Helsinki University of Technology Signal Processing Laboratory
Teknillinen korkeakoulu Signaalinkäsittelytekniikan laboratorio
Espoo 2003 Report 43

ITERATIVE RECEIVERS AND MULTICHANNEL EQUALISATION FOR TIME DIVISION MULTIPLE ACCESS SYSTEMS

Markku Pukkila

Dissertation for the degree of Doctor of Science in Technology to be presented with due permission for public examination and debate in Auditorium S1 at Helsinki University of Technology (Espoo, Finland) on the 10th of October, 2003, at 12 o'clock noon.

Helsinki University of Technology
Department of Electrical and Communications Engineering
Signal Processing Laboratory

Teknillinen korkeakoulu
Sähkö- ja tietoliikennetekniikan osasto
Signaalinkäsittelytekniikan laboratorio

Distributor:
Helsinki University of Technology
Signal Processing Laboratory
P.O. Box 3000
FIN-02015 HUT
Tel. +358-9-451 3211
Fax. +358-9-452 3614
Email: Mirja.Lemetyinen@hut.fi

ISBN 951-22-6716-0
ISSN 1458-6401

Otamedia Oy
Espoo 2003

ABSTRACT

The thesis introduces receiver algorithms improving the performance of TDMA mobile radio systems. Particularly, we consider receivers utilising side information, which can be obtained from the error control coding or by having a priori knowledge of interference sources. Iterative methods can be applied in the former case and interference suppression techniques in the latter.

Convolutional coding adds redundant information into the signal and thereby protects messages transmitted over a radio channel. In the coded systems the receiver is usually comprised of separate channel estimation, detection and channel decoding tasks due to complexity restrictions. This suboptimal solution suffers from performance degradation compared to the optimal solution achieved by optimising the joint probability of information bits, transmitted symbols and channel impulse response. Conventional receiver utilises estimated channel state information in the detection and detected symbols in the channel decoding to finally obtain information bits. However, the channel decoder provides also extrinsic information on the bit probabilities, which is independent of the received information at the equaliser input. Therefore it is beneficial to re-perform channel estimation and detection using this new extrinsic information together with the original input signal.

We apply iterative receiver techniques mainly to Enhanced General Packet Radio System (EGPRS) using GMSK modulation for iterative channel estimation and 8-PSK modulation for iterative detection scheme. Typical gain for iterative detection is around 2 dB and for iterative channel estimation around 1 dB. Furthermore, we suggest two iteration rounds as a reasonable complexity/performance trade-off. To obtain further complexity reduction we introduce the soft trellis decoding technique that reduces the decoder complexity significantly in the iterative schemes.

Cochannel interference (CCI) originates from the nearby cells that are reusing the same transmission frequency. In this thesis we consider CCI suppression by joint detection (JD) technique, which detects simultaneously desired and interfering signals. Because of the complexity limitations we only consider JD for two binary modulated signals. Therefore it is important to find the dominant interfering signal (DI) to achieve the best performance. In the presence of one strong DI, the JD provides major improvement in the receiver performance.

The JD requires joint channel estimation (JCE) for the two signals. However, the JCE makes the implementation of the JD more difficult, since it requires synchronised network and unique training sequences with low cross-correlation for the two signals.

PREFACE

The work for this thesis has been done in the Radio Communications Laboratory of Nokia Research Center and the Signal Processing Laboratory of Helsinki University of Technology.

I wish to thank my supervisor Prof. Timo Laakso, who has guided my post-graduate studies and given very good advice and suggestions for my thesis. Also I like to express my gratitude for Prof. Iiro Hartimo for giving me an opportunity to finalise my thesis in the Signal Processing Laboratory. The pre-examiners, Prof. Tony Ottosson and Prof. Tadashi Matsumoto, are gratefully acknowledged for their efforts to review the thesis.

I am very grateful for my current employer Nokia Research Center, which has supported my post-graduate studies in many ways. I thank the Head of Radio Communications Laboratory Mr. Jukka Soikkeli, the former Head of Laboratory Dr. Steven Gray, the former Head of Laboratory Dr. Pekka Soininen and the former Head of Laboratory Mr. Heikki Huomo for the opportunity to make my thesis while I have been working for Nokia. I like to thank also my superiors Dr. Giridhar Mandyam and Mr. Eero Nikula for their support for my post-graduate studies.

I wish to sincerely thank my very good colleague Dr. Pekka Ranta, who has guided me from the beginning of my Nokia career. He has been an excellent teacher for me and showed me an encouraging example by preparing his own Ph.D. thesis while working. Finally, he has shared his deep technical understanding and given me several suggestions to improve my thesis. I also like to thank my colleague Dr. Nikolai Nefedov, from who I have learned a lot of theoretical and practical solutions to communications problems. Pekka and Nikolai have a significant role to compose this thesis as they have co-authored several journal and conference publications with me. I am grateful for the other co-authors, Dr. Raphael Visoz, Dr. Antoine Berthet, Dr. Jan Olivier and Mr. Markku Heikkilä, for their important contributions. Furthermore, I like to express my gratitude to Dr. Gian Paolo Mattellini and Dr. Kimmo Kettunen for giving valuable comments and suggestions for this thesis.

I like to thank my jolly good colleagues Vuokko and Peter for sharing many pleasant and funny moments outside the working hours. I also thank my parents, sister and brother for their support during the past few years. Finally, I wish to thank my dear friend Annika for her encouragement and bringing joy and happiness to my life.

The financial support of Nokia Foundation is thankfully acknowledged.

Helsinki, October 2003

Markku Pukkila

TABLE OF CONTENTS

LIST OF PUBLICATIONS

[P1] Nikolai Nefedov, Markku Pukkila, Raphael Visoz, and Antoine Berthet, ”Iterative receiver concept for TDMA packet data systems,” *European Trans. Telecommun.*, accepted as a regular paper for publication in Jan. 2003.

[P2] Nikolai Nefedov, Markku Pukkila, Raphael Visoz, and Antoine Berthet, “Iterative data detection and channel estimation for advanced TDMA systems,” *IEEE Trans. Commun.*, vol. 51, no. 2, pp. 141-144, Feb. 2003.

[P3] Markku Pukkila, “Turbo equalisation for the enhanced GPRS system,” in *Proc. 11th IEEE Int. Symp. Personal, Indoor and Mobile Radio Commun.*, London, UK, Sept. 2000, pp. 893-897.

[P4] Nikolai Nefedov and Markku Pukkila, “Iterative channel estimation for GPRS,” in *Proc. 11th IEEE Int. Symp. Personal, Indoor and Mobile Radio Commun.*, London, UK, Sept. 2000, pp. 999-1003.

[P5] Nikolai Nefedov and Markku Pukkila, “Turbo equalization and iterative (turbo) estimation techniques for packet data transmission,” in *Proc. 2nd Int. Symp. Turbo Codes*, Brest, France, Sept. 2000, pp. 423-426.

[P6] Markku Pukkila and Jan C. Olivier, “Turbo equalization with low complexity decoder,” in *Proc. 54th IEEE Vehicular Technology Conf.*, Atlantic City, NJ, Oct. 2001, vol. 2, pp. 1048-1052.

[P7] Markku Pukkila, Markku J. Heikkilä, and Pekka A. Ranta, “Space-time trellis coding with turbo equalisation for the EGPRS system,” in *Proc. 12th IEEE Int. Symp. Personal, Indoor and Mobile Radio Commun.*, San Diego, CA, Sept.-Oct. 2001, vol. 1, pp. 129-133.

[P8] Pekka A. Ranta and Markku Pukkila, “Interference suppression by joint demodulation of cochannel signals,” in *GSM: Evolution Towards 3rd Generation Systems*. Z. Zvonar, P. Jung, and K. Kammerlander, Ed. Dordrecht, Netherlands: Kluwer Academic Publishers, 1999, pp. 153-186.

[P9] Markku Pukkila and Pekka A. Ranta, “Channel estimator for multiple co-channel demodulation in TDMA mobile systems,” in *Proc. 2nd European Personal Mobile Commun. Conf.*, Sept. 1997, pp. 327-333.

ABBREVIATIONS

ACF	Autocorrelation function
A/D	Analog-to-digital
APM	Additive path metric
APP	*A posteriori* probability
AWGN	Additive white Gaussian noise
BCJR	Bahl, Cocke, Jelinek and Raviv
BEP	Bit error probability
CCI	Cochannel interference
C/I	Carrier-to-interference
CMF	Channel matched filter
CRLB	Cramer-Rao lower bound
CS	Channel sounding
DD	Delay diversity
DFE	Decision feedback equaliser
DFSE	Decision feedback sequence estimation
DI	Dominant interfering signal
DIR	Dominant interference ratio
EDGE	Enhanced data rates for global evolution
(E)GPRS	(Enhanced) general packet radio system
EM	Expectation-maximisation
GMSK	Gaussian minimum shift keying
GSM	Global system for mobile communications
ICE	Iterative channel estimation
ISI	Intersymbol interference
JCE	Joint channel estimation
JD	Joint detection
LLR	Log-likelihood ratio
(L)MMSE	(Linear) minimum mean square error
LMS	Least mean squares

LPF	Low-pass filter
LS	Least Squares
MAP	Maximum a posteriori
MCS	Modulation and coding scheme
ML	Maximum-likelihood
MLSE	Maximum-likelihood sequence estimation
MSE	Mean squared error
MVU	Minimum variance unbiased
OSR	Optimum sequential receiver
PCE	Pairwise channel estimation
PSK	Phase shift keying
RLS	Recursive least squares
RSSE	Reduced state sequence estimation
SC	Soft cancelling
SCCC	Serially concatenated convolutional codes
SISO	Soft-In/Soft-Out
SN(I)R	Signal-to-noise (and interference) ratio
SOVA	Soft output Viterbi algorithm
SOVE	Soft output Viterbi equaliser
ST	Space-time
STD	Soft trellis decoding
STTC	Space-time trellis coding
SQRC	Square root raised cosine
TDMA	Time division multiple access
TE	Turbo equalisation
VA	Viterbi algorithm

SYMBOLS

$\mathbf{A}$	Burst of data symbols
$\hat{\mathbf{a}}$	Sequence of estimated data symbols
$\mathbf{a}^{(n)}$	Data symbols from n^{th} cochannel source
$\mathbf{a}_k^{(s)}$	Candidate symbol vector for trellis state s
$\tilde{a}_k$	Soft estimate of data symbol a_k
α	M-ary symbol value
α	Rolloff factor of the raised cosine filter
$\alpha_k(s)$	State probability in forward recursion
$\beta_k(s)$	State probability in backward recursion
$b_{k,j}$	j^{th} bit of data symbol
β_j	Bit value for j^{th} bit of data symbol
$\mathbf{c}$	Sequence of encoded data bits
$\hat{\mathbf{c}}$	Estimated coded data bits
C	Power of desired signal
$\mathbf{C}_{\hat{h}}$	Covariance of channel estimator
$\mathbf{C}_{hy}$	Cross-covariance between $\mathbf{h}$ and $\mathbf{y}$
$\mathbf{C}_w$	Matrix of noise covariance
$\mathbf{C}_y$	Covariance of observations $\mathbf{y}$
$\gamma_k(\cdot)$	Additive branch (transition) metric
$\gamma^a(\cdot)$	Transition metric with *a priori* information
$\gamma^{ext}(\cdot)$	Extrinsic transition metric
Γ	Convolutional encoder
$\Gamma_k(\cdot)$	Additive path metric
$\mathbf{d}^{(n)}$	ST coded symbol sequence from antenna n
D	Decision delay
$D(\cdot)$	Euclidean distance metric
δ	Probability threshold in soft trellis decoding
$\delta(t)$	Dirac's delta function
$E[\cdot]$	Expected value

$\mathbf{\Phi}_k$	Correlation matrix
$g_{ch}(t)$	Impulse response of physical channel
$g_{rx}(t)$	Impulse response of receive filter
$g_{tx}(t)$	Impulse response of transmit filter
$G(f)$	Frequency response of a filter
$\mathbf{h}$	Symbol-spaced channel impulse response
$\hat{\mathbf{h}}$	Estimate of channel impulse response
$\hat{\mathbf{h}}^{(n)}$	Estimated channel for signal n
$\hat{\mathbf{h}}^{(s)}$	Estimated channel for trellis state s
$\mathbf{h}'$	Channel impulse responses of currently estimated cochannels
$\mathbf{h}''$	Channel impulse responses of unestimated cochannels
$\mathbf{H}$	Channel matrix
$\mathbf{I}$	Identity matrix
J_n	Number of subsets in set partitioning
K	Length of burst
K_0	Number of information bits in data block
L	Length of channel memory
λ	Forgetting factor of the RLS algorithm
$\lambda^a(a_k)$	*a priori* information for a_k
$\lambda_{eq}(a_k)$	Soft equaliser output for a_k
$\lambda_{eq}^{ext}(a_k)$	Extrinsic information from equaliser
$\lambda_d^{ext}(c_k)$	Extrinsic information from decoder
M	Number of modulation levels
$\mathbf{M}$	Matrix of midamble symbols
$\mathbf{M}^{(n)}$	Matrix of midamble symbols of n^{th} signal
$\mathbf{M}'$	Midamble matrix of currently estimated cochannels
$\mathbf{M}''$	Midamble matrix of unestimated cochannels
μ	Step size in LMS algorithm
N	Number of transmit antennas
N	Number of cochannel signals
N_0	Two-sided power spectral density of noise
N_0	Number of encoded bits in data block
$n(t)$	Thermal noise

ξ_k	Trellis transition
P	Length of training sequence
$\mathbf{P}_k$	Inverse of the correlation matrix
$\Pr(\cdot)$	Probability
Π	Interleaver
$r(t)$	Received analog signal
$R(\tau)$	Periodic autocorrelation
$s(t)$	Transmitted waveform
S_k	Trellis state
σ^2	Noise variance
T	Symbol duration
T_m	Maximum excess delay
T_n	Subset trellis state in RSSE
$\hat{\mathbf{u}}$	Estimated information bits
$\mathbf{w}$	Sequence of white Gaussian noise samples
$\mathbf{w}_k$	Taps of the linear MMSE filter
W	Bandwidth
$\mathbf{y}$	Vector of received samples
$\mathbf{y}^m$	Received midamble symbols
$\chi_k(S_k)$	Survivor metric for state S_k
Ψ	Signal mapping
Ω	Space-time encoder

1 INTRODUCTION

1.1 Motivation

The capacity of Time Division Multiple Access (TDMA) cellular systems is limited by the bandwidth of radio frequencies that is available. Therefore a lot of effort is spent on the enhancements that can provide more capacity with a given frequency band. Receiver algorithms offer a useful method of enhancement, since they rarely require any changes in the transmitter side, which is usually fixed by the standard. The potential improvements can be achieved by means of digital signal processing at the receiver end.

Wireless transmission suffers from signal distortion caused by the objects in the radio path between the transmitter and receiver antennas. Several delayed echoes of the signal are often received causing intersymbol interference (ISI). This is very common especially in urban environments, as there are a lot of large buildings and street canyons, which reflect radio signals back and forth. The signal is further corrupted by thermal noise that is inherently present at the receiver. An equaliser is used to recover the transmitted signal from the received noisy and distorted signal. The Maximum Likelihood Sequence Estimation (MLSE) algorithm introduced by Forney [22] minimises sequence error probability in the presence of ISI and additive white Gaussian noise (AWGN). Signal detection usually requires the knowledge of channel impulse response, which is estimated using the known training symbols inserted into each transmission burst.

Convolutional coding is used to protect the system against transmission errors by adding redundancy in the transmitted signal. The optimum receiver maximises the joint probability of transmitted coded symbols, information bits and channel impulse response. As this is far too complex to implement in practice, the receiver is usually divided into detection, channel decoding and channel estimation operations. Each part can be independently optimised, but still the separated receiver is suboptimal solution. This thesis is motivated by the availability of *side information* in the receiver: the redundant information added by the convolutional coding provides extrinsic (*a priori*) information on the transmitted symbols. By using iterative turbo equalisation technique, this extra piece of information can be effectively utilised to improve performance [15]. Furthermore, iterative data processing can also be extended to channel estimation. Iterative methods are especially suitable for the data applications in the Enhanced Data Rates for Global Evolution (EDGE) system [18] due to the higher-order modulation (8-PSK) and the implementation of the algorithms is straightforward because of the rectangular interleaving structure over four transmission bursts.

Cochannel interference (CCI) often disturbs the transmission in TDMA cellular systems as the same carrier frequency has to be reused in the nearby cells. Since the system capacity is often interference-limited, it is important to improve *receiver resistance against interference* to increase capacity. This motivates us to consider interference cancelling (IC) techniques, from which this thesis focuses on multiple channel equalisation introduced in [19]. The multichannel equaliser detects simultaneously both the desired and interfering signals and hence it resembles multiuser detection in Code Division Multiple Access (CDMA) systems. As the receiver complexity increases exponentially with the number of detected signals and modulation levels, the joint detection (JD) of two binary modulated signals is feasible in practice. Moreover, the CCI can usually originate from a number of base stations, thus the receiver should first find the strongest interference among the candidates, which is then suppressed by the JD receiver. The CCI suppression is suitable for the GSM system, as the receiver complexity is reasonable for the binary GMSK modulation.

The objective of the thesis is to develop robust and efficient receiver algorithms for TDMA systems that utilise the available side information. Receivers based on iterative data processing are studied for communication systems with convolutional error-correcting codes and multichannel equalisation is proposed to suppress cochannel interference at the receiver.

1.2 Literature review

The iterative turbo coding method was introduced by Berrou *et al*. [11] in 1993 gaining fast a lot of attention. They proposed a parallel concatenation of two recursive systematic convolutional codes separated by a bit interleaver and an iterative decoding scheme, where the two decoders utilise extrinsic feedback information from the other decoder. Turbo codes have been found very powerful method reaching performance close to the Shannon limit as shown in [11]. Later Benedetto *et al.* [8] proposed Serially Concatenated Convolutional Codes (SCCC) showing significant benefit if the inner channel decoder utilises extrinsic information from the outer channel decoder in an iterative fashion.

Douillard *et al.* [15] published in 1995 the Turbo Equalisation (TE) scheme, which is modified from the SCCC by considering the multipath transmission channel as an inner convolutional code having coding rate 1. Accordingly the iterative receiver consists of the equaliser and outer channel decoder exchanging extrinsic information. Soft-Output Viterbi Algorithm (SOVA) in the presence of *a priori* information is presented and iterative receiver is suggested to be implemented by a modular pipelined structure. The TE technique provides good mitigation of Intersymbol Interference (ISI) for Gaussian multipath and Rayleigh fading channels [15] and for Typical Urban GSM channel shown by Picart *et al.* [45]. Both publications use 64x64 matrix for interleaving. Bauch *et al.* [4] suggests symbol-by-symbol Maximum a Posteriori (MAP) algorithm to be used both in the detection and channel decoding of the TE receiver and especially the MAP detector accepting *a priori* information is described. A few stopping criteria for the iteration are also discussed.

Bauch and Franz [6] consider the TE for GSM speech channels using inter-block diagonal interleaving. Due to the delay and memory restrictions only partial extrinsic information is available at the receiver, but still improvement is achieved even for the uncoded (Class 2) bits. Performance analyses of the TE for the EDGE system are presented by several authors with different equaliser selections. Strauch *et al.* [62] consider sub-optimal Minimum Mean Square Error - Block Decision Feedback Equaliser (MMSE-BDFE) and Franz and Bauch [23] suggest Max-Log-MAP. Furthermore, the Delayed Decision Feedback Sequence Estimation (DDFSE) algorithm is proposed by Berthet *et al.* [10] and by the author of this thesis [P3], although different design criteria for the coefficients of the feedforward filter are used. The later joint contributions [P1],[P2] are also based on the DDFSE equaliser.

Trellis-based equalisation becomes prohibitive in broadband transmission with large delay spread channels and multilevel modulation. Therefore linear adaptive filters with iterative data processing are recently proposed. The soft cancelling (SC) of the Intersymbol Interference (ISI), which is followed by the Minimum Mean Square Error (MMSE) linear filter is originally proposed for iterative multiuser detection in coded CDMA systems by Wang and Poor [71] in 1999. Reynolds and Wang applied the iterative SC-MMSE receiver to single carrier systems two years later in [53]. Omori and Asai present further reduction for the receiver complexity by approximations in [41] and Abe, Tomisato and Matsumoto apply the SC-MMSE structure to Multiple-Input Multiple-Output (MIMO) systems in [1],[2]. Several MMSE-based equalisation methods are presented with performance analysis by Tüchler *et al.* in [66],[68]. The MMSE turbo equalisation for multilevel modulations are considered by Laot *et al.* in [37].

Iterative channel estimation (ICE) schemes have also been studied by several authors. Chang and Georghiades [13] propose iterative data processing between the ML channel estimator and ML sequence estimator. Tentative hard data decisions are used to improve channel estimation accuracy. The ICE with soft decision feedback from the equaliser or channel decoder is considered for the GSM system by Sandell *et al.* [56] and for the EDGE by Strauch *et al.* [61]. The estimation schemes based on the Least Squares (LS) and Channel Sounding (CS) techniques are discussed in the both publications. Berthet *et al.* [9] combines the TE technique with channel re-estimation by proposing the Expectation-Maximisation (EM) algorithm. Abe *et al.* [1],[2] and Tüchler *et al.* [67] utilise the soft feedback from the decoder in the Recursive Least Squares (RLS) adaptation algorithm to achieve more reliable channel estimation.

There are several receiver techniques to suppress cochannel interference (CCI), e.g., antenna arrays, blind methods and multichannel equalisation. The focus of this thesis is the last method. The original idea of multiple signal equalisation has been proposed by van Etten [19] in 1976 to mitigate both intersymbol and interchannel interference with the same receiver. He considers a transmission system with M transmit and M receive antennas and extends the conventional single signal MLSE algorithm to vector form. Joint equalisers for M simultaneous signals are derived using Forney's [22] and Ungerboeck's [69] MLSE metrics.

During 1990's the interest of CCI suppression in the wireless mobile systems arose. The MLSE detection with the RLS-based adaptive per-survivor processing technique is extended to suppress the CCI by Yoshino *et al.* in [77]. A low-complexity alternative for the RLS adaptation is also considered and a reduced-state solution for the joint MLSE is presented. A prototype of the joint RLS-MLSE equaliser is constructed and measurement results in laboratory environment are published by Yoshino *et al.* in [76] and field trial results are presented in [78] . A blind scheme for the joint RLS-MLSE is introduced by Fukawa and Suzuki in [25]. This scheme is able to suppress the CCI without knowing the training sequence of the interference.

Giridhar *et al.* [28],[29] propose joint detection (JD) of wanted and interfering signals to suppress CCI in TDMA-based systems. Optimum joint ML and MAP detection algorithms are considered and a low-complexity 2-stage joint MAP algorithm, which is based on successive detection of wanted and interfering signals. Wales [70] also presents the joint ML solution and a reduced-state version for that. He discusses also joint channel estimation of two signals. Ranta considers in his Ph.D. thesis [52] multichannel equalisation thoroughly for the GSM system and part of the results are also included in this thesis [P8],[P9]. The independent publications [48],[49],[50],[51] investigate joint demodulation technique for various applications. Joint MLSE performance in the GSM in the presence of two interfering signals is studied in [48]. Only the stronger is jointly detected with the wanted signal. The network level effects of JD-receivers are considered in [49]. The JD combined with frequency hopping technique is discussed in [50] and application of the JD in street canyon environment is suggested in [51].

1.3 Scope of the thesis

The scientific content of this dissertation is collected from the GSM/EDGE research projects of Nokia Research Center during the years 1996-2000. The target system for the iterative receiver structures has been the EDGE system as specified in [79] and for the multichannel equalisation the current GSM system [18]. Since the aim of the industrial research projects has been in developing receiver enhancements without violating the specified system platform, we have to obey several limitations and system restrictions in algorithm development.

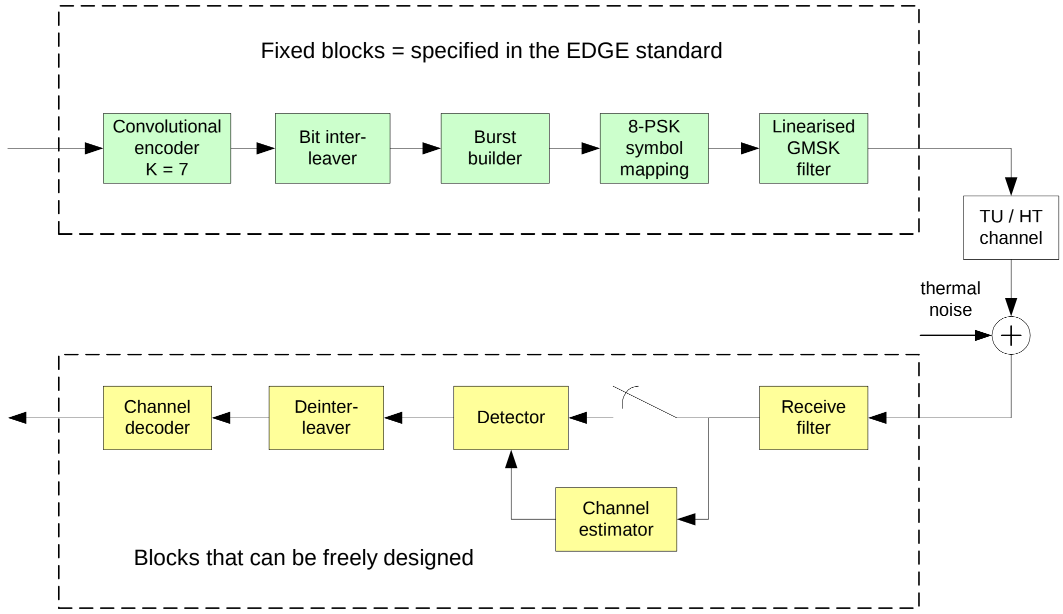

Figure 1. The EDGE system model [79].

