
Cyclostationary Blind Equalisation in Mobile Communications

Jon Altuna



A thesis submitted for the degree of Doctor of Philosophy.

The University of Edinburgh.

- October 1997 -

Abstract

Blind channel identification and equalisation are the processes by which a channel impulse response can be identified and proper equaliser filter coefficients can be obtained, without knowledge of the transmitted signal.

Techniques that exploit cyclostationarity can reveal information about systems which are nonminimum phase; nonminimum phase channels cannot be identified using only second-order statistics (SOS), because these do not contain the necessary phase information. Cyclostationary blind equalisation methods exploit the fact that, sampling the received signal at a rate higher than the transmitted signal symbol rate, the received signal becomes cyclostationary. In general, cyclostationary blind equalisers can identify a channel with less data than higher-order statistics (HOS) methods, and unlike these, no constraint is imposed on the probability distribution function of the input signal. Nevertheless, cyclostationary methods suffer from some drawbacks, such as the fact that some channels are unidentifiable when they exhibit a number of zeros equally spaced around the unit circle.

In this thesis the performance of a cyclostationary blind channel identification algorithm combined with a maximum-likelihood sequence estimation receiver is analysed. The simulations were conducted in the pan-European mobile communication system GSM environment and the performance of the blind technique was compared with conventional channel estimation methods using training. It is shown that although blind equalisation techniques can converge in a few hundred symbols in a time-invariant channel environment, the degradation with respect to methods with training is still considerable. Yet, the fact that a dedicated training sequence is not needed makes blind techniques attractive, because the data used for training purposes can be re-allocated as information data.

In the concluding part of this thesis a new blind channel identification algorithm which combines methods that exploit cyclostationarity implicitly and explicitly is presented. It is shown that the properties of cyclostationary statistics are exploited in the new algorithm, and enhance the performance of the technique that solely exploits fractionally-spaced sampling. The algorithm is robust in the presence of correlated noise and interference from adjacent users.

Declaration of originality

I hereby declare that this thesis and the work reported herein was composed and originated entirely by myself, in the Department of Electrical Engineering at the University of Edinburgh.

Jon Altuna Iraola

Acknowledgements

I would like to thank the following people for their invaluable assistance during the course of this PhD:

- Bernard Mulgrew and Peter Grant, my supervisors, for their continuous support and guidance. Also for reading and checking this thesis.
- The Mondragon E.Polytechnic and the Institution of Electrical Engineers for their financial support.
- To my parents and sister, for their support and understanding.
- To Iraide, for all the special moments that we have spent together and for supporting me in the difficult moments. Also for carefully reading this thesis.
- To the people in the Signals and Systems Group. Especially to Justin Fackrell, Iain Band, John Thompson, Sarat Patra, Iain Scott, Richard Stirling-Gallagher and Achilleas Stoglioglu for spending their precious time in some invaluable discussions. Also to Sarat Patra and John Thompson for reading and checking chapters of this thesis.

Contents

List of Figures	vi
List of Tables	ix
Abbreviations	x
List of Symbols	xii
1 Introduction	1
1.1 Introduction	1
1.2 Motivation	2
1.3 Blind Channel Identification and Equalisation	2
1.4 Cyclostationary Blind Equalisation	3
1.5 Application of Cyclostationary Blind Equalisation to Mobile Radio	5
1.6 Cyclic Subspace Algorithm for Blind Channel Identification	5
1.7 Summary of the Work	6
1.8 Thesis Organisation	7
2 Background	8
2.1 Introduction	8
2.2 Digital Communication System	9
2.3 Supervised versus Unsupervised Equalisation	12
2.4 A Brief History of Blind Equalisation Techniques	14
2.4.1 Bussgang algorithms	15
2.4.2 Higher-order statistics	17
2.4.3 Cyclostationary algorithms	19
2.5 Equaliser Structures for Blind Equalisation	22
2.5.1 Linear equaliser	23
2.5.2 Maximum-likelihood sequence estimator	30
2.6 Conclusions	34
3 Cyclostationarity	36
3.1 Introduction	36
3.2 Fundamental Definitions	36
3.2.1 The cyclic autocorrelation function	39
3.2.2 The cyclic spectrum or spectral correlation density function	40
3.3 Properties of Cyclostationarity	41
3.3.1 Phase recovery	41
3.3.2 Signal selectivity	42

3.4	Cyclostationary Blind Equalisation Algorithms	43
3.4.1	Oversampling system configuration	44
3.4.2	Cyclic statistics methods	45
3.4.3	Multichannel methods	49
3.4.4	Identifiability conditions	53
3.5	Conclusions	55
4	Application of Cyclostationary Blind Equalisation to Mobile Radio	56
4.1	Introduction	56
4.2	The Global System for Mobile	57
4.2.1	Characterisation of the GSM system	57
4.2.2	Modulation scheme	60
4.2.3	The propagation channel	60
4.3	Performance Comparison of Cyclostationary Techniques in Stationary Multipath Channel Environments	62
4.3.1	Moving-average channel parameter estimation	63
4.3.2	Equalisation results	73
4.4	Performance Comparison of Blind Channel Identification Techniques versus Supervised Methods in a Mobile Radio Environment	79
4.4.1	Adaptive and non-adaptive supervised channel estimation methods	79
4.4.2	Blind channel estimation method	82
4.4.3	Typical Urban (TU) environment	82
4.5	Discussion	92
4.6	Conclusions	96
5	A Cyclic Subspace Algorithm for Blind Channel Identification	97
5.1	Introduction	97
5.2	A Cyclic Subspace Algorithm for Blind Channel Identification	98
5.2.1	Problem formulation	98
5.2.2	Channel identification using cyclostationary statistics and subspaces	100
5.3	Asymptotic Performance	109
5.4	Correlated Noise Compensation	111
5.5	Blind Multiuser Detection	112
5.5.1	Multiuser blind equalisation using subspace decomposition	114
5.6	Simulation Results	116
5.6.1	Experiment I	117
5.6.2	Experiment II	118
5.7	Discussion	120
5.8	Conclusions	127
6	Conclusions	128
6.1	Introduction	128
6.2	Achievements of the Work	128
6.3	Limitations of the Work	130
6.4	Areas for Future Work	130

A Identifiability Conditions	132
A.1 Identifiability condition I	132
A.2 Identifiability condition II	135
A.3 Identifiability condition III	136
B Singular Value Decomposition	138
B.1 Inverse of a square matrix	139
B.2 Relation to eigen-decomposition	139
C Original publications	141
References	168

List of Figures

2.1	Continuous-time description of a base-band digital communications system. . . .	10
2.2	Pulses with raised cosine spectrum	11
2.3	Overlapping raised cosine spectrum pulses for $0 < \beta < 1$ and $-\infty < n < \infty$	11
2.4	Receiver structure of a conventional supervised adaptive equaliser.	13
2.5	Block diagram of a Bussgang blind equaliser.	14
2.6	Cost function $J(\theta)$ of Godard Constant Modulus Algorithm for an ARMA channel $h(z) = 1/(1 + 0.5z^{-1})$ and a binary input data. The two global minima are situated at $(\theta_1, \theta_2) = (1, 0.5)$ and $(\theta_1, \theta_2) = (-1, -0.5)$	17
2.7	Minimum phase and maximum phase channels	21
2.8	Transversal Filter.	23
2.9	Fractionally-spaced equaliser.	24
2.10	a) Typical spectrum for a synchronous equaliser. $P(f)$ is the spectrum of the transmitted pulse and $C(f)$ is the Fourier transform of the tap coefficients. b) Typical spectrum for a fractionally-spaced equaliser with tap spacing $T_s < T/(1 + \alpha)$	25
2.11	a) Zero position of channel with transfer function $H(z) = (1 - 0.2z^{-1})(1 - 0.8z^{-1})(1 - 1.8z^{-1})(1 - 2z^{-1})$; b) unstable inverse filter	29
2.12	Separate equalisation of maximum phase and minimum phase factors of transfer function $H(z) = (1 - 0.2z^{-1})(1 - 0.8z^{-1})(1 - 1.8z^{-1})(1 - 2z^{-1})$; $H_m(z) = 1 - z^{-1} + 0.16z^{-2}$ and $H_M(z) = 1 + 0.2z^{-1} - 3.6z^{-2}$: Inverse impulse response of minimum phase filter (top left), inverse of reversed maximum phase filter (bottom left), inverse of nonminimum phase channel $h(n)$ (bottom right) and peak distortion (top right)	31
2.13	Tree diagram of Viterbi algorithm for a 4-QAM ($M = 4$) signal and $L = 2$. Only the $M^L = 4^2 = 16$ sequences marked with the ellipse survive to the next stage.	33
3.1	Spectral correlation density of one BPSK signals with centre frequency $f_{0,1} = 25$ Hz.	43
3.2	Spectral correlation density of two AM signals with carrier frequencies $f_{0,1} = 20$ Hz and $f_{0,1} = 30$ Hz.	44
3.3	Spectral correlation density of a BPSK signal buried in two AM signals.	44
3.4	Discrete-time oversampled system model; a) fractionally-spaced sampling operation, b) multichannel representation.	46
4.1	GSM frame structure.	59
4.2	Mobile channel simulator.	63
4.3	Estimated, mean and true positions of the zeros for channels CH1, CH2, CH3 and CH4. The estimates of the zeros were obtained for 20 Monte-Carlo realisations of the signal subspace blind channel estimation technique of Moulines <i>et al.</i>	66
4.4	Estimated, mean and true positions of the zeros for channels CH1, CH2, CH3 and CH4 for an oversampling factor $P = 3$. The estimates of the zeros were obtained for 20 Monte-Carlo realisations of the cyclic blind channel estimation technique of Hatzinakos	68
4.5	Estimated, mean and true positions of the zeros for channels CH1, CH2, CH3 and CH4 for an oversampling factor $P = 5$. The estimates of the zeros were obtained for 20 Monte-Carlo realisations of the cyclic blind channel estimation technique of Hatzinakos	69

4.6	Estimated, mean and true positions of the zeros for channels CH1, CH2, CH3 and CH4. The estimates of the zeros were obtained for 20 Monte-Carlo realisations of the blind multichannel estimation technique of Tong <i>et al.</i>	72
4.7	Normalised mean-square error (NMSE) of the channel parameter estimates, conducted over 20 Monte-Carlo runs and different number of symbols used in the estimation of the channel.	74
4.8	The optimum symbol-spaced linear equaliser is represented by the cascade of a matched filter, a sampler and a symbol-spaced transversal filter. The transversal filter represents the cascade of a noise whitening filter and a linear filter matched to the channel.	75
4.9	Equivalent fractionally-spaced equaliser. $F(z_\Delta)$ represents the transfer function evaluated at the fractionally-spaced sampling period.	75
4.10	Maximum-likelihood sequence estimation in a receiver where the received signal after the matched-filter is sampled at twice the symbol rate.	76
4.11	Bit-error-rate performance of fractionally-spaced linear equalisation, using cyclic, subspace and TXK blind equalisation methods, for channel CH1.	78
4.12	Bit-error-rate performance of symbol-spaced linear equaliser and MLSE for different delays used in the decisions.	79
4.13	Magnitude of the time-variant impulse response of the combined transmitter and receiver filter with a Typical Urban propagation channel, for Doppler frequencies of $f_d = 40$ Hz and $f_d = 125$ Hz.	83
4.14	Probability of error in each bit position of a 58-bit GSM information data burst for Doppler spectra a) $f_d = 40$ Hz and b) $f_d = 125$ Hz in a $E_s/N_0 = 23$ dB scenario; statistics on 2225 bursts in a <i>Typical Urban</i> environment.	84
4.15	BER performance of the subspace method for blind channel identification combined with a MLSE receiver in a Typical Urban environment. The Doppler frequency is $f_d = 40$ Hz and a 4-QAM signal has been used.	85
4.16	BER performance of the subspace method for blind channel identification combined with a MLSE receiver in a Typical Urban environment. The Doppler frequency is $f_d = 125$ Hz and a 4-QAM signal has been used.	85
4.17	BER performance of supervised adaptive and non-adaptive methods, and blind channel estimation method in a Typical Urban environment. The Doppler frequencies are $f_d = 40$ Hz and $f_d = 125$ Hz and a 4-QAM signal has been used.	86
4.18	BER performance of supervised adaptive and non-adaptive methods, and blind channel estimation method in a Typical Urban environment for different delays used in the MLSE.	87
4.19	Impulse response of the combined transmitter and receiver filter and a time-invariant channel with Typical Urban environment tap-delay profile.	88
4.20	BER performance of the subspace method for blind channel identification combined with a MLSE receiver in a Typical Urban environment. The Doppler frequency is $f_d = 40$ Hz and a BPSK signal has been used.	89
4.21	BER performance of the subspace method for blind channel identification combined with a MLSE receiver in a Typical Urban environment. The Doppler frequency is $f_d = 125$ Hz and a BPSK signal has been used.	90
4.22	BER performance of supervised adaptive and non-adaptive methods, and blind channel estimation method in a Typical Urban environment. The Doppler frequencies are $f_d = 40$ Hz and $f_d = 125$ Hz and a BPSK signal has been used.	90
4.23	Comparison of the BER results obtained using BPSK and half-rate 4-QAM signals for $f_d = 40$ Hz.	91
4.24	Comparison of the BER results obtained using BPSK and half-rate 4-QAM signals for $f_d = 125$ Hz.	92
4.25	BER performance of the subspace method for blind channel identification combined with a MLSE receiver in a Typical Urban environment. The Doppler frequency is $f_d = 40$ Hz and a double-rate 4-QAM signal has been used.	93
4.26	BER performance of the subspace method for blind channel identification combined with a MLSE receiver in a Typical Urban environment. The Doppler frequency is $f_d = 125$ Hz and a double-rate 4-QAM signal has been used.	94

4.27	BER performance of supervised adaptive and non-adaptive methods, and blind channel estimation method in a Typical Urban environment. The Doppler frequencies are $f_d = 40$ Hz and $f_d = 125$ Hz and a double-rate 4-QAM signal has been used.	94
4.28	Comparison of the BER results obtained using BPSK and double-rate 4-QAM signals for $f_d = 40$ Hz.	95
4.29	Comparison of the BER results obtained using BPSK and double-rate 4-QAM signals for $f_d = 125$ Hz.	95
5.1	Oversampled multichannel system configuration.	99
5.2	Antenna-pattern of a receiver which cancels-out the signals from the undesired mobile users.	113
5.3	Multiple-input multiple-output (MIMO) system representation.	114
5.4	MSE of channel parameter estimates for a transmitter and receiver filters using a roll-off factor a) $\beta = 0.1$, b) $\beta = 0.5$ and c) $\beta = 1$, for different E_b/N_0 scenarios (left column, 1000 symbol used in the estimation), and for different number of symbols used in the estimation (right column, $E_b/N_0 = 25$ dB).	119
5.5	True and estimated channel of desired user for different SIR scenarios and SNR = 100 dB; a) SIR = -3 dB (6000 symbols used in the estimation), b) SIR = 0 dB (1000 symbols used in the estimation) c) SIR = 3 dB (1000 symbols used in the estimation).	121
5.6	True and estimated channel of desired user for a SIR = 3 dB scenario, and a) SNR = 30dB and b) SNR = 15 dB noise conditions. 1000 symbols were used in the estimation.	122
5.7	True, mean and true \pm standard deviation values of the estimation of desired user channel, for different SIR scenarios and SNR = 100 dB; a) SIR = -3 dB (6000 symbols used in the estimation), b) SIR = 0 dB (1000 symbols used in the estimation) c) SIR = 3 dB (1000 symbols used in the estimation).	123
5.8	True, mean and true \pm standard deviation values of the estimation of desired user channel, for a SIR = 3 dB scenario, and a) SNR = 30dB and b) SNR = 15 dB noise conditions. 1000 symbols were used in the estimation.	124
5.9	MSE of desired user channel parameter estimate for different interference levels; a) SIR = -3 dB, SNR = 100 dB, b) SIR = 0 dB, SNR = 100 dB and c) SIR = 3 dB.	125
A.1	Positions of zeros of the components of the spectral correlation density (SCD) function; a) (cross) zeros of $\prod_{k=0}^{P-1} (1 - \rho_k z_{\Delta}^{-1})$, (circle) zeros of $\prod_{k=0}^{P-1} (1 - \rho_k^* z_{\Delta} e^{-j2\pi/T})$; b) and c) (circle) zeros of $\prod_{k=0}^{P-1} (1 - \rho_k^* z_{\Delta})$, (cross) zeros of $\prod_{k=0}^{P-1} (1 - \rho_k z_{\Delta}^{-1} e^{-j2\pi/T})$	133
A.2	Position of the zeros for a single tap channel with a reflector occurring at the symbol period T . a) position of the zero at the symbol rate; b) position of the zeros for an oversampling factor $P = 3$; c) position of the zeros for an oversampling factor $P = 4$	136

List of Tables

4.1	Power-delay profile for a typical (non-hilly) urban environment.	62
4.2	Differential cepstrum parameters for channels CH1, CH2, CH3 and CH4.	70
4.3	Differential cepstrum parameters with interpolated values for channels CH1, CH2, CH3 and CH4 for an oversampling factor $P = 3$. A Newton-Aitken interpolation technique of degree 3 was used which requires knowledge of the 2 previous and 2 posterior values of the differential cepstrum parameters.	71
4.4	Differential cepstrum parameters with interpolated values for channels CH1, CH2, CH3 and CH4 for an oversampling factor $P = 5$. A Newton-Aitken interpolation technique of degree 3 was used which requires knowledge of the 2 previous and 2 posterior values of the differential cepstrum parameters.	71
4.5	Received signal quality vs. channel bit error rate in GSM [83].	93

Abbreviations

AR	Auto regressive
ARMA	Auto regressive moving average
BER	Bit error rate
CMA	Constant modulus algorithm
DOA	Direction of arrival
ESPRIT	Estimation of signal parameters via rotational invariance techniques
FDD	Frequency division duplex
FDMA	Frequency division multiple access
FEC	Forward error correction
FIR	Finite impulse response
FSE	Fractionally spaced equaliser
FOT	Fraction of time
FM	Frequency modulation
GMSK	Gaussian minimum shift keying
GSM	Global system for mobile
HDTV	High definition television
HOS	Higher order statistics
i.i.d.	Independent identically distributed
ISI	Inter symbol interference
LMS	Least mean squares
LOS	Line of sight
LS	Least squares
LTI	Linear time invariant
MA	Moving average
MIMO	Multiple input multiple output
MLSE	Maximum likelihood sequence estimator
MMSE	Minimum mean square error

MSE	Mean square error
MUSIC	Multiple signal classification
NMSE	Normalised mean square error
PdF	Probability density function
PDF	Probability distribution function
PSD	Power spectral density
QAM	Quadrature amplitude modulation
QoS	Quality of service
RLS	Recursive least squares
R/W	Transmission rate to bandwidth ratio
SCD	Spectral correlation density
SCORE	Spectral coherence restoral
SFDMA	Space frequency division multiple access
SIMO	Single input multiple output
SIR	Signal to interference ratio
SNOI	Signal not of interest
SNR	Signal to noise ratio
SOI	Signal of interest
SOS	Second order statistics
SOCS	Second order cyclic statistics
SSE	Symbol spaced equaliser
SVD	Singular value decomposition
TDMA	Time division multiple access
TU	Typical urban
TXK	Tong, Xu and Kailath
WMF	Whitened matched filter
ZF	Zero forcing

List of principal symbols

$\Re\{\cdot\}$	Real part of complex number.
$\Im\{\cdot\}$	Imaginary part of complex number.
$P(\cdot)$	Probability operator.
\otimes	Convolutional operator.
$e^{j\theta}$	Complex phasor.
ω	Angular frequency.
k	Discrete time variable.
σ^2	Variance.
λ_i	Eigenvalues of a given matrix.
$x(k)$	Source signal sample at time k .
$\hat{x}(k)$	Estimated source signal sample at time k .
\underline{x}_k	Vector of source signal samples at time k .
$s(k)$	Pulse-shaped signal sample at time k .
$c(k)$	Propagation channel impulse response.
$h(n)$	Overall combined transmitter filter, propagation channel and receiver filter impulse response sample at time n .
$H(z)$	Overall channel transfer function.
$h^{(i)}(n)$	Impulse response of i th subchannel.
$y(n)$	Received signal sample at time n .
$y^{(i)}(n)$	Received signal at the i th sensor.
$\underline{y}^{(i)}(n)$	Vector of received signal samples at the i th sensor.
$v(n)$	Additive noise sample at time n .
$v^{(i)}(n)$	Noise at the i th sensor.
$d(n)$	Desired or reference signal.
$p(t)$	Continuous-time raised-cosine spectrum filter.
T	Source signal symbol period.
T_s	Sampling period.
P	Oversampling factor.

W	Bandwidth.
β	Roll-off factor of raised-cosine spectrum filter.
α	Cyclic frequency.
$r_x(m, n)$	Autocorrelation function of process x at time m , lag n .
$S_x(f)$	Power spectrum of process x .
$S_x^\alpha(f)$	Cyclic spectrum of process x at cyclic frequency α .
c_k	Transversal filter weights.
$\hat{\underline{c}}_{opt}$	Vector of optimum estimated filter weights.
$R_x^\alpha(m)$	Cyclic autocorrelation function of process x at cyclic frequency α .
$\delta(\cdot)$	Dirac delta.
$A(k), B(k)$	Differential cepstrum parameters.
R_x	Autocorrelation matrix of process x .
R_x^α	Cyclic autocorrelation matrix of process x .
Σ	Diagonal matrix with singular values.
f_d	Doppler frequency.
I	Identity matrix.
μ	Step size parameter in steepest gradient and least mean squares algorithms.
$E\{\cdot\}$	Expectation operator.
$tr()$	Matrix trace.
$(\cdot)^T$	Simple transposition.
$(\cdot)^*$	Simple conjugation.
$(\cdot)^H$	Complex conjugate transpose.
$(\cdot)^\dagger$	Pseudo inverse.

Chapter 1

Introduction

1.1 Introduction

With the development of new technologies in digital communications, new services have emerged and new problems have arisen. In digital high-definition television systems currently being developed for terrestrial broadcast applications [99], the transmitted signal is limited both in bandwidth and power, and the channel is characterised by multipath and interference from other stations. The conventional solution against multipath is the use of adaptive equalisers with training periods, in the presence of not very strong interference. During the transmission of the training sequence, a replica of the transmitted signal is available at the receiver, so that proper adjustments can be made at the receiver. When the interference is severe, the transmission of a training signal becomes impractical, because the adaptive equalisers have to be re-initialised constantly. Moreover, since the overall radio spectrum allocated to television broadcasting is to remain unchanged, channels that had previously been designated as unusable, because of the levels of interference, are to be used for HDTV. In these circumstances, it might be desirable to equalise the channel without the use of the training signal. These methods are commonly known as self-recovery or blind equalisation methods.

This thesis is concerned with the study of blind equalisation techniques which exploit cyclostationarity. Cyclostationarity is a statistical property present in most communication signals and it is measured in terms of the spectral correlation exhibited by the signals. Communication signals exhibit spectral correlation at harmonics of their carrier frequencies, baud rates, keying rates etc... Cyclostationarity can be used in many areas of signal processing and communications, and among these in blind equalisation. The relationship between fractionally-spaced sampling and cyclostationarity has led to the development of many blind equalisation methods that sample the received signal at a rate higher than the symbol rate of the transmitted signal [57].

The chapter begins with an exposition of the principal motivations for the work undertaken in this thesis. Then, the problem of blind channel identification using fractionally-spaced sampling is discussed. Following this, the main areas of achievement of this thesis are outlined and finally the organisation of the thesis is described.

1.2 Motivation

Blind equalisation is a difficult task because of the little or null information known about the transmitted signal and the propagation channel characteristics. The complexity of the majority of the existing blind equalisation algorithms and the slow convergence have precluded their application. However, in the last five years, blind equalisation algorithms have improved, and faster techniques have emerged.

The first methods to be used for blind equalisation were known as self-recovery techniques [77] [9] [41], because of the ability to identify a channel with as little information of the transmitted signal as the knowledge of the signal alphabet and the symbol period. These techniques were formulated as gradient-descent type algorithms, where the minimisation of a quadratic cost function was required. Because a nonlinear decision function is involved, these techniques exhibit undesired local minima as well as global minima. The convergence of the algorithms to a global minima is not guaranteed.

Later, methods that exploited higher-order statistics (HOS) were proposed [67] [37] [16] [38]. Unlike methods that use only second-order statistics (SOS), HOS preserve phase information of a system; this information is essential in the identification of mixed phase or nonminimum phase systems. Cumulants are for HOS what the autocorrelation is for SOS. Third- and fourth-order cumulants are commonly used in moving-average (MA) channel parameter estimation. Nevertheless, large amounts of data are required to estimate the cumulants from data samples.

In the attempt to build algorithms which are faster and computationally less expensive, fractionally spaced blind equalisation algorithms have enjoyed a lot of attention. Two are the main objectives of this thesis: *i*) the analysis of the practical aspect of fractionally-spaced blind equalisation; it is important to acknowledge that the development of new techniques needs, undoubtedly, to go in parallel with a thorough analysis of the behaviour of these techniques in real applications; and *ii*) secondly, the description of a new time-domain cyclostationary blind equalisation algorithm.

1.3 Blind Channel Identification and Equalisation

An essential part in the investigation of application-oriented signal processing techniques is an accurate modelling of the system. Systems can be either linear or nonlinear, stationary or time-varying. Signals which obey a certain probabilistic law are treated as random signals. Due to the advanced state of development of digital technology in signal processing and communications, the treatment of signals and in general of the whole system is done in discrete-time, rather than in continuous-time.

The systems found in communications are mostly linear or are treated as linear, because the whole problem of channel identification becomes less complicated. The discrete-time modelling

of the channel classifies the systems in terms of *finite-impulse response* (FIR) or *infinite-impulse response* (IIR) channels. The channels are usually described in polynomial form in the z -domain, by means of their transfer function [72].

Channels which can be expressed only in terms of the numerator of a polynomial are known as *moving-average* (MA) systems. The characteristic of this type of channels is that the output of the system at a specific time instant only depends on the value of the input signal at that time, plus contributions from past and future values of the input signal. The channel can also be expressed as a polynomial with only a denominator. These type of systems are known as *auto-regressive* (AR) systems. The output of an AR model is expressed exclusively in terms of the actual value of the input and past values of the output. Finally, the *auto-regressive moving-average* (ARMA) systems are described in terms of a polynomial, where the actual value of the output depends on the actual and previous values of the input, and past values of the output. In order to guarantee the stability of the AR and ARMA systems, these cannot depend on future values of the output [47].

The term blind equalisation is often associated with blind deconvolution. Blind deconvolution is regarded as the action of inverse filtering of a system, which can either be *minimum phase* (all zeros inside the unit circle), *maximum phase* (all zeros outside the unit circle) or mixed phase (zeros inside and outside the unit circle). The last two are often known as nonminimum phase systems. Difficulties arise in the inverse filtering operation of nonminimum phase systems, since zeros outside the unit circle become poles outside the unit circle. In that case, the inverse filter is unstable, and unless a delay is used in the inverse filtering operation, the system becomes unrealisable [19].

1.4 Cyclostationary Blind Equalisation

Some of the inherent problems of blind equalisation were shown to be solved by exploiting fractionally-spaced sampling at the receiver. It was first shown by Tong *et al.* [87], that the oversampled received sequence contains sufficient information to identify nonminimum phase channels. These problems are associated with the fact that a minimum phase channel and a maximum phase channel can have the same magnitude response in the frequency domain, but different phase response. It is known that second-order statistics can only preserve magnitude information, but that the phase information is lost [72].

In the early years after the emergence of the first fractionally-spaced blind equalisation algorithms, it was shown that exploiting the cyclostationary nature of the oversampled received sequence, either implicitly or explicitly, provides the necessary phase information of the system in the frequency domain. From a statistical point of view, the cyclic spectrum, also known as the spectral correlation density (SCD) function, showed that the spectral correlation of two frequency-shifted versions of the channel frequency response contained the phase information that, for example, the power spectrum did not contain [36]. However, spectral correlation

could only be achieved if the channel were not strictly band-limited to the reciprocal of the transmitted signal symbol period [36] [45].

Statistical methods, either in the time domain using the cyclic autocorrelation function or in the frequency domain using the cyclic spectrum, enjoyed considerable attention during those early years [98] [53] [24]. It was also observed that, because the cyclostationary process can be decomposed in multiple stationary sequences [33], some rich matrix structure could be exploited, by arranging those stationary subchannels in a certain manner. This led to the development of more data-efficient methods. The method proposed by Moulines *et al.* [60] [61], and the method by Slock [81] can be presented in this context. It was the fact that many of the statistical methods needed large amounts of data for the estimation of the particular statistics, while the algebraic methods, also known as multichannel methods, just required the estimation of the autocorrelation function to obtain subspace information.

Nevertheless, the shift in the attention from statistical methods to algebraic methods has inevitably left statistical methods in the ostracism. One of the objectives of this thesis is to demonstrate that cyclostationary statistics contain properties which are complementary to multichannel methods.

There has been considerable debate as to the relative advantages and disadvantages of HOS-based and cyclostationary-based methods [10] [39]. In general, the advantages of cyclostationary blind equalisation can be summarised as follows:

- Convergence within a few hundred symbols,
- identification of nonminimum phase channels,
- no constraint imposed on the probability distribution function of the input signal.

Within this thesis, a comparison of the performance of three cyclostationary blind equalisation methods is presented. Two multichannel methods and a statistical method are used in this comparison. In general, the following channels are unidentifiable using fractionally-spaced blind channel identification methods:

- (i). Channels with P zeros equally-spaced around the unit circle, where P is an integer which represents the oversampling factor. This condition is equivalent to having P subchannels sharing the same zero, in a multichannel representation of an oversampled sequence.
- (ii). Channels with reflections that are multiples of the symbol period T .
- (iii). Channels with reflections occurring at $T/2$ and the oversampling factor P is even.

These channels are strictly unidentifiable by any fractionally-spaced blind equalisation method. The first condition, although critical, is very rare and can be resolved by choosing a different

oversampling factor P . Channel type (ii) is unidentifiable for any oversampling factor P and channel type (iii) can be resolved by selecting an odd oversampling factor P .

Chapter 4 presents some results on the convergence of these algorithms, and the implementation of a fractionally-spaced equaliser structure. Fractionally-spaced blind equalisation leads directly to fractionally-spaced equalisation (FSE), which has been shown to improve the clock recovery in data modems.

1.5 Application of Cyclostationary Blind Equalisation to Mobile Radio

Data transmission over mobile radio channels faces two major difficulties. First, there is severe inter-symbol interference. Second, the channel may vary rapidly with time. To mitigate these effects, an adaptive equaliser is used. However, the coefficients of the equaliser have to be updated continuously, using a periodically transmitted training sequence.

The motivation behind the analysis and comparison of cyclostationary blind equalisation algorithms in a mobile radio environment, is to determine how these perform compared to supervised methods that use training. Several supervised methods are available; they are classified according to the algorithm used in the estimation of the channel during the training period. The *least-squares* (LS), *least-mean squares* (LMS) and *recursive-least squares* (RLS) methods are some of the standard techniques. The channel estimate can be kept fixed during the recovery of the data (non-adaptive supervised method), or it can constantly be updated during equalisation (adaptive supervised method), by means of some adaptive channel estimation technique (LMS or RLS).

One of the constraints of the channel estimation techniques is the assumption of quasi-stationarity of the channel, over the duration of the training sequence. Blind channel estimation techniques would require at least the data from one time-slot to estimate the channel correctly, so that an acceptable bit-error rate is achieved by the equaliser. The assumption of quasi-stationarity is weaker in that case, and it is accentuated as the maximum Doppler frequency of the system increases. The issue is to determine whether the performance of the blind equaliser is good enough to switch to a decision-directed mode, with the guarantee that the decisions made by the equaliser are good enough, so that the decision-directed mode can be kept in operation.

1.6 Cyclic Subspace Algorithm for Blind Channel Identification

Multiuser blind channel identification and multiuser signal detection have concentrated considerable attention, because it is often the case that it is preferable to detect and separate a desired signal, without knowledge of the training sequence of the interfering users. The blind channel

identification methods based on the multichannel representation of the oversampled sequence, have been extended to accommodate multiple signal sources. However, in most of the cases, all the channels of all the users have to be estimated, and in general, the rank of the matrix grows with the number of signals present in the system. Moreover, because the autocorrelation function of the received signal will have contributions from all the desired and undesired signals, it is difficult to remove the ambiguity present in the subspaces and recover the channel of the desired user.

This thesis presents a new blind channel identification algorithm based on subspace decomposition of the cyclic autocorrelation matrix of an oversampled received sequence. The proposed cyclic subspace algorithm treats the interfering users as interference or noise; then, the contribution of the noise, whether this is white or coloured, is null in theory, because it does not exhibit cyclostationarity at any particular cyclic frequency. If the interfering users do not exhibit cyclostationarity at the cyclic frequency of the desired user, the effect of them in the cyclic autocorrelation function of the received signal will be negligible. In this respect, the cyclic subspace algorithm is less sensitive to correlated noise and co-channel interference than conventional subspace methods. The price to pay for this is that generally, a longer data set is necessary for the estimation of the cyclic autocorrelation function under white noise conditions.

Numerical results are presented to show the capability of the cyclic subspace algorithm with respect to correlated noise and co-channel interference. The sufficient and necessary identifiability conditions are established and an equation to calculate the asymptotic variance of the channel estimation is derived.

1.7 Summary of the Work

This thesis examines the problem of cyclostationary blind equalisation in mobile radio communications and presents a new algorithm for blind channel identification using cyclostationary statistics and subspace methods. Subspace decomposition is widely used in direction-of-arrival (DOA) estimation [85]. In fact, the new cyclic subspace algorithm is reminiscent of the Cyclic MUSIC [78] algorithm for DOA estimation.

It is shown that a subspace decomposition of the cyclic autocorrelation function of an oversampled received sequence is possible, and that it contains the information which enables the identification of a FIR channel impulse response. It is crucial, however, that the received signal at a particular subchannel exhibits cyclostationarity, but that can be achieved by means of oversampling the signal received at the subchannel. The cyclic autocorrelation function of additive white or coloured Gaussian noise is ideally zero at a particular cyclic frequency α . The cyclic autocorrelation is especially useful in the presence of multiple interfering users, because the cyclic autocorrelation of the interfering users is negligible (provided that the interfering users exhibit cyclostationarity at a different cyclic frequency from the desired user).

1.8 Thesis Organisation

The remainder of this thesis is organised as follows: Chapter 2 discusses the problem of blind equalisation and the different approaches are summarised briefly. Two equaliser structures used in blind equalisation are also described: i) the linear equaliser and by extension the fractionally-spaced linear equaliser, and ii) the maximum-likelihood sequence estimator.

Chapter 3 is devoted to the theory of cyclostationarity and its application to blind equalisation. General equations for the cyclic autocorrelation function and the cyclic spectrum are derived and the properties of cyclostationarity described. The link between cyclostationarity and fractionally-spaced sampling is established. Some of the most salient fractionally-spaced sampling blind channel identification and equalisation methods are reviewed in a unified framework; the general identifiability conditions, as well as the specific of each of the algorithms, are established.

In Chapter 4, the issue of cyclostationary blind equalisation in a mobile radio environment is considered. The pan-European standard for Mobile GSM is first described. The convergence properties and tracking difficulties of a number of fractionally-spaced blind channel identification algorithms are shown by means of simulations, using stationary multipath channels. Finally, some new results on the comparison of the performance of one particular blind channel identification algorithm and supervised adaptive and non-adaptive channel identification algorithms, in a mobile radio environment, are presented.

Chapter 5 introduces a new blind channel identification algorithm based on the combination of cyclostationary statistics and multichannel methods. The necessary and sufficient conditions for identifiability are established and an expression for the asymptotic variance of the channel estimation is derived.

Chapter 6 summarises the conclusions of this work and gives an insight for areas for further research.

Chapter 2

Background

2.1 Introduction

During the transmission time in a typical communication process, the original signal is exposed to severe distortion caused by the transmission medium such as multipath fading, inter-symbol interference, electromagnetic interference or additive thermal noise. The aim of an equaliser is to recover the transmitted signal by means of removing distortion from it. The two unknown variables at the receiver are the sequence of transmitted symbols and the propagation characteristics, while the only available information is the received signal. This converts the equalisation problem into a single equation problem with two unknowns. The conventional approach to equalisation involves the transmission of a training or pilot signal periodically interposed with the samples of data which are also available at the receiver. During that period only the channel characteristics are unknown and these can be properly identified with the knowledge of the transmitted and received signals.

The basic linear equaliser consists of a weighted tap-delayed line whose weights are the FIR approximation of the inverse channel filter [19]. The conventional equaliser procedure is made adaptive in order to reduce the errors produced in recovering the transmitted signal. The coefficients of the equaliser are constantly updated in order to reduce the error produced in the recovery of the transmitted signal. This error is in part caused by the thermal noise and in part by the impairments caused by the propagation channel, known as inter-symbol interference (ISI). The error signal is constructed using the training signal as the reference signal during the training period or using the known decisions made by the equaliser when no training signal is available (decision directed mode). However, when the propagation channel characteristics vary very quickly in time the adaptive equaliser cannot update the weights quickly enough and the equaliser is simply unable to follow the channel.

In broadcast systems, not only the receiver suffers from the variations on environmental conditions which cause unwanted reflections from clouds, hills and buildings, but it is also affected by electromagnetic interference in the form of storms and electrical discharges from lightnings. In a broadcast system, because of the nature of the system, the receivers can be switched on at any time during the transmission and unless a reference signal is sent periodically, initial connection can be quite difficult. Multipoint data networks highlight another classic problem,

that of re-synchronising. If the *data terminal* is unable to recognise data correctly during the transmission, re-synchronisation is not made possible until the *control-unit* is polled to that data terminal. This could be caused because of the distortion in the channel characteristics or simply because it was not switched on.

Thus, channel characteristics variation and synchronisation highlight the need for some kind of untrained or unsupervised equalisation procedure which will be used to start-up or re-tune the equaliser. This type of equalisers are known as blind equalisers. This chapter presents some background information about existing blind equalisation techniques and equaliser structures used in blind equalisation. The chapter is organised as follows: In section 2.2, a model of a digital communication system is described; in section 2.3, a discussion about supervised and unsupervised methods is presented; in section 2.4, the existing blind equalisation techniques are summarised; in section 2.5, some of the equaliser structures used in blind equalisation are presented and finally, section 2.6 draws some conclusions on this background chapter.

2.2 Digital Communication System

Throughout this thesis a discrete-time model of the otherwise continuous-time analog digital communication system will be used. The configuration of the system is shown in Figure 2.1 where the transmission of a digital signal through a band-limited channel subject to multipath effects is described. The digital data with source symbol period T is first modulated according to an appropriate modulation scheme depending on the specific application [71] and then is driven through a low pass filter which will accommodate the modulated signal a_n to the available channel bandwidth W .

This low-pass filter is designed as a pulse shaping filter whose shape will be determined by the constraint imposed by the limitation in bandwidth availability. Figure 2.2 describes some of the most common pulse shapes. These can be expressed as:

$$p(t) = \frac{\sin(\pi t/T)}{\pi t/T} \frac{\cos(\pi \beta t/T)}{1 - (2\beta t/T)^2}. \quad (2.1)$$

These pulses which exhibit a raised-cosine spectrum can be classified according to how tight the limitation in bandwidth availability is. If the source symbol period is equal to half the reciprocal of the channel bandwidth $T = 1/2W$ ¹ the signal is strictly band-limited to the available bandwidth and there is only one unique filter which satisfies this condition. This filter is the *sinc* function which can be expressed as:

$$p(t) = \frac{\sin(\pi t/T)}{\pi t/T}. \quad (2.2)$$

¹This condition is the critical Nyquist criterion and it means that the sampling frequency needs to be at least twice the system bandwidth

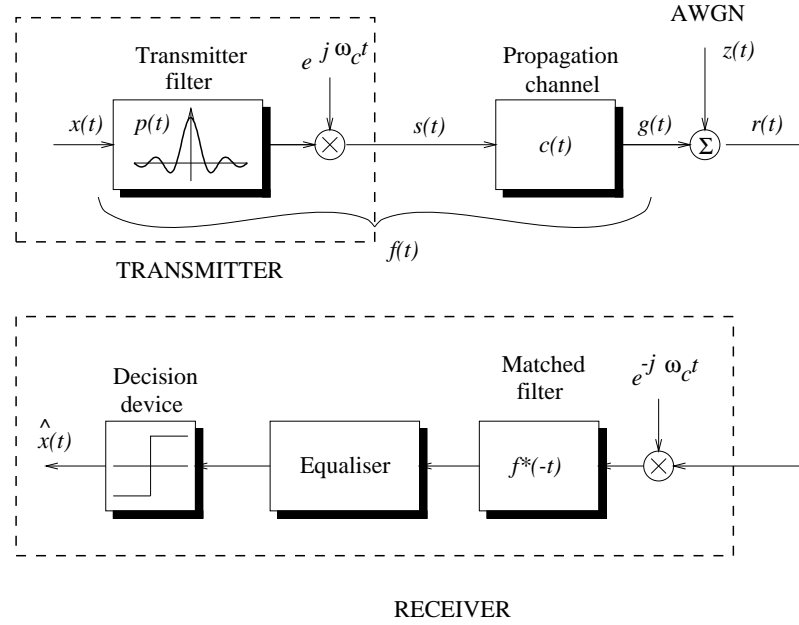


Figure 2.1: Continuous-time description of a base-band digital communications system.

The *sinc* function is in fact an especial case of the filter with raised-cosine spectrum of equation 2.1 for a *roll-off factor* $\beta = 0$. This strict constraint in the frequency domain results in an infinitely long filter in the time domain.

However, when the bandwidth requirements are not so tight and in fact the source symbol period is allowed to be higher than half the reciprocal of the channel bandwidth $T > \frac{1}{2W}$, a series of different raised-cosine pulses can be chosen for different values of β . In this case, the signal spectrum will consist of overlapping replicas of $P(f)$ as shown in Figure 2.3. It follows that the larger the roll-off factor β , the longer the tail of the frequency spectrum is, and as a result the pulse shaping filter has a faster convergence in the time domain.

The raised-cosine pulse filter of equation 2.1 is designed so that it has zero-crossings at the symbol intervals and no interference is produced at adjacent symbols. On the other hand, the amplitude and delay distortion produced by the propagation channel results in pulses with zero-crossings displaced from the symbol intervals. These pulses, as a result, overlap and give rise to ISI. Hence, the inter-symbol interference will be enhanced as the constraint on the bandwidth available is more strict. The *sinc* function for instance will lead to inter-symbol interference in an infinite number of adjacent samples caused by the infinitely long impulse response.

The discrete-time output of the transmitter filter can be expressed as:

$$s(k) = \sum_{n=-\infty}^{\infty} x(n)p(k - nT), \quad (2.3)$$

where $p(k)$ represents the transmitter filter impulse response and $x(n)$ is the sequence of generally complex discrete samples of the modulated source binary sequence. Similarly, the effect

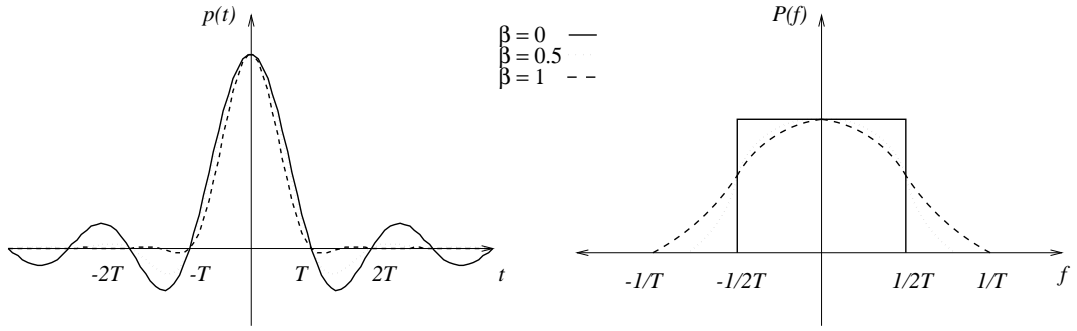


Figure 2.2: Pulses with raised cosine spectrum

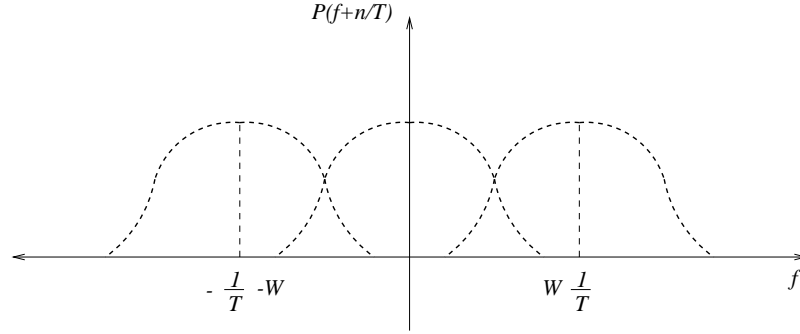


Figure 2.3: Overlapping raised cosine spectrum pulses for $0 < \beta < 1$ and $-\infty < n < \infty$.

that the propagation channel impairments and the additive Gaussian noise produce in the sequence of the pulse-shaped signal can be expressed as:

$$r(k) = c(k) * s(k) + z(k) = g(k) + z(k) = \sum_{n=-\infty}^{\infty} x(n)f(k - nT) + z(k), \quad (2.4)$$

where $c(k)$ is the propagation channel discrete-time impulse response, $z(k)$ is the additive white Gaussian noise process present in the received signal and $f(k) = p(k) * c(k)$ is the overall channel impulse response obtained through the convolution of the transmitter filter and the propagation channel. The input signal is applied to the modulator and transmitter, which converts it into an analog waveform and modulates it onto an appropriate frequency for transmission. As the signal propagates to the receiver, it is delayed, attenuated, and sometimes frequency-distorted. All these effects are represented by the *propagation channel* block. The receiver, which consists of a demodulator and an equaliser, receives the distorted signal corrupted by additive noise and removes both the ISI produced by the propagation channel and the noise in order to recover the transmitted sequence of symbols. The received analog signal is driven through a matched filter and then sampled normally at the symbol period T to produce the matched filter output:

$$y(n) \triangleq y(nT) = \sum_{k=-\infty}^{\infty} r(k)f^*(n - kT)$$

$$= \sum_{k=-\infty}^{\infty} g(k)f^*(n-kT) + \sum_{k=-\infty}^{\infty} z(k)f^*(n-kT) = u(n) + v(n). \quad (2.5)$$

The matched filter, if it is matched to the overall channel impulse response $f^*(-n)$, optimises the output signal-to-noise ratio (SNR) [71] at the sampling instant $t = T$:

$$SNR_{out} = \frac{u^2(T)}{E[v^2(T)]}, \quad (2.6)$$

where $E[v^2(T)]$ simply denotes the variance of the additive coloured Gaussian noise process. The transmitter filter $p(k)$, propagation channel impulse response $c(k)$ and the matched filter at the receiver $f^*(-k)$ are generally represented by an overall transversal filter $h(n)$. Equation 2.5 can be rewritten as:

$$y(n) = \sum_{k=-\infty}^{\infty} x(k)h(n-kT) + v(n), \quad (2.7)$$

where

$$h(n) = \sum_{k=-\infty}^{\infty} f(k)f^*(n+kT). \quad (2.8)$$

It is important to note that the AWGN process $z(k)$, which has a two-sided power spectral density of $\frac{1}{2}N_0$, is no longer white at the output of the matched filter $\{v(n)\}$. In fact, because the matched filter has a duration of more than one symbol period, the consecutive samples of the noise at the output of the matched filter are correlated. In order to cope with this, and because it is much easier to deal with white noise in the equaliser, the equaliser is preceded by a noise whitening filter.

2.3 Supervised versus Unsupervised Equalisation

The removal of the distortion present in the transmitted signal has conventionally been carried out by transmitting a *training* or *pilot* signal periodically during the transmission of information data. A replica of the training signal is available in the receiver so that the coefficients of the equaliser can be updated according to some adaptive filtering algorithm. This kind of algorithms are known as *supervised* methods because during the training period, knowledge of the transmitted sequence is required. Once the training of the equaliser has been completed it can be switched to a *decision-directed* mode of operation in which the equaliser tries to learn from its own decisions [47]. The structure of this receiver using a supervised adaptive equalisation algorithm is shown in Figure 2.4, where an error signal is produced as the difference between the current decisions made by the equaliser $\hat{d}(n)$ and a *desired signal* $d(n)$:

$$e(n) = d(n) - \hat{d}(n). \quad (2.9)$$

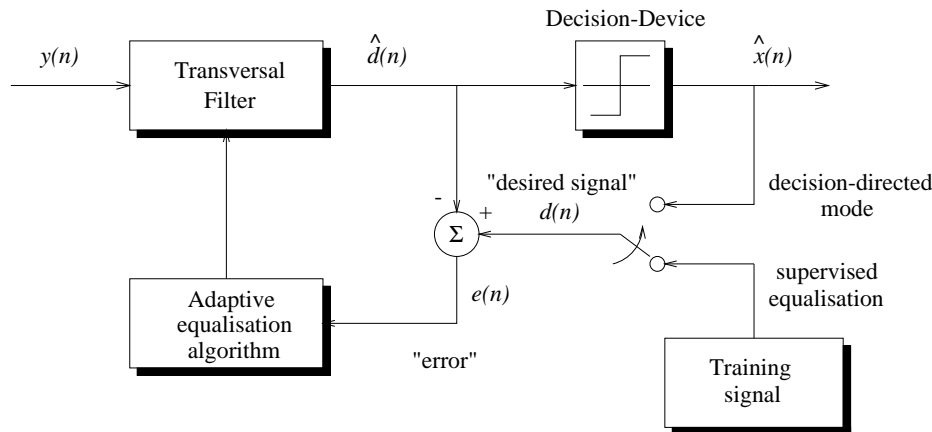


Figure 2.4: Receiver structure of a conventional supervised adaptive equaliser.

During the training period, the desired signal is a replica of the transmitted data stored in the receiver whereas in the decision-directed mode the equaliser's own decisions are used to adjust the coefficients of the transversal filter. When these decisions are good enough the equaliser will tend to reduce the error signal, but if the decisions are not correct the equaliser will make undesired *corrections* to the equaliser weights and the decisions that are made will be worse every time. In decision directed mode, the initial values of the weights are fed to the equaliser by means of a switch, so that the equaliser can converge to a desired global minimum.

However, when the nature of the communication system makes an equalisation procedure based on a training signal impractical, the equaliser needs some means of *unsupervised*, *self-recovery* method to adjust its coefficients. This type of algorithms are referred to as *blind equalisation algorithms* in the literature. The blind equalisation algorithms were conceived as an alternative to conventional adaptive equalisation methods where the synchronisation or re-tuning of the equaliser could be achieved without a training signal. Once the equaliser achieved an acceptable performance in terms of a low enough level of errors produced, it could be switched to a decision-directed mode. However, there are certain scenarios where, due to severe distortion from the propagation channel or simply because the characteristics of the propagation channel are changing rapidly, a constant re-tuning is required. In such cases blind equalisation algorithms can be made adaptive so that they are in operation throughout the reception.

The loss in efficiency due to sending a training signal periodically interposed between the data is normally high. For instance, the GSM standard (Global System for Mobile) allows 26 bits out of 156 bits of a time slot for training. However, the issue of how quick a blind equaliser is able to converge to a global minimum and *open the eye* of the equaliser² is on the other hand a matter of great importance. In a rapidly fading environment it is unrealistic to use a blind equaliser which will take a few thousand symbols to converge. The quality of service that a blind equaliser is able to provide is marked by its convergence speed, i.e. the number of

²An equaliser is said to *open the eye* when it is able to remove distortion from the received signal and the errors produced by the decision function are below a certain threshold [47]

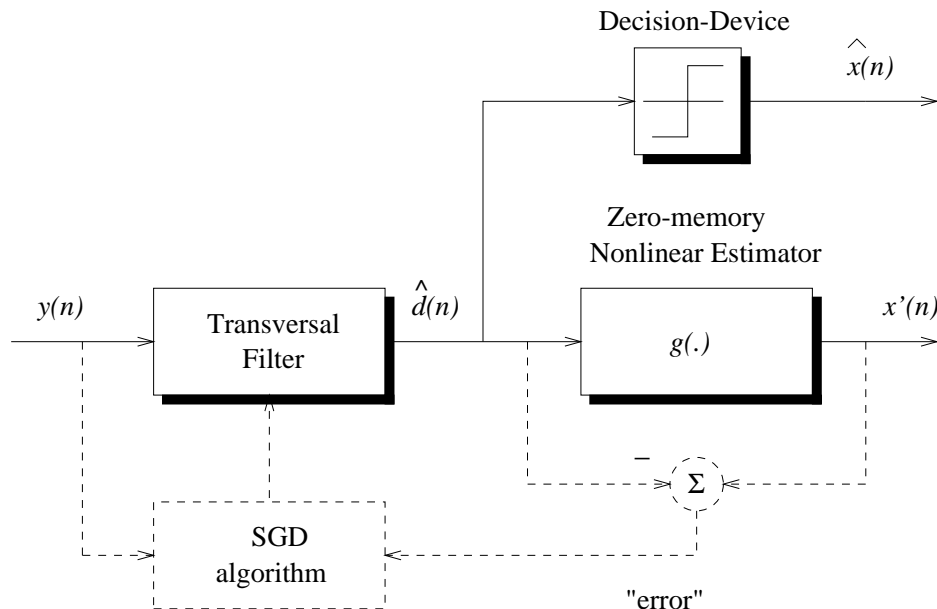


Figure 2.5: Block diagram of a Bussgang blind equaliser.

received samples that they need to provide good enough estimates of the channel characteristics. In mobile radio communication systems such as GSM, the information transmission rate could be increased if the bits allocated for training are re-allocated for instance, as information data bits. This is possible if blind equalisation is used. However, there is always the trade off between the increase in the information transmission rate (which results in the improvement in spectral efficiency) and the degree of performance degradation from using blind equalisation instead of a supervised equalisation method.

2.4 A Brief History of Blind Equalisation Techniques

Blind equalisation algorithms first appeared in the literature under the name of self-recovering equalisation algorithms after the pioneering work of Yoiki Sato [77]. This technique was presented for the case of a *frequency-division multiplexed* (FDM) digital communication network. It led to the development of new techniques because it was shown that a training or pilot signal was not required to equalise multilevel amplitude-modulated signals. The algorithm was simple because it was a modification of the conventional mean-square adaptive equalisation algorithm [47] but it suffered from a very slow convergence.

In the early years of blind equalisation the algorithms that appeared in the literature followed the same idea of Sato's technique [77] [9] [41] [8] [69]: the minimisation of an ensemble-averaged quadratic cost function. These methods were based on the theory that a stochastic process $\tilde{d}(n)$ is a Bussgang process, if the autocorrelation of the process is equal to the cross-correlation between the process and the output of a zero-memory nonlinear estimator shown in Figure 2.5.

$$E\{\tilde{d}(n)\tilde{d}(n-k)\} = E\{\tilde{d}(n)g[\tilde{d}(n-k)]\}. \quad (2.10)$$

These techniques were adaptive in the sense that the performance index was reached through a gradient-type algorithm. Some of the drawbacks of the Bussgang algorithms are: slow convergence, often converged to local undesired minimum rather than to a global minimum etc... [26]. The fact that many communications systems are nonminimum phase, highlighted the need to preserve the phase information of the channel. Nonminimum phase channels have zeros outside and/or inside the unit circle and these systems can only be identified if the magnitude and phase information is preserved. In that case, statistics higher than second-order (autocorrelation and power spectrum) are required [66].

The late 80's and especially the early 90's were dominated by higher-order statistics (HOS) methods [67] [37] [16] [38] [46] [3] [84] [94] [95] [50]. The advantage of HOS methods is that the phase information of the channel can be preserved and therefore, nonminimum phase channels can be identified. HOS was widely used in seismic data deconvolution and nonlinearity detection [66] but they suffered from very slow convergence and some constraints imposed on the input data. In fact, because higher-order cumulants are blind to all kind of Gaussian processes, they require the input signal to be non-Gaussian. The impractical amount of data necessary in the estimation of the cumulants ruled out HOS methods almost entirely for equalisation in a mobile radio environment. However, an eigenvalue identification method based on HOS has recently been proposed for mobile communications with a very fast convergence and a small degradation with respect to conventional supervised methods [11].

Blind equalisation has an inherent ambiguity associated with the phase of the input signal. If the probability distribution of the input signal is symmetric, $\{s_k\}$ and $\{s_k e^{j\theta}\}$ have identical autocorrelations. Such an ambiguity cannot be eliminated without further information of the source. A common practice is to employ differential encoding of the information sequence. Another procedure is to set the first tap-weight of the equaliser to a certain sign which requires knowledge of the sign of a single source symbol [9]. In some algorithms the phase and gain also need to be adjusted before threshold decisions are made, and this can be achieved using an automatic gain tracking algorithm [45] [44].

2.4.1 Bussgang algorithms

The block diagram of a Bussgang blind equaliser is depicted in Figure 2.5. The equaliser makes decisions based on the output of a transversal filter $\hat{d}(n)$, whose weights are updated by a *stochastic gradient descent* algorithm. This algorithm minimises some special (non-MSE) mean cost function that does not require the knowledge of the original input. The mean cost function is defined by:

$$\begin{aligned}
J(n) &= E \{ e^2(n) \} \\
&= E \left\{ (x'(n) - \hat{d}(n))^2 \right\} \\
&= E \left\{ [g[\hat{d}(n)] - \hat{d}(n)]^2 \right\},
\end{aligned} \tag{2.11}$$

where $J(n)$ is a real-valued scalar cost function of $\hat{d}(n)$. The mean cost function will converge to a global minimum if the decisions that are being made are good enough. However, the fact that the cost function is dependent on its own decisions can make the equaliser converge to a local minimum. Figure 2.6 shows an example of the mean cost function of one such Bussgang algorithm presented by Godard [41]. The cost function $J(n)$ is non-convex with two global minima and two local minima and it is not guaranteed that the algorithm will converge to a global minimum. The ill-convergence of Godard algorithms was analysed by Ding in [26]. When the algorithm converges to one of the global minima, the sequence $\hat{d}(n)$ assumes Bussgang statistics for a sufficiently large number of received samples $\hat{d}(n)$. The family of gradient-descent algorithms are known as Bussgang algorithms because of this property [77] [9] [41] [8] [69].

The decision directed algorithm [47] can be considered as an especial case of the Bussgang algorithm, where the zero-memory nonlinear estimator is substituted by a decision function which makes decisions in favour of one or another value of the alphabet of the input signal $x(n)$. The Bussgang algorithms, however, cannot initialise the tap-weights of the equaliser using a training signal and they have to rely on an initial global convergence. The sufficient and necessary conditions for convergence of the Bussgang algorithm were summarised by Benveniste *et al.* [9] for an infinitely large equaliser.

The structure of some of the Bussgang algorithms is described in Figure 2.5. The particularity of each of them is the type of zero-memory nonlinear estimator employed. For instance, the self-recovery algorithm proposed by Sato [77] for M -ary PAM signals uses the following nonlinear estimator:

$$g[\hat{d}(n)] = \gamma \text{sgn}[\hat{d}(n)] \quad \gamma = \frac{E\{x^2(n)\}}{E\{|x(n)|\}} \quad \text{and } \text{sgn}[\hat{d}(n)] \begin{cases} 1 & \hat{d}(n) > 0 \\ -1 & \hat{d}(n) < 0 \end{cases} \tag{2.12}$$

It should be noted that the decision-directed LMS algorithm is an especial case of the Sato algorithm for $\gamma = 1$.

Godard [41] proposed a family of *constant modulus* blind equalisation algorithms (CMA) for two-dimensional digital communication systems. The Godard algorithm minimises the following cost function:

$$J(n) = E \left\{ \left(R_p - |\hat{d}(n)|^p \right)^2 \right\}, \tag{2.13}$$

where p is a positive integer and R_p is defined as:

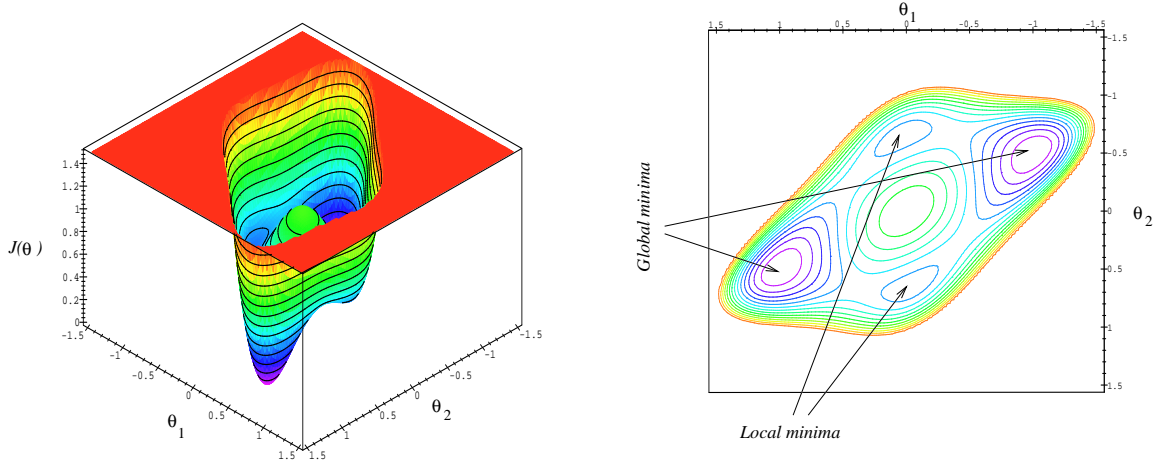


Figure 2.6: Cost function $J(\theta)$ of Godard Constant Modulus Algorithm for an ARMA channel $h(z) = 1/(1+0.5z^{-1})$ and a binary input data. The two global minima are situated at $(\theta_1, \theta_2) = (1, 0.5)$ and $(\theta_1, \theta_2) = (-1, -0.5)$.

$$R_p = \frac{E\{|x(n)|^{2p}\}}{E\{|x(n)|^p\}}. \quad (2.14)$$

The Godard algorithm is designed to penalise deviations of the blind equaliser output $x(n)$ from a constant modulus. The constant R_p is chosen in such a way that the gradient of the cost-function $J(\theta)$ is zero when there is perfect equalisation.

The Godard algorithm includes the following cases:

(i). $p=1$:

$$J(n) = E\{(R_1 - |y(n)|)^2\} \quad R_1 = \frac{E\{|x(n)|^2\}}{E\{|x(n)|\}}. \quad (2.15)$$

This case is the Sato algorithm.

(ii). $p=2$:

$$J(n) = E\{(R_2 - |y(n)|^2)^2\} \quad R_2 = \frac{E\{|x(n)|^4\}}{E\{|x(n)|^2\}}. \quad (2.16)$$

This case is known as the *constant modulus algorithm* (CMA) and it is probably the most widely used blind equalisation algorithm in practice.

2.4.2 Higher-order statistics

During the late 80's and the early 90's, Higher-order statistics methods enjoyed considerable attention in the fields of linear and nonlinear signal processing, where non-Gaussianity, non-

minimum phase systems and coloured noise were encountered [59]. Higher-order statistics are described in terms of *cumulants* and its Fourier transform the *polyspectra*. A feature which differs from the autocorrelation is that cumulants are blind to all kinds of Gaussian processes; in other words, the cumulants of Gaussian processes are zero, making cumulant-based methods suitable when measuring non-Gaussian signals corrupted by Gaussian noise. On the other hand, while the second-order statistics of a process are blind to phase information, it is shown in [67] that higher-order statistics preserve phase information and therefore, they are able to deal with nonminimum phase processes.

