NOTE TO USERS

Page(s) not included in the original manuscript and are unavailable from the author or university. The manuscript was scanned as received.

iii

This reproduction is the best copy available.

UMI[®]

FREQUENCY DOMAIN PROCESSING FOR MULTIPLE INPUT MULTIPLE OUTPUT CHANNELS

By
Reza Kalbasi

A thesis submitted to the Faculty of Graduate Studies and Research in
partial fulfillment of the requirements of the degree of

Doctor of Philosophy

Ottawa-Carleton Institute for Electrical and Computer Engineering
Faculty of Engineering
Department of Systems and Computer Engineering

CARLETON UNIVERSITY
OTTAWA, ON, K1S 5B6
SEPTEMBER 2005

© Copyright by Reza Kalbasi, 2005

Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-494-08335-2

Our file Notre référence

ISBN: 0-494-08335-2

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

The undersigned hereby recommend to
the Faculty of Graduate Studies and Research
Acceptance of the thesis

**Frequency Domain Processing for Multiple Input
Multiple Output Channels**

Submitted by

Reza Kalbasi

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Co-supervisors:

Professor David D. Falconer

Professor Amir H. Banihashemi

External Examiner:

Professor Steven D. Blostein

Department Chair:

Professor Rafik A. Goubran

CARLETON UNIVERSITY

SEPTEMBER 2005

*To My Parents,
My sister and brother*

Table of Contents

Table of Contents	v
List of Tables	ix
List of Figures	ix
Abstract	xiv
Acknowledgements	xvi
Abbreviations	xix
Notation	xxiii
1 Introduction	1
1.1 Multiple-Input-Multiple-Output Wireless Communications Systems .	1
1.2 Thesis Objective	3
1.3 Thesis Overview	3
1.4 Thesis Contributions	5
2 Multiple Input Multiple Output Systems for Non-Frequency Selective Channels	6
2.1 Multiple Antenna Systems: Definition and Notations	7
2.2 MIMO Channel Capacity	8
2.2.1 Ergodic Capacity	8
2.2.2 Outage Capacity	10
2.3 BLAST Architecture	13
2.3.1 Diagonal BLAST	14
2.3.2 Successive Interference Cancellation	16
2.4 Vertical-BLAST (V-BLAST)	17

2.4.1	Ordered Successive Interference Cancellation Algorithm	18
2.5	Combined ML and OSIC detection for V-BLAST	21
2.6	Adaptive Group Detection Algorithm	22
2.7	Efficient Algorithm for BLAST	24
2.7.1	Square Root Algorithm	24
2.8	Space-Time Codes	27
2.8.1	Design Criterion	27
2.8.2	Space-Time Trellis Codes	29
2.8.3	Space-Time Block Codes	31
2.9	Summary	34
3	MIMO Channel Equalization	36
3.1	System Model	37
3.2	Finite length MMSE MIMO-DFE	39
3.3	Wideband Layered Space-Time Receiver	44
3.3.1	System and Channel Model	45
3.4	Generalized Wideband Layered Space-Time Receivers	50
3.5	Infinite-Length	51
3.6	Zero-Forcing (ZF) Design	54
3.7	Remarks on the Finite-Length MMSE MIMO-DFE	55
3.8	Summary	56
4	Frequency Domain Equalization for MIMO Channels	58
4.1	OFDM	59
4.1.1	Channel Equalization	61
4.2	Single-Carrier Frequency Domain Equalization	63
4.3	Hybrid Time-Frequency Domain MIMO Receivers	65
4.3.1	System Model	66
4.3.2	Hybrid Time-Frequency LST Receiver	68
4.3.3	Hybrid Time-Frequency MIMO-DFE Receiver	73
4.4	Uncoded Simulation Results	77
4.4.1	Simulation Results with Genie Receiver	82
4.5	Comparison to MIMO-OFDM	89
4.6	Complexity Analysis	98
4.7	Summary	99
5	Iterative Block DFE (IB-DFE)	101
5.1	System Model	102
5.2	Receiver Structure	103

5.2.1	Iterative LST with Successive Interference Cancellation(SIC)	104
5.2.2	Iterative LST with Parallel Interference Cancellation (PIC)	105
5.2.3	Derivation of the Equalizer Coefficients	107
5.2.4	Computation of the Correlation Coefficient	113
5.3	Complexity Comparison between SIC and PIC Receivers	115
5.4	Performance Results	115
5.4.1	Multiuser Scenarios	119
5.5	Performance of IB-DFE for the 6-Ray channel	122
5.6	Summary	125
6	Joint Transmitter and Receiver Design	127
6.1	System Model	130
6.2	Optimization of MIMO Unlimited bandwidth System	132
6.3	Solution	133
6.4	Optimal Frequency Assignments	139
6.5	Algorithm for Finding the Optimal Sets	141
6.6	Numerical Results	143
6.7	Summary	146
7	Frequency Domain Pre-Equalization (Precoding)	148
7.1	System Model	149
7.2	Spatial Correlation Model	157
7.3	System Performance	158
7.4	Summary	164
8	Optimal Power Allocation for a two transmitter V-BLAST System	174
8.1	Proposed Scheme for Power Allocation	175
8.2	$2 \times N$ system:	177
8.3	Simulation Results	180
8.4	Conclusions	182
9	Conclusions	184
A	MMSE Expression for Hybrid-Time Frequency LST Receiver	189
B	MMSE Expression for Hybrid-Time Frequency MIMO-DFE	191
C	Chapter 6 Proofs	193
C.0.1	Proof of Lemma 1	193
C.0.2	Proof of Lemma 2	194

C.0.3	Proof of Diagonality of Φ	194
C.0.4	Proof of (6.3.14)	195
D	Peak Power of the Output Signal	196
E	Proof of Independence of Noise at Different Detection Stages	198
	Bibliography	200

List of Tables

4.1	UMTS-A Channel: Delay Spread Models	78
4.2	Signal Processing Complexity	98
4.3	Initialization Complexity	99
5.1	Signal Processing Complexity for the first iteration	115
5.2	Signal Processing Complexity for the Remaining Iterations	115
5.3	Power delay profile type C for HIPERLAN/2	119

List of Figures

2.1	Block diagram of a (P, N) MIMO systems	9
2.2	Outage Capacity of (N, N) System with $N = 1, 2, 4, 8, 16$	13
2.3	Diagonal layering	15
2.4	Block diagram of a $(4, 4)$ D-BLAST Architecture	16
2.5	V-BLAST block diagram	18
2.6	4-PSK 4-State Space-Time Code with 2 Tx antennas	30
2.7	STBC (Alamouti) with two transmit antenna	33
2.8	Receiver structure for STBC (Alamouti)	34
3.1	Block Diagram of MIMO-DFE	40
3.2	Block Diagram of Partially Connected OSIC-Receiver	46
3.3	Block Diagram of Fully Connected OSIC-Receiver	49
3.4	Block Diagram of generalized layered space-time receiver with $m_i^j =$ $P - 2(j - 1)$ and $m_o^j = 2$	51
3.5	Block Diagram of the j -th stage	52
4.1	Block Diagram of OFDM	61
4.2	Block Diagram of SC-FDE SC-FDE	64
4.3	Block diagram of LST-FDE receiver	69
4.4	Block diagram of MISO-DFE receiver	70
4.5	Block diagram of MIMO DFE	74
4.6	BER for 4×4 6-ray channel for different receivers	80
4.7	BER for 4×6 6-ray channel for different receivers	81

4.8	BER for 4×4 UMTS-A channel for different receivers	82
4.9	BER for 4×6 UMTS-A channel for different receivers	83
4.10	BER for different feedback taps for LST receiver 6-ray channel	84
4.11	BER for different feedback taps for MIMO-DFE receiver	85
4.12	BER for 4×4 UMTS-A channel for different receivers and high symbol rate	86
4.13	Comparison with genie receiver 6-ray channel	87
4.14	Comparison with genie receiver UMTS-A channel	88
4.15	Block Diagram of LST-OFDM	91
4.16	Coded BER for 4×4 6-ray channel for different receivers	92
4.17	Coded BER for 4×4 6-ray channel for different receivers	93
4.18	Coded BER for 4×4 UMTS-A channel for different receivers	95
4.19	Coded BER for 4×4 UMTS-A channel for different receivers	96
4.20	Coded BER for 4×4 UMTS-A channel for different receivers and high symbol rate	97
5.1	System model	103
5.2	Detection of the p -th layer for a given iteration	106
5.3	Layers BER for different iterations, and matched filter bound (MFB) performance for channel with independently fading tap at each fre- quency, $P = 4$, $N = 4$	117
5.4	Average BER for SIC receiver (solid line), and PIC receiver (dashed line), 4th iteration and matched filter bound performance (MFB) for channel with independently fading tap at each frequency, $P = 4$, $N = 4$	118
5.5	BER for different users (layers) and different iterations, and matched filter bound performance (MFB) for Hiperlan/2 power delay profile C, $P=4$, $N=4$	120
5.6	Average BER for SIC (solid line), and PIC receiver (dashed line), 4th iteration and matched filter bound performance (MFB) for Hiperlan/2 power delay profile C, $P=4$, $N=4$	121

5.7	Average BER for different number of antennas and users ,4th iteration and matched filter bound performance (MFB) for channel with independently fading tap at each frequency	122
5.8	BER for the high power users (a) and the low power users (b) for different iterations (it.), as well as the corresponding MFBs for channel with independently fading taps at each frequency, $P = 4$, $N = 4$. . .	123
5.9	Average BER for the high power and the low power users for different iterations (it.) for channel with independently fading tap at each frequency, $P = 4$, $N = 4$	124
5.10	Average BER for SIC (solid line), and PIC receiver (dashed line), 4th iteration and matched filter bound performance (MFB) for the 6-rays channel, $P = 4$, $N = 4$	125
6.1	Block Diagram of MIMO system	130
6.2	Block Diagram of Optimum MIMO System	135
6.3	Block Diagram of equivalent diagonal channel	137
6.4	Maximum eigenvalue and optimal support for two substreams, b_1 and b_2 are spectra of λ_{max} corresponding to substreams 1 and 2, respectively.	145
6.5	Comparison of MSE for 2×2 MIMO and different transmitter pulses.	146
6.6	Comparison of MSE for 3×3 MIMO and different transmitter pulses.	147
7.1	Block diagram of frequency domain transmitter	150
7.2	Comparison of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4$, $N = 4$, $K = 4$, $\rho = 0.9$	160
7.3	Comparison of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4$, $N = 4$, $K=4$, DOA= 50° , $\Delta = 40^\circ$.	161
7.4	Comparison of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4$, $N = 4$, $K=4$, DOA= 50° , $\Delta = 40^\circ$.	162
7.5	Comparison of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4$, $N = 4$, $K=4$, DOA= 50° , $\Delta = 60^\circ$.	163

7.6	Comparison of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4, N = 4, K=4$, DOA= $50^\circ, \Delta = 70^\circ$.	164
7.7	Comparison of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4, N = 4, K=4$, DOA= $50^\circ, \Delta = 80^\circ$.	165
7.8	Performance of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4, N = 4, K=4$, DOA= $50^\circ, \Delta = 30^\circ$.	166
7.9	Performance of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4, N = 4, K=4$, DOA= $50^\circ, \Delta = 20^\circ$.	167
7.10	Performance of the proposed precoding scheme for different number of substreams , $P = 4, N = 4, K=4, 3, 2$ DOA= $50^\circ, \Delta = 30^\circ$	168
7.11	Performance of the proposed precoding scheme for different substreams , $P = 4, N = 4, K=4, 3, 2$ DOA= $50^\circ, \Delta = 20^\circ$	169
7.12	Performance of the proposed precoding scheme for different number of substreams , $P = 4, N = 4, K=3, 2$ DOA= $50^\circ, \Delta = 10^\circ$	170
7.13	Performance of the proposed precoding scheme for different number of substreams , $P = 4, N = 4, K=3, 2$ DOA= $50^\circ, \Delta = 5^\circ$	171
7.14	Performance of the proposed precoding scheme for different correlations SNR=25dB, , $P = 4, N = 4, K=4$, DOA= 50°	172
7.15	Performance of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4, N = 4, K=4$, DOA= 0°	173
8.1	Block Diagram of New Scheme	180
8.2	Performance comparison of 2×2 V-BLAST	181
8.3	Performance Comparison of 2×3 V-BLAST	182
8.4	BER of the second detection 2×2	183

Abstract

In the past two decades wireless communication technology has developed tremendously. Therefore, the techniques that can support high data rates and provide high spectral efficiencies are of great interest. Employing multiple antennas at both receiver and transmitter and space-time processing techniques has been shown to be able to provide high spectral efficiencies. The fundamental idea is to use multiple antennas combined with signal processing techniques at the receiver in order to increase the spectral efficiencies. The focus of this thesis is on frequency domain transmitter and receiver design for severe frequency selective channels which can effectively mitigate the effect of ISI and interference for single carrier wireless systems with multiple transmit and receive antennas.

First, two new frequency-domain receivers for severe frequency-selective MIMO channels are introduced. These receivers have a frequency domain feedforward and a time domain feedback filter. Assuming the availability of perfect knowledge of the channel at the receiver, the performance of the proposed receivers are evaluated for different channel models. However, these receivers are non-linear and suffer from error propagation. The comparison of the performance of the proposed receivers with OFDM is also given.

Second, a new iterative receiver is proposed for severe time-dispersive MIMO channels. This receiver has a frequency domain feedforward and feedback filter. Since the feedback loop takes into account not just the hard-decisions for each block but also the overall block reliability, the error propagation problem is significantly reduced.

The problem of joint transmitter and receiver design for MIMO channels under the MSE criterion and with the perfect knowledge of the channel at the transmitter when the system does not have a priori bandwidth constraint is considered. It is proven that the optimum space-time filters are strictly bandlimited to a frequency set. The aggregated bandwidth of this set is at most the number of the substreams

times the symbol rate.

Next, we propose a linear frequency domain pre-equalization scheme for spatially correlated MIMO channels. In this scheme partial knowledge of the channel is available at the transmitter. To reduce the complexity the proposed precoder is implemented in the frequency domain.

Finally, we introduce a power optimization algorithm for V-BLAST system with two transmit antennas. This algorithm minimizes the exact probability of vector error.

Acknowledgements

There are many people I would like to thank to for their invaluable help during the process of writing of this thesis. First and foremost, I would like to thank my Ph.D. supervisor Professor David D. Falconer. He inspired me to think critically and build physical intuition on research problems. I am indebted to David for his mentorship, advice, moral support, patience, and for the great academic inspiration that he provided me with. He also created an excellent research environment which was flexible enough to accommodate with my interests to allow me to build my own research.

I would also like to thank my supervisor Professor Amir H. Banihashemi for the many helpful discussions regarding his thorough comments in writing papers and his support during my Ph.D. program. He always encouraged me to publish my research results and write papers. His precise feedback on the research manuscript helped me tremendously.

I am grateful to Professor Rui Dinis at CAPS-IST, Technical university of Lisbon, Portugal, for his guidance and helpful discussions. I enjoyed several long discussions with Rui, mostly while he was a visiting professor at Carleton and by email when he was in Portugal.

I would like to express my gratitude to Professor Saeed Gazor, my Master supervisor, who is currently teaching at Queen's university, Kingston, Ontario, Canada, for his invaluable help and advice. For me, Saeed has been a source of inspiration and motivation. He inspired me to think creatively, positively and to be confident in the early stage of my research career. I have greatly benefitted from having several scientific discussions with Saeed.

I would also thank Professor Halim Yanikomeroglu for agreeing to be on my thesis committee and his helpful advice during my years at Carleton, my external examiner Prof. Steven Blostein and members of my Ph.D. defense committee Prof. Ian

Marsland, Prof. Sergey Loyka and Prof. Len MacEachern for their suggestions and comments.

I thank my friends Eliana, Farhad , Keyvan and Payam for making my stay at Carleton a memorable one. I will always cherish their friendship, bike trips to Gatineau park, playing tennis, skating on the Canal and skiing. Moreover, I am appreciative of the staff at the department of systems and computer engineering, Blazanka Power, Darlene Hebert, Anna Lee and Jenifer Poll who provided much support during my stay at Carleton.

Finally, I would like to acknowledge my beloved parents, my sister and brother for their continuous love and support. They have helped me in every possible way to pursue and achieve my goals. They took care of all little things, and made sure that I had the time, space and resources to work hard with an open mind. I am forever indebted to them.

Ottawa, ON

Canada

June, 2005

Reza Kalbasi

Abbreviations

AGD	Adaptive Group Detection
BER	Bit Error rate
BLAST	Bell-Labs Layered Space-Time Architecture
D-BLAST	Diagonal-BLAST
BS	Base Station
CCI	Co-Channel Interference
CSI	Channel State Information
CDMA	Code Division Multiple Access
CP	Cyclic Prefix
DDFSE	Delay Decision Feedback Sequence Estimator
DOA	Direction of Arrival
DFE	Decision Feedback Equalizer

DFT	Discrete Fourier Transform
DSP	Digital Signal Procesing
FFT	Fast Fourier Transform
FDD	Frequency Division Duplex
FDE	Frequency Domain Equalization
FDMA	Frequency Division Multiple Access
FIR	Finite Impulse Response
GSM	Global System for Mobile Communications
IBI	Inter-Block Interference
IB-DFE	Iterative-Block Decision Feedback Equalizer
IDFT	Inverse Discrete Fourier Transform
IFFT	Inverse Fast Fourier Transform
i.i.d	Independent and Identically Distributed
IIR	Infinite Impulse Response
ISI	Inter-Symbol Interference
LAN	Local Area Network

LST	Layered Space-Time
MFB	Matched Filter Bound
MIMO	Multiple-Input Multiple-Output
MLSE	Maximum-Likelihood Sequence Estimation
MMSE	Minimum Mean-Square Error
MSE	Mean-Square Error
ML	Maximum Likelihood
MU	Mobile Unit
PAR	Peak-to-Average Ratio
OFDM	Orthogonal Frequency Division Multiplexing
OSIC	Ordered Successive Interference cancelation
PIC	Parallel Interference Cancellation
TDMA	Time Division multiple Access
QPSK	Quadrature Phase Shift Keying

SC	Single Carrier
SDMA	Space-Division-Multiple-Access
SIC	Successive Interference Cancellation
SISO	Single Input Single Output
SNR	Signal-To-Noise Ratio
STBC	Space-Time Block Code
STTC	Space-Time Trellis Code
TDD	Time Division Duplex
V-BLAST	Vertical-Blast
VA	Viterbi Algorithm
VLST	Very Large Integrated Circuit
ZF	Zero Forcing

Notation

- a is a scalar.
- $\mathbf{a} = [a_1, a_2, \dots, a_n]$ is a column vector.
- $\mathbf{A} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ is a matrix.
- $(\cdot)^*$ represents the complex conjugate.
- $(\cdot)^{-*}$ denotes the $((\cdot)^-)^*$.
- $(\cdot)^T$ represents transpose.
- $\mathbf{a}_i$ is the i -th column of matrix $\mathbf{A}$.
- A_{ij} is the ij -th element of matrix $\mathbf{A}$.
- $\mathbf{I}_n$ is an $n \times n$ identity matrix.
- $a(n)$ is a as a function of n .
- $\det(\mathbf{A})$ is the determinant of matrix $\mathbf{A}$.
- $\text{tr}(\mathbf{A})$ is the trace of matrix $\mathbf{A}$.
- $\text{rank}(\mathbf{A})$ is the rank of matrix $\mathbf{A}$.
- $\mathbf{A}^\dagger$ is the Moore-Penrose Pseudo-inverse of matrix $\mathbf{A}$.
- $|a| = \sqrt{(\text{Re}(a))^2 + (\text{Im}(a))^2}$ is the absolute value of the complex scalar a .
- $\|\mathbf{a}\| = \sqrt{|a_1|^2 + \dots + |a_n|^2}$ is the L_2 -norm of the vector $\mathbf{a}$.

- $Re(a)$ is real part of a .
- $Im(a)$ is imaginary part of a .
- Lower case denotes time domain.
- upper case denotes frequency domain.
- $vec(.)$ operator is defined as $vec([v_1, \dots, v_k]) = [v_1^T, \dots, v_k^T]^T$.
- $\mathbf{e}_i$ denotes the i -th unit vector.
- $\otimes$ denotes Kronecker product.
- $\mathcal{E}(a)$ or $E(a)$ is the expected value of a .
- $Q(\cdot) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt$.

Chapter 1

Introduction

1.1 Multiple-Input-Multiple-Output Wireless Communications Systems

In the last 20 years, the rapid progress in wireless communications has triggered a communication revolution, making it possible for people to communicate reliably everywhere. Nowadays, high data rate wireless communication systems constitute a major research area both in academia and industry. The growth is driving technology to develop communication techniques which can offer high quality of service at high data rates, with large numbers of users, under difficult propagation conditions. Wireless communication channels feature multipath propagation, a time varying environment and limited frequency spectrum. Therefore, the major challenges in wireless communications research are to develop techniques to deal with multipath propagation and limited frequency spectrum efficiently.

Multiple-Input-Multiple-Output (MIMO) systems and space-time processing techniques have been introduced to overcome some of these challenges. A recent information theory result reported in a pioneering work by Foschini and Gans [1] and at the same time by Telatar [2] showed impressive spectral efficiency can be achieved

through the use of MIMO systems. In particular, for a Rayleigh flat fading environment, comparable to a single-input single-output (SISO) system which by Shannon's classical formula capacity scales as one more bit for every 3dB increase of signal-to-noise ratio, capacity of an MIMO system with n transmit and receive antennas scales n more bits for each 3dB increase in signal-to-noise ratio (SNR), provided that the propagation paths between all transmit and receive antennas are statistically independent and known to the receiver. Space-time processing and space-time coding which combine signal processing techniques both in time and space have been introduced to improve wireless systems performance. Such space-time processors can improve network capacity, coverage, and quality by reducing co-channel interference (CCI) while enhancing diversity and array gain [3].

Future-generation wireless systems are required to support very high data rates. However, high data rates wireless communication systems, whether using space-time processing or not face inter-symbol-interference (ISI), which is caused by frequency selectivity of wireless channel. The maximum data rate is limited by ISI. For such high data rates we can have severe time-dispersion effects associated with multipath propagation. In this case, conventional time domain equalization schemes are not practical since signal processing requirements can be very high. This can be more serious when time-domain equalization methods are employed with multiple antennas since using multiple antennas make the system more complex.

The block transmission techniques, with appropriate cyclic prefix and frequency-domain equalization schemes, were shown to be a better alternative for severe time-dispersive channels, allowing superior performance with lower implementation complexity. the use of orthogonal frequency-division multiplexing (OFDM) modulation is

the most popular technique. Single-Carrier (SC) modulations using frequency-domain equalization (FDE) are an attractive alternative based on this principle. This SC approach with FDE has the same overall envelope fluctuations as the corresponding OFDM schemes, while offering similar, or even better performance.

1.2 Thesis Objective

The focus of this thesis is to develop frequency domain processing techniques for wireless communications systems with multiple transmit/receive antennas. The frequency domain techniques are employed to design low complexity transmitters and receivers which can mitigate ISI and co-channel interferences in severe time-dispersive MIMO channels.

In particular, the approach is to introduce new MIMO transmitter and receiver structures also known as pre-equalizers/equalizers and investigate the performance of these transceivers in different scenarios. While the results herein can readily be applied to any frequency selective channel, the emphasis is on severely-time dispersive MIMO channels. The proposed systems are a suitable candidate for broadband wireless communication systems.

1.3 Thesis Overview

Chapter 2 provides the current literature review on multiple antennas systems. This chapter discusses multiple-antenna systems including information-theoretic aspects and a signal processing point of view. The multiple-antenna architecture, detection algorithm and space-time coding are also discussed.

Chapter 3 discusses existing channel equalization methods for MIMO systems.

Different linear and non-linear wide-band receivers are introduced in this chapter. Finite-length, infinite-length, MMSE and zero-forcing designs are considered. This chapter discusses traditional time-domain equalization methods and is an introduction for the next chapters which develop new frequency-domain approaches.

Chapter 4 briefly discusses frequency-domain equalization methods for SISO systems. Two new hybrid time-frequency receiver structures for MIMO channels are developed. The performance of the proposed receivers in different scenarios assuming perfect channel estimation is evaluated.

Chapter 5 presents a new iterative frequency domain decision-feedback equalizer (DFE) for MIMO channels. This iterative receiver employs the successive interference cancellation principle and possesses excellent performance.

Chapter 6 addresses the problem of joint transmitter-receiver design for MIMO channels when Channel State Information (CSI) is available at the transmitter. This chapter presents a novel analysis for joint MIMO transmitter and receiver design with the inclusion of aliasing and without any restriction on bandwidth.

Chapter 7 introduces frequency domain space-time linear precoders when only the second order statistics of the channel is available at the transmitter. These precoders are designed for spatially correlated severe-time dispersive channels. The proposed designs have low complexity in severe time-dispersive channels.

Chapter 8 develops a power optimization method for the Vertical Bell labs layered space-time (V-BLAST) systems employing two antennas at the transmitter. Error propagation is considered both in analysis and simulation. In this approach CSI is only available at the receiver. The transmit power for each substream determined at the receiver is transferred to the transmitter.

Chapter 9 includes conclusions and suggestions for future research.

1.4 Thesis Contributions

- Two new low complexity non-linear DFE structures for SC-MIMO systems are developed.
- A new iterative frequency domain receiver for MIMO channels is introduced.
- A novel analysis for joint transmitter-receiver optimization for MIMO channels with the inclusion of aliases is developed.
- A new class of space-time frequency domain precoders for spatially correlated channels is introduced.
- A new power optimization algorithm for V-BLAST systems considering error propagation with relatively low overhead feedback is derived.

Chapter 2

Multiple Input Multiple Output Systems for Non-Frequency Selective Channels

The last two decades have seen tremendous growth in wireless communications. Especially in recent years the introduction of multiple-antenna techniques has become very popular which promises extremely high spectral efficiency for wireless communications systems.

There are two main issues in multiple-antenna systems

- Information-theoretic aspects of multiple antennas systems with emphasis on channel capacity.
- Signal processing issues with emphasis on receiver design and transmission techniques to realize at least part of the significant theoretical spectral efficiency gains promised by information theory.

In this chapter we first review the information theoretic aspects of MIMO wireless communication systems. We briefly introduce the definitions of outage and ergodic

capacity. Then we review MIMO System from a signal processing point of view. In this context we present the BLAST architectures which were proposed to attain a substantial portion of the promised theoretical spectral efficiencies. First, we present Diagonal BLAST (D-BLAST) and it is shown that D-BLAST architecture can asymptotically achieve Shannon's capacity bound. Next, a simplified version of layered space-time architecture called V-BLAST is introduced. Finally we review the space-time coding techniques. Blind Space-time coding techniques without CSI at the transmitter are presented in this chapter. Nonblind space-time coding with CSI at the transmitter are presented in the following chapters.

2.1 Multiple Antenna Systems: Definition and Notations

We consider a discrete time complex baseband MIMO wireless communications systems with P antennas at the transmitter and N antennas at the receiver, the system is denoted by descriptor (P, N) . Figure 2.1 shows the block diagram of an (P, N) system, using the following notation:

- $\mathbf{a}(m) = [a_1(m), a_2(m), \dots, a_P(m)]^T$ is a P dimensional vector of transmitted complex symbols selected from a given constellation (e.g., a QPSK constellation) at discrete time m .

The total transmitted power is constrained to P_0 regardless of the value of P . The bandwidth is narrow enough that we can model the channel as flat over frequency.

- $\mathbf{H}(m)$ is a $P \times N$ normalized channel matrix. Normalization is done such that

each element of $\mathbf{H}$ has unit average power and it is drawn from an independent and identically distributed (i.i.d) complex Gaussian random variable with zero mean and unit variance. The channel is assumed to be Rayleigh flat fading over the bands of interest.

- $\mathbf{n}(m) = [n_1(m), n_2(m), \dots, n_N(m)]$ is an N -dimensional complex additive white Gaussian noise with statistically independent components of identical power σ^2 at each of N receive antenna branches.
- $\mathbf{r}(m) = [r_1(m), r_2(m), \dots, r_P(m)]$ is an N -dimensional received signal vector and the average SNR at each receive antenna input is: $\rho = P_0/\sigma^2$. The received signal vector is:

$$\mathbf{r}(m) = \sqrt{\frac{P_0}{P}} \mathbf{H} \mathbf{a}(m) + \mathbf{n}(m) \quad (2.1.1)$$

2.2 MIMO Channel Capacity

The expression for channel capacity of MIMO system was derived by Foschini [1] and Telatar [2]. Telatar derived the Ergodic capacity and the error exponent bounds while Foschini derived the outage capacity.

2.2.1 Ergodic Capacity

The maximum mutual information has the meaning of channel capacity when each use of the channel employs an independent realization of $\mathbf{H} \in \mathcal{C}^{P \times N}$ or when the process that generates $\mathbf{H}$ is ergodic [2]. The realization of $\mathbf{H}$ is assumed to be known

Figure 2.1: Block diagram of a (P, N) MIMO systems

at the receiver but not at the transmitter. The channel capacity in bits per channel use is defined as the maximum mutual information $\mathcal{I}$ between input and output. Since the receiver knows the realization of $\mathbf{H}$, the channel output is the pair $(\mathbf{r}, \mathbf{H}) = (\mathbf{H}\mathbf{a} + \mathbf{n}, \mathbf{H})$, the mutual information between input and output is then [2]:

$$\begin{aligned}
 C &= \max_{\mathbf{Q} > 0, \text{tr}(\mathbf{Q}) \leq P_0} \mathcal{E}_{\mathbf{H}}[\mathcal{I}(\mathbf{a}; (\mathbf{r}, \mathbf{H}))] \\
 &= \max_{\mathbf{Q} > 0, \text{tr}(\mathbf{Q}) \leq P_0} \mathcal{E}_{\mathbf{H}}[\mathcal{I}(\mathbf{a}; \mathbf{H}) + \mathcal{I}(\mathbf{a}; \mathbf{r}|\mathbf{H})] \\
 &= \max_{\mathbf{Q} > 0, \text{tr}(\mathbf{Q}) \leq P_0} \mathcal{E}_{\mathbf{H}}[\mathcal{I}(\mathbf{a}; \mathbf{r}|\mathbf{H})]; \tag{2.2.1}
 \end{aligned}$$

where $\mathbf{Q} = \mathcal{E}[\mathbf{a}(m)\mathbf{a}^*(m)]$. It is known that if $\mathbf{a}$ is constrained to have covariance $\mathbf{Q}$, the optimal distribution for $\mathbf{a}$ is Gaussian [2]. It was shown in [2] that the ergodic capacity formula of this channel is:

$$C = \mathcal{E}_{\mathbf{H}} \left[\log_2 \det \left(\mathbf{I}_N + \frac{\rho}{P} \mathbf{H} \mathbf{H}^* \right) \right] \quad (2.2.2)$$

with $\mathbf{Q} = \frac{P_0}{P} \mathbf{I}_P$.

For fixed N , when $P \rightarrow \infty$ due to central limit theorem $\frac{1}{P} \mathbf{H} \mathbf{H}^* \rightarrow \mathbf{I}_N$ (every element of $\mathbf{H}$ is a complex Gaussian random variable). Thus the capacity in the limit is

$$C = N \log_2(1 + \rho) \quad (2.2.3)$$

2.2.2 Outage Capacity

We remarked in the previous section that the maximum mutual information has the meaning of Shannon's capacity when the channel is memoryless and when each use of the channel employs an independent realization of $\mathbf{H}$ or if the channel is ergodic. In contrast, if the channel $\mathbf{H}$ is chosen randomly at the beginning of a burst and held constant during a burst (quasi-static), the maximum mutual information in general is not equal to the channel capacity. However small the rate, there is a non-zero probability that the realization of $\mathbf{H}$ is incapable of supporting it no matter how long the code length is. In this case the Shannon capacity of the channel is zero. This leads to a new definition of capacity based on outage capacity [2].

Definition 2.2.1. For a given rate R and power P_0 , there exists an outage probability $P_{out}(R, P_0)$ such that for a rate $R_1 < R$ and any $\delta > 0$, there exists a code, for which the error probability $P_r(E) \leq \delta$ for all but a set of quasi-static fading $\mathbf{H}$ whose probability is less than P_{out} .

Foschini and Gans derived the formula for outage capacity of quasi-static MIMO channels. They investigated the case of independent Rayleigh faded path between antenna elements. They found out that the Shannon channel capacity of the above MIMO channel can be increased to n times that of the SISO system, when n is the number of transmit and receive antennas, assumed to be the same. The capacity of MIMO system increases n bits per transmission for every 3dB of SNR increase for a large but practical number n and for large SNRs, whereas the Shannon's capacity of SISO increases 1 bit per transmission for every 3dB [1]. The capacities of Rayleigh fading channels can be expressed using chi-squared variates. Next we present an example of the channel capacities for a given channel realization and with maximum-ratio-combining of the following MIMO systems [1]:

No diversity, $P = N = 1$

$$C = \log_2[1 + \rho\chi_2^2] \quad (2.2.4)$$

where χ_2^2 is a chi-squared random variable with two degrees of freedom ¹.

Receive diversity, $P = 1, N = n$

$$C = \log_2[1 + \rho\chi_{2n}^2] \quad (2.2.5)$$

where χ_{2n}^2 is a chi-squared random variable with $2n$ degrees of freedom. The receive diversity is obtained by maximum-ratio combining [4]. In general, application of receive diversity is limited to base station deployment in wireless communication systems.

Transmit diversity, $P = n, N = 1$

$$C = \log_2[1 + \frac{\rho}{n}\chi_{2n}^2] \quad (2.2.6)$$

¹It represents the received SNR at the receiver.

A simple form of transmit diversity was introduced by Alamouti [5]. The capacity increase of receive diversity with n antennas at the receiver is greater than the capacity increase of transmit diversity with n antennas at the transmitter [1]. Space-time coding techniques are needed to achieve full benefit of diversity [6, 7].

Combined transmit and receive diversity (P, N) $P, N > 1$,

The capacity bounds for flat Rayleigh fading channels with multiple antennas at the transmitter and multiple antennas at the receiver was derived in² [1]

$$\sum_{i=P-(N-1)}^P \log_2 \left(1 + \frac{\rho}{P} \chi_{2i}^2 \right) < C \leq \sum_{i=1}^P \log_2 \left(1 + \frac{\rho}{P} \chi_{2N,i}^2 \right) \quad (2.2.7)$$

where χ_{2i} are independent chi-squared random variable with $2i$ degrees of freedom. $\chi_{2N,i}$ represents chi-squared random variables with $2N$ degrees of freedom where i is used to index statistically independent variables. The upper bound represents an artificial case where each transmitted signal is received by a separate set of N receive antennas where each signal component is received without interference from the other signals. In other words when P transmitted signals are transmitted over an uncoupled channel and each one has a receive diversity order of N . For $P = N$ the lower bound on the capacity is

$$C_L = \sum_{i=1}^P \log_2 \left[1 + \frac{\rho}{P} \chi_{2i}^2 \right] \quad (2.2.8)$$

Since χ_{2i}^2 represents a fading channel with diversity order of i , this lower bound is the sum of the P capacities of independent channels with increasing diversity order from 1 to P . Foschini proposed a layered space-time architecture which achieves the lower bound in (2.2.8), assuming a zero-forcing receiver for interference nulling and cancellation of already detected layers. Ariyavisitakul first showed that Foschini's

²The inequality means almost sure.

Figure 2.2: Outage Capacity of (N, N) System with $N = 1, 2, 4, 8, 16$

lower bound is equal to capacity when the output SNR of each layer is represented by the corresponding matched filter bound [8].

Figure 2.2 depicts the outage channel capacity for a system with an equal number of transmit and receive antennas N and $\rho = 18dB$. This figure shows the significant capacity improvement by doubling the number of transmit and receive antennas.