The model of the EDGE system is described in Figure 1. The transmitter characteristics are defined in very detail in the standard [79] and thereby manufacturers have to make their products according to these specifications. Therefore we consider *the transmitter blocks as fixed* throughout the thesis. These blocks are collected in the upper row in Figure 1. However, the receiver implementation is not specified and can be designed freely. Hence, our studies are concentrated on these receiver blocks that are collected in the lower row in Figure 1.

A convolutional code with the constraint length K = 7 is specified for the EDGE system [79]. This has a clear impact on designing the iterative receiver structure as the decoding complexity is significant compared to the equalisation even with the 8-PSK modulation. Because the equalisation and decoding are repeated in the turbo equalisation schemes, both of them need to have feasible complexity. Since reduced-state trellis based equalisers offer reasonable performance-complexity tradeoff for the equalisation part, we put more effort in this thesis on reducing the decoding complexity.

Commonly used channel models for the narrowband GSM and EDGE systems are Typical Urban (TU) and Hilly Terrain (HT) channel profiles [79], for which we optimise our equaliser design. The delay spread in the HT channel is around 15.0 µs and since the GSM symbol duration is 3.69 µs, no more than *six symbol-spaced channel taps* are considered. The delay spread of the TU channel profile is even shorter. During the research work the Delayed Decision Feedback Sequence Estimation (DDFSE) has been considered a suitable equaliser for the EDGE system as proposed e.g., in [10],[27]. The author recognises that as the DDFSE utilises decision feedback information, non-minimum phase channels cause performance degradation. Therefore we always use a prefilter in the front of the DDFSE in order to convert the channel impulse response into minimum phase [26].

In the GSM system and its derivatives the user usually receives only one time slot from the frame of 8 time slots [18]. Furthermore, if frequency hopping technique is used, the next burst may be received at the different carrier frequency. Therefore the estimation of the channel impulse response has to be repeated for each burst. In this thesis we are restricted to *block fading channel characteristics*, i.e., the channel is constant during the burst, but changes from burst to burst, which is a reasonable assumption for slowly moving mobiles [42]. As a consequence, we do not consider channel tracking and per-survivor approaches, but concentrate on the one-shot channel estimation techniques like Maximum Likelihood (ML) or Linear Minimum Mean Square Error (LMMSE) methods.

The thesis comprises of the collection of international publications (two journal papers, six conference publications and a book chapter) and the present summarising text. The main objective in the work has been to create a deep understanding on the current receiver solutions and finding new improved solutions for the GSM system and its derivative EDGE. We focus on two areas of improvements: utilisation of iterative receiver structures and CCI suppression by the joint demodulation technique. We cover techniques related to turbo equalisation (TE), i.e., iterative detection and channel estimation methods, as they fit into the framework of the current GSM and EDGE standards. Due to the rectangular interleaving in EDGE it is easy to construct the TE receiver and performance improvement can be expected.

CCI can be suppressed by several receiver techniques like adaptive antennas, blind or semi-blind equalisers or multichannel equalisers. The most effective interference suppression methods are based on antenna arrays, but it is not straightforward to implement multiple antennas in the mobile terminal. Blind methods can suppress interference based on the constant envelope property. However, when the envelope is fluctuating, e.g., due to multipath channel, these methods become less reliable. This thesis concentrates on multichannel equalisation, which is straightforward to implement also in the mobile handset and offers very good performance.

The performance analysis is mostly based on link simulations, which are easy to perform even for complicated systems. Theoretical analysis is usually too difficult for multipath fast fading channels that are most important in practice. The thesis considers also implementation issues in terms of computational complexity and system requirements that are posed by the receiver algorithms.

1.4 Outline of the thesis

The thesis is organised as follows. The transmission system is first described with discussion about channel modelling and receive filter structure. The communication system is also extended to cover multiple transmit antennas.

Channel estimation and detection algorithms are presented in the next chapter. Channel estimation using Maximum-Likelihood (ML) and Linear Minimum Mean Square Error (LMMSE) approaches is discussed. Adaptive linear filtering is also considered and Recursive Least Squares (RLS) solution is presented. Furthermore, estimation in the presence of feedback information is proposed, which is utilised later in the iterative receiver schemes. Finally joint channel estimation of several channels is presented. The detection part describes first the optimum Maximum Likelihood Sequence Estimation (MLSE) algorithm and two modifications providing soft information. Low-complexity detection is considered by introducing Reduced State Sequence Estimation (RSSE) approach and per-survivor processing is used in the RLS-MLSE detection scheme. Then the optimum Maximum A Posteriori (MAP) detection is presented and the modification accepting *a priori* information is described in detail. After that joint detection of multiple signals using the ML and MAP algorithms is considered. Finally the cancellation of Intersymbol Interference (ISI) followed by the linear Minimum Mean Squared Error (MMSE) filter is introduced.

Iterative receiver techniques for coded systems are described in Chapter 4. The principle of turbo equalisation is described using the Soft-In/Soft-Out (SISO) equaliser introduced in the previous chapter. Also SISO decoding is considered and a low-complexity algorithm is proposed. Furthermore, the trade-off between performance and complexity is emphasised. Finally, space-time equalisation is considered and an iterative structure for that is described.

In Chapter 5 we extend the equaliser to detect two signals simultaneously. The joint channel estimation and joint detection algorithms presented in Chapter 3 are utilised. Also the interference cancellation scheme by the joint RLS-MLSE is presented. Suitable training sequences to be used with the JCE are proposed and the problem of finding the dominant interfering signal among the candidates is discussed.

Chapter 6 summarises the author's publications with a discussion about theoretical contents and main results. The author's contribution to the publications is given in detail. Finally, the conclusions on the considered techniques and discussion about the usability are given in Chapter 7.

2 TRANSMISSION SYSTEM

2.1 Single antenna system

2.1.1 Transmitter

A communications system with error-correcting coding and M-ary modulation is illustrated in Figure 2 by a block diagram. A block of information bits $\mathbf{u} \equiv \left(u_1, u_2, \ldots, u_{K_0}\right)^T$ is protected by a convolutional encoder Γ and punctured to achieve an appropriate data rate given by K_0/N_0. The encoded data $\mathbf{c} \equiv \left(c_1, c_2, \ldots, c_{N_0}\right)^T$ is interleaved over several successive transmission bursts. A burst is formed from the data bits and known midamble bits $\mathbf{m} \equiv \left(m_1, m_2, \ldots, m_P\right)^T$ that are inserted in the middle of the burst. Each burst consists of K complex-valued symbols denoted by $\mathbf{a} \equiv \left(a_1, a_2, \ldots, a_K\right)^T$ where every M-ary symbol a_k corresponds to $\log_2 M$ input bits according to the signal mapping Ψ. The transmitted complex waveform at the baseband is given by [46]

$$s(t) = \sum_k a_k g_{tx}(t - kT) \ , \tag{2.1}$$

where $g_{tx}(t)$ is the impulse response of lowpass equivalent transmitter filter and the symbol rate is $1/T$.

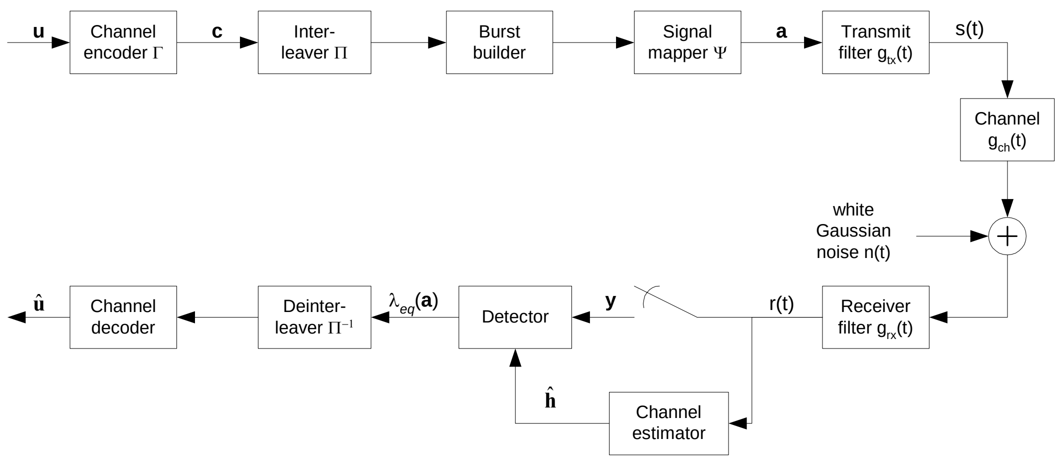

Figure 2. Block diagram of the communications system [46].

2.1.2 Channel modelling

The reliability of the mobile communication depends on the radio propagation channel between the transmitter and receiver. The radio signal is distorted and attenuated due to different mechanisms that are discussed in this section.

In typical cellular environment there are a lot of large buildings and other obstacles that are reflecting the transmitted radio signal. As a consequence, the received signal consists of several delayed echoes of the original signal that have independent signal power attenuations and phases. This phenomenon is called *multipath propagation*. The channel causes delay spread in the received signal as a transmitted impulse spreads in time domain. Because the structure of the transmission medium varies in time, the multipath components are also time-variant, hence the phases and amplitudes vary from time instant to another. Moreover, the variations cannot be predicted and therefore the time-variant channel is characterised statistically [46].

The multipath components can add up in the receiver destructively, i.e., the independent echoes cancel out each other if they arrive with opposite phases. This causes *fast fading* in the received signal power, thus the signal power fluctuates rapidly. If the impulse response is modelled by zero-mean complex-valued Gaussian process, the channel is Rayleigh fading channel. If there is a line-of-sight between the transmitter and receiver, the impulse response is not anymore zero-mean and the channel model is Ricean fading channel [46].

For a transmitted impulse, the time between the first and last received component is called maximum excess delay T_m. For typical mobile channels in the GSM system this delay is larger than the symbol period, i.e., $T_m > T$. Such a channel is *frequency-selective*, which means that the signal distortion due to the channel depends on the frequency. For small delay spread ($T_m << T$) the channel exhibits *flat fading*, i.e., all spectral components of the signal are affected equally and no intersymbol interference (ISI) is introduced [59].

The equivalent low-pass time-variant impulse response $g_{ch}(\tau,t)$ describes the channel response at time *t* to an impulse applied at time *t*-τ. So τ denotes the delay (elapsed time) variable. The response can be represented as [46]

$$g_{ch}(\tau,t) = \sum_k g_k(t)\delta(\tau - kT) \quad , \tag{2.2}$$

where $g_k(t)$ denotes the k^{th} element of the symbol-spaced tapped delay line at time *t*, $\delta(t)$ is the Dirac delta function and *T* is the symbol period. However, the GSM burst is only 577 μs long [39] and therefore the channel variation during a burst is small. Therefore *block fading* characteristics can be used, i.e., the channel is assumed constant during the burst, but varying between bursts. The time-invariant impulse response $g_{ch}(\tau)$ (for a certain burst) is given by

$$g_{ch}(\tau) = \sum_k g_k \delta(\tau - kT) \quad . \tag{2.3}$$

2.1.3 Receive filtering and sampling

A continuous baseband signal $s(t)$ is band-limited if $S(f) = 0$ for $|f| > W$, where $S(f)$ is the Fourier transform of $s(t)$ and W is the highest frequency in $s(t)$. The band-limited signal can be uniquely represented by a discrete-time signal with the minimum sampling rate $f_s \geq 2W$ (Nyquist rate) [46], since lower sampling rates lead to frequency aliasing. In practice communication signals are not usually strictly band-limited due to implementation impairments and finite length pulse shapes. By assuming ideal low-pass filter (LPF) with cut-off frequency W we can achieve truly band-limited signal. Moreover, by selecting large enough W the signal energy can be mostly preserved, since the out-of-band energy becomes negligible. Sufficient statistics for the data detection is then obtained by sampling at the rate $1/T_s = 2W$, i.e., all the necessary information for the detection is preserved in the analog-to-digital (A/D) conversion.

From the implementation point of view it is convenient to use sampling rate which is multiple of the symbol rate $1/T$. Forney suggests in [21] to use channel matched filter (CMF) that is followed by a symbol-rate sampler to provide sufficient statistics. However, in the mobile radio applications the channel is not known in advance, because of which the CMF solution as such is not feasible and fractionally spaced processing is needed to guarantee statistical sufficiency. Fortunately in the GSM system it is feasible to use narrow receive filter $(W < 1/T)$ and therefore in practice symbol-spaced sampling is reasonable despite the insufficient statistics.

Pulse transmission through band-limited channel is possible without intersymbol interference (ISI) if the overall time response $g(kT_s)$ satisfies the Nyquist's first criterion

$$g(kT_s) = \begin{cases} 1, & k = 0 \\ 0, & k = \pm 1, \pm 2, \ldots \end{cases}, \tag{2.4}$$

where T_s is the sampling period. In the frequency domain the Nyquist's first criterion is given by [46]

$$\sum_{n=-\infty}^{\infty} G(f - n/T_s) = T_s \ , \tag{2.5}$$

where $G(f)$ is known as a Nyquist filter. A particular Nyquist filter widely used in practical applications is the raised cosine (RC) filter

$$G_{RC}(f) = \begin{cases} T_s & |f| < \dfrac{1-\alpha}{2T_s} \\ \dfrac{T_s}{2}\left\{1 + \cos\left[\dfrac{\pi T_s}{\alpha}\left(|f| - \dfrac{1-\alpha}{2T_s}\right)\right]\right\}, & \dfrac{1-\alpha}{2T_s} \leq |f| \leq \dfrac{1+\alpha}{2T_s} \\ 0 & |f| > \dfrac{1+\alpha}{2T_s} \end{cases}, \tag{2.6}$$

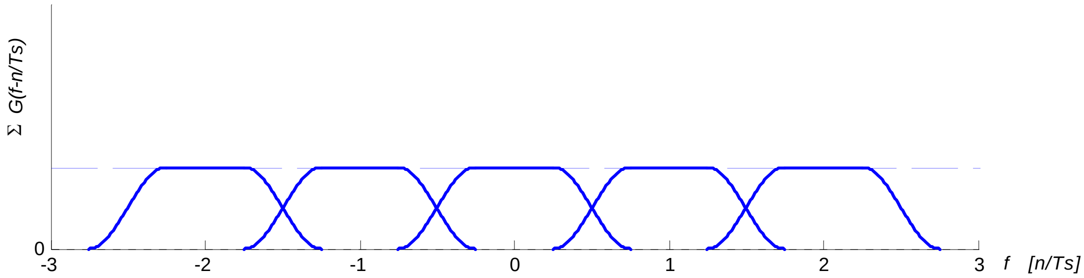

Figure 3. Spectrum of folded raised-cosine pulses, i.e., $\sum_{n=-\infty}^{\infty} G_{RC}\left(f - n/T_s\right)$.

where α is called the rolloff factor having values between zero and one. The bandwidth occupied by the signal beyond the Nyquist frequency $1/2T_s$ is called the excess bandwidth, which is determined by the rolloff factor. The spectrum of folded RC pulses with rolloff $\alpha = 0.5$ is illustrated in Figure 3 , which shows that the Nyquist's criterion (2.5) holds. The pulse with raised cosine spectrum is

$$g(t) = \frac{\sin(\pi t / T_s)}{\pi t / T_s} \frac{\cos(\pi \alpha t / T_s)}{1 - (2\alpha t / T_s)^2} \quad . \tag{2.7}$$

Figure 4 depicts a RC pulse in the time domain with the rolloff factor $\alpha = 0.5$.

In practical implementations the frequency response of the raised cosine filter is split between transmitter and receiver filters as follows [46]

$$\begin{aligned} G_{RC}(f) &= G_{tx}(f) G_{rx}(f) \\ G_{tx}(f) &= \sqrt{|G_{RC}(f)|} \exp(-j2\pi f \tau_0) \quad , \\ G_{rx}(f) &= G_{tx}^*(f) \end{aligned} \tag{2.8}$$

where τ_0 is the nominal delay to ensure physical realizability of the filter. The receiver filter $G_{rx}(f)$ has square root raised cosine (SQRC) spectrum and it is matched to the transmitting filter to maximise the SNR at the sampling instants.

The cascade of the transmitting filter $g_{tx}(t)$, the physical channel $g_{ch}(t)$, the receive filter $g_{rx}(t)$ and the A/D converter (sampler) is represented by an equivalent complex-valued and symbol-spaced tapped delay line $\mathbf{h} \equiv (h_0, h_1, \ldots, h_L)^T$. At the output of the receive filter a sequence of digital samples $\mathbf{y} \equiv (y_1, y_2, \ldots, y_K)^T$ is obtained, where each sample can be written as

$$y_k = \sum_{l=0}^{L} a_{k-l} h_l + w_k \quad . \tag{2.9}$$

The discrete-time channel model is depicted in Figure 5.

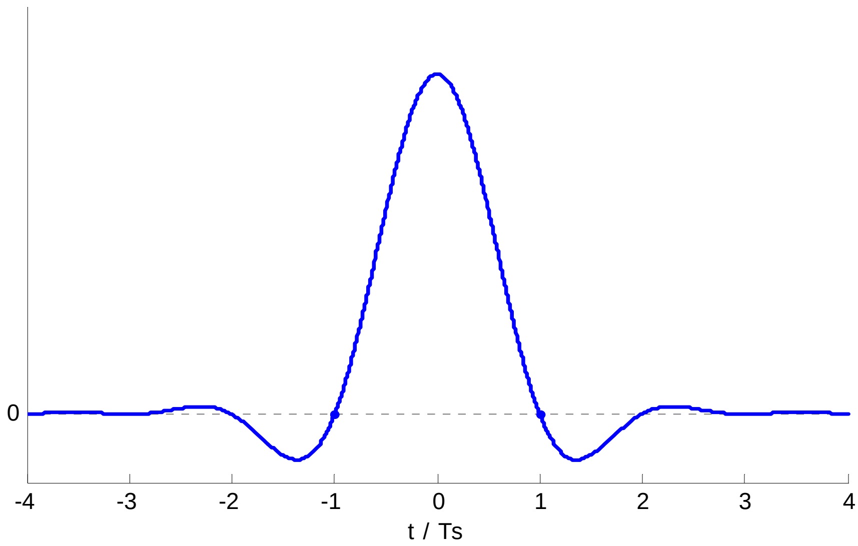

Figure 4. Raised cosine pulse shape with rolloff $\alpha = 0.5$.

The thermal noise $n(t)$ in the receiver is modelled as additive zero-mean white Gaussian noise. It is filtered by the SQRC filter $G_{rx}(f)$ and sampled at the symbol resolution to obtain noise samples w_k . The two-sided power spectral density (PSD) of the input noise $n(t)$ is $N_0/2$ and the PSD of the filtered noise is [58]

$$G_w(f) = G_n(f)\left|G_{rx}(f)\right|^2 = \frac{N_0}{2}\left|G_{rx}(f)\right|^2 \quad . \tag{2.10}$$

The autocorrelation function $R_w(\tau)$ of the filtered noise is obtained as an inverse Fourier transform from the PSD $G_w(f)$ [58], and by assuming SQRC spectrum for the receive filter we obtain

$$R_w(\tau) = \mathcal{F}^{-1}\left[\frac{N_0}{2}\left|G_{rx}(f)\right|^2\right] = \frac{N_0}{2} g_{RC}(\tau) \quad . \tag{2.11}$$

As the RC pulse satisfies the Nyquist's criterion (2.4), the autocorrelation of the noise samples at the symbol rate is given as

$$R_w(kT_s) = \begin{cases} N_0/2, & k = 0 \\ 0, & k = \pm 1, \pm 2, \ldots \end{cases} \tag{2.12}$$

and therefore the noise samples are uncorrelated. Moreover, as the Gaussian noise process $n(t)$ is applied to the time-invariant linear filter $G_{rx}(f)$, the random noise process w_k is also Gaussian [58]. The additive zero-mean white Gaussian noise samples w_k have statistically independent real and imaginary parts and the noise variance is given by

$$\sigma^2 = \int_{-\infty}^{\infty} G_w(f) = \frac{N_0}{2}\int_{-\infty}^{\infty}\left|G_{rx}(f)\right|^2 = \frac{N_0}{2} \quad . \tag{2.13}$$

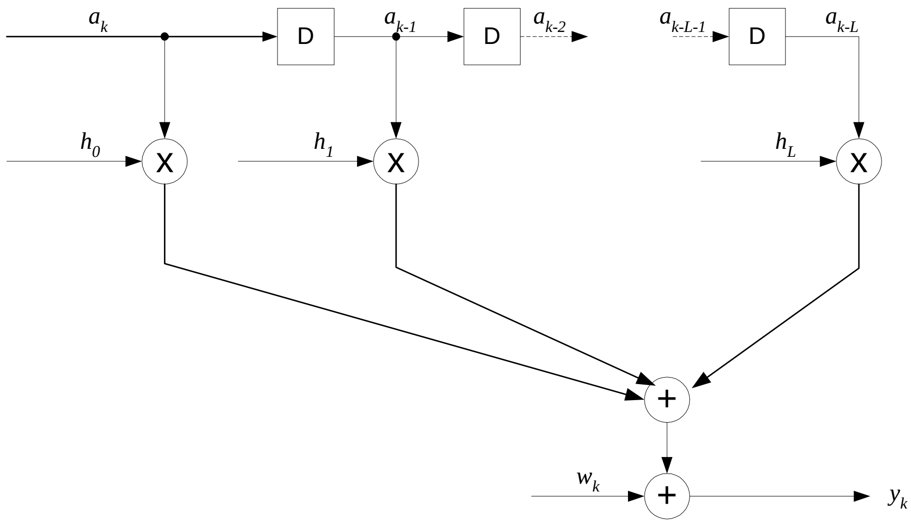

Figure 5. Equivalent discrete-time channel model [46].

The probability density function (pdf) of the noise is given by [34]

$$f(w_k) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{|w_k|^2}{2\sigma^2} \right) \quad . \tag{2.14}$$

Equation (2.9) can be represented in matrix form as

$$\mathbf{y} = \mathbf{Ah} + \mathbf{w} \quad , \tag{2.15}$$

where $\mathbf{A}$ is the symbol matrix

$$\mathbf{A} \equiv \begin{bmatrix} a_{L+1} & a_L & \cdots & a_1 \\ a_{L+2} & a_{L+1} & & a_2 \\ \vdots & & \ddots & \vdots \\ a_K & a_{K-1} & \cdots & a_{K-L} \end{bmatrix} \quad . \tag{2.16}$$

Figure 6 presents the whole discrete-time system model. At the receiver side the received signal $\mathbf{y}$ is equalised with the aid of the channel estimator. The soft outputs $\lambda_{eq}(\mathbf{a})$ are deinterleaved and provided for the channel decoder, which finally decodes the information bits $\hat{\mathbf{u}}$.

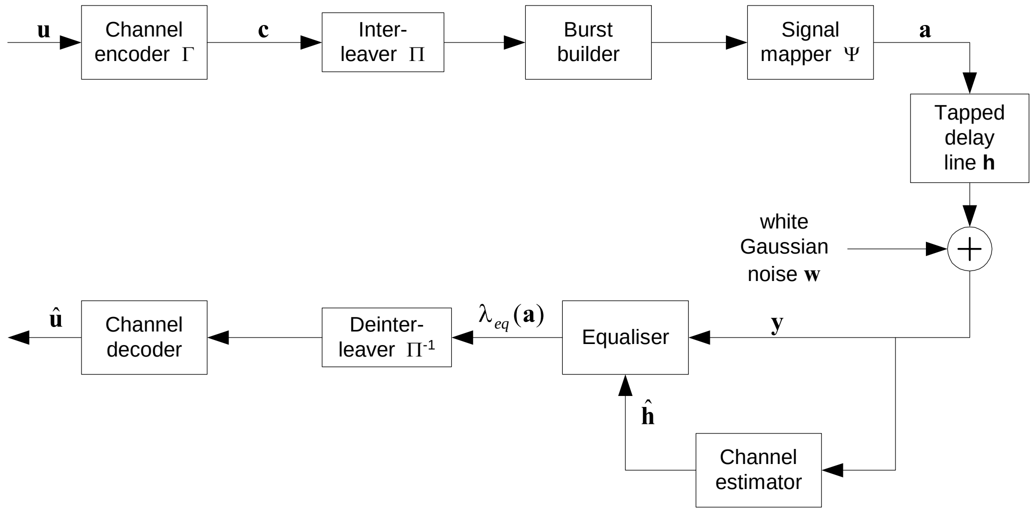

Figure 6. Equivalent discrete-time system model [46].

2.2 Systems with multiple transmit antennas

Figure 7 models a system with two transmit antennas and one receive antenna, although in general Multiple Input Single Output (MISO) system consists of N transmit antennas. Each transmit signal consists of independent data stream $\mathbf{a}^{(n)} \equiv \left(a_1^{(n)}, a_2^{(n)}, ..., a_K^{(n)}\right)^T$, $n = 1,...,N$ and each signal is associated with a known midamble sequence $\mathbf{m}^{(n)} \equiv \left(m_1^{(n)}, m_2^{(n)}, \ldots, m_P^{(n)}\right)^T$ to assist channel estimation at the receiver end. The signals are transmitted over independent frequency-selective fading channels with impulse responses $g_{ch}^{(n)}(t)$.

At the receiver the Square Root Raised Cosine (SQRC) filter and symbol-rate sampler are used. An equivalent discrete-time channel models are represented by the complex-valued and symbol-spaced tapped delay lines $\mathbf{h}^{(n)} \equiv \left(h_0^{(n)}, h_1^{(n)}, \ldots, h_L^{(n)}\right)^T$ including the transmitting filter, physical channel and receive filter for the cochannel signal n. The received signal is the following superposition

$$y_k = \sum_{n=1}^{N} \sum_{l=0}^{L} a_{k-l}^{(n)} h_l^{(n)} + w_k \quad , \tag{2.17}$$

where the noise samples w_k are white Gaussian with zero mean and variance $\sigma^2 = N_0/2$.