Moments and Cumulants

Let us consider a zero-mean stationary stochastic process $u(n)$

$$E[u(n)] = 0 \quad \forall n \quad (2.17)$$

and let $u(n), u(n + \tau_1), \dots, u(n + \tau_{k-1})$ denote the random variables obtained by observing this stochastic process at times $n, n + \tau_1, \dots, n + \tau_{k-1}$ respectively. The k th order cumulant of the process $u(n)$ is thus defined in terms of its joint moments of order up to k . Specifically, the second-, third- and fourth- order cumulants are given respectively by:

$$c_2(\tau) = E[u(n)u(n + \tau)] \quad (2.18)$$

$$c_3(\tau_1, \tau_2) = E[u(n)u(n + \tau_1)u(n + \tau_2)] \quad (2.19)$$

and

$$\begin{aligned} c_4(\tau_1, \tau_2, \tau_3) = & E[u(n)u(n + \tau_1)u(n + \tau_2)u(n + \tau_3)] \\ & - E[u(n)u(n + \tau_1)]E[u(n + \tau_2)u(n + \tau_3)] \\ & - E[u(n)u(n + \tau_2)]E[u(n + \tau_3)u(n + \tau_1)] \\ & - E[u(n)u(n + \tau_3)]E[u(n + \tau_1)u(n + \tau_2)] \end{aligned} \quad (2.20)$$

From the definitions given in 2.18, 2.19 and 2.20 it can be established that:

- *i)* The second-order cumulant $c_2(\tau)$ is the same as the autocorrelation function $r(\tau)$.
- *ii)* The third-order cumulant $c_3(\tau_1, \tau_2)$ is the same as the third-order moment $E[u(n)u(n + \tau_1)u(n + \tau_2)]$.
- *iii)* The fourth-order cumulant $c_4(\tau_1, \tau_2, \tau_3)$ is different from the fourth-order moment $E[u(n)u(n + \tau_1)u(n + \tau_2)u(n + \tau_3)]$. In order to generate the fourth-order cumulant, we need to know the fourth-order moment and six different values of the autocorrelation function.

With regard to moving-average (MA) channel identification, different solutions are available. Some of the techniques are formulated as the combination of autocorrelations and third-order cumulants [3]. In the presence of additive Gaussian noise, theoretically, the third-order cumulants of the noise process are zero but second-order cumulants still contain the effect of the noise. By using only third- (or fourth-) order cumulants based method this problem can be overcome. Some methods use only third-order cumulant slices for the MA parameter estimation [63] [103] but they only use two cumulant slices and therefore ignore the information contained in the unused slices. However, since third-order cumulants are identically zero for non-skewed processes, and in fact most communication signals are non-skewed, fourth-order cumulants should be used. Some methods using fourth-order cumulants also require the use of second-order statistics and the algorithms are affected by the influence of additive Gaussian noise.

2.4.3 Cyclostationary algorithms

The problem of blind channel identification using second-order statistics of the *oversampled* channel output has recently drawn considerable attention, following the pioneering work by W.A. Gardner [34] and L.Tong, G.Xu and T.Kailath [87]. In this last decade, many of the inherent problems of blind equalisation have found an extra degree of freedom from oversampling. The oversampling operation is equivalent to sampling the received signal at a rate higher than the baud rate of the source symbols. The oversampling period T_s is usually an integer fraction P of the baud-rate and thus, the oversampling operation is also known as fractionally spaced sampling. The zero-mean discrete-time signal $x(n)$ sampled at a rate higher than the symbol rate is a sequence of discrete pulses padded with zeros $x(nT_s) = [x(0) 0 0 \cdots 0 x(T_s) 0 0 \cdots 0 x(2T_s) 0 \cdots]$ and the autocorrelation of this sequence is periodic [36]:

$$r(mT_s, nT_s) = E \{x(mT_s)x^*(nT_s)\} \begin{cases} \neq 0 & \text{for } m - n = kP, k \in \mathcal{Z}, m \bmod P = 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.21)$$

Thus, a stationary process $x(n)$ can become a *wide-sense cyclostationary* process, if the mean and autocorrelation become periodic by oversampling. Then, a whole new concept of cyclostationary statistics apply to these processes. The cyclic autocorrelation function is defined as the Fourier series coefficients of the periodic autocorrelation function and its Fourier transform, the cyclic spectrum, is defined as the spectral correlation function of two frequency shifted versions of $x(n)$ ³ [36]. The cyclic second-order statistics (cyclic autocorrelation and cyclic spectrum), unlike conventional second-order statistics, are able to preserve the phase information of a *linear time-invariant* (LTI) filter. It is the lack of conjugate symmetry in the cyclic domain

³An introduction to the principles of cyclostationarity and the definition of the cyclic autocorrelation is presented in the next chapter

that allows the preservation of phase information. The power spectrum of the output of the LTI filter is defined as [47]:

$$S_y(z) = H(z)S_x(z)H^*(z) = |H(z)|^2S_x(z) \quad z = e^{j2\pi f}, \quad (2.22)$$

which is phase blind. Similarly, the cyclic spectrum allows [36]

$$S_y^\alpha(z) = H(z e^{j\pi\alpha})S_x^\alpha(z)H^*(z e^{-j\pi\alpha}) \quad \alpha = k/T, \quad k \in \mathcal{Z}. \quad (2.23)$$

Figure 2.7 shows a maximum phase and minimum phase channel with transfer functions:

$$H_1(z) = 1 - az^{-1} \quad \text{and} \quad H_2(z) = 1 - a^*z \quad |a| < 1. \quad (2.24)$$

From equation 2.22, and assuming that the zero-mean input signal has a constant power spectrum, the power spectrum of the two output signals $y_1(n)$ and $y_2(n)$ are defined as:

$$\begin{aligned} S_{y_1}(z) &= \sigma_x^2(1 - az^{-1})(1 - a^*z) \\ S_{y_2}(z) &= \sigma_x^2(1 - a^*z)(1 - az^{-1}), \end{aligned} \quad (2.25)$$

which are identical. This shows that the power spectrum can only account for the magnitude of the LTI channel transfer function $H(z)$. On the other hand, following 2.23, and assuming that the input signal has a finite and constant cyclic spectrum at a particular frequency α , the cyclic spectrum of the two output signals $y_1(n)$ and $y_2(n)$ are defined as:

$$\begin{aligned} S_{y_1}^\alpha(z) &= q_x(1 - ae^{-j\pi\alpha}z^{-1})(1 - a^*e^{-j\pi\alpha}z) \\ S_{y_2}^\alpha(z) &= q_x(1 - a^*e^{j\pi\alpha}z)(1 - ae^{j\pi\alpha}z^{-1}). \end{aligned} \quad (2.26)$$

The cyclic spectrum can therefore distinguish between two channels equal in magnitude but different in phase. In the presence of a nonminimum phase system, i.e. with zeros inside and outside the unit circle, and assuming that the input signal is stationary, if the received signal is sampled at the symbol period T , HOS will be required to identify the correct phase response of the channel. If, on the other hand, the received signal is sampled at a rate higher than the symbol rate, the received signal becomes wide-sense cyclostationary and phase information can also be preserved. Since the estimation of second-order statistics requires less data compared to the estimation of higher-order statistics, faster algorithms can be implemented by exploiting the cyclostationarity of the channel output. Moreover, cyclostationary statistics do not single out Gaussian or near-Gaussian input processes.

Most blind equalisation schemes to date have sampled the channel output at the baud-rate to produce a stationary channel output, so as to extract useful statistical information that will

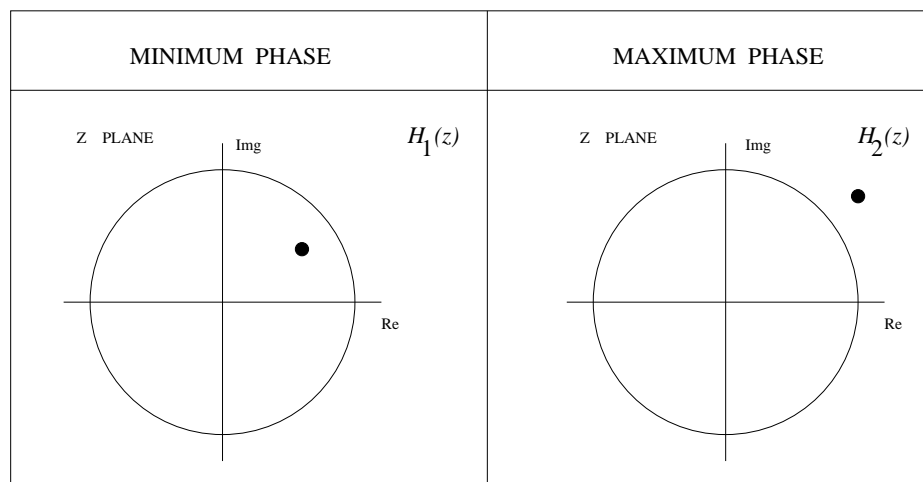


Figure 2.7: Minimum phase and maximum phase channels

be used to identify the channel. In this case, higher-order statistics need to be applied to the stationary output process. However, Gardner [34] investigated the use of the second-order cyclic spectrum in channel identification but proposed a not truly blind system identification procedure, since it was based on sending a training signal, either periodically during message transmission or superposed on top of the message signal. However, his work can be considered as the pioneer in the family of blind channel identification algorithms exploiting second-order cyclostationary statistics, either in the time domain or in the frequency domain.

Based on a certain rank condition of the convolution matrix of the received signal sampled at a rate higher than the baud-rate, Tong, Xu and Kailath [87] [88] earlier proposed an effective channel identification procedure, giving way to the investigation of array processing techniques in blind equalisation.

Motivated by [34], other algorithms that exploit second-order cyclic statistics have been proposed. Dimitrios Hatzinakos [45] proposed a time-domain nonminimum phase channel identification algorithm based on cyclic autocorrelation estimates of the oversampled received signal. This method, although computationally complex, is efficient since it can identify the minimum phase and maximum phase characteristics of the system using the complex cepstrum of the cyclic autocorrelation. Shankar Prakriya and Dimitrios Hatzinakos [70] proposed some methods for identification of MA and ARMA linear models using only the fractionally-spaced output sequence. The method requires only the cyclic autocorrelation to be computed. However, the model orders were assumed to be known. Luis. A. Baccala and Sumit Roy [6] proposed a time-domain approach based on oversampling the channel output and employing a discrete cyclic autocorrelation function. Zhi Ding and Ye Li also proposed a parametric and a non-parametric approach for ARMA system identification [54] based in earlier work presented in [36].

Tong, Xu and Kailath also presented another approach [90] as an extension of the method proposed in [87] for antenna arrays. The beauty of these two approaches is that they can be

combined giving way to a new approach that exploits both oversampling and spatial diversity. Following this, another technique of the same family was presented by M. Moulines, P. Duhamel, J. Cardoso and S. Mayrargue [60] [61], based on the orthogonality between the *signal* and *noise* subspaces. Dirk T.M. Slock [81] also proposed another subspace method. This approach is able to identify the channel from second-order statistics of the received signal by means of linear-prediction in the noise-free case and using the Pisarenko method in the presence of additive noise. Also, Hui Liu and Guanghan Xu [56] showed that blind symbol estimation can be achieved based on some rich structure information of the channel matrix that can be used to identify the symbol subspace from the received signal subspace.

Overall, blind equalisation algorithms using fractionally-spaced sampling can be classified in terms of cyclic statistics based algorithms [34] [45] [6], algebraic methods [88] [61] [81] [100] or maximum-likelihood methods [51] [48]. The large amount of work that has appeared in the literature (15 related papers in the February 1997 issue of IEEE Transactions on Signal Processing followed by tutorial papers) indicate the *good health* of these techniques. The performance analysis of some of the most interesting approaches is one of the contributions of this thesis with especial focus on the application of these methods in mobile radio environments. Other comparative results are also available in [57] [73] [43] [101]. Connections between cyclic statistics based methods and algebraic methods are presented in [57] and between maximum-likelihood methods and cyclic statistics in [51].

2.5 Equaliser Structures for Blind Equalisation

The transmission of digital data through a band-limited channel is often distorted by interference caused by the propagation channel which causes ISI. Over many practical channels, the most important signal distortion is caused partly by the band-limiting of the signal in transmission and partly by the transmission path. Infinitely narrow pulses would certainly eliminate ISI but that would require an infinitely wide bandwidth. The aim of the equaliser is to reduce the ISI at the output of the receiver filter as far as possible; and at the same time, minimise the *bit error rate* (BER).

Detection processes for distorted digital signals may be classified into two separate groups. In the first group, the input signal is coded and the decision process itself is modified to take account of the signal distortion that has been introduced by the channel, and no attempt needs to be made to reduce the signal distortion prior to the decision process. Although no equaliser is required, the decision process may be considerably more complex than that when an equaliser is used. In the second scheme, the received sampled digital signal is fed through an equaliser that corrects the distortion introduced by the channel and restores the received signal, neglecting for the moment the effects of noise. The equaliser acts as the inverse of the channel, so that the channel and equaliser together introduce no signal distortion. Each data symbol is detected as it arrives, independently of the others, by comparing the corresponding sample value with the

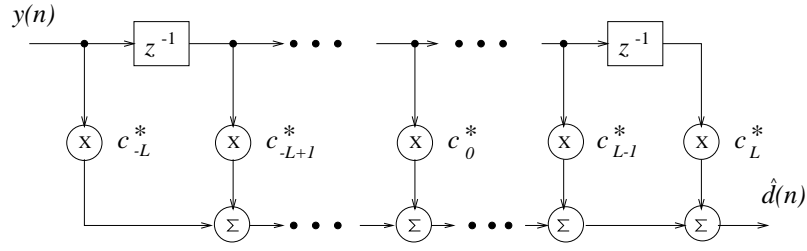


Figure 2.8: Transversal Filter.

appropriate threshold level (or levels).

Equalisers may be *linear* or *nonlinear* depending on their decision function. In this section two equaliser structures for blind equalisation are described:

- Linear equaliser.
- Nonlinear equaliser: maximum-likelihood sequence estimator (MLSE).

These structures are the most widely used in blind equalisation. There are several parameters to be considered at the time of selecting an appropriate equaliser structure for blind equalisation. The trade-off between computational cost and performance gain is without doubt the key parameter in that selection. However, other issues such as the time-varying nature of the propagation channel, play an important role.

2.5.1 Linear equaliser

The basic structure of a linear equaliser is a transversal filter, which consists of a finite set of tapped-delay lines, as shown in Figure 2.8. These tapped-lines are multiplied by some tap-weights c_n^* [47], where $*$ denotes conjugation and it is used for notational convenience. Then they are added up to produce an output signal $\hat{d}(n)$. Decisions will be made on the output of the equaliser based on certain threshold values, in order to obtain an estimate of the transmitted sequence $\hat{x}(n)$. The tap-spacing of the linear transversal filter is the reciprocal of the symbol rate. The optimum receiver is built as a transversal filter with weights matched to the transmitted pulse filter corrupted by the propagation channel.

The output of the equaliser is given by:

$$\begin{aligned} \hat{d}(n) &= \sum_{k=-\infty}^{\infty} c_k^* y(n-k) \\ &= \sum_{k=-\infty}^{-L-1} c_k^* y(n-k) + \sum_{k=-L}^L c_k^* y(n-k) + \sum_{k=L+1}^{\infty} c_k^* y(n-k) \end{aligned}$$

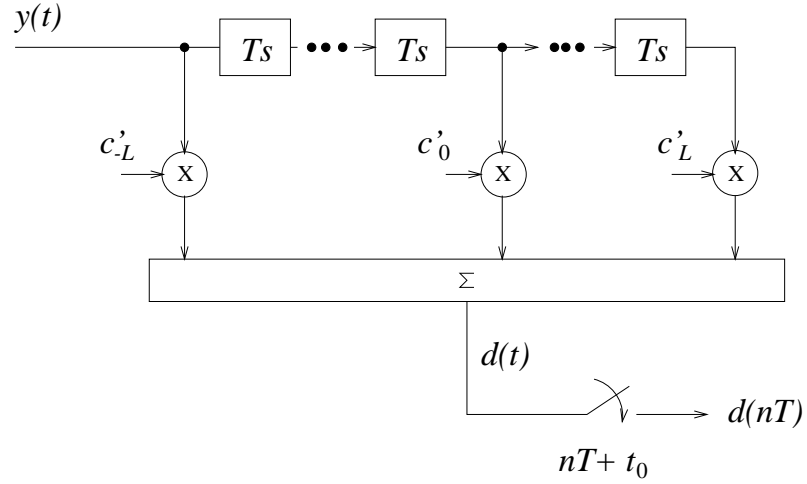


Figure 2.9: Fractionally-spaced equaliser.

$$= \sum_{k=-L}^L c_k^* y(n-k) + v_c(n) \quad n = 1, \dots, N. \quad (2.27)$$

The equalisation process is information lossless in the sense that it provides sufficient information to estimate the source symbols. However, the convolution sum in (2.27) is not infinite for computational purposes and the difference between the optimum convolution sum and the approximation in (2.27) is known as the *convolutional noise*, representing the residual ISI that results from the use of an approximate inverse filter [47].

Fractionally-spaced linear equaliser

The tap spacing of the conventional linear equaliser is the same as the source symbol period T , at which the signal has been sampled. This type of equaliser is known as the *synchronous equaliser* or *symbol-spaced equaliser* (SSE). The advantage of the SSE is that the equaliser works with the same clock-symbol period T . However, the synchronous equaliser cannot be considered optimum in the sense that it suffers from the following drawbacks:

Sensitivity to initial timing or phase errors: As soon as a timing error occurs in the received signal, this results in the loss of unrecoverable information. One particular type of equaliser which overcomes in part this sensitivity to timing is the *fractionally-spaced equaliser* (FSE) [96] [49]. The tap-spacing T_s of the FSE, as shown in Figure 2.9, is generally an integer factor of the symbol period T . A typical and most popular value is $T_s = T/2$. Although the input at the equaliser is sampled at T_s , the equaliser output is still sampled at T , since the data decisions are made at the symbol interval. The FSE is less sensitive than the SSE in terms of timing errors, in the sense that the performance of the SSE can change substantially depending on the timing mismatch between the transmitter and the receiver [74] [40]. This is

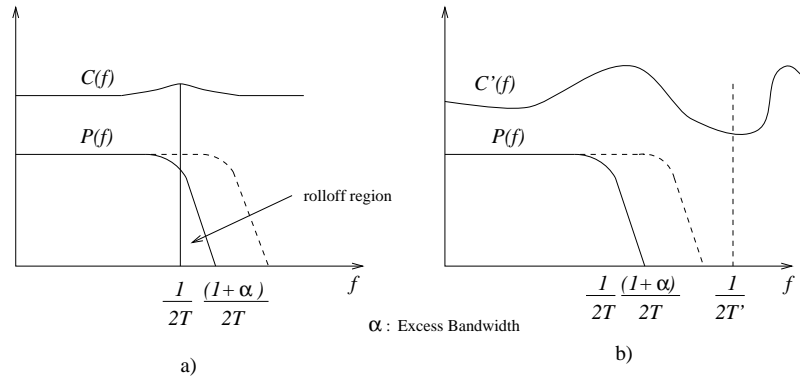


Figure 2.10: a) Typical spectrum for a synchronous equaliser. $P(f)$ is the spectrum of the transmitted pulse and $C(f)$ is the Fourier transform of the tap coefficients. b) Typical spectrum for a fractionally-spaced equaliser with tap spacing $T_s < T/(1+\alpha)$.

due to the fact that the channel observed at one particular time instant differs from the channel observed at another time instant. The FSE, on the other hand, does not depend so much on the selection of the sampling phase because of the multichannel structure of the system. The observed channel is represented by multiple subchannels and the FSE is a combination of the SSE of each of the subchannels [92]. The resulting equalised sequence is the summation of the equalised sequences of the individual subchannel equalisers.

The sensitivity to timing error has an interesting connotation in the frequency domain. The presence of a delay of t_0 is interpreted as a multiplying complex exponential term $e^{-j2\pi ft_0}$ in the frequency domain. The spectrum of the received signal $y(n)$ is expressed as:

$$y_T(f) = \frac{1}{T} \sum_{n=-\infty}^{\infty} S\left(f - \frac{n}{T}\right) e^{j2\pi\left(f - \frac{n}{T}\right)t_0}. \quad (2.28)$$

The timing error t_0 , shown in Figure 2.10, causes the effect of spreading or widening the frequency range of the signal $S(f)$, which is obtained after the convolution of the modulated signal and the transmitter pulse shaping filter $P(f)$. This timing error breaks the Nyquist criterion which fixes the symbol rate to the channel bandwidth, making the sampling rate at least twice as fast as the available channel bandwidth. Due to the frequency shift produced by the timing error, an increase in the sampling rate is required in order to avoid aliasing. Thus, a significant property of the FSE is that it can compensate for *any* delay distortion *without* noise enhancement.

The synchronous equaliser cannot compensate for spectral-nulls caused by certain phase components of the input signal over the aliasing region: For a channel with bandwidth $W > \frac{1}{2T}$ we have seen that the signal can be designed as a raised-cosine pulse in order to satisfy this condition. Such condition results in spectrally overlapping replicas of $X(f)$, as shown in Figure 2.3. Under certain phase conditions of the input signal, the overlapping

roll-off regions can be nulled resulting in a spectral null. In order to account for the noise in the region of the spectral null, the equaliser will tend to have an infinite magnitude component. The FSE prevents this problem, because it accounts for the entire bandwidth of the signal including the roll-off region and avoids spectrally-overlapping roll-off regions. Therefore, the minimum sampling rate before the equaliser needs to be at least $T_s < T/(1 + \beta)$. For example, for a roll-off factor $\beta = 0.5$, it is found that at least a $\frac{2}{3}T$ -spaced equaliser will be required.

In summary, the advantage of the FSE over the SSE is that it can compensate more efficiently for amplitude and phase distortion. It is also interesting to note that the FSE does not require a matched filter preceding it because it can compensate for noise components outside the Nyquist band $\frac{1}{2T}$. In terms of digital implementation the SSE has an obvious advantage because the equaliser will work at the symbol-clock period T . For a $\frac{2}{3}T$ -spaced equaliser, for instance, a three-times faster A/D conversion rate will be required, three-times more memory cells will be needed in the equaliser delay line and $\frac{3}{2}$ more multiplications than the SSE over the delay length of the equaliser. Because of the excess bandwidth present in the system, in order to account for the noise in the entire bandwidth, the sampling rate has to be increased. However, increasing the sampling rate generally means more computational complexity and in general, the tendency is not to sample the received signal at more than twice the symbol rate.

In the case of fractionally-spaced equalisers, the most usual $T/2$ oversampling factor is selected as a compromise between the performance improvement in terms of timing-error recovery and the increase in the number of computations. The length of the FSE however is not directly proportional to the oversampling factor. In general, typical communications systems use low pass filters in their receivers/transmitters, and oversampling, generally means that, the over-sampled discrete channel model shows more zeros, either on the unit circle or very close to the unit circle. In this case a longer equaliser is required to remove ISI. The procedure to estimate the correct delay is the same as the SSE. Experiments were carried out to show the effect of the selection of different equaliser lengths for FSE and were reported by Treichler *et al.* [92], where the delay of the equaliser was preselected.

Linear equalisation algorithms

The selection of the appropriate tap-weights of a linear equaliser can be done according to some linear equalisation algorithm. The optimum coefficients are generally chosen, either to minimise the *mean-square error* (MSE) between the original transmitted sequence and the *equalised* sequence (least-squares criterion) or to maximise the output *signal-to-noise ratio* (Zero-forcing criterion).

Method of least-squares The method of least-squares (LS) is based on the minimisation of the quadratic estimation error $e_{min}(n)$ cost function which is defined as:

$$J_w = \sum_{n=L}^N e(n)e^*(n), \quad (2.29)$$

and the error signal as in (2.9) is defined as:

$$e(n) = d(n) - \hat{d}(n) = d(n) - \sum_{k=-L}^L c_k^* y(n-k) \quad n = L, \dots, N. \quad (2.30)$$

The coefficients of the linear equaliser are selected, so that they satisfy the *minimum mean-square error* (MMSE) criterion, in which case the equaliser reduces the effect of ISI and noise jointly. The coefficients of the linear equaliser are given by the Wiener equation [47]:

$$\hat{\underline{c}}_{opt} = \Phi_{yy}^{-1} \underline{z}_{yd}, \quad (2.31)$$

where matrix Φ_{yy} represents the autocorrelation matrix of the input to the equaliser; \underline{z}_{yd} is the cross-correlation vector between the equaliser input and the desired or reference signal and $\hat{\underline{c}}_{opt} = [c_{-L}^*, \dots, c_L^*]^T$ is the vector with the optimum tap-weights of the linear equaliser, where the superscript T denotes transpose. The inverse of matrix Φ_{yy} exists, if and only if, it is non-singular and can be obtained using *singular-value decomposition* (SVD) (see Appendix B). However, if matrix Φ_{yy} is singular or very close to singular, it is common practice to add energy to the main diagonal in order to avoid getting nearly zero eigenvalues [47].

In fractionally-spaced linear equalisation, the minimisation of the MSE leads to the following tap-weight optimisation formula:

$$\hat{\underline{c}}'_{opt} = A^{-1} \underline{\psi}, \quad (2.32)$$

where A is again the covariance matrix of the equaliser input and $\underline{\psi}$ is the vector of the cross-correlations between the input of the equaliser and the reference signal. The difference from the formulation of the SSE is that the diagonal elements of matrix A are periodic [75] and as a result, some of the eigenvalues of matrix A will be nearly zero. Thus, matrix A becomes singular and non-invertible, but one could add energy to the main diagonal, as explained earlier, in order to overcome this problem. The other problem that might arise with fractionally-spaced linear equalisation, is the fact that oversampled noise is correlated and it is known that the performance of the Wiener filter degrades in the presence of correlated noise [47].

In the presence of cyclostationary signals, an alternative formulation to the Wiener filtering technique, called the *cyclic Wiener filtering* technique was proposed [33] [35] and has been shown to outperform the conventional time-invariant Wiener Filter in the presence of a single cyclostationary signal with a single cyclic period $1/\alpha$. For the more general case of multiple cyclostationary signals with a single/or multiple periodicities, the cyclic Wiener filtering technique

gives way to a polyperiodic filtering structure known as *FREquency-SHift (FRESH) filtering*. The FRESH filters have the ability to separate temporally and spectrally overlapping signals by exploiting the spectral coherence inherent in cyclostationary signals. The theory of spectral coherence and spectrally-overlapping signal separation are presented in Chapter 3.

Zero forcing In a linear equaliser with *zero forcing (ZF)* criterion, the equaliser taps are set so that the peak distortion in the equalised signal is minimised, assuming that the equaliser allows infinite number of taps [58]. In practice, in most of the situations, it is the case that the transversal filter spans only over a finite number of taps. This is a sub-optimum solution which not only minimises the peak distortion, but also leads to some simple iterative approaches to hold the equaliser taps correctly adapted to the slow variations of the propagation channel characteristics. However, these simple forms fail in the presence of severe time-variation in the channel. The peak-distortion is minimised by finding the equaliser taps, which after convolution with the impulse response of the channel, produce a single tap with value 1 and the rest are 0:

$$q_n = \sum_{k=-\infty}^{\infty} c_k h(n-k) = \begin{cases} 1 & (n=0) \\ 0 & (n \neq 0), \end{cases} \quad (2.33)$$

where c_k are the equaliser coefficients and $h(n)$ the channel impulse response. The peak distortion is given by

$$Q(\underline{c}) = \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} |q_n| = \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \left| \sum_{k=-\infty}^{\infty} c_k h(n-k) \right|. \quad (2.34)$$

The peak distortion $Q(\underline{c})$ is a variable of the vector of equaliser tap coefficients $\underline{c} = [c_{-L}, \dots, c_L]^T$ and reaches its minimum when all the values of q_n are zero except for $n=0$. Equation 2.33, in the z domain results in

$$Q(z) = C(z)H(z) = 1, \quad (2.35)$$

or

$$C(z) = \frac{1}{H(z)}, \quad (2.36)$$

where $C(z)$ and $H(z)$ denote the z transforms of the equaliser coefficients c_k and the channel coefficients $h(n)$, respectively. This implies that the zero forcing criterion requires the equaliser coefficients to match the inverse of the linear impulse response of the channel. As mentioned above, in practice, the equaliser is normally of finite length, having $2L+1$ taps and the equaliser cannot remove all the inter-symbol interference. Instead of minimising (2.33) with respect to

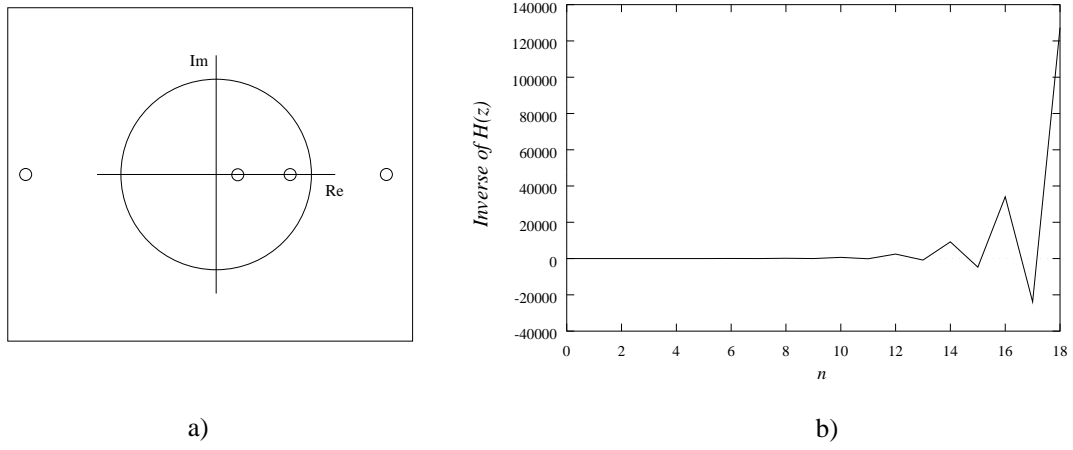


Figure 2.11: a) Zero position of channel with transfer function $H(z) = (1 - 0.2z^{-1})(1 - 0.8z^{-1})(1 - 1.8z^{-1})(1 - 2z^{-1})$; b) unstable inverse filter

c_k using a steepest-descent method, a particular solution for this minimisation is known in the case when the peak distortion at the input to the equaliser is less than unity

$$\mathcal{Q}' = \frac{1}{h(0)} \sum_{n=1}^M |h(n)| < 1. \quad (2.37)$$

This condition means that the ISI present at the input of the equaliser is not severe enough and the eye is open prior to the equaliser. Consequently, the equaliser coefficients are selected so that $\mathcal{Q}(\underline{c})$ is minimised by forcing $q_n = 0$ in the range $1 \leq |n| \leq L$ and q_0 is set to 1. However, the values of q_n outside the range $1 \leq |n| \leq L$ are normally nonzero, and they result in *residual inter-symbol interference* or *convolutional noise*.

$$\mathcal{Q}(\underline{c}) = [\overbrace{0, \dots, 0}^L, 1, \overbrace{0, \dots, 0}^L, q_{L+1}, \dots, q_{L+M-1}]. \quad (2.38)$$

Separate equalisation of the two factors of the channel response The ZF linear equaliser conceived as the inverse of the channel impulse response $h(n)$ has a particular problem when the channel has zeros outside the unit circle. In this case, the inverse converts these zeros in poles outside the unit circle and as a result, the inverse filter is unstable as shown in Figure 2.11. In order to deal with channels with zeros inside and outside the unit circle, both the minimum phase zeros and maximum phase zeros need to be treated separately.

The transfer function of the nonminimum phase channel $H(z)$ consists of a minimum phase part $H_m(z)$ and a maximum phase part $H_M(z)$, which can be represented in terms of a polynomial

$$H_m(z) = 1 + h_m(1)z^{-1} + h_m(2)z^{-2} + \dots + h_m(g)z^{-g}, \quad (2.39)$$

for the minimum phase part with all the zeros inside the unit circle and

$$H_M(z) = h_M(0) + h_M(1)z^{-1} + h_M(2)z^{-2} + \dots + h_M(f)z^{-f}, \quad (2.40)$$

is the maximum phase transfer function with all its zeros outside the unit circle. When the zeros of a minimum phase impulse response are close to the unit circle, a longer equaliser will be required [19]; if the zero is located on the unit circle, the channel cannot be equalised using a finite-length feed-forward linear equaliser. If, on the other hand, the zeros are outside the unit circle, the channel will still be equalisable using a finite delay in the equaliser. In this case, instead of considering a causal impulse response, which is divergent, we might alternatively consider a non-causal impulse response with an appropriate delay, which is convergent, and therefore invertible. The design technique of a ZF linear equaliser in the presence of a nonminimum phase channel is:

- i) Separate zeros inside and outside the unit circle and obtain the minimum and maximum phase impulse responses respectively.
- ii) Obtain the inverse impulse response of the causal minimum phase impulse response using a long division.
- iii) Reverse the order of the coefficients of the maximum phase impulse response to obtain an equivalent convergent impulse response.

$$H'_M(z) = h_M(f) + h_M(f-1)z^{-1} + h_M(f-2)z^{-2} + \dots + h_M(0)z^{-f}. \quad (2.41)$$

- iv) Invert the non-causal impulse response and reverse the order of the coefficients. The length d of the inverse filter is also the delay associated with the inverse of the nonminimum phase filter.
- v) Resolve the inverse filter via convolution of the inverse impulse responses of the minimum and maximum phase parts. The resulting FIR filter is the inverse of the nonminimum phase channel with a delay d associated.

Figure 2.12 shows the procedure to equalise a channel with two zeros inside and two zeros outside the unit circle. Note that the closer the zeros are to the unit circle, the longer the length of the equaliser that is needed to remove the ISI.

2.5.2 Maximum-likelihood sequence estimator

The *maximum-likelihood sequence estimator* (MLSE) is the optimum receiver for the equalisation of channels with memory, if it is preceded by a whitened matched filter [29]. The matched filter $f^*(-t)$ matched to the transmitter filter and the propagation channel impulse response has the objective to account for the noise outside the Nyquist frequency range. However, it is the actual symbol-spaced matched filter which correlates the otherwise white noise process $z(t)$

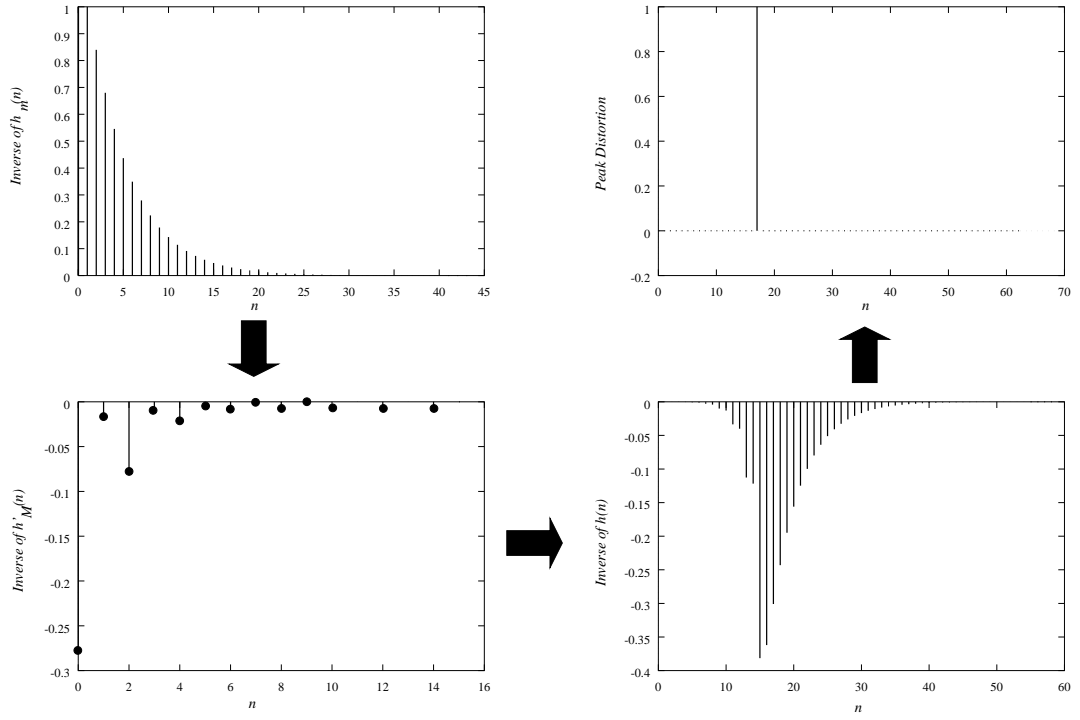


Figure 2.12: Separate equalisation of maximum phase and minimum phase factors of transfer function $H(z) = (1-0.2z^{-1})(1-0.8z^{-1})(1-1.8z^{-1})(1-2z^{-1})$; $H_m(z) = 1-z^{-1} + 0.16z^{-2}$ and $H_M(z) = 1+0.2z^{-1}-3.6z^{-2}$: Inverse impulse response of minimum phase filter (top left), inverse of reversed maximum phase filter (bottom left), inverse of nonminimum phase channel $h(n)$ (bottom right) and peak distortion (top right)

and the performance of the MLSE degrades in the presence of correlated noise [71], because the noise is not identically distributed around the unequalised symbols. In order to account for correlated noise, the matched filter is substituted by a cascade of a matched filter and a whitening filter; the combination of both is known as *whitened matched filter* [29]. The cascade of the transmitter pulse-shaping filter, the propagation channel and the matched filter, denoted by $h(t)$ in section 2.2 is the autocorrelation of the combined transmitter pulse/propagation channel $f(t)$.

The combined channel transfer function $H(z)$, which has N zeros, can be decomposed into two polynomials:

$$H(z) = \Gamma(z)\Gamma^*(z^{-1}). \quad (2.42)$$

The whitening filter which attempts to account for the correlation produced by the matched filter has the form $1/\Gamma^*(z^{-1})$. However, from the 2^N different possibilities, $\Gamma^*(z^{-1})$ is chosen, so that all its zeros lie inside the unit circle and can be inverted without any delay. Thus, the output of the whitened matched filter is expressed as:

$$r_k = \sum_{n=0}^L \gamma_n x_{k-n} + \mu_k, \quad (2.43)$$

where μ_k is a zero-mean Gaussian noise process with variance $\sigma_n^2 = N_0/2$ and γ_n is the inverse z transform of the polynomial $\Gamma(z)$. The conditional probability distribution of a particular received sample r_k for a certain sequence of transmitted input symbols $\{x_k, x_{k-1}, \dots, x_{k-L}\}$ is:

$$p(r_k | x_k, x_{k-1}, \dots, x_{k-L}) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp \left[-\frac{(r_k - \sum_{n=0}^L \gamma_n x_{k-n})^2}{2\sigma_n^2} \right]. \quad (2.44)$$

The joint probability density function of a sequence of received signals r_1, r_2, \dots, r_K may be expressed as a product of K marginal PDFs, if and only if, the samples of the noise sequence μ_k are uncorrelated:

$$\begin{aligned} p(r_1, r_2, \dots, r_K | \underline{x}_L) &= \prod_{k=1}^K p(r_k | x_k, x_{k-1}, \dots, x_{k-L}) \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma_n} \right)^K \exp \left[-\sum_{k=1}^K \frac{(r_k - \sum_{n=0}^L \gamma_n x_{k-n})^2}{2\sigma_n^2} \right]. \end{aligned} \quad (2.45)$$

The detector which maximises the joint probability distribution function $p(r_1, r_2, \dots, r_K | \underline{x}_L)$ with respect to a sequence \underline{x}_L of transmitted symbols is called the *maximum-likelihood sequence estimator* (MLSE). In practice, the MLSE finds the sequence of input symbols which minimises the Euclidean distance $\sum_{k=1}^K (r_k - \sum_{n=0}^L \gamma_n x_{k-n})^2$ between the received signal r_k and the estimate $\sum_{n=0}^L \gamma_n x_{k-n}$, which can be obtained by taking the natural logarithm in equation 2.45 and rejecting the constants which are not associated with the received sequence r_k . Alternatively, the sequence that maximises the negative Euclidean distance metric can be found. Although the MLSE makes decisions on a symbol-by-symbol basis, the actual decisions are made based on an observation of a few received signal samples. All the possible paths or sequences of transmitted symbols can be represented by a tree diagram as shown in Figure 2.13 for a 4-QAM signal ($M = 4$).

The search always commences by computing $L + 1$ symbols in the tree diagram, where $L + 1$ is the length of the impulse response γ_n , corresponding to the combined transmitter filter, propagation channel and whitened matched filter. The Viterbi algorithm [29] [71], often used as a decoding algorithm for convolutional codes, provides a good technique to reduce the number of computations needed in the search for the most likely path through the tree diagram given the output from the whitened matched filter. This reduction in the number of states in each stage is necessary in order to avoid the tree diagram expanding exponentially.

At each stage, M^{L+1} metrics are computed but only M^L metrics survive to the next stage. At stage $L+1$, M groups are formed and each group is associated with a sequence $\{x_{L+1}, x_L, \dots, x_2\}$

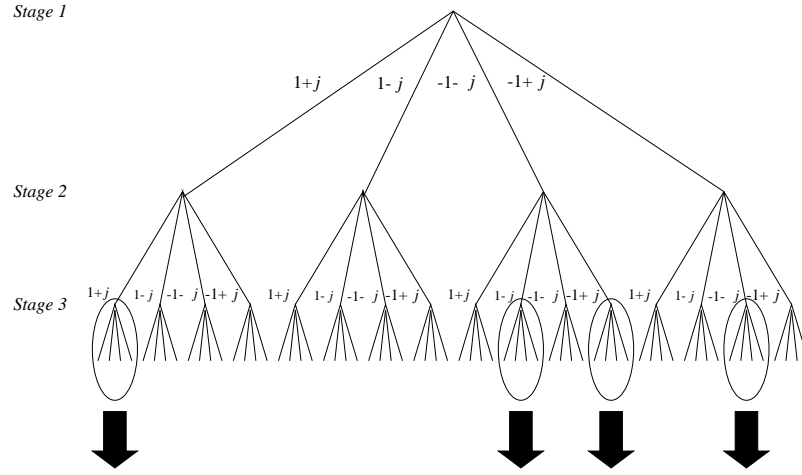


Figure 2.13: Tree diagram of Viterbi algorithm for a 4-QAM ($M = 4$) signal and $L = 2$. Only the $M^L = 4^2 = 16$ sequences marked with the ellipse survive to the next stage.

that differ from each other in x_1 . For each of the M^L sequences $\{x_{L+1}, x_L, \dots, x_2\}$ in the group, the sequence of the M groups which maximises the probability with respect to x_1 is selected and the rest $M - 1$ sequences are discarded. As a result, only M^L paths survive to the next stage. The most likely sequence is obtained by calculating the sequence which maximises the following metric:

$$PM_1(x_{L+1}, x_L, \dots, x_2) = \max_{x_1} \left[- \sum_{k=1}^L \left(r_k - \sum_{j=0}^{L-1} \gamma_j x_{k-j} \right)^2 \right]. \quad (2.46)$$

In general, equation 2.46 can be rewritten as:

$$PM_1(x_{L+1}, x_L, \dots, x_2) = \max_{x_1} \sum_{k=1}^{L+1} \ln p(r_k | x_k, x_{k-1}, \dots, x_{k-L}). \quad (2.47)$$

Upon the reception of more samples r_k the metrics are calculated subsequently using:

$$PM_k(x_{L+k}, x_{L+k-1}, \dots, x_{k+1}) = \max_{x_k} [\ln p(r_{L+k} | x_{L+k}, \dots, x_k) + PM_{k-1}(x_{L+k-1}, \dots, x_k)] \quad k > 1. \quad (2.48)$$

It follows that the metric for the next stage of the tree diagram of Figure 2.13 can be expressed as:

$$PM_2(x_{L+2}, x_L, \dots, x_2) = PM_1(x_{L+1}, x_L, \dots, x_2) - \left(r_{L+2} - \sum_{j=0}^{L-1} \gamma_j x_{k-j} \right)^2. \quad (2.49)$$

The decisions on each symbol are usually taken with a fixed delay. The M^L surviving sequences should agree on the earliest symbols if a sufficiently long delay $q = 5L$ is considered. Then, it might be guaranteed that the symbol x_{k-q} in each of the M^L sequences is the same. If it is not, the most likely symbol of the M^L sequences is selected.

2.6 Conclusions

Equalisation is the process of removing distortion created either by thermal noise or by the limitation in bandwidth availability and multipath effects of the propagation channel. In this chapter, the conventional equalisation procedure of transmitting a training or pilot signal periodically in order to update the coefficients has been reviewed. Under certain circumstances, it is impractical to employ a training signal in equalisation, and other procedures which do not involve prior knowledge of the transmitted sequence might be necessary. The convergence speed of blind equalisation algorithms will determine if these constitute a valid alternative to existing supervised methods.

Several blind equalisation techniques have appeared in the literature. All of them have in common that no prior knowledge of the transmitted sequence is assumed. Therefore, equalisation needs to be achieved using the only information that is available at the receiver; the received signal and its statistics. The first methods which appeared in the literature came under the name of self-recovery methods. A group of algorithms, known as Bussgang algorithms, provided a first approach to how blind equalisation might be achieved. These were methods which minimised a quadratic non-MSE cost function. The error signal was formed as the difference between the output of a transversal filter and the output of the same signal driven through a zero-memory nonlinear estimator. The different algorithms differ in the nonlinear estimator but overall, they suffer from a very slow convergence, and they exhibit local minima apart from global minima.

In order to account for the problem of convergence to a global minimum, it has been suggested that HOS would guarantee a global convergence. In general, HOS methods are found reliable but they suffer from a slow convergence, mainly caused by the large amount of data necessary in the estimation of the higher-order cumulants. On the other hand, HOS are sensitive to phase information and it has been shown that they can account for nonminimum phase channels. These channels, with zeros inside and outside the unit circle, require not only the magnitude response of the channel but the phase response too; SOS are phase-blind in that respect. On the other hand, since higher-order cumulants are blind to Gaussian processes, they are appropriate to remove the effect of additive Gaussian noise but they restrain the input signal not to be Gaussian or nearly Gaussian.

It was clear that if an extra degree of freedom was added to blind equalisation methods using only second-order statistics many of the inherent problems of Bussgang algorithms and HOS methods could be overcome. Fractionally-spaced sampling resulted in a series of blind equalisation algorithms, which were able to identify nonminimum phase channels. The meth-

ods exploited either implicitly or explicitly the cyclostationary nature of a signal sampled at a rate higher than the symbol rate and these resulted in algorithms with a faster convergence. Chapter 3 presents some features of cyclostationarity interesting for application in blind equalisation and describes some blind equalisation algorithms using fractionally-spaced sampling.

Finally, two equalisation structures have also been described in this chapter. The optimum linear equaliser, built as the cascade of a matched filter and a transversal filter, can account for amplitude distortion of the signal efficiently by using a LS or a ZF criterion in the calculation of the coefficients of the transversal filter. If the tap-spacing of the linear equaliser is a fraction of the symbol period, it has been shown that the fractionally-spaced linear equaliser can also account for phase distortion. The MLSE, on the other hand, is the optimum receiver if it is preceded by a matched filter. It has a nonlinear decision-boundary and it is usually implemented using the Viterbi algorithm, which reduces the number of computation to the computation of a fixed number of likelihood functions. The MLSE finds the most likely sequence from all possible combinations.

Chapter 3

Cyclostationarity

3.1 Introduction

Many engineering problems which treated processes as stationary have found an extra degree of freedom in the fact that most of these processes are cyclostationary in nature. A process is said to exhibit cyclostationarity if its statistics are periodic in time. Most man-made processes and systems of mechanical, meteorological and natural origin do, in fact, show this cyclic or rhythmic characteristic and by exploiting the cyclostationarity inherent in them, new techniques with enhanced performance have been developed. The idea of cyclostationarity is not new but William A. Gardner is believed to be the pioneer in this field, as he set the fundamental mathematical descriptions of cyclostationary processes in a series of articles in collaboration with other authors in [36] and previously in [35] and [34].

This chapter sets out the fundamental theory of second-order cyclostationarity with especial emphasis on the application of cyclostationary statistics in blind channel identification and equalisation. Some of the properties of cyclostationarity are inherent in most digital modulated signals and signals sampled in the receiver at a rate higher than the baud-rate of the transmitted signal. Among some of the properties which will be discussed in this chapter and employed in later chapters, are phase sensitivity or the possibility to distinguish systems equal in magnitude but different in phase, or the possibility to separate spectrally overlapping signals which exhibit spectral correlation at different frequencies.

This chapter is organised as follows: first, in section 3.2 the fundamental theory of cyclostationarity will be presented to follow with the definition of second-order cyclostationary statistics. The properties of cyclostationary signals and cyclostationary statistics will then be discussed in section 3.3 and some of the most salient blind equalisation algorithms that exploit cyclostationarity will be described in section 3.4. The issue of the channel identifiability conditions for the correct estimation of the channel characteristics will also be discussed in this section. Finally, section 3.5 draws some conclusions.

3.2 Fundamental Definitions

There are several ways to define a cyclostationary process. The fundamental definition is that a signal exhibits cyclostationarity if its statistics are periodic. This definition embraces all

systems that show up periodical variation in their characteristics due to vibrating or rhythmic phenomena . Cyclostationary processes are also known to exhibit spectral correlation between two frequency-shifted versions of the signal. If the signal exhibits spectral correlation for some frequency-shift α , the signal is said to exhibit cyclostationarity with cyclic frequency α [36]. At this point, it is necessary to define the mathematical framework in which a cyclostationary processes can be understood and consequently, derive the expressions for cyclostationary statistics.

Stochastic processes are known to be random phenomena which are functions of time. At any time instant, the value of a stochastic process is given by a random variable and in essence, stochastic processes are described in terms of random variables with a time-translation parameter t . The *probability distribution function* (PDF) of a stochastic process $X(t)$ is given by

$$F_{X(t)}(x) = P(X(t) < x), \quad (3.1)$$

where x is a real number. The PDF indicates the probability of the event $X(t) < x$ and it is defined in the range $0 \leq F_{X(t)}(x) \leq 1$ [71]. When the observations of the stochastic process $X(t)$ are done at different time instants t_1, t_2, \dots, t_n , we might expect to obtain different values of $X(t)$ every time sample and the PDF of $X(t)$ is given by the joint probability event operator

$$F_{X(t)}(x_{t_1}, x_{t_2}, \dots, x_{t_n}) = P(X(t_1) < x_{t_1}, X(t_2) < x_{t_2}, \dots, X(t_n) < x_{t_n}). \quad (3.2)$$

Similarly, the joint *probability density function* (PdF) of the stochastic process $X(t)$ is defined as:

$$p(x_{t_1}, x_{t_2}, \dots, x_{t_n}) = \frac{\partial^n}{\partial x_{t_1} \dots \partial x_{t_n}} F_{X(t)}(x_{t_1}, x_{t_2}, \dots, x_{t_n}). \quad (3.3)$$

Following the definition of the PDF it can be established that:

Definition 1: $X(t)$ is a stationary (S) process if and only if $F_{X(t)}(x_{t_1}, x_{t_2}, \dots, x_{t_n})$ is independent of the time-translation parameter t , i.e. $F_{X(t)}(x_{t_1}, x_{t_2}, \dots, x_{t_n}) = F_{X(t+\tau)}(x_{t_1+\tau}, x_{t_2+\tau}, \dots, x_{t_n+\tau})$ for any arbitrary τ .

Definition 2: $X(t)$ is a cyclostationary (CS) process with period T , if and only if $F_{X(t)}(x_{t_1}, x_{t_2}, \dots, x_{t_n})$ is periodic in t with period T , i.e. $F_{X(t)}(x_{t_1}, x_{t_2}, \dots, x_{t_n}) = F_{X(t+T)}(x_{t_1+T}, x_{t_2+T}, \dots, x_{t_n+T})$.

The joint PdF $p(x_{t_1}, x_{t_2}, \dots, x_{t_n})$ introduced in equation 3.3 is the partial derivative of the joint probability distribution function $F_{X(t)}(x_{t_1}, x_{t_2}, \dots, x_{t_n})$ and it can be used to determine

the statistical expectation of a stochastic random variable $X(t)$ or of a function of a random variable. The *mean* or expected value of the stochastic process $X(t)$ is defined as:

$$m_x(t_i) = E\{X(t_i)\} = \int_{-\infty}^{\infty} x_{t_i} p(x_{t_i}) dx_{t_i}. \quad (3.4)$$

For stationary stochastic processes, the PDF and consequently the Pdf are independent of the time-translation parameter t and the mean is independent too, i.e. $m_x(t + \tau) = m_x(t)$.

Similarly, the *correlation* of the stochastic process $X(t)$ can be defined as the expected joint value of two observations $X(t_1)$ and $X(t_2)$:

$$E\{X(t_1)X(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_{t_1} x_{t_2} p(x_{t_1}, x_{t_2}) dx_{t_1} dx_{t_2}. \quad (3.5)$$

This joint moment is commonly known as the autocorrelation function of the stochastic process $X(t)$. The autocorrelation of a stationary process is independent of the time instants t_1 and t_2 but dependent on the time difference $t_1 - t_2$. In practice, many processes which are non-stationary can show a stationary behaviour in its mean and autocorrelation. Such processes are known as *wide-sense stationary* because its mean and autocorrelation are stationary in time. Bearing this in mind, since only a finite set of samples from the process are normally available, the statistical average or expectation operator, for which the Pdf of the stochastic process needs to be known, can be substituted by the temporal-averaging operator:

$$\hat{E}\{X(t_i)\} = \lim_{Z \rightarrow \infty} \frac{1}{2Z} \int_{-Z}^Z X(t_i + \tau) d\tau = \int_{-\infty}^{\infty} x_{t_i} \hat{p}(x_{t_i}) dx_{t_i}, \quad (3.6)$$

and $\hat{p}(x_{t_i})$ is known as the *fraction-of-time* (FOT) Pdf.

From the definition of a cyclostationary process given in Definition 2, the periodic probability distribution function $F_{X(t)}(x_{t_1}, x_{t_2}, \dots, x_{t_n})$ can be expressed as the summation of periodic components:

$$F_{X(t)}(x_{t_1}, x_{t_2}, \dots, x_{t_n}) = \sum_{\alpha} F_{X(t)}^{\alpha}(x_{t_1}, x_{t_2}, \dots, x_{t_n}), \quad (3.7)$$

or equivalently by a set of *Fourier coefficients* :

$$F_{X(t)}(x_{t_1}, x_{t_2}, \dots, x_{t_n}) = \sum_{\alpha} F_{X(t)}^{\alpha}(x_{t_1}, x_{t_2}, \dots, x_{t_n}) e^{j2\pi\alpha t}. \quad (3.8)$$

A sine-wave component with a specific frequency α can be extracted using the following Fourier transform:

$$\begin{aligned}
F_{X(t)}^\alpha(x_{t_1}, x_{t_2}, \dots, x_{t_n}) &\triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} F_{X(t+\tau)}(x_{t_1+\tau}, x_{t_2+\tau}, \dots, x_{t_n+\tau}) e^{-j2\pi\alpha\tau} d\tau \\
&= \hat{E}\{F_{X(t)}(x_{t_1}, x_{t_2}, \dots, x_{t_n}) e^{-j2\pi\alpha t}\} e^{j2\pi\alpha t}.
\end{aligned} \tag{3.9}$$

3.2.1 The cyclic autocorrelation function

Identically to the definition of a wide-sense stationary process given in the previous section, a continuous-time process $x(t)$ is *cyclostationary in the wide-sense*, if its mean and autocorrelation are periodic in t with some period T .

$$m_x(t+T) = m_x(t) \tag{3.10}$$

$$r_x\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) = E\left\{x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right)\right\} = r_x\left(t + \frac{\tau}{2} + T, t - \frac{\tau}{2} + T\right), \tag{3.11}$$

for all $\tau \in (-T, T)$. From 3.8 the periodic autocorrelation function can be expressed in terms of a set of Fourier coefficients $R_x^\alpha(\tau)$, known as the *cyclic autocorrelation function* [36]:

$$r_x\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) = \sum_{\alpha} R_x^\alpha(\tau) e^{j2\pi\alpha t}, \quad \alpha = \frac{n}{T}, \quad n \in \mathcal{Z} \tag{3.12}$$

and

$$\begin{aligned}
R_x^\alpha(\tau) &\triangleq \hat{E}\left\{r_x\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) e^{-j2\pi\alpha t}\right\} \\
&\triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} r_x\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) e^{-j2\pi\alpha t} dt,
\end{aligned} \tag{3.13}$$

where α is known as the *cyclic frequency* or *cycle frequency*. In practice, the cyclic autocorrelation function can be determined using:

$$\begin{aligned}
R_x^\alpha(\tau) &\triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \left[\lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) dt \right] e^{-j2\pi\alpha t} dt \\
&= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi\alpha t} dt.
\end{aligned} \tag{3.14}$$

This property is known as *cyclo-ergodicity* [36]. For notational convenience, throughout this thesis the cyclic autocorrelation function will be expressed as:

$$R_x^\alpha(\tau) \triangleq \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T x(t+\tau) x^*(t) e^{-j2\pi\alpha(t+\frac{\tau}{2})} dt. \quad (3.15)$$

This notational change requires a time shift $\tau/2$ in the complex sine-wave component in order to match the cyclic autocorrelation function of equation 3.14¹.

3.2.2 The cyclic spectrum or spectral correlation density function

The cyclic autocorrelation function introduced in the previous section has an interesting connotation in the frequency domain:

$$S_x^\alpha(f) = \sum_{\tau=-\infty}^{\infty} R_x^\alpha(\tau) e^{-j2\pi f\tau}, \quad (3.16)$$

where $S_x^\alpha(f)$ is the *cyclic spectrum* defined as the Fourier transform of the cyclic autocorrelation function $R_x^\alpha(\tau)$. Substituting (3.14) in (3.16):

$$\begin{aligned} S_x^\alpha(f) &= \sum_{\tau=-\infty}^{\infty} \left\{ \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x\left(t + \frac{\tau}{2}\right) x^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi\alpha t} dt \right\} e^{-j2\pi f\tau} \\ &= \sum_{\tau=-\infty}^{\infty} \left\{ \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \left[x\left(t + \frac{\tau}{2}\right) e^{-j\pi\alpha t} \right] \left[x^*\left(t - \frac{\tau}{2}\right) e^{-j\pi\alpha t} \right] dt \right\} e^{-j2\pi f\tau}. \end{aligned} \quad (3.17)$$

If the following substitution of variables is carried out:

$$\begin{cases} u(t) = x(t) e^{-j\pi\alpha t} \\ v(t) = x(t) e^{j\pi\alpha t} \end{cases} \Rightarrow \begin{cases} x\left(t + \frac{\tau}{2}\right) e^{-j\pi\alpha t} = u\left(t + \frac{\tau}{2}\right) e^{j\pi\alpha\tau/2} \\ x^*\left(t - \frac{\tau}{2}\right) e^{-j\pi\alpha t} = v^*\left(t - \frac{\tau}{2}\right) e^{-j\pi\alpha\tau/2} \end{cases} \quad (3.18)$$

The cyclic spectrum function can be expressed as:

$$\begin{aligned} S_x^\alpha(f) &= \sum_{\tau=-\infty}^{\infty} \left\{ \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} u\left(t + \frac{\tau}{2}\right) v^*\left(t - \frac{\tau}{2}\right) dt \right\} e^{-j2\pi f\tau} \\ &= \sum_{\tau=-\infty}^{\infty} r_{uv}\left(t + \frac{\tau}{2}, t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} = S_{uv}(f). \end{aligned} \quad (3.19)$$

The cyclic spectrum $S_x^\alpha(f)$ is also known as the *spectral correlation density function* (SCD), because as indicated in equation 3.19, it represents the correlation of two frequency shifted

¹Some authors have omitted this time shift in the cyclic autocorrelation function [45] [6] and although strictly necessary to match the theory developed in [36], it does not affect the final results of the algorithms.

versions of $x(t)$, namely $u(t)$ and $v(t)$. This concept of spectral correlation provides another definition of a cyclostationary process, where a signal $x(t)$ is said to exhibit cyclostationarity if it shows spectral correlation for some non-zero cyclic frequency α .

If the baseband signal $x(t)$ is, on the other hand, applied to a low-pass filter with impulse response $p(t)$ the filtered output signal can be expressed as:

$$s(t) = p(t) \otimes x(t), \quad (3.20)$$

where \otimes denotes convolution. Following (3.18), the frequency translation variables $u(t)$ and $v(t)$ are now expressed as:

$$\begin{cases} u(t) = s(t)e^{-j\pi\alpha t} = [p(t) \otimes x(t)] e^{-j\pi\alpha t} \\ v(t) = s(t)e^{j\pi\alpha t} = [p(t) \otimes x(t)] e^{j\pi\alpha t} \end{cases} \quad (3.21)$$

According to (3.21), each term of the convolution is translated in the frequency domain by a factor $f \pm \alpha/2$ and the SCD function of the output of the lowpass filter can then be expressed as [36]:

$$S_s^\alpha(f) = P(f + \alpha/2)S_x^\alpha(f)P^*(f - \alpha/2) \quad f \in (-W, W), \quad (3.22)$$

where W denotes the bandwidth of the low-pass filter. According to (3.22), the region of support of the SCD function of the output of the filter lies in the region $|f| < W - 1/2T$.

3.3 Properties of Cyclostationarity

Some of the properties that derive from the formulation of the spectral correlation density function described in the previous section will lead to improvements in many areas of digital communications. Blind equalisation, as the main area of research in this thesis, is one of these areas, where the early stage of development of these techniques require extra degrees of freedom in order to overcome some of the existing problems. Only a subset of the properties of cyclostationary signals are discussed here, those related to the topic of the thesis, but a full list of the properties can be found in [36].

3.3.1 Phase recovery

Consider the cyclostationary signal $x(t)$ which is transformed by a linear filter with transfer function $H(z)$. The transfer function output of the filter is expressed as:

$$Y(z) = H(z)X(z). \quad (3.23)$$

According to 3.23, the power spectrum of the output signal is given by [71]:

$$S_y(f) = H(f)S_x(f)H^*(f) = |H(f)|^2S_x(f). \quad (3.24)$$

With the power spectrum of the output signal as the only information available and with knowledge of the power spectrum of the input signal, only the magnitude of the filter can be obtained. On the other hand, according to (3.22), the cyclic spectrum of the filter output is given by:

$$S_y^\alpha(f) = H(f + \alpha/2)S_x^\alpha(f)H^*(f - \alpha/2). \quad (3.25)$$

The objective in this case is to recover the phase of the frequency response $H(f)$ of the filter because the magnitude can be determined using 3.24. If the input signal's cyclic spectrum $S_x^\alpha(f)$ is constant for a non-zero α , the phase of the received signal cyclic spectrum $S_y^\alpha(f)$ is given by

$$\Psi_\alpha(f) = \Phi(f + \alpha/2) - \Phi(f - \alpha/2) \neq 0 \quad \text{for } \alpha \neq 0, \quad (3.26)$$

where $\Psi_\alpha(f) = \angle H(f)$ denotes the phase response of the filter frequency response $H(f)$.

3.3.2 Signal selectivity

One of the principal characteristics of cyclostationary statistics is the ability to identify a signal buried in high noise or overlapping in time and frequency with other signals from other users. Most man-made signals exhibit some kind of spectral correlation and when multiple signals-of-interest (SOI) and signals-not-of-interest (interference) overlap both in time and frequency, their spectral correlation functions do not overlap, because they exhibit spectral correlation at different cyclic frequencies as a result of signals having different carrier frequencies and/or pulse rates. This feature, known as spectral redundancy, enables the distinction of signals according to a modulation dependent spectral correlation pattern. Several patterns for different modulation schemes are available in [31].

