm now by reordering the channel matrix of Equation (2.107) yields $\boldsymbol{H}^{(o)}$, where the superscript $^{(o)}$ stands for ordered

$$\boldsymbol{H}^{(o)} = \begin{bmatrix} 0.50 & 0.55 & 0.75 \\ 0.25 & 0.70 & 0.95 \end{bmatrix}. \tag{2.108}$$

It may be readily shown that the corresponding upper triangular decomposition of $\boldsymbol{H}^{(o)}$ into the upper triangular matrix $\mathbf{U}^{(o)}$ and the MMSE solution for the ordered system are given as

$$\mathbf{U}^{(o)} = \begin{bmatrix} 1.01 & -0.39 & -0.15 \\ 0 & 1.14 & -0.16 \\ 0 & 0 & 1.32 \end{bmatrix} \tag{2.109}$$

and

$$\hat{\boldsymbol{x}}^{(o)}_{\text{MMSE}} = [-0.002 - 0.12 + 0.33]^T, \tag{2.110}$$

respectively. Note that since both the system matrix of Equation (2.107) and the transmitted signal are real-valued, the imaginary part of the received sequence may be omitted and the resultant MMSE solution also becomes real-valued.

The algorithm commences at node 0 of Figure 2.18 by evaluating the cost function of the hypothetical solutions $\check{x}^{(o)}_{N_b} = \check{x}^{(o)}_3$ associated with the ordered channel matrix of Equation (2.108),

according to Equation (2.103), which yields

$$J_3([\check{x}_3^{(o)}]) = |u_{33}^{(o)}(\check{x}_3^{(o)} - \hat{x}_{\mathrm{MMSE},3}^{(o)})|^2$$
$$J_3([\check{x}_3^{(o)}] = +1) = |1.32 \cdot (+1 - 0.33)|^2 = 0.77$$
$$J_3([\check{x}_3^{(o)} = -1]) = |1.32 \cdot (-1 - 0.33)|^2 = 3.09.$$

The corresponding two values of $J_3(\check{x}_3^{(o)})$ can be seen at the second hierarchical level of Figure 2.18 as nodes 2 and 9 together with the associated hypothetical BPSK solutions indicated along the branches. Based on the two cost-function values seen within node 2 and 9 we select node 2, since it has a value of $J_3(\check{x}_3^{(o)} = +1) = 0.77$, which is the lower cost-function value. The associated symbol value is $\check{x}_3^{(o)} = +1$. In the next step of the algorithm we proceed from node 2 of Figure 2.18 by calculating the cost function of Equation (2.103) for the next two potential values of $\check{x}_2^{(o)} = \pm 1$ as follows:

$$\begin{aligned} a_2 &= \textstyle\sum_{j=3}^{3} u_{2j}^{(o)} \cdot (\check{x}_j^{(o)} - \hat{x}_{\mathrm{MMSE},j}^{(o)}) \\ &= -0.16 \cdot (+1 - 0.33) = -0.11 \\ J_2([\check{x}_2^{(o)} \ -1]^T) &= J_1 + |a_2 + u_{22}^{(o)}(\check{x}_2 - \hat{x}_{\mathrm{MMSE},2}^{(o)})|^2 \\ J_2([+1 \ +1]^T) &= 0.77 + |-0.11 + 1.14 \cdot (+1 + 0.12)|^2 = 2.11 \\ J_2([-1 \ +1]^T) &= 0.77 + |-0.11 + 1.14 \cdot (-1 + 0.12)|^2 = 2.01. \end{aligned}$$

The resultant two values of $J_2(\check{x}_2^{(o)})$ are associated with node 3 and 6, respectively, which are seen at the third hierarchical level of Figure 2.18. The node from which the algorithm is further evolving is node 3, where the associated symbol is $\check{x}_2^{(o)} = -1$, which has a lower cost-function value than node 6.

The value of $\check{x}_2^{(o)} = [-1 \ +1]^T$ is now used for the calculation of the cost-function values of Equation (2.103) associated with $\check{x}_1^{(o)}$. The cost-function values $J_1(\check{x}_1^{(o)})$ are illustrated at the forth hierarchical level of Figure 2.18 within nodes 4 and 5.

Upon arriving at $J_1(\check{x}_1^{(o)})$ at the bottom of the graph, we have calculated the first potential solution of our optimization problem described by Equation (2.93), which is constituted by the left-most branch of the search tree illustrated in Figure 2.18. This potential solution is given by $\check{x}^{(o)} = [-1 \ -1 \ +1]^T$.

The recursive optimization continues from the bottom (node 4) to the top of the tree in Figure 2.18 with the objective of finding the specific branch terminating at the bottom of the search tree (hierarchical level 4), while having the minimum cost-function value. The symbol vector $\check{x}$ associated with this branch constitutes the ML solution. Considering now the flipping of

symbol $\check{x}_1^{(o)}$ at hierarchical level 3 is not beneficial, since the cost-function associated with this change would result in a higher value of 3.59 than the cost function value of 2.6 at node 4.

When the bottom-to-top recursive process initiated at node 4 arrives at node 3, however, it can be seen that changing the value of $\check{x}_2^{(o)}$ from -1 to $+1$ results in a cost function value of 2.11 at node 6, which is lower than the cost-function value of 2.6 recorded at node 4. Pursuing now the path from node 6 further down, this results in two new branches terminating at $\check{x}_1^{(o)}$ in node 7 and 8, which constitute additional potential solutions for $\boldsymbol{x}^{(o)}$. Pursuing this path from node 7 and 8 backwards, returning to the left-most branch and moving recursively up to node 0 will result in no further branches terminating at the bottom of the tree. Hence node 4 has the lowest cost-function value of 2.35.

The desired solution is described by the specific branch terminating at the bottom of the search tree that has the lowest cost-function value. Again, the associated symbol sequence is given by

$$\tilde{\boldsymbol{x}}^{(o)} = \check{\boldsymbol{x}}^{(o)} = [+1 \; +1 \; +1]^T,$$

which can be obtained by tracing the branch back from node 7 to node 1. By contrast, the identically ordered MMSE solution is given by

$$\tilde{\boldsymbol{x}}_{\text{MMSE}}^{(o)} = \texttt{sign}\left(\hat{\boldsymbol{x}}_{\text{MMSE}}^{(o)}\right) = [-1 \; -1 \; +1]^T.$$

The final step of obtaining the desired ML solution is to reverse the ordering of the $\check{\boldsymbol{x}}^{(o)}$. This yields

$$\begin{aligned} \tilde{\boldsymbol{x}} = \check{\boldsymbol{x}} &= [+1 \; +1 \; +1]^T \\ \tilde{\boldsymbol{x}}_{\text{MMSE}} = \texttt{sign}\left(\hat{\boldsymbol{x}}_{\text{MMSE}}\right) &= [-1 \; +1 \; -1]^T, \end{aligned}$$

which highlights the decision errors made by the MMSE detector. In order to apply this scheme for detecting 4QAM signals, the real-valued binary system model of Section 2.1.1.2 has to be employed. The algorithm can be applied as outlined in the previous section.

The search strategy outlined in Example 2.4 has the advantage that when implementing the algorithm, only the minimum cost-function value of all branches terminating at the bottom has to be stored in the memory and is used during the optimization-process as the reference value, when carrying out a decision as to whether particular nodes of the search tree are or are not used for a further evaluation of the search tree. More explicitly, if the first path terminating at the bottom of the search tree has been evaluated, then any node which is discarded owing to its higher cost-function value will never be re-visited. This property of avoiding re-visiting the nodes in the search tree and the fact that only one cost-function value has to be

stored in the memory as reference value renders the implementation of the algorithm particularly attractive. In the next example, we will discuss a different search-strategy, which is implementationally more complex, but may be more efficient in terms of reducing the number of visited nodes.

In addition to Example 2.4 discussing the conventional tree-search based optimization [61], an alternative search tree algorithm, which will be referred to as the Min-Path based optimization is proposed in this paragraph. Let us refer to the set of nodes whose children have not been evaluated as the 'search-front' $\mathcal{F}$. Then the new search strategy follows the simple principle of always pursuing the evaluation of the search tree according to Equation (2.103) from that particular node of the search-front, which is associated with the lowest cost-function value - hence the terminology Min-Path search. The search is terminated, once there is no node left belonging to the search-front, which has a lower cost-function value than the minimum cost-function value of the branches terminating at the bottom. This search-strategy is summarized in terms of pseudo-code in Algorithm 2.1.

Algorithm 2.1: Min-Path Tree-Search

Initialize :

$\mathcal{F} = \{\check{\mathbf{x}}_{\mathcal{F}}^{(1)} = [+1]^T, \check{\mathbf{x}}_{\mathcal{F}}^{(2)} = [-1]^T\}$;

calculate $J_{\mathcal{F}}^{(1)} = J_{N_b}(\check{\mathbf{x}}_{N_b} = \check{\mathbf{x}}_{\mathcal{F}}^{(1)})$, $J_{\mathcal{F}}^{(2)} = J_{N_b}(\check{\mathbf{x}}_{N_b} = \check{\mathbf{x}}_{\mathcal{F}}^{(2)})$ using Equation (2.103);

find index i of node with min. cost $-$ function value : $J_{\min} = \min(J_{\mathcal{F}}^{(i)})$;

while 1 do

$\check{\mathbf{x}}^{(a)} = [+1\ (\check{\mathbf{x}}_{\mathcal{F}}^{(i)})^T]^T$, $\check{\mathbf{x}}^{(b)} = [-1\ (\check{\mathbf{x}}_{\mathcal{F}}^{(i)})^T]^T$, $N_l = N_b - \text{length}(\check{\mathbf{x}}^{(a)})$;

evaluate $J_a = J_{N_l}(\check{\mathbf{x}}_{N_l} = \check{\mathbf{x}}^{(a)})$, $J_b = J_{N_l}(\check{\mathbf{x}}_{N_l} = \check{\mathbf{x}}^{(b)})$;

remove $\check{\mathbf{x}}_{\mathcal{F}}^{(i)}$ from $\mathcal{F}$;

add $\check{\mathbf{x}}^{(a)}$ and $\check{\mathbf{x}}^{(b)}$ to $\mathcal{F}$;

find index i of node with min. cost $-$ function value : $J_{\min} = \min(J_{\mathcal{F}}^{(i)})$;

if $(\text{length}(\check{\mathbf{x}}^{(i)}) == N_b)$

$\hat{\mathbf{x}}^{(o)} = \check{\mathbf{x}}^{(i)}$

break

end

end

Example 2.5: Min-Path Tree-Search Based Optimization

The initial stages of the algorithm, including the Cholesky factorization of Equation (2.95) as well as the re-ordering of the columns of the CIR matrix is identical to that in Example 2.4.

Then, we commence again by evaluating the cost-function value of the first nodes of Figure 2.18, resulting in the cost-function value of node 2 as well as node 9. These two nodes now form the search-front $\mathcal{F} = \{2, 9\}$.

The search continues from the node associated with the lower of the cost-function values, i.e node 2, resulting in the cost-function of node 3 as well as node 6 of Figure 2.18. The new search-front is given by the nodes $\mathcal{F} = \{3, 6, 9\}$.

The specific node, which is considered as a starting point for evaluating the search tree further is node 3 of Figure 2.18, since it is associated with the lowest cost-function. This results in the evaluation of nodes 4 and 5, yielding $\mathcal{F} = \{4, 5, 6, 9\}$.

In the next step the search continues from node 6, which has the lowest cost-function value, yielding $\mathcal{F} = \{4, 5, 7, 8, 9\}$. The search-process is concluded when a specific branch is found, which terminates at the bottom level of the search tree and has a lower cost-function value than all the other nodes in the search-front $\mathcal{F}$. In our example this requirement is satisfied for node 7.

Note that for this specific example both tree-search algorithms result in visiting and evaluating the same nodes in the search tree. This is however not necessarily the case for different channel conditions, as it will be seen in Sections 2.3.4.1 and 3.4.4.

The novel Min-Path tree-search algorithm potentially requires a larger amount of memory, since the complete search-front of the search tree has to be stored and its size increases steadily, as the search-process evolves. The exact effect of the two different search strategies will be illustrated in the next section, where the computational complexity imposed by the OHRSA is discussed. It may be stated in conclusion that the conventional search strategy of Example 2.4 first evaluates a single branch from the top to the bottom of the search tree, which is then used as a reference, when deciding whether it is worth continuing the search process from any given node or not. By contrast, the Min-Path search strategy of Algorithm 2.1 always continues the search process from the node associated with the lowest cost-function, regardless of its hierarchical level.

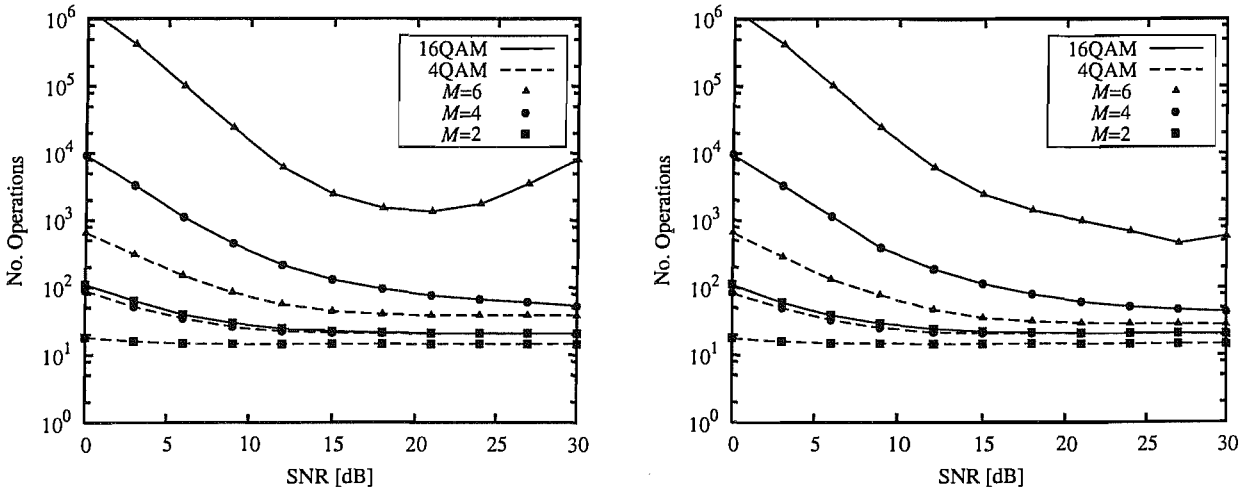
2.3.4.1 Computational Complexity of the OHRSA Detector

The computational complexity of the OHRSA can again be divided into two contributions, namely the cost imposed by the calculation of the MMSE solution of Equation (2.97) and the factorization of the covariance matrix of Equation (2.96) into an upper triangular matrix as well as the evaluation of the search tree according to Equation (2.103).

The calculation of the MMSE solution of Equation (2.97) is composed of $N_b^2(4N-1)+N_b$ number of real-valued addition plus multiplications required for the evaluation of the covariance matrix plus $O(N_b^3)$

Table 2.6: The normalized computational complexity (real-valued additions and multiplications) of the OHRSA detector, when considering the detection of a single transmitted bit.

MMSE Solution and factorization of Equation (2.97) and Equation (2.96)	$8NN_b + 2N - N_b + O(N_b^2)$
Tree evaluation of Equation (2.102) and Equation (2.103) at hierarchical level i	$3(N_b - i) + 1$



(a) Conventional tree-search (b) Min-Path tree-search

Figure 2.19: The normalized computational complexity required for evaluating the search tree as a function of the SNR for a different number of transmit AEs as well as for different modulation schemes when considering the detection of a single transmitted bit calculated according to Table 2.6. The number of receive AEs was set to $N = 4$.

number of real-valued operations required for the calculation of its inverse. For the final calculation of the weight matrix in Equation (2.97) an additional $(2N_b - 1)N_b 2N$ number of real-valued additions plus multiplications are required for the multiplication of the inverse of the covariance matrix with the cross-correlation matrix. The multiplication of the MMSE weight matrix with the received signal contributes a further $(4N - 1)N_b$ operations. The Cholesky decomposition of the covariance matrix requires typically $O(N_b^3)$ number of real-valued operations.

In addition to the calculation of the MMSE solution of Equation (2.97) and the Cholesky factorization of Equation (2.96) , the complexity of the OHRSA is determined by the number of operations required for the evaluation of Equation (2.103) associated with the search tree C^{Tree}. Assume, the search tree operates at the hierarchical level i, the number of real-valued additions and multiplications for the evaluation of Equation (2.103) as well as Equation (2.102) is given as $3 \cdot (N_b - i - 1) + 4 = 3 \cdot (N_b - i) + 1$. The exact computational complexity imposed by the search tree algorithm has to be evaluated through simulations. The complexity of the OHRSA algorithm is summarized in Table 2.6, normalized by the number of bits N_b.

The normalized computational complexity defined as the number of real-valued additions plus multiplications required for the evaluation of the search tree is illustrated in Figure 2.19 as a function of the SNR for

both 4QAM and 16QAM signals, when considering $M = 2, 4$ and 6 transmit AEs, while employing $N = 4$ receive AEs. It can be observed that in contrast to the MMSE detector of Section 2.3.1, to the MBER detector of Section 2.3.2 as well as to the Bayesian detector of Section 2.3.3, the complexity associated with the evaluation of the OHRSA search tree is dependent on both the SNR as well as on the ratio between the number of transmit AEs and the number of receive AEs. When considering a rank-deficient system, the complexity imposed at low SNR values was increased. The terminology rank-deficient in this context implies that the $(2N \times N_{\text{bit}}M)$-dimensional real-valued channel matrix $\underline{\underline{H}}$ has more columns than rows. It is referred to as rank-deficient, since the channel's covariance matrix does not have full rank. When comparing Figure 2.19a, which shows the complexity associated with the conventional tree-search discussed in both Example 2.4 and in Figure 2.19b, which shows the complexity associated with the novel Min-Path search strategy proposed in Algorithm 2.1, it can be observed that if the system is not particularly rank-deficient, both search strategies impose a similar computational cost. However, for highly rank-deficient systems operating at high SNRs, when transmitting for example 16QAM signals used in conjunction with a (4×6)-antenna MIMO system, the Min-Path algorithm of Example 2.5 clearly imposes a lower computational cost than the conventional search of Example 2.4.

Figure 2.20 shows the normalized complexity required for the evaluation of the search tree as a function of the number of transmit AEs M for both BPSK modulated as well as for 4QAM and 16QAM signals, when considering $N = 2, 4$ and 6 receive AEs at SNR=20 dB. Considering the case of $N = 2$ receive antennas, the system becomes rank-deficient when we have $M > 4$ for BPSK modulated signals, $M > 2$ for 4QAM signals as well as in case of $M > 1$ for 16QAM signals. It is exactly for these cases, when the complexity starts to increase quite considerably compared to a low number of transmit AEs M, as it can be seen from Figure 2.20.

In Figure 2.19 and 2.20 we only considered the complexity associated with the evaluation of the search tree. By contrast, in Figure 2.21 the computational complexity is illustrated under the assumption that the MMSE solution and the Cholesky decomposition of Equation (2.97) and Equation (2.96), respectively, have to be evaluated for each received signal vector. The proportionality factor of Table 2.6 was set to 3. It can be seen from Figure 2.21 that the complexity contribution associated with the calculation of the MMSE solution as well as with the Cholesky decomposition, has a dominant effect only for a low number of transmitters. As the number of transmitters increases and the system becomes more rank-deficient, the complexity of the OHRSA detector is dominated by the associated tree-search.

2.3.4.2 Performance of the OHRSA Detector

The MIMO system considered is identical to the system used in Section 2.3.1.2, which is characterized by $M = 2$ transmit AEs at the MS and a BS employing a N receive antennas. A block-fading channel

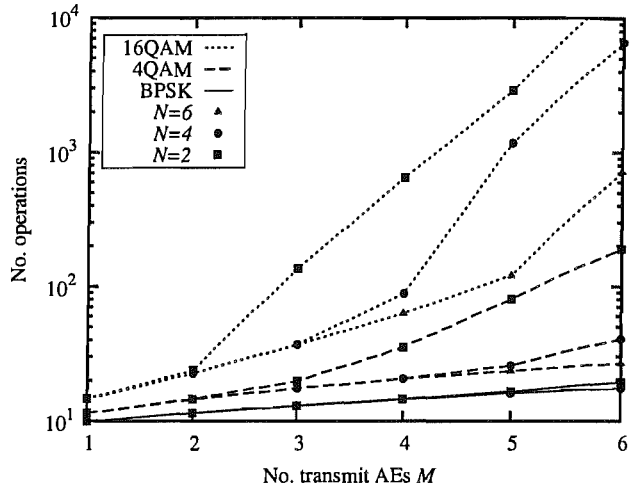


Figure 2.20: The normalized computational complexity required for evaluating the search tree as a function of the number of transmit AEs for different modulation schemes and a different number of MIMO array weights, when considering the detection of a single transmitted bit calculated according to Table 2.6 at SNR=20 dB.

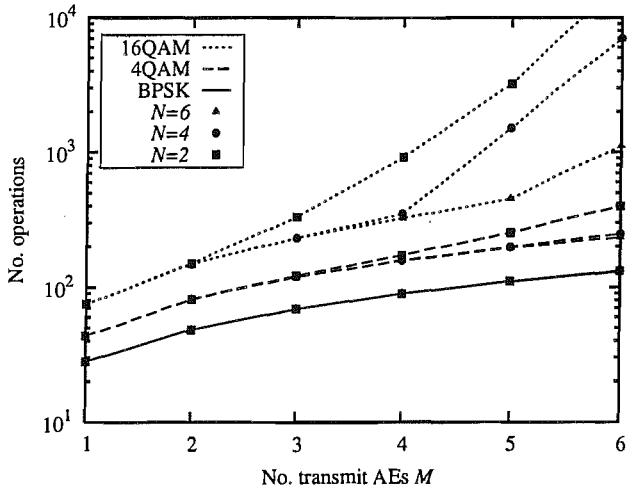


Figure 2.21: The normalized computational complexity required for evaluating the search tree plus the evaluation of the MMSE solution as a function of the number of transmit AEs for different modulation schemes and a different number of MIMO array weights, when considering the detection of a single transmitted bit calculated according to Table 2.6 at SNR=20 dB. The proportionality factor associated with the MMSE solution was set to 3.

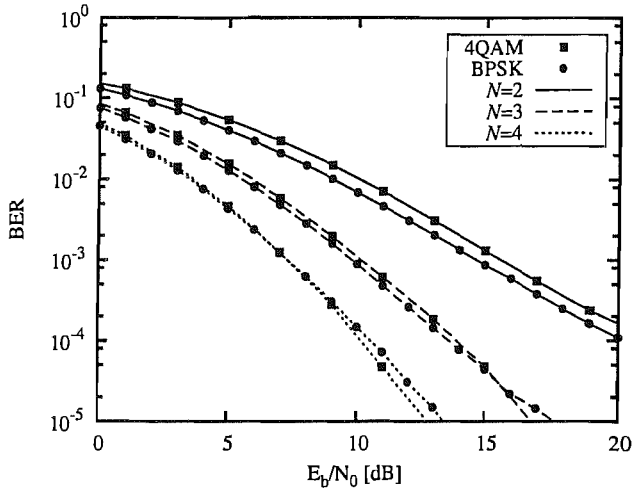


Figure 2.22: BER versus E_b/N_0 performance for a MIMO system employing $M = 2$ transmit AEs and $N = 2, 3$ as well as 4 receive AEs. BPSK and 4QAM modulated signals as well as having perfect channel knowledge were considered.

with $\mathrm{E}\left[|h_{nm}|^2\right] = 1$ was assumed between all antenna links, which was perfectly known by the receiver. Figure 2.22 shows the BER versus E_b/N_0 performance for the uncoded system using the OHRSA MIMO detector and transmitting both BPSK and 4QAM signals. It can be observed that the performance of the OHRSA algorithm is only marginally inferior compared to that of the Bayesian detector discussed in Section 2.3.3.

2.3.5 MIMO Detector Complexity versus Channel Condition

Before studying the performance of the algorithms introduced in this section, let us briefly consider their suitability for different channel conditions. The suitability of the algorithms under three different channel conditions will be evaluated.

2.3.5.1 Complexity Under Block-Fading Conditions

The first channel model considered is the block-fading channel. Under the corresponding very slow-fading conditions, both the linear MMSE and the liner MBER detector exhibit a relatively low computational complexity since once the array weight matrix has been calculated, the detection procedure imposes a low computational cost as shown in Table 2.3 and in Table 2.4 for the MMSE and the MBER algorithm, respectively. The reduced complexity ML algorithm introduced in Section 2.3.4 is also attractive, since the calculation of the MMSE solution and the Cholesky decomposition of the covariance matrix only have to be carried out once. The detection complexity is then only determined by the complexity required for the evaluation of the cost-function throughout the detection tree. The Bayesian detector appears less attractive, since many Euclidean distance calculations are necessary in order to detect each bit.

2.3.5.2 Complexity Under Slow Fading Conditions

If the channel is slowly time-varying, so that the MBER and the MMSE array weights can be adjusted adaptively using the LMS [113] or the LBER [96] algorithm, respectively, the complexity of the linear algorithms remains low and both algorithms retain their benefits. Under slowly time-varying channel conditions the Cholesky decomposition of the covariance matrix can be directly tracked rather than being regularly recomputed and thus also the OHRSA is attractive under such channel conditions. The computational complexity of the Bayesian detector still remains high compared to the previously mentioned algorithms.

2.3.5.3 Complexity Under Rapidly Fading Conditions

If the channel experienced by the transmitted signal is varying rapidly it may have to be tracked using a complex channel estimator such as for example a Kalman estimator. Then the MMSE detector requires a matrix inversion for each detected symbol. Similarly the MBER algorithm would have to calculate the MBER array weights for each symbol to be detected. Compared to the MBER algorithm, the MMSE algorithm still retains a relatively low computational complexity. The MBER algorithm becomes unsuitable for such channel conditions, since its complexity may exceed the complexity of the Bayesian detector as it can be seen by comparing Table 2.4 and Table 2.5. Similarly to the MMSE algorithm, the complexity of the OHRSA increases moderately for rapidly time-varying channels, since the MMSE array weights and the Cholesky decomposition have to be calculated for each symbol as it can be seen when comparing Figure 2.20 and Figure 2.21. Generally the complexity associated with the reduced complexity ML detector of Section 2.3.4 remains far lower than that associated with the Bayesian detector of Section 2.3.3.

2.3.5.4 MIMO Detector Complexity Summary

In Table 2.7 the computational complexity associated with the different algorithms is summarized. It can be observed that for slowly time-varying channels the linear MMSE and MBER detector have a clear advantage over the Bayesian and the OHRSA aided MIMO detector. For rapidly time-varying channels the advantage of the MBER MIMO receiver over the Bayesian detector diminishes, since the computational complexity associated with both detectors increases exponentially with the total number of bits $N_{\text{bit}}M$ to be detected.

2.3.6 Turbo-Coded Performance of MIMO Detectors

In this subsection the performance of the different algorithms is compared for a SDM system supporting M transmit AEs and a receiver employing a different number of receive AEs N. The transmitter employs Forward Error Correction (FEC) coding to improve the achievable performance. The channel code is a

Table 2.7: Summary of the normalized computational complexity associated with the MMSE, the MBER, the Bayesian and the OHRSA MIMO detector under different channel conditions.

	MMSE of Table 2.3	MBER of Table 2.4	Bayesian of Table 2.5	OHRSA of Table 2.6
Block Fading Channel	$O(N)$	$O(N)$	$O(2^{N_{\text{bit}}M})$	$O(N) < \bullet < O(2^{N_{\text{bit}}M})$
Slowly Fading Channel	$O(N)$	$O(N)$	$O(2^{N_{\text{bit}}M})$	$O(N) < \bullet < O(2^{N_{\text{bit}}M})$
Rapidly Fading Channel	$O\left(\frac{N^3}{N_{\text{bit}}M}\right)$	$O(2^{N_{\text{bit}}M})$	$O(2^{N_{\text{bit}}M})$	$O\left(\frac{N^3}{N_{\text{bit}}M}\right) < \bullet < O(2^{N_{\text{bit}}M})$

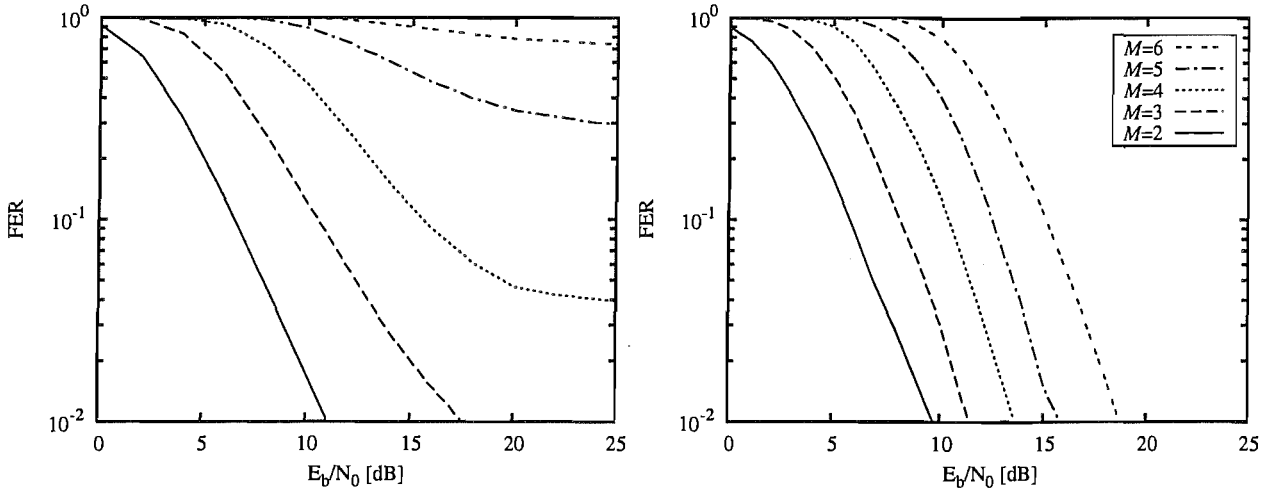
$R = 1/2$ rate punctured turbo code using Recursive Systematic Convolutional (RSC) component codes described by the polynomial $\boldsymbol{g} = [7\ 5]$ [28], which was fed with the hard-decision estimates of the MIMO detector .

2.3.6.1 Performance of MIMO Detectors for Block-Fading Channels

In this first study we compare the achievable performance of the two linear detectors, namely that of the MMSE detector discussed in Section 2.3.1 and that of the MBER detector of Section 2.3.2.

For a first comparison we assume the fading envelope to be constant over an entire transmission burst equivalent to 1000 information bits, where the channel was assumed to be perfectly known by the receiver. Since we consider a half-rate code for our simulations, the interleaver employed by the turbo code had a length of 1000 bits and the channel interleaver of Figure 2.1 was of length 2000 bits. A total of 10 000 bursts were transmitted in order to acquire sufficiently reliable fading statistics and a reliable FER estimate. Note that since we are considering block-fading channels, the performance measure is not the BER but rather the FER, which in this specific setting is equivalent to the codeword error-rate.

In Figure 2.23 the achievable FER versus E_b/N_0 performance is shown for a system employing BPSK modulated signals and $N = 3$ receive AEs as well as $M = 2, 3, 4, 5$ and 6 transmit AEs. It can be observed that the MBER detector achieves a significantly lower FER, especially when the number of transmit AEs is higher than the number of receive AEs, as it can be seen by comparing Figure 2.23a and Figure 2.23b, where the FER versus E_b/N_0 performance is illustrated for the MMSE and the MBER detector, respectively. To understand this phenomenon, we consider the real-valued binary system model of Section 2.1.1.2, where the channel matrix dimensions are $2N \times \log_2(\mathcal{M})M$. Hence a system considering BPSK modulated signals and using $M = 6$ transmit AEs as well as $N = 3$ receive AEs is described by a (6×6)-dimensional channel matrix. Since the number of channel matrix columns and the number of receivers in this system is identical, it is referred to as a *fully-loaded* system. If the number of columns of the real-valued channel matrix is larger than the number of rows, we refer to the system as rank-deficient , which may also be termed as an *over-loaded* system. As already mentioned in Section 2.3.3, the performance difference between the



(a) MMSE Detector (b) MBER Detector

Figure 2.23: FER versus E_b/N_0 performance for a scenario of $N = 3$ receive AEs as well as a transmitter employing $M = 2, 3, 4, 5$ and 6 AEs and benefiting from perfect channel knowledge. The system employed a $R = 1/2$-rate punctured turbo-code and BPSK modulated signals were considered.The channel was assumed to be constant over one codeword determined by the length of the 2000 bit channel interleaver.

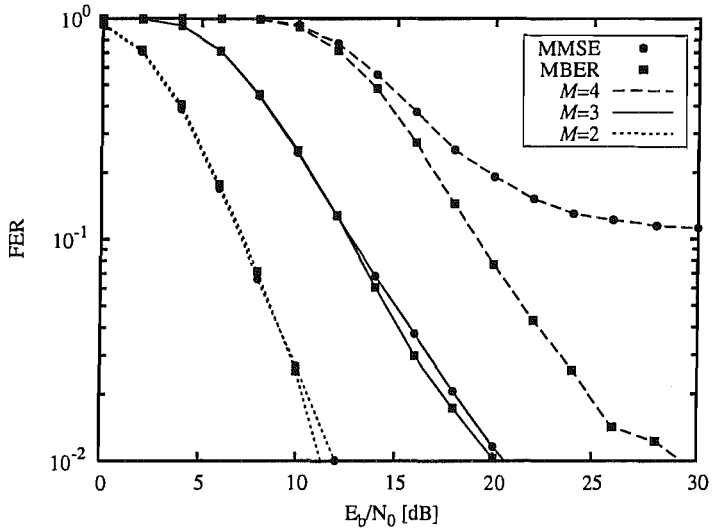


Figure 2.24: FER versus E_b/N_0 performance for a scenario of $N = 3$ receive AEs as well as a transmitter employing $M = 2, 3$ and 4 AEs and benefiting from perfect channel knowledge. The system employed a $R = 1/2$-rate punctured turbo-code and 4QAM signals were considered.The channel was assumed to be constant over one codeword determined by the length of the 2000 bit channel interleaver.

MBER and the MMSE algorithm recorded for BPSK modulated signals is mainly imposed by the fact that the MMSE algorithm attempts to minimize the MSE associated with the real- and the imaginary-part of the detector output, whereas the MBER detector only optimizes the relevant real-part. Since the MMSE detector aims for minimizing the MSE composed by the real- and the imaginary-part of the filter output, it looses out in terms of the associated degree of freedom and hence the MMSE detector may become over-loaded for $M > N$ even when the signal is transmitted using the lowest-throughput BPSK modulation.

The achievable FER performance recorded for 4QAM signals considering the same scenario is illustrated in Figure 2.24. For this system a moderate number of $M = 3$ transmit AEs yield a (6×6)-dimensional real-

valued system matrix associated with a fully loaded system. It can be observed, that for $M = 2$ and 3 both the MBER and the MMSE algorithm exhibit a similar FER performance. If the system is rank-deficient, however, for example when using $M = 4$ transmit AEs, the MBER detector exhibits a significantly lower error-floor than the MMSE detector.

The fact that the MBER detector fails to outperform the MMSE detector more significantly has the following two reasons

- The MBER and the MMSE algorithm perform similarly for BERs higher than 0.01 in the context of the uncoded system. Since it was observed that for BERs below 0.04 the powerful turbo-code employed by the system is capable of correcting the majority of errors, only a marginal performance difference is visible between the considered MMSE and MBER detector;
- In the context of Rayleigh fading channels the FER is determined by a low number of error events. It appears from the simulations that when a severe fade is encountered, which may induce linearly non-separable channel output states, neither the MBER nor the MMSE detector is capable of reliably detecting the received signals.

Let us now consider a scenario, which is characterized by BPSK modulated signals, $M = 2$ transmit AEs and $N = 2$ as well as 4 receive AEs. Figure 2.25 shows the FER exhibited by the different detectors together with the outage probability discussed in Section 2.3. The outage-probability is recorded for this system for a transmission rate of $R = R_c M \log_2 \mathcal{M} = \frac{1}{2} \cdot 2 \cdot 1 = 1$ and constitutes the lower bound for the codeword error-rate. It can be observed that for a low number of receive AEs, such as $N = 2$, the performance difference between the different detectors is more significant than for a higher number of receive AEs, namely for $N = 4$. For $N = 2$ receive AEs and a FER of 10^{-2} both the Bayesian and the OHRSA detector operate with an E_b/N_0 discrepancy of approximately 5 dB with respect to the channel capacity limit.

2.3.6.2 Performance of MIMO Detectors for Uncorrelated Fading Channels

In Section 2.3.5 where the computational complexity imposed by the different MIMO detectors was discussed it was emphasized that the computational cost associated with the linear MBER detector exceeds the computational cost associated with the optimum Bayesian detector for uncorrelated fading channels as it can be seen from Table 2.7. Therefore the MBER detector in not considered in this subsection, where the achievable detector performance is discussed when operating in a fast-fading environment.

The BER performance results recorded for the same system configuration as that used for generating the results shown in Figure 2.25 but assuming an uncorrelated fading channel and 4QAM modulated signals

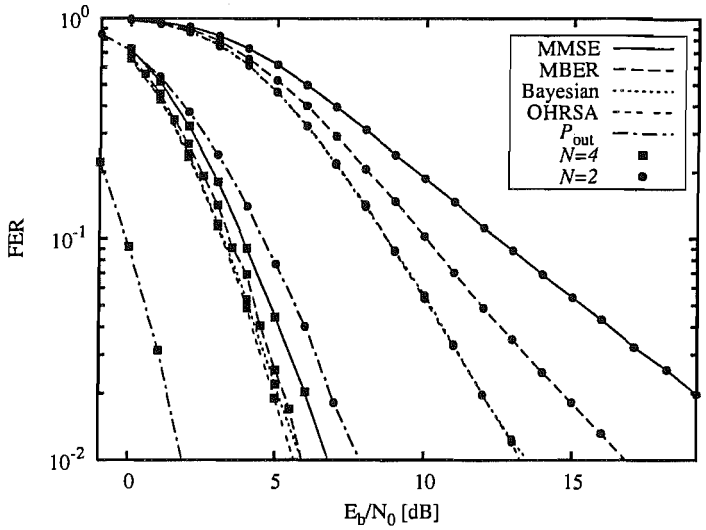


Figure 2.25: FER versus E_b/N_0 performance for a scenario of $M = 2$ transmit AEs as well as a BS receiver employing $N = 2$ and 4 receive AEs and benefiting from perfect channel knowledge. The system employed a $R = 1/2$-rate punctured turbo-code.The channel was assumed to be constant over one codeword determined by the length of the 2000 bit channel interleaver.

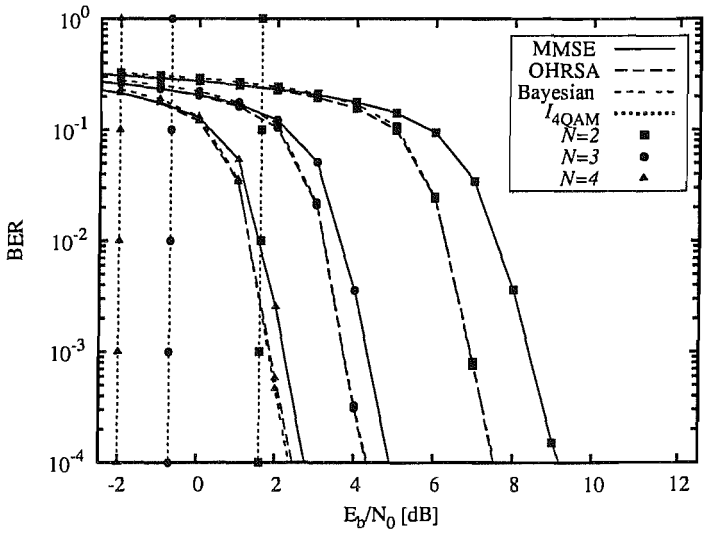


Figure 2.26: BER versus E_b/N_0 performance for a scenario of $M = 2$ receive AEs as well as a BS receiver employing $N = 2,3$ and 4 receive AEs and benefiting from perfect channel knowledge. The system employed a $R = 1/2$-rate punctured turbo-code. The channel was assumed to be an uncorrelated fading channel.

are illustrated in Figure 2.26 together with the achievable information rate bounds, which were obtained as outlined in Example 2.2. It can be observed that for $N = 4$ receive AEs the algorithms operate at approximately 4-5 dB away from capacity. As the detection task becomes more challenging, i.e a lower number of receiver AEs is employed by the BS, it can be observed that the difference between the capacity bound and the performance of the detection algorithms increases. Furthermore, the performance difference between the MMSE and OHRSA detector increases as the number of receive AEs decreases.

2.3.7 Conclusion

In this section we have provided a set of linear and non-linear MIMO detection algorithms. Their performance was compared using a simple channel setting and the differences were highlighted. It was shown that even the more sophisticated algorithms are still far from approaching the lower capacity bound. Furthermore, the computational complexity of the different detection strategies was analyzed considering different channel conditions. In the next section the performance of two of the algorithms, namely the linear MMSE receiver and the non-linear OHRSA detector is investigated under the influence of imperfect, i.e estimated CSI.

2.4 Channel Estimation for Narrowband Channels

In Figure 2.27 we have depicted how a training based channel estimator can be combined with a MIMO detector. The initial channel estimate can be obtained by using a sequence of training symbols, which are known by the receiver and can by employed by the channel estimator. If the detector operates at a relatively low BER and the channel is time-varying and hence requires channel tracking, decision directed channel estimation indicated by the dotted line can be used. In decision directed channel estimation, the symbols returned by the detector are assumed to be correct and fed back to the channel estimator for updating the channel estimate. If the channel estimate needs to be refined further, the detected and decoded transmission burst can be encoded and re-modulated again for re-estimating the channel. The improved channel estimate is then employed for detecting the received symbols again.

Most estimation algorithms discussed in the literature are designed for estimating a single stochastic variable or a vector of stochastic parameters rather than a matrix, as it would be required in the context of MIMO channel estimation. Therefore we first introduce an alternative form of the system model provided in Equation (2.4) representing the channel matrix as a vector. On this basis different channel estimation algorithms known from the literature may be directly applied to MIMO channel estimation. Let us define the channel

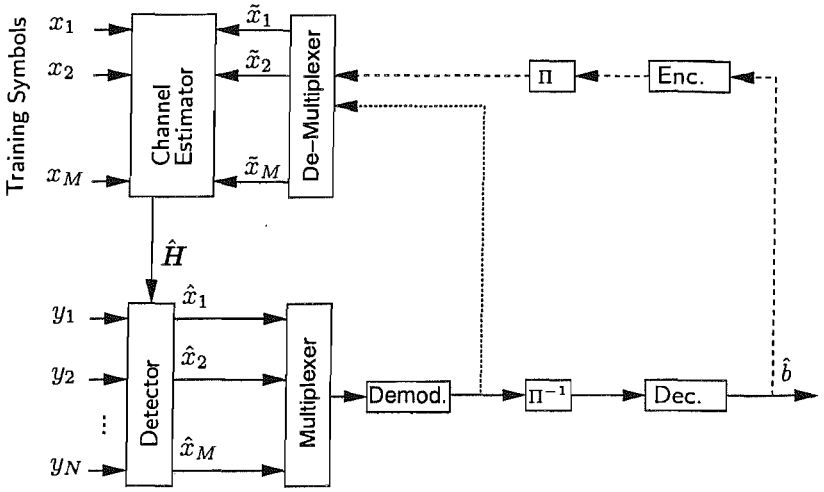


Figure 2.27: MIMO channel estimation.

determining the detector's input signal in form of a (NM)-element vector, which is given as

$$\boldsymbol{h}(k) = \mathrm{vec}\,(\boldsymbol{H}(k)) \tag{2.111}$$

$$= [h_{11}(k),\ldots,h_{N1}(k),\ldots,h_{1M}(k),\ldots,h_{NM}(k)]^T, \tag{2.112}$$

where vec () represents the column-wise vector operator concatenating all columns of $\boldsymbol{H}(k)$. To form the desired system equation, the transmitted symbol vector influencing the channel output is expressed as a matrix, given as

$$\mathbf{X}(k) = \boldsymbol{x}(k)^T \otimes \boldsymbol{I}_N, \tag{2.113}$$

where $\otimes$ is the Kronecker product. The system matrix given in Equation (2.4) can now be re-written as

$$\boldsymbol{y}(k) = \mathbf{X}(k)\boldsymbol{h}(k) + \boldsymbol{\eta}(k). \tag{2.114}$$

Before discussing channel estimation algorithms for the system described by Equation (2.112) it is necessary to understand the time variant behaviour of the channel to be estimated. The rate of channel variations over time depends on the normalized Doppler frequency, which is defined as [28]

$$f_D T_s = \overline{f}_D = \frac{v f_c}{c} T_s, \tag{2.115}$$

where v is the velocity of the MS traveling in direction perpendicular to the signal propagation path, f_c is the carrier frequency, c is the speed of light and T_s is the symbol period. Considering for example a SDMA system operating at $f_c = 2\ GHz$ at a symbol rate of $2\ \frac{\mathrm{MBaud}}{\mathrm{s}}$, a normalized Doppler frequency of $\overline{f}_D \approx 1 \cdot 10^{-4}$ corresponds to a vehicular velocity of $v = 100\frac{\mathrm{km}}{\mathrm{h}}$. This simple example illustrates that in single carrier systems $\overline{f}_D = 1 \cdot 10^{-4}$ corresponds to a high velocity of the mobile user. The channel

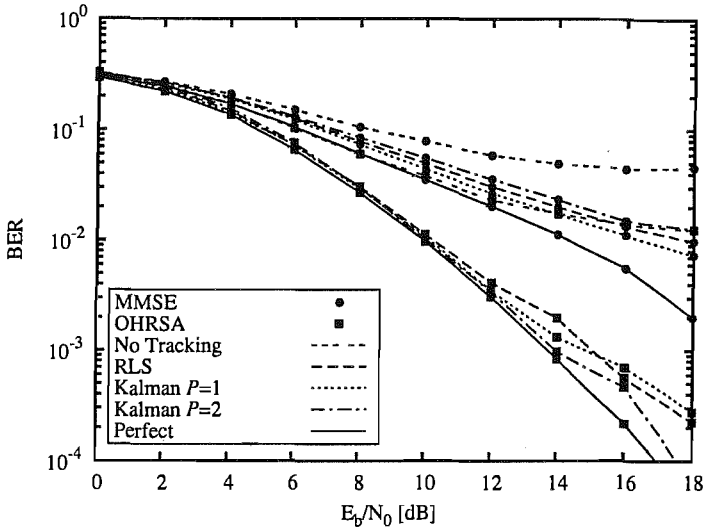


Figure 2.28: BER versus E_b/N_0 performance for a system employing $M = 2$ transmit AEs and a BS using $N = 2$ number of AEs. The Doppler frequency was assumed to be $\overline{f}_D = 0.0001$ and BPSK modulated signals were assumed.

estimation algorithms applied in the next subsection, aim for tracking and estimating channels having a normalized Doppler frequency of this order of magnitude.

Based on Equation (2.114) standard channel estimation algorithm might be applied as outlined in [28]. In this section we consider the RLS channel estimation algorithm [28, 113, 121], which is characterized by the forgetting factor λ and the Kalman channel estimator of [122], which is characterized by the order P [123]. The training sequences were chosen randomly and have not been optimized as for example proposed in [123].

2.4.1 Detector Performance Using Estimated Channel State Information

To evaluate the effect of channel estimation errors on the different MIMO detection algorithms introduced earlier, a system consisting of $M = 2$ transmit AEs and a BS using $N = 2$ receive AEs is considered. The normalized Doppler frequency was assumed to be $\overline{f}_D = 0.0001$. The training sequences were randomly generated BPSK modulated signals for each transmit antenna and consisted of 150 symbols. The 'payload' per transmit antenna was chosen to be 1050 data symbols. The forgetting factor of the RLS was set to $\lambda = 0.98$ and a $P = 1^{st}$ and $P = 2^{nd}$ order Kalman channel estimator were considered.

If Decision Directed Channel Estimation (DDCE) was used, the channel estimate was updated with a delay of one symbol assuming that the decision made by the detector was correct. In other words, when operating in decision directed mode it was assumed by both the detector and the channel estimator that we had $\boldsymbol{h}(k) = \boldsymbol{h}(k-1)$.

In Figure 2.28 the performance of the MMSE algorithm using the estimated rater than perfect CSI is illus-

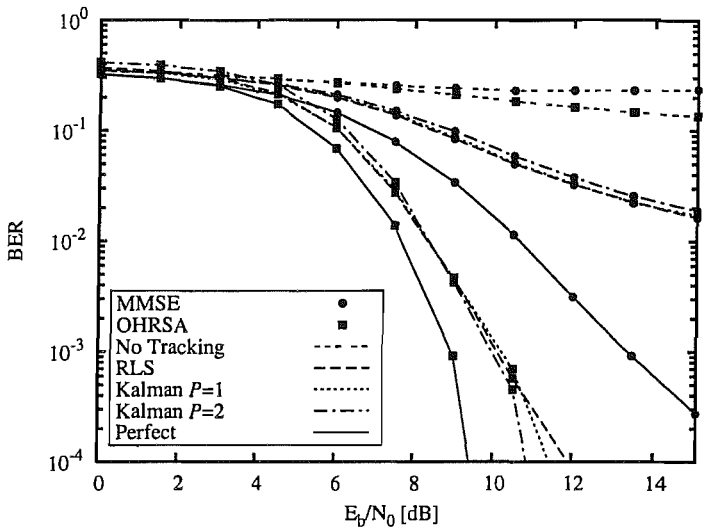


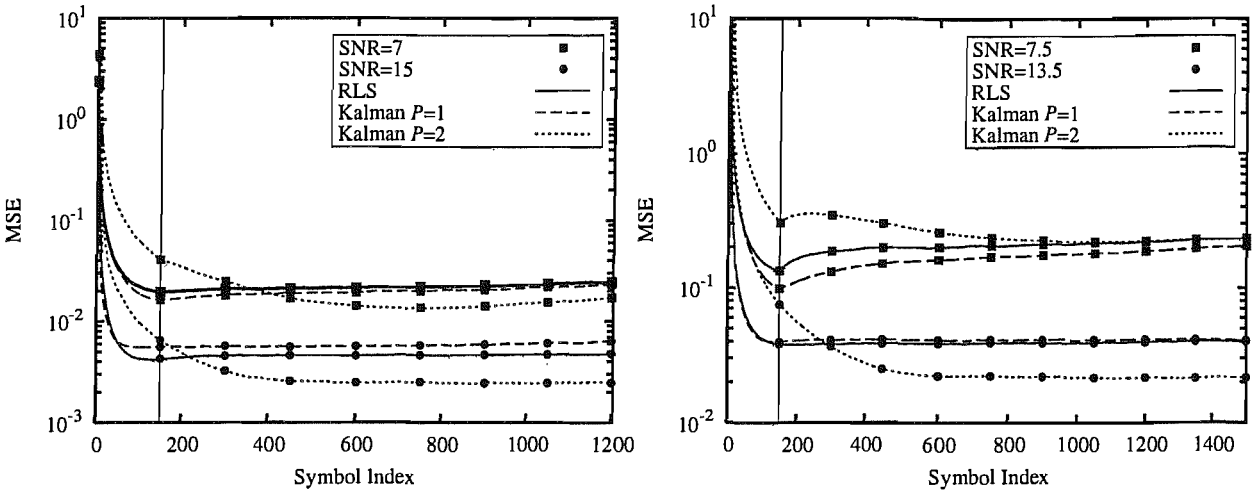
Figure 2.29: BER versus E_b/N_0 performance for a system employing $M = 4$ transmit AEs and a BS using $N = 4$ AEs. The Doppler frequency was assumed to be $\overline{f}_D = 0.0001$ and 4QAM signals were assumed.

trated. The label 'No Tracking' indicates that the channel envelope was assumed to be constant over the whole transmission frame and the channel gain value was updated at the end of the training sequence. It can be seen that if no channel tracking is used, both the MMSE and OHRSA detectors exhibit a relatively high BER. The BER of the OHRSA aided detector operating without channel tracking however is significantly lower than that of the MMSE detector also refraining form channel tracking.

If DDCE is used for the OHRSA assisted detector, it can be observed that the performance of the system using RLS and first-order Kalman channel estimators is identical and is only marginally worse than that associated with the second-order Kalman filter. Compared to the 'genie-aided' detector using perfect CSI, the associated performance degradation is negligible. It can be observed that in the high SNR region, the performance of the first-order Kalman filter and that of the RLS channel estimator aided detector degrades with respect to the detector using perfect CSI. This is due to the fact that at higher SNRs the MSE of the channel estimator may exceed the noise-level and thus contributes to the performance degradation. Considering the performance of the MMSE detector, it can be observed that at the same BER, its performance is significantly more degraded than that of the OHRSA-based detector, when considering realistically estimated CSI.

Let us now extend the MIMO system to having $M = 4$ transmit AEs, $N = 4$ receive AEs and 4QAM signals.The normalized Doppler frequency $\overline{f}_D$ was assumed to be $\overline{f}_D = 0.0001$ and a half-rate punctured turbo code was employed in the system considered. The number of training symbols per transmit AE was chosen to be 150 and the payload per transmit AE was set to 1350 symbols. It can be seen from Figure 2.29, that the performance degradation due to channel estimation errors is now considerably higher since the training overhead is only about 11%. Similar to Figure 2.28 it can be seen that the ML detection based OHRSA assisted MIMO receiver is more robust against channel estimation errors.

In Figure 2.30 the MSE of the channel estimator is plotted as a function of the symbol index at different



(a) $M = 2$, $N = 2$, BPSK modulated signals

(b) $M = 4$, $N = 4$, 4QAM signals

Figure 2.30: Channel estimation MSE versus symbol index for an OHRSA aided MIMO detector employing M transmit AEs and a BS using N number of AEs. The normalized Doppler frequency was assumed to be $\overline{f}_D = 0.0001$.

SNRs for the two systems considered in Figure 2.28 and Figure 2.29, respectively. It can be observed how the RLS and the first-order Kalman filter converge considerably faster at the cost of an increased steady-state MSE. The lower steady-state MSE associated with the second-order Kalman estimator explains its superior performance at a low BER.

In order to benefit from the rapid convergence of the first order Kalman filter and from the low steady-state error of the second-order filter, one might consider reconfigurable channel estimators, while using the estimate of the first-order filter at the beginning and switching to the second-order filter, once it has reached its steady-state.

2.5 Conclusion

In this chapter different MIMO detection algorithms, namely the MMSE algorithm of Section 2.3.1, the MBER algorithm of Section 2.3.2, the Bayesian detector of Section 2.3.3 as well as the OHRSA detector of Section 2.3.4 have been introduced. Their different optimization functions as well as their associated complexity were characterized in terms of the number of real-valued multiplications plus additions and are summarized in Table 2.8.

As detailed in Section 2.3.5 and as it becomes clear from Table 2.8, the MMSE algorithm retains its low computational complexity over both slowly as well as rapidly fading channels. The MBER algorithm however has a low computational complexity for slowly fading channels, whereas in rapidly fading environments its computational cost increases exponentially with the number of transmit AEs as well as with the number of bits per symbol. For slowly fading channels the MBER algorithm has the capability of outperforming the

MMSE detector, especially when the number of transmitters is higher than the number of receive AEs, as shown in both Figure 2.23 as well as Figure 2.24.

The nonlinear Bayesian detector of Section 2.3.3 constitutes the optimum minimum error-rate non-linear receiver, which significantly outperforms the linear MMSE and MBER algorithm. The associated computational cost however increases exponentially with the number of transmit AEs as well as with the number of bits per symbol for all kinds of channel conditions, as shown in Table 2.8. We have therefore introduced a reduced complexity ML detector in Section 2.3.4, which achieves the performance of the Bayesian detector at a reduced computational cost, as shown in Table 2.8.

When quantitatively comparing the different detectors' performance, we investigate their performance in terms of the distance in dB with respect to the theoretical channel capacity limits discussed in Section 2.2. Under the assumption of block-fading channels the distances $\Delta_{P_{\text{out}}|P_{\text{out}}=0.02}$ [dB] measured from the outage probability bound and given in Table 2.8 were obtained for an outage probability of $P_{\text{out}} = 2\%$ from Figure 2.25. Under the assumption of uncorrelated fading channels the distances $\Delta_{I|\text{BER}=10^{-4}}$ [dB] from the channel capacity limit given in Table 2.8 were obtained for a BER of 10^{-4} from Figure 2.26. The values of $\Delta_{P_{\text{out}}|P_{\text{out}}=0.02}$ and $\Delta_{I|\text{BER}=10^{-4}}$, respectively, show that the performance difference between the non-linear and the linear detectors is more pronounced when the ratio between the number of transmit and receive AEs is large. The smaller this ratio becomes the more modest the performance advantage of the non-linear detector over the linear MMSE and MBER receiver.

Furthermore, different MIMO channel estimators have been employed in order to investigate the detectors' performance using imperfect CSI. It was shown that for slowly time-varying channels the OHRSA detector experiences only marginal performance degradation, which is on the order of 1 dB for a (2×2)-dimensional system and on the order of 2 dB for a (4×4)-dimensional system as it can be seen from Figure 2.28 and Figure 2.29, respectively.

Throughout this chapter it was assumed that the channel is a frequency-flat fading channel. In the next chapter the channel will be extended to a single-carrier wideband channel and the previously introduced detectors as well as channel estimators are applied to the resultant dispersive communication environment. They are also appropriately modified in order to limit the complexity of the algorithms.

Table 2.8: Summary of the MIMO detectors studied in Chapter 2.

			MMSE of Section 2.3.1	**MBER** of Section 2.3.2	**Bayesian** of Section 2.3.3	**OHRSA** of Section 2.3.4
Optimization			$\hat{x} = W^H y$, where $W = \arg \min_{W \in \mathbb{C}^{N \times M}} \mathrm{E}\left[\|x - W^H y\|^2\right]$	$\hat{x} = W^H y$, where $W = \arg \min_{W \in \mathbb{C}^{N \times M}} P_e(W)$	$\tilde{x}_m = \arg\max_{x_m \in \{+1,-1\}} p(y\|x_m)$	$\tilde{x} = \arg\max_{x \in \mathcal{X}} p(y\|x)$
Complexity [a]	block-fading		$O(N)$	$O(N)$	$O(2^{N_{\text{bit}}M})$	$O(N) < \bullet < O(2^{N_{\text{bit}}M})$
	slow fading		$O(N)$	$O(N)$	$O(2^{N_{\text{bit}}M})$	$O(N) < \bullet < O(2^{N_{\text{bit}}M})$
	rapid fading		$O\left(\frac{N^3}{MN_{\text{bit}}}\right)$	$O(2^{N_{\text{bit}}M})$	$O(2^{N_{\text{bit}}M})$	$O\left(\frac{N^3}{MN_{\text{bit}}}\right) < \bullet < O(2^{N_{\text{bit}}M})$
Performance [b]	$\Delta_{P_{\text{out}}\|P_{\text{out}}=0.02}$ [dB]	(2×2)	12.1	7.9	5.1	5.1
	for block-fading	(4×2)	5.6	4.6	3.6	3.6
	$\Delta_{I\|\text{BER}=10^{-4}}$ [dB]	(2×2)	7.1	not evaluated	5.6	5.6
	for uncorrelated	(3×2)	5.4	not evaluated	3.9	3.9
	fading	(4×2)	4.5	not evaluated	4.2	4.2

[a] Values obtained from Table 2.7

[b] Distance from the theoretical capacity limit obtained from Figure 2.25 and Figure 2.26, respectively.

Chapter 3

Space-Time Equalization for the SDMA Uplink

This chapter commences with the evolution of the narrowband system model presented in Chapter 2 in order to accommodate wideband channel conditions. Furthermore, we highlight the challenges which a Space-Time Equalizer (STE) faces in a single-carrier uplink Space Division Multiple Access (SDMA) scenario. In contrast to Chapter 2, multiple MIMO users communicating in a dispersive fading environment are considered. In Section 3.3 of this chapter the theoretical channel capacity limits are presented for the system model considered, which constitute an extension of the narrowband capacity limits presented in Section 2.2. In Section 3.4 a set of different finite length STEs, namely the Minimum Mean Squared Error (MMSE), the Minimum Bit Error Rate (MBER) and the Bayesian STE as well as an approximate Maximum Likelihood (ML) STE are introduced. All the algorithms constitute a logical evolution of the narrowband detection algorithms discussed in Section 2.3. In Section 3.5 the STEs considered are extended to Decision Feedback (DF) assisted STEs. At the end of Section 3.6 the achievable performance of the various methods is evaluated both when assuming the availability of perfect channel knowledge as well as when using the estimated Channel State Information (CSI).

3.1 System Model

The system considered in this chapter consists of Q number of perfectly synchronized Mobile-Stations (MSs), each employing an M-element transmit antenna array and a Base-Station (BS) receiver, which has N number of Antenna Elements (AEs). The MSs' transmitters channel encode the input bit-stream at a code-rate R_c, interleave the encoded bits, modulate the signals and map them to the M different transmit AEs, as illustrated in Figure 3.1. The modulated symbols are transmitted to the BS over a frequency-selective

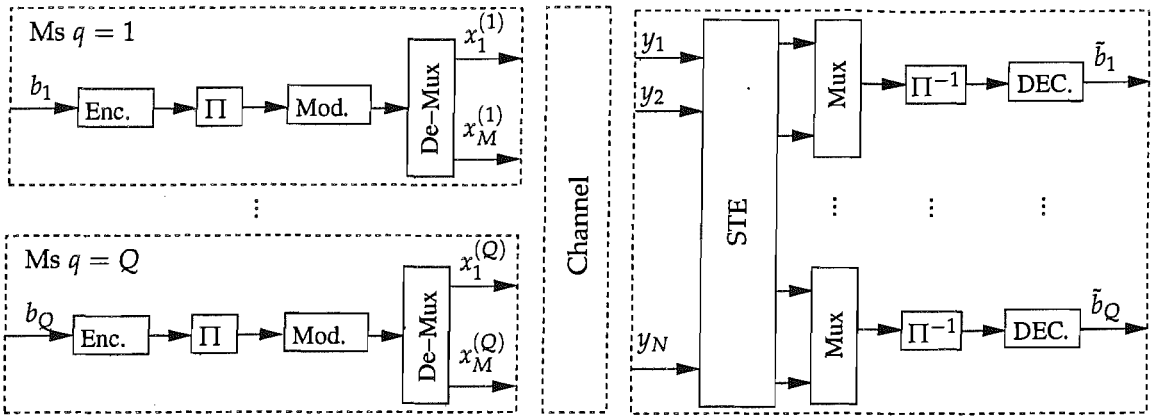


Figure 3.1: Baseband system model.

fading channel having a symbol-spaced Channel Impulse Response (CIR) characterized by the channel coefficients $h_{nm,l}^{(q)}$ as illustrated in Figure 3.1. The channel coefficient $h_{nm,l}^{(q)}$ represents the complex-valued channel coefficient of the l^{th} multipath component of the channel between the q^{th} MS's AE m and the n^{th} BS receiver AE. Given the transmitted symbol $x_m^{(q)}(k)$, which is associated with the q^{th} MS's transmit AE m, the output signal of the n^{th} AE of the BS receiver at time instant k can be written as

$$y_n(k) = \sum_{q=1}^{Q} \sum_{m=1}^{M} \sum_{l=0}^{L-1} h_{nm,l}^{(q)}(k) x_m^{(q)}(k-l) + \eta_n(k). \tag{3.1}$$

Furthermore, L is the number of symbol-spaced multipath components described by the CIR and $\eta_n(k)$ is the complex-valued Additive White Gaussian Noise (AWGN) having a variance of $\mathrm{E}\left[|\eta_n(k)|^2\right] = 2\sigma_{\mathrm{n}}^2$.

Assuming that a MS transmits the modulated symbols to the BS over a frequency-selective channel at a power $\sigma_{\mathrm{Tx},q}^2$, the resultant Signal to Noise Ratio (SNR) is given as

$$\mathrm{SNR}_q = \frac{\sigma_{\mathrm{Tx},q}^2 \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{l=0}^{L-1} \mathrm{E}\left[|h_{nm,l}^{(q)}(k)|^2\right]}{NM2\sigma_{\mathrm{n}}^2}. \tag{3.2}$$

In many cases we will define the CIR and the powers of the different users so that we satisfy

$$\mathrm{SNR} = \mathrm{SNR}_q \; \forall q. \tag{3.3}$$

For a given modulation scheme having $\mathcal{M}$ number of modulation levels, the associated energy per information bit to noise ratio E_{b}/N_0 can be written as

$$\frac{E_{\mathrm{b}}}{N_0} = \frac{\sigma_{\mathrm{TX},q}^2 \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{l=0}^{L-1} \mathrm{E}\left[|h_{nm,l}^{(q)}(k)|^2\right]}{R_{\mathrm{c}} \log_2(\mathcal{M}) MN2\sigma_{\mathrm{n}}^2}. \tag{3.4}$$

The number of bits per transmitted symbol is denoted as $N_{\mathrm{bit}} = \log_2(\mathcal{M})$. The Signal to Interference Ratio

(SIR_q) is defined as

$$\mathrm{SIR}_q = \frac{\sigma^2_{\mathrm{Tx},q} \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{l=0}^{L-1} \mathrm{E}\left[|h^{(q)}_{nm,l}(k)|^2\right]}{\sum_{i \neq q} \sigma^2_{\mathrm{Tx},i} \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{l=0}^{L-1} \mathrm{E}\left[|h^{(i)}_{nm,l}(k)|^2\right]}. \tag{3.5}$$

Under the assumption of perfectly synchronized transmitters the relation between the signal transmitted by the MSs' AEs and the channel's output for CIR tap l is described by a $(N \times QM)$-dimensional matrix $\boldsymbol{H}_l(k)$, where the $(n, (q-1)M+m)^{\text{th}}$ element of the matrix is given by $h^{(q)}_{nm,l}$. The channel output vector $\boldsymbol{y}(k)$ can now be expressed as

$$\boldsymbol{y}(k) = [\boldsymbol{H}_1(k) \ldots \boldsymbol{H}_L(k)] \left[\boldsymbol{x}(k)^T, \ldots, \boldsymbol{x}(k-L-1)^T\right]^T + \boldsymbol{\eta}(k), \tag{3.6}$$

where the column vector $\boldsymbol{x}(k) = \left[x^{(1)}_1(k), \ldots, x^{(1)}_M(k), \ldots, x^{(Q)}_1(k), \ldots, x^{(Q)}_M(k)\right]^T$ contains the symbols transmitted by the Q MSs' AEs and $\boldsymbol{\eta}(k) = [\eta_1(k), \ldots, \eta_N(k)]^T$. For the derivation of the algorithms presented in this chapter it is irrelevant whether the interference experienced by a signal is caused by an AE of the same user or by the signals transmitted by other users. In order to keep the notation simple, we introduce the index $\mathrm{m} = (q-1)M + m$, which ranges from 1 to $\mathfrak{M}$, where $\mathfrak{M} = QM$ is the total number of AEs in the system. We can now rewrite the signal vector containing all the $x_{\mathrm{m}}(k)$ components as $\boldsymbol{x}(k) = [x_1(k), \ldots, x_{\mathfrak{M}}(k)]^T$.

3.2 Problem Definition

In this section we provide a brief introduction to STEs designed for combating both Inter-Symbol Interference (ISI) as well as Co-Channel Interference (CCI). The scenario considered in this section is an evolution of the narrowband scenario discussed in Section 2.1.2, where only the effects of CCI were considered. The discussion of spatial and temporal processing provided in this section does not depend on the actual implementation of the receiver according to a specific cost-function, it rather treats the detection as a general classification problem.

Ultimately, the detection of all transmitters' signals when considering a wideband channel, is substantially more complex than that under the narrowband channel conditions discussed in Section 2.1.2. In Section 2.3 different receiver structures have been proposed in order to solve the narrowband detection problem. In order to solve the wideband detection problem, different receiver structures have been proposed [17]. Some of these structures treat the spatial and classic channel equalization as two separate, concatenated blocks. This two-stage system structure has the advantage of a relatively low complexity. However, when aiming for an optimum design, spatial and temporal equalization have to be performed in an amalgamated structure, as illustrated in Figure 3.2, where N_{F} is the feed-forward order of the STE and Δ is the decision delay at which

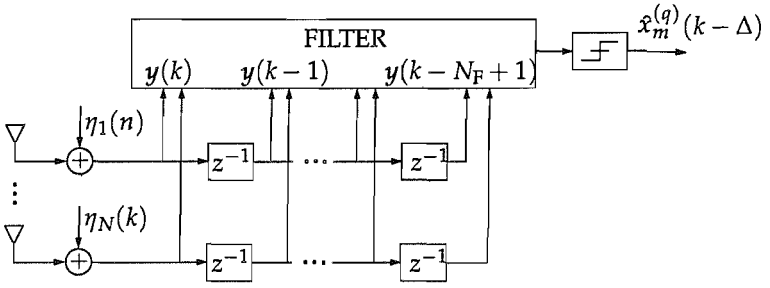


Figure 3.2: STE having a feed-forward order of N_F and employing N receive antennas.

the STE operates. However, the complexity of the detection problem becomes higher, since the number of possible channel output states is given by

$$N_s = (\log_2(\mathcal{M}))^{QM(L+N_F-1)} = N_{bit}^{QM(L+N_F-1)}. \tag{3.7}$$

Example 3.1: Space-Time Equalization

In order to illustrate the complexity associated with joint temporal and spatial equalization, we consider the scenario of two MSs communicating over a dispersive two-tap channel using a two-AE receiver at the BS. The CIRs of the users were chosen to be

$$\begin{aligned} \text{User 1}: \quad & h_{11}^{(1)}(z) = \sqrt{0.20} + \sqrt{0.80}z^{-1} \quad \text{and} \\ & h_{21}^{(1)}(z) = \sqrt{0.15} + \sqrt{0.85}z^{-1}; \\ \text{User 2}: \quad & h_{21}^{(2)}(z) = 0.5(\sqrt{0.10} + \sqrt{0.90}z^{-1}) \quad \text{and} \\ & h_{21}^{(2)}(z) = 0.5(\sqrt{0.35} + \sqrt{0.65}z^{-1}), \end{aligned} \tag{3.8}$$

implying that the signal of the second user is received at a 6 dB lower power than that of the first user. The legitimate channel output phasors can be calculated using Equation (3.1) taking into account all legitimate transmitted symbol sequences of length $QM(L+N_F-1) = 6$ and the CIRs given in Equation (3.8).

The illustration of the channel's output phasor constellation is no longer possible in a single graph, since it would become four dimensional even in this basic channel and user scenario. More explicitly, the four dimensions would correspond to $y_1(k)$, $y_1(k-1)$, $y_2(k)$ and $y_2(k-1)$. We have therefore projected the channel's output associated with user $q = 1$ into the two dimensional plane associated with AE $n = 1$ in Figure 3.3a, where the two coordinates are $y_1(k)$ and $y_1(k-1)$. Similarly, for the plane associated with AE $n = 2$ we plotted the channel's output in Figure 3.3b. Finally, the channel's output in the plane associated with the first STE tap of the first and second AE are portrayed in Figure 3.3c, which would correspond to a narrowband beamformer not benefiting from a tapped-delay-line .