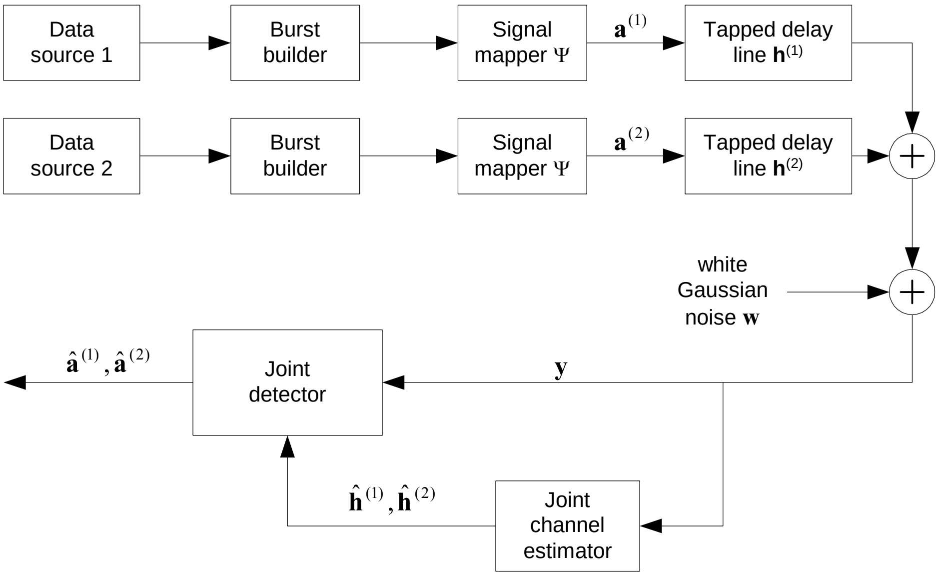

Figure 7. Discrete-time communication system with two transmitted signals.

3 RECEIVER ALGORITHMS

3.1 Channel estimation

Signal detection algorithms require the knowledge of channel impulse response, which is usually estimated by using the known training (midamble) symbols in the middle of the transmission burst [39]. In mobile environment the channel is time-variant, which makes the estimation task more difficult. In the GSM system and its derivatives the time period between the bursts is so long [18] that the channel changes significantly from burst to burst and thus a separate channel estimation is needed for each burst. On the other hand the change during the burst for slowly moving mobiles is rather limited, hence it is reasonable to assume block fading channel characteristics, i.e., the channel is constant during the burst, but is changing between them [42]. In this section we present two basic approaches to estimate the channel impulse response that are Maximum Likelihood (ML) and Linear Minimum Mean Square Error (LMMSE) estimation methods. We also consider adaptive linear filter approach by presenting Recursive Least Squares (RLS) solution for the parameter estimation. Then estimation in the presence of feedback information is discussed and finally extension to multiple channel estimation is considered.

3.1.1 Maximum Likelihood (ML) channel estimation

The received samples y_k^m corresponding to the known midamble symbols m_k are given by

$$y_k^m = \sum_{l=0}^{L} m_{k-l} h_l + w_k \tag{3.1}$$

or in matrix form

$$\mathbf{y}^m = \mathbf{M}\mathbf{h} + \mathbf{w} \tag{3.2}$$

where **M** denotes the following $P \times (L+1)$ midamble symbol matrix

$$\mathbf{M} = \begin{bmatrix} m_{L+1} & m_L & \cdots & m_1 \\ m_{L+2} & m_{L+1} & & m_2 \\ \vdots & & \ddots & \vdots \\ m_P & m_{P-1} & \cdots & m_{P-L} \end{bmatrix} \tag{3.3}$$

and **w** is $P \times 1$ Gaussian noise vector with zero mean and covariance $\mathbf{C}_w$. The pdf of the observations $\mathbf{y}^m$ is given as [34]

$$f(\mathbf{y}^m, \mathbf{h}) = \frac{1}{\sqrt{(2\pi)^P \det(\mathbf{C}_w)}} \exp\left[-\frac{1}{2} \left(\mathbf{y}^m - \mathbf{M}\mathbf{h}\right)^H \mathbf{C}_w^{-1} \left(\mathbf{y}^m - \mathbf{M}\mathbf{h}\right) \right] \quad . \tag{3.4}$$

The ML estimate is found by maximising Eq. (3.4) with respect to **h**, which gives [34]

$$\hat{\mathbf{h}} = \left(\mathbf{M}^H \mathbf{C}_w^{-1} \mathbf{M}\right)^{-1} \mathbf{M}^H \mathbf{C}_w^{-1} \mathbf{y}^m \quad . \tag{3.5}$$

This estimator is unbiased with the covariance

$$\mathbf{C}_{\hat{h}} = \left(\mathbf{M}^H \mathbf{C}_w^{-1} \mathbf{M}\right)^{-1} \quad . \tag{3.6}$$

The variance of a certain channel tap estimator is given by the corresponding diagonal element of the covariance matrix, i.e., $Var(\hat{h}_k) = \left(\mathbf{C}_{\hat{h}}\right)_{kk}$. The same estimator (3.5) is also minimum variance unbiased (MVU) estimator for the given linear model [34]. In the special case of white Gaussian noise the ML estimate reduces to

$$\hat{\mathbf{h}} = \left(\mathbf{M}^H \mathbf{M}\right)^{-1} \mathbf{M}^H \mathbf{y}^m \quad . \tag{3.7}$$

As the midamble symbols are constant from burst to burst, the matrix inverse can be computed in advance and saved into memory. Hence, the channel estimation (3.7) is a fairly simple task to implement in the receiver.

3.1.2 Linear Minimum Mean Square Error (LMMSE) estimation

LMMSE channel estimator is restricted to the linear form

$$\hat{h}_i = a_{i,0} + \sum_{k=1}^{P} a_{i,k} y_k^m \quad , i = 0,1,\ldots,L \quad , \tag{3.8}$$

where all $a_{i,k}$ denote constants. The estimator minimises the Bayesian MSE

$$Bmse(\hat{\mathbf{h}}) = E\left[(\mathbf{h} - \hat{\mathbf{h}})^2\right] \quad , \tag{3.9}$$

where the expectation is with respect to pdf $f\left(\mathbf{y}^m, \mathbf{h}\right)$ given by Eq. (3.4). In general the LMMSE solution is [34]

$$\hat{\mathbf{h}} = E(\mathbf{h}) + \mathbf{C}_{hy} \mathbf{C}_y^{-1} \left(\mathbf{y} - E(\mathbf{y})\right) \tag{3.10}$$

where $\mathbf{C}_y$ is a $P \times P$ covariance matrix of **y** and $\mathbf{C}_{hy}$ is a $(L+1) \times P$ cross-covariance matrix. For the given linear model (3.2) where **h** and **w** are uncorrelated, the estimator is given by

$$\hat{\mathbf{h}} = E(\mathbf{h}) + \left(\mathbf{C}_h^{-1} + \mathbf{M}^H \mathbf{C}_w^{-1} \mathbf{M}\right)^{-1} \mathbf{M}^H \mathbf{C}_w^{-1} \left(\mathbf{y} - \mathbf{M}E(\mathbf{h})\right) \quad , \tag{3.11}$$

where $\mathbf{C}_h$ and $\mathbf{C}_w$ denote the channel tap covariance and noise covariance, respectively. The estimator covariance is given by [34]

$$\mathbf{C}_{\hat{h}} = \left(\mathbf{C}_h^{-1} + \mathbf{M}^H \mathbf{C}_w^{-1} \mathbf{M}\right)^{-1} \tag{3.12}$$

and a single channel tap estimator has variance $Var(\hat{h}_k) = \left(\mathbf{C}_{\hat{h}}\right)_{kk}$. If we assume that the expectation value for fading channel is zero, i.e., $E(\mathbf{h}) = \mathbf{0}$, and noise samples are white Gaussian, then LMMSE estimator further reduces to

$$\hat{\mathbf{h}} = \left(\sigma^2 \mathbf{C}_h^{-1} + \mathbf{M}^H \mathbf{M}\right)^{-1} \mathbf{M}^H \mathbf{y} \quad . \tag{3.13}$$

When compared to the ML solution (3.7) there is only one extra weighting factor $\sigma^2 \mathbf{C}_h^{-1}$.

As the linear signal model (3.1) is applied, there is no degradation due to the linearity constraint (3.8) of the estimator, but the LMMSE achieves the same solution as the optimal Bayesian MMSE [34]. The benefit of the Bayesian estimators in general is that they can incorporate *a priori* information easily into the estimation. The LMMSE channel estimator utilises the *a priori* knowledge of the channel covariance matrix $\mathbf{C}_h$, which with the assumption of independent channel coefficients is diagonal, i.e., $\mathbf{C}_h = \mathrm{diag}\left(h_0^2, h_1^2, \ldots, h_L^2\right)$. Consequently, these positive diagonal terms make the LMMSE estimator covariance (3.12) smaller than the corresponding ML estimator covariance (3.6).

The LMMSE is very useful in practice, as the formula is rather easy to implement in applications and good improvement in performance is often achieved. There is some complexity increase as the extra weighting factor $\sigma^2 \mathbf{C}_h^{-1}$ is needed, which can be obtained from a prior ML estimation, for instance. As the weighting factor is time-variant, the matrix inversion in Eq. (3.13) has to be performed for each burst separately, unlike the constant precomputed matrix inversion of the ML estimator (3.7).

3.1.3 Recursive Least Squares (RLS) estimation

The optimal estimate of the channel impulse response in the least square (LS) sense is

$$\hat{\mathbf{h}} = \arg\min_{\mathbf{h}} \left\| y_k - \sum_{l=0}^{L} a_{k-l} h_l \right\|^2 \quad . \tag{3.14}$$

The recursive least squares (RLS) algorithm can be used to solve this optimisation problem. By defining the channel taps $\mathbf{h}(k) \equiv \left(h_0(k), h_1(k), \ldots, h_L(k)\right)^T$ as the weights of the transversal filter at time instant *k* and the input vector into the filter as $\mathbf{a}_k \equiv \left(a_k, a_{k-1}, \ldots, a_{k-L}\right)^T$ the error function at time *k* is given as

$$e_k = y_k - \hat{\mathbf{h}}^H(k)\mathbf{a}_k \quad . \tag{3.15}$$

The cost function that is minimised by the RLS algorithm is [32]

$$J_k = \sum_{i=1}^{k} \lambda^{k-i} \left|e_i\right|^2 \quad , \tag{3.16}$$

where $\lambda \in \left]0,1\right[$ is the forgetting factor. The optimal weights $\hat{\mathbf{h}}(k)$ can be solved by the normal equations [32]

$$\mathbf{\Phi}_k \hat{\mathbf{h}}(k) = \mathbf{z}_k \quad , \tag{3.17}$$

where the correlation matrix and cross-correlation vector are defined as

$$\underset{(L+1)\times(L+1)}{\mathbf{\Phi}_k} \equiv \sum_{i=1}^{k} \lambda^{k-i} \mathbf{a}_i \mathbf{a}_i^H \tag{3.18}$$

$$\underset{(L+1)\times 1}{\mathbf{z}_k} \equiv \sum_{i=1}^{k} \lambda^{k-i} \mathbf{a}_i y_i^* \quad . \tag{3.19}$$

To calculate $\hat{\mathbf{h}}(k)$ using Eq. (3.17) the matrix inverse $\mathbf{P}_k \equiv \mathbf{\Phi}_k^{-1}$ is needed. By applying the matrix inversion lemma the actual inverse calculation can be avoided and computationally more efficient algorithm can be derived. The RLS parameter updating algorithm can be summarised as follows [32]

$$\hat{\mathbf{h}}(k) = \hat{\mathbf{h}}(k-1) + K_k \varepsilon_k^* \tag{3.20}$$

$$K_k = \frac{\lambda^{-1} \mathbf{P}_{k-1} \mathbf{a}_k}{1 + \lambda^{-1} \mathbf{a}_k^H \mathbf{P}_{k-1} \mathbf{a}_k} \tag{3.21}$$

$$\varepsilon_k = y_k - \hat{\mathbf{h}}^H (k-1) \mathbf{a}_k \tag{3.22}$$

$$\mathbf{P}_k = \lambda^{-1} \mathbf{P}_{k-1} - \lambda^{-1} K_k \mathbf{a}_k^H \mathbf{P}_{k-1} \quad . \tag{3.23}$$

The initial values can be set as

$$\mathbf{P}_0 = \delta^{-1} \mathbf{I} \tag{3.24}$$

$$\hat{\mathbf{h}}(0) = (0,0,...,0)^T \quad , \tag{3.25}$$

where δ is a small positive number.

The RLS algorithm gives more weight to the current observations due to the forgetting factor. If the factor $\lambda = 1$, the solution reduces to the normal least squares (LS) solution.

3.1.4 Channel estimation using feedback information

As shown in the previous sections the conventional ML, LMMSE or RLS estimators rely on the known midamble symbols. However, as the length P is typically rather short in mobile systems the estimation reliability remains low, which in turn can significantly deteriorate receiver performance. Therefore by using decoded symbol decisions $\hat{\mathbf{u}}$ as an extended training sequence the channel estimation accuracy can be improved.

According to Eq. (2.6) the received signal for the whole burst can be represented as

$$\mathbf{y} = \mathbf{A}\mathbf{h} + \mathbf{w} \quad . \tag{3.26}$$

If all the data symbols were known, the ML channel estimator is given by

$$\hat{\mathbf{h}} = \left(\mathbf{A}^H \mathbf{A}\right)^{-1} \mathbf{A}^H \mathbf{y} \quad . \tag{3.27}$$

As the matrix $\mathbf{A}^H\mathbf{A}$ consists of data symbols, it has to be inverted separately for each burst. Moreover, the estimator (3.27) requires a lot of multiplications as the whole burst of data is utilised. For these reasons stochastic adaptation algorithms like the Least Mean Square (LMS) [32] are preferred. The LMS adaptation rule for updating the channel estimate is as follows

$$\hat{\mathbf{h}}^{(k+1)} = \hat{\mathbf{h}}^{(k)} - \mu\mathbf{A}^H\left(\mathbf{A}\hat{\mathbf{h}}^{(k)} - \mathbf{y}\right) \ . \tag{3.28}$$

where $\hat{\mathbf{h}}^{(k)}$ is the channel estimate of k^{th} iteration and μ is the constant step size of the adaptive algorithm. The covariance matrix of the estimator based on the extended training sequence is

$$\mathbf{C}_{\hat{\mathbf{h}}} = \left(\mathbf{A}^H\mathbf{C}^{-1}\mathbf{A}\right)^{-1} \ , \tag{3.29}$$

where the covariance term $\mathbf{C}$ accumulates both received noise and noise enhancement due to incorrect data decisions. The variance of the estimator can be bounded by Cramer-Rao Lower Bound (CRLB) [P4]

$$\mathrm{var}\left[\hat{h}_i\right] \geq \frac{\sigma^2}{P + \dfrac{2N_d\sigma^2}{4p - 4p^2 + \sigma^2}} \ , \ i = 0,1,\ldots,L \ , \tag{3.30}$$

where σ^2 denotes the noise variance, p is the Bit Error Probability (BEP) and P and N_d stand for the number of training and data symbols. This new estimator decreases the estimator variance even with very unreliable data decisions due to the increased number of symbols that are used in the estimation [P4]. In fact, the best improvement is achieved in the low SNR region and the achieved gain gradually disappears when the signal quality improves. Figure 8 shows the CRLB bounds for the estimator variance using the GSM parameters, i.e., P = 20 and N_d = 58. Different BEP values after decoding are considered and a curve for training sequence based estimation is also given. This bound is obtained by setting N_d = 0 in Eq. (3.30).

Also the RLS estimation can utilise the feedback information. The *a priori* information $\lambda^a(a_k)$ obtained from the decoder is used to create soft estimates

$$\tilde{a}_k = \tanh\left[\frac{\lambda_d(a_k)}{2}\right] \ . \tag{3.31}$$

By thresholding these soft values, the most reliable decisions can be identified and those bits can be used with the known training bits in the new RLS channel estimation round according to Eq. (3.20) – (3.23) [1]. Due to the extra information available the estimation accuracy is improved.

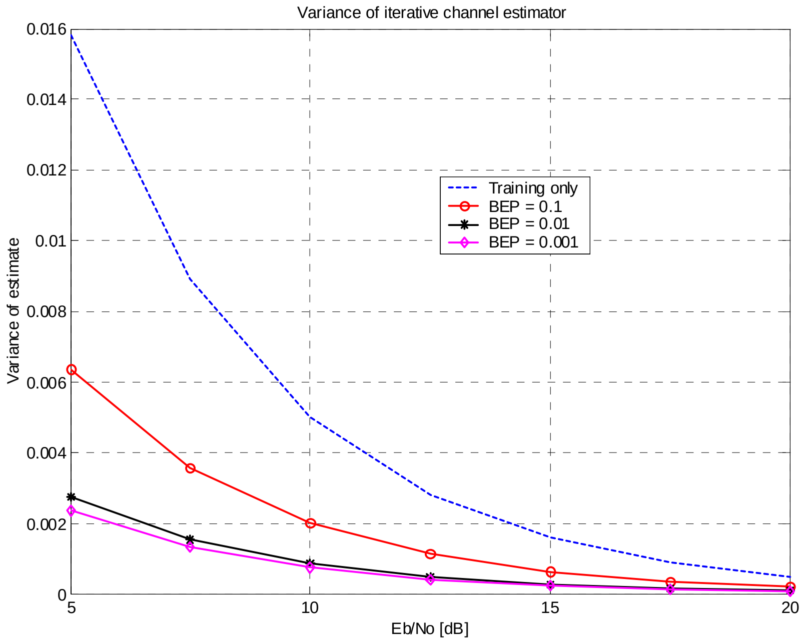

Figure 8. Variance of estimate vs. $E_b / N_0 = 10\log\left(1/\sigma^2\right)$. Curves are shown for various Bit Error Probabilities (BEP) and for the estimator based on training sequence only.

3.1.5 Joint ML estimation of multiple channels

Following the signal model for multiple transmit antennas in Section 2.2 the received samples $\mathbf{y}^m$ corresponding to midamble symbols are given by

$$\mathbf{y}^m = \mathbf{M}\mathbf{h} + \mathbf{w} \tag{3.32}$$

where the radio channels are organised as

$$\mathbf{h} \equiv \begin{bmatrix} \mathbf{h}^{(1)} \\ \mathbf{h}^{(2)} \\ \vdots \\ \mathbf{h}^{(N)} \end{bmatrix} \tag{3.33}$$

and midamble symbol matrices as $\mathbf{M} \equiv \left[\mathbf{M}^{(1)} \quad \mathbf{M}^{(2)} \quad \cdots \quad \mathbf{M}^{(N)}\right]$ and $\mathbf{w}$ is $P \times 1$ Gaussian noise vector with zero mean and covariance $\mathbf{C}_w$. The joint ML estimate is given as [34]

$$\hat{\mathbf{h}} = \left(\mathbf{M}^H \mathbf{C}_w^{-1} \mathbf{M}\right)^{-1} \mathbf{M}^H \mathbf{C}_w^{-1} \mathbf{y} \ . \tag{3.34}$$

By using the square root raised cosine receive filtering the noise samples are uncorrelated, when sampled at the symbol rate as shown in Section 2.1.3. Thereby the covariance matrix $\mathbf{C}_w$ reduces to diagonal matrix and the ML estimator is simplified to

$$\hat{\mathbf{h}} = \left(\mathbf{M}^H \mathbf{M}\right)^{-1} \mathbf{M}^H \mathbf{y} \ . \tag{3.35}$$

The estimation is executed in the same form as the single channel estimation in Eq. (3.7), but the matrix $\mathbf{M}$ is now redefined to represent several training sequences. Hence, the resulting channel estimate $\hat{\mathbf{h}}$ also includes several impulse responses.

3.2 Signal detection

In the following sections a number of detection algorithms are presented including the optimal MLSE with its soft-output modifications and per-survivor approach, the optimal MAP algorithm, low-complexity detectors (reduced-state equalisers and linear filters) and multichannel detection. Extensive reviews on the detector structures can be found e.g., in [5],[17],[65].

3.2.1 Trellis definitions

Multipath transmission channels can be described by a trellis diagram, which illustrates states and transitions of the Markov source [35]. For channels with memory the *state* S_k is defined by L successive symbols $S_k \equiv (a_{k-L+1}, a_{k-L+2}, \ldots, a_k)$, where L is the channel memory length $(L \geq 1)$. The *transition* ξ_k is related to the current symbol a_k determining the transition from the old state S_{k-1} to the new state S_k, i.e., $\xi_k \equiv (S_{k-1}, S_k) = (a_{k-L}, a_{k-L+1}, \ldots, a_k)$. *Additive branch metric* $\gamma_k(\xi_k)$ is a metric related to a single input symbol y_k computed in log-domain as follows

$$\gamma_k(\xi_k) = \left\| y_k - \sum_{l=0}^{L} a_{k-l} h_l \right\|^2 , \tag{3.36}$$

where the transition ξ_k determines the states S_{k-1} and S_k and hence the symbols $a_{k-L}, a_{k-L+1}, \ldots, a_k$. *Additive path metric* (APM) $\Gamma_k(S_0, S_1, \ldots, S_k)$ is the superposition of the branch metrics that belong to the certain trellis path through the states $S_0, S_1, \ldots, S_k$. Hence, the APM is calculated by

$$\Gamma_k(S_0, S_1, \ldots, S_k) = \sum_{j=1}^{k} \left\| y_j - \sum_{l=0}^{L} a_{j-l} h_l \right\|^2 . \tag{3.37}$$

3.2.2 Maximum Likelihood Sequence Estimation (MLSE)

The MLSE problem is to find the sequence of data symbols $\hat{\mathbf{a}}$ which maximises the likelihood function as follows

$$\hat{\mathbf{a}} = \arg\max_{\mathbf{a}} \; p(\mathbf{y}|\mathbf{a}) . \tag{3.38}$$

The receive filter is a square root raised cosine (SQRC) filter followed by a symbol-rate sampler, which provides uncorrelated noise samples [46]. The received discrete-time symbol-spaced signal is denoted by $\mathbf{y}$. The MLSE problem can be recursively solved by the Viterbi algorithm (VA) [21],[73], which selects the data sequence $\mathbf{a} \equiv (a_1 \quad a_2 \quad \cdots \quad a_K)^T$ minimising the Euclidean distance metric

$$D(\mathbf{y},\mathbf{a}) = \sum_k \left\| y_k - \sum_{l=0}^{L} a_{k-l} h_l \right\|^2 \quad . \tag{3.39}$$

Using the definition (3.37) of the APM in trellis the Euclidean distance metric can be expressed as

$$D(\mathbf{y},\mathbf{a}) = \Gamma_K(S_0, S_1, \ldots, S_K) \quad . \tag{3.40}$$

The APM values are updated recursively by adding the metrics of the starting state and corresponding transition as follows

$$\Gamma_k(S_0, S_1, \ldots, S_k) = \chi_{k-1}(S_{k-1}) + \gamma_k(S_{k-1}, S_k) \tag{3.41}$$

after which the survivor metric $\chi_k(S_k)$ for each state S_k is selected as

$$\chi_k(S_k) = \min_{S_{k-1}} \Gamma_k(S_0, S_1, \ldots, S_{k-1}, S_k) \quad . \tag{3.42}$$

After the final time instant k=K the minimum APM of all trellis paths, $\min_{S_K} \chi_K(S_K)$, is selected. This minimum path corresponds to the most likely sequence of transmitted symbols. As the original VA was introduced for decoding convolutional codes [73], it only generates a sequence of hard decisions as output. In the next sections we present two modifications of the original VA that produce also reliability information.

3.2.3 Soft Output Viterbi Algorithm (SOVA)

The SOVA is a modified VA, which delivers the ML-path (hard decisions), but also soft decisions about the symbols along that path [30],[33]. The SOVA consists of the conventional hard-output VA and a soft-deciding unit working with a decision delay D, which is usually chosen as D=5L. The reliability value is calculated by comparing metrics of the two most likely merging paths that have different bits at a particular time instant, i.e., the ML-path and the closest competing path. It usually takes much longer than the trellis memory L before all survivors merge. Therefore in the SOVA survivor paths with D latest symbols for each state are saved and soft outputs within that window have to be updated due to different path histories. This increases the memory requirement and complexity of the algorithm compared to the basic VA.

3.2.4 Soft Output Viterbi Equaliser (SOVE)

The SOVA can be further simplified to the SOVE, where the decision delay is set equal to trellis memory, i.e., D=L [36],[54]. Thus the SOVE operates like the hard-decision VA without any extra memory requirements as it always assumes the shortest error path. The soft output is given in the form of log-likelihood (LLR)

$$\lambda(a_k) = \ln \frac{p(\mathbf{y} \mid a_k = +1)}{p(\mathbf{y} \mid a_k = -1)} \quad . \tag{3.43}$$

The SOVE output is the difference between two metric minima as follows

$$\lambda(a_{k-L}) = \min_{S_k \mid a_{k-L}=+1} \left[\chi_k(S_k) + \gamma_{k+1}(S_k, S_{k+1}) \right] - \min_{S_k \mid a_{k-L}=-1} \left[\chi_k(S_k) + \gamma_{k+1}(S_k, S_{k+1}) \right] \tag{3.44}$$

and the classical survivor selection is

$$\chi_{k+1}(S_{k+1}) = \min_{S_k} \Gamma_{k+1}\left(S_0, S_1, \ldots, S_{k+1}\right) \quad . \tag{3.45}$$

Eq. (3.44) can be extended for *M*-ary symbols by calculating separate differences for each bit within the symbol.