The utility of the spectral redundancy in the case of noise present in a signal can be enhanced by intentionally making the signal exhibit spectral redundancy at a particular cyclic frequency α . The distinctive character of the spectral redundancy is very useful for signal selectivity. Taking the general case of multiple signals arriving at a receiver in the presence of noise,

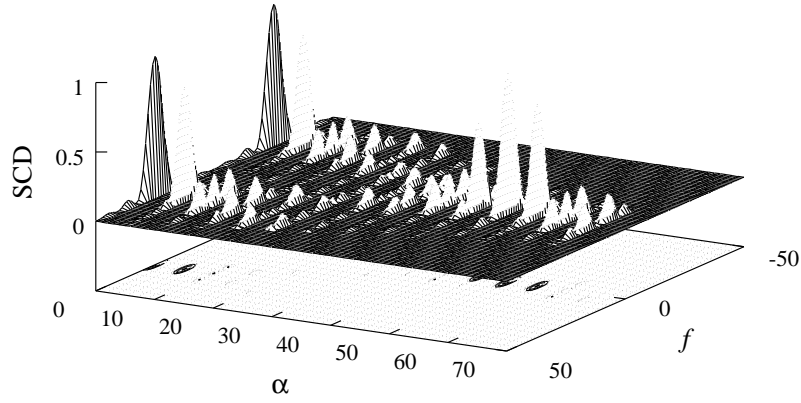


Figure 3.1: Spectral correlation density of one BPSK signals with centre frequency $f_{0,1} = 25$ Hz.

$$x(t) = \sum_{l=1}^L s_l(t) + n(t), \quad (3.27)$$

where $\{s_l(t)\}_1^L$ is the set of signals of interest and interference, and $n(t)$ the background noise. The spectral correlation function of the received signal is expressed as:

$$S_x^\alpha(f) = \sum_{l=1}^L S_{s_l}^\alpha(f) + S_n^\alpha(f), \quad (3.28)$$

but if the only signal that exhibits spectral redundancy at $\alpha = \alpha_k$ is $s_k(t)$, then we have,

$$S_x^\alpha(f) = S_{s_k}^\alpha(f). \quad (3.29)$$

For example, the SCD function of a BPSK signal with centre frequency 25 Hz, and the SCD functions of two AM signals with carrier frequencies of 20 Hz and 30 Hz are shown in Figure 3.1 and Figure 3.2 respectively. The SCD of the received signal is shown in Figure 3.3. The distinct features of the SCD functions of BPSK signals and AM signals enable us to determine the exact carrier frequencies of the AM signals and the centre frequency of the BPSK signal for $\alpha \neq 0$. In the case of $\alpha = 0$, the SCD function reduces to the PSD function and Figure 3.3 shows in fact that spectrally overlapping signals cannot be easily distinguished using the PSD of the received signal.

3.4 Cyclostationary Blind Equalisation Algorithms

In this section some of the cyclostationary blind equalisation algorithms summarised in the previous chapter will be described, in order to understand the analysis and comparisons carried out for different cyclostationary blind equalisation algorithms in the next chapter. The analysis reduces to two of the main cyclostationary blind equalisation approaches:

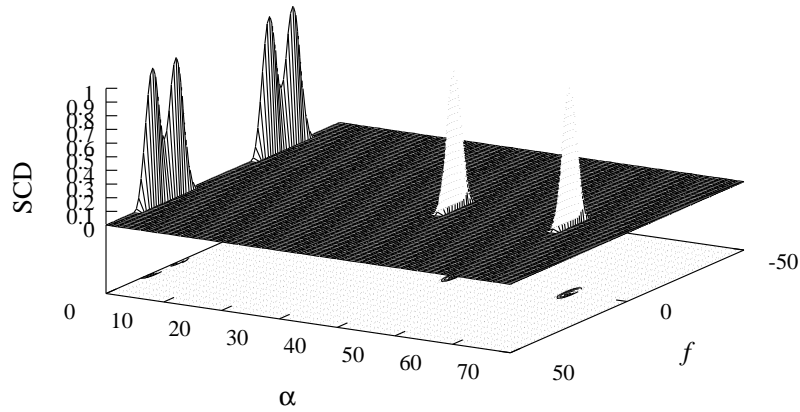


Figure 3.2: Spectral correlation density of two AM signals with carrier frequencies $f_{0,1} = 20$ Hz and $f_{0,1} = 30$ Hz.

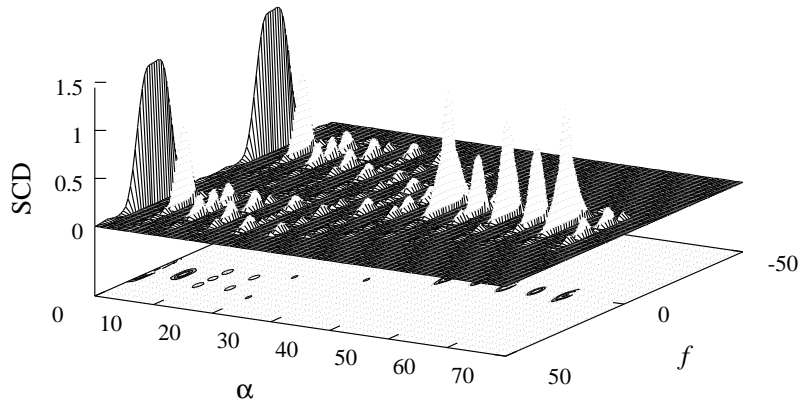


Figure 3.3: Spectral correlation density of a BPSK signal buried in two AM signals.

- cyclic statistics based algorithms: Hatzinakos [45]
- multichannel methods: Tong *et al.* [87] and Moulines *et al.* [61].

These algorithms will be presented using common notation, so that the equivalence between the algorithms can be easily established.

3.4.1 Oversampling system configuration

The oversampling operation, which consists of an increase in the sampling rate of the received signal can be represented in terms of either a single sensor sampled at a rate faster than the symbol or baud rate, or otherwise, as a signal which has gone through different propagation paths and is received by multiple sensors (antenna array) [33]. The configuration of the system is shown in Figure 3.4, where the discrete-time input sequence $x(k)$ is convolved with a different subchannel $h^{(i)}(n)$ as a result of oversampling. The received signal $y^{(i)}(t)$ at the i th sensor, sampled at T , can be formulated as:

$$y^{(i)}(n) = \sum_{k=-\infty}^{\infty} x(k)h^{(i)}(n-k) + v^{(i)}(n), \quad (3.30)$$

or equivalently

$$y(n) = \sum_{l=-\infty}^{\infty} x_o(l)h(n-lP) + v(n) \quad P \in Z, \quad (3.31)$$

where T is the symbol-period of the source signal, $x_o(l) = \sum_k x(k)\delta(l-kP)$ is the oversampled input sequence which takes nonzero values every P samples, $\delta(l)$ is the Dirac delta,

$$\delta(l) = \begin{cases} 1 & \text{if } l = 0 \\ 0 & \text{otherwise,} \end{cases} \quad (3.32)$$

$x(k)$ is the generally complex, independent identically-distributed (i.i.d.) sequence of symbols at the symbol-rate and $v(n)$ is a wide-sense stationary noise sequence statistically independent from $x(k)$. Since $P-1$ zeroes are interpolated between any two symbols of the input sequence $x(k)$ to form the oversampled input sequence $x_o(l)$, the non-zero values are multiplied by the channel coefficients $h(0), h(P), \dots, h(PM)$ at one time instant and by $h(1), h(P+1), \dots, h(PM+1)$ at another instant. The subchannel $h^{(i)}(n)$ associated with the i th sensor can be formulated as:

$$h^{(i)}(n) = h(i+nP), \begin{cases} i = 0, 1, \dots, P-1 \\ n = 0, 1, \dots, M \end{cases} \quad (3.33)$$

where M is the degree of ISI of the subchannels.

3.4.2 Cyclic statistics methods

A discrete-time blind deconvolution algorithm proposed by Hatzinakos in [45] is described in this section as an example of the family of blind identification using SOCS. The output of the channel is sampled at a rate higher than the baud-rate and then, the complex cepstrum of the cyclic autocorrelation is obtained.

Hatzinakos [45] Following the representation of an oversampled received signal given in (3.31), the input sequence $x_o(l)$ padded with zeros is a wide-sense cyclostationary process, i.e., its mean $E\{x_o(l)\}$ and autocorrelation $r_{x_o}(l, m) = E\{x_o(l+m)x_o^*(l)\}$ are periodic in l with period P .

$$r_{x_o}(l, m) = r_{x_o}(l+P, m). \quad (3.34)$$

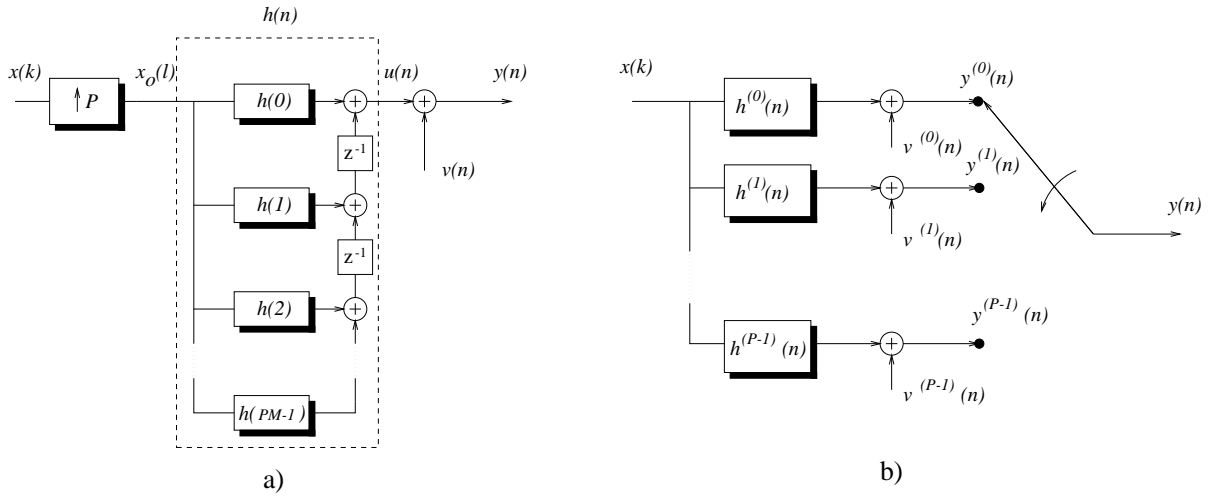


Figure 3.4: Discrete-time oversampled system model; a) fractionally-spaced sampling operation, b) multichannel representation.

According to (3.12), the cyclic autocorrelation function of the output sequence $y(n)$ can be defined as the Fourier coefficients of the discrete periodic autocorrelation function [45] [20]:

$$r_y(n, m) = \sum_{\alpha} R_y^{\alpha}(m) e^{i2\pi\alpha n} \quad \alpha = \frac{l}{P}, \quad l \in Z, \quad (3.35)$$

and

$$R_y^{\alpha}(m) = \lim_{L \rightarrow \infty} \frac{1}{LP} \sum_{n=0}^{LP-1} E\{y(n+m)y^*(n)\} e^{-i2\pi\alpha n}, \quad (3.36)$$

where L is the number of symbols used in the estimation.

The additive noise $v(n)$ is a wide-sense stationary sequence, statistically independent from $x_o(l)$. It will be assumed that the transfer function of the channel allows the following factorisation [45] [67]

$$H(z) = Az^r I(z^{-1}) O(z), \quad (3.37)$$

where, A is a constant gain, r is a constant delay in time, $I(z^{-1}) = \frac{\prod_{i=1}^{L_1} (1 - a_i z^{-1})}{\prod_{i=1}^{L_3} (1 - c_i z^{-1})}$ is a minimum phase polynomial and $O(z) = \prod_{i=1}^{L_2} (1 - b_i z)$ is a maximum phase polynomial, where $|a_i| < 1$, $|c_i| < 1$ and $|b_i| < 1$.

The coefficients $A(n)$ and $B(n)$ are the differential cepstrum parameters which contain the minimum phase and maximum phase information of the channel $h(n)$, respectively. They are defined as:

$$A(n) = \sum_{i=1}^{L_1} a_i^n - \sum_{i=1}^{L_3} c_i^n, \quad n = 1, 2, \dots, p \quad B(n) = \sum_{i=1}^{L_2} b_i^n, \quad n = 1, 2, \dots, q, \quad (3.38)$$

for some positive integers L_1 , L_2 and L_3 . The objective of the blind channel identification is to identify the unknown impulse response of the channel from a finite set of output samples and only partial information about $x(k)$. Then, by applying inverse filtering to the sequence $y(n)$, the original input $x(k)$ can be recovered.

Channel identification using the complex cepstrum of the cyclic autocorrelation

A homomorphic approach was proposed in [45] exploiting the ability of the complex cepstrum to separate maximum and minimum phase parts of the channel in (3.37). The complex cepstrum of the cyclic autocorrelation is defined as the inverse Fourier transform of the cepstrum [45] [67]:

$$c_y^\alpha(m) = \mathcal{Z}^{-1} [C_y^\alpha(z)], \quad (3.39)$$

and the cepstrum function is the natural logarithm of the cyclic spectrum of the received signal $y(n)$, i.e.:

$$C_y^\alpha(z) = \ln [S_y^\alpha(z)]. \quad (3.40)$$

Similarly to the continuous-time representation of the cyclic spectrum of the received signal of equation 3.25, the cyclic spectrum of the discrete received signal $y(n)$ can be expressed as:

$$S_y^\alpha(z) = \begin{cases} q_x H(z) H^*(z e^{-i2\pi\alpha}) + S_v(z), & \text{for } \alpha = 0 \\ q_x H(z) H^*(z e^{-i2\pi\alpha}), & \text{for } \alpha = \frac{l}{P}, \quad l \neq nP, \quad n \in \mathcal{Z}, \end{cases} \quad (3.41)$$

where $q_x = \frac{1}{P} E\{|x(n)|^2\}$. Note that as mentioned in the theoretical development of cyclostationarity, additive Gaussian noise does not exhibit spectral correlation at any particular frequency and thus, $S_v^\alpha(z) = 0$. Substituting (3.37) in (3.25) and applying to (3.40) we obtain:

$$\begin{aligned} C_y^\alpha(z) = & \ln(q_x |A|^2) + i2\pi r\alpha + \sum_{i=1}^{L_1} \ln(1 - a_i z^{-1}) - \sum_{i=1}^{L_3} \ln(1 - c_i z^{-1}) + \sum_{i=1}^{L_2} \ln(1 - b_i z) \\ & + \sum_{i=1}^{L_1} \ln(1 - a_i^* e^{-i2\pi\alpha} z) - \sum_{i=1}^{L_3} \ln(1 - c_i^* e^{-i2\pi\alpha} z) + \sum_{i=1}^{L_2} \ln(1 - b_i^* e^{i2\pi\alpha} z^{-1}). \end{aligned} \quad (3.42)$$

This relation shows that a homomorphic model which separates minimum phase and maximum phase zeros of a channel can be constructed. The differential cepstrum parameters are derived from (3.42) [45]. Provided that the complex cepstrum of the cyclic autocorrelation $c_y^\alpha(m)$ is well defined or equivalently, provided that the channel has no zeros on the unit circle [67], a set of linear overdetermined system of equations can be formulated.

$$R_\alpha \underline{a} = \underline{r}_\alpha, \quad (3.43)$$

where R_α is a $2w \times 2\max(p, q)$ $\{w \geq \max(p, q)\}$ matrix with entries from the cyclic autocorrelation lags of the form

$$\begin{bmatrix} R_y^\alpha(-w-1) & \dots & R_y^\alpha(-w-\max(p, q)) & R_y^\alpha(-w+1) & \dots & R_y^\alpha(-w+\max(p, q)) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ R_y^\alpha(-1-1) & \dots & R_y^\alpha(-1-\max(p, q)) & R_y^\alpha(-1+1) & \dots & R_y^\alpha(-1+\max(p, q)) \\ R_y^\alpha(1-1) & \dots & R_y^\alpha(1-\max(p, q)) & R_y^\alpha(1+1) & \dots & R_y^\alpha(1+\max(p, q)) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ R_y^\alpha(w-1) & \dots & R_y^\alpha(w-\max(p, q)) & R_y^\alpha(w+1) & \dots & R_y^\alpha(w+\max(p, q)) \end{bmatrix} \quad (3.44)$$

and p and q are two positive integers [45].

The $2\max(p, q) \times 1$ vector $\underline{a} = [C_\alpha(1), \dots, C_\alpha(\max(p, q)), \dots, D_\alpha(1), \dots, D_\alpha(\max(p, q))]$ has entries from the differential cepstrum parameters and $\underline{r}_\alpha = [wR_y^\alpha(-w), \dots, R_y^\alpha(-1), -R_y^\alpha(1), \dots, -wR_y^\alpha(w)]$ is a $2w \times 1$ vector.

In practice, the cyclic autocorrelation function can be estimated using only sample estimates [45] according to:

$$\hat{R}_y^\alpha(m) = \frac{1}{LP} \sum_{n=0}^{LP-1} y(n+m)y^*(n)e^{-i2\pi\alpha n}. \quad (3.45)$$

The solution to equation 3.43 exists and it is unique provided that the over-determined matrix R_α has linearly independent columns and it is given as a set of values $C_\alpha(n)$ and $D_\alpha(n)$ which are defined as

$$\left. \begin{array}{l} C_\alpha(n) = A(n) + B^*(n)z_\alpha^{-n} \\ D_\alpha(n) = A^*(n)z_\alpha^n + B(n) \end{array} \right\}, \quad n = 1, 2, \dots, \max(p, q). \quad (3.46)$$

The solution to this linear system determines the differential cepstrum parameters $A(n)$ and $B(n)$ of equation 3.38 and from there, the impulse response of the channel can be obtained using [45]:

$$h(k) = i(k) \otimes o(k), \quad (3.47)$$

where

$$i(k) = -\frac{1}{k} \sum_{n=2}^{k+1} [A(n-1)] i(k-n+1), \quad k = 1, \dots, N_1, \quad (3.48)$$

$$o(k) = \frac{1}{k} \sum_{n=k+1}^0 [B(1-n)] \cdot o(k-n+1), \quad k = -1, \dots, -N_2. \quad (3.49)$$

$i(k)$ and $o(k)$ are the minimum phase and maximum phase parts of the impulse response $h(k)$, respectively.

The selection of the integers p and q is carried out so that these are large enough and the differential cepstrum parameters which decay exponentially go below a certain constant threshold:

$$A(I) < C, \quad I > p = [\ln(C)/\ln(a)] \quad B(J) < C, \quad J > q = [\ln(C)/\ln(b)], \quad (3.50)$$

where $\max(|a_k|, |c_k|) < a < 1$ and $\max(|b_k|) < b < 1$ and C is a small constant. A common choice for this constant is $C = 10^{-4}$.

3.4.3 Multichannel methods

Inspired by *direction of arrival* (DOA) estimation techniques, several algebraic methods have been developed. In this section, a subspace method [61] reminiscent of the MUSIC algorithm [80] [85] for DOA estimation is described. It exploits the orthogonality property between the *noise* and *signal* subspaces and the result is given as the minimisation of a quadratic form. It is computationally more efficient than [87] because a single eigenvector decomposition is needed.

Moulines et al. [61] From the oversampling operation depicted by equation 3.30, where the discrete input sequence $x(k)$ is convolved with a different *subchannel* $h^{(i)}(n)$ every T/P seconds, a $N \times (N + M)$ filtering matrix $\mathcal{H}_N^{(i)}$ associated with a particular *subchannel* can be defined as:

$$\mathcal{H}_N^{(i)} = \begin{pmatrix} h^{(i)}(0) & \dots & h^{(i)}(M) & 0 & \dots & \dots & 0 \\ 0 & h^{(i)}(0) & \dots & h^{(i)}(M) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & h^{(i)}(0) & \dots & h^{(i)}(M) \end{pmatrix}, \quad (3.51)$$

where $\underline{h}^{(i)} = [h^{(i)}(0), \dots, h^{(i)}(M)]^T$ are the coefficients of the i th subchannel and N is the length of the observation window. Equation 3.30 can be rearranged in terms of the filtering matrix $\mathcal{H}_N^{(i)}$ as:

$$\underline{y}_k^{(i)} = \mathcal{H}_N^{(i)} \underline{x}_k + \underline{w}_k^{(i)} \quad \text{for } k = 0, \dots, L - 1, \quad (3.52)$$

where L is the number of symbols used in the estimation of the channel and $\underline{x}_k = [x(k), x(k-1), \dots, x(k-N-M+1)]^T$ is the $(N+M) \times 1$ vector which contains the input sequence $x(k)$.

In the same way, we can define the $N \times 1$ noise vector $\underline{w}_k^{(i)} = [w^{(i)}(k), w^{(i)}(k-1), \dots, w^{(i)}(k-N+1)]^T$ and the received signal vector $\underline{y}_k^{(i)} = [y^{(i)}(k), y^{(i)}(k-1), \dots, y^{(i)}(k-N+1)]^T$.

By expanding equation 3.52 to the P subchannels, the received signal vector \underline{y}_k can be written as:

$$\underline{y}_k \stackrel{\text{def}}{=} [\underline{y}_k^{(0)T}, \dots, \underline{y}_k^{(P-1)T}]^T, \quad (3.53)$$

and

$$\underline{y}_k = \mathcal{H}_N \underline{x}_k + \underline{w}_k, \quad (3.54)$$

where

$$\mathcal{H}_N \stackrel{\text{def}}{=} \Phi(\underline{h}, P, M+1, N) = [\mathcal{H}_N^{(0)T}, \dots, \mathcal{H}_N^{(P-1)T}]^T \quad (3.55)$$

is the $PN \times (N+M)$ filtering matrix.

The linear filtering representation of equation 3.54 gives way to the following representation of the received signal autocorrelation matrix

$$R_y = \mathcal{H}_N R_x \mathcal{H}_N^H + R_w, \quad (3.56)$$

where the superscript H denotes conjugate transpose. $R_x = E \{ \underline{x}_k \underline{x}_k^H \}$ is the source signal autocorrelation matrix and $R_w = E \{ \underline{w}_k \underline{w}_k^H \} = \sigma^2 I$ is the autocorrelation matrix of the white noise sequence with variance σ^2 and I , the identity matrix. It follows that provided that the unknown source autocorrelation matrix R_x is full-rank and matrix \mathcal{H}_N is full column rank, the received signal autocorrelation matrix admits the eigenvalue decomposition $R_y = U \Sigma U^H$, where matrix Σ is a diagonal matrix which contains the eigenvalues $\lambda_0, \dots, \lambda_{PN-1}$ of matrix R_y , and

$$\begin{aligned} \lambda_i &> \sigma^2 && \text{for } i = 0, \dots, M+N-1, \\ \lambda_i &= \sigma^2 && \text{for } i = M+N, \dots, PN-1. \end{aligned} \quad (3.57)$$

The eigenvectors of matrix U associated with the first $N+M$ eigenvalues λ_i form the so-called $PN \times (N+M)$ signal eigenvectors matrix $S = [\underline{s}_0, \dots, \underline{s}_{N+M-1}]$, and the eigenvectors associated with the last $PN - N - M$ conform the $PN \times (PN - N - M)$ noise eigenvector matrix $G = [\underline{g}_0, \dots, \underline{g}_{PN-N-M-1}]$. The $N+M$ columns of matrix S span the *signal subspace* whereas the columns of matrix G generate the so-called *noise subspace* instead. By orthogonality between the columns of matrix S and G , it can be concluded that any vector from the noise subspace is orthogonal to any column vector in the signal subspace and by extension to any column of the filtering matrix \mathcal{H}_N

$$\underline{g}_i^H \mathcal{H}_N = 0 \quad 0 \leq i \leq PN - N - M - 1. \quad (3.58)$$

When only sample estimates of the received signal autocorrelation matrix R_y are available, the set of linear equations of equation 3.58 can be solved in the least squares sense, as the minimisation of the following quadratic form:

$$q = \sum_{i=0}^{PN-N-M-1} |\hat{\underline{g}}_i^H \mathcal{H}_N|^2 = \sum_{i=0}^{PN-N-M-1} \hat{\underline{g}}_i^H \mathcal{H}_N \mathcal{H}_N^H \hat{\underline{g}}_i, \quad (3.59)$$

or equivalently

$$q = \sum_{i=0}^{PN-N-M-1} \underline{h}^H \mathcal{G}_i \mathcal{G}_i^H \underline{h} = \underline{h}^H \left(\sum_{i=0}^{PN-N-M-1} \mathcal{G}_i \mathcal{G}_i^H \right) \underline{h} = \underline{h}^H Q \underline{h}, \quad (3.60)$$

where $\underline{h} = [\underline{h}^{(0)T}, \dots, \underline{h}^{(P-1)T}]^T$ and $\mathcal{G}_i = \Phi(\hat{\underline{g}}_i, P, N, M + 1)$, as defined in (3.55).

The estimate of the channel parameter vector \underline{h} is obtained by picking up the eigenvector associated with the smallest eigenvalue of matrix Q . The estimate of the channel impulse response is non-coherent and proper gain and phase adjustments have to be made. The uniqueness of the estimates up to a scalar factor relies on the special structure of the filtering matrix \mathcal{H}_N which makes it full-rank provided that the following conditions are met:

- *i)* The polynomials $H^{(i)}(z) \stackrel{\text{def}}{=} \sum_{j=0}^M h^{(i)}(j)z^j$ have no common zero.
- *ii)* N is greater than the maximum degree M of the polynomials $H^{(i)}(z)$, i.e. $N \geq M$.
- *iii)* At least one polynomial $H^{(i)}(z)$ has degree M .

Tong et al. [88] The blind identification and equalisation approach of Tong *et al.* [88] was the first method which exploited fractionally-spaced sampling to identify nonminimum phase channels. Previous work [87] presented a method reminiscent of the ESPRIT algorithm for DOA estimation [76]. This approach can not only estimate the coefficients of a possibly nonminimum phase channel but it also provides an exact reconstruction of the source symbols.

The formulation of the algorithm is similar to the subspace method of Moulines *et al.* [61], where a vector of the noise-free oversampled received sequence \underline{y}'_k is expressed as:

$$\underline{y}'_k = \mathcal{H}(K) \underline{x}'_k \quad \text{for } k = 0, \dots, L - 1, \quad (3.61)$$

and the $KP \times (K + M)$ linear channel matrix $\mathcal{H}(K)$ is expressed as:

$$\mathcal{H}(K) = \begin{pmatrix} \underline{h}_M & \cdots & \underline{h}_0 & 0 & \cdots & \cdots & 0 \\ 0 & \underline{h}_M & \cdots & \underline{h}_0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & \underline{h}_M & \cdots & \underline{h}_0 \end{pmatrix}, \quad (3.62)$$

and $\underline{h}_k = [h^{(0)}(k), \dots, h^{(P-1)}(k)]^T$. The parameter K is known as the *smoothing factor* and if the received signal is oversampled, a sufficiently large value of K can be found, so that the block-Toeplitz matrix $\mathcal{H}(K)$ has more rows than columns. The $(K + M) \times 1$ input symbol vector is defined as $\underline{x}'_k = [x(k - K - M + 1), \dots, x(k)]^T$ and the $KP \times 1$ received signal vector takes the form $\underline{y}'_k = [y^{(0)}(k - K + 1), \dots, y^{(P-1)}(k - K + 1), \dots, y^{(0)}(k), \dots, y^{(P-1)}(k)]^T$.

The necessary and sufficient condition for identifiability states that the channel matrix $\mathcal{H}(K)$ has to be full-column rank²; in that case, provided that the sequence of input symbols $x(k)$ is a zero-mean stationary process with autocorrelation function $R_x(n) = E\{\underline{x}'_k \underline{x}'_{k-n}{}^H\}$, the channel matrix $\mathcal{H}(K)$ can be determined up to a scalar factor from measures of the autocorrelation lags $R_y(0)$ and $R_y(1)$.

The $KP \times KP$ autocorrelation function $R_y(0)$ in the same fashion as the subspace algorithm admits the eigenvalue decomposition $R_y(0) = \mathcal{U}\Sigma\mathcal{U}^H$, where the vector $\Sigma = \text{diag}(\lambda_0, \dots, \lambda_{K+M-1}, 0, \dots, 0)$ contains the eigenvalues of matrix $R_y(0)$. Similarly, we will denote as $\mathcal{S} = [\underline{s}_0, \dots, \underline{s}_{K+M-1}]$ the $K + M$ eigenvectors of matrix \mathcal{U} associated with the first $K + M$ eigenvalues $\lambda_0, \dots, \lambda_{K+M-1}$ and we will define a matrix

$$\mathcal{F} = \Sigma'^{-1}\mathcal{S}^H, \quad (3.63)$$

where $\Sigma' = \text{diag}(\sqrt{\lambda_0}, \dots, \sqrt{\lambda_{K+M-1}})$. From there we can obtain

$$R = \mathcal{F}R_y(1)\mathcal{F}^H. \quad (3.64)$$

Matrix R also admits an eigenvalue decomposition of the form $R = \mathcal{W}\text{diag}(\gamma_0, \dots, \gamma_{K+M-1})\mathcal{Z}$, where $\mathcal{W} = [\underline{w}_0, \dots, \underline{w}_{K+M-1}]^T$ and $\mathcal{Z} = [\underline{z}_0, \dots, \underline{z}_{K+M-1}]^T$. From there, the channel matrix $\mathcal{H}(K)$ can be identified using

$$\mathcal{H}(K) = \mathcal{S}\Sigma'Q, \quad (3.65)$$

where

$$Q = [\underline{w}_{K+M-1}, R\underline{w}_{K+M-1}, \dots, R^{(K+M-1)}\underline{w}_{K+M-1}], \quad (3.66)$$

²This condition is also present in the subspace algorithm

or equivalently,

$$Q = \left[(R^\dagger)^{(K+M-1)} \underline{z}_{K+M-1}, (R^\dagger)^{(K+M-2)} \underline{z}_{K+M-1}, \dots, \underline{z}_{K+M-1} \right], \quad (3.67)$$

where the superscript \dagger denotes Moore-Penrose generalised inverse [68]. In practice, the auto-correlation function lags need to be estimated from the available data using time averaging; also, the received signal is generally corrupted by additive noise. In this case, the eigenvalues of matrix $R_y(0)$ take up the form $\Sigma = \text{diag}(\lambda_0 + \sigma^2, \dots, \lambda_{K+M-1} + \sigma^2, \sigma^2, \dots, \sigma^2)$. The dimension of the signal subspace and the noise variance can be estimated using a threshold decision. Once these two parameters have been estimated the effect of the noise eigenvalues needs to be subtracted from $R_y(0)$ using

$$R_0 = \hat{R}_y(0) - \hat{\sigma}^2 I, \quad (3.68)$$

where I is the identity matrix. From this, \mathcal{S} consists of the eigenvectors associated with the largest $K + M$ eigenvalues of matrix R_0 . Matrix R , on the other hand, is obtained using

$$R = \mathcal{F} \left(\hat{R}_y(1) - R_w(1) \right) \mathcal{F}^H, \quad (3.69)$$

where $R_w(1) = \hat{\sigma}^2 J^P$ is the lag 1 of the noise autocorrelation function and,

$$J = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}. \quad (3.70)$$

Finally, the estimate of the channel matrix is obtained according to (3.65) where it is preferable to obtain matrix Q as a combination of (3.66) and (3.67):

$$Q = \left[\underline{w}_{K+M-1}, \dots, R^{(K+M)/2-1} \underline{w}_{K+M-1}, (R^\dagger)^{(K+M)/2-1} \underline{z}_{K+M-1}, \dots, \underline{z}_{K+M-1} \right], \quad (3.71)$$

when $K + M$ is even, or otherwise for odd $K + M$:

$$Q = \left[\underline{w}_{K+M-1}, \dots, R^{(K+M)/2} \underline{w}_{K+M-1}, (R^\dagger)^{(K+M)/2-1} \underline{z}_{K+M-1}, \dots, \underline{z}_{K+M-1} \right]. \quad (3.72)$$

3.4.4 Identifiability conditions

Since Tong *et al.* [87] presented their approach based solely on second-order statistics, several techniques have been proposed; not only they present new algorithms, but also they provide a

fundamental analysis of both cyclic statistics algorithms and multichannel algorithms [14] [25] [57] [45]. In those articles and the references therein, it has been shown that cyclostationary blind equalisation has the following advantages over HOS methods:

- Fewer data samples are required in the estimation of the channel.
- No constraint is imposed on the PDF of the input data.
- Fractionally-spaced sampling is less sensitive to timing errors than synchronous sampling.

On the other hand, it has been recognised that some channels are unidentifiable using cyclostationary blind equalisation methods. These channels are known as *singular* channels. The necessary and sufficient digital communications channels identifiability conditions for blind equalisation methods are presented in [89] [93], where it is shown that the FIR channel impulse response $h(n)$ is unidentifiable from the cyclostationary correlation sequence of the received sequence $y(n)$ oversampled by a factor P if:

- (i). The oversampled channel transfer function: $H(z) \triangleq \sum_n h(n)z^{-n}$ has zeros uniformly spaced around the unit circle with separation of $2\pi/T$. This is equivalent to the polynomials $H^{(i)}(z) \stackrel{\text{def}}{=} \sum_{j=0}^M h^{(i)}(j)z^j$ having a zero in common. The latter is one of the identifiability condition of the multichannel methods.
- (ii). The channel consists of time delays that are integer multiples of the symbol period T .
- (iii). The oversampling factor P is even and the channel consists of time delays which are integer multiples of $T/2$.

The proofs are provided in Appendix A. Condition (i) in the case of the multichannel methods relates to the need of the channel matrix \mathcal{H}_N , in the case of the subspace method of Moulines *et al.* , and $\mathcal{H}(K)$, for Tong's algorithm, to be full column rank [91]. In fact, if all the subchannels of the oversampled channel transfer function share at least one zero means that the subchannels are linearly dependent and the column full rank condition is not fulfilled [89].

These *critical* or *singular* channels do not have a direct physical interpretation to the kind of digital communications channels that might present these properties. If the order q of the channel is overestimated and the overestimation evaluates to $\hat{q} - q \geq P$, $(\hat{q} - q)/P$ common zeros will be introduced in $z = 0$ into the subchannels and condition (i) will stand. However, if the order of the channel is correctly estimated, condition (i) is rare in communications channels and it will only occur for specific values of the oversampling factor P . Moreover, even if conditions (i) and (iii) are met, the channels should be identifiable using a different oversampling factor P [93].

On the other hand, the limitation in the available bandwidth is very important for SOCS. A serious drawback of cyclic statistics methods is the inability to identify channels which are

strictly band-limited to a total bandwidth less than $2\pi/T$ [14]. Since it is known from (3.22) that the support band of the SCD function $S_y^{1/T}(\omega)$ lies in the band $|\omega| < 2\pi(W - 1/2T)$ for $W > 1/2T$, no information can be gained with the SCD function if the channel is strictly band-limited to $W \leq 1/2T$ Hz. Following (3.22),

$$S_y^{1/T}(\omega) = H\left(\omega + \frac{\pi}{T}\right) S_x^{1/T}(\omega) H^*\left(\omega - \frac{\pi}{T}\right), \quad (3.73)$$

which means that if the channel $H(\omega)$ is band-limited to $|\omega| < \pi/T$ the SCD function evaluated at $\alpha = 1/T$ will be 0 for all values of ω in that band. This condition for identifiability of band-limited channels was explained by Z. Ding [25] as an especial case of a more specific condition which states that channels with frequency nulls at $\pm \pi/T$ are unidentifiable from SOCS. In general, it will be assumed that the channel has excess bandwidth with a total bandwidth of $W = 2\pi(1 + \alpha)/T$. Other classes of channels, which are unidentifiable from SOCS, are channels with frequency nulls in $[-\pi(1 - \alpha)/T, \pi(1 - \alpha)/T]$ and channels with one null at $\omega_0 > 0$ and one at $\omega_0 - 2\pi/T$. A number of examples of *singular* channels unidentifiable using SOCS can be seen in [25] [57].

3.5 Conclusions

Cyclostationarity is inherent in many communication processes, as for instance in digital communication signals. Properties such as the ability to identify nonminimum phase channels due to the lack of symmetry of the spectral correlation density function in the frequency domain, or the ability to distinguish signals which are spectrally and temporally overlapping, are some of the interesting features of cyclostationarity. The application of these features to blind channel equalisation are discussed in the next chapters.

In this chapter a description of three representative algorithms of cyclic statistics methods and multichannel methods has been presented using common notation; the fundamental identifiability conditions have also been established. It has been shown, that under certain circumstances, cyclostationary techniques are not able to identify *certain* channels. If the system is for instance strictly band-limited, methods based on cyclostationary statistics cannot identify them because the spectral correlation density function is only defined in the excess-bandwidth area. On the other hand, if the channel exhibits zeros uniformly spaced around the unit circle with separation $2\pi/T$, neither cyclic statistics methods nor multichannel methods can identify the channel. This condition is equivalent to having P subchannels sharing a common zero.

Application of Cyclostationary Blind Equalisation to Mobile Radio

4.1 Introduction

The design of new systems for wireless communications has highlighted the need for higher transmission rates within the current available spectral bandwidth. In this scenario, blind equalisation can constitute a valid alternative because the systems are conceived without a training sequence and the convergence speed is not a determinant factor. Initial results in stationary channel environments suggest that the convergence of cyclostationary algorithms is satisfactory [61] [88] [45].

In this chapter, different issues related to the topic of cyclostationary blind equalisation are studied. The identifiability conditions of the algorithms described in the previous chapter are tested by means of computer simulations. It is also shown that fractionally-spaced blind equalisation algorithms lead directly to a fractionally-spaced equaliser (FSE) structure. Nevertheless, these techniques are also suitable for a MLSE receiver, where the tap coefficients of the equaliser are updated by estimates of the channel produced by the cyclostationary blind channel identification algorithm.

In the final part of this chapter, we will deal with the issue of blind equalisation in a wireless mobile radio environment. The GSM system environment was modelled and simulation results carried out to test the performance of cyclostationary blind equalisation algorithms are presented. The current GSM structure allocates 17% of a TDMA time slot for training which are used by the supervised channel estimation techniques to update the coefficients of the equaliser. With new mobile services that will soon be available, i.e. mobile broadcasting, not only difficulties arise from the fact that the channel conditions vary but also synchronisation is made difficult when clients switch on and off their mobile television sets randomly. Furthermore, the high transmission rates necessary for broadcasting applications suggest that training periods should be avoided and other means of unsupervised blind equalisation techniques would be desirable.

4.2 The Global System for Mobile

The GSM system or Global System for Mobile is the second-generation mobile communications standard adopted in Europe in the 1990s. It specifies a cellular structure and uses a TDMA/FDMA/FDD approach. The *time-division multiple access* (TDMA) approach is adopted to separate users in one cell sharing the same spectral band. The *frequency-division multiple access* (FDMA) approach, on the other hand, separates multiple channels and cells. Finally, the *frequency-division duplex* (FDD) separates uplink (*mobile station to base station*) from downlink (*base station to mobile station*).

The original purpose of cellular mobile radio systems was to provide telephone service to a large population of users in their analogue mobiles. The cellular systems showed an unlimited capacity in early stages, but soon, the industries were to find physical limitations from making the cells smaller and smaller. The first analogue cellular radio systems that emerged in Europe were incompatible, making it impossible to a user to use the mobile telephone in a different country. Thus, a second generation of cellular system had among its goals the creation of a single digital system against the existing collection of incompatible analogue systems. The second generation of cordless telephone system was intended to provide networks of base stations. The cordless telephone would be automatically assigned a channel which offered the least amount of interference. With the objective to develop a unified standard for a cellular pan-European mobile communication system for the 900 MHz band, the **Groupe Speciale Mobile** (GSM) was created [1].

In order to answer the huge demand for mobile communications, new approaches emerged so as to increase the capacity of existing cellular systems. Physical aspects and the expensiveness of new frequency bands were the most important limitations. The TDMA/FDMA approach provides an efficient utilisation of frequency bands that were currently in use. Each cell is allocated to a certain frequency band, so that they do not interfere with neighbouring cells. Inside the cell, data corresponding to each user is assigned a different time slot. These slots are multiplexed in time so that several users can communicate in the same spectral band. However, there is a limitation in the number of users that can be accommodated in one specific cell, considering how rapidly the channel characteristics may change.

4.2.1 Characterisation of the GSM system

The first generation of cellular systems used an analogue FM with a single-channel-per-carrier (SCPC) FDMA scheme. The channel bandwidth was either 25 or 30 kHz. The second generation GSM cellular system is based on a narrowband TDMA structure with 8 channels multiplexed in time and a FDMA system to separate signals corresponding to different cells. The frequencies at which the GSM operates are 890-915 Mhz for uplink and 935-960 MHz for downlink. This means that the duplex channels are spaced 45 MHz between uplink and downlink directions. A

guard band of 200 kHz is allowed between the bottom edge of each band and the first carrier, and 124 duplex channels fill in the rest of the available spectrum, with a separation of 200 kHz between carriers. The transmission rate in GSM is 271 kbits/s; at this rate, the multipath effect results in ISI. In order to combat ISI and the effect of *Gaussian minimum-shift keying* (GMSK) modulation, adaptive equalisation is used [83]. The GMSK modulation is widely used in fading channels due to its robustness against signal fading and interference [83], but it spreads each bit over more than one bit interval and results in additional ISI. Adaptive equalisation techniques are applicable to time-variant multipath channels under the condition that the time variations in the channel are relatively slow in comparison to the total channel bandwidth or, equivalently, to the symbol transmission rate over the channel [64].

The GSM frame structure is shown in Figure 4.1. The bottom level shows the structure of the *time slot*, containing 156.25 bits and with a duration of 576.92 μ s. There are 2 bursts of coded data each containing 58 bits, 26 bits used for training the adaptive equaliser, 3 start bits, 3 stop bits and 8.25 bits guard period. This gives a data bit interval of 3.69 μ s and a transmission rate of 270.833 kb/s.

At a higher level, there is the *frame* with a duration of 4.615 ms. Each FDMA channel is divided into 8 user time slots so that these 8 time slots are multiplexed in time and users of that cell can be assigned to one of these 8 slots. Users assigned to these time slots share the same carrier frequency.

At the next level, there is the *multiframe* with a duration of 120 ms. Each multiframe is composed of 26 frames, 24 of which are for user information and they are carried through traffic channels, while the remaining 2 are reserved for access control information.

Next, there is the *superframe*, which contains 51 multiframe (duration 6.12 s), and handles higher levels of the control structure such as handover decisions. Finally, at the higher level, there is the *hyperframe*, with a duration of 3 hr, 28 min, 54 s. It contains 2048 superframes, and it is used to provide sufficient security for encryption algorithms.

Access control channels

The *access control channels* (ACCH) are virtual channels which correspond to specific frames being allocated for network control purposes. In early implementations of GSM, only one of the available access control channels was used. The ACCH carries the status of the mobile station (MS) in the uplink direction, commands to set the transmitted power level from the base station (BS) to the mobile station (MS) in the downlink direction, and control messages relative to quick handovers. While first generation cellular systems interrupted the user voice signal to send a control signal resulting in audible clicks in the user's handset, second generation systems use the same channel used in the transmission of voice.

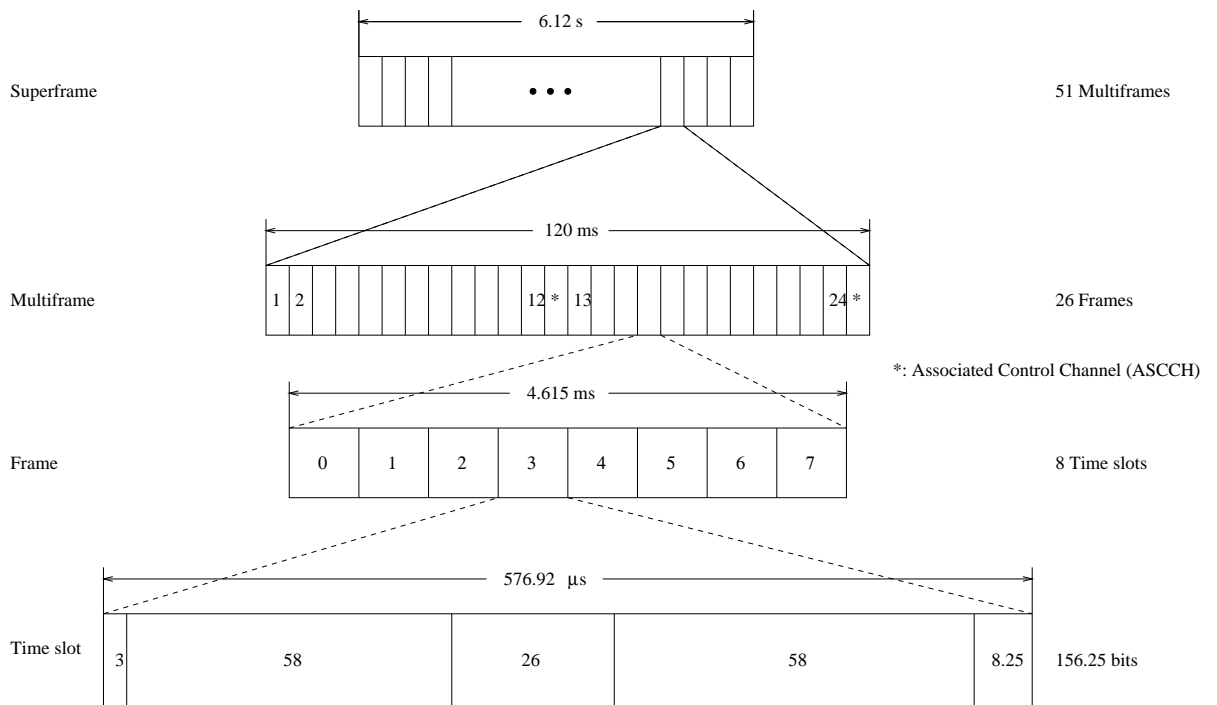


Figure 4.1: GSM frame structure.

The ACCCHs can be categorised in two types: the *fast associated control channel* (FACCH) provides quick handover commands and channel re-assignment in intra-cell handovers, and it is based on the blank-and-burst mode of the first generation systems [42], where the necessary space for the messages is stolen from traffic channels. The second type is the *slow associated control channel* (SACCH), which provides commands from the BS to the MS to set its output power level in downlink direction and in the uplink direction the MS responds with measured RF received signal power levels.

The SACCHs are accommodated within the traffic channels so that interruptions can be avoided, but it causes delays in the transmission of normal data. As shown in Figure 4.1, the 13th and 26th frames of the superframe are dedicated SACCHs. In early implementations with eight full rate channels, only the 13th frame was used, while the 26th frame was a dummy control frame. This would mean that a multiframe containing 26 frames would be transmitted in $26 \times 4.615 = 120$ ms. The SACCH transmission requires $\frac{1}{26} \times 24.7 = 950$ bps and reduces the traffic to $\frac{24}{26} \times 24.7 = 22.8$ kbps, wasting 950 bps in the dummy control frame. The full rate eight channel SACCH frame has eight time slots, each dedicated to the eight users sharing the same carrier frequency. A SACCH message contains 184 information bits which are expanded to 456 channel bits by error control coders, and these are distributed over four time slots. For one dedicated SACCH frame per superframe, a total of $4 \times 120 = 480$ ms are required to transmit a SACCH message at a rate of 950 bps.

Enhanced spectrum efficiency was later added to the GSM system allowing two sets of 12 frames for half rate channels in a multiframe. Hence, both dedicated SACCH frames are used,

increasing the channel capacity to 16 users and reducing the traffic rate to 11.4 kbps.

4.2.2 Modulation scheme

GSM specifies a GMSK modulator; GMSK modulation derives from the *minimum shift keying* (MSK) modulation scheme, where phase variations between adjacent bit periods are linear; this results in instantaneous changes in frequency. This clearly widens the spectrum, but by smoothing the phase transitions using a Gaussian filter this problem is diminished. The key parameter in GMSK is the product BT which represents the factor for which the bandwidth is reduced. Increasing BT improves interference resistance, but at the cost of more bandwidth occupancy. The typical values of BT are between 0.2 and 0.5, but the best compromise has been found at $BT = 0.3$ [83]. The advantage of using a smaller filter bandwidth is the reduction in spectral occupancy of the modulated signal.

When the transmission rate is such that the transmission rate to bandwidth ratio is $R/W \approx 1$, bandwidth-efficient modulation scheme such as M -ary QAM, PAM, or PSK are appropriate [71]. In contrast to the M -ary orthogonal signals, which are not bandwidth efficient, the M -ary QAM, PAM, or PSK increase the transmission rate to bandwidth ratio R/W as a result of increasing M , but at the expense of an increase in energy-per-bit ratio E_b/N_0 . However, with the M -ary orthogonal signals, the R/W ratio decreases as M increases, but at the same time a lower energy-per-bit ratio is achieved. In conclusion, modulation schemes such as M -ary QAM, PAM or PSK are appropriate in situations where the communication channel is band-limited and the transmission rate-to-bandwidth is $R/W > 1$, i.e. telephone channels and mobile radio channels. The M -ary orthogonal signals, on the other hand, are appropriate for power-limited channels with no restriction in channel bandwidth $R/W \leq 1$.

4.2.3 The propagation channel

The propagation channel in a mobile radio environment is affected by highly dispersive multipath effects caused by large reflectors, such as buildings and surrounding topographic accidents and scattering caused by reflectors close to the mobile [7]. As a consequence, several versions of the signal arrive at the antenna with different angles of arrival and times of arrival. The non-principal paths are attenuated with respect to the principal line-of-sight (LOS) path, but in some cases, different non-principal paths can add constructively, producing a path with higher energy than the main path. A serious problem can occur when the different paths add vectorially and produce a resultant vector with a very small scalar value. In this case, indistinctive of what the power of the signal is, the received signal will be negligible and it is said that the signal is in a deep fade. This effect is known as multipath fading, and in this condition, the received signal is totally dependent on the additive noise and in some cases it can lead to negative SNRs. Furthermore, since the source of the signal is moving, the angles of arrival of

the different paths change and this affects the Doppler shift of a particular path. As a result, deep fades will occur approximately every half wavelength along its route.

The propagation channel can be described by a set of discrete taps occurring at a specific delay, and with a particular amount of power associated with it. Channels are characterised by the probability of incurring in a deep fade. In this sense, channels may be classified as: Gaussian, Rician or Rayleigh, in ascending order of probability of incurring in a deep fade. A practical parameter to identify channels is the Rician parameter K , defined as:

$$K = \frac{\text{power of the principal path}}{\text{power of the scattered paths}}. \quad (4.1)$$

For Rayleigh channels, K is 0 or negligible, and Gaussian channels are characterised by very large values of K , ideally ∞ , but in practice a threshold value of $K = 32$ is used [83].

According to the distribution of the taps in the propagation channel impulse response, the amplitude of each of the taps will vary according to a Doppler spectrum $S(\tau_i, f)$, which is a function of the particular path τ_i .

The COST207 model [28], created by the Commission of the European Communities for simulating a GSM system, specifies a power-delay profile for a *rural (non-hilly) area* (RA), *typical (non-hilly) urban* (TU) area, *bad (hilly) urban* (BU) area and a *hilly terrain* (HT). In this chapter, the TU environment will be considered, and the implementation of this environment will be carried out according to the model proposed in [52]. The power-delay profile of the TU environment proposed by the COST207 model is given by 2 general classes of Doppler spectra, which are used to describe the variation of the channel weights at specific time delays of the impulse response. These are the:

Classical Doppler: The classical Doppler distribution is given by

$$S(\tau_i, f) = \frac{a}{\sqrt{[1 - (f/f_d)^2]}}, \quad (4.2)$$

where $f_d = v/\lambda$ is the maximum Doppler frequency of the channel, v (m/s) represents the vehicle speed, λ is the wavelength of the carrier and a is a gain factor. A cascade of two second-order filters was used in the implementation of the Doppler spectrum, consisting of a standard Butterworth low-pass filter, and a modified Butterworth filter that *rings* at the cut-off frequency [72]. The filters are described in the Laplace domain as:

$$H(s) = \frac{1}{s^2 + \sqrt{2}s + 1} \frac{1}{s^2 + 0.02s + 1}. \quad (4.3)$$

Gaussian: The Gaussian spectrum is given by the formula:

$$G(f) = G(A, f_1, f_2) = A \exp\left(-\frac{(f - f_1)^2}{2f_2^2}\right), \quad (4.4)$$

Delay (μs)	Power (dB)	Doppler Category
0.0	-3	Classical
0.2	0	Classical
0.6	-2	GAUS1
1.6	-6	GAUS1
2.4	-8	GAUS2
5.0	-10	GAUS2

Table 4.1: Power-delay profile for a typical (non-hilly) urban environment.

where f_1 and f_2 control the centre and the width of the spectrum respectively. The filter that simulates the Gaussian spectrum is a fourth order Bessel filter [72], described in the Laplace domain as:

$$H(s) = \frac{105}{s^4 + 10s^3 + 45s^2 + 105s + 105}. \quad (4.5)$$

This filter specifies the basic Gaussian filter. Two kinds of Gaussian spectra called GAUS1 and GAUS2 will be needed to describe the power-delay profiles of the reflectors with longer delays. Both GAUS1 and GAUS2 consist of two basic Gaussian spectrum generators with different gain, centre frequency and spectrum width. GAUS1 is given by:

$$S(\tau_i, f) = G(A, -0.8f_d, 0.005f_d) + G(0.1A, 0.4f_d, 0.1f_d), \quad (4.6)$$

where A is the gain factor. Similarly, GAUS2 can be expressed as:

$$S(\tau_i, f) = G(B, 0.7f_d, 0.1f_d) + G(0.0316B, -0.4f_d, 0.15f_d). \quad (4.7)$$

A set of complex weights will then be generated as shown in Figure 4.2. These describe the variation of the different paths according to a particular mobile radio channel environment. The complex channel weights are generated by passing AWGN through the filters described earlier. These tap weights are then multiplied by their corresponding power factor s_k . The COST207 model specifies that the power-delay profile of the 6 tap TU is as shown in Table 4.1.

As it can be seen, the TU environment exhibits one path at $0.2 \mu s$, with higher power than the LOS path. The multipath effect extends to $5 \mu s$, more than one bit period ($T = 3.69 \mu s$), and decreases in exponential form. The TU environment exhibit Rayleigh fading, with coefficient $K = 0.61$.

4.3 Performance Comparison of Cyclostationary Techniques in Stationary Multipath Channel Environments

In this section, the performance of three cyclostationary algorithms representatives of cyclic statistics based methods and algebraic methods are analysed and compared. Channel estimation results are shown, where the algorithms are simulated in environments with multipath effects

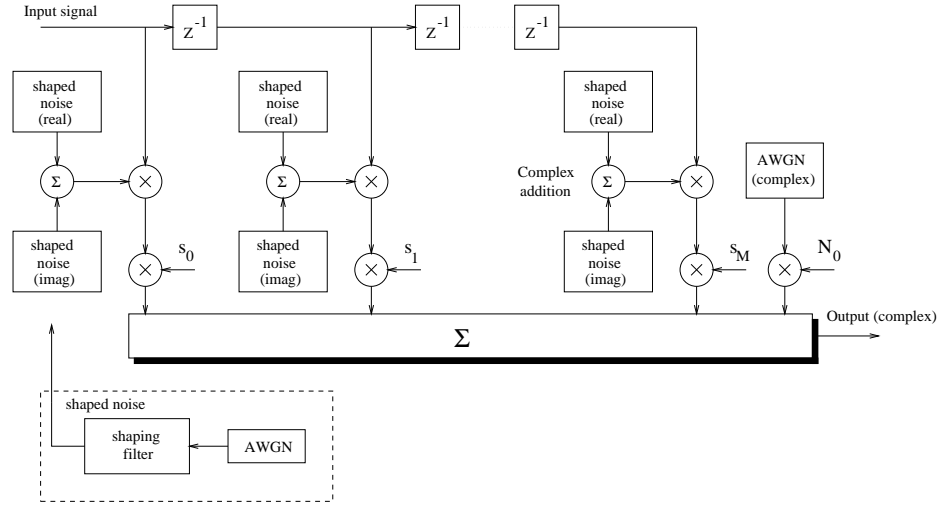


Figure 4.2: Mobile channel simulator.

and high SNRs. Particularly, the algorithms are tested with a channel that is unidentifiable using cyclostationary techniques.

4.3.1 Moving-average channel parameter estimation

Blind equalisation techniques are essentially channel estimation techniques. Therefore, in order to study their performance, the accuracy of the channel estimation needs to be tested first; the identifiability problems of each of the algorithms have to be analysed by means of computer simulations. These simulations were carried out using four models of fixed multipath channels with the following transfer functions:

Mixed Phase I (CH1)

$$H(z) = 1 + 0.9z^{-1} + 1.385z^{-2} - 0.771z^{-3}. \quad (4.8)$$

The zeros of the system are located at 0.403, $-0.651 \pm j1.219$ in the z -plane.

Mixed Phase II (CH2)

$$H(z) = 1 + 0.602z^{-1} + 0.9984z^{-2} - 1.3368z^{-3}. \quad (4.9)$$

The zeros of the system are located at 0.7, $-0.651 \pm j1.219$ in the z -plane.

Minimum Phase I (CH3)

$$H(z) = 1 - 0.9z^{-1} - 0.17z^{-2} + 0.265z^{-3}. \quad (4.10)$$

The zeros of the system are located at $-0.5, 0.7 \pm j0.2$ in the z -plane.

Minimum Phase II (CH4)

$$H(z) = 1 + 0.125z^{-3}. \quad (4.11)$$

The zeros of the system are located at $-0.5, 0.25 \pm j0.433$ in the z-plane ¹.

The first two channels are mixed phase channels, with two zeros outside the unit circle and one zero inside the unit circle. The difference between CH1 and CH2 is the location of the minimum phase zero, which in the case of CH2 is located closer to the unit circle. The other two channels, CH3 and CH4, are minimum phase channels with zeros inside the unit circle. Note that CH4 has three zeros located symmetrically around the unit circle.

In all cases, the simulations were conducted using a bipolar binary input signal $x(n)$ corrupted by white Gaussian noise (AWGN) at the output of the channel filter:

$$y(n) = h(n) \otimes x(n) + v(n), \quad (4.12)$$

where \otimes denotes convolution and $v(n)$ is the AWGN process. The oversampling factor is in all cases $P = 3$, and an *energy-per-bit-to-noise ratio* of $E_b/N_0 = 9$ dB was considered. The E_b/N_0 specifies the variance of the noise present at the received signal for the amount of energy required to transmit one input bit [71], where the variance of the noise is expressed as $\sigma^2 = N_0/2$.

Subspace method

The identification of the channel parameters, in the case of the subspace method, involves the following stages:

- (i) select a sufficiently large number of symbols in the observations N , i.e. $\underline{x}_k = [x(k), x(k-1), \dots, x(k-N-M+1)]^T$,
- (ii) estimation of the autocorrelation matrix of the oversampled received signal \underline{y}_k , by time-averaging,

$$\hat{R}_y = \frac{1}{L} \sum_{k=1}^L \underline{y}_k \underline{y}_k^H, \quad (4.13)$$

- (iii) compute eigenvalue decomposition of the $PN \times PN$ autocorrelation matrix using the *singular-value decomposition* (SVD) algorithm ²,

$$\hat{R}_y = U \Sigma U^H = U \text{diag}(\lambda_0, \dots, \lambda_{PN-1}) U^H, \quad (4.14)$$

¹The transfer functions of the channels are expressed at the sampling interval, rather than at the symbol interval

²see Appendix B for details of the algorithm

(iv) by orthogonality between the signal and noise subspaces, the coefficients of the channel filtering matrix \mathcal{H}_N can be identified by minimising

$$q = \sum_{i=0}^{PN-N-M-1} |\hat{\underline{g}}_i^H \mathcal{H}_N|^2, \quad (4.15)$$

or alternatively, by maximising,

$$q' = \sum_{i=0}^{M+N-1} |\hat{\underline{s}}_i^H \mathcal{H}_N|^2. \quad (4.16)$$

The first approach leads to a noise subspace identification approach, where the channel vector $\underline{h} = [\underline{h}^{(0)T}, \dots, \underline{h}^{(P-1)T}]^T$ is identified by computing the SVD of matrix $Q = \left(\sum_{i=0}^{PN-N-M-1} \mathcal{G}_i \mathcal{G}_i^H \right)$, i.e. $Q = U_q \Sigma_q U_q^H$, and selecting the column eigenvector of U_q associated with the smallest eigenvalue of Σ_q . The solution $\hat{\underline{h}} = [u_{1,P(M+1)}, \dots, u_{P(M+1),P(M+1)}]^T$ is a unit norm estimate of the channel vector \underline{h} , because it gives the smallest value of the cost function q . Ideally, the smallest eigenvalue of matrix Σ_q is zero.

$$q = \underline{h}^H \begin{bmatrix} u_{1,1} & \dots & u_{1,P(M+1)} \\ \vdots & \ddots & \vdots \\ u_{P(M+1),1} & \dots & u_{P(M+1),P(M+1)} \end{bmatrix} \begin{bmatrix} \lambda_1 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \lambda_{P(M+1)-1} & 0 \\ 0 & \dots & 0 & 0 \end{bmatrix} \times \begin{bmatrix} v_{1,1}^* & \dots & v_{P(M+1),1}^* \\ \vdots & \ddots & \vdots \\ v_{1,P(M+1)}^* & \dots & v_{P(M+1),P(M+1)}^* \end{bmatrix} \underline{h}. \quad (4.17)$$

Alternatively, one might choose the signal subspace to minimise $q = N|\underline{h}|^2 - \underline{h}^H \tilde{Q} \underline{h}$, where $\tilde{Q} = \left(\sum_{i=0}^{N+M-1} \mathcal{S}_i \mathcal{S}_i^H \right)$ because the computation of \tilde{Q} requires the multiplication of $M+N-1$ vectors, compared to the $PN-N-M-1$ of matrix Q in the noise subspace.

In these simulations, the oversampling factor was considered to be $P=3$ and the degree of ISI was the minimum value possible for the four channels described earlier $M=1$. The observation window was $N=5$ which represents the number of rows of matrix $\mathcal{H}_N^{(i)}$ in (3.51). Figure 4.3 shows the estimated positions of the zeros for CH1, CH2, CH3 and CH4, as well as the mean value over 20 Monte-Carlo simulations. In all the cases, the subspace algorithm has been able to identify the zeros correctly, except for CH4, where the presence of 3 zeros equally-spaced around the unit circle does not allow the accurate identification of the channel unless another oversampling factor is chosen.

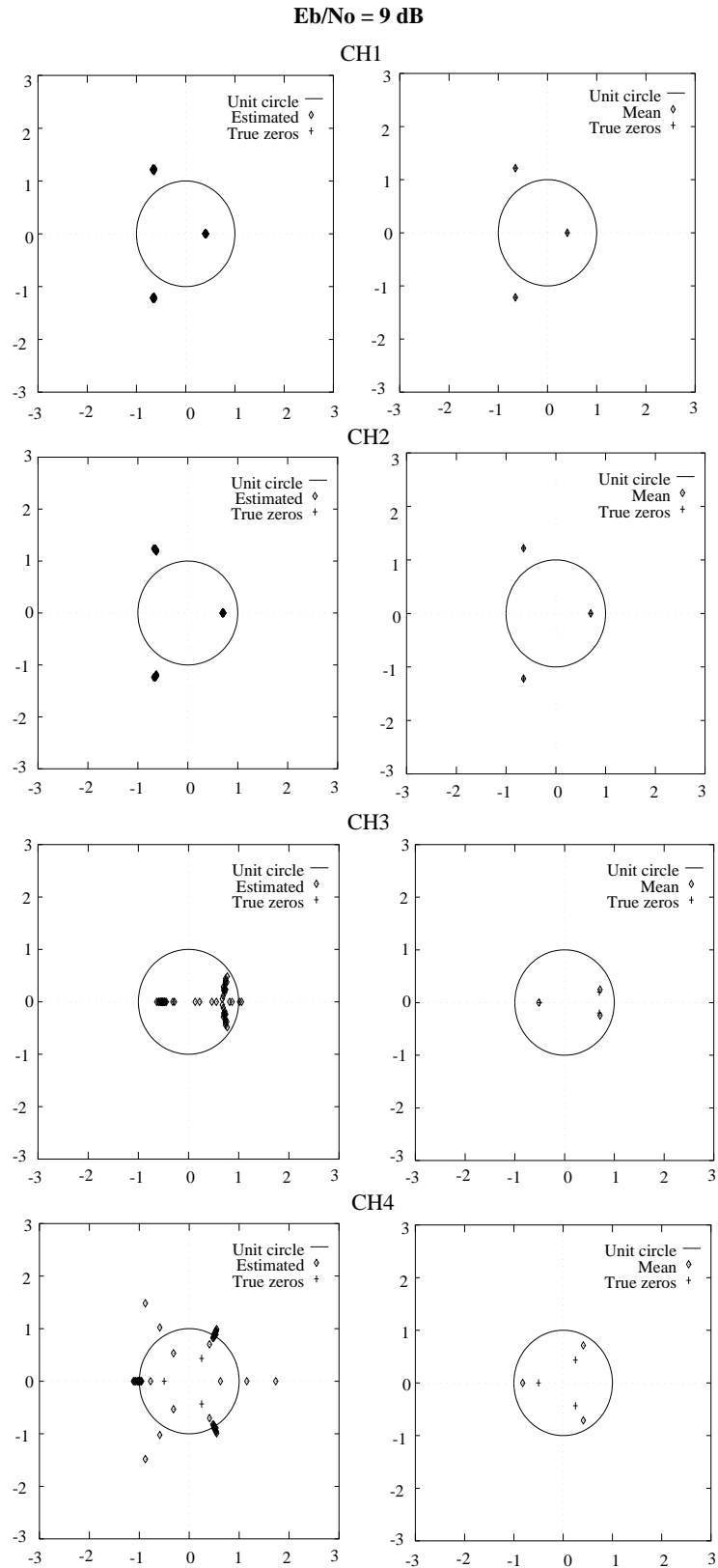


Figure 4.3: Estimated, mean and true positions of the zeros for channels CH1, CH2, CH3 and CH4. The estimates of the zeros were obtained for 20 Monte-Carlo realisations of the signal subspace blind channel estimation technique of Moulines *et al.* .