2.3 BLAST Architecture

Bell labs layered space-time (BLAST) is an architecture which uses multiple antennas at the transmitter and receiver. In BLAST information-bearing data is divided into multiple substreams and each substream is transmitted by different antennas simultaneously using the same frequency bandwidth. At the receiver, signals are received by

an array of antennas. Each antenna receives all transmitted signals from all antennas. In BLAST the signals are transmitted in the same band, therefore each signal is scattered through different paths. At the receiver, they still contain different information, which helps to de-correlate the transmitted signals. However, simultaneous use of the same frequency bands leads to co-channel interference which is the main impairment in BLAST. Foschini raised the following fundamental question and then proposed a layered space-time architecture to answer the question [9]:

Can one construct an (n, n) MIMO system whose capacity scales linearly with the number of transmit antennas, using building blocks of n separately coded one-dimensional subsystems of equal capacity?

The motivation for proposing BLAST is that it is an architecture for processing 2-D signals in multiple antenna systems using already developed 1-D codec (encoding across time only) technology. Using this architecture we can achieve channel capacity [9].

2.3.1 Diagonal BLAST

The structure of diagonally-layered space-time (D-BLAST) architecture is constrained to use spatially 1-D codecs as building blocks. This constraint was motivated by large number of antennas P where it is important to avoid an explosion of processing complexity in spatial domain. The aim of D-BLAST is to do so by leveraging the already developed 1-D codec technology in the last half-century. D-BLAST achieves the lower bound on the channel capacity of MIMO system in (2.2.7). It is shown in [8, 10] that the lower bound in (2.2.7) is equal to the channel capacity when the output SNR of each layer is replaced by the corresponding matched filter bound. The

Figure 2.3: Diagonal layering

novel feature of D-BLAST is the space-time encoding structure by using P 1-D coded substream of equal capacity across diagonal of the space-time system (see Figure 2.3), which permits the decoding complexity to grow linearly with the number of transmit antennas. However, with the diagonal layering, the space-time is wasted at the start and end of a burst. For practical burst length of few hundreds symbols, this wasted space-time is significant.

The block diagram of a D-BLAST is shown in Figure 2.4. A data stream is de-multiplexed into P substreams of equal rate. Each data substream is encoded independently. Instead of assigning each of P encoded substreams to an antenna, the encoded block associated with each antennas is periodically cycled [10]. The diagonal layered codes are illustrated in Figure 2.4. Each diagonal constitute one coded block. One may note the wasted space-time shown in this figure. In D-BLAST, due to the periodic cycling of the substreams all p subchannels have the same capacity, i.e.

Figure 2.4: Block diagram of a (4, 4) D-BLAST Architecture

$C_{Di} = C_D/P$. The capacity would then be obtained by the sum of the capacities of all layers. Thus the capacity of D-BLAST (for a fixed realization of the channel) reduces to:

$$C_D = \sum_{p=1}^P \log_2[1 + SNR_p]; \quad (2.3.1)$$

where SNR_p is the signal-to-noise ratio of the p -th substream.

2.3.2 Successive Interference Cancellation

The diagonal decoding and encoding makes D-BLAST receiver a very complex algorithm for real implementation. It uses a successive interference cancellation algorithm which nulls out the interference from yet-to-be detected signals and then cancels out the interference from already detected signals. We consider the detection of layer B on the right side of Figure 2.3. Detection of this layer is comprised of 4 parts. All layers which are located underneath this layer are assumed to be already detected

and their interference has been cancelled out while all layers located above B are yet to be detected. The processing of this layer at the receiver has 3 stages. First a preprocessing step cancels out the interference from the yet to be detected layers. After the entire layer is enhanced, it is decoded. Finally, after decoding the detected layer is reconstituted to subtract this layer as source of interference to the remaining substreams. Each of the P layers is processed in turn until the last substream is decoded.

2.4 Vertical-BLAST (V-BLAST)

The diagonal decoding and encoding makes D-BLAST architecture a very complex algorithm for real implementation. This motivation leads to introduce V-BLAST. V-BLAST is an uncoded structure with optimally ordered successive cancellation. Real time implementation of V-BLAST has demonstrated spectral efficiencies of 20-40 b/s/Hz in an indoor propagation environment at average SNRs ranging from 24-34 dB [11, 12].

A block diagram of V-BLAST is shown in Figure 2.5. A single data stream is demultiplexed into P substreams and each substream is then independently encoded into symbols and transmitted by different antennas. At the receiver end the transmitted vector is collected by an array of N antennas.

The major difference between D-BLAST and V-BLAST is in the encoding scheme. In D-BLAST the diagonal layering of coded substreams introduces redundancy between substreams. This diagonal layering leads to higher spectral efficiencies for D-BLAST than V-BLAST with the same number of antennas at the transmitter and

Figure 2.5: V-BLAST block diagram

receiver. However, the vector encoding process in V-BLAST is simply a demultiplexing operation followed by independent bit-to-symbol mapping of each substream. No intersubstream coding is required in V-BLAST, although a conventional 1-D channel coding of any kind may be applied in each layer.

2.4.1 Ordered Successive Interference Cancellation Algorithm

In the original design of V-BLAST an ordered successive detection interference cancellation (OSIC) was proposed at the receiver end. Many authors have proposed improved algorithms for V-BLAST [13–17]. In the following we briefly describe the OSIC algorithm which was proposed in [11]. OSIC is a nonlinear algorithm which uses symbol cancellation as well as linear nulling. When symbol cancellation is used, the order in which the components of the received vector are detected affects the overall performance of the system.

Let the detection order

$$\mathcal{S} = \{k_1, k_2, \dots, k_P\} \quad (2.4.1)$$

be a permutation of the integers $1, 2, \dots, P$. The detection algorithm proceeds three steps as follows:

1. Using the nulling vector $\mathbf{w}_{k_i}$, $i = 1, \dots, P$ form the soft decision statistics

$$y_{k_i} = \mathbf{w}_{k_i}^* \mathbf{r}_i \quad (2.4.2)$$

The weight vectors $\mathbf{w}_{k_i}$'s are obtained by using zero-forcing (ZF) or minimum mean-squared error (MMSE) criteria.

2. Obtain the hard decision estimate $\hat{a}_{k_i}$

$$\hat{a}_{k_i} = Q(y_{k_i}) \quad (2.4.3)$$

3. Cancel the signal associated with the detected substream a_{k_i} from the received vector $\mathbf{r}_i$

$$\mathbf{r}_{k_{i+1}} = \mathbf{r}_{k_i} - \hat{a}_{k_i} \mathbf{h}_{k_i} \quad (2.4.4)$$

ZF solution:

It can be shown that ZF nulling vectors are given by

$$\mathbf{w} = \mathbf{H}^\dagger \quad (2.4.5)$$

where $\mathbf{H}^\dagger = (\mathbf{H}^* \mathbf{H})^{-1} \mathbf{H}^*$ is pseudo-inverse [18] and $\mathbf{w} = [w_1, w_2, \dots, w_P]$. Note that ZF solution has a noise enhancement problem.

MMSE solution:

The corresponding MMSE weight vectors are

$$\mathbf{w} = (\mathbf{H}^* \mathbf{H} + \sigma^2 \mathbf{I}_P)^{-1} \mathbf{H}^H \quad (2.4.6)$$

Optimal ordering

As mentioned earlier, when symbol cancellation is used the overall performance of the system is affected by the detection ordering. The post detection SNR for the k_i -th detected substream is given by

$$\rho_{k_i} = \frac{\mathcal{E}[|a_{k_i}|^2]}{\sigma^2 \|\mathbf{w}_{k_i}\|^2} \quad (2.4.7)$$

where the expectation in the numerator is taken over the constellation set. It is assumed that all components of $\mathbf{a}$ have the same constellation. Under this assumption the substream with smallest SNR will dominate the overall performance of the system. Thus, an obvious criterion is the maximization of the worst SNR. It was proved in [11] that simply choosing the best ρ_k at each stage in the detection process leads to the optimum solution.

The ZF V-BLAST detection algorithm with optimal ordering can be described as:

- Initialization

$$i = 1 \quad (2.4.8)$$

$$\mathbf{G}_1 = \mathbf{H}^\dagger \quad (2.4.9)$$

- Recursion

$$k_i = \arg \min_{j \notin k_1, k_2, \dots, k_{i-1}} \|(\mathbf{G}_i)_j\|^2 \quad (2.4.10)$$

$$\mathbf{w}_{k_i} = ((\mathbf{G}_i)_{k_i})^T \quad (2.4.11)$$

$$y_{k_i} = \mathbf{w}_{k_i}^* \mathbf{r}_i \quad (2.4.12)$$

$$\hat{a}_{k_i} = Q(y_{k_i}) \quad (2.4.13)$$

$$\mathbf{r}_{i+1} = \mathbf{r}_i - \hat{a}_{k_i} \mathbf{h}_{k_i} \quad (2.4.14)$$

$$\mathbf{G}_{i+1} = \mathbf{H}_{k_i}^\dagger \quad (2.4.15)$$

$$i = i + 1 \quad (2.4.16)$$

where $(\mathbf{G}_i)_{k_i}$ is the k_i -th row of $\mathbf{G}_i$, $\mathbf{H}_{k_i}$ denotes the matrix obtained by zeroing columns $k_1, k_2, \dots, k_i$ of $\mathbf{H}$. The MMSE V-BLAST algorithm is the same but at each stage the substream with minimum MSE is detected.

2.5 Combined ML and OSIC detection for V-BLAST

V-BLAST suffers from error propagation. Due to the error propagation, the overall performance of V-BLAST is determined by the first subchannel to be detected. In [14] the authors proposed to combine Maximum-Likelihood (ML) detection with OSIC to reduce the error propagation effect. Roughly speaking, the worst p channels are detected using ML algorithm, and then the rest of the subchannels are detected using OSIC. It was shown in [14] that the detection of p subchannels simultaneously attains a diversity order of p .

1. Substreams are ordered according to the post detection SNR. Let the set $\{k_1, k_2, \dots, k_P\}$ denote the ordered detection set.

2. The channel matrix $\mathbf{H}$ is decomposed as

$$\mathbf{H} = \mathbf{Q} \begin{bmatrix} \mathbf{R} & \mathbf{H}_a \\ \mathbf{0} & \mathbf{H}_b \end{bmatrix} \quad (2.5.1)$$

where $\mathbf{R}$ is an upper triangular matrix of size $(P - p) \times (P - p)$, $\mathbf{Q}$ forms an orthonormal matrix and $\mathbf{H}_b$ is a square matrix of size $p \times p$.

3. Using this decomposition first $x_{k_{P-p+1:P}}$ are detected jointly by ML decoding of size p , then the interference caused by these symbols are cancelled and then $x_{k_{1:P}}$ are detected by OSIC.

It was shown in [14] that the block error probability P_b is bounded by

$$P_b \leq \left[\frac{M}{1 + d_{min}^2/4\sigma^2} \right]^p \quad (2.5.2)$$

where M is the constellation size and σ^2 is the noise variance. From (2.5.2) it is observed that ML-OSIC algorithm increases the diversity order to p .

2.6 Adaptive Group Detection Algorithm

To reduce the complexity of ML detection for BLAST, a channel-based adaptive group detection (AGD) algorithm was proposed in [16]. In this algorithm the ZF/MMSE interference suppression is replaced with the AGD algorithm. The AGD algorithm partitions the substream into small groups. Linear subspace projection is used to null/suppress the interference outside the groups and local ML search is performed within the group. By partitioning to small groups the complexity of the ML search is limited. To further improve the performance, adaptive grouping which utilizes the channel information can be employed.

The AGD algorithm contains three steps:

1. Several substreams are grouped together. The group size may vary from group to group. Furthermore the groups can overlap.
2. The second step of group detection is the interference suppression or subspace projection. For simplicity, we assume the groups do not overlap and the transmit antennas have been ordered so that the first $\mathcal{K}_1$ antennas are in $\mathcal{G}_1$, the following $\mathcal{K}_2$ antennas are in $\mathcal{G}_2$, and so on. The main objective of the subspace projection is to find a projection matrix P such that

$$\mathbf{P}\mathbf{H} = \mathbf{D} \quad (2.6.1)$$

where

$$\mathbf{D} = \text{diag}\{\mathbf{D}_1, \mathbf{D}_2, \dots\} \quad (2.6.2)$$

is a block diagonal matrix with squared subblocks $D_1 (\mathcal{K}_1 \times \mathcal{K}_1)$, $D_2 (\mathcal{K}_2 \times \mathcal{K}_2)$, and so on. A similar development can also be made for MMSE. Generally the goal of subspace projection is to null/suppress interference outside the group.

3. ML detection is performed within each group.

These steps are repeated until all layers have been processed. The complexity of ML detection over all P antennas requires M^P searches, where M is the constellation size. The computational complexity is beyond the limit of most systems today. However the group detection is carried out within a group of $N_g \ll P$ antennas, therefore the complexity is proportional to M^{N_g} , and can be made low by using a small group size. The AGD algorithm can be combined with OSIC. Specifically, layers are grouped to be processed together and ML detection is performed within each group to get the improved decisions. After that, the detected decisions or part of them are subtracted

from the received signal and another round of AGD detection would continue until all the layers have been processed.

2.7 Efficient Algorithm for BLAST

The main computational complexity of BLAST algorithm is the optimal ordering for sequential detection of transmitted substreams, as well as the computation of the nulling vectors. This algorithm has the computational complexity $\mathcal{O}(P^4)$ where P is the number of transmit antennas [19].

2.7.1 Square Root Algorithm

To reduce the computational complexity of BLAST an efficient square root algorithm was developed in [19]. This algorithm reduces the computational cost by an order of magnitude, effectively from $\mathcal{O}(P^4)$ to $\mathcal{O}(P^3)$. The MMSE solution (2.4.6) can be expressed as

$$\hat{\mathbf{a}} = (\mathbf{H}^H \mathbf{H} + \sigma^2 \mathbf{I}_P)^{-1} \mathbf{H}^H \mathbf{r} = \begin{bmatrix} \mathbf{H} \\ \sigma \mathbf{I}_P \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{r} \\ \mathbf{0} \end{bmatrix} \quad (2.7.1)$$

The covariance matrix of the estimation error $\mathbf{a} - \hat{\mathbf{a}}$ can be easily seen to be

$$\mathbf{P} = E((\mathbf{a} - \hat{\mathbf{a}})(\mathbf{a} - \hat{\mathbf{a}})^*) = \begin{bmatrix} \mathbf{H} \\ \sigma \mathbf{I}_P \end{bmatrix}^\dagger \left(\begin{bmatrix} \mathbf{H} \\ \sigma \mathbf{I}_P \end{bmatrix}^\dagger \right)^* \quad (2.7.2)$$

In order to avoid squaring and inverting, let us define the **QR** decomposition:

$$\begin{bmatrix} \mathbf{H} \\ \sigma \mathbf{I}_P \end{bmatrix} = \mathbf{Q} \mathbf{R} = \begin{bmatrix} \mathbf{Q}_\sigma \\ \mathbf{Q}_2 \end{bmatrix} \mathbf{R} \quad (2.7.3)$$

where $\mathbf{Q}$ is an $(P + N) \times P$ orthonormal matrix and $\mathbf{R}$ is $M \times M$ nonsingular matrix.

It is not hard to see that

$$\mathbf{P}^{1/2} = \mathbf{R}^{-1} \text{ and } \mathbf{H}_\sigma^\dagger = \mathbf{P}^{1/2} \mathbf{Q}_\sigma^* \quad (2.7.4)$$

where

$$\mathbf{H}_\sigma = \begin{bmatrix} \mathbf{H} \\ \sigma \mathbf{I}_P \end{bmatrix} \quad (2.7.5)$$

and $\mathbf{P}^{1/2} \mathbf{P}^{*1/2} = \mathbf{P}$. Thus, given $\mathbf{P}^{1/2}$ and $\mathbf{Q}_\sigma$ both the pseudoinverse and the error covariance matrix can be computed. The computation of $\mathbf{P}^{1/2}$ and $\mathbf{Q}_\sigma$ via the $\mathbf{QR}$ decomposition is unattractive since it requires inversion of $\mathbf{R}$ to obtain $\mathbf{P}^{1/2}$. To prevent matrix inversions, $\mathbf{P}^{1/2}$ and $\mathbf{Q}_\sigma$ can be computed using the following recursion, initialized with $\mathbf{P}_{|0}^{1/2} = \frac{1}{\sigma} \mathbf{I}_P$ and $\mathbf{Q}_0 = \mathbf{0}_{P \times N}$ which is an extension of the traditional square-root algorithm that was introduced in [19].

$$\begin{bmatrix} 1 & \mathbf{H}_i \mathbf{P}_{|i-1}^{1/2} \\ 0 & \mathbf{H}_i^{1/2} \\ -e_i & \mathbf{Q}_{i-1} \end{bmatrix} \boldsymbol{\Theta}_i = \begin{bmatrix} r_{e,i}^{1/2} & 0 \\ K_{p,i} & \mathbf{P}_{|i}^{1/2} \\ A_i & \mathbf{Q}_i \end{bmatrix} \quad (2.7.6)$$

$\boldsymbol{\Theta}_i$ is any unitary transformation that transforms the first row of the pre-array to lie along the direction of the first row vector. After N steps the algorithm yields the desired quantities i.e.,

$$\mathbf{P}^{1/2} = \mathbf{P}_{|N}^{1/2} \text{ and } \mathbf{Q}_\sigma = \mathbf{Q}_N \quad (2.7.7)$$

Description of Algorithm

The root-square algorithm can be summarized as

1. Compute $\mathbf{P}^{1/2}$ and $\mathbf{Q}_\sigma$ by the recursive algorithm which was described above.
2. Find the minimum length of $\mathbf{P}^{1/2}$ and permute it to the last P row.

3. Find a unitary transform Σ such that $\mathbf{P}^{1/2}\Sigma$ is block upper triangular which rotates the P -th row of $\mathbf{P}^{1/2}$ to lie along the direction of M -th unit vector:

$$\mathbf{P}^{1/2}\Sigma = \begin{bmatrix} \mathbf{P}^{(M-1)/2} & \mathbf{P}_M^{(M-1)/2} \\ \mathbf{0} & P_M^{1/2} \end{bmatrix} \quad (2.7.8)$$

4. Update $\mathbf{Q}$ to $\mathbf{Q}\Sigma$.
5. The nulling vector for the M -th substream is given by $p_M^{1/2} \mathbf{q}_m^*$ where $\mathbf{q}_M$ is the M -th row of $\mathbf{Q}$.
6. Go back to step 3, but now with $\mathbf{P}^{(M-1)/2}$ and $\mathbf{Q}^{(M-1)/2}$ (the first $M-1$ column of $\mathbf{Q}$).

It can be shown that the computational complexity of this algorithm is

$$\frac{2}{3}P^3 + 7NP^2 + 2N^2P \quad (2.7.9)$$

while the computational complexity of the original algorithm is

$$\frac{15}{4}P^4 + 2NP^3 + N^2P^2 \quad (2.7.10)$$

When $P = N$, the computational complexity is thus reduced from $27M^4/4$ to $29M^3/3$, roughly by a factor of $0.7M$

Another efficient algorithm with low computational complexity for V-BLAST was developed in [20]. The proposed algorithm avoids directly inverting a matrix and finds the MMSE filter coefficients via induction with help of the inverse block partitioned matrix and the Sherman-Morrison formula. The proposed algorithm requires $\mathcal{O}(P^3 + P^2N)$ operations, which is a complexity reduction by a factor of P . To be exact, when $P = N$, the proposed algorithm in [20] reduces the number of operations by

a factor of $P + 2$. It was shown that for large number of antennas this algorithm is more efficient than the Hassibi's root-square algorithm [19]. Also, in [21] the authors presented an efficient algorithm for V-BLAST systems which exploits a symmetry property in the triangularization process and requires $O(\frac{1}{2}M^3 + 2M^2N)$ operations.

2.8 Space-Time Codes

Tarokh et al. were the first to introduce the space-time codes [6]. In the last few years design of the space-time codes for improving the data rate and the reliability of communication systems have attracted many researchers. In general, space-time codes are a combination of coding, modulation, transmit and receive diversity to provide high performance. Space-time codes fall into two general categories.

- Blind schemes which do not require the CSI at the transmitter.
- Schemes with feedback which require CSI or partial CSI at the transmitter.

Both of these schemes require CSI at the receiver. Feedforward information is required from transmitter to the receiver to estimate the channel at the receiver. At the receiver the transmitted symbols are decoded by linear processing or ML decoding techniques. In this section, we briefly review the design criterion for space-time codes. Then the space-time trellis and block codes are introduced.

2.8.1 Design Criterion

We assume that the channel is quasi-static. For every input symbol s_l , a space-time encoder generates P code symbols $c_{l1}, c_{l2}, \dots, c_{lP}$. These P code symbols are

transmitted simultaneously from the P transmit antennas. We define the code vector $\mathbf{c}_l = [c_{l1}, c_{l2}, \dots, c_{lP}]$. The transmitted code matrix is

$$\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_L]. \quad (2.8.1)$$

We consider the probability that the decoder decides erroneously in favor of a code vector sequence

$$\tilde{\mathbf{C}} = [\tilde{\mathbf{c}}_1, \tilde{\mathbf{c}}_2, \dots, \tilde{\mathbf{c}}_L]. \quad (2.8.2)$$

Assuming perfect CSI, the probability of transmitting $\mathbf{C}$ and deciding in favor of $\tilde{\mathbf{C}}$ is upper bounded for a Rayleigh channel by [6]

$$P(\mathbf{C} \rightarrow \tilde{\mathbf{C}} | \mathbf{H}) \leq \exp(-d^2(\mathbf{C}, \tilde{\mathbf{C}})) \quad (2.8.3)$$

where $d^2(\mathbf{C}, \tilde{\mathbf{C}})$ defines the distance between $\mathbf{C}$ and $\tilde{\mathbf{C}}$:

$$d^2(\mathbf{C}, \tilde{\mathbf{C}}) = \frac{\rho}{4P} \sum_{i=1}^P \mathbf{h}_i \mathbf{A}(\mathbf{C}, \tilde{\mathbf{C}}) \mathbf{h}_i^* \quad (2.8.4)$$

where $\mathbf{h}_i$ is the i -th row of the channel matrix $\mathbf{H}$, and the squared error matrix $\mathbf{A}$ is defined as

$$\mathbf{A}(\mathbf{C}, \tilde{\mathbf{C}}) = \sum_{l=1}^L (\mathbf{c}_l - \tilde{\mathbf{c}}_l)(\mathbf{c}_l - \tilde{\mathbf{c}}_l)^* \quad (2.8.5)$$

The error matrix

$$\mathbf{B}(\mathbf{C}, \tilde{\mathbf{C}}) = \begin{bmatrix} c_1(1) - \tilde{c}_1(1) & c_1(2) - \tilde{c}_1(2) & \dots & c_1(L) - \tilde{c}_1(L) \\ c_2(1) - \tilde{c}_2(1) & c_2(2) - \tilde{c}_2(2) & \dots & c_2(L) - \tilde{c}_2(L) \\ \vdots & \vdots & \ddots & \vdots \\ c_P(1) - \tilde{c}_P(1) & c_P(2) - \tilde{c}_P(2) & \dots & c_P(L) - \tilde{c}_P(L) \end{bmatrix} \quad (2.8.6)$$

is clearly a square root of $\mathbf{A}(\mathbf{C}, \tilde{\mathbf{C}})$. Thus the eigenvalues of $\mathbf{A}(\mathbf{C}, \tilde{\mathbf{C}})$ are nonnegative real numbers. Let r denote the rank of matrix $\mathbf{A}$ and $\lambda_1, \lambda_2, \dots, \lambda_r$ are the nonzero

eigenvalues of A , then it is possible to show that by averaging over the Rayleigh fading variables, the pairwise error probability (PEP) (2.8.3) is given by

$$P(\mathbf{C} \rightarrow \tilde{\mathbf{C}}|\mathbf{H}) \leq (\prod_{i=1}^r \lambda_1 \lambda_2 \dots \lambda_r)^{-N} \left(\frac{\rho}{4P}\right)^{-rN} \quad (2.8.7)$$

Thus a diversity gain of Nr and a coding gain of $\lambda_1 \lambda_2 \dots \lambda_r$ is achieved. Moreover, it is easy to see that the ranks of $\mathbf{A}(\mathbf{C}, \tilde{\mathbf{C}})$, and $\mathbf{B}(\mathbf{C}, \tilde{\mathbf{C}})$ are equal.

- *The Rank Criterion:* In order to achieve the maximum diversity PN , the matrix $\mathbf{B}(\mathbf{C}, \tilde{\mathbf{C}})$ has to be of full rank for any pair of codeword vector sequences $\mathbf{C}$ and $\tilde{\mathbf{C}}$. If $\mathbf{B}(\mathbf{C}, \tilde{\mathbf{C}})$ has minimum rank r over the set of two tuples of distinct codewords, then a diversity of rN is achieved.
- *The Determinant Criterion:* If the diversity advantage of code is r , to obtain the maximum coding gain, the minimum of the r -th roots of the sum of determinants of all $r \times r$ principal cofactors of $\mathbf{A}$ taken over all pairs of distinct code vector sequences $\mathbf{C}$ and $\tilde{\mathbf{C}}$ must be maximized.

2.8.2 Space-Time Trellis Codes

The criteria derived in the previous section is used to design trellis codes for a wireless communication system that employs P transmit antennas and N receive antennas. The encoding of these trellis codes is obvious, with the exception that at the beginning and the end of each frame, the encoder is required to be in the zero state. At each time l , depending on the state of the encoder and the input symbol, a transition branch is chosen. If the label of this branch is $c_1(l), c_2(l), \dots, c_P(l)$, then transmit antenna p is used to send the symbols $c_p(l)$, $p = 1, \dots, P$. Figure 2.6 provides a

Figure 2.6: 4-PSK 4-State Space-Time Code with 2 Tx antennas

4-PSK, 4-state space-time trellis code (STC) for the transmission of 2b/s/Hz for two antennas. Each row in the matrix shown in Figure 5 provides the labelling of the 4-PSK constellation and the trellis diagram for this code. The edge label $a_1 a_2$ indicates that symbol a_1 is transmitted over the first antenna and the symbol a_2 is transmitted over the second antenna. The input bit stream to the ST encoder is divided into groups of 2 bits and each group is mapped into one of the four constellation points. Assuming, one receive antenna, this code has a diversity advantage of two and a bandwidth efficiency of 2 bits per channel use. Assuming perfect CSI at the receiver and given the received signal $\mathbf{r}(l)_{l=1}^L$, the ML decoder choose the code matrix $\tilde{\mathbf{C}}$ which minimizes the following space-time metric:

$$\tilde{\mathbf{C}} = \arg \min_{\tilde{\mathbf{C}}} \sum_{l=1}^L \left\| \mathbf{r}(l) - \sqrt{\frac{\rho}{P}} \tilde{\mathbf{c}}(l) \right\| \quad (2.8.8)$$

The space-time ML decoding is done using Viterbi algorithm, where the trellis path with smallest accumulated metric is chosen. For suboptimal decoding algorithm the

interested author may refer to [22].

2.8.3 Space-Time Block Codes

Space-time trellis codes perform extremely well in slow fading environments and come within 2-3 dB of the outage capacity [7]. The space-time codes presented in [6] provide the best possible tradeoff between constellation size, data rate, diversity gain. When the number of antennas is fixed, the decoding complexity of space-time trellis coding increases exponentially as a function the diversity level and transmission rate. In addressing the issue of decoding complexity, Alamouti discovered a remarkable space-time block coding scheme over two transmit antennas [5]. This scheme supports simple decoding complexity based on linear processing at the receiver. This code provides maximum diversity and half of the maximum possible transmission rate allowed by space-time codes. The very simple structure of the encoder and the linear decoding of Alamouti code has led to it be part of W-CDMA and CDMA-2000 standards. STBC technique, introduced by Alamouti was later generalized to an arbitrary number of transmit antennas and is able to achieve the full diversity promised by transmit and receive antennas [7, 23]. However, the limitation of these space-time codes is that they are not designed to achieve an additional coding gain. To achieve coding gain, they have to be concatenated with an outer code. Here, we briefly review the basics of STBCs. Figure 2.7 represents the block diagram of Alamouti STBC with two transmit antennas. The input data symbols are divided into two groups. At a given symbol period, the two symbols $\{a_1, a_2\}$ from two groups are transmitted simultaneously from antennas 1 and 2 respectively. In the next symbol period, the signal $-a_2^*$ is transmitted from antenna 1 and the signal a_1^* is transmitted

from antenna 2. If the channel is time-invariant for at least two symbol periods, then the received signal over two consecutive symbol periods can be expressed as

$$r(l) = h_1 a_1 + h_2 a_2 + n(l) \quad (2.8.9)$$

$$r(l+1) = -h_1 a_2^* + h_2 a_1^* + n(l+1) \quad (2.8.10)$$

where h_1 and h_2 represent the channel from the first and second transmit antennas to the receive antenna, respectively. $n(l)$ is the AWGN with mean zero and variance σ^2 . We define the received vector $\mathbf{r} = [r_1, r_2^*]^T$, the code symbol vector $\mathbf{a} = [a_1, a_2]^T$, the noise vector $\mathbf{n} = [n_1, n_2^*]^T$. Equation (2.8.10) can be rewritten in matrix form as

$$\begin{bmatrix} r(l) \\ r^*(l+1) \end{bmatrix} = \begin{bmatrix} h_1 & h_2 \\ h_2^* & -h_1^* \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} n(l) \\ n(l+1) \end{bmatrix} = \mathbf{H}\mathbf{a} + \mathbf{n} \quad (2.8.11)$$

Note that $\mathbf{H}$ is now only a virtual channel matrix. The ML decoding can be simplified by realizing that the channel matrix $\mathbf{H}$ is orthogonal. Hence, multiplying the receive vector $\mathbf{r}$ by $\mathbf{H}^*$, yields

$$\tilde{\mathbf{r}} = \mathbf{H}^* \mathbf{r} = (h_1^2 + h_2^2) \mathbf{a} + \tilde{\mathbf{n}} \quad (2.8.12)$$

where the noise vector $\tilde{\mathbf{n}} = \mathbf{H}^* \mathbf{n}$ is orthogonal. In this case, the ML decoding simplifies to separate and much simpler decoding rule for a_1 and a_2 .

$$\hat{a}_1 = \arg \min_{a_1 \in \mathcal{A}} \|\tilde{r}_1 - (h_1^2 + h_2^2) a_1\| \quad (2.8.13)$$

$$\hat{a}_2 = \arg \min_{a_2 \in \mathcal{A}} \|\tilde{r}_2 - (h_1^2 + h_2^2) a_2\| \quad (2.8.14)$$

where $\tilde{\mathbf{r}} = [\tilde{r}_1, \tilde{r}_2]^T$ and $\tilde{\mathbf{n}} = [\tilde{n}_1, \tilde{n}_2]^T$. Assuming a signaling constellation with 2^b points, this reduces the decoding complexity from 2^{2b} to 2×2^b .

MIMO Extension: STBC can be easily extended to more than 1 receive antennas. When the receiver uses N receive antennas, the received vector r_n at the receive

Figure 2.7: STBC (Alamouti) with two transmit antenna

antenna n is

$$\mathbf{r}_n = \mathbf{H}_n \mathbf{a} + \mathbf{n}_n \quad (2.8.15)$$

where $\mathbf{H}_n$ is the channel matrix from the two transmit antenna to the n -th receive antenna. In this case the ML decoding rule is

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a} \in \mathcal{A}} \sum_{n=1}^N \|\mathbf{r}_n - \mathbf{H}_n \mathbf{a}\| \quad (2.8.16)$$

As in two receive antennas, the decoding rule can be simplified by premultiplying the received signal r_n by $\mathbf{H}_n^*$, then the decision rule in (2.8.16) amounts to performing a hard decision on $\tilde{\mathbf{r}} = \sum_{n=1}^N \mathbf{H}_n^* \mathbf{r}_n$. Figure 2.8 shows the block diagram of receiver with two transmit antennas. The extension of the above STBC to more than two transmit antennas was discussed in [23]. There, a general technique for constructing STBCs for $P > 2$ was developed which achieves full diversity promised by multiple antennas and uses a simple ML decoding.

Figure 2.8: Receiver structure for STBC (Alamouti)

2.9 Summary

In this chapter, we discussed the fundamental concepts of MIMO wireless communication systems. The first category of these multi-antenna systems are designed to increase the data rate and are known as spatial multiplexing systems. The BLAST architecture, first proposed by Foschini, is a spatial multiplexing system and it was shown to achieve a hefty part of the promised theoretical spectral efficiencies. Originally two different BLAST architecture were proposed:

- The D-BLAST architecture theoretically achieves MIMO channel capacity with 1-D codec technology but is very complex to implement.
- The V-BLAST architecture has an efficient receiver but it suffers from error propagation and deep fades. It was shown that it can achieve 40 – 50% of the promised theoretical channel capacity [11].

The second category of MIMO systems include space-time codes. These space-time codes introduce diversity and coding gain to MIMO wireless communication systems. Two different kind of space-time codes were discussed:

- STTC were proposed by Tarokh et al combines signal processing at the receiver with coding technique at the transmitter and provides significant diversity and coding gains. When the number of transmit antennas is fixed the ML decoding complexity of STTC increases with diversity level and transmission rate.
- To reduce the decoding complexity of STTC, STBCs were introduced by Alamouti and Tarokh. STBC can attain maximum diversity gain promised by multi-antenna systems but it suffers from coding gain.

Chapter 3

MIMO Channel Equalization

Wireless communications systems require signal processing techniques to improve the system performance in mobile radio environment. Multipath propagation and Doppler spread are two main impairments of mobile radio channels. These impairments cause distortion which can degrade system performance. Equalization techniques compensate for ISI created by multipath within time-dispersive channels. If the signal bandwidth exceeds the coherence bandwidth of the channel ISI occurs and transmitted pulses are spread in time into adjacent symbols. The problem of channel equalization has been already studied in the literature [24–27]. However, the introduction of multi-element antenna arrays at both transmitter and receiver has motivated new research on equalization techniques for MIMO channels. The challenge in MIMO channel equalization is the design of robust and low complexity receivers to mitigate ISI and interference simultaneously, while obtaining good performance. Optimum maximum likelihood sequence estimation for MIMO channel was developed in [28]; however, its exponential complexity increases with the number of transmit antennas and channel memory makes its implementation costly over severely time-dispersive channels. The suboptimal receiver for MIMO channel equalization has been already

studied in the literature [29–34]. In this chapter we review MMSE channel equalization for the MIMO channel. Finite length and infinite length design are considered. It is assumed that the channel is quasi-static during a block and it is known at the receiver. The system of interest has P transmit antennas and N receive antennas ($P \leq N$). Each component of channel matrix is considered to be frequency selective.