The drawback of the SOVE is that the detection performance is degraded due to truncated survivors. However, in fading multipath channels the simple SOVE achieves performance close to optimum MAP performance, because the survivors merge quite quickly in the trellis [36]. Theoretically, the SOVA and SOVE produce the same outputs if there is no more than single bit difference in the competing trellis paths that are compared. This arises in the high SNR range, but with low SNR the competing paths can differ during several bits and the SOVE becomes suboptimal compared to the SOVA [54]. To overcome this suboptimality improved SOVE scheme is proposed in [40], where the memory length is slightly expanded to achieve more reliable soft information.

3.2.5 Reduced State Sequence Estimation (RSSE)

The MLSE-VA requires M^L signal states to detect *M*-ary symbols and therefore it becomes impractical for multilevel modulations due to the excessive trellis size. However, a suboptimal solution with much lower complexity is achieved by the RSSE algorithm using *set partitioning* and/or *decision feedback*, which reduce the number of trellis states and truncate the trellis memory, respectively.

As given in Section 3.2.1 the MLSE trellis state is $S_k \equiv \left(a_{k-L+1}, a_{k-L+2}, \ldots, a_k\right)$, where a_k denotes a complex-valued *M*-ary symbol. In the set partitioning the *M* possible constellation points are collected into subsets, thus more than one symbol value can be within each subset. Furthermore, the n^{th} MLSE state element a_{k-n} is partitioned into J_n subsets, so that $J_n \in [1, M]$ and $J_0 \geq J_1 \geq \cdots \geq J_L$. Hence, the reduced number of states in the subset trellis is $J_0 \times J_1 \times \cdots \times J_L$. The subset trellis state is defined as $T_k \equiv \left(t_{k-L+1}, t_{k-L+2}, \ldots, t_k\right)$, where each element $t_{k-n} \in [1, J_n]$. When the trellis size is actually reduced, i.e., $J_0 < M$, there are several parallel transitions arriving at a state, each corresponding to a certain subset. The branch metric for the RSSE is slightly modified as the subset trellis state does not correspond to unique symbol history due to parallel transitions. The branch metric is [20]

$$\gamma_k(\xi_k) = \left\| y_k - a_k h_0 - \sum_{l=1}^{L} \hat{a}_{k-l}^{(s)} h_l \right\|^2 \quad , \tag{3.46}$$

where $\hat{a}_{k-l}^{(s)}$ is the symbol estimate of a_{k-l} in the subset state *s*. Hence, each subset trellis node should have the state estimate $\hat{S}_k \equiv (\hat{a}_{k-L+1}, \hat{a}_{k-L+2}, \ldots, \hat{a}_k)$ available and therefore the history information is to be stored in all subset states at each time instant.

The set partitioning should optimise the RSSE trade-off between complexity and performance. Because subset states include multiple MLSE states, certain reduced state paths merge earlier than MLSE paths. It is beneficial to be able to separate the merging paths reliably, which is achieved by maximising the intrasubset Euclidean distance [20]. The structure of the RSSE is flexible, as one can control the complexity/performance trade-off by selecting suitable number of states for a certain application. If $J_n = 1$ for all *n*, the RSSE reduces to the Decision Feedback Equaliser (DFE) and with $J_n = M$ for all *n*, the equaliser becomes the MLSE.

For binary modulation the possible receiver structures obtained by the set partitioning are the MLSE $(J_0 = 2)$ and the DFE $(J_0 = 1)$. Hence, the partitioning is useless in binary transmission [20]. However, the complexity can be reduced by truncating the trellis memory by the Decision Feedback Sequence Estimation (DFSE) algorithm, which considers only *D*+1 most recent symbols in the trellis and the earlier *L-D* symbols are taken into account through the embedded decision-feedback structure [16]. Hence, the number of states can be reduced to M^D. In the MLSE-VA the trellis is spanned by all the *L* symbols in (3.39), but in the DFSE branch metric is composed from two parts so that (3.39) is replaced by [16]

$$\gamma_k(\xi_k) = \left\| y_k - \sum_{l=0}^{D} a_{k-l} h_l - \sum_{l=D+1}^{L} \hat{a}_{k-l}^{(s)} h_l \right\|^2 \quad . \tag{3.47}$$

The last term is the feedback information of the trellis state *s*, which is extracted from the survival path $\hat{\mathbf{a}}^{(s)}$ leading to the state. As a result, the trellis is spanned over *D*+1 symbols and the feedback information is updated for each state as the recursion proceeds.

Since only first few channel taps form the trellis structure, the DFSE works properly only with minimum phase channel, where the energy is concentrated to the beginning of the impulse response. The same applies to set partitioning as the RSSE selects the subsets based on the first few channel taps and the state estimates cannot be changed later. However, practical fading channels in the GSM and EDGE systems may instantaneously be non-minimum phase and therefore a separate feedforward transversal filter in front of the detector is needed to convert the channel into minimum phase [16]. The prefilter can be designed by the same methods as the DFE feedforward filter, e.g., by minimising the mean-squared error [46] or maximising the signal-to-noise ratio [26] at the input of the DFE decision device.

3.2.6 MLSE with Recursive Least Square (RLS) adaptation

This section presents an equalisation scheme that is a combination of Recursive Least Squares (RLS) adaptive algorithm and Maximum-Likelihood Sequence Estimation (MLSE), which is depicted in Figure 9. The scheme is introduced in detail in [24],[75]. This method is capable of tracking fast time-varying mobile channels during transmission bursts. The RLS-MLSE minimises the distance metric

$$D(\mathbf{y},\mathbf{a}) = \sum_k \left\| y_k - \sum_{l=0}^{L} a_{k-l} h_l \right\|^2 \quad , \tag{3.48}$$

where both the data symbols a_k and channel taps h_l are unknown for the receiver. The equalisation problem is to detect the symbol sequence and simultaneously estimate the time-varying channel parameters.

In the RLS-MLSE the initial channel estimates are first obtained in the training mode by using the known training symbols in the RLS channel estimation algorithm as described in Section 3.1.3. After that the parameter estimation continues during the MLSE detection in the tracking mode. The adaptation is done separately for each trellis state *s*, hence the candidate symbol vectorsn $\mathbf{a}_k^{(s)} \equiv \left(a_k^{(s)}, a_{k-1}^{(s)}, ..., a_{k-L}^{(s)}\right)^T$ are used to update the channel estimate $\hat{\mathbf{h}}^{(s)}$ with the RLS algorithm [24]. The transversal filter generates the replica of the desired signal using the candidate symbols and newly estimated channel taps for the current state *s*. The replica is subtracted from the received signal to obtain the estimation error

$$e_k^{(s)} = y_k - \sum_{l=0}^{L} a_{k-l}^{(s)} \hat{h}_l^{(s)} \quad . \tag{3.49}$$

The MLSE uses the squared error term $\left|e_k^{(s)}\right|^2$ as the branch metric. In order to update the channel parameters the MLSE has to output the candidate decisions $\mathbf{a}_k^{(s)}$ for each state. The RLS-MLSE structure removes ISI effectively and it can be extended to remove Co-Channel Interference (CCI) as well [77].

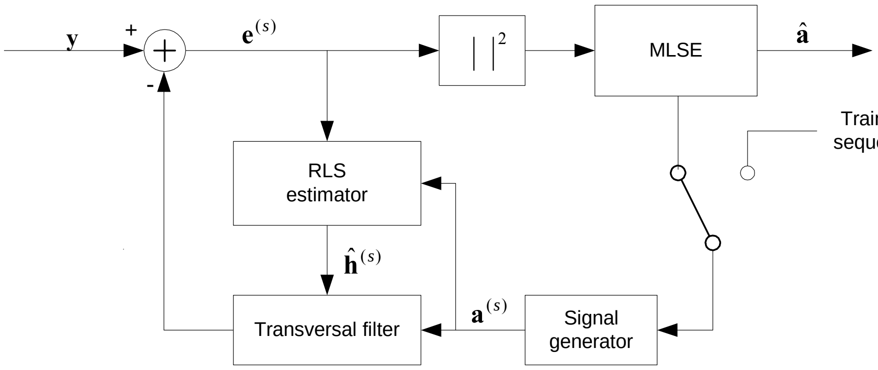

Figure 9. Block diagram of RLS-MLSE equaliser [77].

3.2.7 Maximum a Posteriori (MAP) detection

The MAP algorithms calculate *a posteriori* probabilities (APP) of states and transitions of the Markov source using information from the whole received signal sequence. Maximising the joint APP of the whole sequence would require the calculation of $\Pr(a_1, a_2, \ldots, a_K \mid \mathbf{y})$ for each M^K symbol combinations and selecting the highest probability, which is impossible in practice. Chang and Hancock present in [12] a MAP equaliser algorithm for removing ISI, which provides optimum joint decision for a trellis state $S_k \equiv (a_{k-L+1}, a_{k-L+2}, \ldots, a_k)$, hence L symbol decisions are obtained simultaneously. The joint APP $\Pr(a_{k-L+1}, a_{k-L+2}, \ldots, a_k \mid \mathbf{y})$ is calculated for every M^L state and the maximum APP is selected. Furthermore, a recursive formulation is given for the APP calculation, which the authors call as the Optimum Sequential Receiver (OSR).

There are two reasons why the OSR is not very useful in practical mobile radio applications.

1) The OSR is optimised to provide good hard decisions. However, the MLSE-VA provides better hard output performance with lower complexity.

2) The OSR does not provide APP for each bit separately. The bitwise APP is of great importance when detector is concatenated with the channel decoder as the coded performance is clearly improved by using soft output detection instead of hard outputs.

The MAP algorithm providing bitwise APP is introduced by Bahl, Cocke, Jelinek and Raviv in [3] and it is named according to the authors as the BCJR algorithm. The original proposal is for decoding linear codes, but the BCJR can be used in the detection as well. The objective of the BCJR-MAP algorithm is to minimise the bit error probability.

The BCJR algorithm selects the symbol a_k at time instant k, which maximises the following APP

$$\hat{a}_k = \arg\max_{a_k} \left[\Pr\left(a_k \middle| \mathbf{y}\right) \right] \quad . \tag{3.50}$$

The algorithm utilises the joint probabilities for each transition, which are split as follows [3]

$$\Pr\left(S_{k-1} = s', S_k = s, \mathbf{y}_1^K\right) = \alpha_{k-1}(s')\gamma_k(s', s)\beta_k(s) \tag{3.51}$$

where the probability functions are

$$\alpha_k(s) = \Pr\left(S_k = s, \mathbf{y}_1^k\right) \tag{3.52}$$

$$\beta_k(s) = \Pr\left(\mathbf{y}_{k+1}^K \middle| S_k = s\right) \tag{3.53}$$

$$\gamma_k(s', s) = \Pr\left(S_k = s, y_k \mid S_{k-1} = s'\right) \quad . \tag{3.54}$$

Furthermore, in the presence of AWGN and ISI channel the last term is given by

$$\gamma_k(s',s)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2\sigma^2}\left\|y_k-\sum_{l=0}^{L}a_{k-l}h_l\right\|^2\right\} \tag{3.55}$$

where the trellis states (s',s) determine the hypothetical symbols $a_{k-L},...,a_k$. Probabilities $\alpha_k(s)$ and $\beta_k(s)$ can be calculated recursively in forward and backward directions as follows

$$\alpha_k(s)=\sum_{s'}\alpha_{k-1}(s')\gamma_k(s',s) \tag{3.56}$$

$$\beta_k(s)=\sum_{s'}\beta_{k+1}(s')\gamma_{k+1}(s,s') \quad . \tag{3.57}$$

Finally, the BCJR calculates APP values by adding up transition probabilities as follows [3]

$$\Pr\left(a_k=i\,|\,\mathbf{y}_1^K\right)=\sum_{\substack{(s',s)\\ a_k=i}}\alpha_{k-1}(s')\gamma_k(s',s)\beta_k(s)\ ,\ i\in\{0,1\} \quad . \tag{3.58}$$

In practice the backward recursion is first calculated for a block of data and state probabilities $\beta_k(s)$ are stored in memory. Then forward recursion is performed, where probabilities $\alpha_k(s)$ are computed and finally decisions (APP) can be calculated.

The optimum MAP algorithm saves multiplicative transition probabilities in the trellis, which is computationally difficult. The log-MAP algorithm computes all branch metrics in the log-domain, which leads to cumulative state metrics [38],[55]. There is no loss in the performance, but the implementation is much easier. The further simplified Max-Log-MAP provides a straightforward implementation at the cost of slight degradation compared to the optimum MAP [55]. It utilises approximation

$$\ln\left(e^{\delta_1}+\cdots+e^{\delta_n}\right)\approx\max_i\ \delta_i \tag{3.59}$$

to calculate the state metrics, where $\delta_1,\delta_2,...,\delta_n$ represent transition log-probabilities merging into a trellis state. Hence, for a certain state Max-Log-MAP selects the maximum incoming transition metric instead of computing the total log-probability. At high SNR region the correct transition dominates the others, i.e., one of the transition probabilities is clearly higher than the others, and the approximation (3.59) is very accurate. The output of the Max-Log-MAP algorithm is given as [55]

$$\begin{aligned}\lambda(a_k)=&\max_{\substack{(s',s)\\ a_k=+1}}\left\{\ln\alpha_{k-1}(s')+\ln\gamma_k(s',s)+\ln\beta_k(s)\right\}\\ &-\max_{\substack{(s',s)\\ a_k=-1}}\left\{\ln\alpha_{k-1}(s')+\ln\gamma_k(s',s)+\ln\beta_k(s)\right\}\end{aligned} \tag{3.60}$$

where

$$\ln\alpha_k(s)=\max_{s'}\left\{\ln\gamma_k(s',s)+\ln\alpha_{k-1}(s')\right\} \tag{3.61}$$

$$\ln \beta_k(s) = \max_{s'}\{\ln \beta_{k+1}(s') + \ln \gamma_{k+1}(s, s')\} \ . \tag{3.62}$$

Detailed descriptions of the Log-MAP and Max-Log-MAP are given in Appendix I.

Due to the backward recursion the optimum MAP algorithm requires roughly two times more metrics calculations than the MLSE-VA. However, as the MAP calculates APP values, it provides more reliable soft information, which is beneficial in coded systems. Therefore the choice between MLSE and MAP depends on the available processing power and required performance. Moreover, there are log-MAP algorithms omitting the backward recursion altogether [36] and sophisticated VA versions with improved soft outputs [30]. These offer further compromises between complexity and performance.

3.2.8 MAP detection using *a priori* information

In the iterative receiver schemes that are introduced in Chapter 4 detection in the presence of *a priori* information is needed. The MAP structure is suitable for this purpose, as it inherently calculates *a posteriori* probabilities and it is easy to incorporate *a priori* probabilities $\lambda^a(\mathbf{a})$ into the algorithm. The MAP equaliser is maximising the conditional probability $\Pr(a_k \mid \mathbf{y})$ for the data symbol a_k, which can be expressed after Bayes' rule as

$$\Pr(a_k \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid a_k)\Pr(a_k)}{p(\mathbf{y})} \ . \tag{3.63}$$

An equivalent problem is to solve

$$\hat{a}_k = \arg\max_{a_k} p(\mathbf{y} \mid a_k)\Pr(a_k) \tag{3.64}$$

as $p(\mathbf{y})$ does not depend on a_k. Thus the additive branch metric γ_k^a in the presence of *a priori* information is defined by

$$\gamma_k^a(s', s) = \gamma_k^a(a_k) = p(y_k \mid a_k)\Pr(a_k) \ . \tag{3.65}$$

The conditional observation probability $p(y_k \mid a_k)$ is computed in the Max-Log-MAP equaliser (see Appendix I) by

$$\ln p(y_k | a_k) = \ln \gamma_k(a_k) \cong -\frac{1}{2\sigma^2}\left\| y_k - \sum_{l=0}^{L} a_{k-l} h_l \right\|^2 \ . \tag{3.66}$$

Furthermore, in the case of multilevel M-ary modulation the *a priori* probability is given by

$$\Pr(a_k = \alpha) = \Pr\left(b_{k,1} = \beta_1, b_{k,2} = \beta_2, \ldots, b_{k,\log_2 M} = \beta_{\log_2 M}\right) = \prod_{m=1}^{\log_2 M} \Pr\left(b_{k,m} = \beta_m\right) \tag{3.67}$$

assuming ideal decorrelation between individual bits within one symbol. Using the notation in Eq. (3.66) and relation (3.67) we can represent Eq. (3.65) in the log-domain by

$$\ln \gamma_k^a(a_k) = \ln \gamma_k(a_k) + \sum_{m=1}^{\log_2 M} \ln \Pr\left(b_{k,m}\right) \ . \tag{3.68}$$

The equaliser output for the single bit $b_{k,j}$ is

$$\lambda_{eq}(b_{k,j}) = \max_{\substack{(s',s)\\ b_{k,j}=1}}\left\{\ln\alpha_{k-1}(s') + \ln\gamma_k^a(s',s) + \ln\beta_k(s)\right\} - \max_{\substack{(s',s)\\ b_{k,j}=0}}\left\{\ln\alpha_{k-1}(s') + \ln\gamma_k^a(s',s) + \ln\beta_k(s)\right\} \quad . \tag{3.69}$$

By using the relation (3.68) we obtain the general form

$$\lambda_{eq}(b_{k,j}) = \max_{\substack{(s',s)\\ b_{k,j}=1}}\left\{\ln\alpha_{k-1}(s') + \ln\gamma_k(s',s) + \sum_{m=1}^{\log_2 M}\ln\Pr\left(b_{k,m}\right) + \ln\beta_k(s)\right\} - \max_{\substack{(s',s)\\ b_{k,j}=0}}\left\{\ln\alpha_{k-1}(s') + \ln\gamma_k(s',s) + \sum_{m=1}^{\log_2 M}\ln\Pr\left(b_{k,m}\right) + \ln\beta_k(s)\right\} \quad . \tag{3.70}$$

In conventional equalisation there is no *a priori* information available, thus $\Pr(b_{k,m}) = 0.5$ for all *k,m*. In Eq. (3.70) the terms $\sum \ln\Pr\left(b_{k,m}\right)$ are constants, which do not depend on the trellis transitions and therefore they cancel out each other. The output is then

$$\lambda_{eq}(b_{k,j}) = \max_{\substack{(s',s)\\ b_{k,j}=1}}\left\{\ln\alpha_{k-1}(s') + \ln\gamma_k(s',s) + \ln\beta_k(s)\right\} - \max_{\substack{(s',s)\\ b_{k,j}=0}}\left\{\ln\alpha_{k-1}(s') + \ln\gamma_k(s',s) + \ln\beta_k(s)\right\} \quad , \tag{3.71}$$

which is similar to Eq. (3.60). However, when bits are not equally likely *a priori*, Eq. (3.70) takes the form

$$\lambda_{eq}(b_{k,j}) = \lambda_{eq}^{ext}(b_{k,j}) + \lambda^a(b_{k,j}) \tag{3.72}$$

where the extrinsic information from the equaliser is

$$\lambda_{eq}^{ext}(b_{k,j}) = \max_{\substack{(s',s)\\ b_{k,j}=1}}\left\{\ln\alpha_{k-1}(s') + \ln\gamma_k(s',s) + \sum_{m=1,m\neq j}^{\log_2 M}\ln\Pr\left(b_{k,m}\right) + \ln\beta_k(s)\right\} - \max_{\substack{(s',s)\\ b_{k,j}=0}}\left\{\ln\alpha_{k-1}(s') + \ln\gamma_k(s',s) + \sum_{m=1,m\neq j}^{\log_2 M}\ln\Pr\left(b_{k,m}\right) + \ln\beta_k(s)\right\} \tag{3.73}$$

and *a priori* information is

$$\lambda^a(b_{k,j}) = \ln\frac{\Pr\left(b_{k,j}=1\right)}{\Pr\left(b_{k,j}=0\right)} \quad . \tag{3.74}$$

According to Eq. (3.61) the forward recursion for the state probability is calculated as

$$\begin{aligned}\ln \alpha_k(s) &= \max_{s'}\left\{\ln \gamma_k^a(s',s) + \ln \alpha_{k-1}(s')\right\} \\ &= \max_{s'}\left\{\ln \gamma_k(s',s) + \sum_{m=1}^{\log_2 M} \ln \Pr\left(b_{k,m}\right) + \ln \alpha_{k-1}(s')\right\}\end{aligned} \quad . \tag{3.75}$$

Respectively, the backward recursion is given by

$$\ln \beta_k(s) = \max_{s'}\left\{\ln \beta_{k+1}(s') + \ln \gamma_{k+1}(s,s') + \sum_{m=1}^{\log_2 M} \ln \Pr\left(b_{k+1,m}\right)\right\} \quad . \tag{3.76}$$

The effect of *a priori* information $\Pr(b_{k,m})$ on the detection process is interesting. Firstly, it has an influence on the path selection, when the state probabilities are calculated according to Eq. (3.75) and (3.76). Hence, *a priori* information makes certain trellis paths more dominant. Secondly, it appears in the extrinsic output (3.73) directly in the form of $\sum_{m=1,m\neq j}^{\log_2 M} \Pr\left(b_{k,m}\right)$. This term only arises in higher level modulations (*M*>2), but disappears in the binary case. For instance, since an 8-PSK modulated symbol consists of three bits, the extrinsic output for the first bit is affected by the two other bits of the symbol. As part of the *a priori* information is present in the extrinsic output, the higher level modulations are expected to gain more from the iterative processing.

3.2.9 Joint detection of multiple signals

The MLSE solution for a single signal as given in Section 3.2.2 is extended in [19] to the problem of detecting N signals simultaneously. The joint MLSE problem is to find those sequences of data symbols $\hat{\mathbf{a}}^{(n)}$, $1 \le n \le N$, which maximise the following likelihood function

$$\hat{\mathbf{a}} = \underset{\mathbf{a}^{(1)},\mathbf{a}^{(2)},\ldots,\mathbf{a}^{(N)}}{\arg\max} \Pr\left(\mathbf{y} \middle| \mathbf{a}^{(1)},\mathbf{a}^{(2)},\ldots,\mathbf{a}^{(N)}\right) \quad . \tag{3.77}$$

The recursive Viterbi Algorithm [22] selects the data sequences $\mathbf{a}^{(n)} = \left(a_1^{(n)} \;\; a_2^{(n)} \;\; \cdots \;\; a_K^{(n)}\right)^T$, $1 \le n \le N$ that minimise the Euclidean distance metric

$$D(\mathbf{y},\mathbf{a}) = \left\| \mathbf{y} - \sum_{n=1}^{N} \mathbf{A}^{(n)}\mathbf{h}^{(n)} \right\|^2 = \sum_k \left| y_k - \sum_{n=1}^{N}\sum_{l=0}^{L} a_{k-l}^{(n)} h_l^{(n)} \right|^2 \quad . \tag{3.78}$$

Using this distance metrics the recursion proceeds as described in Section 3.2.2 for single signal. Assuming that the desired user transmits symbols $\mathbf{a}^{(1)}$, only decisions $\hat{\mathbf{a}}^{(1)}$ are of interest and other decisions $\hat{\mathbf{a}}^{(2)},\ldots,\hat{\mathbf{a}}^{(N)}$, which describe the interfering signals, can be discarded.

The joint BCJR algorithm selects simultaneously the symbols $a_k^{(1)}, a_k^{(2)},\ldots,a_k^{(N)}$ at time instant *k*, which maximise the following APP

$$\hat{a}_k^{(n)} = \arg\max_{a_k^{(n)}} \left[\Pr\left(a_k^{(n)} \middle| \mathbf{y}\right)\right], \; 1 \le n \le N \quad . \tag{3.79}$$

The joint Max-Log-MAP equaliser calculates transition metrics as follows

$$\ln\gamma_k(s',s) \cong -\frac{1}{2\sigma^2}\left\| y_k - \sum_{n=1}^{N}\sum_{l=0}^{L} a_{k-l}^{(n)} h_l^{(n)} \right\|^2 \quad . \tag{3.80}$$

The output of the detector is given as the logarithmic ratio of the APP values of symbols $a_k^{(1)}$ as follows

$$\lambda(a_k^{(1)}) = \ln\frac{\Pr(a_k^{(1)} = +1 \mid \mathbf{y})}{\Pr(a_k^{(1)} = -1 \mid \mathbf{y})} \quad . \tag{3.81}$$

Furthermore, the output is expressed as

$$\begin{aligned} \lambda\left(a_k^{(1)}\right) = & \max_{\substack{(s',s) \\ a_k^{(1)}=+1}} \left\{\ln\alpha_{k-1}(s') + \ln\gamma_k(s',s) + \ln\beta_k(s)\right\} \\ & - \max_{\substack{(s',s) \\ a_k^{(1)}=-1}} \left\{\ln\alpha_{k-1}(s') + \ln\gamma_k(s',s) + \ln\beta_k(s)\right\} \end{aligned} \tag{3.82}$$

with

$$\ln\alpha_k(s) = \max_{s'}\left\{\ln\gamma_k(s',s) + \ln\alpha_{k-1}(s')\right\} \tag{3.83}$$

$$\ln\beta_k(s) = \max_{s'}\left\{\ln\beta_{k+1}(s') + \ln\gamma_{k+1}(s,s')\right\} \quad . \tag{3.84}$$

Since the joint detector takes simultaneously N signals into account, the trellis contains M^{NL} states, where M is the number of modulation levels and L is the length of the channel memory in symbol periods. As a consequence, the joint detection with full trellis is feasible for two binary modulated signals or otherwise suboptimal reduced-state algorithms are needed to keep the receiver complexity at a reasonable level.