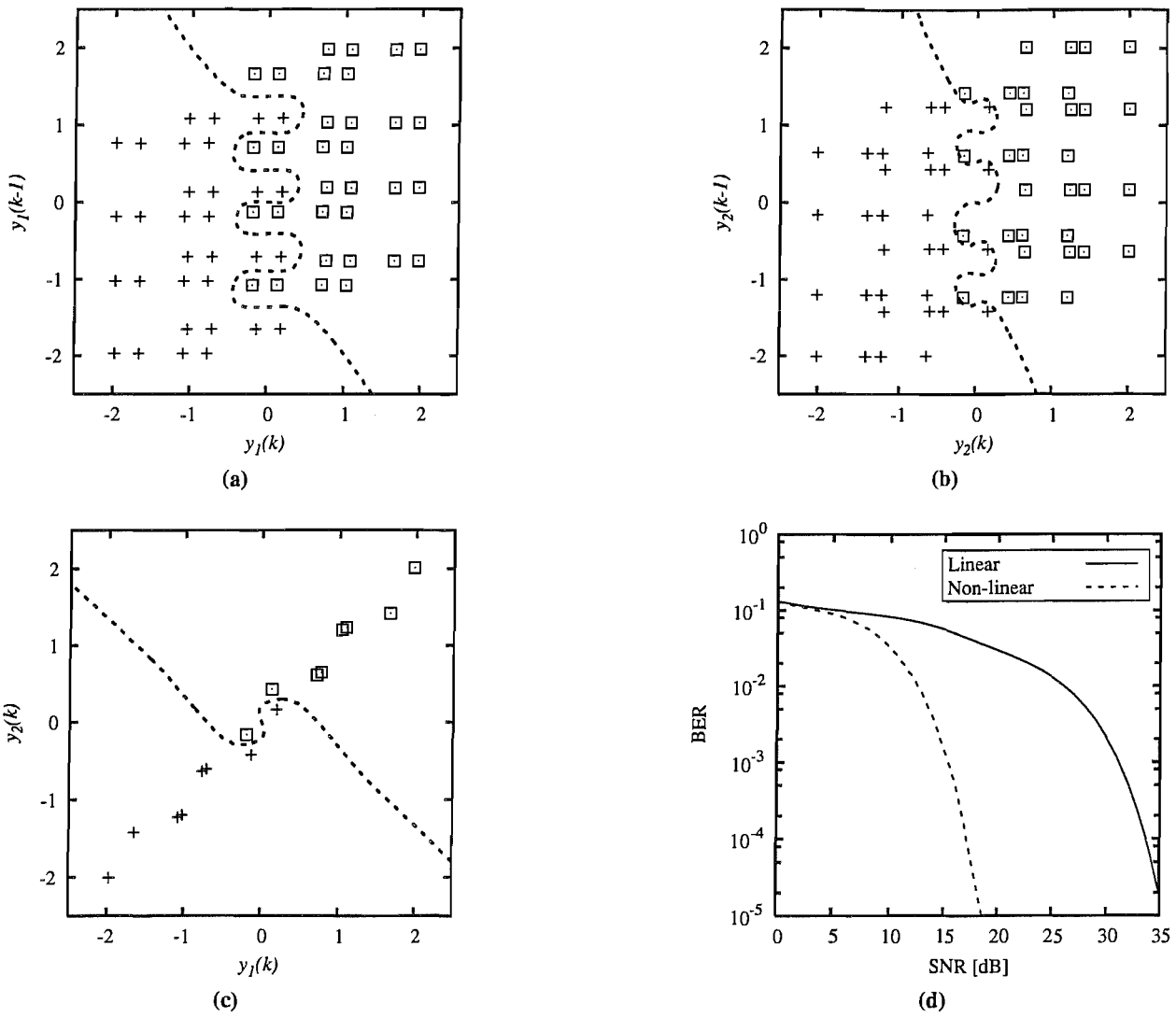


Figure 3.3: Projected channel output in the absence of noise (a-c) and the BER versus SNR performance (d) for BPSK modulated sources, where user $q = 1$ is assumed to be the desired user. The channel coefficients are given in Equation (3.8). The noise-free channel output is plotted for a decision delay $\Delta = 1$, where the legend $\square$ corresponds to a transmitted symbol of $\check{x}_1^{(1)}(k-\Delta) = +1$ and the legend $+$ to a transmitted symbol of $\check{x}_1^{(1)}(k-\Delta) = -1$.

In Figure 3.3a and Figure 3.3b all the $N_s = 2^{Q(L+N_F-1)} = 64$ number of different projected channel output states can be observed. It can be seen that the two subsets of the channel's output corresponding to a transmitted $+1$ and -1 of user $q = 1$ recorded for $\Delta = 1$ is linearly non-separable in all graphs but can be readily separated by the Bayesian decision boundary shown.

Note, that although that the channel output states shown in the projected planes are linearly non-separable, this does not imply non-separability in the higher-dimensional decision space. What can be stated however, is that even a simple Bayesian channel equalizer, which is associated with the decision boundary seen in Figure 3.3a and Figure 3.3b would be capable of detecting the signal of user 1. The same can be stated concerning the beamformer associated with Figure 3.3c, where the channel's output at the input of a hypothetical narrowband detector was portrayed.

Finally, it becomes clear from Figure 3.3d, which illustrates the attainable Bit Error Rate (BER) of user $q = 1$ versus SNR performance that the non-linear receiver is capable of separating the two users at a significantly lower SNR than the linear MMSE scheme. Since the BER curve associated with the MMSE STE does not exhibit an error-floor, it can be concluded that even though the channel output states associated with a narrowband beamformer seen in Figure 3.3a and Figure 3.3b is linearly non-separable, the 4-dimensional decision space considered by the STE is linearly separable by a hyperplane.

In this section we have provided a brief introduction to spatio-temporal equalization from a classification problem perspective. It was shown for this simple example, how the detection problem might become linearly non-separable and thus requires the employment of a non-linear receiver structure. In the next section first an information-theoretic benchmark is derived for the system model considered, which will serve as our ultimate performance bound for all the STE algorithms discussed.

3.3 Benchmarking - Channel Capacity and Information Rates

When analyzing the channel capacity of dispersive multi-user MIMO systems, where no CSI is available at the transmitter, we will interpret the multi-user system as a single-user system having a large, combined antenna array. The channel capacity for all users can then be expressed as

$$C(k) = \mathrm{E}\left[\max I\left(\boldsymbol{x},\boldsymbol{y}\right)\right], \tag{3.9}$$

where I is the mutual information between $\boldsymbol{x}$ and $\mathbf{y}$, respectively, as it was discussed in Section 2.2.

3.3.1 Gaussian Input Signal

If the MIMO system's input signals $x_m^{(q)}$ are assumed to be Gaussian distributed, which is reasonable well approximated for example by higher-order QAM signals, then $I(\boldsymbol{x},\boldsymbol{y})$ is maximized [110]. For a Gaussian distributed input signal $x_m^{(q)}$ the channel capacity for dispersive MIMO channels can be calculated using the Discrete Fourier Transform (DFT), as proposed in [124]. Physically this can be justified based on the following two points:

- The frequency-selective fading channel's impulse response is transformed into N_{DFT} number of parallel frequency-domain channels using the DFT, following a philosophy similar to that of multi-carrier based transceiver schemes [17]. The resultant channel may be viewed as a MIMO channel.
- The Gaussian distributed time-domain input signal remains Gaussian distributed in the frequency-domain, since a linear system's Gaussian input results in a Gaussian output. Therefore the MIMO capacity formulae derived for narrowband channels can be applied for each frequency-domain bin of index i, where $1 \leq i \leq N_{\text{DFT}}$.

The channel capacity for a given Frequency-Domain Channel Transfer Function (FDCTF) $\mathbf{F} = [\boldsymbol{F}_1, \ldots, \boldsymbol{F}_{N_{\text{DFT}}}]^T$ associated with the realization of the L-tap CIR $[\boldsymbol{H}_1(k) \ldots \boldsymbol{H}_L(k)]$ can then be written as

$$C(k) = \mathrm{E}\left[\frac{1}{N_{\text{DFT}}} \sum_{i=0}^{N_{\text{DFT}}-1} \log_2\left(\det\left(\boldsymbol{I}_N + \text{SNR}\, \boldsymbol{F}_i \boldsymbol{F}_i^H\right)\right)\right], \tag{3.10}$$

where

$$\boldsymbol{F}_i = \sum_{l=1}^{L} \boldsymbol{H}_l(k) \exp\left(-\mathrm{j}\frac{2\pi}{N_{\text{DFT}}} i(l-1)\right). \tag{3.11}$$

The channel capacity C of a Rayleigh fading channel rather than that of a single unfaded channel realization may be obtained on the basis of Equation (3.10) by taking the expectation value with respect to the FDCTF $\mathbf{F}$ yielding

$$C = \mathrm{E}[C(k)], \tag{3.12}$$

which may be evaluated using Monte Carlo simulations.

3.3.2 Discrete Input Signal

If the input signal of the channel is a discrete-valued signal, which can assume only a limited number of $\mathcal{M}$ different modulated signal values depending on the modulation scheme chosen, there exists no general closed form solutions for the achievable information rate for transmission over a frequency-selective MIMO channel. Extending the frequency-flat solution to the wideband case using the DFT based method described

in Section 3.3.1 cannot be accomplished in the same way, since for non-Gaussian signals of the channel input symbols the distribution will no longer be identical in the time-domain and in the frequency-domain. In other words, if the signal in the time-domain is BPSK modulated, the input signal of the sub-channels characterized by the FDCTF is generally not a real-valued binary wave-form but rather a complex-valued signal.

In order to characterize the capacity of a frequency-selective MIMO channel for non-Gaussian input signals, we therefore have to return to the definition of the mutual information outlined in Equation (2.13). The entropy of the noise required in Equation (2.13) can readily be calculated as [110]

$$H(\eta) = \log_2 \left(\det \left(\pi e 2 \sigma_n^2 \boldsymbol{I}_N \right) \right). \tag{3.13}$$

Considering an optimum detector which would take into account the total received symbol sequence for the detection of the desired symbols, the entropy of the received signal sequence $\boldsymbol{y}(K),\ldots,\boldsymbol{y}(1)$ can be defined as

$$H(\boldsymbol{y}(K),\ldots,\boldsymbol{y}(1)) = -\lim_{K\to\infty} \frac{1}{K} \mathrm{E}\left[\log_2 \left(p(\boldsymbol{y}(K),\ldots,\boldsymbol{y}(1))\right)\right], \tag{3.14}$$

where $p(\boldsymbol{y}(K),\ldots,\boldsymbol{y}(1))$ is the joint probability of the channel output, which may be calculated using the forward recursion of the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [125], as proposed in [126]. The exact procedure for initializing the BCJR algorithm for the system model considered is discussed in detail in Appendix C.

3.3.3 Results

In our further discourse we consider the overall system-capacity, i.e when all users' transmit AEs form a single combined antenna array, where all users transmit at identical data rates and generally experience identical average channels. If different throughputs were to be considered, it would be necessary to introduce so-called capacity regions [12], which are beyond the scope of this treatise.

Let us first assume the asymptotic case of an ergodic channel, where the time average of Equation (3.12) over different CIR realization of the channel is equal to the corresponding ensemble average. In other words, the random effects of the CIR tap fluctuations can be averaged out over time. For this specific scenario error-free transmission is theoretically possible, if the overall data rate R of all transmitters is smaller than or equal to the channel capacity C. For our later discourse, we also adopt the definition of the achievable throughput I of a system introduced in Section 2.2 as the data-rate supported by the channel under the constraint of a specific modulation scheme considered.

Figure 3.4 shows both the capacity and the achievable information rates of BPSK modulated signals for a

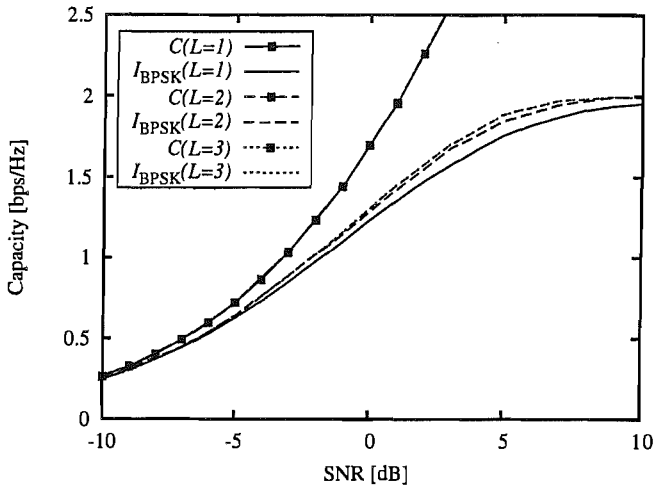
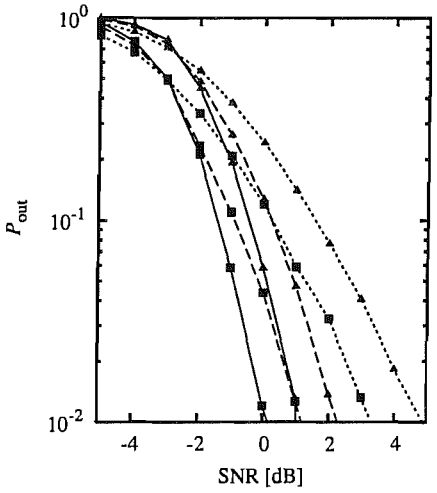


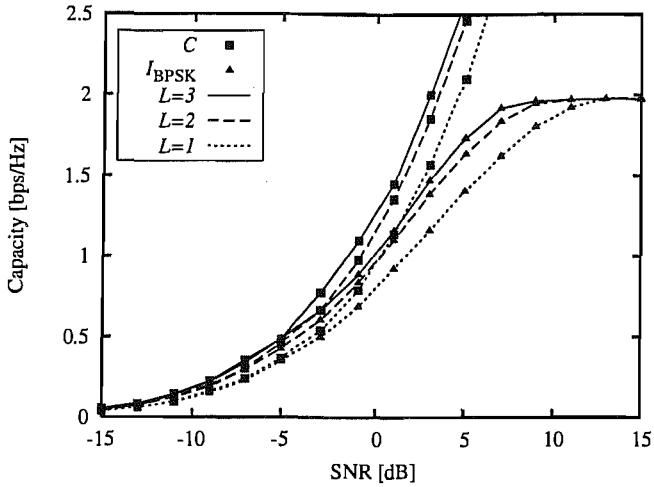
Figure 3.4: Channel capacity and achievable information rate of an ergodic channel for BPSK modulated input signals for a different number of equal-power independently Rayleigh faded CIR taps L, $N = 2$ receive AEs and a single user employing $M = 2$ transmit AEs. The curves associated with $C(L = 2)$ and $C(L = 3)$ are not visible since they are identical to the curve associated with $C(L = 1)$.

system supporting a single user employing a two-element transmit antenna array. The capacity of this system may be expected to be identical to that of a two-user system employing single antennas shifted by 3 dB SNR. The CIR was assumed to be an equal-power L-tap channel response between all antenna links and the BS was assumed to employ two receive AEs. The illustrated achievable throughput was plotted with the aid of Equation (2.13) as well as Equation (3.13) and Equation (3.14), respectively. It can be observed that for an increased number of CIR taps L the achievable throughput increases, but it has a tendency to saturate as the number of CIR taps increases. Only a moderately increased achievable throughput can be observed, when increasing the number of channel taps from $L = 2$ to $L = 3$, whereas it increases more substantially, when the number of CIR taps is increased from $L = 1$ to $L = 2$. The unconstrained capacity of Equation (3.12), which was generated on the basis of having a Gaussian channel input, does not increase for an increased number of CIR taps. To understand this phenomenon characterized by Equation (3.12) in more detail, the frequency-selective ergodic channel is viewed as N_{DFT} number of parallel ergodic narrowband channels. Since these sub-channels are ergodic, they are all associated with the same channel capacity and hence the expected value is equal to the narrowband channel capacity. This duality of the time and frequency observed for ergodic Gaussian channels has also been discussed in [14].

In analogy to the narrowband MIMO channels discussed in Chapter 2, the channel capacity and the achievable throughput of wideband channels can be characterized using the outage probability defined in Equation (2.19). In Figure 3.5a the outage probability is plotted as a function of the SNR for a different number of independently faded CIR taps L assuming a single-user MIMO system, $M = 2$ transmit AEs and $N = 2$ receive AEs. The outage probability was calculated by assuming an information rate $R = 1$ as outlined in detail in Section 2.2. It can be seen in Figure 3.5a that the outage probability associated with both the unconstrained capacity and the achievable information rate substantially decreases for an increasing number of



(a) Outage probability for a transmission rate $R = 1$ bps/Hz .



(b) Outage capacity associated with an outage probability $P_{out} = 0.1$.

Figure 3.5: Outage probability and outage capacity evaluated as outlined in the text, for a single user system using $M = 2$ transmit AEs and $N = 2$ receive AEs. The channel was an $L = 1$ and 2-tap equal-power channel with independently faded CIR matrix elements.

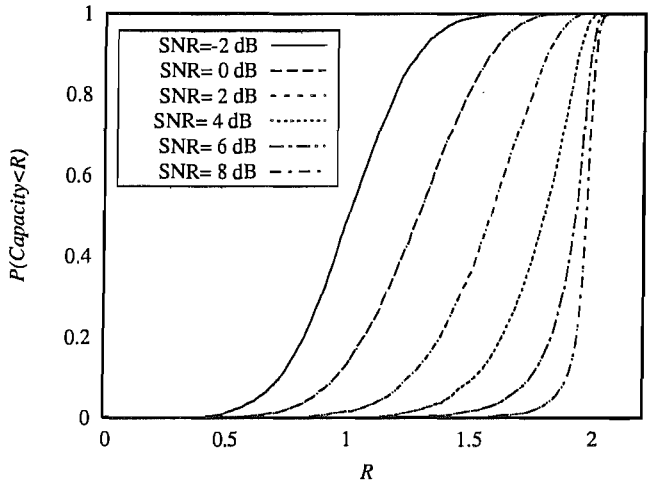


Figure 3.6: CDF for BPSK modulated input signals, a single user employing $M = 2$ transmit AEs, two receive AEs at different SNR. The number of independently Rayleigh-faded CIR taps was $L = 2$.

CIR taps. There is also a significant performance gap between the outage probability recorded for Gaussian and BPSK modulated inputs.

By fixing the outage probability to $P_{out} = 0.1$ in Figure 3.5a we can obtain the outage capacity curves shown in Figure 3.5b again following the procedure outlined in Section 2.2. We sometimes refer to this value as the 10%-outage-capacity. When comparing the 10%-outage-capacity results of Figure 3.5b to the ergodic capacity plotted in Figure 3.4, it can be observed that the capacity gain provided by having a richer independently faded multipath channel is greater for the non-ergodic channel. Furthermore, the capacity of the non-ergodic channel increases for a larger number of channel taps, since the receiver benefits from the independently faded taps of the dispersive channel.

A third way of characterizing the block-constant channel for the scenario considered is shown in Figure 3.6,

which shows its Cumulative Distribution Function (CDF) for different SNRs when considering $L = 2$ CIR taps. The CDF was obtained as outlined in Section 2.2.

3.4 Finite Length Space-Time Equalizers

In this section different finite length STEs, namely the MMSE, the MBER, the Bayesian STE and a novel reduced complexity ML STE will be presented. All the algorithms are derived on the basis of the system model in matrix notation, which is introduced in the following.

Considering a finite-length STE having a feed-forward order of N_F, the CIR super-matrix $\mathbf{H}$, which represents the total system is obtained by concatenating the $(N \times QM)$=$(N \times \mathfrak{M})$-dimensional matrices $\boldsymbol{H}_l$, which have been defined in the context of Equation (3.6), yielding:

$$\mathbf{H}(k) = \begin{bmatrix} \boldsymbol{H}_1(k) & \cdots & \boldsymbol{H}_L(k) & 0 & \cdots & 0 \\ 0 & \boldsymbol{H}_1(k-1) & \cdots & \boldsymbol{H}_L(k-1) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \boldsymbol{H}_1(k-N_F+1) & \cdots & \boldsymbol{H}_L(k-N_F+1) \end{bmatrix},$$

Let us denote the N-element channel output vector of the MIMO system as $\boldsymbol{y}(k)$. Then the channel output super-vector $\mathbf{y}(k) = [\boldsymbol{y}(k)^T, \ldots, \boldsymbol{y}(k-N_F+1)^T]^T$ can be expressed as

$$\begin{aligned} \mathbf{y}(k) &= \mathbf{H}(k)\left[\boldsymbol{x}(k)^T, \ldots, \boldsymbol{x}(k-L-N_F+2)^T\right]^T \\ &\quad + \left[\boldsymbol{\eta}(k)^T, \ldots, \boldsymbol{\eta}(k-N_F+1)^T\right]^T \\ &= \mathbf{H}(k)\mathbf{x}(k) + \boldsymbol{\eta}(k) \\ &= \check{\mathbf{y}}(k) + \boldsymbol{\eta}(k), \end{aligned} \tag{3.15}$$

where $\boldsymbol{x}(k) = [x_1(k), \ldots, x_{\mathfrak{M}}(k)]^T$ is a column vector containing the symbols transmitted by the $\mathfrak{M} = QM$ number of AEs present in the system and $\boldsymbol{\eta}(k) = [\eta_1(k), \ldots, \eta_N(k)]^T$ is the complex-valued Additive White Gaussian Noise (AWGN) vector having a variance of $\mathrm{E}\left[|\eta_n(k)|^2\right] = 2\sigma_n^2$.

Equivalently, the system may be described using the binary real-valued system representation introduced in

Section 2.1.1 which yields

$$\begin{bmatrix} \Re(\mathbf{y}(k)) \\ \Im(\mathbf{y}(k)) \end{bmatrix} = \begin{bmatrix} \Re(\mathbf{H}(k)\boldsymbol{Q}) \\ \Im(\mathbf{H}(k)\boldsymbol{Q}) \end{bmatrix} \begin{bmatrix} \underline{x}(k) \\ \vdots \\ \underline{x}(k-L-N_F+2) \end{bmatrix} + \begin{bmatrix} \underline{\eta}(k) \\ \vdots \\ \underline{\eta}(k-N_F+1) \end{bmatrix} \tag{3.16}$$

$$\underline{\mathbf{y}}(k) = \underline{\mathbf{H}}(k)\underline{\mathbf{x}}(k) + \underline{\eta}(k) \tag{3.17}$$

$$= \underline{\check{\mathbf{y}}}(k) + \underline{\eta}(k), \tag{3.18}$$

where the transformation matrix is given as

$$\mathbf{Q} = \boldsymbol{I}_{\mathfrak{M}(N_F+L-1)} \otimes \boldsymbol{q} \tag{3.19}$$

and where the row-vector $\boldsymbol{q}$, the column-vectors $\underline{x}$ as well as $\underline{\eta}$ have been defined in the context of Equation (2.7).

When referring to a delayed vector such as for example $\boldsymbol{x}(k-\Delta)$, this is indicated by $\boldsymbol{x}_{\Delta+1}$, which suggests that $\boldsymbol{x}_{\Delta+1}$ is the $(\Delta+1)^{\text{th}}$ sub-vector of the super-vector $\mathbf{x}(k)$ defined in Equation (3.15). The vector notation used may be summarized as follows:

- $\mathbf{x}(k) = \mathbf{x}$;
- $\boldsymbol{x}(k-i) = \boldsymbol{x}_{i+1}$ which stands for the $(i+1)^{\text{th}}$ sub-vector of $\mathbf{x}$;
- $x(k-i) = x_{i+1}$ which stands for the $(i+1)^{\text{th}}$ element of $\mathbf{x}$.

For notational simplicity the time-index k will be dropped in the remainder of this chapter, where this is possible without ambiguity.

3.4.1 Minimum Mean Squared Error STE

The linear MMSE STE aims for equalizing the channel's phasor output states as for example illustrated in Figure 3.3, by multiplying the channel's output with a weight matrix $\boldsymbol{W}$. Each column of the weight matrix describes a linear hyperplane, which is designed for separating the channel's output phasors according to their associated symbol value. In mathematical terms, this operation can be expressed as

$$\hat{\boldsymbol{x}} = \boldsymbol{W}^H \mathbf{y}, \tag{3.20}$$

where $\boldsymbol{W}$ is the $(NN_F \times \mathfrak{M})$-dimensional complex-valued array weight matrix, where the different columns of $\boldsymbol{W}$ are associated with the different transmitters' symbols and $\hat{\boldsymbol{x}}$ is the $(\mathfrak{M} \times 1)$-dimensional filter output

vector generated for the detection of the associated transmitted $\mathfrak{M}$-symbol signal vector. The filter output of the corresponding corresponding real-valued system model of Equation (3.18) is given as

$$\underline{\hat{x}} = \underline{W}^T \underline{y}, \tag{3.21}$$

where $\underline{W}$ is the $(2NN_F \times N_{bit}\mathfrak{M})$-dimensional real-valued weight matrix, where the different columns of $\underline{W}$ are associated with the different transmitters conveying BPSK symbols, while $\underline{\hat{x}}$ is the $(N_{bit}\mathfrak{M} \times 1)$-dimensional filter output vector for the detection of the associated transmitted BPSK signal vector.

Assuming that the STE operates with a decision lag of Δ, which is often also referred to as the decision delay, the squared error at the filter's output can then be written as

$$\begin{aligned} \epsilon^2 &= ||x_{\Delta+1} - W^H \mathbf{y}||^2 = ||x_{\Delta+1} - \hat{x}||^2 && (3.22) \\ &= ||\underline{x}_{\Delta+1} - \underline{W}^H \underline{\mathbf{y}}||^2 = ||\underline{x}_{\Delta+1} - \underline{\hat{x}}||^2. && (3.23) \end{aligned}$$

The corresponding MMSE STE weight optimization problem is now defined as

$$W = \arg \min_{W \in \mathbb{C}^{NN_F \times \mathfrak{M}}} \mathrm{E}\left[||\hat{x}_{\Delta+1} - W^H \mathbf{y}||^2\right] \tag{3.24}$$

or

$$\underline{W} = \arg \min_{\underline{W} \in \mathbb{R}^{2NN_F \times \log_2(\mathcal{M})\mathfrak{M}}} \mathrm{E}\left[||\underline{\hat{x}}_{\Delta+1} - \underline{W}^T \underline{\hat{y}}||^2\right], \tag{3.25}$$

respectively. Following the derivation of the array weight vector for the narrowband channel, the MMSE receiver can be characterized as

$$W = \left(\mathbf{H}\mathbf{H}^H + 2\sigma_n^2 \mathbf{I}_{NN_F}\right)^{-1} [\boldsymbol{h}_{\mathfrak{M}\Delta+1}, \ldots, \boldsymbol{h}_{\mathfrak{M}(\Delta+1)}]^T, \tag{3.26}$$

where $\boldsymbol{h}_i$ is a vector of scalars corresponding to the i^{th} columns of the super-matrix $\mathbf{H}$. The STE array weight vector for the corresponding real-valued BPSK system is given as

$$\underline{W} = \left(\underline{\mathbf{H}}\underline{\mathbf{H}}^T + \sigma_n^2 \boldsymbol{I}_{2NN_F}\right)^{-1} [\boldsymbol{h}_{N_{bit}\mathfrak{M}\Delta+1}, \ldots, \boldsymbol{h}_{N_{bit}\mathfrak{M}(\Delta+1)}]^T. \tag{3.27}$$

For BPSK modulated transmitters, the detected symbol vector is given as

$$\tilde{x}_{\Delta+1} = \mathtt{sign}\left(\Re\left(W^H \mathbf{y}\right)\right) = \underline{\tilde{x}}_{\Delta+1} = \mathtt{sign}\left(\underline{W}^T \underline{\mathbf{y}}\right) \tag{3.28}$$

Table 3.1: The normalized computational complexity (real-valued additions and multiplications) of the MMSE STE when considering the detection of a single transmitted bit.

Weight vector calculation of Equation (3.27)	$(2N_{\text{bit}}\mathfrak{M}(N_{\text{F}}+L-1)-1)4N^2N_{\text{F}}^2/(N_{\text{bit}}\mathfrak{M})$ $+2NN_{\text{F}}/(N_{\text{bit}}\mathfrak{M})$ $+(4NN_{\text{F}}-1)2NN_{\text{F}}+O(N^3N_{\text{F}}^3/(N_{\text{bit}}\mathfrak{M}))$
Detection according to Equation (3.30)	$4NN_{\text{F}}-1$

and for 4QAM modulated sources, the detected symbol is given as

$$\tilde{x}_{\Delta+1} = \frac{1}{\sqrt{2}}\text{sign}\left(\Re\left(\boldsymbol{W}^H\mathbf{y}\right)\right) + \frac{\text{j}}{\sqrt{2}}\text{sign}\left(\Im\left(\boldsymbol{W}^H\mathbf{y}\right)\right), \tag{3.29}$$

which is identical to

$$\underline{\tilde{x}}_{\Delta+1} = \text{sign}\left(\underline{\boldsymbol{W}}^T\underline{\mathbf{y}}\right). \tag{3.30}$$

3.4.1.1 Computational Complexity of the Real-Valued MMSE STE

For the evaluation of the computational complexity associated with the MMSE based STE, the real-valued system model of Equation (3.18) is considered similar to the complexity evaluation of the real-valued MMSE MIMO detector of Section 2.3.1.1. This will facilitate its fair comparison to other algorithms, since for real-valued operations, the complexity imposed by a multiplication and an addition can be considered identical. The complexity of a linear STE might be divided into two parts, namely the calculation of the weight vector and the detection of the symbol. Upon introducing the variables $A = N_{\text{bit}}\mathfrak{M}\cdot(N_{\text{F}}+L-1)$ and $B = 2NN_{\text{F}}$, and assuming that the array weight vector is calculated using direct matrix inversion rather than an iterative algorithm, furthermore, assuming that the CIR matrix is known to the receiver, the associated complexity is calculated as follows:

- The calculation of the covariance matrix according to Equation (3.27) invokes $(2A-1)B^2+B$ real-valued additions and multiplications.
- The matrix inversion of Equation 3.27 may be carried out at a computational cost, which is proportional to the complexity order of $O(B^3)$.
- The matrix multiplication seen in Equation (3.27) and required for the final calculation of the weight matrix can be preformed at a cost of $(2B-1)B\mathfrak{M}N_{\text{bit}}$ number of real-valued multiplications and additions.

The total number of operations constituted by the number of multiplications plus the number of additions can be obtained by evaluating A as well as B. The resultant number of operations required for the calculation

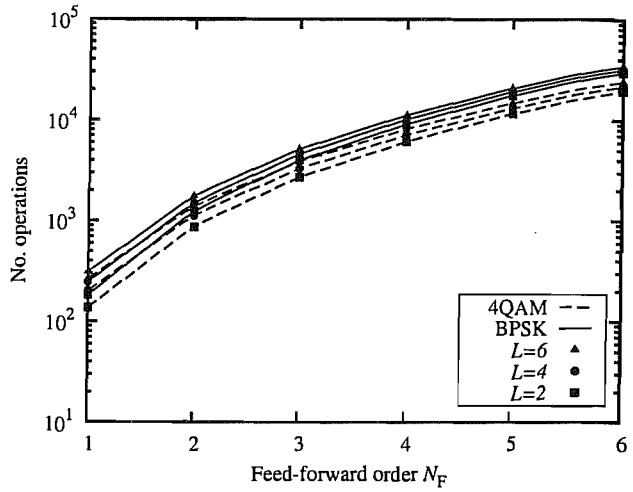


Figure 3.7: The normalized computational complexity required for evaluating the MMSE array weights as a function of the feed-forward order N_F for different modulation schemes and a different number of CIR taps L, when considering the detection of a single transmitted bit calculated according to Table 3.1. The proportionality factor was set to 3 and the total number of transmit AEs was chosen to be $\mathfrak{M} = 2$, while the number of receive AEs was $N = 2$.

of the weight-matrix $\underline{W}$ in Equation (3.27) is given as

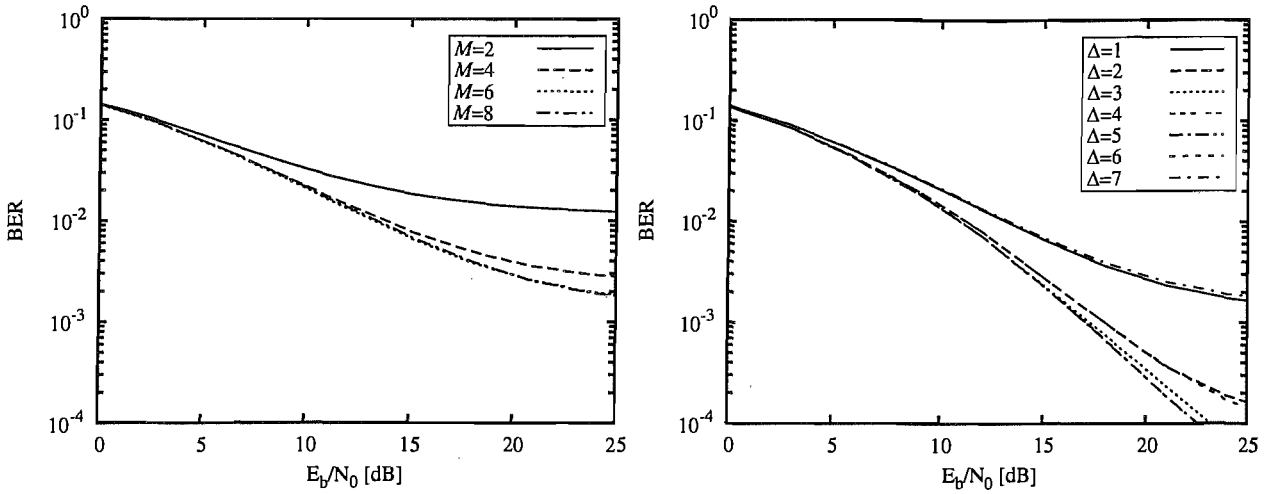
$$C = (2A-1)B^2 + B + (2B-1)B\mathfrak{M}N_{\text{bit}} + O(B^3) \tag{3.31}$$

The complexity imposed by the detection of a single transmitted bit only involves B number of multiplications and $(B-1)$ number of additions. In conclusion, the total complexity is summarized in Table 3.1.

In Figure 3.7 the normalized computational complexity defined as the number of real-valued additions plus multiplications required for the evaluation of the MMSE weight vector as summarized in Table 3.1 is illustrated as a function of the feed-forward order N_F for both BPSK modulated as well as 4QAM signals, when considering $L = 2, 4$ and 6 CIR taps for a system characterized by $N = 2$ receive AEs as well as $\mathfrak{M} = 2$ transmit AEs. The proportionality factor associated with the matrix inversion was set to 3. It can be seen in Figure 3.7 that the complexity associated with the MMSE STE designed for 4QAM signals is lower than that for BPSK signals. The same behaviour was observed in Section 2.3.1 for the narrowband MMSE detector.

3.4.1.2 Performance of the MMSE STE

Let us first consider a simple single-user SDM system, which is characterized by $M = 2$ transmit AEs at the MS, a BS employing $N = 2$ receive AEs with an equalizer characterized by the feed-forward order N_F and a decision delay Δ. The channel was assumed to be a symbol-spaced two-path equal-power block-fading channel between all antenna links and the signals were assumed to be BPSK modulated. All channel CIR components were assumed to be subject to independent Rayleigh fading. Figure 3.8a shows the BER



(a) Raw BER versus E_b/N_0 for different feed-forward order N_F. The decision delay was chosen to $\Delta = N_F - 1$.

(b) Raw BER versus E_b/N_0 for feed-forward order $N_F = 8$ and different decision delays Δ.

Figure 3.8: BER versus E_b/N_0 performance of the MMSE STE for a scenario supporting $Q = 1$ user employing $M = 2$ transmit AEs, a two-path equal-power independent Rayleigh channel and having perfect channel knowledge. The BS employed a two-element antenna array.

versus E_b/N_0 performance for the uncoded system and for a set of different feed-forward orders N_F as specified in the graph. The decision delay was chosen to be $\Delta = N_F - 1$. It can be seen that for a larger STE feed-forward order N_F, the error floor is reduced which may be explained as follows: By increasing the feed-forward order of the STE the dimensionality of the decision space is increased. In the higher-dimensional decision space, phasor constellations which have been linearly non-separable in the lower-dimensional decision space associated with a small feed-forward order might become linearly separable in the higher-dimensional decision space as demonstrated in [118] for classic channel equalization. The linear separability is defined with respect to the hyper-plane described by the MMSE weights. Hyperplanes obtained on the basis of other criteria than the MMSE might be capable of achieving a better separability. However, the beneficial effect of increasing the feed-forward order of the STE saturates and increasing the order beyond $N_F = 6$ hardly gives any further improvement.[1]

If, however, the decision delay Δ is chosen differently and the feed-forward order is fixed to $N_F = 8$, it can be seen in Figure 3.8b that the BER floor is reduced further, provided that the decision delay is chosen appropriately. In our system these appropriate decision delays would be $\Delta = 3, 4$ or 5. For a larger or a smaller decision delay, the performance is degraded. The beneficial effect of optimally choosing the decision delay was also discussed in [128], where a specific algorithm was proposed for finding the optimal decision delay. However, the discussions of [128] only considered the conventional single-user, single-antenna based channel equalization case and did not take into account Rayleigh fading channels. It is plausible, however, that in the case of a static channel an optimally chosen decision delay results in a better performance than

[1]Note that in [127] a set of algorithms was discussed, which allow the adaptive choice of the feed-forward order depending on the channel.

all the other delays. It is interesting that in the case of the independent Rayleigh fading channel considered in Figure 3.8 a similar effect can be seen. Again, that certain delay choices result in a better average BER performance than others. For block-fading independent Rayleigh CIRs, one might consider the choice of an adaptive decision delay based on the method proposed in [128]. The complexity imposed by finding the optimum decision delay can however hardly be justified, because in Section 3.5.1 we will extend the MMSE algorithm by a decision feedback structure which solves the problem of choosing the correct decision delay as a by-product.

3.4.2 Minimum Bit Error Rate STE

Rather than considering the MSE as the optimization criterion for the determination of the STE's weights, we aim in this section for minimizing the BER directly in analogy to Section 2.3.2, where the direct minimization of the BER was considered for narrowband MIMO systems. Directly minimizing the BER will always result in a lower or identical BER than the MMSE approach. More specifically the two approaches become equivalent, if the conditional PDF of the linear STE is Gaussian distributed irrespective of the weights. The MBER solution may be defined as

$$\boldsymbol{W} = \arg \min_{\boldsymbol{W} \in \mathbb{C}^{NN_F \times \mathfrak{M}}} P_e(\boldsymbol{W}) \tag{3.32}$$

or equivalently as

$$\underline{\boldsymbol{W}} = \arg \min_{\boldsymbol{W} \in \mathbb{R}^{2NN_F \times N_{bit}\mathfrak{M}}} P_e(\underline{\boldsymbol{W}}). \tag{3.33}$$

The first step to derive the STE weights, which minimize the BER, is again to find an expression for the BER at the output of a linear receiver by integrating over the conditional PDF of the STE's output. For a STE of feed-forward order N_F, a CIR of length L and a system employing a total of $\mathfrak{M}$ transmit AEs the transmitted symbol vector may assume $N_s = \mathcal{M}^{\mathfrak{M}(L+N_F-1)}$ different values, where $\mathcal{M}$ is the number of modulated phasor constellation points.

In the remainder of this section, we will discuss the MBER STE only on the basis of the real-valued binary system model of Equation (3.18). The complex-valued MBER STE for 4QAM modulated signals can readily be obtained by following the procedure outlined in Section 2.3.2.2 for the complex-valued MBER MIMO detector discussed in the context of narrowband channels.

The N_s number of legitimate transmitted symbol vectors $\underline{\mathbf{x}}$ form the set

$$\underline{\mathcal{X}} = \{\underline{\check{\mathbf{x}}}^{(i)},\ 1 \leq i \leq N_s\}. \tag{3.34}$$

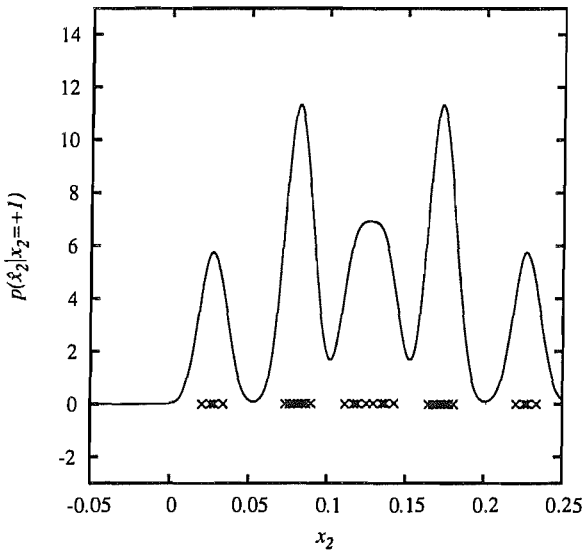


Figure 3.9: Conditional PDF of the filter output of the MMSE-based STE for the example of Section 3.2 for user two at SNR=34 dB

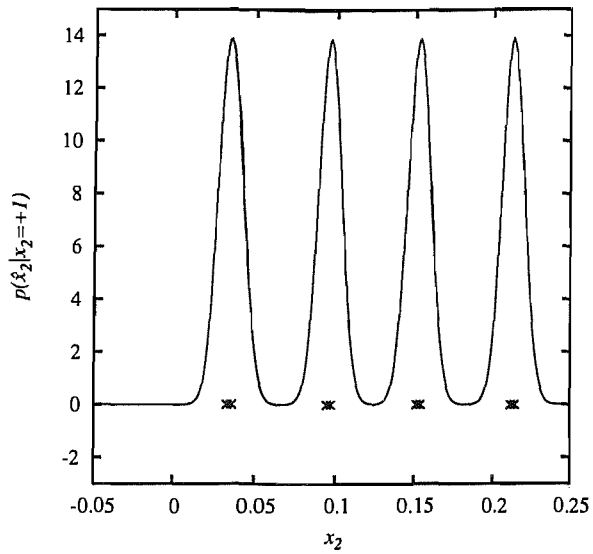


Figure 3.10: Conditional PDF of the filter output of the MBER-based STE for the example of Section 3.2 for user two at SNR=34 dB

The set of legitimate channel output vectors associated with the set $\underline{\mathcal{X}}$ may be defined as

$$\underline{\mathcal{Y}} = \{\underline{\check{y}}^{(i)} = \underline{\mathbf{H}}\underline{\check{x}}^{(i)},\ 1 \leq i \leq N_s\}, \tag{3.35}$$

which can be partitioned into the two subsets corresponding to the transmitted symbols by:

$$\underline{\mathcal{Y}}_m^{\pm} = \{\underline{\check{y}}^{(i)} \in \underline{\mathcal{Y}} \mid \check{x}_{\mathfrak{M}\Delta+m}^{(i)} = \pm 1\}. \tag{3.36}$$

In a similar manner, the noise-free part of the STE's output signal may assume only values of the set

$$\underline{\hat{\mathcal{X}}}_m = \{\check{\underline{x}}_m^{(i)} = \underline{w}_m^T \underline{\check{y}}^{(i)},\ 1 \leq i \leq N_s\}, \tag{3.37}$$

where $\underline{w}_m$ is the m^{th} column of the STE's weight matrix $\underline{W}$. This set, again, can be partitioned into two subsets defined as

$$\underline{\hat{\mathcal{X}}}_m^{\pm} = \{\check{\underline{x}}_m^{(i)} \in \underline{\hat{\mathcal{X}}}_m \mid \check{x}_{\mathfrak{M}\Delta+m}^{(i)} = \pm 1\}. \tag{3.38}$$

The conditional PDF of $\hat{\underline{x}}_m(k)$, given $\underline{x}_m(k-\Delta) = \underline{x}_{\mathfrak{M}\Delta+m}$ can now be expressed as

$$p(\hat{\underline{x}}_m | \underline{x}_{\mathfrak{M}\Delta+m} = +1) = \frac{1}{N_{sb}\sqrt{2\pi\sigma_n^2 \underline{w}_m^T \underline{w}_m}} \sum_{\check{\underline{x}}_m^{(i)} \in \underline{\hat{\mathcal{X}}}_m^+} \exp\left(-\frac{(\hat{\underline{x}}_m - \check{\underline{x}}_m^{(i)})^2}{2\sigma_n^2 \underline{w}_m^T \underline{w}_m}\right), \tag{3.39}$$

where $N_{sb} = \frac{N_s}{2}$ is the number of possible filter output states without noise in the set $\underline{\hat{\mathcal{X}}}_m^+$. It can be seen, that the conditional PDF of the filter output is the weighted sum of the Gaussian functions over all STE output sates in $\underline{\hat{\mathcal{X}}}_m^+$.

Example 3.2: STE MBER Conditional PDF

In Figure 3.10 the conditional PDF of the STE's output given by Equation (3.39) is plotted for the second user of Example 3.1 considered in Section 3.2 at SNR=34 dB. It can be seen from Figure 3.9 and Figure 3.10 that the filter output is not Gaussian distributed regardless whether the MMSE solution or the MBER solution is considered. It also can be observed by comparing Figure 3.9 and Figure 3.10 that the minimum distance between the legitimate STE output states indicated as small crosses and the decision boundary at $\underline{x}_2 = 0$ is slightly larger for the MBER solution. This will result in a lower BER for the MBER solution.

Following the derivations of Section 2.3.2, the gradient of the BER associated with the real-valued binary system model is given as

$$\nabla P_e(\underline{\boldsymbol{w}}_m) = C \sum_{\underline{\check{x}}_m^{(i)} \in \underline{\mathcal{X}}_m^+} \exp\left(-\frac{\left(\underline{\check{x}}_m^{(i)}\right)^2}{2\sigma_n^2 \underline{\boldsymbol{w}}_m^H \underline{\boldsymbol{w}}_m}\right) \text{sign}\left(\underline{\check{x}}_{\mathfrak{M}\Delta+m}^{(i)}\right) \left(\frac{\underline{\check{x}}_m^{(i)} \underline{\boldsymbol{w}}_m}{\underline{\boldsymbol{w}}_m^H \underline{\boldsymbol{w}}_m} - \underline{\check{y}}_m^{(i)}\right), \tag{3.40}$$

where we have

$$C = \frac{1}{N_{sb}\sqrt{2\pi\sigma_n^2 \underline{\boldsymbol{w}}_m^H \underline{\boldsymbol{w}}_m}}. \tag{3.41}$$

Assuming a normalized weight vector, where $\underline{\boldsymbol{w}}_m^T \underline{\boldsymbol{w}}_m = 1$ this is further simplified to

$$\nabla P_e(\underline{\boldsymbol{w}}_m) = C \sum_{\underline{\check{x}}_m^{(i)} \in \underline{\mathcal{X}}_m^+} \exp\left(-\frac{\left(\underline{\check{x}}_m^{(i)}\right)^2}{2\sigma_n^2}\right) \left(\underline{\check{x}}_m^{(i)} \underline{\boldsymbol{w}}_m - \underline{\check{y}}_m^{(i)}\right),$$

where we have

$$C = \frac{1}{N_{sb}\sqrt{2\pi\sigma_n^2}}. \tag{3.42}$$

The MBER weights again can be found using the simplified conjugate gradient algorithm of Appendix B in order to calculate the optimum point of $\nabla P_e(\underline{\boldsymbol{w}}_m) = 0$ commencing for example from the MMSE STE weight solution.

3.4.2.1 Computational Complexity of the Real-Valued MBER STE

The computational cost of the MBER algorithm evaluated on the basis of the real-valued binary system model and under the assumption that the CIR matrix is known to the receiver can again be divided into two parts similarly to the complexity evaluation of the MMSE algorithm provided in Section 3.4.1.1 .

Let us first quantify the computational complexity required for the evaluation of the gradient of the BER surface for a normalized weight vector. The evaluation of the expression in the sum of Equation (3.42)

Table 3.2: The normalized computational complexity (real-valued additions and multiplications) of the MBER STE when considering the detection of a single transmitted bit.

Weight vector calculation	$N_{\text{grad}}N_{\text{sb}}(2+16NN_{\text{F}})+$ $(2N_{\text{sb}}N_{\text{F}}N(2N_{\text{bit}}\mathfrak{M}(N_{\text{F}}+L-1)-1))/(\mathfrak{M}N_{\text{bit}})$
Detection	$4NN_{\text{F}}-1$

requires $3+3\cdot 2NN_{\text{F}} = 3+6NN_{\text{F}}$ number of operations and one evaluation of the exp function. This expression has to be evaluated N_{sb} times and summed up, which requires further $(N_{\text{sb}}-1)$ additions. The normalization of the weight vector contributes $2NN_{\text{F}}+(2NN_{\text{F}}-1)+2NN_{\text{F}} = 6NN_{\text{F}}-1$ number of operations. Additionally, the receiver has to determine the legitimate channel output states defined Equation (3.36), which requires $2N_{\text{sb}}N_{\text{F}}N(2N_{\text{bit}}\mathfrak{M}(N_{\text{F}}+L-1)-1)$. The evaluation of the legitimate filter outputs imposes $N_{\text{sb}}(4N-1)$. The computational complexity difference between the detection of BPSK and 4QAM signals is associated with the factor N_{sb} operations.

When considering the complexity of the conjugate gradient algorithm presented in Equation (B.1) in Appendix B, then it can be seen that the computational complexity associated with the evaluation with the calculation of the gradient compromises the main contribution. We therefore approximate the complexity required for the calculation of the weight vector associated with a single bit as $N_{\text{grad}}C^{\nabla P_{\text{e}}}$ where N_{grad} is the number of iterations of the conjugate gradient algorithm and

$$C^{\nabla P_{\text{e}}} \approx N_{\text{sb}}(2+16NN_{\text{F}}) \tag{3.43}$$

is the computational complexity required for the calculation of the BER surface gradient for a given weight vector. The total complexity required for the detection of one bit is summarized in Table 3.2 where the evaluation of the exp function has been neglected.

In Figure 3.11 the normalized computational complexity defined as the number of real-valued additions plus multiplications required for the evaluation of the MBER weight vector as summarized in Table 3.2 is illustrated as a function of the feed-forward order N_{F} for both BPSK modulated as well as 4QAM signals, when considering $L = 2, 4$ and 6 CIR taps for a system characterized by $N = 2$ receive AEs as well as $\mathfrak{M} = 2$ transmit AEs. It can be observed in Figure 3.11 that in contrast to the computational complexity imposed by the MMSE detector illustrated in Figure 3.7, the complexity required for determining the MBER STE weights is substantially higher. The high computationally complexity associated with determining the MBER STE weights is owing to the fact that the number of legitimate channel output states is proportional to $O(N_{\text{sb}})$ and thus grows exponentially with the number of transmitters, with the length of the CIR and with the feed-forward order of the STE. However, once the array weight matrix has been determined, the MMSE and the MBER STE impose the same complexity when detecting a single bit.

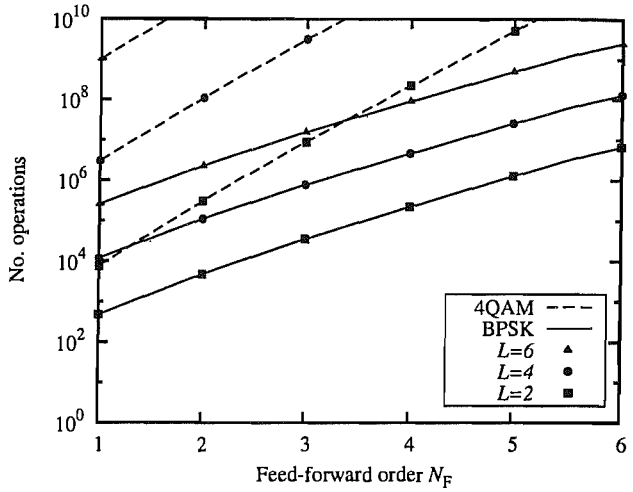
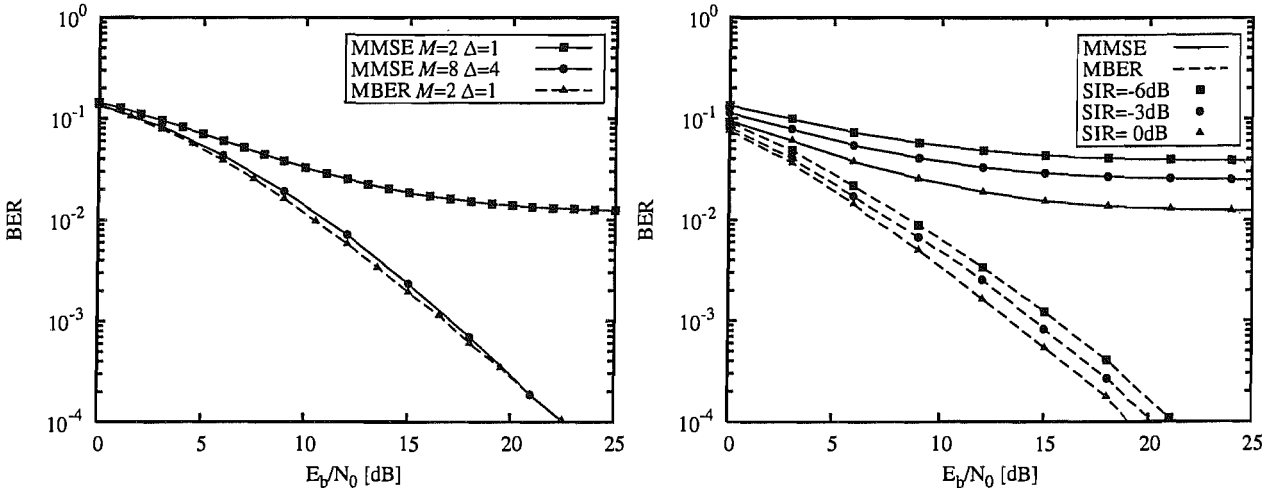


Figure 3.11: The normalized computational complexity required for evaluating the MBER array weights of a system characterized by $\mathfrak{M} = 2$ and $N = 2$ as a function of the feed-forward order N_F for different modulation schemes and a different number of CIR taps, when considering the detection of a single transmitted bit calculated according to Table 3.2. The number of conjugate gradient algorithm iterations N_{grad} was set to 1.



(a) Raw BER versus E_b/N_0 for a scenario supporting $Q = 1$ user employing $M = 2$ transmit AEs.

(b) Raw BER of one MS versus E_b/N_0 . The scenario considered supported $Q = 2$ users, each employing $M = 1$ transmit AE.

Figure 3.12: BER versus E_b/N_0 performance of the MMSE and the MBER STE for different user constellations each experiencing a two-path equal-power independent Rayleigh channel between all antenna links and having perfect channel knowledge. The BS employed a two-element antenna array.

3.4.2.2 Performance of the MBER STE

The SDM system considered is identical to the system used in Section 3.3.2, which is characterized by $M = 2$ transmit AEs at the MS, a BS employing $N = 2$ receive antennas associated with an equalizer characterized by the feed-forward order N_F and a decision delay Δ. The channel was assumed to be a symbol-spaced two-path equal-power block-fading independent Rayleigh channel between all antenna links and the transmitted signals were assumed to be BPSK modulated. Figure 3.12a shows the BER versus E_b/N_0 performance for the uncoded system employing a MBER STE using a feed-forward order of $N_F = 2$

and a decision delay of $\Delta = 1$. As a reference, the BER curve for the MMSE STE is also plotted for identical system parameters, as well as for the optimized parameters of $N_F = 8$ associated with a delay of $\Delta = 1$. It can be observed in Figure 3.12a that for identical parameters, the MBER algorithm outperforms the MMSE algorithm. This indicates that the MMSE algorithm is incapable of confidently separating two linearly separable data sets owing to its MSE-based optimization function, whereas the MBER algorithm achieves this goal. Note in Figure 3.12a that for identical STE parameters, both the MBER and the MMSE algorithm have to perform the same classification task by constructing a linear hyperplane defined by the STE weights. The MMSE algorithm however constructs the hyperplane by minimizing the MSE of the points from the decision boundary and hence certain phasor points are always falsely classified, yielding a BER floor. Only in the higher-dimensional space, i.e for $N_F = 8$ it is capable of correctly separating the two data sets associated with a transmitted $+1$ and a -1.

It was indicated at the beginning of this section that a particular advantage of the MBER algorithm is that it does not rely on the assumption of having a Gaussian distributed conditional PDF of the STE's output. In Rayleigh fading channels, where the powers of the CIR taps vary and the number of transmitters in the system as well as the number of channel path is moderately low, the Gaussian STE output assumption has limited validity, yielding the performance difference between the MBER and the MMSE algorithm observed in Figure 3.12a. This effect can be further explained by considering a system constellation consisting of two MSs, each employing a single transmit AE, which corresponds to a low-dimensional SDMA system. The channel was, again, assumed to be an equal-power two-path block-fading independent Rayleigh channel between all antenna links and a STE associated with $N = 2$, $N_F = 2$ and $\Delta = 1$ was considered. The signals of the two users were assumed to be received by the BS at a different power, as defined by the SIR. In Figure 3.12b the uncoded BER performance of user two versus E_b/N_0 is shown, where it was assumed that the second user was received at a lower average power in order to characterize the 'near-far' behaviour of the system. It can be observed that the BER performance of user two detected with the aid of the MBER detector is hardly degraded as the interferer's power is increased, whereas the MMSE algorithm suffers from a severe BER performance degradation.

3.4.3 Bayesian STE

As seen from Equation (3.20), the MMSE and MBER algorithms presented in Section 3.4.1 and Section 3.4.2, respectively, both constitute so-called linear receivers. The optimum receiver however is not defined by a linear hyperplane obtained with the aid of the MMSE or the MBER criterion, but by a non-linear structure associated with non-linear decision boundaries as indicated in Section 3.2. The receiver realizing the non-linear MBER detector [39] is the Bayesian receiver.

For the derivation of a non-linear STE we will commence from the ML decision for single bit $x_{m\Delta+m}$ of a

BPSK modulated signal, which is defined as

$$\hat{x}_{\mathfrak{m}\Delta+\mathrm{m}} = \arg \max_{x_{\mathfrak{m}\Delta+\mathrm{m}} \in \{+1,-1\}} p(\mathbf{y}|x_{\mathfrak{m}\Delta+\mathrm{m}}), \tag{3.44}$$

where the vector $x_{\mathfrak{m}\Delta+m}$ is a scalar element of the super-vector $\mathbf{x}$ depending on the specific decision delay Δ, as defined in Equation (3.15).

In analogy to the MBER STE also the Bayesian STE may readily be derived for the real-valued system model presented in Equation (3.18). Again, the complex-valued Bayesian STE may be obtained following the derivation of the complex-valued narrowband MIMO detector of Section 2.3.3. Given the set of legitimate real-valued channel output states $\underline{\mathcal{Y}}_{\mathrm{m}}^{\pm}$ defined in Equation (3.36), the Bayesian decision function $f_{\mathrm{B,m}}(\cdot)$ may be expressed as

$$\begin{aligned} f_{\mathrm{B,m}}\left(\underline{\mathbf{y}}\right) &= p\left(\underline{\mathbf{y}} \mid \underline{x}_{\mathfrak{m}\Delta+\mathrm{m}} = +1\right) - p\left(\underline{\mathbf{y}} \mid \underline{x}_{\mathfrak{m}\Delta+\mathrm{m}} = -1\right) \\ &= \sum_{\underline{\check{\mathbf{y}}}_{\mathrm{m}}^{(i,+)} \in \underline{\mathcal{Y}}_{\mathrm{m}}^{+}} \frac{1}{(2\pi\sigma_{\mathrm{n}}^2)^{NN_{\mathrm{F}}}} \exp\left(-\frac{||\underline{\mathbf{y}} - \underline{\check{\mathbf{y}}}_{\mathrm{m}}^{(i,+)}||^2}{2\sigma_{\mathrm{n}}^2}\right) \\ &\quad - \sum_{\underline{\check{\mathbf{y}}}_{\mathrm{m}}^{(i,-)} \in \underline{\mathcal{Y}}_{\mathrm{m}}^{-}} \frac{1}{(2\pi\sigma_{\mathrm{n}}^2)^{NN_{\mathrm{F}}}} \exp\left(-\frac{||\underline{\mathbf{y}} - \underline{\check{\mathbf{y}}}_{\mathrm{m}}^{(i,-)}||^2}{2\sigma_{\mathrm{n}}^2}\right). \end{aligned} \tag{3.45}$$

Again, the introduced notation for the presented real-valued binary interpretation of the Bayesian STE is independent of the modulation scheme and only real-valued operations have to be performed at the receiver.

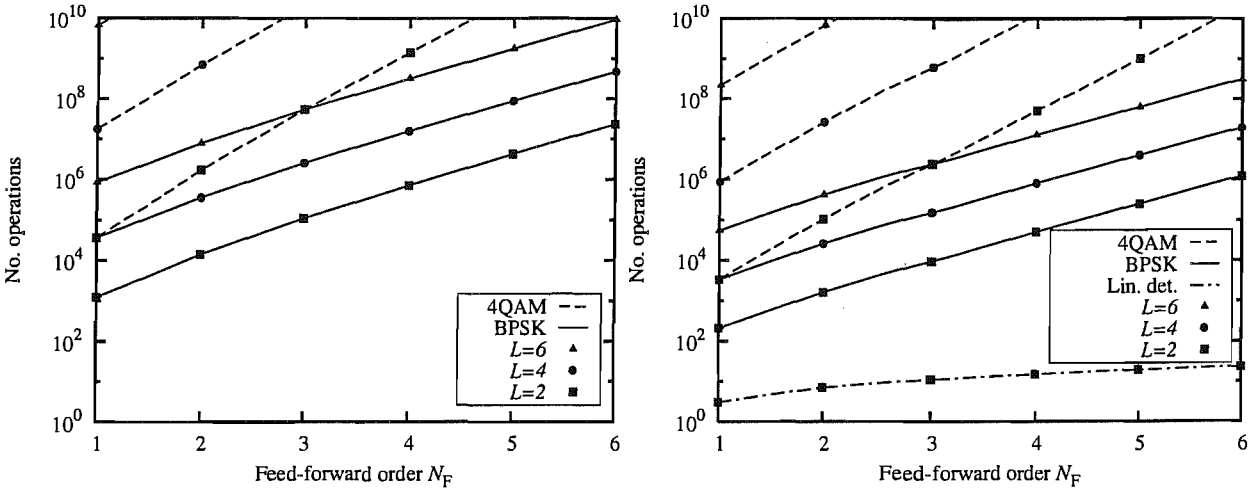
3.4.3.1 Computational Complexity of the Bayesian STE

Similar to the linear MMSE and MBER STE discussed in Section 3.4.1 and Section 3.4.2 respectively, the computational complexity required by the Bayesian STE can be separated into two parts, namely the calculation of the possible channel output states and the detection of the transmitted bits.

The computation of the possible channel output states requires $2N_{\mathrm{s}}N_{\mathrm{F}}N(2N_{\mathrm{bit}}\mathfrak{M}(N_{\mathrm{F}}+L-1)-1)$ real-valued operations. The detection of one binary symbol requires $N_{\mathrm{s}}(6NN_{\mathrm{F}}+1)$ operations plus N_{s} evaluations of the exp function. The total complexity is summarized in Table 3.3 and plotted in Figure 3.13 for both the block-constant fading channel as well as for the time-varying fading channel. It can be seen from Figure 3.13 that the computational complexity associated with the Bayesian STE is exponentially increasing with the number of CIR taps as well as the feed-forward order. This relationship can be observed for both the time-varying as well as for the block-fading channel in Figure 3.13a and Figure 3.13b, respectively. When comparing the complexity of the Bayesian STE and that associated with a linear STE for the block-fading channels illustrated in Figure 3.13b, it can be observed that the Bayesian STE exhibits a significantly higher

Table 3.3: The normalized computational complexity (real-valued additions and multiplications) of the Bayesian STE when considering the detection of a single transmitted bit.

Calculation of the legitimate channel output states	$2N_sN_FN(2N_{bit}\mathfrak{M}(N_F+L-1)-1)/N_{bit}/\mathfrak{M}$
Detection of a single bit	$N_s(6NN_F+1)$ plus N_s exp function evaluations



(a) Number of real-valued additions plus multiplications required for the calculation of the legitimate channel output states plus the the real-valued operations required for the detection of a single bit.

(b) Number of real-valued additions plus multiplications required for the detection of a single bit once the legitimate channel output states have been determined. The complexity considered for the linear detectors only includes the detection and not the computation of the weight vector according to Table 3.1.

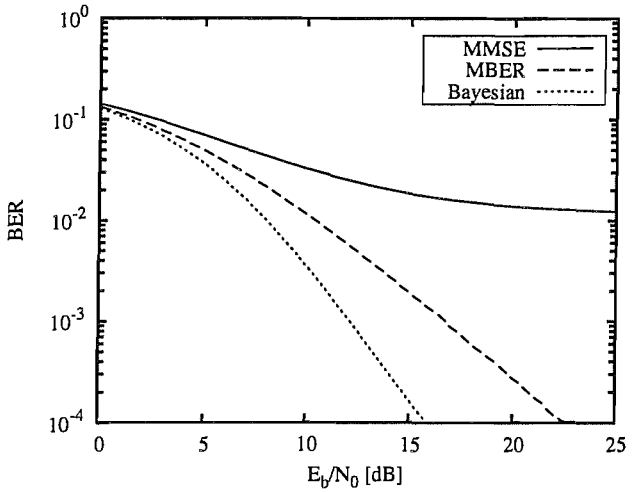
Figure 3.13: The normalized computational complexity imposed by the Bayesian STE for the detection of a single transmitted bit as a function of the number of feed-forward order N_F evaluated for different numbers of CIR taps L as well as different modulation schemes. The number of transmit AEs was $\mathfrak{M} = 2$ and the number of receive AEs was chosen to be $N = 2$. The values were calculated according to Table 3.3.

computational complexity.

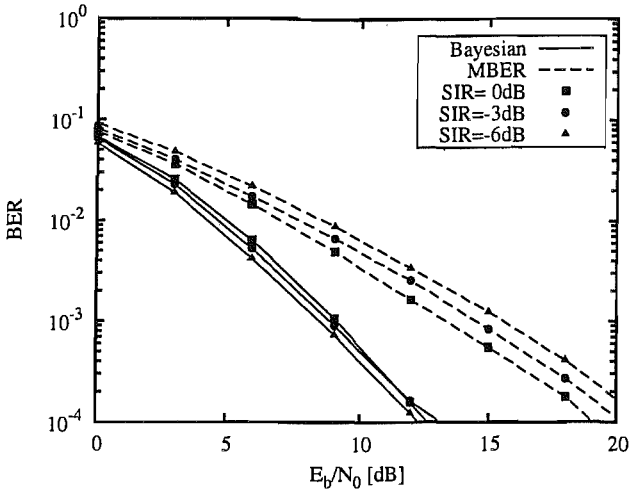
3.4.3.2 Performance of the Bayesian STE

The SDM system considered in this section is identical to that used in Section 3.4.1.2, which is characterized by $M = 2$ transmit AEs at the MS, a BS employing $N = 2$ receive antennas with a STE characterized by the feed-forward order N_F and a decision delay Δ. The channel was assumed to be a symbol-spaced two-path equal-power block-fading independent Rayleigh channel between all antenna links and the signals were assumed to be BPSK modulated. Figure 3.14a shows the BER versus E_b/N_0 performance recorded for the uncoded MBER STE using a feed-forward order $N_F = 2$ and a decision delay of $\Delta = 1$. It can be observed that the non-linear detection strategy constituted by the Bayesian detector outperforms the optimum linear detector represented by the MBER algorithm by more than 5 dB at a BER of 10^{-4}.

When considering the low-dimensional reference SDMA system employed for the creation of Figure 3.12b,



(a) BER versus E_b/N_0 performance for a scenario supporting $Q = 1$ user employing $M = 2$ transmit AEs.



(b) BER of versus E_b/N_0 performance for a scenario with $Q = 2$ users, each employing $M = 1$ transmit AE. The BER of the user received at a lower power is considered.

Figure 3.14: BER versus E_b/N_0 performance of the MMSE, the MBER and the Bayesian STE for different user constellations each experiencing a two-path equal-power independent Rayleigh channel between all antenna links and having perfect channel knowledge. The BS employed a two-element antenna array and a STE characterized by $N_F = 2$ and $\Delta = 1$.

where two users each employing one transmit AE were considered, we parametrized the average power of the signals received by the BS from the two MSs. The channel was assumed to be an equal-power two-path block-fading independent Rayleigh channel between all antenna links and the STE parameters were set to $N = 2$, $N_F = 2$ and $\Delta = 1$. It can be seen from Figure 3.14b that the uncoded BER versus E_b/N_0 performance of the lower-power user is only marginally degraded by the near-far problem imposed by the higher-power user.