3.1 System Model

We start this section by describing the input-output model of the MIMO system over a frequency selective channel. We consider the general case of linear, dispersive digital communication system with P transmit antennas and N receive antennas. The MIMO channel is assumed to be quasi-static within a block where each element of the channel matrix is considered to be frequency selective with memory L . The channel is also assumed to be AWGN. We consider the standard discrete-time complex-valued baseband signal model. Assuming an oversampling factor of I , the sampled received signal at n th receive antenna ($1 \leq n \leq N$) is

$$\mathbf{r}_n(m) = \sum_{p=0}^{P-1} \sum_{k=0}^L \mathbf{h}_{pn}(k) a_p(m-k) + \mathbf{n}_n(m) \quad (3.1.1)$$

$$\begin{aligned} &\triangleq \sum_{p=1}^P \mathbf{h}_{pn} [a_p(k), \dots, a_p(k-L)]^T + \mathbf{n}_n(m) \\ &m = 0, \dots, \pm I, \pm 2I, \dots, \pm \infty. \end{aligned} \quad (3.1.2)$$

where $\mathbf{r}_n(m)$ is the received signal at the n -th receive antenna, $\mathbf{h}_{pn}(k)$ $k = 0, I, LI-1$ is the channel impulse response between the p -th transmit antenna and the n -th receive antenna, whose memory is denoted by L , and $\mathbf{n}_n(m)$ is the noise at the receive antenna n . All of these vectors are $I \times 1$ column vector corresponding I

samples per symbol i.e, $\mathbf{v}(m) = [v(m - T + T/I), v(m - T + 2T/I), \dots, v(m)]$ and T is the symbol rate. Furthermore the overall channel impulse response between the p -th transmit antenna and n -th receive antenna is represented by the vector $\mathbf{h}_{pn} \triangleq [\mathbf{h}_{pn}(0) \ \mathbf{h}_{pn}(1) \ \dots \ \mathbf{h}_{pn}(L)]$. By stacking the received samples from all P transmit antennas at symbol time m into an $IN \times 1$ column vector $\mathbf{r}(m)$, the (3.1.2) can be written as

$$\mathbf{r}(m) = \sum_{k=0}^L \mathbf{H}_k \mathbf{a}(m - k) + \mathbf{n}(m) \quad (3.1.3)$$

where

$$\mathbf{H}_k = \begin{bmatrix} \mathbf{h}_{11}(k) & \dots & \mathbf{h}_{1P}(k) \\ \mathbf{h}_{12}(k) & \dots & \mathbf{h}_{2P}(k) \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{N1}(k) & \dots & \mathbf{h}_{PN}(k) \end{bmatrix} \quad (3.1.4)$$

is the $IN \times P$ MIMO channel matrix, and $\mathbf{a}(m - k)$ is an $P \times 1$ input vector at time $m - k$. Assuming that the channel matrix is time-invariant within a block of length K_f , equation (3.1.3) can be expressed in matrix notation as follows

$$\begin{bmatrix} \mathbf{r}(m + K_f - 1) \\ \mathbf{r}(m + K_f - 2) \\ \vdots \\ \mathbf{r}(m) \end{bmatrix} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_1 & \dots & \mathbf{H}_v & 0 & \dots & 0 \\ 0 & \mathbf{H}_0 & \mathbf{H}_1 & \dots & \mathbf{H}_v & 0 & \dots \\ \vdots & & \ddots & & \ddots & & \vdots \\ 0 & \dots & 0 & \mathbf{H}_0 & \mathbf{H}_1 & \dots & \mathbf{H}_v \end{bmatrix} \begin{bmatrix} \mathbf{a}(m + K_f - 1) \\ \mathbf{a}(m + K_f - 2) \\ \vdots \\ \mathbf{a}(m - v) \end{bmatrix} + \begin{bmatrix} \mathbf{n}(m + K_f - 1) \\ \mathbf{n}(m + K_f - 2) \\ \vdots \\ \mathbf{n}(m) \end{bmatrix} \quad (3.1.5)$$

or compactly, it can be written as

$$\mathbf{r}_{m:m+K_f-1} = \mathbf{H} \mathbf{a}_{m:m+K_f-1} + \mathbf{n}_{m:m+K_f-1} \quad (3.1.6)$$

where $\mathbf{r}_{m:m+K_f-1} = [\mathbf{r}(m+K_f-1)^T, \dots, \mathbf{r}(m)^T]^T$, $\mathbf{a}_{m:m+K_f-1} = [\mathbf{a}(m+K_f-1)^T, \dots, \mathbf{a}(m-v)^T]^T$, $\mathbf{n}_{m:m+K_f-1} = [\mathbf{n}(m+K_f-1)^T, \dots, \mathbf{n}(m-v)^T]^T$ and the channel matrix $\mathbf{H}$ is defined as

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_0 & \mathbf{H}_1 & \dots & \mathbf{H}_v & 0 & \dots & 0 \\ 0 & \mathbf{H}_0 & \mathbf{H}_1 & \dots & \mathbf{H}_v & 0 & \dots \\ \vdots & & \ddots & & \ddots & & \vdots \\ 0 & \dots & 0 & \mathbf{H}_0 & \mathbf{H}_1 & \dots & \mathbf{H}_v \end{bmatrix} \quad (3.1.7)$$

Defining the $P(K_f + v) \times P(K_f + v)$ input autocorrelation matrix

$$\mathbf{R}_{aa} \triangleq \mathbb{E}[\mathbf{a}_{m:m+K_f-1} \mathbf{a}_{m:m+K_f-1}^*] \quad (3.1.8)$$

and the $NIK_f \times NIK_f$ noise auto-correlation matrix

$$\mathbf{R}_{nn} \triangleq \mathbb{E}[\mathbf{n}_{m:m+K_f-1} \mathbf{n}_{m:m+K_f-1}^*] \quad (3.1.9)$$

the input-output cross-correlation and the output auto-correlation matrices are given by

$$\mathbf{R}_{ra} \triangleq \mathbb{E}[\mathbf{r}_{m:m+K_f-1} \mathbf{a}_{m:m+K_f-1}^*] = \mathbf{H} \mathbf{R}_{aa} \quad (3.1.10)$$

$$\mathbf{R}_{rr} \triangleq \mathbb{E}[\mathbf{r}_{m:m+K_f-1} \mathbf{r}_{m:m+K_f-1}^*] = \mathbf{H} \mathbf{R}_{aa} \mathbf{H}^* + \mathbf{R}_{nn} \quad (3.1.11)$$

3.2 Finite length MMSE MIMO-DFE

The optimal setting for finite length MIMO-DFE was derived in [30]. In this section we review the different receiver structures for MIMO-DFE and then derive the optimal settings for the equalizers.

The FIR MIMO-DFE consists of a feedforward (FFF) and feedback filter (FBF) matrix with K_f and K_B taps respectively. The feedforward filter matrix is

$$\mathbf{W}^* \triangleq [\mathbf{W}_0^* \quad \mathbf{W}_1^* \quad \dots \quad \mathbf{W}_{K_f-1}^*] \quad (3.2.1)$$

Figure 3.1: Block Diagram of MIMO-DFE

where each tap $\mathbf{W}_i$ is of size $NI \times P$ and the feedback filter matrix is equal to

$$[\mathbf{I}_P - \mathbf{B}_0^* \quad -\mathbf{B}_1^* \quad \dots \quad -\mathbf{B}_{K_B-1}^*] \quad (3.2.2)$$

where $\mathbf{B}_i$ is a $P \times P$ matrix. Furthermore, we define the matrix

$$\tilde{\mathbf{B}}^* \triangleq [\mathbf{0}_{P \times P\Delta} \quad \mathbf{I}_P - \mathbf{B}_0^* \quad -\mathbf{B}_1^* \quad \dots \quad -\mathbf{B}_{K_B-1}^*] \quad (3.2.3)$$

where $0 \leq \Delta \leq K_f + v - 1$ is the decision delay that satisfies the condition $(\Delta + K_B + 1) = (K_f + v)$. Assuming correct previous decisions, it can be shown that the MMSE error vector at time m is given by

$$\mathbf{E}(m) = \tilde{\mathbf{B}}\mathbf{a}(m) - \mathbf{W}^*\mathbf{r}(m) \quad (3.2.4)$$

Therefore, the $P \times P$ auto-correlation matrix is equal to

$$\mathbf{R}_{ee} = \mathbb{E}\{\mathbf{E}(m)\mathbf{E}(m)^*\} \quad (3.2.5)$$

$$\begin{aligned} &= \tilde{\mathbf{B}}^*(\mathbf{R}_{aa} - \mathbf{R}_{ar}^{-1}\mathbf{R}_{ra})\tilde{\mathbf{B}} + (\mathbf{W}^* - \tilde{\mathbf{B}}^*\mathbf{R}_{ar}\mathbf{R}_{rr}^{-1}) \\ &\quad \mathbf{R}_{rr}(\mathbf{W}^* - \tilde{\mathbf{B}}\mathbf{R}_{ar}\mathbf{R}_{rr}^{-1}) \\ &= \tilde{\mathbf{B}}^*\mathbf{R}^\perp\tilde{\mathbf{B}} + \mathbf{G}^*\mathbf{R}_{rr}\mathbf{G} \end{aligned} \quad (3.2.6)$$

where we introduced the matrices $\mathbf{R}^\perp = \mathbf{R}_{aa} - \mathbf{R}_{ar}^{-1}\mathbf{R}_{ra}$ and $\mathbf{G} = \mathbf{W}^* - \tilde{\mathbf{B}}\mathbf{R}_{ar}\mathbf{R}_{rr}^{-1}$. The trace of $\mathbf{R}_{ee}$ denotes the MMSE. Now, assuming ideal channel estimation at the receiver, we derive the optimum setting of MIMO-DFE [30]. Applying the orthogonality principle, i.e, $\mathbb{E}[\mathbf{e}(m)\mathbf{r}^*(m)] = 0$, it can be shown that the optimum matrix feedforward and feedback filters are related by

$$\mathbf{W}_{opt} = \tilde{\mathbf{B}}_{opt}^*\mathbf{R}_{ra}\mathbf{R}_{rr}^{-1} \quad (3.2.7)$$

It can be shown that this solution minimizes the MSE. Replacing (3.2.7) in (3.2.8), we can write

$$\begin{aligned} \mathbf{R}_{ee} &= \tilde{\mathbf{B}}^*\mathbf{R}^\perp\tilde{\mathbf{B}} \\ &= \tilde{\mathbf{B}}^*(\mathbf{R}_{aa} - \mathbf{R}_{ra}\mathbf{R}_{aa}^{-1}\mathbf{R}_{ra})\tilde{\mathbf{B}} \\ &= \tilde{\mathbf{B}}^*(\mathbf{R}_{aa} - \mathbf{R}_{aa}\mathbf{H}^*(\mathbf{H}\mathbf{R}_{aa}\mathbf{H}^* + \mathbf{R}_{nn})^{-1}\mathbf{H}\mathbf{R}_{aa})\tilde{\mathbf{B}} \\ &= \tilde{\mathbf{B}}^*(\mathbf{R}_{aa}^{-1} + \mathbf{H}^*\mathbf{R}_{nn}^{-1}\mathbf{H})^{-1}\tilde{\mathbf{B}} \\ &= \tilde{\mathbf{B}}^*\mathbf{R}^{-1}\tilde{\mathbf{B}} \end{aligned} \quad (3.2.8)$$

In [30] three MIMO-DFE receiver were proposed. Here we briefly review these structures.

Receiver Structure 1: In this structure only previous decision of other substreams

are available for any substream at present time and hence their interference can be suppressed, i.e., $\mathbf{B}_0 = \mathbf{I}_{n_i}$. We define the block Cholesky factorization

$$\begin{aligned}\mathbf{R} &= \mathbf{R}_{aa} + \mathbf{H}^* \mathbf{R}_{nn}^{-1} \mathbf{H} \\ &= \begin{bmatrix} \mathbf{L}_1 & 0 \\ \mathbf{L}_2 & \mathbf{L}_3 \end{bmatrix} \begin{bmatrix} \mathbf{D}_1 & 0 \\ 0 & \mathbf{D}_2 \end{bmatrix} \begin{bmatrix} \mathbf{L}_1^* & \mathbf{L}_2^* \\ 0 & \mathbf{L}_3^* \end{bmatrix} \\ &\triangleq \mathbf{L} \mathbf{D} \mathbf{L}^*\end{aligned}\tag{3.2.9}$$

where $\mathbf{L}_1$ is an $P(\Delta + 1) \times P(\Delta + 1)$ matrix and $\mathbf{L}$ has $P \times P$ identity block entries on the diagonal. Also, $\mathbf{D}$ is formed from $P \times P$ diagonal. The optimum feedback filter matrix and the minimum error matrix can be written as

$$\begin{aligned}\tilde{\mathbf{B}}_{opt} &= \begin{bmatrix} \mathbf{I}_{P(\Delta+1)} \\ \mathbf{L}_2 \mathbf{L}_1^{-1} \end{bmatrix} \mathbf{C} = \begin{bmatrix} \mathbf{I}_P \\ \mathbf{L}_2 \mathbf{L}_1^{-1} \mathbf{C} \end{bmatrix} \\ &= \mathbf{L} \begin{bmatrix} \mathbf{L}_1^{-1} \\ 0 \end{bmatrix} [\mathbf{e}_{P\Delta_{opt}}, \dots, \mathbf{e}_{P(\Delta_{opt}+1)-1}] \\ \mathbf{R}_{ee,min} &= \mathbf{C}^* \mathbf{L}_1^{-*} \mathbf{D}_1^{-1} \mathbf{L}_1^{-1} \mathbf{C} \\ &= \text{diag}(d_{P\Delta_{opt}}^{-1}, \dots, d_{P(\Delta_{opt}+1)-1}^{-1})\end{aligned}\tag{3.2.10}$$

where, the decision delay Δ_{opt} ($0 \leq \Delta_{opt} \leq N_f + v - 1$) is chosen to minimize the trace of $\mathbf{R}_{ee,min}$. The optimum feedforward filter matrix can be expressed as follows

$$\begin{aligned}\mathbf{W}_{opt}^* &= \tilde{\mathbf{B}}_{opt}^* \mathbf{R}_{aa} \mathbf{H}^* (\mathbf{H} \mathbf{R}_{aa} \mathbf{H}^* + \mathbf{R}_{nn})^{-1} \\ &= \tilde{\mathbf{B}}_{opt}^* (\mathbf{R}_{aa}^{-1} + \mathbf{H}^* \mathbf{R}_{nn}^{-1} \mathbf{H})^{-1} \mathbf{H}^* \mathbf{R}_{nn}^{-1} \\ &= \begin{bmatrix} d_{P\Delta_{opt}}^{-1} \mathbf{e}_{P\Delta_{opt}}^* \\ \vdots \\ d_{P\Delta_{(opt+1)-1}}^{-1} \mathbf{e}_{P\Delta_{(opt+1)-1}}^* \end{bmatrix} \mathbf{L}^{-1} \mathbf{H}^* \mathbf{R}_{nn}^{-1}\end{aligned}\tag{3.2.11}$$

Receiver Structure 2: The substreams are ordered so that lower index substreams are detected first, then the current decision from lower-indexed substreams are used by

higher indexed users, $\mathbf{B}_0$ is now a lower-triangular matrix. If we define the partitioning $\mathbf{R} \triangleq \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{12}^* & \mathbf{R}_{22} \end{bmatrix}$, where $\mathbf{R}_{11}$ is of size $P(\Delta + 1) \times P(\Delta + 1)$ and

$$\mathbf{R}_{11}^{-1} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{R}_2 \\ \mathbf{R}_2^* & \mathbf{R}_3 \end{bmatrix} \quad (3.2.12)$$

where $\mathbf{R}_1$ is $P\Delta \times P\Delta$ and $\mathbf{R}_3$ is $P \times P$. It can be shown that the error covariance matrix simplifies to [30]:

$$\mathbf{R}_{ee,min} = \mathbf{B}_0 \mathbf{R}_3 \mathbf{B}_0 \quad (3.2.13)$$

It can be easily shown that the optimum lower-triangular matrix $\mathbf{B}_0$ which minimizes trace ($\mathbf{R}_{ee,min}$) is given by

$$\begin{aligned} \mathbf{R}_3^{-1} &\triangleq \mathbf{L}_3 \mathbf{D}_3 \mathbf{L}_3^* \\ \mathbf{B}_0^{opt} &= \mathbf{L}_3 \\ \mathbf{R}_{ee,min} &= \mathbf{D}_3^{-1} \end{aligned} \quad (3.2.14)$$

A second approach for computing the optimum MIMO-DFE filter setting is by performing the standard Cholesky factorization of matrix $\mathbf{R} = \mathbf{R}_{aa}^{-1} + \mathbf{H}^* \mathbf{R}_{nn}^{-1} \mathbf{H} = \mathbf{L} \mathbf{D} \mathbf{L}^*$. Then it can be shown that the optimum feedback filter matrix is given by the P adjacent column of $\mathbf{L}$ that correspond to a diagonal matrix with the smallest trace.

Receiver Structure 3: The multistage detectors was proposed in [35], in which current decision from all other substreams, obtained from a previous detection stage, are available for the detection of the desired substream. Therefore, suppressing their interfering effects would improve performance. This structure has the same mathematical formulation as structures 1 and 2 except that $\mathbf{B}_0$ is now only constrained to have unit diagonal elements, i.e., $\mathbf{e}_i \mathbf{B}_0 \mathbf{e}_i = 1$ for all $0 \leq i \leq P - 1$. The optimum

feedback filter matrix $\mathbf{B}_0$ is the solution of the following optimization problem:

$$\begin{aligned} & \min_{\mathbf{B}_0} \text{trace}(\mathbf{B}_0^* \mathbf{R}_3 \mathbf{B}_0) \\ & \text{subject to} \\ & \mathbf{e}_i \mathbf{B}_0 \mathbf{e}_i = 1 \text{ for all } 0 \leq i \leq (P-1) \end{aligned} \quad (3.2.15)$$

Using the calculus of variation Lagrange multiplier technique it can be shown that the optimum $\mathbf{B}_0$ is given by:

$$\mathbf{B}_0^{opt} = \frac{\mathbf{R}_3^{-1}}{\mathbf{e}_i^* \mathbf{R}_3^{-1} \mathbf{e}_i} : 0 \leq p \leq P-1 \quad (3.2.16)$$

It is shown in [30] that structure 3 achieves lower MMSE than structure 2 and structure 2 achieves lower MMSE than structure 1.

Computation of the optimum feedforward and feedback filter is complex. An efficient algorithm for the computation of these optimum setting based on the displacement structure theory was presented in [30].

3.3 Wideband Layered Space-Time Receiver

We reviewed the layered space-time architecture concepts in the previous chapter. However, this layered space-time architecture was originally proposed for narrowband signals. In order to enable high data rate transmission, it might be necessary to transmit wide band signals. This brings in the issues of frequency selective channels into the design of wideband layered space-time receivers. Two different receiver structures that operate on the basis of layered space-time architecture were introduced in [36]. Here, we briefly review these receiver architectures.

3.3.1 System and Channel Model

To simplify the analysis, we consider a different channel model than (3.1.7). We consider a discrete-time complex baseband model for a single-user link, the received signals are sampled at the symbol rate and perfect channel estimation is assumed at the receiver. Also, for simplicity it is assumed that the transmitted signal at time m and the receiver additive white gaussian noise is spatially uncorrelated with covariance matrix

$$E[\mathbf{a}(m)\mathbf{a}^*(m)] = \frac{P_T}{P}\mathbf{I}_P \quad (3.3.1)$$

$$E[\mathbf{n}(m)\mathbf{n}^*(m)] = \sigma^2\mathbf{I}_N \quad (3.3.2)$$

where P_T is the total average transmit power. We stack the $\mathbf{h}_{pn}$ vectors into P separate matrices as follows ¹:

$$\mathbf{H}_p = \begin{bmatrix} \mathbf{h}_{p1}^T \\ \mathbf{h}_{p2}^T \\ \vdots \\ \mathbf{h}_{pN}^T \end{bmatrix}, \quad p = 1, 2, \dots, P. \quad (3.3.3)$$

where $\mathbf{h}_{pn} = [h_{pn}(0), \dots, h_{pn}(L-1)]^T$ with this definition, we can express the $N \times 1$ received $\mathbf{a}(m)$ as

$$\mathbf{r}(m) = \sum_{p=1}^P \mathbf{H}_p \mathbf{a}_p(m) + \mathbf{n}(m) \quad (3.3.4)$$

The sequence of $v+1$ symbols transmitted by the m -th antenna denoted by

$$\mathbf{a}_p(m) = \begin{bmatrix} a_p(m) \\ a_p(m-1) \\ \vdots \\ a_p(m-L) \end{bmatrix}. \quad (3.3.5)$$

¹To simplify the receiver equations a different model from the model in page 38 is used.

Figure 3.2: Block Diagram of Partially Connected OSIC-Receiver

A. Partially Connected OSIC Wideband Receiver

The receiver structure is shown in Figure 3.2 which consists of M successive stages (MISO-DFE), each having a feedforward filter $\mathbf{W}$ with K_f taps and a feedback filter with K_b taps. The K_f N -dimensional symbols spanned by the feedforward filter can be grouped as

$$\mathbf{R}(m) = \begin{bmatrix} \mathbf{r}(m) & \mathbf{r}(m-1) & \dots & \mathbf{r}(m-K_f+1) \end{bmatrix} \quad (3.3.6)$$

the N -dimensional received vector can be expressed as

$$\mathbf{r}(m) = \sum_{p=1}^P \bar{\mathbf{H}}_p \bar{\mathbf{a}}_p(m) + \mathbf{n}(m) \quad (3.3.7)$$

where $\mathbf{r}(m) = \text{vec}(\mathbf{R}(m))$ and $\bar{\mathbf{a}}_p(m)$ are $\bar{\mathbf{H}}_p$ defined as

$$\bar{\mathbf{a}}_p(m) = \begin{bmatrix} a_p(m) \\ a_p(m-1) \\ \vdots \\ \mathbf{a}_p(m-v-k_f) \end{bmatrix} \quad (3.3.8)$$

$$\bar{\mathbf{H}}_p = \begin{bmatrix} \mathbf{H}_p & \dots & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & \dots & \mathbf{H}_p \end{bmatrix} \quad (3.3.9)$$

In addition, we have

$$\mathbf{n}(m) = \text{vec}([\mathbf{n}(m) \quad \mathbf{n}(m-1) \quad \dots \quad \mathbf{n}(m-K_f+1)]). \quad (3.3.10)$$

At each stage, the best data stream in the MMSE sense is extracted, detected, and cancelled out. With the decision delay Δ , the output of every stage at time m is an estimate $\hat{\mathbf{a}}_p(m-\delta)$ for symbol $\mathbf{a}_p(m-\delta)$. The sequence of K_b most recent decisions produced at that stage for data substream p is:

$$\hat{\mathbf{a}}_p(k-\Delta-1) = \begin{bmatrix} \hat{\mathbf{a}}_p(k-\Delta-1) \\ \vdots \\ \hat{\mathbf{a}}_p(k-\Delta-K_b) \end{bmatrix} \quad (3.3.11)$$

which is an input to the feedback filter at each stage. Let us $\mathbf{W}_p$ and $\mathbf{b}_p$ denote the feedforward and feedback filter matrix that extract substream p as

$$\begin{aligned} y_p(m) &= \mathbf{w}_p^* \mathbf{r}(m) - \mathbf{b}_p^* \hat{\mathbf{a}}_p(m-\Delta-1) \\ &= \begin{bmatrix} \mathbf{w}_p \\ \mathbf{b}_p^* \end{bmatrix}^* \begin{bmatrix} \mathbf{r} \\ -\hat{\mathbf{a}}_p(m-\Delta-1) \end{bmatrix}. \end{aligned} \quad (3.3.12)$$

where $\mathbf{w}_p^* = \text{vec}(\mathbf{W}_p)$. Assuming correct decisions at all stages, the optimum MMSE setting for $\mathbf{w}_p$ and $\mathbf{b}_p$ is

$$\begin{bmatrix} \mathbf{w}_p^* \\ \mathbf{b}_p \end{bmatrix} = \mathbf{R}_p^{-1} \mathbf{p}_p. \quad (3.3.13)$$

with [36]

$$\begin{aligned}\mathbf{b}_p &= \mathbb{E} \left[\begin{array}{c} \mathbf{r} \\ -\mathbf{a}_p(m - \Delta - 1) \end{array} \right] \mathbf{a}_p^*(m - \Delta) \\ &= \frac{P_T}{M} \left[\begin{array}{c} (\bar{\mathbf{H}}_p)_{\Delta+1} \\ 0 \end{array} \right].\end{aligned}\quad (3.3.14)$$

and

$$\begin{aligned}\mathbf{R}_p &= \mathbb{E} \left(\left[\begin{array}{c} \mathbf{r}(m) \\ -\mathbf{a}_p(m - \Delta - 1) \end{array} \right] \left[\begin{array}{c} \mathbf{r}(m) \\ -\mathbf{a}_p(m - \Delta - 1) \end{array} \right]^* \right) \\ &= \frac{P_T}{M} \left[\begin{array}{cc} \sum_i \bar{\mathbf{H}}_i \bar{\mathbf{H}}_i^* + \sigma^2 \mathbf{I}_{K_f N} & -(\bar{\mathbf{H}}_p)_{\Delta+2 \dots \Delta_{K_b}+1} \\ -(\bar{\mathbf{H}}_p)_{\delta+2 \dots \Delta+K_b+1}^* & \mathbf{I}_{K_b} \end{array} \right]\end{aligned}\quad (3.3.15)$$

where the summation is over yet-to-be-detected substreams. The MSE for substream p can be written as

$$\mathbb{E}|y_p(m) - a_p(m - \Delta)|^2 = 1 - \mathbf{p}_p^* \mathbf{R}_p^{-1} \mathbf{p}_p \quad (3.3.16)$$

where at each stage the undetected substreams with the smallest MSE in (3.3.16) is selected and extracted. Finally, once the substreams have been detected, its interference from the undetected substream can be removed from the input signal $\mathbf{r}(m)$ as follows:

$$\begin{aligned}\mathbf{r}(m) &\leftarrow \mathbf{r}(m) - (\mathbf{H}_m)_1 \hat{\mathbf{a}}_p(m) \\ \mathbf{r}(m+1) &\leftarrow \mathbf{r}(m+1) - (\mathbf{H}_m)_2 \hat{\mathbf{a}}_p(m) \\ &\vdots \\ &\vdots\end{aligned}\quad (3.3.17)$$

$$\mathbf{r}(m+v) \leftarrow \mathbf{r}(m+v) - (\mathbf{H}_m)_{v+1} \hat{\mathbf{a}}_p(m) \quad (3.3.18)$$

B. Fully Connected OSIC Wideband Receiver

The fully connected OSIC receiver is displayed in Figure 3.3. It has P stages, in which

Figure 3.3: Block Diagram of Fully Connected OSIC-Receiver

each stage consists of a MIMO-DFE. Similar to the partially connected receiver, the P substreams are ordered so that the data streams are detected in the sequence of decreasing MSE. Unlike the partially connected case, though, all the substreams are detected at each stage and their interferences are canceled out, but only the substreams with the lowest MSE are retained at each stage while the others are discarded and detected again in a later stage. Thus, the fully connected receiver has the advantage of increased feedback at every stage. At the same time, this powerful feedback makes the receiver more vulnerable to error propagation, in particular the tentative decisions from weaker streams are utilized to estimate the symbols of the strong streams. Since, the fully connected receiver is a straightforward cascade of P MIMO-DFE receivers, the optimum feedback and feedforward filter setting can be easily derived as in Section 2.2.

3.4 Generalized Wideband Layered Space-Time Receivers

The wideband layered space-time receiver in [36] utilizes a MISO or MIMO equalizer at each stage. To achieve tradeoffs between performance and complexity the authors in [37] replaced the MISO-DFE or MIMO-DFE at each stage with a MIMO delayed decision feedback sequence estimator (MIMO-DDFSE) which detect a group of selected data streams. With the proposed architecture in [37] the number of tentatively detected and output streams can be different at different stages.

The block diagram of the receiver for $m_o^{(j)} = 2$ and $m_i^{(j)} = P^{(j)}$ is shown in Figure 3.4. The required number of the stages is $P/2$. At the j -th stage, there are $P^{(j)}$ undetected substreams and the best $m_I^{(j)}$ of these substreams in the MMSE sense are selected for tentative decisions by the equalizer. The tentative decisions of the best $m_o^{(j)}$ substreams out of $m_I^{(j)}$ in the MMSE sense are retained and then cancelled from the received signal before they are passed to the next stage. Note that at each stage when tentatively detecting $m_I^{(j)}$ substreams, the other $P^{(j)} - m_I^{(j)}$ are regarded as interferers, and the number of undetected substreams can be expressed as $K^{(j)} = P - \sum_{p=1}^{j-1} m_o^{(p)}$. The advantage of this approach is that by selecting various $m_I^{(j)}$ and $m_o^{(j)}$ combinations, different tradeoffs between performance and complexity can be achieved. For instance, selecting $m_I^{(j)} = m_o^{(j)} = 1$ and $m_i^{(j)} = 1$, $m_o^{(j)} = j$ for each stage reduces to partially connected and fully connected receiver respectively.

The j -th stage of this architecture is shown in Figure 3.5, where a DDFSE equalizer is employed. At each stage, the received signals are passed through a feedforward filter of length K_f . This feedforward filter partially suppresses the interference and ISI,

Figure 3.4: Block Diagram of generalized layered space-time receiver with $m_i^j = P - 2(j - 1)$ and $m_o^j = 2$

yielding m_I^j outputs. The feedforward filter outputs are input to a vector DDFSE [38]. The DDFSE is composed of a vector Viterbi algorithm sequence estimator of memory μ^j and a feedback filter of length $m_i^j(B - \mu^j)$, which uses the tentative decisions of the VA sequence estimator. Assuming M-ary modulated signals the number of VA states grows exponentially as $M^{m_I^j \mu^j}$. Therefore, the DDFSE is a reduced-state MLSE, since μ_j can be chosen to be smaller than the channel memory. Note that when $\mu^j = 0$, DDFSE reduces to DFE.

3.5 Infinite-Length

In the infinite-length case, the length of the DFE filters are infinitely long (IIR); while in the finite-length case, these filters have finite length [31, 32, 39].

matrix filter $\mathbf{W}(z)$ and a feedback filter represented by the linear, causal matrix filter $\mathbf{I} - \mathbf{B}(z)$.

The same as finite-length DFE the optimum feedforward and feedback filters $\mathbf{W}$ and $\mathbf{B}$ minimize the MSE. If the output of feedforward filter is denoted by $\tilde{\mathbf{A}}(z)$, and the output to the decision elements is denoted by $\mathbf{A}(z)$ then, the MSE at the input of the decision device is

$$\text{MSE} \min_{\mathbf{W}, \mathbf{B}} E(\|\mathbf{a}(m) - \tilde{\mathbf{a}}(m)\|^2). \quad (3.5.4)$$

The error signal can be written as

$$\begin{aligned} \mathbf{e} &= \mathbf{a}(m) - \tilde{\mathbf{a}}(m) \\ &= \mathbf{a}(m) - \mathbf{W}_m * \mathbf{r}(m) - (\mathbf{I} - \mathbf{B}_m) * \mathbf{a}(m) \\ &= \mathbf{B}_m * \mathbf{a}(m) - \mathbf{W}_m * \mathbf{r}(m). \end{aligned} \quad (3.5.5)$$

Applying the orthogonality principle to minimize (3.5.5)

$$E[\mathbf{e}(m)\mathbf{r}(m)^*] = 0 \text{ for all } k. \quad (3.5.6)$$

we have the following relationship in the z-domain

$$\mathbf{B}(z)\mathbf{S}_{ar}(z) - \mathbf{W}(z)\mathbf{S}_{rr}(z) = 0 \quad (3.5.7)$$

where $\mathbf{S}_{ar}(z)$ and $\mathbf{S}_{rr}(z)$ are defined as

$$\mathbf{S}_{ra} = E[\mathbf{A}(z)\mathbf{R}^*(\frac{1}{z^*})] \quad (3.5.8)$$

$$\mathbf{S}_{rr} = E[\mathbf{R}(z)\mathbf{R}^*(\frac{1}{z^*})] \quad (3.5.9)$$

Assuming that the transmitted symbols are uncorrelated, i.e. $\mathbf{S}_a(z) = \sigma_a^2 \mathbf{I}$ where σ_a^2 is the average transmitted symbol energy then

$$\mathbf{W}(z) = \mathbf{B}(z)[\mathbf{H}(z) + \frac{\sigma_n^2}{\sigma_a^2} \mathbf{I}]^{-1} \quad (3.5.10)$$

where the $\sigma_n^2 \mathbf{I}$ is the noise covariance matrix. Now we define

$$\tilde{\mathbf{Q}}(z) = [\mathbf{H}(z) + \frac{\sigma_n^2}{\sigma_a^2} \mathbf{I}] \quad (3.5.11)$$

and let us define the spectral factorization of $\tilde{\mathbf{Q}}(z)$ by [31]

$$\tilde{\mathbf{Q}}(z) = k_0 \mathbf{G}(z) \mathbf{G}^* \left(\frac{1}{z^*} \right) \quad (3.5.12)$$

where $\mathbf{G}(z)$ is causal and stable matrix filter with causal inverse. The filter $\mathbf{G}^* \left(\frac{1}{z^*} \right)$ is stable if $\tilde{\mathbf{Q}}(z)$ is nonsingular on the unit circle. Then the feedback and feedforward filters are given by [31]

$$\mathbf{B}(z) = \mathbf{G}(z) \quad (3.5.13)$$

$$\mathbf{W}(z) = \mathbf{G}(z) \tilde{\mathbf{Q}}(z)^{-1}. \quad (3.5.14)$$

A solution for the spectral factorization (3.5.13) was given in [39].

3.6 Zero-Forcing (ZF) Design

The MMSE criterion minimizes the total distortion in the presence of noise. ZF criterion removes all ISI and interference in the absence of noise. The ZF-MIMO equalizer was derived in [32]. Here, we briefly review the ZF criterion.

Definition 3.6.1. Considering a finite length channel of size L in (3.5.1), and a MIMO-DFE with feedforward and feedback filter of degree K_f and K_B which estimates $\hat{\mathbf{a}}(\Delta - m)$ from the sequence of transmitted symbol vector $\mathbf{a}(m)$. If

$$\hat{\mathbf{a}}(m - \Delta) = \mathbf{a}(m - \Delta) + \mathbf{n}(m) \quad (3.6.1)$$

then the equalizer is said to be ZF.

It can be easily shown that that the ZF condition implies the relation

$$\mathbf{W}(z^{-1})\mathbf{H}(z^{-1})\mathbf{a}(m) - (\mathbf{B} - \mathbf{I}_P)\mathbf{a}(m - \Delta - 1) = \mathbf{a}(m - \Delta) \quad (3.6.2)$$

or a DFE will be in ZF if and only if $\mathbf{W}(z^{-1})$ and $\mathbf{B}(z^{-1})$ constitutes a solution to the polynomial matrix equation [32]:

$$\mathbf{W}(z^{-1})\mathbf{H}(z^{-1}) - z^{\Delta-1}(\mathbf{B} - \mathbf{I}_P) = z^{-\Delta}\mathbf{I}_P \quad (3.6.3)$$

In the absence of noise, the MMSE equalizer reduces to a ZF equalizer, and the estimation error will vanish. In this case perfect equalization is possible. If the ZF equalizer does not exist, all ISI and interference cannot be removed and the estimation error will not vanish.

When a ZF equalizer exists, good performance can be achieved when the signal-to-noise ratio (SNR) goes to infinity. When the ZF problem is ill-conditioned, the corresponding MMSE equalizer will work well also at realistic SNR's. However, when the ZF problem is ill-conditioned, the corresponding MMSE solution may not provide adequate performance, despite the fact that the ZF equalizer does exist [32].

3.7 Remarks on the Finite-Length MMSE MIMO-DFE

For a given channel, the DFE structure is given by the degrees of the feedforward and feedback filters as well as the decision delay.

- The degree K_f of the feedforward filter should be chosen to be as large as possible, at least equal to the decision delay Δ .

- The degree K_B of the feedback filter should be large enough to cancel all postcursor taps (taps with delay $> \Delta$). The number of postcursor taps after linear equalizing by feedforward filter is equal to $L + K_f - \Delta$, so we conclude that $K_B = L + K_f - \Delta - 1$.
- The decision delay Δ is chosen as a tradeoff between complexity and performance. However, choosing Δ larger than the delay spread L only leads to minor improvements.

3.8 Summary

In this chapter we briefly reviewed the MIMO frequency selective channels when the symbol rate is greater than the delay spread of the channel. In MIMO frequency selective channels, each transmitted symbol at each receive antenna is received while superimposed with ISI and interference. To detect the incoming transmitted signal different types of linear and nonlinear decision feedback equalizers have been proposed in the literature. We briefly reviewed existing MIMO-DFE receiver architectures and derive the optimal setting for this receiver.

Generally, equalizer design includes two categories: finite-length and infinite length. The infinite length procedure assumes that the feedforward and feedback filters are infinite-length, while in finite length design they are assumed to have finite length which is specified by the designer. The finite-length structure is by no means optimal, however the optimal feedforward and feedback filter setting can be derived by MMSE or ZF criterion.