3.2.10 Soft Canceller followed by Minimum Mean Square Error filter

The soft cancelling (SC) of the Intersymbol Interference (ISI), which is followed by the Minimum Mean Square Error (MMSE) linear filter is introduced by Reynolds and Wang in [53]. The receiver complexity is further reduced by approximations in [41] and the SC-MMSE structure is applied to Multiple-Input Multiple-Output (MIMO) systems in [1],[2].

The SC-MMSE receiver utilises diversity channels and therefore either oversampling or multiple antennas are required. Assuming that M samples per symbol are taken we define the following vectors and channel matrix

$$\underset{M(L+1)\times 1}{\mathbf{y}_k} \equiv \left(y_{k,1}, \ldots, y_{k,M}, \ldots, y_{k-L,1}, \ldots, y_{k-L,M}\right)^T, \tag{3.85}$$

$$\underset{(2L+1)\times 1}{\mathbf{a}_k} \equiv \left(a_{k+L}, a_{k+L-1}, \ldots, a_{k-L}\right)^T \tag{3.86}$$

$$\underset{M\times 1}{\mathbf{h}_i} \equiv \left(h_{i,0}, h_{i,1}, \ldots, h_{i,M-1}\right)^T \tag{3.87}$$

$$\underset{M(L+1)\times(2L+1)}{\mathbf{H}} \equiv \begin{bmatrix} \mathbf{h}_0 & \mathbf{h}_1 & \cdots & \mathbf{h}_L & & & \mathbf{0} \\ & \mathbf{h}_0 & \mathbf{h}_1 & & \mathbf{h}_L & & \\ & & \ddots & & & \ddots & \\ \mathbf{0} & & & \mathbf{h}_0 & \mathbf{h}_1 & \cdots & \mathbf{h}_L \end{bmatrix} . \tag{3.88}$$

Using the *a priori* information $\lambda^a(a_k)$ obtained from the channel decoder, the MMSE detector calculates a soft estimate for the symbol *k* as follows [53]

$$\widetilde{a}_k = \tanh\left[\frac{\lambda^a(a_k)}{2}\right] , \tag{3.89}$$

which are used to create soft replica of the ISI components $\sum_{l=1}^{L} \widetilde{a}_{k-l} h_l$. The soft estimates are collected into the vector $\widetilde{\mathbf{a}}_k \equiv \left(\widetilde{a}_{k+L}, \ldots, \widetilde{a}_{k+1}, 0, \widetilde{a}_{k-1}, \ldots, \widetilde{a}_{k-L}\right)^T$. The estimated ISI part is removed from the received signal giving the oversampled residual signal vector [53]

$$\mathbf{e}_k = \mathbf{y}_k - \mathbf{H}\widetilde{\mathbf{a}}_k \ . \tag{3.90}$$

This method is called Soft Cancellation (SC) of Intersymbol Interference (ISI) and it can be extended to the cancellation of multiple access interference (MAI) as well [1],[2]. The SC-MMSE receiver structure is illustrated in Figure 10.

The adaptive linear MMSE filter is used to suppress ISI residuals after soft cancellation step. The values for the $M(L+1)\times 1$ filter tap vector $\mathbf{w}_k$ are determined by minimising the mean square error (MSE) between the filter output and the desired symbol a_k as follows

$$\mathbf{w}_k = \arg\min_{\mathbf{w}_k} \left\| \mathbf{w}_k^H \mathbf{e}_k - a_k \right\|^2 \ . \tag{3.91}$$

The solution is given by [2]

$$\mathbf{w}_k = \left[\mathbf{H}\mathbf{\Lambda}_k \mathbf{H}^H + \sigma^2 \mathbf{I}\right]^{-1} \overleftarrow{\mathbf{h}} \ , \tag{3.92}$$

where

$$\overleftarrow{\mathbf{h}} \equiv \left(h_{L,0}, \ldots, h_{L,M-1}, \ldots, h_{0,0}, \ldots, h_{0,M-1}\right)^T \tag{3.93}$$

$$\mathbf{\Lambda}_k \equiv diag\left(1-\widetilde{a}_{k+L}^2, \ldots, 1-\widetilde{a}_{k+1}^2, 1, 1-\widetilde{a}_{k-1}^2, \ldots, 1-\widetilde{a}_{k-L}^2\right) \ . \tag{3.94}$$

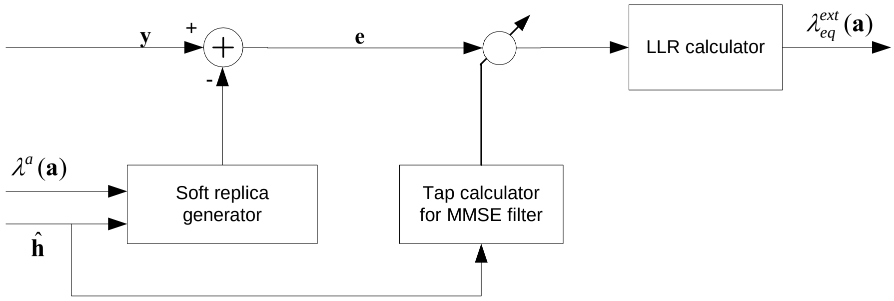

Figure 10. SC-MMSE signal detection [1].

The extrinsic information from the MMSE output can be derived as [53]

$$\lambda_{eq}^{ext}(a_k) = \frac{4\,\mathrm{Re}[z_k]}{1-\mu_k} \quad , \tag{3.95}$$

where the filter output is

$$z_k = \mathbf{w}_k^H \mathbf{e}_k \tag{3.96}$$

and

$$\mu_k = \bar{\mathbf{h}}^H \left[\mathbf{H}\boldsymbol{\Lambda}_k\mathbf{H} + \sigma^2\mathbf{I}\right]^{-1}\bar{\mathbf{h}} \quad . \tag{3.97}$$

The computational complexity of the MMSE filter is mainly due to the matrix inverse in the filter tap calculation (3.91), hence the equaliser complexity is $\mathrm{O}(L^3)$ with the channel memory length L [53]. By assuming that the *a priori* information $\lambda^a(a_k)$ obtained from the decoder is perfect and applying the matrix inverse lemma [32] to the calculation of the filter taps the MMSE filter can be approximated by the channel matched filter (CMF) from the 2nd iteration onwards [41]. The complexity of this approximated SC-MMSE equaliser is $\mathrm{O}(L^2)$ [41].

Significant performance improvement is achieved by the iterative SC-MMSE structure and ISI can be almost eliminated in fast-varying Rayleigh fading channel [53]. Although the performance of the SC-MMSE is worse compared to the SOVA at the first iteration, the performance loss can be compensated by further iterations [41]. Furthermore, even the CMF approximation of the MMSE filter is comparable to the original SC-MMSE equaliser after a few iterations [41].

3.3 Summary of receiver algorithms

In this chapter we have presented a number of channel estimation and detection algorithms. Conventional channel estimation methods based on the known training sequence are first described. The Maximum-Likelihood (ML) solution maximises the probability of the received signal conditioned on the transmitted training bits and channel impulse response. The ML estimator is simple to implement and offers reasonable performance. Improved accuracy is obtained by the Linear Minimum Mean Square Error (LMMSE) approach, which minimises the estimator variance. However, the complexity is increased as the LMMSE requires separate matrix inversion for each burst. The channel estimate can also be considered as the taps of the linear filter. The Recursive Least Squares (RLS) algorithm operates symbol-by-symbol adjusting the filter taps based on the received samples and known training bits. We propose channel re-estimation method utilising decoded bits as an extended training sequence. Cramer-Rao Lower Bound (CRLB) is derived for the iterative estimator, which shows that the estimator variance decreases even with unreliable data decisions. Finally the ML channel estimation is extended to estimate simultaneously multiple channels.

The optimum Maximum Likelihood Sequence Estimation (MLSE) algorithm is presented and the recursive solution by the Viterbi Algorithm (VA) is formulated. The original MLSE provides only hard outputs and therefore we present two modifications with the capability of producing soft information. The Soft Output Viterbi Algorithm (SOVA) calculates the bit reliability by comparing the metrics of the ML-path and the second best merging path. As the decision delay in the SOVA is usually quite long, the Soft Output Viterbi Equaliser (SOVE) reduces the delay to the channel memory. As the SOVE considers only the shortest error path, it is suboptimal with low SNR, but achieves the SOVA performance with high SNR. The Reduced State Sequence Estimation (RSSE) provides complexity reduction by means of merging the trellis states (set partitioning) or truncating the channel memory (decision feedback). Per-survivor processing is added to the MLSE by adjusting the channel parameters of each trellis state separately with the RLS algorithm. The channel impulse response is re-calculated during the detection after each received data symbol and thereby along the correct trellis path the estimation accuracy improves.

The optimum Maximum A Posteriori (MAP) detection is discussed and especially the BCJR-MAP algorithm providing symbol-by-symbol soft decisions is presented in detail. Furthermore, practical modifications Log-MAP and Max-Log-MAP are described. The MAP structure accepting *a priori* information is covered in detail, since it is essential in the iterative detection schemes. The next section presents joint detection of multiple signals using the ML and MAP structures. Finally the linear filter equaliser based on the Minimum Mean Squared Error (MMSE) criterion is considered. The MMSE filter adjusts the transversal filter taps to minimise the mean squared error between the filter output and the desired symbol. The

MMSE filter can be effectively utilised in the iterative equaliser structure by cancelling the Intersymbol Interference (ISI) from the received signal. The replica of the ISI components can be generated by using the decoder feedback information. The performance loss of the linear filtering compared to trellis-based equalisation can be compensated by iterations.

4 ITERATIVE RECEIVER CONCEPT FOR CODED SYSTEMS

4.1 Background

The optimum receiver for a coded communication system maximises the joint probability of transmitted coded symbols, information bits and channel impulse response, which is too complex to implement in practice. Therefore a suboptimal solution with separate detection, channel decoding and channel estimation parts is commonly used. However, it is possible to improve the suboptimum performance by exchanging *a priori* information between the independent receiver parts in an iterative fashion. The redundant information added by the convolutional coding provides extrinsic information on the transmitted symbols. This information can be extracted from the decoder output and used as *a priori* information in the re-equalisation in the Turbo Equalisation (TE) technique introduced by Douillard et al. [15]. The TE method is derived from the famous turbo codes [11] by considering the frequency selective channel as a real field convolutional code with rate 1. Since the feedback information is related to symbol probabilities, the feedback consists of soft values and Soft-In/Soft-Out (SISO) channel decoder is therefore needed.

The TE usually assumes ideal or fixed channel state information during the iteration. Nevertheless, iterative data processing can also be extended to channel estimation. The conventional channel estimation is based on the short midamble sequence, thereby the estimation accuracy is not very high. Improved performance is achieved by applying decision-directed iterative channel estimation (ICE) technique, where decoder decisions are used as an extended training sequence in the channel re-estimation. The ICE and TE techniques can be combined in the receiver as considered e.g., in the publications [P1-P2], [P4-P5] of this thesis.

4.2 Principle of turbo detection

The iterative receiver structure is presented in Figure 11. It is convenient to use the Max-Log-MAP algorithm as the SISO equaliser to produce soft outputs in the form of log-likelihood ratio (LLR), which are denoted by $\lambda_{eq}(\mathbf{a})$. The SISO equaliser utilises the *a priori* information from the decoder as described in detail in Section 3.2.8. At the initial iteration there is no feedback information from the channel decoder available, so the LLR values $\lambda_{eq}(\mathbf{a})$ are based only on the received samples $\mathbf{y}$ from the channel. In the transmitter the encoded bits $\mathbf{c}$ are scrambled by the interleaver Π (see Figure 6 in Section 2.1), thus the transmitted symbols are obtained by $\mathbf{a}=\Pi(\mathbf{c})$.

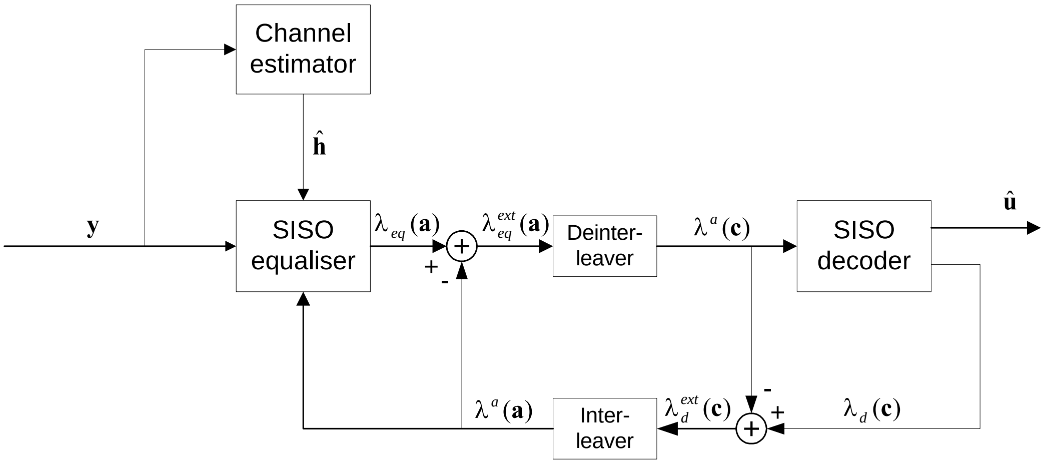

Figure 11. Turbo equaliser structure [15].

At the receiver end the equaliser operates with the symbols **a**, whereas the channel decoder uses the bits **c**. Hence, the LLR values $\lambda_{eq}(\mathbf{a})$ are deinterleaved before providing for the SISO decoder as *a priori* information $\lambda^a(\mathbf{c})$. The decoder calculates new LLR values $\lambda_d(\mathbf{c})$ for the coded bits **c**, since they are needed in the feedback branch to the equaliser.

The decoder output consists of two parts: *intrinsic and extrinsic information* [15]. The intrinsic part is the same as the input stream $\lambda^a(\mathbf{c})$, which is already known beforehand. It is useless to feed the intrinsic part back to the equaliser, as the intrinsic information corresponds to the equaliser output $\lambda_{eq}(\mathbf{a})$ and therefore it only makes the earlier equaliser decisions – right or false – more dominant without correcting any errors. The extrinsic part $\lambda_d^{ext}(\mathbf{c})$ is created by the decoder, which utilises the redundancy due to the coding structure, and it is obtained from the decoder output $\lambda_d(\mathbf{c})$ by bitwise subtraction [15]

$$\lambda_d^{ext}(c_k) = \lambda_d(c_k) - \lambda^a(c_k)\,. \tag{4.1}$$

The turbo equalisation technique is based on the utilisation of this extrinsic information at the next iteration round [15]. So it is passed through the interleaver to the equaliser as *a priori* information $\lambda^a(\mathbf{a})$ on the bit reliabilities. By exploiting this side information in the detection, more reliable decisions are achieved. In the equaliser output the extrinsic information $\lambda_{eq}^{ext}(\mathbf{a})$ is extracted from the output as follows

$$\lambda_{eq}^{ext}(a_k) = \lambda_{eq}(a_k) - \lambda^a(a_k)\,. \tag{4.2}$$

This equaliser information is again used in the SISO decoder to produce new soft outputs and furthermore the new extrinsic information. As soon as this feedback information becomes available, the new iteration round can be started. The number of iterations may depend on the processing power available or the wanted performance improvement. At the final stage, there is no need for the SISO decoder, since only hard decisions $\hat{\mathbf{u}}$ on the information bits are needed.

4.3 Soft-In/Soft-Out (SISO) decoder

At the conventional receiver the channel decoder operates with soft input information, but provides hard outputs. However, the TE scheme requires that probabilistic values are fed back from the decoder to the equaliser and therefore the SISO decoder is needed. The optimum soft outputs are obtained by the MAP decoding algorithms, as they calculate *a posteriori* probabilities while processing data [3].

The BCJR-MAP decoder follows the same principles as the BCJR detection algorithm introduced in Section 3.2.7. In the decoding process the MAP algorithm selects the bit $u_k \in \{0, 1\}$ at time instant *k*, which maximises the following APP [3]

$$\hat{u}_k = \arg\max_{u_k} \Pr\left(u_k \middle| \lambda^a(\mathbf{c})\right) \tag{4.3}$$

using the *a priori* information $\lambda^a(\mathbf{c})$ on the coded bit probabilities provided by the equaliser. The Max-Log-MAP decoder uses the following transition probabilities

$$\gamma_k(s', s) = \Pr\left(S_k = s, \lambda^a(\mathbf{c}_k) \mid S_{k-1} = s'\right) \ , \tag{4.4}$$

which are calculated in the log-domain as follows

$$\ln \gamma_k(s', s) = \sum_{i=1}^{N} \frac{1}{2} \lambda^a\left(c_{k,i}\right) c_{k,i} \ , \tag{4.5}$$

where $c_{k,i}$ is the i^{th} code bit for the information bit u_k and the coding rate is 1/*N*. Finally, the soft feedback information from the decoder consists of log-likelihood values on the coded bits as follows

$$\lambda_d(c_k) = \ln \frac{\Pr(c_k = 1 \mid \lambda^a(\mathbf{c}))}{\Pr(c_k = 0 \mid \lambda^a(\mathbf{c}))} \ . \tag{4.6}$$

Figure 12 shows the inputs and outputs of the SISO decoder. The BCJR decoder uses only the input $\lambda^a(\mathbf{c})$ on coded bits (from the equaliser). The low-complexity algorithm Soft Trellis Decoder (STD) utilises also the feedback information $\lambda^a(\mathbf{u})$ on information bits to control the trellis size adaptively. The STD is introduced in the next section in more detail.

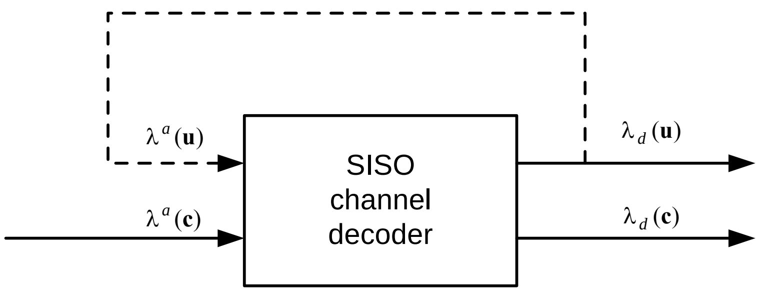

Figure 12. Input/output model of SISO channel decoder. The dashed feedback line from $\lambda_d(\mathbf{u})$ is used in Soft Trellis Decoding (STD).

4.4 Soft Trellis Decoder (STD)

The TE receiver improves the performance at the cost of higher complexity, since the equalisation and decoding are performed several times and the SISO modules are required. There are several possibilities for reducing the trellis-based equaliser complexity like Reduced-State Sequence Estimation (RSSE) [20] based on set partitioning, Decision Feedback Sequence Estimation (DFSE) [16] or Decision Feedback Equaliser (DFE) [14]. All of these can provide significant complexity reduction with only slight performance loss. Even the linear Minimum Mean Squared Error (MMSE) filter can be used as the equaliser if the TE iteration is used to compensate the initial performance loss [53]. However, the main burden for the TE receiver in the EGPRS system is the complex SISO decoder due to the convolutional code with constraint length seven [79].

4.4.1 Algorithm

The decoder complexity can be reduced by adapting the size of the trellis depending on the reliability of the current symbol. Hence, in the presence of symbols with high probability the decoding trellis is significantly reduced. This method is called Soft Trellis Decoding (STD). It is essentially the same as the adaptive T-algorithm, but the STD exploits the soft values from the previous iteration instead of the short-term impulse response of the channel, which is used in the original T-algorithm [57].

The STD requires *a priori* knowledge on the information bit reliabilities

$$\lambda^{a}(u_k) = \ln \frac{\Pr(u_k = 1)}{\Pr(u_k = 0)} \tag{4.7}$$

to adjust the trellis size accordingly. Since the TE receiver performs multiple decoding steps during iterative processing, this *a priori* information can be obtained as a feedback from the previous decoding step as shown in Figure 12. Hence, *a priori* information at iteration round $n+1$ is given by

$$\lambda^{a,(n+1)}(\mathbf{u}) = \lambda_d^{(n)}(\mathbf{u}) \quad , \tag{4.8}$$

where $\lambda_d^{(n)}(\mathbf{u})$ denotes the soft output for the information bits at n^{th} iteration defined as follows

$$\lambda_d^{(n)}(u_k) = \ln \frac{\Pr(u_k = 1 \mid \lambda^{a,(n)}(\mathbf{c}))}{\Pr(u_k = 0 \mid \lambda^{a,(n)}(\mathbf{c}))} \quad , \tag{4.9}$$

where $\lambda^{a,(n)}(\mathbf{c})$ is the *a priori* coded information coming into the channel decoder at n^{th} iteration. From the ratio (4.7) we extract the actual *a priori* probabilities

$$\Pr(u_k = 1) = \frac{\exp[\lambda^a(u_k)]}{1+\exp[\lambda^a(u_k)]} \tag{4.10}$$

$$\Pr(u_k = 0) = \frac{1}{1+\exp[\lambda^a(u_k)]} \quad . \tag{4.11}$$

The STD algorithm requires a predetermined threshold probability δ, by which we can control the trade-off between the receiver performance and complexity reduction. Once *a priori* probability exceeds the threshold δ, i.e., $\Pr(u_k = i) > \delta$, only transition metrics $\gamma_k(s', s)$ corresponding to bit $u_k = i$ need to be calculated and the other metrics are neglected as highly improbable. Also if during the previous time instants some transitions are neglected, the starting state either in forward or backward direction may have probability zero and there is no need to compute the following transition metrics $\gamma_k(s', s)$. Equivalently, we may give the metric value $-\infty$ for a neglected transition in the Max-Log-MAP decoder and summarise the STD rules as follows:

$$\text{IF} \quad \Pr(u_k = i) < 1-\delta \quad \text{THEN} \quad \gamma_k(s', s) = \gamma_k(u_k = i) = -\infty \tag{4.12}$$

$$\text{IF} \quad \alpha_k(s') = -\infty \quad \text{THEN} \quad \gamma_k(s', s) = -\infty \tag{4.13}$$

$$\text{IF} \quad \beta_k(s) = -\infty \quad \text{THEN} \quad \gamma_k(s', s) = -\infty \tag{4.14}$$

$$\text{OTHERWISE} \quad \ln\gamma_k(s', s) = \sum_{i=1}^{N} \frac{1}{2}\lambda^a(c_{k,i})c_{k,i} \quad . \tag{4.15}$$

Figure 13 shows an example of the trellis structure, where at three stages highly improbable bits are encountered. As it can be seen, only a fraction of the original metrics is calculated. The reduction depends both on the signal quality and the threshold δ. For example, if the threshold is set to value 0.99, the decoder assumes all bits having the probability 0.99 or greater as correct and only recalculates the other bits with lower probability. The decoder complexity reduces even further in the following iterations, since the SNR improves during the TE process.

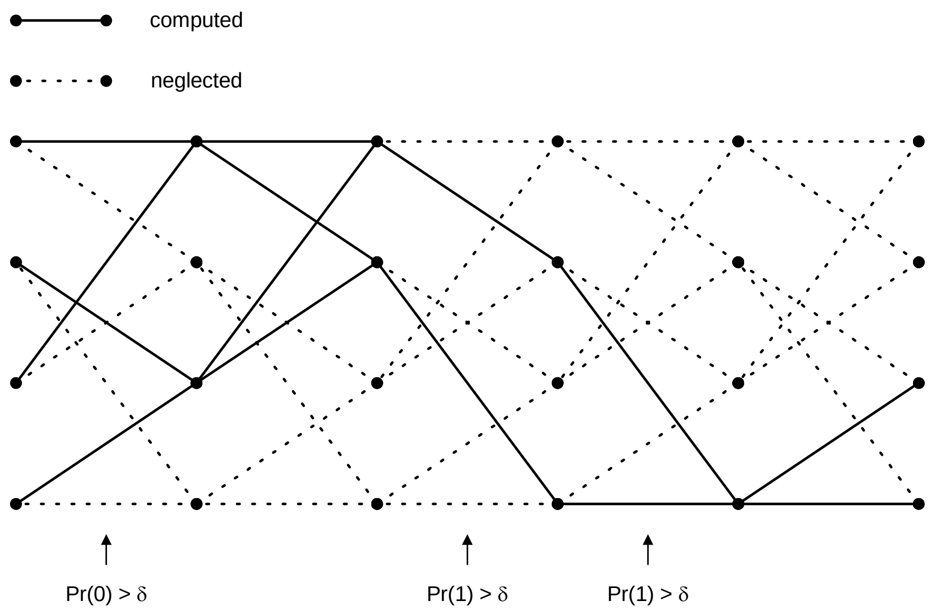

Figure 13. Example of STD trellis structure where only a part of the metrics is computed.

4.4.2 Complexity

The threshold can be set beforehand and the same value is then used in various situations to reduce the average computing complexity. Since the threshold value is not adjusted burst-by-burst, the actual peak complexity can be rather high (close to full trellis complexity) during bad channel conditions. Although a lot of metrics calculations can be avoided by the STD, it also requires some extra software and memory. The comparison (4.12) has to be done for each stage during decoding, where the threshold δ can be represented in log-domain to avoid conversions (4.10) and (4.11). Also the required soft outputs $\lambda_d^{(n)}(\mathbf{u})$ need to be calculated and stored to be used in the next iteration round.

By adjusting the threshold properly we can control the trade-off between the receiver performance and complexity reduction. The complexity reduction with a certain threshold depends on the E_b/N_0 value and channel conditions. As the input information $\lambda^a(c_{k,i})$ from the equaliser is less reliable with low E_b/N_0 or with long delay spread, the decoded bits are also less reliable and thus less complexity reduction by the STD can be achieved. Moreover, the code parameters also affect the complexity. Coding is weaker if the coding rate is higher or the constraint length is shorter. Consequently, less reliable decoding results and thereby less complexity reduction is obtained by the STD.