3.4.4 Reduced Complexity Maximum Likelihood STE

Due to the structure of the presented algorithm, let us formulate the ML detection of the complete transmitted signal vector given in Equation (3.15) as

$$\hat{\mathbf{x}} = \arg\max_{\check{\mathbf{x}} \in \mathcal{X}} p(\mathbf{y}|\check{\mathbf{x}}), \tag{3.46}$$

where $\mathcal{X}$ is the set of potentially transmitted symbol vectors $\mathbf{x}$. Note that in this formulation the super-vector $\mathbf{x}$ is detected, rather than only the sub-vector of interest, namely $\boldsymbol{x}_{\Delta+1}$. Equation (3.46) may be re-written as

$$\hat{\mathbf{x}} = \arg\min_{\check{\mathbf{x}} \in \mathcal{X}} \{||\mathbf{y} - \mathbf{H}\check{\mathbf{x}}||^2\}. \tag{3.47}$$

Let us assume for the derivation of the algorithm that the CIR matrix $\mathbf{H}$ as well as the noise are real-valued and that the transmitted signal is Binary Phase Shift Keying (BPSK) modulated. It was shown in [17] that

the solution to the problem defined by Equation (3.47) is identical to solving

$$\hat{\mathbf{x}} = \arg\min_{\check{\mathbf{x}} \in \mathcal{X}} ||\mathbf{U}(\check{\mathbf{x}} - \hat{\mathbf{x}}_{\text{MMSE}})||^2, \tag{3.48}$$

where the upper triangular matrix $\mathbf{U}$ is defined by

$$\mathbf{U}^H\mathbf{U} = \mathbf{H}^H\mathbf{H} + \sigma_n^2\mathbf{I}_{\mathfrak{M}(N_F+L-1)}, \tag{3.49}$$

while

$$\hat{\mathbf{x}}_{\text{MMSE}} = (\mathbf{H}^H\mathbf{H} + \sigma_n^2\mathbf{I}_{\mathfrak{M}(N_F+L-1)})^{-1}\mathbf{H}^H\mathbf{y}. \tag{3.50}$$

Let us now first define $N_b = \mathfrak{M}(N_F + L - 1)$ as the number of symbols in $\mathbf{x}$ considered by the OHRSA. Exploiting the fact that the matrix $\mathbf{U}$ has an upper triangular structure, it can be shown that the objective function used for the detection of the transmitted symbol vector $\check{\mathbf{x}}$ may be written as [61]

$$\begin{aligned} J(\check{\mathbf{x}}) &= ||\mathbf{U}(\check{\mathbf{x}} - \hat{\mathbf{x}}_{\text{MMSE}})||^2 && (3.51)\\ &= (\check{\mathbf{x}} - \hat{\mathbf{x}}_{\text{MMSE}})^H\mathbf{U}^H\mathbf{U}(\check{\mathbf{x}} - \hat{\mathbf{x}}_{\text{MMSE}}) && (3.52)\\ &= \sum_{i=1}^{N_b} |\sum_{j=i}^{N_b} u_{ij}(\check{x}_j - \hat{x}_{\text{MMSE},j})|^2 && (3.53)\\ &= \sum_{i=1} \phi_i(\check{\mathbf{x}}_i), && (3.54) \end{aligned}$$

where u_{ij} is the $(i,j)^{\text{th}}$ element of $\mathbf{U}$, $\check{\mathbf{x}}_i = [\check{x}_i, \ldots, \check{x}_{N_b}]^T$ and $\phi_i(\check{\mathbf{x}}_i)$ may be expressed as [61]

$$\begin{aligned} \phi_i(\check{\mathbf{x}}_i) &= |u_{ii}(\check{x}_i - \hat{x}_{\text{MMSE},i}) \\ &\quad + \underbrace{\sum_{j=i+1}^{N_b} u_{ij}(\check{x}_j - \hat{x}_{\text{MMSE},j})}_{a_i}|^2, \end{aligned} \tag{3.55}$$

and where the second term a_i of Equation (3.55) is independent of the specific symbol's value of $\check{x}_i$. The cost-function given in Equation (3.54) may now be re-written in a recursive manner as

$$J_i(\check{\mathbf{x}}_i) = J_{i+1}(\check{\mathbf{x}}_{i+1}) + \phi_i(\check{\mathbf{x}}_i), \quad i = N_b - 1, \ldots 1, \tag{3.56}$$

where we have $J_{N_b}(\check{\mathbf{x}}_{N_b}) = |u_{N_bN_b}(\check{x}_{N_b} - \hat{x}_{\text{MMSE}N_b})|^2$. The cost-function has the essential property that [61]

$$J(\check{\mathbf{x}}) = J_1(\check{\mathbf{x}}_1) > J_2(\check{\mathbf{x}}_2) > \cdots > J_{N_b}(\check{\mathbf{x}}_{N_b}) = J_{N_b}(\check{x}_{N_b}) > 0. \tag{3.57}$$

Due to the fact that the cost-function $J_i(\check{\mathbf{x}}_i)$ only depends on $\{\check{x}_i, \ldots, \check{x}_{N_b}\}$, we introduce the notation

$$J_i(\check{\mathbf{x}}_i) = J_i([\check{x}_i, \ldots, \check{x}_{N_b}]^T) = J_i([\check{x}_i\ \check{\mathbf{x}}_{i+1}]^T). \tag{3.58}$$

Similarly to the narrowband MIMO detector of Section 2.3.4 we may now employ a tree-search algorithm for finding the ML solution, which are discussed in the following example.

Example 3.3: Reduced Complexity ML STE

Consider a simple real-valued candidate system supporting $Q = 2$ MSs, each employing $M = 1$ AE, a $L = 2$-path channel and a BS configuration given as $N_F = 2$, $\Delta = 1$ and $N = 2$. This system configuration results in a $(NN_F \times QM(N_F + L - 1)) = (4 \times 6)$-dimensional channel matrix constructed as outlined in the context of Equation (3.15). The $QM(N_F + L - 1) = 6$-element signal vector contains $N_F + L - 1 = 3$ number of consecutively transmitted vectors $\boldsymbol{x}$, each hosting $QM = 2$ number of elements. The $(N_F N)$-element channel output $\mathbf{y}$ is given as the product of the CIR matrix $\mathbf{H}$ and the transmitted symbol vector $\mathbf{x}$ plus the AWGN. The characteristic quantities of such a system are given for example as

$$\mathbf{H} = \begin{bmatrix} 0.45 & 0.16 & 0.89 & 0.47 & 0 & 0 \\ 0.39 & 0.30 & 0.92 & 0.40 & 0 & 0 \\ 0 & 0 & 0.45 & 0.16 & 0.89 & 0.47 \\ 0 & 0 & 0.39 & 0.30 & 0.92 & 0.40 \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} -1.37 \\ -1.20 \\ 0.04 \\ 0.44 \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} +1 \\ -1 \\ -1 \\ +1 \\ +1 \\ -1 \end{bmatrix}, \tag{3.59}$$

which were recorded at $E_b/N_0 = 12$ dB. This corresponds to a scenario, where the first MS's signal is received at 6 dB higher power than the signal of the second MS.

As mentioned earlier, the convergence of the algorithm can be improved by reordering the columns of the CIR matrix $\mathbf{H}$, so that the columns of the reordered CIR matrix have a monotonously increasing ordered norm [17, 61]. Commencing the algorithm now by reordering the channel

matrix of Equation (3.59) yields $\mathbf{H}^{(o)}$, where the superscript $^{(o)}$ stands for ordered

$$\mathbf{H}^{(o)} = \begin{bmatrix} 0.16 & 0.45 & 0 & 0.47 & 0 & 0.89 \\ 0.30 & 0.39 & 0 & 0.40 & 0 & 0.92 \\ 0 & 0 & 0.47 & 0.16 & 0.89 & 0.45 \\ 0 & 0 & 0.40 & 0.30 & 0.92 & 0.39 \end{bmatrix}. \tag{3.60}$$

It may be readily shown that the corresponding upper triangular decomposition of $\mathbf{H}^{(o)}$ into $\mathbf{U}^{(o)}$ and the required MMSE solution for the ordered system are given as

$$\mathbf{U}^{(o)} = \begin{bmatrix} 0.42 & 0.44 & 0 & 0.46 & 0 & 0.99 \\ 0 & 0.47 & 0 & 0.35 & 0 & 0.69 \\ 0 & 0 & 0.67 & 0.29 & 1.19 & 0.55 \\ 0 & 0 & 0 & 0.38 & 0.19 & 0.33 \\ 0 & 0 & 0 & 0 & 0.52 & 0.06 \\ 0 & 0 & 0 & 0 & 0 & 0.45 \end{bmatrix} \tag{3.61}$$

and

$$\hat{\mathbf{x}}^{(o)}_{\text{MMSE}} = [-0.11 \; -0.61 \; 0.10 \; -0.12 \; 0.69 \; -1.01]^T, \tag{3.62}$$

respectively.

Conventional Search: Note that since both the system matrix and the transmitted signal are real-valued, the imaginary part of the received sequence may be omitted and the MMSE solution is also real-valued. The conventional search algorithm commences at node 0 of Figure 3.15 by evaluating the cost-function of the hypothetical solutions $\breve{x}^{(o)}_{N_b} = \breve{x}^{(o)}_6$ associated with the ordered channel matrix of Equation (3.60), according to Equation (3.56), which yields

$$\begin{aligned} J_6(\breve{x}^{(o)}_6) &= |u^{(o)}_{66}(\breve{x}^{(o)}_6 - \hat{x}^{(o)}_{\text{MMSE},j})|^2 \\ J_6(\breve{x}^{(o)}_6 = +1) &= |.45 \cdot (+1 + 1.01)|^2 = 0.79 \\ J_6(\breve{x}^{(o)}_6 = -1) &= |.45 \cdot (-1 + 1.01)|^2 = 0.00. \end{aligned}$$

The corresponding two values of $J_6(\breve{x}^{(o)}_6)$ can be seen at the second hierarchical level of Figure 3.15 as nodes 1 and 20 together with the associated hypothetical BPSK solutions indicated along the branches. Based on the two cost-function values seen within node 1 and 20 we select node 1, since it has a value of $J_6(\breve{x}^{(o)}_6 = -1) = 0.0$, which is the lower cost-function value. The associated symbol value is $\breve{x}^{(o)}_6 = -1$. In the next step of the algorithm we proceed from node

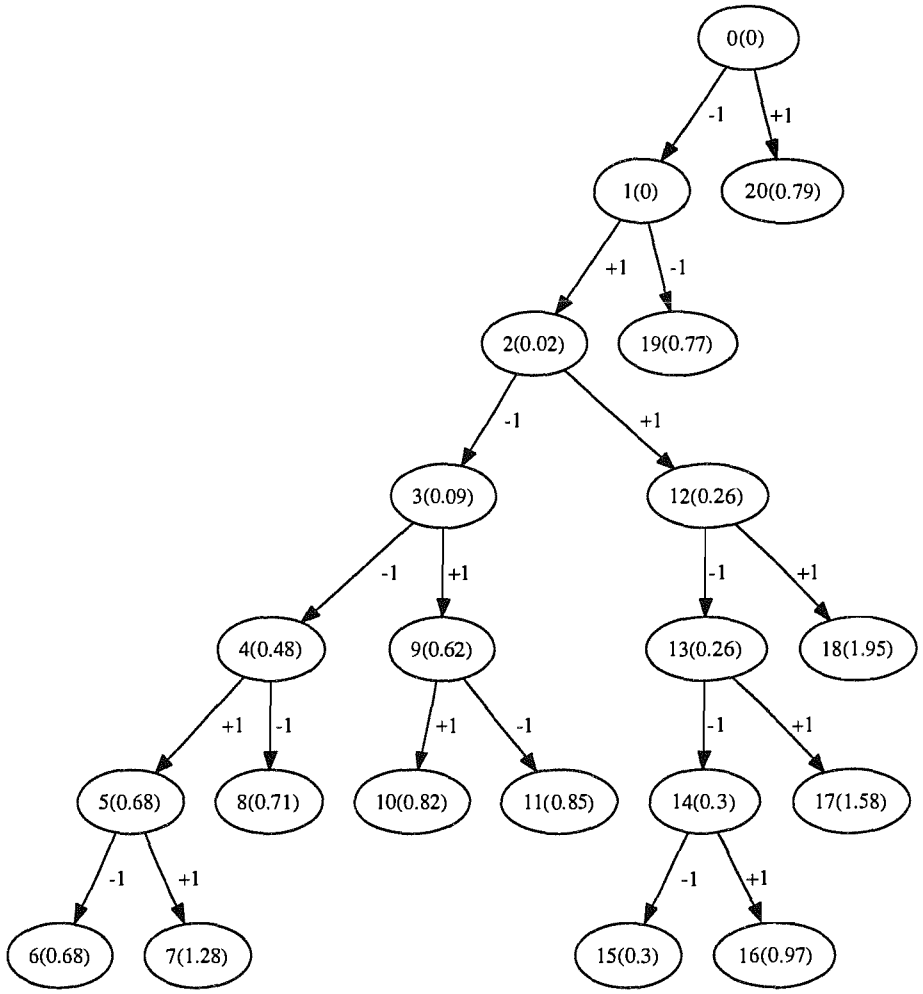


Figure 3.15: Example of a search tree formed by the OHRSA based STE in a scenario employing BPSK modulation, $N = 2, N_F = 2, \Delta = 1, L = 2, Q = 2$ and $M = 1$ and encountering $E_{b/N_0} = 12$ dB. The exact received signal vector $\mathbf{y}$, the channel matrix $\mathbf{H}$ and the transmitted sequence $\mathbf{x}$ are given in Equation (3.59). The labels indicate the search node index, while the value of the cost-function of Equation (3.56) is given in brackets.

1 of Figure 3.15 by calculating the cost function of Equation (3.56) for the next two potential values of $\check{x}_5^{(o)} = \pm 1$ as follows:

$$\begin{aligned} a_5 &= \sum_{j=6}^{6} u_{5j}^{(o)} \cdot (\check{x}_j^{(o)} - \hat{x}_{j,MMSE}^{(o)}) \\ &= 0.85 \cdot (-1 + 1.01) = 0.00 \\ J_5([\check{x}_5^{(o)} \; -1]^T) &= |a_5 + u_{55}^{(o)}(\check{x}_5 - \hat{x}_{\mathrm{MMSE},5}^{(o)})|^2 \\ J_5([+1 \; -1]^T) &= |0.00 + .52 \cdot (+1 - .68)|^2 = 0.02 \\ J_5([-1 \; -1]^T) &= |0.00 + .52 \cdot (-1 - .68)|^2 = 0.77. \end{aligned}$$

The resultant two values for $J_5(\check{\mathbf{x}}_5^{(o)})$ are associated with node 2 and 19, respectively, which are seen at the third hierarchical level of Figure 3.15. The node from which the algorithm is further evolving is node 2, where the associated symbol is $\check{x}_5^{(o)} = +1$, which has a lower cost-function value than node 19.

The value of $\check{\mathbf{x}}_5^{(o)} = [+1 \; -1]^T$ is now used for the calculation of the cost-function values of Equation (3.56) associated with $\check{\mathbf{x}}_4^{(o)}$. The cost-function values $J_4(\check{\mathbf{x}}_4^{(o)})$ are illustrated at the fourth hierarchical level of Figure 3.15 within node 3 and 12. Repeating this procedure will result in the calculation of the cost-function value of $J_3(\check{\mathbf{x}}_3^{(o)})$ provided within node 4 and 9 at the fifth hierarchical level of Figure 3.15, the calculation of $J_2(\check{\mathbf{x}}_2^{(o)})$ given in node 5 and 8 at the sixth hierarchical level of Figure 3.15 and finally $J_1(\check{\mathbf{x}}_1^{(o)})$ given in node 6 and 7 at the seventh and last hierarchical level of Figure 3.15.

Upon arriving at $J_1(\check{\mathbf{x}}_1^{(o)})$ at the bottom of the graph, we have calculated the first potential solution of our optimization problem described by Equation (3.46), which is constituted by the left-most branch of the search tree illustrated in Figure 3.15. This potential solution is given by $\check{\mathbf{x}}^{(o)} = [-1 \; +1 \; -1 \; -1 \; +1 \; -1]^T$.

The recursive optimization continues from the bottom (node 6) to the top of the tree of Figure 3.15 with the objective of finding the specific branch terminating at the bottom of the search tree (hierarchical level 6) while having the minimum cost-function value. The symbol vector x associated with this branch constitutes the ML solution. Considering now the flipping of symbol $\check{x}_1^{(o)}$ or $\check{x}_2^{(o)}$ at hierarchical level 5 or 6 is not beneficial, since the cost-function associated with these changes would result in a higher value, namely 1.28 and 0.71, than the cost function of 0.68 at node 6. Hence, if these tentative branches would be pursued to the bottom of the tree, it would ultimately result in an increased cost-function value as a consequence of the ordering property outlined in Equation (3.57).

When the bottom-to-top recursive process initiated at node 6 arrives at node 3, however, it can

be seen that changing the value of $\check{x}_4^{(o)}$ from -1 to $+1$ results in a cost function value of 0.62 at node 9, which is lower than the cost-function value of 0.68 recorded at node 6. Pursuing now the path from node 9 further down the tree will however again result in a cost function value that is higher than that of node 6. The corresponding branch is therefore not pursued further.

Returning to the left-most branch of Figure 3.15 and changing the value of $\check{x}_4^{(o)}$ at the hierarchical level 3 from -1 to $+1$, results in a cost-function value of 0.26, which is lower than that of node 6 and the corresponding path is therefore pursued further down through nodes 13 and 14. Ultimately, this results in two new branches through nodes 12, 13, 14 and 15 as well as 12, 13, 14 and 16, terminating at $\check{\mathbf{x}}_1^{(o)}$, which constitute additional potential solutions for $\mathbf{x}^{(o)}$. Pursuing this path from node 15 and 16 backwards, returning to the most-left branch and moving recursively up to node 0 will result in no further branches terminating at the bottom of the tree. Hence node 15 has the lowest cost-function value of 0.3.

The desired solution is described by the specific branch terminating at the bottom of the search tree that has the lowest cost-function value. Again, the associated symbol sequence is given by

$$\check{\mathbf{x}}^{(o)} = [-1\ -1\ -1\ +1\ +1\ -1]^T,$$

which can be obtained by tracing the branch back from node 15 to node 0. By contrast, the identically ordered MMSE solution is given by

$$\mathtt{sign}\left(\hat{\mathbf{x}}_{\mathrm{MMSE}}^{(o)}\right) = [-1\ -1\ +1\ -1\ +1\ -1]^T.$$

The final step to obtain the desired ML solution is to reverse the ordering of the obtained $\check{\mathbf{x}}^{(o)}$ solution and to choose the desired sub-vector $\check{\boldsymbol{x}}_{\Delta+1}$ according to the decision delay. This yields

$$\begin{aligned} \check{\mathbf{x}} &= [-1\ -1\ \underbrace{-1\ +1}_{\boldsymbol{x}_{\Delta+1}}\ +1\ -1]^T \\ \mathtt{sign}\left(\hat{\mathbf{x}}_{\mathrm{MMSE}}\right) &= [-1\ -1\ \underbrace{-1\ -1}_{\boldsymbol{x}_{\Delta+1}}\ +1\ +1]^T \end{aligned},$$

which highlights the decision errors made by the MMSE detector.

Min-Path Search: In Example 2.5 and Algorithm 2.1, the recursive search algorithm of Example 2.18 was further developed to the Min-Path strategy. The same principle may also be applied in the context of STE, which results in the following search strategy for the example considered.

Again, the search-process commences from node 0 of Figure 3.15 by evaluating the cost-

function value of nodes 1 and 20 in Figure 3.15 according to Equation (3.56) in analogy to the above-mentioned conventional search process. The nodes associated with the search-front $\mathcal{F}$ are now node 1 with a cost-function value of 0 and node 20 with a cost-function value of 0.79. The search-front is written as $\mathcal{F} = \{1, 20\}$.

According to the Min-Path search strategy, the search process continues from that specific node of the search-front, which has the lowest cost-function value. In our case this yields the evaluation of the 'children' of node 1 in Figure 3.15, which are given as nodes 2 and 19. Their cost-function can be evaluated according to Equation (3.56). These nodes, namely nodes 2 and 19, are now added to the search-front, while node 1 is removed. The new search-front is then represented by $\mathcal{F} = \{20, 2, 19\}$.

Of all the nodes in the search-front, node 2 has now the lowest cost-function value, as seen in Figure 3.15 and is therefore considered for the continuation of the optimization process. Adding the children of node 2 of Figure 3.15 given as nodes 3 and 12 and removing node 2 from the search-front results in $\mathcal{F} = \{20, 19, 3, 12\}$.

When the search process is evolved further following this principle, i.e. by always continuing the evaluation of the search-process from that specific node, which is associated with the lowest cost-function, the following sequence of search-fronts is obtained:

$\mathcal{F} = \{20, 19, 12, 4, 9\} \Longrightarrow$ continue from node 12

$\mathcal{F} = \{20, 19, 4, 9, 13, 18\} \Longrightarrow$ continue from node 13

$\mathcal{F} = \{20, 19, 4, 9, 18, 14, 17\} \Longrightarrow$ continue from node 14

$\mathcal{F} = \{20, 19, 4, 9, 18, 17, 15, 16\}$.

When arriving at node 15, the search process is terminated, since that particular node, which has the lowest cost-function value of all nodes forming the search-front is a node terminating at the bottom of the search tree in Figure 3.15, i.e. node 15. The ML solution associated with node 15 of Figure 3.15 can now be obtained in analogy to the conventional tree-search discussed in the previous paragraph of this example.

When comparing the conventional and the Min-Path search strategy, it can be observed that for this particular example the Min-Path search is associated with a lower computational cost, since it does not visit nodes $\{5, 10, 8, 11, 6, 7\}$ of Figure 3.15.

Despite the fact that the Min-Path algorithm requires a lower number of additions and multiplications for finding the ML solution, we opted for using the conventional tree-search in our simulations. This decision was motivated by the fact that for most scenarios considered in this treatise tracking the search-front of the Min-Path approach requires an excessive number of memory allocation and deallocation operations, which resulted in a longer run-time of the algorithm compared to that of the conventional tree-search.

Table 3.4: The normalized computational complexity (real-valued additions and multiplications) of the OHRSA STE, when considering the detection of a single transmitted bit.

MMSE Solution and factorization of Equation (3.50) and Equation (3.49)	$(8N_b^2NN_F + N_b - N_b^2 - 2N_bNN_F)/(\mathfrak{M}N_{bit})$ $+N_b(4NN_F - 1)/(\mathfrak{M}N_{bit}) + O(N_b^3/(\mathfrak{M}N_{bit}))$
Tree evaluation of Equation (3.55) and Equation (3.56) at hierarchical level i	$3(N_b - i) + 1$

The fact, that only a subset of the detector outputs is actually of interest, namely the sub-vector $\boldsymbol{x}_{\Delta+1}$ of $\mathbf{x}$ is a major difference between the OHRSA applied to the non-dispersive OFDM sub-channels in [61] and the wideband scenarios considered in this subsection.

In order to apply this scheme for detecting 4-level Quadrature Amplitude Modulated (4QAM) signals the real-valued binary system model of Equation (3.18) has to be employed.

3.4.4.1 Computational Complexity of the OHRSA Detector

The computational complexity of the OHRSA STE can again be divided into two contributions, namely the cost associated with the calculation of the MMSE solution and the factorization of the covariance matrix of Equation (3.49) into an upper triangular matrix as well as the evaluation of the cost-function along the search tree. Let us first define $B = 2NN_F$.

The calculation of the MMSE solution is constituted by $N_b^2(2B-1) + N_b$ number of operations imposed by the evaluation of the covariance matrix of Equation (3.49), plus $O(N_b^3)$ number of operations associated with the calculation of its inverse. For the final calculation of the weight vector, an additional $(2N_b - 1)N_bB$ number of operations are required for carrying out the multiplication of the inverse of the covariance matrix of Equation (3.50) with the cross-correlation matrix, as suggested by Equation (3.50). The multiplication of the MMSE weight matrix with the received signal contributes a further $(2B-1)N_b$ operations. The Cholesky decomposition of the covariance matrix typically requires $O(N_b^3)$ number of operations. In addition to the MMSE solution of Equation (3.50) and the Cholesky factorization of Equation (3.49), the complexity of the OHRSA STE is determined by the number of operations required for the evaluation of the cost-function along the search tree. At hierarchical level i, the evaluation of the cost-function characterized by Equation (3.56) as well as Equation (3.55) requires $3 \cdot (N_b - i - 1) + 4 = 3(N_b - i) + 1$ operations. In Table 3.4 the complexity normalized by the number of bits $\mathfrak{M}N_{bit}$, which have to be detected is summarized.

Figure 3.16 shows the normalized complexity required for the evaluation of the search tree as a function of the feed-forward order N_F for both BPSK as well as 4QAM, when considering $L = 2, 4$ and 6 CIR taps at SNR=20 dB for $\mathfrak{M} = 2$ and $N = 2$. The complexity associated with the search tree evaluation increases only moderately with the feed-forward order N_F. This specific behaviour is mainly due to the fact that

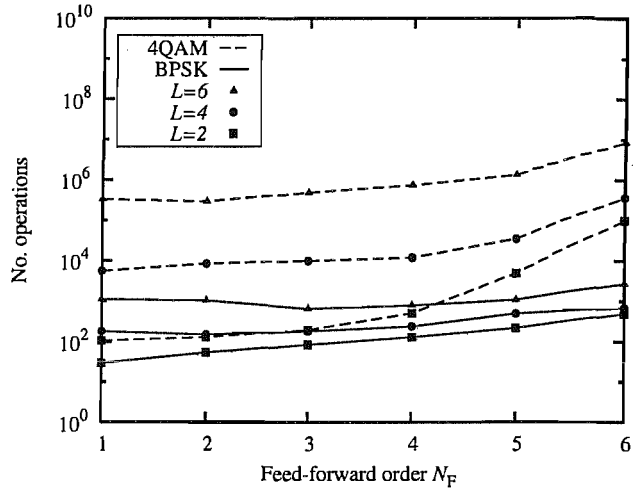


Figure 3.16: The normalized computational complexity required for evaluating the search tree as a function of the number of transmit AEs for different modulation schemes and a different number of MIMO array weights, when considering the detection of a single transmitted bit calculated according to Table 3.4 at SNR=20 dB.

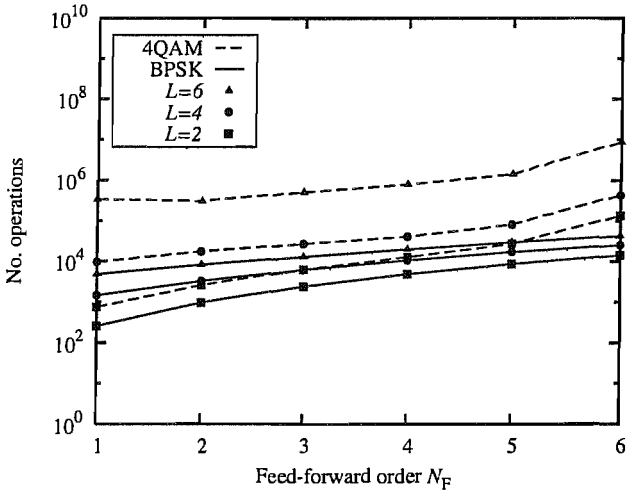
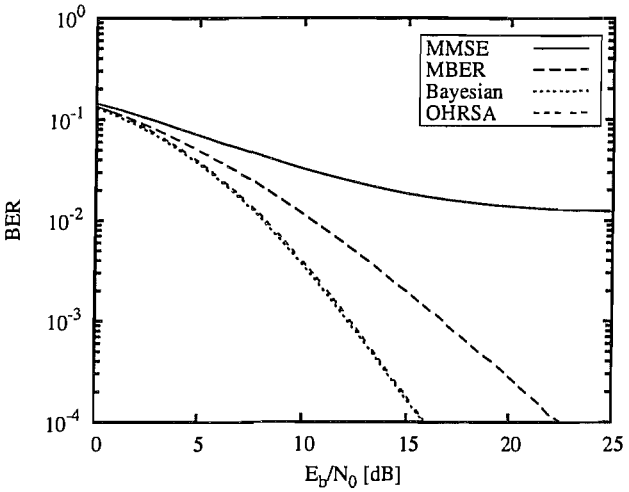
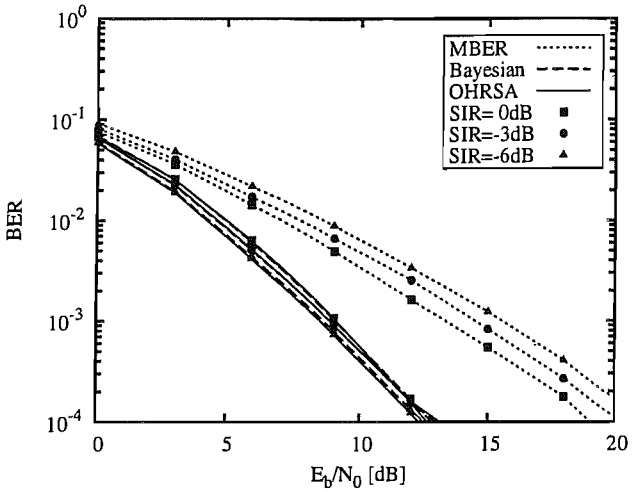


Figure 3.17: The normalized computational complexity required for evaluating the search tree plus the evaluation of the MMSE solution as a function of the number of transmit AEs for different modulation schemes and a different number of MIMO array weights, when considering the detection of a single transmitted bit calculated according to Table 3.4 at SNR=20 dB. The proportionality factor associated with the MMSE solution was set to 3.



(a) Raw BER versus E_b/N_0 for a scenario supporting $Q = 1$ user employing $M = 2$ transmit AEs.



(b) Raw BER of one MS versus E_b/N_0 . The scenario considered supported $Q = 2$ users, each employing $M = 1$ transmit AE.

Figure 3.18: BER versus E_b/N_0 performance of the MMSE, the MBER and the Bayesian STE for different user constellations each experiencing a two-path equal-power independent Rayleigh channel between all antenna links and having perfect channel knowledge. The BS employed a two-element antenna array.

even though the search-space increases upon increasing the feed-forward order, the problem of finding the ML solution generally becomes less difficult, i.e less branches of the search tree have to be evaluated. If in addition to the search tree-evaluation also the computation of the MMSE solution is taken into account, it can be observed from Figure 3.17 that the computational complexity associated with the OHRSA STE increases only moderately upon increasing the feed-forward order N_F. Finally, the computational cost associated with the OHRSA STE illustrated in Figure 3.16 and Figure 3.17 is significantly lower than that of the Bayesian STE characterized in Figure 3.13a and Figure 3.13b, respectively.

3.4.4.2 Performance of the OHRSA STE

The SDM system considered is identical to that used in Section 3.4.1.2 which is characterized by $M = 2$ transmit AEs at the MS, a BS employing $N = 2$ receive antennas in conjunction with an equalizer characterized by the feed-forward order N_F and a decision delay Δ. The channel was assumed to be a symbol-spaced two-path equal-power block-fading independent Rayleigh channel between all antenna links and the transmitted signals were assumed to be BPSK modulated. Figure 3.12a portrays the BER versus E_b/N_0 curve for the uncoded MBER, MMSE, Bayesian and OHRSA STE using a feed-forward order of $N_F = 2$ and a decision delay of $\Delta = 1$. It can be observed that the BER performance of the OHRSA algorithm is identical to that of the Bayesian detector. The same agreement between the performance of the OHRSA and the Bayesian detector can also be observed in Figure 3.18b, which shows the uncoded BER versus E_b/N_0 performance for a SDMA system, where the channel was again assumed to be an equal-power two-path block-fading independent Rayleigh channel between all antenna links and a STE associated with

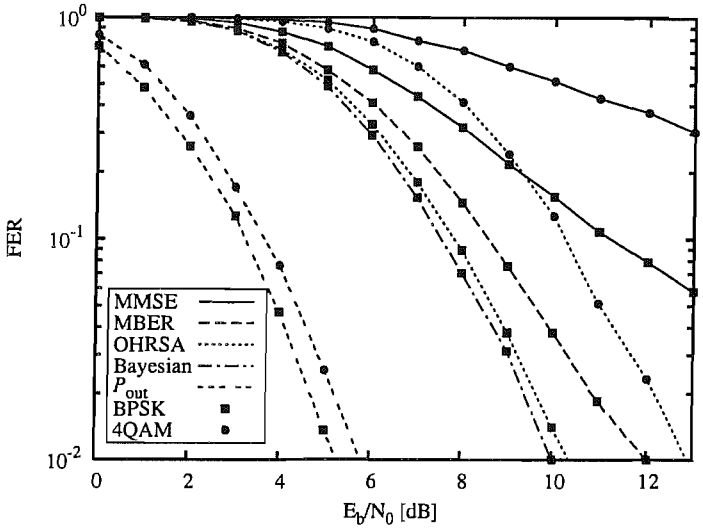


Figure 3.19: FER versus E_b/N_0 performance of the MMSE, MBER, Bayesian and OHRSA STE for a scenario supporting $Q = 1$ user employing $M = 2$ transmit AEs, a two-path equal-power independent block-fading Rayleigh channel, a BS associated with $N = 2$, $N_F = 2$, $\Delta = 1$ and having perfect channel knowledge. The system employed a $R = 1/2$-rate punctured turbo-code.

$N = 2$, $N_F = 2$ and $\Delta = 1$ was considered. The different users' signals were received at a different power at the BS. The BER is shown for the specific user, whose signal was received at a lower power.

3.4.5 Turbo-Coded Performance of STEs

In this subsection the performance of the different algorithms is compared for a SDM system supporting $Q = 1$ user, which employs $M = 2$ transmit AEs[2]. The user communicates with the BS over a dispersive equal-power two-tap independent Rayleigh channel and the BS is equipped with $N = 2$ AEs and a STE with feed-forward order $N_F = 2$ using a decision delay $\Delta = 1$. In order to render the transmitter more realistic, we use Forward Error Correction (FEC) to improve its achievable performance. The channel code used is a $R = 1/2$-rate punctured turbo-code having Recursive Systematic Convolutional (RSC) component codes described by the generator polynomial $\boldsymbol{g} = [7\ 5]$, which was fed with the hard-decision estimates of the STE. For a first comparison we assume the fading envelope to be constant over one transmission burst equivalent to 1000 information bits and the channel envelope was assumed to be perfectly known by the receiver. In total 5000 bursts were transmitted in order to acquire sufficient fading statistics and a reliable BER and FER estimate.

Since for a block-fading channel the most pertinent performance measure is the code-word error rate, which is identical to the FER in our case, Figure 3.19 shows the FER versus E_b/N_0 performance for the different algorithms introduced in this chapter. As a reference the outage probability for an information rate of $R = 1$ for BPSK modulated signals and of $R = 2$ for 4QAM modulated signals is also given in Figure 3.19, which

[2]If no channel coding is used, there is no difference between a Q user SDMA system having a single AE and a single-user SDM system employing Q antennas, except for their difference in transmit power.

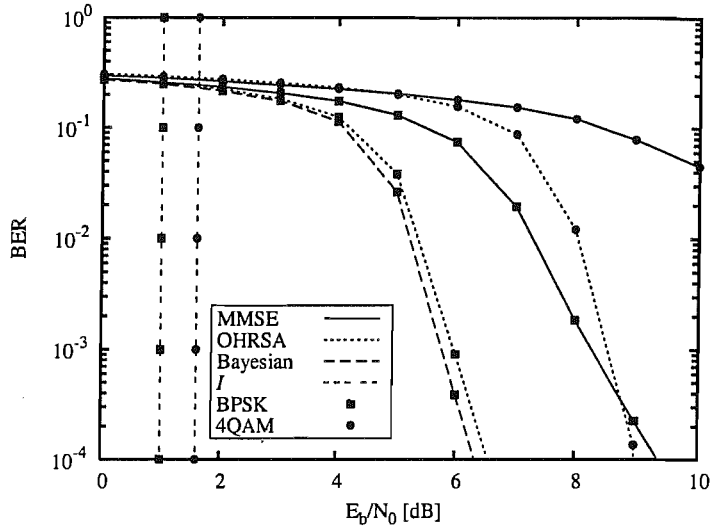


Figure 3.20: BER versus E_b/N_0 performance of the MMSE, Bayesian and OHRSA STE for a scenario supporting $Q = 1$ user employing $M = 2$ transmit AEs, a two-path equal-power independent uncorrelated Rayleigh channel, a BS associated with $N = 2$, $N_F = 2$, $\Delta = 1$ and having perfect channel knowledge. The system employed a $R = 1/2$-rate punctured turbo-code.

was calculated in Section 3.3 as a performance bound for all algorithms.

When comparing the achievable performance of the MBER and the MMSE algorithm for BPSK modulated signals, a substantial performance difference can be observed. Compared to the optimum non-linear detectors, namely to the Bayesian and the OHRSA STE, the MBER algorithm exhibits a relatively moderate performance degradation, which increased with an increasing SNR. For 4QAM signals only the performance of the MMSE and the OHRSA assisted STE is depicted, since the complexity imposed by the Bayesian as well as the MBER STE was considerably high, i.e the number of legitimate channel output states to be considered was $N_s = 4^{2\cdot(2+2-1)} = 4^6$.

In Figure 3.20 the achievable BER versus E_b/N_0 performance is illustrated for the same system but assuming an uncorrelated rather than a block-fading channel. The performance associated with both the MBER as well as the Bayesian STE designed for 4QAM signals is not included owing to their excessive computational cost. Similarly to the achievable performance for the block-fading channel shown in Figure 3.19, it can be seen in Figure 3.20 that the MMSE detector performs considerably worse than the Bayesian and the OHRSA STE also for the uncorrelated fading channel scenario. Furthermore, it can be observed that even the ML STEs benefiting from turbo-coding operate at a distance of more than 4 dB from the channel capacity.

3.4.6 Conclusion

In this section, we have introduced different approaches for implementing STE algorithms. Their performance was compared in Figure 3.19 as well as Figure 3.18 using a very simple channel and interferer

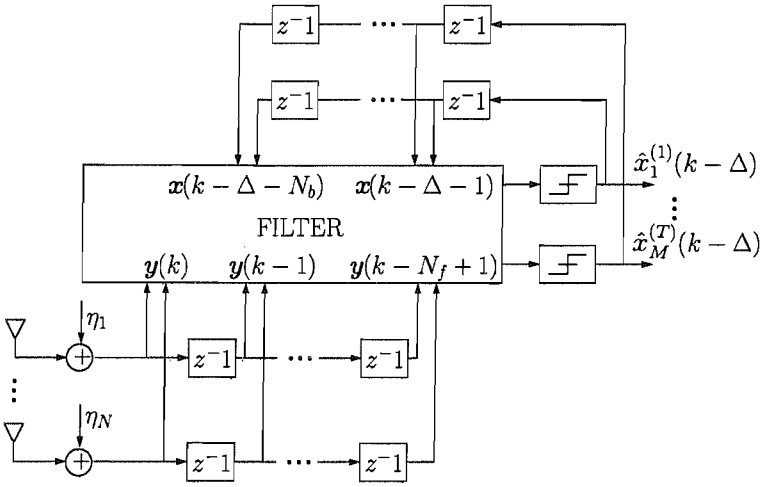


Figure 3.21: Space-time equalizer with feed-forward order N_F and feedback order N_B employing N receive antennas.

setting and the differences were highlighted. It was shown that even the more sophisticated algorithms are far from approaching the lower capacity bound. Furthermore, the computational complexity of the different detection strategies has to be reduced in order to apply them in larger-scale systems. These two topics are addressed in the next section by introducing the decision feedback structure, which aims for tackling both above-mentioned problems.

3.5 Decision Feedback Assisted Finite-Length STE

The performance of both linear and non-linear equalizers can be enhanced by incorporating a decision feedback structure [118] in the receiver, as shown in Figure 3.21. In addition to the feed-forward order N_F and the decision delay parameter Δ we introduce the decision feedback order N_B. Note, that the oldest symbol vector, which still influences the detected symbol $\hat{\boldsymbol{x}}_{\Delta+1}$ is $\boldsymbol{x}_{N_F+L}$. Furthermore, the oldest feedback symbol vector is $\boldsymbol{x}_{\Delta+N_B+1}$. Without loss of generality we therefore chose $N_B = N_F + L - 2 - \Delta$ for the derivation of the proposed decision feedback aided STEs. In order to describe the feedback structure, we first divide the system matrix $\mathbf{H}$ into two sub-matrices [118]

$$\mathbf{H} = [\mathbf{H}_1 \; \mathbf{H}_2], \tag{3.63}$$

where $\mathbf{H}_1$ hosts the first $\mathfrak{M}(\Delta+1)$ columns of $\mathbf{H}$ and $\mathbf{H}_2$ represents the last $\mathfrak{M}N_B$ columns in $\mathbf{H}$. The array output can then be written as

$$\begin{aligned} \mathbf{y} &= \mathbf{y}_1 + \mathbf{y}_2 \\ &= \check{\mathbf{y}}_1 + \check{\mathbf{y}}_2 + \boldsymbol{\eta} \\ &= \mathbf{H}_1\mathbf{x}_1 + \mathbf{H}_2\mathbf{x}_2 + \boldsymbol{\eta}, \end{aligned} \tag{3.64}$$

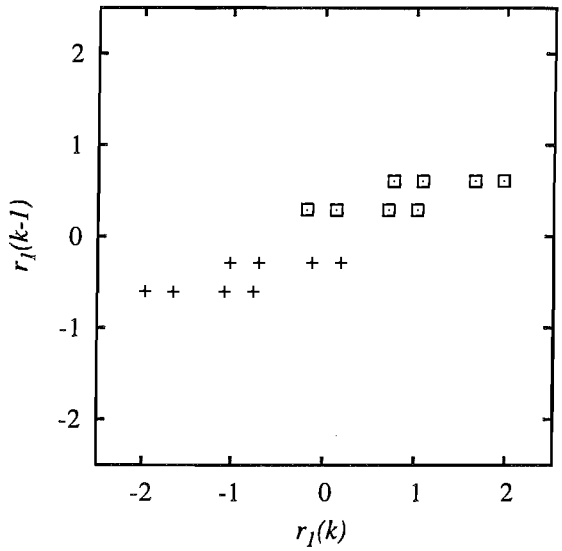


Figure 3.22: STE complexity reduction by space-translation for the DF-aided STE depicted in Figure 3.21. Its counterpart using no DF was characterized in Figure 3.3a.

where $\mathbf{x}_1 = \left[\boldsymbol{x}(k)^T \ldots \boldsymbol{x}(k-\Delta)^T\right]^T = \left[\boldsymbol{x}_1^T \ldots \boldsymbol{x}_{\Delta+1}^T\right]^T$ indicates the symbols in the feed-forward shift register and $\mathbf{x}_2 = \left[\boldsymbol{x}(k-\Delta-1)^T \ldots \boldsymbol{x}(k-\Delta-N_B)^T\right]^T = \left[\boldsymbol{x}_{\Delta+2}^T \ldots \boldsymbol{x}_{\Delta+N_B+1}^T\right]^T$ denotes the symbols in the feedback register. The complexity reduction achieved by the decision feedback is due to the fact that the previously received symbols of all transmitters have already been decided upon and hence their ambiguity imposed on the constellation may be eliminated.

Equation (3.65) can be interpreted as a space translation [129], where a received signal vector $\mathbf{y}$ is translated into the new observation space $\mathbf{r}$ by subtracting the product of the decision feedback sequence $\mathbf{x}_2$ and the channel matrix $\mathbf{H}_2$. The space translation is given as

$$\mathbf{r} = \mathbf{y} - \mathbf{H}_2\mathbf{x}_2. \tag{3.65}$$

For the given choice of N_B the space translation can be implemented as shown in [129]. This implementation has the advantage that for a feedback assisted STE still only the feed-forward filter has to be designed and the space-translation due to the fed back symbols only requires the knowledge of the feedback channel matrices. The design of the feed-forward filter now has to be done in the translated space $\mathbf{r}$ rather than in the channel output space $\mathbf{y}$. Compared to the computational complexity of the detection algorithms, the computational complexity associated with the space-translation employed to implement the decision feedback is negligible since it is linear in all variables.

Example 3.4: Space-Translation Approach to DF aided STE

In Figure 3.22 the translated decision space is illustrated for the system considered in Example 3.1 in Section 3.2 and for a feedback order of $N_B = 1$. In contrast to its counterpart characterized in Figure 3.3a, which does not benefit from having decision feedback, Figure 3.22

Table 3.5: The normalized computational complexity (real-valued additions and multiplications) of the MMSE DF-STE when considering the detection of a single transmitted bit. The decision delay was chosen to be $\Delta = N_F - 1$.

Weight Vector Calculation	$(2N_{bit}\mathfrak{M}(\Delta+1)-1)4N^2N_F^2/(N_{bit}\mathfrak{M})$ $+2NN_F/(N_{bit}\mathfrak{M})$ $+(4NN_F-1)2NN_F+O(N^3N_F^3/(N_{bit}\mathfrak{M}))$
Detection	$4NN_F-1$

illustrates a lower-dimensional, linearly separable decision space. In the following sections we will characterize the algorithms discussed in Section 3.3 in the context of the new observation space $\mathbf{r}$ of Equation (3.65).

3.5.1 Minimum Mean Squared Error Decision-Feedback STE

Under the assumption that the detected feedback sequence $\tilde{\mathbf{x}}_2$ is correct, the output of a linear filter designed for a system incorporating feedback for supporting space translation can be written as

$$\begin{aligned} \hat{\boldsymbol{x}} &= \boldsymbol{W}^H(\mathbf{y}-\mathbf{H}_2\tilde{\mathbf{x}}_2) \\ &= \boldsymbol{W}^H\mathbf{r}, \end{aligned} \tag{3.66}$$

where $\boldsymbol{W}$ is again the STE's array weight matrix, where the columns are associated with the different transmitters, $\hat{\boldsymbol{x}}$ is the associated filter output and $\tilde{\mathbf{x}}_2$ are the detected feedback symbols. The squared error vector at the filter output can then be written as

$$\epsilon^2 = ||\boldsymbol{x}_{\Delta+1} - \boldsymbol{W}^H\mathbf{r}||^2 = ||\boldsymbol{x}_{\Delta+1} - \hat{\boldsymbol{x}}||^2. \tag{3.67}$$

In analogy to the derivation of the MMSE weights for the STE operating without decision feedback, the weights minimizing the MSE at the detector's output can be shown to obey

$$\boldsymbol{W} = \left(\mathbf{H}_1\mathbf{H}_1^H + 2\sigma_n^2\mathbf{I}_{NN_F}\right)^{-1}\boldsymbol{R}_{yx}. \tag{3.68}$$

The cross-correlation matrix $\mathbf{R}_{yx}$ is given as the columns ranging from the $(\mathfrak{M}\Delta+1)^{\text{th}}$ to the $(\mathfrak{M}(\Delta+1))^{\text{th}}$ column of the channel matrix $\mathbf{H}_1$. It can be shown that this solution is identical to the MMSE solution presented in [75], which realizes the DF-STE using a linear feed-forward and feedback filter.

Following the evaluation of the computational complexity of the MMSE STE dispensing with feedback given in Section 3.4.1, it can be shown that the computational complexity of the DF-aided STE is as given in Table 3.5. The operations required for performing the space-translation are neglected. Comparing Table 3.1

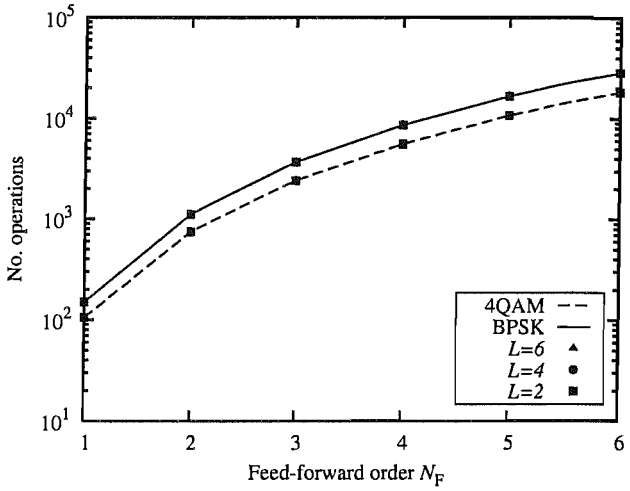


Figure 3.23: The normalized computational complexity required for evaluating the MMSE array weights as a function of the number of feed-forward order for different modulation schemes and a different number of CIR taps, when considering the detection of a single transmitted bit calculated according to Table 3.5. The proportionality factor was set to 1 and the total number of transmit AEs was chosen to be $QM = \mathfrak{M} = 2$, while the number of receive AEs was $N = 2$. The curves associated with $L = 2, 4$ and 6 are indistinguishable.

and Table 3.5 it can be seen that the DF structure removes the dependency of the complexity on the length of the CIR L and substituting it by the chosen decision delay.

In Figure 3.23 we illustrated the computational complexity imposed by the DF-MMSE detector according to Table 3.5 as a function of the feed-forward order of the STE considering $\mathfrak{M} = 2$ transmit AEs, $N = 2$ receive AEs as well as $L = 2, 4$ and 6 CIR taps. The modulation schemes considered were BPSK as well as 4QAM and the decision delay was chosen to be $N_F - 1$. When comparing the complexity illustrated in Figure 3.23 to that of Figure 3.7, which is associated with a STE dispensing with DF, then it can be observed that the employment of DF reduced the computationally cost imposed by the MMSE STE for all the different scenarios considered.

Furthermore, it can be seen in Figure 3.24, which shows the BER versus E_b/N_0 performance for a simple SDM system as given in the caption that with the advent of DF, the error-floor compared to the STE not benefiting from DF is removed. Increasing the feed-forward order to $N_F = 4$ increases the performance further at the cost of an increased complexity. Making the feed-forward order even larger is not beneficial as it can be seen from Figure 3.24. The difference between the curve associated with detected feedback and the curve assuming that the symbols which are fed back are correctly detected is in the order of approximately 2 dB at a BER of 10^{-3}.

3.5.2 Minimum Bit Error Rate Decision-Feedback STE

The derivation of the MBER DF-STE is based on the knowledge of the legitimate noise-free channel output states in the translated space $\mathbf{r}$. For the DF-assisted receiver the noise-free channel output states observed in

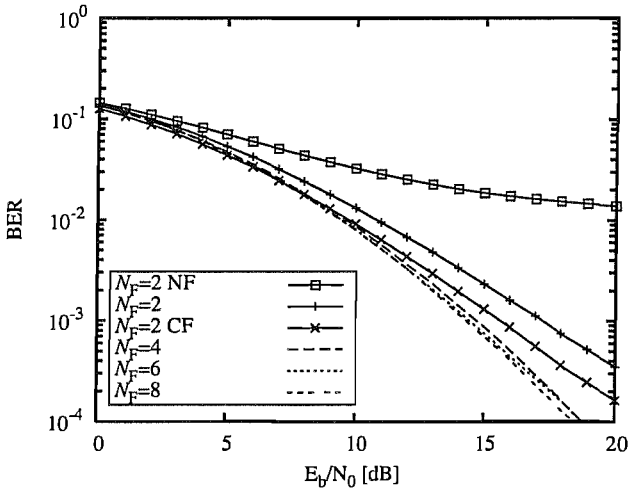


Figure 3.24: BER versus E_b/N_0 performance of the MMSE STE for a scenario supporting $Q = 1$ user employing $M = 2$ transmit AEs, a two-path equal-power independent Rayleigh channel and having perfect channel knowledge. The BS employed a two-element antenna array. The label NF indicates that no feedback was used and the label CF represents correct feedback. BPSK modulated signal were considered.

the translated observation space are given as

$$\mathcal{R} = \{\check{\mathbf{r}}^{(i)} = \mathbf{H}_1 \check{\mathbf{x}}_1^{(i)},\ 1 \leq i \leq N_s\}, \tag{3.69}$$

where in the case of BPSK modulated sources N_s is given as

$$N_s = 2^{\mathfrak{M}(\Delta+1)}. \tag{3.70}$$

For each transmitter $\mathfrak{m}$ the legitimate channel output set $\mathcal{R}$ is partitioned into two subsets, depending on the value of $x_\mathfrak{m}(k-\Delta)$ yielding

$$\mathcal{R}_\mathfrak{m}^{\pm} = \{\check{\mathbf{r}}^{(i)} \in \mathcal{R} \mid \check{x}_{\mathfrak{M}\Delta+\mathfrak{m}}^{(i)} = \pm 1\}. \tag{3.71}$$

Following the derivation of the MBER STE operating without DF, the noise-free part of the linear STE's output is limited to have values from the set

$$\hat{\mathcal{X}}_\mathfrak{m} = \{\check{\hat{x}}_\mathfrak{m}^{(i)} = \boldsymbol{w}_\mathfrak{m}^H \check{\mathbf{r}}^{(i)},\ 1 \leq i \leq N_s\}, \tag{3.72}$$

which again can be partitioned into two subsets defined as

$$\hat{\mathcal{X}}_\mathfrak{m}^{\pm} = \{\check{\hat{x}}_\mathfrak{m}^{(i)} \in \hat{\mathcal{X}}_\mathfrak{m} \mid \check{x}_{\mathfrak{M}\Delta+\mathfrak{m}}^{(i)} = \pm 1\}. \tag{3.73}$$

Table 3.6: The normalized computational complexity (real-valued additions and multiplications) of the MBER DF-STE when considering the detection of a single transmitted bit. The decision delay was chosen to be $\Delta = N_F - 1$.

Weight Vector Calculation	$N_{grad}N_{sb}(2+16NN_F)+$ $2N_{sb}N_FN(2N_{bit}\mathfrak{M}(\Delta+1)-1)/(\mathfrak{M}N_{bit})$
Detection	$4NN_F-1$

The conditional PDF of the DF-aided STE's output real part denoted as $\hat{x}_{R,m} = \Re(\hat{x}_{R,m})$ can now be expressed as

$$p(\hat{x}_{R,m}|x_{\mathfrak{M}\Delta+m}=+1) = \frac{1}{N_{sb}\sqrt{2\pi\sigma_n^2 w_m^H w_m}} \sum_{\check{x}_m^{(i)}\in\hat{\mathcal{X}}_m^+} \exp\left(-\frac{(\hat{x}_{R,m}-\check{x}_{R,m}^{(i)})^2}{2\sigma_n^2 w_m^H w_m}\right), \tag{3.74}$$

where we have $\check{x}_{R,m}^{(i)} = \Re\left(\check{x}_m^{(i)}\right)$ and $N_{sb} = \frac{N_s}{2}$ is the number of legitimate channel output states recorded in the absence of noise in the set $\hat{\mathcal{X}}_m^+$. Following the discussion of Section 3.4.2, the BER can now be obtained by integrating each of the conditional PDF terms seen in Equation (3.74) in order generate the error probability expression of

$$P_e(w_m) = \frac{1}{N_{sb}} \sum_{\check{x}_m^{(i)}\in\hat{\mathcal{X}}_m^{(+)}} \mathcal{Q}\left(g_i(w_m)\right), \tag{3.75}$$

where we have

$$g_i(w_m) = \frac{\mathtt{sign}\left(\check{x}_{\mathfrak{M}\Delta+m}^{(i)}\right)\check{x}_{R,m}^{(i)}}{\sigma_n\sqrt{W_m^H W_m}}. \tag{3.76}$$

The MBER DF-STE solution is now defined as

$$W = \arg\min_{W\in\mathbb{C}^{NN_F\times\log_2(\mathcal{M})\mathfrak{M}}} P_e(W). \tag{3.77}$$

As outlined in Section 3.3.2, the solution of problem can be found by taking the derivative of Equation (3.77) which may be calculated following the derivation of the BER gradient for the STE operating without DF, as presented in Section 3.3.2. The corresponding 4QAM MBER DF-STE formulae may be obtained following the derivation provided in Section 3.3.2 and applying it to the translated signal space, as illustrated for the BPSK case. The minimization of the BER cost-function can again be carried out using a suitable algorithm such as the conjugate gradient algorithm of Appendix B or a genetic algorithm of [130].

The computational complexity imposed by the detection of a single transmitted binary symbol according to Equation (3.77) is summarized in Table 3.6. The values have been obtained following the method employed in order to obtain the values summarized in Table 3.2, but assuming the reduced matrix dimensions facilitated by the DF. Similar to the complexity reduction achieved for the MMSE algorithm, the dependency of

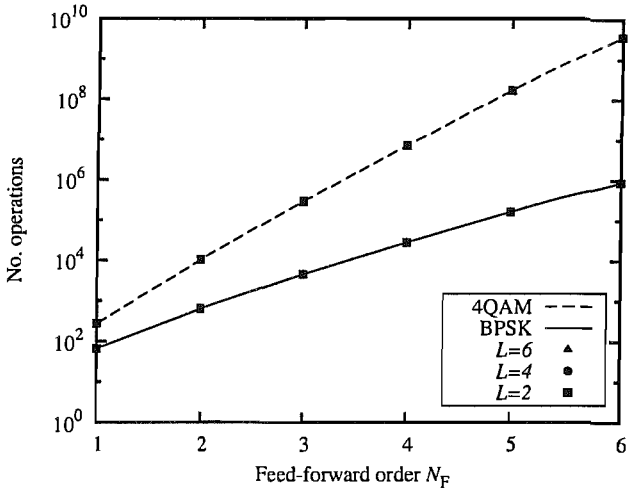


Figure 3.25: The normalized computational complexity required for evaluating the DF-STE MBER array weights as a function of the number of feed-forward order for different modulation schemes and a different number of CIR taps, when considering the detection of a single transmitted bit calculated according to Table 3.1. The proportionality factor was set to 1 and the total number of transmit AEs was chosen to be $\mathfrak{M} = 2$ while the number of receive AEs was $N = 2$. The curves associated with $L = 2, 4$ and 6 are indistinguishable.

the MBER STE's complexity on both the length of the CIR and the feed-forward order has been replaced by that on the decision delay Δ. This becomes particularly important, since the number N_{sb} of the STE's output states not benefiting from DF depends exponentially on the factor $(L + N_F - 1)$, whereas that operating with the advent of DF exponentially depends on the lower factor of $(\Delta + 1)$.

The reduced complexity of the DF-aided MBER STE suggested by Table 3.6 can be seen more explicitly in Figure 3.25, which shows the number of real-valued additions and multiplications required for determining the MBER DF-STE's weight matrix as a function of the feed-forward order N_F. The system considered for generating Figure 3.25 according to Table 3.6 supported $\mathfrak{M} = 2$ transmit AEs, $N = 2$ receive AEs as well as $L = 2, 4$ and 6 CIR taps. The decision delay was chosen to be $\Delta = N_F - 1$. When comparing the complexity imposed by the DF-aided STE to that of the STE dispensing with DF in Figure 3.11, it can seen that the DF-STE requires significantly less operations for the evaluation of the STE's weight matrix. Furthermore, it can be observed from Figure 3.25 that with the advent of DF, the complexity is no longer a function of the number of CIR taps L, since it is unambiguously determined by the feed-forward order of the STE.

Figure 3.26 shows the BER versus E_b/N_0 performance of our simple SDM system as characterized in the caption. It can be observed that the DF structure considerably improves the BER performance, but not as dramatically as for the MMSE algorithm of Figure 3.26. The difference between the curve associated with realistically detected error-prone feedback and the curve assuming that all the symbols which are fed back are correctly detected, decreases upon increasing the SNR. In fact, there is no discernible difference between the two curves at high SNRs, i.e at low BERs, which is a characteristic of the DF assisted MBER STE. A

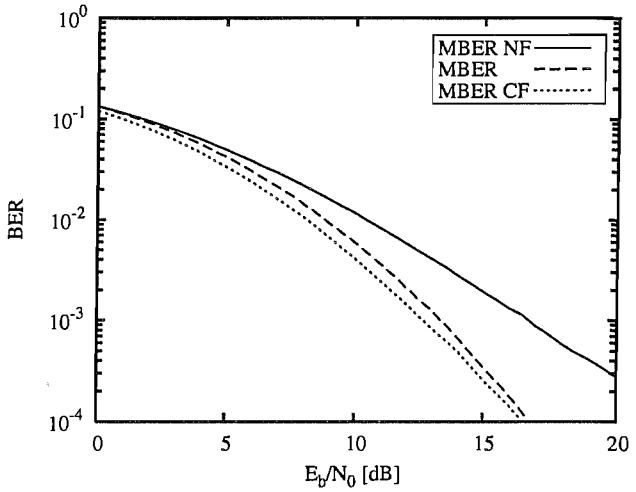


Figure 3.26: BER versus E_b/N_0 performance of the MBER DF-STE for a scenario supporting $Q = 1$ user employing $M = 2$ transmit AEs, a two-path equal-power independent Rayleigh channel and having perfect channel knowledge. The BS employed a two-element antenna array. The label NF indicates that no feedback was used while the label CF represents correct feedback.

similar robustness against error-propagation has also been observed for classic DF-aided MBER channel equalization [129].

3.5.3 Bayesian Decision-Feedback STE

In this subsection two different interpretations of the Bayesian DF aided STE based on the noise-free channel output states are presented in the context of the translated observation space $\mathbf{r}$ of Equation (3.65). Both the proposed space-translation approach and the sub-center [118] selection approach realize the same STE. To elaborate a little further, the sub-center selection technique is defined as follows: Depending on the feedback symbols sequence, the detector chooses a subset of $\mathcal{M}^{\mathfrak{M}(\Delta+1)}$ of the total $\mathcal{M}^{\mathfrak{M}(N_F+\Delta-1)}$ legitimate output states of the STE for the detection of the desired bits.

In the context of the space translation approach we proceed as already outlined in Section 3.5.2 for the MBER algorithm. Under the assumption of having a correct feedback vector $\mathbf{x}_2$, the Bayesian decision function given in Equation (3.45) transforms to

$$\begin{aligned} f_{B,m}(\mathbf{r}) &= p(\mathbf{r} \mid x_{\mathfrak{M}\Delta+m} = +1) - p(\mathbf{r} \mid x_{\mathfrak{M}\Delta+m} = -1) \\ &= \sum_{\check{\mathbf{r}}_q^{(i,+)} \in \mathcal{R}_m^+} p^{(i,+)} \frac{1}{(2\pi\sigma_n^2)^{NN_F}} \exp\left(-\frac{||\mathbf{r} - \check{\mathbf{r}}_m^{(i,+)}||^2}{2\sigma_n^2}\right) \\ &\quad - \sum_{\check{\mathbf{r}}_m^{(i,-)} \in \mathcal{R}_m^-} p^{(i,-)} \frac{1}{(2\pi\sigma_n^2)^{NN_F}} \exp\left(-\frac{||\mathbf{r} - \check{\mathbf{r}}_m^{(i,-)}||^2}{2\sigma_n^2}\right). \end{aligned} \tag{3.78}$$

Another way of interpreting the DF structure is to view it as a subset selection procedure. The binary subsets $\mathcal{Y}_m^+$ and $\mathcal{Y}_m^-$ given in Equation (3.35) for the STE operating without DF can be further partitioned into $N_{s,fb}$

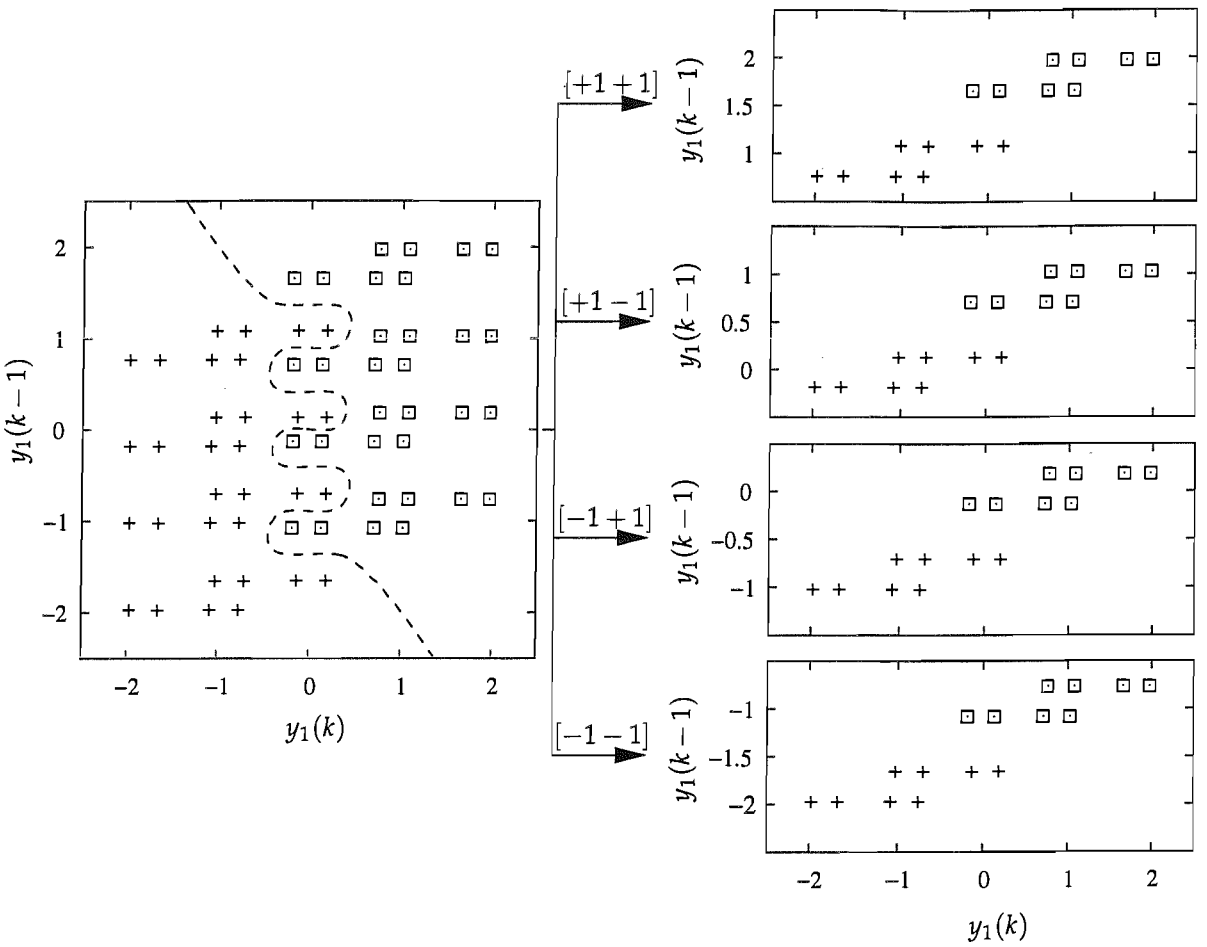


Figure 3.27: Sub-center selection

subsets $\mathcal{Z}_{\mathrm{m}}^{j,+}$ and $\mathcal{Z}_{\mathrm{m}}^{j,-}$ according to the $N_{\mathrm{s,fb}}$ possible feedback states $\check{\mathbf{x}}_{\mathrm{s,fb}}^{(j)}$, $1 \leq j \leq N_{\mathrm{sb}}$ of the feedback shift register, where we have

$$N_{\mathrm{s,fb}} = \mathcal{M}^{\mathfrak{M}N_{\mathrm{B}}}. \tag{3.79}$$

The union of all the $N_{\mathrm{s,fb}}$ feedback states associated with the two legitimate binary symbols transmitted by user m can then be written as

$$\mathcal{Y}_{\mathrm{m}}^{\pm} = \bigcup_{1 \leq j \leq N_{\mathrm{s,fb}}} \mathcal{Z}_{\mathrm{m}}^{j,\pm}, \tag{3.80}$$

where $\mathcal{Z}_{\mathrm{m}}^{j,\pm}$ is the set of legitimate values of the noise-free channel output vectors $\check{\mathbf{y}}$ associated with the delayed transmitted symbol $\check{x}_{\mathfrak{M}\Delta+\mathrm{m}}$ of the desired transmitter m and with the feedback symbol sequence $\check{\mathbf{x}}_2$. This yields

$$\mathcal{Z}_{\mathrm{m}}^{j,\pm} = \{\check{\mathbf{y}}(k) | \check{x}_{\mathfrak{M}\Delta+\mathrm{m}} = \pm 1 \wedge \mathbf{x}_2 = \check{\mathbf{x}}_2^{(j)}\}. \tag{3.81}$$

In this case, the receiver's feedback structure is used for selecting a reduced-size subset of legitimate channel output states for a given feedback vector $\check{\mathbf{x}}_2$ on the basis of the already decided symbols, since this allows us to reduce the detector's complexity. Explicitly, if the feedback vector is $\mathbf{x}_2 = \check{\mathbf{x}}_2^{(j)}$, then the reduced-

Table 3.7: The normalized computational complexity (real-valued additions and multiplications) of the DF-aided Bayesian STE when considering the detection of a single transmitted bit.

Calculation of the legitimate channel output states	$2N_sN_FN(2\mathfrak{M}(\Delta+1)-1)/N_{bit}/\mathfrak{M}$
Detection of a single bit	$N_s(6NN_F+1)$ plus N_s exp function evaluations

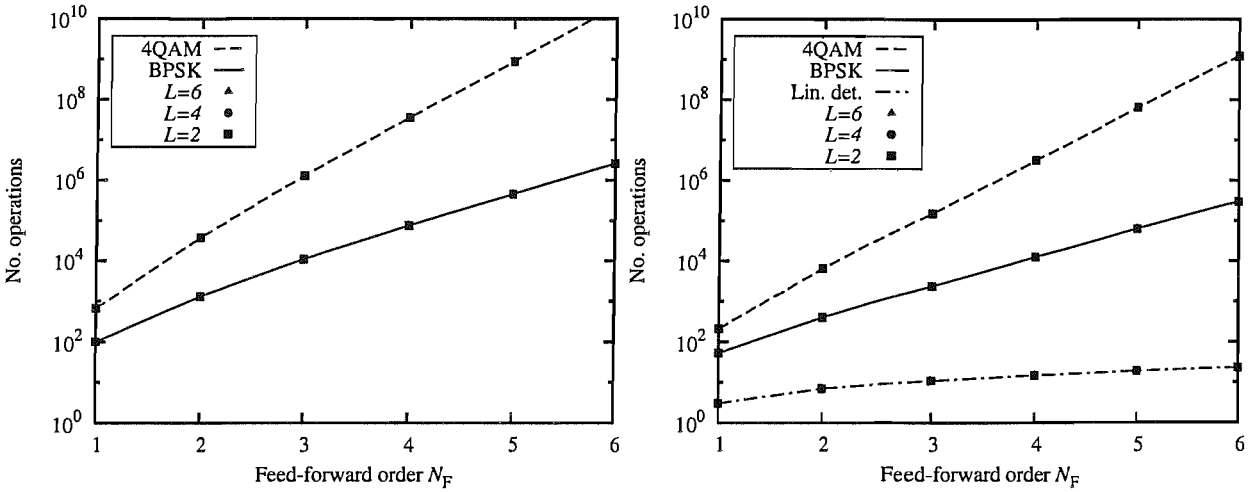
dimensional noiseless channel output vectors $\mathbf{r}_m^{(i,\pm)} \in \mathcal{Z}_m^{j,\pm}$ are selected as hypothetical channel output states for the detection of the desired symbol vector.

Example 3.5: Sub-center Selection DF aided STE

The space translation and the sub-center selection technique constitute different interpretations of the same Bayesian DF-STE. In Figure 3.22 the noise-free channel outputs of a DF assisted STE are illustrated for the rudimentary example used in Section 3.2. This phasor-diagram corresponds to Figure 3.3a recorded for the STE operating without DF. The space translation approach suggests that each received signal vector is transformed to the new observation space $\mathbf{r}$ before detection is performed. The hypothetical channel output states remain unchanged and each received signal vector is transformed depending on the specific feedback sequence already detected. The translated space $\mathbf{r}$ illustrated in Figure 3.22 has a reduced number of $N_s = 2^{m(\Delta+1)} = 16$ legitimate states.

By contrast, the sub-center selection based interpretation of the decision feedback divides the channel output space $\mathcal{Y}$ recorded in the absence of noise into $\mathcal{M}^{\mathfrak{M}N_B} = 4$ subsets, where $\mathcal{Z}_2^1$ corresponds to a previously detected feedback sequence $\mathbf{x}_1 = [+1\ +1]^T$, $\mathcal{Z}_1^2$ to $\mathbf{x}_2 = [+1\ -1]^T$, $\mathcal{Z}_1^3$ corresponds to $\mathbf{x}_2 = [-1\ +1]^T$ and $\mathcal{Z}_1^4$ to $\mathbf{x}_2 = [-1\ -1]^T$, as indicated in Figure 3.27. We can see that both interpretations of the decision feedback structure yield the same decision space to be considered by the STE.