Cyclic method

The channel estimation using the cyclic algorithm involves the following stages:

- (i) estimation of the sample cyclic autocorrelation function using (3.45),
- (ii) select parameters p and q according to the position of the zeros of the channel. In practice, one could choose those values of p and q that minimise the MSE of the channel estimation. Then select w so that $w \geq \max(p, q)$,
- (iii) form the $2w \times 2w$ matrix R_α and $2w \times 1$ vector \underline{r}_α as shown in section 3.4.2,
- (iv) solve (3.43) as a least-squares solution:

$$\underline{a} = R_\alpha^\dagger \underline{r}_\alpha, \quad (4.18)$$

where \dagger denotes pseudo-inverse.

- (v) solve the set of linear equations (3.46), and obtain the differential cepstrum parameters $A(k)$ and $B(k)$. For an oversampling factor of $P = 3$, the values of $A(k)$ and $B(k)$ for $k = \pm P, \pm 2P, \dots$ cannot be recovered, unless some kind of interpolation is carried out [45].
- (vi) once the differential cepstrum parameters have been identified the channel parameters can be obtained using (3.47).

The true values of the differential cepstrum parameters for the channels used in the simulations have been listed in Table 4.2. It can be observed that as the zeros get closer to the unit circle, bigger values of p and q must be selected. Note the difference in the number of differential cepstrum parameters that contribute in $A(k)$ for CH1, with a minimum phase zero located at 0.403 in the unit circle, compared to CH2, with a minimum phase zero located at 0.7.

Figure 4.4 shows the identification of the zeros of CH1, CH2, CH3 and CH4, for an oversampling factor $P = 3$. The identification is very poor in all the cases. The cyclic algorithm can only estimate those differential cepstrum parameters $A(k)$ and $B(k)$ which are not multiples of $P = 3$, because the equations of the linear system (3.46) are linearly dependent for $k = P, 2P, \dots$. In these cases, those values of $A(k)$ and $B(k)$ can be estimated using Newton-Aitken interpolation [45]. Table 4.2 shows the true values of $A(k)$ and $B(k)$, and Table 4.3 shows the values which were interpolated. It can be seen that the interpolated values (denoted by an asterisk), differ considerably from the true ones shown in Table 4.2. CH4 is clearly unidentifiable using an oversampling factor $P = 3$, because the only non-zero values of $A(k)$ and $B(k)$ are the ones that need to be interpolated ($A(2), A(5), \dots$ and $B(2), B(5), \dots$).

Nevertheless, if an oversampling factor of $P = 5$ is chosen, the identification of the zeros is very good for all the channels, as shown in Figure 4.5. In this case, the interpolation of those values which cannot be identified (denoted by an asterisk) is far better than with $P = 3$ with as few as 200 symbols used in the estimation of the channel.

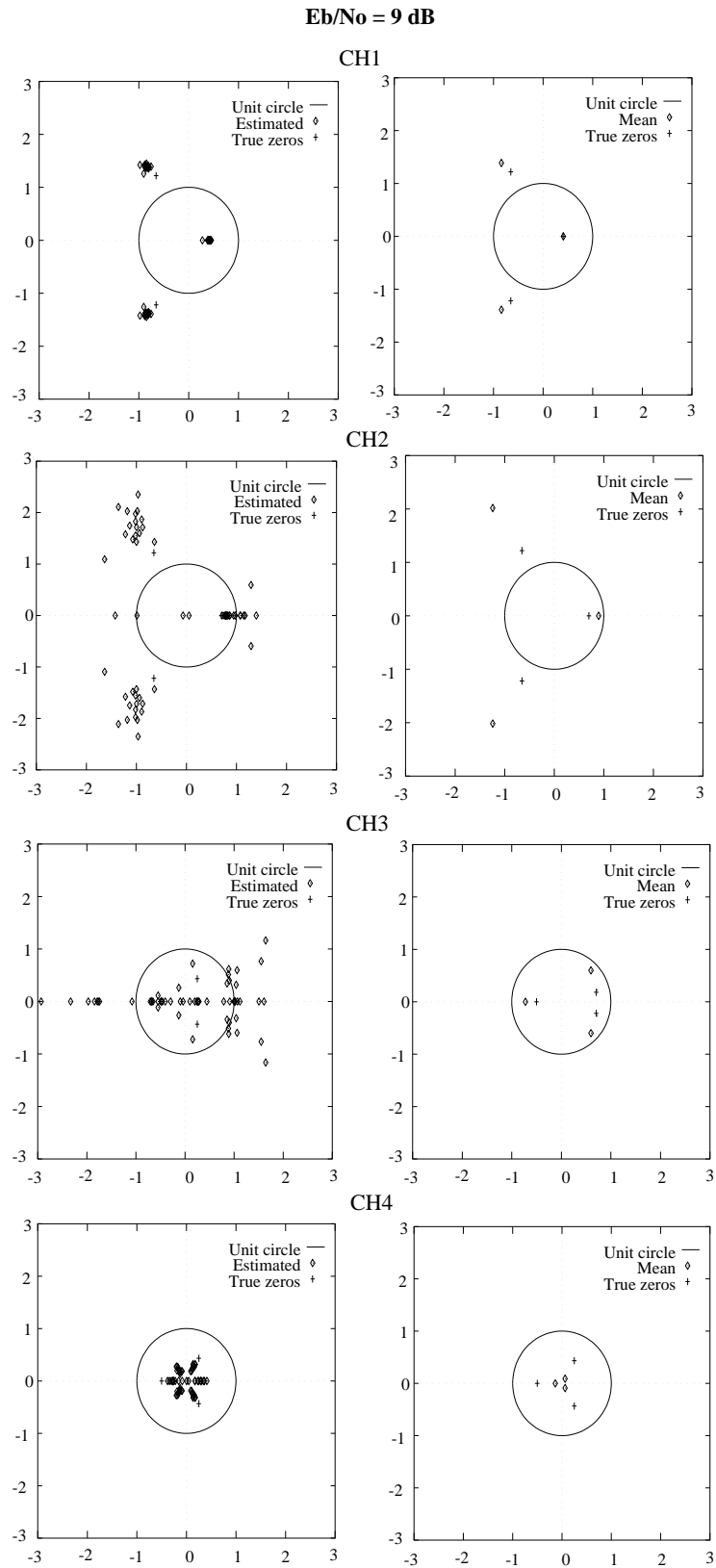


Figure 4.4: Estimated, mean and true positions of the zeros for channels CH1, CH2, CH3 and CH4 for an oversampling factor $P = 3$. The estimates of the zeros were obtained for 20 Monte-Carlo realisations of the cyclic blind channel estimation technique of Hatzinakos .

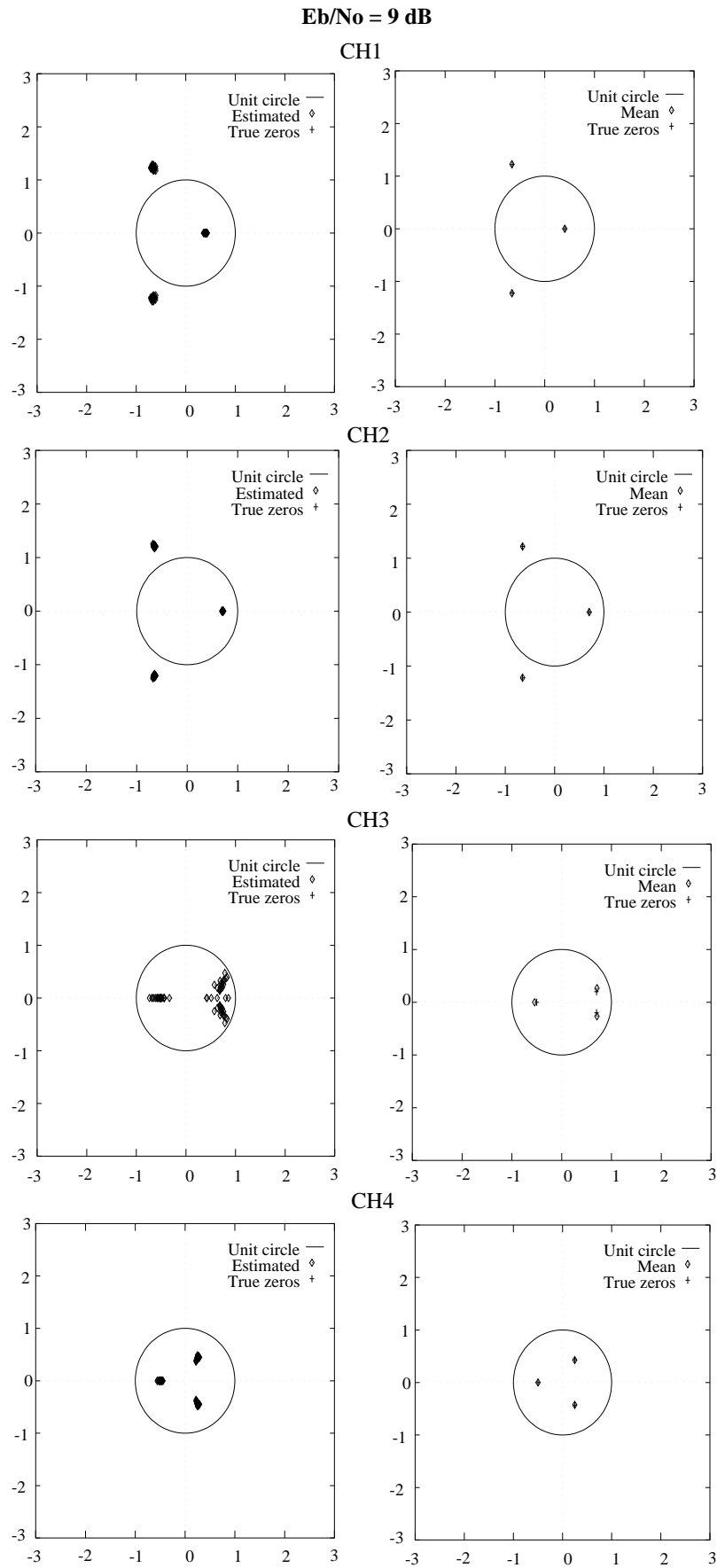


Figure 4.5: Estimated, mean and true positions of the zeros for channels CH1, CH2, CH3 and CH4 for an oversampling factor $P = 5$. The estimates of the zeros were obtained for 20 Monte-Carlo realisations of the cyclic blind channel estimation technique of Hatzinakos .

k	CH1		CH2		CH3		CH4	
	$A(k)$	$B(k)$	$A(k)$	$B(k)$	$A(k)$	$B(k)$	$A(k)$	$B(k)$
0	0.402	-0.6817	0.7	-0.6817	0.9	0.0	0.0000	0.0
1	0.1616	-0.5825	0.49	-0.5825	1.15	0.0	0.0000	0.0
2	0.0649	0.7541	0.343	0.7541	0.393	0.0	-0.375	0.0
3	0.0261	-0.2091	0.2401	-0.2091	0.3107	0.0	0.0000	0.0
4	0.0105	-0.2523	0.1681	-0.2523	0.0417	0.0	0.0000	0.0
5	0.0042	0.2815	0.1176	0.2815	-0.0138	0.0	0.0469	0.0
6	0.0017	-0.0598	0.0823	-0.0598	-0.0877	0.0	0.0000	0.0
7	0.0007	-0.1066	0.0576	-0.1066	-0.0923	0.0	0.0000	0.0
8	0.0003	0.10401	0.0403	0.10401	-0.0943	0.0	-0.0059	0.0
9	0.0001	-0.0151	0.0282	-0.0151	-0.0773	0.0	0.0000	0.0
10	0.0000	-0.0442	0.0198	-0.0442	-0.0612	0.0	0.0000	0.0
11	0.0000	0.0380	0.0138	0.0380	-0.0432	0.0	0.0007	0.0
12	0.0000	-0.0028	0.0097	-0.0028	-0.0288	0.0	0.0000	0.0
13	0.0000	-0.0180	0.0068	-0.0180	-0.0171	0.0	0.0000	0.0
14	0.0000	0.01373	0.0047	0.01373	-0.0088	0.0	0.0000	0.0

Table 4.2: Differential cepstrum parameters for channels CH1, CH2, CH3 and CH4.

TXK method

The method proposed by Tong *et al.* [87], commonly known as the TXK method, is another of the multichannel blind channel identification methods. The procedure to identify the channel impulse response is as follows:

- (i) Estimate $\hat{R}_y(0)$ and $\hat{R}_y(1)$ from the oversampled received signal, using time-averaging:

$$\hat{R}_y(0) = \frac{1}{L} \sum_{k=1}^L \underline{y}_k \underline{y}_k^H, \quad \hat{R}_y(1) = \frac{1}{L} \sum_{k=1}^L \underline{y}_k \underline{y}_{k-1}^H. \quad (4.19)$$

- (ii) Carry out eigenvalue decomposition of the $KP \times KP$ autocorrelation matrix $\hat{R}_y(0)$, and using threshold decision determine the dimension d of the signal space. Then, knowing that the last $KP - d$ eigenvalues of matrix $\hat{R}_y(0)$ denote the variance of the noise, estimate the variance σ^2 by averaging the $KP - d$ smallest eigenvalues.
- (iii) Compute the SVD of the $(K + M) \times (K + M)$ matrix R (3.64).
- (iv) Finally, determine the channel block-Toeplitz matrix $\hat{\mathcal{H}}(K)$ using (3.65).

Figure 4.6 shows the identification of the zeros of channels CH1, CH2, CH3 and CH4 using only 200 symbols in the estimation. The smoothing factor K , defined in section 3.4.3, was set to 5. The results suggest that in the case of CH1, CH2 and CH3, the identification is quite good. Compared to the subspace method of Moulines *et al.*, the estimated values show more deviation in the case of the TXK method. Moreover, from these 3 channels, the TXK method experiences problems with CH3. In the case of CH4, the identification is not possible in theory, but in practice a large variance in the mean estimated value is observed. Nevertheless, in this case, the subspace algorithm showed higher variance. It should be noted, however, that CH4 would be identifiable by changing the oversampling factor.

k	CH1		CH2		CH3		CH4	
	$A(k)$	$B(k)$	$A(k)$	$B(k)$	$A(k)$	$B(k)$	$A(k)$	$B(k)$
0	0.402	-0.6817	0.7	-0.6817	0.9	0.0	0.0000	0.0
1	0.1616	-0.5825	0.49	-0.5825	1.15	0.0	0.0000	0.0
2	0.0564*	-0.3721*	0.342*	-0.3721*	0.8168*	0.0*	0.0000*	0.0*
3	0.0261	-0.2091	0.2401	-0.2091	0.3107	0.0	0.0000	0.0
4	0.0105	-0.2523	0.1681	-0.2523	0.0417	0.0	0.0000	0.0
5	0.0037*	-0.1554*	0.1681*	-0.1554*	-0.0671*	0.0*	0.0000*	0.0*
6	0.0017	-0.0598	0.0823	-0.0598	-0.0877	0.0	0.0000	0.0
7	0.0007	-0.1066	0.0576	-0.1066	-0.0923	0.0	0.0000	0.0
8	0.0002*	-0.0638*	0.0402*	-0.0638*	-0.0883*	0.0*	0.0000*	0.0*
9	0.0001	-0.0151	0.0282	-0.0151	-0.0773	0.0	0.0000	0.0
10	0.0000	-0.0442	0.0198	-0.0442	-0.0612	0.0	0.0000	0.0
11	0.0000*	-0.0258*	0.0138*	-0.0258*	-0.0443*	0.0*	0.0000*	0.0*
12	0.0000	-0.0028	0.0097	-0.0028	-0.0288	0.0	0.0000	0.0
13	0.0000	-0.0180	0.0068	-0.0180	-0.0171	0.0	0.0000	0.0
14	-0.0000*	-0.0115*	0.0029*	-0.0115*	-0.0066*	0.0*	0.0000*	0.0*

Table 4.3: Differential cepstrum parameters with interpolated values for channels CH1, CH2, CH3 and CH4 for an oversampling factor $P = 3$. A Newton-Aitken interpolation technique of degree 3 was used which requires knowledge of the 2 previous and 2 posterior values of the differential cepstrum parameters.

k	CH1		CH2		CH3		CH4	
	$A(k)$	$B(k)$	$A(k)$	$B(k)$	$A(k)$	$B(k)$	$A(k)$	$B(k)$
0	0.402	-0.6817	0.7	-0.6817	0.9	0.0	0.0000	0.0
1	0.1616	-0.5825	0.49	-0.5825	1.15	0.0	0.0000	0.0
2	0.0649	0.7541	0.343	0.7541	0.393	0.0	-0.375	0.0
3	0.0261	-0.2091	0.2401	-0.2091	0.3107	0.0	0.0000	0.0
4	0.0091*	-0.0674*	0.1675*	-0.0674*	0.1470*	0.0*	0.0938*	0.0*
5	0.0042	0.2815	0.1176	0.2815	-0.0138	0.0	0.0469	0.0
6	0.0017	-0.0598	0.0823	-0.0598	-0.0877	0.0	0.0000	0.0
7	0.0007	-0.1066	0.0576	-0.1066	-0.0923	0.0	0.0000	0.0
8	0.0003	0.10401	0.0403	0.10401	-0.0943	0.0	-0.0059	0.0
9	0.0001*	0.0513*	0.0282*	0.0513*	-0.0811*	0.0	-0.0040*	0.0*
10	0.0000	-0.0442	0.0198	-0.0442	-0.0612	0.0	0.0000	0.0
11	0.0000	0.0380	0.0138	0.0380	-0.0432	0.0	0.0007	0.0
12	0.0000	-0.0028	0.0097	-0.0028	-0.0288	0.0	0.0000	0.0
13	0.0000	-0.0180	0.0068	-0.0180	-0.0171	0.0	0.0000	0.0
14	0.0000*	-0.0115*	0.0029*	-0.0115*	-0.0066*	0.0*	0.0000*	0.0*

Table 4.4: Differential cepstrum parameters with interpolated values for channels CH1, CH2, CH3 and CH4 for an oversampling factor $P = 5$. A Newton-Aitken interpolation technique of degree 3 was used which requires knowledge of the 2 previous and 2 posterior values of the differential cepstrum parameters.

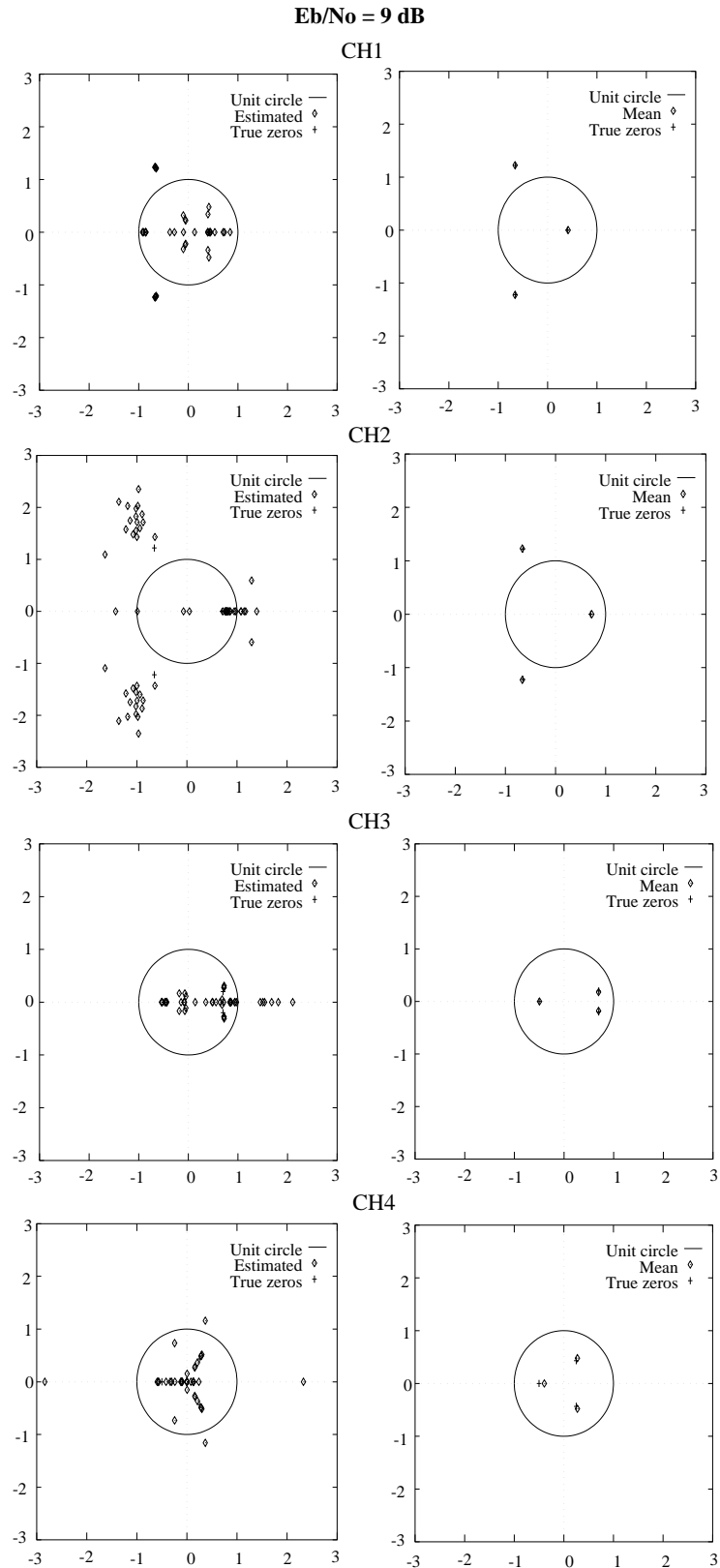


Figure 4.6: Estimated, mean and true positions of the zeros for channels CH1, CH2, CH3 and CH4. The estimates of the zeros were obtained for 20 Monte-Carlo realisations of the blind multichannel estimation technique of Tong *et al.*.

Performance comparison

The performance of the 3 algorithms was tested in terms of the accuracy of the channel estimations using the *normalised mean-square error* (NMSE) as a measurement. The NMSE for a general case of complex tap channels is defined as:

$$\text{NMSE} \triangleq \frac{1}{\text{MC}} \sum_{i=1}^{\text{MC}} \frac{\sum_{j=1}^{P(M+1)} |\hat{h}_i(j) - h(j)|^2}{\sum_{j=1}^{P(M+1)} |h(j)|^2}, \quad (4.20)$$

where MC denotes the number of Monte-Carlo simulations conducted, and $P(M + 1)$ is the length of the channel impulse response. The term $\hat{h}_i(j)$ denotes the estimate of the j th sample of the channel impulse response at the i th realisation of the experiment.

Figure 4.7 shows the results of the experiment in terms of the NMSE. The NMSE was evaluated for different lengths of data used in the estimation, and an *energy-per-bit ratio* of 9 dB. As discussed before, the performance of the cyclic algorithm is particularly bad for an oversampling factor $P = 3$, and as a consequence, both the subspace and TXK methods were tested for $P = 3$, but $P = 5$ was used in the case of the cyclic algorithm. For channels CH1, CH2 and CH3, both the subspace and the cyclic algorithms perform similarly, with a slight performance improvement of the subspace algorithm with respect to the cyclic algorithm for 200 symbols used in the estimation. The TXK method is outperformed in the case of the 3 channels, but especially for CH3, where particular problems with the identification of the zeros were reported earlier. As far as CH4 is concerned, the advantage of using an oversampling factor of $P = 5$ in the cyclic method is apparent. The degradation of both multichannel methods is also apparent.

4.3.2 Equalisation results

This section extrapolates the results on channel identification to the recovery of the transmitted sequence by means of inverse filtering. The procedures that cyclostationary blind equalisation algorithms use to implement a FSE will be described.

Coloured noise compensation: A whitened-matched filter (WMF) structure

The optimum symbol-spaced linear receiver described in Chapter 2 is essentially the cascade of a matched filter, a symbol-spaced sampler, a noise-whitening filter and a transversal equaliser, as shown in Figure 4.8. The matched filter is ideally matched to the combined transmitter filter $p(n)$ and propagation channel $g(n)$, which can be represented by a discrete-time impulse response $h(n)$. The optimum symbol-spaced linear equaliser can, as a result, be expressed as the combination of the noise whitening filter $1/F^*(z^{-1})$ and a transversal equaliser $1/F(z)$. The aim of the noise-whitening filter is to remove the correlation caused by the matched filter

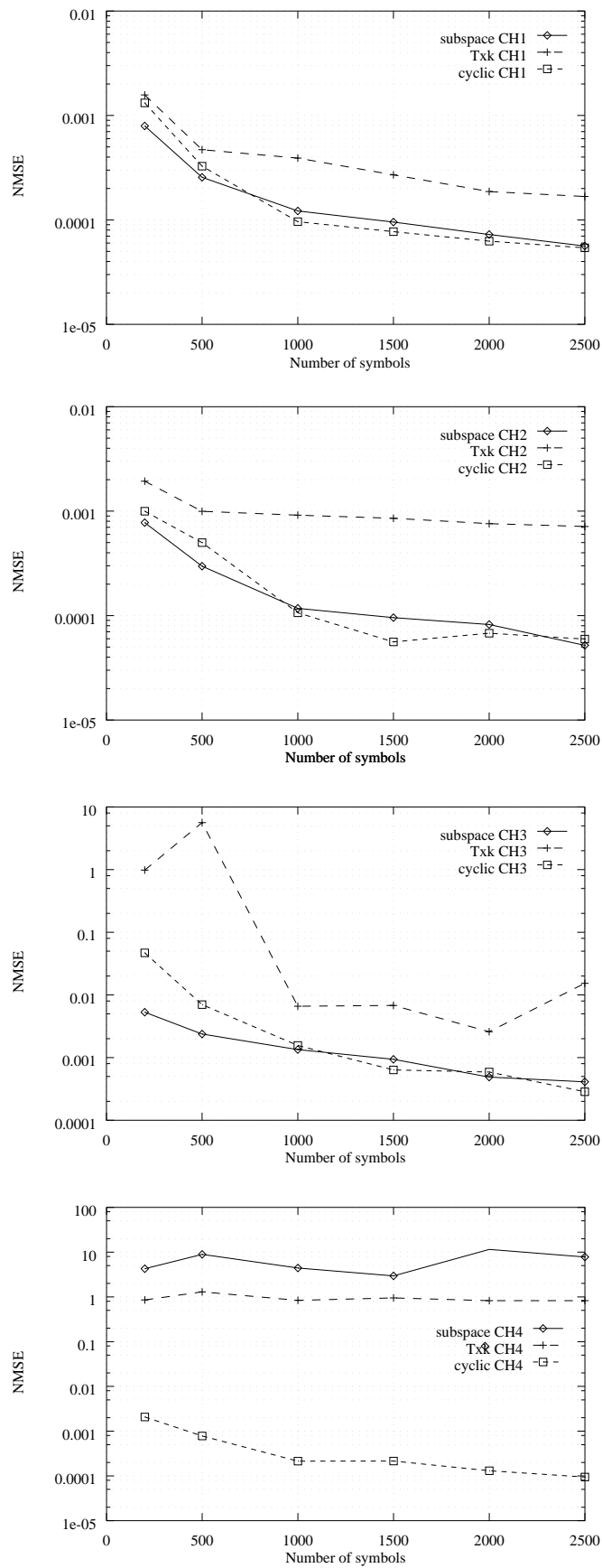


Figure 4.7: Normalised mean-square error (NMSE) of the channel parameter estimates, conducted over 20 Monte-Carlo runs and different number of symbols used in the estimation of the channel.

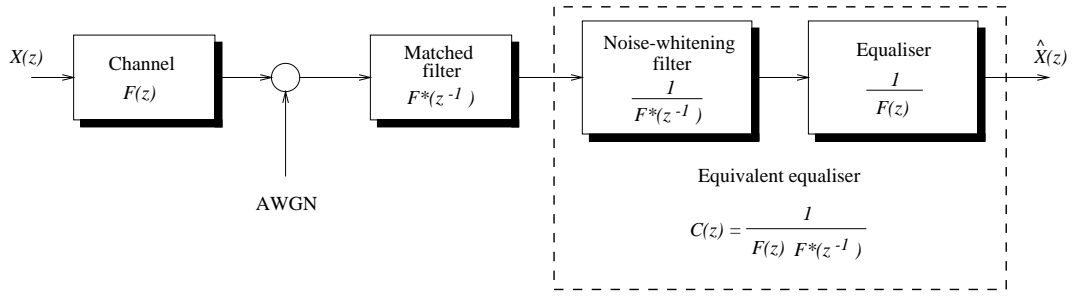


Figure 4.8: The optimum symbol-spaced linear equaliser is represented by the cascade of a matched filter, a sampler and a symbol-spaced transversal filter. The transversal filter represents the cascade of a noise whitening filter and a linear filter matched to the channel.

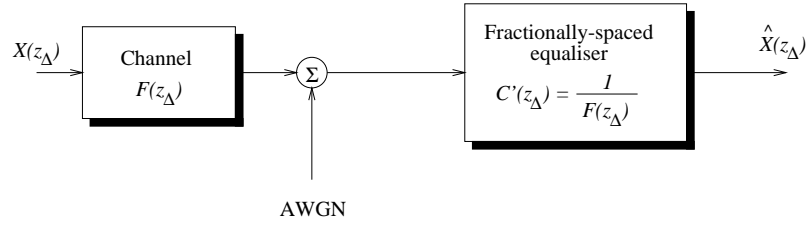


Figure 4.9: Equivalent fractionally-spaced equaliser. $F(z_\Delta)$ represents the transfer function evaluated at the fractionally-spaced sampling period.

$F^*(z^{-1})$ in the AWGN process, and the transversal filter will remove the ISI caused by the channel $F(z)$. However, since the FSE can compensate for the noise in the received signal in the whole of the channel bandwidth, the optimum FSE is equivalent to the cascade of the matched filter and a symbol-spaced linear equaliser.

The maximum-likelihood sequence estimator (MLSE) is the optimum receiver, if it is preceded by a matched filter and a sampler [29]. Ideally, the matched filter should be matched to the combined transmitter filter and propagation channel impulse response, but in practice, because the propagation channel characteristics are not known, the matched filter is matched to the transmitter filter. As a consequence, the matched filter is kept fixed and the complexity of having to estimate the overall channel and update the matched filter continuously, due to variations in the propagation channel, is avoided. The cascade of the transmitter filter $p(n)$, propagation channel $g(n)$ and the matched filter $p^*(-n)$ can be expressed as a single filter with impulse response $h(n)$. If the received signal after the receiver filter is sampled at the symbol interval T , no whitening filter is required because the receiver filter, which is a raised-cosine spectrum type filter, does not introduce ISI at the symbol intervals. However, if as shown in Figure 4.10, the received signal is sampled at $T/2$, the subsequent samples of the noise will be correlated, and the appropriate noise whitening filter can be built, because the receiver filter is known.

The coefficients of the whitening filter can be obtained using:

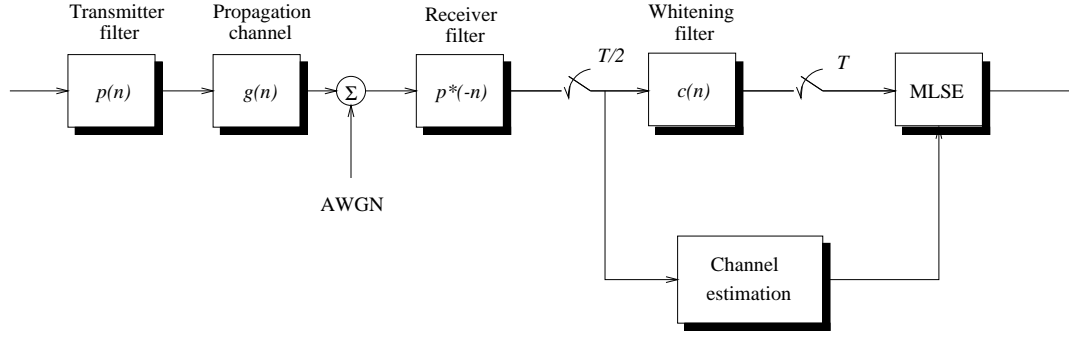


Figure 4.10: Maximum-likelihood sequence estimation in a receiver where the received signal after the matched-filter is sampled at twice the symbol rate.

$$\sum_{j=-K}^0 \psi_{nj} c_j = p(n), \quad n = -K_1, \dots, -1, 0, \quad (4.21)$$

where

$$\psi_{nj} = \sum_{m=0}^{-n} p^*(m) p(m+n-j), \quad n, j = -K_1, \dots, -1, 0. \quad (4.22)$$

The coefficients of the MLSE can then be obtained using [71]:

$$q_k = - \sum_{j=-K_1}^0 c_j \hat{h}(k-j), \quad k = 1, 2, \dots, K_2 \quad (4.23)$$

Fractionally-spaced equalisation algorithms using cyclostationary blind techniques

Fractionally-spaced blind equalisation algorithms lead implicitly to a FSE structure. In most cases, they produce elegant solutions of a FSE. The FSE solution for the three cyclostationary blind equalisation algorithms described in Chapter 3 are summarised next.

Hatzinakos [45] The cyclic statistics method of Hatzinakos gives the weights of a FSE, provided that the differential cepstrum parameters are correctly estimated using equation 3.46. Since $A(k)$ expresses the differential cepstrum components of the minimum phase part and $B(k)$ of the maximum phase part, the overall T/P -spaced transversal filter is given as the convolution of a minimum phase inverse filter $i_{inv}(k)$ and a maximum phase inverse filter $o_{inv}(k)$, i.e.:

$$c(k) = i_{inv}(k) \otimes o_{inv}(k), \quad (4.24)$$

where,

$$i_{inv}(k) = -\frac{1}{k} \sum_{n=2}^{k+1} [-A(n-1)] i_{inv}(k-n+1), \quad k = 1, \dots, N_1, \quad (4.25)$$

$$o_{inv}(k) = \frac{1}{k} \sum_{n=k+1}^0 [-B(1-n)] o_{inv}(k-n+1), \quad k = -1, \dots, -N_2. \quad (4.26)$$

ZF can then be applied with the filter coefficients obtained using (4.24), as explained in Chapter 2.

Moulines et al. [61] In [61], equalisation is carried out by a multichannel equaliser equivalent to a FSE [92], expressed by a $(M+N) \times PN$ equalisation matrix of the form

$$\Gamma = (\mathcal{H}_N^H \mathcal{H}_N)^{-1} \mathcal{H}_N^H. \quad (4.27)$$

Each row of matrix Γ contains a linear equaliser of length PN and each channel is filtered by a N tap filter, and the P resulting outputs are then added. The sequence of source information symbols can then be recovered by means of inverse filtering:

$$\hat{\underline{x}}_k = \Gamma \underline{y}_k \quad \text{for } k = 0, \dots, L-1, \quad (4.28)$$

where L is the number of input symbols available for equalisation.

Tong et al. [87] The solution to the channel identification problem leads directly to the estimation of the sequence of transmitted sequence $x(n)$. Recall equation 3.65 where the block-Toeplitz filtering matrix $\mathcal{H}(K)$ was identified. The information symbols can then be recovered using inverse filtering:

$$\hat{\underline{x}}'_k = \mathcal{H}(K)^\dagger \underline{y}'_k \quad \text{for } k = 0, \dots, L-1, \quad (4.29)$$

where the superscript \dagger denotes the pseudo-inverse, $\hat{\underline{x}}'_k = [\hat{x}(k-K-M+1), \dots, \hat{x}(k)]^T$ is the $(K+M) \times 1$ estimated input symbol vector, and $\underline{y}'_k = [y^{(0)}(k-K+1), \dots, y^{(P-1)}(k-K+1), \dots, y^{(0)}(k), \dots, y^{(P-1)}(k)]^T$ is the $KP \times 1$ received signal vector.

Figure 4.11 shows the results of the equalisation process of channel CH1. The number of symbols used in the estimation of the channel was 200, and the oversampling factor $P = 3$ for both multichannel methods and $P = 5$ for the cyclic algorithm. In accordance with the channel estimation results, the subspace method of Moulines *et al.* [61] gives the best performance, followed by the cyclic algorithm. It should be noted, however, that the oversampling factor P was carefully chosen not to be 3, because of the identification problems of the cyclic algorithm reported earlier.

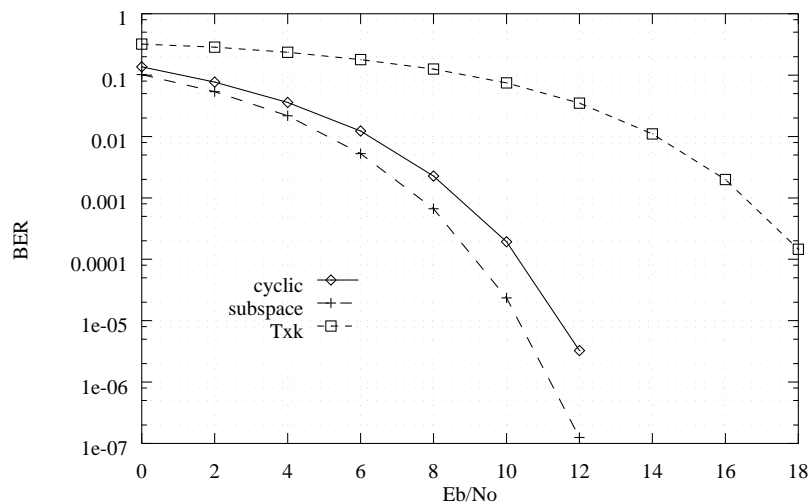


Figure 4.11: Bit-error-rate performance of fractionally-spaced linear equalisation, using cyclic, subspace and TXK blind equalisation methods, for channel CH1.

Optimum delay estimation

As shown in Chapter 2, the performance of either the symbol-spaced linear equaliser or the MLSE is dependent on the delay at which decisions are made at the receiver. If a ZF criterion is used in the linear equaliser, a certain delay needs to be associated with the inverse filter, so that the mixed-phase channel might be invertible. The value of this delay becomes critical in linear equalisation. If a particular delay has been used in the calculation of the inverse filter, the decisions must be made using the same delay; otherwise, they lead to incorrect decisions, as shown in Figure 4.12. Although the delay is not critical in MLSE, by increasing the delay in making the decisions, the overall performance will improve. This way, the probability that the M^L surviving sequences will agree in their earliest symbol, and the probability of making a correct decision will increase. In general, it has been found that the MLSE achieves an acceptable performance if the delay is set to $5L$ [71], where L is the duration of the multipath channel. Figure 4.12 shows the effect of increasing the delay in the decisions for two channels with transfer functions *a*) $H_1(z) = -0.0668 + 0.1314z^{-1} + 0.5198z^{-2} - 0.1612z^{-3} + 0.0744z^{-4} - 0.0465z^{-5}$ and *b*) $H_2(z) = -0.0065 + 0.0103z^{-1} + 0.5366z^{-2} - 0.4171z^{-3} + 0.0197z^{-4} - 0.0098z^{-5}$ for a SNR=5 dB scenario.

Nevertheless, it is interesting to note that both fractionally-spaced equalisation solutions given by the two algebraic methods of Moulines *et al.* [61] and Tong *et al.* [88], do not introduce any delay in the decisions if the associated subchannels are minimum phase. The pseudo-inverse of the respective channel filtering matrices has a set of P multichannel symbol spaced equalisers [92] and each multichannel equaliser attempts to remove the distortion caused by the corresponding subchannel. If the transfer functions of the subchannels are minimum phase, their inverses are physically realisable and no delay might be required.

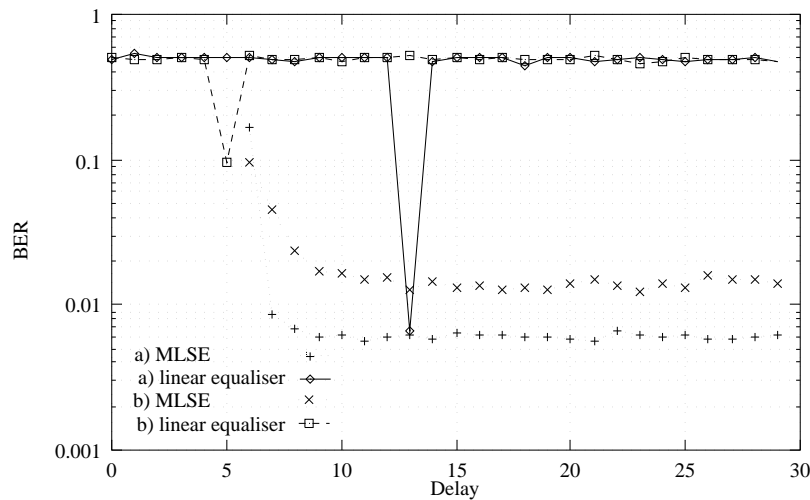


Figure 4.12: Bit-error-rate performance of symbol-spaced linear equaliser and MLSE for different delays used in the decisions.

4.4 Performance Comparison of Blind Channel Identification Techniques versus Supervised Methods in a Mobile Radio Environment

In this section, some numerical results are presented on the convergence and tracking properties of cyclostationary blind equalisation techniques in a mobile radio environment. The objective is also to study how this convergence compares to conventional supervised methods in time-varying environments. It has been shown in the previous section, that the performance of the subspace method of Moulines *et al.* [61], both in terms of channel estimation error and stability of the algorithm, is representative of cyclostationary blind channel identification methods. This method was selected for the simulations.

4.4.1 Adaptive and non-adaptive supervised channel estimation methods

The GSM system described earlier in this chapter is especially designed for supervised equalisation methods, where a dedicated training period is allocated for the training of the equaliser. Several supervised channel estimation techniques are available in the literature [47] [13]. This section describes two standard supervised approaches:

Supervised non-adaptive method

The supervised non-adaptive channel estimation method is widely used in GSM systems [13]. Following the structure of a GSM time slot shown in Figure 4.1, the data from one GSM time

slot are read first from left to right, and an estimate of the channel is carried out using a *least-squares* (LS) estimation technique. This channel estimate is made available to the MLSE as shown in Figure 4.10. The MLSE will use it to equalise the burst of 58 information data bits situated at the right hand side of the training data. The channel estimate is kept fixed during the equalisation of the 58 data bits; thus, it is understandable that the probability of error in those bits which are further away from the training burst would be higher than those which are closer [13]. In fact, if the Doppler frequency is high, more errors would be expected in these positions.

The equalisation of the first burst of 58 information data bits is carried out next, by reading the training bits from right to left, and obtaining a new channel estimate using LS technique again. The channel estimate is passed to the MLSE and kept fixed during the equalisation of the data bits on the left hand side of the training bits.

The linear LS technique [47] for channel estimation uses only the 26 training bits to produce a reliable channel estimate under the assumption that the channel remains time-invariant during those 26 bits. An error signal $e(n)$ can be formed as the difference between the noise corrupted received signal $y(n)$, sampled at the symbol rate, and the reconstructed received signal obtained through the convolution of the training signal and an *estimate* of the channel,

$$e(n) = y(n) - \hat{\underline{h}}_n^T \underline{x}_n, \quad (4.30)$$

where $\hat{\underline{h}}_n = [\hat{h}_n(0), \dots, \hat{h}_n(M)]^T$ denotes the estimate of the overall channel impulse response at time n , and $M + 1$ is the length of the symbol-spaced channel impulse response. On the other hand, $\underline{x}_n = [x(k), \dots, x(k - M)]^T$ denotes the sequence of known transmitted symbols at time n .

Multiplying (4.30) by the conjugate of the input signal and taking the expectation on both sides:

$$E \{e_n \underline{x}_n^*\} = E \{y(n) \underline{x}_n^*\} - \hat{\underline{h}}_n^T E \{\underline{x}_n \underline{x}_n^*\}. \quad (4.31)$$

The error signal in (4.30) represents the additive Gaussian noise present in the received signal; because the noise is orthogonal to the input signal \underline{x}_n , $E\{e_n \underline{x}_n^*\} = 0$ and (4.31) reduces to [64]:

$$R_{yx} = \hat{\underline{h}}_n^T R_{xx}, \quad (4.32)$$

and

$$\hat{\underline{h}}_n^T = R_{yx} R_{xx}^\dagger, \quad (4.33)$$

where the superscript \dagger denotes pseudo-inverse. The estimate of the channel is then obtained using the correlation of the input sequence, available during the training period, and the cross-correlation between the received signal and the training sequence. It should be noted, however, that the channel estimate that is obtained this way is sub-optimum, in the sense that it tends to reduce the error signal to zero. Other adaptive techniques can be used during the training period. The *least-mean squares* (LMS) and *recursive least-squares* (RLS) methods [47] are two examples. The problem of LMS is that 26 bits might not be enough for the algorithm to converge and the RLS technique, on the other hand, shows a faster convergence, but at the expense of more computational complexity.

Supervised adaptive method

The supervised adaptive method is essentially the same as the non-adaptive technique, with the difference that the channel estimates that were kept fixed during the equalisation of the two bursts of information data, are now constantly updated using a LMS technique. The LMS technique is a stochastic gradient adaptive filtering algorithm [13] which is formulated as:

$$\hat{\underline{h}}_{n-d+1} = \hat{\underline{h}}_{n-d} + \mu e(n) \hat{\underline{x}}_{n-d}^* \quad (4.34)$$

where $\hat{\underline{h}}_{n-d}$ and $\hat{\underline{h}}_{n-d+1}$ are the estimates of the channel impulse response at time $n-d$ and $n-d+1$ respectively; $\hat{\underline{x}}_{n-d}$ is the vector of the estimated transmitted symbols, and $e(n)$ is the error signal defined as:

$$e(n) = y(n-d) - \hat{\underline{h}}_{n-d}^T \hat{\underline{x}}_{n-d}. \quad (4.35)$$

Since the decisions in the MLSE receiver are delayed by d , and these decisions are used as the *desired signal* to create the error signal, only a channel estimate at time $n-d+1$ is available. Increasing the delay d improves the performance of the MLSE, but the channel estimates obtained using the equaliser's own decisions will not correspond to the actual channel; this will lead to a degradation in the performance of the LMS channel estimation/MLSE receiver structure. This problem is accentuated when the Doppler frequency of the propagation channel are high.

The step-size parameter μ controls the convergence speed of the algorithm and stability. For high μ , although a high convergence speed is achieved, the LMS algorithm can diverge. Thus, a compromise needs to be reached; in most cases an optimum step-size parameter can be obtained [47].

The LMS technique is one of the standard channel estimation techniques and although it does not constitute an optimum channel estimation procedure, it is very simple. Other techniques,

such as the RLS technique, can produce a better performance. The *minimum-survivor* LMS and RLS techniques proposed in [13] for a MLSE receiver structure have shown a similar performance of the standard LMS and RLS techniques but with a simpler structure.

4.4.2 Blind channel estimation method

In this section, the procedure for blind channel estimation in a mobile radio environment is described. The simulations are conducted using the subspace algorithm for blind channel identification proposed by Moulines *et al.* [61]. As shown in the previous section, its performance in a stationary channel environment is good for a short data length used in the estimation of the channel (200 symbols). Due to the computational complexity of the algorithm compared to the conventional LS method or similar, the channel estimate is updated only once every time slot. Channel estimation can be classified according to the type of observation window employed. The observation window contains a number of received samples which are used in the estimation of the channel. The length of the observation window is variable in order to test the performance of the blind channel estimation techniques for different lengths of data used in the estimation. Two window types were used in the simulations:

- The exponential window is used in applications such as the *recursive least-squares* (RLS) algorithm, where a weighting factor λ is introduced [47] in order to give more weight to the actual values of the estimation, and less to past values. This window type can be useful especially in the case of time-varying channels. The weighted autocorrelation function is defined as:

$$R_{yy}(k) = \sum_{n=0}^k \underline{y}_k \underline{y}_k^H \lambda^{k-n}, \quad (4.36)$$

where $R_{yy}(k)$ denotes the autocorrelation function of the sequence $y(k)$ at time k and λ is the *forgetting factor*. The values of λ range in $0 \leq \lambda \leq 1$. Typical values of λ are those in the range $0.98 \leq \lambda \leq 1$.

- Another alternative is to use a rectangular window, where past and actual values of the estimates have the same weight. This approach is suitable when the variations in the simulated environment conditions are not high. The rectangular window constitutes an especial case of the exponential window for a forgetting factor of $\lambda = 1$.

4.4.3 Typical Urban (TU) environment

The *Typical Urban* (TU) environment characterised by the power-delay profile shown in Table 4.1, has a main reflector occurring very close to the main path, with more power than the *line-of-sight* path. The multipath effect extends to $5 \mu\text{s}$, with a bit period of $3.69 \mu\text{s}$. The combined

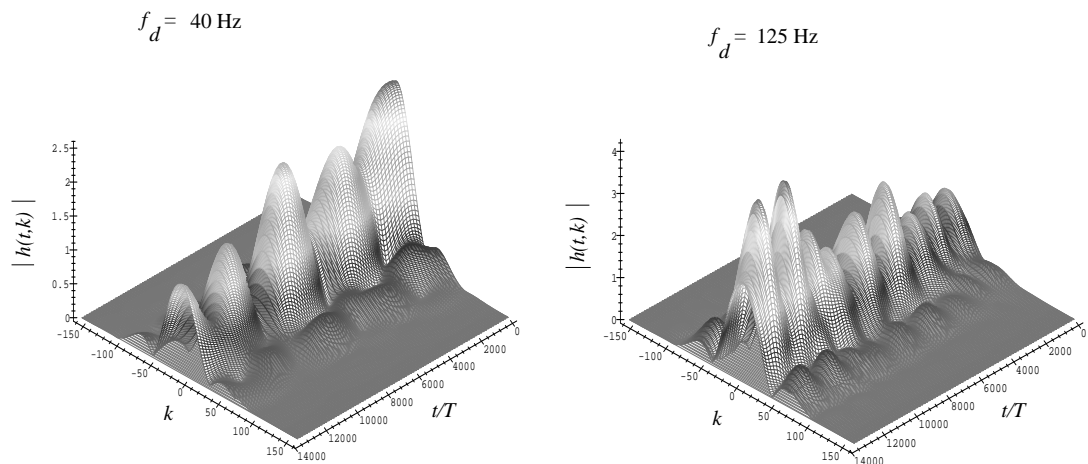


Figure 4.13: Magnitude of the time-variant impulse response of the combined transmitter and receiver filter with a Typical Urban propagation channel, for Doppler frequencies of $f_d = 40$ Hz and $f_d = 125$ Hz.

raised-cosine transmitter/receiver filters plus the time-varying propagation channel impulse response is shown in Figure 4.13 for Doppler frequencies of $f_d = 40$ Hz and $f_d = 125$ Hz.

Computer simulations were conducted to test the performance of blind equalisation in a TU environment, and to compare these with the results from conventional supervised adaptive and non-adaptive methods, using a MLSE in the receiver. The simulations were carried out using a linear 4-QAM modulation scheme. The complex data symbols $x(n) = \Re\{x(n)\} + i\Im\{x(n)\}$ take up values ± 1 with equal probability from the 4-QAM signal constellation. The real part $\Re\{x(n)\}$ and imaginary part $\Im\{x(n)\}$ are independent, identically distributed (i.i.d.) sequences. Two bits are transformed into one 4-QAM symbol, and the assignment of the values of the alphabet is done using differential Gray encoding [71].

Figure 4.14 shows the evolution of the errors along one burst of information data for a *energy-per-symbol-to-noise ratio* of $E_s/N_0 = 20$ dB. The E_s/N_0 measures the power of the noise for the energy required to transmit a single input symbol. The total energy transmitted is equivalent to the energy of one input symbol multiplied by the power of the channel. At 40 Hz Doppler frequency, the adaptive method can update the channel estimates along the burst. With the non-adaptive method, on the other hand, the probability of error is higher in those bits on the right hand side of the burst, which are further away from the training sequence. When the Doppler frequency is as high as 125 Hz, the increase in the number of errors in the bits on the right hand side is notorious. In fact, not even the adaptive method can *correct* the errors, because it cannot keep track of the variations in the channel characteristics.

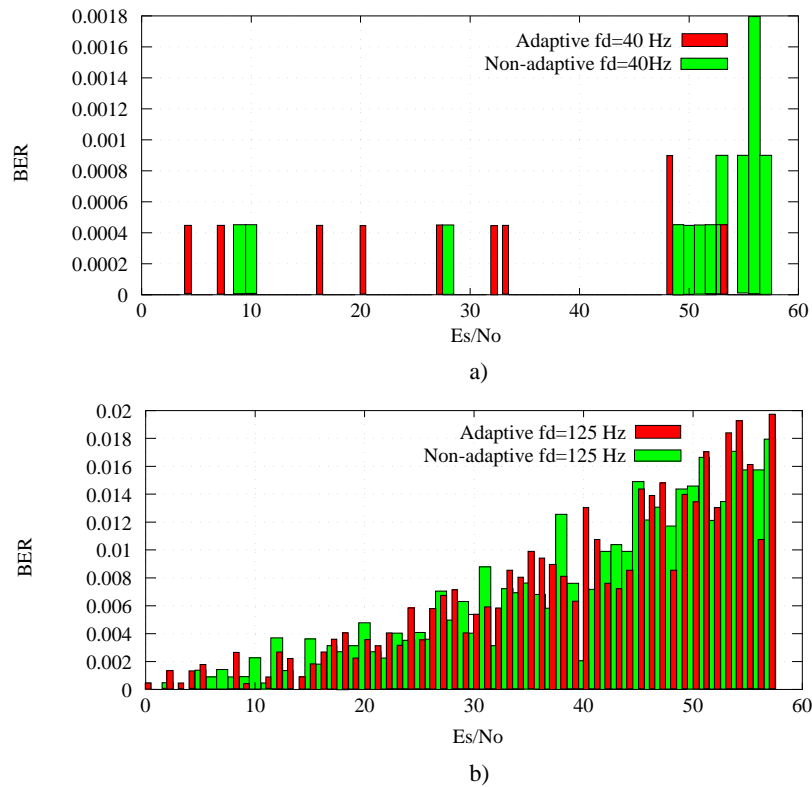


Figure 4.14: Probability of error in each bit position of a 58-bit GSM information data burst for Doppler spectra a) $f_d = 40$ Hz and b) $f_d = 125$ Hz in a $E_s/N_0 = 23$ dB scenario; statistics on 2225 bursts in a Typical Urban environment.

The performance of the subspace method for blind channel identification is shown in Figures 4.15 and 4.16 for Doppler frequencies of $f_d = 40$ Hz and $f_d = 125$ Hz respectively. The performance was measured in terms of the *bit-error-rate* (BER), for different lengths of data used in the estimation of the channel. The length of the observation window is a multiple of the length of one TDMA time-slot. Data lengths of 156 bits (78 4-QAM symbols), 312 bits (156 4-QAM symbols) and 624 bits (312 4-QAM symbols) were used. In each of the cases, a rectangular window and an exponential window ($\lambda = 0.98$) were used. In general, for a Doppler frequency of $f_d = 40$ Hz, increasing the length of the observation window improves the performance. The assumption that the channel remains quasi-stationary during the length of the observation window is weak, especially at high Doppler frequencies; that is why at a Doppler frequency of 125 Hz, the improvement from using a larger observation window is almost non-existing. The channel estimate obtained this way is the average of the channel during the duration of the observation window.

In Figure 4.15, it is also observed that using an exponential window improves the overall performance. However, this is true for short data lengths (1 slot and 2 slots); for 4 slots used in the estimation of the channel, a slight degradation is observed. The results shown in Figures 4.15 and 4.16 suggest that 78 4-QAM symbols (1 slot) are insufficient to estimate the channel correctly, because even increasing the length of the observation window improves the performance, although the assumption of quasi-stationarity is weaker.

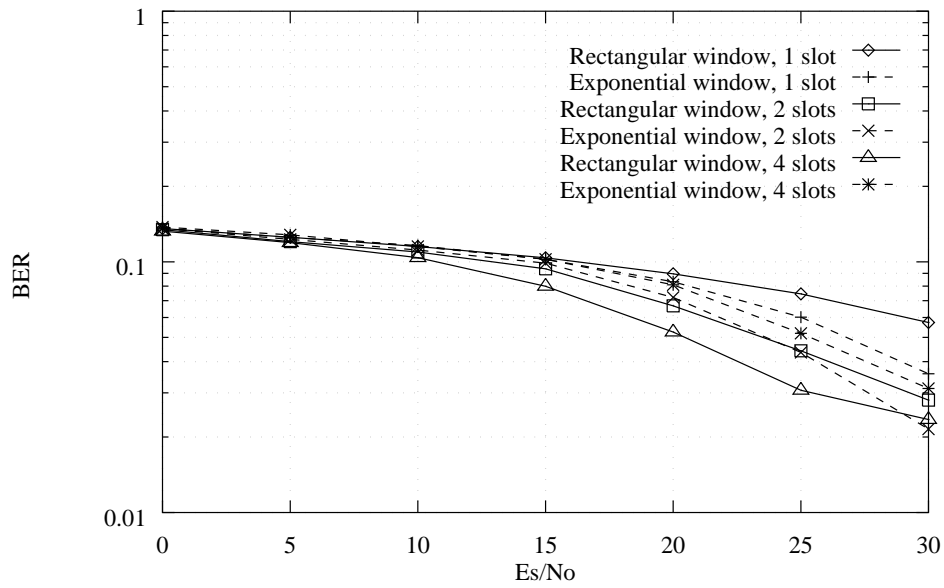


Figure 4.15: BER performance of the subspace method for blind channel identification combined with a MLSE receiver in a Typical Urban environment. The Doppler frequency is $f_d = 40$ Hz and a 4-QAM signal has been used.

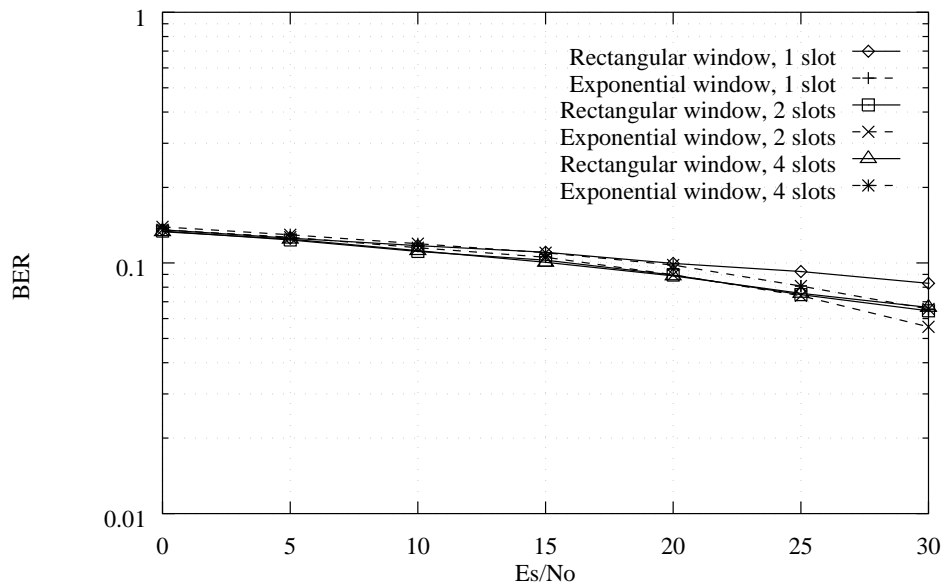


Figure 4.16: BER performance of the subspace method for blind channel identification combined with a MLSE receiver in a Typical Urban environment. The Doppler frequency is $f_d = 125$ Hz and a 4-QAM signal has been used.

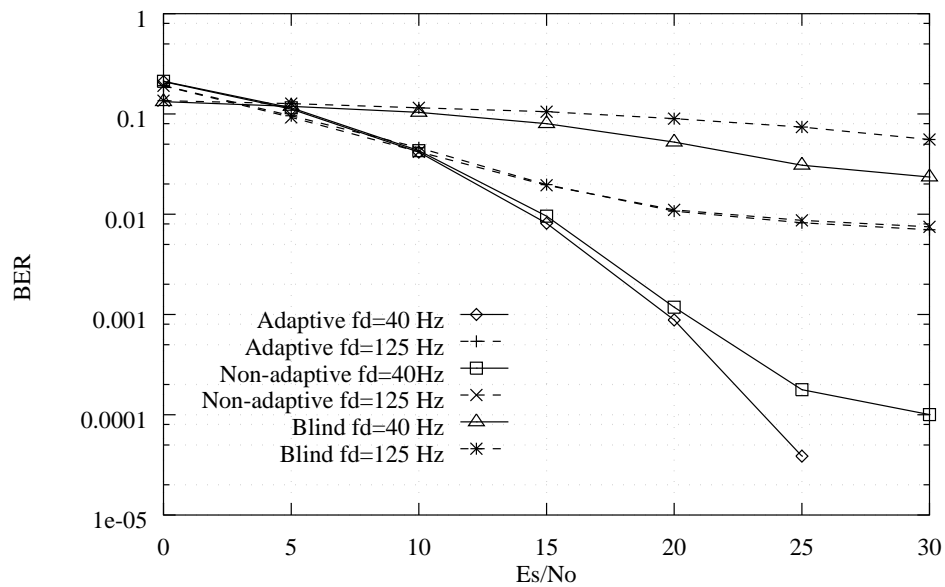


Figure 4.17: BER performance of supervised adaptive and non-adaptive methods, and blind channel estimation method in a Typical Urban environment. The Doppler frequencies are $f_d = 40$ Hz and $f_d = 125$ Hz and a 4-QAM signal has been used.

The performance of both supervised methods is shown in Figure 4.17. As it can be seen, the performance of the adaptive supervised method improves as the E_s/N_0 increases. The reason is that the LMS algorithm can keep track of the channel variations, and proper adjustments can be made to the coefficients of the MLSE. When the E_s/N_0 is low or when the Doppler frequency is high, the adaptive method does not provide performance gain over the non-adaptive method. The step-size parameters used were $\mu = 0.02$ and $\mu = 0.04$ for Doppler frequencies of 40 Hz and 125 Hz respectively. In Figure 4.17, for Doppler frequencies of 40 Hz and 125 Hz, it is observed that the performance of the blind channel estimation method is particularly bad with respect to methods using training. Similarly to what happens with the supervised methods for $f_d = 125$ Hz, the BER cannot be reduced any further for both $f_d = 40$ Hz and $f_d = 125$ Hz. In fact, it is said that an irreducible BER performance pattern is reached.