In publication [P6] we use as an example the threshold value 0.998, which reduces the STD complexity to 10-20 % of the full trellis decoder at the typical E_b/N_0 operation point. Moreover, we observe hardly any degradation due to the STD in the receiver performance in the simulations. The first iteration is performed with a conventional full trellis decoder, as there is no feedback information available yet, and further iterations exploit the reduced complexity decoder.

4.5 Iterative space-time receivers

The reliability of the radio transmission can be improved by spatial diversity techniques, where several transmission and/or receiver antennas are used. In this thesis we concentrate on transmit diversity in the EGPRS system with several transmit antennas and a single receive antenna. Two specific Space-Time (ST) schemes, i.e., Delay Diversity (DD) and Space-Time Trellis Codes (STTC) are presented in this section. We describe the receiver structure, which performs ST processing for the received signal utilising the ST code structure of the system. Moreover, an iterative receiver structure to improve receiver performance is introduced.

Detailed descriptions on the ST coded systems and the turbo equalisation technique are given in the book published by Hanzo, Liew and Yeap [31]. The book contains also extensive performance analysis on the considered receiver algorithms.

4.5.1 System model

Space-time coded system with two transmit antennas and a single receive antenna is represented in Figure 14. A block of data bits $\mathbf{u} \equiv \left(u_1, u_2, ..., u_{K_0}\right)^T$ is protected by a convolutional encoder Γ and punctured to achieve appropriate coding rate. The coded bits $\mathbf{c} \equiv \left(c_1, c_2, ..., c_{N_0}\right)^T$ are interleaved and provided for a space-time (ST) encoder Ω, which maps the input bits to *N* separate symbol vectors $\mathbf{d}^{(n)} \equiv \left(d_1^{(n)}, d_2^{(n)}, ..., d_K^{(n)}\right)^T$, where $n = 1, \ldots N$ is the antenna index and *N* is the number of transmit antennas. A burst builder inserts midamble symbols into the middle of each burst to assist channel estimation and an *M*-ary signal mapper associates each input symbol $d_k^{(n)}$ with a complex-valued transmission symbol $a_k^{(n)}$.

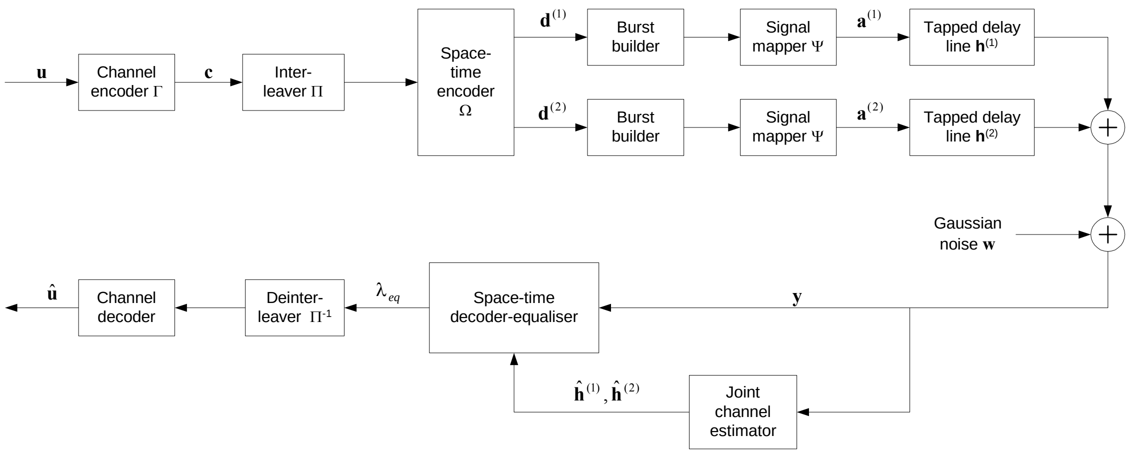

Figure 14. Transmission system model with two transmit antennas [64].

The ST coded signal is transmitted over a frequency selective fading channel using N uncorrelated antennas. The symbol-spaced received signal $\mathbf{y}$ is the superposition of N signals and noise as follows

$$y_k = \sum_{n=1}^{N}\sum_{l=0}^{L} a_{k-l}^{(n)} h_l^{(n)} + w_k \quad , \tag{4.16}$$

where $a_k^{(n)}$ denotes the ST coded symbols transmitted from antenna n and the total channel impulse response from antenna n to the receive antenna is described by the symbol-spaced tapped delay line $\mathbf{h}^{(n)} \equiv \left(h_0^{(n)}, h_1^{(n)}, \ldots, h_L^{(n)}\right)^T$. The white Gaussian noise samples are denoted by $\mathbf{w}$ with two-sided power spectral density N_0.

4.5.2 Delay Diversity (DD)

The simple DD scheme is depicted in Figure 15. It consists of ½-rate repetition code and demultiplexer, which divides the same signal into two transmission branches. Furthermore, one of the branches is delayed by one symbol period T. Hence, the transmitted symbols are

$$\begin{aligned} a_k^{(1)} &= a_k \\ a_k^{(2)} &= a_{k-1} \end{aligned} \quad . \tag{4.17}$$

The received signal can be simplified due to the same transmitted symbols as follows

$$\begin{aligned} y_k &= \sum_{n=1}^{2}\sum_{l=0}^{L} a_{k-l}^{(n)} h_l^{(n)} + w_k \\ &= \sum_{l=0}^{L} a_{k-l} h_l^{(1)} + \sum_{l=0}^{L} a_{k-l-1} h_l^{(2)} + w_k \\ &= \sum_{l=0}^{L+1} a_{k-l} h_l^{eff} + w_k \end{aligned} \tag{4.18}$$

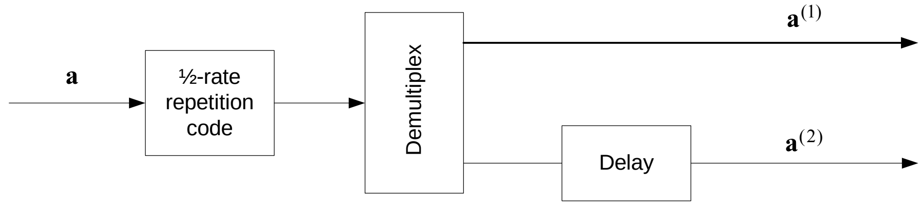

Figure 15. Transmission in delay diversity scheme [72].

where the effective l^{th} channel tap is given by

$$h_l^{eff} = h_l^{(1)} + h_{l-1}^{(2)} \quad . \tag{4.19}$$

The same midamble code is used in the both transmission antennas, which enables to use the conventional receiver, since the received signal corresponds to that of single antenna transmission. However, the delay element increases the effective length of the impulse response as shown by (4.18) and (4.19).

4.5.3 Optimised Space-Time Trellis Codes (STTC)

In the optimised ST trellis codes [64] more complicated trellis structures than the repetition code of the DD scheme are used in the ST encoder to provide coding advantage on top of the spatial diversity. Figure 16 shows an example of the trellis of 8-state ST code for the 8-PSK modulation with two transmission antennas [64] where the labels beside the trellis define the transmitted symbols corresponding to each trellis transition.

The receiver requires the knowledge of the ST trellis structure as the superposition of independent symbol streams from the antennas is considered. Therefore the channel impulse responses cannot be merged but must be estimated jointly as given in Section 3.1.5, hence a unique training sequence need to be transmitted from each antenna. The ST decoder-equaliser algorithm is described in the following section.

4.5.4 Space-time decoder-equaliser

The trellis diagram in the ST decoder-equaliser is determined by the input and transmitted symbols, since each input symbol defines a transition ξ_k from the current state S_k to the next state S_{k+1}. The transition label is defined by the ST encoded symbols $d_k^{(1)} d_k^{(2)} \cdots d_k^{(N)}$, which determine the symbols $a_k^{(1)} a_k^{(2)} \cdots a_k^{(N)}$ to be transmitted simultaneously from the N antennas.

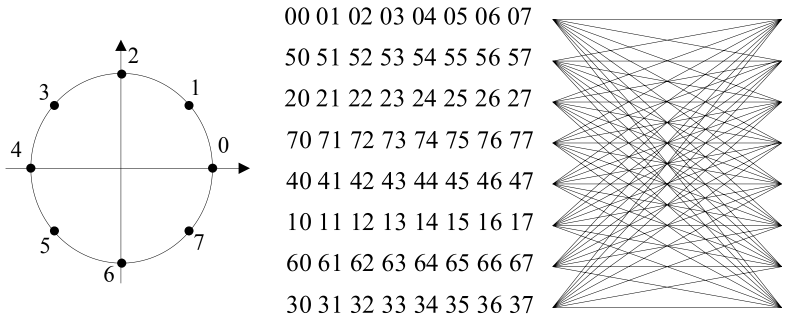

Figure 16. Symbol constellation and 8-state STTC for 8-PSK and 2 transmit antennas [64].

The additive branch metric related to the transition is given as follows

$$\gamma_k(\xi_{k-1}) = \left\| y_k - \sum_{n=1}^{N}\sum_{l=0}^{L} a_{k-l}^{(n)} h_l^{(n)} \right\|^2 \quad , \tag{4.20}$$

where $h_l^{(n)}$ denotes l^{th} tap of total impulse response from transmit antenna n. The decoding of the ST trellis code is in practice combined with the equalisation task at the receiver. The ML solution for this kind of receiver is given as

$$\hat{\mathbf{c}} = \arg\min_{c} \sum_{k} \left| y_k - \sum_{n=1}^{N}\sum_{l=0}^{L} a_{k-l}^{(n)} h_l^{(n)} \right|^2 \quad , \tag{4.21}$$

where the transmitted symbols $\mathbf{a}^{(n)}$ depend on the trellis structure $\Omega^{(n)}$, symbol mapping Ψ and interleaver Π as follows

$$\mathbf{a}^{(n)} = \Psi(\mathbf{d}^{(n)}) = \Psi\left(\Omega^{(n)}\left[\Pi(\mathbf{c})\right]\right) \quad . \tag{4.22}$$

4.5.5 Iterative ST receiver

The iterative ST receiver structure is illustrated in Figure 17 below, which corresponds to Figure 11 for the single antenna transmission. The log-likelihood (LLR) from the ST decoder-equaliser in the presence of *a priori* information is given as

$$\lambda_{eq}(b_{k,j}) = \lambda^a(b_{k,j}) + \lambda_{eq}^{ext}(b_{k,j}) \tag{4.23}$$

where *a priori* information is obtained as a feedback from the outer decoder

$$\lambda^a(b_{k,j}) \approx \lambda_d^{ext}(b_{k,j}) \quad . \tag{4.24}$$

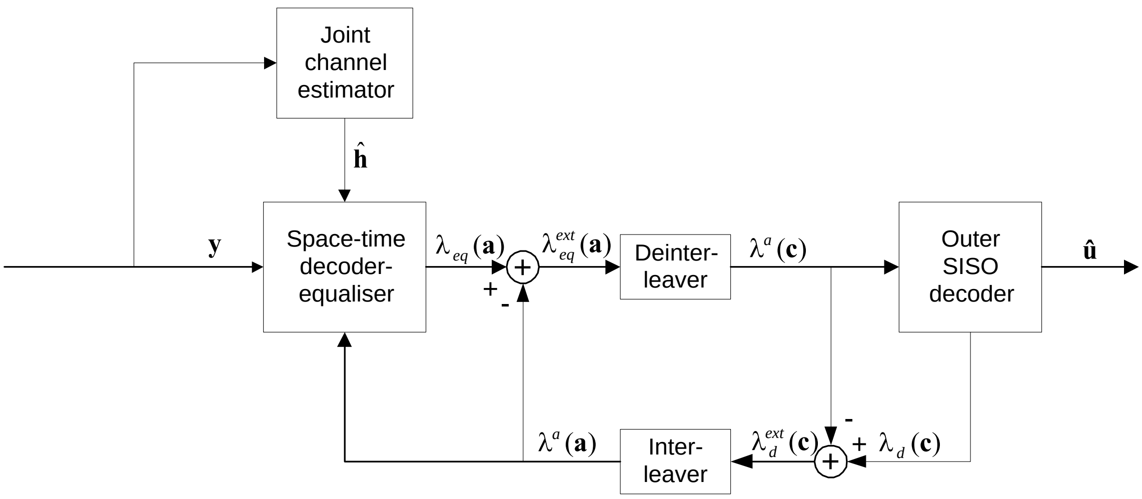

Figure 17. Structure of iterative ST receiver [15].

Furthermore, the extrinsic information is

$$\lambda_{eq}^{ext}(b_{k,j}) = \max_{\substack{(s',s)\\ b_{k,j}=1}} \left\{ \ln \alpha_{k-1}(s') + \ln \gamma_k(s',s) + \sum_{\substack{m=1\\ m\neq j}}^{\log_2 M} \ln \Pr\left(b_{k,m}\right) + \ln \beta_k(s) \right\}$$
$$- \max_{\substack{(s',s)\\ b_{k,j}=0}} \left\{ \ln \alpha_{k-1}(s') + \ln \gamma_k(s',s) + \sum_{\substack{m=1\\ m\neq j}}^{\log_2 M} \ln \Pr\left(b_{k,m}\right) + \ln \beta_k(s) \right\} \tag{4.25}$$

where the transition probability $\gamma_k(s',s)$ is calculated by

$$\gamma_k(s',s) = \left\| y_k - \sum_{n=1}^{N} \sum_{l=0}^{L} a_{k-l}^{(n)} h_l^{(n)} \right\|^2 \quad . \tag{4.26}$$

Even though there are several transmit antennas and the received signal is a superposition of several signals, they are all based on the same information source. As the feedback from the decoder is related to this unique signal, the effect of *a priori* information resembles the case of single antenna transmission. So the ST receiver merely calculates the trellis metrics (4.26) differently, but *a priori* information is used similarly, because there is no *a priori* information available on separate transmission streams.

4.6 Summary of iterative schemes

This section discusses about iterative receiver schemes for convolutionally coded systems. The principle of the iterative Turbo Equalisation (TE) is described and the exchange of the extrinsic information between the detector and channel decoder is emphasised. The detector performance is improved by utilising *a priori* information on the transmitted symbols as described in Section 3.2.8. The *a priori* information is obtained as a feedback from the decoder output. Due to soft feedback information we consider the BCJR-MAP decoder that produces soft outputs. We also propose the Soft Trellis Decoder (STD) algorithm, which reduces the decoding complexity of the TE receiver significantly with only a slight performance loss. The STD utilises the feedback information from the previous decoding step to adjust the trellis size. The STD is also a useful tool to control the trade-off between performance and complexity. The last section considers Space-Time (ST) coded systems, which are utilising both spatial and time domains in data transmission. We introduce simple Delay Diversity (DD) scheme and Space-Time Trellis Codes (STTC). The combined ST decoder-equaliser based on the ML solution is described and an extension using *a priori* knowledge is given. Finally, the iterative ST receiver is presented, where the extrinsic information is changed between the ST decoder-equaliser and outer convolutional decoder.

5 MULTICHANNEL EQUALISATION

5.1 Background

In TDMA systems the cochannel interference (CCI) originates from the neighbouring cells due to the reusage of the transmission frequencies. The CCI can be suppressed by joint detection (JD) technique, where the receiver equalises both the wanted signal and interfering signal at the same time and thus the detection reliability improves. The conventional structure of the MLSE equaliser is extended to handle two signals instead of one and therefore the complexity grows exponentially, which limits the usability of the JD technique. Furthermore, the training sequences of the two signals should overlap (timing offset is not more than a few bits) to enable accurate channel estimation and for that reason a synchronised mobile network is required. With binary modulation, like in the GSM, the JD is found beneficial and also implementable.

The complexity of the JD is substantially high, hence in practice we detect jointly only two cochannel signals (wanted signal and one interfering signal). The CCI can usually originate from several possible sources and of course it is most beneficial to detect the strongest interfering signal among them. Two factors affect the receiver performance: how strong the dominant interfering signal (DI) is compared to other interference and how often the correct DI is identified by the receiver. In this thesis we study typical interference distributions and evaluate the performance degradation due to failed DI identifications.

In order to perform the JD, we first have to perform joint channel estimation (JCE). The detector needs the channel state information for both the wanted and interfering signals. Since there are two independent channels, but only one received signal (one receive antenna), it is beneficial to estimate those channel impulse responses jointly. The conventional one-shot channel estimators like the ML or LMMSE estimators can be extended to estimate two channel impulse responses instead of one as described in Section 3.1.5. It is also possible to extend the RLS-MLSE equaliser (Section 3.2.6) for two signals and update the channel estimates separately and perform per-survivor processing simultaneously for the two signals [77].

The performance of JCE relies heavily on the training sequence properties, especially autocorrelation functions and cross-correlation between the two sequences. For instance, in the GSM the training sequences are originally designed for single signal detection and therefore the autocorrelations are optimised, but the cross-correlations are more or less random. When these sequences are applied to JCE the performance can sometimes be very poor if an unfavourable pair of sequences happens to be in use.

5.2 Frequency reuse

Since the radio frequencies are a scarce natural resource, the same transmission frequencies have to be reused in cellular systems. This frequency reuse is measured by reuse pattern, which tells the number of different frequencies used after which the same set of frequencies is used again. Figure 18 shows an example of a GSM network utilising reuse pattern three. From the mobile operator point of view the reuse pattern should be as low as possible to maximise the network capacity, i.e., more mobile users are served with a given frequency bandwidth. On the other hand, the cochannel interference becomes a severe problem, as there are users at the same carrier frequency in the nearby cells. Therefore methods to make communication less sensitive against cochannel interference by means of signal processing at the receiver are valuable.

Figure 18 depicts the CCI problem in the GSM network with ideal hexagonal cells. In downlink direction the mobile receives the desired signal but also the superposition of the CCI signals from the neighbouring cells at the same frequency. Although CCI from cells farther away also exist, their contribution is usually rather low due to the greater path loss. Hence, the number of essential CCI sources is quite limited.

The signal to noise and interference ratio (SNIR) at the receiver is defined as

$$\mathrm{SNIR} = \frac{C}{\sum_{n=1}^{N} I_n + \sigma^2} \tag{5.1}$$

where C, I_n and σ^2 denote the power of desired signal, the power of n^{th} cochannel signal and the noise variance with two-sided power spectral density N_0, respectively. It is assumed that the cochannel signals propagate through independent frequency selective channels that involve effects of fast Rayleigh fading, slow fading (shadowing), attenuation loss and power control. The shadowing follows a long-tailed lognormal distribution, which is likely to dominate the power distribution of each cochannel signal. Thereby the total interference experienced by a receiver could be approximated by a superposition of lognormally distributed signals. The authors in [7],[63] conclude that this kind of superposition of lognormal signals is itself close to lognormal. So we can assume that in frequency hopping system the interference level follows lognormal distribution at each frequency separately.

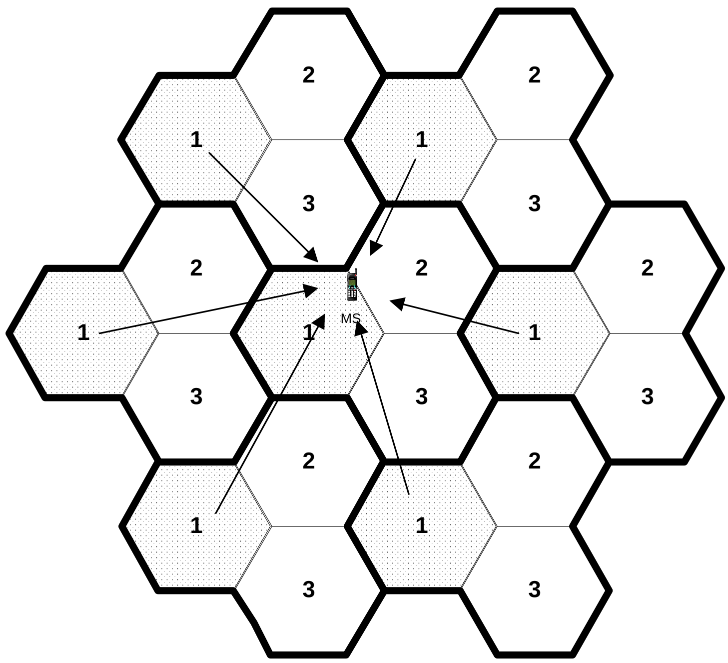

Figure 18. Cochannel interference problem in the GSM system. Downlink direction and reuse pattern three are assumed.

5.3 Equalisation algorithms for multiple signals

5.3.1 Joint detection

Joint detection (JD) algorithms are extended versions from conventional single-signal detection algorithms. The same detection principles are used, but several (usually two) signals are detected simultaneously. In the presence of CCI the joint detection of the wanted and interfering signals suppresses interference significantly, which improves receiver performance. However, the complexity of the JD is substantially high, since the number of trellis states increases exponentially with the number of detected signals. The joint MLSE and joint MAP detection algorithms are discussed in detail in Section 3.2.9.

5.3.2 Multichannel estimation

The estimation of the channel impulse response for both desired and dominant interfering signals is a crucial matter for joint detection. In the conventional GSM receiver, channel estimation is based on *a priori* known training sequence. Evidently, joint channel estimation (JCE) can also exploit the knowledge of training sequences carried by cochannel signals.

However, the cross-channel interference between cochannel signals makes the task very challenging.

The most straightforward method to solve the channel estimates fast, reliably and accurately is to use a one-shot channel estimation based on a solution of a system of linear equations [48],[60], which is described in Section 3.1.5. However, this method requires a synchronous system, i.e., all training sequences are received simultaneously from the cochannel sources. Slight asynchronism (a few bits) is allowed, but the midamble symbols should be mainly overlapping to enable reliable estimation [P8]. In the presence of greater asynchronism blind methods for channel estimation have to be applied, which degrades performance. Another drawback is that the interference may change during the burst and therefore the interference can be only partly suppressed. Also the training sequences must be unique and with low cross-correlation at least within the closest cochannel signals. In practice there are enough sequences (eight) in the current GSM system, but some of the sequences are too correlated with each other. Better sequence alternatives from the JCE point of view are presented in the next section.

5.3.3 Joint RLS-MLSE

The RLS-MLSE method introduced in Section 3.2.6 can be extended to suppress Co-Channel Interference (CCI) by jointly detecting the desired and interfering signals as presented by Yoshino *et al.* in [77]. The joint RLS-MLSE scheme is described in Figure 19.

The RLS estimator described in Section 3.1.3 can be extended for two signals $a_k^{(1)}$ and $a_k^{(2)}$ by redefining the vectors

$$\mathbf{a}_k \equiv \left(a_k^{(1)},\ldots,a_{k-L}^{(1)},a_k^{(2)},\ldots,a_{k-L}^{(2)}\right)^T \tag{5.2}$$

$$\mathbf{h}(k) \equiv \left(h_0^{(1)}(k),\ldots,h_L^{(1)}(k),h_0^{(2)}(k),\ldots,h_L^{(2)}(k)\right)^T \quad , \tag{5.3}$$

after which Eq. (3.20) updates the both channel estimates $\hat{\mathbf{h}}^{(1)}$ and $\hat{\mathbf{h}}^{(2)}$ simultaneously. In the training mode the known training symbols $m_k^{(1)}$ and $m_k^{(2)}$ of the desired and interfering signals are utilised in obtaining the initial channel estimates. In the tracking mode the channel adaptation is done for each trellis state *s* by using the candidate symbol vector

$$\mathbf{a}_k^{(s)} \equiv \left(a_k^{(s,1)},\ldots,a_{k-L}^{(s,1)},a_k^{(s,2)},\ldots,a_{k-L}^{(s,2)}\right)^T \tag{5.4}$$

to update the channel parameters $\hat{\mathbf{h}}^{(s)}(k)$ for the state *s*.

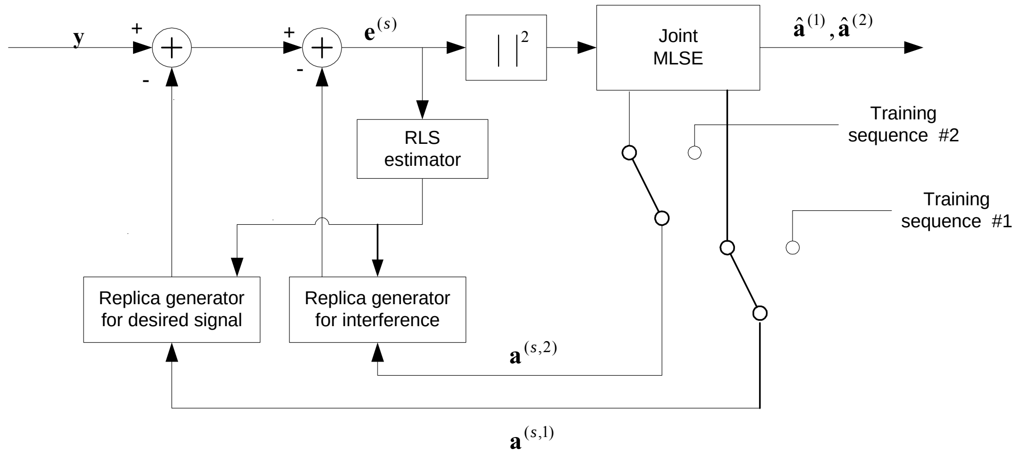

Figure 19. Joint RLS-MLSE receiver structure for two cochannel signals [77].