The computational complexity associated with the DF aided Bayesian STE is summarized in Table 3.7, which was obtained on the basis of the complexity evaluation of the Bayesian STE dispensing with DF, as discussed in Section 3.4.3. Similarly to the DF assisted MBER STE, the most significant change in computational cost imposed by the detection is that the number of channel output states no longer depends exponentially on $L+N_F-1$, but rather on $(\Delta+1)$. In Figure 3.28a the computational cost associated with a system as described in the caption is illustrated as a function of the feed-forward order N_F. When comparing Figure 3.28a to Figure 3.13, which illustrates the complexity associated with a system not benefiting from DF, then it can be observed that the computational cost is reduced dramatically. However, the cost for detecting a single bit remains considerable higher than that of the linear detector as it can be seen from



(a) Number of real-valued additions plus multiplications required for the calculation of the legitimate channel output states plus the the real-valued operations required for the detection of a single bit.

(b) Number of real-valued additions plus multiplications required for the detection of a single bit once the legitimate channel output states have been determined. The complexity considered for the linear detectors only includes the detection and not the computation of the weight vector according to Table 3.5.

Figure 3.28: The normalized computational complexity imposed by the Bayesian DF-STE for the detection of a single transmitted bit as a function of the number of feed-forward order N_F evaluated for different numbers of CIR taps L as well as different modulation schemes. The values were calculated according to Table 3.7 and the scenario considered supported $QM = \mathfrak{M} = 2$ transmit as well as $N = 2$ receive AEs. The curves associated with $L = 2, 4$ and 6 are indistinguishable.

Figure 3.28b.

Considering the same SDM system it can be observed in Figure 3.26 characterizing the achievable BER versus E_b/N_0 performance that the feedback structure improves the BER performance only slightly. The difference between the curve associated with realistically detected feedback and the curve assuming that all the symbols which are fed back are correctly detected is almost negligible.

3.5.4 Reduced Complexity Maximum Likelihood Decision-Feedback STE

The reduced complexity ML DF-STE concerned may be defined using two different approaches similar to the Bayesian STE, which we refer to as the space-translation based and the tree-based DF structure. The space-translation based DF-aided OHRSA STE may readily be implemented by applying the OHRSA algorithm to the translated decision space $\mathbf{r}$.

A second way of incorporating DF in the proposed receiver structure is to include the symbols, which have already been decided upon in the top section of the search tree of Figure 3.15 on a bit-by-bit basis. The binary real-valued system incorporating DF may be characterized as

$$\underline{y} = \underline{H}_1 \underline{x}_1 + \underline{H}_2 \underline{x}_2 + \underline{\eta}. \tag{3.82}$$

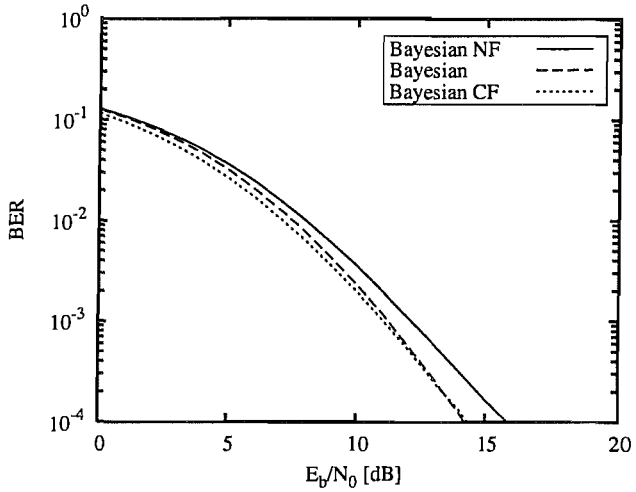


Figure 3.29: BER versus E_b/N_0 performance of the Bayesian STE for a scenario supporting $Q = 1$ user employing $M = 2$ transmit AEs, a two-path equal-power independent Rayleigh channel and having perfect channel knowledge. The BS employed a two-element antenna array. The label NF indicates that no feedback was used while the label CF represents correct feedback.

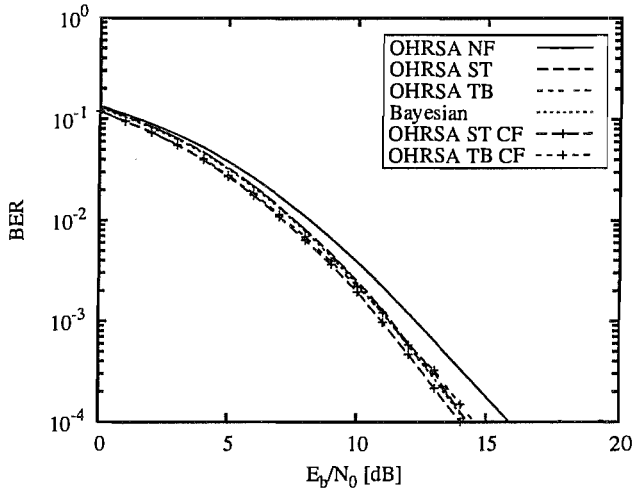


Figure 3.30: BER versus E_b/N_0 performance of the OHRSA DF-STE for a scenario supporting $Q = 1$ user employing $M = 2$ transmit AEs, a two-path equal-power independent Rayleigh channel and having perfect channel knowledge. The BS employed a two-element antenna array. The label NF indicates that no feedback was used while the label CF represents correct feedback.

It can now be seen from Equation (3.82) that allocating the feedback symbols to the top section of the search tree may be ensured by invoking the bit-wise CIR sub-matrix $\underline{\mathbf{H}}_1$ for carrying out the reordering according to the squared norm of the columns and leaving the bit-wise CIR matrix associated with the feedback bits $\underline{\mathbf{H}}_2$ unchanged of its columns yielding $\underline{\mathbf{H}}^{(o)} = [\underline{\mathbf{H}}_1^{(o)}\ \underline{\mathbf{H}}_2]$.

When evaluating the OHRSA outlined in Section 3.4.4, the last $\log_2(\mathcal{M})N_B\mathfrak{M}$ number of symbols of the hypothetical solution $\check{\mathbf{x}}$ are set to the bit-wise feedback sequence $\underline{\mathbf{x}}_2$. The feedback reduces the size of the search space and thus reduces the computational cost imposed by searching through the entire detection space.

Figure 3.30 shows the BER versus E_b/N_0 for a SDM system characterized in the caption. It can be seen that

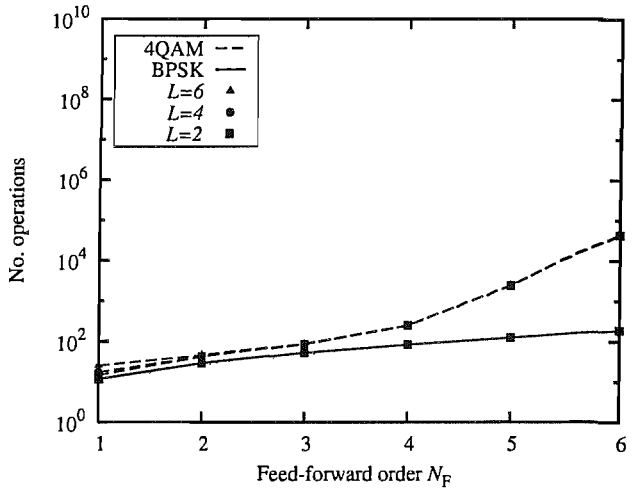


Figure 3.31: The normalized computational complexity required for evaluating the search tree as a function of the feed-forward order for different modulation schemes and a different number of MIMO array weights, when considering the detection of a single transmitted bit recorded during simulations at SNR=20 dB. The curves associated with $L = 2, 4$ and 6 are indistinguishable.

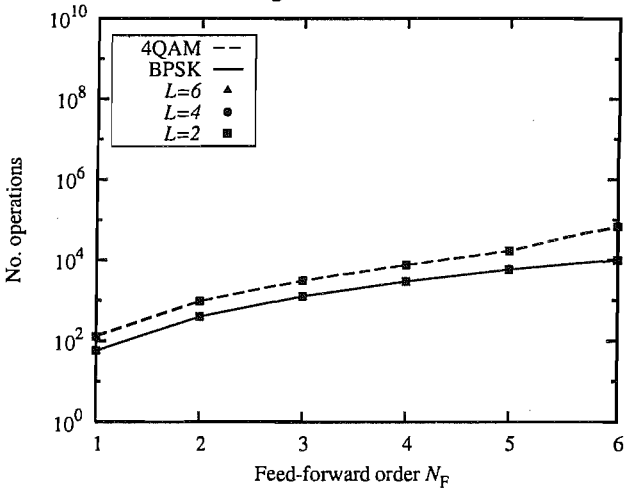


Figure 3.32: The normalized computational complexity required for evaluating the search tree plus the evaluation of the MMSE solution as a function of the feed-forward order for different modulation schemes and a different number of MIMO array weights, when considering the detection of a single transmitted bit at SNR=20 dB. The proportionality factor associated with the MMSE solution was set to 3.The curves associated with $L = 2, 4$ and 6 are indistinguishable.

the feedback structure improves the achievable BER performance only slightly, as also experienced for the Bayesian detector. It can also be observed that there is no performance difference between the two possible ways of incorporating feedback into the OHRSA, namely between the space-translation and the tree-based methods. Both methods perform identically and both exhibit the same robustness against DF-induced error-propagation.

Figure 3.32 shows the computational cost associated with the DF-aided OHRSA STE as a function of the feed-forward order N_F for a system described in the caption. Similar to the other DF-aided STEs discussed in Section 3.5.1-3.5.3 also the OHRSA STE imposes a reduced computational cost, if DF is employed. Again, the computational cost is no longer a function of the number of channel taps L and the resultant

complexity imposed by the OHRSA DF-STE is significantly lower than that associated with the DF-aided Bayesian STE of Figure 3.28.

In contrast to the application of the OHRSA for the detection of narrowband OFDM sub-carriers as in Chapter II of [60], the detection of all symbols hosted by the vector $\mathbf{x}$ is not required in the context of wideband systems, since we mentioned earlier that we are only interested in the detection of the desired sub-vector $\boldsymbol{x}_{\Delta+1}$ of Equation (3.46).

An attractive way of exploiting this potential complexity reduction technique to arrange the columns of the CIR matrix $\mathbf{H}$ not only according to the ascending order of received signal energy of the columns as suggested in [61], but also by ensuring that the columns of specific interest appear at the end of the ordered CIR matrix $\underline{\mathbf{H}}^{(o)}$. In other words, we aim for ensuring that the bits associated with $\underline{\boldsymbol{x}}_{\Delta+1}$ appear at the top of the search tree and are therefore detected first.[3]

Let us assume that in order to incorporate a DF structure in the receiver, the space-translation approach is considered. The new CIR matrix re-ordering scheme invoked for beneficially arranging the CIR matrix $\underline{\mathbf{H}}_1$ can then be divided into two steps as follows:

1. Reorder the last $\mathfrak{M}\log_2(\mathcal{M})$ number of columns of the bit-wise feed-forward CIR matrix given by

$$(\underline{\mathbf{H}}_1)_i \; i \in \{\log_2(\mathcal{M})\Delta\mathfrak{M}+1,\ldots,\log_2(\mathcal{M})(\Delta+1)\mathfrak{M}\}$$

   according to their energy in ascending order. This ensures that the symbols of interest contained by the desired sub-vector $\boldsymbol{x}_{\Delta+1}$ are detected first according to the energy of the associated columns. The associated symbols appear in the top section of the search tree of Figure 3.15.

2. Reorder the remaining columns of the CIR matrix according to the ascending energy order of the columns.

From a physical point of view, the CIR matrix columns having a very low energy may contribute little to the final cost-function value. Hence, the search tree of Figure 3.15 is subject to further branching at its bottom levels due to the low-power, high-delay CIR taps, which has little influence on the final cost-function value. This imposes a substantial increase in required computational complexity since the cost-function contribution of each of these branches has to be evaluated. Due to the fact that the desired symbols appear now in the top section of the search tree, a certain number of search tree levels at the bottom might be neglected, in order to avoid the evaluation of the cost function associated with the tree-branches induced by the high-delay, low-power CIR taps.

[3] If the tree-based feedback method is employed, the desired but yet undetected symbols are arranged so that they appear directly after the feedback symbols.

The recursive cost-function of Equation (3.56) can then be rewritten as

$$J_i(\check{\mathbf{x}}_i) = J_{i+1}(\check{\mathbf{x}}_{i+1}) + \phi(\check{\mathbf{x}}_i),\ i = N_b - 1, \ldots, N_{\text{trunc}} + 1, \tag{3.83}$$

where N_{trunc} indicates the number of layers of the search tree that have been discarded. Note however that this tree-truncation will always result in a controllable performance degradation, because even though we are not interested in the final decision concerning some of the ISI induced symbols, they still mildly influence the channel's output vector $\mathbf{y}$ and the corresponding cost-function, as it becomes explicit from Equation (3.48). Under certain circumstances, however, when the power associated with these symbols is very low (i.e. the associated column of the channel matrix exhibits a low energy), their influence becomes marginal and hence they might be judiciously neglected.

Let us now define the number of columns of the ordered bit-wise CIR matrix satisfying the condition of

$$\text{SIR}_i < \gamma \tag{3.84}$$

as N_γ, where we have

$$\text{SIR}_i = \frac{||\underline{\boldsymbol{h}}_i^{(o)}||^2}{N2\sigma_n^2 + \sum_{j \neq i} ||\underline{\boldsymbol{h}}_j^{(o)}||^2} \tag{3.85}$$

and γ is a BER versus complexity control factor. The number of levels in the search tree, which are neglected is then given as

$$N_{\text{trunc}} = \begin{cases} N_{\max} & \text{if } N_\gamma \geq N_{\max} \\ N_\gamma & \text{otherwise} \end{cases}, \tag{3.86}$$

where $N_{\max}$ has to be chosen for ensuring that the specific levels of the search tree, which are associated with the desired symbols are not truncated, yielding $N_{\max} = N_b - \log_2(\mathcal{M})\mathfrak{M}$.

3.5.5 Turbo-Coded Performance

For the evaluation of the DF-assisted receiver we consider the same parameter setting, as in Section 3.4.5, where the different algorithms were compared without DF for $Q = 1$, $M = 2$, $N = 2$, $N_F = 2$, $\Delta = 1$ and for an equal-power two-tap independent Rayleigh channel. Additionally, we opted now for the feedback order of $N_B = 1$.

The FER of BPSK modulated signals is illustrated in Figure 3.33 for all the different algorithms considered, including the performance bound constituted by the outage probability. It can be observed that the BER curves of the OHRSA as well as the Bayesian DF assisted STE are approximately 5 dB from the theoretical limit. The MBER algorithm experiences an additional 2 dB performance loss. Finally, the MMSE based

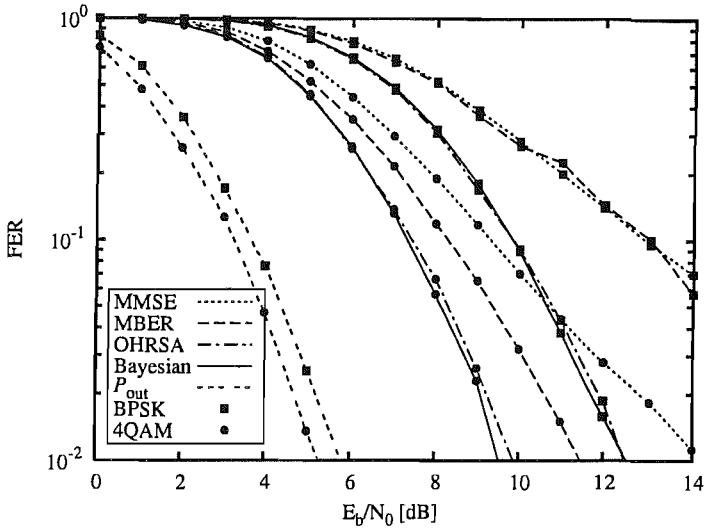


Figure 3.33: FER versus E_b/N_0 performance of the MMSE, MBER, Bayesian and OHRSA DF-STE for a scenario supporting $Q = 1$ user employing $M = 2$ transmit AEs, a two-path equal-power independent block-fading Rayleigh channel, a BS associated with $N = 2$, $N_F = 2$, $\Delta = 1$ and having perfect channel knowledge. The system employed a $R = 1/2$-rate punctured turbo-code.

detector operates approximately 9 dB away from the outage capacity at a code-word error rate of 10^{-2}.

When considering 4QAM, it can be observed in Figure 3.33, that the difference between the MBER and the MMSE algorithm is only marginal. Furthermore, it was observed during our simulations that the MBER algorithm had to be carefully initialized and even then may have experienced convergence problems under certain channel conditions. The FER performance of the MBER DF-STE recorded for 4QAM signals in Figure 3.33 was obtained using a genetic algorithm [130] rather than the simplified conjugate gradient algorithm of Appendix B.

Let us now change the assumption of experiencing slow or block-fading, where the fading amplitude was constant over an entire transmission burst to the other extreme choice, namely to that of encountering uncorrelated fading. The corresponding results shown in Figure 3.34 were obtained for a frame-length of 1000 bits and the same SDM settings as used for the evaluation of the block-constant fading were considered.

The employment of the MBER algorithm was not considered for this rapidly time-varying scenario owing to its high computational complexity. For the other algorithms a similar E_b/N_0 discrepancy can be observed with respect to the theoretical capacity bound as that recorded for the block-fading a scenario of Figure 3.33. Again, the discrepancy observed in Figure 3.34 with respect to the capacity bound is smaller for BPSK modulated signals than for 4QAM.

All STEs discussed so far have been based on using a finite observation interval for the received signal vector. It is well documented however, that the longer the observation interval and the decision delay, the better the STEs perform, resulting in the optimum detector setting, when considering an infinite observation interval.

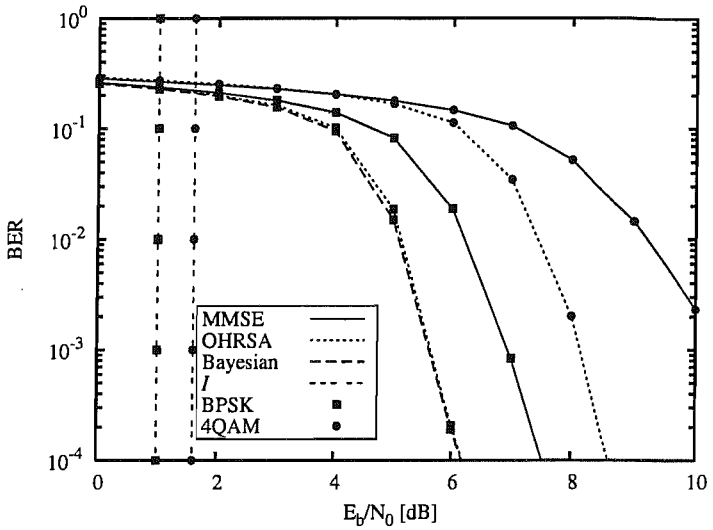


Figure 3.34: BER versus E_b/N_0 performance of the MMSE, MBER, Bayesian and OHRSA DF-STE for a scenario supporting $Q = 1$ user employing $M = 2$ transmit AEs, a two-path equal-power independent uncorrelated Rayleigh channel, a BS associated with $N = 2$, $N_F = 2$, $\Delta = 1$ and having perfect channel knowledge. The system employed a $R = 1/2$-rate punctured turbo-code.

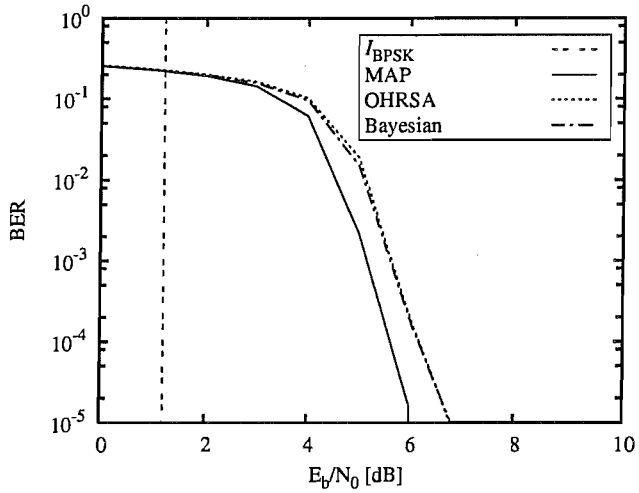
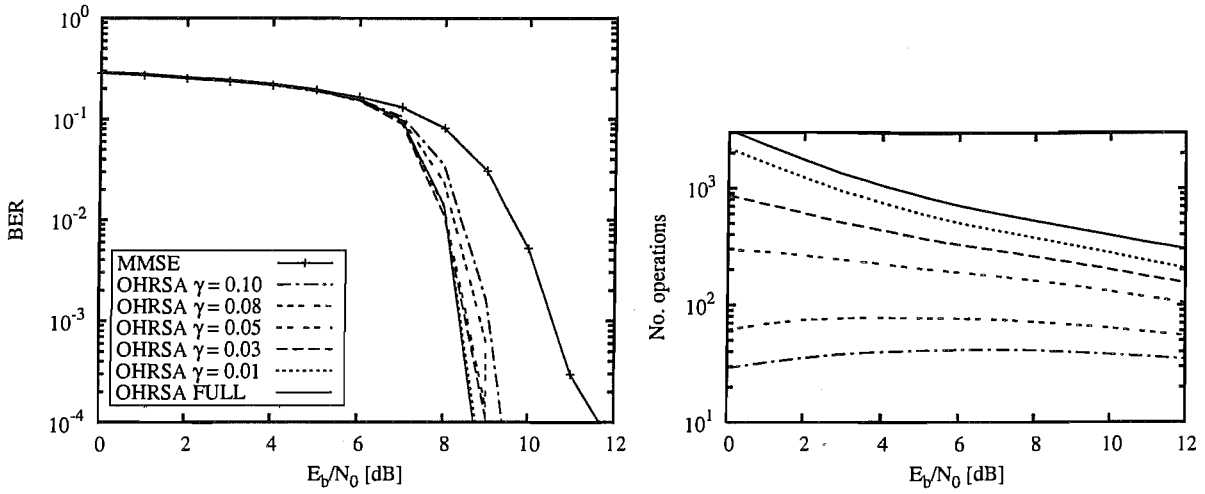


Figure 3.35: BER versus E_b/N_0 performance of the Bayesian and the OHRSA DF-STE as well as the MAP STE for a scenario supporting $Q = 1$ user employing $M = 2$ transmit AEs, a two-path equal-power independent Rayleigh channel and having perfect channel knowledge. The BS employed a two-element antenna array. The Bayesian and the OHRSA DF-STE used $N_F = 2$, $\Delta = 1$, $N_B = 1$.

In Figure 3.35 the BER versus E_b/N_0 performance of the MAP STE using an observation interval spanning over the complete data-frame is illustrated for the same scenario as used in Section 3.4.5 for BPSK modulated signals. The trellis associated with the MAP STE was constructed as outlined in Appendix C. The performance of both the Bayesian and the OHRSA DF-STE are plotted as a reference. It can be seen that the MAP based STE slightly outperforms both finite length STEs slightly at the cost of a potentially more complex block processing of the received symbols sequence.

The systems considered above were of moderate complexity in order to allow us to compare the different algorithms. In the remainder of this subsection we will investigate the performances of the OHRSA and the MMSE DF-STE, when considering higher throughput communication scenarios.



(a) BER versus E_b/N_0.

(b) Number of real-valued operations required for the detection of one bit versus E_b/N_0.

Figure 3.36: A scenario supporting $Q = 1$ user employing $M = 2$ uplink transmit AEs, a four-path equal-power ergodic independent fading channel, a BS equipped with $N = 2$, $N_F = 4$, $\Delta = 3$ and benefiting from perfect channel knowledge. 4QAM modulation was considered.

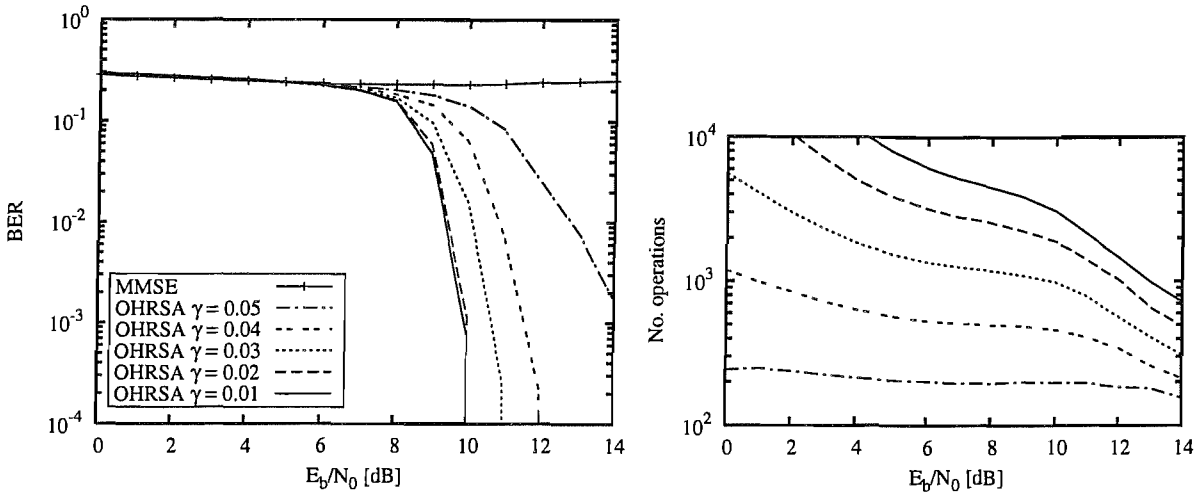
Let us now consider a propagation scenario having four equal-power paths and a DF-STE associated with $N_F = 4$, $\Delta = 3$, $N_B = 3$ and $N = 2$ receive antennas. Uncorrelated fast fading was assumed for the simulations. A single MS $Q = 1$ employing $M = 2$ transmit AEs and 4QAM was considered. For this modest-size system the full ML detector would have to consider $N_s = 4^{2\cdot4} = 65\,536$ states for the detection of each bit.

Figure 3.36a shows the average BER versus E_b/N_0 performance of the system for different values of γ. It can be observed that the reduced-complexity OHRSA using tree-truncation outperforms the MMSE based detector and exhibits only a modest performance degradation in comparison to the full-complexity OHRSA.

Let us now quantify the computational complexity of the detector, which is defined as the number of real-valued operations required for the evaluation of Equation (3.56) normalized by the $\mathfrak{M}\log_2(\mathcal{M})$ number of binary symbols detected[4]. The computational complexity associated with the BER performance curves of Figure 3.36a is shown in Figure 3.36b. It can be seen that especially at low Signal to Noise Ratio (SNR) values, the complexity of the full OHRSA algorithm is substantially lower than that of the full ML detector, but nonetheless, it becomes excessive. With the advent of truncation we are however capable of eliminating the large complexity variations as a function of the SNR. It can also be seen that the number of operations imposed is dramatically reduced, when applying tree truncation.

The associated high number of operations required for the conventional OHRSA assisted DF-STE at low SNR values is due to the fact that the low power associated with some columns of the CIR matrix results in numerous extra branches at the bottom of the search tree, all of which have to be considered by the algorithm. It is exactly this set of low-power tree branches, which can be truncated without any significant

[4]The complexity associated with the decomposition of the covariance matrix in not considered.



(a) BER versus E_b/N_0.

(b) Number of real-valued operations required for the detection of one bit versus E_b/N_0.

Figure 3.37: A scenario supporting $Q = 2$ user employing $M = 2$ uplink transmit AEs, a four-path equal-power ergodic independent fading channel, a BS equipped with $N = 3$, $N_F = 4$, $\Delta = 3$ and benefiting from perfect channel knowledge. 4QAM modulation was considered.

loss of performance.

If detection is performed with the aid of the truncated OHRSA, rank-deficient systems may also be considered. Adding another MS and an additional receive AE at the BS to the above-mentioned system results in a propagation scenario having four equal-power paths and a DF-STE associated with $Q = 2$, $N_F = 4$, $\Delta = 3$, $N_B = 3$ and $N = 3$ receive AEs as well as 4QAM. For the joint detection of all MSs' signals, the full ML or Bayesian DF aided STE would have to consider $N_s = 4^{4\cdot4} = 4\,294\,967\,296$ states. Assuming, that the full ML detector's complexity is proportional to the number of states N_s, the proposed detector is capable of achieving a complexity reduction, which is several orders of magnitude.

In Figure 3.37a the BER versus E_b/N_0 performance of the OHRSA as well as of the MMSE based DF-STE is shown for different values of γ. It can be seen that the MMSE based STE fails, whereas the OHRSA aided detector is capable of identifying all MSs' signals. The complexity reduction owing to truncation of the search tree is considerable, as it becomes clear from Figure 3.37b, which shows the number of numerical operations required for the detection of a single bit. Again, at low SNRs the truncation of the search tree assists in preventing the detector from becoming excessively complex. At higher SNRs this complexity reduction however imposes a modest performance degradation.

So far the performance of different STE algorithms was only investigated under the assumption of perfect channel knowledge. In the next section a set of different channel estimation algorithms will be employed, which will assist in investigating the effect of channel estimation errors on the attainable performance of the discussed STEs.

3.6 Channel Estimation for Dispersive Channels

Let us commence with an alternative form of the system Equation (3.6) representing the channel matrix as a vector, on which basis different channel estimation algorithms are introduced. Let us define the channel determining the STE input signal in form of a $(N\mathfrak{M}L)$-element vector, which is given as

$$\mathbf{h} = \mathtt{vec}\left([\boldsymbol{H}_1 \ldots \boldsymbol{H}_L]\right) \tag{3.87}$$

$$= [h_{1,11}, \ldots, h_{1,N1}, \ldots, h_{1,NQ}, \ldots, h_{L,11}, \ldots, h_{L,N1}, \ldots, h_{L,N\mathfrak{M}}]^T, \tag{3.88}$$

where $\mathtt{vec}\,(\cdot)$ represents the column-wise vector operator. To form the desired system equation, the symbol vector influencing the channel output is expressed as a matrix given as

$$\mathbf{X} = \mathbf{x}^T \otimes \mathbf{I}_N, \tag{3.89}$$

where $\otimes$ is the Kronecker product. The system matrix given in Equation (3.6) can now be re-written as

$$\boldsymbol{y}(k) = \mathbf{X}(k)\boldsymbol{h}(k) + \boldsymbol{\eta}(k). \tag{3.90}$$

The channel estimation algorithms discussed in the context of narrowband systems of Section 2.4 can now be applied directly to STE as depicted in Figure 2.27.

3.6.1 Results Using Estimated Channel State Information

To evaluate the effect of channel estimation errors on the different DF aided STE algorithms introduced earlier, a system consisting of Q number of MSs, each employing $M = 2$ number of transmit AEs and a BS using $N = 2$ number of receive AEs was considered. The DF assisted STE employed by the BS was characterized by $N_{\mathrm{F}} = 2$, $\Delta = 1$ and $N_{\mathrm{F}} = 1$. The channel was assumed to be have two equal power path and its variation over time was characterized by a normalized Doppler frequency $\overline{f}_{\mathrm{D}} = 0.0001$. The training length per transmit AE was chosen to be 150 symbols and the payload per transmit antenna was chosen to be 900 symbols. If Decision Directed Channel Estimation (DDCE) was used, the channel estimate was updated with a delay Δ assuming the decision made by the DF aided STE was correct. In other words, when operating in decision directed mode, is was assumed by the STE that $\boldsymbol{h}(k) = \boldsymbol{h}(k - \Delta - 1)$.

In Figure 3.38 the performance of the MMSE algorithm under consideration of estimated CSI is illustrated. It can be seen that if no channel tracking is used the algorithm is unable to operate adequately.

If however DDCE is switched on, it can be observed that the performance using RLS and first order Kalman

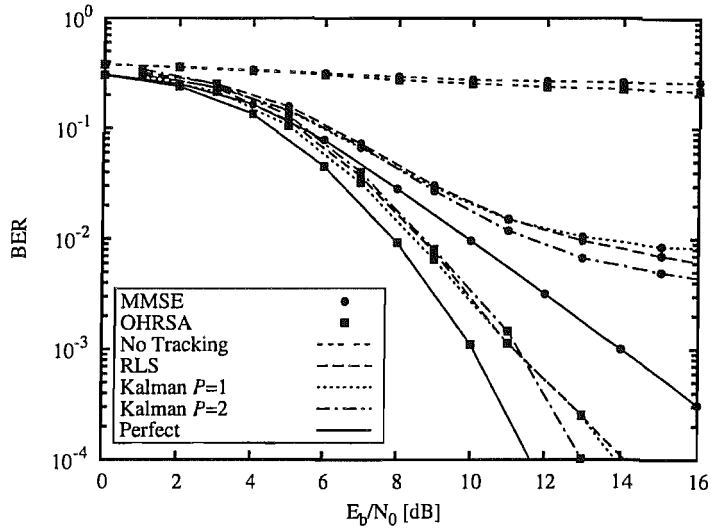


Figure 3.38: BER versus E_b/N_0 performance of the MMSE, MBER, Bayesian and OHRSA DF-STE for a scenario supporting $Q = 1$ user employing $M = 2$ transmit AEs, a two-path equal-power independent block-fading Rayleigh channel, a BS associated with $N = 2$, $N_F = 2$, $\Delta = 1$ using estimated CSI. The system employed a $R = 1/2$-rate punctured turbo-code and the Doppler frequency was $\overline{f}_D = 0.0001$. BPSK modulated signals were considered.

channel estimators is slightly worse than that of a STE benefiting form a second order Kalman channel estimator. Compared to the genie aided STE using perfected CSI, the error-floor of the receiver is increased significantly. This is due to the fact that at high SNR the variance of the channel estimation error exceeds the noise-variance and thus results in a degraded BER performance.

In contrast to the relatively large performance degradation experienced by the MMSE algorithm, the BER performance associated with the OHRSA aided STE is relatively little degraded by channel estimation errors as it can be seen in Figure 3.38. Especially at higher SNR the second order Kalman channel estimator outperforms the first order Kalman and the RLS channel estimator aided STE. If no DDCE is used the OHRSA based STE can not detect the transmitted signal correctly.

3.7 Conclusion

In this chapter the narrowband detection strategies discussed in Chapter 2 have been further developed to STE algorithms resulting in the MMSE STE of Section 3.4.1, the MBER STE of Section 3.4.2, the Bayesian STE of Section 3.4.3 as well as the OHRSA aided STE of Section 3.4.4. In order to improve the achievable performance and to reduce the computational complexity associated with the different STEs a DF structure has been introduced in Section 3.5. The different optimization functions associated with the STEs discussed as well as their associated complexity expressed as the number of real-valued multiplications pus additions is given in Table 3.8.

As it becomes clear from Table 3.8, the MMSE algorithm retains its low computational complexity over both

slow as well as rapid fading channels. The MBER algorithm however has a low computational complexity for slowly fading channels, whereas in rapid fading environments its computational cost increases exponentially with the number of CIR taps, with the number of bits per symbol as well as with the feed-forward order. For slow fading channels the MBER algorithm has the capability of outperforming the MMSE STE, as shown in Table 3.8.

The nonlinear Bayesian detector of Section 3.4.3 constitutes the optimum minimum error-rate non-linear receiver, which significantly outperforms the linear MMSE and MBER STE. The associated computational cost however increases exponentially with the number of CIR taps as well as with the number of bits per symbol and with the number of transmit AEs $\mathfrak{M}$ for all the channel conditions considered in Table 3.8. We have therefore introduced a reduced complexity ML detector in Section 3.4.4, which achieves the performance of the Bayesian STE at a reduced computational cost, as evidenced by Table 3.8.

When quantitatively comparing the different detectors' performance, we investigated their performance in terms of their distance expressed in terms of dB with respect to the theoretical channel capacity limits discussed in Section 3.3. Under the assumption of block-fading channels the distances $\Delta_{P_{\text{out}}|P_{\text{out}}=0.02}$ [dB] from the outage probability bound given in Table 3.8 were obtained for an outage probability of $P_{\text{out}} = 2\%$ from Figure 3.19 as well as Figure 3.33.

Under the assumption of uncorrelated fading channels the distance from the channel capacity limit in $\Delta_{I|\text{BER}=10^{-4}}$ [dB] given in Table 3.8 were obtained for a BER of 10^{-4} from Figure 3.20 as well as Figure 3.34. The values of $\Delta_{P_{\text{out}}|P_{\text{out}}=0.02}$ and $\Delta_{I|\text{BER}=10^{-4}}$, respectively, show that the performance difference between the non-linear and the linear STEs is more pronounced, when no DF is employed.

It can also be seen from Table 3.8 that when employing DF, the achievable performance of the MMSE STE is improved and its complexity is also reduced marginally. The MBER, the Bayesian as well as the OHRSA aided STE exhibit a significantly lower performance improvement with the advent of the DF structure. However, their complexity was reduced more substantially than for the MMSE STE.

Furthermore, different channel estimators have been employed in order to investigate the STEs performance using imperfect CSI. It was shown that for slowly time-varying channels the OHRSA detector experiences only a slight performance degradation, which is on the order of 1.5 dB for a (2×2) system, as it can be seen from Figure 3.38.

Table 3.8: Summary of the STEs studied in Chapter 3.

			MMSE of Section 3.4.1	**MBER** of Section 3.4.2	**Bayesian** of Section 3.4.3	**OHRSA** of Section 3.4.4
Optimization			$\hat{\mathbf{x}} = \boldsymbol{W}^H\mathbf{y}$, where $\boldsymbol{W}=\arg\min_{\boldsymbol{W}\in\mathbb{C}^{N\times M}} \mathrm{E}[\|\mathbf{x}-\boldsymbol{W}^H\mathbf{y}\|^2]$	$\hat{\mathbf{x}}=\boldsymbol{W}^H\mathbf{y}$, where $\boldsymbol{W}=\arg\min_{\boldsymbol{W}\in\mathbb{C}^{N\times M}} P_e(\boldsymbol{W})$	$\hat{x}_{\mathfrak{M}\Delta+\mathfrak{m}} = \arg\max_{x_{\mathfrak{M}\Delta+\mathfrak{m}}\in\{+1,-1\}} p(\mathbf{y}, x_{\mathfrak{M}\Delta+\mathfrak{m}})$	$\hat{\mathbf{x}} = \arg\max_{\hat{\mathbf{x}}\in\mathcal{X}} p(\mathbf{y}\|\mathbf{x})$
Complexity [a]	block-fading	No DF	$O(NN_F)$	$O(NN_F)$	$O(2^{N_{bit}\mathfrak{M}(L+N_F-1)})$	$O(NN_F) < \bullet < O(2^{N_{bit}\mathfrak{M}(L+N_F-1)})$
		DF	$O(NN_F)$	$O(NN_F)$	$O(2^{N_{bit}\mathfrak{M}(\Delta+1)})$	$O(NN_F) < \bullet < O(2^{N_{bit}\mathfrak{M}(\Delta+1)})$
	slow fading	No DF	$O(NN_F)$	$O(NN_F)$	$O(2^{N_{bit}\mathfrak{M}(L+N_F-1)})$	$O(NN_F) < \bullet < O(2^{N_{bit}\mathfrak{M}(L+N_F-1)})$
		DF	$O(NN_F)$	$O(NN_F)$	$O(2^{N_{bit}\mathfrak{M}(\Delta+1)})$	$O(NN_F) < \bullet < O(2^{N_{bit}\mathfrak{M}(\Delta+1)})$
	rapid fading	No DF	$O\left(\frac{N^3N_F^3}{\mathfrak{M}N_{bit}}\right)$	$O(2^{N_{bit}\mathfrak{M}(L+N_F-1)})$	$O(2^{N_{bit}\mathfrak{M}(L+N_F-1)})$	$O\left(\frac{N^3N_F^3}{\mathfrak{M}N_{bit}}\right) < \bullet < O(2^{N_{bit}\mathfrak{M}(L+N_F-1)})$
		DF	$O\left(\frac{N^3N_F^3}{\mathfrak{M}N_{bit}}\right)$	$O(2^{N_{bit}\mathfrak{M}(\Delta+1)})$	$O(2^{N_{bit}\mathfrak{M}(\Delta+1)})$	$O\left(\frac{N^3N_F^3}{\mathfrak{M}N_{bit}}\right) < \bullet < O(2^{N_{bit}\mathfrak{M}(\Delta+1)})$
Performance [b]	$\Delta_{P_{out}\|P_{out}=0.02}$ [dB]	No DF	> 10	6.2	4.8	4.8
	for block-fading	DF	8.2	6.0	4.5	4.5
	$\Delta_{I\|BER=10^{-4}}$ [dB]	No DF	8.3	not evaluated	5.5	5.5
	for uncorrelated fading	DF	6.5	not evaluated	5.2	5.2

[a] Values obtained from Table 3.1-3.7

[b] Distance from the theoretical capacity limit obtained from Figure 3.19 and Figure 3.20 as well as Figure 3.33 and Figure 3.34.

Chapter 4

Iterative Narrowband Single-Carrier Multiple-Input Multiple Output Detection

4.1 Introduction

Iterative processing was introduced by Berrou in [131, 132] in the context of iteratively decoding two parallel concatenated convolutional codes referred to as a turbo-code. His work has later been extended to serially concatenated codes [133] and then found its way gradually into iterative detector designs, such as for example iterative equalizers [134] or iterative multi-user detectors [135]. Strictly speaking the term 'turbo' refers to the detection of parallel concatenated convolutional codes. All other schemes are referred to as iterative schemes. This definition is also adopted in this treatise. In the literature, however, the term 'turbo' is often also used in the context of diverse iterative detector or decoding schemes [118].

In this chapter we will investigate the performance of a single-user narrowband MIMO system employing iterative detection. The remainder of this chapter is organized as follows. In Section 4.2 the system model including the iterative detector is introduced and the information flow during the iterative detection process is outlined. In Section 4.3 EXtrinsic Information Transfer-function (EXIT) Charts are introduced, which are employed throughout the chapter for analyzing the characteristic behaviour of different Soft-Input Soft-Output (SISO) detectors. The algorithms introduced in Chapter 2 for non-iterative MIMO detection are extended in Section 4.4 to iterative SISO detectors and their performance is analyzed using EXIT charts.

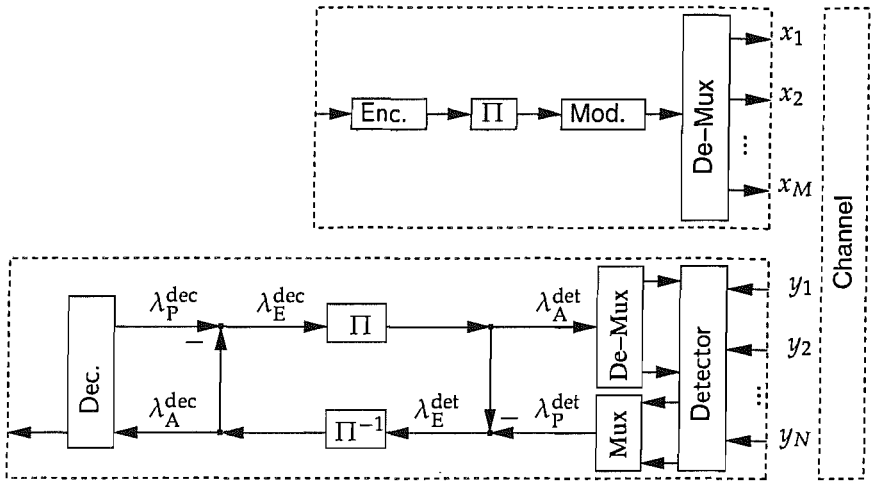


Figure 4.1: Baseband MIMO system model for iterative detection.

4.2 System Model

The transmitter of the iterative system is identical to that used in Chapter 2 for the non-iterative MIMO system where a single MS employing an M-element transmit antenna array and a BS receiver, which has N number of AEs was considered. The MS's transmitter channel encodes the input bit-stream at a code-rate R_c, passes the encoded bits through an interleaver, modulates the signal and maps the symbols to the M different transmit AEs, as illustrated in Figure 4.1. The modulated symbols are transmitted to the BS over a narrowband independent Rayleigh fading channel characterized by the channel coefficients h_{nm}, as illustrated in Figure 4.1.

The joint iterative detection scheme employed by the BS is depicted in Figure 4.1. In the context of iterative detection the information generated by the different receiver components is exchanged between them in the form of Log Likelihood Ratios (LLRs), which are defined as [136]

$$\lambda_m(k) = \lambda(\underline{x}_m(k)|\underline{y}(k)) = \log\left(\frac{P(\underline{x}_m(k) = +1|\underline{y}(k))}{P(\underline{x}_m(k) = -1|\underline{y}(k))}\right), \tag{4.1}$$

where $\underline{x}_m(k)$ and $\underline{y}(k)$ are the transmitted and received signal vector, respectively, when considering the real-valued binary system model of Section 2.1.1.1 given as

$$\underline{y}(k) = \underline{\boldsymbol{H}}(k)\underline{x}(k) + \underline{\eta}(k). \tag{4.2}$$

The transmitted symbol index m ranges from 1 to MN_{bit}, where M is the number of transmit AEs and N_{bit} is the number of bits per symbol, as discussed in Section 2.1.1.1.

Let us now denote a vector of LLR values as

$$\lambda(k) = [\lambda_1(k), \ldots, \lambda_{N_{\text{bit}}M}(k)]^T, \tag{4.3}$$

as well as the superscript $^{\text{det}}$ and $^{\text{dec}}$ indicating that the LLR value is associated with the detector and the channel decoder, respectively. In our further discussions the explicit conditioning on the channel output as well as the time index k are neglected for notational simplicity and accordingly we define

$$\lambda_m = \lambda(\underline{x}_m) = \lambda(\underline{x}_m(k)|\underline{y}(k)). \tag{4.4}$$

Considering now the joint detection of all M AEs' signals, the detector first performs a soft detection of the transmitted signal vector $\boldsymbol{x}(k)$ on the basis of the received signal $\boldsymbol{y}(k)$ and returns the *a posteriori* LLR of the interleaved channel coded bits $\lambda_{\text{P}}^{\text{det}}$, as seen in Figure 4.1, where the superscript $^{\text{det}}$ represents the detector. The *a priori* LLR $\lambda_{\text{A}}^{\text{det}}$ generated from the extrinsic LLR $\lambda_{\text{E}}^{\text{dec}}$ by the interleaver Π of Figure 4.1 is then removed from the *a posteriori* LLR $\lambda_{\text{P}}^{\text{det}}$, resulting in the extrinsic information $\lambda_{\text{E}}^{\text{det}}$, which is passed through the deinterleaver denoted by Π^{-1} in Figure 4.1 to the channel decoder. The channel decoder carries out a soft decision using the deinterleaved extrinsic information provided by the detector as *a priori* information $\lambda_{\text{A}}^{\text{dec}}$, where the superscript $^{\text{dec}}$ denotes the channel decoder. After convolutional decoding the decoder calculates the *a posteriori* LLR $\lambda_{\text{A}}^{\text{dec}}$ of the coded symbols and subtracts the *a priori* LLR $\lambda_{\text{A}}^{\text{dec}}$ in order to obtain the extrinsic information $\lambda_{\text{E}}^{\text{dec}}$, as seen in Figure 4.1. The extrinsic information of the decoder is interleaved again and used by the detector as *a priori* information $\lambda_{\text{A}}^{\text{det}}$ for the next iteration. In the first iteration the detector assumes identical *a priori* probabilities for all bits, yielding $\lambda_{\text{A}}^{\text{det}} = 0$ for all transmitters. For a more detailed description of iterative detection the interested readers are referred to [118].

4.3 Extrinsic Information Transfer Chart Analysis of SISO Components

In order to analyze the characteristic behaviour of SISO detectors we highlight the concept of EXIT charts, which were first introduced by ten Brink in [137] for the analysis of turbo codes. EXIT charts depict the mutual information I_{E} at the output of a SISO component as a function of the mutual information associated with the *a priori* information I_{A} at the input of the SISO component. Let us therefore introduce the mutual information between a transmitted binary symbol x and the associated *a priori* information as [137]

$$I_{\text{A}} = \frac{1}{2} \sum_{x \in \{+1,-1\}} \int_{-\infty}^{\infty} p_{\text{A}}(\zeta|X = x) \cdot \log_2 \frac{2p_{\text{A}}(\zeta|X = x)}{p_{\text{A}}(\zeta|X = -1) + p_{\text{A}}(\zeta|X = 1)} \partial\zeta, \tag{4.5}$$

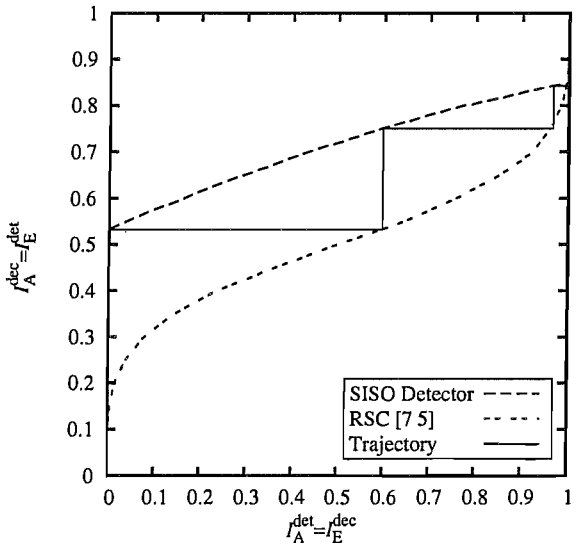


Figure 4.2: EXIT chart example for an iterative detector.

where p_A is the PDF of the *a priori* LLRs. The mutual information associated with the extrinsic LLR values at the output of the SISO component is similarly defined as [137]

$$I_E = \frac{1}{2} \sum_{x \in \{+1,-1\}} \int_{-\infty}^{\infty} p_E(\zeta|X=x) \cdot \log_2 \frac{2p_E(\zeta|X=x)}{p_E(\zeta|X=-1)+p_E(\zeta|X=1)} \partial\zeta, \tag{4.6}$$

where p_E is the PDF of the extrinsic LLR values. The integration required for the calculation of the mutual information in Equation (4.5) and Equation (4.6) can be carried out numerically and the required distribution of the LLR values can be obtained using one of the following two methods [137] :

- Generate a histogram based approximation of $p_A(\zeta|X=x)$;
- Approximate the distribution of the LLR values using a single-parameter Gaussian distribution.

Note that EXIT charts are based on the assumption that the components exchanging extrinsic information have independent information. Furthermore, the statistics of all random variables are sufficiently accurately modeled. This implies that if the EXIT chart of a detector is recorded, it has to be ensured that fading statistics are accurately modeled. This unfortunately renders the classical EXIT chart analysis unsuitable for frame-invariant fading channels.

4.3.1 Two Concatenated Components

The simplest way of concatenating two SISO components is to use serial concatenation, as illustrated in Figure 4.1, where the superscript indicates the index of the component and the subscript indicates, whether the mutual information is associated with the *a priori* or the extrinsic information, as defined in Equation (4.5)

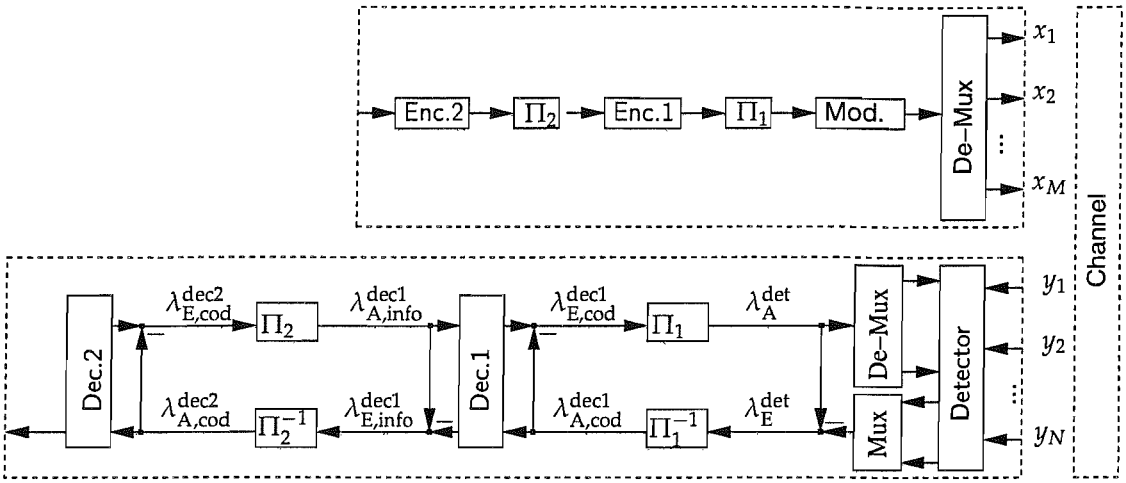


Figure 4.3: Baseband MIMO system model for iterative detection.

and Equation (4.6), respectively. In order to discuss the nature of EXIT charts, we consider a serially concatenated system consisting of a SISO detector and a SISO decoder associated with a Recursive Systematic Convolutional (RSC) code.

The exchange of information between the two system components is completely described by the evolution of the mutual information at the output of the SISO components as a function of the mutual information associated with the *a priori* LLRs. In order to illustrate this evolution, the EXIT chart transfer function of both receiver components is plotted in a joint graph using the output of one component as the input to the other component. In other words, the EXIT chart function of the second component is mirrored with respect to the first-quarter diagonal of the EXIT chart by swapping the two axis, as illustrated in Figure 4.2, where the label RSC [7 5] indicates the EXIT function of the RSC code using the octally represented generator polynomials $\boldsymbol{g} = [7\,5]$, for swapped axes.

The iterative process commences with an initial detection without any available *a priori* information, where we have $\lambda_{\mathrm{A}}^{\mathrm{det}} = 0\ \forall\ m$ and therefore $I_{\mathrm{A}}^{\mathrm{det}} = I_{\mathrm{E}}^{\mathrm{dec}} = 0$. In the second step, the $I_{\mathrm{A}}^{\mathrm{dec}} = I_{\mathrm{E}}^{\mathrm{det}}$ is provided as the input of the channel decoder. The decoder output is then fed back to the detector and so forth. This procedure is illustrated in Figure 4.2. The solid line resembling a stair case in Figure 4.2 illustrates the iterative decoding process. If the mutual information at the output of the decoder is one, the receiver is capable of operating at a BER approaching zero. Hence, it has to be the ultimate goal of iterative detector design to enable the trajectory to approach this point in the EXIT chart.

4.3.2 Three Concatenated Components

In many applications not only two, but rather three SISO blocks are concatenated, for example when a soft detector operating in conjunction with two serially concatenated codes is employed, as illustrated in

Figure 4.3. It can be seen that decoder 1 of Figure 4.3 takes two inputs, namely the *a priori* information associated with the information bits as well as the *a priori* information associated with the coded bits. Thus, in order to analyze the behaviour of such components, while aiming for their joint optimization, 3-Dimensional (3D) EXIT charts [138] would have to be employed. 3D-EXIT charts however have the drawback that they are often difficult to portray graphically and hence it is difficult to characterize the iterative behaviour of the detector.

Assuming however a specific iteration pattern, one might group two of the three SISO components of the detector into a joint component and represent its EXIT characteristic by a more readily interpreted two-dimensional function. In our specific case we are interested in characterizing the behaviour of different SISO detectors when employed in iterative receiver structures. Let us assume that the channel code employed by the detector is composed of two serially concatenated convolutional codes. For a complete characterization of the convergence behaviour a 3D-EXIT chart would have to be considered, as illustrated in Figure 4.4a, which shows a 3D-EXIT chart for the two serially concatenated convolutional codes specified in Table 4.1. The subscript $_{\text{info}}$ indicates that the mutual information is associated with the information bits, while the subscript $_{\text{cod}}$ indicates that the mutual information is associated with the coded bits. Furthermore, the mesh of rectangles indicated the EXIT-plane of decoder 1 and the mesh of crosses that of the outer-most decoder. The former appears parallel to the $I_{\text{E}}^{\text{det}}$ axis and hence is independent of this value. Decoder 2 is not directly linked to the detector of Figure 4.3 and hence the EXIT chart of channel code two (C2) only depends on $I_{\text{A,cod}}^{\text{dec2}} = I_{\text{E,info}}^{\text{dec1}}$, since no *a priori* information is available about the source bits. To elaborate a little further, it can be observed in Figure 4.4a that the EXIT characteristic of decoder 1 in the middle of Figure 4.3 depends on both the *a priori* information $I_{\text{A,info}}^{\text{dec1}} = I_{\text{E,cod}}^{\text{dec2}}$ provided by the outer-most decoder and on the *a priori* information $I_{\text{A,cod}}^{\text{dec1}} = I_{\text{E}}^{\text{det}}$ provided by the detector.

The relation between the 3D-EXIT charts and their conventional two-dimensional representations can also be understood by considering Figure 4.4b and Figure 4.4c, which illustrates the 2D projections of the 3D-EXIT chart. The side-view shown in Figure 4.4c portrays slices of the $I_{\text{E,info}}^{\text{dec1}} = I_{\text{A,cod}}^{\text{dec2}}$ versus $I_{\text{A,info}}^{\text{dec1}} = I_{\text{E,cod}}^{\text{dec2}}$ surface for different values of $I_{\text{E}}^{\text{det}}$. It can be observed that for $I_{\text{E}}^{\text{det}} < 0.6$ the EXIT functions of the two codes intersect and hence only a moderate iteration gain can be achieved. For $I_{\text{E}}^{\text{det}} > 0.6$, however, convergence to the $I_{\text{E,info}}^{\text{dec2}} = 1$ point becomes possible and as a result error-free transmissions are supported. Note again that only the first intersection of the two EXIT functions is relevant, since it determines how low a BER can

Table 4.1: Channel code parameters.

Encoder 2 (C2)	RSC Code with generator polynomial $\boldsymbol{g} = [6\ 5]$
Encoder 1 (C1)	Generator polynomial 1/(1+D)
Overall Code-Rate	0.5

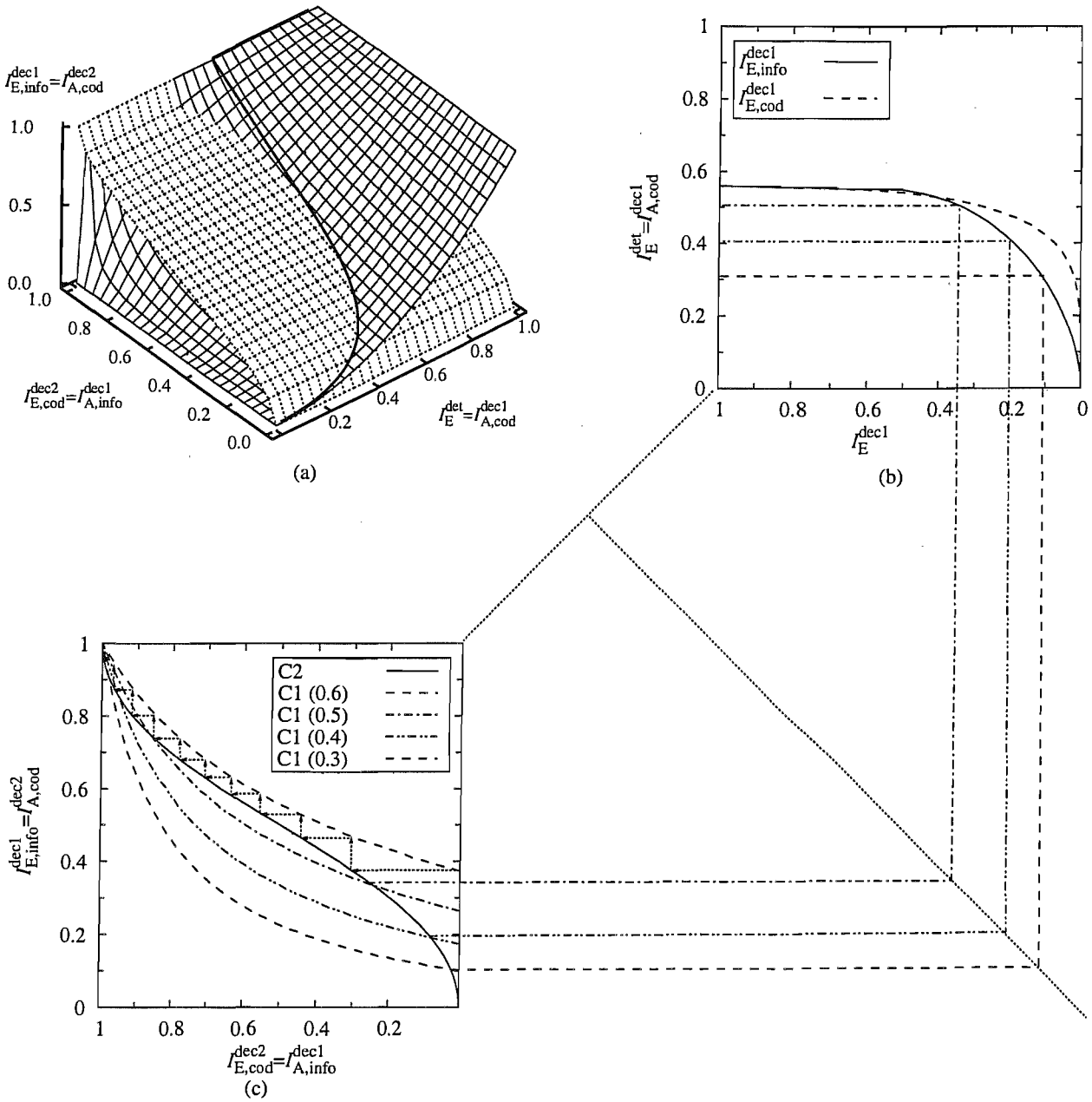


Figure 4.4: 3D- EXIT chart for a Rate-1 intermediate code (C1) and an outer half-rate RSC code (C2) together with its two dimensional projections. The exact code parameters are summarized in Table 4.1.

be achieved. In Figure 4.4c it can be observed that the curve associated with $I_E^{det} = 0.5$ and labeled C1 (0.5) intersects twice with the curve C2. This second intersection results in an overlap between the two 3D EXIT charts. In the 2D projection the 3D EXIT chart shown in Figure 4.4b we have illustrated the projection of the first intersection between the two 3D-EXIT curves.

Since we are interested in generating a two-dimensional EXIT function, which is independent of $I_{E,cod}^{dec2}$, we consider the special case when, $I_{E,cod}^{dec2}$ has converged to its maximum value for a given I_E^{det} value. This assumption implies that the iterations between the two decoders of Figure 4.3 are continued, until no further performance improvements can be achieved by further iterating. In the 3D-EXIT chart of Figure 4.4a this condition is represented by the *first* intersection of the two 3D-EXIT functions, where $I_{A,cod}^{dec2} = I_{E,info}^{dec1}$, which is marked by the bold curve. The area of Figure 4.4a, which is located left of the bold line and associated with the EXIT plane of code 1 represents the second intersection between the two 3D-EXIT planes.

In order to obtain the corresponding 2D-EXIT chart, we consider the *first* intersection trajectory of the two 3D-EXIT planes when commencing iterations from the point $I_{E,cod}^{dec2} = 0$ for a given I_E^{det} at the bottom of Figure 4.4a. This intersection is projected to the plane spanned by the coordinate $I_{E,info}^{dec1}$ and $I_{A,cod}^{dec1}$. The 2D-EXIT function is then obtained by interpolating the values of $I_{E,cod}^{dec1}$ for this projected intersection values. This interpolated 2D-EXIT function is illustrated in Figure 4.4b as the dashed line.

The 2D-views of Figure 4.4a seen in Figures 4.4c and 4.4b can be connected using the standard projection technique inherited from 3D blue-prints used in mechanical engineering. In order to interpret the classic 2D EXIT chart further, if, for example, the intersection between the two EXIT curves of Figure 4.4c associated with $I_E^{det} = 0.3$ is projected into Figure 4.4b, it can be observed that the intersection between the projection line and the projected EXIT curve occurs at exactly $I_E^{det} = 0.3$.

4.4 Soft-Input Soft-Output MIMO Detection

In this section the algorithms introduced in Section 2.3 for hard-decision outputs are extended to SISO detectors. All algorithms are presented for the real-valued binary system model of Section 2.1.1.1.

4.4.1 MMSE SISO Detection

The iterative MMSE detector has first been proposed in [29] for CDMA multi-user detection and BPSK modulated signals. The proposed scheme in-fact is not a classical SISO detector in the sense that it carries out a Maximum A Posteriori (MAP) decision, but rather uses the soft information provided by the channel decoder in order to perform soft interference cancellation. The soft-output of the MMSE detector is then produced upon stipulating the assumption that the output of the MMSE detector is Gaussian distributed.

The original iterative MMSE multi-user detector has in [85, 139] later been applied to channel equalization. Furthermore, the soft-interference cancellation scheme was extended to higher-order modulation schemes.

The MMSE SISO MIMO detector for the real-valued binary symbol $\underline{x}_m$ might be derived as follows. Assuming a given *a priori* probability soft estimate in the form of [136]

$$\overline{\underline{x}}_m = \tanh\left(\frac{\lambda_{\mathrm{A},m}^{\mathrm{det}}}{2}\right) \tag{4.7}$$

and defining the vector

$$\underline{\boldsymbol{u}}_m = [1-|\overline{\underline{x}}_1|^2,\ldots,1-|\overline{\underline{x}}_{m-1}|^2,0,1-|\overline{\underline{x}}_{m+1}|^2,\ldots,1-|\overline{\underline{x}}_{N_{\mathrm{bit}}M}|^2]^T \tag{4.8}$$

we can write the received signal vector cleaned with the aid of soft interference cancellation as

$$\begin{aligned}\underline{\boldsymbol{y}}_m &= \underline{\boldsymbol{y}} - \underline{\boldsymbol{H}}\underline{\boldsymbol{u}}_m \\ &= \underline{\boldsymbol{y}} - \underline{\boldsymbol{H}}\overline{\underline{\boldsymbol{x}}} + \overline{\underline{x}}_m\underline{\boldsymbol{h}}_m,\end{aligned} \tag{4.9}$$

where we have $\overline{\underline{\boldsymbol{x}}} = [\overline{\underline{x}}_1,\ldots,\overline{\underline{x}}_{MN_{\mathrm{bit}}}]^T$. The optimization criterion used for determining the weight vector can now be expressed as

$$\underline{\boldsymbol{w}}_m = \arg\min_{\underline{\boldsymbol{w}}_m\in\mathbb{R}^{2N\times 1}} \mathrm{E}\left[|\underline{x}_m - \underline{\boldsymbol{w}}_m^T(\underline{\boldsymbol{y}} - \underline{\boldsymbol{H}}\overline{\underline{\boldsymbol{x}}} + \overline{\underline{x}}_m\underline{\boldsymbol{h}}_m)|^2\right], \tag{4.10}$$

which in physical terms means that the MSE is minimized for the signal after soft-canceling all the interfering signals. Let us now define the output of the MMSE detector as

$$\begin{aligned}\hat{\underline{x}}_m &= \underline{\boldsymbol{w}}_m^T\underline{\boldsymbol{y}}_m \\ &= \underline{\boldsymbol{w}}_m^T(\underline{\boldsymbol{y}} - \underline{\boldsymbol{H}}\overline{\underline{\boldsymbol{x}}} + \overline{\underline{x}}_m\underline{\boldsymbol{h}}_m),\end{aligned} \tag{4.11}$$

then following the derivations in [29], the MMSE weight vector can be obtained as

$$\underline{\boldsymbol{w}}_m = (\underline{\boldsymbol{H}}\mathtt{diag}(\underline{\boldsymbol{u}})\underline{\boldsymbol{H}}^H + \sigma_{\mathrm{n}}^2\boldsymbol{I}_N)^{-1}\underline{\boldsymbol{h}}_m. \tag{4.12}$$

Assuming furthermore that $\hat{\underline{x}}_m$ obeys a Gaussian distribution with mean $\mu_{\mathrm{MMSE},m}$, and variance $\sigma^2_{\mathrm{MMSE},m}$

Table 4.2: The normalized computational complexity (real-valued additions and multiplications) of the SISO MMSE detector when considering the detection of a single transmitted bit.

Weight vector calculation of Equation (4.12)	$4NMN_{\text{bit}} + 8N^2MN_{\text{bit}} + 2N + O(2N^3)$
Detection according to Equation (4.13)	$14N + 4NN_{\text{bit}}M + 3$

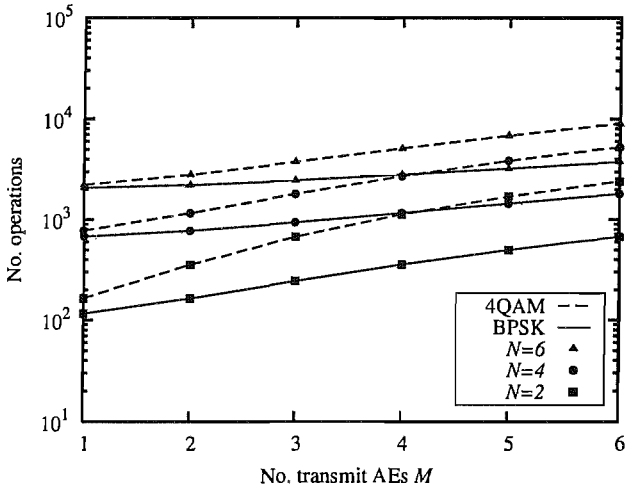


Figure 4.5: The normalized computational complexity required for evaluating the SISO MMSE array weights as a function of the number of transmit AEs for different modulation schemes and a different number of MIMO array weights, when considering the detection of a single transmitted bit calculated according to Table 4.2. The proportionality factor was set to 3.

the desired LLR value may be expressed as [29]

$$\begin{aligned} \lambda_{\text{E}}^{\text{det}}(\underline{x}_m) &= \frac{(\hat{\underline{x}}_m - \mu_{\text{MMSE},m})^2}{\sigma^2_{\text{MMSE},m}} - \frac{(\hat{\underline{x}}_m + \mu_{\text{MMSE},m})^2}{\sigma^2_{\text{MMSE},m}} \\ &= \frac{4\underline{w}^H(\underline{y} - \underline{\underline{H}}\underline{\bar{x}} + \bar{\underline{x}}_m\underline{h}_m)}{1 - \underline{h}_m^T\underline{w}_m}. \end{aligned} \tag{4.13}$$

When evaluating the computational complexity associated with the MMSE detector, we have to distinguish between the first and all subsequent iterations. During the first iteration the computational complexity imposed by iterative the MMSE detector is identical to that associated with the classic non-iterative MMSE detector dispensing with *a priori* information as discussed in Section 2.3.1.

The number of real-valued multiplications plus additions required for the evaluation of the weight vector according to Equation (4.12) can be shown to be $4NMN_{\text{bit}} + 8N^2MN_{\text{bit}} + 2N + O(2N^3)$, where the assumption was made that we have $(2A - 1) \approx 2A$. Once the weight vector is calculated according to Equation (4.12), the calculation of the LLR value in Equation (4.13) requires $14N + 4NN_{\text{bit}}M + 3$ real-valued additions plus multiplications.