The derivation of equalizer tap setting with ZF criterion is less complex than the

MMSE solution, however the ZF solution does not always exist and its performance is worse than the MMSE solution. When a ZF solution exists, ideal performance can be achieved when the SNR goes to infinity.

Chapter 4

Frequency Domain Equalization for MIMO Channels

Future generation wireless systems are required to support high quality of service at high data rates. For such high data rates we can have severe time-dispersion with very long ISI spans. In this case, the conventional time-domain equalization schemes are not practical since the signal processing requirements are very high. This can be more serious when conventional time domain equalization methods are employed in high data rate MIMO systems.

Frequency domain equalization techniques (FDE) in which the signal processing operations is done on a block by block basis, using fast Fourier transform (FFT) techniques have been shown to be a better alternative for severe time dispersive channels, allowing superior performance with lower implementation complexity. Orthogonal frequency division multiplexing (OFDM) modulation is a popular frequency domain technique. Single carrier (SC) modulation using FDE techniques is another valuable

candidate for highly dispersive channels in broadband wireless communications [40].

It is well-known that DFE can significantly outperform linear equalizers [41], provided that the error propagation problem is limited. However, time domain MIMO-DFE techniques are not practical for high data rate MIMO systems since they are complex [30, 36]. In this chapter we briefly review the frequency domain techniques and then we introduce two new hybrid time-frequency MIMO receivers based on the ideas of [42]. We design a frequency domain feedforward filter for joint interference and ISI cancellation and a time domain feedback filter for post-cursor residual ISI cancellation. We evaluate the performance of the proposed receivers in different scenarios assuming perfect channel estimation at the receiver. This part was the subject of a conference paper [43].

4.1 OFDM

OFDM is a popular, standardized technique for broadband wireless communication systems. It is used in digital audio/video broadcasting (DAB/DVB) standards in Europe, in high-speed digital subscriber line (DSL) modems in United States [44], wireless local area network (LAN) and digital cable television. IEEE802.11a, 802.16 and HIPERLAN/2 standards adopt it completely or partially [45].

By implementing an inverse fast Fourier transform (IFFT) at the transmitter, the transmitter modulation is carried out on a block of data symbols. The received block at the receiver is demodulated by implementing an FFT. To eliminate the inter-block interference (IBI) between successive transmitted blocks, a cyclic prefix (CP) of length greater than channel memory is inserted at the beginning of each block. At

the receiver the CP is discarded.

A simplified block diagram of OFDM is shown in Fig. 4.1. The serial-to-parallel and the parallel-to-serial conversions are dropped for convenience. The data stream $a(m)$ at the transmitter is converted from serial to parallel then, the transmitted sequence $b(l), l = 0, \dots, M$ is obtained through an M -point IFFT,

$$b(l) = \frac{1}{M} \sum_{m=0}^{M-1} a(m) \exp\left(\frac{j2\pi lm}{M}\right), \quad l = 1 \dots M \quad (4.1.1)$$

where $a(m)$ represents the m -th input symbol of IFFT block, and $b(l)$ is the l output of the transmitted block. After this transformation, the M parallel outputs are converted to serial, lowpass filtered, and passed to a quadrature modulator which shifts the signal spectrum to a center frequency f_0 .

On the receiver side, the received signal is coherently demodulated, sampled at the symbol rate, and then transformed to the frequency domain by taking FFT. The demodulator includes a lowpass filter which limits the noise and interference. The continuous time transmitted signal can be represented as

$$s(t) = \left\{ \sum_{l=-\infty}^{\infty} b(l) f(t - lT) e^{j(\omega_0 t + \phi)} \right\} \quad (4.1.2)$$

$f(t)$ denotes the transmit filter impulse response, T is the symbol period, ω_0 is the carrier frequency and ϕ is the carrier phase. The output spectrum of IFFT is composed of M sinc pulses with a frequency separation of $\frac{1}{MT}$. Thanks to the efficient FFT and IFFT implementation, the computational complexity of OFDM is approximately $\log_2 M$ operations per data symbol. Assuming an $\mathcal{M}$ -ary modulation for each symbol, OFDM can achieve a bandwidth efficiency of $\log_2 \mathcal{M}$ bit/s/Hz [46]. Given that the symbol rate is $1/T$, the corresponding bit rate of $\mathcal{M}$ -ary systems is $\log_2 \mathcal{M}/T$, and

Figure 4.1: Block Diagram of OFDM

each subchannel transmits at a lower rate $\log_2 \mathcal{M}/(MT)$. So, the total bandwidth of the OFDM system is

$$B = f_{M-1} - f_0 + 2\delta \quad (4.1.3)$$

where f_m $m = 0, \dots, M-1$ is the m -th subcarrier and δ is the one-sided bandwidth of the subchannel. Since the subcarriers are uniformly spaced $\delta = \frac{1}{2}\Delta_f$ then we have

$$\begin{aligned} f_{M-1} - f_0 &= (M-1)\Delta_f \\ &= (1 - 1/M)(1/T) \end{aligned} \quad (4.1.4)$$

with $\delta = \frac{1}{2}\Delta_f$, therefore the bandwidth efficiency β becomes

$$\beta = \frac{\log_2 \mathcal{M}}{(1 - \frac{1}{M}) + 2\delta T} \quad (4.1.5)$$

or $\beta = \log_2 \mathcal{M}$ bits/s/Hz.

4.1.1 Channel Equalization

Let $h(t)$ denote the channel impulse response and let $H(f)$ denote its Fourier transform. If the number of carriers is sufficiently large, then the channel transfer function is flat fading within the bandwidth of each subchannel. So, the influence of multipath

fading on each subcarrier is just the channel gain at each frequency and a phase rotation. The fact that each subchannel in OFDM has a different gain and phase rotation implies that the channel fading is not still flat and it needs to be equalized. Assuming $H(l)$, $l = 1 \dots M$ designates the frequency response of the channel at carrier l , ZF linear equalization of the channel requires that after FFT processing at the receiver, the k -th carrier signal be multiplied by a complex coefficient.

$$C_l = 1/H(l) \quad (4.1.6)$$

This is the result of the ZF criterion. The MMSE criterion yields

$$C_l = \frac{H^*(l)}{|H(l)|^2 + \sigma_n^2/\sigma_a^2} \quad (4.1.7)$$

where σ_n^2 is the variance of AWGN, and σ_a^2 is the variance of the transmitted symbol. Given that $R_{ZF}(l)$ and $R_{MMSE}(l)$, $l = 1 \dots M$ are the output samples after ZF and MMSE equalizer respectively, it can be easily shown that the signal and noise terms in $R_{MMSE}(l)$ are scaled versions of the signal and noise terms in $R_{ZF}(l)$, through a multiplication by the same real factor. Although the MSE is reduced under the MMSE criterion, the uncoded error probability at subchannel l is the same for ZF and MMSE criteria, provided that we assume optimum decision thresholds with both criterions [47]. Therefore there is no good reason for using MMSE criterion with OFDM, since the ZF criterion provides an identically optimum performance, while keeping the same decision thresholds. However, the ZF coefficients cannot be employed for a very bad subchannel, when $H(l) = 0$.

Although OFDM simplifies equalization at the receiver, it has the following disadvantages

- Since the transmitted signal is a sum of a large number of modulated subcarriers, it has a high Peak-to-Average power ratio (PAPR), necessitating transmitter power amplifier back-off, unless signal processing techniques are employed to control this ratio [48].
- Because the information data are transmitted on subcarriers, OFDM can exhibit sensitivity to transmit-receive oscillators' mismatch, Doppler effects and phase noise. [49].
- Uncoded OFDM is virtually unusable on multipath fading channels with deep notches occurring in the signal spectrum. The performance of uncoded OFDM is essentially dictated by the lowest SNR value of the subchannels [49]. Thus coding is essential.

4.2 Single-Carrier Frequency Domain Equalization

Frequency domain equalization, using efficient FFT and IFFT implementation can be applied to the the traditional single carrier modulation schemes. Figure 4.2 shows the block diagram of this single carrier system. The difference from Figure 4.1 is that in OFDM systems, both channel equalization and receiver decisions are performed in the frequency domain, whereas in single-carrier systems the receiver decisions are made in the time domain. As this figure shows, the IFFT operation at the transmitter is moved to the receiver. Similar to OFDM modulations, the data blocks are preceded by a CP, long enough to cope with the channel length. The CP is used to avoid the IBI between successive blocks and to make the linear convolution associated with the

Figure 4.2: Block Diagram of SC-FDE SC-FDE

channel equivalent to cyclic convolution, with respect to useful, M -length, part of the block. At the receiver the CP is discarded. The received symbols at the receiver are converted from serial to parallel, transformed to frequency domain by taking the FFT, equalized in frequency domain and then transformed back to the time domain by taking IFFT. Due to the circular convolution properties, the frequency domain equivalent of convolution of the channel and transmitted symbols, is converted into a simple complex multiplication per tone, which simplifies the equalization process, i.e., equalization is done with a 1-tap equalizer per tone. Although the operation of a single-carrier system with FDE is similar to OFDM, the receiver decisions in single carrier are based on the signal energy transmitted over the entire bandwidth. In other words, the average SNR of the channel dictates the system performance. Consequently, the performance degradation due to a deep null in the channel remains small compared to OFDM. The overall transmission plus reception implementation complexities for SC schemes with FDE and OFDM schemes are similar. However, the SC schemes have lower envelope fluctuation than the corresponding OFDM schemes, while offering similar, or even better performance. [42, 50, 51].

4.3 Hybrid Time-Frequency Domain MIMO Receivers

We briefly reviewed the frequency domain equalization concepts for SISO channels. Kadel explored the possibility of combining the frequency domain equalization with diversity techniques [52]. Later, Clark presented the potential performance of frequency domain adaptive algorithm with diversity techniques [53].

When multipath causes severe nulls in the channel frequency response linear equalizers have a high MSE. In this case DFE yields better performance, if the error propagation is limited [41]. When the channel has a long impulse response, the DFE needs a long feedforward and feedback filter, which is too complex to implement. Another alternative for severe time dispersive channel is to truncate the channel impulse response with a pre-filter and then use MLSE or DFE [54]. Recently a reduced complexity time domain DFE for high dispersive SISO channel was proposed in [55]. This reduced complexity receiver uses a long feedback filter whose length is as long as the channel impulse response and a short feedforward filter to capture the multipath energy. The feedback filter is multiplication free since it can be implemented by adders and lookup tables. However, this DFE still requires a very long feedback filter for highly dispersive channels.

In conventional DFE equalizers, the information data is filtered by the feedforward filter symbol-by-symbol and immediately fed back to the feedback filter to remove the postcursor ISI. Because of the delay in FFT processing, the feedback cannot be implemented in the frequency domain (except in iterative DFE). For this reason,

a hybrid time frequency SC-FDE was proposed in [42, 51], employing a frequency-domain feedforward filter and a short time-domain feedback filter. The feedback filter is relatively simple to implement since it performs only multiplication on a limited number of filtered symbols. Therefore, the hybrid time-frequency domain DFE has a better performance than a linear FDE. To limit the error propagation effect, the feedback filter length should be made as short as possible. In this section we extend the idea of hybrid DFE receiver to MIMO channel and develop two new low complexity MIMO-DFE receiver. Then, the performance of this receiver is evaluated in different scenarios.

4.3.1 System Model

We consider a discrete-time complex baseband model for a SC-MIMO system with P transmit and N receive antennas. The data is transmitted in blocks of M data symbols $a_p(m)$ ($m = 0, \dots, M-1$) from transmitter p at a symbol rate $\frac{1}{T}$. Each block is preceded by a cyclic prefix (CP) to make the linear convolution of the channel equivalent to a circular convolution. At the receiver, the cyclic prefix is discarded and the received signals are sampled at I/T , where $I > 1$ gives a fractionally-spaced receiver whose performance is not sensitive to sampling phase. It is assumed that the channel is quasi-stationary; furthermore, perfect channel estimation and synchronization are assumed at the receiver. The N dimensional received signal at the sampling instant m can be expressed as:

$$\mathbf{r}(m) = \sum_{p=0}^{P-1} \sum_{k=0}^{M-1} \mathbf{h}_p(mT/I - kT) a_p(k) + \mathbf{n}(mT/I), m = 0, \dots, MI-1. \quad (4.3.1)$$

where $\mathbf{r}(m) = [r_1(m), \dots, r_N(m)]^T$ and $\mathbf{h}_p(mT/I - kT) = [h_{p1}(mT/I - kT), \dots, h_{pN}(mT/I - kT)]^T$ ($p = 1, \dots, P$) with h_{pn} denoting the sampled channel impulse response between transmitter p and receiver n . $\mathbf{n}(\cdot)$ is the N dimensional noise vector with mean zero and variance $\sigma^2 \mathbf{I}_N$. The data symbols are assumed to be uncorrelated complex random variables which are selected from a given constellation (e.g., a QPSK constellation). After the FFT operation on a block of MI samples (4.3.1), the corresponding MI frequency domain samples can be expressed as:

$$\mathbf{R}(l + kM) = \sum_{p=1}^P \mathbf{H}_p(l + kM) A_p(l) + \mathbf{N}(l + kM) \quad l = 0, \dots, M-1, k = 0, \dots, I-1. \quad (4.3.2)$$

where:

$$A_p(l) = \sum_{m=0}^{M-1} a_p(m) \exp\left(\frac{-j2\pi ml}{M}\right) \quad (4.3.3)$$

$$\mathbf{N}(l + kM) = \sum_{m=0}^{MI-1} \mathbf{n}(l + kM) \exp\left(\frac{-j2\pi(l + kM)m}{M}\right) \quad l = 0, \dots, M-1, k = 0, \dots, I-1. \quad (4.3.4)$$

$$\mathbf{H}_p(l + kM) = \sum_{m=0}^{MI-1} \mathbf{h}_p(l + kM) \exp\left(\frac{-j2\pi(l + kM)m}{M}\right) \quad l = 0, \dots, M-1, k = 0, \dots, I-1. \quad (4.3.5)$$

We define the NI -dimension vectors $\mathbf{R}_l = \text{vec}([\mathbf{R}(l), \mathbf{R}(l+M), \dots, \mathbf{R}(l+(I-1)M)])$; $\mathbf{H}_l^p = \text{vec}([\mathbf{H}_p(l), \dots, \mathbf{H}_p(l+(I-1)M)])$; $\mathbf{N}_l^p = \text{vec}([\mathbf{N}(l), \dots, \mathbf{N}(l+(I-1)M)])$, then equation (4.3.2) can be expressed as:

$$\mathbf{R}_l = \sum_{p=1}^P \mathbf{H}_l^p A_p(l) + \mathbf{N}_l, \quad l = 0, \dots, M-1. \quad (4.3.6)$$

4.3.2 Hybrid Time-Frequency LST Receiver

The hybrid time-frequency domain layered space-time (LST) receiver for detecting the P streams of data symbol, shown in Figure 4.3, consists of P successive Multiple-Input-Single-Output (MISO) hybrid time-frequency domain DFEs. At each stage, the “best” substream data block, in the Minimum Mean Square Error (MMSE) sense, is selected, detected by a MISO-DFE, transformed to frequency domain by an FFT operation, is subtracted from the received signal in the frequency domain and the residual signal is passed to the next step for equalization and detection of the “next best” data block. The p th stage of the MISO-DFE, shown in Figure 4.4, has a frequency domain feedforward filter with MI taps $w_{pn}(l + kM)$ ($l = 0, \dots, M - 1; k = 0, \dots, I - 1$) at receive antenna n and B sparse time domain feedback taps. The feedforward filter suppresses the ISI and interference. The residual ISI can be further removed by the feedback filter. The m th time domain output of MISO-DFE, at a certain stage which extracts substream p , obtained after decimation by $1/I$ and M -point IFFT operation, is

$$y_p(m) = \frac{1}{MI} \sum_{l=0}^{M-1} \mathbf{W}_l^{p*} \mathbf{R}_l \exp\left(\frac{j2\pi lm}{M}\right) \quad (4.3.7)$$

where we define the NI -dimensional vector $\mathbf{W}_l^p = \text{vec}([\mathbf{W}_p(l), \mathbf{W}_p(l+M), \dots, \mathbf{W}_p(l+(I-1)M)])$ and $\mathbf{W}_p(l) = [w_{p1}(l), \dots, w_{pN}(l)]$. The soft estimate of the m th symbol at the p th stage can be written as

$$\tilde{a}_p(m) = y_p(m) - \sum_{k \in F_B(p)} f_p(k) \hat{a}_p(m - k) \quad (4.3.8)$$

where $\{f_p(k), k \in F_B(p)\}$ denotes the feedback taps for the p -th stage and $\hat{a}(m)$ is the hard estimate of $a(m)$. For the hybrid time-frequency LST receiver, the only feedback

Figure 4.3: Block diagram of LST-FDE receiver

at stage p consists of the previous data decisions of the p th data streams. The number of interfering signals diminished by 1 at each stage due to the interference cancellation.

The m -th symbol error $\varepsilon_p(m)$ for substream p can be written as:

$$\varepsilon_p(m) = y_p(m) - \sum_{k \in F_B(p)} f_p(k) \hat{a}_p(m-k) - a_p(m) \quad (4.3.9)$$

or:

$$\varepsilon_p(m) = \frac{1}{MI} \sum_{l=0}^{M-1} \mathbf{W}_l^{p*} \mathbf{R}_l \exp\left(\frac{j2\pi lm}{M}\right) - \sum_{k \in F_B(p)} f_p(k) \hat{a}_p(m-k) - a_p(m); \quad (4.3.10)$$

Assuming no error propagation, the MSE per symbol for substream p can be written as:

$$MSE = \frac{1}{M} E \left(\sum_{m=0}^{M-1} \varepsilon_p(m) \varepsilon_p^*(m) \right) \quad (4.3.11)$$

Figure 4.4: Block diagram of MISO-DFE receiver

Using Parseval's theorem, (4.3.11) can be expressed as:

$$MSE = \frac{1}{M^2} E \left(\sum_{l=0}^{M-1} E_p(l) E_p^*(l) \right) \quad (4.3.12)$$

we have:

$$E_p(l) = \frac{1}{I} \mathbf{W}_l^* \mathbf{R}_l - \left[\sum_{k \in F_B(p)} f_p(k) \exp\left(\frac{-j2\pi kl}{M}\right) + 1 \right] A_p(l) \quad (4.3.13)$$

where the block $\{A_p(l), l = 0, \dots, M-1\}$ is the DFT of the block $\{a_p(m), m = 0, \dots, M-1\}$. Replacing (4.3.13) in (4.3.12) results in

$$MSE = \frac{1}{M^2} E \left(\sum_{l=0}^{M-1} \left[\frac{1}{I} \mathbf{W}_l^{p*} \mathbf{R}_l - F_l^p A_p(l) \right] \left[\frac{1}{I} \mathbf{W}_l^{p*} \mathbf{R}_l - F_l^p A_p(l) \right]^* \right) \quad (4.3.14)$$

where we define:

$$F_l = 1 + \sum_{k \in F_B} f_p(k) \exp\left(\frac{-j2\pi kl}{M}\right) \quad (4.3.15)$$

Taking the ensemble average we get:

$$E(A_i(l)A_j(l)^*) = M\delta(i-j) \quad (4.3.16)$$

$$E(\mathbf{N}(l+k_1M)\mathbf{N}(l+k_2M)^*) = MI\sigma^2\mathbf{I}_N\delta(k_1-k_2) \quad (4.3.17)$$

$$E(\mathbf{N}_l\mathbf{N}_l^*) = MI\sigma^2\mathbf{I}_{NI} \quad (4.3.18)$$

$$E(\mathbf{R}_l\mathbf{A}_p^*(l)) = M\mathbf{H}_l^p \quad (4.3.19)$$

$$E(\mathbf{R}_l\mathbf{R}_l^*) = M\sum_{j=1}^d \mathbf{H}_l^{ij}\mathbf{H}_l^{ij*} + MI\sigma^2\mathbf{I}_{N(I-1)} \quad (4.3.20)$$

Where d is the number of undetected substreams and $\{i_1, \dots, i_d\}$ are the ordered yet-to-be-detected substreams. Using the above equations the MSE can be written as

$$\begin{aligned} MSE &= \frac{1}{M^2} \sum_{l=0}^{M-1} \left(\frac{1}{I^2} \mathbf{W}_l^{p*} \left[M \sum_{j=1}^d \mathbf{H}_l^{ij}\mathbf{H}_l^{ij*} + MI\sigma^2\mathbf{I}_{NI} \right] \mathbf{W}_l^p \right. \\ &\quad \left. - \frac{M}{I} F_l^{p*} \mathbf{W}_l^{p*} \mathbf{H}_l^p - \frac{M}{I} F_l^p \mathbf{H}_l^{p*} \mathbf{W}_l^{p*} + M F_l F_l^* \right) \end{aligned} \quad (4.3.21)$$

The optimum feedforward filter taps $\mathbf{W}_l^p$ can be derived by minimizing (4.3.21), i.e.

$$\frac{\partial MSE}{\partial \mathbf{W}_l^{p*}} = \sum_{l=0}^{M-1} \left(\frac{1}{I^2} \left[M \sum_{j=1}^d \mathbf{H}_l^{ij}\mathbf{H}_l^{ij*} + MI\sigma^2\mathbf{I}_{NI} \right] \mathbf{W}_l^p - \frac{M}{I} F_l^* \mathbf{H}_l^p \right) = 0 \quad (4.3.22)$$

with some simplifications we have:

$$\mathbf{W}_l^p = F_l^{p*} \left[\frac{1}{I} \sum_{j=1}^d \mathbf{H}_l^{ij}\mathbf{H}_l^{ij*} + \sigma^2\mathbf{I}_{NI} \right]^{-1} \mathbf{H}_l^p \quad (4.3.23)$$

Substituting (4.3.23) in (4.3.21), $MMSE$ can be expressed as

$$MMSE = \frac{1}{M} \sum_{l=0}^{M-1} F_l^p F_l^{p*} \left[1 - \frac{1}{I} \mathbf{H}_l^{p*} \left[\frac{1}{I} \sum_{j=1}^d \mathbf{H}_l^{ij}\mathbf{H}_l^{ij*} + \sigma^2\mathbf{I}_{NI} \right]^{-1} \mathbf{H}_l^p \right] \quad (4.3.24)$$

(4.3.24) can be written as

$$MMSE = \frac{1}{M} \sum_{l=0}^{M-1} F_l^p F_l^{p*} \mathbf{e}_p^T [\mathbf{I}_{NI} - \frac{1}{I} \mathbf{H}_l^{p*} \left[\frac{1}{I} \sum_{j=1}^d \mathbf{H}_l^{ij}\mathbf{H}_l^{ij*} + \sigma^2\mathbf{I}_{NI} \right]^{-1} \mathbf{H}_l^p] \mathbf{e}_p \quad (4.3.25)$$

where $\mathbf{e}_p$ denotes the p -th unit vector (i.e. its p -th element is equal to one and the other elements are equal to zero). Applying the matrix inversion lemma to (4.3.25) gives us

$$MMSE = \frac{\sigma^2}{M} \sum_{l=0}^{M-1} F_l^p F_l^{p*} \mathbf{e}_p^T [\sigma^2 \mathbf{I}_d + \frac{1}{I} \mathbf{H}_l^* \mathbf{H}_l]^{-1} \mathbf{e}_p. \quad (4.3.26)$$

where we define $\mathbf{H}_l = [\mathbf{H}_l^{i1}, \dots, \mathbf{H}_l^{id}]$. The MMSE can be written as

$$\begin{aligned} MMSE &= \frac{1}{M} \sum_{l=0}^{M-1} \left(1 + \sum_{k \in F_B(p)} f_p(k) \exp\left(\frac{-j2\pi lk}{M}\right) \right) \left(1 + \sum_{k \in F_B(p)} f_p^*(k) \exp\left(\frac{j2\pi lk}{M}\right) \right) \\ &\quad \mathbf{e}_p^T \left[\mathbf{I}_d + \frac{1}{\sigma^2 I} \mathbf{H}_l^* \mathbf{H}_l \right]^{-1} \mathbf{e}_p \end{aligned} \quad (4.3.27)$$

Minimizing (4.3.27) in f_k^* we get a set of B linear equations for the optimum feedback taps $\{f^p(k); k \in F_B(p)\}$.

$$\begin{aligned} \sum_{k' \in F_B} f_p(k') \sum_{l=0}^{M-1} \exp\left(\frac{j2\pi l(k-k')}{M}\right) \mathbf{e}_p^T \left[\mathbf{I}_d + \frac{1}{\sigma^2 I} \mathbf{H}_l^* \mathbf{H}_l \right]^{-1} \mathbf{e}_p \\ = - \sum_{l=0}^{M-1} \exp\left(\frac{j2\pi lk}{M}\right) \mathbf{e}_p^T \left[\mathbf{I}_d + \frac{1}{\sigma^2 I} \mathbf{H}_l^* \mathbf{H}_l \right]^{-1} \mathbf{e}_p \end{aligned} \quad (4.3.28)$$

(4.3.28) can be expressed in vector form:

$$\mathbf{V}^p \mathbf{f}_p = -\mathbf{v}^p \quad (4.3.29)$$

where:

$$\mathbf{f}_p = [f_{k1}^p, f_{k2}^p, \dots, f_{kB}^p]^T \quad (4.3.30)$$

$$\mathbf{v}^p = [v_{k1}^p, v_{k2}^p, \dots, v_{kB}^p]^T \quad (4.3.31)$$

$$\mathbf{V}^p = \begin{bmatrix} v_0^p & v_{k1-k2}^p & \dots & v_{k1-kB}^p \\ v_{k2-k1}^p & v_0^p & \dots & v_{k2-kB}^p \\ v_{k3-k1}^p & v_{k3-k2}^p & \dots & v_{k3-kB}^p \\ \dots & \dots & \dots & \dots \\ v_{kB-k1}^p & \dots & \dots & v_0^p \end{bmatrix} \quad (4.3.32)$$

and

$$v_k^p = \sum_{l=0}^{M-1} \exp\left(\frac{j2\pi lk}{M}\right) \mathbf{Q}_l^p(1, 1) \quad (4.3.33)$$

$$\mathbf{Q}_l^p = [\sigma^2 \mathbf{I}_d + \frac{1}{I} \mathbf{H}_l^* \mathbf{H}_l]^{-1}. \quad (4.3.34)$$

In appendix A it is shown that MSE can be written as:

$$MMSE = \sigma^2 v_0 + \sigma^2 \sum_{k \in F_B(p)} v_k f_p(k) = \frac{\sigma^2}{M} \sum_{l=0}^{M-1} \mathbf{Q}_l^p(1, 1) F_l^p. \quad (4.3.35)$$

It can be easily shown that (4.3.23) can be written as

$$\mathbf{W}_l^p = F_l^{p*} \sum_i b_l^{i,p} \mathbf{H}_l^p \quad (4.3.36)$$

where $\mathbf{Z}_l \mathbf{b}_l^p = \mathbf{e}_p$ and $\mathbf{Z}_l = [\frac{1}{I} \sum_{j=1}^d \mathbf{H}_l^{ij} \mathbf{H}_l^{ij*} + \sigma^2 \mathbf{I}_{NI}]$, $\mathbf{b}_l^p = [b_l^{i_1,p}, \dots, b_l^{i_d,p}]$. From (4.3.23), one may note that the optimum feedforward filter at each frequency, contains a matched filter to the channels of all substreams followed by F_l^p .

4.3.3 Hybrid Time-Frequency MIMO-DFE Receiver

The MIMO-DFE, shown in Figure 4.5, consists of a frequency domain fractionally spaced feedforward filter and a time domain feedback filter with a temporal span of K_B taps. It is fully connected, where cross-feedbacks are implemented, to feed back all the past decisions from all substreams into the detection of each substream. The use of cross feedback in multi-user systems to couple the decision from other users to the desired user was originally proposed in [35] in a Code-Division-Multiple-Access (CDMA) context and subsequently in [56, 57] in an Space-Division-multiple-Access (SDMA) context.

Figure 4.5: Block diagram of MIMO DFE

In MIMO-DFE all substreams are detected simultaneously, each having a feed-forward filter $w_{pn}(l + kM)$ ($l = 0, \dots, M - 1, k = 0, \dots, I - 1$) at antenna n and a feedback filter $\mathbf{f}_p(k) = [f_{p1}(k), \dots, f_{pP}(k)]^T$ ($p = 1, \dots, P$). With the same definitions and notations from section (4.3.2), the m -th time domain output of the p th substream obtained after decimation by $1/I$ and M -point IFFT operation, is

$$y_p(m) = \frac{1}{MI} \sum_{l=0}^{M-1} \mathbf{W}_l^{p*} \mathbf{R}_l \exp\left(\frac{j2\pi lm}{M}\right). \quad (4.3.37)$$

The soft estimate of m th symbol of the p th substream, can be written as

$$\tilde{a}_p(m) = y_p(m) - \sum_{k \in F_B(p)} \mathbf{f}_p^*(k) \hat{\mathbf{a}}(m - k) \quad (4.3.38)$$

where $\mathbf{a}(m) = [a_1(m), \dots, a_P(m)]^T$, $\hat{\mathbf{a}}$ is the hard estimate of $\mathbf{a}$ and $\{\mathbf{f}_p(k), k \in F_B(p)\}$. ($F_B(p)$ is the set of feedback taps for substream p). The m -th symbol error for

substream p can be written as

$$\varepsilon_p(m) = y_p(m) - \sum_{k \in F_B(p)} \mathbf{f}_p^*(k) \hat{\mathbf{a}}(m - k) - a_p(m) \quad (4.3.39)$$

or:

$$\varepsilon_p(m) = \frac{1}{MI} \sum_{l=0}^{M-1} \mathbf{W}_l^{p*} \mathbf{R}_l \exp\left(\frac{j2\pi lm}{M}\right) - \sum_{k \in F_B(p)} \mathbf{f}_p^*(k) \hat{\mathbf{a}}(m - k) - a_p(m); \quad (4.3.40)$$

Using Parseval's theorem the MSE in frequency domain can be written as

$$MSE = \frac{1}{M^2} E \left(\sum_{l=0}^{M-1} E_p(l) E_p^*(l) \right) \quad (4.3.41)$$

we have:

$$E_p(l) = \frac{1}{I} \mathbf{W}_l^{p*} \mathbf{R}_l - \left[\sum_{k \in F_B} \mathbf{f}_p^*(k) \exp\left(\frac{-j2\pi kl}{M}\right) + \mathbf{e}_p^T \right] \mathbf{A}(l) \quad (4.3.42)$$

where $\mathbf{A}(l) = [A_1(l), \dots, A_P(l)]^T$ is the DFT of the block $\{\mathbf{a}(m), m = 0, \dots, M-1\}$ at frequency tone l and $\mathbf{e}_p$ denotes the p -th unit vector. MSE can be written as:

$$\begin{aligned} MSE &= \frac{1}{M} E \left(\sum_{l=0}^{M-1} |\tilde{a}_p(m) - \hat{a}_p(m)|^2 \right) = \\ &= \frac{1}{M^2} E \left(\sum_{l=0}^{M-1} \left[\frac{1}{I} \mathbf{W}_l^{p*} \mathbf{R}_l - \mathbf{F}_l^{p*} \mathbf{A}(l) \right] \left[\frac{1}{I} \mathbf{W}_l^{p*} \mathbf{R}_l - \mathbf{F}_l^p \mathbf{A}_p(l) \right]^* \right). \end{aligned} \quad (4.3.43)$$

where $\mathbf{F}_l^p = \mathbf{e}_p + \sum_{k \in F_B} \mathbf{f}_p(k) \exp\left(\frac{j2\pi kl}{M}\right)$.

Computing the correlation function we have

$$E(\mathbf{A}(l) \mathbf{A}(l)^*) = M \mathbf{I}_P \quad (4.3.44)$$

$$E(\mathbf{R}_l \mathbf{A}^*(l)) = \sum_{p=1}^P \mathbf{H}_l^p A_p(l) [A_1^*(l) \dots A_P^*(l)] \quad (4.3.45)$$

$$= M \sum_{p=1}^P \mathbf{H}_l^p \mathbf{e}_p^T = M \mathbf{H}_l \quad (4.3.46)$$

Using (4.3.46), (4.3.43) can be simplified as:

$$\begin{aligned} MMSE &= \frac{1}{M^2} \sum_{l=0}^{M-1} \left(\frac{1}{I^2} \mathbf{W}_l^{p*} \left[M \sum_i \mathbf{H}_l^i \mathbf{H}_l^{i*} + MI\sigma^2 \mathbf{I}_{NI} \right] \mathbf{W}_l^p \right. \\ &\quad \left. - \frac{M}{I} \mathbf{F}_l^{p*} \mathbf{W}_l^{p*} \mathbf{H}_l^p - \frac{M}{I} \mathbf{F}_l^p \mathbf{H}_l^{p*} \mathbf{W}_l^{p*} + M \mathbf{F}_l^{p*} \mathbf{F}_l^p \right) \end{aligned} \quad (4.3.47)$$

The optimum feedforward filter $\mathbf{W}_l^p$ can be derived by minimizing $MMSE$,

$$\frac{\partial MMSE}{\partial \mathbf{W}_l^{p*}} = \sum_{l=0}^{M-1} \left(\frac{1}{I^2} \left[M \sum_i \mathbf{H}_l^i \mathbf{H}_l^{i*} + MI\sigma^2 \mathbf{I}_{NI} \right] \mathbf{W}_l^p - \frac{M}{I} \mathbf{F}_l^* \mathbf{H}_l^p \right) = 0 \quad (4.3.48)$$

or:

$$\mathbf{W}_l^p = \left[\frac{1}{I} \sum_{i=1}^P \mathbf{H}_l^i \mathbf{H}_l^{i*} + \sigma^2 \mathbf{I}_{NI} \right]^{-1} \mathbf{H}_l \mathbf{F}_l^p \quad (4.3.49)$$

Substituting (4.3.49) in (4.3.47), we get

$$MMSE = \frac{1}{M} \sum_{l=0}^{M-1} \mathbf{F}_l^{p*} \left[\mathbf{I}_P - \frac{1}{I} \mathbf{H}_l^* \left[\frac{1}{I} \mathbf{H}_l \mathbf{H}_l^* + \sigma^2 \mathbf{I}_{NI} \right]^{-1} \mathbf{H}_l \right] \mathbf{F}_l^p \quad (4.3.50)$$

where we define $\mathbf{H}_l = [\mathbf{H}_l^1, \dots, \mathbf{H}_l^P]$. Using matrix inversion lemma, (4.3.50) can be expressed as

$$MMSE = \frac{\sigma^2}{M} \sum_{l=0}^{M-1} \mathbf{F}_l^{p*} \left[\sigma^2 \mathbf{I}_P + \frac{1}{I} \mathbf{H}_l^* \mathbf{H}_l \right]^{-1} \mathbf{F}_l^p. \quad (4.3.51)$$

In Appendix B it is shown that $MMSE$ can be written as

$$MMSE = \sigma^2 v_0(p, p) + \sigma^2 \sum_{k \in F_B(p)} \mathbf{e}_p^T \mathbf{v}_k \mathbf{f}_p(k) \mathbf{e}_p = \frac{\sigma^2}{M} \sum_{l=0}^{M-1} \mathbf{e}_p^* \mathbf{F}_l^p \mathbf{Q}_l \mathbf{e}_p \quad (4.3.52)$$

where

$$\mathbf{Q}_l = \left[\sigma^2 \mathbf{I}_P + \frac{1}{I} \mathbf{H}_l^* \mathbf{H}_l \right]^{-1}. \quad (4.3.53)$$

The optimum feedback taps which minimize (4.3.51) satisfy a set of $K_B \times P$ linear equations

$$\begin{aligned} &\sum_{l=0}^{M-1} \sum_{k' \in F_B(p)} \exp\left(\frac{j2\pi l(k' - k)}{M}\right) \left[\mathbf{I}_P + \frac{1}{\sigma^2 I} \mathbf{H}_l^* \mathbf{H}_l \right]^{-1} \mathbf{f}_p(k') \\ &= - \sum_{l=0}^{M-1} \exp\left(\frac{j2\pi l k}{M}\right) \left[\mathbf{I}_P + \frac{1}{\sigma^2 I} \mathbf{H}_l^* \mathbf{H}_l \right]^{-1} \mathbf{e}_p \end{aligned} \quad (4.3.54)$$

which can be represented in vector form:

$$\mathbf{V}^p \mathbf{f}_p = -\mathbf{v}^p \quad (4.3.55)$$

where:

$$\mathbf{f}_p = [\mathbf{f}_{k1}^p, \mathbf{f}_{k2}^p, \dots, \mathbf{f}_{kB}^p]^T \quad (4.3.56)$$

$$\mathbf{v}^p = [\mathbf{v}_{k1}^p, \mathbf{v}_{k2}^p, \dots, \mathbf{v}_{kB}^p]^T \mathbf{e}_p \quad (4.3.57)$$

$$\mathbf{V}^p = \begin{bmatrix} \mathbf{v}_0^p & \mathbf{v}_{k1-k2}^p & \dots & \mathbf{v}_{k1-kB}^p \\ \mathbf{v}_{k2-k1}^p & \mathbf{v}_0^p & \dots & \mathbf{v}_{k2-kB}^p \\ \mathbf{v}_{k3-k1}^p & \mathbf{v}_{k3-k2}^p & \dots & \mathbf{v}_{k3-kB}^p \\ \dots & \dots & \dots & \dots \\ \mathbf{v}_{kB-k1}^p & \dots & \dots & \mathbf{v}_0^p \end{bmatrix} \quad (4.3.58)$$

and

$$\mathbf{v}_k^p = \sum_{l=0}^{M-1} \exp(-\frac{j2\pi lk}{M}) \mathbf{Q}_l \quad (4.3.59)$$

4.4 Uncoded Simulation Results

In the previous sections we reviewed the mathematical analysis of the proposed frequency domain receivers. Here we evaluate receiver performance in different scenarios¹. A rigorous comparison of LST-DFE and MIMO-DFE needs extensive simulations, considering the error propagation effects. We consider uncoded average BER (Bit Error Rate) performance and perfect channel estimation. Each data block consists of M=64 QPSK data symbols, plus CP. In our simulations up to 35000 data blocks were transmitted. We consider a quasi-static channel which is time-invariant within the transmission of a block. For each received block the equalizer coefficients are

¹For simulations the equalizer coefficients are computed from the equations derived in section 4.3 assuming perfect channel estimation. The MMSE expressions are not used in the simulations.

computed from (4.3.23), (4.3.29), (4.3.49) and (4.3.55). The SNR is defined as the average received SNR at each antenna. The symbol duration is $T = 0.25\mu s$ and the channel models used in these simulations are based on the following power delay profiles:

- 1) Six-Ray profile: The six rays have equal power, fade independently and uniformly spaced by interval T (the symbol period). The normalized rms delay spread τ_{rms}/T for this channel is equal to 1.7078, where τ_{rms} is the rms delay spread. This channel model was used in GSM [55].
- 2) UMTS-A: This channel corresponds to a hilly terrain environment with heavy tree densities [58]. The channel power delay profile is shown in table 4.1. The channel has 6 independently faded Rayleigh taps with the rms delay spread $0.3704 \mu s$, corresponding to a normalized rms delay spread τ_{rms}/T of 1.4816 (The ISI spans is 11 symbols).