The issue of the delay associated with the MLSE has been discussed both in Chapter 2 and in this chapter. It is known that the performance of the MLSE will improve as the delay associated with the decisions increases. This does not constitute a problem to blind channel estimation methods and supervised non-adaptive methods; in the case of adaptive channel estimation, however, because the adjustment of the channel estimate is done using the decisions made by the MLSE, the bigger the delay, the more old-fashioned the estimate of the channel that will be obtained; consequently, the performance of the equaliser will degrade. Figure 4.18 shows the effect that an increase in the delay has in the overall performance of each system for a $E_s/N_0 = 15$ dB scenario. The performance of the blind channel estimation does not improve but it does not degrade. The same happens to the supervised non-adaptive method. In fact, a slight improvement is shown between delays of 35 and 40 symbols.

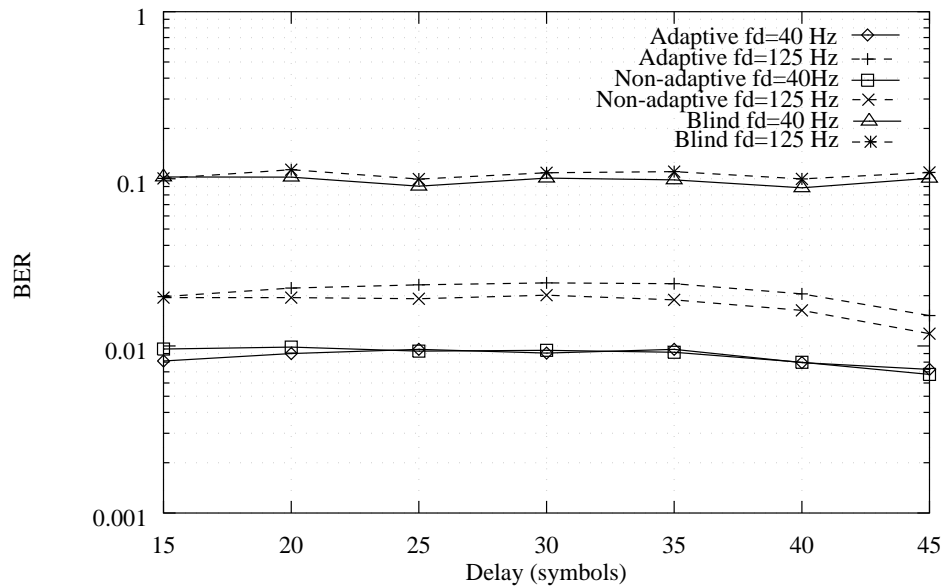


Figure 4.18: BER performance of supervised adaptive and non-adaptive methods, and blind channel estimation method in a Typical Urban environment for different delays used in the MLSE.

On the other hand, the supervised adaptive method begins to degrade slightly between delays of 15 and 35. It should be kept in mind that the first channel estimate produced in the supervised adaptive method is by the non-adaptive LS technique. Therefore, in the region where the performance of the LS technique is stable (15 to 35 symbols), a degradation is observed in the performance of the adaptive method, but from 35 to 45, the overall performance improves due to the improvement of the initial LS estimation.

Another simulation was conducted using a real-valued *binary-phase shift keying* (BPSK) signal. The transmission rate is double the transmission-rate of the 4-QAM signal, because in the case of the 4-QAM signal, two bits can be transmitted in one symbol in two orthogonal real and imaginary components. However, the 4-QAM signal has a double-sided power spectrum due to the in-phase and quadrature components. The transmitter filter in 4-QAM was matched to a symbol period of ($T = 7.4 \mu\text{s}$), but because it is transmitted in both sides of the spectral band, the total bandwidth is the same as the case of the BPSK signal. Consequently, the *transmission-rate-to-bandwidth ratio* (R/W) is double in the case of the BPSK signal.

The complex-valued 4-QAM signal and the real-valued BPSK signal, having different symbol periods, $T = 7.4 \mu\text{s}$ and $T = 3.7 \mu\text{s}$ respectively, have an effect in the way in which each of the symbols are spread by the propagation channel. The particular delay spread profile of the TU environment, shown in Table 4.1, indicates that in the case of the complex-valued 4-QAM signal, with double the symbol period compared to the real-valued BPSK signal, the delay-spread of the propagation channel does not extend to more than one symbol. Figure 4.19 shows the magnitude of the combined transmitter filter, propagation channel and receiver filter. Although in the case of the BPSK signal the reflections of the propagation channel extend to more than one symbol period $T = 3.7 \mu\text{s}$, the power of the reflectors is small and has little effect

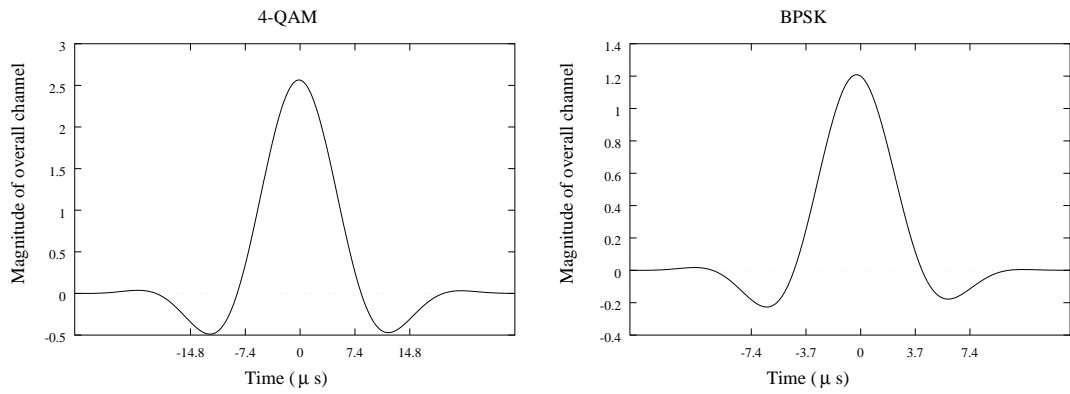


Figure 4.19: Impulse response of the combined transmitter and receiver filter and a time-invariant channel with Typical Urban environment tap-delay profile.

on the overall channel. In fact, the channels sampled at their respective symbol periods look very similar. At most, more ISI is expected in the case of the BPSK signal, as the delay-spread of the propagation channel extends beyond one symbol period.

Another thing to be considered in this comparison is the length of the MLSE. This parameter will determine the computational complexity of the receiver and the effect on the removal of the ISI. The more coefficients that are taken into account, the better the removal of ISI, but at the expense of more computational complexity. At each stage, the MLSE has to compute M^q metrics, where M is the alphabet of the signal, and q is the length of the channel. A channel length of 3 seems to be appropriate for a TU environment. For the 4-QAM signal, the number of states to be computed by the Viterbi algorithm at each stage is $4^3 = 64$. Ideally, a length of 5 would be convenient, but that would require the computation of $4^5 = 1024$ states. In the case of the BPSK signal, a channel length of 5 would not constitute a serious drawback, because the number of states that need to be calculated are $2^5 = 32$. In [13], a channel length of 5 is used for a TU environment for a GMSK signal. Nevertheless, for the sake of the comparison with the 4-QAM signal, a channel length of 3 was used in the MLSE for the BPSK signal.

The performance of the subspace method for blind channel estimation using a BPSK signal is shown in Figure 4.20 for a Doppler frequency of 40 Hz and in Figure 4.21 for a Doppler frequency of 125 Hz. Data lengths equivalent to one time-slot (156 bits) and two time-slots (312 bits) were used, and rectangular and exponential windows were tested. It can be noted that the rectangular window, using data from two time-slots, improves the overall performance, but similarly to the 4-QAM case, for an exponential window, increasing the window length results in a slight degradation in the performance. Interestingly, whereas the rectangular window with 2 time slots does not reduce the error significantly for a E_b/N_0 greater than 25 dB, the exponential window with data from one time slot improves progressively.

For a Doppler frequency of 125 Hz, the use of data from one time slot is beneficial, because of the large variations in the channel characteristics. Figure 4.22 shows the performance of both supervised adaptive and non-adaptive methods and compares them with the best performance

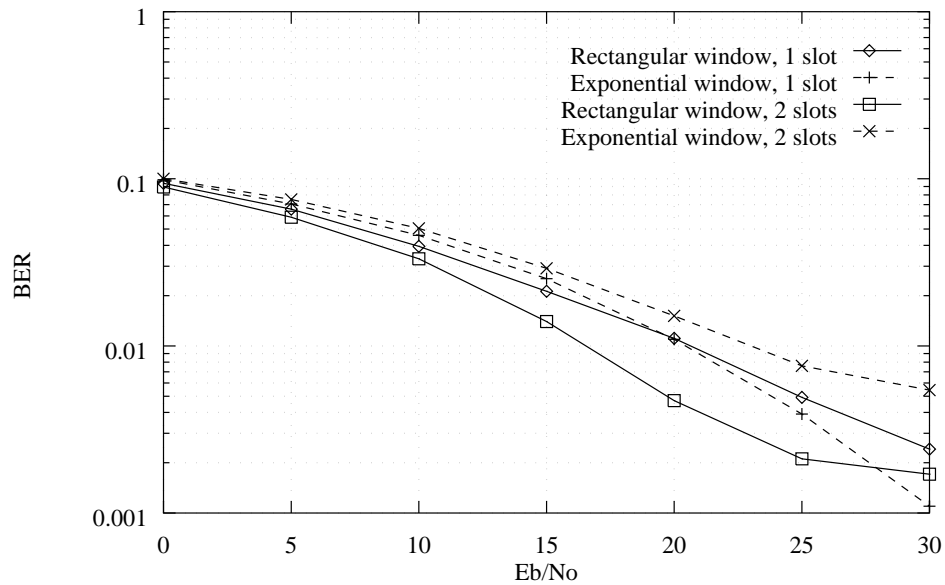


Figure 4.20: BER performance of the subspace method for blind channel identification combined with a MLSE receiver in a Typical Urban environment. The Doppler frequency is $f_d = 40$ Hz and a BPSK signal has been used.

from the blind method. The adaptive method with a step-size parameter of $\mu = 0.02$ does not improve the performance of the non-adaptive method considerably, but at $f_d = 125$ Hz, the adaptive method with step-size $\mu = 0.04$ reduces the error considerably for values of E_b/N_0 greater than 15 dB. It is interesting to note that although the blind method performs poorly, the difference is smaller than in the previous example, where a 4-QAM signal has been used. In fact, for $f_d = 125$ Hz, at high E_b/N_0 s (30 dB), the BER achieved by the supervised non-adaptive method and the blind method is 0.0025 and 0.0057 respectively.

The comparison of the results shown for the 4-QAM signal and the BPSK signal is not easy. However, it is believed that this comparison will highlight some singularities, particularly about the behaviour of the blind channel estimation method. These are the factors that differentiate the results of both simulations:

- Different symbol periods and different energy transmitted per symbol.
- Transmission in double-sided spectrum in 4-QAM, for single-sided in BPSK.
- Reflections of the propagation channel occurring beyond the symbol period for BPSK and within one symbol period for 4-QAM.
- Channel estimation carried out with twice as much real-valued data in the case of BPSK compared to 4-QAM.
- The same step-size parameter μ has been used in the supervised adaptive methods. $\mu = 0.02$ and $\mu = 0.04$ were optimised for the 4-QAM signal, but these are not necessarily the optimum for BPSK.

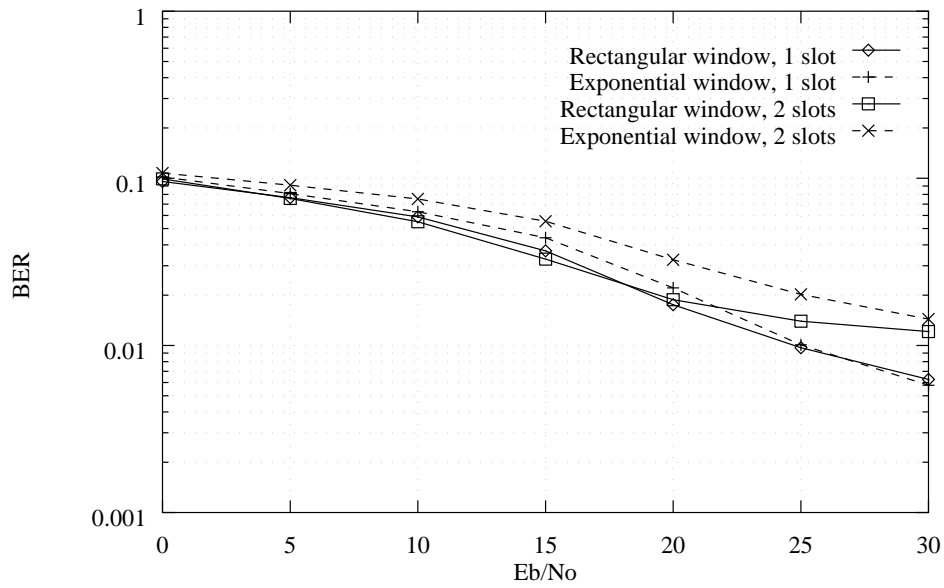


Figure 4.21: BER performance of the subspace method for blind channel identification combined with a MLSE receiver in a Typical Urban environment. The Doppler frequency is $f_d = 125$ Hz and a BPSK signal has been used.

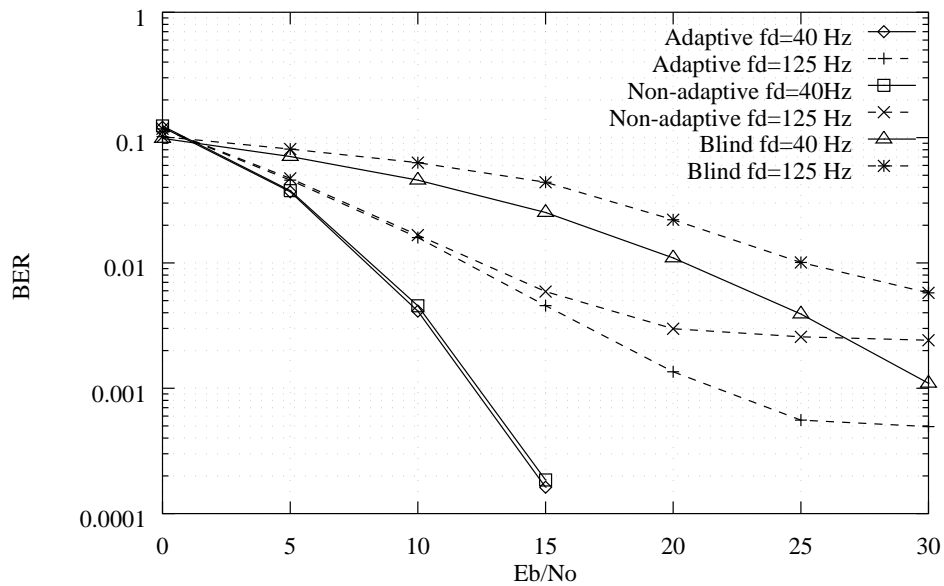


Figure 4.22: BER performance of supervised adaptive and non-adaptive methods, and blind channel estimation method in a Typical Urban environment. The Doppler frequencies are $f_d = 40$ Hz and $f_d = 125$ Hz and a BPSK signal has been used.

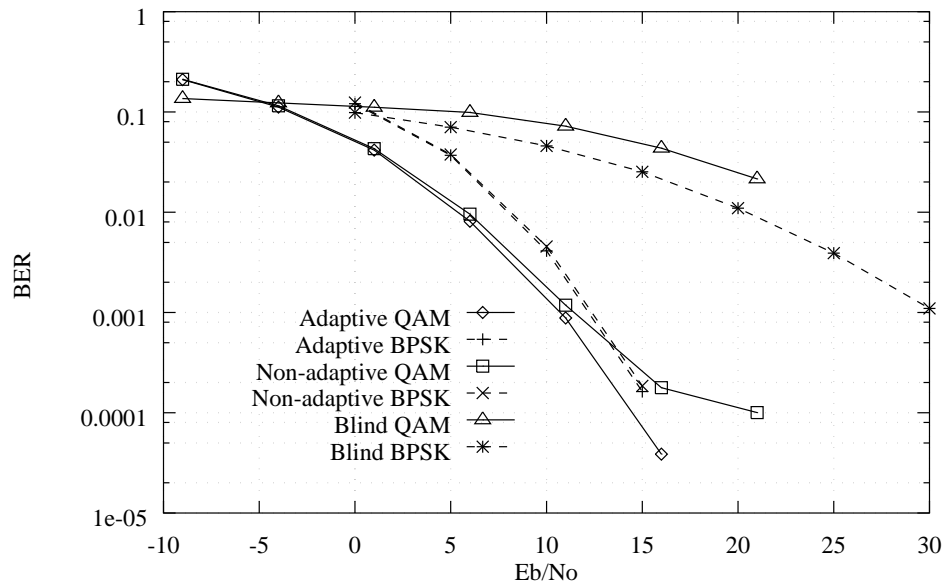


Figure 4.23: Comparison of the BER results obtained using BPSK and half-rate 4-QAM signals for $f_d = 40$ Hz.

- MLSE is carried out for two different signal alphabets, $M = 4$ and $M = 2$. The distance between the metrics is $d = \sqrt{2}$ and $d = 1$ for 4-QAM and BPSK respectively.
- The same channel length L has been considered in both cases in the MLSE. For the same number of computations, a bigger channel length could have been considered for BPSK.

Figure 4.23 and 4.24 show the comparison of the performance of both systems taking into account the above factors and after proper normalisation of the power. Although it may seem that due to the list of factors that differentiate both systems these cannot be compared, the results shown in Figures 4.23 and 4.24 lead to some conclusions. In the case of the supervised methods, the 4-QAM system performs slightly better than with the BPSK signal, but the difference reduces as the power of the noise diminishes. It is interesting to note that the blind method performs better for BPSK. This can only be understood by the fact that on the same time-duration, double the data are available for the estimation of the autocorrelation of the same size.

Some of the factors that contribute in the comparison can be made comparable, especially that concerning the propagation channel, by considering a 4-QAM system, with double the transmission rate of the simulation carried out earlier for the same signal alphabet. Effectively, the reflections now happen at the same time instants as in the BPSK case.

The results of blind channel estimation of Doppler frequencies of 40 Hz and 125 Hz are shown in Figures 4.25 and 4.26. These are also compared to the supervised non-adaptive and adaptive channel estimation methods in Figure 4.27. It is seen again that blind equalisation approximates to the performance of the supervised non-adaptive technique for $f_d = 125$ Hz, without using a training sequence.

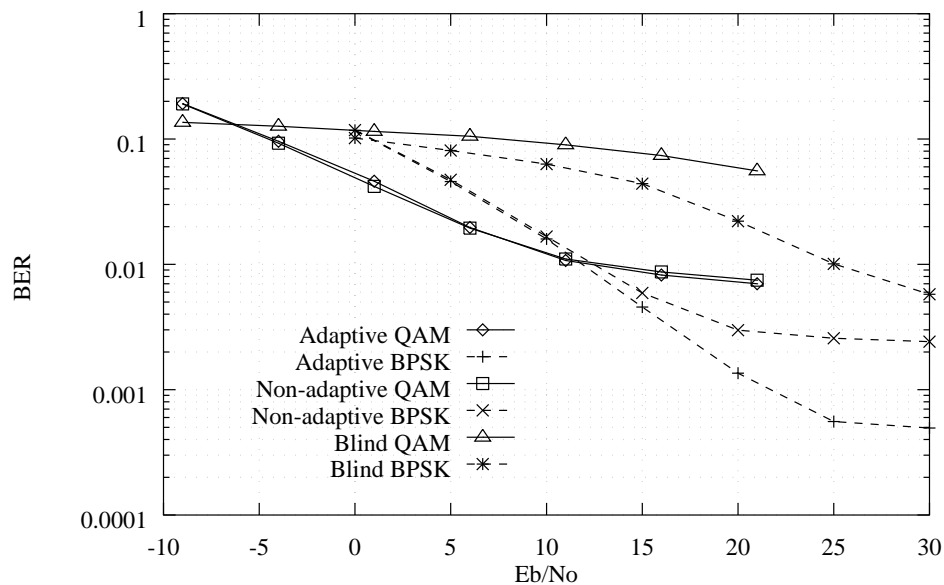


Figure 4.24: Comparison of the BER results obtained using BPSK and half-rate 4-QAM signals for $f_d = 125$ Hz.

After appropriate normalisation of E_b/N_0 measures, the results of Figure 4.27 for the double-rate 4-QAM signal and those of Figure 4.22 for BPSK are compared in Figures 4.28 and 4.29, for $f_d = 40$ Hz and $f_d = 125$ Hz respectively. The results for Doppler frequency 40 Hz reveal that the performance is very similar in both cases. This suggests, that, the amount of data available for estimation, as well as the time-instants that channel reflections occur, are determining factors.

4.5 Discussion

The *quality of service* (QoS) in a GSM system is provided by the radio link control. When due to a poor performance of the receiver results in an unacceptable voice or data quality, a radio link failure occurs. The reasons for radio link failure may be loss of radio coverage or very high interference levels. The radio link control measures the power of the received signal, the quality of the signal and the absolute distance between the base station and the mobile. In the analysis presented in this chapter, the power control and absolute distance are not taken into account, as the main concern is the quality of the received signal. The quality of the signal is measured in terms of the BER achieved by the MLSE. Table 4.5 shows the different quality levels specified in GSM for GMSK modulation and using coding [83]. The lower bound in terms of quality is marked by BER=0.002. The fact that a longer channel length can be used in the MLSE receiver, combined with the fact that *forward error correction* (FEC) coding can be used in conjunction with equalisation, can reduce the BER performance [12]. The question that remains to be answered is whether blind equalisation can be used in mobile radio channels. The answer is that it depends. At low Doppler frequencies, the supervised methods perform considerably better than the blind method. At high Doppler frequencies, blind equalisation might be considered. The QoS figures established for GSM suggest that cyclostationary blind

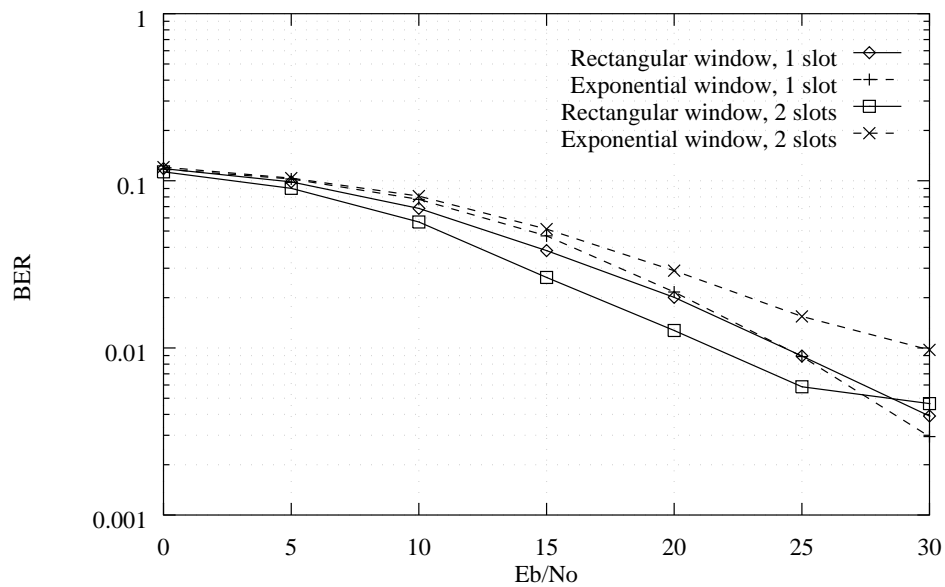


Figure 4.25: BER performance of the subspace method for blind channel identification combined with a MLSE receiver in a Typical Urban environment. The Doppler frequency is $f_d = 40$ Hz and a double-rate 4-QAM signal has been used.

equalisation methods might not be adequate to be used on their own. However, the fact that a BER performance of 10^{-2} is expected in order to open the eye of the equaliser in a digital radio application [30] indicates, that only under certain circumstances can a switch to decision-directed mode be guaranteed. Only at relatively high SNR scenarios the use of cyclostationary blind equalisers could provide a suitable alternative.

Quality	BER
0	< 0.2%
1	0.2% to 0.4%
2	0.4% to 0.8%
3	0.8% to 1.6%
4	1.6% to 3.2%
4	3.2% to 6.4%
4	6.4% to 12.8%
5	> 12.8%

Table 4.5: Received signal quality vs. channel bit error rate in GSM [83].

A problem associated with supervised adaptive channel estimation schemes is the propagation of channel errors. The information data are transmitted in time slots which are allocated to a different user and multiplexed in time. The training sequence divides the information data in 2 bursts. Since the training sequence is often not long enough to obtain an estimate of the channel using adaptive channel estimation techniques (i.e. LMS), if the channel estimate is constantly updated using an adaptive decision directed algorithm, there is the risk that at high interference conditions the channel error might expand. At the next training sequence, the LMS algorithm cannot *correct* the channel quickly enough and the errors continue to expand. At one point, the channel estimation has to be re-initialised. This problem is only associated to supervised adaptive channel estimation techniques such as LMS. However, in our simulations,

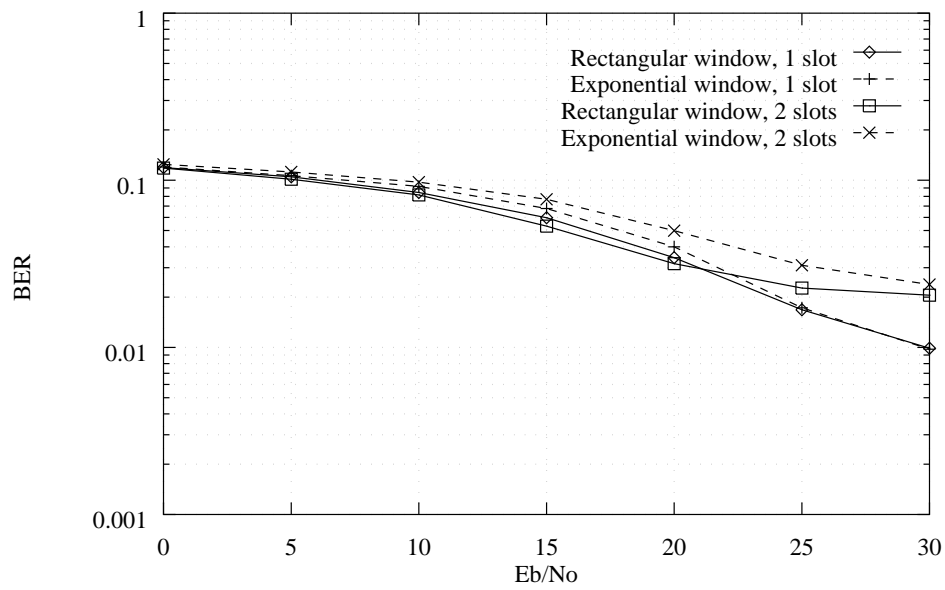


Figure 4.26: BER performance of the subspace method for blind channel identification combined with a MLSE receiver in a Typical Urban environment. The Doppler frequency is $f_d = 125$ Hz and a double-rate 4-QAM signal has been used.

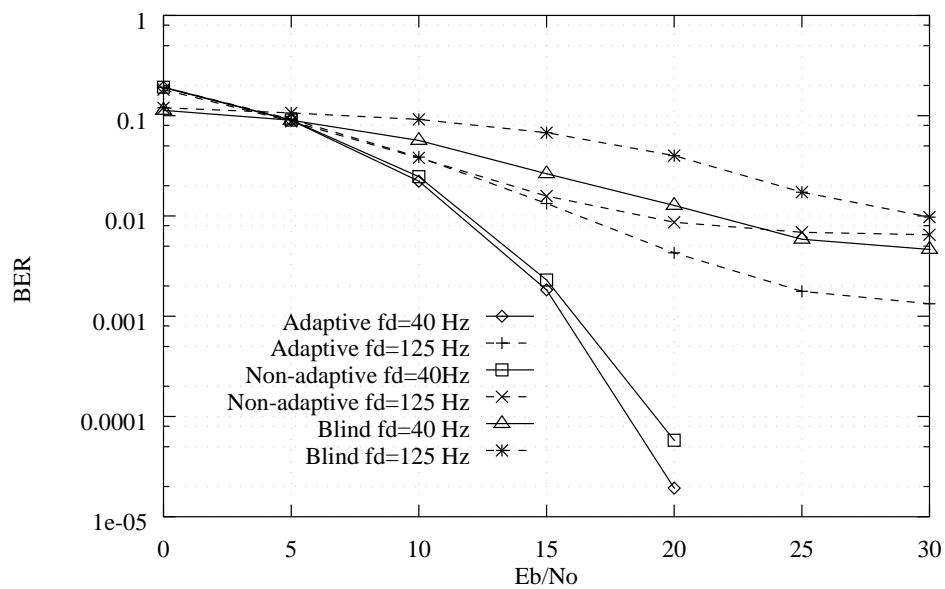


Figure 4.27: BER performance of supervised adaptive and non-adaptive methods, and blind channel estimation method in a Typical Urban environment. The Doppler frequencies are $f_d = 40$ Hz and $f_d = 125$ Hz and a double-rate 4-QAM signal has been used.

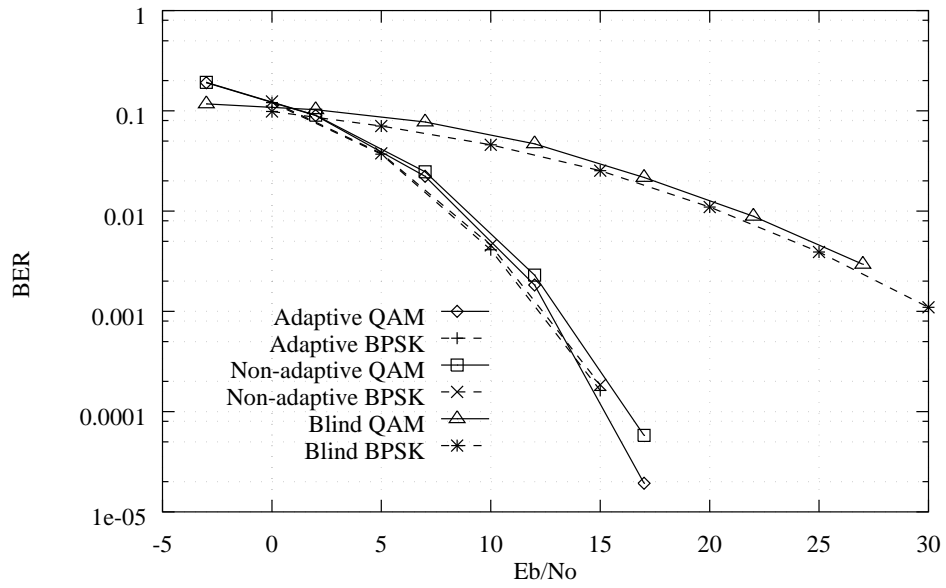


Figure 4.28: Comparison of the BER results obtained using BPSK and double-rate 4-QAM signals for $f_d = 40$ Hz.

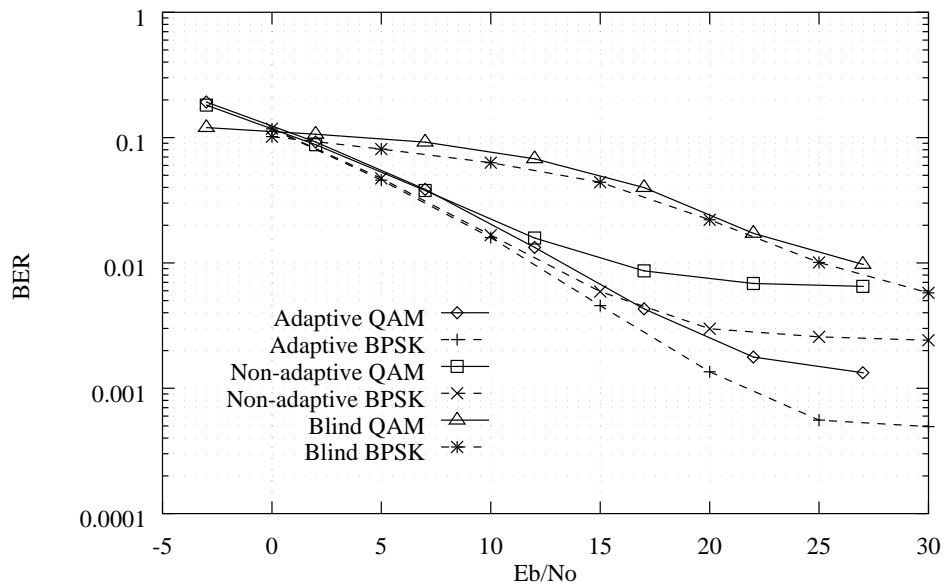


Figure 4.29: Comparison of the BER results obtained using BPSK and double-rate 4-QAM signals for $f_d = 125$ Hz.

a LS approach has been considered and a completely new channel estimate is obtained in every burst.

4.6 Conclusions

Most current cyclostationary blind equalisation techniques are not adequate for mobile communication channels that require channel identification and equalisation within 156 symbols. The convergence of algorithms such as [88] [61] [45] [81] lies around the few hundred symbols, for stationary channels. Simulation studies carried out in this chapter, in a GSM typical urban environment, indicate that in most cases, the performance of blind channel estimation is not close to the performance of supervised methods. However, as the variation in the channel characteristics increases, under high SNR scenarios the gap in the performance is not so important.

In this chapter, the fundamental identification conditions of cyclostationary blind equalisation methods have been tested. Although there are channels strictly unidentifiable using fractionally-spaced blind equalisation, these are rare, and in most of the cases can be avoided by carefully choosing the sampling frequency. Model order overestimation can lead to subchannels sharing zeros; as a consequence, an accurate model order estimation is crucial so that the channel can be identified.

It has been suggested that the GMSK nonlinear modulation used in GSM may not be used in most existing blind channel identification methods [18], because due to the memory of the nonlinear modulator, the input is not expressed as an independent, identically-distributed signal. This condition is fundamental, for example in [88], but it does not affect [61]. More importantly, the nonlinear nature of the modulator can represent a serious drawback to most channel estimation techniques.

The computational cost of cyclostationary blind equalisation methods plays an important role too. In the subspace method for blind channel estimation of Moulines *et al.* [61] used in these simulations, two eigenvalue-decompositions were needed for every channel estimate. This constitutes a serious drawback compared, for example, with either the supervised adaptive LMS algorithm or the supervised non-adaptive channel estimation method, where only the calculation of the cross-correlation vector was required.

A Cyclic Subspace Algorithm for Blind Channel Identification

5.1 Introduction

Motivated by the good performance of the subspace methods for blind channel identification and the potential of cyclostationary statistics, a new cyclic subspace algorithm combining both subspace methods and cyclostationary statistics is proposed in this chapter. Two of the multichannel blind channel identification algorithms studied in the previous chapter [87] [61] are derived from *direction-of-arrival* (DOA) estimation techniques; i.e. ESPRIT [76] and MUSIC [85] respectively. The new approach presented in this chapter is in the line of another DOA estimation technique known as the Cyclic MUSIC algorithm [78]. The Cyclic MUSIC algorithm is an extension of the conventional MUSIC algorithm to the cyclic statistics domain. It has been shown to improve the performance of the conventional MUSIC algorithm in terms of the estimation accuracy of the DOA for a desired user. The cyclic subspace method is the extension of the conventional subspace method for blind channel identification [61] to the cyclic statistics domain. The advantages are robustness against correlated noise and signal selectivity, because it enables to separate signals from users which are temporally and spectrally overlapping.

This chapter is organised as follows: First, the cyclic subspace algorithm for blind channel identification is introduced. In section 5.3, an asymptotic variance estimator is derived. Section 5.4 and section 5.5 are devoted to two properties of the algorithm: correlated noise compensation and blind multiuser detection. Section 5.6 presents some simulation results, and section 5.7 discusses these results and establishes the connections and differences of the cyclic subspace algorithm with other blind channel identification algorithms using cyclostationary statistics. Finally, some conclusions are drawn in section 5.8.

5.2 A Cyclic Subspace Algorithm for Blind Channel Identification

5.2.1 Problem formulation

The single-user system considered in this section consists of one input signal source and an array of K subchannels. The signal at each subchannel is then sampled at a rate higher than the source signal symbol-rate $1/T$ ¹.

The source signal $x(t)$ is convolved with a different *subchannel* impulse response $h^{(i)}(k)$ as represented in Figure 5.1. The received signal $y^{(i)}(t)$ at the i th subchannel, sampled at $T_s = \frac{T}{P}$, can be expressed as:

$$y^{(i)}(k) = x(k) * h^{(i)}(k) + w^{(i)}(k) = \sum_{m=-\infty}^{\infty} a_m h^{(i)}(k - mP) + w^{(i)}(k) \quad P \in \mathcal{Z}, \quad (5.1)$$

where P is the oversampling factor, $x(k) = \sum_m a_m \delta(k - mP)$ is the oversampled input sequence which takes non-zero values every P samples, $\delta(k)$ is the Dirac delta, a_m is the generally complex, i.i.d. sequence of symbols at the baud-rate, $h^{(i)}(k)$ is the impulse response of the i th subchannel and $w^{(i)}(k)$ is an additive Gaussian noise process, which is generally coloured temporally by the receiver filter. The sequence $x(k)$ filled with zeros is a wide-sense cyclostationary process because its mean $E\{x(k)\}$ and autocorrelation function $R_x(k, m) = E\{x(k+m)x^*(k)\}$ are periodic in k with period P , for each value of m [36] [45]. Since digital communication signals sampled at a rate higher than the symbol rate exhibit cyclostationarity at cyclic frequencies which are harmonics of the symbol period T , the cyclic autocorrelation of the received sequence $y^{(i)}(k)$ is non-zero for some cyclic frequency $\alpha = l/P$, $l \in \mathcal{Z}$, where $T_s = 1$ was considered for normalisation.

A $N \times (N + M)$ filtering matrix $\mathcal{H}_N^{(i)}$ associated with the impulse response of the i th subchannel can be defined as:

$$\mathcal{H}_N^{(i)} = \begin{pmatrix} h^{(i)}(0) & \cdots & h^{(i)}(M) & 0 & 0 & \cdots & 0 \\ 0 & h^{(i)}(0) & \cdots & h^{(i)}(M) & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & h^{(i)}(0) & \cdots & h^{(i)}(M) \end{pmatrix}. \quad (5.2)$$

The subchannel impulse response vector $\underline{h}^{(i)}$ is defined as:

$$\underline{h}^{(i)} = [h^{(i)}(0), \dots, h^{(i)}(M)]^T \quad \text{for } i = 0, \dots, K - 1. \quad (5.3)$$

¹Although the system configuration is presented for a single-user case, the multi-user case will also be considered later in this chapter.

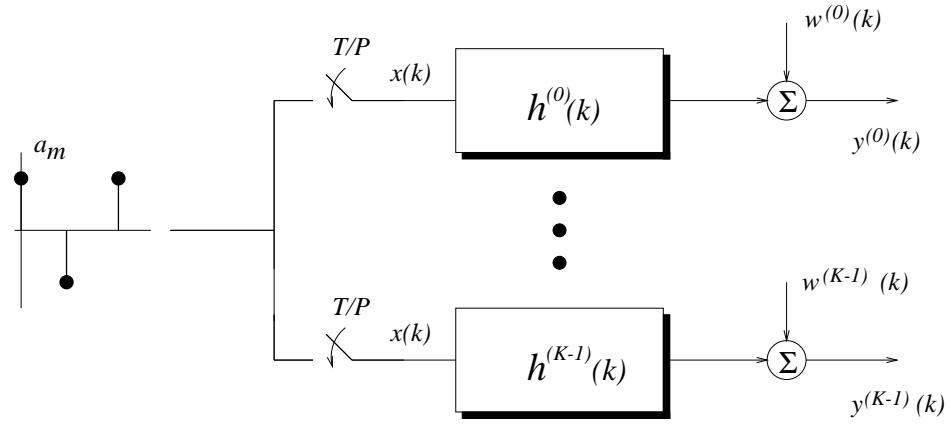


Figure 5.1: Oversampled multichannel system configuration.

The parameter M is the degree of ISI of the subchannel impulse response, measured in terms of the number of oversampled received samples that the subchannel is spread, and N is the length of the received samples observation window. Both parameters are multiples of the oversampling factor P .

Equation 5.1 can be re-arranged in terms of the filtering matrix $\mathcal{H}_N^{(i)}$ as:

$$\underline{y}_k^{(i)} = \mathcal{H}_N^{(i)} \underline{x}_k + \underline{w}_k^{(i)}, \quad (5.4)$$

where $\underline{x}_k = [x(k), x(k-1), \dots, x(k-N-M+1)]^T$ is the $(N+M) \times 1$ vector which contains the oversampled input sequence $x(k)$.

Similarly, $\underline{w}_k^{(i)} = [w^{(i)}(k), w^{(i)}(k-1), \dots, w^{(i)}(k-N+1)]^T$ is the $N \times 1$ noise vector, and $\underline{y}_k^{(i)} = [y^{(i)}(k), y^{(i)}(k-1), \dots, y^{(i)}(k-N+1)]^T$ the $N \times 1$ received signal vector.

By arranging the vectors of the oversampled signal at each subchannel $\underline{y}_k^{(i)}$ in a single vector:

$$\underline{y}_k = \left[\underline{y}_k^{(0)T}, \dots, \underline{y}_k^{(K-1)T} \right]^T, \quad (5.5)$$

the received signal vector \underline{y}_k can be expressed as:

$$\underline{y}_k = \mathcal{H}_N \underline{x}_k + \underline{w}_k, \quad (5.6)$$

where $\mathcal{H}_N \stackrel{\text{def}}{=} \Phi(\underline{h}, K, M+1, N) = \left[\mathcal{H}_N^{(0)T}, \dots, \mathcal{H}_N^{(K-1)T} \right]^T$ is a $KN \times (N+M)$ filtering matrix. The transformation $\Phi(\underline{h}, K, M+1, N)$ takes the form:

$$\mathcal{H}_N = \Phi(\underline{h}, K, M+1, N)$$

$$= \begin{pmatrix} h^{(0)}(0) & \dots & h^{(0)}(M) & 0 & 0 & \dots & 0 \\ 0 & h^{(0)}(0) & \dots & h^{(0)}(M) & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & h^{(0)}(0) & \dots & h^{(0)}(M) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h^{(K-1)}(0) & \dots & h^{(K-1)}(M) & 0 & 0 & \dots & 0 \\ 0 & h^{(K-1)}(0) & \dots & h^{(K-1)}(M) & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & h^{(K-1)}(0) & \dots & h^{(K-1)}(M) \end{pmatrix}, \quad (5.7)$$

where $\underline{h} = [\underline{h}^{(0)T}, \dots, \underline{h}^{(P-1)T}]^T$.

5.2.2 Channel identification using cyclostationary statistics and subspaces

In Chapter 3, the autocorrelation function of a wide-sense cyclostationary process $x(k)$ was presented as:

$$r_x(k, m) = E\{x(k+m)x^*(k)\} = r_x(k+P, m), \quad (5.8)$$

which is periodic in k with period P . Because of its periodicity, some Fourier series coefficients of the autocorrelation function of the input sequence $x(k)$ exist, and are known as the cyclic autocorrelation function $R_x^\alpha(m)$. These can be defined as:

$$r_x(k, m) = \sum_{\alpha} R_x^\alpha(m) e^{i2\pi\alpha k} \quad \alpha = \frac{l}{P}, \quad l \in Z, \quad (5.9)$$

or

$$R_x^\alpha(m) = \frac{1}{P} \sum_{k=0}^{P-1} r_x(k, m) e^{-i2\pi\alpha k}, \quad (5.10)$$

where α is a set of frequencies known as the cyclic frequencies [36]. From equation 5.10, the signal $x(k)$ is said to be wide-sense cyclostationary with period P , if a non-zero cyclic frequency α exists, for which the Fourier coefficients $R_x^\alpha(m)$ are non-zero. Equation 5.10 can also be written as [36] [45]:

$$R_x^\alpha(m) \triangleq \frac{1}{P} \sum_{k=0}^{P-1} E\{x(k+m)x^*(k)\} e^{-i2\pi\alpha k}$$

$$\begin{aligned}
 &= \frac{1}{P} \sum_{k=0}^{P-1} \left\{ \lim_{L \rightarrow \infty} \frac{1}{LP} \sum_{k=0}^{LP-1} x(k+m)x^*(k) \right\} e^{-i2\pi\alpha k} \\
 &= \lim_{L \rightarrow \infty} \frac{1}{LP} \sum_{k=0}^{LP-1} x(k+m)x^*(k)e^{-i2\pi\alpha k},
 \end{aligned} \tag{5.11}$$

where $*$ denotes conjugation and L is the number of symbols used in the estimation. In practice, for a large number of symbols L , (5.11) could be used instead of (5.10) to estimate the sample cyclic autocorrelation function $\hat{R}_x^\alpha(m)$.

The $n \times n$ autocorrelation matrix $R_x(k) = E\{\underline{x}_k \underline{x}_k^H\}$ of the zero-padded input sequence \underline{x}_k defined in (5.4), at time k , is a diagonal matrix with $P-1$ zeros interspersed between two non-zero elements:

$$\begin{aligned}
 R_x(0) &= \text{diag} \left[\sigma_x^2 \overbrace{0 \ \dots \ 0}^{P-1 \ \text{zeros}} \ \sigma_x^2 \ 0 \ \dots \ 0 \ \sigma_x^2 \ 0 \ \dots \right] \\
 R_x(1) &= \text{diag} \left[0 \ \sigma_x^2 \ \dots \ 0 \ 0 \ \sigma_x^2 \ \dots \ 0 \ 0 \ \sigma_x^2 \ \dots \right] \\
 &\vdots \\
 R_x(k) &= \text{diag} \left[\underbrace{0 \ \dots \ 0}_{k \ \text{zeros}} \ \sigma_x^2 \ \overbrace{0 \ \dots \ 0}^{P-1 \ \text{zeros}} \ \sigma_x^2 \ \overbrace{0 \ \dots \ 0}^{P-1 \ \text{zeros}} \ \sigma_x^2 \ 0 \ \dots \right],
 \end{aligned} \tag{5.12}$$

and following equation 5.10, the cyclic autocorrelation matrix can be calculated using:

$$R_x^\alpha = \frac{1}{P} \sum_{k=0}^{P-1} R_x(k) e^{-i2\pi\alpha k}. \tag{5.13}$$

This results in a diagonal matrix of the form

$$R_x^\alpha = \text{diag} \left[\underbrace{1 \ e^{-i2\pi\alpha} \ \dots \ e^{-i2\pi\alpha(P-1)} \ 1 \ \dots \ e^{-i2\pi\alpha(P-1)} \ \dots}_{n} \right] \frac{\sigma_x^2}{P}. \tag{5.14}$$

where n is a multiple of P . Since the magnitude of each element on the leading diagonal is non-zero, matrix R_x^α is full rank for all α , except $\alpha = 0$.

On the other hand, following the expression of the received signal vector \underline{y}_k given in (5.6), the output autocorrelation matrix can be expressed as:

$$\begin{aligned}
 R_y(k) &= E\{\underline{y}_k \underline{y}_k^H\} \\
 &= \mathcal{H}_N R_x(k) \mathcal{H}_N^H + R_w,
 \end{aligned} \tag{5.15}$$

where R_w is the autocorrelation matrix of the stationary noise process. By definition:

$$R_y^\alpha = \frac{1}{P} \sum_{k=0}^{P-1} R_y(k) e^{-i2\pi\alpha k}. \quad (5.16)$$

Using (5.15),

$$\begin{aligned} R_y^\alpha &= \frac{1}{P} \sum_{k=0}^{P-1} \mathcal{H}_N R_x(k) \mathcal{H}_N^H e^{-i2\pi\alpha k} + \frac{1}{P} \sum_{k=0}^{P-1} R_w e^{-i2\pi\alpha k} \\ &= \frac{1}{P} \mathcal{H}_N \left\{ \sum_{k=0}^{P-1} R_x(k) e^{-i2\pi\alpha k} \right\} \mathcal{H}_N^H + R_w \frac{1}{P} \sum_{k=0}^{P-1} e^{-i2\pi\alpha k} \\ &= \mathcal{H}_N R_x^\alpha \mathcal{H}_N^H + R_w \frac{1}{P} \sum_{k=0}^{P-1} e^{-i2\pi\alpha k}. \end{aligned} \quad (5.17)$$

For $\alpha = 0$,

$$R_y^0 = \mathcal{H}_N R_x^0 \mathcal{H}_N^H + \underline{R}_w \quad (5.18)$$

is the received signal average autocorrelation matrix, but more importantly, for $\alpha \neq 0$,

$$R_y^\alpha = \mathcal{H}_N R_x^\alpha \mathcal{H}_N^H \quad \alpha = \frac{l}{P} \quad l \neq nP, \quad l, n \in Z, \quad (5.19)$$

where R_x^α is the $(M + N) \times (M + N)$ cyclic autocorrelation matrix of the oversampled input sequence, as defined in (5.14), and R_y^α is the $KN \times KN$ cyclic autocorrelation matrix of the received signal $y(k)$.

Subspace decomposition

Similarly to the formulation of the conventional subspace method of Moulines *et al.* [61], the identification of the channel coefficients, in this case, is based on the subspace decomposition of the cyclic autocorrelation matrix of the received signal, instead of the autocorrelation matrix. The main difference between the formulation of a multichannel system sampled at T using the autocorrelation function [61] and the formulation of an oversampled multichannel system using the cyclic autocorrelation function (5.19) is the fact that the cyclic autocorrelation matrix of the additive noise is theoretically zero for $\alpha \neq 0$; the Fourier coefficients of the autocorrelation of the noise process, which is not periodic, are zero for any cyclic frequency α . This condition is robust in the presence of white or coloured noise, whereas in [61], the condition that the noise autocorrelation has to be $R_w = \sigma^2 I$ is not satisfied when the noise is coloured.

The source cyclic autocorrelation matrix R_x^α is unknown but full-rank. Hence, provided that the filtering matrix \mathcal{H}_N in (5.19) is full column rank, the received signal cyclic autocorrelation matrix R_y^α admits the following SVD:

$$R_y^\alpha = U \Sigma U^H. \quad (5.20)$$

Matrix Σ is a diagonal matrix which contains the eigenvalues $\lambda_0, \dots, \lambda_{KN-1}$ of matrix R_y^α , where

$$\begin{aligned} \lambda_i &> 0 && \text{for } i = 0, \dots, N + M - 1 \\ \lambda_i &= 0 && \text{for } i = N + M, \dots, KN - 1. \end{aligned} \quad (5.21)$$

The eigenvectors of matrix $U = [S \ G]$ associated with the first $N + M$ eigenvalues in equation 5.21 form the so called signal eigenvectors matrix $S = [\underline{s}_0, \dots, \underline{s}_{N+M-1}]$ of size $KN \times (N + M)$, whereas the eigenvectors associated with the last $KN - N - M$ for $\alpha \neq 0$ conform the null-space eigenvector matrix $G = [\underline{g}_0, \dots, \underline{g}_{KN-N-M-1}]$ of size $KN \times (KN - N - M)$. The $(N + M)$ columns of matrix S span the *signal subspace*, while the columns of matrix G generate the so-called *null subspace*. As a difference from the conventional subspace method, the vectors not associated with the signal subspace, form the null subspace in the cyclic subspace method, and not the noise subspace as in [61]. In [81], the null-space concept was used to denote the noise subspace.

By orthogonality between the columns of matrix S and G , it can be concluded that any vector from the null subspace is orthogonal to any column vector in the signal subspace, and by extension to any column of the filtering matrix \mathcal{H}_N :

$$\underline{g}_i^H \mathcal{H}_N = 0 \quad 0 \leq i < KN - N - M. \quad (5.22)$$

The solution to this linear system of equations results in the identification of the subchannel impulse responses, provided that matrix \mathcal{H}_N is full column rank. The necessary conditions to meet this criterion are [61]:

- i)** The oversampled polynomials $H^{(i)}(z) = \sum_{j=0}^M h^{(i)}(j)z^j$ have no common zero,
- ii)** The length of the observation N is greater or equal to the maximum degree M of the polynomials $H^{(i)}(z)$, i.e. $N \geq M$,
- iii)** One polynomial $H^{(i)}(z)$ has at least degree M .
- iv)** At least one of the subchannels should have a non-zero M th tap. This will avoid introducing a zero at $z = 0$ in each of the subchannels.

In practice, the linear system of equations of (5.22) can be determined using (3.59), as shown in section 3.4.3 for the conventional subspace method.

The uniqueness of the estimate of the set of subchannel impulse responses $\underline{h}^{(i)}$, up to a scalar factor, is determined by the special structure of the filtering matrix \mathcal{H}_N , which establishes that \mathcal{H}_N has to be full-column rank ².

Proof: Consider 2 full-column matrices, \mathcal{H}_N and \mathcal{H}'_N . It is shown that provided that both matrices have the same range, both matrices are proportional, i.e. a complex scalar value $\gamma = |\gamma| \exp(\theta)$ exists where,

$$\mathcal{H}_N = \gamma \mathcal{H}'_N. \quad (5.23)$$

In this case, the K subchannels associated with \mathcal{H}_N and \mathcal{H}'_N are proportional. This implies that the solution to the linear system of equations (5.22) is the true channel matrix or some scaled version of the true matrix.

By permuting the rows of matrix \mathcal{H}_N , a block-Toeplitz matrix can be formed:

$$\underline{\mathcal{H}}_N = \begin{pmatrix} h^{(0)}(0) & \cdots & h^{(0)}(M) & 0 & \cdots & 0 & \cdots & 0 \\ h^{(1)}(0) & \cdots & h^{(1)}(M) & h^{(0)}(M) & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ h^{(K-1)}(0) & \cdots & h^{(K-1)}(M) & h^{(K-2)}(M) & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & h^{(0)}(0) & \cdots & h^{(0)}(M) \\ 0 & \cdots & 0 & 0 & \cdots & h^{(1)}(0) & \cdots & h^{(1)}(M) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & h^{(K-1)}(0) & \cdots & h^{(K-1)}(M) \end{pmatrix}. \quad (5.24)$$

The column space of matrix $\underline{\mathcal{H}}_N$ is equivalent to that of matrix \mathcal{H}_N , i.e. the determinants of both matrices have the same absolute value. Equation 5.24 can also be formulated in terms of a matrix of lower rank as:

$$\underline{\mathcal{H}}_N = \left(\begin{array}{c|cccccc} \underline{h}(0) & \underline{h}(1) & \cdots & \underline{h}(M) & \cdots & 0_K & \cdots & 0_K \\ 0_K & \underline{h}(0) & \cdots & \underline{h}(M-1) & \cdots & 0_K & \cdots & 0_K \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0_K & 0_K & \cdots & 0_K & \cdots & \underline{h}(0) & \cdots & \underline{h}(M) \end{array} \right) \quad N > M, \quad (5.25)$$

where $\underline{h}(i) = [h^{(0)}(i) \cdots h^{(K-1)}(i)]^T$ denotes the vector containing the i th tap of each of the subchannels and 0_K denotes a vector with K zeros:

²The proof of uniqueness was first established in [61]

$$0_K = [\underbrace{0 \cdots 0}_K]^T. \quad (5.26)$$

Equivalently, (5.25) can be re-written as:

$$\underline{\mathcal{H}}_N = \begin{pmatrix} \underline{h}(0) & p_{N-1} \\ 0_{K(N-1)} & \underline{\mathcal{H}}_{N-1} \end{pmatrix}, \quad (5.27)$$

where

$$p_{N-1} = [\underbrace{\underline{h}(1) \cdots \underline{h}(M)}_{M+N-1} \cdots 0_K] \quad (5.28)$$

is a $K \times (M + N - 1)$ matrix.

Similarly, it can be established that:

$$\underline{\mathcal{H}}_N = \left(\begin{array}{cccccc|c} \underline{h}(0) & \underline{h}(1) & \cdots & \underline{h}(M) & \cdots & 0_K & \cdots & 0_K \\ 0_K & \underline{h}(0) & \cdots & \underline{h}(M-1) & \cdots & 0_K & \cdots & 0_K \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0_K & 0_K & \cdots & 0_K & \cdots & \underline{h}(0) & \cdots & \underline{h}(M) \end{array} \right) \quad N > M, \quad (5.29)$$

and accordingly,

$$\underline{\mathcal{H}}_N = \begin{pmatrix} \underline{\mathcal{H}}_{N-1} & 0_{K(N-1)} \\ q_{N-1} & \underline{h}(0) \end{pmatrix}. \quad (5.30)$$

Because the range of matrix $\underline{\mathcal{H}}'_N$ lies in the range of $\underline{\mathcal{H}}_N$, the column space of $\underline{\mathcal{H}}'_N$ is spanned by the column space of $\underline{\mathcal{H}}_N$. Thus, the 1st column of $\underline{\mathcal{H}}'_N$ can be expressed as:

$$\begin{pmatrix} \underline{h}'(0) \\ 0_{K(N-1)} \end{pmatrix} = \begin{pmatrix} \underline{h}(0) & p_{N-1} \\ 0_{K(N-1)} & \underline{\mathcal{H}}_{N-1} \end{pmatrix} \begin{pmatrix} \alpha_0 \\ 0_{M+N-1} \end{pmatrix}, \quad (5.31)$$

where α_0 is a complex scalar factor. Consequently, the linear system of equations of (5.31) reduces to:

$$\underline{h}'(0) = \alpha_0 \underline{h}(0) \quad (5.32)$$

which indicates that the vectors $\underline{h}'(0)$ and $\underline{h}(0)$ are proportional. The same relation used in (5.31) can be applied to the 2nd column of matrix $\underline{\mathcal{H}}'_N$:

$$\begin{pmatrix} \underline{h}'(1) \\ \underline{h}'(0) \\ 0_{K(N-2)} \end{pmatrix} = \begin{pmatrix} \underline{h}(0) & p_{N-1} \\ 0_{K(N-1)} & \underline{\mathcal{H}}_{N-1} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_0 \\ 0_{M+N-2} \end{pmatrix}, \quad (5.33)$$

and

$$\begin{aligned} \underline{h}'(1) &= \alpha_1 \underline{h}(0) + \alpha_0 \underline{h}(1) \\ \underline{h}'(0) &= \alpha_0 \underline{h}(0) \end{aligned} \quad (5.34)$$

Following the same procedure for the $M + 1$ th column of matrix $\underline{\mathcal{H}}'_N$, the following relationship can be established:

$$\begin{pmatrix} \underline{h}'(M) \\ \vdots \\ \underline{h}'(0) \\ 0_{K(N-M-1)} \end{pmatrix} = \begin{pmatrix} \underline{h}(0) & p_{N-1} \\ 0_{K(N-1)} & \underline{\mathcal{H}}_{N-1} \end{pmatrix} \begin{pmatrix} \alpha_M \\ \vdots \\ \alpha_0 \\ 0_{N-1} \end{pmatrix}, \quad (5.35)$$

and

$$\begin{aligned} \underline{h}'(M) &= \alpha_M \underline{h}(0) + \alpha_{M-1} \underline{h}(1) + \cdots + \alpha_0 \underline{h}(M) \\ &\vdots \\ \underline{h}'(1) &= \alpha_1 \underline{h}(0) + \alpha_0 \underline{h}(1) \\ \underline{h}'(0) &= \alpha_0 \underline{h}(0) \end{aligned} \quad (5.36)$$

This linear system of equations can be written in matrix form as:

$$\begin{pmatrix} \underline{h}'(M) \\ \vdots \\ \underline{h}'(1) \\ \underline{h}'(0) \end{pmatrix} = \begin{pmatrix} \underline{h}(0) & \underline{h}(1) & \cdots & \underline{h}(M) & \cdots & 0 \\ 0 & \underline{h}(0) & \cdots & \underline{h}(M-1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \underline{h}(0) & \cdots & \underline{h}(M) \end{pmatrix} \begin{pmatrix} \alpha_M \\ \vdots \\ \alpha_0 \\ 0_M \end{pmatrix}, \quad (5.37)$$

and

$$\underline{h}' = \underline{\mathcal{H}}_{M+1} \underline{\alpha}, \quad (5.38)$$

where $\underline{h}' = [\underline{h}'(M)^T \cdots \underline{h}'(1)^T \underline{h}'(0)^T]^T$ is the oversampled channel impulse response, as defined in section 5.2.1, and $\underline{\alpha} = [\alpha_M \cdots \alpha_0 0_M^T]^T$.

Now, the same procedure used with (5.27) for columns $1 \cdots M + 1$ of matrix $\underline{\mathcal{H}}'_N$ can be used with (5.30), for columns $M + 2 \cdots 2M + 2$ of $\underline{\mathcal{H}}'_N$. Consequently, the $M + 2$ th column of matrix $\underline{\mathcal{H}}'_N$ can be expressed as:

$$\begin{pmatrix} \underline{h}'(M) \\ \vdots \\ \underline{h}'(0) \\ 0_{K(N-M-1)} \end{pmatrix} = \begin{pmatrix} \underline{\mathcal{H}}_{N-1} & 0_{K(N-1)} \\ q_{N-1} & \underline{h}(0) \end{pmatrix} \begin{pmatrix} 0_M \\ \beta_0 \\ \vdots \\ \beta_M \end{pmatrix}, \quad (5.39)$$

and

$$\underline{h}'(M) = \beta_0 \underline{h}(M). \quad (5.40)$$

Similarly, for the $(2M+2)$ th column of matrix $\underline{\mathcal{H}}'_N$, we have:

$$\begin{pmatrix} 0_{K(N-M-1)} \\ \underline{h}'(M) \\ \vdots \\ \underline{h}'(0) \end{pmatrix} = \begin{pmatrix} \underline{\mathcal{H}}_{N-1} & 0_{K(N-1)} \\ q_{N-1} & \underline{h}(0) \end{pmatrix} \begin{pmatrix} 0_M \\ \beta_0 \\ \vdots \\ \beta_M \end{pmatrix}, \quad (5.41)$$

and

$$\begin{aligned} \underline{h}'(0) &= \beta_M \underline{h}(M) + \beta_{M-1} \underline{h}(M-1) + \cdots + \beta_0 \underline{h}(0) \\ \underline{h}'(1) &= \beta_{M-1} \underline{h}(M) + \cdots + \beta_0 \underline{h}(1) \\ &\vdots \\ \underline{h}'(M) &= \beta_0 \underline{h}(M) \end{aligned} \quad (5.42)$$

This system of linear equations can also be written in matrix form as:

$$\begin{pmatrix} \underline{h}'(M) \\ \vdots \\ \underline{h}'(1) \\ \underline{h}'(0) \end{pmatrix} = \begin{pmatrix} \underline{h}(0) & \underline{h}(1) & \cdots & \underline{h}(M) & \cdots & 0 \\ 0 & \underline{h}(0) & \cdots & \underline{h}(M-1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \underline{h}(0) & \cdots & \underline{h}(M) \end{pmatrix} \begin{pmatrix} 0_M \\ \beta_0 \\ \vdots \\ \beta_M \end{pmatrix}, \quad (5.43)$$

and

$$\underline{h}' = \underline{\mathcal{H}}_{M+1} \underline{\beta}, \quad (5.44)$$

where $\underline{\beta} = [0_M^T \ \beta_M \ \beta_0]^T$.

As a result, both equations (5.38) and (5.38) equal to:

$$\underline{\mathcal{H}}_{M+1} \underline{\alpha} = \underline{\mathcal{H}}_{M+1} \underline{\beta}. \quad (5.45)$$

From this relationship, it is clear that $\underline{\alpha} = \underline{\beta}$, or equivalently:

$$\begin{pmatrix} \alpha_M \\ \vdots \\ \alpha_0 \\ 0_M \end{pmatrix} = \begin{pmatrix} 0_M \\ \beta_M \\ \vdots \\ \beta_0 \end{pmatrix}. \quad (5.46)$$

This equivalence indicates that there is only one single nonzero element in both vectors $\underline{\alpha}$ and $\underline{\beta}$, i.e. $\alpha_0 = \beta_M$. The rest of the elements are simply zero,

$$\begin{aligned} \beta_i &= 0 & i = 0, \dots, M-1 \\ \alpha_i &= 0 & i = 1, \dots, M. \end{aligned} \quad (5.47)$$

Finally, it can be established that in fact (5.23) applies, and matrix $\underline{\mathcal{H}}'_N$ and $\underline{\mathcal{H}}_N$ are proportional, where

$$\gamma = \alpha_0 \quad (5.48)$$

This concludes the proof.