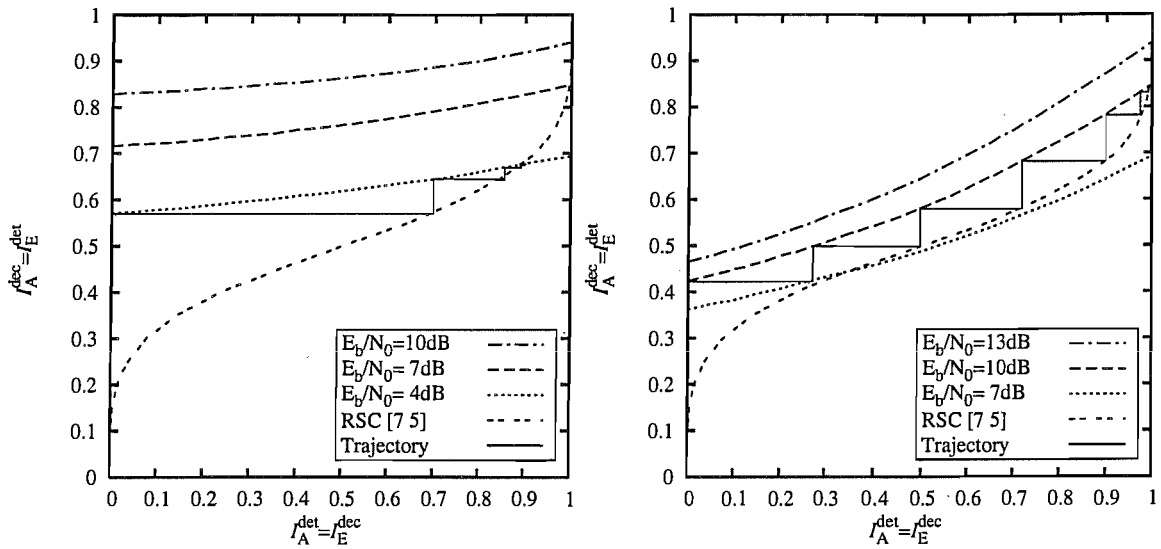
In Figure 4.5 we have illustrated the computational complexity versus the number of transmit AEs imposed by the MMSE detector benefiting from soft-information for a different number of receive AEs as well as for BPSK and 4QAM signals. It can be seen in Figure 4.5 that in contrast to the MMSE detector dispensing with

a priori information and only providing hard outputs, which was discussed in Section 2.3.1, the complexity of the SISO detector is slightly higher. Furthermore, the normalized computational complexity imposed by the SISO detector increases for 4QAM signals compared to BPSK modulated signals. This is owing to the fact that when using the real-valued binary system model of Section 2.1.1.2, calculating the soft-output for a given soft-input involves the matrix inversion of Equation (4.12) for each bit to be detected. The complexity illustrated in Figure 4.5 and summarized in Table 4.2 is only an upper bound. When taking into consideration the specific structure of the covariance matrix of Equation (4.12) its inversion can be carried out at a lower complexity using the matrix inversion lemma discussed in [29].

Furthermore, the complexity associated with the first iteration is lower than that of the subsequent iterations, since only one matrix inversion has to be performed for the detection of all bits, as it becomes clear, when setting the *a priori* soft information $\overline{x}_m$ of Equation (4.12) for all transmit AEs to zero. The weight vector for the first iteration is identical to the weight vector of the hard-output MMSE detector of Section 2.3.1.

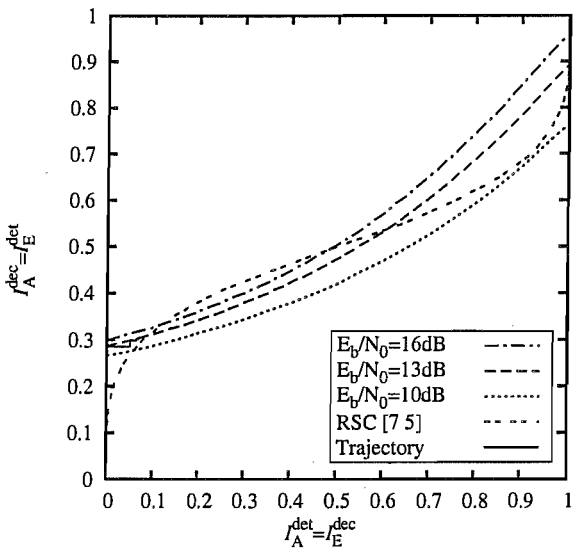
In Figure 4.6 the EXIT functions of an MMSE MIMO detector supporting different number of transmit AEs M, while employing $N = 2$ receive AEs are plotted. The modulation scheme considered was 4QAM and the channel was assumed to be an uncorrelated, independent Rayleigh fading channel with $\mathrm{E}\left[|h_{nm}|^2 = 1\right]$. Additionally, the EXIT function of a RSC code having an octally represented generator polynomial of $g = [7\ 5]$ is also illustrated. It can be seen from Figure 4.6a that for this non-overloaded system the EXIT function of the MMSE detector is rather flat and its starting point at $\lambda_{\mathrm{A}}^{\mathrm{det}} = 0$ increases upon increasing the SNR. The EXIT characteristic associated with four transmit AEs shown in Figure 4.6b has a similar behaviour as the two transmit AE assisted system only with a slightly increased slope of the EXIT function. Naturally a higher SNR is required to create a tunnel between the detector's and the decoder's EXIT function for allowing the detector to iterate. However, if the number of transmitters is increased to $M = 6$ transmit AEs as illustrated in Figure 4.6c, the starting point of the EXIT function at $\lambda_{\mathrm{A}}^{\mathrm{det}} = 0$ only increases marginally upon increasing the SNR. This clearly indicates that the MMSE system is interference limited for strongly rank-deficient systems. The detector is incapable of creating an open EXIT function tunnel and thus convergence to an infinitesimally low BER becomes impossible. Hence, the receiver exhibits an error-floor.

The BER versus E_{b}/N_0 performance for the same system as considered for generating Figure 4.6 is shown in Figure 4.7 for different numbers of iterations. It can be observed that for the case of $M = 2$ transmit AEs no further BER performance improvement is achieved after a few iterations. For four transmit AEs the number of iterations required for attaining convergence is increased, but convergence is still possible. The shape of the BER curve associated with four transmit AEs after $N_{\mathrm{iter}} = 7$ iterations can be explained by considering the EXIT chart of Figure 4.6b. In the region of $E_{\mathrm{b}}/N_0 = 5$ dB initial convergence is limited by the absence of an open EXIT tunnel. Around $E_{\mathrm{b}}/N_0 = 8$ dB a small increase in SNR suddenly opens up a



(a) $M = 2$ transmit AEs

(b) $M = 4$ transmit AEs



(c) $M = 6$ transmit AEs

Figure 4.6: MMSE EXIT function for a system using M number of transmit AEs and $N = 2$ of receive AEs. 4QAM was considered. Additionally, the EXIT function of a RSC code using the generator polynomial of $g = [7\ 5]$ is illustrated. The trajectories were recorded for an interleaver length of 10 000 bits.

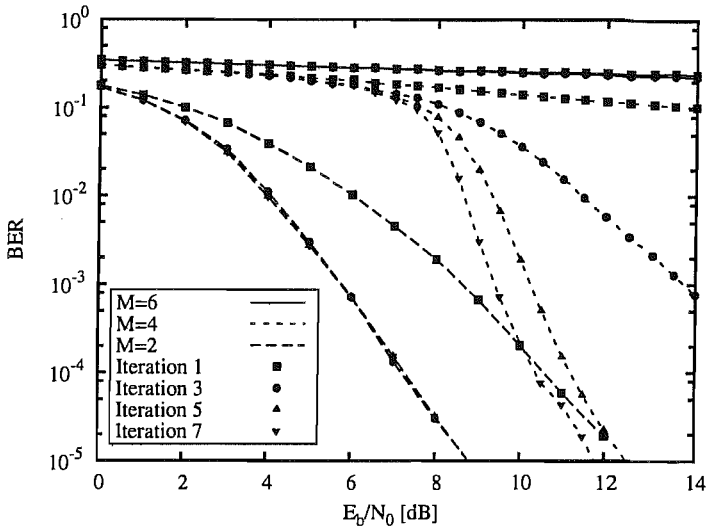


Figure 4.7: BER versus E_b/N_0 performance of a MMSE detector for a system using M number of transmit AEs and $N = 2$ receive AEs. 4QAM and an interleaver length of 10 000 bits were considered.

tunnel resulting in a large BER improvement in exchange for a relatively small E_b/N_0 increase. At about $E_b/N_0 = 10$ dB the achievable iteration gain is reduced again. Considering the case of $M = 6$ number of transmit AEs, it can be seen that the detector is incapable of detecting the transmitted signal, since the system is interference limited.

4.4.2 Bayesian SISO Detection

The Bayesian SISO detector is a manifestation of the exact MAP solution of the MIMO detection problem considered. Recalling the definition of the set of legitimate transmitted symbol vectors

$$\underline{\mathcal{X}} = \{\underline{\check{x}}^{(i)},\ 1 \leq i \leq N_s\} \tag{4.14}$$

and the associated legitimate channel output states

$$\underline{\mathcal{Y}} = \{\underline{\check{y}}^{(i)} = \underline{\underline{H}}\underline{\check{x}}^{(i)},\ 1 \leq i \leq N_s\} \tag{4.15}$$

of Section 2.3.2, the LLR associated with the m^{th} bit can be expressed as

$$\begin{aligned} \lambda_P^{\text{det}}(\underline{x}_m) &= \log\left(\frac{P(\underline{x}_m = +1 \mid \underline{y}(k))}{P(\underline{x}_m = -1 \mid \underline{y}(k))}\right) \\ &= \log\left(\frac{\sum_{\underline{\check{y}}_m^{(i,+)} \in \underline{\mathcal{Y}}_m^+} p^{(i,+)} \exp\left(-\frac{\|\underline{y}(k) - \underline{\check{y}}_m^{(i,+)}\|^2}{2\sigma_n^2}\right)}{\sum_{\underline{\check{y}}_m^{(i,-)} \in \underline{\mathcal{Y}}_m^-} p^{(i,-)} \exp\left(-\frac{\|\underline{y}(k) - \underline{\check{y}}_m^{(i,-)}\|^2}{2\sigma_n^2}\right)}\right), \end{aligned} \tag{4.16}$$

Table 4.3: The normalized computational complexity (real-valued additions and multiplications) of the SISO Bayesian detector when considering the detection of a single transmitted bit.

Evaluation of the legitimate channel output states of Equation (4.15)	$2N_sN(2MN_{bit}-1)/(MN_{bit})$
Euclidean distance calculation in Equation (4.16)	$N_s6N/(MN_{bit})$
Legitimate output states *a priori* information of Equation (4.18)	$(5MN_{bit}-1)N_s/(MN_{bit})$
LLR calculation without Euclidean distance calculation according to Equation (4.16)	$2N_s-1$

where $\underline{\check{y}}_i^{m,\pm} \in \underline{\mathcal{Y}}^{m,\pm}$ and $p^{(i,+)}$ as well as $p^{(i,-)}$ are the a-priori probabilities of $\underline{y}_m^{(i,+)}$ and $\underline{y}_m^{(i,-)}$, respectively. Assuming that the symbols in the sequence $\underline{x}$ are statistically independent of each other, the *a priori* probability of the channel state $\underline{\check{y}}^{(i)}$ can be obtained from the *a priori* bit LLRs as follows:

$$P(\underline{\check{y}}^{(i)}) = P(\underline{\check{x}}^{(i)}) \tag{4.17}$$

$$= \prod_{m=1}^{N_{bit}M} \frac{\exp\left(\lambda_A^{det}(\underline{x}_m^{(i)})/2\right)}{1+\exp\left(\lambda_A^{det}(\underline{x}_m^{(i)})\right)} \exp\left(\underline{\check{x}}_m^{(i)}\lambda_A^{det}(\underline{x}_m^{(i)})/2\right), \tag{4.18}$$

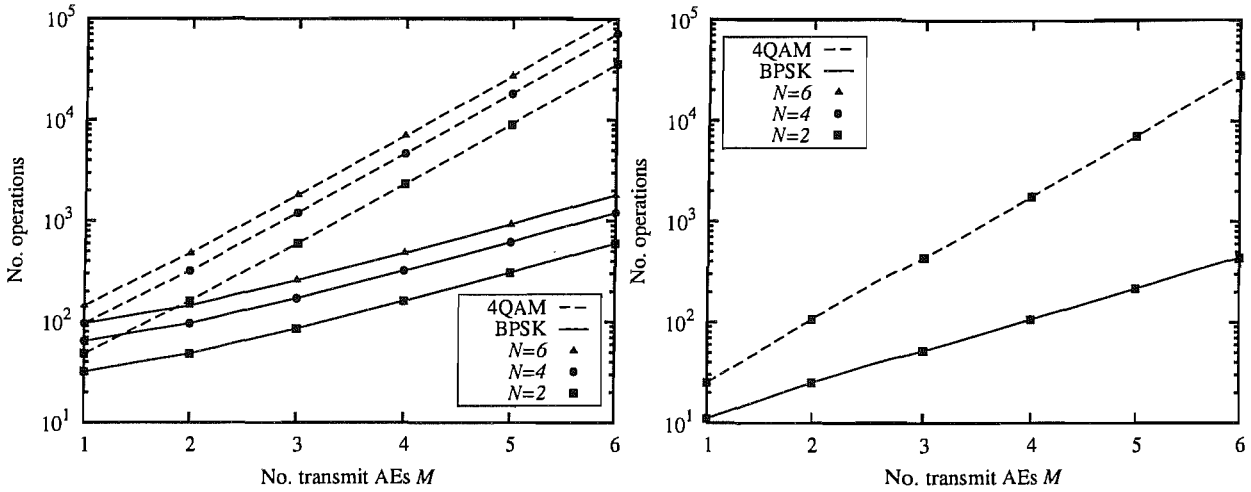
where $\lambda_A^{det}(\underline{x}_m^{(i)})$ is the *a priori* LLR of the binary symbol associated with the m^{th} AE. Equation (4.16) weights each possible channel output state with its *a priori* probability.

The computational complexity imposed by the Bayesian detector, which is capable of processing soft-information may be divided into four steps:

1. The calculation of the legitimate channel output states according to Equation (4.15).
2. The calculation of the *a priori* information of each state according to Equation (4.18).
3. The Euclidean distance evaluation of Equation (4.16).
4. The calculation of the LLR values according to Equation (4.16).

The calculation of the legitimate channel output states according to Equation (4.15) has been discussed in Section 2.3.3 and requires $2N_sN(2MN_{bit}-1)$ real-valued additions plus multiplications. The calculation of the legitimate output states' *a priori* information according to Equation (4.18) requires $(5MN_{bit}-1)N_s$ number of real-valued additions plus multiplications. Finally, the Euclidean distance evaluations of Equation (4.16), including the division by the noise power, require N_s6N number of real-valued operations. A further $(2N_s-1)$ number of operations are required in order to evaluate the desired LLR value according to Equation (4.16).

Note that the number of exp and log function evaluations have not been considered, since in practice Equation (4.16) as well as Equation (4.18) are evaluated in the logarithmic domain using the Jacobian approxima-



(a) Computational complexity imposed by the Euclidean distance calculations plus the calculation of the legitimate channel output states.

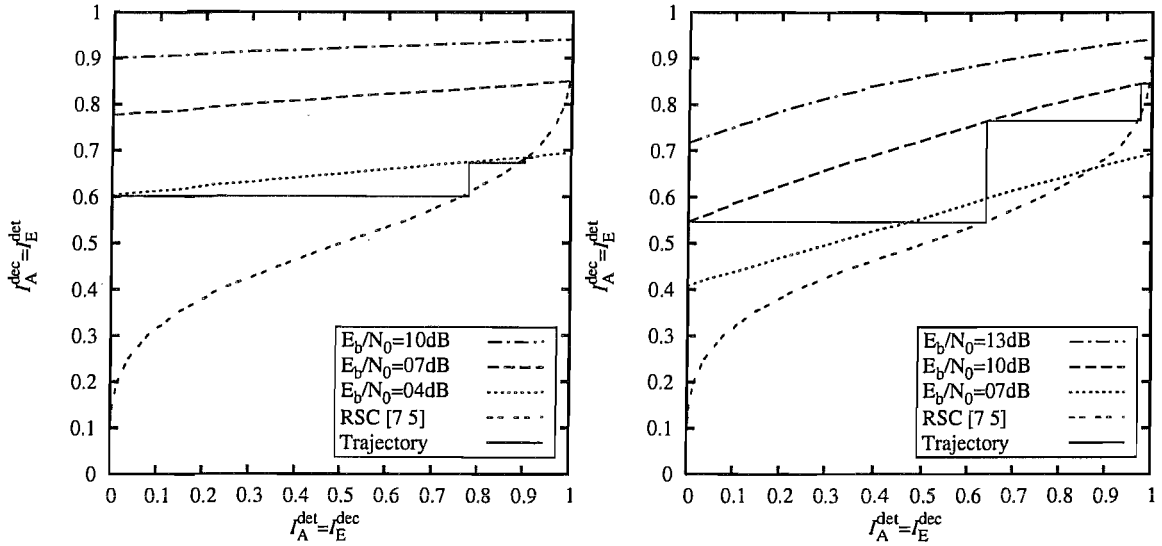
(b) Computational complexity required in order to evaluate the channel output states' *a priori* probability plus the final evaluation of the LLR values.

Figure 4.8: The normalized computational complexity imposed by the Bayesian SISO detector for the detection of a single transmitted bit as a function of the number of transmit AEs, evaluated for different modulation schemes and a different number of MIMO array elements. The values were calculated according to Table 4.3.

tion [136]. Hence, no exp evaluations are required. Furthermore, the evaluation of the legitimate channel output states as well as the Euclidean distance calculations are only necessary during the first iteration and then can be stored for subsequent iterations, which suggests that the total computational complexity is dominated by the contribution of the first iteration. In Table 4.3 we have summarized the normalized computational complexity imposed by the Bayesian detector.

In Figure 4.8 we have illustrated the computational complexity imposed by the Bayesian SISO detector, versus the number of transmit AEs for BPSK and 4QAM signals as well as for a different number of receive AEs. It can be observed that only the evaluation of the legitimate channel output states as well as the calculation of the Euclidean distances is dependent on the number of receive AEs. These operations only have to be performed during the first iteration. In all subsequent iterations the computational complexity is independent of the number of receive AEs, as it becomes clear from Figure 4.8b. Furthermore, it can be observed that for a low number of transmit AEs the Bayesian detector exhibits a lower computational complexity than the MMSE SISO detector of Section 4.4.1. However, the complexity associated with the Bayesian detector remains exponentially proportional to the number of transmit AEs.

In Figure 4.9 the EXIT functions plotted for a Bayesian MIMO detector using a different number of transmit AEs M and employing $N = 2$ receive AEs is shown. The modulation scheme considered was 4QAM and the channel was assumed to be an uncorrelated independent Rayleigh fading channel. It can be observed that the general behaviour of an increased starting point of the detector EXIT function at $\lambda_A^{det} = 0$ for an increased SNR is identical to that of the MMSE detector's EXIT charts of Figure 4.6. The difference



(a) $M = 2$ transmit AEs (b) $M = 4$ transmit AEs

Figure 4.9: Bayesian EXIT function for a system using M number of transmit AEs and a $N = 2$ receive AEs. 4QAM was considered. Additionally, the EXIT function of a RSC code using the generator polynomial of $\boldsymbol{g} = [7\ 5]$ is illustrated. The trajectories were recorded for an interleaver length of 10 000 bits.

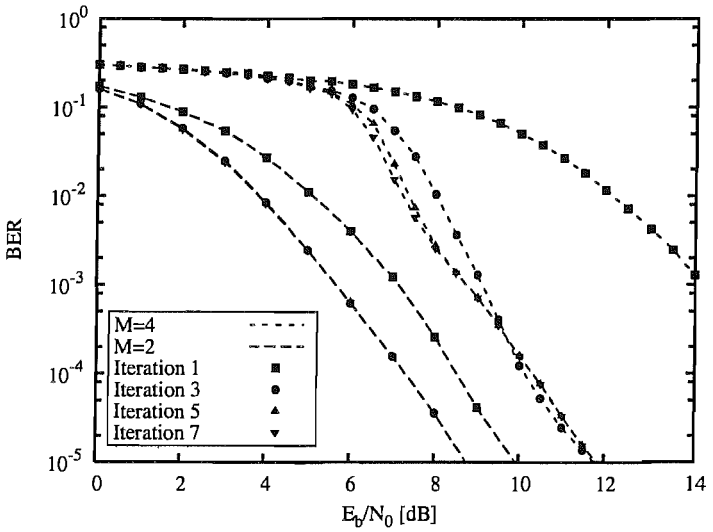


Figure 4.10: BER versus E_b/N_0 performance of a Bayesian detector for a system using M number of transmit AEs and $N = 2$ receive AEs. 4QAM and an interleaver length of 10 000 bits were considered.

however is that at an identical SNR the Bayesian detector's EXIT function emerges from a higher point and thus requires less iterations in order to converge. Furthermore, it is important to note that both detectors EXIT charts' terminate at the same point. This means that for *a priori* LLRs associated with the mutual input information of $\lambda_{\mathrm{A}}^{\mathrm{det}} = 1$ both detectors achieve the same BER performance. It does however not mean that both detectors exhibit an identical BER performance. Their BER will only be identical, if there exists an open EXIT tunnel between the detector's and the decoder's EXIT function, allowing the iterative process to converge.

The BER versus E_{b}/N_0 performance associated with the parameters used for generating Figure 4.9 are shown in Figure 4.10. It can be observed that the BER performance achieved by the iterative Bayesian detector is identical to that of the MMSE detector at a lower number of iterations. The results recorded for $M = 6$ transmit AEs are not shown here, since the complexity exhibited by the detector was excessive.

4.4.3 Reduced-Complexity Max-Log Detection

In this section a Reduced-complexity Max-Log (RML) algorithm designed for the detection of the transmitted signal is highlighted [61]. The Max-Log approximation of the desired LLRs required for the detection of the transmitted signal outlined in the context of Equation (4.1) can be formulated as

$$\lambda_{\mathrm{P}}^{\mathrm{det}}(\underline{x}_m) \approx \log\left(\frac{\max_{\underline{\check{x}}\in\underline{\mathcal{X}}^+} P(\underline{\check{x}})p(\underline{y}\mid\underline{\check{x}})}{\max_{\underline{\check{x}}\in\underline{\mathcal{X}}^-} P(\underline{\check{x}})p(\underline{y}\mid\underline{\check{x}})}\right), \tag{4.19}$$

where $\underline{\mathcal{X}}^{\pm}$ is the set of potentially transmitted vectors $\underline{\check{x}}$, $1 \leq m \leq MN_{\mathrm{bit}}$ associated with $\underline{\check{x}}_m = \pm 1$ and $P(\underline{\check{x}})$ is the *a priori* probability of the given hypothetical signal vector being considered, which may be obtained from the *a priori* bit LLRs $\lambda_{\mathrm{A}}^{\mathrm{dec}}(x_m)$ as outlined in [48] upon assuming independence of the individual bits. Equation (4.19) may now be re-written as

$$\begin{aligned}\lambda_{\mathrm{A}}^{\mathrm{det}}(\underline{x}_m) &= J_{\min}(\underline{\check{x}}_m = +1) - J_{\min}(\underline{\check{x}}_m = -1) + \\ &= -\min_{\underline{\check{x}}^{(i)}\in\underline{\mathcal{X}}^+}\{-\log(P(\underline{\check{x}}^{(i)})) + \frac{||\underline{y} - \underline{\boldsymbol{H}}\underline{\check{x}}^{(i)}||^2}{2\sigma_{\mathrm{n}}^2}\} \\ &\quad \min_{\underline{\check{x}}^{(i)}\in\underline{\mathcal{X}}^-}\{-\log(P(\underline{\check{x}}^{(i)})) + \frac{||\underline{y} - \underline{\boldsymbol{H}}\underline{\check{x}}^{(i)}||^2}{2\sigma_{\mathrm{n}}^2}\},\end{aligned} \tag{4.20}$$

where J represents the cost-function to be optimized and $J_{\min}(\underline{\check{x}}_i = j)$ denotes the minimum cost-function value under the constraint of $\underline{\check{x}}_i = j$. The algorithm presented in this section has the capability of solving the optimization problem associated with $J_{\min}(\underline{\check{x}}_m = \pm 1)$ at a moderate computational cost. Based on the derivations of Section 2.3.4 it can be seen that the solution to the problem defined by Equation (4.20) is

identical to solving

$$J_{\min}(\underline{\check{x}}_m = +1) = C + \min_{\underline{\check{x}}^{(i)} \in \underline{\mathcal{X}}^+} \left(\frac{||\boldsymbol{U}(\underline{\check{x}}^{(i)} - \underline{\hat{x}}_{\mathrm{MMSE}})||^2}{2\sigma_{\mathrm{n}}^2} - \log(P(\underline{\check{x}}^{(i)})) \right), \tag{4.21}$$

where the upper triangular matrix $\mathbf{U}$ is defined by

$$\boldsymbol{U}^H \boldsymbol{U} = \underline{\boldsymbol{H}}^H \underline{\boldsymbol{H}} + \sigma_{\mathrm{n}}^2 \mathbf{I}_{N_{\mathrm{b}}}, \tag{4.22}$$

while

$$\underline{\hat{x}}_{\mathrm{MMSE}} = (\underline{\boldsymbol{H}}^T \underline{\boldsymbol{H}} + \sigma_{\mathrm{n}}^2 \mathbf{I}_{N_{\mathrm{b}}})^{-1} \underline{\boldsymbol{H}}^H \underline{\boldsymbol{y}}, \tag{4.23}$$

where $\mathbf{I}_{N_{\mathrm{b}}}$ is the $(N_{\mathrm{b}} \times N_{\mathrm{b}})$-dimensional identity matrix. The constant C is identical for both $J_{\min}(\underline{\check{x}}_m = +1)$ as well as $J_{\min}(\underline{\check{x}}_m = -1)$ and hence it can be dropped from the evaluation of the *a posteriori* LLRs given in Equation (4.19).

Exploiting the fact that the matrix $\mathbf{U}$ has an upper triangular structure, it can be shown that the objective function used for the detection of the transmitted symbol vector $\check{\mathbf{x}}$ may be written as [61]

$$\begin{aligned} J(\underline{\check{x}}) &= \frac{||\boldsymbol{U}(\underline{\check{x}} - \underline{\hat{x}}_{\mathrm{MMSE}})||^2}{2\sigma_{\mathrm{n}}^2} - \log(P(\underline{\check{x}})) \\ &= \frac{(\underline{\check{x}} - \underline{\hat{x}}_{\mathrm{MMSE}})^T \boldsymbol{U}^T \boldsymbol{U} (\underline{\check{x}} - \underline{\hat{x}}_{\mathrm{MMSE}})}{2\sigma_{\mathrm{n}}^2} - \log(P(\underline{\check{x}})) \\ &= \sum_{i=1}^{N_{\mathrm{b}}} \frac{|\sum_{j=i}^{M} u_{ij}(\underline{\check{x}}_j - \underline{\hat{x}}_{\mathrm{MMSE},j})|^2}{2\sigma_{\mathrm{n}}^2} - \log(P(\underline{\check{x}}_i)) \\ &= \sum_{i=1}^{N_{\mathrm{b}}} \phi_i(\underline{\check{x}}_i), \end{aligned} \tag{4.24}$$

where u_{ij} is the $(i,j)^{th}$ element of $\mathbf{U}$, $\underline{\check{x}}_i = [\underline{\check{x}}_i, \ldots, \underline{\check{x}}_{N_{\mathrm{b}}}]^T$ and $\phi_i(\check{\mathbf{x}}_i)$ may be expressed as [61]

$$\phi_i(\underline{\check{x}}_i) = \frac{|u_{ii}(\underline{\check{x}}_i - \underline{\hat{x}}_{\mathrm{MMSE},i}) + a_i)|^2}{2\sigma_{\mathrm{n}}^2} - \log(P(\underline{\check{x}}_i)), \tag{4.25}$$

where

$$a_i = \sum_{j=i+1}^{N_{\mathrm{b}}} u_{ij}(\underline{\check{x}}_j - \underline{\hat{x}}_{\mathrm{MMSE},j}). \tag{4.26}$$

The term a_i of Equation (4.25) is independent of the specific symbol's value of $\underline{\check{x}}_i$. The cost-function given in Equation (4.24) may now be re-written in a recursive manner as

$$J_i(\underline{\check{x}}_i) = J_{i+1}(\underline{\check{x}}_{i+1}) + \phi_i(\underline{\check{x}}_i), \quad i = N_{\mathrm{b}} - 1, \ldots 1, \tag{4.27}$$

where we have $J_{N_{\mathrm{b}}}(\underline{\check{x}}_{N_{\mathrm{b}}}) = |u_{N_{\mathrm{b}}N_{\mathrm{b}}}(\underline{\check{x}}_{N_{\mathrm{b}}} - \underline{\hat{x}}_{\mathrm{MMSE},N_{\mathrm{b}}})|^2$. The cost-function retains the essential property

that

$$J(\underline{\check{x}}) = J_1(\underline{\check{x}}_1) > J_2(\underline{\check{x}}_2) > \cdots > J_{N_b}(\underline{\check{x}}_{N_b}) = J_{N_b}(\underline{\check{x}}_{N_b}) > 0. \tag{4.28}$$

Based on Equations (4.24)-(4.28) a low-complexity tree-search algorithm may be employed, which is outlined in Section 2.3.4. A further property of the recursive cost-function imposed by the *a priori* information available is that we have

$$J_{N_b}(\underline{\check{x}}) > J_i(\underline{\check{x}}_i) - \sum_{j=1}^{i+1} \max\{\log(P(\underline{x}_j = -1)), \log(P(\underline{x}_j = +1))\}, \tag{4.29}$$

which might be exploited by the tree-search algorithm in order to improve its convergence rate.

The tree-search algorithm is guaranteed to find the ML solution $\boldsymbol{x}_{\mathrm{ML}}$ satisfying the cost-function value $J_{\mathrm{ML}} = \min(J_{\min})$. This constitutes an effective way of calculating the $2N_b$ cost-function values required for obtaining the desired N_b number of LLRs. The proposed method is summarized in Algorithm 4.1, where $J_{\min}(\underline{\check{x}}_i = j)$ is the minimum cost-function value under the constraint of $\underline{\check{x}}_i = j$.

Algorithm 4.1: Max-Log Metric calculation

```
Initialize:
    J_ML = 10^10 (Value higher than true J_ML)
    J_min(x̌_i = ±1) := a constraint > J_ML   for   1 ≤ i ≤ M

calculate J_ML and x̌_ML evaluating search tree
for  p = 1  to  M  do
    J_min(x̌_p = x̌_ML,p) = J_ML
end

for  i = 1  to  M  do
    for  j ∈ {−1, 1}  do
        if  J_min(x̌_i = j) > J_ML  then
            calculate J_min(x̌_i = j) evaluating search tree
    end
    i++
end
```

The proposed algorithm requires $(N_b + 1)$ search tree passes for the calculation of the $2N_b$ cost-function values. The algorithm exploits the fact that the cost-function value J_{ML} associated with the ML solution is the lowest possible cost-function value and therefore already provides the first N_b values $J_{\min}(\underline{\check{x}}_i = x_{\mathrm{ML},i}) =$

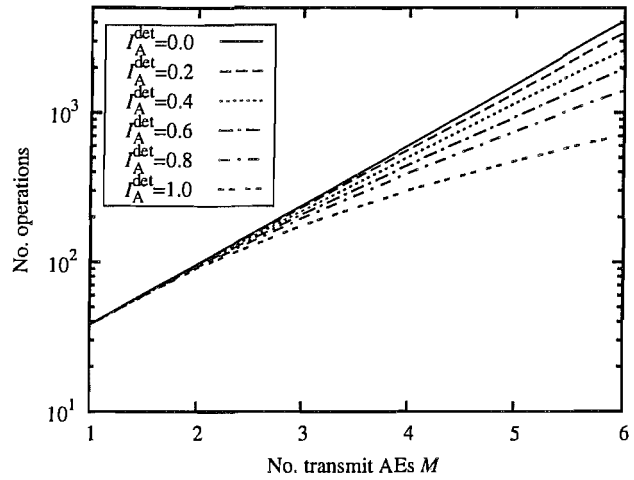
J_{ML}. The algorithm may also readily be used for the detection of higher order modulation schemes by representing a complex-valued symbol as a string of independent bits, as outlined in [61]. The algorithm presented in this section extends the original hard-output OHRSA of [61] to a SISO algorithm, which may be employed in iterative systems for calculating the required $2N_b$ cost-function values.

The computational complexity associated with the RML detector may be divided into two contributions, namely the calculation of the MMSE solution and the Cholesky factorization, both of which have been discussed in Section 2.3.4 plus the additional complexity imposed by the search-tree evaluation of the RML detector. In Figure 4.11, the computational complexity imposed by the search-tree evaluation, when embedded in Algorithm 4.1 is illustrated as a function of the number of transmit AEs for four receive AEs. The computational complexity is illustrated for different values of the mutual information, associated with the *a priori* LLRs available at the detector. The system considered employed 4QAM signals.

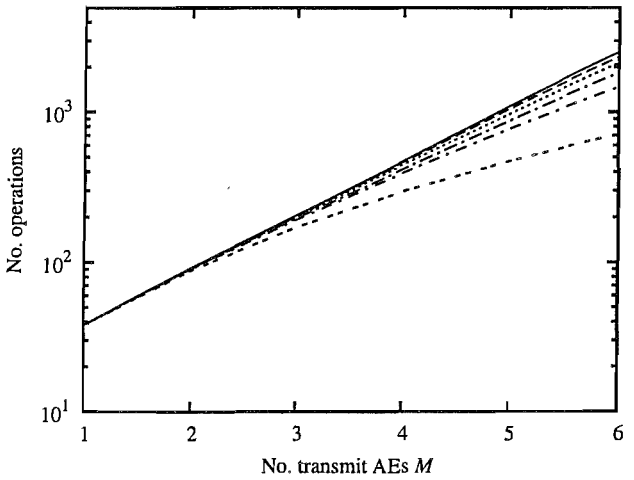
It can be seen from Figure 4.11a that as the number of transmit AEs increases, the computational complexity is reduced for a higher mutual information, which suggests that the *a priori* LLRs assist beneficially in the search-tree evaluation. However, when the SNR increases, this effect becomes less pronounced, as it can be seen from Figure 4.11b. When the SNR is increased even further, the benefit of having *a priori* information becomes even less beneficial in terms of the associated computational complexity reduction, as it can be seen from Figure 4.11c. This phenomenon indicates that at lower SNRs the *a priori* LLR becomes the dominant factor for the search-tree evaluation, when considering a high *a priori* mutual information. By contrast, at higher SNRs the available *a priori* information appears to have a more modest conductive effect, since the evaluation of the cost-function of Equation (4.25) is dominated by the detector's contribution. When the complexity is compared to that of the Bayesian detector of Section 4.4.2 it can be observed that the complexity of the RML detector is almost two magnitudes lower for $M = 6$ transmit AEs.

In Figure 4.12 the EXIT charts of the RML detector are illustrated for $M = 2, 4$ and 6 transmit AEs, while employing $N = 2$ receive AEs and considering 4QAM signals. It can be observed from Figure 4.12a and Figure 4.12b that for $M = 2$ and $M = 4$ the RML detector has a similar EXIT function characteristic, as the Bayesian detector illustrated in Figure 4.9. It appears that the Max-Log approximation has a relatively modest effect on the shape of the EXIT function. When considering $M = 6$ transmit AEs, it can be observed from Figure 4.12c that in contrast to the SISO MMSE detector of Figure 4.6c the performance of the RML detector is not interference limited. In fact, for this rank-deficient scenario, which is associated with a relatively steep exit characteristic a high iteration gain can be achieved.

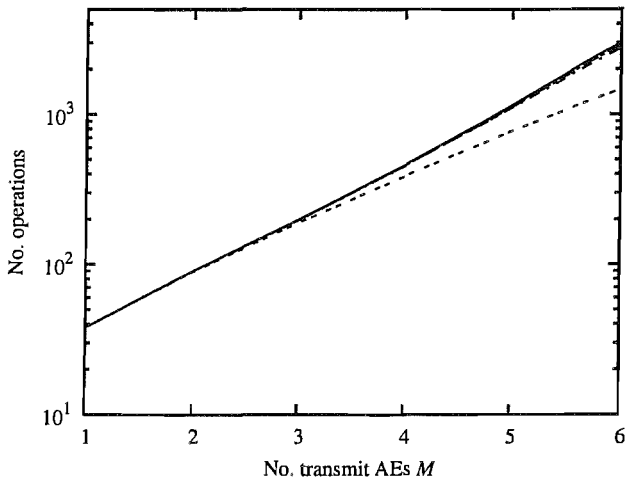
The same phenomenon can be observed in Figure 4.13, where the achievable BER versus E/N_0 performance is illustrated for the same scenario. The steep EXIT characteristic associated with $M = 6$ AEs results in a water-fall shape BER characteristic. For a lower number of transmit AEs, the RML, the Bayesian and the MMSE SISO detector exhibit a similar BER performance in the context of the system considered.



(a) E_b/N_0=5 dB

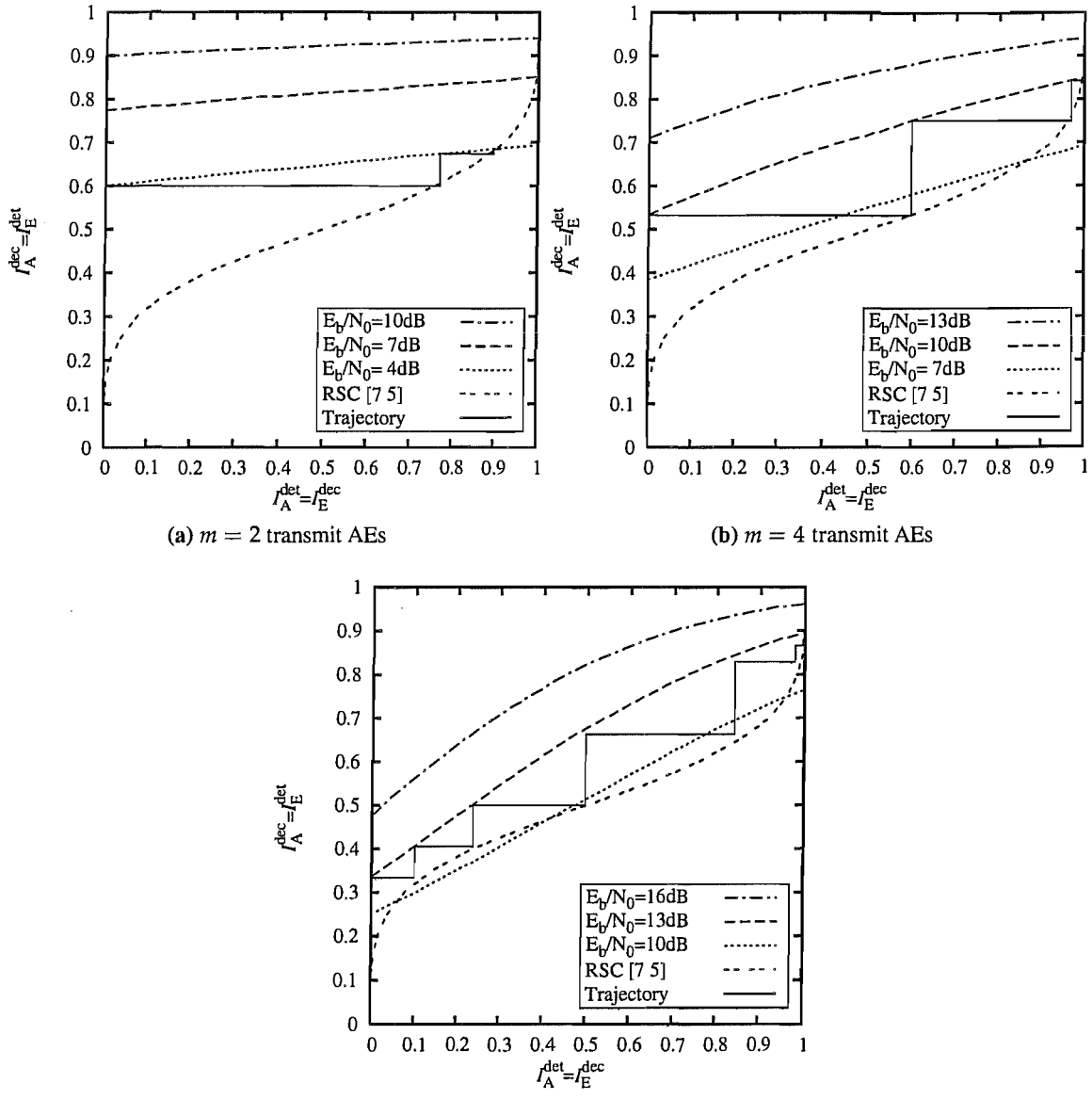


(b) E_b/N_0=10 dB.



(c) E_b/N_0=20 dB.

Figure 4.11: The normalized computational complexity imposed by the search-tree evaluation of the RML detector when detecting a single transmitted bit as a function of the number of transmit AEs, evaluated for $N = 4$ receive AEs and 4QAM signals. Different values of mutual information associated with the *a priori* LLRs were considered.



(a) $m = 2$ transmit AEs

(b) $m = 4$ transmit AEs

(c) $m = 6$ transmit AEs

Figure 4.12: RML EXIT function for a system using M transmit AEs and $N = 2$ receive AEs. 4QAM signals were considered. In addition the EXIT function of a RSC code using the generator polynomial of $g = [7\ 5]$ is illustrated. The trajectories were recorded for an interleaver length of 10 000 bits.

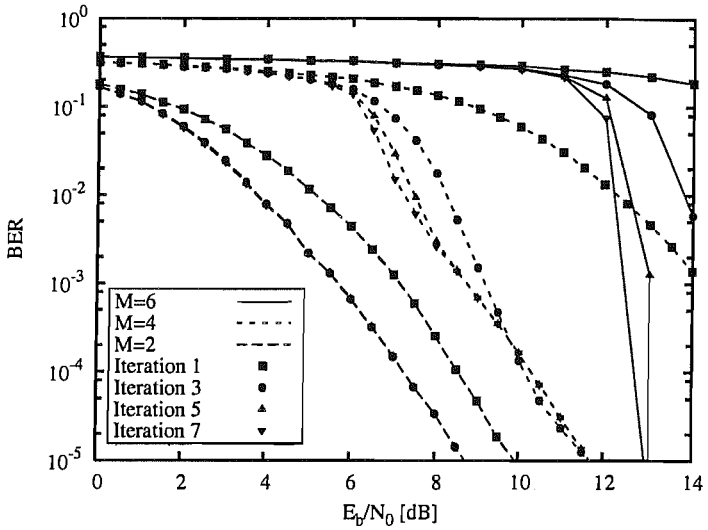


Figure 4.13: BER versus E_b/N_0 performance of a RML detector for a system using M number of transmit AEs and $N = 2$ receive AEs. 4QAM and an interleaver length of 10 000 bits were considered.

4.4.4 MBER SISO Detection

Simply applying the iterative detection principle to the MBER algorithm may not be very attractive, since its complexity exceeds that of the optimum Bayesian detector. Let us consider the case of a block-constant fading channel. It was argued in Chapter 2 for this scenario that the computational complexity associated with the calculation of the MBER weights may be justified, because once the weight-vector has been evaluated, the detection itself can be achieved at a low computational cost.

Following the approach presented in Section 2.3.2, the BER at the detector output may be expressed as

$$P_e(\underline{w}_m) = \frac{1}{N_{sb}} \sum_{\underline{\check{x}}_m^{(i)} \in \underline{\hat{\mathcal{X}}}_m^+} P(\underline{\check{x}}_m) Q\left(\frac{\mathtt{sign}\left(\underline{\check{x}}_m^{(i)}\right) \underline{\check{x}}_m^{(i)}}{\sigma_n \sqrt{\underline{w}_m^T \underline{w}_m}} \right), \tag{4.30}$$

where $\underline{\hat{\mathcal{X}}}_m^+$ is the set of hypothetically transmitted symbols associated with $\underline{\check{x}}_m = +1$ and

$$P(\underline{\check{x}}_m^{(i)}) = \prod_{j=1}^{MN_{bit}} P(\underline{\check{x}}_m^{(i)}) \tag{4.31}$$

is the *a priori* probability of the hypothetical solution considered, assuming that the transmitted bits are independent. Similarly to Equation (2.68), the BER gradient can be expressed as

$$\nabla P_e(\underline{w}_m) = C \sum_{\underline{\check{x}}_m^{(i)} \in \underline{\hat{\mathcal{X}}}_m^{(+)}} P(\underline{\check{x}}_m^{(i)}) \exp\left(-\frac{\left(\underline{\check{x}}_m^{(i)}\right)^2}{2\sigma_n^2 \underline{w}_m^T \underline{w}_m} \right) \mathtt{sign}\left(\underline{\check{x}}_m^{(i)}\right) \left(\frac{\underline{\check{x}}_m^{(i)} \underline{w}_m}{\underline{w}_m^T \underline{w}_m} - \underline{\check{y}}_m^{(i)} \right), \tag{4.32}$$

where we have

$$C = \frac{1}{2N_{\text{sb}}\sqrt{2\pi\sigma_{\text{n}}^2 \underline{w}_m^T \underline{w}_m}}. \tag{4.33}$$

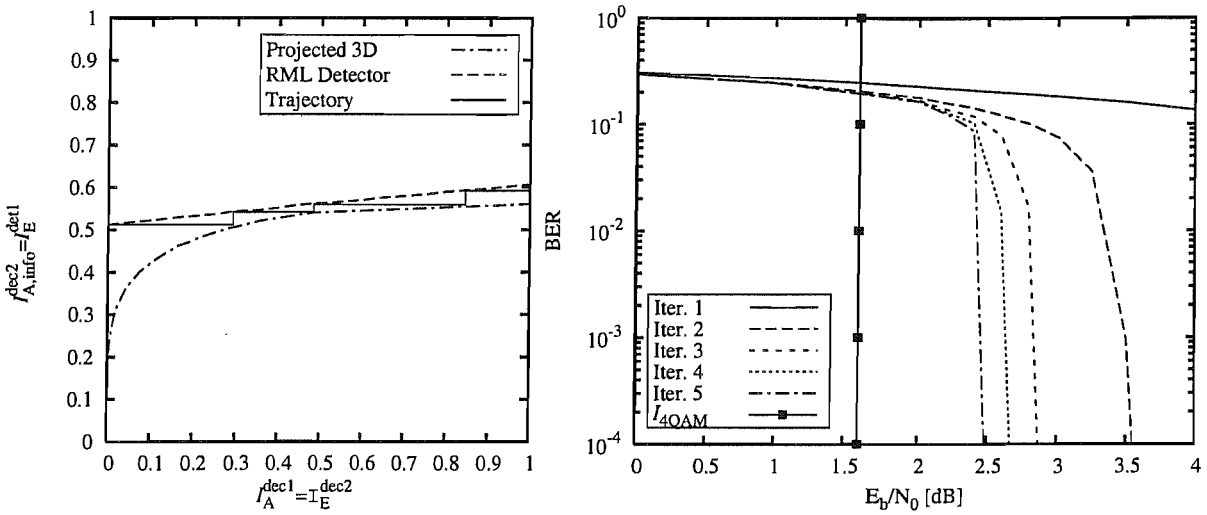
It can be seen from Equation (4.32) that during the first iteration when no *a priori* information is available at the detector, the SISO MBER solution is equivalent to the MBER algorithm not benefiting from soft input, which was discussed in Section 2.3.3. Assuming, that the channel is constant over one transmission burst, the weight vector for each transmit AE only has to be calculated once for the whole transmission burst. If, however, *a priori* information is available for the MBER detector, the BER surface and thus the BER-gradient is changed for each symbol differently according to the associated *a priori* information. Thus a BER-gradient based algorithm would have to be run for each symbol to be detected. The complexity of this scheme would be several orders of magnitudes higher than that of the full Bayesian detector. Furthermore, applying the conjugate gradient algorithm for each symbol conflicts with the philosophy of MBER detection since P_{e} is a statistical measure. As a result of these problems, the iterative MBER SISO detector is not considered in the remainder of this chapter.

4.4.5 Iterative MIMO Detection Simulation Results

The system considered employed two serially concatenated channel codes and used $M = 2$ transmit AEs as well as $N = 2$ receive AEs, where 4QAM signals and an interleaver length of 64 000 bits was considered. All channel code parameters are summarized in Table 4.1. In order to be able to employ EXIT chart analysis for the proposed detector, we consider independent and uncorrelated fading channels for all antenna links. Since the system consists of three concatenated SISO components, a projected EXIT chart discussed in Section 4.3 is considered. In the derivation of the 2D-projected 3D-EXIT chart it was assumed that the two codes iterate to the point of convergence. We therefore consider $N_{\text{inner}} = 10$ iterations between the two channel codes after each detection by the RML detector.

In Figure 4.14a the BER is shown as a function of E_{b}/N_0 after a different number of RML-aided detection stages together with the theoretically achievable information rate obtained as outlined in Section 2.2.3. It can be observed that the BER curve follows waterfall shape typical for iterative detectors and comes as close as 0.8 dB to the theoretically achievable information rate bound. It can also be observed that during the first and second iteration the gain is high and the incremental gains decreased upon increasing the iteration index. This behaviour can be explained by considering the EXIT chart of the system recoded at $E_{\text{b}}/N_0 = 2.6$ dB, which shows a wide-open EXIT tunnel for low $I_{\text{E}}^{\text{det}}$ values, before the tunnel narrows. It also becomes evident from the good agreement between the recoded decoding trajectory and the 2D-projected EXIT function that projected 3D-EXIT describe the system accurately, given a sufficiently long interleaver.

For the same system we have also illustrated the E_{b}/N_0 versus iteration index in Figure 4.15, required



(a) EXIT Chart at $E_b/N_0 = 2.6$ dB.

(b) BER versus E_b/N_0.

Figure 4.14: The system considered employed two serially concatenated channel codes and used $M = 2$ transmit AEs as well as a $N = 2$ receive AEs. 4QAM signals and an interleaver of length 64 000 bits were considered. All system parameters are summarized in Table 4.1. The iteration index indicated is associated with the RML aided detector. Between two consecutive detection stages, the receiver iterated $N_{\text{inner}} = 10$ times between the two channel codes.

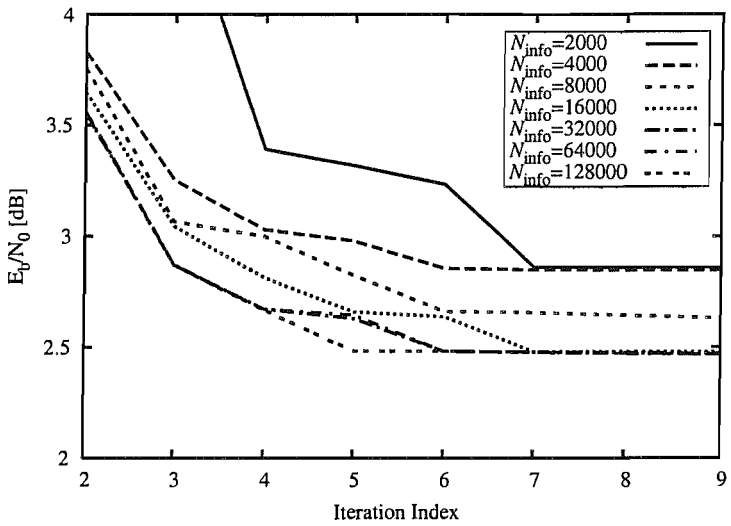


Figure 4.15: E_b/N_0 versus iteration index required in order to achieve a target BER of $10^{-}3$ for different interleaver length. The system considered supported M number of transmit AEs and a different number of receive AEs and was identical to the system considered for generating Figure 4.14. The iteration index indicated is associated with the RML aided detector. Between tow detection, the receiver iterated 10 times between the two channel codes.

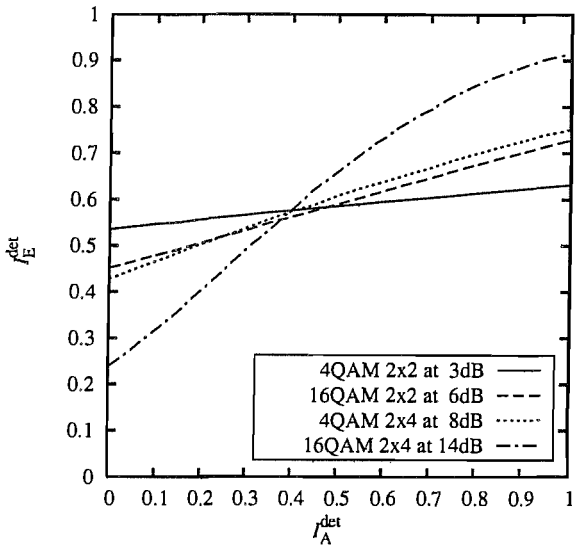


Figure 4.16: Two-dimensional EXIT chart [137] for a MIMO system associated with $M = 2$ or 4 transmit antennas and $N = 2$ receive antennas at different E_b/N_0 values, as defined in Equation (4.4) for a code-rate of $R = 1/2$. The modulation schemes considered were 4QAM and 16QAM.

in order to achieve a target BER of 10^{-3} for different interleaver lengths. It can be clearly observed in Figure 4.15 that there exists a trade-off between the complexity associated with a given iteration index and the delay associated with a given interleaver length. For a relatively short interleaver a higher number of iterations has to be performed, in order to achieve an identical performance.

The performance of the RML detector has been evaluated considering both 4QAM as well as 16QAM signals and a receiver employing $N = 2$ antennas for receiving from M number of transmit AEs at different E_b/N_0 values. The random bit interleaver employed had a length of 2^{16} bits and the fading process imposed by the channel was assumed to be uncorrelated and independent between the different AEs.

In Figure 4.16 the proposed RML detector is characterized with the aid of EXIT charts [137] for both $M = 2$ as well as $M = 4$ transmit AEs. It can be observed that for 4QAM signals as well as for the (2×2)-antenna aided 16QAM scenario considered, the EXIT function is almost horizontal[1]. As a benefit, in conjunction with a well-designed channel code, which matches this EXIT characteristic with the aid of a near-parallel EXIT function the iterative detector has a low complexity, since it needs only a few detection iterations in order to converge. By contrast, the EXIT function associated with the (2×4)-antenna based rank-deficient 16QAM scenario shows a higher gradient, which implies that an open EXIT tunnel can be readily created in conjunction with an appropriately designed channel code EXIT function, but iterating through this steeper tunnel may involve numerous detection iterations and thus a high computational complexity. Furthermore, it can be observed in Figure 4.16 that the rank-deficient (4×2)-antenna 4QAM scheme and the fully-loaded (2×2)-antenna 16QAM arrangement have similar EXIT-functions and they both support identical system throughputs of $4 \cdot 2 = 2 \cdot 4 = 8 \frac{\text{bit}}{\text{MIMO symbol}}$. The fully-loaded 16QAM system however has a 2 dB E_b/N_0

[1] In contrast to the typical turbo channel code EXIT characteristics, the EXIT functions of iterative detectors generally do not reach the point of perfect convergence at $(I_A^{det}, I_E^{det}) = (1,1)$.

Table 4.4: SCC Codes

	Component Code 1	Component Code 2
SCC Code 1	RSC, $R_c = 1, g = [3\ 1]$	RSC, $R_c = 1/2, g = [6\ 5]$
SCC Code 2	RSC, $R_c = 2/3, g = [6\ 4]$, $P_{\text{punct}} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$	RSC, $R_c = 3/4, g = [7\ 4]$, $P_{\text{punct}} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

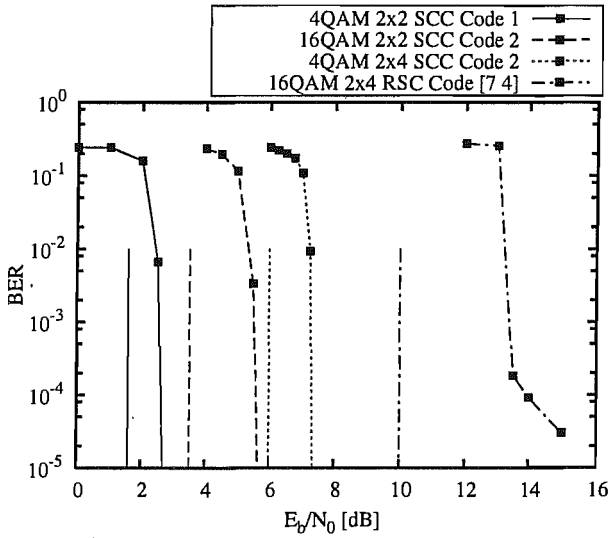


Figure 4.17: BER versus E_b/N_0 for a MIMO system having $M = 2$ or 4 transmit antennas and $N = 2$ receive antennas and using 4QAM or 16QAM. The channel codes employed are the various serially concatenated codes of Table 4.4.

advantage in comparison to the rank-deficient 4QAM system.

In order to achieve a performance close to the channel capacity bound, in most cases it is insufficient to consider an iteratively detected serially concatenated RSC code and a SISO detector. In order to support communication close to the channel capacity we therefore consider a channel code consisting of two serially concatenated codes, which are separated by an interleaver in order to facilitate iterative channel decoding. In our simulations, these two codes were considered to be a single code, where the number of inner iterations between the two codes was chosen to be $N_{\text{inner}} = 10$. The exact code parameters are provided in Table 4.4.

In Figure 4.17 the achievable BER versus E_b/N_0 performance of the RML detector benefiting from iterative detection is illustrated together with the associated capacity bounds represented by the vertical lines. It can be observed that the performance of all scenarios is less than 2 dB away from the channel capacity, except for the (2×4)-antenna 16QAM scenario. This can be explained by the fact that for this specific scenario the code's EXIT function and the RML detector's EXIT function match poorly and thus the detector fails to approach the capacity limit. Communication closer to the capacity limit could be supported using for

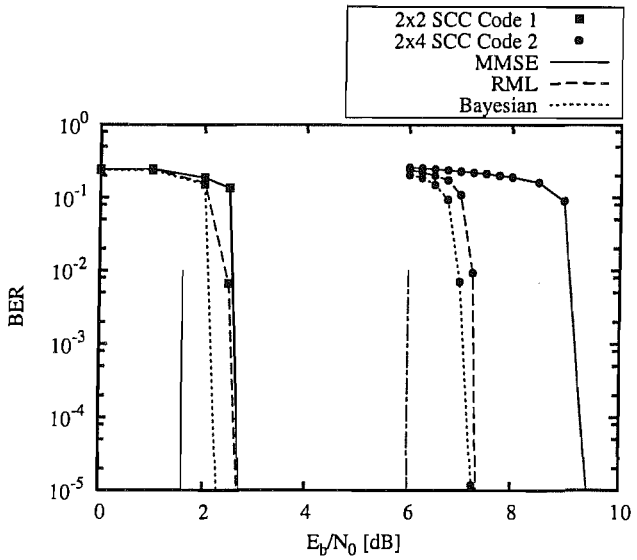


Figure 4.18: BER versus E_b/N_0 for a MIMO system having $M = 2$ or 4 transmit antennas and $N = 2$ receive antennas and using 4QAM. The channel codes employed are the various serially concatenated codes of Table 4.4.

example irregular codes specifically designed to match the shape of the EXIT functions [103], although this is achieved at the cost of a higher number of detector iterations.

In Figure 4.18 we consider the same system as that used for generating the results of Figure 4.17, but now consider additionally the Bayesian and the MMSE detector. It can be observed from Figure 4.18 that for the system supporting only $M = 2$ transmit AEs, there is only a marginal performance difference between the three algorithms. If, however, the number of transit AEs is increased to $M = 4$, then the Bayesian as well as the RML detector perform significantly better than the MMSE aided iterative detector.

It was observed in Figure 4.18 that for certain channel conditions, both the MMSE and the RML detector exhibit a similarly BER versus E_b/N_0 performance, despite having different EXIT functions. In this respect it might be more practical to compare different detection algorithms not only according to their BER performance, but also according to the area under their EXIT function. If different detectors are compared using the BER versus E_b/N_0 performance, the achievable performance depends on how well the EXIT function of the detector matches that of the decoder. A different approach would be to compare the various detectors according to the area under their EXIT functions, which allows their comparison, regardless of the specific channel code used. To elaborate a little further, it has to be emphasized that one minus the area under the EXIT function of the channel code corresponds to its code rate [140]. This was only proved for the AWGN and the Binary Erasure Channel (BEC), but observations in the literature suggest that it also approximately holds for Rayleigh fading channels. As a result, it follows that if we assume having a channel code EXIT function, which is perfectly matched to the EXIT function of the detector, a detector having a larger area under its EXIT function is capable of converging in conjunction with a higher rate code. Comparing the EXIT functions of the different detectors shown in Figure 4.6, Figure 4.9 and Figure 4.12, respectively, it

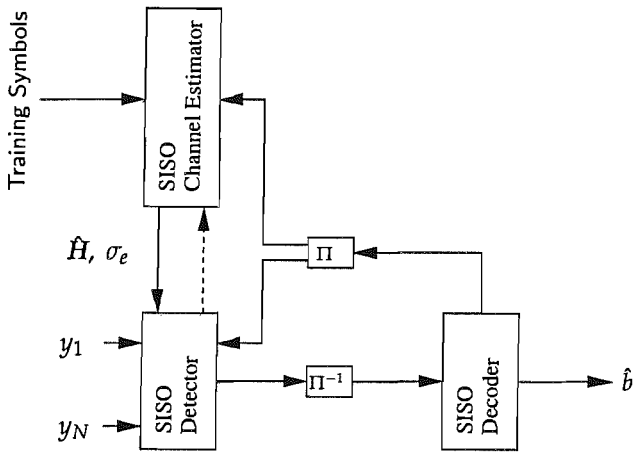


Figure 4.19: Block-diagram illustrating a joint channel estimation and detection structure.

can be observed that even though there is no direct difference in achievable BER performance for some of the systems considered, the potentially achievable performance determined by the area under the EXIT function is always higher for both the Bayesian and the RML aided detector than for the MMSE based detector. In other words, the Bayesian and the RML detector are capable of supporting a higher effective data-rate at the same SNR.

In line with the classic lessons of Shannon, all the observations made in this section lead to a single bold conclusion. Explicitly, the closer we want to approach the capacity bound, the higher the complexity and the delay.

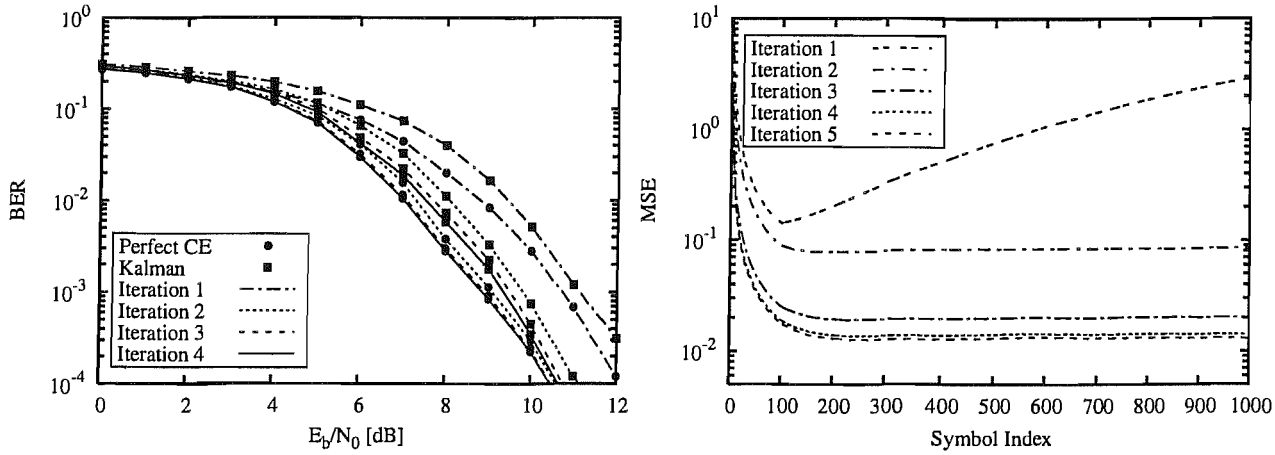
4.5 Training Aided Iterative Channel Estimation

In this section we will discuss joint iterative channel estimation using a simple first-order Kalman channel estimator. The channel estimator considered in this section is a Kalman channel estimator introduced in [141], which can readily be used for iterative detection and channel estimation as depicted in Figure 4.19.

The channel estimator depicted in Figure 4.19 might operate in three different modes:

1. Training based channel estimation;
2. Decision directed channel estimation;
3. SISO channel estimation using the decoder's LLR values.

The training based as well as the decision directed channel estimation schemes used are identical to the non-iterative channel estimator discussed in Section 2.4. If the channel is varying rapidly, it might be required that during the first detection stage the channel estimator tracks the channel in a decision directed mode. By contrast, if the channel is slowly time-varying, it might acceptable to assume the channel to be constant for



(a) BER versus E_b/N_0

(b) MSE versus symbol index

Figure 4.20: A system using $M = 4$ transmit AEs and $N = 4$ receive AEs was considered and the normalized Doppler frequency was assumed to be $\overline{f}_D = 0.0001$. The modulation scheme employed was 4QAM.

the first detection stage and to dispense with decision directed channel estimation. During all subsequent detection stages the channel estimator updates the estimated CIR entirely based on the decoder's extrinsic information.

As it becomes clear from Figure 4.19, the channel estimator has to have the capability of estimating the channel using soft-information and return the channel estimation error in form of an estimation error variance σ_e^2. The detector then considers not only the Gaussian noise for the detection but also the estimation error provided by the channel estimator. The total noise variance considered by the detector is then given as $\sigma_n^2 + \sigma_e^2$.

In order to highlight the performance gains that can be achieved by iterative channel estimation, we consider a system supporting $M = 4$ transmit AEs and $N = 4$ receive AEs. The normalized Doppler frequency was chosen to be $\bar{f} = 0.0001$ and 4QAM signals were considered. The detector used was a RML aided detector and the channel code considered was a half-rate RSC code having a octally represented generator polynomial $\boldsymbol{g} = [7\ 5]$. The number of training symbols per transmit AE was set to 100 and the useful payload per transmit AE was chosen to be 900 symbols. After the first training stage the channel estimator operated in decision directed mode. For the subsequent iterations the soft information provided by the channel decoder was used for estimating the channel.

It can be seen from Figure 4.20a that the BER performance difference for the considered system is less than 1 dB compared to the system benefiting from perfect channel knowledge although this is achieved at the cost of a higher number of iterations. Figure 4.20b shows the MSE recorded at $E_b/N_0 = 10$ dB as a function of the symbol index. It can be observed how the MSE is reduced gradually for an increased number of iterations, which allows the use of relatively short training sequences.

4.6 Conclusion

In this chapter we have discussed different SISO detection algorithms, namely the MMSE technique of Section 4.4.1, the Bayesian solution of Section 4.4.2 as well as the novel RML detector of Section 4.4.3. All algorithms have been characterized using EXIT charts and the complexity imposed by the different algorithms has been discussed. In order to enable us to characterize concatenated systems consisting of three serially cascaded SISO components, we have introduced a novel 2D-projected EXIT chart. This 2D-projected EXIT function enabled us to analyze systems consisting of three SISO blocks with the aid of a single 2D EXIT chart.

Furthermore, it was discussed in Section 4.4.4 why simply extending the MBER algorithm of Section 2.3.3 to a SISO algorithm is not practical, since it results in a computational cost associated with the SISO MBER detector, which exceeds that of the optimum Bayesian detector.

In Table 4.5 we have summarized the optimization criteria of the various algorithms discussed and have portrayed the computational complexity associated with the SISO detection strategies considered. It can be observed that the MMSE SISO detector exhibits the lowest computational complexity during the first iteration, while the Bayesian detector exhibits a lower computational complexity during subsequent iterations. The complexity of the RML detector depends on the SNR, but decreases slightly for subsequent iterations, as it can be seen from Figure 4.11. For large systems, the complexity of the RML detector is between that of the MMSE and that of the Bayesian detector.

When quantitatively comparing the different detectors' performance, we investigate their performance in terms of the distance expressed in dB from the theoretical channel capacity limits discussed in Section 4.2. Under the assumption of uncorrelated fading channels the distance $\Delta_{I|\text{BER}=10^{-4}}$ from the channel capacity limit in [dB] is given in Table 4.5. The values were obtained from Figure 4.18 for a BER of 10^{-4} . It can be observed that the MMSE detector performs almost as well as the Bayesian detector, when the (2×2)-antenna system is considered. If, however, the number of transmit AEs is increased to four, resulting in a (2×4)-antenna system, the MMSE detector's performance is significantly degraded. By contrast, the RML detector approaches the performance of the Bayesian detector also for the rank-deficient system. The performance difference between the RML and the Bayesian detector is a consequence of using the Max-Log approximation for the RML detector.

Finally, in Section 4.5 we have discussed the performance of the RML detector, when performing joint channel estimation and detection using a SISO Kalman channel estimator. It was shown that with the aid of SISO channel estimation, the performance degradation of the system considered owing to using the estimated rather than perfect CSI is less than 1 dB as shown in Figure 4.16.

Table 4.5: Summary of the MIMO SISO detectors studied in Chapter 4.

		MMSE of Section 4.4.1	**Bayesian** of Section 4.4.2	**RML** of Section 4.4.3
Optimization		$\lambda_{\mathrm{E}}^{\mathrm{det}}(\underline{x}_m) = \frac{4\underline{w}^H(\underline{y}-\underline{\underline{H}}\underline{\bar{x}}+\bar{x}_m\underline{h}_m)}{1-\underline{h}_m^T\underline{w}_m}$, where $\underline{w}_m = \arg\min_{\underline{w}_m\in\mathbb{R}^{2N\times 1}} \|\underline{x}-\underline{w}_m^H(\underline{y}-\underline{\underline{H}}\underline{\bar{x}}+\bar{x}_m\underline{h}_m)\|^2$	$\lambda_{\mathrm{P}}^{\mathrm{det}}(\underline{x}_m) = \log\left(\frac{P(x_m=+1\|y)}{P(x_m=-1\|y)}\right)$	$\lambda_{\mathrm{P}}^{\mathrm{det}}(\underline{x}_m) \approx \log\left(\frac{\max_{\underline{\check{x}}\in\underline{\mathcal{X}}_m^+} P(\underline{\check{x}})p(\underline{y}\|\underline{\check{x}})}{\max_{\underline{\check{x}}\in\underline{\mathcal{X}}_m^-} P(\underline{\check{x}})p(\underline{y}\|\underline{\check{x}})}\right)$
Comp. [a]	1. Iteration	$O\left(\frac{N^3}{N_{\mathrm{bit}}M}\right)$	$O(2^{N_{\mathrm{bit}}M})$	$O\left(\frac{N^3}{N_{\mathrm{bit}}M}\right) < \bullet < O(2^{N_{\mathrm{bit}}M})$
	Subsequent Iterations	$O(N^3)$	$O(2^{N_{\mathrm{bit}}M})$	$O(N^3) < \bullet < O(2^{N_{\mathrm{bit}}M})$
Perf. [b]	(2×2)	≈ 1	< 1	≈ 1
	(2×4)	3.5	1.2	1.4

[a] Values obtained from Table 4.2, Table 4.3 and Figure 4.11, respectively.
[b] Distance from the theoretical capacity limit obtained from Figure 4.18.