Table 4.1: UMTS-A Channel: Delay Spread Models

	Delay (μs)	Relative Power (dB)
1	0	0
2	0.31	-1
3	0.71	-9
4	1.09	-10
5	1.73	-15
6	2.51	-20

The CP is assumed to be longer than the overall channel impulse response and we used a raised cosine pulse with roll-off factor 0.3 and $I = 2$ samples per symbol. To reduce error propagation and complexity we choose a few sparse feedback taps. For the LST receiver we select B feedback taps corresponding to the largest received

power for each stage. For the MIMO-DFE receiver, they correspond to the largest aggregate received power for all substreams. $B = 0$ corresponds to linear MMSE equalization.

First we evaluate the performance of different receivers for the 6-ray channel. Figure 4.6 shows the BER performances for $P = 4$ transmit antennas and $N = 4$ receive antennas. The first curve corresponds to *MMSE* linear equalization, LST corresponds to the LST receiver of Section 4.3.3 without any feedback filter ($B = 0$) for all substreams, LST-5 taps corresponds to the LST receiver with 5 feedback taps for each substream and, MIMO-DFE 5 taps corresponds to the MIMO-DFE receiver of Section 4.3.4 with 5 feedback taps. As this figure shows, the performances of all of the receivers are much better than that of the linear receiver, with a 2dB difference for the LST receiver at $\text{BER}=10^{-2}$. LST with 5 taps has about 1dB gain over LST at $\text{BER}=10^{-3}$. The performance of MIMO-DFE and LST with 5 taps are similar for low SNR, as SNR increases MIMO-DFE performance is slightly better. This is because of the more powerful feedback for interference cancellation (For this channel the ISI components form interferences are very large). Figure 4.7 shows the BER performances for $P = 4$ transmit and $N = 6$ receive antennas. It can be observed that LST with 5 feedback taps performs slightly better than the MIMO-DFE 5 taps, since the two more antennas provide more degrees of freedom for interference cancellation. The other results are similar to those of Figure 4.6

Next, we evaluate the performance of the different receivers for the UMTS-A channel. As Figure 4.8 shows the performance of LST and MIMO-DFE receivers for this channel are different from 6-rays channel. Both LST and LST with 5 taps

Figure 4.6: BER for 4×4 6-ray channel for different receivers

perform better than the MIMO-DFE with 5 taps per substream signal (really MIMO-DFE has 5P feedback taps) , with about 1 dB gain for LST and about 2dB gain for LST with 5 taps at BER 10^{-4} . It is observed that since the interfering channels (for the detection of the p th substream) do not have large ISI components, we do not need a powerful feedback for ISI cancellation from interference. Similar results were observed for 4×6 systems (Figure 4.9). Figure 4.10 and Figure 4.11 show the effect of different numbers of feedback taps for LST and MIMO-DFE receivers for 6-ray channels with $P = 4$ transmit and $N = 4$ receive antennas, respectively. From Figure 4.11, we can observe that for MIMO-DFE, by increasing the number of feedback taps

Figure 4.7: BER for 4×6 6-ray channel for different receivers

from 1 to 2 we get a 0.5 dB performance improvement at a BER of 10^{-3} and as we increase the number of the feedback taps to 5, the performance improvement increases to 3.8 dB. Figure 4.10 shows the same results for LST receiver. We may note that the performance improvement from increasing the number of feedback taps for LST receiver is less than that of the MIMO-DFE receiver, with 1 dB gain from increasing the number of the feedback taps from 1 to 5 at a BER of 10^{-4} . Once again, it is observed that since the interfering channels for the detection of a particular substream have large ISI components, the more powerful feedback of the MIMO-DFE can cancel the interferers more effectively.

Figure 4.8: BER for 4×4 UMTS-A channel for different receivers

Figure 4.12 shows the performance of different receivers for the UMTS-A channel with symbol duration $T = 0.05\mu s$ and block length $M=512$. With this symbol duration, the ISI span is 50 symbols. This figure shows similar results as in Figure 4.8. The proposed frequency domain receivers can effectively mitigate ISI in channels with large ISI spans.

4.4.1 Simulation Results with Genie Receiver

Figure 4.13 shows the BER performance comparison for LST and MIMO-DFE receiver for 6-ray channel, where both receivers have 5 taps with a "genie" receiver. A

Figure 4.9: BER for 4×6 UMTS-A channel for different receivers

genie receiver is defined as one that has perfect interference and ISI cancellation, i.e., whenever an error occurs it corrects the error (there is no error propagation in genie receiver). For this channel it is observed that both receivers have almost the same sensitivity to error propagation, almost 2dB for high SNR and 1 dB or less for low SNR.

Fig 4.14 shows the same BER performance as in Figure 4.14 but for UMTS-A channel. For this channel the MIMO-DFE has slightly less performance degradation due to error propagation at low SNR. For high SNR both receivers have almost 1 dB performance degradation compared to genie receiver.

Figure 4.10: BER for different feedback taps for LST receiver 6-ray channel

Figure 4.11: BER for different feedback taps for MIMO-DFE receiver

Figure 4.12: BER for 4×4 UMTS-A channel for different receivers and high symbol rate

Figure 4.13: Comparison with genie receiver 6-ray channel

Figure 4.14: Comparison with genie receiver UMTS-A channel

4.5 Comparison to MIMO-OFDM

For SISO systems several authors have compared the performance of SC modulations employing frequency domain equalization with OFDM [42, 49–51]. In the absence of error control coding, frequency domain equalization employing SC modulations can substantially outperform OFDM modulations. Comparing OFDM and SC requires powerful channel coding and interleaving, where it has been shown that they offer similar BER performance [42]. With CSI available at the transmitter adaptive OFDM modulation outperforms the SC alternative. However SC systems with DFE have been shown to have the same performance and capacity of OFDM systems, if error propagation is limited [59]. In this section we compare the performance of the proposed SC receivers with MIMO-OFDM and LST-OFDM shown in Figure 4.15.

The convolution of the channel and the transmitted symbol can be considered as a convolutional encoder, so the optimum ML decoding algorithm can be achieved by jointly decoding the channel and the convolutional code. However, the complexity increases exponentially in the sum of channel memory and convolutional code constraint length. Suboptimal decoders split the equalization and decoding into two stages. In the first stage, the coded symbols are equalized by linear or DFE receiver to obtain an estimate of the coded symbols. In the equalization stage the coded symbols are assumed to be independent and identically distributed (i.i.d.) Then, these estimates are passed through a Viterbi decoder. We consider 6-ray and UMTS-A channel models defined in Section 4.6. A 64-state rate 1/2 convolutional code with octal generator polynomials (133, 171) is employed. The encoder is assumed to both start and end at the all zero state. Code rate 3/4 is realized by puncturing this standard rate 1/2

code. We select the puncturing matrix as [41]

$$P_u = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad (4.5.1)$$

The coded bits are randomly interleaved within 10 blocks² and mapped into QPSK data symbols using Gray mapping. The QPSK data symbols are transmitted in block of length $M = 128$, plus CP. After processing the received block at the receiver hard decisions are made for FDE feedback loop. The noisy version of transmitted symbols (soft symbols) are passed through a demodulator, metric computation units and metric deinterleaver. Finally, soft symbols are passed to the Viterbi decoder in soft form, that is, before they are quantized and the decisions on the transmitted sequence are taken. Error propagation is considered in simulations. SNR is defined as the average received SNR per information bit at each receive antenna. The SNR definition does not include the effect of cyclic prefix overhead.

Figure 4.16 shows the average BER for 4×4 system using a rate 1/2 code for the 6-ray channel. As this shows for coded systems and at low SNRs the performance of the non-linear receivers is worse than that of the linear receiver. The main reason behind the slight non-linear receiver inferiority in low SNR is its suboptimality in decoding caused by three effects: error propagation, non-Gaussianity of the error due to error propagation and residual interference and ISI, and non-ML decisions in DFE. Also for coded system at BERs below 10^{-2} MIMO-DFE performance is 3 db worse than the LST receiver with 5 taps, while for the uncoded system they have almost the same performance. This is due to the fact that in MIMO-DFE and for the detection of each substream, the non-ML decision from all the other streams

²For each time slot 8-10 data blocks were used in different papers. For instance in [60], each time slot consists of 10 blocks and each block contains 252 symbols.

Figure 4.15: Block Diagram of LST-OFDM

contribute in the detection process. We can observe that both linear MMSE-SC and LST-SC receivers outperform the corresponding OFDM receivers³. The reason is that since in MIMO-OFDM the residual interference and ISI is not Gaussian, the decoding algorithm is suboptimum, so we are not able to take advantage of the modified channel metrics which takes into account the different SNRs in each OFDM subchannel in the optimum ML decoding of OFDM symbols, as it is used in SISO channels.

Figure 4.17 depicts the performance of different receivers for 6-rays channel and code rate 3/4. For rate 3/4 SC exhibits a large gain in SNR over OFDM. The reason is that SC already enjoys frequency diversity and rate 3/4 code is mainly used to deal with residual ISI and additive noise. In OFDM this weak error control code tries to

³The standard Viterbi algorithm was used for the decoding of OFDM symbols.

Figure 4.16: Coded BER for 4×4 6-ray channel for different receivers

improve diversity and deals with interference and additive noise. The situation for OFDM-LST is worse, the rate $3/4$ code has to cope with error propagation as well. As in uncoded systems, the MIMO-DFE performance is very close to that of LST. The reason is that the rate $1/2$ code mitigates the error propagation more effectively than the rate $3/4$ code for each layer, so the detection of the next layer can enjoy more diversity. For both code rates LST receiver with 5 feedback taps has an advantage of about 1 dB in SNR over LST receiver for BER below 10^{-3} .

Next we evaluate the performance of different receivers for UMTS-A channel. Figure 4.18 presents the coded BER for different receivers for rate $1/2$ error-control

Figure 4.17: Coded BER for 4×4 6-ray channel for different receivers

code. We can see that for this channel SC linear MMSE outperforms OFDM at high SNR (2 dB at BER 10^{-3}). For this channel SC-LST has an advantage of about 1 dB in SNR over OFDM-LST. As in uncoded system, we observe that since in this channel ISI components from interference are not very big the MIMO-DFE receiver, which has a powerful feedback, is not very effective compared with LST receivers. This feedback reinforces error propagation and the suboptimality of decoding algorithm. Figure 4.19 shows the same results for rate 3/4. The results for this rate are similar to the 6-rays channel. We can see that the performance of both linear MMSE OFDM OFDM-LST are much worse than their single carrier equivalent. For low SNR MIMO-DFE

performance is better than LST while for high SNR their performance are very close. At low SNRs, the LST receivers suffer more from error propagation. At high SNR which error propagation is limited, the LST receivers enjoy diversity from successive interference cancellation algorithm. Also, we observe that for this channel and for both rates $1/2$ and $3/4$ the performance of LST receiver with 5-taps is worse than the LST in low SNRs. The reason is that since in this channel the taps have small values, the feedback reinforces the error propagation in low SNRs and with error propagation the decoding algorithm is not optimal.

Figure 4.20 shows the coded performance of the proposed receivers for 4×4 UMTS-A channel and a rate $1/2$ code for Symbol rate $T = 0.05\mu s$ and block length $M = 512$. With this symbol rate ISI spans 50 symbols duration. These results are similar to the results in Figure 4.16. So, The proposed frequency domain receivers are robust in severe time-dispersive channels with large ISI span.

Figure 4.18: Coded BER for 4×4 UMTS-A channel for different receivers

Figure 4.19: Coded BER for 4×4 UMTS-A channel for different receivers

Figure 4.20: Coded BER for 4×4 UMTS-A channel for different receivers and high symbol rate

4.6 Complexity Analysis

The signal processing complexity of the proposed receivers and their time-domain equivalent in the number of complex multiplications per block are given in Table (4.2). The complexity cost of performing interference cancellation for LST archi-

Table 4.2: Signal Processing Complexity

Receiver	FFT/IFFT	Feedforward	Feedback
LST	$(P + 2PN - 1)\frac{MI}{2} \log_2 MI$	$MIPN$	$MP \sum_i B_i$
MIMO-DFE	$NMI \log_2 MI$	$MIPN$	$MP^2 \sum_i B_i$
Linear-MMSE SC	$NMI \log_2 MI$	$MIPN$	-
MIMO-OFDM	$NMI \log_2 MI$	MPN	-
OFDM-LST	$NMI \log_2 MI$	MPN	-

itecture is comparatively small and is not accounted for. The signal processing cost for equivalent time domain equalization for a block of length M is of $O(NPM^2I^2)$ complex multiplications.

A comparison of receiver initialization complexities in the number of complex multiplications is presented in table (4.3). These complexity orders were computed assuming that the computation of the inverse of an $L \times L$ matrix generally requires $O(L^3)$ operations by using classical Gaussian elimination techniques. The initial complexity of time domain LST and MIMO-DFE equalization is in the order $O(\frac{P^5}{2}M^3I^3)$ and $O(P^3M^3I^3)$ respectively [30,36]. The increase in initialization complexity of LST architecture compared to MIMO-DFE is due to the ordering process.

Table 4.3: Initialization Complexity

Receiver	Feedforward	Feedback
Linear MMSE-SC	$O(MPN^3I^3)$	-
LST	$O(\frac{MP^3}{2}N^3I^3)$	$O(PK_B^3)$
MIMO-DFE	$O(MPN^3I^3)$	$O(P^3K_B^3)$
MIMO-OFDM	$O(MPN^3)$	-
OFDM-LST	$O(\frac{MP^3}{2}N^3)$	-

4.7 Summary

In this chapter we introduced two new low complexity receiver architectures for severe time-dispersive MIMO channels. These receivers employ single carrier modulation. The feedforward filter is designed and implemented in the frequency domain, so it has a long length and thanks to FFT implementation it has low complexity. The feed back filters are designed in the time-domain. To reduce error propagation and complexity the feedback taps are limited to a few sparse taps. We evaluated the complexity costs of the proposed receiver. We observed that compared to their time-domain equivalent, they have strong decrease in both equalization and initialization complexity.

Next, the performance of the proposed receiver was evaluated in different scenarios. We compared the performance of these receivers with a genie receiver in which its feedback filter cancels the interference and ISI perfectly.

Finally, we compared the performance of the proposed receiver with their corresponding OFDM receivers for different code rates. We observed that the simulated scenarios, SC outperforms OFDM, with a slightly higher complexity paid for DFE equalization. The non ML decoding and non-Gaussian noise ⁴ (due to residual interference and ISI and error propagation) renders the MIMO-OFDM and MIMO-OFDM-LST architectures inferior to MIMO-SC schemes in low code rates. Due to this property we are not able to take advantage of modified Viterbi decoder in which branch metrics are computed based on the SNR for each sub-channel. We expect MIMO-OFDM to become comparable to MIMO-SC for near optimum/iterative decoding. When the coding rate is low SC has an advantage of up to 1-2 dB for BER below 10^{-2} and for the simulated scenarios over OFDM. In high code rate SC substantially outperforms OFDM.

We observed that for the simulated scenarios non-linear receivers always outperform linear receivers for uncoded systems, while for coded systems non-linear receivers are better for high SNRs, due to hard decision feed back and error propagation phenomena. In uncoded systems and when the channel taps are large the MIMO-DFE which has a more powerful feedback is a better candidate than LST architecture. For low code rate LST architecture can enjoy diversity more than MIMO-DFE. Adding feedback taps to LST architecture can always improve the performance, especially for high code rates and high SNRs.

⁴In the optimum Viterbi decoding algorithm Gaussian noise is assumed.

Chapter 5

Iterative Block DFE (IB-DFE)

In the previous chapter we briefly reviewed the frequency domain equalization methods and introduced two new low complexity MIMO receiver structures. We observed that although the performance of the proposed non-linear receivers are better than that of linear receivers, they suffer from error propagation, especially for long feedback filters.

To reduce the error propagation effect, a promising IB-DFE approach for SC transmission was proposed in [61,62], where both the feedforward and feedback filters are implemented in frequency domain. Since the feedback loop takes into account not just the hard-decisions for each block, but also the overall block reliability, the error propagation is significantly reduced. Consequently, the IB-DFE techniques offer much better performance than the non-iterative methods [61,62]. These IB-DFE techniques can be regarded as low complexity turbo equalization [63], since the feedback loop uses the equalizer outputs instead of the channel decoder outputs.

In this section, we extend the IB-DFE principle to MIMO systems and define a

class of frequency domain MIMO receivers (single user system) with iterative interference cancellation. Our approach is to consider each substream as a layer and then to cancel the ISI and the multiuser interference, layer by layer. At each iteration, we use the estimated layers from previous iterations, as well as the corresponding block-level reliability. This receiver can also be used at the BS of an SDMA system for multiuser detection. This chapter was published in [64–66].

5.1 System Model

We consider a single carrier MIMO system with P transmit antenna (or an SDMA system with P users), transmitting independent data blocks, and N receive antennas as depicted in Figure 5.1. The time-domain block transmitted by the p -th antenna is $\{a_p(m); m = 0, 1, \dots, M - 1\}$ with $a_p(m)$ denoting the m -th data symbol of the p -th substream (or users in the SDMA system), which is selected from a given constellation (e.g, a QPSK constellation) under an appropriate mapping rule. A CP preceding each block is used to make the linear convolution associated with the channel equivalent to a cyclic convolution. At the receiver the CP is discarded.

Given that the time-domain block at the n th receive antenna is $\{r_n(m); m = 0, 1, \dots, M - 1\}$, the corresponding frequency-domain block, obtained after an appropriate DFT operation, is $\{R_n(l); l = 0, 1, \dots, M - 1\}$, where

$$R_n(l) = \sum_{p=1}^P A_p(l) H_{pn}(l) + N_n(l), \quad (5.1.1)$$

with block $\{A_p(l); l = 0, 1, \dots, M - 1\}$ denoting the DFT of the block $\{a_p(l); l = 0, 1, \dots, M - 1\}$. $\{H_{pn}(l)\}$ denoting the overall channel frequency response from the

Figure 5.1: System model

p -th transmitting antenna to the n -th receive antenna, for the l -th frequency, and $\{N_n(l); l = 0, 1 \dots, M - 1\}$ denoting the corresponding channel noise.

5.2 Receiver Structure

At the transmitter, the data stream is multiplexed into P blocks, each one with length M , which are simultaneously transmitted by P antennas. A severe time-dispersive channel is considered and SC modulation, with appropriate CP, is employed for each block (layer).

We consider the frequency-domain iterative receiver proposed in [64] which combines IB-DFE principles [61,62] and LST interference cancellation schemes [11]. Each iteration consisting of P detection stages, one for each layer. The interference cancellation algorithm can be done successively or in parallel.

5.2.1 Iterative LST with Successive Interference Cancellation(SIC)

To detect a particular layer, the other layers are considered as interferers. For the first iteration, the interference from previously detected layers is cancelled when detecting a given layer, as with conventional LST receivers. However, contrary to conventional LST receivers, this interference cancellation takes into account the reliability of each of the previously detected layers. For the remaining iterations, we cancel the interference from all users (using the most updated version of each layer), as well as the residual ISI from the layer that is being detected. This means that the proposed receiver structure can be regarded as a serial multiuser detection scheme combating the ISI and interference.

For a given iteration, the receiver structure for the detection of the p th layer is illustrated in Figure 5.2, where we have N frequency-domain feedforward filters (one for each antenna) and P frequency-domain feedback filters (one for each layer). The feedforward filters are designed to minimize both the ISI and interference that cannot be cancelled by the feedback filters, due to decision errors in the previous detection steps. This structure can be regarded as an equalizer capable of combating the ISI and with interference suppression properties. After an IFFT operation, the corresponding time-domain outputs are passed through a decision device so as to estimate the transmitted layer. At the next iteration, these steps are repeated with a-priori knowledge of the estimated layers from the previous detection steps. The proposed receiver requires N FFT operations, one for each receiver antenna, and a pair DFT/IDFT for the detection of each layer, at each iteration. The detection procedure can be summarized as follows:

- First iteration:
 - (1) Detect layer 1.
 - (2) Detect layer 2 by removing the interference from layer 1.
 - (3) Detect layer 3 by removing the interference from layers 2 and 1.
 - (4) Proceed until the detection of layer P .
- Remaining iterations
 - (1) Repeat the detection of layer 1, now removing the interference from all layers obtained from the most recent iteration ($p \neq 1$) and the residual ISI ($p = 1$).
 - (2) Repeat the detection of layer 2, now removing the interference from all layers ($p \neq 2$) and the residual ISI ($p = 2$).
 - (3) Proceed until the detection of layer P .

5.2.2 Iterative LST with Parallel Interference Cancellation (PIC)

The second architecture combines IB-FDE principles with parallel interference cancellation. Each iteration consists of P parallel detection stages. At the first iteration, all of the layers are detected simultaneously. For the remaining iterations, we use the detected layers from the previous iteration to cancel the interference and residual ISI. Although the complexity of this structure is almost the same as the previous one, a parallel design can be advantageous from the implementation point of view. The detection procedure can be summarized as follows:

Figure 5.2: Detection of the p -th layer for a given iteration

- First iteration:

(1) Detect layer ($p = 1, \dots, P$) using linear processing.

- Remaining iterations:

(1) Repeat the detection of each layer ($p = 1, \dots, P$), now using the layers' estimate from iteration $i - 1$ to remove the interference from the other layers ($p' \neq p$) and the residual ISI ($p' = p$).

5.2.3 Derivation of the Equalizer Coefficients

The frequency-domain samples associated with the p -th layer at the output of the equalizer are given by

$$\begin{aligned}\tilde{A}_p(l) &= \sum_{n=1}^N W_{pn}(l)R_n(l) - \sum_{p'=1}^P F_{pp'}(l)\hat{A}_{p'}(l) = \\ &\sum_{n=1}^N W_{pn}(l)R_n(l) - F_{pp}(l)\hat{A}_p(l) - \sum_{p' \neq p} F_{pp'}(l)\hat{A}_{p'}(l)\end{aligned}\quad (5.2.1)$$

where $W_{pn}(l)$ and $F_{pp'}(l)$ ($l = 0, 1, \dots, M-1$) denote the feedforward and feedback coefficients, respectively. It will be shown that the coefficients $F_{pp}(l); l = 0, \dots, M-1$ are used for residual ISI cancellation and the coefficients $F_{pp'}(l); l = 0, \dots, M-1$ ($p' \neq p$) are used for interference cancellation. The frequency domain block $\{\hat{A}_{p'}(l); l = 0, \dots, M-1\}$ is the DFT of the block $\{\hat{a}_{p'}(m); m = 0, \dots, M-1\}$, with $\hat{a}_{p'}(m)$ denoting the hard decision estimates for the p' -th layer. For the SIC receiver, these estimates are associated with i iteration for $p' < p$ and to the $(i-1)$ -th iteration for $p' \geq p$. For PIC receiver, the estimates are associated with $(i-1)$ th iteration for any p (for the first iteration, $\hat{a}_{p'}(m) = 0$). The time-domain estimates, $\hat{a}_p(m)$, can be written as

$$\hat{a}_p(m) = \rho_p a_p(m) + \delta_{k,p}, \quad (5.2.2)$$

The frequency-domain estimates $\hat{A}_p(l)$ can be written as

$$\hat{A}_p(l) = \rho_p A_p(l) + \Delta_{l,p} \quad (5.2.3)$$

where the $\{\Delta_{l,p}, l = 0, \dots, M-1\}$ is the DFT of the block $\{\hat{a}_p(m), m = 0, \dots, M-1\}$.

The correlation coefficient ρ_p is given by

$$\rho_p = \frac{E[\hat{a}_p(m)a_p^*(m)]}{E_{a,p}} = \frac{E[\hat{A}_p(l)A_p^*(l)]}{ME_{a,p}}. \quad (5.2.4)$$

and $\Delta_{l,p}(l)$ denotes a zero-mean error term and $E_{a,p} = E[|a_p(m)|^2]$ denotes the average symbol energy, common to all layers. It is assumed that $E[\Delta_{l,p}A_p^*(l')] \approx 0$, for any l' . Therefore,

$$E[|\Delta_{l,p}|^2] = (1 - \rho_p^2)ME_{a,p}. \quad (5.2.5)$$

By combining (5.1.1), (5.2.1) and (5.2.2), we obtain

$$\begin{aligned} \tilde{A}_p(l) &= \sum_{l=1}^N W_{pn}(l) \left(\sum_{p'=1}^P H_{p'n}(l) A_{p'}(l) + N^n(l) \right) \\ &\quad - F_{pp}(l) (\rho_p A_p(l) + \Delta_{l,p}) - \sum_{p' \neq p} F_{pp'}(l) (\rho_{p'} A_{p'}(l) + \Delta_{l,p'}) \\ &= \gamma_p A_p(l) + \left(\sum_{n=1}^N W_{pn}(l) H_{pn}(l) - \gamma_p - \rho_p F_{pp}(l) \right) A_p(l) \\ &\quad + \sum_{p' \neq p} \left(\sum_{n=1}^N W_{pn}(l) H_{p'n}(l) - \rho_{p'} F_{pp'}(l) \right) A_{p'}(l) \\ &\quad - \sum_{p'=1}^P F_{pp'}(l) \Delta_{l,p'} + \sum_{n=1}^N W_{pn}(l) N^{(n)}(l). \end{aligned}$$

with

$$\gamma_p = \frac{1}{M} \sum_{l=0}^{M-1} \sum_{n=1}^N W_{pn}(l) H_{pn}(l). \quad (5.2.6)$$

Clearly, γ_p can be regarded as the average overall channel frequency response for the p -th user, after combining the outputs of the N feedforward filters. It is also the sample at time zero of the overall forward-equalized impulse response. This means that $\tilde{A}_p(l)$ has five components: a “signal” component, $\gamma_p A_p(l)$, and four “noise” components. The second term in the last equality of (5.2.6) is the residual ISI; the third term is the residual co-channel interference; the fourth term accounts for the errors in $\hat{a}_{p'}(m)$ ($p' = 1, m = 1, \dots, P$); and the final term comes from the channel noise.

The forward and backward coefficients, $\{W_{pn}(l); l = 0, 1, \dots, M-1\}$, $n = 1, \dots, N$, and $\{F_{pp'}(l); l = 0, 1, \dots, M-1\}$, $p' = 1, \dots, P$, respectively, are chosen so as to maximize the "signal-to-interference plus noise ratio" (SINR) for the p -th substream, defined as

$$SINR_p = \frac{|\gamma_p|^2 E_a}{\frac{1}{M} \sum_{m=0}^{M-1} E[|\varepsilon_{m,p}^{eq}|^2]}, \quad (5.2.7)$$

with

$$\varepsilon_{m,p}^{eq} = \tilde{a}_{m,p} - \gamma_p a_p(m) \quad (5.2.8)$$

where $\tilde{a}_{m,p}$ is the soft symbol at the equalizer output. The maximization of SINR is equivalent to the minimization of $MSE = \frac{1}{M} \sum_{m=0}^{M-1} E(|\tilde{a}_p(m) - a_p(m)|^2)$ with the constraint $\gamma_p = 1$. The minimization of the MSE in the time-domain samples is equivalent to the minimization of MSE in the corresponding frequency-domain samples. The MSE in frequency domain can be expressed as:

$$\begin{aligned} MSE &= \frac{1}{M^2} \sum_{l=0}^{M-1} E[|\tilde{A}_p(l) - A_p(l)|^2] = \\ &= \frac{E_a}{M} \left| \sum_{n=1}^N W_{pn}(l) H_{pn}(l) - 1 - \rho_p F_{pp}(l) \right|^2 \\ &\quad + \frac{E_a}{M} \sum_{p' \neq p} \left| \sum_{n=1}^N W_{pn}(l) H_{p'n}(l) - \rho_{p'} F_{pp'}(l) \right|^2 \\ &\quad + \frac{E_a}{M} \sum_{p'=1}^P |F_{pp'}(l)|^2 (1 - \rho_{p'}^2) + \frac{\sigma_N^2}{M} \sum_{n=1}^N |W_{pn}(l)|^2, \end{aligned} \quad (5.2.9)$$

where

$$\sigma_N^2 = \frac{1}{M} E[|N^{(n)}(l)|^2] \quad (5.2.10)$$

denotes the variance of complex Gaussian channel noise. To minimize the MSE

subject to the constraint,

$$\frac{1}{M} \sum_{l=0}^{M-1} \sum_{n=1}^N W_{pn}(l) H_{pn}(l) = 1 \quad (5.2.11)$$

we form the Lagrangian function:

$$L(\mu, W_{pn}^*(l), F_{pp'}^*(l)) = MSE + \mu \left(\frac{1}{M} \sum_{l=0}^{M-1} \sum_{n=1}^N W_{pn}^*(l) H_{pn}^*(l) - 1 \right) \quad (5.2.12)$$

where μ is the Lagrangian multiplier. The optimum feedforward and feedback coefficients $W_{pn}(l)$, $F_{pn}(l)$, $l = 0, \dots, M-1$ must satisfy the following conditions:

$$\begin{aligned} \frac{\partial L(\mu, W_{pn}^*(l), F_{pp'}^*(l))}{\partial W_{pn}^*(l)} &= 0 \\ \frac{\partial L(\mu, W_{pn}^*(l), F_{pp'}^*(l))}{\partial F_{pp'}^*(l)} &= 0 \\ \frac{\partial L(\mu, W_{pn}^*(l), F_{pp'}^*(l))}{\partial F_{pp'}^*(l)} &= 0 \end{aligned} \quad (5.2.13)$$

The above equations can be expressed as:

$$\begin{aligned} \frac{\partial L(\mu, W_{pn}^*(l), F_{pp'}^*(l))}{\partial W_{pn}^*(l)} &= \frac{2E_a}{M} H_{pn}(l)^* \left(\sum_{n'=1}^N W_{pn'}(l) H_{pn'}(l) - 1 - \rho_p F_{pp}(l) \right) \\ &+ \frac{2E_a}{M} \sum_{p' \neq p} H_{p'n}^*(l) \left(\sum_{n'=1}^N W_{pn'}(l) H_{pn'}(l) - \rho_{p'} F_{pp'}(l) \right) \\ &+ \frac{2\sigma_N^2}{M} F_{pn}(l) + \frac{\mu}{M} H_{pn}^*(l) = 0, \quad n = 1, 2, \dots, N, \end{aligned} \quad (5.2.14)$$

$$\begin{aligned} \frac{\partial L(\mu, W_{pn}^*(l), F_{pp'}^*(l))}{\partial F_{pp}^*(n)} &= -\frac{2E_a \rho_p}{M} \left(\sum_{n'=1}^N W_{pn'}(l) H_{pn'}(l) - 1 - \rho_p F_{pp}(l) \right) \\ &+ \frac{2E_a}{M} (1 - \rho_p^2) F_{pp}(l) = 0 \end{aligned} \quad (5.2.15)$$

$$\begin{aligned} \frac{\partial L(\mu, W_{pn}^*(l), F_{pp'}^*(l))}{\partial F_{pp'}^*(l)} &= -\frac{2E_a \rho_{p'}}{M} \left(\sum_{n'=1}^N W_{pn'}(l) H_{pn'}(l) - \rho_{p'} F_{pp'}(l) \right) \\ &+ \frac{2E_a}{M} (1 - \rho_{p'}^2) F_{pp'}(l) = 0, \quad p' \neq p. \end{aligned} \quad (5.2.16)$$

These equations can be rewritten in the following form

$$\begin{aligned}
H_{pn}^*(l) & \left(\sum_{n'=1}^N W_{pn'}(l) H_{pn'}(l) - 1 - \frac{\mu}{E_a} - \rho_p F_{pp}(n) \right) \\
& + \sum_{p' \neq p} H_{p'n}^*(l) \left(\sum_{n'=1}^N W_{pn'}(l) H_{p'n'}(l) - \rho_{p'} F_{pp'}(l) \right) \\
& + \frac{W_{pn}(l)}{SNR} = 0, \quad n = 1, 2, \dots, N,
\end{aligned} \tag{5.2.17}$$

$$\begin{aligned}
\rho_p & \left(\sum_{n'=1}^N W_{pn'}(l) H_{pn'}(l) - 1 - \rho_p F_{pp}(l) \right) \\
& = (1 - \rho_p^2) F_{pp}(l)
\end{aligned} \tag{5.2.18}$$

and

$$\begin{aligned}
\rho_{p'} & \left(\sum_{n'=1}^N W_{pn'}(l) H_{p'n'}(l) - \rho_{p'} F_{pp'}(l) \right) \\
& = (1 - \rho_{p'}^2) F_{pp'}(l), \quad p' \neq p,
\end{aligned} \tag{5.2.19}$$

where $SNR = \frac{E_p}{\sigma_N^2}$. From (5.2.18), the optimum values of $F_{pp}(l)$ are

$$F_{pp}(l) = \rho_p \left(\sum_{n'=1}^N W_{pn'}(l) H_{pn'}(l) - 1 \right); \tag{5.2.20}$$

from (5.2.19), the optimum values of $F_{pp'}(l)$, $p' \neq p$, are

$$F_{pp'}(l) = \rho_{p'} \sum_{n'=1}^N W_{pn'}(l) H_{p'n'}(l). \tag{5.2.21}$$

By replacing (5.2.20) and (5.2.21) in (5.2.17), we get the set of N equations

$$\begin{aligned}
\sum_{p'=1}^P (1 - \rho_{p'}^2) H_{p'n}^*(l) \sum_{n'=1}^N W_{pn'}(l) H_{p'n'}(l) & + \frac{W_{pn}(l)}{SNR} \\
& = (1 - \rho_p^2 + \frac{\mu}{E_a}) H_{pn}^*(l), \quad n = 1, 2, \dots, N.
\end{aligned} \tag{5.2.22}$$

These feedforward coefficients can be used in (5.2.20) and (5.2.21) for obtaining the feedback coefficients $F_{pp}(l)$ and $F_{pp'}(l)$, $p' \neq p$, respectively. We define $\mathbf{H}_{p'}(l) = [H_{p'1}, \dots, H_{p'N}]$ and $\mathbf{W}_{p'}(l) = [W_{p'1}, \dots, W_{p'N}]$, then the set of equations (5.2.22) can be written as:

$$\left[\sum_{p'=1}^P \mathbf{H}_{p'}^*(l) \mathbf{H}_{p'}(l) + \sigma^2 \mathbf{I} \right] \mathbf{W}_p(l) = (1 - \rho^2 + \frac{\mu}{E_a}) \mathbf{H}_p^*(l) \quad (5.2.23)$$

We define $\mathbf{A}_p(l) = [\sum_{p'=1}^P \mathbf{H}_{p'}^*(l) \mathbf{H}_{p'}(l) + \sigma^2 \mathbf{I}]$ and $\mathbf{B}_p(l) = \mathbf{H}_p(l)^*$, then we have:

$$\mathbf{W}_p(l) = (1 - \rho^2 + \frac{\mu}{E_a}) \mathbf{A}_p^{-1}(l) \mathbf{B}_p(l) \quad (5.2.24)$$

We define $\mathbf{C}_p(l) = \mathbf{A}_p^{-1}(l) \mathbf{B}_p(l)$ then the optimum feedforward coefficients $W_{pn}(l)$, $l = 1, \dots, M-1$ are

$$W_{pn}(l) = (1 - \rho^2 + \frac{\mu}{E_a}) c_{pn}(l), l = 1, \dots, M-1. \quad (5.2.25)$$

where $c_{pn}(l)$ is the n -th element of the $\mathbf{C}_p(l)$. Replacing 5.2.25 in 5.2.11 we get

$$\mu = \frac{M}{\sum_{l=1}^{M-1} \sum_{n=1}^N c_{pn}(l) H_{pn}(l)} \quad (5.2.26)$$

For the special case where $P = 1$, it can be shown by substitution that

$$W_{pn}(l) = \frac{(1 - \rho_p^2 + \frac{\mu}{E_a}) SNR \cdot H_{pn}^*(l)}{1 + SNR(1 - \rho_p^2) \sum_{n'=1}^N |H_{pn'}(l)|^2}, \quad n = 1, 2, \dots, N. \quad (5.2.27)$$

It should be noted that, for the first iteration ($i = 0$), we do not have any information about $A_{p'}(l)$ for $p' \leq p$ and the corresponding correlation coefficient ρ_p is zero, leading to $F_{pp'}(l) = 0$ for $p' \leq p$. After the first iteration, and if the residual BER (Bit Error Rate) is not too high, $\hat{a}_{p'}(m) = a_{p'}(m)$ for most of the data symbols, and $\hat{A}_{p'}(l) \approx A_{p'}(l)$; this means that we can use the feedback coefficients to eliminate a

significant part of the residual intersymbol interference, as well as the residual co-channel interference. Naturally, when $\hat{a}_{p'}(m) = a_{p'}(m)$ for all data symbols, $\hat{A}_{p'}(l) = A_{p'}(l)$ and $\rho_{p'} = 1$, ($p = 1, \dots, P$), leading to

$$W_{pn}(l) = \frac{\mu}{E_a} SNR \cdot H_{pn}^*(l), n = 1, 2, \dots, N, \quad (5.2.28)$$

$$F_{pp}(l) = \left(\frac{\mu}{E_a} SNR \sum_{n'=1}^N |H_{pn'}(l)|^2 - 1 \right) \quad (5.2.29)$$

and

$$F_{pp'}(l) = \frac{\mu}{E_a} SNR \sum_{n'=1}^N |H_{p'n'}(l)|^2, p' \neq p. \quad (5.2.30)$$

which corresponds to the total elimination of the ISI and interference cancellation.