Implementation of the cyclic subspace algorithm

In practice, by assuming that the wide-sense cyclostationary process \underline{y}_k is cyclo-ergodic [36], the cyclic autocorrelation function can be estimated using [45]:

$$\hat{R}_y^\alpha = \frac{1}{LP} \sum_{k=0}^{LP-N-M} \underline{y}_k \underline{y}_k^H e^{-i2\pi\alpha k}. \quad (5.49)$$

When only sample estimates of the received signal cyclic autocorrelation matrix \underline{R}_y^α are available, the set of linear equations of equation 5.22 can be solved in the least-squares sense, as the minimisation of the following quadratic form:

$$q = \sum_{i=0}^{KN-M-N-1} | \hat{\underline{g}}_i^H \underline{\mathcal{H}}_N |^2. \quad (5.50)$$

The solution to this quadratic form is formed using a unit norm constraint on the impulse response of the channel, as explained in section 4.3.1. The given solution of the channel impulse response estimate is non-coherent, and proper gain and phase adjustments need to be made. However, this does not constitute a serious drawback, since it is recognised as part of the blind equalisation problem.

5.3 Asymptotic Performance

In this section, an asymptotic variance estimator for the cyclic subspace method will be derived. The mathematical development is reminiscent of the asymptotic variance estimator for the conventional subspace method given in [73].

The channel estimation error, denoted $\Delta \underline{h}$, is the difference between the estimated and true values of the multichannel impulse response \underline{h} , which is defined in (5.7):

$$\Delta \underline{h} = \hat{\underline{h}} - \underline{h}. \quad (5.51)$$

Alternatively, as shown in [73], the channel estimation error can also be expressed as:

$$\Delta \underline{h} = -(Q^\dagger)^H \hat{Q} \underline{h}. \quad (5.52)$$

Matrix Q , defined in section 4.3.1, is expressed as;

$$Q = \sum_{i=0}^{KN-N-M-1} \mathcal{G}_i \mathcal{G}_i^H, \quad (5.53)$$

where $\mathcal{G}_i = \Phi(\underline{g}_i, K, N, M+1)$ and \underline{g}_i denote the null-space eigenvectors. The transformation $\Phi(\cdot)$ was introduced in (5.7). Thus, \hat{Q} and Q^\dagger denote the sample estimate of matrix Q obtained using the sample cyclic autocorrelation in (5.49) and the pseudo-inverse of matrix Q respectively. The k th entry of the channel estimation error, denoted $\Delta h(k) = \hat{h}(k) - h(k)$, can be expressed as:

$$\begin{aligned} \Delta h(k) &= -q_k^H \hat{Q} \underline{h} = \sum_{i=0}^{KN-N-M-1} -q_k^H \hat{\mathcal{G}}_i \hat{\mathcal{G}}_i^H \underline{h} = \sum_{i=0}^{KN-N-M-1} -\hat{\underline{g}}_i^H \mathcal{H}_N Q_k^H \hat{\underline{g}}_i \\ &= \sum_{i=0}^{KN-N-M-1} -\hat{\underline{g}}_i \hat{\underline{g}}_i^H \mathcal{H}_N Q_k^H, \end{aligned} \quad (5.54)$$

where q_k is the k th column of Q^\dagger and $Q_k = \Phi(\underline{q}_k, K, M+1, N)$. According to [85],

$$\hat{\underline{g}}_i \hat{\underline{g}}_i^H \approx -\underline{g}_i \underline{g}_i^H \Delta S S^H \mathcal{H}_N, \quad (5.55)$$

where $\Delta S = \hat{S} - S$ denotes the difference between the sample estimate of matrix S and the true value. Consequently, (5.54) can be expressed as:

$$\Delta h(k) = - \sum_{i=0}^{KN-N-M-1} \hat{\underline{g}}_i \hat{\underline{g}}_i^H \mathcal{H}_N Q_k^H \approx \sum_{i=0}^{KN-N-M-1} \underline{g}_i \underline{g}_i^H \Delta S S^H \mathcal{H}_N Q_k^H.$$

$$= \sum_{i=0}^{KN-N-M-1} \underline{g}_i^H \Delta S S^H \mathcal{H}_N Q_k^H \underline{g}_i \quad (5.56)$$

Also according to [85], one can express (5.56) in terms of the received signal sample estimate autocorrelation matrix \hat{R}_y^α as:

$$\Delta h(k) \approx \sum_{i=0}^{KN-N-M-1} \underline{g}_i^H \hat{R}_y^\alpha S \Sigma_s^{-1} S^H \mathcal{H}_N Q_k^H \underline{g}_i, \quad (5.57)$$

where

$$\underline{g}_i^H \Delta S \approx \underline{g}_i^H \hat{R}_y^\alpha S \Sigma_s^{-1} \quad (5.58)$$

has been used. Σ_s is the matrix with the eigenvalues corresponding to the signal subspace, i.e. $\Sigma_s \triangleq \text{diag} [\lambda_0 \cdots \lambda_{N+M-1}]$. Matrix Σ_s differs from the definition given in [73] for the conventional subspace method. There, Σ_s was defined as $\Sigma_s \triangleq \text{diag} [(\lambda_0 - \sigma^2) \cdots (\lambda_{N+M-1} - \sigma^2)]$, whereas here, because the effect of the noise in the theoretical received signal cyclic autocorrelation matrix is null, it takes the form defined in this paragraph.

The variance of the estimate of an unit-norm channel is given by [73]

$$\sigma_{\underline{h}}^2 = E \{ \|\Delta \underline{h}\|^2 \} = \frac{\sigma^2}{L} (a + \sigma^2 b), \quad (5.59)$$

where L is the number of symbols used in the estimation of the cyclic autocorrelation matrix R_y^α and

$$a = \sum_{k=1}^{K(M+1)} \sum_{i=-(N-1)}^{(N-1)} \frac{L-|i|}{L} \times \text{tr} \left(G^H (I_P \otimes J_i) G \tilde{G}_k^H \mathcal{H}_N R_x^\alpha(-i) \mathcal{H}_N^H \tilde{G}_k \right), \quad (5.60)$$

$$b = \sum_{k=1}^{P(M+1)} \sum_{i=-(N-1)}^{(N-1)} \frac{L-|i|}{L} \times \text{tr} \left(G^H (I_P \otimes J_i) G \tilde{G}_k^H (I_P \otimes J_{-i}) \tilde{G}_k \right), \quad (5.61)$$

where I_P is a $P \times P$ identity matrix,

$$(J_k)_{ij} = \begin{cases} 1 & \text{if } i = j + k \text{ and } |k| \leq N - 1, \\ 0 & \text{otherwise} \end{cases}$$

and

$$\tilde{G}_k \triangleq [\tilde{g}_{1,k}, \dots, \tilde{g}_{PN-N-M,k}],$$

where

$$\tilde{g}_{i,k} = S\Sigma_s S^H \mathcal{H}_N Q_k^H \underline{g}_i. \quad (5.62)$$

For a sufficiently large number of symbols used in the estimation ($L \gg W$), a and b are independent of L , and the lower bound of the variance of the channel parameter estimates can be obtained in a computationally efficient manner.

5.4 Correlated Noise Compensation

In [61], at least partial knowledge of the noise process is required in order to build a whitening filter to account for correlated noise, because the algorithm degrades considerably in the presence of correlated noise. Buisán *et al.* [22] recently proposed a method to deal with correlated noise in the subspace method proposed by Moulines *et al.* [61], where no knowledge of the noise autocorrelation was required.

In the presence of correlated noise, the noise covariance matrix R_w given in (3.56) is not diagonal,

$$R_w \neq \sigma^2 I. \quad (5.63)$$

The solution to the problem given in [61] is to whiten the received signal. The whitening filter can be constructed using a matrix $R'_w = \frac{1}{\sigma^2} R_w$. The whitening filter, then, consists of the inverse of the Hermitian square root of R'_w :

$$F_w = (R'_w)^{-1/2} \quad (5.64)$$

The filter can be obtained using the SVD algorithm³ in the following way:

$$R'_w = U \Sigma V^H, \quad (5.65)$$

where $\Sigma = \text{diag}(\sigma_0 \cdots \sigma_{KN-1})$. It can then be established that

$$(R'_w)^{1/2} = U \Sigma^{1/2} V^H, \quad (5.66)$$

and $\Sigma^{1/2} = \text{diag}(\sigma_0^{1/2} \cdots \sigma_{KN-1}^{1/2})$. Finally, as shown in Appendix B, the inverse of $(R'_w)^{1/2}$ can be calculated using:

³see Appendix B

$$(R'_w)^{-1/2} = U\Sigma^{-1/2}U^H, \quad (5.67)$$

where $\Sigma^{-1/2} = \text{diag}(1/\sigma_0^{1/2} \dots 1/\sigma_{KN-1}^{1/2})$. As some of the singular values $\sigma_0^{1/2} \dots \sigma_{KN-1}^{1/2}$ will be small at low noise power scenarios, the inverse matrix will become ill-conditioned. It is common practice, then, to add energy to the diagonal elements of matrix $(R'_w)^{1/2}$ so that the matrix might be invertible [47].

Finally, the autocorrelation function of the whitened signal becomes:

$$R'_y = F_w R_y F_w^H = F_w \mathcal{H}_N R_x \mathcal{H}_N^H F_w^H + \sigma^2 I \quad (5.68)$$

Since the span of the product $F_w \mathcal{H}_N$ is in the range of \mathcal{H}_N , a case similar to the white noise case holds, with the exception that now, equation 3.58 takes the form:

$$\underline{g}_i^H F_w \mathcal{H}_N = 0 \quad 0 \leq i \leq PN - N - M - 1 \quad (5.69)$$

The solution to this system yields an estimate of the filtered matrix $F_w \mathcal{H}_N$, rather than \mathcal{H}_N .

As shown in (5.19), the cyclic autocorrelation matrix of the additive Gaussian noise process $w(n)$ is zero for some non-zero cyclic frequency α . This implies that the received signal cyclic autocorrelation matrix is insensitive in theory to whether the additive Gaussian noise process is correlated or not.

The cyclic subspace algorithm does not require either the estimation and subtraction of the covariance matrix of the noise as in [87] or the noise to be uncorrelated from one sensor to another. Moreover, it can account for correlated noise without any pre-processing, as in [22]. However, it does require a long enough integration time, in order to ensure that the noise components in the cyclic autocorrelation matrix are sufficiently small; this affects either white or coloured noise. The principal effect is that when a sufficiently large data set is not available for estimation, the cyclic subspace algorithm becomes more sensitive to white noise than the subspace algorithm. The integration time is normally longer than that required by the conventional subspace algorithm [78].

5.5 Blind Multiuser Detection

Most of the blind channel identification algorithms in the literature consider a *single-input-multiple-output* (SIMO) system. However, most of the times, it is the case that other users are also transmitting alongside the desired user. The signal of the desired user is known as the *signal-of-interest* (SOI), whereas the signals from other users are denoted as *signals-not-of-interest* (SNOI).

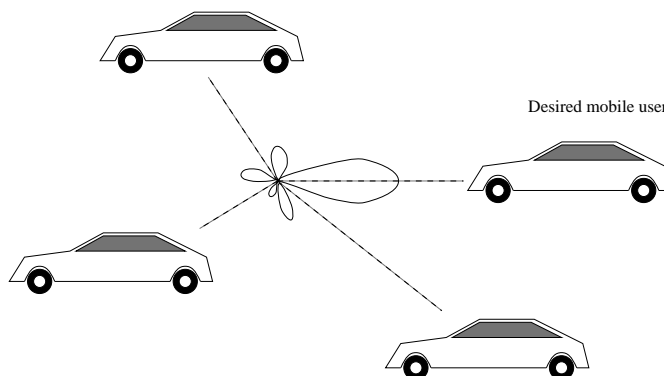


Figure 5.2: Antenna-pattern of a receiver which cancels-out the signals from the undesired mobile users.

Thanks to *direction-of-arrival* (DOA) estimation techniques, the DOA of the SOI can be determined. The appropriate weights of a spatial filter can then be calculated [5] to cancel out the interfering signals, which are temporarily and spectrally overlapping, as shown in Figure 5.2. Many algorithms are available for DOA estimation; two of the best known are the MUSIC [85] and ESPRIT [76] algorithms, and also the extensions of these algorithms to accommodate cyclostationary signals, described in [32]. Nevertheless, in general, all DOA estimation techniques are known to have poor performance in the presence of multipath propagation.

On the other hand, the issue of temporal equalisation of users, which are relatively closely spaced spatially and cannot be separated by a spatial filter, remains a problem. Several adaptive filtering techniques are available [62] [79] to remove the effect of co-channel interference but they need training. This requires extra capacity to store the training signal and transmit it periodically, and the need to synchronise the received signal and the training signal stored locally. Although the current mobile radio system has adopted a configuration where the training signal is transmitted periodically, interposed between the transmission of information data, other services such as the digital terrestrial HDTV standard [65], establish the need to adaptively increase the duration of the training signal to accommodate varying levels of interference. The *spectral coherence restoral* (SCORE) algorithms [36] are blind adaptive filtering techniques, which exploit cyclostationary properties, that require neither a training or pilot signal, nor knowledge of the spatial characteristic of the SOI.

In general, those signals which are spatially separable can be assigned overlapping spectral bands because the interfering signals can be cancelled using a spatial filter. This configuration is known as space-frequency-division-multiple-access (SFDMA). The SFDMA scheme can be complemented using multiuser blind equalisation techniques, which are able to equalise signals which cannot be separated temporally and spatially, without using a training signal.

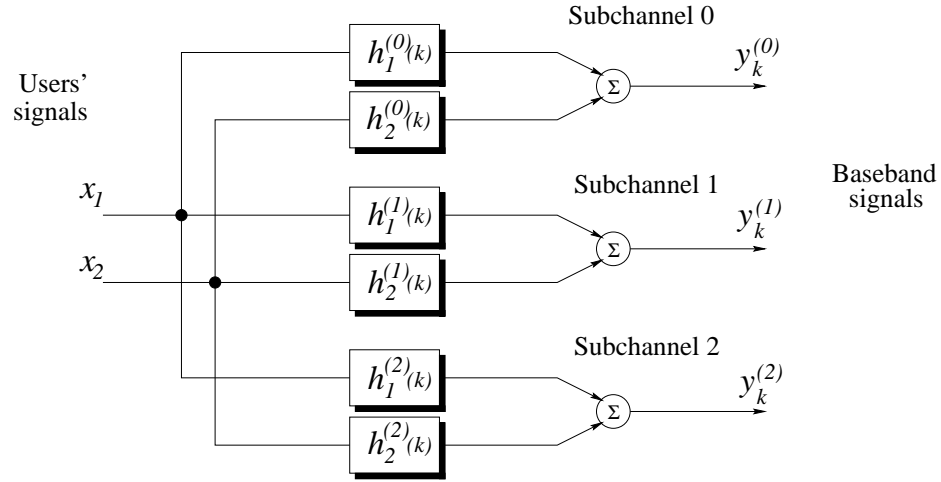


Figure 5.3: Multiple-input multiple-output (MIMO) system representation.

5.5.1 Multiuser blind equalisation using subspace decomposition

Several methods have recently been proposed which deal with the problem of multiuser blind channel identification and signal detection [82] [97] [4] [55] [2]. Some of the algorithms are in fact extensions of multichannel SIMO blind channel identification schemes transformed to accommodate multiple signals. The multichannel structure of fractionally-spaced blind channel identification algorithms, such as those described in Chapter 3 [61] [87], as well as [81], has a special structure which leads directly to a multiuser system configuration.

Figure 5.3 represents a system with two signals $x_1(k)$ and $x_2(k)$, which are transmitted simultaneously. These signals go through $K = 3$ subchannels each and the received signal at any particular subchannel i is the combination of the convolution of each of the transmitted signals with their respective subchannels. Denote $h_1^{(0)}(k)$ and $h_2^{(0)}(k)$, the subchannel 0 impulse responses of the signals $x_1(k)$ and $x_2(k)$ respectively.

Similarly to the formulation of the single-user system shown in (3.54) for the conventional subspace algorithm of Moulines *et al.* [61], in the multi-user system, the received signal vector \underline{y}_k can then be expressed as:

$$\underline{y}_k = [\mathcal{H}_{N,1} \quad \mathcal{H}_{N,2}][\underline{x}_{1,k}^T \quad \underline{x}_{2,k}^T]^T + \underline{w}_k \quad \text{for } k = 0, \dots, L-1 \quad (5.70)$$

This representation of the multiuser system indicates that the rank of the received signal autocorrelation matrix grows proportionally to the number of users, assuming that the column span of both matrices $\mathcal{H}_{N,1}$ and $\mathcal{H}_{N,2}$ is the same. As a consequence, the subchannels of both signals have to be estimated, instead of only the subchannel of the SOI. The received signal autocorrelation matrix can be expressed as:

$$R_y = [\mathcal{H}_{N,1} \quad \mathcal{H}_{N,2}]R_x[\mathcal{H}_{N,1} \quad \mathcal{H}_{N,2}]^H + R_w \quad (5.71)$$

where R_x represents the $2(M+N) \times 2(M+N)$ matrix of autocorrelations and cross-correlations of the transmitted signals. The eigenvalue decomposition of matrix R_y indicates that the eigenvalues of matrix R_y take the form:

$$\begin{aligned} i) \quad \lambda_i &= \sigma_i + \sigma^2 & \text{for } i = 0, \dots, 2(M+N) - 1 \\ ii) \quad \lambda_i &= \sigma^2 & \text{for } i = 2(M+N), \dots, KN - 1 \end{aligned} \quad (5.72)$$

the first $2(M+N)$ eigenvalues will correspond to the signal subspace and the last $KN - 2(M+N)$ to the noise subspace. Because matrix R_y is of dimension $KN \times KN$, the number of channels K has to be equal or higher than the number of interfering users, i.e. $K > \frac{(N_u+1)}{N}(M+N)$.

However, problems are expected due to the fact that the signal subspace is a combination of the two transmitted signals. Since the noise subspace vectors are orthogonal to any of the column vectors of matrix $\mathcal{H}_N \stackrel{\text{def}}{=} [\mathcal{H}_{N,1} \quad \mathcal{H}_{N,2}]$, the identification of the channel impulse response of a particular signal is formulated as:

$$\underline{g}_i^H \mathcal{H}_{N,1} = 0 \quad 0 \leq i \leq KN - 2(M+N) - 1, \quad (5.73)$$

for signal 1, and

$$\underline{g}_i^H \mathcal{H}_{N,2} = 0 \quad 0 \leq i \leq KN - 2(M+N) - 1, \quad (5.74)$$

for signal 2. Although the minimisation of both (5.73) and (5.74) can still be achieved, the information contained in the noise subspace vectors or equivalently the signal subspace vectors⁴ is insufficient to independently determine the impulse filtering matrices $\mathcal{H}_{N,1}$ and $\mathcal{H}_{N,2}$ [57]. Nevertheless, it has been shown in [82] [86], that the best it can be done under these circumstances is to estimate \mathcal{H}_N and recover the input signal vector $\underline{x} \stackrel{\text{def}}{=} [\underline{x}_{1,k}^T \quad \underline{x}_{2,k}^T]^T$ up to an ambiguity factor W . In other words, if a non-singular $2(M+N) \times 2(M+N)$ matrix W is selected, so that:

$$\underline{y}_k = \mathcal{H}_N W W^{-1} \underline{x} + \underline{w}_k \quad \text{for } k = 0, \dots, L-1, \quad (5.75)$$

$\mathcal{H}_N W$ and $W^{-1} \underline{x}$ are one of the possible solution obtained from the subspace decomposition of matrix R_y . It was proven in [57] that the ambiguity of matrix W can be removed by exploiting the finite-alphabet property of the transmitted signal.

However, going back to the original problem, assuming that despite the presence of interfering users only the first $M+N$ eigenvalues of matrix R_y are considered to correspond to the signal

⁴It was shown in Chapter 3 that the minimisation with respect to the noise subspace vectors is equivalent to the maximisation using signal subspace vectors

subspace, and the rest to the noise subspace, this would be equivalent to a representation of (5.70) of the form:

$$\underline{y}_k = \mathcal{H}_{N,1}\underline{x}_{1,k} + \mathcal{H}_{N,2}\underline{x}_{2,k} + \underline{w}_k \quad \text{for } k = 0, \dots, L-1, \quad (5.76)$$

where the rank of the received signal matrix R_y is $M + N$, and the effect of the interfering users comes in the form of correlated noise. As shown in the previous section for correlated noise, unless knowledge of the interfering user signal is known, the subspace method will be especially affected by the interference.

However, in the case of the cyclic subspace algorithm, for a given cyclic frequency α , the contribution from interfering users not exhibiting cyclostationarity at that particular cyclic frequency is negligible. As the number of received samples used in the estimation increases, the cyclic correlation of the interfering users, and possibly correlated noise, converges to zero. In that case, the channel impulse response of the SOI can be identified without any ambiguity factor, and the signal will be recovered. The received signal cyclic autocorrelation matrix can be expressed as:

$$R_y^\alpha = \mathcal{H}_{N,1} R_{x_1}^\alpha \mathcal{H}_{N,1}^H + \mathcal{H}_{N,2} R_{x_2}^\alpha \mathcal{H}_{N,2}^H. \quad (5.77)$$

Ideally, because the interfering user does not exhibit cyclostationarity at cyclic frequency α , $R_{x_2}^\alpha = 0$ and equation 5.77 reduces to:

$$R_y^\alpha = \mathcal{H}_{N,1} R_{x_1}^\alpha \mathcal{H}_{N,1}^H. \quad (5.78)$$

The cyclic subspace algorithm exploits the cyclic cross-correlation between the signals received and sampled at each of the subchannels. If the condition of cyclostationarity of the signals at each subchannel does not hold, the cyclic cross-correlation would be zero. The proposed cyclic subspace method would therefore only be applicable in those circumstances where:

- The SOI exhibits cyclostationarity at a particular cyclic frequency α and,
- no other interfering signal exhibits cyclostationarity at that particular cyclic frequency. This could be achieved, for instance, when signals have different symbol rates.

5.6 Simulation Results

In this section, some numerical results are presented. These results demonstrate the robustness of the cyclic subspace algorithm with respect to correlated noise and interfering users.

5.6.1 Experiment I

In this experiment a system represented by $y(k) = p(k) \otimes [f(k) \otimes x(k) + v(k)]$ is considered, where $f(k) = p(k) \otimes c(k)$ is the combined impulse response of transmitter filter and multipath propagation channel, $x(k)$ is a binary i.i.d. BPSK signal, and $v(k)$ is additive white Gaussian noise. The low-pass filter at the transmitter and the receiver, denoted $p(k)$, is defined as:

$$p(k) = \frac{\sin \pi k / K'}{\pi k / K'} \frac{\cos \beta \pi k / K'}{1 - (2\beta k / K')^2} \quad k = -3 \cdots 3, \quad (5.79)$$

where β is the roll-off factor. The period K' corresponds to the overall oversampling factor, i.e. $K' = KP$, which is the product between the number of subchannels $K = 2$ and the oversampling factor at each subchannel $P = 3$. The noise process $v(k)$ is correlated at the output of the filter $p(k)$. The impulse response of the propagation channel is given by:

$$c(k) = \delta(k) - 0.8\delta(k - 0.5) + 0.4\delta(k - 3). \quad (5.80)$$

The symbol period is $T = 0.75$ and the sampling period is $T_s = T/K'$.

The channel model considered here is a FIR channel of length q (at the symbol rate). This channel is then oversampled by a factor K' to generate the coefficients of the subchannels for the conventional subspace algorithm:

$$h^{(i)}(n) = h(i + nK'), \begin{cases} i = 0, 1, \dots, K' - 1 \\ n = 0, 1, \dots, q - 1 \end{cases} \quad (5.81)$$

where $h(k) = p(k) \otimes f(k)$ is the overall channel impulse response which comprises the transmitter and receiver filters plus the propagation channel. The channel model for the cyclic subspace algorithm on the other hand was selected so that:

$$h^{(i)}(n) = h(i + nK), \begin{cases} i = 0, 1, \dots, K - 1 \\ n = 0, 1, \dots, (q - 1)(KP - 1) \end{cases} \quad (5.82)$$

System identification results are shown in Figure 5.4. The simulations were performed for three different low-pass filters using roll-off factors of $\beta = 0.1$, $\beta = 0.5$ and $\beta = 1$. The performance of the channel identification is measured in terms of the *mean-square error* (MSE) of the channel coefficients estimates for 50 Monte-Carlo trials and is defined as:

$$MSE(dB) = 10 \log_{10} \frac{1}{M_c} \sum_{i=1}^{M_c} \sum_{k=1}^{K(M+1)} \|\hat{h}_i(k) - h(k)\|^2, \quad (5.83)$$

where M_c is the number of Monte-Carlo trials and $\hat{h}_i(\cdot)$ is the vector of the channel parameter estimates from the i th trial. The results presented in the first column correspond to the MSE associated with the channel estimation for different values of E_b/N_0 with a fixed length of 1000 samples used for the estimation of the channel. In the subspace algorithm, an overall oversampling factor of $K' = 6$ is chosen and in the cyclic subspace algorithm $K = 2$ subchannel elements are considered, and each subchannel is oversampled by a factor $P = 3$. This effectively means that the received signal is also oversampled by a factor of 6. The results show that the performance of the cyclic subspace algorithm improves with respect to the conventional subspace algorithm, with a reduction in noise power.

The second column of Figure 5.4 presents the results for a $E_b/N_0 = 25$ dB scenario, for different lengths of samples used in the estimation of the channel. It shows that, as mentioned in the theory, the cyclic algorithm needs a larger time to converge, but it can account for correlated noise better than the subspace algorithm of Moulines *et al.* .

Computational Cost

In terms of computational cost, both algorithms need a similar amount of computations except for the calculation of the autocorrelation/cyclic autocorrelation function. The algorithms involve three stages: *i*) estimation of the autocorrelation/cyclic autocorrelation matrix; *ii*) eigenvalue decomposition of a matrix of size $KN \times KN$ and ; *iii*) minimisation of the quadratic cost function subject to the unit norm constraint. In our case the minimisation was carried out by finding the unit-norm eigenvector associated with the smallest eigenvalue of a matrix of size $K(M + 1) \times K(M + 1)$ [61]. Stage *ii*) and *iii*) involve the computation of an EVD of a matrix of size 24×24 and 24×24 respectively. However, in stage *i*) the estimation of the cyclic autocorrelation matrix for a non-zero cyclic frequency α generally involves more complex multiplications than the estimation of the autocorrelation matrix. For example, the computation of the autocorrelation matrix requires $(L - N' - M' + 1)^{(2K'N')}$ complex multiplications where L is the number of source symbols required in the estimation of the channel. On the other hand, the estimation of the cyclic autocorrelation matrix involves $(2PL - 2N - 2M + 2)^{(2KN)}$ multiplications. This means that in general the estimation of the cyclic autocorrelation requires $\approx 2P^{2KN}$ times more complex multiplications than the estimation of the autocorrelation matrix.

5.6.2 Experiment II

In this experiment the performance of the cyclic subspace algorithm is tested with respect to co-channel interference. Two signals with symbol periods $T = 6$ and $T' = 8$ are transmitted simultaneously. The signals after passing through their respective propagation channels are sampled with a period $T_s = 1$. The overall oversampling factor of the SOI is $K' = T/T_s = 6$

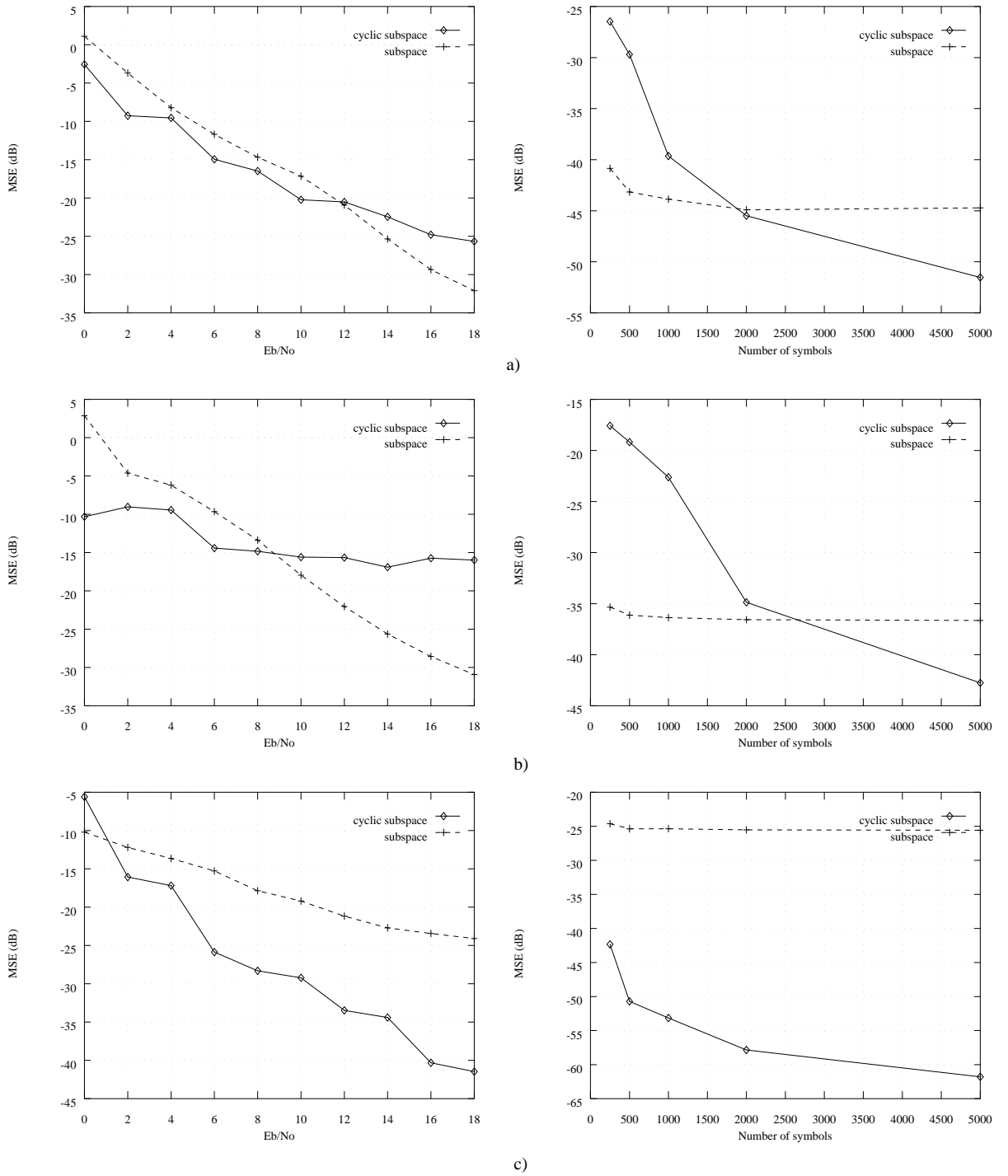


Figure 5.4: MSE of channel parameter estimates for a transmitter and receiver filters using a roll-off factor a) $\beta = 0.1$, b) $\beta = 0.5$ and c) $\beta = 1$, for different E_b/N_0 scenarios (left column, 1000 symbol used in the estimation), and for different number of symbols used in the estimation (right column, $E_b/N_0 = 25$ dB).

and $K'' = T'/T_s = 8$ for the SNOI. The subspace method is formulated taking into account the oversampling factor $K' = 6$ of the SOI and the cyclic subspace method is simulated so that $K = 2$ subchannels are oversampled by a factor $P = 3$ like in the previous experiment. The output samples of the SNOI are added to the appropriate output samples of the SOI to generate the oversampled received sequence. Simulations were conducted for different interference scenarios. The co-channel interference was measured in terms of the *signal-to-interference ratio* (SIR), defined as:

$$\text{SIR (dB)} = 10 \log \frac{\|\underline{h}_1\|^2 \sigma_1}{\|\underline{h}_2\|^2 \sigma_2}, \quad (5.84)$$

and the noise in terms of the *signal-to-noise ratio* (SNR), defined as [73]:

$$\text{SNR (dB)} = 10 \log \frac{\|\underline{h}_1\|^2 \sigma_1}{K' \sigma_n}, \quad (5.85)$$

where \underline{h}_1 and \underline{h}_2 are the channel impulse responses of the desired user and interferer respectively. σ_1 , σ_2 and σ_n are the variances of the SOI, SNOI and noise respectively. The true and estimated channels are shown in Figure 5.5 and Figure 5.6 for different SIR and SNR conditions. The mean value of the estimates over 20 realisations of the experiment and the true \pm standard deviation value are represented in Figures 5.7 and 5.8 respectively. Also, the MSE of the channel parameter estimates are shown in Figure 5.9.

It can be observed, that under co-channel interference effects, the performance of the conventional subspace algorithm of Moulines *et al.* degrades considerably, because as expected, the interference comes in the form of correlated noise. Unlike the conventional subspace method, the cyclic subspace method has been able to identify the channel of the desired user correctly even when the power of the interfering user is equal to that of the desired user. When the desired signal is buried in co-channel interference (SIR=-3 dB), the cyclic subspace algorithm, remarkably, manages to trace the channel, although the estimates of the channel exhibit high variance.

In terms of MSE, it is interesting to note that for SIR = -3 dB the cyclic subspace algorithm requires 3300 samples to achieve a better MSE than the conventional subspace method, as shown in Figure 5.9.a. It can also be seen that a smaller MSE can be achieved by increasing the data length with the cyclic subspace algorithm, whereas the subspace method cannot achieve a lower MSE. In the case of SIR = 0 dB and SIR = 3 dB conditions, the cyclic subspace algorithm requires less amount of data to estimate the channel correctly.

5.7 Discussion

The blind channel identification and equalisation approach presented in this chapter is one of many blind channel estimation techniques that exploits the cyclostationary nature of an over-

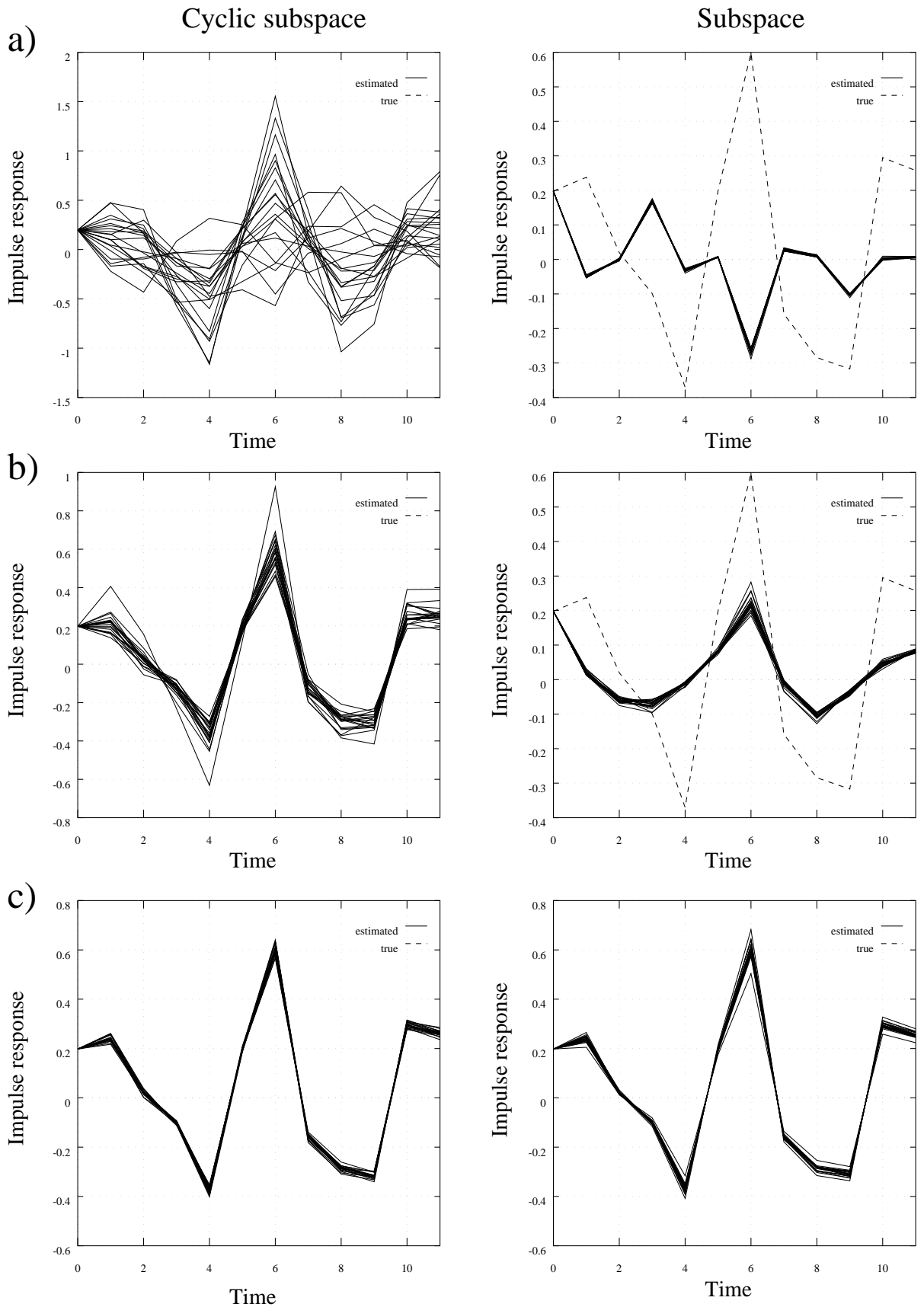


Figure 5.5: True and estimated channel of desired user for different SIR scenarios and $SNR = 100$ dB; a) $SIR = -3$ dB (6000 symbols used in the estimation), b) $SIR = 0$ dB (1000 symbols used in the estimation) c) $SIR = 3$ dB (1000 symbols used in the estimation).

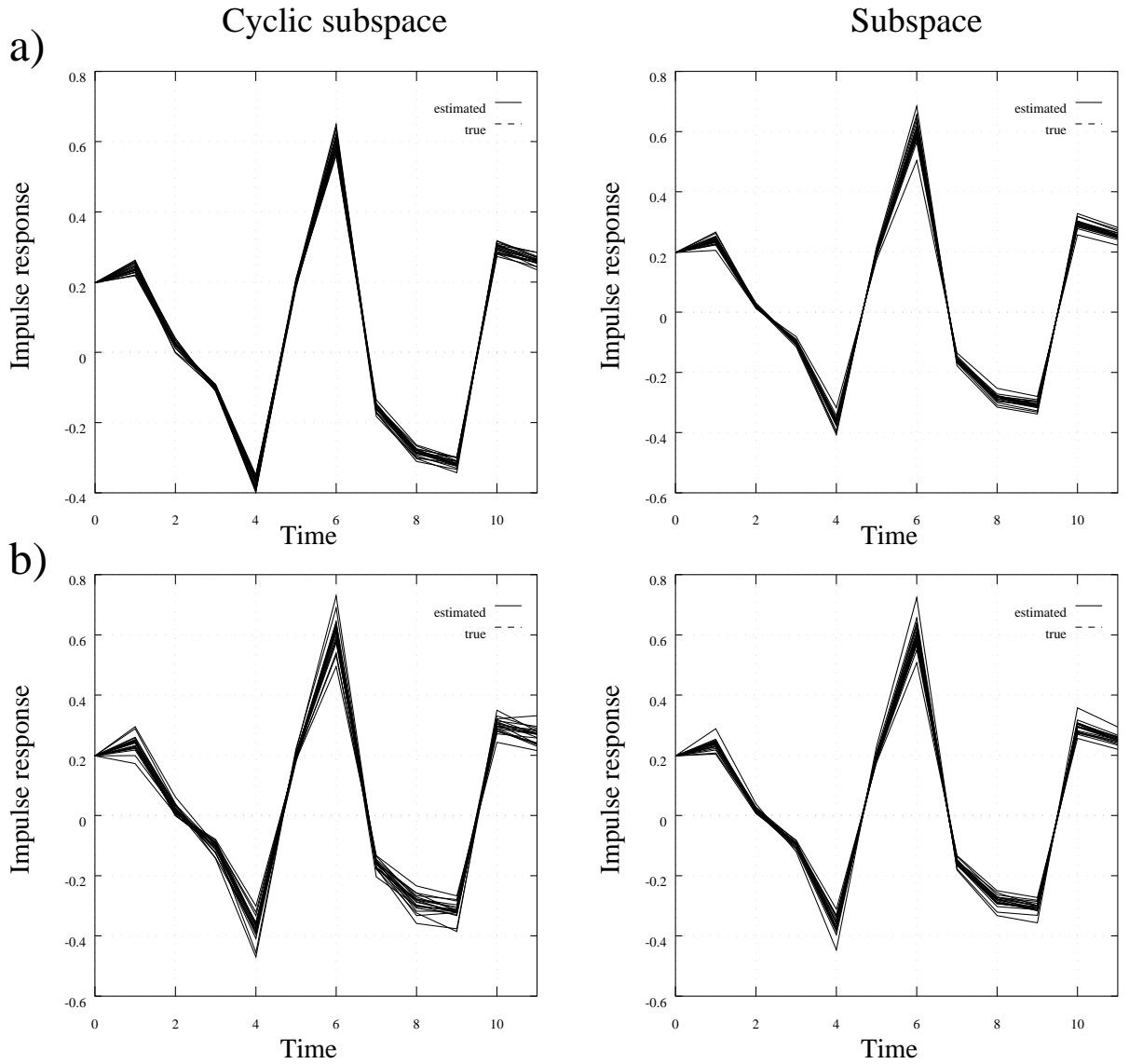


Figure 5.6: True and estimated channel of desired user for a $SIR = 3$ dB scenario, and a) $SNR = 30$ dB and b) $SNR = 15$ dB noise conditions. 1000 symbols were used in the estimation.

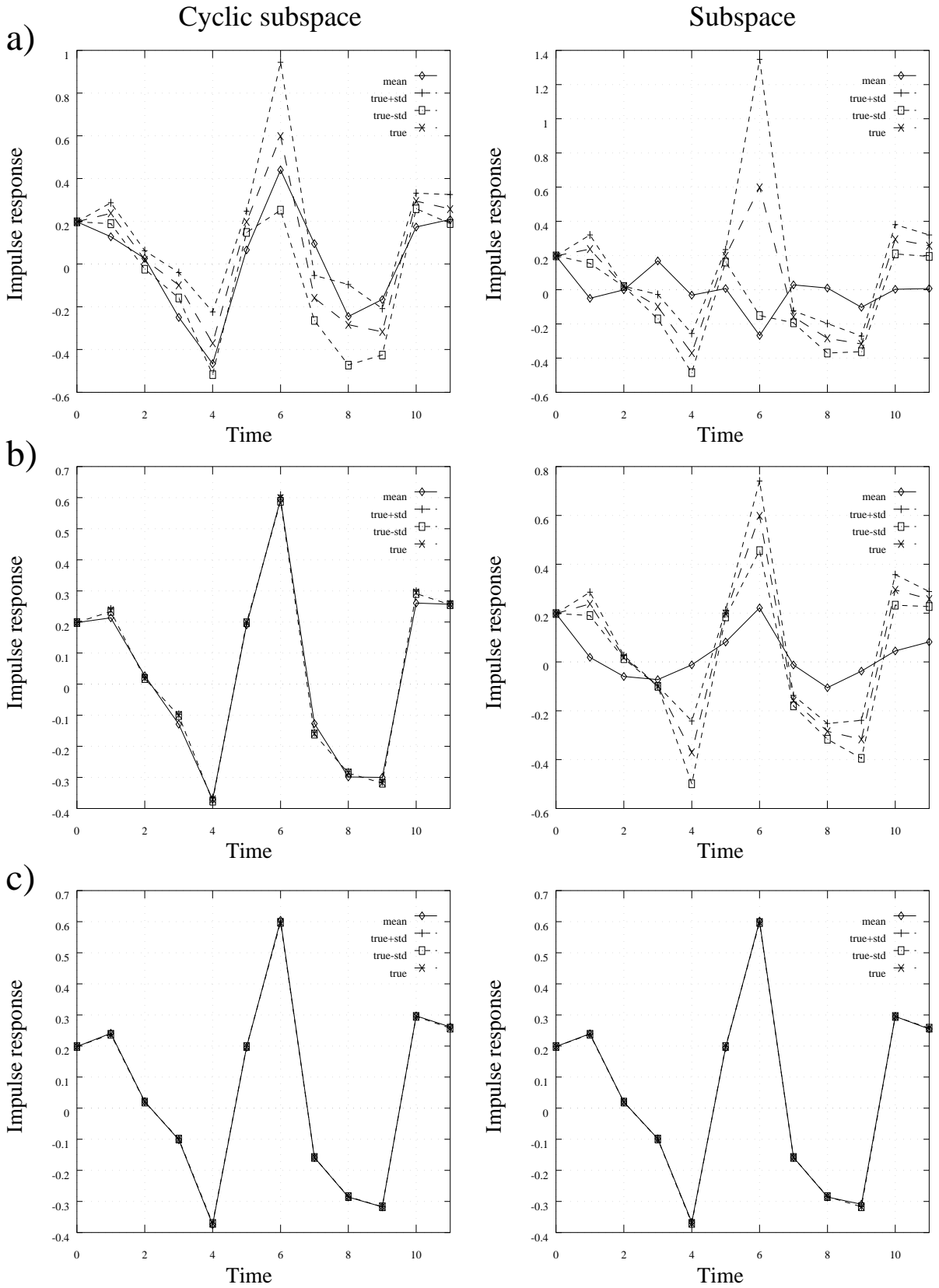


Figure 5.7: True, mean and true \pm standard deviation values of the estimation of desired user channel, for different SIR scenarios and SNR = 100 dB; a) SIR = -3 dB (6000 symbols used in the estimation), b) SIR = 0 dB (1000 symbols used in the estimation) c) SIR = 3 dB (1000 symbols used in the estimation).

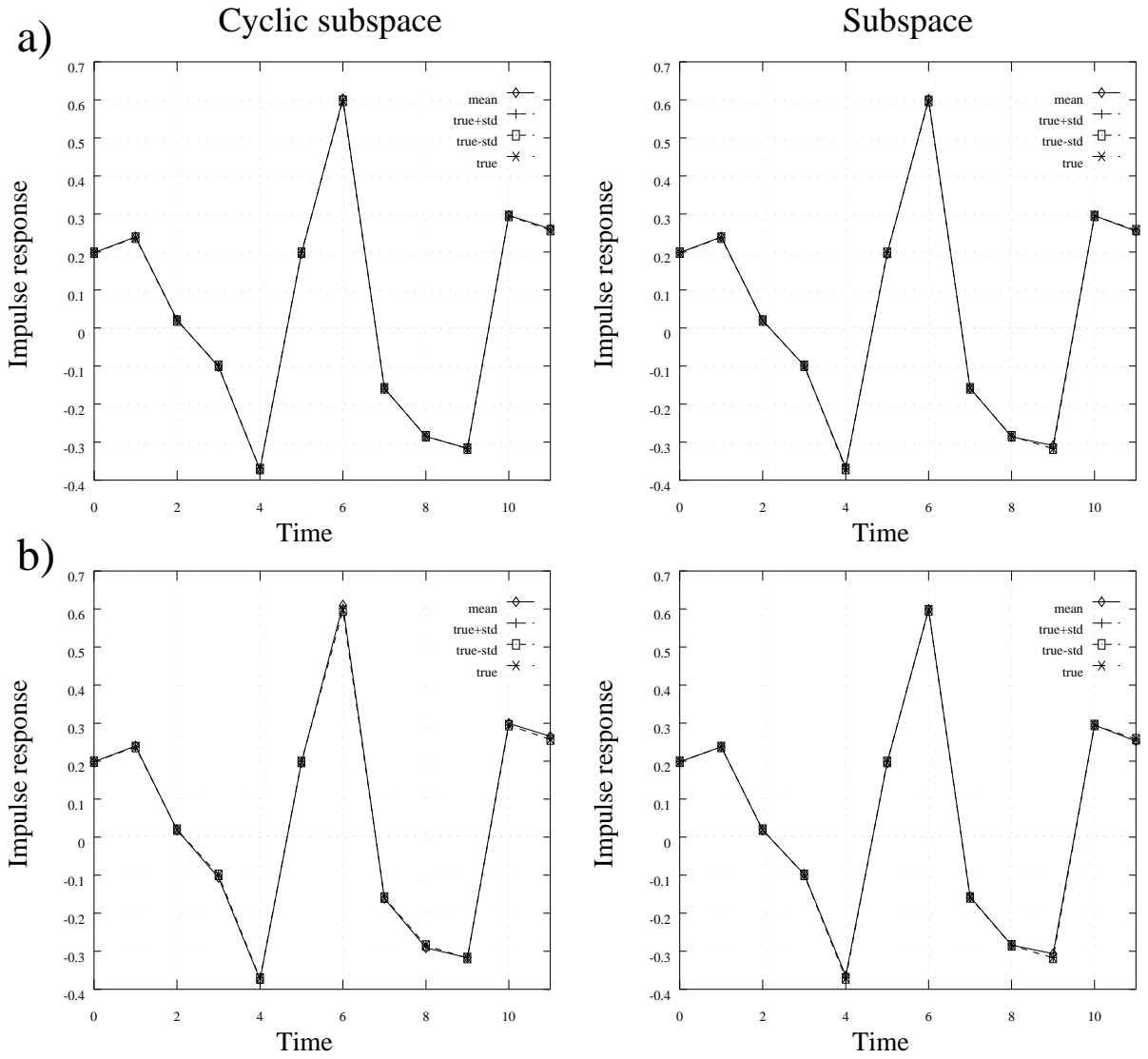


Figure 5.8: True, mean and true \pm standard deviation values of the estimation of desired user channel, for a SIR = 3 dB scenario, and a) SNR = 30dB and b) SNR = 15 dB noise conditions. 1000 symbols were used in the estimation.

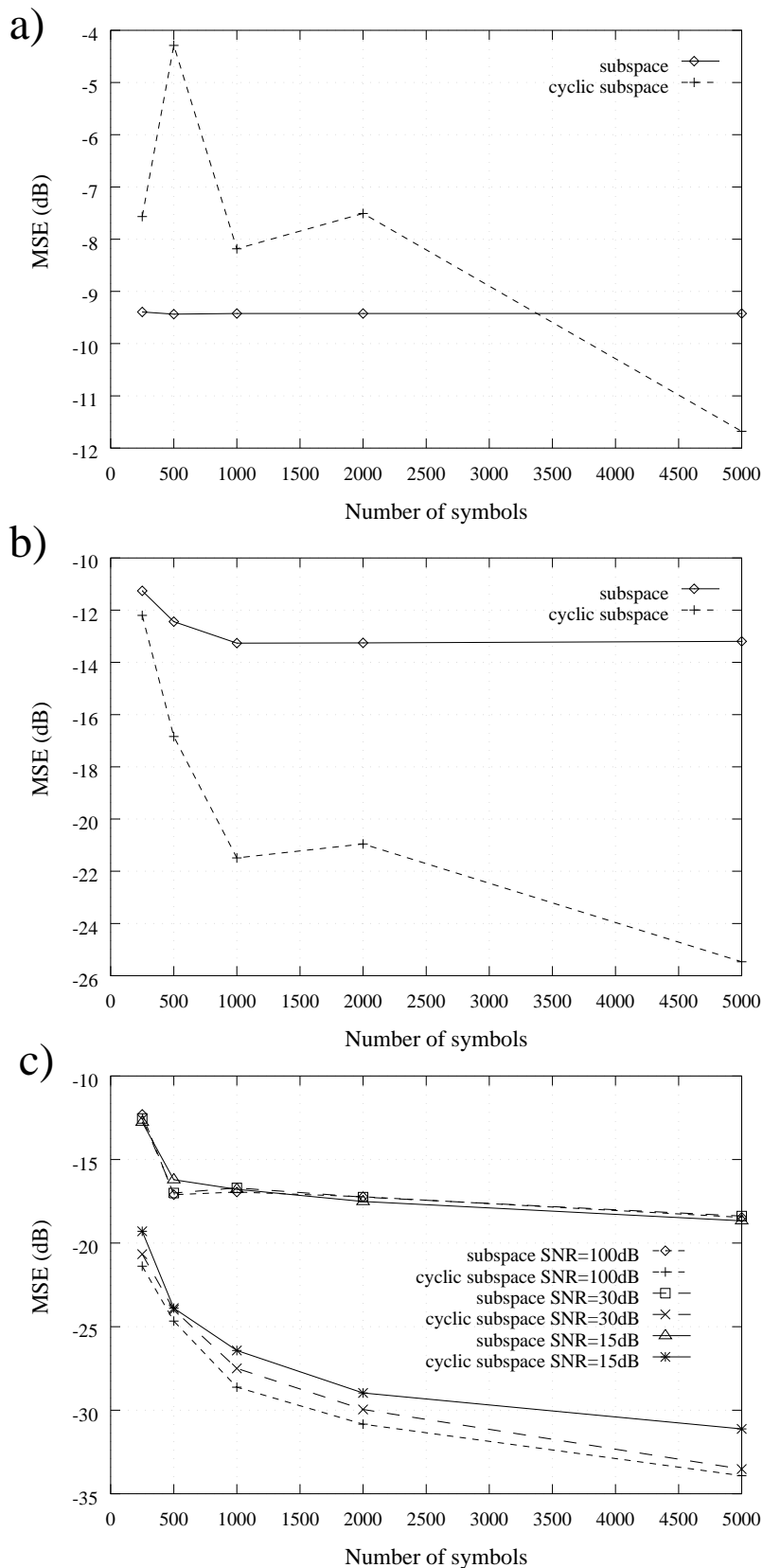


Figure 5.9: MSE of desired user channel parameter estimate for different interference levels; a) SIR = -3 dB, SNR = 100 dB, b) SIR = 0 dB, SNR = 100 dB and c) SIR = 3 dB.

sampled sequence. In fact, the cyclic subspace approach exploits cyclostationarity both implicitly and explicitly, since it is expressed as the combination of multichannel methods [87] [61] [81] and cyclostationary statistics based methods [45] [6]. The advantages and drawbacks of the proposed approach can be summarised as follows:

Advantages

- The cyclic subspace method can account for coloured and white noise.
- Provided that the desired user exhibits cyclostationarity and no other SNOI exhibits cyclostationarity at the cyclic frequency α of the SOI, the proposed approach is less sensitive to co-channel interference than the conventional subspace methods.

Drawbacks

- Channels are unidentifiable if the subchannels share common zeros. The main source for this critical condition is the overestimation of the channel model order. However, this condition is present in all fractionally-spaced blind equalisation algorithms.
- Due to the use of the measured cyclic autocorrelation function, a longer integration time might be required so that the effect of noise and other interference is negligible.

Recently, a number of subspace methods which exploit the cyclostationarity nature of an over-sampled sequence or modulated signals have been proposed. In [23], a cyclic statistics based blind channel identification method was presented using subspace fitting. This approach is derived from the conventional subspace fitting approach of Slock [81], whereas the cyclic subspace method proposed in this thesis is derived from the subspace method of Moulines *et al.* [61]. Both methods exploit the fact that sufficient subspace information is contained in the received signal cyclic autocorrelation.

In [15], another subspace method was proposed. Instead of exploiting the cyclostationarity of signals exhibiting different baud or symbol rates, they force the cyclostationarity by *modulating* the input signal by a deterministic, almost periodic sequence. This technique exploits the symmetry property of two cyclic spectra functions of the channel, and can be solved using a structured subspace technique. The advantage of the proposed method is that it is not affected by any of the critical conditions of the position of the zeros of the channel, shared by all fractionally-spaced blind equalisation method. In fact, it can achieve channel identification even if the channel model order is over-determined. The disadvantage in forcing the cyclostationarity of the input signal is lower spectral efficiency.

5.8 Conclusions

This chapter presents a new approach for blind channel identification in the presence of correlated noise. The use of cyclostationary statistics, which have been shown to be less sensitive than conventional second-order statistics to correlated noise, is believed to be a factor which will be exploited in the development of new blind channel identification algorithms. The algorithm is reminiscent of [61] and it has shown to be not as sensitive to correlated noise as the conventional subspace algorithm for a similar computational complexity. Thus, no pre-processing or whitening filter is required at the expense of a small increase in the number of complex multiplications required to estimate the cyclic autocorrelation matrix compared to the autocorrelation matrix.

The approach not only improves the transmission efficiency of the system by eliminating the training sequence, but it also accounts for co-channel interference in an efficient manner. If the signal from the desired user exhibits cyclostationarity, and no other interfering signal exhibits cyclostationarity at the same cyclic frequency α , the effect of the interfering users in the measured cyclic autocorrelation and cyclic spectrum functions becomes negligible. Nevertheless, it becomes imperative that the source signals exhibit cyclostationarity. This does not constitute a serious problem since modulated signals exhibit cyclostationarity at harmonics of their carrier frequencies and baseband signals also exhibit cyclostationarity at harmonics of their baud-rates.

Chapter 6

Conclusions

6.1 Introduction

The work described in this thesis is primarily concerned with the analysis of cyclostationary techniques for blind equalisation in the mobile radio environment, and with the development of a new blind channel identification method based on cyclostationary statistics and subspace methods. In the analysis of cyclostationary blind equalisation techniques in the mobile radio environment, the convergence properties and tracking difficulties of existing methods have been discussed, and this performance has been compared to supervised equalisation techniques. The proposed blind channel identification and equalisation technique estimates the channel without a training or pilot signal, with the information obtained through the measurement of the cyclic correlation function of the oversampled received signal. The method is less sensitive to correlated noise and co-channel interference than conventional subspace methods using correlation measures. This chapter summarises the main conclusions of the work, establishes the limitations found in the experimental analysis and points out the advantages and limitations of the method proposed. Finally, the chapter concludes with some pointers towards future work.

6.2 Achievements of the Work

For the problem of cyclostationary blind channel identification and equalisation, statistical approaches and algebraic methods have been studied in Chapter 3 and Chapter 4. The statistical methods, based on cyclostationary statistics, and algebraic methods, which exploit some rich matrix structure obtained from the multichannel representation of an oversampled received sequence, are described in an unified notational framework. The important issue of identifiability was also addressed in Chapter 3, and the basic identifiability conditions of each family of algorithms have been found to be equivalent. Chapter 4 considers the convergence behaviour of these techniques. The analysis shows that the convergence behaviour of these techniques is in fact very similar, but fundamental identifiability problems arise, in particular in the cyclic statistics based algorithm analysed, which can be overcome by selecting an appropriate oversampling factor. The differential cepstrum parameters of the channel transfer function have been proposed, as an efficient measure of identifiability for a particular oversampling factor.

Chapter 4 is primarily devoted to the comparison of cyclostationary blind equalisation techniques with supervised equalisation methods in a mobile radio environment, described by the

COST207 model and used for simulating a GSM system. Different modulation schemes and different transmission rates have been considered. In general, it has been shown that the 156 bits of one TDMA time-slot are insufficient to estimate the time-varying channel accurately using cyclostationary blind channel identification techniques. Unless the convergence speed of the blind algorithms reduces to within 100 symbols in this environment, supervised methods have shown to provide a better performance. Yet, under high Doppler frequency conditions, these methods can constitute a valid alternative. The computational complexity of cyclostationary blind channel identification methods, combined with a MLSE in the receiver is also a matter of concern, since apart from the estimation of the correlation functions, one or two eigenvalue-decompositions are sometimes required. It has also been shown that using a weighted autocorrelation function improves the performance of the estimation of the channel, to the point that a switch to a decision-directed mode can be established at high Doppler frequencies.

In Chapter 5, a new blind channel identification algorithm was presented. This method combines cyclostationary statistics and subspace methods for blind channel identification. The algorithm is known as the cyclic subspace method and it is shown that, the cyclic autocorrelation matrix contains the sufficient subspace information to identify the channel. It exploits the orthogonality between the signal and null subspaces, similarly to the conventional subspace methods. As a difference from other subspace methods, the received signal at each of the sensors must exhibit cyclostationarity. This can easily be achieved by means of oversampling at each of the subchannels. The interesting feature of the cyclic subspace algorithm is the fact that the cyclic autocorrelation function of the noise and interference from adjacent users is negligible, at the specific cyclic frequency of the desired user. White or coloured noise, indistinctively, does not exhibit cyclostationarity, and if a sufficiently large amount of data is available for estimation, its effect on the received signal cyclic autocorrelation function will disappear. This indicates that the proposed cyclic subspace approach is less sensitive than conventional subspace methods to correlated noise. Since most digital communication systems use a low-pass filter at the receiver, no whitening filter would be required to account for the correlated noise, with the consequent reduction in the computational complexity of the algorithm. Numerical results using different receiver filters indicate that for a similar computational complexity, the cyclic subspace algorithm achieves a more than acceptable performance.

As far as the problem of blind multiuser channel identification is concerned, most adaptive blind and supervised spatial filtering techniques and multiuser signal detection techniques fail in the presence of multipath propagation. The fractionally-spaced blind channel identification techniques studied in this thesis are adequate to combat the ISI present due to multipath. Provided that the signals from interfering users do not exhibit cyclostationarity at the same cyclic frequency of the SOI, the effect of the co-channel interferers will be again negligible. This can be achieved when the signals have different baud rates or carrier frequencies. The conventional blind multiuser channel identification methods would generally require the estimation of the channels of the desired users and those of the undesired users too. In the way these are

formulated, an ambiguity factor is present in the estimated channel, due to the fact that the subspaces have contributions from all the signals. I have proposed a formulation of the problem, where co-channel interference is treated as noise. The necessary and sufficient identifiability conditions have been established and a measure of the asymptotic statistical performance has been proposed.

6.3 Limitations of the Work

The analysis of the performance of cyclostationary blind equalisation techniques gives a proof of the fundamental unidentifiability situations that are encountered. Channels that have a number of zeros equal to the oversampling factor, symmetrically located in the unit circle, are unidentifiable using fractionally-spaced blind equalisation algorithms; equivalently, if the subchannels of the multichannel representation of the oversampled received sequence share common zeros, the channel is unidentifiable. Although this situation is rare, and can be overcome by selecting a different oversampling factor, this condition may arise if the model order of the channel is overestimated by more than the oversampling factor.

The blind channel estimation method proposed in this thesis combines spatial and temporal diversity. The approach considered describes a system of an array of oversampled subchannels. The total oversampling factor is the multiplication of the number of subchannels and the oversampling factor of each of the subchannels. In general, channel identification could be achieved with a smaller oversampling factor in the case of conventional subspace methods. Furthermore, because cyclostationary statistics make no distinction between white and coloured noise, in general, more data would be required for the effect of the noise to disappear from the cyclic autocorrelation function. Under white noise conditions, conventional subspace methods perform better than the cyclic subspace method.

6.4 Areas for Future Work

To conclude this thesis, I provide some pointers to further areas of development and suggest some alternative applications.

The work presented in this thesis has concentrated on the analysis of cyclic statistics-based and multichannel blind channel identification methods. A series of Maximum-Likelihood (ML) methods have recently been proposed [48] [51] [18] [21]. The solution provided by the ML methods is optimum, but the computational complexity of the likelihood function is tremendous. However, gradient algorithms can be used instead, with a significant reduction in computational complexity.

Another issue that requires further investigation is the problem with singular channels which cannot be identified using fractionally-spaced blind equalisation techniques. This problem can be overcome with the use of spectral factorisation [15]. It requires the estimation of the spectral

correlation function at two cyclic frequencies, and the symmetry that derives from it, allows the identification of the channel irrespective of the position of the zeros of the channel, even if the channel order is overestimated.

The results presented on the performance of cyclostationary blind equalisers in the mobile radio environment have dealt with a series of problems associated with equalisation of time-varying channels, using an approximate model for GSM. Other modulation schemes, especially nonlinear ones, such as GMSK have to be analysed and its impact on blind equalisation algorithms tested. Already, some methods have been proposed for nonlinearly modulated signals that can operate with short data lengths [18]. Other simulation environments, besides the typical urban environment, need to be investigated too. For instance, the bad urban environment, where the presence of multipath effects are more severe than in the TU environment would be an interesting case.