Chapter 5

Soft-Input Soft-Output Space-Time Equalization

This chapter commences with the evolution of the narrowband Soft-Input Soft-Output (SISO) detection algorithms discussed in Chapter 4 in order to facilitate their employment to Single-Carrier (SC) Space-Time Equalization (STE). In contrast to Chapter 4, where a single-user narrowband Multiple-Input Multiple Output (MIMO) system was considered for iterative detection, the focus of this chapter is on a system supporting multiple MIMO users, who are communicating over a frequency-selective Rayleigh fading channel with the Base-Station (BS). In Sections 5.1 and 5.2 a recipe is provided for extending the narrowband MIMO detection algorithms of Chapter 3 to STE algorithms. These principles are essentially based on the derivations presented in Chapter 3 as well as on the formulation of the Minimum Mean Squared Error (MMSE) technique and the Bayesian as well as the Reduced complexity Max-Log (RML) algorithm of Chapter 4. In addition to the Decision Feedback (DF) methods discussed in Section 3.5 in the specific context of hard DF, here a more sophisticated soft DF method will be introduced. In order to circumvent the deficiencies observed in the context of the algorithms discussed in Section 5.2, a novel scheme referred to as the Markov Chain Monte Carlo (MCMC) aided detector will be introduced in Section 5.3.

5.1 System Model and Decision Feedback Methods

The multi-user system model considered in this chapter is illustrated in Figure 5.1, where Q number of perfectly synchronized Mobile-Stations (MSs), each employing an M-element transmit antenna array, for communicating over a frequency-selective channel to a BS receiver, which has N number of Antenna Elements (AEs) are considered. Each MS's transmitter channel encodes the input bit-stream at a code-rate R_c, interleaves the encoded bits, maps them to the M transmit AEs and modulates the signals as seen in

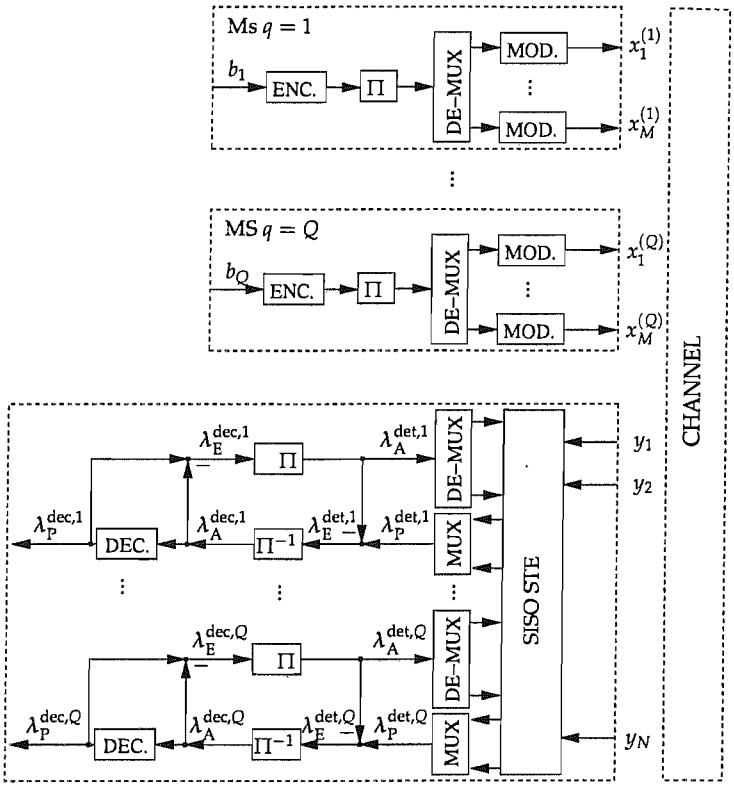


Figure 5.1: Baseband system model.

Figure 5.10. The channel encoder considered might be a simple convolutional scheme or a more sophisticated structure consisting of, for example, two serially concatenated codes. When comparing Figure 4.1 illustrating a narrowband single-user iterative MIMO detector to the system model presented in Figure 5.1, the presence of multiple users might yield near-far channel conditions, which make the detection process more challenging. Furthermore, the soft information considered by the STE is not provided by a single channel decoder but by Q number of channel decoders. In contrast to the narrowband MIMO detector of Chapter 4, it also will become evident during the course of this section that owing to the dispersive nature of the channel and the feed-forward order of the STE, a substantially increased search-space has to be explored by the receiver.

The channel is represented by the complex-valued coefficient $h_{nm,l}^{(q)}$ of the l^{th} multipath component between the q^{th} MS's AE m and the n^{th} BS receiver AE. Given the transmitted symbol $x_m^{(q)}(k)$, which is associated with the q^{th} MS's transmit AE m, the output signal of the n^{th} AE of the BS receiver at time instant k can be written as

$$y_n(k) = \sum_{q=1}^{Q} \sum_{m=1}^{M} \sum_{l=0}^{L-1} h_{nm,l}^{(q)} s_m^{(q)}(k-l) + \eta_n(k), \tag{5.1}$$

where L is the number of symbol-spaced multipath components. After transmission over an L-path MIMO channel, the SISO STE jointly calculates the *a posteriori* information associated with all users' signals. The

vector hosting the *a posteriori* Log Likelihood Ratio (LLR) values of the q^{th} user's signal can be written as

$$\boldsymbol{\lambda}_{\text{P}}^{\text{det},q} = [\lambda_{1,1}^{(q)},\ldots,\lambda_{1,N_{\text{bit}}}^{(q)},\ldots,\lambda_{M,1}^{(q)},\ldots,\lambda_{M,N_{\text{bit}}}^{(q)}]^T, \tag{5.2}$$

where $\lambda_{m,j}^{(q)}$ is the *a posteriori* LLR of the j^{th} symbol transmitted from the m^{th} transmit AE of the q^{th} MS. The subscript $_{\text{P}}$ as well as the superscript $^{\text{det}}$ have been omitted from the scalar elements of Equation (5.2) for the sake of notational simplicity.

For the reader's convenience we will now repeat the required system equations discussed in more detail in Section 3.4 in order to show how the LLRs required for iterative detection fit into this model. When considering a STE with feed-forward order N_{F} in addition to the CIR as part of the system model, the super-vector representing the new channel output can be written as

$$\mathbf{y}(k) = \mathbf{H}(k)\mathbf{x}(k) + \boldsymbol{\eta}(k), \tag{5.3}$$

where

$$\mathbf{y}(k) = [y_1(k),\ldots,y_{NN_{\text{F}}}(k)]^T$$

with $\boldsymbol{y}_i(k) = [y_1(k-i+1),\ldots,y_N(k-i+1)]^T$,

$$\boldsymbol{\eta}(k) = [\eta_1(k),\ldots,\eta_{NN_{\text{F}}}(k)]^T$$

with $\boldsymbol{\eta}_i(k) = [\eta_1(k-i+1),\ldots,\eta_N(k-i+1)]^T$ and

$$\mathbf{x}(k) = [\boldsymbol{x}_1(k),\ldots,\boldsymbol{x}_{L+N_{\text{F}}-1}(k)]^T$$

with $\boldsymbol{x}_i(k) = [x_1(k-i+1),\ldots,x_{MQ}(k-i+1)]^T$. The exact structure of $\mathbf{H}(k)$ has been defined in the context of Equation (3.15). The vector hosting the *a priori* LLR values provided for the detector by all users' channel decoders - which is associated with the super-vector $\mathbf{x}(k)$ of the transmitted symbols - can be written as

$$\boldsymbol{\lambda}_{\text{A}}^{\text{det}} = [(\boldsymbol{\lambda}_{\text{A},1}^{\text{det}})^T,\ldots,(\boldsymbol{\lambda}_{\text{A,L+N}_{\text{F}}-1}^{\text{det}})^T]^T, \tag{5.4}$$

where $\boldsymbol{\lambda}_{\text{A},i}^{\text{det}}$ is the QMN_{bit}-element *a priori* LLR vector associated with the symbol vector $\boldsymbol{x}_i$, which is defined in analogy to the *a posteriori* information of Equation (5.2). The LLR vectors associated with the extrinsic information and the *a priori* information can be obtained by substituting the subscript $_{\text{P}}$ with $_{\text{A}}$ and $_{\text{E}}$, respectively. The iterative detection process exchanging extrinsic information between the SISO STE and the channel decoders associated with the different users is carried out in analogy to the narrowband iterative detector discussed in Section 4.1.

Based on this system model, different DF methods might be employed within the STE. The first two design options are identical to the STE having no DF and to the hard DF aided STE discussed in Sections 3.4 and 3.5, respectively. These options will be discussed in the context of SISO STE in Sections 5.1.1 and 5.1.2, respectively. The above-mentioned novel soft DF approach is introduced in Section 5.1.3.

5.1.1 Using No Decision Feedback

If no DF is considered, the SISO algorithms discussed in Chapter 3 may readily be converted to SISO STE algorithms in analogy to the modifications performed in Chapter 4, when evolving from hard output MIMO schemes designed for non-dispersive channels to hard output STE algorithms contrived for dispersive environments. The corresponding conversion of the SISO single-user narrowband MIMO detectors of Chapter 4 to SISO STEs can be summarized as follows:

- The narrowband channel matrix $\boldsymbol{H}$ of Equation (4.2) is substituted by the system matrix $\mathbf{H}$ of Equation (5.3).
- The transmitted symbol vector $\boldsymbol{x}$ of Equation (4.2) is substituted by the super-vector $\mathbf{x}$ consisting of multiple delayed transmitted signal vectors as defined in Equation (5.3).
- The signal vector $\boldsymbol{y}$ of Equation (4.2) received over a narrowband channel is substituted by the super-vector $\mathbf{y}$, which constitutes the input to the STE as defined in Equation (5.3).
- The *a priori* information $\lambda_{\mathrm{A}}^{\mathrm{det}}$ of Equation (4.2) is substituted by the super-vector $\lambda_{\mathrm{A}}^{\mathrm{det}}$ of Equation (5.4).

SISO STE may now be readily implemented by using the SISO algorithms discussed in Sections 4.4.1-4.4.3 and applying them in the context of the new system model.

5.1.2 Hard Decision Feedback

Let us now contrast the hard DF aided STE to the procedure outlined in Section 5.1.1, while following the same principles:

- The narrowband channel matrix $\boldsymbol{H}$ of Equation (4.2) is substituted by the feed-forward channel matrix $\mathbf{H}_1$, which hosts the first $\mathfrak{M}(\Delta+1)$ columns of the system matrix $\mathbf{H}$ seen in Equation (5.3), where Δ is the decision delay at which the STE operates.
- The transmitted symbol vector $\boldsymbol{x}$ of Equation (4.2) is substituted by the super-vector $\mathbf{x}_1$, which consists of the first $\mathfrak{M}(\Delta+1)$ elements of $\mathbf{x}$. Again, $\mathbf{x}$ consists of multiple delayed transmitted signal vectors as defined in Equation (5.3).

- The signal vector $\boldsymbol{y}$ of Equation (4.2) received over a narrowband channel is substituted by the super-vector $\mathbf{r} = \mathbf{y} - \mathbf{H}_2\mathbf{x}_2$, where $\mathbf{H}_2$ as well as $\mathbf{x}_2$ represent the feedback matrix and the feedback symbol, respectively, which have been defined in Equation (3.64).

- The *a priori* information λ_A^{det} of Equation (4.2), which was generated in the narrowband scenario is substituted by the super-vector $\lambda_A^{det,DF} = [\lambda_{A,1}^{det} \ldots \lambda_{A,\Delta+1}^{det}]^T$ associated with the symbols in the feed-forward shift register of the SISO STE.

The benefit of hard DF is the reduction of the search space by removing the ambiguity of the already decided symbols, as discussed in Section 3.5. It was shown in Table 3.8 that for the MMSE STE of Section 3.5.1 hard DF yields a significant performance improvement, while for the classification based algorithms of Sections 3.5.3 and 3.5.4, such as the Bayesian and the RML detection algorithms, the associated complexity is reduced significantly.

5.1.3 Soft Decision Feedback

Since in the context of iterative detection soft information is provided by the detector, soft DF becomes possible. The soft information gleaned from the symbols, which have already been decided upon is exploited by incorporating the additional knowledge provided by the *a priori* information available at the STE. Assume that the *a priori* information provided by the channel decoder for the STE not benefiting from soft DF is given as $\lambda_A^{det} = [(\lambda_{E,1}^{dec})^T, \ldots, (\lambda_{E,L+N_F-1}^{dec})^T]^T$. The *a priori* information used by the iterative STE incorporating soft DF can then be written as

$$\lambda_A^{det,SDF} = [(\lambda_{E,1}^{dec})^T, \ldots, (\lambda_{E,\Delta+1}^{dec})^T, (\lambda_{E,\Delta+2}^{det} + \lambda_{E,\Delta+2}^{dec})^T, \ldots, (\lambda_{E,L+N_F-1}^{det} + \lambda_{E,L+N_F-1}^{dec})^T]^T. \quad (5.5)$$

In the proposed soft DF scheme, the extrinsic information obtained by the SISO STE for bits associated with the symbols of the different antennas at time k are directly used as *a priori* information for the detection of the symbols at time $k+1$. Let us now briefly contrast the soft DF STE procedure to the procedures outlined in Sections 5.1.1 and 5.1.2:

- The narrowband channel matrix $\boldsymbol{H}$ of Equation (4.2) is substituted by the system matrix $\mathbf{H}$ of Equation (5.3).

- The transmitted symbol vector $\boldsymbol{x}$ of Equation (4.2) is substituted by the super-vector $\mathbf{x}$ consisting of multiple delayed transmitted signal vectors as defined in Equation (5.3).

- The signal vector $\boldsymbol{y}$ of Equation (4.2) received over a narrowband channel is substituted by the super-vector $\mathbf{y}$ of Equation (5.3).

Table 5.1: Extension of SISO MIMO detection algorithms to STE algorithms, when considering different DF methods.

	no DF	hard DF	soft DF
Channel matrix $\boldsymbol{H}$	$\mathbf{H}$	$\mathbf{H}_1$	$\mathbf{H}$
Received signal $\boldsymbol{y}$	$\mathbf{y}$	$\mathbf{r} = \mathbf{y} - \mathbf{H}_2\mathbf{x}_2$	$\mathbf{y}$
Transmitted signal $\boldsymbol{x}$	$\mathbf{x}$	$\mathbf{x}_1$	$\mathbf{x}$
A priori LLR λ_A^{det}	λ_A^{det}	$\lambda_A^{det,DF}$	$\lambda_A^{det,SDF}$

- The *a priori* information λ_A^{det} of Equation (4.2), which was valid for the narrowband scenario is substituted by the super-vector $\lambda_A^{det,SDF}$.

In Table 5.1 the three different DF approaches of Sections 5.1.1-5.1.3 have been summarized. When comparing soft and hard DF, it can be observed that as a benefit of the smaller system matrix generated after the symbol decision of hard DF, the dimensionality of the search space is reduced at the cost of a potential error propagation. By contrast, soft DF does not offer this complexity reduction, but as a benefit of taking into account the reliability of the previously detected symbols, further performance gains are expected when compared to both the STE dispensing with DF as well as over the hard DF aided STE.

5.2 Conventional Soft-Input Soft-Output STE

In this section the performance of classic detection methods, namely the MMSE, the Bayesian as well as the RML algorithms discussed earlier in the context of narrowband MIMO detection in Chapter 4, are investigated in iterative STEs using EXIT charts. Furthermore, the computational cost imposed by the different SISO STEs is discussed. The evolution from narrowband SISO MIMO detectors to MMSE SISO STEs operating in dispersive environments, has already been proposed for iterative MMSE detectors in [76] using no DF.

5.2.1 Minimum Mean Squared Error STE

The computational complexity of the MMSE STE during the first iteration is identical to that of the hard output STE dispensing with *a priori* information, which was quantified in Table 3.1. For all subsequent iterations, the complexity imposed can be evaluated by substituting NN_F for N and $QM(N_F + L - 1)$ for M in Table 4.2, when refraining from the employment of DF. Similarly, we have to substitute NN_F for N and $QM(\Delta + 1)$ for M for the STE of Section 5.1.2 benefiting from hard DF in order to appropriately adopt the

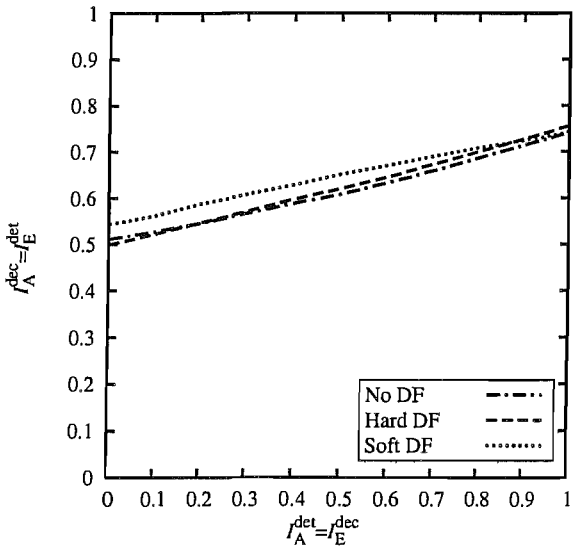
complexity evaluation of the iterative narrowband MMSE MIMO detector of Section 4.4.1. Quantitatively, for the STE using no DF and for the soft DF aided detector of Sections 5.1.1 and 5.1.3 , this yields a computational cost imposed by the MMSE detector expressed as:

$$C = 4NN_{\mathrm{F}}MQ(N_{\mathrm{F}} + L - 1)N_{\mathrm{bit}} + 8N^2N_{\mathrm{F}}^2MQ(N_{\mathrm{F}} + L - 1)N_{\mathrm{bit}} + 2NN_{\mathrm{F}} + O(4N^3N_{\mathrm{F}}^3) \quad (5.6)$$

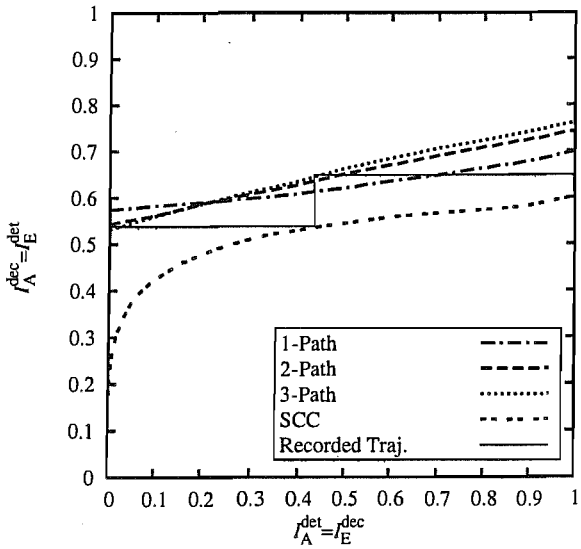
in terms of the number of real-valued additions and multiplications, which are imposed by the evaluation of the MMSE weight vector in Equation (4.12) under consideration of the new system model. Note that in our complexity evaluations the real-valued binary STE system model corresponding to the complex-valued system model of Equation (5.3) was considered, which was evolved from the context of Equation (4.12). Similarly, for hard DF we obtain

$$C = 4NN_{\mathrm{F}}MQ(\Delta + 1)N_{\mathrm{bit}} + 8N^2N_{\mathrm{F}}^2MQ(\Delta + 1)N_{\mathrm{bit}} + 2NN_{\mathrm{F}} + O(4N^3N_{\mathrm{F}}^3), \quad (5.7)$$

which, again, was derived form Table 4.2. Taking into account that the feed-forward order N_{F} is generally chosen to be the same as the length of the CIR L, it can be seen from Equations (5.6) and (5.7) that the complexity of all feedback aided approaches depends cubically on the length of the CIR.



(a) Two-path equal-power independent Rayleigh fading channel and a STE characterized by $N_{\mathrm{F}} = 2$, $\Delta = 1$ and $N_{\mathrm{B}} = 1$ was used, when feedback was considered.



(b) L-path equal-power independent Rayleigh fading channel and a STE using soft DF was employed, which was characterized by $N_{\mathrm{F}} = L$, $\Delta = L - 1$ and $N_{\mathrm{B}} = L - 1$.

Figure 5.2: MMSE EXIT function at $E_{\mathrm{b}}/N_0 = 4$ dB for a system using $M = 2$ transmit AEs and $N = 2$ receive AEs. 4QAM was considered. A random interleaver having a length of 10 000 bits was used.

In Figure 5.2a the EXIT function of the MMSE detector is shown for a single-user (2×2) MIMO system communicating over a two-path equal-power channel CIR characterizing all antenna links. The STE was operating with a feed-forward order of $N_{\mathrm{F}} = 2$ and a decision delay of $\Delta = 1$. The feedback order for

both the hard as well as soft DF STE was chosen to be $N_B = 1$. It can be observed that for the scenario considered, soft-feedback offers the best performance, while the employment of hard DF as well as no DF result in a slightly lower mutual information at the STE's output. It can also be observed that for $I_A^{det} = 0$ the utilization of hard DF results in a lower mutual information at the output of the detector than in the scenario where the STE refrains from using DF.

The EXIT functions illustrated in Figure 5.2b have been recorded for the same system as in Figure 5.2a using soft DF and a varying number of equal-power CIR taps. The STE parameters were chosen to be $N_F = L$, $\Delta = L - 1$ and $N_B = N_F - 1$. It can be observed that if the number of paths is increased from one to two, the area under the EXIT transfer function becomes larger and thus the system becomes capable of supporting a higher throughput. Furthermore, a steeper EXIT function slope can be observed. If the number of paths is increased to three, the theoretically achievable throughput increases only marginally. This behaviour seems to follow the channel capacity trends observed in Figure 3.4 for the same system, where it was found that the channel capacity increases more substantially upon increasing the number of paths from one to two, than it does when increasing the number of paths from two to three. The steeper EXIT function of the detector recorded for an increased number of paths implies that a higher number of iterations is required for reaching the point of convergence in the EXIT chart. Consequently, as expected, a higher complexity is imposed, when communicating at the highest possible data-rate supported by the detector. In addition to the detector's EXIT function, the projected EXIT function of a Serially Concatenated Convolutional (SCC) code, consisting of a rate-1/2 RSC code using the octally represented generators of $\boldsymbol{g} = [6\ 5]$ and and a unity-rate code having the octal generators $\boldsymbol{g} = [3\ 1]$ using $N_{\text{inner}} = 10$ inner iterations, is also shown in Figure 5.2b. It can be seen that the trajectory recorded for the soft DF STE at $E_b/N_0 = 4$ dB and at an interleaver length of 10 000 bits follows the performance trends predicted by the EXIT curves closely. Recall from Figure 4.4 that for the observed scenario, the system would operate at an infinitesimally low BER owing to the beneficial characteristics of the SCC code considered, as discussed earlier in Section 4.3.

5.2.2 Bayesian STE

The structure of the Bayesian SISO detector is identical to that of the Bayesian STE discussed in Section 3.4.3, which was returning hard outputs, with the only difference being that in this section soft values are generated. Rather than calculating a hard estimate of the transmitted symbol on the basis of Equation (3.44), the Bayesian SISO STE calculates the output LLR values, as demonstrated in Equation (4.16) for the narrowband Bayesian SISO detector. For convenience here we repeat the complexity formulae of Tables 3.3 and Table 3.7 derived in Section 3.4.3.1. Explicitly, the calculation of the legitimate channel output states of Equation (3.36) imposes a computational cost expressed in terms of real-valued multiplications

and additions formulated as

$$\mathcal{C} = 2N_sN_FN(2\mathfrak{M}(N_F + L - 1) - 1)/N_{bit}/\mathfrak{M} \text{ for } N_s = 2^{N_{bit}\mathfrak{M}(N_F+L-1)} \quad (5.8)$$

for the STE either using no DF or considering soft DF, where $\mathfrak{M} = QM$. By contrast, the number of real-valued additions and multiplications in Equation (3.65) is reduced to

$$\mathcal{C} = 2N_sN_FN(2\mathfrak{M}(\Delta + 1) - 1)/N_{bit}/\mathfrak{M} \text{ for } N_s = 2^{N_{bit}\mathfrak{M}(\Delta+1)} \quad (5.9)$$

when using hard DF. Additionally, the evaluation of the LLR values formulated in Equation (4.16) requires

$$\mathcal{C} = N_s(6NN_F + 1) \text{ for } N_s = 2^{N_{bit}\mathfrak{M}(N_F+L-1)} \quad (5.10)$$

number of real-valued additions and multiplications for the STE not benefiting from DF or considering soft DF and

$$\mathcal{C} = N_s(6NN_F + 1) \text{ for } N_s = 2^{N_{bit}\mathfrak{M}(\Delta+1)} \quad (5.11)$$

for the hard DF STE. Note that the calculation of the *a priori* probability of each legitimate channel state has not been considered in these complexity evaluations.

It can be seen from Equation (5.8) as well as Equation (5.9) that the complexity of the Bayesian SISO STE is exponentially proportional to both the length of the CIR as well as to the number of bits per symbol and to the number of transmit antennas in the system. Therefore the resultant complexity becomes excessive for relatively low values of the above-mentioned parameters. Consider for example a single-user (2×2)-antenna MIMO system employing 4QAM for transmission over a three-path channel. For this system the Bayesian STE would have to consider $N_s = 2^{2\cdot2(3+3-1)} = 1\,048\,576$ legitimate channel output states, which exceeds any realistic practical limitations at the current state-of-the-art.

The system considered for generating the EXIT functions of Figure 5.3 was therefore a (2×2)-antenna single-user MIMO system employing a STE characterized by $N_F = 2$, $\Delta = 1$ and $N_B = 1$, when DF was considered. The channel was assumed to be a two-path equal-power independently fading Rayleigh channel. It can be seen in Figure 5.3 that the EXIT functions of the Bayesian STE have a higher starting point for $I_A^{det} = 0$ than the MMSE STE even for this full-rank system. Furthermore, it can be observed that the Bayesian detector appears to be more robust against DF errors induced by hard DF, when compared to the MMSE based STE. Similar to the MMSE STE, the Bayesian STE exhibits its best performance when soft DF is employed. In addition to the Bayesian STE's EXIT functions, Figure 5.3 shows the EXIT function of a SCC code consisting of a rate-1/2 RSC code employing the octally represented generators of $\boldsymbol{g} = [6\ 5]$ and a unity-rate code with octal generators of $\boldsymbol{g} = [3\ 1]$ using $N_{inner} = 10$ inner iterations. The trajectory

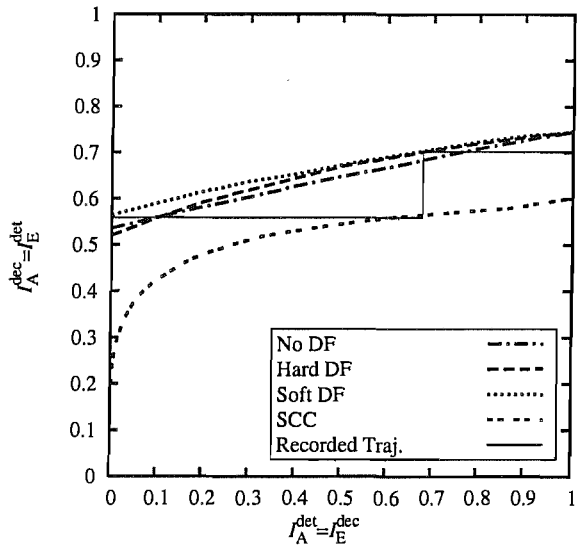


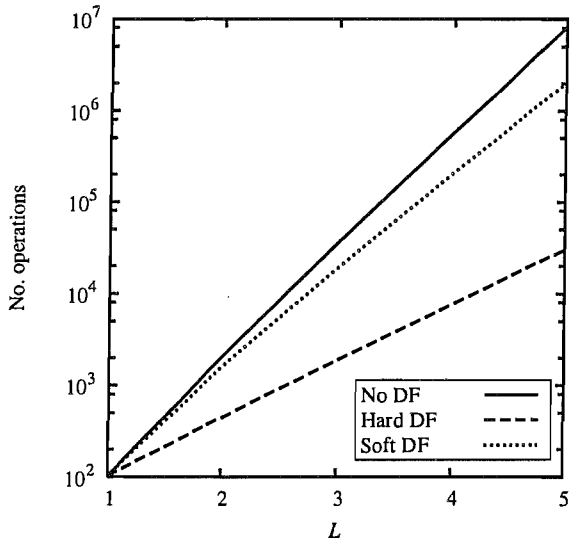
Figure 5.3: Bayesian EXIT function at $E_b/N_0 = 4$ dB for a system using $M = 2$ transmit AEs and $N = 2$ receive AEs. 4QAM was considered. A two-path equal-power independent Rayleigh fading channel and a STE characterized by $N_F = 2$, $\Delta = 1$ as well as $N_B = 1$ were considered for the feedback-aided scenarios. The interleaver-depth was chosen to be 10 000 bits.

recorded at $E_b/N_0 = 4$ dB for an interleaver length of 10 000 bits shows a good agreement with the predicted performance.

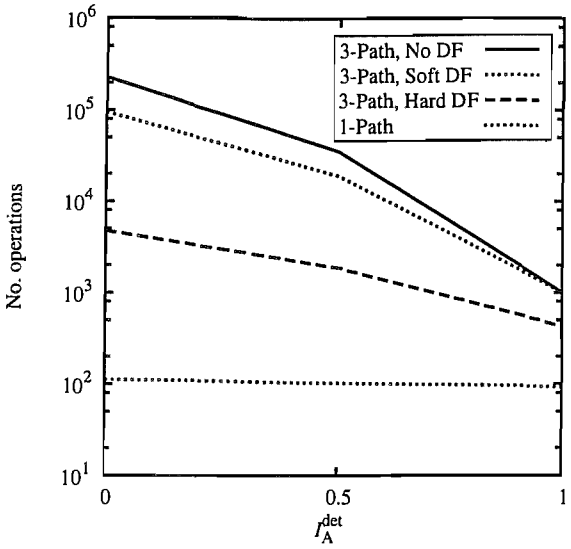
5.2.3 Reduced Complexity Max-Log STE

The RML STE's structure is also based on that of the narrowband MIMO detector introduced in Section 4.4.3 and its complexity is, again, quantified in terms of the number of real-valued multiplications and additions required for the detection of a single bit. Note that in the scenario the truncation scheme discussed in Section 3.5.4 has not been considered.

Figure 5.4a shows the number of real-valued operations required for the evaluation of the soft-information of a single bit for a single-user (2×2)-antenna MIMO system at $E_b/N_0 = 4$ dB and an *a priori* information at the STE associated with $I_A^{det} = 0.5$ recorded as a function of the number of CIR path L. The STE parameters were chosen to be $N_F = L$ and $\Delta = L - 1$. When DF was considered, the feedback order was set to $N_B = L - 1$. It can be observed in Figure 5.4a that in contrast to the hard-detection based STE of Sections 3.5.3 and 3.5.4, the complexity of the iterative STE increases exponentially as a function of the number of CIR taps, i.e. with the channel-induced dispersion, which has two main reasons. Firstly, the hard-output STE evaluates the ML search tree only once, while the RML based STE has to evaluate the search tree $(N_{bit}\mathfrak{M} + 1)$ number of times in order to generate the soft information based on the Max-Log approximation [118]. Secondly, in the context of the iterative STE the receiver operates at a relatively low SNR, where we have observed in Sections 3.4.4 and 2.3.4 that the tree-search imposes a higher computational cost than that imposed at higher SNR values. Furthermore, it can be seen in Figure 5.4a that upon using soft DF the



(a) Computational complexity imposed by the RML STE as a function of the CIR length. The *a priori* mutual information was fixed to $I_A^{dec} = 0.5$.



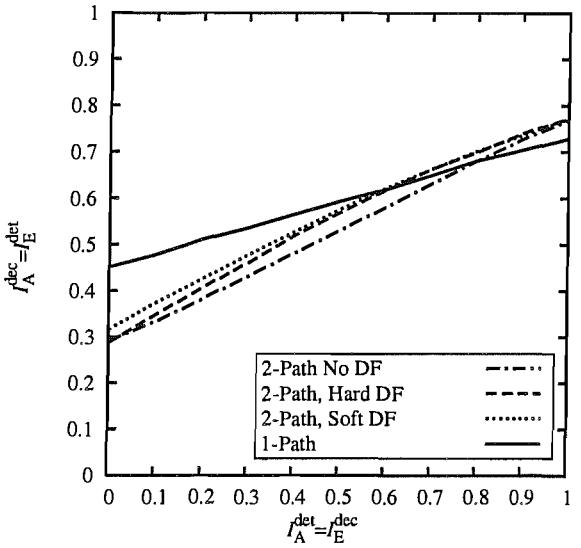
(b) Complexity as a function of *a priori* information for a three-path channel.

Figure 5.4: Computational complexity in terms of real-valued additions plus multiplications imposed by the RML detector, which was recorded during simulations according to Table 3.4 for a single-user (2×2)-antenna MIMO system operating at $E_b/N_0 = 4$ dB and for a STE characterized by $N_F = L$, $\Delta = L - 1$ and $N_B = L - 1$ when feedback-aided detection was considered.

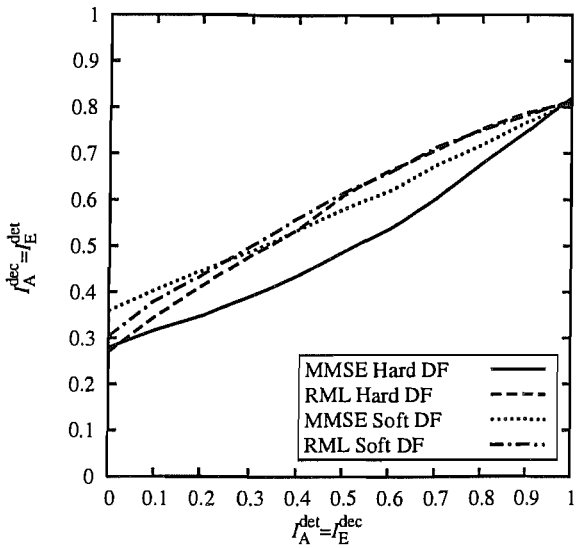
associated computational cost is reduced compared to the STE dispensing with DF, although both algorithms have to explore identical search spaces.

The graphs seen in Figure 5.4b quantify the complexity imposed by the RML STE as a function of the *a priori* information available at the detector, which was plotted for a single-user (2×2)-antenna MIMO system at $E_b/N_0 = 4$ dB, when communicating over a three-path equal-power MIMO channel and using a STE characterized by $N_F = 3$, $\Delta = 2$ as well as $N_B = 2$. Additionally, the complexity of the corresponding single-user (2×2)-antenna narrowband MIMO system is also illustrated. It can be observed that in contrast to the narrowband system, where the complexity is little influenced by the mutual information at the input of the detector, the complexity imposed by the RML aided SISO STE decreases upon increasing the mutual information associated with the *a priori* LLRs at the input of the STE.

The performance of the RML-aided SISO STE is, again, evaluated using EXIT chart analysis, as already seen in Figures 5.2 and 5.3 for the MMSE and for the Bayesian SISO STE, respectively. In Figure 5.5a we have illustrated the EXIT function of a (2×2)-antenna single-user MIMO system using 16QAM. The channel was assumed to be a two-path equal-power transmission medium and the STE was characterized by $N_F = 2$, $\Delta = 1$ as well as $N_B = 1$. As a reference, the EXIT function associated with the corresponding narrowband system was also plotted. It can be observed that owing the effects of the frequency-selective channel, the EXIT function of the narrowband scenario is further tilted, which results in a higher number of iterations required for approaching the maximum achievable data-throughput that may be supported by



(a) 2 transmit AEs, $E_b/N_0 = 6$ dB, 16QAM was considered



(b) 4 transmit AEs, $E_b/N_0 = 8$ dB, 4QAM was considered

Figure 5.5: RML EXIT function for $N = 2$ receive AEs.Two-path equal-power independent Rayleigh fading and a STE characterized by $N_F = 2$, $\Delta = 1$ and $N_B = 1$ when feedback-aided detection was considered.

the detector. Furthermore, it can be observed that even though the multipath channel has a higher channel capacity, the RML detector is unable to exploit this fact. Quite the contrary, it actually supports a lower effective throughput, since the area under the EXIT function is decreased for transmission over a frequency-selective channel.

The effects of rendering the system rank-deficient have been investigated in Figure 5.5b, where the EXIT function is shown for the same system as the one that was considered for Figure 5.5a, but now employing $M = 4$ and 4QAM. It can be observed that in this rank-deficient system, both the MMSE and the RML aided STE using soft DF perform similarly. Interestingly, the RML-aided detector performs worse than the MMSE STE for $I_A^{det} = 0$, which is owing to the employment of the Max-Log approximation invoked by the RML detector. This phenomenon will be discussed in more detail in the next paragraph. It can also be seen that if hard DF rather than soft DF is used, the performance of the RML detector is substantially reduced and in fact, it becomes inferior to that of the MMSE STE. This observation highlights again the sensitivity of the hard DF-aided SISO STEs to error-propagation.

It was mentioned in the previous paragraph that using the Max-Log approximation in the context of iterative STEs is associated with a significant performance degradation at low SNR. This performance degradation becomes more pronounced for channels, which are subject to a higher degree of dispersion. The reason for this becomes clear when recalling the assumption made for deriving the Max-Log detection scheme. When neglecting the influence of the *a priori* information and assuming Binary Phase Shift Keying (BPSK) modulation, it transpires from the Max-Log principle introduced in Equation (4.19) that the following ap-

proximations are made:

$$\begin{aligned} P(x_m = \pm 1|\mathbf{y}) &= \sum_{\check{\mathbf{x}}^{(i)} \in \mathcal{X}_m^{\pm}} P(\check{\mathbf{x}}^{(i)}|\mathbf{y}) \\ &\propto \sum_{\check{\mathbf{x}}^{(i)} \in \mathcal{X}_m^{\pm}} \exp\left(-\frac{||\mathbf{y} - \mathbf{H}\check{\mathbf{x}}^{(i)}||^2}{2\sigma_n^2}\right) \\ &\approx \max_{\check{\mathbf{x}}^{(i)} \in \mathcal{X}_m^{\pm}} \exp\left(-\frac{||\mathbf{y} - \mathbf{H}\check{\mathbf{x}}^{(i)}||^2}{2\sigma_n^2}\right), \end{aligned} \tag{5.12}$$

where $\mathcal{X}_m^{\pm}$ is the set of hypothetically transmitted symbol vectors $\check{\mathbf{x}}^{(i)}$, with $\check{x}_m = \pm 1$ and the symbol $\propto$ indicates 'proportional to'. An approximation of the *a posteriori* probability based on considering only the dominant hypothetical channel output states will hold only, if the effect of all other states becomes negligible, which is indeed the case at high SNR values. By contrast, at low SNR values the Max-Log approximation of Equation (5.12) will always be sub-optimum, since the *a posteriori* probabilities of all states are likely to be of the same order of magnitude. When moving from a narrowband scenario to a wideband scenario, the number of dispersed multi-user MIMO constellation points in the search-space increases and hence the hypothetical channel output states are less well separated, although the same noise power is observed. The Max-Log approximation therefore discards a higher number of points from the channel output space, namely those which lay close to the specific state associated with the maximum *a posteriori* probability and thus is more likely to return exaggeratedly high, but potentially flawed LLR values. By contrast, even though the MMSE algorithm uses only a linear filter, it always takes into account the complete set of hypothetical channel output states.

If we assume that similarly to all our previous examples, all transmit AEs are associated with identical transmit powers and the number of super-imposed MIMO signals is sufficiently high, the MMSE STE's output may accurately be approximated by the Gaussian distribution. If, however, not all transmit AEs belong to the same user, i.e. a multi-user scenario is considered, the power of the signals received form the different mobile transmitters might substantially differ. In this case the output of the MMSE detector may no longer be accurately approximated as being Gaussian [31].

In Figure 5.6 we therefore consider a two-user MIMO system, each user having two transmit AEs. The received power of user one was fixed to $E_b/N_0 = 4$ dB, while the second users' power was varied from 4 to 19 dB. The receiver also used a two-element antenna array, resulting in a rank-deficient scenario associated with a total of $\mathfrak{M} = 4$ transmit AEs and $N = 2$ receive AEs. When the channel had two paths and was hence frequency-selective, the STE parameters were chosen to be $N_F = L - 1$, $\Delta = L - 1$ and $N_B = L - 1$. Soft DF was employed. Figure 5.6 shows the extrinsic information at the output of the STE computed on the basis of the LLR values of both users as a function of the higher-power user's received signal strength. Observe in Figure 5.6 that at a low received power the MMSE detector outperforms the RML detector.

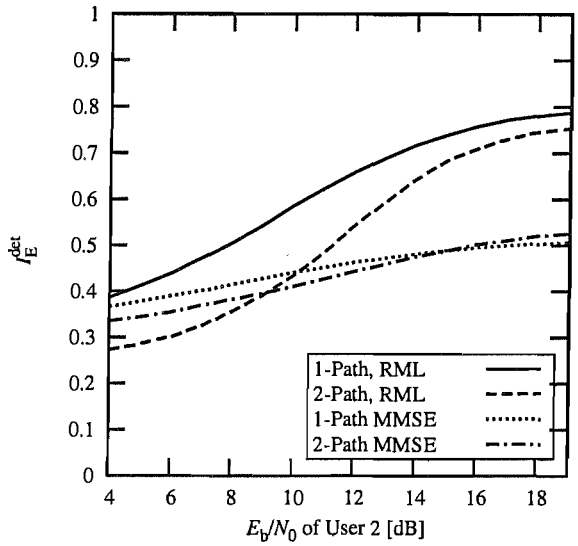


Figure 5.6: Joint mutual information performance of a two-user MIMO system, where the signal of user 1 was received at $E_b/N_0 = 4$ dB, while the second user's power was varied from 4 to 19 dB. Each user employed a two-element transmit AE, while the receiver used a two-element receive antenna array. If the channel was frequency selective, the STE parameters where chosen to be $N_F = L-1$, $\Delta = L-1$ and $N_B = L-1$. Soft DF was employed.

When the power of the high-power user is increased, however, the RML detector substantially outperforms the MMSE based STE. When comparing the single-path RML STE scenario to the two-path RML STE, it can be seen that for the dispersive channel the RML aided STE gradually approaches the performance of the corresponding narrowband scenario. By contrast the MMSE STE is incapable of exploiting the additional diversity provided by the dispersive channel for a high power of the stronger user.

5.2.4 Notes on Rank-Deficiency

From our observations in Chapter 4 as well as from the results presented in this section, we surmise that the slope of the EXIT transfer function is closely related to the rank-deficiency of the system. In this context we refer to the system as being rank-deficient, if the number of rows in the channel matrix is smaller than the number of linearly independent columns, where the ratio of columns and rows is denoted as

$$\Gamma = \frac{\mathfrak{M}(L+N_F-1)}{NN_F}. \tag{5.13}$$

From the slopes of the EXIT functions seen in Figure 5.5 and Figure 4.12 it might be concluded that increasing the number of bits per symbol also increases the slope of the EXIT chart. This however is only valid for the specific symbol mapping used for the hierarchical 16QAM constellation considered. Other mapping schemes such as Gray mapping may result in almost horizontal EXIT functions, as it will be shown in Section 5.3.4. Therefore the number of bits per symbol has not been included in the definition of rank-deficiency, despite the fact that when considering a real-valued binary-system the associated grade of

rank-deficiency would be increased.

It can be observed from Equation (5.13) that when choosing the feed-forward order of the STE to be equal to the number of CIR taps: $N_F = L$, then the grade of rank-deficiency of the system matrix is increased for an increasing number of CIR taps. When opting for $N_F = L$, we can re-write Equation (5.13) as

$$\Gamma = \frac{\mathfrak{M}}{N} \cdot \frac{2L-1}{L}, \tag{5.14}$$

which approaches $2\frac{M}{N}$ for a high number of channel taps L. It can be seen from Equation (5.14) that by encountering a higher number of CIR taps, the system becomes more rank-deficient and therefore renders the detection more challenging.

5.2.5 Concluding Remarks

In this section it was shown that the MMSE SISO equalizer constitutes a powerful iterative STE, while the Bayesian STE's practical applicability is limited by its complexity. The RML STE succeeds in outperforming the MMSE detector for the 4QAM cases considered only in near-far scenarios. The employment of the MMSE SISO detector in its present form is limited to 4QAM signals, while higher-order QAM can only be detected with the aid of the RML detector, which is however limited to employment in hierarchical 16QAM. Furthermore, we have seen that the RML algorithm has its limitations owing to the Max-Log approximation.

It is therefore desirable to employ specifically designed algorithms for iterative STE, which allow us to approximate the Maximum A Posteriori (MAP) solution and do not rely on the Max-Log approximation, while remaining capable of detecting arbitrary QAM constellations for the sake of supporting different mapping schemes. Furthermore a desirable algorithm should have the capability of operating efficiently in large search spaces. An attractive set of such detection algorithms is constituted by the family of Markov Chain Monte Carlo based algorithms discussed in the next section.

5.3 Markov Chain Monte Carlo Aided Detection

In the previous section we have seen that the full Bayesian STE, as well as the RML aided STE impose a computational cost on the receiver, which may limit their employment in practical systems at the time of writing. In this section we therefore introduce a novel scheme referred to as the Markov Chain Monte Carlo (MCMC) aided detector, which is known for its capability of exploring large search spaces at a moderate computational cost. Ideally, the specific fraction of the ML search-space that will be explored should still contain the most likely set of solutions.

MCMC detection combines two different techniques. Firstly, it uses the technique of Monte Carlo integration [69] in order to evaluate integrals over high-dimensional spaces. More explicitly, Monte Carlo integration aims for solving an integral over a high-dimensional space $\mathcal{X}$ on the basis of generating random samples $\check{x}$ of $\mathcal{X}$ and then approximating the integral of interest using these random samples. This method will be discussed in detail in Section 5.3.2. Originally, Monte Carlo integration was introduced in nuclear physics [142], where it was used in the early 1950s for calculating the properties of the newly discovered neutrons. Later it has found its way into signal-processing algorithms and since the early 1990s it has also been applied to wireless communications [69].

The second term 'Markov chain' refers to the way the random samples $\check{x}$ necessary for applying Monte Carlo integration are generated. The main idea is to represent the observed process, in our case the detection of a symbol vector, as a Markov chain which in turn is used to draw samples from the space of interest [143]. Markov chain representation of detection processes will be discussed In Section 5.3.3.

Over the last few years MCMC detection has been applied in several different communication problems, such as for example in [144], where a MCMC aided blind OFDM detector was proposed. In [145] MCMC-aided detection was considered in the context of Multi-Carrier CDMA. These seminal solutions have been further developed in [58, 69], where MCMC has been applied to both MIMO systems as well as CDMA MUDs. Other applications include for example wideband beamforming [146]. In [57] MCMC-detectors were compared to other detection methods, such as Sphere Decoding (SD). The convergence behaviour of MCMC aided detection has been analyzed in [147].

5.3.1 Introduction

In this section we will first introduce our basic system model and then discuss the underlying relationships between the different variables observed from a detection theory perspective. *For notational simplicity we will consider only the detection of narrowband MIMO systems in this section. The extension of this solution to the STE system model may be analogously achieved as outlined in Section 5.1 for the MMSE, the Bayesian as well as the RML aided detectors.*

Under the assumption of perfect synchronization, the relation between the signal transmitted by the MS's AEs and the channel's output is described by an $(N \times M)$-dimensional matrix $\boldsymbol{H}$, where the $(n,m)^{\text{th}}$ element of the matrix is given by h_{nm}, while N and M represent the number of receive and transmit AEs, respectively. The channel's output vector $\boldsymbol{y}(k)$ can now be expressed as

$$\boldsymbol{y}(k) = \boldsymbol{H}(k)\boldsymbol{x}(k) + \boldsymbol{\eta}(k), \tag{5.15}$$

where the column vector $\boldsymbol{x}(k) = [x_1(k), \ldots, x_M(k)]^T$ contains the symbols transmitted by the MS's AEs

and the associated AWGN is given by $\eta(k) = [\eta_1(k), \ldots, \eta_N(k)]^T$. In our forthcoming discourse the time index k is neglected, where this is possible without ambiguity.

When detecting the transmitted signal $\boldsymbol{x}$ using an iterative receiver structure, the quantity required for generating soft-information is the *a posteriori* probability of the transmitted symbol $P(\boldsymbol{x}|\boldsymbol{y}, \boldsymbol{\lambda}_A^{det})$, conditioned on the received symbol vector of Equation (5.15) and the *a priori* LLR provided by the channel decoder given as $\boldsymbol{\lambda}_A^{det} = [\lambda_{A,1}^{det}, \ldots, \lambda_{A,MN_{bit}}^{det}]^T$. Let us now assume that the legitimate values of the transmitted symbol x_m are ± 1.

By using Bayes' theorem the *a posteriori* probability of the transmitted symbol vector $\boldsymbol{x}$ defined in the context of Equation (5.15) can be written as

$$\begin{aligned} P(\boldsymbol{x}|\boldsymbol{y}, \boldsymbol{\lambda}_A^{det}) &= \frac{p\left(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{\lambda}_A^{det}\right) P\left(\boldsymbol{x}|\boldsymbol{\lambda}_A^{det}\right)}{p(\boldsymbol{y}|\boldsymbol{\lambda}_A^{det})} \\ &= \frac{p\left(\boldsymbol{y}|\boldsymbol{x}\right) P\left(\boldsymbol{x}|\boldsymbol{\lambda}_A^{det}\right)}{p(\boldsymbol{y})}, \end{aligned} \tag{5.16}$$

where the second line follows from the fact that the channel output $\boldsymbol{y}$ and the *a priori* LLRs $\boldsymbol{\lambda}_A^{det}$ are independent random variables. The PDF of the channel output can be obtained with the aid of integration or summation:

$$p(\boldsymbol{y}) = \sum_{\check{\boldsymbol{x}}^{(i)} \in \mathcal{X}} p(\boldsymbol{y}|\check{\boldsymbol{x}}^{(i)}) = \text{const}, \tag{5.17}$$

where $\mathcal{X}$ represents the set of all hypothetically transmitted MIMO symbol vectors $\check{\boldsymbol{x}}^{(i)}$ with $1 \leq i \leq 2^M$, given the 2^M-ary composite signal of the M transmit AEs. Since $p(\boldsymbol{y})$ is constant and does not depend on the transmitted symbol vector $\boldsymbol{x}$, we can rewrite Equation (5.16) as

$$P(\boldsymbol{x}|\boldsymbol{y}, \boldsymbol{\lambda}_A^{det}) \propto p\left(\boldsymbol{y}|\boldsymbol{x}\right) P\left(\boldsymbol{x}|\boldsymbol{\lambda}_A^{det}\right),$$

where $\propto$ represents 'proportional to'. Finding an expression for the *a posteriori* probability of Equation (5.16) was only the first step. In order to generate soft-information, the marginalization of $P(\boldsymbol{x}|\boldsymbol{y}, \boldsymbol{\lambda}_A^{det})$ given as $P(x_m = \pm 1|\boldsymbol{y}, \boldsymbol{\lambda}_A^{det})$ is required. More explicitly, the marginalization of the MIMO symbol vector probability is defined as the integration over all legitimate MIMO symbol vectors, which can be expressed with the aid of their sums as

$$P(x_m = \pm 1|\boldsymbol{y}, \boldsymbol{\lambda}_A^{det}) \propto \sum_{\check{\boldsymbol{x}}^{(i)} \in \mathcal{X}_m^{\pm}} p\left(\boldsymbol{y}|\check{\boldsymbol{x}}^{(i)}\right) P\left(\check{\boldsymbol{x}}^{(i)}|\boldsymbol{\lambda}_A^{det}\right), \tag{5.18}$$

where $\mathcal{X}_m^{\pm}$ is the set of all hypothetically transmitted symbol vectors $\check{\boldsymbol{x}}^{(i)}$, where the m^{th} symbol is equal

± 1. Alternatively, Equation (5.18) can also be written as

$$P(x_m = \pm 1 | \boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}) \propto \sum_{\check{\boldsymbol{x}}^{(i)} \in \mathcal{X}_{-m}} p\left(\boldsymbol{y} | x_m = \pm 1, \check{\boldsymbol{x}}_{-m}^{(i)}\right) P\left(x_m = \pm 1, \check{\boldsymbol{x}}_{-m}^{(i)} | \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}\right), \tag{5.19}$$

where $\boldsymbol{x}_{-m}^{(i)} = [\check{x}_1^{(i)}, \ldots, \check{x}_{m-1}^{(i)}, \check{x}_{m+1}^{(i)}, \ldots, \check{x}_M^{(i)}]^T$ is a vector containing all symbols, except for the m^{th} symbol and $\mathcal{X}_{-m}$ is the set of hypothetically possible transmitted symbol vectors $\check{\boldsymbol{x}}_{-m}^{(i)}$.

Finally, the quantities required in all equations given above, namely the conditional PDF of the received signal and the *a priori* probability can be written as

$$\begin{aligned} p\left(\boldsymbol{y} | \check{\boldsymbol{x}}^{(i)}\right) &= \frac{1}{(2\pi\sigma_{\mathrm{n}}^2)^N} \exp\left(-\frac{\| \boldsymbol{y} - \boldsymbol{H}\check{\boldsymbol{x}}^{(i)} \|^2}{2\sigma_{\mathrm{n}}^2}\right) \\ &\propto \exp\left(-\frac{\| \boldsymbol{y} - \boldsymbol{H}\check{\boldsymbol{x}}^{(i)} \|^2}{2\sigma_{\mathrm{n}}^2}\right) \end{aligned} \tag{5.20}$$

and

$$P(\check{\boldsymbol{x}}^{(i)} | \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}) = \prod P(\check{x}_m^{(i)} | \lambda_{\mathrm{A}}^{\mathrm{det}}) \tag{5.21}$$

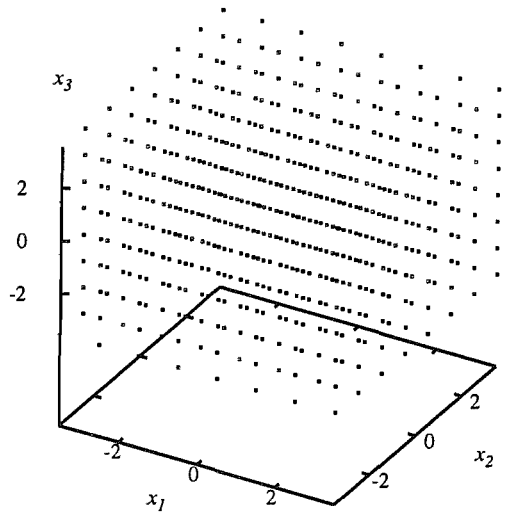
$$= \prod \frac{\exp\left(\frac{1}{2}\lambda_{\mathrm{A,m}}^{\mathrm{det}}\right)}{1 + \exp(\lambda_{\mathrm{A,m}}^{\mathrm{det}})} \exp\left(\mathtt{sign}\left(\check{x}_m^{(i)}\right) \frac{1}{2}\lambda_{\mathrm{A,m}}^{\mathrm{det}}\right), \tag{5.22}$$

respectively.

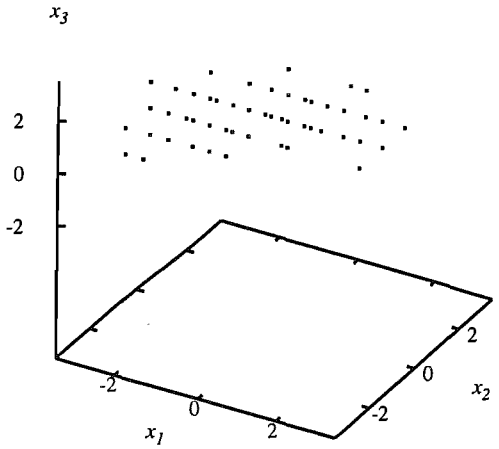
Upon careful reflection it may be observed that all equations given in this section correspond to a different notation of the Bayesian detector, which has already been discussed in Section 2.3.3. It can be seen that the computational cost associated with the calculation of the conditional probabilities is proportional the size of $\mathcal{X}_{-m}$. In the next subsection, we will show how the marginalization, which is in-fact, again, an integration over all possible MIMO symbol sequences relates to classic Monte Carlo integration.

5.3.2 Monte Carlo Integration

Calculating the *a posteriori* probability of a symbol according to Equation (5.19) might become computationally costly, when the number of dimensions of $\mathcal{X}$ determined by the number of transmit AEs M is high, as seen in Section 2.3.3 for the full Bayesian detector. It is therefore desirable to calculate the *a posteriori* probability of a symbol based on a set of hypothetically transmitted symbols $\mathcal{X}_{\mathrm{MC}}$, which has a lower number of elements than the full set $\mathcal{X}$.

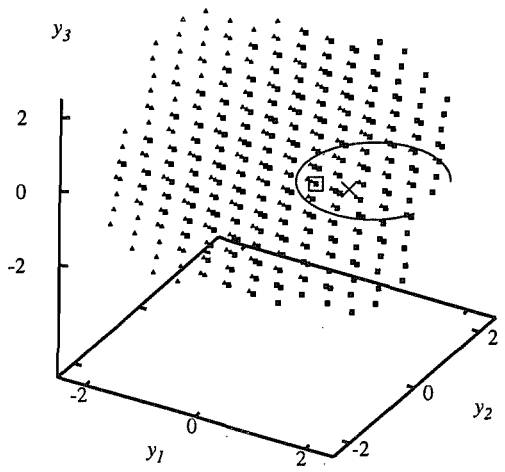


(a) Complete set of hypothetically transmitted symbols $\mathcal{X}$.

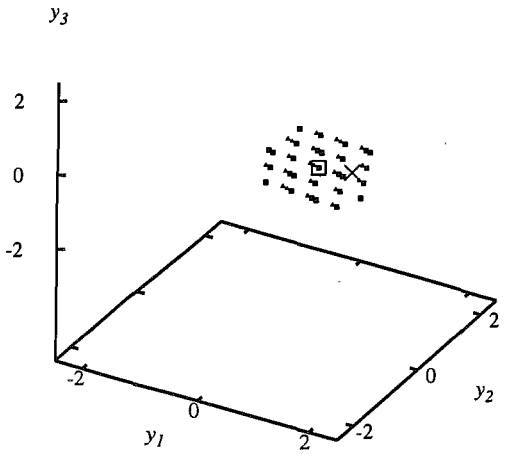


(b) Set of hypothetically transmitted symbols $\mathcal{X}_{\text{MC}}$ used for Monte Carlo integration.

Figure 5.7: Set of hypothetically transmitted MIMO symbols $\boldsymbol{x} = [x_1\ x_3\ x_3]^T$ for a three transmit AE system using 8-PAM.



(a) Complete set of received legitimate symbols.



(b) Set of received legitimate symbols used for Monte Carlo integration.

Figure 5.8: Set of hypothetically received MIMO symbols $\boldsymbol{y} = [y_1\ y_3\ y_3]^T$ three transmit AEs using 8-PAM and $N = 3$ receive AEs. The squares are associated with a transmitted bit $\check{x}_1 = +1$, while the triangles are associated with a transmitted -1. The cross indicates a received noise-contaminated signal vector associated with the corresponding truly received signal vector indicated by the box.

Example 5.1: Monte Carlo Integration

For the sake of illustration, let us consider the hypothetically transmitted symbol vectors of a transmitter employing $M = 3$ AEs and using 8-level Pulse-Amplitude-Modulation (PAM) signals. The set of legitimate transmitted symbol sequences $\mathcal{X}$ is illustrated in Figure 5.7a, which has $2^{3 \cdot 3} = 512$ elements. The full *a posteriori* or Bayesian detector has to evaluate the objective function of Equation (5.19) for all the 512 points seen in Figure 5.7a. By contrast rather than solving the objective function of Equation (5.19) for all these points, the rationale of Monte Carlo integration is to use only a small subset of these points defined as $\mathcal{X}_{\text{MC}}$, which is illustrated in Figure 5.7b and will be calculated using the so-called Gibbs-Sampler to be derived in Section 5.3.3.

Assume now further that the legitimate MIMO symbols are transmitted over a real-valued random channel to a receiver employing $N = 3$ receive AEs, which results in the observation space illustrated in Figure 5.8a. For each of the transmitted bits, the full Bayesian detector would group these points into two subsets corresponding to a transmitted plus one indicated as squares as well as minus one indicated by triangles and calculate the *a posteriori* probability according to Equation (5.19).

By contrast, a sphere decoding aided detector would search the received signal space around a noisy received symbol vector indicated by the cross in Figure 5.8a and attempt to find the channel output states associated with the transmitted symbol, which is marked by the box. It now becomes clear that Bayesian detection considers the distribution of all legitimate channel output states for calculating the a posteriori probability, while in its nature sphere decoding algorithms are designed to find the ML solution associated with a single constellation point. They are therefore particularly suited for generating hard-output. Only by modifying sphere decoding algorithms such as discussed in Section 4.4.3, they can be adapted to return soft-information.

The MCMC aided detector only considers a subset of received channel output states illustrated in Figure 5.8b, which are associated with a subset of legitimate transmitted symbol vectors $\mathcal{X}_{\text{MC}}$. These samples naturally should be distributed around the specific channel output state, which is associated with the transmitted symbol indicated by the box.

To elaborate a little further, Genetic Algorithms (GAs) are in this sense reminiscent of sphere decoding algorithms since they also aim for finding the ML solution and have to be specifically modified in order to produce soft-output [148].

We will now introduce two different methods of determining the *a posteriori* probabilities of Equation (5.18) without considering all the 2^M possible transmitted symbol vectors in $\mathcal{X}$.

5.3.2.1 Empirical Average Calculation

Let us commence by formulating a simple mathematical portrayal of the problem, which will then be related to our detection problem in the next section. A problem often considered in the context of Monte Carlo integration is the evaluation of the weighted mean of a function $g(x)$ of the random variable x, given the weighting function $f(x)$, which is formulated as [69]

$$\mathrm{E}\left[g(X)\right]_{f(x)} = \int_X g(x)f(x)\partial x. \tag{5.23}$$

According to classic Monte Carlo integration [142] an estimate of $\mathrm{E}\left[g(X)\right]_{f(x)}$ may be obtained by gradually constructing the empirical average

$$\mathrm{E}\left[g(X)\right] \approx \frac{1}{N_{\mathrm{MC}}} \sum g(x), \tag{5.24}$$

where N_{MC}, is the number of random samples considered for the integration.

Example 5.2: Monte Carlo Integration

Consider the simple case, where for the weighting function we have $f(x) = 1$, while the function to be integrated is $g(x) = x$ and we would like to calculate the integral of $g(x)$ weighted by $f(x) = 1$ and integrated between zero and one. According to Monte Carlo integration this can be carried out by generating random samples of $g(x) = x$ and averaging them. For this example we would have to generate independent, uniformly distributed samples of $g(x) = x$ between 0 as well as 1 and then average them. In a randomized trial, where we generated 100 samples, we obtained: $\int_X g(x)f(x)\partial x = 0.50466$, which is close to the true expected or average value of 0.5. As anticipated, the more samples are used for approximating the integral, the more accurate the result becomes.

The problem of finding the integral or expected value of a given distribution can now be related to our MIMO detection problem, where we would like to calculate the *a posteriori* probability of the desired symbol x_m associated with transmit AE m. Let us first assume that we are able to generate N_{MC} samples $\check{\boldsymbol{x}}^{(i)}$ from the distribution $P(\boldsymbol{x}|\boldsymbol{y}, \lambda_{\mathrm{A}}^{\mathrm{det}})$, where the samples form the set $\mathcal{X}_{\mathrm{MC}}$. The samples in this set are not necessarily distinct. From the set $\mathcal{X}_{\mathrm{MC}}$ we can now construct the subset $\mathcal{X}_{\mathrm{MC},-m}$, which contains all $N_{\mathrm{MC},-m}$ elements $\check{\boldsymbol{x}}^{(i)}$, where the subscript $_{-m}$ indicates that the m^{th} symbol associated with the m^{th} transmit AE was removed. Furthermore, we may also construct the two subsets $\mathcal{X}_{\mathrm{MC},m}^{+}$ and $\mathcal{X}_{\mathrm{MC},m}^{-}$, which contain all elements of $\mathcal{X}_{\mathrm{MC}}$, where we have $\check{x}_m^{(i)} = +1$ and $\check{x}_m^{(i)} = -1$. More explicitly, these two subsets correspond to the m^{th} transmit antennas bit being $+1$ or -1.

By substituting $P\left(x_m|\check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}\right)$ for $g(x)$ as well as $P\left(\check{\boldsymbol{x}}_{-m}^{(i)}|\boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}\right)$ for $f(x)$ in Equation (5.23) we arrive at

$$P\left(x_m|\boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}\right) = \mathrm{E}\left[P\left(x_m|\check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}\right)\right]_{P\left(\check{\boldsymbol{x}}_{-m}^{(i)}|\boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}\right)} \tag{5.25}$$

$$\approx \frac{1}{N_{\mathrm{MC},-m}} \sum_{\check{\boldsymbol{x}}_{-m}^{(i)} \in \mathcal{X}_{\mathrm{MC},-m}} P\left(x_m|\check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}\right). \tag{5.26}$$

The desired probability $P(x_m = +1|\boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}})$ may then be approximated as follows

$$P(x_m = +1|\boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}) = \sum_{\check{\boldsymbol{x}}^{(i)} \in \mathcal{X}_m^+} P(\check{\boldsymbol{x}}^{(i)}|\boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}})$$

$$\approx \frac{1}{N_{\mathrm{MC},-m}} \sum_{\check{\boldsymbol{x}}_{-m}^{(i)} \in \mathcal{X}_{\mathrm{MC},-m}} P(x_m = +1|\check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}) \tag{5.27}$$

where

$$P(x_m = +1|\boldsymbol{x}_{-m}^{(i)}, \boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}) = \frac{1}{1 + \frac{P(x_m=-1, \check{\boldsymbol{x}}_{-m}^{(i)}|\boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}) p(\boldsymbol{y}|x_m=-1, \check{\boldsymbol{x}}_{-m}^{(i)})}{P(x_m=+1, \check{\boldsymbol{x}}_{-m}^{(i)}|\boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}) p(\boldsymbol{y}|x_m=+1, \check{\boldsymbol{x}}_{-m}^{(i)})}}. \tag{5.28}$$

Equation (5.27) follows directly from Equation (5.26) by substituting $x_m = +1$ for x_m. Depending on the accuracy required for the *a posteriori* probability $P(x_m = +1|\boldsymbol{x}_{-m}^{(i)}, \boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}})$ Equation (5.27) might be further simplified to

$$P(x_m = +1|\boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}) \approx \frac{1}{N_{\mathrm{MC}}} \sum_{\check{\boldsymbol{x}}^{(i)} \in \mathcal{X}_{\mathrm{MC}}} \delta(\check{x}_m = +1), \tag{5.29}$$

where δ represents the delta-Dirac function [110]. The approach of Equation (5.27) averages the *a posteriori* probabilities conditioned on each sample of the set $\mathcal{X}_{\mathrm{MC},-m}$, while the approach of Equation (5.29) evaluates the desired *a posteriori* probability by counting how many symbol vectors in $\mathcal{X}_{\mathrm{MC}}$ are associated with $\check{x}_m^{(i)} = +1$, which is computationally more attractive. A graphical illustration of the two approaches will be presented in Example 5.3. Calculating the *a posteriori* probability according to Equation (5.27) was originally proposed in [58], while the approach of statistical inference represented by Equation (5.29) has been considered for example in [58, 69, 145].

5.3.2.2 Integration Based on Importance Sampling

A different approach to the approximation used in Equation (5.27) and Equation (5.29) for Monte Carlo integration is to use importance sampling [69] in order to calculate the desired integral over a function $g(x)$ given the weighting function $f(x)$ as defined in Equation (5.23). In importance sampling each sample x drawn from the space $\mathcal{X}$ in order to calculate the desired integral is weighted with an importance weight $f_a(x)$. For a discrete space such as the legitimate combinations of the 2^M-ary signals of the M transmit AEs

this yields [58]

$$\mathrm{E}\left[g(X)\right]_{f(x)} \approx \frac{1}{N_{\mathrm{MC}}} \sum \frac{f(x)}{f_a(x)} g(x). \tag{5.30}$$

In the specific case, when $f_a(x)$ is equal to $f(x)$, Monte Carlo integration based on importance sampling and integration based on the empirical averaging method become identical, as it can be seen by comparing Equation (5.24) and Equation (5.30). The calculation of the optimum importance weight function f_a is challenging to calculate and involves itself an integration over the complete decision space [58]. Therefore often a more practical approximation is used [58]

$$\mathrm{E}\left[g(X)\right]_{f(x)} \approx \frac{\sum \frac{f(x)}{f_a(x)} g(x)}{\sum \frac{f(x)}{f_a(x)}}. \tag{5.31}$$

If the weighting function f_a is assumed to obey a uniform distribution, we arrive at [58]

$$\mathrm{E}\left[g(X)\right]_{f(x)} \approx \frac{\sum f(x) g(x)}{\sum f(x)}. \tag{5.32}$$

When applied to our MIMO detection problem, where we want to weight the samples $\check{\boldsymbol{x}}^{(i)} \in \mathcal{X}_{\mathrm{MC}}$, a uniform weighting function $f_a(\check{\boldsymbol{x}}^{(i)})$ implies that the vectors $\check{\boldsymbol{x}}^{(i)}$, which make a significant contribution to the *a posteriori* probability have a discrete uniform distribution over the space $\mathcal{X}$. The consideration of a uniform distribution has been proposed in [58], where it was proposed furthermore that all duplicate[1] entries from the set $\mathcal{X}_{\mathrm{MC}}$ should be removed before evaluating the *a posteriori* information. Removing double entries will ensure that the samples remain approximately uniformly distributed over $\mathcal{X}$. It will become clear in Section 5.3.2.3 that this assumption yields an importance sampling based integration whose formulation is identical to that of the full Bayesian detector discussed in Section 2.3.3, where the set of all hypothetically transmitted 2^M-ary symbol vectors $\mathcal{X}$ is replaced by the reduced-size sampled set $\mathcal{X}_{\mathrm{MC}}$.