5.2.4 Computation of the Correlation Coefficient

In the following we will show how to estimate the correlation coefficient ρ_p when the transmitted symbols a_m belong to a QPSK constellation (for the sake of simplicity, we omit the dependence with the user number p). In this case,

$$a(m) = a^I(m) + ja^Q(m) = \pm d \pm jd \quad (5.2.31)$$

where $a^I(m) = \text{Re}\{a(m)\}$ and $a^Q(m) = \text{Im}\{a(m)\}$ are the in-phase and quadrature components of $a(m)$, respectively. The minimum Euclidian distance within the constellation is $D = 2d$. In this case, the average symbol energy is

$$E_a = E[|a(m)|^2] = \frac{D^2}{4}. \quad (5.2.32)$$

The time-domain samples at the output of the equalizer are

$$\tilde{a}(m) = \tilde{a}^I(m) + j\tilde{a}^Q(m) = a(m) + \nu(m), \quad (5.2.33)$$

where $\tilde{a}^I(m) = \text{Re}\{\tilde{a}(m)\}$, $\tilde{a}^Q(m) = \text{Im}\{\tilde{a}(m)\}$ and $\nu(m)$ is the overall noise component. $\nu(m)$ is assumed Gaussian with mean 0 and the variance $2\sigma_\nu^2$. We will assume that ν_m is approximately Gaussian-distributed, with $E[\nu(m)] = 0$ and the variance of both $\text{Re}\{\nu_m\}$ and $\text{Im}\{\nu_m\}$ given by σ_ν^2 .

The symbol estimates can be written as

$$\hat{a}(m) = a(m) + \varepsilon_m^I + j\varepsilon_m^Q \quad (5.2.34)$$

where the error coefficients ε_m^I (or ε_m^Q) are zero if there is no error in $a^I(m)$ (or $a^Q(m)$) and $\pm D$ otherwise. This means that ε_m^I and ε_m^Q are random variables with

$$\varepsilon_m^I = \begin{cases} \pm D, & \text{probability } P_{e_m^I} \\ 0, & \text{probability } 1 - P_{e_m^I} \end{cases} \quad (5.2.35)$$

and

$$\varepsilon_m^Q = \begin{cases} \pm D, & \text{probability } P_{e_m^Q} \\ 0, & \text{probability } 1 - P_{e_m^Q} \end{cases} \quad (5.2.36)$$

where $1 - P_{e_m^I}$ and $1 - P_{e_m^Q}$ denote the error probability in $a^I(m)$ and $a^Q(m)$, respectively.

We have

$$E[\hat{a}(m)a^*(m)] = E_a + E[a^I(m)\varepsilon_m^I] + E[a^Q(m)\varepsilon_m^Q] \quad (5.2.37)$$

where

$$E[a^I(m)\varepsilon_m^I] = -\frac{D^2}{2}P_{e_m^I} = -E_a P_{e_m^I} \quad (5.2.38)$$

and

$$E[a^Q(m)\varepsilon_m^Q] = -\frac{D^2}{2}P_{e_m^Q} = -E_a P_{e_m^Q}. \quad (5.2.39)$$

By replacing (5.2.38) and (5.2.39) in (5.2.37) and using (5.2.4), we get

$$\rho = 1 - 2P_e \quad (5.2.40)$$

where $P_e = P_{e_m^I} = P_{e_m^Q}$ denotes the average BER.

Table 5.1: Signal Processing Complexity for the first iteration

Receiver	FFT/IFFT	Equalization
SIC	$(P + N)M/2 \log M$	$MP(2N + P - 1)/2$
PIC	$(P + N)M/2 \log M$	NPM

Table 5.2: Signal Processing Complexity for the Remaining Iterations

Task	Remaining Iterations
FFT/IFFT	$(P + N)M/2 \log M$
Equalization	$MP(N + P)$

5.3 Complexity Comparison between SIC and PIC

Receivers

The equalization signal processing complexity of both SIC and PIC receivers, assuming perfect channel estimation, in terms of the number of complex multiplications per block, is given in Tables 5.1 and 5.2. It can be observed that the SIC structure is more complex for the first iteration. For the remaining iterations, both receivers have the same complexity. The initialization complexity to solve the set of equations (5.2.20), (5.2.21) and (5.2.22) for each iteration is of order $N^3PM + 2NMP$.

5.4 Performance Results

In this section, we evaluate the performance of the proposed receivers using extensive simulations in different scenarios.

We consider uncoded BER performances, under perfect synchronization and channel estimation conditions. Each data block consists of $M = 64$ QPSK data symbols, plus CP. The SNR is defined as the average received SNR at each antenna (The effect of extra power due to cyclic prefix is not considered).

First we evaluate the performance of the proposed receivers for a channel that has independent fading taps at each frequency. This channel corresponds to a severe time-dispersive scenario with rich multipath propagation. Figure (5.3) shows the BER performances for different layers (substreams) and different iterations, for $P = 4$ transmit and $N = 4$ receive antennas, and for a receiver with successive interference cancellation. From this figure, we can observe that, for the first iteration, the layers have very different performances: 7dB from layer 0 to layer 3, at $\text{BER}=10^{-4}$. This difference decreases as we increase the number of iterations, with all layers having almost the same performance after 3 iterations. Moreover, the resulting performance can be very close to the matched filter bound (MFB) after four iterations: the required SNRs for average $\text{BER}=10^{-4}$ are 19.8dB, 12.8dB, 11.7dB, 11.7dB and 11.4dB, for iterations 1 to 4 and for the MFB, respectively. This shows that the proposed receiver is able to eliminate a significant part of the ISI and interlayer interference.

Figure 5.4 shows the average BER performances for the parallel and successive receiver for different iterations. After the first iteration the performance of the parallel receiver is almost 2dB worse than the performance of the successive receiver. After iteration 3 this gap reduces to less than 1dB and after iteration 4 to 0.2 dB. Once again, the BER performances after four iterations are very close to the corresponding MFB for both structures. It should be mentioned that for the parallel receiver, all layers have the same average BER. Next, we evaluate the performance of the systems

Figure 5.3: Layers BER for different iterations, and matched filter bound (MFB) performance for channel with independently fading tap at each frequency, $P = 4$, $N = 4$.

for a more realistic channel. We consider the HYPERLAN/2 power delay profile type C (see Table 5.3) [67], which corresponds to a typical large open space environment for Non-Line-of-Sight (NLOS) conditions, 150ns r.m.s delay spread and a maximum delay 1050ns . The symbol rate is assumed to be 62.5ns, and as usual, it is assumed that the CP is longer than the overall channel impulse response. Figure 5.5 illustrates the BER performance for different layers. Once again, the BER performances for different layers are substantially different for the first iteration (almost 6 dB); however, after four iterations the BER performances for all layers are very similar. For this channel,

Figure 5.4: Average BER for SIC receiver (solid line), and PIC receiver (dashed line), 4th iteration and matched filter bound performance (MFB) for channel with independently fading tap at each frequency, $P = 4$, $N = 4$.

the BER performances after four iterations is still about 1.5dB from the MFB. It was observed that the richer the multipath propagation the closer we can be to the matched filter bound.

Figure 5.6 shows the average BER for both SIC and PIC receivers. Once again, the PIC receiver has worse performance after the first iteration; however, the performance of the SIC and PIC receivers are similar after four iterations.

Table 5.3: Power delay profile type C for HIPERLAN/2

	Delay (ns)	Relative Power (dB)
1	0	-3.3
2	10	-3.6
3	20	-3.9
4	30	-4.2
5	50	-0
6	80	-0.9
7	110	-1.7
8	140	-2.6
9	180	-1.5
10	230	-3.0
11	280	-4.4
12	330	-5.9
13	400	-5.3
14	490	-7.9
15	600	-9.4
16	730	-13.2
17	880	-16.3
18	1050	-21.2

5.4.1 Multiuser Scenarios

In this section, we consider the use of the proposed receiver in an SDMA system where each user has 1 transmit antenna and the base station has a number of antennas N , equal or greater than the number of users P . The data block consists of $M = 64$ QPSK data symbols, plus CP. The channel has uncorrelated Rayleigh fading on different frequencies. We consider uncoded BER performance, under perfect synchronization and channel estimation conditions. The horizontal axis shows the transmitted E_b/N_0 , where E_b is the transmitted energy per bit and $N_0/2$ is the noise power spectral density.

Figure 5.7 shows the average BER performance after four iterations, for different

Figure 5.5: BER for different users (layers) and different iterations, and matched filter bound performance (MFB) for Hiperlan/2 power delay profile C, $P=4$, $N=4$.

N and P . This figure illustrates that the BER performances after four iterations are close to the corresponding MFBs. Let us consider now a scenario where the received powers are not the same for all users, due to a wrong power control and/or different uncoded BER requirements for different users. We will consider $N = 4$ receive antennas and $P = 4$ users, with the received power for users 0 and 1 being 6dB larger than the received powers for users 2 and 3. For a given iteration, the receiver detects first the high-power users and then the low-power ones. Figure 5.8 shows the BER performance for the different users. Once again, the iterative detection procedure allows significant performance gains and after four iterations, the BER

Figure 5.6: Average BER for SIC (solid line), and PIC receiver (dashed line), 4th iteration and matched filter bound performance (MFB) for Hiperlan/2 power delay profile C, $P=4$, $N=4$

performances are similar for users with the same average power. We can also note that the performances of the low power users are asymptotically close to the MFB. However, for the high power users, the performances are about 1.5dB away from the MFB, at $\text{BER}=10^{-4}$. This results from the fact that the BER is much lower for the high-power users, allowing an almost perfect cancellation of their effects on the low power users, which can have performance close to the MFB. However, the higher BERs for the low-power users preclude an appropriate interference cancellation on the high-power users, at least for moderate BERs (see also Figure 5.9, where all BER

Figure 5.7: Average BER for different number of antennas and users ,4th iteration and matched filter bound performance (MFB) for channel with independently fading tap at each frequency

are expressed as a function of the E_b/N_0 of the low power users, and MFB was shifted 6dB shift to the left for the high power users).

5.5 Performance of IB-DFE for the 6-Ray channel

In this section we evaluate the performance of the IB-DFE receiver with SIC and PIC interference cancellation for 6-rays channel used in Chapter 4. As Figure 5.10 shows the results are similar to Hyperlan type-C channel. For the first iteration

Figure 5.8: BER for the high power users (a) and the low power users (b) for different iterations (it.), as well as the corresponding MFBs for channel with independently fading taps at each frequency, $P = 4$, $N = 4$

the performance of PIC receiver is about 2 dB worse than the SIC receiver, while after the 4th iteration they have the same performance, which is 2.2 dB far from the matched filter bound. By comparing between these results and the results in Figure 4.6 of Chapter 4 we observe that after 4th iteration and at probability of error 10^{-3} the IB-DFE receiver has about 3dB advantage in SNR over LST with 5-taps which has the best performance in Figure 4.6. However, this IB-DFE is 4 times more complex than the non-linear receivers of Chapter 4. For channels with large numbers of taps where all the taps have equal power and are faded independently, IB-DFE

Figure 5.9: Average BER for the high power and the low power users for different iterations (it.) for channel with independently fading tap at each frequency, $P = 4$, $N = 4$

approaches matched filter bound for high SNR's [68]. When the number of the taps is small the performance of IB-DFE is realization-dependent and does not conform to its asymptotic performance [68]. The different simulation scenarios of this chapter for channels with different numbers of taps confirm this fact.

Figure 5.10: Average BER for SIC (solid line), and PIC receiver (dashed line), 4th iteration and matched filter bound performance (MFB) for the 6-rays channel, $P = 4$, $N = 4$

5.6 Summary

In this chapter we developed a new iterative FDE layered space-time receiver for single carrier modulation. Since the feedback loop takes into account not just the hard decisions for each block, but also the block reliability, the error propagation is significantly reduced. As a result of efficient FFT/IFFT implementation, this receiver has low computational complexity, which is comparable to the linear and non-linear equalizers introduced in Chapter 4.

Finally, the performance of the proposed receiver was evaluated for different types

of channel delay profile. It was observed that the proposed receiver has very good performance, which can be very close to the matched filter bound for severe time-dispersive channels, after just a few iterations. This makes the proposed receiver suitable for broadband wireless communications systems.

Chapter 6

Joint Transmitter and Receiver Design

In the previous chapters we considered the problem of receiver optimization for a given transmitter in a spatial multiplexing MIMO system. In spatial multiplexing independent substreams are transmitted across multiple antennas to maximize throughput, whereas in diversity systems the same data stream is mapped across space and time to get diversity gain. Both schemes assume no CSI at the transmitter. In some application CSI is also available at the transmitter. The channel can be estimated at the receiver and feedback to the transmitter.

The design of MIMO communication systems based on joint optimization of the transmitter and receiver has attracted much interest due to its potential to improve the performance [69–71]. Joint transmitter and receiver optimization where a power constraint is placed at the output of transmit filter will be a considerably more difficult problem than receiver optimization.

Much of the recent interest in joint transmitter receiver optimization can be linked to the pioneering work of Berger and Tufts [72], who demonstrated the key properties of optimal transmitter design for SISO channel. The principal conclusion of their work was that a jointly optimized system is generally bandlimited to a frequency set i.e. the generalized Nyquist set. They realized that, in general this set is not continuous and extends over more than the first Nyquist interval $[-\frac{1}{2T}, \frac{1}{2T}]$, where T is the symbol rate. Later it was shown in [73], that the optimum transmitter spectrum is always bandlimited under any criterion. The optimization of pre and post filter was first considered in [74, 75] for a MIMO system distorted by additive noise. Salz was the first author who reported joint transmitter receiver optimization in a frequency selective MIMO channel distorted with noise [76]. In his derivation he assumed that the number of outputs is equal to the number of inputs and to simplify the analysis he limited the bandwidth to the first Nyquist interval. However, he mentioned that the general optimization problem with the inclusion of aliasing appears to be extremely complicated to solve. More recently a solution for jointly optimized MIMO system was derived in [69]. In [69] the authors gave a solution when the system is bandlimited to the first Nyquist interval and, when the system is bandlimited to L (integer) Nyquist intervals, they converted the optimization problem to an equivalent optimization problem similar to Nyquist bandlimited case with augmented dimensions. In practice the design of FIR transmitter and receiver structure is desirable due to their better numerical properties and suitability for VLSI and DSP implementation. When the transmitter and receiver filters are FIR a closed form solution for the optimization problem is not available. A random search method to jointly optimize the MSE was proposed in [77]. These results are all in discrete-time domain

where a finite bandwidth is assumed for corresponding continuous-time signals. In none of these works is the frequency support for each substream derived. We derive these optimum frequency supports when in general the MIMO system has no a priori restriction on bandwidth. In this chapter we present a novel analysis for joint transmitter receiver design in a MIMO system. The analytical contribution of this chapter is the derivation of the optimized transmitter and receiver without any a priori restriction on bandwidth. In order to obtain a general solution we do not a priori restrict either the MIMO system or the transmitted signal to be limited to any specific bandwidth. A nice feature of our generalized design is that we are now able to provide a general framework for optimized transmitter and receiver filters with arbitrary bandwidth with the inclusion of aliases.

We summarize the main results of this chapter:

- We first introduce the system model and the optimization criteria assuming the knowledge of CSI and a total power constraint at the transmitter for a fixed channel realization. In order to give a general solution we do not restrict the bandwidth of the MIMO system.
- Next, with the above assumption we prove that the optimum space-time filters are bandlimited to a frequency set, which in general is not continuous. For SISO case this result was reported in [72].
- Next, we derive the optimum bandwidth and the optimum frequency sets for different substreams. An iterative algorithm is proposed to find the optimum frequency sets.
- Next, we show that for a fixed channel realization, the different frequency sets

Figure 6.1: Block Diagram of MIMO system

corresponding to different substreams might be disjoint if two channel eigenmodes are not equal for the same Nyquist translate.

- Next, we show that Nyquist bandlimited or excess bandwidth system have the same transmit and receive filter structure.

This chapter was published in [78, 79].

6.1 System Model

A block diagram of the system is shown in Figure 6.1. The input data stream is demultiplexed to K substreams processed by $P \times K$ dimensional transmit filter matrix $\mathbf{T}(f)$ and outputs M symbols that are launched into a $N \times P$ MIMO channel $\mathbf{H}(f)$. The transmit filter is channel dependent. The received signal is processed by a $K \times N$ dimensional receive filter $\mathbf{G}(f)$ where $P \leq N$ and $K \leq \text{rank}(\mathbf{H}(f))$, and then fed through the decision device. In this chapter, we will focus on *boxed* sections of the

communication system Figure 6.1. We will not consider coding or signal design, but rather focus on the design of $\mathbf{T}(f)$ and $\mathbf{G}(f)$.

Let $\mathbf{x}(t) = \mathbf{t}(t) * \mathbf{h}(t)$, where $\mathbf{t}(t)$ and $\mathbf{h}(t)$ are the time domain response matrices of $\mathbf{T}(f)$ and $\mathbf{H}(f)$ respectively. Then, the continuous-time received signal can be written as:

$$\mathbf{r}(t) = \sum_{k=-\infty}^{\infty} \mathbf{x}(t - kT) \mathbf{a}(k) + \mathbf{n}(t) \quad (6.1.1)$$

where $\mathbf{n} = [n_1(m), \dots, n_N(m)]$ is the noise vector. $\mathbf{a}(m) = [a_1(m), \dots, a_K(m)]$ is the data vector at time mT . From Figure (6.1) the error between the transmitted vector and received vector can be written as:

$$\begin{aligned} \varepsilon(m) &= \mathbf{r}(m) - \mathbf{a}(m) \\ &= \sum_{k=-\infty}^{\infty} \mathbf{x}(mT - kT) \mathbf{a}(k) - \mathbf{a}(m) + \int_{-\infty}^{\infty} \mathbf{g}(mT - \tau) \mathbf{n}(\tau) d\tau \end{aligned} \quad (6.1.2)$$

using (6.1.2) MSE can be written as:

$$\begin{aligned} MSE &= E(|\varepsilon(m)|^2) = \text{tr} E \left(\left[\sum_{k=-\infty}^{\infty} \mathbf{x}(mT - kT) \mathbf{a}(k) - \mathbf{a}(m) + \int_{-\infty}^{\infty} \mathbf{g}(mT - \tau) \mathbf{n}(\tau) d\tau \right] \right. \\ &\quad \left. \left[\sum_{k'=-\infty}^{\infty} \mathbf{x}(mT - k'T) \mathbf{a}(k') - \mathbf{a}(m) + \int_{-\infty}^{\infty} \mathbf{g}(mT - \tau) \mathbf{n}(\tau) d\tau \right]^* \right) \end{aligned} \quad (6.1.3)$$

Using Parseval's theorem, the MSE expression for the detection of $\mathbf{a}(m)$ in the frequency domain can be written as (for simplicity we drop the dependence on frequency):

$$\begin{aligned} MSE &= \text{tr} \left(T \int_{-\frac{1}{2T}}^{\frac{1}{2T}} \left\{ \left[\frac{1}{T} \sum_k \mathbf{G}_k \mathbf{H}_k \mathbf{T}_k - \mathbf{I}_K \right] \left[\frac{1}{T} \sum_k \mathbf{G}_k \mathbf{H}_k \mathbf{T}_k - \mathbf{I}_K \right]^* \right. \right. \\ &\quad \left. \left. + \frac{1}{T} \sum_k \mathbf{G}_k \mathbf{R}_k \mathbf{G}_k^* \right\} df \right) \end{aligned} \quad (6.1.4)$$

where the use of subscript k is defined later, $\mathbf{R}(f)$ is a $N \times N$ noise power spectrum

matrix $\mathbf{R}(f) = \int E(\mathbf{n}(t + \tau)\mathbf{n}^*(\tau)) e^{-j2\pi f t} dt$ ¹ and that data and noise are zero-mean and uncorrelated.

Note that, for simplicity of analysis, we have assumed uncorrelated input sources, normalized to unit power and $K = \text{rank}(\mathbf{H})$. We will show later how to handle the case when $K < \text{rank}(\mathbf{H})$. The case when $K > \text{rank}(\mathbf{H})$ is not possible, since the transmitter knows the channel and it will not transmit more than $\text{rank}(\mathbf{H})$ independent data streams (only $\text{rank}(\mathbf{H})$ substream is detectable). The subscript k denote the k -th Nyquist translate of the given function, i.e. ($Z_0 = Z(f)$)

$$Z_k = Z(f - \frac{k}{T}), \forall k \quad (6.1.5)$$

6.2 Optimization of MIMO Unlimited bandwidth System

Our goal is to design $\mathbf{T}$ and $\mathbf{G}$ matrices to minimize the MSE (6.1.4) subject to the total power constraint

$$\text{tr}(\sum_k \int_{-\frac{1}{2T}}^{\frac{1}{2T}} \mathbf{T}_k \mathbf{T}_k^* df) \leq P_0 \quad (6.2.1)$$

where P_0 is the total power constraint. This optimization problem can be solved by using Lagrange multiplier method. We first form the Lagrangian (μ is the lagrange multiplier)

$$L(\mu, \mathbf{G}, \mathbf{T}) = MSE(T, G) + \mu[\text{tr}(\sum_k \int_{-\frac{1}{2T}}^{\frac{1}{2T}} \mathbf{T}_k \mathbf{T}_k^* df) - P_0] \quad (6.2.2)$$

$$L(\mu, \mathbf{G}, \mathbf{T}) = MSE(T, G) + \mu[\text{tr}(\sum_k \int_{-\frac{1}{2T}}^{\frac{1}{2T}} \mathbf{T}_k \mathbf{T}_k^* df) - P_0] \quad (6.2.2)$$

¹if the noises at the N antennas outputs are independent, and each has power spectrum $N(f)$, then matrix $\mathbf{R}(f) = N(f)\mathbf{I}_{N \times N}$.

Using (6.1.4) in (6.2.2), we can write:

$$\begin{aligned}
L(\mathbf{T}, \mathbf{G}, \mu) = & \operatorname{tr} \left(T \int_{\frac{1}{2T}}^{\frac{1}{2T}} \left\{ \left[\frac{1}{T} \sum_k \mathbf{G}_k \mathbf{H}_k \mathbf{T}_k - \mathbf{I}_K \right] \left[\frac{1}{T} \sum_k \mathbf{G}_k \mathbf{H}_k \mathbf{T}_k - \mathbf{I}_K \right]^* \right. \right. \\
& \left. \left. + \frac{1}{T} \sum_k \mathbf{G}_k \mathbf{R}_k \mathbf{G}_k^* \right\} df \right) + \mu \left[\operatorname{tr} \left(\sum_k \int_{\frac{1}{2T}}^{\frac{1}{2T}} \mathbf{T}_k \mathbf{T}_k^* df \right) - P_0 \right] \quad (6.2.3)
\end{aligned}$$

The optimal $\mathbf{G}$ and $\mathbf{T}$ can be found from the following conditions: $\mathbf{G}$ and $\mathbf{T}$ are optimal if and only if there is a $\mu > 0$ that together with $\mathbf{G}$ and $\mathbf{T}$ satisfy the following conditions [80]:

$$\nabla_G L(\mu, \mathbf{T}, \mathbf{G}) = 0 \quad (6.2.4)$$

$$\nabla_T L(\mu, \mathbf{T}, \mathbf{G}) = 0 \quad (6.2.5)$$

Using a standard variational techniques from calculus of variation and substituting (6.2.3) in (6.2.5), we get the following relations between $\mathbf{T}$ and $\mathbf{G}$, respectively [80]:

$$\frac{1}{T^2} \mathbf{H} \mathbf{F} \sum_k \mathbf{T}_k^* \mathbf{H}_k^* \mathbf{G}_k^* + \frac{1}{T} \mathbf{R} \mathbf{G}^* = \frac{1}{T} \mathbf{H} \mathbf{T} \quad (6.2.6)$$

$$\frac{1}{T^2} \sum_k \mathbf{T}_k^* \mathbf{H}_k^* \mathbf{G}_k^* \mathbf{G} \mathbf{H} + \mu \mathbf{T}^* = \frac{1}{T} \mathbf{G} \mathbf{H} \quad (6.2.7)$$

To obtain the above equation we have used the $\frac{\partial \mathbf{A} \mathbf{X} \mathbf{B}}{\partial \mathbf{X}} = \mathbf{B} \mathbf{A}$ and $\frac{\partial \mathbf{A} \mathbf{X}^* \mathbf{B}}{\partial \mathbf{X}} = 0$ from the theory of matrix derivatives.

6.3 Solution

We now solve the equations (6.2.6) and (6.2.7) to derive the optimum transmitter $\mathbf{F}$ and receiver $\mathbf{G}$. Note that since spectra foldovers are included in (6.2.6) and (6.2.7) the optimization problem is very complicated. In what follows, we show how to

simplify this complex problem. We define the eigenvalue decomposition:

$$\mathbf{H}^* \mathbf{R}^{-1} \mathbf{H} = (\mathbf{V} \quad \tilde{\mathbf{V}}) \begin{pmatrix} \mathbf{\Lambda} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} (\mathbf{V} \quad \tilde{\mathbf{V}})^* \quad (6.3.1)$$

where $\mathbf{V}$ is an $N \times K$ orthogonal matrix which forms a basis for $\mathbf{H}^* \mathbf{R}^{-1} \mathbf{H}$, $\tilde{\mathbf{V}}$ is a $N \times (P - K)$ orthogonal matrix which forms a basis for orthogonal complement of $\mathbf{V}$ (in the space of P dimensional vectors), and $\mathbf{\Lambda}$ is a diagonal matrix that contains the nonzero eigenvalues whose elements are ordered as follows

$$\lambda_1(f) \leq \lambda_2(f) \leq \dots \leq \lambda_K(f). \quad (6.3.2)$$

if for some f , $K < \text{rank}(\mathbf{H})$ we only need to consider the K largest eigenvalues. It is well known that the MMSE error receiver consists of a matched filter followed by a periodic T -spaced tapped delay line $\mathbf{Y}$ [81]; i.e.

$$\mathbf{G}_k = \mathbf{Y} \mathbf{T}_k^* \mathbf{H}_k^* \mathbf{R}_k^{-1} \quad (6.3.3)$$

$$\mathbf{Y} = \mathbf{Y}_0 = \mathbf{Y}_k \quad \forall k \quad (6.3.4)$$

Lemma 1: Without loss of generality (w.l.o.g) the optimum filter $\mathbf{F}$ has the following structure:

$$\mathbf{T} = \mathbf{V} \mathbf{\Phi}. \quad (6.3.5)$$

where $\mathbf{\Phi}$ is a $K \times K$ matrix.

Proof. See Appendix C. □

Figure (6.2) shows the block diagram of the MIMO system with optimum transmitter and receiver.

Lemma 2: $\mathbf{\Phi} \mathbf{\Phi}^*$ is diagonal and w.l.o.g $\mathbf{\Phi}$ is diagonal.

Figure 6.2: Block Diagram of Optimum MIMO System

Proof. See Appendix C. □

The optimum $\mathbf{T}$, has the same structure as the bandlimited case given in [69]. In this case, $\mathbf{Y}$ is also diagonal as will be shown later. If for some frequency f_o , $\text{rank}(\mathbf{H}) < K$, then some of the eigenvalues (diagonal elements of $\mathbf{\Lambda}$) are zero. In this case the *MSE* expression is invariant to the rows of $\mathbf{\Phi}$ and columns of $\mathbf{Y}$ corresponding to those eigenvalues so, w.l.o.g we can set those rows and columns equal to zero and reduce the optimization problem to a lower dimensional problem. For the most widely used channel model $\mathbf{H}(f) = \mathbf{\Phi}_R \mathbf{H}_w(f) \mathbf{\Phi}_T$, where $\mathbf{\Phi}_R$ and $\mathbf{\Phi}_T$ are fixed matrices that introduce correlation among the entries of $\mathbf{H}(f)$, and $\mathbf{H}_w(f)$ is a circularly symmetric Gaussian random matrix with i.i.d entries which has rank K' . So, at every frequency, $\text{rank}(\mathbf{H}(f)) = \text{rank}(\mathbf{H}_w(f)) = K'$ and we do not send more than K' substreams over the channel (only K' signals can be detected).

Lemma 3: The minimum *MSE* expression for a fixed transmitter is equal to:

$$MSE = T - \text{tr} \left(\int_{-\frac{1}{2T}}^{\frac{1}{2T}} \sum_k \mathbf{T}_k^* \mathbf{H}_k^* \mathbf{G}_k^* df \right). \quad (6.3.6)$$

Proof. If we set the first variations of (6.2.3) with respect to $\mathbf{T}_k$ and $\mathbf{G}_k$ equal to

zeros we will get the following equations similar to (6.2.6), (6.2.7)

$$\frac{1}{T^2} \mathbf{H}_k \mathbf{T}_k \sum_{k'} \mathbf{T}_{k'}^* \mathbf{H}_{k'}^* \mathbf{G}_{k'}^* + \frac{1}{T} \mathbf{R}_k \mathbf{G}_k^* = \frac{1}{T} \mathbf{H}_k \mathbf{T}_k \quad (6.3.7)$$

$$\frac{1}{T^2} \sum_{k'} \mathbf{T}_{k'}^* \mathbf{H}_{k'}^* \mathbf{G}_{k'}^* \mathbf{G}_k \mathbf{H}_k + \mu \mathbf{T}_k^* = \frac{1}{T} \mathbf{G}_k \mathbf{H}_k \quad (6.3.8)$$

If we premultiply (6.3.7) by $\mathbf{G}_k$, postmultiply (6.3.5) by $\mathbf{T}_k$ and add up the new equations over k , we obtain the following equations: $\square$

$$\frac{1}{T^2} \sum_k \mathbf{G}_k \mathbf{H}_k \mathbf{T}_k \sum_{k'} \mathbf{T}_{k'}^* \mathbf{H}_{k'}^* \mathbf{G}_{k'}^* + \frac{1}{T} \sum_k \mathbf{G}_k \mathbf{R}_k \mathbf{G}_k^* = \frac{1}{T} \sum_k \mathbf{G}_k \mathbf{H}_k \mathbf{T}_k \quad (6.3.9)$$

$$\frac{1}{T^2} \sum_k \mathbf{T}_k^* \mathbf{H}_k^* \mathbf{G}_k^* \sum_{k'} \mathbf{G}_{k'} \mathbf{H}_{k'} \mathbf{T}_{k'} + \mu \sum_k \mathbf{T}_k^* \mathbf{T}_k = \frac{1}{T} \sum_k \mathbf{G}_k \mathbf{H}_k \mathbf{T}_k. \quad (6.3.10)$$

using (6.1.4) and (6.3.10) we will reach (6.3.6).

Corollary 1: $\sum_k \mathbf{G}_k \mathbf{H}_k \mathbf{T}_k$ is Hermitian.