The transversal filters are used to replicate the desired and interfering signals based on the symbol candidates and updated channel parameters. These replicas are subtracted from the received signal giving the residual signal for trellis state *s* as follows [77]

$$e_k^{(s)} = y_k - \sum_{l=0}^{L} a_{k-l}^{(s,1)} \hat{h}_l^{(s,1)}(k) - \sum_{l=0}^{L} a_{k-l}^{(s,2)} \hat{h}_l^{(s,2)}(k) \; . \tag{5.5}$$

The trellis state *s* determines uniquely the symbol vector $\mathbf{a}_k^{(s)}$ including the desired signal and interfering signal parts and the joint detection MLSE utilises squared estimation error $\left|e_k^{(s)}\right|^2$ as the branch metric. The joint detection algorithm is described in detail in Section 3.2.9.

In order to reduce the complexity of the RLS estimation the Ensemble-averaged Inverse-matrix Least Squares (EILS) algorithm can be employed to update the both estimates simultaneously as follows [77]

$$\hat{\mathbf{h}}^{(s)}(k) = \hat{\mathbf{h}}^{(s)}(k-1) + P_0 \mathbf{a}_k^{(s)H} e_k^{(s)} \tag{5.6}$$

where P_0 is the inverse of the covariance matrix calculated by

$$P_0 = \lim_{K \to \infty} \left\{ \sum_{k=1}^{K} \lambda^{K-1} \mathbf{a}_k^{(s)} \mathbf{a}_k^{(s)H} \right\}^{-1} , \tag{5.7}$$

where λ is the forgetting factor taking values between zero and one.

For the correct candidate vector $\mathbf{a}_k^{(s)}$ the residual signal (5.5) includes only noise and channel estimation error components, thus the interference (ISI and CCI) can be suppressed. Moreover, the joint RLS-MLSE scheme can be extended to eliminate several interfering signals by designing the RLS estimator and joint MLSE detector for several cochannel signals and adding appropriate number of replica generators to modify the residual signal (5.5) respectively [77].

5.4 Training sequences

The structure of the training sequence in TDMA systems is illustrated in Figure 20. The sequence $\mathbf{m} = (m_1, m_2, ..., m_{P+L+1})^T$ is divided into P symbols of reference part $m^r = (m_1, m_2, ..., m_P)^T$ and L+1 symbols of guard part $m^g = (m_{P+1}, m_{P+2}, ..., m_{P+L+1})^T$. The guard part is designed to cover the channel memory length L. To minimise the periodic autocorrelation values the guard part should replicate the very first reference symbols, i.e.,

$$m_{P+k} = m_k \ , \ 1 \le k \le L+1 . \tag{5.8}$$

When considering a set of midamble codes for joint channel estimation, we may constitute the following design criteria.

- *Number of sequences*: unique midamble codes needed for the closest cochannel signals.
- *Sequence length*: trade-off between estimation accuracy and overhead.
- *Correlation properties*: both auto- and cross-correlation affect performance.

Using these criteria we construct two sets of sequences that are presented next. We first consider solely autocorrelation function (ACF) criterion to obtain a reasonably large set of candidate sequences. From the candidates we then pick subsets, which also have low cross-correlation within the subset.

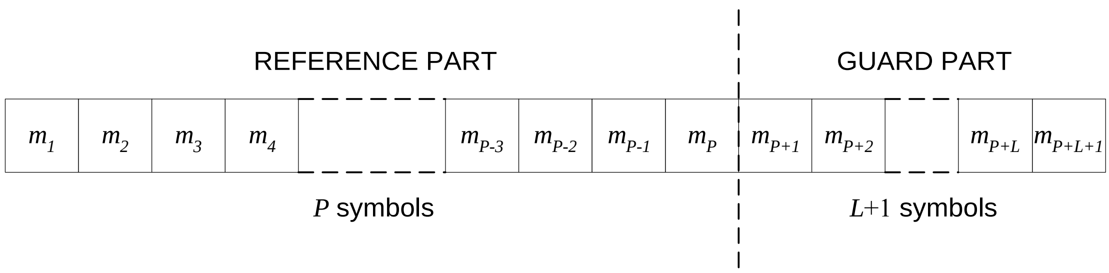

Figure 20. Structure of training sequence, which is divided into reference and guard parts.

Length 20-bit sequences

The current GSM system employs training sequences of 26 bits total length and optimised 16 bits reference length. The novel set of length 20-bit sequences are optimised by their periodic ACF using reference part of 20 bits and the same total length of 26 bits. Hence, the GSM frame structure could be used as such with the new sequences. According to [74] there are two distinguished 20-bit sequences with only one sidelobe in ACF, i.e., the main peak is followed by nine consecutive zeros before the sidelobe. To find more sequences, also ACFs with eight and seven zeros are accepted. This set consists of 22 sequences, which are listed in [P9].

Gold sequences

Maximal length sequence (*m*-sequence) is a special type of sequence with low autocorrelation [43]. It is generated by a shift register with memory N and has length of $K = 2^N - 1$. The periodic autocorrelation is given by

$$R(\tau) = -1/K \ , \ 1 \le \tau \le K-1 . \tag{5.9}$$

When two *m*-sequences of the same length are combined using shift registers, well-known family of Gold sequences is obtained [43]. These sequences have also low cross-correlation values, because of which they are suitable for signature codes in wireless multi-user applications.

The properties of Gold sequences are:

- Generated from two *m*-sequences with shift registers of memory N
- Length $K = 2^N - 1$
- Number of sequences $M = 2^N + 1$
- Set includes two generator *m*-sequences
- Auto- and cross-correlation follow the same three-valued spectrum.

The shortest Gold sequences are of length 31, which are considered hereafter. Evidently, the Gold sequences do not fit into the GSM frame structure, but due to their good correlation properties it is interesting to compare them to other sequences. The set of Gold sequences includes 33 sequences and the three possible correlation values are – 9/31, – 1/31 and 7/31 [44].

Selection of subsets

The 20-bit and Gold sequences have low ACF and subsets with good cross-correlation properties have to be chosen. In publications [P8],[P9] we consider subsets of 7, 10 and 15 sequences from the new sequence families and compare them to the subset of 7 current GSM sequences. The selection criterion is based on the covariance matrix of the estimator $\mathbf{C}_{\hat{\mathbf{h}}}$. The authors in [60] derive a relationship between SNR degradation and estimator variance, which in the presence of AWGN is approximated as follows

$$d_{ce} / dB = 10 \cdot \log_{10} \left[1 + \mathrm{tr}\left(\mathbf{C}_{\hat{\mathbf{h}}} \right) \right] \tag{5.10}$$

where $\mathbf{C}_{\hat{\mathbf{h}}} = \left(\mathbf{M}^H \mathbf{M}\right)^{-1}$. The trace of the estimator covariance matrix collects the sum of the estimated channel tap variances. Thus when the trace term is small, all the estimator variances are low and thereby the estimates are accurate. Since this measure gives the performance degradation due to channel estimation and in particularly the effect of non-ideal cross-correlation between training sequences, it is suitable to evaluate sequence pairs. The smaller the degradation, the better the pair of sequences. The subsets and their properties are discussed in detail in [P9].

5.5 Dominant interfering signal

The practical implementation of the joint detection algorithm prevents us from detecting more than two signals simultaneously, as the receiver complexity increases exponentially with the number of detected signals. Therefore it is necessary to concentrate on the desired signal and the strongest interference to achieve the best results. The channel estimator can assist in the task of finding the dominant interfering signal (DI) among the candidates.

One may consider jointly estimating all channels at the same time and then picking the strongest out of those. However, in the GSM system only a short part of the burst is allocated for training sequence (26 bits out of 148) to maximise the system payload, i.e., the transmission rate of the useful data bits. As the multipath delay requires several channel taps to be estimated, the joint channel estimation is limited to two or three cochannel signals. Therefore sophisticated algorithms are needed to identify the DI, which are introduced in Sec. 5.5.2.

5.5.1 Distribution of interference sources

It is probable that in the GSM system most of the received interference power originates from a single source. This is due to limited number of nearest cochannel transmitters, e.g., for omnidirectional antennas there are half a dozen cells in the nearest cochannel tier. This number is further reduced by techniques like sectorised cells or antenna arrays. Moreover, the signals experience independent fluctuation due to fast fading, shadowing, distance attenuation and power control operations. The transmission can be even cut off due to discontinuous transmission (DTX) or fractional load.

The performance improvement is already significant, when merely the DI is suppressed, since it carries most of the interference power. This power fraction determines the efficiency of the suppression algorithm, so we introduce dominant interference ratio (DIR) given as follows

$$DIR = \frac{I_d}{\sum_{n=1}^{N} I_n - I_d + \sigma^2} \quad , \quad I_d = \max_n \{I_n\} \tag{5.11}$$

where the power of the DI is denoted by I_d and noise variance by σ^2.

We study interference distribution in [P8] for urban cellular networks with hexagonal, omnidirectional cells and reuse pattern three. We find DIR at least 5 dB with 30 % probability over all mobiles, but with 60 % probability for mobiles of bad channel conditions (C/I below 9 dB). It is shown in [48] that interference suppression provides 3 dB gain in the presence of 5 dB DIR. Hence, 60 % of mobiles with bad channel quality can obtain at least 3 dB gain from interference cancellation. In microcell environment, mobiles at street crossings are often subject to heavy interference. In [51] it is shown that suppressing the DI gives remarkable improvement in street crossings areas. The evaluated network capacity increases with the factors from 1.5 up to 4 depending on how the quality criterion of the network is defined.

5.5.2 Identification of dominant interfering signal

The DI identification is based on the pairwise channel estimation (PCE), i.e., each candidate cochannel signal is estimated jointly with the desired signal. Let there be N synchronous cochannel signals and without losing generality we can denote the desired signal with index 1 and interfering signals with indices 2,…,N. We consider the contribution of two cochannel signals at a time, whose midamble codes are denoted by $\mathbf{M}' = \left[\mathbf{M}^{(1)} \quad \mathbf{M}^{(2)}\right]$. The contribution of other midambles $\mathbf{M}'' = \left[\mathbf{M}^{(3)} \quad \mathbf{M}^{(4)} \quad \cdots \quad \mathbf{M}^{(N)}\right]$ is regarded as Gaussian noise. The corresponding radio channels are

$$\mathbf{h}' = \begin{bmatrix} \mathbf{h}^{(1)} \\ \mathbf{h}^{(2)} \end{bmatrix} \quad \text{and} \quad \mathbf{h}'' = \begin{bmatrix} \mathbf{h}^{(3)} \\ \mathbf{h}^{(4)} \\ \vdots \\ \mathbf{h}^{(N)} \end{bmatrix} \quad . \tag{5.12}$$

Using these notations the received signal **y** in the presence of AWGN can be presented as the following superposition

$$\mathbf{y} = \mathbf{M}'\mathbf{h}' + \mathbf{M}''\mathbf{h}'' + \mathbf{w} \ , \tag{5.13}$$

where white Gaussian noise samples are denoted by **w**, whose variance is σ^2. By merging the latter terms in Eq. (5.13) we obtain

$$\mathbf{y} = \mathbf{M}'\mathbf{h}' + \widetilde{\mathbf{w}} \ , \tag{5.14}$$

where $\tilde{\mathbf{w}} = \mathbf{M}''\mathbf{h}'' + \mathbf{w}$. To simplify the estimation problem, Gaussian distribution is assumed for $\tilde{\mathbf{w}}$. The ML solution is then given as

$$\hat{\mathbf{h}}' = \left(\mathbf{M}'^H \mathbf{C}_w^{-1} \mathbf{M}'\right)^{-1} \mathbf{M}'^H \mathbf{C}_w^{-1} \mathbf{y}. \tag{5.15}$$

Further simplification is achieved by assuming white Gaussian distribution for $\tilde{\mathbf{w}}$, thus $\mathbf{C}_w = \sigma^2 \mathbf{I}$, which gives

$$\hat{\mathbf{h}}' = \left(\mathbf{M}'^H \mathbf{M}'\right)^{-1} \mathbf{M}'^H \mathbf{y}. \tag{5.16}$$

Eq. (5.16) gives channel estimates for the desired signal (index 1) and interference with index 2. Furthermore, channel estimate $\hat{\mathbf{h}}^{(k)}$ for the interfering signal k is obtained by interchanging $\mathbf{M}^{(2)}$ and $\mathbf{M}^{(k)}$ in the matrices $\mathbf{M}'$ and $\mathbf{M}''$ and re-calculating (5.16). The midamble of the desired signal $\mathbf{M}^{(1)}$ is always present in $\mathbf{M}'$. After all the channel estimates are available, the DI among the candidates can be identified. We consider two identification approaches: power estimation (PWR) and maximum-likelihood (ML) solution. The PWR simply calculates the power estimates

$$\hat{I}^{(n)} = \sum_{l=0}^{L} \left|\hat{h}_l^{(n)}\right|^2 \quad , \; n = 2,3,...,N. \tag{5.17}$$

and select the largest as dominant

$$I_{dom} = \max_n \left\{\hat{I}^{(n)}\right\}. \tag{5.18}$$

A more sophisticated solution is based on the ML approach, which uses the conditional probability

$$p(\mathbf{y} \mid \mathbf{M}', \mathbf{h}') = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[\frac{-\left\|\mathbf{y} - \mathbf{M}'\mathbf{h}'\right\|^2}{2\sigma^2}\right] \quad . \tag{5.19}$$

The DI identification problem is solved by finding the training sequence matrix $\mathbf{M}' = \left[\mathbf{M}^{(1)} \; \mathbf{M}^{(n)}\right]$ that maximises the probability (5.19). Equivalently one can minimise the Euclidean distance as follows

$$\hat{\mathbf{M}}'_{ML} = \arg\max_{\mathbf{M}^{(n)}} p(\mathbf{y} \mid \mathbf{M}', \mathbf{h}') = \arg\min_{\mathbf{M}^{(n)}} \left\|\mathbf{y} - \mathbf{M}'\mathbf{h}'\right\|^2 \quad . \tag{5.20}$$

Cross-correlation between training sequences causes bias in the PCE and therefore it makes DI identification more unreliable. Especially PWR estimates (5.17) can lead to wrong conclusions, because a weak signal may be highly emphasised in PCE due to heavily correlated sequences. The ML criterion (5.19) takes cross-correlations better into account, which is also shown by simulations in [P9]. Furthermore, publication [P8] shows that the average receiver performance is hardly any degraded due to DI identification by the ML approach. We can conclude that whenever the DI identification by the PCE-ML fails, it is likely due to two (or more) equally strong interfering signals and it is insignificant which one is actually suppressed.

5.6 Summary of the multichannel equalisation

We consider Cochannel Interference (CCI) suppression by the Joint Detection (JD) technique in this chapter. The single signal channel estimator and detector are extended for multiple signals in Sections 3.1.5 and 3.2.9, respectively. The extension of the RLS-MLSE equaliser for two signals is presented in this section. We also discuss about the CCI problem in the GSM system and describe interference distribution at a mobile terminal. Since the GSM training sequences are not designed for multichannel equalisation, we propose improved training sequences for the JD system. Especially the good cross-correlation properties of the new sequences are emphasised. Due to complexity restrictions the JD is best suited for detecting two signals, i.e., the desired signal and one interfering signal. Therefore we suggest a reliable method to identify the strongest interfering signal from the possible candidates. The identification is based on pairwise channel estimation with the ML based selection criterion.

6 SUMMARY OF PUBLICATIONS

The publications are divided into two main groups so that the publications [P1-P7] consider iterative equalisation techniques from various aspects and the publications [P8-P9] concentrate on multichannel equalisation methods. Furthermore, the former category mainly considers the EDGE/EGPRS platform, whereas the latter uses the GSM system.

Detailed overview on the Turbo Equalisation (TE) and Iterative Channel Estimation (ICE) techniques is given in Publication [P1] with algorithm descriptions and simulation results. More compressed presentation with main results is given in [P2]. In [P3] the main emphasis is in the performance analysis of the TE scheme. Also the iterative structure and associated equaliser and decoder algorithms are presented. The ICE technique is considered in [P4] and [P5] with performance analysis for packet data systems. [P4] gives theoretical bounds for the estimator variance, whereas [P5] provides an iterative receiver algorithm combining both the TE and ICE methods. A low complexity algorithm for the decoding in the TE receiver is proposed in [P6] giving the performance and complexity analysis based on simulations. Iterative data processing for Space-Time Trellis Coded (STTC) transmission is studied in [P7]. TE gains in the STTC system and simple diversity transmission are compared in this publication.

Multichannel equalisation by the JD technique is thoroughly discussed in [P8]. Receiver algorithms are presented, performance analysis from both link and network point of view are given and system requirements are discussed. Joint channel estimation issues are considered in [P9] including requirements for training sequences and identifying the DI. Innovations in the CCI suppression research have lead to *a granted patent* [47], which describes a novel method to determine the signal strengths of the co-channel signals at the receiver. This method is essential in the JD technique, as the receiver does not know beforehand which interfering signal is the most powerful.

6.1 Iterative equaliser structures

6.1.1 Publications [P1]-[P5]

The publications [P1]-[P5] consider basic iterative detection and channel estimation techniques for the GPRS and EGPRS packet data systems. The principle of the TE technique is presented in [P1] and [P3] and the both publications give detailed description of the DFSE equaliser using *a priori* information. We show in [P1] that low complexity DFSE equaliser with only forward recursion provides reasonable performance in minimum phase channel.

Furthermore, the Max-Log-MAP decoder providing soft outputs is presented in [P3]. The performance evaluation consists of both GMSK and 8-PSK modulated systems utilising the strongest coding schemes MCS-1 and MCS-5 that are specified for the actual EGPRS. The TE provides only 0.2 dB improvement for the GMSK, whereas 2 dB gain is obtained for the 8-PSK. In [P1] and [P2] further simulation results for the weaker codings (MCS-5 to MCS-9) are reported. The achieved TE gain decreases with weaker coding and only 0.7 dB gain is achieved in MCS-9. We conclude that *the 8-PSK modulated EGPRS schemes (MCS-5 to MCS-9) benefit more from the TE receiver than the binary modulated (MCS-1 to MCS-4)*, as the extrinsic feedback information is richer due to multilevel symbols. As every 8-PSK symbol consist of three bits, there are three times more bits involved in the iteration process than in the GMSK systems. Furthermore, *strong coding is preferable for the TE*, as more redundant information is generated and thereby the extrinsic information is spread over a larger set of bits.

Most of the performance gain is obtained after a few iterations and extra iterations provide less and less improvement as shown in [P1]-[P3],[P5]. For example, the gain in MCS-5 after the second iteration is 2.0 dB and after the fourth iteration 3.0 dB [P3]. The most probable improvement is achieved in the second iteration, since the extrinsic information is zero at the first iteration, but provides some valuable *a priori* knowledge at the second iteration. During further iterations smaller-scale shifts are expected, since the extrinsic information is not likely to change so radically anymore. As the complexity increases linearly with the number of iterations, we suggest that *the TE with two iterations is a practical compromise giving good performance with a reasonable complexity.*

The ICE concept is presented in detail in [P4] giving the LMS adaptation rule and derivation of the Cramer-Rao Lower Bound (CRLB) for the improved channel estimator. Using the CRLB we show in [P4] that *the iterative estimation improves estimator variance even with unreliable data decoding*. Moreover, the iterative improvement is mainly visible at the low SNR region and the training sequence based conventional channel estimation reaches asymptotically the same estimator variance with the high SNR. The estimator variance is inverse proportional to the training sequence length. As the extended sequence in the ICE consists of the whole burst, the estimator accuracy is better than the conventional even in the presence of some incorrect data bits. However, with high SNR the conventional estimator is already very accurate due to small noise variance and no significant improvement can be obtained. Simulation results in [P1],[P4],[P5] show around 1 dB gain for the ICE technique in the both binary and 8-PSK modulated systems.

The combined TE and ICE receiver algorithm is described fully in [P1] and [P5]. That can utilise both elements during the same iteration round or select just one of them. However, the simulations in [P1],[P5] show that the combined receiver structure can provide only a slight extra improvement. Since the TE partly recovers the loss due to the inaccurate channel estimation, the TE benefits less if channel estimation is simultaneously improved by the ICE. Based on the results in publications [P1]-[P5] we conclude that *the best trade-off between complexity and performance is the TE technique for the 8-PSK modulation and the ICE method for GMSK modulated systems.*

6.1.2 Publication [P6]

The TE provides a significant performance gain, but at the cost of more complex receiver. In [P6] we propose a novel soft trellis decoding (STD) method to decrease the decoder complexity of the TE receiver. Basically, the STD utilises reliability information from the previous TE iteration to neglect some of the most unlikely trellis transitions and thereby reduces the number of the metrics calculations needed. The soft output at the previous decoding step tells the reliability of an individual symbol and if it exceeds a predetermined threshold, the symbol is not re-calculated but just regarded as certain. Hence, the decoding concentrates to the symbols and frames with greater uncertainty after initial decoding round. By adjusting the decision threshold of the STD we can easily control the trade-off between decoding performance and complexity.

Simulation results for the EGPRS system exploiting the TE receiver with the STD decoder are shown. The first iteration requires a full trellis decoder, but *during the further iterations the complexity can be reduced by 80-90 % with a negligible performance loss in typical operation point*. The reduction is larger when the signal quality improves or a further iteration round is considered.

6.1.3 Publication [P7]

Publication [P7] considers transmission diversity techniques and iterative receiver structures for them. STTC techniques transmit the signal through several antennas and the transmission streams are also encoded. Thus both spatial diversity and coding gain are utilised. For a comparison, the simple delay diversity (DD) scheme based on a repetition code is also considered. Moreover, the TE receiver for the EGPRS system with the STTC transmission and outer convolutional coding is presented.

The simulations show an equivalent or lower iterative gain for the STTC than for the DD. Hence, the TE is not able to utilise the more complex structure of the STTC, but merely improves the channel equalisation reliability for both systems. *The difference between the DD and the STTC is already present in the first (conventional) iteration and it remains the same during further iterations*.

6.2 Equalisers for multiple channels

6.2.1 Publication [P8]

An overview of CCI suppression in GSM systems by JD is given in [P8]. The CCI problem in cellular systems and distribution of interference sources is discussed in the paper. It is shown in [P8] that especially if the mobile suffers from bad quality (C/I < 9 dB), there is probably a single dominant interfering signal (DI), since with 60 % probability the DI is at least 5 dB higher than the rest of interference. This item is discussed in more detail in Section 5.5.1. JD algorithms using the MLSE and MAP approaches are presented in [P8] and suitable training sequences for JCE with low cross-correlation are discussed. Furthermore, DI identification algorithms are proposed. Since the complexity of the JD receiver grows exponentially, the implementation of only two binary modulated signals can be considered.

The receiver performance is evaluated by a link simulator, which includes a large number of independently fading interfering signals. In the performance analysis of [P8] hexagonal and omnicell layout is assumed and 18 interfering signals are used in the simulator. *Training sequence evaluation shows that the current GSM sequences perform reasonably well in the JD receiver, but 1.3 dB average gain can be achieved by new optimised sets of training sequences*. Moreover, sequence set of seven is enough in omnicell case and possibly even less for sectorised cells. *The DI identification algorithm performs very reliably* as no degradation is observed compared to the ideal DI identification. Base station activity factor affects the JD receiver significantly, since the interference distribution depends on that. The achieved gains vary from 4 dB to 9 dB depending on the activity. When two interfering signals are suppressed, the performance improves only by 1 dB, but the complexity increases significantly. Finally, the JD gain decreases clearly as asynchronism grows, thus a synchronous network is required in practice.

Finally, some requirements that JD receiver poses to the system are summarised. *The base stations should be synchronised, cell sizes and reuse factors should be low to enable accurate JCE*. Furthermore, unique training sequences should be allocated for the nearest CCI sources either manually or automatically.

6.2.2 Publication [P9]

Publication [P9] concentrates on JCE issues to enable successful CCI cancellation by JD. New training sequences are also proposed in the paper to meet JCE requirements for the sequence length and correlation properties. Furthermore, a selection algorithm for finding good subsets from the initial sequence sets is proposed. Novel set of 20-bit sequences and Gold sequences of length 31 bits are considered and optimised subsets of 7, 10 and 15 sequences are suggested. JD receiver performance with different training sequences is analysed by link level simulations using the best and worst sequence pair from each basic set. The best pairs perform rather similarly, but *the worst GSM pair is even 6 dB worse than the worst 20-bit sequence pair*.

The issue of DI identification is also discussed in [P9]. Identification algorithms utilise pairwise channel estimation (PCE) and two approaches, the strongest DI power (PWR) and the ML criterion, are introduced. The reliability of the DI identification is evaluated by simulations, which show that the training sequences and the relative power of the DI affect the reliability very much. However, the PCE-ML algorithm overperforms all the other identification methods.

6.3 Author’s contribution to publications

Author's contribution to all publications [P1-P9] is essential. In publications [P1],[P2] the author contributes in analytical development of the equaliser algorithms (especially iterative DFSE), constructing the iterative structures and providing most of the performance results. Publication [P3] is written by the author and the system performance is evaluated by him. In publications [P4],[P5] the author describes the turbo equalisation scheme and contributes in developing the channel estimation algorithms. Performance analysis is mostly done by the author. In publications [P6],[P7] the author is the responsible writer and he provides the simulation results for the both papers. In [P8] the author is responsible of the joint channel estimation and training sequence parts and the analysis related to them. He also contributes to link performance analysis. In publication [P9] the author contributes in the sections of new training sequences and dominant interfering signal and participates in performance evaluation.

Publications [P8] and [P9] are used by the co-author Dr. Ranta in his Ph.D. thesis [52]. Additionally, the author’s publications [P1-P9] are not included in any other Ph.D. thesis.

7 CONCLUSIONS

This thesis considers methods to enhance the spectrum efficiency of TDMA based cellular systems by means of digital signal processing at the receiver. The optimum receivers theoretically minimise the error probability in the presence of ISI and AWGN. Nevertheless, it is possible to *improve the receiver performance by utilising available side information*: in coded systems by iterative equalisation and in the presence of CCI by multichannel equalisation. These two extensions to the conventional equaliser structure are considered in this thesis. The discovered properties of the three different receiver techniques, i.e., Turbo Equalisation (TE), Iterative Channel Estimation (ICE) and Joint Detection (JD), are summarised in Table 1 from receiver performance, complexity and practical implementation point of views.