Some areas of application have been suggested for blind equalisation. It is interesting to see that the application of blind equalisation has been studied and proposed for HF applications [102], where the environment is characterised by time-varying effects. It has also been proposed for HDTV, because of the loss in efficiency from transmitting a training sequence. The particular channels that are being allocated for HDTV were previously classified as very bad because of the high levels of interference. This suggests that it can result beneficial to use blind equalisation under these circumstances [17].

Appendix A

Identifiability Conditions

In this appendix the necessary and sufficient conditions for blind identifiability of digital communications channels for methods using fractionally-spaced sampling are established.

A.1 Identifiability condition I

This identification condition is fundamental to all blind equalisation methods using fractionally-spaced sampling. It states that a channel is unidentifiable using cyclostationary statistics if it has P zeros symmetrically located on the unit circle with the centre as the reference, where P is the oversampling factor. Lets consider the transfer function of a linear time-invariant channel $H(z_\Delta)$ that has a set of P zeros, i.e. $\rho_0, \dots, \rho_{P-1}$, symmetrically located on the unit circle where $z_\Delta = e^{j\omega}$ indicates that the channel transfer function is evaluated at a sampling rate of $T_s = T/P$:

$$\begin{aligned}
 H(z_\Delta) &= (1 - \rho_0 z_\Delta^{-1})(1 - \rho_1 z_\Delta^{-1}) \cdots (1 - \rho_{P-1} z_\Delta^{-1}) H'(z_\Delta) \\
 &= H'(z_\Delta) \prod_{k=0}^{P-1} (1 - \rho_k z_\Delta^{-1}) \quad |\omega| < \frac{\pi}{T_s}
 \end{aligned} \tag{A.1}$$

The set of symmetric zeros ρ_k can be represented as a complex rotating phasor as:

$$\rho_k = |\rho| e^{j(\theta + 2\pi k)/T} \quad \text{for } k = 0, \dots, P-1 \quad \text{and } |\theta| < \pi \tag{A.2}$$

In Chapter 3 it was shown that the identification of nonminimum phase channels was possible due to the ability of the cyclic spectrum (and in general cyclostationary statistics) to preserve the phase information of the system which was not the case for conventional second-order statistics (autocorrelation and power spectrum). Here, it will be demonstrated that maximum phase and minimum phase zeros cannot be distinguished if a set of P zeros are present uniformly spaced around the unit circle. As shown in section 3.4.2, the transfer function of the cyclic spectrum of the output of a LTI filter sampled at T_s can be expressed as:

$$S_y^\alpha(z_\Delta) = q_x H(z_\Delta) H^*(z_\Delta z_\alpha) \tag{A.3}$$

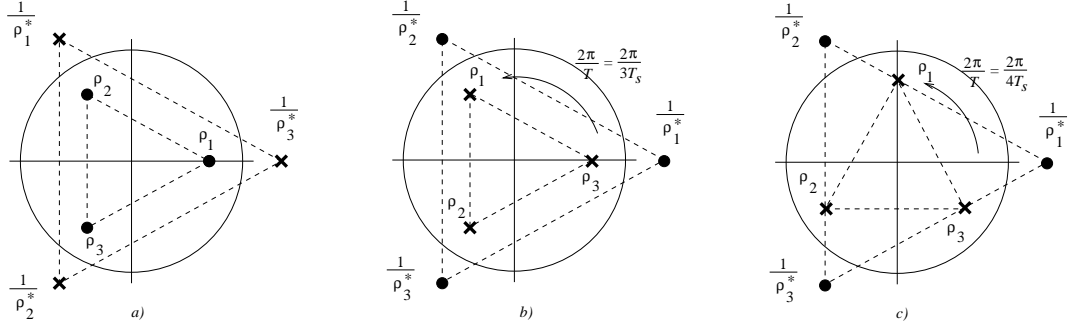


Figure A.1: Positions of zeros of the components of the spectral correlation density (SCD) function; a) (cross) zeros of $\prod_{k=0}^{P-1} (1 - \rho_k z_{\Delta}^{-1})$, (circle) zeros of $\prod_{k=0}^{P-1} (1 - \rho_k^* z_{\Delta} e^{-j2\pi/T})$; b) and c) (circle) zeros of $\prod_{k=0}^{P-1} (1 - \rho_k^* z_{\Delta})$, (cross) zeros of $\prod_{k=0}^{P-1} (1 - \rho_k z_{\Delta}^{-1} e^{-j2\pi/T})$.

where $z_{\alpha} = e^{-j2\pi\alpha}$ and the superscript * conjugates the transfer function and the variables z_{Δ} and z_{α} . Using (A.1) in (A.3) we obtain

$$S_y^{\alpha}(z_{\Delta}) = q_x H'(z_{\Delta}) \prod_{k=0}^{P-1} (1 - \rho_k z_{\Delta}^{-1}) H'^*(z_{\Delta} z_{\alpha}) \prod_{k=0}^{P-1} (1 - \rho_k^* z_{\Delta} z_{\alpha}) \quad (\text{A.4})$$

$$= q_x H'(z_{\Delta}) H'^*(z_{\Delta} z_{\alpha})$$

$$\times \prod_{k=0}^{P-1} (1 - \rho_k z_{\Delta}^{-1}) \prod_{k=0}^{P-1} (1 - \rho_k^* z_{\Delta} e^{-j2\pi l/T}) \quad \text{for } l \neq nP \neq 0 \quad n \in \mathcal{Z} \quad (\text{A.5})$$

The frequency-shift due to the cyclic frequency α , i.e. $e^{-j2\pi l/T}$, rotates the conjugate zero clockwise by the same phase difference between the symmetric zeros. As a result, it is not possible to know from the observation of $S_y^{\alpha}(z_{\Delta})$, whether ρ_k is a zero of the channel or $1/\rho_k^*$ is. In this case, *i*) the product $\prod_{k=0}^{P-1} (1 - \rho_k z_{\Delta}^{-1}) \prod_{k=0}^{P-1} (1 - \rho_k^* z_{\Delta} e^{-j2\pi/T})$ for the case of channel $H(z_{\Delta})$ having zeros at ρ_k and *ii*) the product $\prod_{k=0}^{P-1} (1 - \rho_k^* z_{\Delta}) \prod_{k=0}^{P-1} (1 - \rho_k z_{\Delta}^{-1} e^{-j2\pi/T})$ corresponding to $H(z_{\Delta})$ having zeros at $1/\rho_k^*$ are identical, as shown in Figures A.1.a and A.1.b. However, if the sampling rate is changed the zeros of the channel can still be recovered, as shown in Figure A.1.c.

The product $\prod_{k=0}^{P-1} (1 - \rho_k z_{\Delta}^{-1})$ of the transfer functions of the P zeros of equation A.1, can be written as:

$$H(z_{\Delta}) = H'(z_{\Delta}) \prod_{k=1}^P (1 - \rho_k z_{\Delta}^{-1}) = H'(z_{\Delta}) (1 - \rho_0^P z_{\Delta}^{-P}) = H'(z_{\Delta}) (1 - |\rho|^P e^{jP\theta/T} z_{\Delta}^{-P}) \quad (\text{A.6})$$

Channel $H(z_{\Delta})$ is therefore unidentifiable if $|\rho|^P e^{jP\theta/T} z_{\Delta}^{-P} = 1$, or equivalently:

$$z_{\Delta}^P = e^{j\omega P} = |\rho|^P e^{jP\theta/T} = |\rho|^P e^{j\theta/T_s} \quad (\text{A.7})$$

and

$$\omega = -j \ln |\rho| + \frac{\theta}{T} \quad \omega \in \left[-\frac{\pi}{T}, \frac{\pi}{T} \right] \quad (\text{A.8})$$

which indicates that if the channel $H(\omega)$ has a null for an specific θ in the range $-\pi/T < \omega < \pi/T$, the channel will also exhibit nulls at

$$\omega = -jP \ln |\rho| + \frac{\theta}{T_s} + \frac{2\pi k}{T} \bmod \frac{2\pi P}{T} \quad \text{for } k = 0, \dots, P-1 \quad (\text{A.9})$$

in the range $\omega \in [-\pi/T_s, \pi/T_s]$. This condition means that the zero in the range $\omega \in [-\pi/T, \pi/T]$ repeats itself P times in the range $\omega \in [-\pi/T_s, \pi/T_s]$.

This is the sufficient and necessary condition for band-limited channels to be unidentifiable from SOCS. An alternative formulation of the identifiability condition I can be established from a multichannel representation of the channel transfer function $H(z_\Delta)$, where as shown in section 3.4.1 the P subchannels derived as a result of oversampling the received signal P times faster than the baud-rate can be expressed as:

$$H(z_\Delta) = \sum_{n=0}^{P-1} z_\Delta^{-n} H_n(z) = \sum_{n=0}^{P-1} z_\Delta^{-n} H_n(z_\Delta^P) \quad (\text{A.10})$$

where $z = z_\Delta^P$. Now suppose that each subchannel $H_n(z)$ has a common zero ρ' for every value of $n = 0, \dots, P-1$. Thus, we may write

$$H_n(z) = (1 - \rho' z^{-1}) H'_n(z) \quad (\text{A.11})$$

Substituting (A.11) in (A.10) leads to

$$H(z_\Delta) = \sum_{n=0}^{P-1} z_\Delta^{-n} (1 - \rho' z_\Delta^{-P}) H'_n(z_\Delta^P) \quad (\text{A.12})$$

and

$$H(z_\Delta) = (1 - \rho' z_\Delta^{-P}) \sum_{n=0}^{P-1} z_\Delta^{-n} H'_n(z_\Delta^P) = (1 - \rho' z_\Delta^{-P}) H''(z_\Delta) \quad (\text{A.13})$$

Equation A.13 is identical to A.6 and the term $(1 - \rho' z_\Delta^{-P})$ on the left hand side of (A.13) is analog to the product of the transfer functions of P zeros equally spaced around the unit circle. The result is that a zero $\rho' = e^{j\theta}$ common in the P subchannels $H_0(z_\Delta^P), \dots, H_{P-1}(z_\Delta^P)$ is equivalent to $H(z_\Delta)$ having P zeros $\rho'_k = |\rho'|^{1/P} e^{j(\theta+2\pi k)/T}$ for $k = 0, \dots, P-1$ [93].

A.2 Identifiability condition II

The second identifiability condition is related to the unidentifiability of channels with echoes which are multiples of the fundamental period T . Lets consider the transfer function of a LTI impulse response h_j which takes non-zero values at discrete samples which correspond to the fundamental period T , and are zero elsewhere:

$$H(z_\Delta) = \sum_{j=0}^{PM+P-1} h_j z_\Delta^{-j} \quad \begin{cases} h_j \neq 0 & \text{if } j \bmod P = 0 \\ h_j = 0 & \text{otherwise} \end{cases} \quad (\text{A.14})$$

Alternatively, since the channel transfer function oversampled by a factor P can be represented as a multichannel system, we may write:

$$\begin{aligned} H(z_\Delta) &= \sum_{i=0}^{P-1} \sum_{j=0}^M h_{i+jP} z_\Delta^{-(i+jP)} \\ &= \underbrace{\sum_{j=0}^M h_{jP} z_\Delta^{-jP}}_{H_0(z_\Delta^P)} + z_\Delta^{-1} \underbrace{\sum_{j=0}^M h_{jP+1} z_\Delta^{-jP}}_{H_1(z_\Delta^P)} + \dots + z_\Delta^{-P+1} \underbrace{\sum_{j=0}^M h_{jP+P-1} z_\Delta^{-jP}}_{H_{P-1}(z_\Delta^P)} \end{aligned} \quad (\text{A.15})$$

From (A.14), since only one of the subchannel transfer functions $H_n(z_\Delta^P)$ is non-zero, the channel transfer function can finally be expressed as:

$$H(z_\Delta^P) = H_k(z_\Delta^P) = z_\Delta^{-k} \sum_{j=0}^M h_{jP+k} z_\Delta^{-jP} \quad H_n(z_\Delta^P) = \begin{cases} 0 & \text{for } n \neq k \\ H_k(z_\Delta^P) \neq 0 & \text{for } n = k \end{cases} \quad (\text{A.16})$$

and

$$H(z_\Delta^P) = z_\Delta^{-k} \sum_{j=0}^{M-1} (1 - \rho_j z_\Delta^{-P}) \quad (\text{A.17})$$

This can be interpreted as a system oversampled by a factor of P with P zeros equally spaced around the unit circle with separation $2\pi/T$ and as a result unidentifiable using SOCS. In this case, the problem does not disappear by selecting a different oversampling factor.

As shown in Figure A.2, a channel with one tap at T is represented by a single zero (Figure A.2.a). If the channel is sampled at $T/3$ (Figure A.2.b), replicas of the zero will appear each with separation $2\pi/T$. The channel is clearly unidentifiable because the condition for unidentifiability I is encountered. Similarly, for an oversampling factor of $P = 4$ (Figure A.2.c), 4 zeros are found equally-spaced around the unit circle.

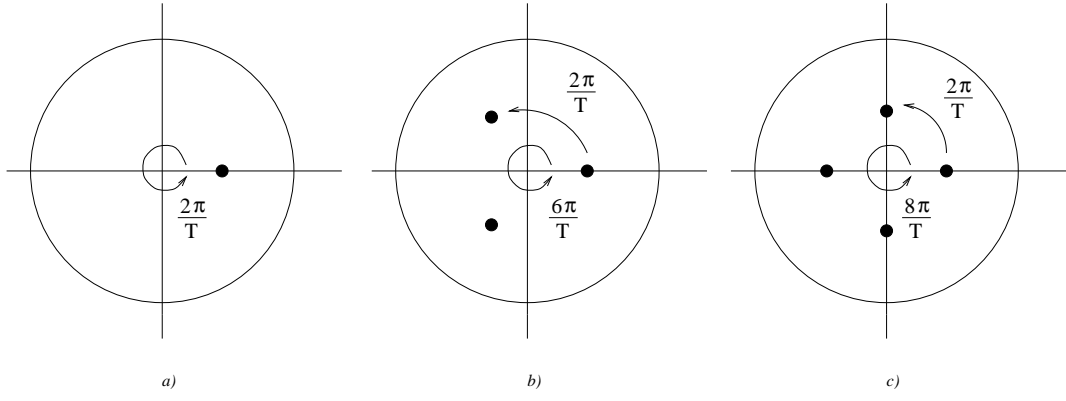


Figure A.2: Position of the zeros for a single tap channel with a reflector occurring at the symbol period T . a) position of the zero at the symbol rate; b) position of the zeros for an oversampling factor $P = 3$; c) position of the zeros for an oversampling factor $P = 4$.

A.3 Identifiability condition III

The third identifiability condition states that channels with echo delays which happen at integer multiples of $T/2$ are unidentifiable when the oversampling factor P is even. The transfer function of this multipath channel is expressed as:

$$\begin{aligned}
 H(z_\Delta) &= \sum_{j=0}^{PM} h_j z_\Delta^{-j} = \sum_{i=0}^{P-1} z_\Delta^{-i} \left(\sum_{k=0}^M h_{ki} z_\Delta^{-k} + \sum_{k=0}^{M-1} h_{P(2k+1)/2} z_\Delta^{-P(2k+1)/2} \right) \\
 &= \sum_{i=0}^{P-1} z_\Delta^{-i} H_i(z_\Delta^P) \quad \begin{cases} h_j \neq 0 & \text{if } j \bmod P/2 = 0 \\ h_j = 0 & \text{otherwise} \end{cases} \quad (\text{A.18})
 \end{aligned}$$

and

$$\begin{aligned}
 H(z_\Delta) &= \sum_{i=0}^{2M} h_{iP/2} z_\Delta^{-iP/2} = \sum_{k=0}^M h_{kP} z_\Delta^{-kP} + \sum_{n=0}^{M-1} h_{P(2n+1)/2} z_\Delta^{-P(2n+1)/2} \\
 &= \sum_{k=0}^M h_{kP} z_2^{-kP/2} + \sum_{n=0}^{M-1} h_{P(2n+1)/2} z_2^{-(2n+1)} = \sum_{j=0}^1 z_2^{-j} H_j(z_2^2) \quad (\text{A.19})
 \end{aligned}$$

$$\begin{aligned}
 H(z_\Delta) &= \sum_{i=0}^{2M} h_{iP/2} z_\Delta^{-iP/2} = \sum_{k=0}^M h_{kP} z_\Delta^{-kP} + \sum_{n=0}^{M-1} h_{P(2n+1)/2} z_\Delta^{-P(2n+1)/2} \\
 &= \sum_{k=0}^M h_{kP} z_2^{-kP/2} + \sum_{n=0}^{M-1} h_{P(2n+1)/2} z_2^{-(2n+1)} = \sum_{j=0}^1 z_2^{-j} H_j(z_2^2) \quad (\text{A.20})
 \end{aligned}$$

where we have used $z_2 = z_\Delta^2$ to denote that the variable z is evaluated at the sampling period $T/2$. If the transfer function of the channel is represented in a multichannel form with an *odd*

and *even* subchannel transfer functions. Following (A.10), these can be defined as:

$$H(z_\Delta) = \sum_{n=0}^1 z_2^{-n} H_n(z_2^2) = H_e(z_\Delta^P) + z_2^{-1} H_o(z_\Delta^P) \quad (\text{A.21})$$

Therefore, if there exists a root ρ common to $H_e(z_\Delta^P)$ and $H_o(z_\Delta^P)$ so that $H_e(\rho) = H_o(\rho) = 0$, and the channel $H(z_\Delta)$ has a pair of roots $\pm \rho_0$, it follows that this pair of zeros becomes unidentifiable, irrespective of the oversampling factor P is, because $H_e(\rho_0^2) = H_o(\rho_0^2) = 0$. Hence, if $H(z_\Delta)$ has a pair of zeros $\pm \rho_0$, this becomes equivalent to the *odd* and *even* subchannels sharing a common zero ρ_0^2 .

Singular Value Decomposition

The *singular value decomposition* (SVD) is one of the most widely used algorithms for solving linear system of equations in the least-squares sense. It is based on the idea that a $M \times N$ matrix X can be expressed as the multiplication of a $M \times N$ column-orthogonal matrix U , a $N \times N$ diagonal matrix Σ and the transpose of a $N \times N$ column-orthogonal matrix V :

$$X = U \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_N \end{bmatrix} V^H, \quad (\text{B.1})$$

where $\Sigma = \text{diag}(\sigma_1 \cdots \sigma_N)$ are some non-negative elements arranged in descending order according to their values, i.e. $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_N$. These elements are known as the *singular values* of matrix X . The columns of both U and V are orthogonal,

$$\sum_{i=1}^M U_{ik}^* U_{in} = \delta_{kn} \quad 1 \leq k, n \leq N, \quad (\text{B.2})$$

$$\sum_{j=1}^N V_{jk}^* V_{jn} = \delta_{kn} \quad 1 \leq k, n \leq N, \quad (\text{B.3})$$

and

$$\delta_{kn} = \begin{cases} 1 & \text{if } k = n \\ 0 & \text{otherwise} \end{cases} \quad (\text{B.4})$$

is the Kronecker delta. Both equations can be written in matrix form as $U^H U = V^H V = I$, where I is the identity matrix.

The good thing about the SVD algorithm is that it always provides a single good solution even if more than one solution exist and in those systems where a single solution is possible it provides the unique and optimum solution. Although the SVD algorithm can be used for

systems of equations with more equations than unknowns and systems with more unknowns than equations, here, the case of the number of equations being equal to the number of unknowns is considered only. This condition is found when matrices are square.

The SVD routine used in this work is the widely recognised LINPACK algorithm [27]. SVD can be useful in the determination of:

B.1 Inverse of a square matrix

The inverse of an square matrix can easily be determined using SVD. It takes the form:

$$X^{-1} = V\Sigma^{-1}U^H, \tag{B.5}$$

where

$$\Sigma^{-1} = \begin{bmatrix} \sigma_1^{-1} & & & \\ & \sigma_2^{-1} & & \\ & & \ddots & \\ & & & \sigma_N^{-1} \end{bmatrix}. \tag{B.6}$$

It is readily seen that,

$$\begin{aligned} XX^{-1} &= U\Sigma \underbrace{V^H V}_I \Sigma^{-1}U^H \\ &= U \underbrace{\Sigma\Sigma^{-1}}_I U^H \\ &= UU^H = I. \end{aligned} \tag{B.7}$$

Problems arise when one of the singular values $\sigma_1 \cdots \sigma_N$ is zero or very small. In that case, the reciprocal becomes infinity or too large.

B.2 Relation to eigen-decomposition

The eigenvalues of a $M \times M$ matrix R are a set of constants λ_i for which the following relation is satisfied:

$$Rq_i = \lambda_i q_i \quad \text{for } i = 0 \cdots M - 1, \tag{B.8}$$

or otherwise

$$R\underline{q} = \Lambda\underline{q}, \tag{B.9}$$

where \underline{q} is the eigenvector associated with the eigenvalues $\Lambda \stackrel{\text{def}}{=} [\lambda_0 \cdots \lambda_{M-1}]$ [47].

Recall the SVD of matrix X in (B.1). If matrix X is multiplied by its Hermitian X^H , the result shows that:

$$XX^H = U\Sigma \underbrace{V^H V}_I \Sigma U^H = U\Sigma^2 U^H = R. \tag{B.10}$$

XX^H represents the correlation matrix. (B.10) can be re-arranged and be written as:

$$RU = U\Sigma^2. \tag{B.11}$$

Note that the left singular vectors $U \stackrel{\text{def}}{=} [\underline{u}_0 \cdots \underline{u}_{M-1}]$ of matrix X are in fact the eigenvector of matrix $R = XX^H$. Equivalently, the right singular vectors $V \stackrel{\text{def}}{=} [\underline{v}_0 \cdots \underline{v}_{M-1}]$ of matrix X represent the eigenvectors of matrix $R^H = X^H X$ associated with the eigenvalues $\Sigma^2 \stackrel{\text{def}}{=} [\sigma_0^2 \cdots \sigma_{M-1}^2]$. The eigenvalues of matrix R and R^H are the square of the singular values of matrix X , denoted $\Sigma \stackrel{\text{def}}{=} [\sigma_0 \cdots \sigma_{M-1}]$.

Original publications

The work described in this thesis has been reported in the following publications:

- † J. Altuna and B. Mulgrew, “A Comparison of Cyclostationary Blind Equalisation Techniques”, *IEE Colloquium on Multipath Countermeasures*, Savoy Place, London, pages 8/1-8/6, May, 1996.
- † J. Altuna and B. Mulgrew, “A Comparison of Fractionally-Spaced Blind Equalisation Techniques for Fixed and Fading Channels”, *International Symposium on Communication Theory and Application*, The Lake District, England, July, 1997.

Accepted for publication:

- † J. Altuna and B. Mulgrew, “A Comparison of Cyclostationary Blind Equalisation Algorithms in the Mobile Radio Environment”, *International Journal of Adaptive Control and Signal Processing*.

Paper submitted:

- J. Altuna and B. Mulgrew, “A Cyclic Subspace Algorithm for Blind Channel Identification”, *IEEE Transactions on Signal Processing*, April 1997.

† Reprinted in this appendix.

A COMPARISON OF CYCLOSTATIONARY BLIND EQUALISATION TECHNIQUES

J. Altuna and B. Mulgrew¹

Abstract: Interest in blind equalisation has grown in recent years due to the fact that in some communication systems the transmission of a training sequence is not physically feasible. Most blind equalisation schemes to date have sampled the channel output at the symbol rate to produce a stationary output sequence. This paper presents a comparison of the performance of two algorithms for nonminimum phase blind channel equalisation using fractionally-spaced sampling in the receiver. The major problem associated with nonminimum phase channel equalisation is that the measured output statistics must preserve the phase characteristics of the channel. Cyclostationary statistics, unlike the conventional second-order statistics, have been shown to be efficient in this respect. Two families of algorithms [5] [1] are available which exploit the cyclostationary nature of the oversampled received signal in different ways. Both philosophies provide algorithms which are superior to existing symbol-spaced blind equalisation techniques in terms of reliability and speed of convergence. To date no comparison of these two approaches has been carried out and this paper examines these techniques by means of simulation using multipath channels.

1 Introduction

Interest in blind equalisation has grown in the recent years due to the fact that in many communication systems the transmission of a training signal is not possible, e.g. broadcast systems. During the transmission time, the original signal is exposed to severe distortion caused by the transmission medium such as multipath fading, commonly called inter-symbol interference. The aim of the equaliser is to remove distortion from the original signal and the main advantage of blind equalisation is that a training-sequence can be omitted.

The first steps in blind equalisation were given by Sato [8] and his work gave rise to a new set of algorithms known as Bussgang algorithms. The main feature is their similarity to the LMS algorithm in terms of minimization of a quadratic form and therefore low computational complexity but on the other hand their initial convergence is slow. However, higher-order statistics methods have been shown to preserve the phase information necessary to identify nonminimum phase channels. These require the channel output to be sampled at the symbol rate, producing a stationary output. The most important drawback of HOS is the large amount of data necessary in the estimation of the cumulants. The resulting delay produced in the estimation makes HOS algorithms impractical for land mobile communications where it would be impossible to track the time-varying channel or at least it results in loss of useful information.

However, if the channel output is sampled at a rate higher than the symbol rate, the output sequence generated is not stationary but cyclostationary and Gardner [4] showed that identification of nonminimum phase channels can be achieved using only second-order statistics exploiting the property of cyclostationarity. Many algorithms have been developed based on the idea of fractionally-spaced sampling (also called oversampling) but the fact that a discrete-time cyclostationary process can be described using a *time series representation* (TSR) in terms of P stationary processes via a time-division demultiplexing operation [3], where P is the oversampling factor, has yielded two families of algorithms. The former is based on the use of second-order cyclostationary statistics applied to the oversampled output and will be referred to as *cyclostationary algorithms* [5]. The latter exploits some rich matrix structure derived from the *multichannel representation* of the TSR [1].

¹J. Altuna and B. Mulgrew are with the Department of Electrical Engineering, The University of Edinburgh

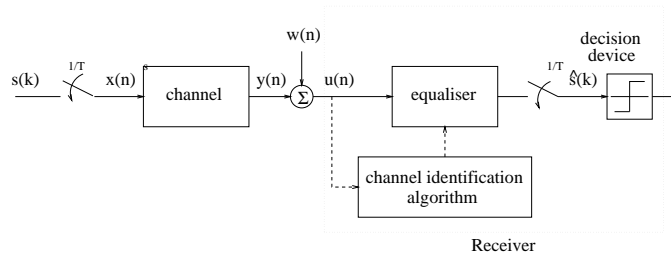


Figure 1: Communication system.

2 Problem Definition

Consider the situation shown in Figure 1 where the output sequence $y(n)$ is written as:

$$y(n) = h(n) * x(n) + w(n) = \sum_{k=-\infty}^{\infty} h(n - kP)x(k) + w(n) \quad (1)$$

where, $*$ denotes linear convolution and the input data sequence is obtained from $x(n) = \sum_k s(k)\delta(n - kP)$ where, $s(k)$ is in i.i.d. process and $\delta(n)$ is the discrete delta function. The filter $h(n)$ represents the generally nonminimum phase time-invariant channel. The oversampling factor P is an integer $\frac{T}{T_s}$, where T is the baud period and T_s the oversampling period. Thus, $x(n)$ is a wide-sense cyclostationary process, i.e., its mean $E\{x(n)\}$ and autocorrelation $R(n, m) = E\{x(n + m)x^*(n)\}$ are periodic in n with period P . The additive noise $w(n)$ is a wide-sense stationary sequence, statistically independent from $x(n)$.

The objective of the blind deconvolution problem is to identify the unknown impulse response of the channel from a finite set of output samples and only partial information about $x(n)$. Then, recover the input sequence $x(n)$ from $y(n)$ by means of inverse filtering.

3 Channel Identification Using The Complex Cepstrum of the Cyclic Autocorrelation

A discrete-time blind deconvolution algorithm proposed by Hatzinakos in [5] will be described in this section as an example of the family cyclostationary algorithms. The output of the channel is fractionally sampled at a rate higher than the baud-rate and then the complex cepstrum of the cyclic autocorrelation obtained.

The cyclic autocorrelation $R_y^\alpha(m)$ is defined as the Fourier coefficients of the classical autocorrelation function $R_y(n, m) = E\{y(n + m)y^*(n)\}$ which is periodic in n with period P .

$$\begin{aligned} R_y(n, m) &= \sum_{\alpha} R_y^\alpha(m) e^{i2\pi\alpha n} \\ R_y^\alpha &= \frac{1}{P} \sum_{n=(P)} E\{y(n + m)y^*(n)\} e^{-i2\pi\alpha n} \end{aligned} \quad (2)$$

A homomorphic approach was proposed in [5] exploiting the ability of the complex cepstrum to separate maximum and minimum phase polynomials from the transfer function of the channel. Provided that the complex cepstrum of the cyclic autocorrelation is well defined or provided that the channel has no zeros on the unit circle, a set of linear overdetermined system of equations can be formulated.

$$\mathbf{R}_\alpha \mathbf{a} = \mathbf{r}_\alpha \quad (3)$$

where, \mathbf{R}_α is a $2w \times 2max(p, q)$ matrix ($w \geq max(p, q)$) with entries from the cyclic autocorrelation lags of the form

$$\begin{bmatrix} R_y^\alpha(-w-1) & \dots & R_y^\alpha(-w-\max(p,q)) & R_y^\alpha(-w+1) & \dots & R_y^\alpha(-w+\max(p,q)) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ R_y^\alpha(-1-1) & \dots & R_y^\alpha(-1-\max(p,q)) & R_y^\alpha(-1+1) & \dots & R_y^\alpha(-1+\max(p,q)) \\ R_y^\alpha(1-1) & \dots & R_y^\alpha(1-\max(p,q)) & R_y^\alpha(1+1) & \dots & R_y^\alpha(1+\max(p,q)) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ R_y^\alpha(w-1) & \dots & R_y^\alpha(w-\max(p,q)) & R_y^\alpha(w+1) & \dots & R_y^\alpha(w+\max(p,q)) \end{bmatrix} \quad (4)$$

$\mathbf{a} = [C_\alpha(1), \dots, C_\alpha(\max(p,q)), \dots, D_\alpha(1), \dots, D_\alpha(\max(p,q))]$ is $2\max(p,q) \times 1$ vector with entries from the differential cepstrum parameters and $\mathbf{r}_\alpha = [wR_y^\alpha(-w), \dots, R_y^\alpha(-1), -R_y^\alpha(1), \dots, -wR_y^\alpha(w)]$ is a $2w \times 1$ vector.

The solution to equation 3 exists and is unique provided that the over-determined matrix \mathbf{R}_α has linearly independent columns [6].

In terms of computational cost, the algorithm described above is computationally quite expensive but however, the computation of the cyclic autocorrelation involves less complex multiplications than in the estimation of higher-order cumulants or moments.

Apart from the cyclic autocorrelation estimation, one singular value decomposition is computed to find the least-squares solution of the overdetermined system of equations shown in equation 3. The size of the matrix in this case is $2w \times 2\max(p,q)$ where $w \geq \max(p,q)$. Parameters p and q are related to the channel model as shown in [5].

4 Subspace Method for the Blind Identification of Multichannel FIR Filters

This section describes a subspace method for blind equalisation proposed by Moulines *et al.* in [1] based on the multichannel representation of equation 1. This is in the same spirit of the pioneering work by Tong *et al.* [7]. It exploits the *orthogonality* property between the *noise* and *signal* subspaces. The result is given as a minimisation of a quadratic form and is believed to be more efficient computationally than [7] due to the fact that a single eigenvector decomposition is needed.

Based on equation 1, for a certain value of n in $x(n)$ the output sequence is obtained convolving a different sequence of $h(\cdot)$ every T_s seconds. However, the input sequence $x(n)$ can be expanded to N successive samples of this sequence and the channel matrix spanned to create the following matrix $\mathbf{H}_N^{(n)}$:

$$\mathbf{H}_N^{(n)} = \begin{pmatrix} h(n) & \dots & \dots & h(n+MP) & 0 & \dots & \dots & 0 \\ 0 & h(n) & \dots & \dots & h(n+MP) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & \dots & 0 & h(n) & \dots & h(n+MP) \end{pmatrix} \quad (5)$$

$$\mathbf{Y}_k^{(n)} = \mathbf{H}_N^{(n)} \mathbf{X}_k + \mathbf{W}_k^{(n)} \quad \text{where } 0 \leq n \leq P-1$$

and

$$\begin{pmatrix} \mathbf{Y}_k^{(0)} \\ \vdots \\ \mathbf{Y}_k^{(P-1)} \end{pmatrix} = \begin{pmatrix} \mathbf{H}_N^{(0)} \\ \vdots \\ \mathbf{H}_N^{(P-1)} \end{pmatrix} \mathbf{X}_k + \begin{pmatrix} \mathbf{W}_k^{(0)} \\ \vdots \\ \mathbf{W}_k^{(P-1)} \end{pmatrix} \quad (6)$$

$$\mathbf{Y}_k = \mathbf{H}_N \mathbf{X}_k + \mathbf{W}_k$$

The identification of the channel matrix $\mathbf{H}_N^{(n)}$ comes from the measurement of the autocorrelation matrix R_y of the oversampled received signal like in [7].

$$R_y = \mathbf{H}_N R_x \mathbf{H}_N^H + R_w \quad (7)$$

For the channel to be identifiable from the correlation matrix, matrix H_N needs to be full column rank. This situation, $rank(H_N) = M + N$, is fulfilled if

1. The polynomials $H^{(n)}(z) \stackrel{\text{def}}{=} \sum_{j=0}^M h_j^{(n)} z^j$ have no common zero.
2. N is greater than the maximum degree M of the polynomials $H^{(n)}(z)$, i.e. $N \geq M$.
3. At least one polynomial $H^{(n)}(z)$ has degree M .

where N is the length of the temporal window and M is the degree of *inter-symbol-interference* (ISI).

The fact that the rank of the product $H_N R_x H_N^H$ is $M + N$ is represented in the first $M + N$ eigenvalues of matrix R_y and the rest $PN - M - N$ eigenvalues correspond solely to the noise covariance matrix R_w .

$$\begin{aligned} \lambda_i &> \sigma^2 & \text{for } i=1, \dots, M+N \\ \lambda_i &= \sigma^2 & \text{for } i=M+N+1, \dots, PN \end{aligned} \quad (8)$$

where σ^2 is the variance of the noise.

The property of orthogonality between the *signal subspace* conformed by the first $M + N$ eigenvalues of matrix R_y and *noise subspaces* gives way to a unique identification of the channel parameters (up to a scalar factor) as a solution of the minimisation of a quadratic form [1].

The subspace algorithm described in [1] involves three stages: (i) estimation of the autocorrelation matrix; (ii) eigenvalue decomposition of that matrix; (iii) minimisation of cost function subject to linear or quadratic constraints. The eigenvalue decomposition is of a PN by PN matrix. Finally a linear constraint requires less computations since it involves solving a linear system compared to the quadratic constraint which requires one further eigenvalue decomposition of a $P(M + 1)$ by $P(M + 1)$ matrix.

5 Simulation Results

To examine the sensitivity of the algorithms to the position of channel zeros, two nonminimum phase channels were considered. The first

$$H_1(z) = (1 + 0.2z^{-1})(1 - 0.35\sqrt{3}z)(1 - 0.25\sqrt{2}z) \quad (9)$$

has one zero inside the unit circle at $z = -0.2$ and two zeros outside the unit circle at $z = \frac{1}{0.35\sqrt{3}}$ and $z = \frac{1}{0.25\sqrt{2}}$. To obtain the second channel the zero inside the unit circle is moved closer to the unit circle to $z = -0.5$. Thus the second channel is:

$$H_2(z) = (1 + 0.5z^{-1})(1 - 0.35\sqrt{3}z)(1 - 0.25\sqrt{2}z) \quad (10)$$

The source symbols $s(k)$ were drawn from bipolar binary data with a uniform distribution. Additive white Gaussian noise was added to the output of the channel with signal-to-noise ratio 15 dB. The SNR is defined as $10 \log_{10} \frac{\sum_n |u(n)|^2}{\sum_n |w(n)|^2}$ where $u(n) = h(n) * x(n)$ is the oversampled channel output uncorrupted by noise and $w(n)$ is the noise sequence.

The following parameters were chosen in each of the algorithms. For the subspace algorithm the number of virtual channels is $P = 3$, the width of the temporal window is $N = 10$ and the degree of ISI is $M = 1$. A quadratic constraint was used to minimise the cost function. For the cyclostationary algorithm the oversampling factor is also $P = 3$ and the value for the parameters p and q is 11 in both cases. In order to compare the performance of both algorithms, channel identification was performed using different data lengths and 1000 symbols were then transmitted and equalised. 16 taps were used in the linear equaliser where a ZF equalisation criterion was chosen. The MSE of the symbol estimates using a linear equaliser was obtained as shown in figure 2 for channel 1 and channel 2. The MSE is defined as:

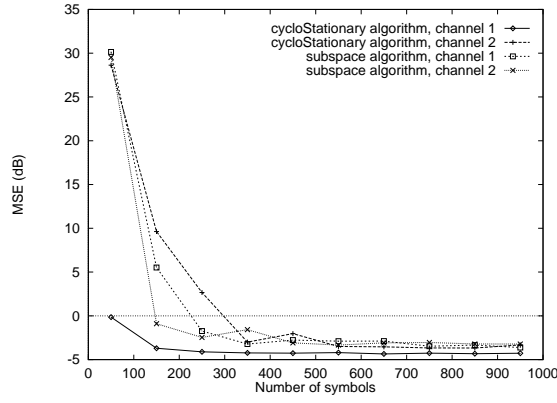


Figure 2: MSE of symbol estimates.

$$MSE = 10 \log_{10} \frac{1}{M} \sum_{i=1}^M \sum_{k=1}^N \|\hat{s}_i(k) - s(k)\|^2 \quad (11)$$

where M is the number of Monte-Carlo trials (20 in our case) and $\hat{s}_i(\cdot)$ is the vector of the symbol estimates from the i th trial. Figure 3 shows the BER or the probability of symbol error at the receiver which has to be less than 0.1 so that one may switch to a conventional decision-directed equaliser with little chance of disastrous error propagation [2]. In figure 2 the interesting thing to note is that for channel 1 the cyclostationary algorithm performs better than the subspace method but for channel 2 the performances are similar in terms of MSE at big data lengths but the convergence in the case of the subspace algorithm is faster. If the performance of each algorithm is analysed with respect to the influence that getting the zero closer to the unit circle has, we see that in the case of the cyclostationary algorithm the performance is clearly poorer. This is understood in terms of the distortion introduced by the zero getting closer to the unit circle. One of the drawbacks of this algorithm is that the complex cepstrum is not well defined when a zero is on the unit circle in which case the channel is not identifiable but Hatzinakos proposes a method to deal with it which consists on moving the zero inside or outside the unit circle. However it is recognised that this method is very sensitive to errors. In the case of the subspace algorithm the first thing we note is that moving the zero closer to the unit circle does not affect the performance so much and in fact the convergence is faster than for channel 1. Compared to the cyclostationary algorithm the convergence of the subspace algorithm is faster for channel 2.

From figure 3 the following conclusions can be drawn. The cyclostationary algorithm opens the eye of the equaliser with as few as 90 symbols used in the estimation of channel 1 and 280 for channel 2 whereas the subspace algorithm takes 220 symbols to open the eye and only 140 for channel 2.

This result suggest that the subspace algorithm is perhaps better suited for channels with zeros close to the unit circle which make them good candidates for a typical mobile radio communication environment where it is quite common to find nonminimum phase zeros close or even on the unit circle. The advantages of the cyclostationary algorithm is that it is less sensitive to correlated noise and exhibits faster convergence when zeros are far from the unit circle. In terms of computational cost, the size of the cyclic autocorrelation matrix of equation 4 is 22×22 ($2w \times 2max(p, q)$) for the cyclostationary algorithm in both cases, whereas in the case of the subspace algorithm the size of the autocorrelation matrix is 30×30 ($P.N \times P.N$). However, it is worth noting that reducing the length of the temporal window (N) in the subspace algorithm gives rise to a smaller autocorrelation matrix and zeros closer to the unit circle carries an increase in the size of the cyclic autocorrelation matrix (equation 4).

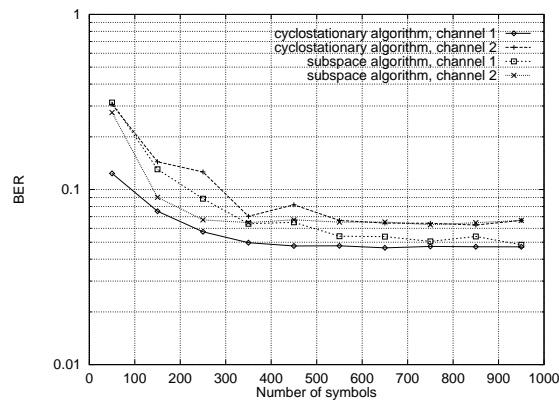


Figure 3: BER of symbol estimates.

6 Conclusions

This paper presents some results of two algorithms for nonminimum phase channel equalisation, which represent cyclostationary statistic based methods [5] and subspace based methods [1] respectively. Thus far, no comparison of these approaches has been carried out and this paper examines these two techniques by means of Monte-Carlo simulations. The authors believe that this analysis will lead to a better understanding of this problem, and therefore give rise to improved equalisation algorithms.

References

- [1] J. Cardoso E. Moulines, P. Duhamel and S. Mayrargue. Subspace methods for the blind identification of multichannel fir filters. *IEEE Trans. on SP*, 43(2):516–525, February 1995.
- [2] G.J. Foschini. Equalizing without altering or detecting data. *B.S.T.J.*, 64:1183–1186, May 1985.
- [3] W.A. Gardner. *Introduction to Random Processes with Applications to Signals and Systems*, chapter 12, pages 323–415. McGraw-Hill, 2nd edition, 1990.
- [4] W.A. Gardner. *Cyclostationarity in Communications and Signal Processing*. IEEE Press, 1st edition, 1994.
- [5] D. Hatzinakos. Nonminimum phase channel deconvolution using the complex cepstrum of the cyclic autocorrelation. *IEEE Trans. on SP*, 42(11):3026–3042, November 1994.
- [6] S. Haykin. *Adaptive Filter Theory*, chapter 20, pages 722–755. Prentice-Hall, 2nd edition, 1991.
- [7] G.Xu L. Tong and T. Kailath. Blind identification and equalization based on second-order statistics: A time domain approach. *IEEE Trans. on IT*, pages 340–349, March 1994.
- [8] Y. Sato. A method of self-recovering equalization for multilevel amplitude-modulation systems. *IEEE Trans. on Comm.*, 23:679–682, June 1975.

A Comparison of Fractionally-Spaced Blind Equalisation Techniques for Fixed and Fading Channels

Jon Altuna and Bernard Mulgrew

Signals and Systems Group, Department of Electrical Engineering, The University of Edinburgh,
Mayfield Road, Edinburgh EH9 3JL, Tel: (0131) 650 5655, Fax: (0131) 650 6554, e-mail: Jon.Altuna@ee.ed.ac.uk

ABSTRACT

Interest in blind equalisation has grown in recent years due to the fact that in many communication systems the transmission of a training signal is not desirable, e.g. broadcast systems. The use of conventional second-order statistics in blind equalisation applied to the received signal sampled at the symbol rate does not preserve phase information and therefore cannot be applied to identify nonminimum phase communication channels. However, it has been shown that phase information can be preserved using only second-order statistics if the received signal is sampled at multiples of the symbol rate. The idea of fractionally-spaced sampling in blind equalisation has led to different approaches which share the same principle but yield different algorithms. Two approaches based on cyclic statistics and subspaces respectively will be compared in a fixed channel environment. The subspace method for blind channel estimation will also be used for a fading environment simulation with the COST207 model for simulating a GSM system. The objective is to determine how the convergence speed of cyclostationary blind channel estimation techniques compare to conventional supervised methods in different fading environments.

1. Introduction

Most blind equalisation schemes to date have sampled the channel output at the symbol rate to produce a stationary output. The system identification procedure based on second-order statistics of the stationary output can be used efficiently to identify channels which are either minimum phase or maximum phase, but they cannot be used for nonminimum phase channel identification since they do not preserve phase information. In this case the higher-order statistics (HOS) of the stationary output need to be applied. The most important drawback of HOS is the large amount of data necessary in the estimation of the cumulants. The resulting delay produced in the estimation makes HOS algorithms impractical for land mobile communications where it would be difficult to track the time-varying channel.

However, in recent years it has been shown that the second-order statistics of an oversampled (fractionally-spaced sampled) sequence can provide the necessary phase information to identify the channel. Further, fewer data samples are needed in the estimation of the channel characteristics compared to HOS methods. These methods were first proposed by Gardner [3] and Tong *et al* [7] and are known as cyclostationary blind channel identification and equalisation algorithms. Gardner [3] investigated the use of the cyclic spectrum [5] in channel identification but he proposed a channel identification algorithm which was not truly blind since it relies on sending a training signal either periodically during message transmission or superposed on top of the message signal. However, his work can be considered as pioneering the family of blind channel identification algorithms exploiting second-order cyclostationary statistics. Motivated by this work, other algorithms that employ second-order cyclostationary statistics (SOCS) have been proposed [1] [6] [8].

This paper presents some results of a comparison of the convergence of cyclostationary blind equalisers in a fixed environment and compares the best algorithm with supervised methods in a mobile communications environment, where existing adaptive equalisers with training take only few hundred data samples to converge. Current research on cyclostationary blind equalisation algorithms has been focussed on the development of new algorithms and little attention has been paid to see whether the convergence is fast enough for a typical mobile communication environment. Blind

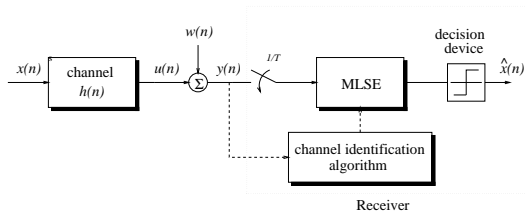


Fig. 1: Communications system model. The received signal is sampled P times faster than the symbol rate and the blind channel identification algorithm estimates the coefficients of the MLSE.

equalisation has been regarded as an alternative to conventional adaptive equalisers in mobile radio communications where the latter often fail to track rapidly varying channels.

2. Problem formulation

Figure 1 shows a typical communications system which is modelled as:

$$\begin{aligned} y(n) &= h(n) * x(n) + w(n) = u(n) + w(n) \\ &= \sum_{k=-\infty}^{\infty} h(n - kP)x(k) + w(n) \end{aligned} \quad (1)$$

where, $*$ denotes linear convolution, the input data sequence $x(n)$ is a complex, i.i.d. process and the filter $h(n)$ represents the generally nonminimum phase time-invariant channel. The oversampling factor P is an integer $\frac{T_s}{T}$, where T is the symbol period and T_s the oversampling period. Thus, $x(n)$ is a wide-sense cyclostationary process, i.e., its mean $E\{x(n)\}$ and autocorrelation $R(n, m) = E\{x(n+m)x^*(m)\}$ are periodic in n with period P . The additive noise $w(n)$ is a wide-sense stationary sequence, statistically independent from $x(n)$. The objective of blind equalisation is to recover the input sequence $x(n)$ using only a finite set of output samples $y(n)$ and only partial information about $x(n)$. The process is therefore unsupervised and no training signal is involved.

As a consequence of the oversampling operation which involves the interpolation of zeros in between samples of the input data, the statistics of the fractionally-spaced sampled output sequence $u(n)$ are periodic with period P [5] and it is then possible to apply second-order cyclostationary statistics to identify the channel [6]. This has the advantage of few data samples being required in the estimation of the channel and the preservation of phase information. On the

other hand, the multichannel representation of the system where the cyclostationary output is decomposed into P stationary sequences has led to algebraic approaches derived from Direction-of-Arrival (DOA) estimation algorithms in Array Signal Processing [2] [7].

Both cyclic statistics based approaches and multichannel methods have shown a fast convergence and this paper compares two representative algorithms of each family. The approach presented by Hatzinakos based on channel identification using the cyclic spectrum of the cyclic autocorrelation [6] and the subspace algorithm by Moulines *et al.* [2] have been selected. Initial results with fixed channels are encouraging and suggest their convergence could be fast enough for a mobile communication environment. The second part considers the convergence and tracking difficulties of cyclostationary algorithms in a mobile communications environment. Especially we are concerned with some drawbacks inherent in cyclostationary blind equalisation algorithms such as the inability to identify channels with:

- zeros on the unit circle [6]
- zeros and poles evenly spaced radially from 0 to 2π with angle $2\pi/P$ [6] [2]
- strictly bandlimited to less than $2\pi/T$ [6].

Compared to HOS approaches, cyclostationary algorithms present the following advantages:

- Fewer data samples are required in the estimation of the channel.
- No constraint is imposed on the p.d.f. of the input data.
- Fractionally-spaced sampling is less sensitive than synchronous sampling to timing errors.

3. GSM Channel Description

The typical communication system description of equation 1 in a mobile environment is expressed in terms of a composite channel (transmitter and receiver filters as well as a time-varying propagation channel) which is given by:

$$h_c(n, k) = c(n, k) * g_{TR}(n) \quad (2)$$

where n indicates the echo delay time lag, k represents the absolute time and $g_{TR}(n)$ is the discrete combined receiver and transmitter filter which is time-invariant.

The propagation channel is modelled here according to the COST207 directive from the European Commission and was implemented using the model developed by [4]. The central carrier frequency of the GSM-TDMA system is 900 MHz with 8 channels per carrier,

Delay μs	Power dB	Doppler category
0	-3	CLASS
0.2	0	CLASS
0.6	-2	GAUSS1
1.6	-6	GAUSS1
2.4	-8	GAUSS2
5.0	-10	GAUSS2

Table 1: Power/delay profile of 6 tap COST207 Typical Urban environment propagation channel.

each spaced by a 200 kHz bandwidth. The channel data rate as in GSM is 270.833 kbits/s, which results from transmitting at a baud-period of $3.69 \mu s$. The structure of a typical GSM time-slot structure comprises two bursts of information data containing 58 bits each with a midamble 26 bits for the training of conventional supervised algorithms as shown in Figure 4. In this scenario a complete time slot is read first and the training sequence from left to right is used to recover the second burst of data and next the training sequence from right to left is used to equalise the first burst of data from right to left. In blind equalisation the 26 bits allocated for training can be reallocated as channel coding bits or they can be used to increase the spectral efficiency of the TDMA system by allocating them as information data bits.

In a typical urban (TU) environment which will be analysed in this paper, the propagation channel is modelled as a multipath fading impulse response with six taps where the multipath delay extends to up to $5 \mu s$ with a significant reflector occurring at $0.2 \mu s$ due to large buildings nearby. Three different Doppler spectra are used to describe the variation of the channel weights at specific time delays of the impulse response: classical Doppler spectra (CLASS) and two spectra based on Gaussian distributions (GAUSS1 and GAUSS2). The power/delay profile of the TU environment is shown in Table 1.

4. Simulation Results

Two experiments were conducted, the first of which compares the performance of cyclic-statistics based methods and subspace methods for blind channel estimation in a fixed channel environment. In a second experiment the subspace method will be compared with conventional supervised methods in a mobile radio environment using a *maximum-likelihood sequence*

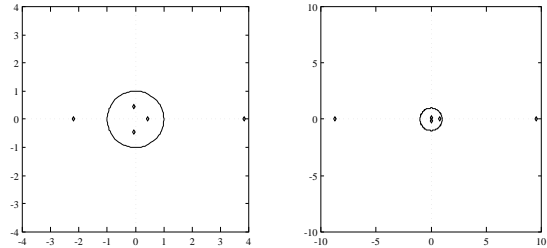


Fig. 2: Position of the zeros for channel 1 and channel 2 respectively.

estimator (MLSE) in the receiver. The question of the delay associated with the MLSE is an interesting one for blind equalisation and will be studied in this section. It is known that the performance of the MLSE improves as the delay introduced in the decisions increases but in the case of supervised adaptive channel estimation algorithms the equaliser coefficient estimates become more and more old-fashioned as the delay increases because the adaptive equaliser learns from the decisions made by the MLSE. On the other hand, the blind equaliser which does not require knowledge of any kind of *reference* signal allows a delay as long as required in the equaliser.

4.1. Simulation I

In this first experiment we compare two representative algorithms of cyclic-statistics based algorithms [6] and subspace algorithms for blind channel identification [2]. Two channels were considered and the position of its zeros are shown in Figure 2. Both channels have two zeros outside the unit circle and the sensitivity of both algorithms to the position of zeros in the unit circle will be examined. The input is a bipolar binary signal and the equaliser is a Viterbi equaliser (VE) with a delay of $5L$ introduced when making the decisions, where L is the memory of the VE. The oversampling factor is $P = 3$ in both cases. Figure 3 shows the average BER performance for different $\frac{E_b}{N_0}$ scenarios over 20 Monte-Carlo realisations for at least 500 errors. 200 symbols were used in the estimation of the channel. The results show that in both cases the subspace algorithm is superior to the cyclic algorithm. This point is essential in mobile radio channels where zeros very close or on the unit circle are often found. It is known that the cyclic algorithm [6] is very sensitive to the position of the zeros and that is illustrated for channel 2 the performance of the cyclic algorithm degrades more with respect to the subspace method than for channel 1. In

fact the cyclic algorithm is not well defined when a zero is positioned on the unit circle and at least knowledge of the amount of zeros inside and outside the unit circle is required because the minimum phase and maximum phase zeros are found separately. This requirement makes the cyclic algorithm unrealistic for a multipath fading environment where the zeros of the channel are constantly moving inside and outside the unit circle.

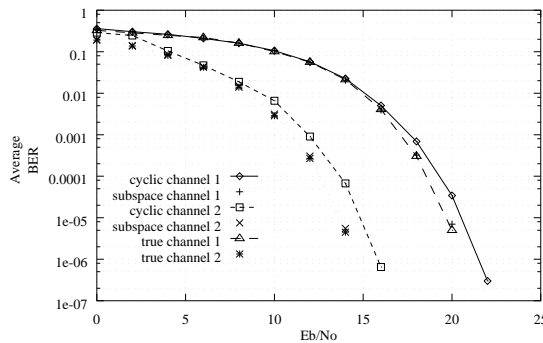


Fig. 3: Comparison of cyclic and subspace algorithms in terms of average BER for different E_b/N_0 scenarios.

4.2. Simulation II

The second experiment is concerned with the performance comparison of the subspace method for blind channel identification and conventional supervised methods in the pan-European digital mobile radio system GSM. The source signal is a 4-QAM signal and the transmitter filter is a raised-cosine spectrum type filter with roll-off factor $\beta = 0.1$. A Doppler frequency of 40 Hz was considered as a result of a vehicular speed of 48 km/h. In a second experiment a Doppler shift of 125 Hz was considered for a vehicular speed of 150 km/h. The receiver comprises an analog receiver filter and the output from the Rx filter is sampled $P = 3$ times faster than the symbol rate. The equaliser is a Viterbi equaliser (VE) with 4^3 states. A delay of 5 times the memory of the equaliser was used in making the decisions.

The simulation was performed by comparing the subspace blind equalisation method, a supervised non-adaptive algorithm and a supervised adaptive LMS algorithm. The non-adaptive method estimates the channel through the training sequence and this estimate is held fixed for the Viterbi algorithm for the entire burst and it is only updated in the next time slot. The *adaptive* receiver uses a *least-mean square* (LMS) algorithm where the training bits are used for start-up

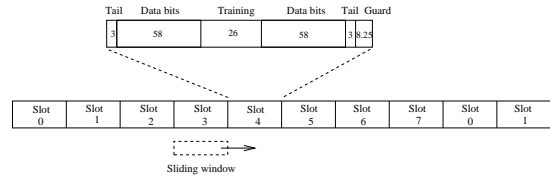


Fig. 4: GSM-TDMA time slot structure with 8 time slots per frame.

purposes and the equaliser parameters are constantly updated through the burst. Figure 5 shows the average BER performance for different E_b/N_0 scenarios, by running the respective algorithms for 5 Doppler periods of the propagation channel in order to avoid locally generated results and averaged over 20 realisations of the thermal noise. The blind channel estimation method uses a sliding window of variable length of multiples of one time slot to estimate the channel and this estimate is kept fixed for one time slot and it is updated for the next time slot. During the length of the sliding window it is assumed that the channel remains *quasi* stationary. This assumption is weak for high Doppler frequencies but the objective here is to see if the BER obtained by the mean channel estimate over 4 time slots applied to MLSE of 142 data bits (blind method) approximates the BER obtained by the mean channel estimate of the LS method over 26 bits and applied to MLSE of 116 data bits (supervised methods). In Figure 6 we note that for both Doppler frequencies data from 1 time slot is insufficient for the blind algorithm to achieve a similar BER performance to the conventional supervised methods. Taking data from adjacent time slots reduces the BER but this reduction is less apparent as the Doppler frequency increases.

The issue of the delay associated with the MLSE is dealt with next. As shown in Figure 7, the BER performance improves slightly as the delay of the MLSE receiver is increased when the channel estimates are made by the blind equalisation method. However, the performance degrades in a similar situation when it is driven by estimates produced by both supervised methods. However, the level of improvement does not take the blind equalisation algorithm any closer to the best supervised performance.

5. Conclusions

This paper presents some new results on the convergence of cyclostationary blind equalisation techniques in a mobile radio environment in an attempt to es-

establish how this compares to conventional supervised methods. The results suggest that new and faster blind equalisation techniques will probably be needed in the current structure of a mobile radio system if blind equalisation is to be considered. It has been shown that increasing the delay of the MLSE receiver generally improves the performance in the case of blind equalisation methods but this improvement does not come close to the best performance of the supervised methods. Other issues such as the effect of the nonlinear nature of the GMSK modulation scheme in GSM will also need to be dealt with in the near future.

References

- [1] L.A. Baccala and S. Roy. 'A new blind time-domain channel identification method based on cyclostationarity', *IEEE Signal Processing Letters*, Vol 1(6), pp 89-91, June 1994.
- [2] J. Cardoso E. Moulines, P. Duhamel and S. Mayrargue. 'Subspace methods for the blind identification of multichannel FIR filters. *IEEE Transactions on Signal Processing*, Vol 43(2), pp 516-525, February 1995.
- [3] W.A. Gardner. 'A new method for channel identification' *IEEE Transactions on Communications*, Vol 39(6), pp 813-817, June 1991.
- [4] D.G.M. Cruickshank D.I. Laurenson and G.J.R. Povey. 'A computationally efficient channel simulator for the COST 207 channel models', *IEE Colloquium on Computer Modelling of Communication Systems*, May 1994.
- [5] W.A. Gardner. *Cyclostationarity in Communications and Signal Processing*. IEEE Press, 1st edition, 1994.
- [6] D. Hatzinakos. 'Nonminimum phase channel deconvolution using the complex cepstrum of the cyclic autocorrelation' *IEEE Transactions on Signal Processing*, Vol 42(11), pp 3026-3042, November 1994.
- [7] G.Xu L. Tong and T. Kailath. 'Blind identification and equalization based on second-order statistics: A time domain approach.' *IEEE Transactions on Information Theory*, pp 340-349, March 1994.
- [8] S. Prakriya and D. Hatzinakos. 'Identification of parametric models with cyclostationary inputs' In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol IV, pp 417-419, Adelaide, Australia, April 1994.

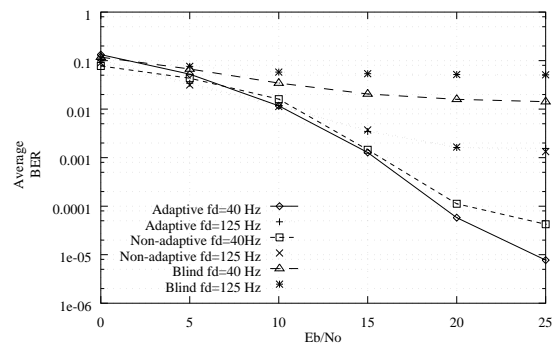


Fig. 5: Average BER performance for different E_b/N_o scenarios throughout the GSM data block for Doppler frequencies of $f_d = 40$ Hz and $f_d = 125$ Hz

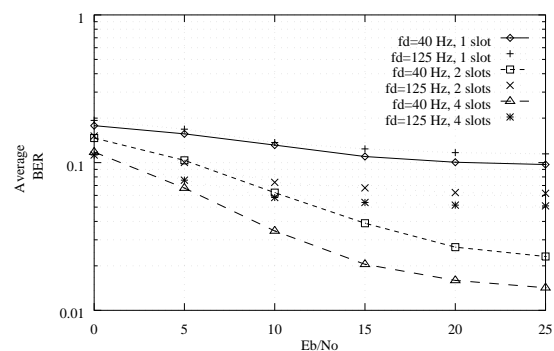


Fig. 6: Average BER performance of the subspace algorithm for blind equalisation for different E_b/N_o scenarios using 1, 2 and 4 TDMA time slots for block channel identification

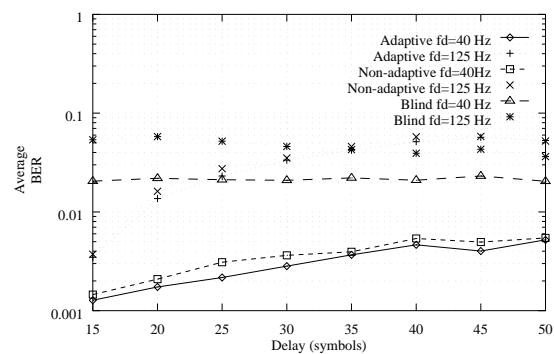


Fig. 7: Average BER performance versus delay associated in the MLSE in a $E_b/N_o = 15$ dB scenario.

A COMPARISON OF CYCLOSTATIONARY BLIND EQUALISATION ALGORITHMS IN THE MOBILE RADIO ENVIRONMENT

JON ALTUNA AND BERNARD MULGREW

*Signals and Systems Group, Department of Electrical Engineering, The University of Edinburgh,
Mayfield Road, Edinburgh EH9 3JL, Tel: (0131) 650 5655, Fax: (0131) 650 6554,
E-mail: Jon.Altuna@ee.ed.ac.uk Bernie.Mulgrew@ee.ed.ac.uk*

SUMMARY

In this paper we address the problem of blind equalisation in the mobile radio environment using cyclostationary techniques, where the channel is generally nonminimum phase and the phase information of the system has to be preserved. Conventional second-order statistics applied to the received signal sampled at the symbol rate do not preserve phase information and therefore cannot be applied to identify separate minimum phase and maximum phase zeros of the channel. It has been shown that phase information can be preserved using only second-order statistics if the received signal is sampled using fractionally-spaced sampling. This paper deals with the issue of how the convergence of the cyclostationary blind equalisation algorithms compares to conventional supervised methods in the pan-European mobile radio system GSM. The results suggest that the convergence speed of some of the actual cyclostationary algorithms is fast for stationary channels but slow compared to conventional supervised non-adaptive (LS) and adaptive (LMS) algorithms in time-varying mobile radio channels.

KEY WORDS Blind Equalisation; Nonminimum Phase Channels; Cyclostationarity; Mobile Radio Channels

1 INTRODUCTION

Interest in blind equalisation has grown in recent years due to the fact that in many communication systems the transmission of a training signal is not possible, e.g. broadcast systems. Most blind equalisation schemes to date have sampled the channel output at the symbol rate to produce a stationary output. The system identification procedure based on second-order statistics of the stationary output can be used efficiently to identify channels which are either minimum phase or maximum phase, but they cannot be used for nonminimum phase channel identification since they do not preserve phase information. To overcome this limitation higher-order statistics (HOS) of the stationary output have to be applied. The most important drawback of HOS is the large amount of data necessary in the estimation of the cumulants. The resulting delay produced in the estimation makes HOS algorithms impractical for land mobile communications where it would be impossible to track the time-varying channel or at least it results in the loss of useful information.