Proof. The left hand side (LHS) of (6.3.9) is Hermitian, so the right hand side (RHS) must be Hermitian. It verifies 6.3.6 since MSE is real. $\square$

When $\mathbf{G} = \mathbf{Y} \mathbf{T}^* \mathbf{H}^* \mathbf{R}^{-1}$ is substituted in (6.2.6) we obtain the following expression for the periodic function $\mathbf{Y}$:

$$\mathbf{Y}^* = \left[\frac{1}{T} \sum_k \mathbf{T}_k^* \mathbf{H}_k^* \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{T}_k + \mathbf{I}_K \right]^{-1} \quad (6.3.11)$$

If (6.3.4) and (6.3.11) are used in (6.3.6) then, (6.2.5) is simplified to

$$MSE = tr \left(T \int_{-\frac{1}{2T}}^{\frac{1}{2T}} \mathbf{Y}^* df \right) \quad (6.3.12)$$

Figure 6.3: Block Diagram of equivalent diagonal channel

Substituting (6.3.5) in (6.3.7), we can further write

$$\mathbf{Y}^* = [\frac{1}{T} \sum_k \Phi_k^* \Lambda_k \Phi_k + \mathbf{I}_K]^{-1}. \quad (6.3.13)$$

It is shown in appendix C that

$$\Phi [\frac{1}{T} \Lambda \mathbf{Y}^* \mathbf{Y} - \mu \mathbf{I}] \Phi^* = \mathbf{0}. \quad (6.3.14)$$

From (6.3.13) and since Φ is diagonal $\mathbf{Y}$ must be diagonal. Hence, the optimum transmitter and receiver diagonalize $\mathbf{H}$. Specifically $\mathbf{GHT} = \mathbf{Y} \Phi^* \Lambda \Phi$. The equivalent diagonalized system is shown in Figure 6.3.

Since $\mathbf{Y}$ is a diagonal matrix then (6.3.13) can be simplified to the following equations:

$$y_{ii} = \sqrt{\frac{\mu}{\lambda_i(f)}} \quad (6.3.15)$$

$$\text{or } \phi_{ii} = 0. \quad (6.3.16)$$

We conclude that the optimized system at each frequency must satisfy either one of the above equations. y_{ii} (since $\mathbf{Y}$ is Hermitian, y_{ii} 's are real) and ϕ_{ii} and

λ_i denote the i th diagonal elements of $\mathbf{Y}$, Φ and Λ respectively. We know that y_{ii} must be periodic, but (6.3.15) is generally aperiodic, i.e., if (6.3.15) is satisfied at some frequency f_0 , it cannot be applied at any other frequency of the form $f = f_0 + \frac{k}{T}$, $k \neq 0$. Therefore, we satisfy (6.3.15) at some f_0 , and set $y_{ii}(f_0) = \sqrt{\frac{\mu}{\lambda_i(f_0)}}$. For any other frequency of the form $f \in \{f_0 + \frac{k}{T}, k \neq 0\}$ we must resort to (6.3.16) and set $\phi_{ii}(f_0 + \frac{k}{T}) = 0$. Since $y_{ii}(f)$ is periodic, for any other frequency $f \in \{f_0 + \frac{k}{T}, k \neq 0\}$ we set $y_{ii}(f_0 + \frac{k}{T}) = y_{ii}(f_0)$. Note that the transmitted vector can be written as $\mathbf{Z} = \sum_{i=1}^K \mathbf{V}_i \phi_{ii} A_i(f)$ where $A_i(f)$ is the Fourier transform of $a_i(m)$. Therefore the main conclusion to be drawn from the above discussion is that the matrix filter $\mathbf{V}_i \phi_{ii}$ corresponding to substream a_i is bandlimited to a frequency set with aggregated bandwidth $\frac{1}{T}$, where no two points coincide under translation $\frac{k}{T}$. This result is similar to the result for Single-Input-Single-Output (SISO) systems in [72]. In SISO we have only one frequency set [72] but it is obvious that in MIMO we have K (number of independent substreams that can be transmitted through channel) frequency sets, where each one has an aggregated bandwidth $\frac{1}{T}$. From (6.3.4) and (6.3.5) and since Φ is bandlimited one can infer that $\mathbf{G}$ is bandlimited, therefore the optimum space-time filters (MIMO system) are bandlimited. Another interesting comparison is with the work in [82]. In that work they show how to apply SISO receiver design to Single-Input-Multiple-Output (SIMO). They decompose the L -Branch linear diversity channel to L non-overlapping SISO channels in time domain. In this work we convert MIMO channel to K SISO channels which may overlap in the frequency domain, as will be shown later.

We observed that the optimum bandwidth for a MIMO system is a frequency set. Actually for each substream there are infinite numbers of frequency set to choose

from. For the optimum design we need to choose the frequency sets that minimize the *MSE* expression. In the following sections of this chapter we focus on these optimum frequency sets. We develop an iterative algorithm to find these optimal sets. When we substitute (6.3.4) and (6.3.5) in (6.2.6) and (6.2.7), and considering that Φ is non-zero only at one value of k then equations (6.2.6), (6.2.7) and (6.2.1) simplify to the same form as in the bandlimited system:

$$\frac{1}{T^2} \mathbf{H} \mathbf{T} \mathbf{T}^* \mathbf{H}^* \mathbf{G}^* + \frac{1}{T} \mathbf{R} \mathbf{G}^* = \frac{1}{T} \mathbf{H} \mathbf{T} \quad (6.3.17)$$

$$\frac{1}{T^2} \mathbf{T}^* \mathbf{H}^* \mathbf{G}^* \mathbf{G} \mathbf{H} + \mu \mathbf{T}^* = \frac{1}{T} \mathbf{G} \mathbf{H} \quad (6.3.18)$$

$$\text{tr} \left(\int_B \mathbf{T} \mathbf{T}^* df \right) \leq P_0 \quad (6.3.19)$$

Where B denotes the optimal frequency set. It can be shown that the solution for the above equation is [83]:

$$|\phi_{ii}|^2 = \max \left\{ 0, \sqrt{\frac{1}{\mu \lambda_i(f)}} - \frac{1}{\lambda_i(f)} \right\}. \quad (6.3.20)$$

From (6.3.20) it is clear that if the following inequality is not satisfied at frequency f , the optimal filter set will have $\phi_{ii} = 0$.

$$\mu \leq \lambda_i(f). \quad (6.3.21)$$

6.4 Optimal Frequency Assignments

In the previous Section the theoretical analysis for the jointly optimized system was presented. We showed how to convert the complex optimization problem with inclusion of spectrum foldover to a simple optimization problem which does not involve aliases. In this section we investigate the characteristics of the optimal sets.

Since $\mathbf{Y}$ is diagonal the expression for MSE in (6.3.12) can be written as:

$$MSE = T \int_{-\frac{1}{2T}}^{\frac{1}{2T}} \sum_{i=1}^K y_{ii} df. \quad (6.4.1)$$

The optimum frequency sets must minimize MSE in (6.4.1). The minimization of MSE in (6.4.1) is equal to the minimization of its integrand at every frequency $f \in I_0 = [-\frac{1}{2T}, \frac{1}{2T}]$. The optimum frequency sets corresponding to different substreams may have overlaps or may be disjoint. The following theorem presents the properties of the optimum frequency sets.

Theorem 1: Let $k_{i,opt} = \arg \max_{k_i} \lambda_i(f + k_i/T)$ and $k_{j,opt} = \arg \max_{k_j} \lambda_j(f + k_j/T)$, $\forall i \neq j, f \in I_0$, then if there are no i, j such that $k_i = k_j$, the optimum frequency assignment is frequency division multiplexing (FDM).

Proof. For filter matrix $\mathbf{V}_l \phi_{ll}$ corresponding to substream l , and each $f \in I_0$ there are two options:

- i) Satisfy (6.3.15) at frequency $f + k_l/T$ for some value of k_l , and satisfy (6.3.16) at $f + j/T$ for all $j \neq k_l$.
- ii) Satisfy (6.3.16) for all frequencies $f + j/T$.

Now we show that option (i) is preferable. To show this, we assume that for filter matrix $\mathbf{V}_l \phi_{ll}$ and for some $f \in I_0$, (6.3.16) is satisfied for all frequencies $f + j/T$, then:

$$\sum_{i=1, i \neq l}^K \frac{1}{\phi_{ii}^2 \lambda_i(f + k_i/T) + 1} > \sum_{i=1}^K \frac{1}{\phi_{ii}^2 \lambda_i(f + k_i/T) + 1}. \quad (6.4.2)$$

where LHS and RHS are the integrand of (6.4.1) for options (i) and (ii), respectively. So, if option (ii) is satisfied for some f , MSE at this f will be greater than that for option (i). Therefore, it is better to satisfy (6.3.15) for some k_l . In order to

choose the optimal sets, we minimize y_{ii} at each $f \in I_0$. Therefore for substreams for $i = K, \dots, 1$ we choose $k_{i,opt}$ such that:

$$y_{ii,opt}(f) = \min\{\min_{k_i} \{\sqrt{\frac{\mu}{\lambda_i(f + k_i/T)}}\}, \min_{k_i} \{\sqrt{\frac{\mu}{\lambda_j(f + k_i/T)}} \mid j > i, k_i \notin B_j\}\}. \quad (6.4.3)$$

where B_i is the optimal frequency set for substream i . From (6.3.15), it is obvious that minimizing y_{ii} is equivalent to maximizing λ_i . If there are at least two eigenvalues such that, $k_{i,opt} \neq k_{j,opt}$ then, $\phi_{ii}(f + k_{i,opt}/T) = 0, \forall k \neq k_{i,opt}$ and $\phi_j(f + k_{j,opt}/T) = 0, \forall k \neq k_{j,opt}$. Therefore, substreams i, j are not active at the same frequency but, if $k_{i,opt} = k_{j,opt}$, then in this case $\phi_{ii}(f + k_{i,opt}/T) = \phi_{jj}(f + k_{i,opt}/T) = 0, \forall k \neq k_{i,opt}$. So, substreams i, j are active at frequency $f + k_{i,opt}/T$. This proves the theorem. $\square$

Furthermore, (6.4.1) is simplified to:

$$MSE = T \sum_{i=1}^K \int_{B_i} \sqrt{\frac{\mu}{\lambda_{B_i}(f + k_{i,opt}/T)}} df. \quad (6.4.4)$$

for every $f \in B_{i,\mu} \cup_{i=1}^K B_i = B$.

6.5 Algorithm for Finding the Optimal Sets

In this section we develop an algorithm to find the optimum frequency supports B_i . The optimum sets B_i for each substream can be chosen such that (6.4.3) is minimized. From (6.4.4), in order to minimize (6.4.4), we select $k_{i,opt}$ such that $\lambda_{B_i}(f + k_i/T)$ is maximized. We propose an algorithm similar to the algorithm in [72] to find the optimal sets B_i . The algorithm for finding the optimal sets can be summarized in the following steps:

- Choose a μ .

- For every f in the Nyquist interval $I_0 = [-\frac{1}{2T}, \frac{1}{2T})$, find $k_{i,opt}$ from (6.4.3).
- Define set i corresponding to substream i :

$$B_{i,\mu} = \{f : f = \omega + \frac{k_{i,opt}(\omega)}{T}, \mu < \lambda_{B_i}(\omega + k_{i,opt}(\omega)/T), \omega \in I_0\}. \quad (6.5.1)$$

- For every $f \notin B_{i,\mu}$, we set $\phi_{ii} = 0$.

After the above steps are done for every frequency in the interval I_0 , the transmitted power is computed from (6.3.19). These steps are repeated for different values of μ until the desired MSE is achieved and the power constraint is satisfied. As it was shown before the optimum design makes the channel diagonal and each subchannel has gain $\lambda_i(f)$. Indeed, our algorithm chooses the best part of the channel, i.e. with the highest gain.

In general, if the conditions in Theorem 1 are satisfied, the two sets B_i and $B_j, i \neq j$ are disjoint and we have K disjoint sets. In [84], for two uncoupled users, the authors showed that when the channel satisfies some relatively weak conditions i.e., the channel for user 1 can not be a constant times channel for user 2 on a set of positive measure, FDM is optimal. Also, in [69] by considering simulation results, it was pointed out that the optimum frequency support is nearly FDM but the authors did not provide a rigorous proof. If the channel satisfies the conditions in theorem 1, and considering the MIMO channel as a multiaccess single user channel, then FDM is the optimal signaling design. We may note, in the worst case, i.e. $\bigcap_{i=1}^K B_i = \emptyset$, the jointly optimized system needs an aggregated bandwidth of K/T that is not generally a continuous set. Therefore this is the best performance we can

obtain by transmitter/receiver design, and any bandwidth beyond the measure of the optimal set is surplus. This result is consistent with the result in [85] that in the absence of noise we need K Nyquist zones to completely cancel the interference with a linear equalizer. It is interesting to mention that the optimum transmitter at each frequency allocates the power to only one Nyquist translate per substream. In [86] the authors proved this fact for two uncoupled users which have the same transmitter filter. When B_i 's are disjoint, the optimum transmitter allocates the power to only one unique Nyquist translate and one unique substream. Note that in the algorithm for finding the optimum set we did not restrict the system to be bandlimited to the Nyquist interval I_0 , so the optimum design for the system with excess bandwidth can be obtained in the same way as Nyquist bandlimited systems. The optimal design for a Nyquist bandlimited and excess bandwidth system has the same transmitter and receiver structure as in (6.3.4), (6.3.5), respectively. The frequency domain analysis presented here assumes that the filters have infinite length impulse response (IIR). However, in practice an FIR implementation are more desirable due to better numerical properties ² and suitability for VLSI and DSP implementation. [87].

6.6 Numerical Results

In this section, we provide some numerical examples. We compare the performance of the optimal transmitter with a MIMO system with uncoupled raised cosine transmitter for a wireless channel. We use a Hyperlan type C channel model with the delay

²IIR are more susceptible to problems of finite-length arithmetic, such as noise generated by calculations, and limit cycles

profile presented in Chapter 3. The fading paths between different transmit and receive antennas are assumed to be uncorrelated. This channel corresponds to a typical indoor multipath scenario for WLAN with 150ns average delay spread and 1050ns maximum delay spread. This channel has 18 taps and these taps are uncorrelated complex random variables. We assume a symbol rate 6.667 Msymbols/s ($T=150\text{ns}$). The channel bandwidth was truncated to $[-3/2T, 3/2T]$. The results are averaged over 1000 random channels. The frequency axis is normalized by $\frac{1}{T}$ ($R = I$).

Example 1: Figure 6.4 shows the maximum eigenvalues and optimized bandwidth for one realization of a 2×2 MIMO system. This optimized bandwidth was found from (6.5). As Figure 6.4 shows the optimal bandwidth assignment is $2/T$ and it is discontinuous. $b1$, $b2$ are spectra of eigenmodes corresponding to substreams 1 and 2 respectively, and are determined from (6.4.3). As this figure shows, in this example $b1$ and $b2$ lie on the spectra of maximum eigenmode.

Example 2: We considered a 2×2 MIMO channel where the channel frequency response was derived from the model above. Figure (6.5) shows the MSE (as defined in (6.1.4) for the minimum MSE system ($MSEOPT$). MSE_{RAC} we define as MSE when two uncoupled raised cosine Nyquist filters with 100% excess bandwidth are used. MSE_{SUB} shows MSE when we use the optimal bandwidth and optimal transmitter for each realization of the channel and we distribute the power equally between different substreams through these sets. This suboptimal approach is less complex than the optimum solution and its performance is almost identical to the optimum design. From this figure it is observed that transmitter optimization can significantly improve the performance.

Figure 6.4: Maximum eigenvalue and optimal support for two substreams, b_1 and b_2 are spectra of $\lambda_{\max}$ corresponding to substreams 1 and 2, respectively.

Example 3: Figure (6.6) shows the MSE for the same scenario as in example 2 for a 3×3 MIMO system. The results are similar to those observed in example 2.

Figure 6.5: Comparison of MSE for 2×2 MIMO and different transmitter pulses.

6.7 Summary

In this chapter we considered the general problem of joint transmitter and receiver design for MIMO channel when the system is not bandlimited to Nyquist band. When the bandwidth of the system is more than the first Nyquist bandwidth the spectrum foldover occurs. The transmitter optimization problem with the inclusion of the spectra foldovers becomes very complex. To the best of our knowledge there was no previous general solution for this transmitter optimization with the inclusion of spectrum foldover. We gave a general solution for this optimization problem which considers spectrum foldover and this solution is valid for system without a priori

Figure 6.6: Comparison of MSE for 3×3 MIMO and different transmitter pulses.

restriction on bandwidth (either Nyquist bandlimited or excess bandwidth). We derived the optimum frequency support for each substream and presented an algorithm to find these frequency support sets.

We observed that the optimum frequency support for each substream is bandlimited to a frequency set which is not necessarily a continuous set, where no two points coincide under the translation $\frac{k}{T}$. Also, we observed that the optimum frequency support set is FDM if the channel satisfies certain conditions provided in Theorem 1.

Finally we provided several numerical examples which illustrate the theoretical results.

Chapter 7

Frequency Domain

Pre-Equalization (Precoding)

In Chapter we presented a joint transmitter and receiver optimization with the availability of perfect knowledge of channel at the transmitter. However, the assumption of perfect CSI knowledge at the transmitter is not realistic. Also, The efficient implementation of the the proposed optimal transmitter and receiver is for sever time-dispersive channels is an important issue.

The idea of transmit precoding in the frequency domain for OFDM systems was previously proposed in [88]. In [88], it has been shown that this approach is asymptotically optimum from the information theory point of view. In order to achieve capacity a waterfilling solution was proposed in [83].

In [70,71], optimal linear space-time precoders and decoders for a flat fading channel were derived using a MMSE criterion with the assumption that CSI is available at the transmitter. CSI can be fed back to the transmitter in frequency division

duplex (FDD) systems or can be estimated in the receiver mode in time division duplex (TDD) systems. For frequency selective channels, time-domain block processing techniques were proposed [70, 71]. However, the computational complexity is very high since the size of transmit matrix is dependent on the data block length. FIR precoders assuming a zero-forcing constraint and subject to a power constraint were presented in [89]. FIR precoders are different from block precoders in that they operate on a continuous stream of data and do not require a guard interval between successive blocks. In [90, 91], Tomlinson-Harashima precoding, a non-linear precoding, is proposed for transmission over MIMO channels.

In this chapter we introduce a new low complexity linear precoder (multiple-beamformer) for FDE-SC receivers introduced in Chapter 4 in a spatially correlated severely time-dispersive environment. These precoders are designed using the MMSE criterion subject to a transmit power constraint. The overall equalization task is shared by the linear precoder and by forward and feedback filters at the receiver. The CSI information at the transmitter is limited to second order statistics of channel and noise. At the receiver the instantaneous realization of the channels is assumed to be known. Due to efficient FFT implementation our design has relatively low complexity in severely frequency selective channels, providing a solution to simplify the complexity of traditional transmitter time-domain processing. This chapter also presents an efficient implementation for transmitter and receiver design of Chapter 6.

7.1 System Model

The configuration of the proposed frequency domain transmitter is shown in Figure 7.1 where P antennas are located at the transmitter. At the transmitter, the data

Figure 7.1: Block diagram of frequency domain transmitter

stream is demultiplexed into $K < P$ substreams which are converted from serial to parallel in block of length M and then transformed to frequency domain by an FFT block. The corresponding frequency domain blocks are processed by the block diagonal matrix $\mathbf{T}$ and converted to P substreams (at each frequency). After that, at each transmit antenna, the transmitted block is converted to time domain by an IFFT block. After parallel to serial conversion, a cyclic prefix is appended to the end of each block and then the signals are launched to the channel. The receiver consists of an array of N antennas, as in Chapter 4. The receiver samples the antenna output at the data symbol rate. Also, all substreams occupy the same frequency band.

We consider the same definitions and notation as in Chapter 4, also for simplicity one sample per symbol ($I=1$) is assumed. The receiver block diagram is shown in

Chapter 5. The $MP \times MK$ frequency domain transmit matrix $\mathbf{T}$ is defined as:

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_2 & \ddots & \\ \vdots & \mathbf{0} & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{T}_M \end{bmatrix} \quad (7.1.1)$$

where $\mathbf{T}_i$ is a $P \times K$ matrix. After the FFT processing at the N -antennas of the receiver, the input-output relation for the l -th frequency tone can be expressed in matrix form as

$$\mathbf{R}_l = \mathcal{H}_l \mathbf{T}_l \mathbf{A}_l + \mathbf{N}_l \quad (7.1.2)$$

where $\mathbf{R}_l = [R_l^1, R_l^2, \dots, R_l^N]^T$ and $\mathbf{N}_l = [N_l^1, N_l^2, \dots, N_l^N]^T$ are the length- N frequency domain received vector and noise, whose elements are i.i.d zero-mean, complex with variance $M\sigma^2$ (assuming white noise, independent from antenna to antenna) respectively. $\mathcal{H}_l$ is $N \times P$ frequency domain channel matrix which can be written as

$$\mathcal{H}_l = \begin{bmatrix} H_{11}(l) & \dots & H_{1P}(l) \\ \vdots & \vdots & \vdots \\ H_{N1}(l) & \dots & H_{NP}(l) \end{bmatrix} \quad (7.1.3)$$

and $\mathbf{A}_l = [A_1(l), \dots, A_K(l)]^T$ is the length- K vector of the FFT of data symbols at frequency l . From Chapter 4 the p -th output at time m of the LST receiver for $p = 1, \dots, K$ can be written as:

$$\tilde{a}_p(m) = \frac{1}{M} \sum_{l=0}^{M-1} \mathbf{W}_l^{p*} \mathbf{R}_l \exp\left(\frac{j2\pi lm}{M}\right) - \sum_{k \in F_B(p)} f_p(k) \hat{a}_p(m-k) \quad (7.1.4)$$

and the MSE expression for the p -th antenna is defined as:

$$MSE^{(p)} = \frac{1}{M^2} E\left(\sum_{l=0}^{M-1} [\mathbf{W}_l^{p*} \mathbf{R}_l - F_l^p A_p(l)] [\mathbf{W}_l^{p*} \mathbf{R}_l - F_l^p A_p(l)]^*\right) \quad (7.1.5)$$

for MIMO-DFE we have:

$$\tilde{a}_p(m) = \frac{1}{M} \sum_{l=0}^{M-1} \mathbf{W}_l^{p*} \mathbf{R}_l \exp\left(\frac{j2\pi lm}{M}\right) - \sum_{k \in F_B(p)} \mathbf{f}_p^*(k) \hat{\mathbf{a}}(m-k) \quad (7.1.6)$$

$$MSE^{(p)} = \frac{1}{M^2} E\left(\sum_{l=0}^{M-1} [\mathbf{W}_l^{p*} \mathbf{R}_l - \mathbf{F}_l^{p*} \mathbf{A}(l)] [\mathbf{W}_l^{p*} \mathbf{R}_l - \mathbf{F}_l^p \mathbf{A}_p(l)]^*\right). \quad (7.1.7)$$

where $\mathbf{W}_l^p$ is an N dimensional feedforward filter as defined in Chapter 4. Let us define the following feedback matrix for the LST receiver configuration:

$$\mathbf{F}_l = \begin{bmatrix} F_l^1 & 0 & \dots & 0 \\ 1 & F_l^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 1 & \dots & 1 & F_l^K \end{bmatrix} \quad (7.1.8)$$

where $F_l^p = 1 + \sum_{k \in F_B} f^p(k) \exp\left(\frac{-j2\pi kl}{M}\right)$.

For MIMO-DFE receiver configuration, we let

$$\mathbf{F}_l = \begin{bmatrix} \mathbf{F}_l^{1*} \\ \vdots \\ \mathbf{F}_l^{K*} \end{bmatrix} \quad (7.1.9)$$

where $\mathbf{F}_l^p = \mathbf{e}_p + \sum_{k \in F_B} \mathbf{f}^p(k) \exp\left(\frac{j2\pi kl}{M}\right)$ is a $K \times 1$ vector.

Considering the above definitions, the average $MSE = \frac{1}{K} \sum_{p=1}^K MSE^{(p)}$ per sub-stream for both receiver structures can be expressed in matrix form as

$$MSE = \frac{1}{P \times M^2} \text{tr} \left[E \left[\sum_{l=0}^{M-1} (\mathbf{W}_l \mathbf{R}_l - \mathbf{F}_l \mathbf{A}_l) (\mathbf{W}_l \mathbf{R}_l - \mathbf{F}_l \mathbf{A}_l)^* \right] \right] \quad (7.1.10)$$

where

$$\mathbf{W}_l = \begin{bmatrix} \mathbf{W}_l^{1*} \\ \vdots \\ \mathbf{W}_l^{K*} \end{bmatrix} \quad (7.1.11)$$

Taking the ensemble average, Assuming the K input substreams are independent and have unit variance we have:

$$E\{\mathbf{A}_l \mathbf{A}_l^*\} = M \mathbf{I}_K \quad (7.1.12)$$

$$E\{\mathbf{N}_l \mathbf{N}_l^*\} = M \mathbf{I}_N \sigma^2 \quad (7.1.13)$$

By substituting (7.1.12) and (7.1.13) into (7.1.10), (7.1.10) can be simplified to

$$MSE = \frac{1}{MP} \text{tr} \left[\sum_{l=0}^{M-1} [(\mathbf{W}_l \mathcal{H}_l \mathbf{T}_l - \mathbf{F}_l)(\mathbf{W}_l \mathcal{H}_l \mathbf{T}_l - \mathbf{F}_l)^* + \sigma^2 \mathbf{W}_l \mathbf{W}_l^*] \right] \quad (7.1.14)$$

The optimum $\mathbf{W}_l$ which minimizes (7.1.14) can be found by setting $\frac{\partial MSE}{\partial \mathbf{W}_l} = 0$ ¹, i.e.,

$$\begin{aligned} (\mathbf{W}_l \mathcal{H}_l \mathbf{T}_l - \mathbf{F}_l) \mathbf{T}_l^* \mathcal{H}_l^* + \sigma^2 \mathbf{W}_l &= 0 \\ \mathbf{W}_l [\mathcal{H}_l \mathbf{T}_l \mathbf{T}_l^* \mathcal{H}_l^* + \sigma^2 \mathbf{I}_N] &= \mathbf{F}_l \mathbf{T}_l^* \mathcal{H}_l^* \\ \mathbf{W}_l = \mathbf{F}_l \mathbf{T}_l^* \mathcal{H}_l^* [\mathcal{H}_l \mathbf{T}_l \mathbf{T}_l^* \mathcal{H}_l^* + \sigma^2 \mathbf{I}_N]^{-1}, l = 0, \dots, M-1. \end{aligned} \quad (7.1.15)$$

(7.1.14) can be written as

$$MSE = \frac{1}{PM} \text{tr} \left[\sum_{l=0}^{M-1} \mathbf{W}_l [\mathcal{H}_l \mathbf{T}_l \mathbf{T}_l^* \mathcal{H}_l^* + \sigma^2 \mathbf{I}_K] \mathbf{W}_l^* - \mathbf{W}_l \mathcal{H}_l \mathbf{T}_l \mathbf{F}_l^* - \mathbf{F}_l \mathbf{T}_l^* \mathcal{H}_l^* \mathbf{W}_l^* + \mathbf{F}_l \mathbf{F}_l^* \right] \quad (7.1.16)$$

Substituting (7.1.15) in (7.1.16) MMSE can be written as

$$\begin{aligned} MMSE &= \frac{1}{PM} \text{tr} \left[\sum_{l=0}^{M-1} (\mathbf{F}_l \mathbf{F}_l^* - \mathbf{W}_l \mathcal{H}_l \mathbf{T}_l \mathbf{F}_l^*) \right] \\ &= \frac{1}{PM} \text{tr} \left[\sum_{l=0}^{M-1} (\mathbf{F}_l \mathbf{F}_l^* - \mathbf{F}_l \mathbf{T}_l^* \mathcal{H}_l^* [\mathcal{H}_l \mathbf{T}_l \mathbf{T}_l^* \mathcal{H}_l^* + \sigma^2 \mathbf{I}_K]^{-1} \mathcal{H}_l \mathbf{T}_l \mathbf{F}_l^*) \right] \end{aligned} \quad (7.1.17)$$

¹It can be shown that MSE is a convex function in $\mathbf{W}_l$ [18].

From the matrix inversion lemma,

$$\begin{aligned} MMSE &= \frac{1}{PM} \text{tr} \left[\sum_{l=0}^{M-1} \mathbf{F}_l \left[\mathbf{I}_p + \frac{1}{\sigma^2} \mathbf{T}_l^* \mathcal{H}_l^* \mathcal{H}_l \mathbf{T}_l \right]^{-1} \mathbf{F}_l^* \right] \\ &= \frac{\sigma^2}{PM} \text{tr} \left[\sum_{l=0}^{M-1} \mathbf{F}_l \left[\sigma^2 \mathbf{I}_K + \mathbf{T}_l^* \mathcal{H}_l^* \mathcal{H}_l \mathbf{T}_l \right]^{-1} \mathbf{F}_l^* \right] \end{aligned} \quad (7.1.18)$$

Assuming that the second order statistics of the channel is given at the transmitter the average MMSE over different realization of the channel can be expressed as:

$$MMSE_{avg} = E\{MMSE\} = \frac{\sigma^2}{PM} \text{tr} E \left[\sum_{l=0}^{M-1} \mathbf{F}_l \left[\sigma^2 \mathbf{I}_K + \mathbf{T}_l^* \mathcal{H}_l^* \mathcal{H}_l \mathbf{T}_l \right]^{-1} \mathbf{F}_l^* \right] \quad (7.1.19)$$

To find the optimum transmit matrix $\mathbf{T}$ under the constraint $\text{tr}(\mathbf{T}^* \mathbf{T}) = \sum_{l=0}^{M-1} (\mathbf{T}_l^* \mathbf{T}_l) \leq P_0$ we use Jensen's inequality to find a lower bound for (7.1.19) and then minimize the lower bound.

Lemma 1: Define $\mathbf{Q}_l = [\sigma^2 \mathbf{I}_K + \mathbf{T}_l^* \mathcal{H}_l^* \mathcal{H}_l \mathbf{T}_l]^{-1}$ for $l = 1, \dots, M-1$. Then $\text{tr}(\mathbf{F}_l \mathbf{Q}_l \mathbf{F}_l^*)$ is convex for $\forall \mathbf{Q}_l$.

Proof. To show the convexity of $\mathbf{F}_l \mathbf{Q}_l \mathbf{F}_l^*$, $l = 0, \dots, M-1$ we show that $\mathbf{x} \mathbf{Q}_l^{-1} \mathbf{x}^*$ is convex for any vector $\mathbf{x}$. It is sufficient to show that for any positive definite matrices $\mathbf{Q}_{l1}$, $\mathbf{Q}_{l2}$ and α , $0 \leq \alpha \leq 1$:

$$\mathbf{x} [\alpha \mathbf{Q}_{l1} + (1 - \alpha) \mathbf{Q}_{l2}]^{-1} \mathbf{x}^* \leq \alpha \mathbf{x} \mathbf{Q}_{l1}^{-1} \mathbf{x}^* + (1 - \alpha) \mathbf{x} \mathbf{Q}_{l2}^{-1} \mathbf{x}^* \quad (7.1.20)$$

or equivalently we have:

$$\mathbf{x} [\alpha \mathbf{Q}_{l1}^{-1} + (1 - \alpha) \mathbf{Q}_{l2}^{-1} - [\alpha \mathbf{Q}_{l1} + (1 - \alpha) \mathbf{Q}_{l2}]^{-1}] \mathbf{x}^* \geq 0 \quad (7.1.21)$$

Equivalently we need to show that the following matrix is positive definite:

$$\alpha \mathbf{Q}_{l1}^{-1} + (1 - \alpha) \mathbf{Q}_{l2}^{-1} - [\alpha \mathbf{Q}_{l1} + (1 - \alpha) \mathbf{Q}_{l2}]^{-1} \geq 0. \quad (7.1.22)$$

Since $\mathbf{Q}_{l1}$ and $\mathbf{Q}_{l2}$ are Hermitian and positive definite matrices there is a non-singular matrix $\mathbf{C}$ which diagonalizes both $\mathbf{Q}_{l1}$ and $\mathbf{Q}_{l2}$ and we can write [92] (th. 7.6.6):

$$\mathbf{Q}_{l1} = \mathbf{C}\mathbf{I}_K\mathbf{C}^* \quad (7.1.23)$$

$$\mathbf{Q}_{l2} = \mathbf{C}\mathbf{D}\mathbf{C}^* \quad (7.1.24)$$

$$\mathbf{D} = \text{diag}(d_1, \dots, d_K) ; d_i > 0 \forall i. \quad (7.1.25)$$

Substituting (7.1.23) and (7.1.24) in (7.1.22), we obtain

$$\alpha\mathbf{C}^{-*}\mathbf{I}_K\mathbf{C}^{-1} + (1 - \alpha)\mathbf{C}^{-*}\mathbf{D}^{-1}\mathbf{C}^{-1} - [\alpha\mathbf{C}\mathbf{I}_K\mathbf{C}^* + (1 - \alpha)\mathbf{C}\mathbf{D}\mathbf{C}^*]^{-1} \geq 0 \quad (7.1.26)$$

(7.1.26) can be simplified to:

$$\mathbf{C}^{-*} [\alpha\mathbf{I}_K + (1 - \alpha)\mathbf{D}^{-1} - [\alpha\mathbf{I}_K + (1 - \alpha)\mathbf{D}]^{-1}] \mathbf{C}^{-1} \geq 0 \quad (7.1.27)$$

It follows that we need to show that the following matrix is positive definite

$$\alpha\mathbf{I}_K + (1 - \alpha)\mathbf{D}^{-1} - [\alpha\mathbf{I}_K + (1 - \alpha)\mathbf{D}]^{-1} \geq 0 \quad (7.1.28)$$

(7.1.28) is a diagonal matrix so we need to show that all diagonal elements are non-negative or

$$\alpha + \frac{(1 - \alpha)}{d_i} - \frac{1}{\alpha + (1 - \alpha)d_i} \geq 0 \quad (7.1.29)$$

After some algebraic manipulations

$$\frac{(1 - \alpha)[\alpha + (1 - \alpha)d_i] + \alpha d_i [\alpha + (1 - \alpha)d_i] - d_i}{d_i [\alpha + (1 - \alpha)d_i]} \quad (7.1.30)$$

$$= \frac{\alpha(1 - \alpha)(d_i - 1)^2}{d_i [\alpha + d_i(1 - \alpha)]} \geq 0 \quad (7.1.31)$$

It completes the proof. $\square$

Jensen's inequality states that for a convex function f

$$E[f(x)] \geq f[E(x)] \quad (7.1.32)$$

Using the Jensen's inequality a lower bound for (7.1.19) is

$$MMSE_{avg} \geq \frac{\sigma^2}{PM} \text{tr} \left[\sum_{l=0}^{M-1} \mathbf{F}_l [\sigma^2 \mathbf{I} + \mathbf{T}_l^* E[\mathcal{H}_l^* \mathcal{H}_l] \mathbf{T}_l]^{-1} \mathbf{F}_l^* \right] \quad (7.1.33)$$

To make the minimization problem tractable, the optimum transmit matrix $\mathbf{T}$ can be derived by minimizing (7.1.33) assuming no feedback filter. Then the optimum $\mathbf{T}_l$ can be used to find the optimum feedback filter $\mathbf{F}_l$. Using the Theorem 2 stated in the Appendix C (7.1.33) can be written as

$$MMSE_{avg} \geq \frac{\sigma^2}{M} \sum_{l=0}^{M-1} \frac{1}{[\sigma^2 \mathbf{I}_K + \mathbf{T}_l^* E[\mathcal{H}_l^* \mathcal{H}_l] \mathbf{T}_l]_{ii}} \quad (7.1.34)$$

The lower bound in (7.1.34) is achieved when $[\sigma^2 \mathbf{I}_K + \mathbf{T}_l^* E[\mathcal{H}_l^* \mathcal{H}_l] \mathbf{T}_l]$ is diagonal.