Table 1. Summary of receiver methods utilising side information.

Receiver method	Performance improvement			Receiver Complexity	Imple-mentation
	GSM Speech (GMSK)	GPRS (GMSK)	EGPRS (8-PSK)		
TE	Small	Small	Large	Medium	Possible
ICE	Medium	Medium	Medium	Low	Possible
JD	Large	Large	--	High	Difficult
TE = Turbo Equalisation ICE = Iterative Channel Estimation JD = Joint Detection					

Error protection coding adds redundancy in the transmitted signal to improve the transmission reliability in mobile radio applications. The optimal receiver optimises the joint probability of information bits, encoded bits and channel impulse response, which is impractical due to the complexity restrictions. Therefore the suboptimal receiver solution with separated equalisation and channel decoding is applied. The accuracy of the Bayesian estimators is generally improved in the presence of *a priori* knowledge on the estimated parameters. The Maximum A Posteriori (MAP) detector is based on the Bayesian approach and therefore *a priori* information on the transmitted bits improves detection reliability. The *a priori* knowledge can be extracted from the channel decoder output, as redundant coded bits contain extrinsic information, which is independent of the received signal at the detector input. The TE scheme introduces the *a priori* feedback signal from the decoder to the detector and thus improves the receiver performance by iterative data processing.

The single most important result of the thesis is that in GSM based systems the TE is especially beneficial for the systems with higher order 8-PSK modulation, but less attractive for binary modulated systems. The first reason is that in the higher order systems there are more bits involved in the exchange of the extrinsic information and more bits contribute to the *a priori* probability of a certain transmitted symbol. Secondly, the potential performance improvement is much lower for the binary transmission, as conventional nonlinear MLSE or MAP based equalisers perform already very reliably without any *a priori* knowledge. Thirdly, the GSM bit interleaver is not random, but deterministic. As the 8-PSK data block contains three times more bits than the binary block, the 8-PSK block is also better randomised by the interleaver. Consequently, the detector and decoder inputs are more independent and better extrinsic information is obtained.

Link simulations show that most of the performance improvement is obtained already after the second iteration and further iterations are less important. For instance, for the strongest EGPRS coding (MCS-5) 2.0 dB gain is achieved by two iterations and 1.0 dB extra gain after fourth iteration. The expected improvement is highest in the second iteration, since there is no *a priori* information available in the first iteration, but the decoder provides useful feedback information for the second iteration. Since the receiver complexity grows rapidly with the number of iterations due to repeated equalisation and decoding, *we propose the TE receiver with two iterations as a feasible trade-off between complexity and performance*. Since the EGPRS system has rectangular interleaving over four bursts, the implementation of the TE is rather attractive. Each data block can be processed separately in the receiver and iterations can be implemented with consecutive modules.

We also consider the adaptive ICE technique, which utilises the decoded symbols in channel re-estimation. We show that the expected estimation error decreases even in the presence of bad signal quality and around 1 dB gain on the average is obtained for both GMSK and 8-PSK modulated systems. It is shown that the improvements of the TE and ICE techniques are partly overlapping, since the TE also recovers the loss due to the non-ideal channel state information. Because the TE benefits less if channel estimation is simultaneously improved by the ICE, the combination of the two techniques is less attractive. Therefore, *we suggest to use the TE with the 8-PSK modulation and the ICE with the GMSK modulation*, respectively. Furthermore, the complexity of the TE scheme can be significantly reduced by the novel STD decoding method. It eliminates the most unlikely trellis paths in the decoder and only very small performance loss is observed.

The CCI problem arises in cellular networks due to the frequency reuse in the transmission links. The JD technique actively suppresses CCI in the receiver by equalising multiple channels simultaneously, one of which is the desired channel and others are related to interference. From the complexity point of view the JD is best suited for two binary modulated cochannel signals. For higher order modulations reduced-state equalisers are required. *In the interference-limited system and in the presence of one dominant interfering (DI) signal the JD improves the receiver performance significantly* by suppressing the DI effect. To find the DI among the interfering candidates we propose an algorithm, which is based on the pairwise channel estimation with the maximum-likelihood criterion (PCE-ML). This method is shown to perform very reliably and it has a negligible effect on the overall receiver performance. Joint channel estimation sets a few requirements for the system such as unique training sequences for the cochannel signals, low cross-correlation between the sequences and synchronised network, which complicate the implementation of the JD.

The methods studied in this thesis are considered in the *standardisation* process of the GSM evolution and the forthcoming packet data systems. The TE technique itself does not require any modifications in the standard, but it may lower the receiver sensitivity requirements if applied in the receiver. Special attention is drawn to power consumption in the receiver, thus low-complexity solutions like the STD are preferred and the number of iterations should be limited. This kind of device is quite possible to implement in the near-future systems. The GSM standard supports quite well also the implementation of the JD technique. Our results show that the current GSM training sequences offer reasonable performance, but further improvement can be achieved by new optimised sequences. However, synchronised base stations, careful training sequence allocation and small cell sizes are required to benefit from the JD as emphasised by Ranta in his Ph.D. thesis [52].

Recently a lot of research interest is drawn to combining iterative methods with multiantenna systems. The data processing in both the time and spatial domains is found useful for combating against ISI or CCI. Higher data rates are also achieved by multiantenna transmission. Also iterative receivers in broadband transmission systems are worth studying, since the very long channels can be equalised by iterative linear filter structures. These aspects are not considered in this thesis.

Further work to suppress unsynchronised CCI is needed. The JD cannot be used, but some semi-blind or blind methods are possible solutions for this problem.

APPENDIX I. SUB-OPTIMAL MAP ALGORITHMS.

A. Log-MAP

The optimum BCJR algorithm selects the symbol a_k at time instant k, which maximises the following APP

$$\hat{a}_k = \arg\max_{a_k} \left[\Pr\left(a_k \middle| \mathbf{y}\right)\right] \quad . \tag{A.1}$$

As stated in Section 3.2.7 the optimum BCJR provides the following APP values [3]

$$\Pr\left(a_k = i \mid \mathbf{y}_1^K\right) = \sum_{\substack{(s',s)\\ a_k=i}} \alpha_{k-1}(s')\gamma_k(s',s)\beta_k(s) \ , \ i \in \{0,1\} \quad , \tag{A.2}$$

where

$$\alpha_k(s) = \sum_{s'} \alpha_{k-1}(s')\gamma_k(s',s) \tag{A.3}$$

$$\beta_k(s) = \sum_{s'} \beta_{k+1}(s')\gamma_{k+1}(s,s') \quad . \tag{A.4}$$

$$\gamma_k\left(s',s\right) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}\left\| y_k - \sum_{l=0}^{L} a_{k-l} h_l \right\|^2\right\} \quad . \tag{A.5}$$

To avoid complicated multiplicative calculations the probabilities can be computed in logarithmic domain without any performance loss [55]. By taking the logarithm of (A.5) and omitting some constant terms (they cancel out later) we get

$$\ln \gamma_k\left(s',s\right) \cong -\frac{1}{2\sigma^2}\left\| y_k - \sum_{l=0}^{L} a_{k-l} h_l \right\|^2 \quad . \tag{A.6}$$

Recursive formulas (A.3) and (A.4) are rewritten as

$$\ln \alpha_k(s) = \ln \sum_{s'} e^{\ln \alpha_{k-1}(s') + \ln \gamma_k(s',s)} \tag{A.7}$$

$$\ln \beta_k(s) = \ln \sum_{s'} e^{\ln \beta_{k+1}(s') + \ln \gamma_{k+1}(s,s')} \tag{A.8}$$

The output of this Log-MAP detector is given as the logarithmic ratio of the APP values of symbols a_k as follows

$$\lambda(a_k) = \ln \sum_{\substack{(s',s)\\ a_k=+1}} e^{\ln \alpha_{k-1}(s') + \ln \gamma_k(s',s) + \ln \beta_k(s)} - \ln \sum_{\substack{(s',s)\\ a_k=-1}} e^{\ln \alpha_{k-1}(s') + \ln \gamma_k(s',s) + \ln \beta_k(s)} \quad . \tag{A.9}$$

Finally, the problem of computing the nonlinear functions $\ln\left(\exp\delta_1 + \ldots + \exp\delta_n\right)$ is recursively solved by using [55]

$$\ln\left(e^{\delta_1} + e^{\delta_2}\right) = \max\left(\delta_1, \delta_2\right) + \ln\left(1 + e^{-|\delta_2 - \delta_1|}\right) \quad , \tag{A.10}$$

where the latter term can be tabulated beforehand.

B. Max-Log-MAP

To further simplify the calculation of Eq. (A.9) the following approximation is used [55]

$$\ln\left(e^{\delta_1} + \cdots + e^{\delta_n}\right) \approx \max_i \ \delta_i \quad , \tag{A.11}$$

which leads to the Max-Log-MAP algorithm. The output is then conveniently expressed as

$$\begin{aligned}\lambda(a_k) = & \max_{\substack{(s',s) \\ a_k=+1}} \left\{\ln \alpha_{k-1}(s') + \ln \gamma_k(s',s) + \ln \beta_k(s)\right\} \\ & - \max_{\substack{(s',s) \\ a_k=-1}} \left\{\ln \alpha_{k-1}(s') + \ln \gamma_k(s',s) + \ln \beta_k(s)\right\}\end{aligned} \tag{A.12}$$

with

$$\ln \alpha_k(s) = \max_{s'} \left\{\ln \gamma_k(s',s) + \ln \alpha_{k-1}(s')\right\} \tag{A.13}$$

$$\ln \beta_k(s) = \max_{s'} \left\{\ln \beta_{k+1}(s') + \ln \gamma_{k+1}(s,s')\right\} \ . \tag{A.14}$$

The approximation (A.11) is rather accurate if one of the arguments δ_i dominates, i.e., if there is a clearly best path metric among the candidates. This happens with a high probability in the presence of high SNR. On the other hand at low SNR there are often several competing paths close to each other and therefore the Max-Log-MAP suffers from larger degradation.

REFERENCES

[1] T. Abe and T. Matsumoto, "Space-time turbo equalization and symbol detection in frequency selective MIMO channels," in *Proc. IEEE 54th Vehicular Technology Conf.*, Atlantic City, Oct. 2001, pp. 1230-1234.

[2] T. Abe, S. Tomisato, and T. Matsumoto, "Performance evaluation of a space-time turbo equalizer in frequency selective MIMO channels using field measurement data," in *Proc. IEE Workshop MIMO Commun. Systems*, London, UK, Dec. 2001, pp. 21/1-21/5.

[3] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 284-287, Mar. 1974.

[4] G. Bauch, H. Khorram, and J. Hagenauer, "Iterative equalization and decoding in mobile communications systems," in *Proc. 2nd European Personal Mobile Commun. Conf.*, Sept. 1997, pp. 307-312.

[5] G. Bauch and V. Franz, "A comparison of soft-in/soft-out algorithms for turbo-detection," in *Proc. Int. Conf. Telecommunications*, Porto Carras, Greece, June 1998, pp. 259-263.

[6] G. Bauch and V. Franz, "Iterative equalization and decoding for the GSM-system," in *Proc. IEEE Vehicular Technology Conf.*, Ottawa, Canada, May 1998, pp. 2262-2266.

[7] N. C. Beaulieu, A. A. Abu-Dayya, and P. J. McLane, "Estimating the distribution of a sum of independent lognormal random variables," *IEEE Trans. Commun.*, vol. 43, no. 12, pp. 2869-2873, Dec. 1995.

[8] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: Performance analysis, design, and iterative decoding," Jet Propulsion Lab., Pasadena, CA, TDA Progress Rep. 42-126, pp. 1-26, Aug. 1996.

[9] A. O. Berthet, B. Ünal, and R. Visoz, "Iterative decoding of convolutionally encoded signals over multipath Rayleigh fading channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1729-1743, Sept. 2001.

[10] A. O. Berthet, R. Visoz, and P. Tortelier, "Sub-optimal turbo-detection for coded 8-PSK signals over ISI channels with application to EDGE advanced mobile system," in *Proc. 11th IEEE Int. Symp. Personal, Indoor and Mobile Radio Commun.*, London, UK, Sept. 2000, pp. 151-157.

[11] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proc. IEEE Int. Conf. Communications*, Geneva, Switzerland, May 1993, pp. 1064-1070.

[12] R. W. Chang and J. C. Hancock, "On receiver structures for channels having memory," *IEEE Trans. Inf. Theory*, vol. 12, no. 3, pp. 463-468, Oct. 1966.

[13] K.-H. Chang and C. N. Georghiades, "Iterative joint sequence and channel estimation for fast time-varying intersymbol interference channels," in *Proc. IEEE Int. Conf. Communications*, Seattle, WA, 1995, pp. 357-361.

[14] J. M. Cioffi, G. P. Dudevoir, M. V. Eyeboglu, and G. D. Forney, "MMSE decision-feedback equalizers and coding – Part I: Equalization results," *IEEE Trans. Commun.*, vol. 43, no. 10, pp. 2582-2594, Oct. 1995.

[15] C. Douillard, M. Jézéquel, C. Berrou, A. Picart, P. Didier, and A. Glavieux, "Iterative correction of intersymbol interference: Turbo-equalization," *European Trans. Telecommunications*, vol. 6, no. 5, pp. 507-511, Sept.-Nov. 1995.

[16] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Trans. Commun.*, vol. 37, no. 5, pp. 428-436, May 1989.

[17] V. Eskelinen, "Reduced-complexity soft output equalization for GSM400 mobile communication systems," M.Sc. thesis, Dept. Electr. Eng., Helsinki Univ. of Technology, Helsinki, Finland, 2001.

[18] *Digital Cellular Telecommunications System (Phase 2+)*, ETSI Standard GSM 05 series, 1999.

[19] W. van Etten, "Maximum-likelihood receiver for multiple channel transmission systems," *IEEE Trans. Commun.*, vol. COM-24, pp. 276-283, Feb. 1976.

[20] M. V. Eyuboglu and S. U. H. Qureshi, "Reduced-state sequence estimation with set partitioning and decision feedback," *IEEE Trans. Commun.*, vol. 36, no. 1, pp.13-20, Jan. 1988.

[21] G. D. Forney Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Trans. Inform. Theory*, vol. IT-18, no. 3, pp. 363-378, May 1972.

[22] G. D. Forney Jr., "The Viterbi algorithm," *Proc. IEEE*, vol. 61, no. 3, pp. 268-278, Mar. 1973.

[23] V. Franz and G. Bauch, "Turbo-detection for an 8-PSK modulation scheme in a mobile TDMA communication system," in *Proc. IEEE Vehicular Technology Conf.*, Amsterdam, Netherlands, Sept. 1999, vol. 4, pp. 2954-2958.

[24] K. Fukawa and H. Suzuki, "Adaptive equalization with RLS-MLSE for frequency-selective fast fading mobile radio channels," in *Proc. IEEE Global Telecommun. Conf.*, Phoenix, AZ, Dec. 1991, vol. 1, pp. 548-552.

[25] K. Fukawa and H. Suzuki, "Blind interference cancelling equalizer for mobile radio communications," *IEICE Trans. Commun.*, vol. E77-B, no. 5, pp. 580-588, May 1994.

[26] W. H. Gerstacker and J. B. Huber, "Maximum SNR decision-feedback equalization with FIR filters: Filter optimization and a signal processing application," in *Proc. IEEE Int. Conf. Communications*, June 1996, pp. 1188-1192.

[27] W. H. Gerstacker and R. Schober, "Equalization concepts for EDGE", *IEEE Trans. Commun.*, vol. 1, no. 1, pp. 190-199, Jan. 2002.

[28] K. Giridhar, S. Chari, J. J. Shynk, R. P. Gooch, and D. Artman, "Joint estimation algorithms for cochannel signal demodulation," in *Proc. IEEE Int. Conf. Communications*, Geneva, Switzerland, 1993, pp. 1497-1501.

[29] K. Giridhar, J. J. Shynk, A. Mathur, S. Chari, and R. P. Gooch, "Nonlinear techniques for the joint estimation of cochannel signals," *IEEE Trans. Commun.*, vol. 45, no. 4, pp. 473-484, Apr. 1997.

[30] J. Hagenauer and P. Hoeher, "A Viterbi algorithm with soft-decision outputs and its applications," in *Proc. IEEE Global Telecommunications Conf.*, Dallas, TX, Nov. 1989, vol. 3, pp. 47.1.1-47.1.7.

[31] L. Hanzo, T. H. Liew, and B. L. Yeap, *Turbo Coding, Turbo Equalisation and Space-Time Coding for Transmission over Fading Channels*. Chichester, UK: Wiley & Sons, 2002.

[32] S. Haykin, *Adaptive Filter Theory*, 3rd ed. New Jersey:Prentice-Hall, 1996.

[33] P. Hoeher, "TCM on frequency-selective fading channels: a comparison of soft-output probabilistic equalizers," in *Proc. IEEE Global Telecommunications Conf.*, 1990, vol. 1, pp. 376-381.

[34] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Vol I. New Jersey: Prentice-Hall, 1993.

[35] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Vol II. New Jersey: Prentice-Hall, 1998.

[36] W. Koch and A. Bair, "Optimum and sub-optimum detection of coded data disturbed by time-varing intersymbol interference," in *Proc. IEEE Global Telecommunications Conf.*, 1990, vol. 3, pp. 1679-1684.

[37] C. Laot, A. Glavieux, and J. Labat, “Turbo equalization: Adaptive equalization and channel decoding jointly optimized,” *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1744-1752, Sept. 2001.

[38] Y. Li, B. Vucetic, and Y. Sato, “Optimum soft-output detection for channels with intersymbol interference,” *IEEE Trans. Inform. Theory*, vol. IT-41, no. 3, pp. 704-713, May 1995.

[39] M. Mouly and M.-B. Pautet, *The GSM System for Mobile Communications*, France: Mouly-Pautet, 1992.

[40] T. Nagayasu, H. Kubo, K. Murakami, and T. Fujino, “A soft-output Viterbi equalizer employing expanded memory length in a trellis,” *IEICE Trans. Commun.*, vol. E80-B, no. 2, pp. 381-385, Feb. 1997.

[41] H. Omori and T. Asai, “A matched filter approximation for SC/MMSE iterative equalizers,” *IEEE Commun. Letters*, vol. 5, no. 7, pp. 310-312, July 2001.

[42] L. H. Ozarow, S. Shamai, and A. D. Wyner, “Information theoretic considerations for cellular mobile radio,” *IEEE Trans. Vehicular Technology*, vol. 43, no. 2, pp. 339-378, May 1994.

[43] W. W. Peterson and E. J. Weldon Jr., *Error-Correcting Codes*, 2nd ed. Cambridge, MA: MIT Press, 1972.

[44] R. L. Peterson, R. E. Ziemer, and D. E. Borth, *Introduction to Spread Spectrum Communications.* New Jersey: Prentice-Hall, 1995.

[45] A. Picart, P. Didier, and A. Glavieux, “Turbo-detection: A new approach to combat channel frequency selectivity,” in *Proc. IEEE Int. Conf. Communications*, 1997, pp. 1498-1502.

[46] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[47] M. Pukkila and P. A. Ranta, “Method for determining strength of co-channel signals, and a receiver,” Finnish Patent no. 103539, July 15, 1999.

[48] P. A. Ranta, A. Hottinen, and Z. C. Honkasalo, “Co-channel interference cancelling receiver for TDMA mobile systems,” in *Proc. IEEE Int. Conf. Communications*, Seattle, WA, 1995, pp. 17-21.

[49] P. A. Ranta, Z. Honkasalo, and J. Tapaninen, “TDMA cellular network application of an interference cancellation technique,” in *Proc. IEEE Vehicular Technology Conf.*, Chicago, Illinois, July 1995, pp. 296-300.

[50] P. A. Ranta, A. Lappeteläinen, and Z. C. Honkasalo, “Interference cancellation by joint detection in random frequency hopping TDMA networks,” in *Proc. IEEE Int. Conf. Universal Personal Commun.*, Cambridge, MA, Sept.-Oct. 1996, pp. 428-432.

[51] P. A. Ranta and A. Lappeteläinen, “Application of dominant interference cancellation in street microcells,” in *Proc. IEEE Int. Conf. Communications*, Montreal, Canada, June 1997, vol. 2, pp. 964-968.

[52] P. A. Ranta, “Cochannel interference cancellation in the GSM system,” Ph.D. dissertation, Dept. Electr. Eng., Helsinki Univ. of Technology, Helsinki, Finland, 1999.

[53] D. Reynolds and X. Wang, “Low-complexity turbo-equalization for diversity channels,” *Signal Processing*, Elsevier Science Publishers, vol. 81, no. 5, pp. 989-995, May 2001.

[54] B. Rislow, T. Maseng, and O. Trandem, “Soft information in concatenated codes,” *IEEE Trans. on Commun.*, vol. 44, no. 3, pp. 284-286, Mar. 1996.

[55] P. Robertson, E. Villebrun, and P. Hoeher, “A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain,” in *Proc. IEEE Int. Conf. Communications*, Seattle, WA, June 1995, pp. 1009-1013.

[56] M. Sandell, C. Luschi, P. Strauch, and R. Yan, “Iterative channel estimation using soft decision feedback,” in *Proc. IEEE Global Telecommunications Conf.*, Dec. 1998, vol. 6, pp. 3728-3733.

[57] S. J. Simmons, “Breadth-first trellis decoding with adaptive effort,” *IEEE Trans. Commun.*, vol. 38, no. 1, pp. 3-12, Jan. 1990.

[58] B. Sklar, *Digital Communications: Fundamentals and Applications.* New Jersey: Prentice-Hall, 1988.

[59] B. Sklar, "Rayleigh fading channels in mobile digital communication systems, Part I: Characterization," *IEEE Communications Mag.*, vol. 35, no. 7, pp. 90-100, July 1997.

[60] B. Steiner and P. Jung, "Optimum and suboptimum channel estimation for the uplink CDMA mobile radio systems with joint detection," *European Trans. Telecommun.*, vol. 5, no. 1, pp. 39-50, Jan.-Feb. 1994.

[61] P. Strauch, C. Luschi, and A. M. Kuzminskiy, "Iterative channel estimation for EGPRS," in *Proc. IEEE Vehicular Technology Conf.*, Sept. 2000, vol. 5, pp. 2271-2277.

[62] P. Strauch, C. Luschi, M. Sandell, and R. Yan, "Turbo equalization for an 8-PSK modulation scheme in a mobile TDMA communication system," in *Proc. IEEE Vehicular Technology Conf.*, 1999, pp. 1605-1609.

[63] G. Stüber, *Principles of Mobile Communications.* Boston, MA: Kluwer Academic Publishers, 1996.

[64] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744-765, Mar. 1998.

[65] D. P. Taylor, B. D. Hart, G. M. Vitetta, and A. Mämmelä, "Wireless channel equalisation," *European Trans. Telecommun.*, vol. 9, no. 2, pp. 117-143, Mar.-Apr. 1998.

[66] M. Tüchler, A. C. Singer, and R. Koetter, "Minimum mean squared error equalization using *a priori* information," *IEEE Trans. Signal Proc.*, vol. 50, no. 3, pp. 673-683, Mar. 2002.

[67] M. Tüchler, R. Otnes, and A. Schmidbauer, "Performance of soft iterative channel estimation in turbo equalization," in *Proc. IEEE Int. Conf. Communications*, New York, Apr.-May 2002, vol. 3, pp. 1858-1862.

[68] M. Tüchler, R. Koetter, and A. C. Singer, "Turbo equalization: principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754-767, May 2002.

[69] G. Ungerboeck, "Adaptive maximum likelihood receiver for carrier modulated data-transmission systems," *IEEE Trans. Commun.*, vol. COM-22, no. 5, pp. 624-636, May 1974.

[70] S. W. Wales, "Technique for cochannel interference suppression in TDMA mobile systems," *IEE Proc. Commun.*, vol. 142, no. 2, pp. 106-114, Apr. 1995.

[71] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA", *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046-1061, July 1999.

[72] J. H. Winters, "The diversity gain of transmit diversity in wireless systems with Rayleigh fading," *IEEE Trans. Vehicular Technology*, vol. 47, no. 1, pp. 119-123, Feb. 1998.

[73] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inform. Theory*, vol. IT-13, no. 2, pp. 260-269, Apr. 1967.

[74] J. Wolfmann, "Almost perfect autocorrelation sequences," *IEEE Trans. Inform. Theory*, vol. 38, no. 4, pp. 1412-1418, July 1992.

[75] H. Yoshino, K. Fukawa, and H. Suzuki, "Adaptive equalization with RLS-MLSE for fast fading mobile radio channels," in *Proc. IEEE Int. Symp. Circuits and Systems*, May 1992, vol. 2, pp. 501-504.

[76] H. Yoshino and H. Suzuki, "In-lab performance evaluation results of interference canceling equalizer (ICE)," in *Proc. IEEE 13th Int. Conf. Digital Signal Processing*, July 1997, vol. 2, pp. 1003-1006.

[77] H. Yoshino, K. Fukawa, and H. Suzuki, "Interference canceling equalizer (ICE) for mobile radio communication," *IEEE Trans. Vehicular Technology*, vol. 46, no. 4, pp. 849-861, Nov. 1997.

[78] H. Yoshino, K. Hirade, and H. Suzuki, "Field trial of interference canceling equalizer (ICE) for TDMA mobile communication systems," in *Proc. IEEE Int. Conf. Communications*, Vancouver, Canada, June 1999, vol. 2, pp. 952-957.

[79] *3rd Generation Partnership Project; Technical Specification Group GSM/EDGE*, 3GPP TS 45 series, 2002.