When applied to our detection problem, the assumption of an underlying discrete uniform distribution yields

$$P(x_m = +1 | \boldsymbol{y}, \lambda_{\mathrm{A}}^{\mathrm{det}}) \approx \frac{\sum_{\check{\boldsymbol{x}}_{-m}^{(i)} \in \mathcal{X}_{\mathrm{MC},-m}} P(x_m = +1 | \check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{y}, \lambda_{\mathrm{A}}^{\mathrm{det}}) P(\check{\boldsymbol{x}}_{-m}^{(i)} | \boldsymbol{y}, \lambda_{\mathrm{A}}^{\mathrm{det}})}{\sum_{\check{\boldsymbol{x}}_{-m}^{(i)} \in \mathcal{X}_{\mathrm{MC},-m}} P(\check{\boldsymbol{x}}_{-m}^{(i)} | \boldsymbol{y}, \lambda_{\mathrm{A}}^{\mathrm{det}})}, \tag{5.33}$$

where $P(x_m = +1 | \check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{y}, \lambda_{\mathrm{A}}^{\mathrm{det}})$ is *a posteriori* probability conditioned on the sample $\check{\boldsymbol{x}}_{-m}^{(i)}$, i.e. $g(x)$ of Equation (5.32) and $P(\check{\boldsymbol{x}}_{-m}^{(i)} | \boldsymbol{y}, \lambda_{\mathrm{A}}^{\mathrm{det}})$ is the weighting function associated with $\check{\boldsymbol{x}}_{-m}^{(i)}$, i.e. $f(x)$ of Equation (5.32).

[1] So far we have not stated how the set $\mathcal{X}_{\mathrm{MC}}$ is generated. Depending on the particular method chosen for generating $\mathcal{X}_{\mathrm{MC}}$ duplicate entries might occur as it will be shown in Section 5.3.3.

5.3.2.3 Summary

In Section 5.3.2.1 as well as Section 5.3.2.2 we have discussed three different ways of calculating the *a posteriori* probability of an M-antenna transmitted symbol, given a set of samples $\mathcal{X}_{\text{MC}}$ and therefore three different ways of calculating the desired LLR values. These may be summarized as follows.

- Using the empirical average according to Equation (5.29) yields

$$\lambda_{\text{P},m}^{\text{det}} = \log\left(\frac{\sum_{\mathcal{X}_{\text{MC}}} \delta(\check{x}_m^{(i)} = +1)}{\sum_{\mathcal{X}_{\text{MC}}} \delta(\check{x}_m^{(i)} = -1)}\right). \tag{5.34}$$

- Using the empirical average calculation approach described by Equation (5.27) leads to

$$\lambda_{\text{P},m}^{\text{det}} = \log\left(\frac{\frac{1}{N_{\text{MC},-m}} \sum_{\mathcal{X}_{\text{MC},-m}} P(x_m = +1|\boldsymbol{y}, \check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{\lambda}_A^{\text{det}})}{\frac{1}{N_{\text{MC},-m}} \sum_{\mathcal{X}_{\text{MC},-m}} P(x_m = -1|\boldsymbol{y}, \check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{\lambda}_A^{\text{det}})}\right). \tag{5.35}$$

- Uniform sampling based on Equation (5.33) leads to

$$\lambda_{\text{P},m}^{\text{det}} = \log\left(\frac{\sum_{\mathcal{X}_{\text{MC},-m}} P(x_m = +1|\check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{y}, \boldsymbol{\lambda}_{\text{A}}^{\text{det}}) P(\check{\boldsymbol{x}}_{-m}^{(i)}|\boldsymbol{y}, \boldsymbol{\lambda}_{\text{A}}^{\text{det}})}{\sum_{\mathcal{X}_{\text{MC},-m}} P(x_m = -1|\check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{y}, \boldsymbol{\lambda}_{\text{A}}^{\text{det}}) P(\check{\boldsymbol{x}}_{-m}^{(i)}|\boldsymbol{y}, \boldsymbol{\lambda}_{\text{A}}^{\text{det}})}\right). \tag{5.36}$$

When taking into account that we can write[2]

$$P(\check{\boldsymbol{x}}_{-m}^{(i)}|\boldsymbol{y}, \boldsymbol{\lambda}_{\text{A}}^{\text{det}}) = \frac{p(\check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{y}|\boldsymbol{\lambda}_{\text{A}}^{\text{det}})}{p(\boldsymbol{y}|\boldsymbol{\lambda}_{\text{A}}^{\text{det}})} = \frac{P(\check{\boldsymbol{x}}_{-m}^{(i)}|\boldsymbol{\lambda}_{\text{A}}^{\text{det}}) p(\boldsymbol{y}|\check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{\lambda}_{\text{A}}^{\text{det}})}{p(\boldsymbol{y}|\boldsymbol{\lambda}_{\text{A}}^{\text{det}})} \tag{5.37}$$

and relating the Bayes' theorem of Equation (5.16) to the transmission of ± 1, we have

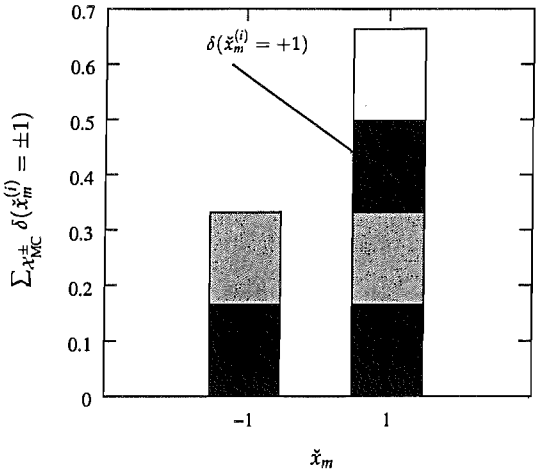
$$P(x_m = \pm 1|\check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{y}, \boldsymbol{\lambda}_{\text{A}}^{\text{det}}) = \frac{p(\boldsymbol{y}|\check{\boldsymbol{x}}_{-m}^{(i)}, x_m = \pm 1) P(x_m = \pm 1|\boldsymbol{\lambda}_{\text{A}}^{\text{det}})}{p(\boldsymbol{y}|\check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{\lambda}_{\text{A}}^{\text{det}})}. \tag{5.38}$$

Finally, Equation (5.36) may be expressed as

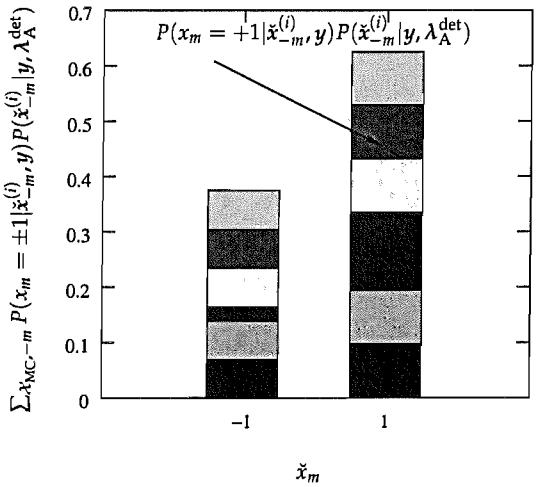
$$\lambda_{\text{P},m}^{\text{det}} = \log\left(\frac{\sum_{\mathcal{X}_{\text{MC},-m}} p(\boldsymbol{y}|x_m = +1, \check{\boldsymbol{x}}_{-m}^{(i)}) P(x_m = +1, \check{\boldsymbol{x}}_{-m}^{(i)}|\boldsymbol{\lambda}_{\text{A}}^{\text{det}})}{\sum_{\mathcal{X}_{\text{MC},-m}} p(\boldsymbol{y}|x_m = -1, \check{\boldsymbol{x}}_{-m}^{(i)}) P(x_m = -1, \check{\boldsymbol{x}}_{-m}^{(i)}|\boldsymbol{\lambda}_{\text{A}}^{\text{det}})}\right). \tag{5.39}$$

This formulation is identical to the full Bayesian detector of Equation (4.16), with the only difference that the M-ary set $\mathcal{X}$ of the Bayesian detector, which contains all possible transmitted symbol sequences of the M transmit AEs, has been replaced by the set $\mathcal{X}_{\text{MC}}$.

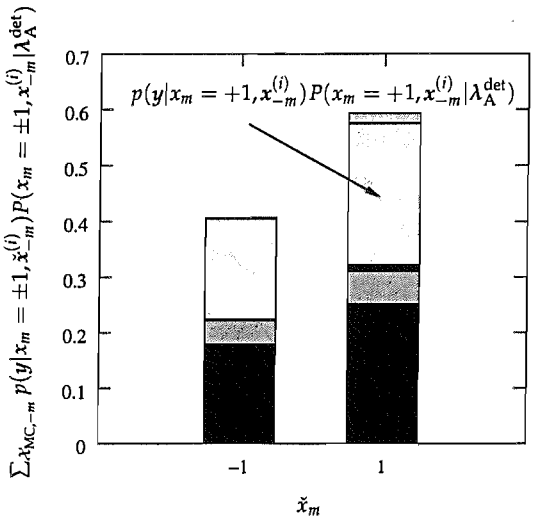
[2]Note that we can write $p(a|b) = p(a, b)/p(b)$.



(a) Empirical average calculation according to Equation (5.34).



(b) Empirical average calculation according to Equation (5.35).



(c) Importance sampling based approach according to Equation (5.36).

Figure 5.9: Illustration of the empirical average and the importance sampling approach to LLR value calculation on the basis of a given set of samples $\mathcal{X}_{MC}$. The graphs illustrated calculated the summation terms on the basis of the same set of samples $\mathcal{X}_{MC}$, which was recoded during a simulation of a (4×4)-antenna narrowband MIMO system employing 16QAM signals and operating at $E_b/N_0 = 6$ dB.

Example 5.3: LLR Calculation Methods

In order to illustrate the characteristics of the three different methods of evaluating the *a posteriori* LLR values, we consider a (4×4)-antenna-aided narrowband MIMO system transmitting 16QAM signals and operating at $E_b/N_0 = 6$ dB. During a randomized run a set $\mathcal{X}_{MC}$ was generated, which was then used for evaluating the sums seen in Equations (5.34)-(5.39). The numerator and the denominator of Equation (5.34) are illustrated in Figure 5.9a for the example considered. The different shades in the stacked bar graph correspond to a specific term of Equation (5.34) or equivalently to a single sample of $\mathcal{X}_{MC}$. It can be observed from Figure 5.9a that each sample of $\mathcal{X}_{MC}$ is given the same weight, i.e all fields in the bar graph have the same area. Using the values of Figure 5.9a, the LLR value of the associated transmitted bit can be calculated using Equation (5.34) as

$$\lambda_{P,m}^{\text{det}} = \log\left(\frac{0.67}{0.33}\right) = 0.71. \tag{5.40}$$

If the weighted averaging method of Equation (5.35) rather than Equation (5.34) is considered, then it can be seen from Figure 5.9b that the samples of $\mathcal{X}_{MC}$ no longer occupy the same area since each of them was assigned a different weight. When calculating the LLR value according to Equation (5.35), while using the values of Figure 5.9b results in

$$\lambda_{P,m}^{\text{det}} = \log\left(\frac{0.62}{0.38}\right) = 0.49. \tag{5.41}$$

The corresponding graph recorded when considering the importance sampling method of Equation (5.39) is illustrated in Figure 5.9c. It can be seen that in contrast to the averaging approaches of Figures 5.9a and 5.9b, the different samples are weighted even stronger according to their importance. Calculating the LLR value according to Equation (5.39) while using the values of Figure 5.9c results in

$$\lambda_{P,m}^{\text{det}} = \log\left(\frac{0.59}{0.41}\right) = 0.36. \tag{5.42}$$

It can be seen from this simple randomized run how differently the three approaches weight the different samples provided, in order to obtain the resultant LLR values. The exact effect of the different weighting techniques on the achievable performance will be studied in Section 5.3.4.

Given a set $\mathcal{X}_{MC}$ and Equations (5.34)-(5.39), we are now ready to obtain the LLR values. In the next section, it will be shown how the desired set $\mathcal{X}_{MC}$ can be generated.

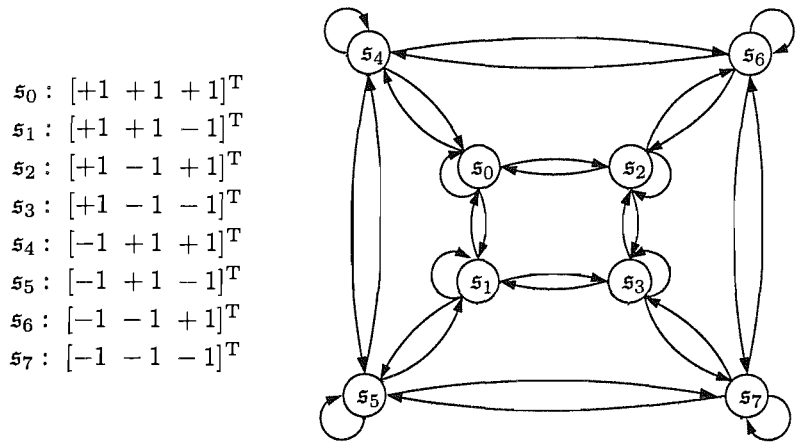


Figure 5.10: Markov chain describing a detection process, when a single bit can be changed.

5.3.3 Markov Chain Representation and Gibbs-Sampler

In this section we will show how the detection of a 2^M-ary MIMO signal can be represented with the aid of a Markov chain and how this property can be exploited by the receiver in order to generate a set of samples $\mathcal{X}_{\mathrm{MC},-\mathrm{m}}$ defined in the context of Equation (5.19), which asymptotically represents the distribution $P\left(x_m|\check{\boldsymbol{x}}_{-m}^{(i)}, \boldsymbol{y}, \lambda_{\mathrm{A}}^{\mathrm{det}}\right)$.

It is important to note that the Markov chain is not used for representing the evolution of the signal vector $\boldsymbol{x}$ over time. In other words, the Markov chain is not employed for the sake of modeling the relation between a symbol vector $\boldsymbol{x}(k)$ transmitted at time k and the symbol vector $\boldsymbol{x}(k+1)$ transmitted at time $k+1$. It is rather used for generating the legitimate hypothetical solutions $\check{\boldsymbol{x}}^{(i)}$ for a specific transmitted symbol vector $\boldsymbol{x}(k)$ at a specific sampling instant k. Generating the legitimate hypothetically transmitted samples with the aid of a Markov chain must not therefore be interpreted as an evolution in time, but rather as efficiently exploring the search space of interest $\mathcal{X}$ and generating a reduced-size subspace $\mathcal{X}_{\mathrm{MC}}$ approximating the full-size search-space appropriately.

Let us now first consider a simple $(N \times M) = (3 \times 3)$-antenna aided narrowband MIMO system and BPSK modulated signals. For this system the legitimate hypothetically transmitted symbol vector $\check{\boldsymbol{x}}$ may assume $N_{\mathrm{s}} = 2^3 = 8$ different values, which represent the 8 states of our Markov chain. Furthermore, in order to keep the algorithm simple we assume that only one symbol $\check{x}_m$ of a hypothetical transmitted symbol vector $\check{\boldsymbol{x}}$ can be changed at a time.[3] Then the Markov chain describing our detection problem is portrayed in Figure 5.10, where 4 different transitions are emerging from each state.

The Markov chain shown in Figure 5.10 has the following properties:

- Since the transition probabilities between different states are jointly determined by the channel matrix $\boldsymbol{H}$ of the system as well as the received signal vector and since we assume the channel matrix to be

[3] In this treatise we assume that only one bit is changed at a time, which imposes the lowest computational cost.

time-invariant over a least one symbol interval, the state transition probabilities do not change during a single symbol interval either.

- There is a non-zero probability that at some time in the future the Markov chain will be in state j, given that the Markov chain originally emerged from state i. In other words, it is possible to reach any state of the Markov chain from any state.
- Aperiodicity: A Markov chain is referred to as aperiodic, if it is possible to return to state i in any number of steps. If it would be possible to return to a state i only in multiples of k, then the state would be referred to as being periodic with a period of k.

Since the Markov chain representing the MIMO detection problem satisfies the above three properties, it converges to a stationary distribution [145]. The word 'stationary' in this context implies that when generating samples using the above Markov chain model, the resultant hypothetical solutions to the detection problem, which are represented by the states of the Markov chain are distributed according to a stationary distribution. This property will be made more explicit in the context of Example 5.4 of this section.

Having discussed the structure of the Markov chain, we now have to quantify the transition probabilities for the sake of its complete characterization. Let us therefore denote the probability of encountering state i of the Markov chain corresponding to a legitimate transmit symbol vector as

$$p_i = P(\check{\boldsymbol{x}}^{(i)}|\boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}), \tag{5.43}$$

which is conditioned on both the channel output $\boldsymbol{y}$ as well as on the *a priori* LLR $\boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}$. Similarly, the Markov chain transition probability between state i and j or, equivalently, between the corresponding MIMO symbols can then be expressed as

$$\pi_{ij} \propto P(\check{\boldsymbol{x}}^{(j)}|\boldsymbol{y}, \boldsymbol{\lambda}_{\mathrm{A}}^{\mathrm{det}}), \tag{5.44}$$

where the proportionality factor can be obtained by using the fact that $\sum_j \pi_{ij} = 1$.

By comparing Equation (5.44) and Equation (5.43), it may be readily seen that we can write

$$p_i \pi_{ij} = p_j \pi_{ji}. \tag{5.45}$$

A Markov chain obeying this property may be referred to as a reversible Markov chain [149]. It was shown for example in [149] that the reversible distribution p_i is also the stationary distribution of state i. In the following example, we will show why it is important that the Markov chain converges to its stationary distribution. Again, evolving one state of the Markov chain to another implies exploring the legitimate transmit symbol combinations of the search-space rather than an evolution in time from the k^{th} to

the $(k+1)^{th}$ MIMO symbol.

Example 5.4: Stationary Distribution of a Markov Chain

Consider a real-valued $(N \times M) = (3 \times 3)$ MIMO system operating at a noise level characterized by $2\sigma_n^2 = 1$ and having a system equation given as

$$\begin{matrix} \boldsymbol{y} & = & \boldsymbol{H} & & \boldsymbol{x} & + & \boldsymbol{\eta} \end{matrix}$$
$$\begin{bmatrix} -0.97 \\ +0.18 \\ -0.69 \end{bmatrix} = \begin{bmatrix} -0.43 & -0.65 & +0.48 \\ +0.22 & -0.69 & -0.61 \\ -0.39 & -0.61 & +0.87 \end{bmatrix} \begin{bmatrix} +1 \\ -1 \\ -1 \end{bmatrix} + \begin{bmatrix} +0.39 \\ -0.25 \\ +0.10 \end{bmatrix}. \quad (5.46)$$

Since the M-bit MIMO symbols are assumed to be independent, the probability of each Markov chain state for the observed system can be directly calculated using Equations (5.19)-(5.23), by considering all eight legitimate hypothetically transmitted M-bit symbol sequences $\check{\boldsymbol{x}}^{(i)}$ as well as the parameters provided in Equation (5.46). Let us assume that this yields the eight states' stationary distributions of

p_0	p_1	p_2	p_3	p_4	p_5	p_6	p_7
≈ 0.0	≈ 0.0	≈ 0.0	0.69	≈ 0.0	≈ 0.0	≈ 0.0	0.31

. (5.47)

It is worth noting that all states have a non-zero probability, but some of them may be potentially infinitesimally low at the current noise variance of $2\sigma_n^2 = 1$ and the observed system of Equation (5.46). If we now calculate the Markov chains transition matrix for the same example according to Equation (5.44) for $2\sigma_n^2 = 1$, we arrive at

$$\Pi = \begin{bmatrix} 0.50 & 0.06 & 0.44 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.77 & 0.08 & 0.0 & 0.0 & 0.0 & 0.14 & 0.0 & 0.0 \\ 0.01 & 0.0 & 0.0 & 0.99 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.69 & 0.0 & 0.0 & 0.0 & 0.31 \\ 0.01 & 0.0 & 0.0 & 0.99 & 0.0 & 0.01 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.01 & 0.0 & 0.99 \\ 0.0 & 0.0 & 0.01 & 0.0 & 0.0 & 0.0 & 0.0 & 0.99 \\ 0.0 & 0.0 & 0.0 & 0.69 & 0.0 & 0.0 & 0.0 & 0.31 \end{bmatrix}. \quad (5.48)$$

For a reversible Markov Chain, the resultant stationary distribution may also be obtained as [149]

$$\mathbf{1}_{N_s} \pi = lim_{z \to \infty} \Pi^z, \quad (5.49)$$

where $\mathbf{1}_{N_s}$ is a column vector containing N_s number of unity values and π is a row vector hosting the stationary distributions. When calculating Π^{10} for $z = 10$, i.e after the 10^{th} Markov

chain state transition obeying the matrix of Equation (5.48), we arrive at

$$\Pi^{10} = \begin{bmatrix} 0.0 & 0.0 & 0.0 & 0.69 & 0.0 & 0.0 & 0.0 & 0.31 \\ 0.0 & 0.0 & 0.0 & 0.69 & 0.0 & 0.0 & 0.0 & 0.31 \\ 0.0 & 0.0 & 0.0 & 0.69 & 0.0 & 0.0 & 0.0 & 0.31 \\ 0.0 & 0.0 & 0.0 & 0.69 & 0.0 & 0.0 & 0.0 & 0.31 \\ 0.0 & 0.0 & 0.0 & 0.69 & 0.0 & 0.0 & 0.0 & 0.31 \\ 0.0 & 0.0 & 0.0 & 0.69 & 0.0 & 0.0 & 0.0 & 0.31 \\ 0.0 & 0.0 & 0.0 & 0.69 & 0.0 & 0.0 & 0.0 & 0.31 \\ 0.0 & 0.0 & 0.0 & 0.69 & 0.0 & 0.0 & 0.0 & 0.31 \end{bmatrix}, \qquad (5.50)$$

where we can see that the rows of Π^{10} approach the stationary distribution π. For this example the Markov chain's state transitions converges rapidly to the true stationary distribution and the corresponding Markov state-machine has already forgotten its initial state after $z = 10$ Markov chain transitions carried out according to the state-transition matrix of Equation (5.48).

Recall that our final goal is that of generating a reduced set of potentially transmitted MIMO symbols, which are considered to be samples from a high-dimensional MIMO decision space obeying the distribution $P(\check{x}_m|y, \lambda_A^{\text{det}})$ in order to enable us to perform Monte Carlo integration. We aim for generating these reduced number of samples representing the potentially transmitted MIMO symbols based on a stationary Markov chain. If we generate our samples or MIMO symbols according to the stationary Markov chain, the samples obtained after a sufficiently high number of Markov chain transitions will ultimately be those of the stationary distribution, regardless of the initial value of $\check{x}_{\text{init}}$. The most popular algorithm, which is based on generating a reduced set of samples according to stationary Markov Chain is the so-called Gibbs-Sampler [69], which may be viewed as a simplified version of the Metropolis-Hastings algorithm [69] and has become popular owing to its implementational simplicity [150]. The algorithm can be summarized with the aid of its pseudo code as follows [69].

Algorithm 5.1: Gibbs-Sampler [69]

```
Initialize $\check{x}^{-N_{\text{burn}}}$ randomly
for $i = -N_{\text{burn}}$ to $N_{\text{MC}}$ do
    draw sample from $P(x_1|\check{x}_{-1}^{(i-1)}, y, \lambda_A^{\text{det}})$;
    draw sample from $P(x_2|\check{x}_{-2}^{(i)}, y, \lambda_A^{\text{det}})$;
        ⋮
    draw sample from $P(x_M|\check{x}_{-M}^{(i)}, y, \lambda_A^{\text{det}})$;
    if $i > 0$ add sample $\check{x}^{(i)}$ to $\mathcal{X}_{\text{MC}}$;
    $i++$
end ,
```

where N_{burn} is the so-called burn-in period and N_{MC} is the number of Markov chain samples generated,

which in our context correspond to the reduced subset of MIMO symbols to be considered by the detector. Note that in order to generate a single sample $\check{\boldsymbol{x}}^{(i)}$, M Markov chain transitions are required each of which is associated with a single bit change and finally yields an M-ary MIMO symbol. The employment of a burn-in period is generally required, since the first samples of the Markov chain will depend on the initial state of the corresponding state-machine of Figure 5.10 and therefore they will be less characteristic of the desired stationary distribution. In the following example, we will show with the aid of a simple example how the Gibbs-Sampler may be employed for MIMO detection.

Example 5.5: Gibbs-Sampler

Let us consider the same system as in Example 5.4, where - again - it was assumed that no *a priori* information is available at the detector and therefore all bits are equiprobable. Using the Gibbs-Sampler of Algorithm 5.1 we are able to generate the desired samples as follows:

1. Initialize $\check{\boldsymbol{x}}$ randomly: $\check{\boldsymbol{x}}_{\text{init}} = [-1\ -1\ +1]^{\text{T}}$ according to line 1 of Algorithm 5.1;
2. Generate new sample according to the loop in Algorithm 5.1
   - Draw x_1 from $P(x_1|\check{\boldsymbol{x}}_{-1}, \boldsymbol{y}, \lambda_{\text{A}}^{\text{det}}) = P(x_1|[\check{x}_2\ \check{x}_3] = [-1\ +1], \boldsymbol{y}, \lambda_{\text{A}}^{\text{det}})$:

$$p(\boldsymbol{y}|\check{\boldsymbol{x}} = [+1\ -1\ +1]^{\text{T}}) \propto \exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{H}[+1\ -1\ +1]^{\text{T}}\|^2}{2\sigma_{\text{n}}^2}\right) = 2.35 \cdot 10^{-3}$$
$$p(\boldsymbol{y}|\check{\boldsymbol{x}} = [-1\ -1\ +1]^{\text{T}}) \propto \exp\left(-\frac{\|\boldsymbol{y}-\boldsymbol{H}[-1\ -1\ +1]^{\text{T}}\|^2}{2\sigma_{\text{n}}^2}\right) = 0.07 \cdot 10^{-7}$$
$$P(x_1 = +1|[x_2\ x_3] = [-1\ +1], \boldsymbol{y}, \lambda_{\text{A}}^{\text{det}}) = \frac{1}{1+\frac{p(\boldsymbol{y}|\boldsymbol{x}=[-1\ -1\ +1]^{\text{T}})}{p(\boldsymbol{y}|\boldsymbol{x}=[+1\ -1\ +1]^{\text{T}})}} = 0.99;$$

   Generate a random variable ξ, which is uniformly distributed within $[0, 1]$: $\xi = 0.78$.
   Since $\xi = 0.78 < P(x_1 = +1|[\check{x}_2\ \check{x}_3] = [-1\ +1]) = 0.99$, set $\check{x}_1 = +1$.
   - Draw x_2 from $P(x_2|\check{\boldsymbol{x}}_{-2}, \boldsymbol{y}, \lambda_{\text{A}}^{\text{det}}) = P(x_2|[\check{x}_1\ \check{x}_3] = [+1\ +1], \boldsymbol{y}, \lambda_{\text{A}}^{\text{det}})$:

$$p(\boldsymbol{y}|\check{\boldsymbol{x}} = [+1\ +1\ +1]^{\text{T}}) = 2.36 \cdot 10^{-3}$$
$$p(\boldsymbol{y}|\check{\boldsymbol{x}} = [+1\ -1\ +1]^{\text{T}}) = 2.02 \cdot 10^{-3}$$
$$P(x_2 = +1|[\check{x}_1\ \check{x}_3] = [+1\ +1], \boldsymbol{y}, \lambda_{\text{A}}^{\text{det}}) = 0.54$$

   Generate a random variable ξ, which is uniformly distributed within $[0, 1]$: $\xi = 0.11$.
   Since $\xi = 0.11 < P(x_2 = +1|[\check{x}_1\ \check{x}_3] = [+1\ +1], \boldsymbol{y}, \lambda_{\text{A}}^{\text{det}}) = 0.54$, set $\check{x}_2 = +1$.
   - Draw Sample from $P(x_3|\check{\boldsymbol{x}}_{-3}, \boldsymbol{y}, \lambda_{\text{A}}^{\text{det}}) = P(x_3|[\check{x}_1\ \check{x}_2] = [+1+1], \boldsymbol{y}, \lambda_{\text{A}}^{\text{det}})$

$$p(\boldsymbol{y}|\check{\boldsymbol{x}} = [+1\ +1\ +1]^{\text{T}}) = 2.35 \cdot 10^{-3}$$
$$p(\boldsymbol{y}|\check{\boldsymbol{x}} = [+1\ +1\ -1]^{\text{T}}) = 2.58 \cdot 10^{-4}$$
$$P(x_3 = +1|[\check{x}_1\ \check{x}_2] = [+1+1], \boldsymbol{y}, \lambda_{\text{A}}^{\text{det}}) = 0.90$$

Generate a random variable ξ, which is uniformly distributed within $[0,1]$: $\xi = 0.30$.

Since $\xi = 0.30 < P(x_3 = +1|[\check{x}_1\ \check{x}_2] = [+1\ +1], \boldsymbol{y}, \lambda_A^{\text{det}}) = 0.90$, set $\check{x}_3 = +1$.

3. Add sample $\check{\boldsymbol{x}} = [+1\ +1\ +1]^T$ to $\mathcal{X}_{\text{MC}}$
4. Goto 2 and generate new sample

Running the Gibbs-Sampler of Algorithm 5.1 through several state-transitions of the Markov chain of Figure 5.10 the following sequence of states is obtained

$$
\begin{array}{ccccccccccccccccc}
s_6 & \rightarrow & s_2 & & s_0 & & s_0 & & s_3 & & s_3 & & s_7 & & s_3 & & s_3 \\
 & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
 & & s_0 & \Rightarrow & s_0 & \Rightarrow & s_2 & \Rightarrow & s_3 & \Rightarrow & s_3 & \Rightarrow & s_7 & \Rightarrow & s_3 & \Rightarrow & s_3 & \cdots \\
 & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
 & & \boxed{s_0} & & \boxed{s_0} & & \boxed{s_3} & & \boxed{s_3} & & \boxed{s_3} & & \boxed{s_7} & & \boxed{s_3} & & \boxed{s_3}
\end{array}
$$

$$
\begin{array}{ccccccccc}
-1 & +1 & +1 & +1 & +1 & +1 & -1 & +1 & +1 \\
-1 & +1 & +1 & -1 & -1 & -1 & -1 & -1 & -1 \\
+1 & +1 & +1 & -1 & -1 & -1 & -1 & -1 & -1
\end{array}
$$

where each column corresponds to a single run through the loop of Algorithm 5.1 and each arrow $\downarrow$ to a single state-transition of the Gibbs-Sampler. A new iteration of the Gibbs sampler is indicated by the horizontal arrows. The first state s_6, indicates the initial state as used above, while the states in the last row within the boxes are the states of the Markov chain at the end of each iteration of the Gibbs-Sampler. Furthermore, the 3-bit symbol sequences given below the state transitions correspond to the boxed states of the Markov chain of Figure 5.10.

Starting from the initial state s_6 of Figure 5.10, the Gibbs-Sampler of Algorithm 5.1 changes the first symbol to $+1$ and therefore moves to state s_2. The second bit is also changed, resulting in state s_0, while the third remains unchanged during the first iteration of the Gibbs-Sampler. Therefore we obtain the first sample, which is $\check{\boldsymbol{x}}^{(1)} = [+1\ +1\ +1]$ associated with state s_0. During the second iteration, the Gibbs-Sampler remains in state s_0, while during the third iteration, only the first bit remains unchanged. The Markov chain then traverses through the state s_2 to state s_3, which is the state of the chain at the end of the third iteration of the Gibbs-Sampler. From there onwards, it can be observed that the Gibbs-Sampler only moves between state s_3 and s_7. These are exactly the two states associated with the highest probability, as already shown in Example 5.4. Based on the samples obtained with the aid of the Gibbs-Sampler of Algorithm 5.1, the *a posteriori* LLRs might be calculated choosing any of the three

methods summarized in Equations (5.34)-(5.34) of Section 5.3.2.3.

A common problem when using a Gibbs-Sampler, which is generally observed at high SNR, is that it might converge prematurely to a state having a high probability $P(\underline{s}_i)$, where it might get trapped without exploring the entire solution space of interest. In order to avoid this problem, which is often also referred to as'stalling', different methods have been proposed [69, 147]. The most powerful ones were shown to be the so-called parallel Gibbs-Samplers [69]. When using parallel Gibbs-Samplers [69], all samples obtained from the different parallel samplers are used for evaluating the *a posteriori* probability. Other methods which might be considered are for example simulated annealing [69], which can be implemented by considering an increased noise variance at the beginning of the sampling process in order to render the transition probabilities to distant solutions more similar to those of the more close solutions. The noise variance is then gradually reduced to the true noise variance, which might be viewed 'as cooling' down the system employing simulated annealing.

In the above examples and derivations only BPSK modulated signals were considered. The extension to complex-valued signals might readily be achieved in two different ways:

- If the mapping of the bits to the modulated symbols can be represented as a linear operation, such as for example used in the construction of 4QAM or hierarchical 16QAM, the real-valued binary system model introduced in Section 2.1.1.2 may be employed.
- If the above-mentioned mapping cannot be represented as a linear operation, the extension to higher-order modulation schemes can be achieved in analogy to the full Bayesian detector discussed in Section 2.3.3.

5.3.4 Simulation Results for Narrowband MIMO Channels

First we compare the three different methods of calculating the *a posteriori* LLRs at the output of the detector, namely the empirical averaging method of Equation (5.34), the averaging method of Equation (5.35) as well as the uniform sampling approach of Equation (5.36), as detailed in Section 5.3.2.3.

The system's schematic was depicted in Figure 4.1. The mutual information provided at the detector's output will be plotted as a function of the number of samples N_{MC} for different burn-in periods. For the results shown in Figure 5.11 an (8×8) MIMO system using 4QAM signals and operating at $E_b/N_0 = 4$ dB was considered. The narrowband channel was assumed to obey independent, uncorrelated Rayleigh fading and was perfectly known by the receiver.

The mutual information at the input of the detector was $I_A^{dec} = 0$. When considering the empirical averaging method of Equation (5.34), it can be seen that there is no further improvement, when increasing the initial

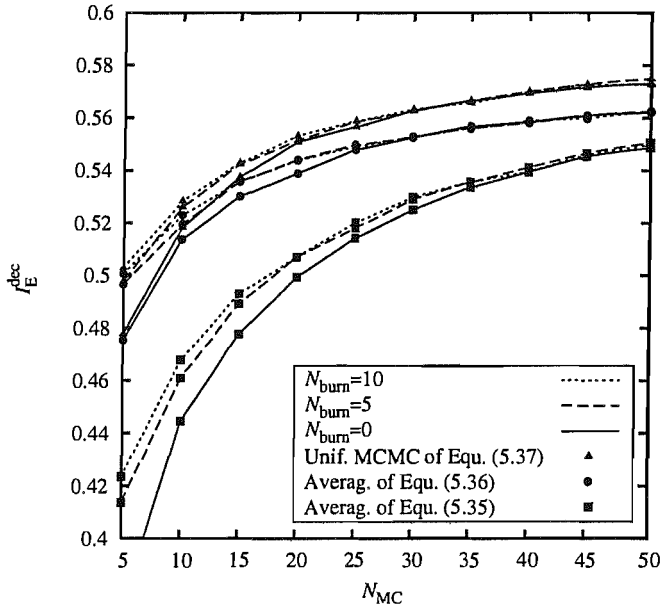


Figure 5.11: Mutual information provided at the detector's output for $I_A^{dec} = 0$ as a function of the number of samples N_{MC} for different burn-in periods and for the three different approaches discussed in Section 5.3.2.3 for calculating the *a posteriori* LLRs. The system considered was a (8×8)-antenna MIMO system using 4QAM modulation operating at $E_b/N_0 = 4$ dB and the channel was assumed to be independently, uncorrelated Rayleigh fading. A single Markov chain was used for drawing the samples from the target distribution.

burn-in period from 5 to 10 samples. The same can be observed for the uniform sampling method of Equation (5.36) as well as for the averaging method of Equation (5.35). When comparing the three different methods, it can be seen that uniform sampling results in the highest mutual information at the output of the MCMC detector, but all three techniques appear to reach a steady state after 50 state-transitions of the Gibbs-Sampler of Algorithm 5.1.

In the context of the same system, we performed simulations also for the MMSE SISO detector, which achieved a mutual information of $I_E^{dec} = 0.54$ at the output of the detector. Similarly, the RML detector of Section 4.4.3 exhibits $I_E^{dec} = 0.57$. Note that the mutual information of the MCMC based detector was recorded, when considering only 50 samples or legitimate symbols carefully selected from the total of $N_s = 4^8 = 65 \cdot 536$ samples in the ML search-space. The MCMC detector performs as well as the RML detector and outperforms the MMSE based detector by exhibiting a mutual information of $I_E^{dec} = 0.57$, when the uniform sampling approach of Equation (5.39) is considered.

Let us now consider a (4×4) narrowband MIMO system and 16-QAM signals as well as independent uncorrelated fading channels. The detector employed was a MCMC aided scheme, which evaluated the *a posteriori* LLRs based on the uniform sampling approach of Equation (5.39). The burn-in period was set to zero, while the number of parallel Gibbs-Samplers employed was 10. Each Gibbs-Sampler was initialized randomly and each generated 10 samples. Thus a total of 100 hypothetically transmitted MIMO symbols were used for evaluating the *a posteriori* LLR. This is only a small fraction of the ML search-space hosting

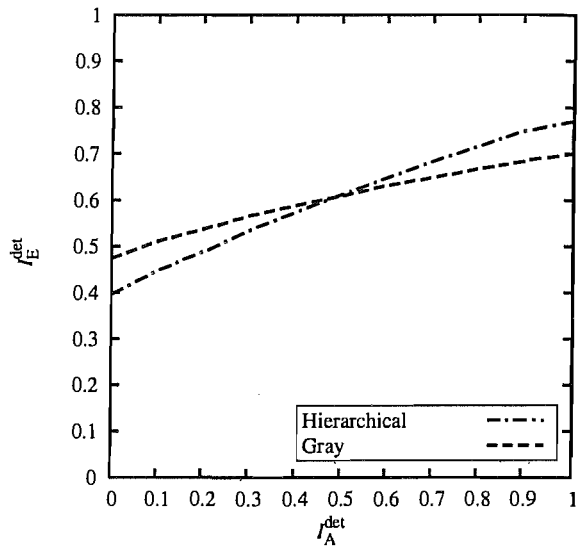


Figure 5.12: EXIT chart of a (4×4) MIMO system using a MCMC detector based on uniform sampling with 10 parallel Markov chains each running 10 iterations. The results were obtained for hierarchical and Gray mapped 16-QAM modulated signals at $E_b/N_0 = 6$ dB.

a total of $N_s = 16^4 = 65\,536$ legitimate hypothetically transmitted MIMO symbols.

In Figure 5.12 the resultant EXIT characteristic is shown for both hierarchical 16 QAM of Figure 2.2b as well as for Gray mapping based 16QAM [115, 151] at $E_b/N_0 = 6$ dB. It can be observed in Figure 5.12 that the area under both curves is identical and therefore both mapping schemes support the same throughput. However, the EXIT function associated with Gray mapping is significantly more flat and thus it is potentially expected to attain a more substantial BER improvement for a given number of iterations between the detector and a channel code as illustrated in Figure 4.1. The EXIT curve of the hierarchically mapped 16-QAM system is identical to that using the Max-Log detector of Figure 4.18.

In Figure 5.13 the achievable BER versus E_b/N_0 performance is illustrated for the same system using Gray mapping as well as a SCC Code using a rate-1 convolutional component code characterized by the octally represented generator $\boldsymbol{g} = [3\ 1]$ and a RSC code with a generator polynomial of $\boldsymbol{g} = [6\ 5]$. The SCC code used $N_{\text{inner}} = 10$ inner iterations. Additionally, the achievable capacity bound is also provided in Figure 5.12. It can be observed that after $N_{\text{iter}} = 6$ iterations between the SCC Code and the MCMC aided detector, the attainable performance approaches the channel capacity within less than 2 dB.

In conclusion of this subsection, we have shown that a MCMC aided detector is capable of efficiently detecting MIMO signals by searching the ML search-space to a small fraction of it. In the next section its applicability to SC STE will be investigated and its detailed complexity analysis is provided.

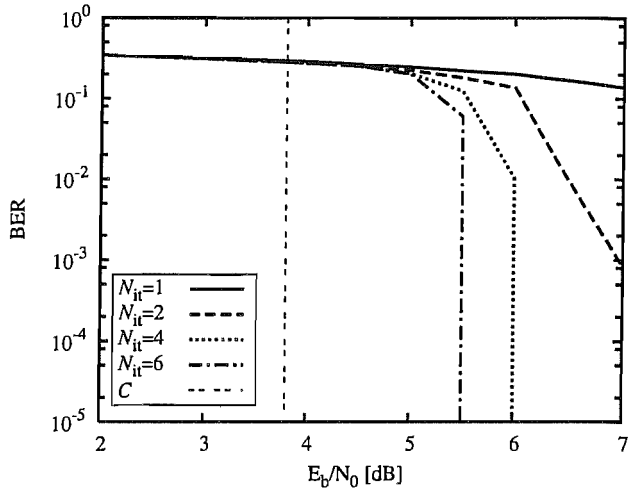


Figure 5.13: BER versus E_b/N_0 performance of a (4×4) MIMO system for a different number of iterations N_{it} using a MCMC detector based on importance sampling with 10 parallel Markov chains each running generating 10 samples. The channel code employed was a serial concatenation of a rate-1 code with polynomial $\boldsymbol{g} = [3\ 1]$ and a RSC code with a generator polynomial in octal form $\boldsymbol{g} = [6\ 5]$. The serially concatenated convolutional code used 10 inner iterations.

5.3.5 Computational Cost

Similar to the RML algorithm of Section 5.2.3, where we have characterized the computational cost imposed in terms of the real-valued operations required for the evaluation of the search tree, here we will characterize the complexity imposed by the MCMC-aided detector quantified by the number of operations required for the set of candidates $\mathcal{X}_{\text{MC}}$ picked by the MCMC-aided detector from the entire search-space $\mathcal{X}$. For this evaluation we directly consider the multi-user STE system model of Equation (5.3), rather than the narrowband MIMO model of Equation (5.15) used for the derivation of the MCMC detector. Again, the extension from narrowband to wideband detection maybe achieved by appropriately replacing the system parameters of Table 5.1, which were applied for the MMSE, the Bayesian as well as the RML detectors in Section 5.2.

Gibbs-Sampler Initialization: When initializing a single chain of the parallel Gibbs-Sampler, we have to calculate

$$\mathbf{y} = \mathbf{H}\mathbf{x} = x_1\mathbf{h}_1 + \cdots + x_{\mathfrak{M}(N_F+L-1)}\mathbf{h}_{\mathfrak{M}(N_F+L-1)}. \tag{5.51}$$

During our complexity evaluation to be discussed in the forthcoming paragraphs only the operations involving non-zero elements are considered. In total, the super-matrix $\mathbf{H}$ of Equation (5.3) hosts $N_F NL\mathfrak{M}$ complex-valued elements.

The initialization of the Gibbs-Sampler commences with the calculation of the channel output $\check{\mathbf{y}}_{\text{init}}$ associated with the first randomly generated hypothetically transmitted signal vector $\check{\mathbf{x}}_{\text{init}}$, which requires a total of $N_F NL\mathfrak{M}$ complex-valued multiplications as well as $NN_F(\mathfrak{M}L-1)$ complex-valued additions. Additionally, the squared Euclidean distance of the initial hypothetical channel output state $\check{\mathbf{y}}_{\text{init}}$ with respect to the

received channel output state $\mathbf{y}$ additionally imposes a cost of $(2N_{\mathrm{F}}N-1)$ complex-valued additions and $N_{\mathrm{F}}N$ complex-valued multiplications. The resultant computational cost associated with the initialization of the Gibbs-Sampler is therefore given as:

- Multiplications: $N_{\mathrm{F}}NL\mathfrak{M}+N_{\mathrm{F}}N \approx NN_{\mathrm{F}}L\mathfrak{M}$;
- Additions: $NN_{\mathrm{F}}(L\mathfrak{M}-1)+2N_{\mathrm{F}}N-1 \approx NN_{\mathrm{F}}L\mathfrak{M}$.

This corresponds to approximately $9NN_{\mathrm{F}}L\mathfrak{M}$ number of real-valued operations.

A Single Gibbs-Sampler Iteration: When employing the Gibbs-Sampler of Algorithm 5.2, a single bit of a given hypothetically transmitted MIMO symbol vector $\check{\mathbf{x}}_{\mathrm{a}}$ is changed in order to obtain a new vector $\check{\mathbf{x}}_{\mathrm{b}}$. Let us assume that the i^{th} bit, which is mapped to the j^{th} symbol $\check{x}_{\mathrm{a},j}$ is considered for change and results in a new hypothetical MIMO symbol vector $\check{\mathbf{x}}_{\mathrm{b}}$. The operation of calculating the hypothetical channel output state $\check{\mathbf{y}}_{\mathrm{b}}$ can then be expressed as

$$\begin{aligned}\check{\mathbf{y}}_{\mathrm{b}}=\mathbf{H}\check{\mathbf{x}}_{\mathrm{b}} &= \sum_{\mathrm{m}=1}^{\mathfrak{M}(N_{\mathrm{F}}+L-1)}\mathbf{h}_{\mathrm{m}}\check{x}_{\mathrm{b},\mathrm{m}}\\ &= \mathbf{h}_j x_{\mathrm{b},j}+\sum_{\mathrm{m}=1,\mathrm{m}\neq j}^{\mathfrak{M}(N_{\mathrm{F}}+L-1)}\mathbf{h}_{\mathrm{m}}x_{\mathrm{b},\mathrm{m}}\\ &= \check{\mathbf{y}}_{\mathrm{a}}-\underbrace{\mathbf{h}_j x_{\mathrm{a},j}}_{\text{known}}+\mathbf{h}_j x_{\mathrm{b},j},\end{aligned} \tag{5.52}$$

where $\check{\mathbf{y}}_{\mathrm{a}}$ is the channel output state associated with $\check{\mathbf{x}}_{\mathrm{a}}$. The third line follows from the fact that only the j^{th} symbol of $\check{\mathbf{x}}_{\mathrm{a}}$ is modified. When considering the fact that $\mathbf{h}_j$ has a maximum of NL non-zero elements, the complexity associated with updating the hypothetical channel output states is given by NL complex-valued multiplications as well $3NL$ complex-valued additions. Let us now define the set, which contains all the indices of the non-zero elements of $\mathbf{h}_j$ as $\mathcal{I}$, which has a maximum cardinality of NL and furthermore, $\check{\mathbf{y}}_{\mathrm{a},j}=\mathbf{h}_j\check{x}_{\mathrm{a},j}=[y_{\mathrm{a},1},\ldots,y_{\mathrm{a},N_{\mathrm{F}}N}]^T$. The squared Euclidean distance calculation can then be expressed as

$$\begin{aligned}||\mathbf{y}-\mathbf{H}\check{\mathbf{x}}_{\mathrm{b}}||^2 &= \sum_{i=1}^{NN_{\mathrm{F}}}|y_i-\check{y}_{\mathrm{b},i}|^2\\ &= \underbrace{\sum_{i=1}^{NN_{\mathrm{F}}}|y_i-\check{y}_{\mathrm{a},i}|^2}_{\text{known}}-\sum_{i\in\mathcal{I}}\underbrace{|y_i-\check{y}_{\mathrm{a},i}|^2}_{\text{known}}+\sum_{i\in\mathcal{I}}|y_i-\check{y}_{\mathrm{b},i}|^2,\end{aligned} \tag{5.53}$$

which imposes a computational cost of $3NL+1$ complex-valued additions as well as NL complex-valued multiplications.

Given the squared Euclidean distances, the probability required for generating a new sample in Algo-

rithm 5.1 is given as

$$\exp\left(-\texttt{jaclog}\left(0, \frac{||\mathbf{y}-\mathbf{H}\check{\mathbf{x}}_a||^2 - ||\mathbf{y}-\mathbf{H}\check{\mathbf{x}}_b||^2}{2\sigma_n^2}\right)\right), \tag{5.54}$$

which imposes a further three real-valued operations plus one Jacobian logarithmic [118] or jaclog[4] as well one exp operation. This cost will be neglected in our further discussion since it is negligible compared to the other operations. Finally, according to Equations (5.52) and (5.53), updating a single bit imposes a computational cost of

- Multiplications: $2NL$;
- Additions: $6NL$.

Since during the course of a single iteration of the Gibbs-Sampler a total of $A = N_{\text{bit}}\mathfrak{M}(N_F + L - 1)$ bits have to be updated, the computational cost imposed by iterating the Gibbs-Sampler once is given as:

- Multiplications: $2NLN_{\text{bit}}\mathfrak{M}(N_F + L - 1)$;
- Additions: $6NLN_{\text{bit}}\mathfrak{M}(N_F + L - 1)$.

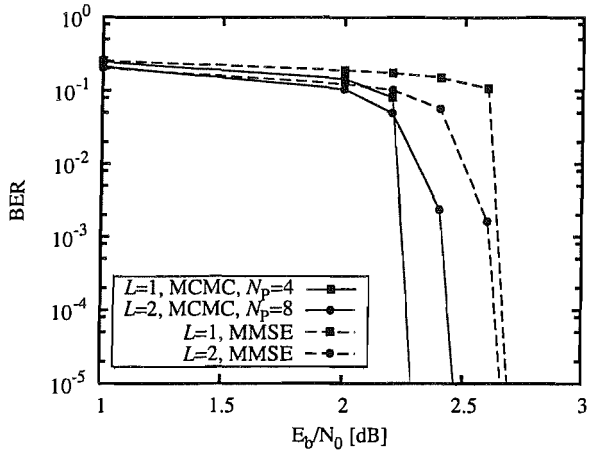
Translating this into real-valued operations by considering one complex-valued addition being equal two real-valued additions and one complex-valued multiplication being equal seven real-valued additions plus multiplications, one iteration of the Gibbs-Sampler imposes a cost of $26NLN_{\text{bit}}\mathfrak{M}(N_F + L - 1)$ real-valued operations.

Note that the complexity imposed by the calculation of the *a priori* probability of a single sample only involves a low number of operations and contributes little to the total computational cost associated with the Gibbs-Sampler. It has therefore not been included in the presented complexity evaluation.

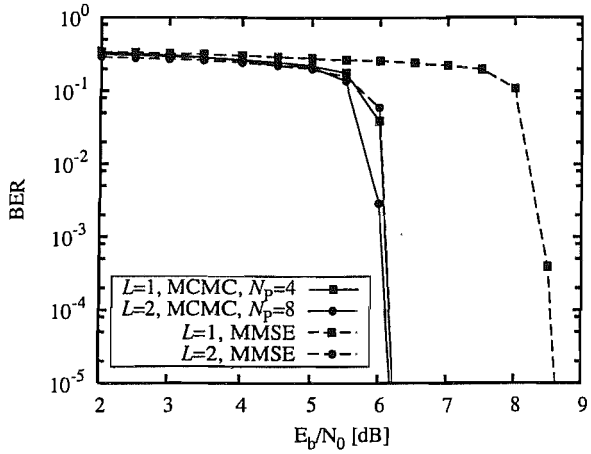
Total Computational Cost: Considering an MCMC aided detector, which uses N_p parallel Gibbs-Samplers, each generating N_{MC} samples, the computational cost expressed in terms of real-valued additions and multiplications, including the initialization of the Gibbs-Sampler, can be summarized as

$$C = N_{MC}N_p 26NLN_{\text{bit}}\mathfrak{M}(N_F + L - 1) + N_p 9NN_F L\mathfrak{M}. \tag{5.55}$$

It can be seen that the only component that does not contribute linearly to the complexity is the channel dispersion, which results in a squared contribution to the computational complexity.



(a) Single-user MIMO system employing $M = 4$ transmit AEs and $N = 4$ receive AEs.



(b) Two-user MIMO system, each user employing $M = 2$ transmit AEs. The BS employed $N = 2$ receive AEs.

Figure 5.14: BER versus E_b/N_0 performance after 10 turbo iterations between the STE and the SCC code for 4QAM modulated which were transmitted over independently fading Rayleigh channels characterized by an L-path equal-power CIR. The STE was characterized by $N = 2$, $N_F = L - 1$ and soft DF. The interleaver-length was chosen to 32 000 bits and the number of samples drawn by one Markov chain of the detector was set to $N_{MC} = 10$. The channel code used for $L = 1$ was SCC Code 1 of Table 4.4, while for $L = 2$ we employed SCC Code 2 of Table 4.4. The number of inner iterations was set to 10.

5.3.6 Simulation Results for Wideband Channels

The first system considered in Figure 5.14 transmitted 4QAM signals over uncorrelated Rayleigh fading channels characterized by an L-path equal-power CIR. The STE was characterized by $N_F = L - 1$ and benefited from soft DF. The number of information bits was chosen to be 32 000 bits and the number of samples drawn by a single Markov chain of the detector was set to $N_{MC} = 10$. The channel code used for $L = 1$ was the SCC Code 1 of Table 4.4, while for $L = 2$ we employed the SCC Code 2 of Table 4.4. The number of inner iterations of the SCC Code was set to $N_{inner} = 10$.

The BER versus E_b/N_0 performance of this system, when considering a single-user MIMO system having $M = 4$ transmit and $N = 4$ receive AEs is depicted in Figure 5.14a. If the channel is assumed to be frequency flat, it can be seen that the MCMC detector outperforms the MMSE detector by approximately 0.4 dB. For an equal-power two-path channel this gap narrows to approximately 0.2 dB.

A similar trend can be observed, if the system is rendered rank-deficient. We have therefore plotted in Figure 5.14b the average BER versus E_b/N_0 performance for a two-user system, where both users' signals were received at an identical average power. The BS employed $N = 2$ receive AEs. It can be seen from Figure 5.14b that for the narrowband system the MCMC detector clearly outperforms the MMSE-aided STE, while for the dispersive channel the corresponding gap narrows again to approximately 0.2 dB.

It becomes clear from Figure 5.14 that for 4QAM signals the additional complexity invested in the detection

[4] $\texttt{jaclog}(a, b) = \log(\exp(a) + \exp(b))$

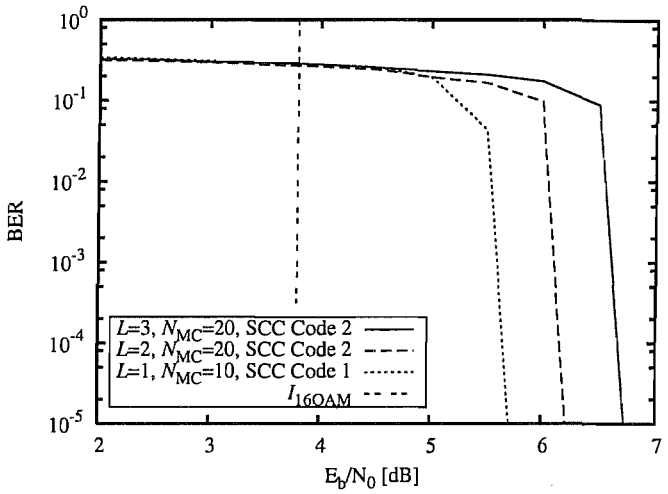


Figure 5.15: BER versus E_b/N_0 performance for a MIMO system employing $M = 4$ transmit AEs and $N = 4$ receive AEs. The signals were 16QAM modulated using Gray mapping and transmitted over L-path equal power independent Rayleigh fading channels. If the channel was dispersive the BS employed a STE characterized by $N_F = L$, $\Delta = L - 1$ and $N_B = L - 1$ while benefiting from soft DF. The illustrated BER performance was obtained for an interleaver length of 48 000 bits, after 10 turbo iterations between a half-rate turbo code and the MCMC detector. The exact channel code parameters for SCC 1 as well as SCC 2 can be found in Table 4.4. The MCMC detector employed N_p=10 parallel Markov chains.

of the MIMO signals by using a MCMC aided rather than a MMSE aided detector, might yield only marginal performance improvements. However, in contrast to the proposed MMSE detector, the MCMC detector discussed is readily capable of detecting higher-order QAM signals.

In Figure 5.15 we have illustrated the BER versus E_b/N_0 performance for a single-user MIMO system employing $M = 4$ transmit AEs and $N = 4$ receive AEs. Gray-coded 16QAM signals were transmitted over L-path equal-power independent Rayleigh fading channels. When the channel was dispersive, the BS employed a STE characterized by $N_F = L$, $\Delta = L - 1$ and $N_B = L - 1$, while benefiting from soft DF. The illustrated BER performance was obtained for a interleaver length determined by 32 000 information, i.e 64 000 encoded bits, after $N_{iter} = 10$ turbo iterations between a half-rate turbo code and the MCMC detector. The exact channel code parameters for SCC 1 as well as SCC 2, each of which employed $N_{inner} = 10$ inner iterations, can be found in Table 4.4. The MCMC detector employed $N_p = 10$ parallel Markov chains, each generating N_{MC} samples. For the evaluation of the *a posteriori* LLR values, the importance sampling approach of Section 5.3.2.2 was considered. Additionally, the theoretically achievable information rate bound is provided as a reference. It can be observed in Figure 5.15, that the MCMC detector is capable of detecting the signals at an infinitesimally low BER within less than 2 dB of the theoretical capacity limit of the narrowband channel, while for the wideband channels considered the distance to capacity was slightly increased owing to the residual ISI.

In Table 5.2 we have provided the number of MIMO-symbol candidates used by the MCMC aided STE as well as the complexity imposed by the Gibbs-Sampler of the MCMC algorithm in terms of real-valued multiplications and additions, calculated according to Equation (5.55). The results provided in Table 5.2

Table 5.2: Computational complexity associated with the BER performance curves provided in Figure 5.15.

	Size of search-space $\mathcal{M}^{M(L+N_F-1)}$	Number of MCMC candidates $N_{MCMC} = N_{MC}N_P$	Complexity calculated according to Equation (5.55)
$L = 1$	65 536	100	167 840
$L = 2$	$2.81 \cdot 10^{14}$	200	2 002 560
$L = 3$	$1.21 \cdot 10^{24}$	200	5 004 960

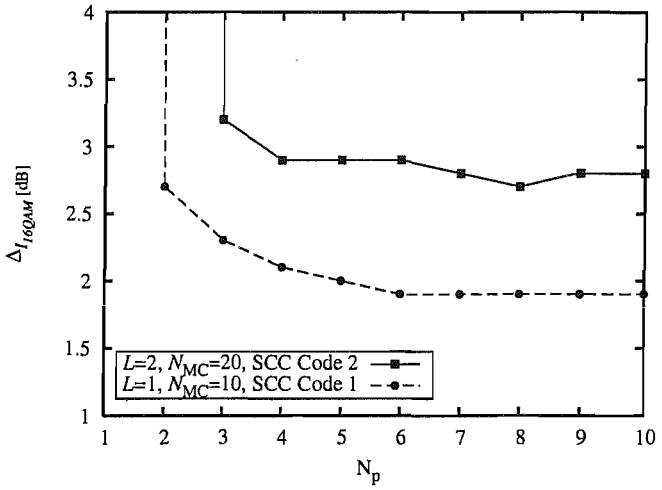


Figure 5.16: Distance to the channel capacity bound for a single-user (4×4)-antenna MIMO system as a function of the number of parallel Markov chains N_P employed by the STE. The signals were 16QAM modulated using Gray mapping and transmitted over a L-path equal power independent Rayleigh fading channels. If the channel was dispersive the BS employed a STE characterized by $N_F = L$, $\Delta = L - 1$ and $N_B = L - 1$ while benefiting from soft DF.

clearly illustrate that the number of candidates considered by the MCMC detector is only a modest fraction of the entire ML search-space. Furthermore, it can also be observed that the actual number of real-valued operations increased at a significantly slower rate than the size of the ML search-space.

In order to illustrate the effect of the number of candidates considered by the MCMC detector on the attainable performance, in Figure 5.16 we have plotted the distance $\Delta_{I_{16QAM}}$ from the theoretically achievable information rate as a function of the number of parallel Markov chains considered by the STE. The system considered was identical to that used for generating the results of Figure 5.15, but we varied the number of parallel chains N_P. The distance from the theoretical limit was calculated for a BER of 10^{-5}. It can be seen in Figure 5.16 that for the narrowband channel associated with $L = 1$, at least two parallel chains have to be considered in order to achieve steady-state operation, but that increasing the number of parallel chains to more than six may not result in any further substantial performance improvements. For the equal-power two-path channel at least three chains are required, while using more than eight parallel chains does not result in a further performance improvement.

5.4 Conclusion

In this chapter we have first introduced a set of SISO STE algorithms, which constitute an evolution of the narrowband single-user MIMO systems discussed in Chapter 4. The classic SISO STE algorithms discussed were the SISO MMSE technique of Section 5.2.1, the Bayesian solution of Section 5.2.2 as well as the novel RML STE of Section 5.2.3. These three different algorithms were also compared in terms of the computational cost they impose, as well as with respect to the achievable performance. It was observed, that while the Bayesian as well as the RML aided detector offer potential performance gains over the SISO MMSE based STE, the complexity imposed by both the RML as well as the Bayesian STE was in excess of state-of-the-art practical complexity limitations at the time of writing. When comparing the MMSE scheme as well as the RML aided SISO STE, it was observed in Section 5.2.3 that in certain scenarios where the STE operates at low SNR values, the MMSE outperforms the RML detector, since the Max-Log approximation is no longer valid.

Therefore, a novel MCMC based detection algorithm was introduced in Section 5.3 in order to cope with the increased search-space dimensions but still being able to perform non-linear, approximate Log-Map detection. Furthermore, it allows the detection of higher-order modulation schemes and supports all types of bit mappings.

In Table 5.3 we have summarized the optimization criteria of the various algorithms discussed and have portrayed the computational complexity associated with the SISO STE strategies considered. Again, as already observed for the narrowband SISO MIMO detection strategies, it can be observed that the MMSE based SISO detector exhibits the lowest computational complexity, while the Bayesian STE exhibits the highest complexity. The complexity of the RML SISO STE depends on the SNR, but on average it is substantially higher than that of the MMSE-aided SISO STE. In fact, it was observed in Figure 5.4 that the complexity tends to grow exponentially with the number of CIR paths. The complexity associated with the novel MCMC aided STE depends on the specific number of candidate samples considered, but it was shown in Equation (5.55) that the complexity of the Gibbs-Sampler used by the MCMC aided detector only increases quadratically with the number of CIR taps. Nevertheless, since for an increased search-space the number of candidates has to be increased, the complexity imposed by the MCMC aided STE is higher than that of the MMSE aided STE, but lower than that of the RML aided STE.

In order to put the performance of the MCMC aided detector into context, the performance measured as the distance from the channel capacity limit, has been included in Table 5.3. It can be observed that for all systems associated with $L = 1$ the MCMC aided detector approaches the performance of the Bayesian detector and slightly exceeds that of the Max-Log based RML SISO detector. When increasing the number of CIR paths to $L = 2$, it can be observed form Table 5.3 that the performance discrepancy between the

non-linear detectors and the MMSE based SISO STE narrows. This effect is even more pronounced when considering rank-deficient systems, as illustrated in Figure 5.14b. However, the MMSE aided SISO STE in its current form did not support the detection of higher-order QAM constellations, while the MCMC aided STE was shown to be capable of detecting 16QAM signals on the basis of a small number of candidates. The performance degradation compared to the narrowband system is approximately 0.5 dB, as it can be seen from Table 5.3.

Table 5.3: Summary of the SISO STE studied in Chapter 5. The performance results were obtained form Figure 4.18, Figure 5.14 as well as Figure 5.15.

		MMSE of Section 5.2.1	**Bayesian** of Section 5.2.2	**RML** of Section 5.2.3	**MCMC** of Section 5.3
Optimization		$\lambda_{\mathrm{E}}^{\mathrm{det}}(\underline{x}_{\mathrm{m}}) = \frac{4\underline{w}^H(\underline{y}-\underline{\mathbf{H}}\underline{\bar{x}}+\bar{x}_{\mathrm{m}}\underline{h}_{\mathrm{m}})}{1-\underline{h}_{\mathrm{m}}^T\underline{w}_{\mathrm{m}}}$, where $\underline{w}_{\mathrm{m}} = \arg\min_{\underline{w}_{\mathrm{m}}\in\mathbb{R}^{2N\times 1}} \|\underline{x}-\underline{w}_{\mathrm{m}}^H(\underline{y}-\underline{\mathbf{H}}\underline{\bar{x}}+\bar{x}_{\mathrm{m}}\underline{h}_{\mathrm{m}})\|^2$	$\lambda_{\mathrm{P}}^{\mathrm{det}}(\underline{x}_{\mathrm{m}}) = \log\left(\frac{P(x_{\mathrm{m}}=+1\|\underline{y})}{P(x_{\mathrm{m}}=-1\|\underline{y})}\right)$	$\lambda_{\mathrm{P}}^{\mathrm{det}}(\underline{x}_{\mathrm{m}}) \approx \log\left(\frac{\max_{\check{\underline{x}}\in\underline{\mathcal{X}}_{\mathrm{m}}^{+}} P(\check{\underline{x}})p(\underline{y}\|\check{\underline{x}})}{\max_{\check{\underline{x}}\in\underline{\mathcal{X}}_{\mathrm{m}}^{-}} P(\check{\underline{x}})p(\underline{y}\|\check{\underline{x}})}\right)$	$\lambda_{\mathrm{P}}^{\mathrm{det}}(\underline{x}_{\mathrm{m}}) \approx \log\left(\frac{\sum_{\check{\underline{x}}\in\underline{\mathcal{X}}_{\mathrm{MC,m}}^{+}} P(\check{\underline{x}})p(\underline{y}\|\check{\underline{x}})}{\sum_{\check{\underline{x}}\in\underline{\mathcal{X}}_{\mathrm{MC,m}}^{-}} P(\check{\underline{x}})p(\underline{y}\|\check{\underline{x}})}\right)$
Complexity		$O\left(N_{\mathrm{F}}^3 N^3\right)$	$O(2^{N_{\mathrm{bit}}M(N_{\mathrm{F}}+L-1)})$	$O(N^3) << \bullet < O(2^{N_{\mathrm{bit}}M(N_{\mathrm{F}}+L-1)})$	$O(NN_{\mathrm{F}}^3) < \bullet << O(2^{N_{\mathrm{bit}}M(N_{\mathrm{F}}+L-1)})$
Perf., $L=1$	(2x2) 4QAM	1.2	0.7	1.1	0.7
	(2x4) 4QAM	3.5	1.2	1.4	1.2
	(4x4) 16QAM	not eval.	not eval.	2.1	1.9
Perf., $L=2$	(2x2) 4QAM	1.1	0.9	1.0	0.9
	(4x4) 4QAM	1.1	not eval.	not eval.	0.9
	(4x4) 16QAM	not eval.	not eval.	not eval.	2.4

Chapter 6

Conclusion and Future Work

The conclusions provided in this chapter constitute an amalgam of our previously drawn conclusions provided at the end of Chapters 2-5 and establishes their logical connection. Our suggestions for future research are outlined thereafter.

6.1 Chapter Summaries

6.1.1 Chapter 1

In Chapter 1 the state-of-the-art and background knowledge of this treatise was discussed. More specifically, in Section 1.1 we have discussed the benefits of Single-Carrier (SC) systems and highlighted their benefits on the basis of the 3rd Generation Partnership Project (3GPP) Long-Term Evolution (LTE) proposals [7]. Furthermore, in Section 1.2 we have provided a discussion and classified the different detection techniques designed both for single as well as multi-carrier systems, which was accompanied by an overview of the open literature on the different detection approaches. In Section 1.3 the novel contributions of this treatise were highlighted and an outline of the following chapters was provided. Finally, in Section 1.4 the basic channel and system model was introduced.

6.1.2 Chapter 2

In Chapter 2 we have provided a comprehensive overview of state-of-the-art Multiple-Input Multiple-Output (MIMO) detection strategies. The novel contributions of this chapter include our discussion of the Minimum Bit Error Rate (MBER) MIMO detector of Section 2.3.2, where we provided novel insights concerning 4-Quadrature Amplitude Modulated (4QAM) signals based on the real-valued binary system model. Furthermore, we have highlighted the difference between the Bayesian detector of Section 2.3.3, which optimizes

the error-probability of each single transmitted bit and the classic Maximum Likelihood (ML) detector of Section 2.3.4 aiming for finding the most likely transmitted MIMO symbol vector. Owing to the high computational cost imposed by the full-complexity ML detector, a novel reduced-complexity ML detector was discussed in Section 2.3.4, which is referred to as the Optimized Hierarchical Recursive Search Algorithm (OHRSA). In addition to the original OHRSA of [97] and Section 2.3.4, we have also provided an alternative search technique, referred to as the Min-Path search algorithm, which imposed a lower computational complexity than the original OHRSA, especially in rank-deficient systems employing higher-order QAM as seen in Figure 2.19.

All the algorithms discussed in Section 2.3, namely the Minimum Mean Squared Error (MMSE), the MBER, the Bayesian as well as the OHRSA aided detector, have been compared with respect to their specific optimization philosophy, their computational complexity summarized in Section 2.3.5 as well as their BER versus E_b/N_0 performance in Figures 2.23-2.25. In Table 6.1 we have summarized the E_b/N_0-distance expressed in dB and associated with a specific algorithm, measured from the theoretical capacity limits derived in Section 2.2. Explicitly, for the independent Rayleigh block-fading channels we have considered the distance $\Delta_{P_{out}|P_{out}=0.02}$ [dB] between the outage probability curve discussed in Section 2.2.3 and the Frame Error Rate (FER) performance of the specific detector at an outage probability of $P_{out} = 0.02$. By contrast for the uncorrelated independent Rayleigh fading channel we considered the distance $\Delta_{I|BER=10^{-4}}$ [dB] between the BER performance curve of the detector and the theoretical channel capacity limit derived in Section 2.2.3 at a BER of 10^{-4}. Furthermore, it can be observed in Figures 2.24 and 2.25 that both the Bayesian as well as the OHRSA aided detector exhibit an almost identical performance.

For the different $(N \times M)$-antenna systems considering Binary Phase-Shift Keying (BPSK) modulated signals as well as half-rate turbo-coding, which was fed with the hard-output of the detector, it can be observed in Table 6.1 that all algorithms operate at a considerable margin from the theoretical capacity limit. However, when comparing the linear MMSE scheme to the linear MBER detector, it can be observed from Table 6.1 that for $N = 2$ receive AEs the MBER detector outperforms the MMSE detector by more than 4 dB. Notice in Table 6.1 that this performance difference decreases upon increasing the number of receive AEs. When employing the optimum non-linear Bayesian detector or the low complexity OHRSA aided detector, it can be observed in Table 6.1 that a further 2 dB performance gain can be achieved. Again, for a higher number of receive AEs the performance difference between the algorithms decreases as a benefit of achieving an increased receive diversity gain. It can also be observed that even the more sophisticated algorithms of Table 6.1 are far from approaching the theoretical channel capacity limit. If an entry in Table 6.1 is marked as 'not evaluated' then either the algorithm imposed an excessive complexity or it was unable to reach the target integrity with the aid of the specific modulation scheme considered. A detailed overview of these performance trends may be found in Table 2.8.

Table 6.1: Summary of the achievable performance of different detection algorithms studied in this treatise.