However, in recent years it has become clear that many of the inherent problems of blind equalisation can be solved by adding an extra degree of freedom to the receiver. This extra degree of freedom could come from oversampling (fractionally-spaced sampling) the received signal by an integer factor P or equivalently by measuring the signal using P antennas. The pioneering work by Gardner [6] and Tong *et al.* [10] suggested faster algorithms using only second-order statistics of the *oversampled* received output

due to the fact that less data is necessary in the estimation process compared to HOS methods. Tong *et al.* are thought to be the pioneers in blind channel equalisation using fractionally-spaced sampling and since then considerable research activity has taken place and many new algorithms have emerged in this area. Gardner [6] investigated the use of the cyclic spectrum [7] in channel identification but he proposed a channel identification algorithm which is not truly blind since it relies on sending a training signal either periodically during message transmission or superposed on top of the message signal. However, his work is considered to be the first in the family of blind channel identification algorithms exploiting second-order cyclostationary statistics. Motivated by his work, other algorithms that exploit second-order cyclic statistics (SOCS) [7] have been proposed [1] [9] [11]. Either the oversampling representation or the multichannel representation suggested the possibility of exploiting some especial matrix structure which resulted in a series of algebraic methods [5] [10] [12] [13] and other methods based on Maximum-Likelihood [14] [16] have also been proposed and comparative results can be found in [15] [8] [17] [18] [19].

Until now there has not been a proper comparison on the convergence of cyclostationary blind equalisers in a mobile communications environment where existing adaptive equalisers with training take only a few hundred data samples to converge. First indications show that convergence properties of cyclostationary blind equalisers could make them good candidates for this task. Besides, blind equalisation techniques could represent a valuable alternative to conventional adaptive equalisers in the presence of severe multipath fading. This paper is concerned with convergence and tracking properties of two cyclostationary algorithms representatives of cyclic statistics based methods [9] and multichannel methods [5] in the digital mobile communications environment. Simulations will be conducted using the COST207 model set up by the European Commission for simulating a GSM system.

The rest of the paper is organised as follows: In Section 2 the general problem of blind equalisation based on a single sensor oversampled or multiple sensors sampled at the symbols rate will be formulated completed with the limitations and advantages of the cyclostationary algorithms. In Section 3 and 4 two representative algorithms will be described. The performance evaluation and comparison of these algorithms will be presented in Section 5 using fixed channels and a simulation will be conducted using the multichannel method in order to compare the performance of blind channel estimation techniques in a multipath fading environment. Finally Section 6 draws some conclusions.

2 PROBLEM FORMULATION

The system configuration is shown in Figure 1 where the discrete-time input sequence $x(k)$ is convolved with a different subchannel $h^{(i)}(n)$ as a result of oversampling. The received signal $y^{(i)}(t)$ at the i th sensor, sampled at $T_s = T/P$ can be formulated as:

$$y^{(i)}(n) = \sum_{k=-\infty}^{\infty} x(k)h^{(i)}(n-k) + w^{(i)}(n) \quad (1)$$

or equivalently

$$y(n) = \sum_{l=-\infty}^{\infty} v(l)h(n-lP) + w(n) \quad P \in Z \quad (2)$$

where T is the baud-period of the source signal, $v(l) = \sum_k x(k)\delta(l-kP)$ is the oversampled input sequence which takes nonzero values every P samples, $\delta(l)$ is the Dirac delta, $x(k)$ is the generally complex, i.i.d. sequence of symbols at the baud-rate and $w(n)$ is a wide-sense stationary noise sequence statistically independent from $x(k)$. Since $P-1$ zeroes are interpolated between any two symbols of the input sequence $x(k)$ to form the oversampled input sequence $v(l)$, the nonzero values are multiplied by

the channel coefficients $h(0), h(P), \dots, h(PM)$ at one time instant and by $h(1), h(P+1), \dots, h(PM+1)$ another instant. The subchannel $h^{(i)}(n)$ associated with the i th sensor can be formulated as:

$$h^{(i)}(n) = h(i + nP), \begin{cases} i = 0, 1, \dots, P-1 \\ n = 0, 1, \dots, M \end{cases} \quad (3)$$

where M is the degree of ISI of the subchannels.

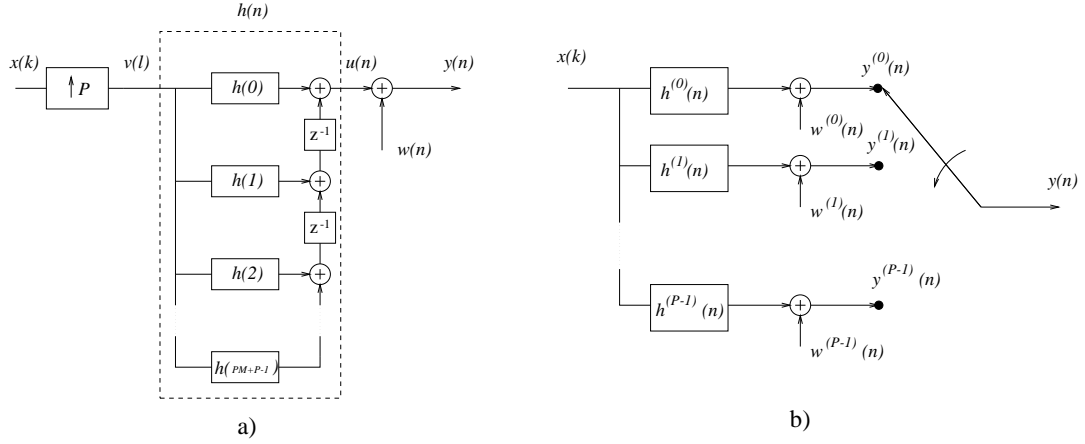


Figure 1: Oversampled output sequence, a) oversampling representation, b) multichannel representation.

Since Tong *et al.* presented their algorithm several articles have appeared in the literature not only presenting new algorithms but also providing a fundamental analysis of both cyclic statistics algorithms and subspace algorithms [2] [4] [8]. There are some inherent ambiguities in all blind equalisation schemes. Sign ambiguity is present at the blind equaliser when the probability distribution of the original signal is symmetric. This problem can be overcome by initialising the deconvolution filter with the desired sign or employing differential encoding in the transmitted sequence. On the other hand, it has been established [7] that both cyclic or multichannel methods share a common critical condition which marks the sufficient identifiability condition:

- **Theorem:** *The oversampled channel $h(n)$ can be identified if and only if the channel transfer function: $H(z) \triangleq \sum_n h(n)z^{-n}$ does not have zeros uniformly spaced around the unit circle with separation of $2\pi/P$ rad.*

This problem, although a critical condition is rare and will only occur for specific values of P only. Particular interest has been shown in the identifiability conditions of bandlimited channels. A serious drawback of cyclic statistics methods is the inability to identify channels which are strictly bandlimited to a total bandwidth less than $2\pi/T$ [2]. Because the support band of the cyclic spectrum $S^{1/T}(w)$ lies in the band $|w| < 2\pi(W - 1/2T)$, where W is the bandwidth of the channel, no information can be gained if the bandwidth falls below $1/2T$ and the identification problem reduces to a classical power spectrum based estimation problem. A series of *singular* channels which are unidentifiable using SOCS can be seen in [4] [8].

Compared to HOS approaches, cyclostationary algorithms present the following advantages:

- Fewer data samples are required in the estimation of the channel.
- No constraint is imposed on the p.d.f. of the input data.
- Fractionally-spaced sampling is less sensitive than synchronous sampling to timing errors.

3 CYCLIC STATISTICS BASED CHANNEL IDENTIFICATION ALGORITHM

A discrete-time blind deconvolution algorithm proposed by Hatzinakos in [9] will be described in this section as an example of the family of blind identification using SOCS. The output of the channel is sampled at a rate higher than the baud-rate and then the complex cepstrum of the cyclic autocorrelation will be obtained.

3.1 Problem Definition

Consider the situation depicted by equation 2. The input sequence $v(l)$ padded with zeros is a wide-sense cyclostationary process, i.e., its mean $E\{v(l)\}$ and autocorrelation $R_v(l, m) = E\{v(l+m)v^*(l)\}$ are periodic in l with period P .

$$R_v(l, m) = R_v(l + P, m) \quad (4)$$

Let us define the cyclic autocorrelation function of the output sequence $y(n)$ as the Fourier coefficients of the periodic autocorrelation function:

$$R_y(n, m) = \sum_{\alpha} R_y^{\alpha}(m) e^{i2\pi\alpha n} \quad \alpha = \frac{l}{P}, \quad l \in Z \quad (5)$$

and

$$R_y^{\alpha}(m) = \lim_{L \rightarrow \infty} \frac{1}{LP} \sum_{n=0}^{LP-1} E\{y(n+m)y^*(n)\} e^{-i2\pi\alpha n} \quad (6)$$

where L is the number of symbols used in the estimation.

It will be assumed that the transfer function of the channel allows the following factorisation [9]

$$H(z) = Az^r I(z^{-1})O(z) \quad (7)$$

where, A is a constant gain, r is a constant delay in time, the $I(z^{-1}) = \frac{\prod_{i=1}^{L_1} (1 - a_i z^{-1})}{\prod_{i=1}^{L_3} (1 - c_i z^{-1})}$ is a minimum phase polynomial and $O(z) = \prod_{i=1}^{L_2} (1 - b_i z)$ is a maximum phase polynomial, where $|a_i| < 1$, $|c_i| < 1$ and $|b_i| < 1$.

The coefficients $A(n)$ and $B(n)$ are the differential cepstrum parameters which contain the minimum phase and maximum phase information of the channel $h(n)$, respectively. They are defined as

$$A(n) = \sum_{i=1}^{L_1} a_i^n - \sum_{i=1}^{L_3} c_i^n, \quad n = 1, 2, \dots, p \quad B(n) = \sum_{i=1}^{L_2} b_i^n, \quad n = 1, 2, \dots, q \quad (8)$$

The objective of the blind channel identification problem is to identify the unknown impulse response of the channel from a finite set of output samples and only partial information about $x(k)$. Then, by applying inverse filtering to the sequence $y(n)$, the original input $x(k)$ can be recovered.

3.2 Channel Identification Using The Complex Cepstrum of the Cyclic Autocorrelation

A homomorphic approach was proposed in [9] exploiting the ability of the complex cepstrum to separate maximum and minimum phase parts of the channel in Equation 7. Provided that the complex cepstrum of the cyclic autocorrelation is well defined or provided that the channel has no zeros on the unit circle, a set of linear overdetermined system of equations can be formulated.

$$\mathbf{R}_{\alpha} \mathbf{a} = \mathbf{r}_{\alpha} \quad (9)$$

where \mathbf{R}_α is a $2w \times 2\max(p, q)$ matrix ($w \geq \max(p, q)$) with entries from the cyclic autocorrelation lags of the form

$$\begin{bmatrix} R_y^\alpha(-w-1) & \dots & R_y^\alpha(-w-\max(p, q)) & R_y^\alpha(-w+1) & \dots & R_y^\alpha(-w+\max(p, q)) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ R_y^\alpha(-1-1) & \dots & R_y^\alpha(-1-\max(p, q)) & R_y^\alpha(-1+1) & \dots & R_y^\alpha(-1+\max(p, q)) \\ R_y^\alpha(1-1) & \dots & R_y^\alpha(1-\max(p, q)) & R_y^\alpha(1+1) & \dots & R_y^\alpha(1+\max(p, q)) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ R_y^\alpha(w-1) & \dots & R_y^\alpha(w-\max(p, q)) & R_y^\alpha(w+1) & \dots & R_y^\alpha(w+\max(p, q)) \end{bmatrix} \quad (10)$$

and p and q are two positive integers.

The vector $\mathbf{a} = [C_\alpha(1), \dots, C_\alpha(\max(p, q)), \dots, D_\alpha(1), \dots, D_\alpha(\max(p, q))]$ has a length of $2\max(p, q) \times 1$ with entries from the differential cepstrum parameters, $\mathbf{r}_\alpha = [wR_y^\alpha(-w), \dots, R_y^\alpha(-1), -R_y^\alpha(1), \dots, -wR_y^\alpha(w)]$ is a $2w \times 1$ vector and $w \geq \max(p, q)$.

In practice, the cyclic autocorrelation function can be estimated using only sample estimates [9]

$$\hat{R}_y^\alpha(m) = \frac{1}{LP} \sum_{n=0}^{LP-1} y(n+m)y^*(n)e^{-i2\pi\alpha n} \quad (11)$$

The solution to equation 9 exists and is unique provided that the over-determined matrix \mathbf{R}_α has linearly independent columns and it is given as a set of values $C_\alpha(n)$ and $D_\alpha(n)$ which are defined as

$$\left. \begin{aligned} C_\alpha(n) &= A(n) + B^*(n)z_\alpha^{-n} \\ D_\alpha(n) &= A^*(n)z_\alpha^n + B(n) \end{aligned} \right\}, \quad n = 1, 2, \dots, \max(p, q) \quad (12)$$

This linear system of equations has a solution for those values of n which are not multiples of the oversampling factor P and consequently, the differential cepstrum parameters for $n = P, 2P, \dots$ have to be estimated through interpolation. The solution to this linear system determines the differential cepstrum parameters $A(n)$ and $B(n)$ of equation 8 and from there the impulse response of the channel using:

$$c(k) = i(k) * o(k) \quad (13)$$

where

$$i(k) = -\frac{1}{k} \sum_{n=2}^{k+1} [A(n-1)]i(k-n+1), \quad k = 1, \dots, N_1, \quad (14)$$

$$o(k) = \frac{1}{k} \sum_{n=k+1}^0 [B(1-n)]o(k-n+1), \quad k = -1, \dots, -N_2. \quad (15)$$

The selection of the integers p and q is carried out so that these are large enough that the differential cepstrum parameters which decay exponentially go below a certain constant threshold:

$$A(I) < C, \quad I > p = \lceil \ln(C)/\ln(a) \rceil \quad B(J) < C, \quad J > q = \lceil \ln(C)/\ln(b) \rceil \quad (16)$$

where $\max(|a_k|, |c_k|) < a < 1$ and $\max(|b_k|) < b < 1$ and C is a small constant like 10^{-4} . It is apparent from equation 16 that the closer the zeros are located to the unit circle, the bigger the parameters p and q .

3.3 Performance analysis

The algorithm fails when there is a zero on the unit circle or very close to it, because ideally we should make $p \rightarrow \infty$ and/or $q \rightarrow \infty$, which becomes impractical. The variance in the estimation of the channel parameters therefore increases as the channel zeros move closer to the unit circle as a result of truncating p and q . When zeros are located on the unit circle a procedure to move the zeros away from the unit circle can be adopted [9]. It is also important to note that among the assumptions made it is required to know at least how many zeros lie inside and outside the unit circle. This requirement is considered unrealistic when we deal with time varying channels where not only zeros can lie on the unit circle or very close to it but it also shows a pattern of zeros moving inside and outside the unit circle constantly.

The algorithm described above is computationally more expensive than other gradient type blind deconvolution algorithms (Bussgang algorithms) but the computation of the cyclic autocorrelation, however, involves less complex multiplications than in the estimation of higher-order cumulants or moments. For example, given N data samples the estimation of the cyclic autocorrelation lag $R_y^\alpha(0)$ requires $2N + 1$ complex multiplications while the estimation of the fourth-order cumulant lag $C_y(0, 0, 0)$ requires $4N + 4$ complex multiplications. Besides, the cyclic autocorrelation estimation, one singular value decomposition is computed to find the least-squares solution of the overdetermined system of equations shown in Equation 9. The size of the matrix in this case is $2w \times 2\max(p, q)$ where $w \geq \max(p, q)$.

4 SUBSPACE METHOD FOR BLIND IDENTIFICATION OF MULTICHANNEL FIR FILTERS

Inspired in Direction of Arrival Estimation (DOAE) techniques, several algebraic methods have been developed. In this section a subspace method [5] which is reminiscent of the MUSIC algorithm for DOAE will be described. It exploits the *orthogonality* property between the *noise* and *signal* subspaces and the result is given as a minimisation of a quadratic form.

4.1 Problem Formulation

Consider the situation depicted by equation 1 where the discrete input sequence $x(k)$ is convolved with a different *subchannel* every T/P seconds. Consequently, we might define a $N \times (N + M)$ filtering matrix $\mathcal{H}_N^{(i)}$ associated with a particular *subchannel* as:

$$\mathcal{H}_N^{(i)} = \begin{pmatrix} h^{(i)}(0) & \dots & h^{(i)}(M) & 0 & \dots & \dots & 0 \\ 0 & h^{(i)}(0) & \dots & h^{(i)}(M) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \dots & 0 & h^{(i)}(0) & \dots & h^{(i)}(M) \end{pmatrix} \quad (17)$$

where $\underline{h}^{(i)} = [h^{(i)}(0), \dots, h^{(i)}(M)]^\dagger$ are the coefficients of the i th subchannel, N is the length of the observation window and the superscript \dagger denotes transpose. Equation 1 can be rearranged in terms of the filtering matrix $\mathcal{H}_N^{(i)}$ as:

$$\underline{y}_k^{(i)} = \mathcal{H}_N^{(i)} \underline{x}_k + \underline{w}_k^{(i)} \quad (18)$$

where $\underline{x}_k = [x(k), x(k-1), \dots, x(k-N-M+1)]^\dagger$ is the $(N+M) \times 1$ vector which contains the input sequence $x(k)$, where the superscript \dagger denotes transpose. In the same fashion, we can define the $N \times 1$ noise vector $\underline{w}_k^{(i)} = [w^{(i)}(k), w^{(i)}(k-1), \dots, w^{(i)}(k-N+1)]^\dagger$ and the received signal vector $\underline{y}_k^{(i)} = [y^{(i)}(k), y^{(i)}(k-1), \dots, y^{(i)}(k-N+1)]^\dagger$.

By expanding equation 18 to the P subchannels, the received signal vector \underline{y}_k can be written as:

$$\underline{y}_k \stackrel{\text{def}}{=} \left[\underline{y}_k^{(0)\dagger}, \dots, \underline{y}_k^{(P-1)\dagger} \right]^\dagger \quad (19)$$

and

$$\underline{y}_k = \mathcal{H}_N \underline{x}_k + \underline{w}_k \quad (20)$$

where $\mathcal{H}_N \stackrel{\text{def}}{=} \Phi(\underline{h}, P, M + 1, N) = \left[\mathcal{H}_N^{(0)\dagger}, \dots, \mathcal{H}_N^{(P-1)\dagger} \right]^\dagger$ is the $PN \times (N + M)$ filtering matrix.

The linear filtering representation of equation 20 gives way to the following representation of the received signal autocorrelation matrix

$$R_y = \mathcal{H}_N R_x \mathcal{H}_N^H + R_w \quad (21)$$

where the superscript H denotes conjugate transpose, $R_x = E \left\{ \underline{x}_k \underline{x}_k^H \right\}$ is the source signal autocorrelation matrix and $R_w = E \left\{ \underline{w}_k \underline{w}_k^H \right\} = \sigma^2 I$ is the autocorrelation matrix of the white noise sequence with variance σ^2 and I the identity matrix. It follows that provided that the unknown source autocorrelation matrix R_x is full-rank and matrix \mathcal{H}_N is full column rank, the received signal autocorrelation matrix admits the eigenvalue decomposition $R_y = U \Sigma V^H$ where matrix Σ is a diagonal matrix which contains the eigenvalues $\lambda_0, \dots, \lambda_{PN-1}$ of matrix R_y , where

$$\begin{aligned} \lambda_i &> \sigma^2 && \text{for } i = 0, \dots, M + N - 1 \\ \lambda_i &= \sigma^2 && \text{for } i = M + N, \dots, PN - 1 \end{aligned} \quad (22)$$

The eigenvectors of matrix U associated with the first $N + M$ eigenvalues λ_i form the so called $PN \times (N + M)$ signal eigenvectors matrix $S = [\underline{s}_0, \dots, \underline{s}_{N+M-1}]$, whereas the eigenvectors associated with the last $PN - N - M$ conform the $PN \times (PN - N - M)$ noise eigenvector matrix $G = [\underline{g}_0, \dots, \underline{g}_{PN-N-M-1}]$. The $N + M$ columns of matrix S span the *signal subspace* whereas the columns of matrix G generate the so-called *noise subspace*. By orthogonality between the columns of matrix S and G it can be concluded that any vector from the noise subspace is orthogonal to any column vector in the signal subspace and by extension to any column of the filtering matrix \mathcal{H}_N

$$\underline{g}_i^H \mathcal{H}_N = 0 \quad 0 \leq i \leq PN - N - M - 1 \quad (23)$$

When only sample estimates of the received signal autocorrelation matrix R_y are available, the set of linear equations of equation 23 can be solved in the least squares sense, as the minimisation of the following quadratic form:

$$q = \sum_{i=0}^{PN-N-M-1} |\hat{\underline{g}}_i^H \mathcal{H}_N|^2 = \sum_{i=0}^{PN-N-M-1} \hat{\underline{g}}_i^H \mathcal{H}_N \mathcal{H}_N^H \hat{\underline{g}}_i \quad (24)$$

or equivalently

$$q = \sum_{i=0}^{PN-N-M-1} \underline{h}^H \mathcal{G}_i \mathcal{G}_i^H \underline{h} = \underline{h}^H \left(\sum_{i=0}^{PN-N-M-1} \mathcal{G}_i \mathcal{G}_i^H \right) \underline{h} = \underline{h}^H Q \underline{h} \quad (25)$$

where $\underline{h} = \left[\underline{h}^{(0)\dagger}, \dots, \underline{h}^{(P-1)\dagger} \right]^\dagger$ and $\mathcal{G}_i = \Phi(\hat{\underline{g}}_i, P, N, M + 1)$.

The estimate of the channel parameter vector \underline{h} is obtained by picking up the eigenvector associated with the smallest eigenvalue of matrix Q . The estimate of the channel impulse response is non-coherent and proper gain and phase adjustments need to be made. The uniqueness of the estimates up to an scalar factor relies on the especial structure of the filtering matrix \mathcal{H}_N which makes it full-rank provided that the following conditions are met:

- i) The polynomials $H^{(i)}(z) \stackrel{\text{def}}{=} \sum_{j=0}^M h^{(i)}(j) z^j$ have no common zero.
- ii) N is greater than the maximum degree M of the polynomials $H^{(i)}(z)$, i.e. $N \geq M$.
- iii) At least one polynomial $H^{(i)}(z)$ has degree M .

4.2 Performance Analysis

An analysis of the performance of this subspace methods was presented in [8]. The variance of the estimate of an unit-norm channel is given by

$$\sigma_{\underline{h}}^2 = E \left\{ \|\Delta \underline{h}\|^2 \right\} = \frac{\sigma^2}{L} (a + \sigma^2 b), \quad (26)$$

where L is the number of symbols used in the estimation of the autocorrelation matrix R_y and

$$a = \sum_{k=1}^{P(M+1)} \sum_{l=-(N-1)}^{(N-1)} \frac{L-|l|}{L} \times \text{tr} \left(G^H (I_P \otimes J_l) G \tilde{G}_k^H \mathcal{H}_N R_x(-l) \mathcal{H}_N^H \tilde{G}_k \right),$$

$$b = \sum_{k=1}^{P(M+1)} \sum_{l=-(N-1)}^{(N-1)} \frac{L-|l|}{L} \times \text{tr} \left(G^H (I_P \otimes J_l) G \tilde{G}_k^H (I_P \otimes J_{-l}) \tilde{G}_k \right).$$

where I_P is a $P \times P$ identity matrix,

$$(J_k)_{ij} = \begin{cases} 1 & \text{if } i = j + k \text{ and } |k| \leq N - 1, \\ 0 & \text{otherwise} \end{cases}.$$

and

$$\tilde{G}_k \triangleq [\tilde{g}_{1,k}, \dots, \tilde{g}_{PN-N-M,k}]$$

where

$$\tilde{g}_{i,k} = S \Sigma_s S^H \mathcal{H}_N Q_k^H \underline{q}_i.$$

Matrix $Q_k = \Phi(\underline{q}_k, P, M + 1, N)$ is obtained using the same transformation $\Phi(\cdot)$ as in equation 17, where \underline{q}_k denotes the k th column of the pseudoinverse of matrix Q and $\Sigma_s = \text{diag}(\lambda_0, \dots, \lambda_{PN-1})$ is the diagonal matrix associated with the signal eigenvector matrix S .

For a sufficiently large number of symbols used in the estimation ($L \gg N$), a and b are independent of L and the lower bound of the variance of the channel parameter estimates can be obtained in a computationally efficient manner.

In terms of computational cost, the subspace algorithm involves three stages: (i) estimation of the autocorrelation matrix, (ii) eigenvalue decomposition of that matrix; (iii) minimisation of a cost function subject to linear or quadratic constraint. The eigenvalue decomposition is of a PN by PN matrix. Finally, a linear constraint requires less computations since it involves solving a linear system compared to the quadratic constraint which requires one further eigenvalue decomposition of a $P(M + 1)$ by $P(M + 1)$ matrix.

5 SIMULATION RESULTS AND DISCUSSION

In this section we compare the performance of the two representative algorithms described in the previous section by means of computer simulations. In the first experiment their performance will be evaluated in the presence of time-invariant multipath channels and the second experiment shows the performance of the subspace method for blind channel identification in a typical urban multipath fading environment modelled using the COST207 model compared to conventional supervised adaptive and non-adaptive algorithms.

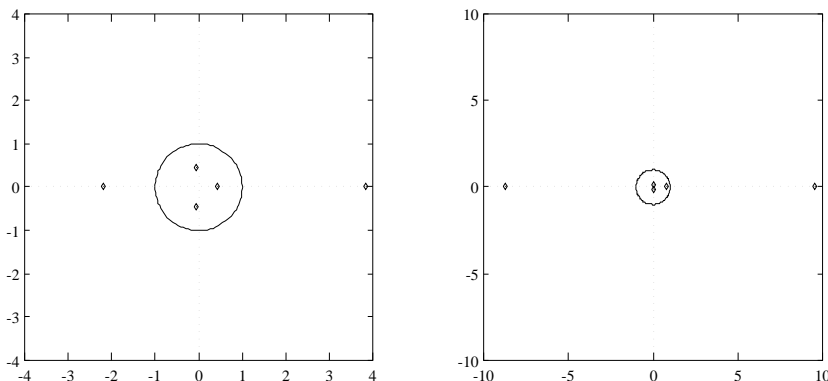


Figure 2: Position of the zeros for channel 1 and channel 2 respectively.

5.1 Experiment 1

In this experiment we first examine the sensitivity of both algorithms to the position of zeros in the unit circle and the choice of the oversampling factor. Two nonminimum phase channels were selected and the positions of the zeros are shown in Figure 2. The input is a bipolar binary signal and additive Gaussian noise was then added to the output of the channel. A symbol-spaced MLSE receiver using the Viterbi algorithm was used in the receiver with a delay of 10 symbols. The oversampling factor is $P = 3$ and in the subspace algorithm the width of the temporal window is $N = 10$ and the degree of ISI is $M = 1$. A signal subspace based quadratic constraint was selected to minimise the cost function in (25) [5]. For the cyclic algorithm both parameters p and q were 11 for channel 1 and channel 2. The performance was measured in terms of BER over 20 Monte-Carlo trials using 200 symbols in the estimation of the channel. The results are shown in Figure 3 and it suggests that in both cases the subspace method performs better.

A second simulation was conducted to evaluate the performance in terms of the variance of the channel estimation and the results are shown in Figure 4. The sample variances of the fractionally-spaced channel estimates are shown for an oversampling factor $P = 3$ and again it shows the better performance of the subspace method for both channels. The variance of the sample estimates was obtained using 20 Monte-Carlo runs and for a $E_b/N_0 = 15$ dB scenario. The asymptotic variance of the subspace algorithm was also simulated. However, a second run of the cyclic algorithm using an oversampling factor of $P = 5$ suggests that the variance of the channel estimates can be brought down to almost the same level as the subspace algorithm. This fact can be understood by looking at the true values of the differential cepstrum parameters $A(k)$ and $B(k)$ as shown in Table 1. Because $A(k)$ and $B(k)$ can only be estimated for those values of k which fulfill the condition $k \neq P, 2P, \dots$ and those *singular* values can only be estimated using interpolation such as the Newton-Aitken interpolation technique [9], we see that for a $P = 5$ scenario the interpolation is closer to the true values of $A(k)$ and $B(k)$ and the performance is better. Despite the

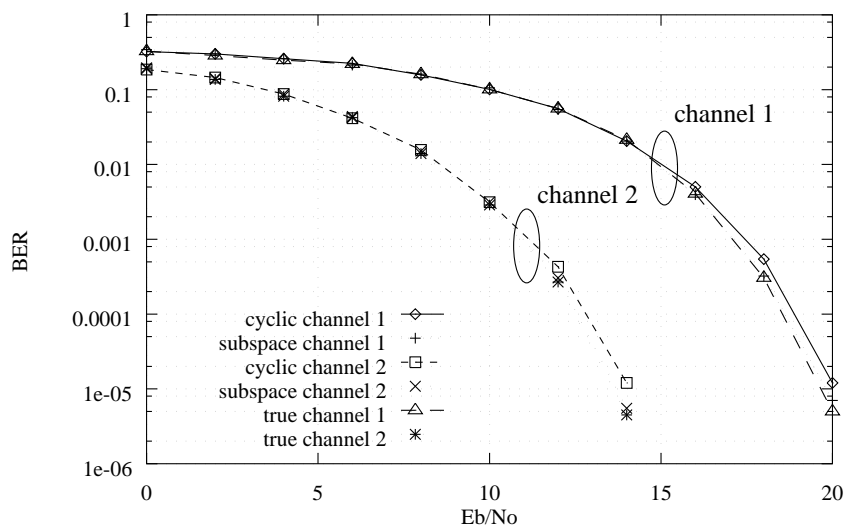


Figure 3: Comparison of cyclic and subspace algorithms in terms of BER for different E_b/N_0 scenarios for an oversampling factor of $P = 3$.

closeness of one of the zeros to the unit circle in the case of channel 2 which traduces in a longer value of p for the $A(k)$ sequence, the performance is better than for channel two for both $P = 3$ and $P = 5$. This is due to better interpolation carried out with both $P = 3$ and $P = 5$ in the case of channel 2 compared to channel 1.

5.2 Experiment 2

In this experiment we evaluate the performance of the subspace method for blind channel identification compared to conventional supervised methods in the pan-European digital mobile radio system GSM. The propagation channel was modelled according to the COST207 directive set up by the European Commission and was implemented using the model developed by [3]. The central carrier frequency of the GSM-TDMA system is 900 MHz with 8 channels (frames) per carrier, each spaced by a 200 kHz bandwidth. The channel data rate as in GSM is 270.833 kbits/s, which results from transmitting at a baud-period of $3.69 \mu\text{s}$. The structure of a typical time-slot is shown in Figure 5 where two bursts of information data containing 58 bits each are transmitted with a midamble 26 bits for the training of conventional supervised algorithms. In this scenario a complete time slot is read first and the training sequence from left to right is used to recover the second burst of data and next the training sequence from right to left is used to equalise the first burst of data from right to left. In blind equalisation the 26 bits allocated for training can be reallocated as information data which as a result gives an information data rate of 246.53 kbps compared to the 201.07 kbps of the conventional supervised method. The procedure for recovering the data requires the data from one time slot to be read and used to update the equaliser and recover the 142 data bits. The source signal is a 4-QAM signal and the transmitter filter is a raised-cosine spectrum type filter with roll-off factor $\beta = 0.1$.

The propagation channel is modelled as a multipath fading impulse response with six taps simulating a typical urban (TU) environment where the multipath effect extends to up to $5 \mu\text{s}$ with a significant reflector occurring at $0.2 \mu\text{s}$ due to large buildings nearby. A Doppler frequency of 40 Hz was considered as a result of a vehicular speed of 48 km/h. In a second experiment a Doppler frequency of 125 Hz was considered for a vehicular speed of 150 km/h. The receiver comprises an analog receiver filter and the output from

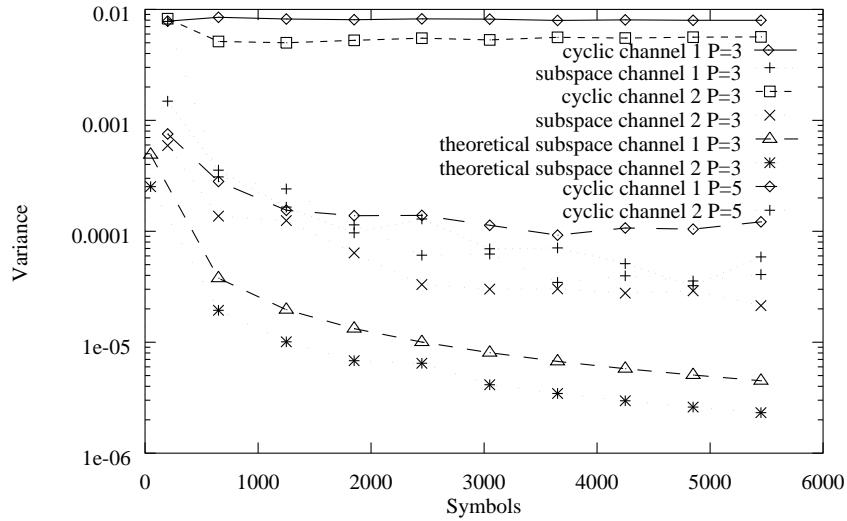


Figure 4: Variance of channel parameter estimates for channel 1 and channel 2; sample estimates and theoretical lower bound for subspace algorithm.

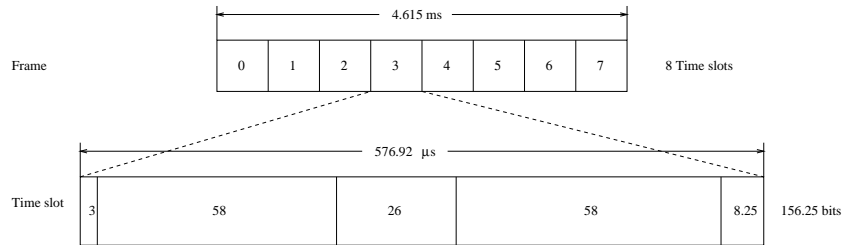


Figure 5: Time slot structure for a 8 channel TDMA system.

the Rx filter is sampled $P = 3$ times faster than the symbol rate. The equaliser is a symbol-spaced Viterbi equaliser (VE) with 4^3 states. A delay of 5 times the memory of the equaliser was used in making the decisions.

The simulation was performed by comparing the subspace blind equalisation method, a supervised non-adaptive algorithm and a supervised adaptive LMS algorithm. The non-adaptive method estimates the channel through the training sequence and this estimate is held fixed for the Viterbi algorithm for the entire burst and it is only updated in the next time slot. The *adaptive* receiver uses a *least-mean square* (LMS) algorithm where the training bits are used for start-up purposes and the equaliser parameters are constantly updated through the burst. Figure 6 shows the BER performance for different E_b/N_0 scenarios, where 20 realisations of the thermal noise for a Doppler spectrum of 40 Hz and 125 Hz respectively. Data from 4 time slots have been used by the blind equalisation method. In Figure 7 we note that taking data from adjacent time slots reduces the BER but this reduction is less apparent as the Doppler spectrum increases. In both Doppler spectrum scenarios, data from 1 time slot is insufficient for the blind algorithm to achieve a similar BER performance of the conventional supervised methods and even with 4 time slots the achieved BER is considerably worse than conventional supervised methods. The fact that the Blind Equaliser (BE) does not need a dedicated training signal means that the spectral efficiency of the system can be improved. All the experiments were conducted by running the respective algorithms for 5 *periods* of the propagation channel in order to avoid locally generated results.

Channel 1						
k	True values		Interpolated $P = 3$		Interpolated $P = 5$	
	$A(k)$	$B(k)$	$A(k)$	$B(k)$	$A(k)$	$B(k)$
0	0.3236	-0.1942	0.3236	-0.1942	0.3236	-0.1942
1	-0.1905	0.2740	-0.1905	0.2740	-0.1905	0.2740
2	0.1377	-0.0762	-0.1141*	0.2495*	0.1377	-0.0762
3	0.0993	0.0472	0.0993	0.0472	0.0993	0.0472
4	-0.0039	-0.0182	-0.0039	-0.0182	0.0389*	0.0508*
5	-0.0046	0.0091	-0.0143*	-0.0226*	-0.0046	0.0091
6	0.0073	-0.0039	0.0073	-0.0039	0.0073	-0.0039

Channel 2						
k	True values		Interpolated $P = 3$		Interpolated $P = 5$	
	$A(k)$	$B(k)$	$A(k)$	$B(k)$	$A(k)$	$B(k)$
0	0.7716	-0.0098	0.7716	-0.0098	0.7716	-0.0098
1	0.5223	0.0242	0.5223	0.0242	0.5223	0.0242
2	0.4290	-0.0004	0.3960*	0.0179*	0.4290	-0.0004
3	0.3259	0.0003	0.3259	0.0003	0.3259	0.0003
4	0.2452	-0.0000	0.2452	-0.0000	0.2458*	0.0002*
5	0.1851	0.0000	0.2023*	-0.0000*	0.1851	0.0000
6	0.1397	-0.0000	0.1397	-0.0000	0.1397	-0.0000

Table 1: Differential cepstrum parameters for channel 1 and channel 2. True values and interpolated for oversampling factors of $P = 3$ and $P = 5$.

5.3 Discussion

The blind equalisation methods described in this paper are applicable to all channels except those

- with zeros on the unit circle [9]
- with zeros evenly spaced radially from 0 to 2π with angle $2\pi/P$ [9] [5]
- strictly bandlimited to less than $2\pi/T$ [9].

Among the advantages and disadvantages of each algorithm there is the fact in the subspace algorithm that the various noise processes $w^{(i)}(n)$ are not necessarily uncorrelated since the frequency $1/T_s$ may be higher than the Nyquist rate. In that case, a whitening operation has to be used which involves knowledge of the noise autocorrelation matrix. The advantages are the fast convergence and the fact that it is not as sensitive as the cyclic algorithm to the position of channel zeros. Equally to the algorithm in [10] the linear system of Equation 21 needs to have a precise rank and difficulties are to be expected when some zeros of any two subchannels get too close.

The cyclic algorithm has the advantage of fast convergence rates when zeros are far from the unit circle but the performance degrades for certain channels when interpolation of some of the values of the differential cepstrum parameters has to be made. The method can work equally well with finite impulse response (FIR) and infinite impulse response (IIR) channels and it is insensitive (in theory) to correlated noise. The problems associated with zeros on the unit circle or very close to it, sensitivity to interpolation depending on the channel and the fact that up to date information of the amount of zeros that are located inside and outside the unit circle must be known suggest that this method becomes impractical for mobile radio communications.

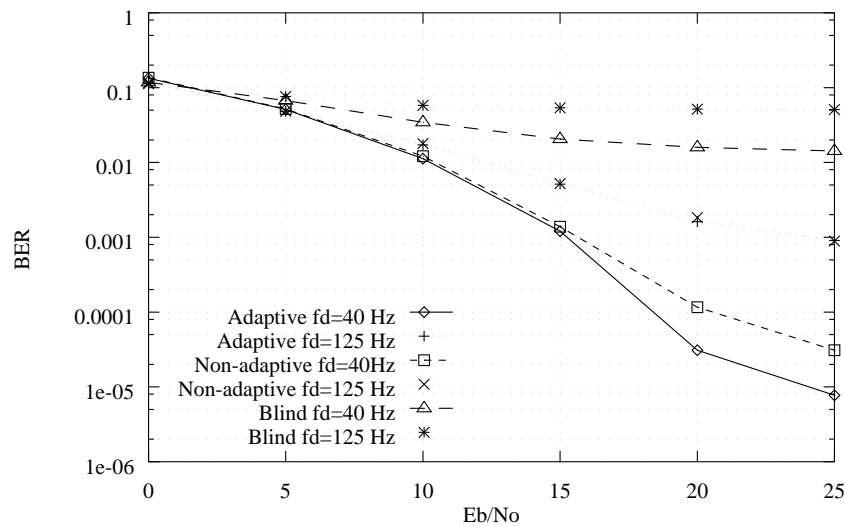


Figure 6: BER performance for different E_b/N_0 scenarios throughout the GSM data block for Doppler frequencies of $f_d = 40$ Hz and $f_d = 125$ Hz

6 CONCLUSION

This paper presents some results on the convergence speed of cyclostationary blind equalisation techniques in stationary and mobile radio environments. Two algorithms which are compared in a stationary environment are *i*) a statistical approach based on second-order cyclic statistics [9] and *ii*) a method which exploits matrix algebra described by a multichannel representation of the fractionally-spaced sampling operation [5]. It has been shown that although the convergence speed of the algorithms is fast for stationary channels, their convergence speed cannot be compared to conventional supervised methods in a mobile radio environment. It was also shown that problems are expected when *i.1*) zeros are present on the unit circle or very close to it or *i.2*) interpolation of some values of the differential cepstrum parameters has to be done for certain channels in the case of the cyclic statistics based algorithm, or *ii*) as a consequence of oversampling, the different subchannels share common zeros in the case of the subspace algorithm.

REFERENCES

- [1] L.A. Baccala and S. Roy. 'A new blind time-domain channel identification method based on cyclostationarity'. *IEEE Signal Processing Letters*, Vol 1(6), pp 89–91, June 1994.
- [2] Y. Chen and C.L. Nikias. 'Blind identifiability of a band-limited nonminimum phase system from its output autocorrelation', *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol IV, pp 444–447, 1993.
- [3] D.G.M. Cruickshank D.I. Laurenson and G.J.R. Povey. 'A computationally efficient channel simulator for the COST 207 channel models', *IEE Colloquium on Computer Modelling of Communication Systems*, May 1994.
- [4] Z. Ding. 'Characteristics of band-limited channels unidentifiable from second-order cyclostationary statistics', *IEEE Signal Processing Letters*, Vol 3(5), pp 150–152, May 1996.

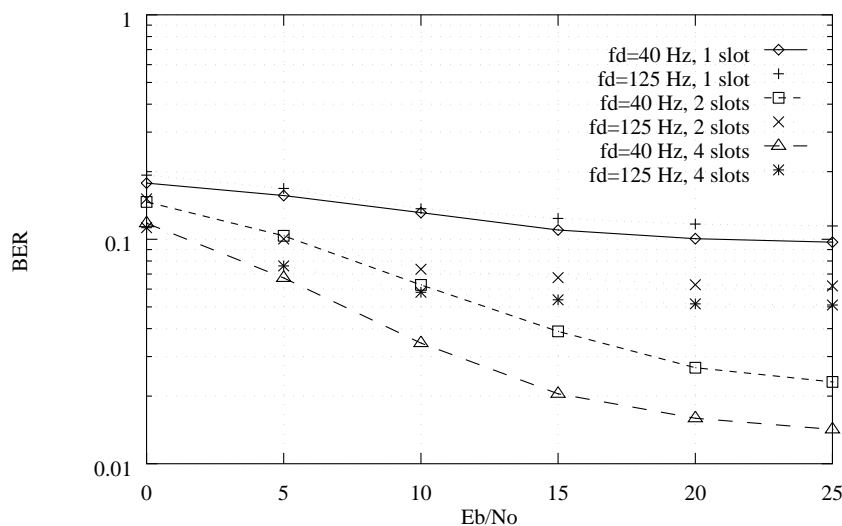


Figure 7: BER performance of the subspace algorithm for blind equalisation for different E_b/N_0 scenarios using 1, 2 and 4 TDMA time slots for block channel identification

- [5] J. Cardoso E. Moulines, P. Duhamel and S. Mayrargue. 'Subspace methods for the blind identification of multichannel FIR filters. *IEEE Transactions on Signal Processing*, Vol 43(2),pp 516–525, February 1995.
- [6] W.A. Gardner. 'A new method for channel identification' *IEEE Transactions on Communications*, Vol 39(6), pp 813–817, June 1991.
- [7] W.A. Gardner. *Cyclostationarity in Communications and Signal Processing*. IEEE Press, 1st edition, 1994.
- [8] L. Tong H. Liu, G. Xu and T. Kailath. 'Recent developments in blind channel equalization: from cyclostationarity to subspaces' *Signal Processing, Elsevier*, pp 83–89, April 1996.
- [9] D. Hatzinakos. 'Nonminimum phase channel deconvolution using the complex cepstrum of the cyclic autocorrelation' *IEEE Transactions on Signal Processing*, Vol 42(11), pp 3026–3042, November 1994.
- [10] G.Xu L. Tong and T. Kailath. 'A new approach to blind identification and equalization of multipath channels' *Proceedings of the 25th IEEE Asilomar Conference on Signals, Systems and Computers*, pp 856–860, Pacific Grove, CA, USA, November 1991.
- [11] S. Prakriya and D. Hatzinakos. 'Identification of parametric models with cyclostationary inputs' In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol IV, pp 417–419, Adelaide, Australia, April 1994.
- [12] D.T.M. Slock and C.B. Papadias. 'Blind fractionally-spaced equalization based on cyclostationarity' In *Proceedings of the Vehicular Technology Conference*, Stockholm, Sweden, June 1994.
- [13] G. Xu, H. Liu, L. Tong and T. Kailath. 'A least-squares approach to blind equalisation' *IEEE Transactions on Signal Processing*, Vol 43(12), December 1995.

- [14] Y. Hua. 'Fast maximum likelihood for blind identification of multiple FIR channels'. *IEEE Transactions on Signal Processing*, Vol 44(3), pp 661–672, March 1996.
- [15] H. Liu and Y. Hua. 'Performance analysis of the subspace method for blind channel identification' *Signal Processing, Elsevier*, pp 71–81, April 1996.
- [16] I. Lakkis, D. McLernon and L. Lopez. 'A novel blind channel identification and equalisation algorithm based on maximum likelihood' *to appear in Wireless Personal Communications, special issue on Synchronization and Equalization in Wireless Communications*.
- [17] H. Zeng and L. Tong. 'Some new results on blind channel estimation: performance and algorithms' *In 29th Annual Conference on Information Sciences and Systems*, March 1995.
- [18] G. Harikumar and Y. Bresler. 'Analysis and comparative evaluation of techniques for multichannel blind deconvolution' *In Proceedings SSAP'96*, pp. 332–335, Corfu, Greece, June 1996.
- [19] D. Yellin and B. Freidlander. 'Blind multichannel system identification and deconvolution: performance bounds' *In Proceedings SSAP'96*, pp. 582–585, Corfu, Greece, June 1996.

References

- [1] *Group Speciale Mobile (GSM) Recommendation*. April 1988.
- [2] K. Abed-Meraim and Y. Hua. Blind identification of multi-input multi-output system using minimum noise subspace. *IEEE Transactions on Signal Processing*, 45(1):254–258, January 1997.
- [3] S.A. Alshebeili, N. Venetsanopoulos, and A.E. Cetin. Cumulant based identification approaches for nonminimum phase FIR systems. *IEEE Transactions on Signal Processing*, 41(4):1576–1588, April 1993.
- [4] K. Anand, G. Mathew, and V.U. Reddy. Blind separation of multiple co-channel BPSK signals arriving at an antenna array. *IEEE SP Letters*, 2(9):176–178, September 1995.
- [5] S. Anderson, M. Millnert, M. Viberg, and B. Wahlberg. An adaptive array for mobile communication systems. *IEEE Transactions on Vehicular Technology*, 40:230–236, February 1991.
- [6] L.A. Baccala and S. Roy. A new blind time-domain channel identification method based on cyclostationarity. *IEEE SP Letters*, 1(6):89–91, June 1994.
- [7] A.S. Bajwa and J.D. Parsons. Small-area characterisation of uhf urban and suburban mobile radio propagation. *IEE Proceedings*, 129(2):102–109, 1982.
- [8] A. Benveniste and M. Goursat. Blind equalizers. *IEEE Transactions on Communications*, 32(8):871–883, August 1984.
- [9] A. Benveniste, M. Goursat, and G. Ruget. Robust identification of a nonminimum phase system: blind adjustment of a linear equalizer in data communications. *IEEE Transactions on AC*, 25(3):385–399, June 1980.
- [10] D. Boss, M. Boe, and K.D. Kammeyer. Exploiting 2nd order cyclostationarity or higher order statistics for the blind identification of mixed phase FIR systems? In *GRETSI*, Juan Les Pins, France, September 1995.
- [11] D. Boss and K. Kammeyer. Blind GSM channel estimation based on higher-order statistics. In *International Conference on Communications*, Quebec, Canada, June 1997.
- [12] A.G. Burr. The multipath problem: An overview. In *Colloquium on Multipath Countermeasures*, London, 1995. IEE.
- [13] G. Castellini, F. Conti, E. Del Re, and L. Pierucci. A continuously adaptive MLSE receiver for mobile communications: algorithms and performance. *IEEE Transactions on Communications*, 45(1):80–89, January 1997.
- [14] Y. Chen and C.L. Nikias. Blind identifiability of a band-limited nonminimum phase system from its output autocorrelation. In *ICASSP*, volume IV, pages 444–447. IEEE, 1993.
- [15] A. Chevreuril and Ph. Loubaton. Blind second-order identification of FIR channels: forced-cyclostationarity and structured subspace method. In *SPAWC*, pages 121–124, Paris, France, April 1997. IEEE.

-
- [16] H. Chiang and C.L. Nikias. Adaptive deconvolution and identification of nonminimum phase FIR systems based on cumulants. *IEEE Transactions on AC*, 35(1):36–47, January 1990.
- [17] Y.S. Choi, H. Hwang, and D.I. Song. Adaptive blind equalization coupled with carrier recovery for hdtv modem. *IEEE Trans. on Consumer Electronics*, 39(3):386–391, August 1993.
- [18] H.A. Cirpan and M.K. Tsatsanis. Blind receivers for nonlinearly modulated signals in multipath. In *SPAWC*, pages 149–152, Paris, France, April 1997. IEEE.
- [19] A.P. Clark. *Equalizers for Digital Modems*. Pentech Press, 1985.
- [20] A.V. Dandawaté and G.B. Ginnakis. Asymptotic theory of mixed time averages and Kth-order cyclic-moment and cumulant statistics. *IEEE Transactions on Information Theory*, 41(1):216–232, January 1995.
- [21] L.M. Davis, I.B. Collings, and R.J. Evans. Constrained maximum likelihood estimation of time-varying linear channels. In *SPAWC*, pages 1–4, Paris, France, April 1997. IEEE.
- [22] J. Buisán Gómez del Moral and E. Biglieri. Blind identification of digital communication channels with correlated noise. *IEEE Transactions on Signal Processing*, 44(12):3154–3156, December 1996.
- [23] L. Deniere and D.T.M. Slock. Second-order cyclic statistics based blind channel identification and equalization. In *SPAWC*, pages 21–24, Paris, France, April 1997. IEEE.
- [24] Z. Ding. Blind channel identification and equalization using spectral correlation measurements, part I: Frequency-domain analysis. In *Cyclostationarity in Communications and Signal Processing*, pages 417–436, 1994.
- [25] Z. Ding. Characteristics of band-limited channels unidentifiable from second-order cyclostationary statistics. *IEE SP Letters*, 3(5):150–152, May 1996.
- [26] Z. Ding, R.A. Kennedy, B.D.O. Anderson, and C.R. Johnson. Ill-convergence of Godard blind equalizers in data communication systems. *IEEE Transactions on Communications*, 39(9):1313–1327, September 1991.
- [27] J.J. Dongarra et al. LINPACK user’s guide. *Society for Industrial and Applied Mathematics*, 1979.
- [28] M. Failli. COST207 - Digital land mobile radio communications. Technical report, Directorate-General Telecommunications, Information, Industries and Innovation, Commission of the European Communities, 1989.
- [29] G.D. Forney. Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference. *IEEE Transactions on Information Theory*, 18(3):363–378, May 1972.
- [30] G.J. Foschini. Equalizing without altering or detecting data. *AT&T Technical Journal*, 64:1885–1911, October 1985.
- [31] W.A. Gardner. *Statistical Spectral Analysis: A Nonprobabilistic Theory*. Prentice-Hall, 1987.
- [32] W.A. Gardner. Simplification of MUSIC and ESPRIT by exploitation of cyclostationarity. *Proceedings of the IEEE*, 76(7):845–847, July 1988.
- [33] W.A. Gardner. *Introduction to Random Processes with Applications to Signals and Systems*, pages 323–415. McGraw-Hill, 2nd edition, 1990.
- [34] W.A. Gardner. A new method for channel identification. *IEEE Transactions on Communications*, 39(6):813–817, June 1991.

-
- [35] W.A. Gardner. Cyclic Wiener filtering: theory and method. *IEEE Transactions on Communications*, 41(1):151–163, January 1993.
- [36] W.A. Gardner. *Cyclostationarity in Communications and Signal Processing*. IEEE Press, 1st edition, 1994.
- [37] G.B. Giannakis and J.M. Mendel. Identification of nonminimum phase systems using higher order statistics. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(3):360–377, March 1989.
- [38] G.B. Giannakis and A. Swami. On estimating noncausal nonminimum phase arma models of non-gaussian processes. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(3):478–495, March 1990.
- [39] G.B. Giannakis and M.K. Tsatsanis. HOS or SOS for parametric modeling? In *ICASSP*, volume V, pages 3097–3100, Toronto, Canada, May 1991. IEEE.
- [40] R.D. Gitlin and S.B. Weinstein. Fractionally-spaced equalization: An improved digital transversal equalizer. *Bell Systems Technical Journal*, 60:275–296, February 1981.
- [41] D. Godard. Self-recovering equalization and carrier tracking in two-dimensional data communication systems. *IEEE Transactions on Communications*, 28(11):1867–1875, November 1980.
- [42] D.J. Goodman. Trends in cellular and cordless communications. *IEEE Communications Magazine*, June 1991.
- [43] G. Harikumar and Y. Bresler. Analysis and comparative evaluation of techniques for multichannel blind deconvolution. In *SSAP*, pages 332–335, Corfu, Greece, June 1996.
- [44] D. Hatzinakos. Carrier phase tracking and tricepstrum-based blind equalization. *Computers and EE*, 18(2):109–118, February 1992.
- [45] D. Hatzinakos. Nonminimum phase channel deconvolution using the complex cepstrum of the cyclic autocorrelation. *IEEE Transactions on Signal Processing*, 42(11):3026–3042, November 1994.
- [46] D. Hatzinakos and C.L. Nikias. Blind equalization using a tricepstrum-based algorithm. *IEEE Transactions on Communications*, 39(5):669–681, May 1991.
- [47] S. Haykin. *Adaptive Filter Theory*. Prentice-Hall, 3rd edition, 1996.
- [48] Y. Hua. Fast maximum likelihood for blind identification of multiple FIR channels. *IEEE Transactions on Signal Processing*, 44(3):661–672, March 1997.
- [49] C.R. Johnson Jr. et al. On fractionally-spaced equalizer design for digital microwave radio channels. In *29th Asilomar Conference on Signals, Systems and Computers*, California, October 1995.
- [50] K. Kammeyer and B. Jelonnek. A new fast algorithm for blind MA-system identification based on higher order cumulants. In *SPIE Advanced Signal Processing: Algorithms, Architectures & Implementations*, pages 9–14, San Diego, USA, July 1994.
- [51] I. Lakkis, D. McLernon, and L. Lopes. A novel blind channel identification and equalisation algorithm based on maximum likelihood. *to be published in Wireless Personal Communications*, 1997.
- [52] D.I. Laurenson, D.G.M. Cruickshank, and G.J.R. Povey. A computationally efficient channel simulator for the COST 207 channel models. In *Colloquium on Computer Modelling of Communication Systems*. IEE, May 1994.
- [53] Y. Li and Z. Ding. Blind channel identification based on second order cyclostationary statistics. In *ICASSP*, volume IV, pages 81–84, Minneapolis, Minnesota, USA, April 1993. IEEE.

- [54] Y. Li and Z. Ding. ARMA system identification based on second-order cyclostationarity. *IEEE Transactions on Signal Processing*, 42(12):3483–3494, December 1994.
- [55] Y. Li and K.J.R. Liu. Blind identification and equalization for multiple-input/multiple-output channels. In *GLOBECOM*, volume 3, pages 1789–1793, London, UK, November 1996. IEEE.
- [56] H. Liu and G. Xu. A deterministic approach to blind symbol estimation. *IEEE SP Letters*, 1(12):205–207, December 1994.
- [57] H. Liu, G. Xu, L. Tong, and T. Kailath. Recent developments in blind channel equalization: from cyclostationarity to subspaces. *Signal Processing, ELSEVIER*, pages 83–89, April 1996.
- [58] R.W. Lucky. Automatic equalization for digital communication. *Bell Systems Technical Journal*, (44):547–588, 1965.
- [59] J. Mendel. Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proceedings of the IEEE*, 79(3):277–304, March 1991.
- [60] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue. Subspace methods for the blind identification of multichannel FIR filters. In *ICASSP*, volume IV, pages 573–576, Adelaide, Australia, April 1994. IEEE.
- [61] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue. Subspace methods for the blind identification of multichannel FIR filters. *IEEE Trans. on SP*, 43(2):516–525, February 1995.
- [62] A.F. Naguib, B. Khalaj, A. Paulraj, and T. Kailath. Adaptive channel equalization for TDMA digital cellular communications using antenna arrays. In *ICASSP*, volume IV, Adelaide, Australia, April 1994. IEEE.
- [63] A. Nandi. Blind identification of FIR systems using third order cumulants. *IEEE Transactions on Signal Processing*, 39(3), September 1994.
- [64] P. Newson. *Adaptive Algorithms for Equalisers and Channel Estimators for Use Within Digital Mobile Radio Systems*. PhD thesis, The University of Edinburgh, September 1992.
- [65] J.J. Nicolas and J.S. Lim. Equalization and interference rejection for the terrestrial broadcast of digital HDTV. In *ICASSP*, volume IV, pages 176–179, Minneapolis, USA, April 1993. IEEE.
- [66] C.L. Nikias and M.R. Raghuveer. Bispectrum estimation: a digital signal processing framework. *IEEE Proceedings*, 75(7):869–891, July 1987.
- [67] R. Pan and C.L. Nikias. The complex cepstrum of the higher order cumulants and nonminimum phase system identification. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 36(2):186–205, February 1988.
- [68] R. Penrose. *A Generalized Inverse for Matrices*, pages 406–413. 1st edition, 1954.
- [69] G. Picchi and G. Prati. Blind equalization and carrier recovery using 'stop-and-go' decision-directed algorithm. *IEEE Transactions on Communications*, 35(9):877–887, September 1987.
- [70] S. Prakriya and D. Hatzinakos. Identification of parametric models with cyclostationary inputs. In *ICASSP*, volume IV, pages 417–419, Adelaide, Australia, April 1994. IEEE.
- [71] J.G. Proakis. *Digital Communications*. New York: McGraw-Hill, 3rd edition, 1995.
- [72] J.G. Proakis and D. Manolakis. *Introduction to Digital Signal Processing*. Macmillan, 1989.

- [73] W. Qiu and Y. Hua. Performance analysis of the subspace method for blind channel identification. *Signal Processing, ELSEVIER*, (50):71–81, April 1996.
- [74] S.U.H. Qureshi and G.D. Forney. Performance and properties of a $T/2$ equalizer. *National Telecommunications Conference Record*, pages 11.1.1–11.1.14, December 1977.
- [75] U.H. Qureshi. Adaptive equalization. *Proceedings of the IEEE*, 73(9):1349–1387, May 1985.
- [76] R. Roy and T. Kailath. ESPRIT- estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37:984–995, July 1989.
- [77] Y. Sato. A method of self-recovering equalization for multilevel amplitude-modulation systems. *IEEE Transactions on Communications*, 23:679–682, June 1975.
- [78] S.V. Schell, R.A. Calabretta, W.A. Gardner, and B.G. Agee. Cyclic MUSIC algorithms for signal-selective direction estimation. In *ICASSP*, volume II, pages 2278–2281, Glasgow, Scotland, May 1989. IEEE.
- [79] S.V. Schell and W.A. Gardner. Maximum likelihood and common factor analysis-based blind adaptive spatial filtering for cyclostationary signals. In *ICASSP*, volume IV, pages 292–295, Minneapolis, USA, 1993. IEEE.
- [80] R.O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas Propagation*, 34:276–280, March 1986.
- [81] D.T.M. Slock. Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction. In *ICASSP*, volume IV, pages 585–588, Adelaide, Australia, April 1994. IEEE.
- [82] D.T.M. Slock. Blind joint equalization of multiple synchronous mobile users using oversampling and/or multiple antennas. In *ASILOMAR*, volume 2, pages 1154–1158, Pacific Grove, CA, USA, 1995. IEEE.
- [83] R. Steele. *Mobile Radio Communications*. Pentech Press Ltd, 1992.
- [84] A. Stogioglou and S. McLaughlin. MA parameter estimation and cumulant enhancement. *IEEE Transactions on Signal Processing*, 44(7):1704–1718, July 1996.
- [85] P. Stoica and A. Nehorai. MUSIC, Maximum Likelihood, and Cramer-Rao bound. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(5):720–741, May 1989.
- [86] S. Talwar, M. Viberg, and A. Paulraj. Blind estimation of multiple co-channel digital signals using an antenna array. *IEEE SP Letters*, 1(2):29–31, February 1994.
- [87] L. Tong, G. Xu, and T. Kailath. A new approach to blind identification and equalization of multipath channels. In *25th Asilomar Conference on Sig., Sys. and Comp.*, pages 856–860, Pacific Grove, CA, USA, November 1991. IEEE.
- [88] L. Tong, G. Xu, and T. Kailath. Blind identification and equalization based on second-order statistics: a time domain approach. *IEEE Transactions on Information Theory*, pages 340–349, March 1994.
- [89] L. Tong, G. Xu, and T. Kailath. Blind identification and equalization of multipath channels. In *ICC*, volume III, pages 1513–1517, Chicago, USA, June 1992. IEEE.
- [90] L. Tong, G. Xu, and T. Kailath. Fast blind equalization via antenna arrays. In *ICASSP*, volume IV, pages 272–275. IEEE, 1993.
- [91] L. Tong, G. Xu, and T. Kailath. *Cyclostationarity in Communications and Signal Processing*, chapter A.5, pages 437–454. IEEE Press, 1994.

-
- [92] J.R. Treichler, I. Fikalkow, and C.R. Johnson Jr. Fractionally spaced equalizers - how long should they really be? *IEEE Signal Processing Magazine*, 13(3):65–81, May 1996.
- [93] J. Tugnait. On blind identifiability of multipath channels using fractional sampling and second-order cyclostationary statistics. In *GLOBECOM*, volume III, pages 2001–2005, Houston, Texas, December 1993. IEEE.
- [94] J.K. Tugnait. Blind equalization and estimation of digital communication FIR channels using cumulant matching. *IEEE Transactions on Communications*, 43(2/3/4):1240–1245, February/March/April 1995.
- [95] J.K. Tugnait and U. Gummadavelli. Blind channel estimation and deconvolution in colored noise using higher-order cumulants. 1995.
- [96] G. Ungerboeck. Fractional tap-spacing equalizer and consequences for clock recovery in data modems. *IEEE Transactions on Communications*, 24(8):856–864, August 1976.
- [97] A. van der Veen. Blind estimation of multiple digital signals transmitted over FIR channels. *IEEE SP Letters*, 2(5):99–102, May 1995.
- [98] M. Webster and N. Tepedelenlioglu. Frequency-domain techniques for the cyclostationary signals encountered in fractionally-spaced equalizers. In *ICASSP*, volume IV, pages 705–708, San Francisco, USA, September 1992. IEEE.
- [99] Y. Wu and B. Caron. Digital television terrestrial broadcasting. *IEEE Communications Magazine*, pages 46–52, May 1994.
- [100] G. Xu, H. Liu, L. Tong, and T. Kailath. A least-squares approach to blind channel identification. *IEEE Transactions on Signal Processing*, 43(12):2982–2993, December 1995.
- [101] D. Yellin and B. Freidlander. Blind multichannel system identification and deconvolution: Performance bounds. In *SSAP*, pages 582–585, Corfu, Greece, June 1996.
- [102] Q. Zhang and L. Tong. Applications of blind channel estimation to HF communications. In *SPAWC*, pages 125–128, Paris, France, April 1997. IEEE.
- [103] X.D. Zhang and Y.S. Zhang. FIR system identification using third order cumulants. *IEEE Transactions on Signal Processing*, 42(10):2854–2858, October 1994.