			MMSE	MBER	Bayesian	OHRSA/RML	MCMC
MIMO [d]	$\Delta_{P_{out}\|P_{out}=0.02}$ [dB]	$(2\times 2)_{BPSK}$	12.1	7.9	5.1	5.1	not eval. [a]
		$(4\times 2)_{BPSK}$	5.6	4.6	3.6	3.6	not eval. [a]
	$\Delta_{I\|BER=10^{-4}}$ [dB]	$(2\times 2)_{BPSK}$	7.1	not eval. [b]	5.6	5.6	not eval. [a]
		$(3\times 2)_{BPSK}$	5.4	not eval. [b]	3.9	3.9	not eval. [a]
		$(4\times 2)_{BPSK}$	4.5	not eval. [b]	4.2	4.2	not eval. [a]
STE No DF [e]	$\Delta_{P_{out}\|P_{out}=0.02}$ [dB]	$(2\times 2)_{BPSK}$	>10	6.2	4.8	4.8	not eval. [a]
	$\Delta_{I\|BER=10^{-4}}$ [dB]		8.3	not eval. [b]	5.5	5.5	not eval. [a]
STE DF [f]	$\Delta_{P_{out}\|P_{out}=0.02}$ [dB]		8.2	6.0	4.5	4.5	not eval. [a]
	$\Delta_{I\|BER=10^{-4}}$ [dB]		6.5	not eval. [b]	5.2	5.2	not eval. [a]
SISO MIMO [g]	$\Delta_{I\|BER=10^{-5}}$ [dB]	$(2\times 2)_{4QAM}$	≈ 1	not eval. [b]	<1	≈ 1	<1
		$(2\times 4)_{4QAM}$	3.5	not eval. [b]	1.2	1.4	1.2
		$(4\times 4)^{Hier.}_{16QAM}$	not eval.[c]	not eval. [b]	not eval. [b]	1.3	1.2
		$(4\times 4)^{Gray}_{16QAM}$	not eval. [c]	not eval. [b]	not eval. [b]	not. eval. [c]	1.9
SISO STE [h]	$\Delta_{I\|BER=10^{-5}}$ [dB]	$(4\times 4)_{4QAM}$	1.1	not eval. [b]	not. eval. [b]	not. eval. [b]	0.9
	Soft DF	$(4\times 4)^{Hier.}_{16QAM}$	not eval. [c]	not eval. [b]	not eval. [b]	not eval. [b]	>8
	$L=2$	$(4\times 4)^{Gray}_{16QAM}$	not eval. [c]	not eval. [b]	not eval. [b]	not eval [b]	2.4

[d] Results obtained from Figure 2.25 and Figure 2.26, respectively.
[a] Algorithm has not been introduced at this point of the treatise.
[b] Imposed complexity too high.
[e] Results obtained from Figure 3.19 and Figure 3.20, respectively.
[f] Results obtained from Figure 3.34 and Figure 3.33, respectively.
[g] Results obtained from Figure 4.17 and Figure 4.18, respectively.
[c] Algorithm in its current form does not support modulation scheme.
[h] Results obtained from Figure 5.14 and Figure 5.15, respectively.

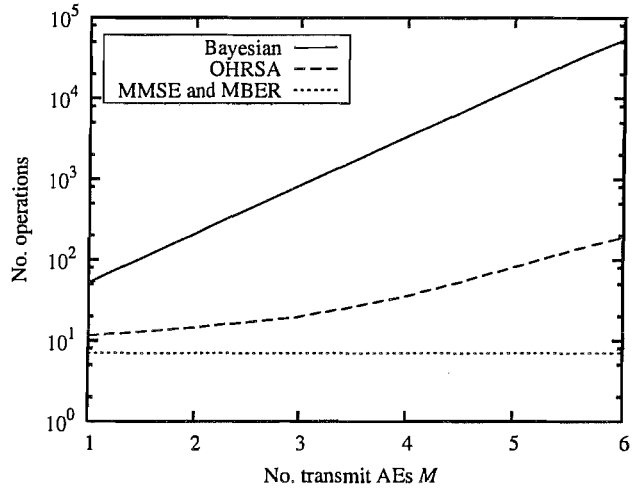
The performance figures provided in Table 6.1 have been obtained for relatively low-dimensional systems, which allowed us to consider more sophisticated but naturally more complex detection algorithms. For these low-dimensional systems the four algorithms considered, namely the MMSE, the MBER, the Bayesian and the OHRSA aided detector impose a computational complexity which is of a similar order of magnitude. In this summary chapter, we are however less interested in the complexity imposed by low-throughput, low-dimensional systems, but rather in identifying the general computational complexity trends when increasing the throughput of the system.

When considering the computational cost associated with the different MIMO detection algorithms, which was detailed in Section 2.3.5, it was observed for independent Rayleigh block-fading channels that both the MMSE as well as the MBER algorithm exhibit a moderate detection complexity, which is linearly proportional to the number of receive AEs. This trend can also be observed in Figure 6.1a, where we have plotted the normalized detection complexity of a single bit for the specific algorithms as a function of the number of transmit AEs M. The number of receive AEs was chosen to be $N = 2$ and the modulation scheme considered was 4QAM. In contrast to the linear MMSE and MBER MIMO schemes, the complexity associated with the Bayesian detector is an exponentially increasing function of the number of transmit AEs, while that of the OHRSA-aided detector is slightly higher than that of the linear MMSE and MBER detectors.

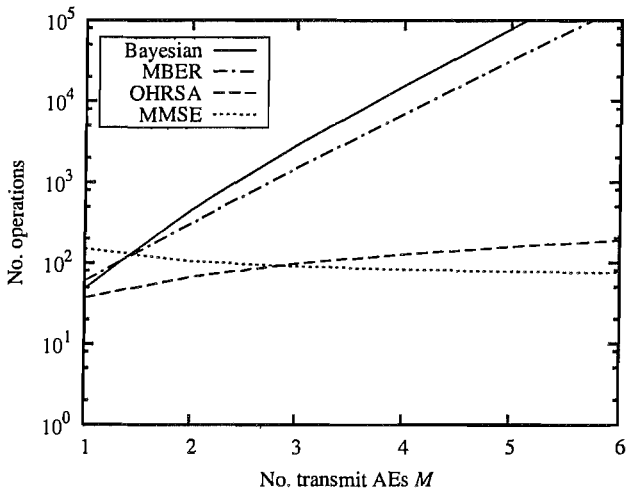
The computational complexity associated with the preprocessing carried out by each algorithm before the bits can be detected is illustrated in Figure 6.1b based on Tables 2.3, 2.4, 2.5 and 2.6, where the graphs were obtained from Figures 2.8, 2.13, 2.16 and 2.19-2.21 of Section 2.3. For the MMSE as well as the MBER algorithm the preprocessing complexity is constituted by the evaluation of the weight vectors, while that of the Bayesian detector is determined by the evaluation of the legitimate channel output states. The preprocessing invoked by the OHRSA algorithm is mainly determined by that of the Cholesky factorization required, as discussed in Section 2.3.4. It can be observed from Figure 6.1b that while the Bayesian as well as the MBER detector exhibit a substantial preprocessing complexity, which is exponentially proportional to the number of transmit AEs, the MMSE as well as the OHRSA-aided MIMO detector's computational complexity increases only moderately upon increasing the number of transmit AEs.

6.1.3 Chapter 3

The algorithms discussed in Chapter 2 in the context of non-dispersive channels have been extended to Space-Time Equalization (STE) algorithms applied in dispersive scenarios in Chapter 3, which additionally may benefit from employing Decision Feedback (DF). In order to physically interpret the rationale of using DF, we have introduced a general space-translation structure in Section 3.5 and have provided further interpretations of using DF as a subset selection approach in Section 3.5.3 as well as a tree-based DF structure in



(a) MIMO detection complexity.



(b) MIMO preprocessing complexity.

Figure 6.1: Normalized computational complexity defined as the number of real-valued multiplications plus additions required for the detection of a single bit which were expressed in Tables 2.3,2.4,2.5 and 2.6 and extracted from Figures 2.8, 2.13, 2.16 and 2.19-2.21 of Section 2.3. All proportionality factors were chosen to be unity and the modulation scheme considered was 4QAM. The receiver was assumed to employ $N = 2$ receive AEs and the computational complexity imposed by the OHRSA was recorded at E_b/N_0=20 dB.

Section 3.5.4. The novel contribution of Chapter 3 was the introduction of the MBER, the Bayesian as well as the OHRSA aided STE. More specifically the OHRSA aided SC STE of Section 3.4.4 was improved with the aid of a tree-truncation scheme, which allowed us to control the complexity imposed by the tree-search at a moderate performance degradation. Furthermore, we have provided Example 3.3 for highlighting the difference between the conventional tree-search algorithm associated with the OHRSA and the novel Min-Path tree-search.

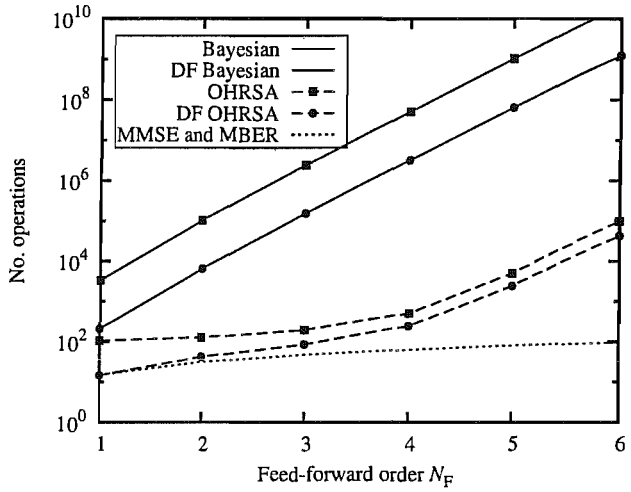
The extension of the narrowband MIMO detection schemes of Chapter 2 to the more realistic wideband scenario of Chapter 3 has been beneficial for all four detection algorithms' performances, which were found to be marginally closer to the channel's capacity limit, as it can be observed by comparing rows 1 and 6 as well as rows 3 and 7 of Table 6.1. Naturally, the wideband schemes exhibit a higher complexity than the corresponding narrowband scenarios. The relative performance difference between the specific algorithms appears to be approximately the same. Again for the OHRSA we proposed an alternative to the classic space-translation based interpretation of the DF, since we advocated in Section 3.5.4 an amalgamation of the DF directly into the tree-search.

It was observed that while DF resulted in a significant performance improvement for the MMSE-based STE, it only marginally improved the performance of the MBER scheme, of the Bayesian as well as of the OHRSA-aided detector, as it can be seen from Table 6.1.

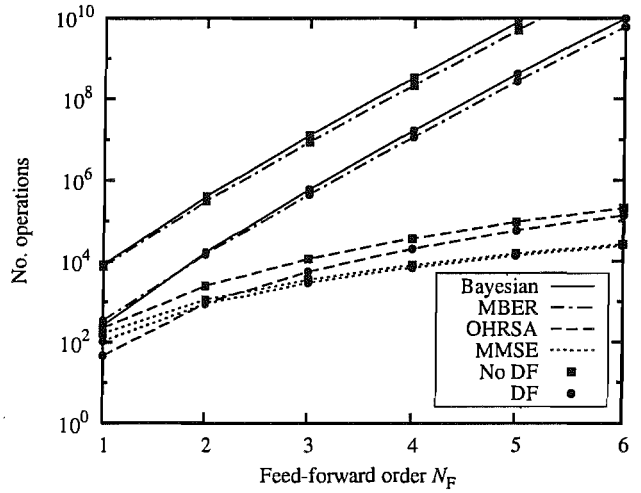
The normalized computational cost imposed by the different STE algorithms is summarized in Figure 6.2a and Figure 6.2b, where we have portrayed the detection complexity and the preprocessing complexity, respectively obtained according to Tables 3.1-3.4 as well as Tables 3.5-3.7. The system considered was assumed to employ $N = 2$ receive AEs, while using $M = 2$ transmit AEs and 4QAM signals. The channel was assumed to have an equal-power two-tap CIR. It can be seen from Figure 6.2a that the detection complexity of the linear MMSE and MBER STE does not depend on whether DF is considered or not. By contrast, the detection complexity of the Bayesian as well as that of the OHRSA aided STE decreased, when DF was employed. Note that if DF is employed, the computational complexity imposed by the detectors is more-or-less independent of the number of channel taps and only depends on the feed-forward order chosen, while that of the STE dispensing with DF depends largely on the number of CIR taps, as discussed throughout Chapter 3.

Similar to the narrowband MIMO detectors considered, the preprocessing complexity of the MBER scheme as well as that of the Bayesian detector is exponentially proportional to the feed-forward-order, while that of OHRSA arrangement as well as that of the MMSE detector increases only moderately upon increasing the feed-forward order.

Furthermore, it is concluded that while DF has improved the performance of the MMSE detector, it only



(a) STE detection complexity for an equal-power two-tap channel and $M = 2$ transmit AEs.



(b) STE preprocessing complexity for an equal-power two-tap channel and $M = 2$ transmit AEs..

Figure 6.2: Normalized computational complexity defined as the number of real-valued multiplications plus additions required for the detection of a single bit, which were obtained according to Tables 3.1-3.4 as well as Tables 3.5-3.7. All proportionality factors were chosen to be unity and the modulation scheme considered was 4QAM. The receiver was assumed to employ $N = 2$ receive AEs and the computational complexity imposed by the OHRSA was recorded at E_b/N_0=20 dB.

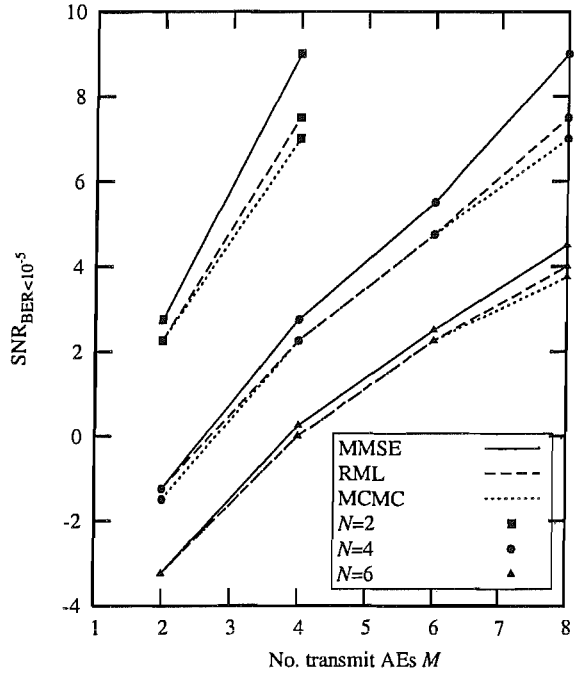
moderately reduced the complexity. By contrast, the employment of DF more substantially reduced the computational cost associated with the non-linear detectors, although it only resulted in a moderate performance improvement. Finally, we have introduced a novel reduced-complexity ML STE based on the OHRSA, which was further enhanced by a novel tree-truncation procedure in order to control the algorithm's computational cost. A detailed summary of Chapter 3 can be found in Table 3.8, where the different algorithms' performances as well as the complexity they impose are portrayed.

6.1.4 Chapter 4

Since the narrowband detection algorithms discussed in Chapter 2 were incapable of approaching the channel capacity bound, we have improved the algorithms further in order to be able to process soft-information. Hence, in Section 4.4.3 we proposed a novel Max-Log SISO detector based on the OHRSA algorithm, which was referred to as the Reduced complexity Max-Log (RML) detector. Furthermore, a discussion was provided in Section 4.4.4 as to why iterative MBER detection might not be attractive. In order to analyze the proposed SISO detectors, in Section 4.3 we invoke their EXIT chart analysis and proposed a novel projection based approach, which facilitates the two-dimensional EXIT chart analysis of systems consisting of three serially concatenated SISO components, namely the detector and two concatenated channel codes. The resultant two-dimensional EXIT chart was plotted in Figure 4.4. The two channel codes may also be interpreted as a single serially concatenated turbo-code.

It can be observed from Table 6.1 that with the advent of iterative MIMO detection, the MMSE detector of Section 4.4.1, the novel RML detector of Section 4.4.2 as well as the Bayesian detector of Section 4.4.3 are capable of approaching the channel capacity limit, provided that the system has full rank. If, however, the system becomes rank-deficient, the MMSE detector fails to approach the channel capacity limit, while the RML as well as the Bayesian detector still remain capable of approaching the capacity limit within about 1.3 dB. Furthermore, it can be seen from Table 6.1 that the RML detector is capable of operating within 1.2 dB of the information theoretic limit even for 4QAM signals.

In Figure 6.3 we have extracted the results form Figures 4.14-4.18 as well as Figures 5.13-5.15 and plotted the SNR required in order to achieve a BER of 10^{-5} as a function of the number of transmit AEs for 4QAM modulated signals and for a different number of receive AEs. For $N \geq M$ the system used the SCC Code 1 of Table 4.4 as a channel code, otherwise SCC Code 2 of Table 4.4 was employed. The SCC code used 10 inner iterations. For missing points additional simulations were performed. For complexity reasons, no results are included for the SISO Bayesian detector. From the graphs seen in Figure 6.3 it appears that the SNR required by a certain system configuration, as well as the relative performance of the different SISO algorithms is scalable. For example, when considering a (4×4) and a (2×2)-antenna system, both require the same SNR in order to achieve a BER of 10^{-5} and even the performance difference between the three



MCMC Parameters

(N_p, N_{MC})	$M=2$	$M=4$	$M=6$	$M=8$
$N=2$	4,10	4,10		
$N=4$	4,10	4,10	7,10	10,10
$N=6$	4,10	4,10	6,10	8,10

Figure 6.3: SNR required in order to achieve a BER of 10^{-5} as a function of the number of transmit AEs for 4QAM modulated signals and a different number of receive AEs. If $N \geq M$ the system used SCC Code 1 of Table 4.4 as a channel code, otherwise SCC Code 2 of Table 4.4 was employed. The SCC code used 10 inner iterations. Most results were obtained from Chapter 4 as well as Chapter 5. For missing points additional simulations were performed. For the considered system the throughput is identical to the number of transmit AEs M, since the throughput can be expressed as $R = M \log_2(\mathcal{M}) R_c = M \cdot 2 \cdot \frac{1}{2} = M$, where R_c is the overall code-rate.

algorithms considered is similar. The same trends can be observed, when considering the rank-deficient (2×4) and (4×8)-antenna systems. As already seen from Table 6.1, the performance difference between the MMSE and the RML detector becomes more pronounced for rank-deficient systems.

The computational complexity associated with the different iterative MIMO detection algorithms was extracted from Figures 4.5,4.8 and 4.11 and is summarized in Figure 6.2b, where a system employing $M = 2$ transmit AEs and 4QAM signals was considered. Firstly, it can be observed that while the computational cost associated with the MMSE detector is lower for the first iteration than for all subsequent iterations, that of the Bayesian as well as of the RML detector is lower for all subsequent iterations than it is for the first iteration.

6.1.5 Chapter 5

In Chapter 5 we have provided a frame work for extending the SISO narrowband MIMO detectors of Chapter 4 to SISO STE algorithms. In addition to the DF methods of Chapter 3 employed for non-iterative STE, we have also introduced a novel soft-DF method. The novel soft-DF method did not offer the same complexity reduction as the hard DF technique of Chapter 3, but by contrast it did not suffer from error-propagation and hence remained capable of delivering a performance gain. This behaviour was observed for

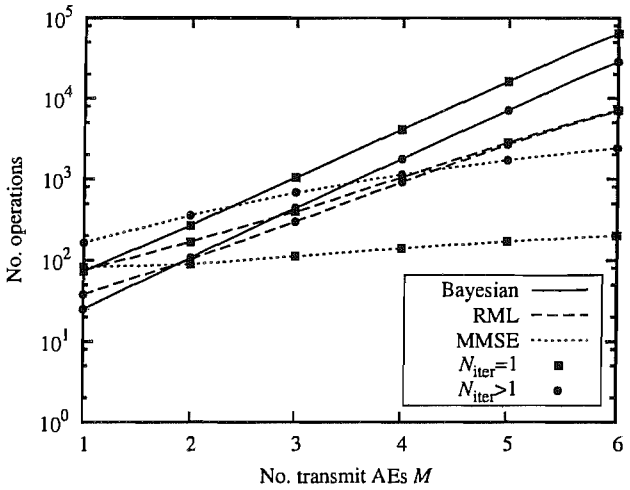


Figure 6.4: Normalized iterative computational complexity defined as the number of real-valued multiplications plus additions required for the detection of a single bit for independent uncorrelated Rayleigh fading channels extracted from Figures 4.5,4.8 and 4.11. All proportionality factors were chosen to be unity and the modulation scheme considered was 4QAM. The receiver was assumed to employ $N = 2$ receive AEs.

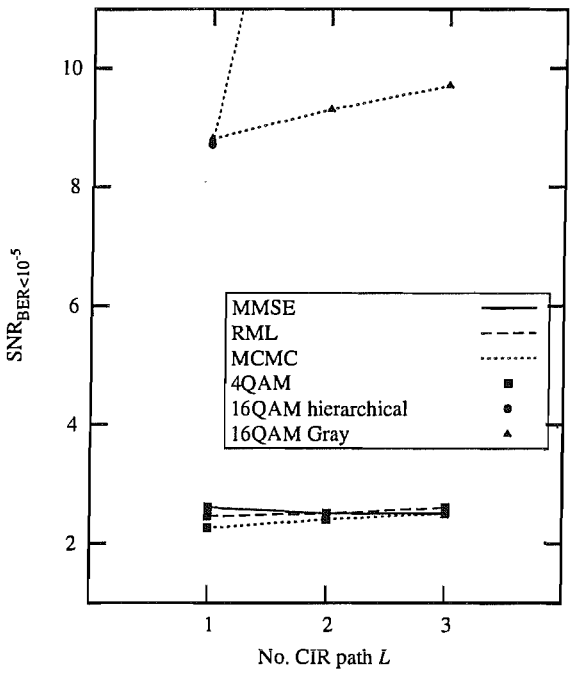
all the three considered SISO algorithms, namely the MMSE SISO STE of Section 5.2.1, the RML SISO STE of Section 5.2.2 as well as the Bayesian SISO STE of Section 5.2.3.

When comparing the RML aided and the MMSE aided STE of Sections 5.2.1 and 5.2.3, respectively, it was observed that the RML algorithm was only capable of providing a performance gain over the MMSE STE for near-far scenarios, as seen from Figure 5.6. In certain cases it was even observed that the Max-Log approximation employed by the RML algorithm resulted in a performance loss in comparison to the low-complexity MMSE algorithm as illustrated in Figure 5.5.

In order to mitigate the above-mentioned deficiencies of the RML aided STE in Section 5.2.3, we proposed a novel Markov-Chain Monte-Carlo (MCMC) aided STE in Section 5.3, which did not rely on the Max-Log approximation and was capable of exploring large search-spaces at a moderate computational cost. As a further benefit the MCMC STE facilitated the detection of arbitrary modulation schemes and bit-to-symbol mappings.

In Figure 6.5 we have illustrated the SNR required for achieving a BER of 10^{-5} as a function of the number of CIR taps for a (4×4)-antenna single-user MIMO system, while considering different modulation schemes as well as different detection algorithms. It can been from Figure 6.5 that if 4QAM signals are considered, the performance difference between the three algorithms considered decreases upon increasing the number of CIR taps L.

It also can be observed from Figure 6.5, that in conjunction with the channel code considered, 16QAM using hierarchical and Gray mapping perform similarly in a narrowband scenario. If the number of CIR taps is increased, the system using 16QAM in conjunction with Gray mapping only experiences a moderate performance degradation. By contrast, the system considering 16QAM combined with hierarchical mapping



MCMC Parameters

(N_p, N_{MC})	$L=1$	$L=2$	$L=3$
4QAM	4,10	8,10	8,10
$16QAM^{Hier.}$	10,10	10,20	10,20
$16QAM^{Gray}$	10,10	10,20	10,20

Figure 6.5: SNR required in order to achieve a BER of 10^{-5} as a function of the number of CIR path L for a (4×4)-antenna system considering different modulation schemes as well as different detection algorithms. If $L = 1$ and the modulation scheme was 4QAM the system used SCC Code 1 of Table 4.4 as a channel code, otherwise SCC Code 2 of Table 4.4 was employed. The SCC code used 10 inner iterations. Most results were obtained from Chapter 4 as well as Chapter 5. For missing points additional simulations were performed.

exhibits a substantial performance degradation, as also seen from lines 12 and 15 of Table 6.1.

The computational complexity of the SISO MMSE, Bayesian as well as RML STE was analyzed in Section 5.2, where it was found that the Bayesian as well as the RML aided SISO STE exhibit an exponentially increasing computational cost upon increasing the number of CIR taps. By contrast, the complexity imposed by the MMSE detector is cubic in the number CIR path. The complexity imposed by the novel MCMC aided STE depends on the number of parallel Markov Chains considered as well as on the number of samples generated by the Gibbs sampler. It was however observed in Table 5.2 as well as in the parameter-list of the MCMC aided STE of Figure 6.5 that the complexity imposed by the MCMC detector increased significantly slower than the exponentially growing size of the search-space.

6.2 Future Work

6.2.1 Non-Binary OHRSA

It was observed in Section 5.3.4 that the convergence of iterative MIMO receivers can be influenced by modifying the bit-mapping of the considered modulation scheme, many of which can not be expressed using the real-valued binary system model of Section 2.1.1.2. It may therefore be interesting to construct a

non-binary OHRSA algorithm, which is capable of detecting signals with arbitrary bit-to-symbol mappings. For a narrowband MIMO detector this results in a search tree with M rather than $M\log_2(\mathcal{M})$ hierarchical levels. The number of branches leaving each node would then be increased from two to $\log_2(\mathcal{M})$ and therefore a higher computational cost would be imposed by the evaluation of a single node. Despite the fact that the evaluation of the search tree on each hierarchical level would become more complex, the convergence of the algorithm would be improved since a higher number of hypothetical solutions to the detection problem are considered jointly.

6.2.2 Non-Binary Gibbs-Sampler

In analogy to the non-binary OHRSA also the Gibbs-Sampler of Section 5.3.3 may be extended to a non-binary algorithm. When introducing the Gibbs-Sampler in Section 5.3.3 we allowed only a single bit of the hypothetical solution to be changed during one evolution of the Gibbs-Sampler. It is however be possible to change an arbitrary number of bits jointly in order to support a faster convergence [69]. Again, similarly to the non-binary OHRSA proposed in Section 6.2.1, the complexity associated with each evaluation step of the Gibbs-Sampler is increased but its convergence rate will be improved and the MCMC aided detector becomes more robust against stalling. Interpreting the non-binary Gibbs-Sampler form a Markov chain perspective, a non-binary Gibbs-Sampler results in a higher number of transitions between the nodes and therefore allows to move more rapidly through the search-space.

6.2.3 Joint Detection and Channel Estimation

In Section 5.3 we have shown that MCMC detection can be efficiently employed in the context of MIMO systems. It was however purely employed for detection purposes. One strength of the MCMC aided detector is that it can be evolved to perform joint channel estimation and data detection as for example outlined in [69]. When considering MIMO systems for joint detection and channel estimation special care has to be taken in order to resolve the inherent ambiguity.

6.2.4 EXIT Charts for Short Interleaver-Depth

The iterative detectors of Chapter 4 and Chapter 5 were only discussed in the context of ergodic Rayleigh fading channels in conjunction with long interleaver length in order to be capable of performing EXIT chart analysis of the system. A scenario of more practical interest would however be to consider slowly fading or block-fading channels in conjunction with short interleavers. For the design and the analysis of such systems a new type of EXIT charts would have to be developed.

6.2.5 Cyclic Prefix Aided Space-Time Equalization

A further focus of our future work should be on cyclic prefix aided STEs. Attaching a cyclic prefix to the transmitted signal renders the transmitted signal to appear periodic to the receiver and therefore facilitates Frequency-Domain (FD) processing. The most commonly used FD STE is the MMSE equalizer. Especially in highly frequency-selective fading channels FD processing is preferable to Time-Domain (TD) STE, which would require a complex high-order filter. However, to the best of our knowledge, FD STE has so far only been considered in the context linear detectors, since the decision space in the FD may appear to be excessive for non-linear detection. We have seen from our previous results that the MMSE equalizer's performance becomes poor when it comes to rank-deficient systems. It may therefore be of interest to investigate the performance of non-linear detectors such as the MCMC aided detector in the context of cyclic-prefix-aided STE.

Appendix A

BER Gradient

The BER at the detector's output for a given weight vector $\boldsymbol{w}_m$ can be written as

$$P_{\mathrm{e}}(\boldsymbol{w}_m) = \frac{1}{N_{\mathrm{sb}}} \sum_{\check{x}_m^{(i)} \in \hat{\mathcal{X}}_m^+} \mathcal{Q}\left(g_i(\boldsymbol{w}_m)\right) \tag{A.1}$$

with

$$g_i(\boldsymbol{w}_m) = \frac{\mathtt{sign}\left(\check{x}_{\mathrm{R},m}^{(i)}\right) \check{y}_{\mathrm{R}m}^{(i)}}{\sigma_{\mathrm{n}} \sqrt{\boldsymbol{w}_m^H \boldsymbol{w}_m}} \tag{A.2}$$

$$= \frac{\mathtt{sign}\left(\check{x}_{\mathrm{R},m}^{(i)}\right) \Re\left(\boldsymbol{w}_m^H \check{\boldsymbol{y}}_m^{(i)}\right)}{\sigma_{\mathrm{n}} \sqrt{\boldsymbol{w}_m^H \boldsymbol{w}_m}} \tag{A.3}$$

The gradient of $P_{\mathrm{e}}(\boldsymbol{w}_m)$ can be calculated make use of the following integration rule

$$\frac{d}{dt} \int_{a(t)}^{c(t)} f(y)dy = f\left(c(t)\right) \frac{dc(t)}{dt} - f\left(a(t)\right) \frac{da(t)}{dt}. \tag{A.4}$$

This yields

$$\nabla P_{\mathrm{e}}(\boldsymbol{w}_m) = -\frac{1}{N_{\mathrm{sb}}\sqrt{2\pi}} \sum_{i=1}^{N_{\mathrm{sb}}} \exp\left(-\frac{\mathtt{sign}\left(\check{x}_{\mathrm{R},m}^{(i)}\right) \Re\left(\boldsymbol{w}_m^H \check{\boldsymbol{y}}_m^{(i)}\right)}{2\sigma_{\mathrm{n}}^2 \boldsymbol{w}_m^H \boldsymbol{w}_m}\right) \nabla f(\boldsymbol{w}_m) \tag{A.5}$$

$$= -\frac{1}{N_{\mathrm{sb}}\sqrt{2\pi}} \sum_{i=1}^{N_{\mathrm{sb}}} \exp\left(-\frac{\Re\left(\boldsymbol{w}_m^H \check{\boldsymbol{y}}_m^{(i)}\right)}{2\sigma_{\mathrm{n}}^2 \boldsymbol{w}_m^H \boldsymbol{w}_m}\right) \nabla f(\boldsymbol{w}_m), \tag{A.6}$$

where

$$\nabla f(\boldsymbol{w}_m) = \frac{\mathtt{sign}\left(\check{x}_{\mathrm{R},m}^{(i)}\right)}{\sigma_{\mathrm{n}}} \nabla \left(\frac{\Re\left(\boldsymbol{w}_m^H \check{\boldsymbol{y}}_m^{(i)}\right)}{\sqrt{\boldsymbol{w}_m^H \boldsymbol{w}_m}} \right) \tag{A.7}$$

$$= \frac{\nabla\left(\Re\left(\boldsymbol{w}_m^H \check{\boldsymbol{y}}_m^{(i)}\right)\right)}{\sqrt{\boldsymbol{w}_m^H \boldsymbol{w}_m}} + \nabla\left(\frac{1}{\sqrt{\boldsymbol{w}_m^H \boldsymbol{w}_m}}\right) \Re\left(\boldsymbol{w}_m^H \check{\boldsymbol{y}}_m^{(i)}\right) \tag{A.8}$$

$$= \frac{\check{\boldsymbol{y}}_m^{(i)}}{\sqrt{\boldsymbol{w}_m^H \boldsymbol{w}_m}} - \frac{\boldsymbol{w}_m}{\left(\boldsymbol{w}_m^H \boldsymbol{w}_m\right)^{-\frac{3}{2}}} \Re\left(\boldsymbol{w}_m^H \check{\boldsymbol{y}}_m^{(i)}\right). \tag{A.9}$$

The gradient of the BER can now be expressed as

$$\nabla P_{\mathrm{e}}(\boldsymbol{w}_m) = \sum_{\check{x}_m^{(i)} \in \hat{\mathcal{X}}_m^+} \exp\left(-\frac{\left(\check{\hat{x}}_{\mathrm{R},m}^{(i)}\right)^2}{2\sigma_{\mathrm{n}}^2 \boldsymbol{w}_m^H \boldsymbol{w}_m} \right) \mathtt{sign}\left(\check{x}_{\mathrm{R},m}^{(i)}\right) \left(\frac{\check{\hat{x}}_{\mathrm{R},m}^{(i)} \boldsymbol{w}_m}{\boldsymbol{w}_m^H \boldsymbol{w}_m} - \check{\boldsymbol{y}}_m^{(i)} \right), \tag{A.10}$$

where

$$C = \frac{1}{N_{\mathrm{sb}} \sqrt{2\pi\sigma_{\mathrm{n}}^2 \boldsymbol{w}_m^H \boldsymbol{w}_m}}. \tag{A.11}$$

Appendix B

Simplified Conjugate Gradient Algorithm

1. Choose the algorithmic parameters:

   Choose a step size of $\mu > 0$ and a termination criterion (upper limit for the gradient) $\beta > 0$.

2. Initialize:

   Set the initial weight vector $\underline{w}_m(1)$, set $\mathbf{d}(1) = -\nabla P_e(\underline{w}_m(1))$ and set the iteration index to $i = 1$.

3. *Loop:*

   If $||\nabla P_e(\underline{w}_m(i))|| < \beta$

$$\begin{aligned} \underline{w}_m(i+1) &= \underline{w}_m(i) + \mathbf{d}(i)\underline{w}_m(i) \\ \underline{w}_m(i+1) &= \frac{\underline{w}_m(i+1)}{||\underline{w}_m(i+1)||} \\ \phi(i) &= \frac{||\nabla P_e(\underline{w}_m(i+1))||^2}{||\nabla P_e(\underline{w}_m(i))||^2} \\ \mathbf{d}(i+1) &= \phi(i)\mathbf{d}(i) - \nabla P_e(\underline{w}_m(i+1)) \end{aligned} \tag{B.1}$$

   for $i = i + 1$, goto *Loop*.

The solution is given in the form of the vector $\underline{w}_m(i)$ when the algorithm has terminated. For the sake of achieving an attractive convergence behavior, the right choice of the initial weight vector is essential.

Appendix C

Maximum A Posteriori Probability STE

Under the assumption of an infinite observation interval, a receiver design similar to those introduced in Section 3.3 would result in an excessive complexity. Therefore in [80] it was proposed to use a trellis based equalizer. We will opt here for employing the Maximum A Posteriori (MAP) equalizer based on the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm, due to its capability of processing soft information.

Let us first introduce the trellis, which can be associated with a frequency-selective multiple antenna scenario. The number of possible states of the trellis is given by

$$N_{\mathrm{s}} = \mathcal{M}^{\mathfrak{M}(L-1)} \tag{C.1}$$

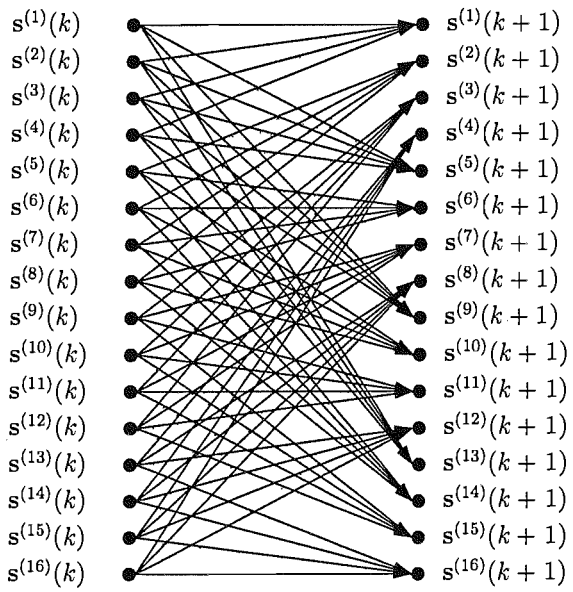


Figure C.1: Trellis segment

and the number of transitions per trellis state can be written as

$$N_t = \mathcal{M}^{\mathfrak{M}}. \tag{C.2}$$

In Figure C.1 a trellis segment corresponding to a two-path channel and either two BPSK modulated transmitters or a single 4QAM transmitter is shown. Furthermore, assume we aim evaluating the *a posteriori* probabilities of the transmitted symbols based on a sequence $\mathcal{Y}$ of received channel output vectors $\mathbf{y}(k)$ defined in Equation (3.6). Assuming further, that we are interested in evaluating the *a posteriori* probability associated with the symbol vector associated with the channel output $\mathbf{y}(k)$, the sequence $\mathcal{Y}$ may be split into three sections: $\mathcal{Y}_k$ which is the channel output state associated with transition at time k, $\mathcal{Y}_{<k}$ associated with the priori received channel output vectors as well as $\mathcal{Y}_{>k}$ associated with the received channel output states after the present transition k.

The joint probability may now be written as [118]

$$\begin{aligned} p(\mathbf{s}', \mathbf{s}, \mathcal{Y}) &= p(\mathbf{s}', \mathcal{Y}_{<k}) \cdot p(\mathcal{Y}_k, \mathbf{s}|\mathbf{s}') \cdot p(\mathcal{Y}_{>k}, \mathbf{s}') \\ &= \alpha_{k-1}(\mathbf{s}') \cdot \gamma_k(\mathbf{s}, \mathbf{s}) \cdot \beta_k(\mathbf{s}), \end{aligned} \tag{C.3}$$

where the probability that the trellis is in state $\mathbf{s}'$ at time k is given as [118]

$$\alpha_k(\mathbf{s}) = \sum_{\text{all } \mathbf{s}'} \gamma_k(\mathbf{s}', \mathbf{s}) \alpha_{k-1}(\mathbf{s}'), \tag{C.4}$$

while the probability that the trellis is in state $\mathbf{s}$ given the future received sequence is given as [118]

$$\beta_{k-1}(\mathbf{s}') = \sum_{\text{all } \mathbf{s}} \gamma_k(\mathbf{s}', \mathbf{s}) \beta_k(\mathbf{s}) \tag{C.5}$$

and finally [118]

$$\gamma_k(\mathbf{s}', \mathbf{s}) = p(\mathcal{Y}_k|\mathbf{s}', \mathbf{s}) P(\mathbf{s}|\mathbf{s}'). \tag{C.6}$$

Furthermore, $p(\mathcal{Y}_k|\mathbf{s}', \mathbf{s})$ may be written as

$$p(\mathcal{Y}|\mathbf{s}', \mathbf{s}) = \frac{1}{(2\pi\sigma_n^2)^N} \exp\left(-\frac{||\mathbf{y}(k) - \check{\mathbf{y}}||^2}{2\sigma_n^2}\right), \tag{C.7}$$

where $\check{\mathbf{y}}$ is the noiseless channel output associated with the transition from state $\mathbf{s}'$ to $\mathbf{s}$. Assuming that no *a priori* knowledge is available, all transitions are equally likely, hence resulting in $P(\mathbf{y}(k)|\mathbf{s}^{(i)}|\mathbf{s}^{(j)}) = \frac{1}{N_t}$. By marginalizing the joint probability, one may readily obtain *a posteriori* bit estimates [118].

Glossary

3D	3-Dimensional
3GPP	3rd Generation Partnership Project
A/D	Analog-to-Digital
AEs	Antenna Elements
AWGN	Additive White Gaussian Noise
BCJR	Bahl-Cocke-Jelinek-Raviv
BEC	Binary Erasure Channel
BEP	Bit Error Probability
BER	Bit Error Rate
BPSK	Binary Phase Shift Keying
BS	Base-Station
CCI	Co-Channel Interference
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CIR	Channel Impulse Response
CSI	Channel State Information
D/A	Digital-to-Analog
DDCE	Decision Directed Channel Estimation
DF	Decision Feedback
DFT	Discrete Fourier Transform
DPSK	Differential Binary Phase Shift Keying
FD	Frequency-Domain
FDCTF	Frequency-Domain Channel Transfer Function
FEC	Forward Error Correction
FIR	Finite Impulse Response
GAs	Genetic Algorithms

GSD	Generalized Sphere Decoder
IIR	Infinite Impulse Response
ISI	Inter-Symbol Interference
LDPC	Low-Density Parity-Check
LLR	Log Likelihood Ratio
LMS	Least Mean Square
LOS	Line-of-Sight
LS	Least Square
LSD	List Sphere Decoder
LTE	Long-Term Evolution
MAP	Maximum A Posteriori
MBER	Minimum Bit Error Rate
MC	Multi-Carrier
MCMC	Markov Chain Monte Carlo
MIMO	Multiple-Input Multiple-Output
ML	Maximum Likelihood
MMSE	Minimum Mean Squared Error
MS	Mobile-Station
MSE	Mean Squared Error
MSER	Minimum Symbol Error Rate
OFDM	Orthogonal Frequency-Division Multiplexing
OHRSA	Optimized Hierarchy Reduced Search Algorithm
PAM	Pulse-Amplitude-Modulation
PAPR	Peak-to-Average Power Ratio
PDF	Probability Density Function
PDP	Power Delay Profile
PIC	Parallel Interference Cancellation
PSD	Power Spectral Density
QAM	Quadrature-Amplitude Modulation
RBFN	Radial Basis Function Network
RF	Radio Frequency
RLS	Recursive Least Squares
RML	Reduced-complexity Max-Log
RSC	Recursive Systematic Convolutional
SC	Single-Carrier
SD	Sphere Decoding

SDM	Space Division Multiplexing
SDMA	Space Division Multiple Access
SIC	Serial Interference Cancellation
SINR	Signal to Interference plus Noise Ratio
SIR$_q$	Signal to Interference Ratio
SISO	Soft-Input Soft-Output
SNR	Signal to Noise Ratio
STC	Space-Time Coding
STE	Space-Time Equalizer
TD	Time-Domain
TDD	Time Division Duplex
V-BLAST	Vertical Bell LAbs Layered Space Time
WSS	Wide Sense Stationary

Index

T

V

W

Bibliography

[1] "The Global Market for High Speed Packet Access (HSPA): Quantitative and Qualitative analysis - Report 39," tech. rep., UMTS Forum, 2006.

[2] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bolcskei, "An overview of MIMO communications - a key to gigabit wireless," *Proceedings of the IEEE*, vol. 92, pp. 198–218, Feb. 2004.

[3] G. Foschini and M. Gans, "On limits of wireless communication in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, March 1998.

[4] C. E. Shannon, "A mathematical theory of communication — PART I," *Bell Systems Technical Journal*, pp. 379–405, 1948.

[5] C. E. Shannon, "A mathematical theory of communication — PART II," *Bell Systems Technical Journal*, pp. 405–423, 1948.

[6] C. E. Shannon, "A mathematical theory of communication — PART III," *Bell Systems Technical Journal*, vol. 27, pp. 623–656, 1948.

[7] H. Ekstrom, A. Furuskar, J. Karlsson, M. Meyer, S. Parkvall, J. Torsner, and M. Wahlqvist, "Technical solutions for the 3G long-term evolution," *IEEE Communications Magazine*, vol. 44, pp. 38–45, Mar. 2006.

[8] A. Ghosh, D. R. Wolter, J. G. Andrews, and R. Chen, "Broadband wireless access with WiMax/802.16: current performance benchmarks and future potential," *IEEE Communications Magazine*, vol. 43, pp. 129–136, Feb. 2005.

[9] J. Gozalvez, "Mobile WiMAX rollouts announced," *IEEE Vehicular Technology Magazine*, vol. 1, pp. 53–59, Sept. 2006.

[10] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 45, pp. 139–157, Jan. 1999.

[11] A. Goldsmith, S. J. N., Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, 2003.

[12] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "MIMO cellular, SNR, Shannon capacity, achievable rate region, antenna array, capacity limits of MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 5, pp. 684–702, 2003.

[13] S. Shamai and T. Marzetta, "Multiuser capacity in block fading with no channel state information," *IEEE Transactions on Information Theory*, vol. 48, no. 4, pp. 938–942, 2002.

[14] E. Biglieri, J. Proakis, and S. Shamai, "Fading channels: information-theoretic and communications aspects," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2619–2692, 1998.

[15] T. S. Rappaport, *Wireless Communications : Principles and Practice*. Prentice Hall, 1996.

[16] Technical Specification Group Radio Access Network, "Physical layer aspects for evolved universal terrestrial radio access (UTRA) (release 7)," tech. rep., 3rd Generation Partnership Project (3GPP), 2006.

[17] L. Hanzo, M. Münster, B. Choi, and T. Keller, *OFDM and MC-CDMA for Broadcasting Multi-User Communications, WLANs and Broadcasting*. John Wiley and IEEE Press, July 2003.

[18] J. A. C. Bingham, "Multicarrier modulation for data transmission: an idea whose time has come," *IEEE Communications Magazine*, vol. 28, pp. 5–14, May 1990.

[19] T. Jiang and G. Zhu, "Complement block coding for reduction in peak-to-average power ratio of OFDM signals," *IEEE Communications Magazine*, vol. 43, pp. 17–22, Sept. 2005.

[20] C.-L. Wang and Y. Ouyang, "Low-complexity selected mapping schemes for peak-to-average power ratio reduction in OFDM systems," *IEEE Transactions on Signal Processing*, vol. 53, pp. 4652–4660, Dec. 2005.

[21] S. Zhou, Y. Li, M. Zhao, X. Xu, J. Wang, and Y. Yao, "Novel techniques to improve downlink multiple access capacity for Beyond 3G," *IEEE Communications Magazine*, vol. 43, pp. 61–69, Jan. 2005.

[22] P. Zhang, X. Tao, J. Zhang, Y. Wang, L. Li, and Y. Wang, "A vision from the future: beyond 3G TDD," *IEEE Communications Magazine*, vol. 43, pp. 38–44, Jan. 2005.

[23] A. Jamalipour, T. Wada, and T. Yamazato, "A tutorial on multiple access technologies for beyond 3G mobile networks," *IEEE Communications Magazine*, vol. 43, pp. 110–117, Feb 2005.

[24] G. Wetzker, "Definition of spatial multiplexing gain," *IEE Electronics Letters*, vol. 41, pp. 656–657, May 2005.

[25] D. Tse, P. Viswanath, and L. Zheng, "Diversity-multiplexing tradeoff in multiple-access channels," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 1859–1874, 2004.

[26] S. X. Ng and L. Hanzo, "On the MIMO channel capacity of multidimensional signal sets," *IEEE Transactions on Vehicular Technology*, vol. 55, pp. 528–536, Mar. 2006.

[27] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, pp. 1451–1458, Oct. 1998.

[28] J. Proakis, *Digital communications*. McGraw-Hill, 1995.

[29] X. Wang and H. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Transactions on Wireless Communications*, vol. 47, no. 7, pp. 1046–1061, 1999.

[30] S. Chen, "Adaptive minimum bit-error-rate filtering," *IEE Proceedings of Vision, Image and Signal Processing*, vol. 151, no. 1, pp. 76–85, 2004.

[31] S. Chen, N. Ahmad, and L. Hanzo, "Adaptive minimum bit-error rate beamforming," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 341–348, Mar. 2005.

[32] D. Gesbert, "Robust linear MIMO receivers: a minimum error-rate approach," *IEEE Transactions on Signal Processing*, vol. 51, no. 11, pp. 2863–2871, 2003.

[33] S. Chen, B. Mulgrew, and L. Hanzo, "Least bit error rate adaptive nonlinear equalisers for binary signalling," *IEE Proceedings on Communications*, vol. 150, pp. 29–36, Feb. 2003.

[34] S. Chen, L. Hanzo, N. Ahmad, and A. Wolfgang, "Adaptive minimum bit error rate beamforming assisted receiver for QPSK wireless communication," *Digital Signal Processing*, vol. 15, pp. 545–567, 2004.

[35] C.-C. Yeh and J. Barry, "Adaptive minimum symbol-error rate equalization for quadrature-amplitude modulation," *IEEE Transactions on Signal Processing*, vol. 51, no. 12, pp. 3263–3269, 2003.

[36] S. Chen, H. Q. Du, and L. Hanzo, "Adaptive minimum symbol error rate beamforming assisted receiver for quadrature amplitude modulated systems," in *Proccedings of IEEE 63rd Vehicular Technology Conference VTC*, vol. 5, pp. 2236–2240, Spring 2006.

[37] J. Li, G. Wei, and F. Chen, "On minimum-BER linear multiuser detection for DS-CDMA channels," *IEEE Transactions on Wireless Communications*, vol. 55, pp. 1093–1103, Mar. 2007.

[38] N. Wang and S. D. Blostein, "Approximate Minimum BER Power Allocation for MIMO Spatial Multiplexing Systems," *IEEE Transactions on Communications*, vol. 55, pp. 180–187, Jan. 2007.

[39] S. M. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*. Prentice Hall, New Jersey, 1993.

[40] U. Fincke and M. Pohst, "Improved method for calculating vector of short length in a lattice, including a complexity analysis," *Math. Comput.*, vol. 44, pp. 463–471, April 1985.

[41] E. Viterbo and J. Bouros, "A universal lattice code decoder for fading channels," *IEEE Transactions on Information Theory*, vol. 45, no. 5, pp. 1639–1642, 1999.

[42] L. Brunel and J. Boutros, "Lattice decoding for joint detection in direct-sequence CDMA systems," *IEEE Transactions on Information Theory*, vol. 49, no. 4, pp. 1030–1037, 2003.

[43] B. Hassibi and H. Vikalo, "On the sphere-decoding algorithm I. Expected complexity," *IEEE Transactions on Signal Processing*, vol. 53, pp. 2806–2818, 2005.

[44] M. Damen, K. Abed-Meraim, and J.-C. Belfiore, "Generalised sphere decoder for asymmetrical space-time communication architecture," *Electronics Letters*, vol. 36, no. 2, pp. 166–167, 2000.

[45] D. Pham, K. R. Pattipati, P. K. Willet, and J. Luo, "An improved complex sphere decoder for V-BLAST Systems," *IEEE Signal Processing Letters*, vol. 11, pp. 748–751, September 2004.

[46] Z. Yang, C. Liu, and J. He, "A new approach for fast generalized sphere decoding in MIMO systems," *IEEE Signal Processing Letters*, vol. 12, no. 1, pp. 41–44, 2005.

[47] Z. Guo and P. Nilsson, "Algorithm and implementation of the K-best sphere decoding for MIMO detection," *IEEE Journal on Selected Areas in Communications*, vol. 24, pp. 491–503, March 2006.

[48] B. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Transactions on Communications*, vol. 51, no. 3, pp. 389–399, 2003.

[49] H. Vikalo, B. Hassibi, and T. Kailath, "Iterative decoding for MIMO channels via modified sphere decoding," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 2299–2311, 2004.

[50] R. Wang and G. Giannakis, "Approaching MIMO channel capacity with soft detection based on hard sphere decoding," *IEEE Transactions on Communications*, vol. 54, no. 4, pp. 587–590, 2006.

[51] V. Pauli, L. Lampe, and R. Schober, "Turbo DPSK using soft multiple-symbol differential sphere decoding," *IEEE Transactions on Information Theory*, vol. 52, pp. 1385–1398, Apr. 2006.

[52] S. Thoen, L. Deneire, L. V. der Perre, M. Engels, and H. D. Man, "Constrained least squares detector for OFDM/SDMA-based wireless networks," *IEEE Transactions on Wireless Communications*, vol. 2, pp. 129–140, Jan. 2003.

[53] G. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. Autumn, pp. 41–59, 1996.

[54] H. Lee, B. Lee, and I. Lee, "Iterative detection and decoding with an improved V-BLAST for MIMO-OFDM systems," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 504–513, 2006.

[55] R. Gallager, "Low-density parity-check codes," *IEEE Transactions on Information Theory*, vol. 8, pp. 21–28, Jan. 1962.

[56] R. S. Mozos and M. J. F.-G. Garcia, "Efficient complex sphere decoding for MC-CDMA systems," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 2992–2996, Nov. 2006.

[57] H. Zhu, B. Farhang-Boroujeny, and R.-R. Chen, "On performance of sphere decoding and Markov chain Monte Carlo detection methods," *IEEE Signal Processing Letters*, vol. 12, pp. 669–672, Oct. 2005.

[58] B. Farhang-Boroujeny, H. Zhu, and Z. Shi, "Markov chain Monte Carlo algorithms for CDMA and MIMO communication systems," *IEEE Transactions on Signal Processing*, vol. 54, no. 5, pp. 1896–1909, 2006.

[59] P. Aggarwal and X. Wang, "Multilevel sequential Monte Carlo algorithms for MIMO demodulation," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 750–758, Feb. 2007.

[60] J. Akhtman and L. Hanzo, "Optimized-Hierarchy RSA-Aided Space-Time Processing," in *OFDM and MC-CDMA: A Primer by L. Hanzo and T. Keller*, pp. 259–292, New York, USA: John Wiley, 2005. (URL: http://www-mobile.ecs.soton.ac.uk/comms/Books.htm).

[61] J. Akhtman and L. Hanzo, "Reduced-complexity maximum-likelihood detection in multiple-antenna-aided multicarrier systems," in *Proceedings of 5th Int'l Workshop on Multi-Carrier Spread Spectrum (in press)*, (Oberpfaffenhofen, Germany), 2005.

[62] P. A. Hoeher, S. Badri-Hoeher, W. Xu, and C. Krakowski, "Single-antenna co-channel interference cancellation for TDMA cellular radio systems," *IEEE Wireless Communications*, vol. 12, pp. 30–37, Apr. 2005.

[63] A. Zanella, M. Chiani, and M. Win, "MMSE reception and successive interference cancellation for MIMO systems with high spectral efficiency," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 1244–1253, May 2005.

[64] S. W. Kim and K. P. Kim, "Log-likelihood-ratio-based detection ordering in V-BLAST," *IEEE Transactions on Communications*, vol. 54, no. 2, pp. 302–307, 2006.

[65] M. Ghotbi and M. Soleymani, "Multiuser detection of DS-CDMA signals using partial parallel interference cancellation in satellite communications," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 3, pp. 584–593, 2004.

[66] Y. Cho and J. Lee, "Analysis of an adaptive SIC for near-far resistant DS-CDMA," *IEEE Transactions on Communications*, vol. 46, pp. 1429–1432, November 1998.

[67] A. Wolfgang, S. Chen, and L. Hanzo, "Parallel interference cancellation based turbo space-time equalization in the SDMA uplink," *IEEE Transactions on Wireless Communications*, vol. 6, pp. pp. 609–616., 2007.

[68] A. Wolfgang, S. Chen, and L. Hanzo, "Decision feedback aided Bayesian turbo space-time equalizer for parallel interference cancellation in SDMA systems," in *Proceedings of IEEE Vehicular Technology Conference*, vol. 4, pp. 2112–2116, 2005.

[69] A. Doucet and X. Wang, "Monte Carlo methods for signal processing: a review in the statistical signal processing context," *IEEE Signal Processing Magazine*, vol. 22, no. 6, pp. 152–170, 2005.

[70] C. Ergun and K. Hacioglu, "Multiuser detection using a genetic algorithm in CDMA communications systems," *IEEE Transactions on Communications*, vol. 48, no. 8, pp. 1374–1383, 2000.

[71] S. Chen, L. Hanzo, and A. Wolfgang, "Kernel-based nonlinear beamforming construction using orthogonal forward selection with Fisher ratio class separability measure.," *IEEE Signal Processing Letters*, vol. 11, pp. 478–481, 2004.

[72] S. Chen, A. Wolfgang, and L. Hanzo, "A robust nonlinear beamforming assisted receiver for BPSK signalling," in *Proceedings of IEEE Vehicular Technology Conference*, vol. 3, pp. 1921–1925, 2005.

[73] N. Al-Dhahir and A. Sayed, "The finite-length multi-input multi-output MMSE-DFE," *IEEE Transactions on Signal Processing*, vol. 48, pp. 2921–2936, Oct. 2000.

[74] R. Merched and N. Yousef, "Fast techniques for computing finite-length MIMO MMSE decision feedback equalizers," *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 701–711, 2006.

[75] C. Tidestav, M. Sternad, and A. Ahlen, "Reuse within a cell-interference rejection or multiuser detection?," *IEEE Transactions on Communications*, vol. 47, pp. 1511–1522, Oct. 1999.

[76] T. Abe and T. Matsumoto, "Space-time turbo equalization in frequency-selective MIMO channels," *IEEE Transactions on Vehicular Technology*, vol. 52, no. 3, pp. 469–475, 2003.

[77] S. Chen, A. Livingstone, and L. Hanzo, "Minimum bit-error rate design for spacetime equalization-based multiuser detection," *IEEE Transactions on Communications*, vol. 54, no. 5, pp. 824–832, 2006.

[78] A. M. Tonello, "MIMO MAP equalization and turbo decoding in interleaved space-time coded systems," *IEEE Transactions on Communications*, vol. 51, pp. 155–160, Feb. 2003.

[79] H. Vikalo, B. Hassibi, and U. Mitra, "Sphere-constrained ML detection for frequency-selective channels," *IEEE Transactions on Communications*, vol. 54, pp. 1179–1183, July 2006.

[80] A. Viterbi, "Optimum detection and signal selection for partially coherent binary communication," *IEEE Transactions on Information Theory*, vol. 11, pp. 239–246, Apr. 1965.

[81] G. Forney, "The Viterbi algorithm," *Proceedings of the IEEE*, vol. 61, pp. 268–278, March 1973.

[82] D. Falconer, S. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Communications Magazine*, vol. 40, pp. 58–66, April 2002.

[83] J. Louveaux, L. Vandendorpe, and T. Sartenaer, "Cyclic prefixed single carrier and multicarrier transmission: bit rate comparison," *IEEE Communications Letters*, vol. 7, no. 4, pp. 180–182, 2003.

[84] X. Li and X. Cui, "Application of lattice code decoder to SC-CP for short block length," *IEE Electronics Letters*, vol. 40, no. 15, pp. 954–955, 2004.

[85] M. Tuchler, R. Koetter, and A. Singer, "Turbo equalization: principles and new results," *IEEE Transactions on Communications*, vol. 50, no. 5, pp. 754–767, 2002.

[86] B. Dong and X. Wang, "Sampling-based soft equalization for frequency-selective MIMO channels," *IEEE Transactions on Communications*, vol. 53, pp. 278–288, Feb. 2005.

[87] G. Dietl and W. Utschick, "Complexity reduction of iterative receivers using low-rank equalization," *IEEE Transactions on Signal Processing*, vol. 55, pp. 1035–1046, Mar. 2007.

[88] M. W. Kwan and C. W. Kok, "MMSE Equalizer for MIMO-ISI Channel With Shorten Guard Period," *IEEE Transactions on Signal Processing*, vol. 55, pp. 389–395, Jan. 2007.

[89] H. Duan, B. P. Ng, C. M. S. See, and J. Fang, "Broadband beamforming using TDL-form IIR filters," *IEEE Transactions on Signal Processing*, vol. 55, pp. 990–1002, Mar. 2007.

[90] X. Li and T. F. Wong, "Turbo equalization with nonlinear Kalman filtering for time-varying frequency-selective fading channels," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 691–700, Feb. 2007.

[91] S. Roy and T. M. Duman, "Soft input soft output Kalman equalizer for MIMO frequency selective fading channels," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 506–514, Feb. 2007.

[92] N. Benvenuto, F. Boccardi, and G. Carnevale, "Frequency domain realization of space-time receivers in dispersive wireless channels," *IEEE Transactions on Signal Processing*, vol. 55, pp. 313–326, Jan. 2007.

[93] B. Ng, C.-T. Lam, and D. Falconer, "Turbo frequency domain equalization for single-carrier broadband wireless systems," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 759–767, Feb. 2007.

[94] A. Wolfgang, N. Ahmad, S. Chen, and L. Hanzo, "Genetic algorithm assisted error probability optimisation for beamforming," *IEE Electronics Letters*, vol. 40, no. 5, pp. 320–322, 2004.

[95] A. Wolfgang, N. Ahmad, S. Chen, and L. Hanzo, "Genetic algorithm assisted minimum bit error rate beamforming," in *Proccedings of IEEE Vehicular Technology Conference, 2004.*, vol. 1, pp. 142–146, 2004.

[96] S. Chen, L. Hanzo, N. Ahmad, and A. Wolfgang, "Adaptive minimum bit error rate beamforming assisted QPSK receiver," in *Proceedings of IEEE International Conference on Communications*, vol. 6, pp. 3389–3393, 2004.

[97] J. Akhtman, A. Wolfgang, S. Chen, and L. Hanzo, "An optimized-hierarchy-aided approximate Log-MAP detector for MIMO systems." To appear in IEEE Transactions on Wireless Communications.

[98] A. Wolfgang, S. Chen, and L. Hanzo, "Radial basis function aided space-time equalization in dispersive fading uplink environments," in *Proceedings of IEEE Vehicular Technology Conference*, vol. 3, pp. 1552–1556, 2005.

[99] A. Wolfgang, S. Chen, and L. Hanzo, "Radial basis function network assisted space-time equalisation for dispersive fading environments," *IEE Electronics Letters*, vol. 40, no. 16, pp. 1006–1007, 2004.

[100] A. Wolfgang, S. Chen, and L. Hanzo, "Radial basis function network assisted wide-band beamforming," in *Proceedings of IEEE Vehicular Technology Conference*, pp. 266–270, Sept. 2004.

[101] A. Wolfgang, J. Akhtman, S. Chen, and L. Hanzo, "Reduced complexity single-carrier maximum-likelihood detection for decision feedback assisted space-time equalization," in *Proceedings of IEEE WCNC'06, Las Vegas, Nevada, USA.*, vol. 1, pp. 142–146, 2006.

[102] A. Wolfgang, J. Akhtman, S. Chen, and L. Hanzo, "Reduced complexity near maximum-likelihood uplink detection for decision feedback assisted space-time equalization." To appear in IEEE Transactions on Wireless Communications.

[103] J. Wang, S. X. Ng, A. Wolfgang, L. L. Yang, S. Chen, and L. Hanzo, "Near-capacity three-stage MMSE turbo equalization using irregular convolutional codes," in *Proceedings of The 4th International Symposium on Turbo Codes*, (Munich, Germany), April 2006.

[104] A. Wolfgang, J. Akhtman, S. Chen, and L. Hanzo, "Iterative MIMO detection for rank-deficient systems," *IEEE Signal Processing Letters*, vol. 13, pp. 699–702, Nov. 2006.

[105] H. Asplund, A. A. Glazunov, A. F. Molisch, K. I. Pedersen, and M. Steinbauer, "The COST 259 directional channel model: Part II: Macrocells," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 3434–3450, Dec. 2006.

[106] T. Nagell, *Lattice Points and Point Lattices*, ch. 11, pp. 32–34. Introduction to Number Theory, New York: Wiley, 1951.

[107] H. Vikalo and B. Hassibi, "On the sphere-decoding algorithm II. Generalizations, second-order statistics, and applications to communications," *IEEE Transactions on Signal Processing*, vol. 53, pp. 2819–2834, 2005.

[108] J. Blogh and L. Hanzo, *Third-Generation Systems and Intelligent Networking*. John Wiley and IEEE Press, 2002. (For detailed contents, please refer to http://www-mobile.ecs.soton.ac.uk.).

[109] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, June and October 1948.

[110] T. M. Cover and J. A. Thomas., *Elements of Information Theory*. John Wiley and Sons, 1991.

[111] E. Baccarelli, "Evaluation of the reliable data rates supported by multiple-antenna coded wireless links for QAM transmissions," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 295–304, Feb. 2001.

[112] H. Zhu, B. Farhang-Boroueny, Z. Shi, and C. Schlegel, "An efficient statistical approach for calculation of capacity of mimo channels," in *Proceedings of 3rd IASTED Int. Con. On Wireless and Optical Communication*, 2002.

[113] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1996.

[114] S. Chen, S. Tan, and L. Hanzo, "Linear beamforming assisted receiver for binary phase shift keying modulation systems," in *Proceedings of IEEE WCNC'06*, 2006. Las Vegas, Nevada, USA.

[115] L. Hanzo, W. Webb, and T. Keller, *Single- and Multi-carrier Quadrature Amplitude Modulation*. New York, USA: John Wiley and IEEE Press, April 2000.

[116] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, Massachusetts: Addison-Wesley, 1989.

[117] S. Chen, A. Wolfgang, and L. Hanzo, "Constant modulus algorithm aided soft decision-directed blind space-time equalization for SIMO channels," in *Proceedings of IEEE Vehicular Technology Conference*, vol. 3, pp. 1718–1722, 2004.

[118] L. Hanzo, C. H. Wong, and M. S. Yee, *Adaptive Wireless Transceivers: Turbo-Coded, Turbo-Equalized and Space-Time Coded TDMA, CDMA, and OFDM Systems*. John Wiley and IEEE Press, Feb. 2002.

[119] M. S. Yee, T. H. Liew, and L. Hanzo, "Burst-by-burst adaptive turbo-coded radial basis function-assisted decision feedback equalization," *IEEE Transactions on Communications*, vol. 49, pp. 1935–1945, Nov. 2001.

[120] M. S. Yee and L. Hanzo, "A wide-band radial basis function decision feedback equalizer-assisted burst-by-burst adaptive modem," *IEEE Transactions on Communications*, vol. 50, pp. 693–697, May 2002.

[121] A. Sayed and B. Hassibi, *Linear Estimation*. Prentice-Hall, 2000.

[122] C. Komninakis, C. Fragouli, A. Sayed, and R. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1065–1076, May 2002.

[123] S. A. Yang and J. Wu, "Optimal binary training sequence design for multiple-antenna systems over dispersive fading channels," *IEEE Transactions on Vehicular Technology*, vol. 51, pp. 1271–1276, Sept. 2002.

[124] H. E. Gamal, A. Hammons, Y. Liu, M. Fitz, and O. Takeshita, "On the design of space-time and space-frequency codes for MIMO frequency-selective fading channels," *IEEE Transactions on Information Theory*, vol. 49, no. 9, pp. 2277–2292, 2003.

[125] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. IT-20, pp. 284–287, Mar. 1974.

[126] Z. Zhang and T. M. Duman, "Achievable information rates of multi-antenna systems over frequency-selective fading channels with constrained inputs," *IEEE Communications Letters*, vol. 7, no. 6, pp. 260–262, 2003.

[127] X. Wei, D. G. M. Cruickshank, B. Mulgrew, and F. Riera-Palou, "A unified approach to dynamic length algorithms for adaptive linear equalizers," *IEEE Transactions on Signal Processing*, vol. 55, pp. 908–920, Mar. 2007.

[128] E. Chng and S. Chen, "Determining the optimal decision delay parameter for the linear equalizer," *International Journal Automation and Computing*, vol. 1, pp. 25–29, 2005.

[129] S. Chen, B. Mulgrew, E.-S. Chng, and G. J. Gibson, "Space translation properties and the minimum-BER linear-combiner DFE," *IEE Proceedings on Communications*, vol. 145, pp. 316–322, October 1998.

[130] M. Alias, S. Chen, and L. Hanzo, "Multiple-antenna-aided OFDM employing genetic-algorithm-assisted minimum bit error rate multiuser detection," *IEEE Transactions on Vehicular Technology*, vol. 54, no. 5, pp. 1713–1721, 2005.

[131] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo-codes," *IEEE Transactions on Communications*, vol. 44, no. 10, pp. 1261–1271, 1996.

[132] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and decoding: Turbo-codes. 1," in *IEEE International Conference on Communications*, vol. 2, (Geneva), pp. 1064–1070, 1993.

[133] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Analysis, design, and iterative decoding of double serially concatenated codes with interleavers," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 2, pp. 231–244, 1998.

[134] C. Douillard, A. Picart, M. Jézéquel, P. Didier, C. Berrou, and A. Glavieux, "Iterative correction of inter-symbol interference: Turbo-equalization," *European Transactions on Communications*, vol. 6, pp. 507–511, September-October 1995.

[135] J. Li, K. Letaief, and Z. Cao, "Space-time turbo multiuser detection for coded mc-cdma," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 538–549, March 2005.

[136] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Transactions on Information Theory*, vol. 42, no. 2, pp. 429–445, 1996.

[137] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Transactions on Communications*, vol. 49, no. 10, pp. 1727–1737, 2001.

[138] S. Ng, J. C. L., and Hanzo, "Turbo-detected unequal protection mpeg-4 wireless video telephony using multi-level coding, trellis coded modulation and space-time trellis coding," in *IEE Proceedings on Communications*, vol. 152, pp. 1116–1124, 9 Dec. 2005.

[139] M. Tuchler, A. Singer, and R. Koetter, "Minimum mean squared error equalization using a priori information," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 673–683, 2002.

[140] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: model and erasure channel properties," *Information Theory, IEEE Transactions on*, vol. 50, pp. 2657–2673, Nov. 2004.

[141] S. Song, A. Singer, and K.-M. Sung, "Turbo equalization with an unknown channel," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP).*, vol. 3, 2002.

[142] N. Metropolis and S. Ulam, "The Monte Carlo method," *Journal of the American Statistical Association*, vol. 44, p. 335, 1949.

[143] J. Spall, "Estimation via Markov chain Monte Carlo," *IEEE Control Systems Magazine*, vol. 23, pp. 34–45, Apr. 2003.

[144] B. Lu and X. Wang, "Bayesian blind turbo receiver for coded OFDM systems with frequency offset and frequency-selective fading," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 2516–2527, Dec. 2001.

[145] Z. Yang, B. Lu, and X. Wang, "Bayesian Monte Carlo multiuser receiver for space-time coded multicarrier CDMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 1625–1637, Aug. 2001.

[146] W. Ng, J. Reilly, T. Kirubarajan, and J.-R. Larocque, "Wideband array signal processing using MCMC methods," *IEEE Transactions on Signal Processing*, vol. 53, pp. 411–426, Feb. 2005.

[147] R. Chen, J. Liu, and X. Wang, "Convergence analyses and comparisons of Markov chain Monte Carlo algorithms in digital communications," *IEEE Transactions on Signal Processing,*, vol. 50, pp. 255–270, Feb. 2002.

[148] M. Jiang, J. Akhtman, F. Guo, and L. Hanzo, "Iterative joint channel estimation and symbol detection for multi-user MIMO OFDM," in *Proceedings of IEEE Vehiclular Technology Conference*, (Melbourne, Australia), pp. 221–225, Spring 2006.

[149] J. R. Norris, *Markov Chains*. Cambridge University Press, 1997.

[150] G. Casella and E. I. George, "Explaining the Gibbs Sampler," *The American Statistician*, vol. 46, pp. 167–174, 1992.

[151] F. Gray, "Pulse code communications." U. S. Patent 2632058, March 1953.

Author Index

I

J

K

L

M

N