Following the same procedure as in chapter 5 it can be shown that

$$\mathbf{T}_l = \mathbf{V}_l \Phi_l \quad (7.1.35)$$

where the eigenvalue decomposition of $E[\mathcal{H}_l^* \mathcal{H}_l]$ is

$$E[\mathcal{H}_l^* \mathcal{H}_l] = (\mathbf{V}_l \quad \tilde{\mathbf{V}}_l) \begin{pmatrix} \Lambda & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} (\mathbf{V}_l \quad \tilde{\mathbf{V}}_l)^* \quad (7.1.36)$$

$$|\phi_{ii}|^2 = \max \left\{ 0, \frac{\sigma}{\sqrt{\mu \lambda_i}} - \frac{\sigma^2}{\lambda_i} \right\}. \quad (7.1.37)$$

The feedback filters can be derived from (4.3.29) and (4.3.55) considering the equivalent channel matrix $\mathcal{H}_l \mathbf{T}_l$ instead of $\mathcal{H}_l$. In Appendix D it is shown that in this system, the FFT and IFFT cancel each other so, the transmitted signals to the SC signals, and therefore lower peak to average power ratios than comparable OFDM signals.

7.2 Spatial Correlation Model

Assuming that the channel length is L then the time domain received vector can be written as

$$\mathbf{r}(m) = \sum_{v=0}^{L-1} \mathbf{H}_v \mathbf{a}(m-v) \quad (7.2.1)$$

where $\mathbf{H}_v$ represents the $N \times P$ channel matrix tap. In general the spatially correlated channel can be modeled by [93–95]

$$\text{vec}(\mathbf{H}_v) = \mathbf{R}_v^{1/2} \text{vec}(\mathbf{H}_{w,v}). \quad (7.2.2)$$

where $\mathbf{H}_{w,v}$ is the spatially white $N \times P$ MIMO channel whose component are i.i.d complex with mean zero and unit variance. $\mathbf{R}_v$ is $NP \times NP$ covariance matrix defined as:

$$\mathbf{R}_v = E[\text{vec}(\mathbf{H}_v) \text{vec}(\mathbf{H}_v)^*] \quad (7.2.3)$$

Although the model described above is able to represent the correlation effect between the elements of $\mathbf{H}$, a simpler model was derived in [93–95]

$$\mathbf{H}_v = (\mathbf{R}_v^r)^{1/2} \mathbf{H}_w (\mathbf{R}_v^t)^{1/2} \quad (7.2.4)$$

where $\mathbf{R}_v^t$ and $\mathbf{R}_v^r$ are the transmit and receive fade correlation matrices respectively. It is easily seen that the relation between $\mathbf{R}_v$, $\mathbf{R}_v^t$ and $\mathbf{R}_v^r$ is

$$\mathbf{R}_v = \mathbf{R}_v^t \otimes \mathbf{R}_v^r. \quad (7.2.5)$$

The frequency response of the correlated channel matrix can be written as

$$\mathcal{H}_l = \sum_{m=0}^{M-1} (\mathbf{R}_m^t)^{1/2} \mathbf{H}_{w,m} (\mathbf{R}_m^r)^{1/2} \exp\left(\frac{-j2\pi ml}{M}\right) \quad (7.2.6)$$

or in the vector form

$$\text{vec}(\mathcal{H}_l) = \sum_{m=0}^{M-1} \mathbf{R}_m^{1/2} \text{vec}(\mathbf{H}_{w,m}) \exp\left(\frac{-j2\pi ml}{M}\right) \quad (7.2.7)$$

The second order statistics of $\mathcal{H}$ can be computed as follows:

$$\begin{aligned} E[\text{vec}(\mathcal{H}_l) \text{vec}(\mathcal{H}_l)^*] &= \\ &= E \left[\sum_{m=0}^{M-1} \mathbf{R}_m^{1/2} \text{vec}(\mathbf{H}_{w,m}) \exp\left(\frac{-j2\pi ml}{M}\right) \sum_{m'=0}^{M-1} \text{vec}(\mathbf{H}_{w,m'})^* \mathbf{R}_{m'}^{1/2} \exp\left(\frac{j2\pi m'l}{M}\right) \right] \\ &= \sum_{m=0}^{M-1} \mathbf{R}_m^{1/2} \mathbf{R}_m^{1/2} = \sum_{m=0}^{v-1} \mathbf{R}_m \\ &= \sum_{m=0}^{v-1} \mathbf{R}_m^r \otimes \mathbf{R}_m^t = \mathbf{R} \end{aligned} \quad (7.2.8)$$

Since (7.2.8) $\mathcal{H}_l$ does not depend on l , the transmit matrices $\mathbf{T}_l, l = 0, \dots, M-1$ must be equal, i.e.,

$$\mathbf{T}_0 = \mathbf{T}_1 = \dots = \mathbf{T}_{M-1} = \mathbf{V}\Phi \quad (7.2.9)$$

7.3 System Performance

The performance of the proposed pre-equalization method with different receiver structures is evaluated in terms of MSE for $P = 4$ transmit, $N = 4$ receive antennas and $K = 4$. We consider the 6-ray channel model from chapter 4'. In our simulations each data block consists of 64 data symbols, plus CP. The MSE is averaged over 10000 fading channels. We assume that the mobile unit (MU) is surrounded by local scatterers so that fading at the MU antennas are uncorrelated. The BS, however, is located high off the ground so that it is unobstructed and no local scattering occurs. Therefore spatial correlation exists between different antennas at the BS. The following two models are considered for the fading correlation matrix R_t :

- 1) The correlation matrix R_t is given by $[R_t]_{ij} = \rho^{|i-j|}$ where $0 \leq \rho \leq 1$ [96].
- 2) The fading correlation matrix R_t is considered as a function of the mean angle of arrival (DOA) ϕ , beamwidth Δ and antenna spacing, where the probability density function of the angle of arrival is assumed to be uniform [97]:

$$p(\phi_i) = \begin{cases} \frac{1}{2\Delta} & -\Delta + \phi \leq \phi_i \leq \Delta + \phi \\ 0 & \text{elsewhere} \end{cases} \quad (7.3.1)$$

and $[R_t]_{ij}$ is given by [97]:

$$[R_t]_{ij} = J_0(z(i-j)) + 2 \sum_{k=1}^{\infty} J_{2k}(z(i-j)) \cos(2k\phi) \frac{\sin(2k\Delta)}{2k\Delta} \quad (7.3.2)$$

$$- 2 \sum_{k=1}^{\infty} J_{2k+1}(z(i-j)) \sin(2m\phi) \frac{\sin((2k+1)\Delta)}{(2k+1)\Delta} \quad (7.3.3)$$

where $z(k) = 2k\pi \frac{d}{\lambda}$ and J_m 's are Bessel Functions of integer order.

In our simulations the antenna spacing at the transmitter and DOA ϕ are assumed to be half wavelength and 50° respectively. Given the fading correlations the spatially correlated MIMO channel is generated as explained in [98]. Figure 7.2 shows the comparison of the proposed pre-equalization method with the optimal precoders in [70] for $\rho = 0.9$ and correlation model 1. This scenario corresponds to a highly spatial correlated channel. With perfect knowledge of the channel at the transmitter the optimal precoder in [70] obtains 3dB improvement at an $MSE = 10^{-2}$ compared to the case without any pre-processing at the transmitter. The proposed pre-equalization scheme with the partial knowledge of the channel at the transmitter can obtain 1 dB improvement at $MSE = 10^{-2}$. Figure 7.3 shows the same comparison for the correlation model 2 and for $\Delta = 40^\circ$ which corresponds to a partially correlated channel. For this scenario, the performance of the proposed pre-equalization scheme

Figure 7.2: Comparison of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4$, $N = 4$, $K = 4$, $\rho = 0.9$.

is close to the optimum design (almost 1 dB at $MSE = 10^{-2}$). Figures 7.4-7.7 shows the comparison of the proposed pre-equalization scheme with the system without any pre-processing for different degrees of spatial correlation (DOA= 80° and DOA= 50° corresponds to the lowest and highest correlation respectively). For high DOA (on the order of 70°), the improvement due to the proposed pre-processing is slight. For these high angle spreads, the channel has a low correlation so the spatial fading correlation matrix R_t is close to the identity matrix. Therefore, knowing the fading correlation at the transmitter could not significantly improve the performance. In low correlation and high SNR pre-processing can still improve the performance of

Figure 7.3: Comparison of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4$, $N = 4$, $K=4$, DOA= 50°, $\Delta = 40^\circ$.

the LST receiver. These figures show that in low correlation and high SNR the MIMO-DFE with pre-processing and without pre-processing has only a slight improvement compared to the LST architecture. This observation was described in chapter 5. Figures 7.8 and 7.9 show the performance of the proposed pre-processing scheme for high correlations. At MSE= 10^{-1} linear and MIMO-DFE with pre-processing have about 5dB improvement compared to the case without preprocessing. For these scenarios the improvement for the LST receiver is slight. With precoding the different substreams are transmitted with different powers so the ordering scheme in [11] is not optimal for this case. In high correlation, the ordering process tries to decrease the

Figure 7.4: Comparison of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4$, $N = 4$, $K=4$, $\text{DOA} = 50^\circ$, $\Delta = 40^\circ$.

gap between pre-coded and the system without precoding.

Figures 7.10-7.14 demonstrate the performance of the proposed pre-processing scheme for different number of substreams K , $P = 4$ and $N = 4$ for highly correlated fading channel. In Figures 7.10 and 7.11 the correlation matrix R_t has two major eigenvalues and this is why the set of the curve for $K = 2$ and $K = 3$ have much better performance than the set of the curve for $K = 4$. In Figures 7.12 and 7.13 the correlation matrix R_t has 3 non-zero eigenvalues so, the proposed method is not possible to send 3 substreams through the channels. One of these two eigenvalues is small, therefore as it is observed the performance of the set of the curves corresponding

Figure 7.5: Comparison of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4$, $N = 4$, $K=4$, $\text{DOA} = 50^\circ$, $\Delta = 60^\circ$.

to $K = 3$ substream is much worse than $K = 2$.

Figures 7.10 shows the performance of the proposed precoding scheme as a function of DOA. Figure 7.15 shows the performance of the proposed precoding scheme for $\text{DOA} = 0^\circ$ and Beam width= 40° . For this scenario the correlation between transmit antennas is very low so, the knowledge of fading correlation matrix is not very helpful. The results in this figure are very similar to the results in Chapter 4 for uncorrelated 6-ray channels.

Figure 7.6: Comparison of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4$, $N = 4$, $K=4$, $\text{DOA} = 50^\circ$, $\Delta = 70^\circ$.

7.4 Summary

In this chapter we developed a low complexity frequency domain precoder (pre-equalizer) for severely time-dispersive channels using the knowledge of second-order statistics of the channel at the transmitter. The performance of the proposed precoder with different linear and non-linear receiver structures was evaluated. The proposed design is suitable when due to channel variation or delay in feedback, full channel knowledge cannot be fed back to the transmitter but a form of slowly-adapting beam-forming is achievable. This pre-equalization technique has moderate peak-to-average

Figure 7.7: Comparison of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4$, $N = 4$, $K=4$, $\text{DOA} = 50^\circ$, $\Delta = 80^\circ$.

ratio comparable to SC signals.

Figure 7.8: Performance of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4$, $N = 4$, $K=4$, DOA= 50°, $\Delta = 30^\circ$.

Figure 7.9: Performance of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4$, $N = 4$, $K=4$, $\text{DOA} = 50^\circ$, $\Delta = 20^\circ$.

Figure 7.10: Performance of the proposed precoding scheme for different number of substreams , $P = 4, N = 4, K=4, 3, 2$ DOA= 50°, $\Delta = 30^\circ$.

Figure 7.11: Performance of the proposed precoding scheme for different substreams , $P = 4, N = 4, K = 4, 3, 2$ DOA= 50°, $\Delta = 20^\circ$.

Figure 7.12: Performance of the proposed precoding scheme for different number of substreams , $P = 4, N = 4, K=3, 2$ DOA= 50°, $\Delta = 10^\circ$.

Figure 7.13: Performance of the proposed precoding scheme for different number of substreams , $P = 4, N = 4, K=3, 2$ DOA= $50^\circ, \Delta = 5^\circ$.

Figure 7.14: Performance of the proposed precoding scheme for different correlations SNR=25dB, , $P = 4$, $N = 4$, $K=4$, DOA= 50° .

Figure 7.15: Performance of the proposed precoding scheme with the optimal precoding for correlated channel, $P = 4$, $N = 4$, $K=4$, $\text{DOA} = 0^\circ$.

Chapter 8

Optimal Power Allocation for a two transmitter V-BLAST System

In Chapter 3 we reviewed the BLAST architecture and the detection algorithms. In the original BLAST algorithm, the different substreams are transmitted with the same power. The BLAST algorithm suffers from error propagation. The exact analysis of $P \times N$ V-BLAST system is very complicated. In [99], the authors developed a unified analysis for a $2 \times N$ V-BLAST system. In this chapter, we present a power allocation scheme for $2 \times N$ V-BLAST systems. In our approach we assign different energies to the two transmitted substreams to minimize the probability of vector error (a vector means that the detected transmitted vector has at least one error). Assuming perfect channel estimation at the receiver, the transmit energies for different substreams are determined at the receiver and then fed back to the transmitter. This power optimization improves the performance not only by adapting the transmit powers to the channel conditions, but also by restricting the effect of error propagation. We observe that this new algorithm can mitigate the error propagation. Using a different

performance criterion, the authors of [100] and [57] have developed algorithms for power optimization in V-BLAST. In [100], the authors minimize the average BER for every symbol. As the exact expression for average BER is complex, they then use an approximation at high signal-to-noise ratios (SNR). In [101], an approximate solution which minimizes the BER without considering error propagation in the analysis has been derived. In this work, to tackle the problem of power allocation analytically, we use the probability of vector error as the performance measure. We then derive an exact expression for the probability of vector error for $2 \times N$ systems (considering error propagation). Then using numerical methods the transmit powers are optimized through minimizing the probability of vector error. This chapter was published in [102, 103].

8.1 Proposed Scheme for Power Allocation

We consider a system, which consists of P transmitters and N receivers. $\mathbf{a} = (a_1, a_2, \dots, a_P)$ denotes the vector of the transmitted symbols. The corresponding received N -dimensional vector is:

$$\mathbf{r}_1 = \mathbf{H}\mathbf{a} + \mathbf{n} = \sum_{i=1}^P \mathbf{h}_i a_i + \mathbf{n} \quad (8.1.1)$$

where $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_P]$ is the $N \times P$ channel matrix and $\mathbf{n}$ is $N \times 1$ is assumed to be i.i.d complex white Gaussian with zero-mean and variance σ^2 . Each transmitter is sending independent BPSK symbols with energy E_i , $i = 1, \dots, P$ and the total power available at the transmitter is fixed. It is assumed that the channel is known at the receiver and it is flat fading. For a simple case that consists of 2 transmitters, a closed form for the probability of vector error (the transmitted vector is received in error)

is derived. For each transmission, optimal values for E_i 's are found to minimize the probability of vector error (the probability that the transmitted vector has at least one error).

Lemma 1: Suppose $\mathbf{w}_i$ are ZF nulling vectors for successive detection of the i -th substream¹. Define $\nu_i = \mathbf{w}_i^* \mathbf{n}$, $i = 1, \dots, P$. Then ν_P is independent of $\nu_i, i = 1, \dots, P - 1$ ².

Proof. It can be easily seen that $\nu_i, i = 1, \dots, P - 1$ is a complex random variable with mean zero and variance $\|\mathbf{w}_i\|^2 \sigma^2$. Assuming ZF criteria $\mathbf{w}_i$ has to satisfy the following conditions:

$$\mathbf{w}_i^* \mathbf{h}_i = 1, i = 1, \dots, P \quad (8.1.2)$$

$$\mathbf{w}_i^* \mathbf{h}_j = 0, j = i + 1, \dots, P. \quad (8.1.3)$$

In particular, $\mathbf{w}_P$ satisfies:

$$\mathbf{w}_P^* \mathbf{h}_P = 1 \quad (8.1.4)$$

This is the Maximum Ratio Combining (MRC), so the optimum value for $\mathbf{w}_P$ is:

$$\mathbf{w}_P^* = \frac{\mathbf{h}_P^*}{\|\mathbf{h}_P\|^2} \quad (8.1.5)$$

from (8.1.3) and (8.1.5) we get:

$$\mathbf{w}_i^* \mathbf{h}_P = \mathbf{w}_i^* \mathbf{w}_P \|\mathbf{h}_P\|^2 = 0; i = 1, \dots, P - 1 \quad (8.1.6)$$

which results in:

$$\mathbf{w}_i^* \mathbf{w}_P = 0; i \leq P - 1 \quad (8.1.7)$$

¹For simplicity we assume that the detection order is $\{1, 2, \dots, P\}$

²It can be proved that ν_i is independent of $\nu_j, \forall i \neq j$, see Appendix E.

or on the other hand:

$$\begin{aligned} \mathbb{E}\{\nu_i \nu_P^*\} &= \mathbb{E}\{\mathbf{w}_i^* \mathbf{n} \mathbf{n}^* \mathbf{w}_P\} = \mathbf{w}_i^* \mathbb{E}\{\mathbf{n} \mathbf{n}^*\} \mathbf{w}_P \\ &= \mathbf{w}_i^* \sigma^2 \mathbf{I}_P \mathbf{w}_P \end{aligned} \quad (8.1.8)$$

Combining (8.1.7) and (8.1.8) we have $\mathbb{E}\{\nu_i \nu_P^*\} = 0$, and since ν_i and ν_P , are zero-mean random variables they must be independent. $\square$

Let Pe be the probability that a vector symbol has at least one error and the event $e_i = 0$ corresponds to the i -th detection stage being correct. One could write:

$$\begin{aligned} Pe &= 1 - P(e_P = 0, \dots, e_1 = 0) \\ &= 1 - P(e_P = 0 | e_{P-1} = 0, \dots, e_1 = 0) \dots P(e_2 = 0 | e_1 = 0) \\ &\quad P(e_1 = 0) \end{aligned} \quad (8.1.9)$$

Pe is a function of channel parameters and SNR_i $1 \leq i \leq P$, where SNR_i is the SNR of i -th substream. For a fixed channel, SNR_i is proportional to E_i , the energy of i -th substream. The idea is to minimize Pe under the constraint that the total transmit power is fixed. In what follows we focus on $2 \times N$ system since, the analysis for $2 \times N$ is tractable.

8.2 $2 \times N$ system:

Using nulling vector $\mathbf{w}_1$ the decision variable for the detection of the first symbol is:

$$\begin{aligned} y_1 &= \mathbf{w}_1^* r_1 = \mathbf{w}_1^* \mathbf{h}_1 a_1 + \mathbf{w}_1^* \mathbf{h}_2 a_2 + \mathbf{w}_1^* \mathbf{n} \\ &= a_1 + \mathbf{w}_1^* \mathbf{n} = a_1 + \nu_1 \end{aligned} \quad (8.2.1)$$

To derive (8.2.1) we used (8.1.2) and (8.1.3). Probability of error for the first detection assuming symbols are equiprobable and encoded by BPSK³ is:

$$Pe_1 = 1/2P(\hat{a}_1 = \sqrt{E_1}|a_1 = -\sqrt{E_1}) + 1/2P(\hat{a}_1 = -\sqrt{E_1}|a_1 = \sqrt{E_1}) \quad (8.2.2)$$

$$P(\hat{a}_1 = \sqrt{E_1}|a_1 = -\sqrt{E_1}) = P(\hat{a}_1 = -\sqrt{E_1}|a_1 = \sqrt{E_1}) = Q\left(\frac{\sqrt{E_1}}{\sigma\|\mathbf{w}_1\|}\right). \quad (8.2.3)$$

From (8.2.2) and (8.2.3):

$$Pe_1 = Q\left(\frac{\sqrt{E_1}}{\sigma\|\mathbf{w}_1\|}\right). \quad (8.2.4)$$

Similarly, we get the decision variable for the second symbol if the first symbol was detected correctly:

$$y_2 = \mathbf{w}_2^* \mathbf{h}_2 a_2 + \mathbf{w}_2^* n^* = a_2 + \nu_2. \quad (8.2.5)$$

Using y_2 the probability of error for the second detection is:

$$\begin{aligned} Pe_2(e_2 \neq 0|e_1 = 0) &= \\ &= 1/2P(Re(y_2) < 0|e_1 = 0, a_2 = \sqrt{E_1}) + 1/2P(Re(y_2) > 0|e_1 = 0, a_2 = -\sqrt{E_1}) \\ &= 1/2P(Re(\nu_2) < \sqrt{E_1}) + 1/2P(Re(\nu_2) > -\sqrt{E_1}) = Q\left(\frac{\sqrt{E_2}}{\sigma\|\mathbf{w}_1\|}\right) \end{aligned} \quad (8.2.6)$$

where (8.2.6) resulted from Lemma 1. Now we can compute the probability of vector error:

$$\begin{aligned} Pe &= 1 - P(e_2 = 0, e_1 = 0) = 1 - P(e_2 = 0|e_1 = 0)P(e_1 = 0) \\ &= 1 - (1 - P(e_2 \neq 0|e_1 = 0))(1 - P(e_1 \neq 0)) \\ &= 1 - (1 - Q\left(\frac{\sqrt{E_1}}{\sigma\|\mathbf{w}_1\|}\right))(1 - Q\left(\frac{\sqrt{E_2}}{\sigma\|\mathbf{w}_2\|}\right)). \end{aligned} \quad (8.2.7)$$

For $P \times N$ system it is not easy to find a closed form for the probability of error but it might be possible to derive some tight bound.

³For simplicity BPSK is assumed, the same steps can be followed for higher order modulation schemes.

If the total transmit energy is $2E$ and the transmitter knows which symbol is going to be detected first, we assign αE to the symbol to be detected first and $(2 - \alpha)E$ to the second symbol thus (8.2.7) can be written as (the detection ordering assumed the same as optimum ordering for equal energy transmission):

$$Pe = 1 - (1 - Q(\frac{\sqrt{(2 - \alpha)E}}{\sigma\|\mathbf{w}_2\|})) (1 - Q(\frac{\sqrt{\alpha E}}{\sigma\|\mathbf{w}_1\|})) \quad (8.2.8)$$

To find the optimal power allocation we propose to minimize Pe and find the optimum value of α_{min} for a given channel, for which $\mathbf{w}_1$ and $\mathbf{w}_2$ are known. Numerical methods may be used to find α_{min} .

The transmission algorithm is described as:

$$\mathbf{G} = \mathbf{H}^+ \quad (8.2.9)$$

$$K_1 = \underset{j}{\operatorname{argmin}} \|(\mathbf{G})_j\|^2 \quad (8.2.10)$$

$$K_2 = \underset{j}{\operatorname{argmax}} \|(\mathbf{G})_j\|^2 \quad (8.2.11)$$

$$\alpha_{min} = \underset{\alpha}{\operatorname{argmin}} Pe; \quad 0 \leq \alpha_{min} \leq 2 \quad (8.2.12)$$

$$E_1 = \sqrt{\alpha_{min} E} \quad (8.2.13)$$

$$E_2 = \sqrt{2 - \alpha_{min} E} \quad (8.2.14)$$

where $\mathbf{G}$ is the pseudoinverse of $\mathbf{H}$, K_1 is the symbol that is going to be detected first, K_2 is the signal to be detected second and $(\mathbf{G})_j$ is the j -th row of $\mathbf{G}$. A block diagram of this system is shown in Figure 8.1.

Figure 8.1: Block Diagram of New Scheme

8.3 Simulation Results

We used the FMINCON function from Matlab which uses Sequential Quadratic Programming⁴ to find the optimal value $0 \leq \alpha_{\min} \leq 2$ that minimize Pe given by (8.2.8) for each realization of 2×2 and 2×3 flat Rayleigh fading MIMO systems (with variance of each tap $\sigma^2 = 2$). Figure 8.2 shows the performance improvement possible for the 2×2 case by using power allocation method we introduced here. We see that at a probability of error 3×10^{-3} we obtain 2dB improvement. Figure 8.3 shows the same result for 2×3 case and as we see the new scheme gives a 3dB improvement at probability 5×10^{-5} compared to sending the symbols with the same energy all the time ($SNR_{avg} = 15dB$). The performance improvement in Figure 8.2 over 8.3 can be attributed to the increase in receiver diversity order. Our results are obtained by averaging over 10000 channel realizations (For each realization, simulation stopped when a certain number of errors were counted) . Gain improvement is significant at

⁴SQP methods are preferred for small dimensional, highly nonlinear problems, where evaluation of model functions and gradients is expensive.

Figure 8.2: Performance comparison of 2×2 V-BLAST

high SNR but converges to zero as SNR decreases. Figure 8.4 shows bit error rate for the second detection for original V-BLAST, new scheme and V-BLAST with a genie (if an error occurs in the detection of the first symbol the genie corrects it). We see that at BER 2×10^{-3} ($SNR_{avg} = 15dB$) the new scheme performs 2 db better than original V-BLAST but we are still 2db from the ideal case with genie (no error propagation). The horizontal axis in Figure (8.2) and (8.3) is the averaged received SNR, i.e., $1/2(SNR_1 + SNR_2)$ where SNR_i is the ratio of received signal power at the i -th receiver. Implementation of this scheme requires a feedback channel. Via the feedback channel, the receiver sends the transmitter α_{min} .

Figure 8.3: Performance Comparison of 2×3 V-BLAST

8.4 Conclusions

In this chapter we derived the analytical equation for probability of vector error for $2 \times N$ V-BLAST system when signals are transmitted with different powers. Using this equation we found the optimal energy distribution between different sub-channels provided that probability of vector error is minimized. Finally simulation results demonstrate performance gain of the new technique.

Figure 8.4: BER of the second detection 2×2

Chapter 9

Conclusions

Employing multiple antennas at both transmitter and receiver and space-time processing techniques are prime candidates for high quality and high data rate future wireless communications systems. However, transmitter and receiver design for MIMO systems are more challenging and more complex than the SISO systems, since it should effectively mitigate both interference and ISI. ML receiver is the optimum receiver for MIMO channels, however its implementation complexity grows exponentially with the number of transmit antennas and channel length. The suboptimal linear and non-linear receivers are less complex while offering performance comparable to ML receiver. In a number of applications, perfect channel or partial channel knowledge can be available at the transmitter. If the channel variation is slow, the channel can be estimated at the receiver and then fed back to the transmitter. In faster varying channels, the rate of the variations of the channel statistics is usually still slower than the rate of the channel variations, so the channel statistics can be estimated at the receiver and fed back to the transmitter.

In this thesis, we consider frequency domain methods for the design of the low

complexity receiver and transmitter for severe time-dispersive MIMO channels. In addition, we derive a power optimization algorithm for V-BLAST system with two transmit antennas.

First, we introduced two new non-linear receiver structures employing SC modulations, with frequency domain feedforward and time domain feedback filters. Due to the efficient FFT implementation, the proposed receiver can have very low complexity in severely time dispersive channels. To reduce the error propagation effects in these receivers, a small number of feedback taps are selected. The performance of the proposed receivers depends on the channel model. We evaluated the performance of coded and uncoded systems. We observed that designing the decoder and DFE receiver separately is not very promising, due to non-linear effects in the DFE. We compared the performance of the SC modulation with OFDM systems. Our simulation results show that for the 4×4 simulated scenario, coded MIMO-SC outperforms coded MIMO-OFDM. A similar observation has been reported in [104]. For rate-1/2 convolutional coding, depending on the channel type and receiver structure, sometimes OFDM performance approaches SC performance. DFEs suffer from error propagation effects. To reduce error propagation an iterative DFE was introduced in Chapter 5. This iterative method employs frequency domain feedforward and feedback filters for ISI and interference cancellation, which can be implemented efficiently by FFT. The feedback filters at each iteration are optimized based on the reliability function of the detected block in the previous iteration. The interference cancellation can be done in parallel or successively. Simulation results show that the proposed receiver can approach the matched filter bound after few iterations. The complexity of the iterative receiver is more than that of the linear and non-linear receivers of the

Chapter 4 but it is much less than the complexity of time domain methods.

The problem of joint transmitter-receiver design for MIMO channels with the inclusion of aliasing has been mentioned in the literature to be extremely complicated. In the past, some people attempted to simplify or solve this problem. To the best of our knowledge there has not been completely general solution for this optimization problem. In Chapter 6 we presented a new general solution when there is no bandwidth constraint. We derived the optimum frequency support for each substream and we observed that the optimum required bandwidth is equal to the number of substreams times the symbol rate. This optimized bandwidth may not be continuous.

In the following chapter, we developed a linear frequency domain precoder (pre-equalizer) using the knowledge of the second order statistics of the channel at the transmitter. This precoder is implemented in the frequency domain, so it has a very low complexity in severely time-dispersive channels. The proposed structure was derived by minimizing an upper bound of the MSE. The proposed design forced eigenmode transmission of the correlation channel matrix. It was shown that the precoder matrices for different frequencies are all equal.

Finally, we developed an algorithm for power optimization in V-BLAST system with two transmit antennas. We derived an analytical equation for probability of vector error for $2 \times N$ V-BLAST systems when the signals are transmitted with different powers. Assuming the total power is fixed, this equation is minimized to find the optimal energy distributions for different substreams. This procedure is done at the receiver and only the optimum energy for each substream is transmitted back to the transmitter, so the proposed method has low over-head feedback compared to the transmitter optimization techniques of Chapter 6 which require feed back of the

whole channel matrix for each realization.

In solving the above problems, we introduced several novel mathematical techniques that can be very useful in solving many different communication theory problems especially those related to MIMO systems.

Suggestions for Future research

- In practice, the channel cannot be estimated perfectly at the receiver. The performance of the proposed methods needs to be evaluated in the presence of channel estimation errors [105, 106].
- The analysis of the proposed receivers in Chapter 4 in the presence of error propagation needs to be investigated. Also, to reduce the error propagation some new techniques need to be developed [107–117].
- Turbo equalization and joint iterative decoding and equalization in frequency domain need to be considered to extend the results of Chapter 5, and performance should be evaluated with respect to independent equalization and decoding [8, 63, 118].
- The performance of the IB-DFE with channel coding needs to be characterized. Theoretical analysis of the IB-DFE combined with channel codes and channel capacity analysis of IB-DFE are interesting research problems [25, 26].
- Design of MIMO receivers and space-time codes with more than Nyquist bandwidth is an interesting research problem which has not been investigated yet.
- Exact power optimization for BLAST algorithm with more than two transmitters considering error propagation is a very challenging and interesting problem.

- Power optimization methods with partial channel knowledge at the transmitter and non ideal channel estimation need to be developed.
- The design of the efficient frequency domain receivers without cyclic prefix is an open research problem [119].
- Efficient implementation issues of the optimum transmitter design of Chapter 6 and its application in multi-band transmission [120] is an interesting research problem.

Appendix A

MMSE Expression for Hybrid-Time Frequency LST Receiver

The MMSE expression in (4.3.27) can be written as:

$$\begin{aligned} MMSE = & \frac{\sigma^2}{M} \sum_{l=0}^{M-1} \left(1 + \sum_{k \in F_B(p)} f_p(k) \exp\left(\frac{-j2\pi lk}{M}\right) + \sum_{k \in F_B(p)} f_p^*(k) \exp\left(\frac{j2\pi lk}{M}\right) \right. \\ & \left. + \sum_{k \in F_B(p)} \sum_{k' \in F_B(p)} f^*(k) f(k') \exp\left(\frac{j2\pi l(k - k')}{M}\right) \right) \mathbf{Q}_l^p(1, 1) \end{aligned} \quad (\text{A.0.1})$$

(A.0.1) can be written as:

$$\begin{aligned} MMSE = & \frac{\sigma^2}{M} \sum_{l=0}^{M-1} \left(1 + \sum_{k \in F_B(p)} f_p(k) \exp\left(\frac{-j2\pi lk}{M}\right) + \sum_{k \in F_B(p)} f_p^*(k) \exp\left(\frac{j2\pi lk}{M}\right) \right. \\ & \left. + \sum_{k \in F_B(p)} f^*(k) \sum_{k' \in F_B(p)} f(k') \exp\left(\frac{j2\pi l(k - k')}{M}\right) \right) \mathbf{Q}_l^p(1, 1) \end{aligned} \quad (\text{A.0.2})$$

Using (4.3.28) (A.0.2) can be written as:

$$\begin{aligned}
 MMSE = & \frac{\sigma^2}{M} \sum_{l=0}^{M-1} \left(1 + \sum_{k \in F_B(p)} f_p(k) \exp\left(\frac{-j2\pi lk}{M}\right) + \right. \\
 & \left. \sum_{k \in F_B(p)} f_p^*(k) \exp\left(\frac{j2\pi lk}{M}\right) \right) \mathbf{Q}_l^p(1, 1) \\
 & - \frac{\sigma^2}{M} \sum_{k \in F_B(p)} f_p^*(k) \sum_{l=0}^{M-1} \exp\left(\frac{j2\pi lk}{M}\right) \mathbf{Q}_l^p(1, 1)
 \end{aligned} \tag{A.0.3}$$

Or:

$$MMSE = \frac{\sigma^2}{M} \sum_{l=0}^{M-1} \left(1 + \sum_{k \in F_B(p)} f_p(k) \exp\left(\frac{-j2\pi lk}{M}\right) \right) \mathbf{Q}_l^p(1, 1) = \frac{\sigma^2}{M} \sum_{l=0}^{M-1} F_l \mathbf{Q}_l^p(1, 1) \tag{A.0.4}$$

Using Parseval's theorem A.0.4 can be written as:

$$MMSE = \sigma^2 v_0 + \sigma^2 \sum_{k \in F_B(p)} v_k f_p(k) \tag{A.0.5}$$

Appendix B

MMSE Expression for Hybrid-Time Frequency MIMO-DFE

The MMSE expression in 4.3.27 can be written as:

$$\begin{aligned} MMSE = & \frac{\sigma^2}{M} \sum_{l=0}^{M-1} \left(\mathbf{e}_p^* \mathbf{Q}_l \mathbf{e}_p + \mathbf{e}_p^* \mathbf{Q}_l \sum_{k \in F_B(p)} \mathbf{f}_p(k) \exp\left(\frac{j2\pi lk}{M}\right) + \right. \\ & \sum_{k \in F_B(p)} \mathbf{f}_p^*(k) \exp\left(\frac{-j2\pi lk}{M}\right) \mathbf{Q}_l \mathbf{e}_p + \\ & \left. \sum_{k \in F_B(p)} \sum_{k' \in F_B(p)} \mathbf{f}_p^*(k) \mathbf{Q}_l \mathbf{f}_p(k') \exp\left(\frac{j2\pi l(k' - k)}{M}\right) \right) \end{aligned} \quad (\text{B.0.1})$$

(B.0.1) can be written as:

$$\begin{aligned}
 MMSE = & \frac{\sigma^2}{M} \sum_{l=0}^{M-1} \left(\mathbf{e}_p^* \mathbf{Q}_l \mathbf{e}_p + \mathbf{e}_p^* \mathbf{Q}_l \sum_{k \in F_B(p)} \mathbf{f}_p(k) \exp\left(\frac{j2\pi l k}{M}\right) + \right. \\
 & \sum_{k \in F_B(p)} \mathbf{f}_p^*(k) \exp\left(\frac{-j2\pi l k}{M}\right) \mathbf{Q}_l \mathbf{e}_p + \\
 & \left. \sum_{k \in F_B(p)} \mathbf{f}_p^*(k) \sum_{k' \in F_B(p)} \mathbf{Q}_l \mathbf{f}_p(k') \exp\left(\frac{j2\pi l (k' - k)}{M}\right) \right) \quad (B.0.2)
 \end{aligned}$$

Using (4.3.52) (B.0.1) can be written as:

$$\begin{aligned}
 MMSE = & \frac{\sigma^2}{M} \sum_{l=0}^{M-1} \left(\mathbf{e}_p^* \mathbf{Q}_l \mathbf{e}_p + \mathbf{e}_p^* \mathbf{Q}_l \sum_{k \in F_B(p)} \mathbf{f}_p(k) \exp\left(\frac{j2\pi l k}{M}\right) + \right. \\
 & \sum_{k \in F_B(p)} \mathbf{f}_p^*(k) \exp\left(\frac{j2\pi l k}{M}\right) \mathbf{Q}_l \mathbf{e}_p + \\
 & \left. - \sum_{k \in F_B(p)} \mathbf{f}_p^*(k) \sum_{l=0}^{M-1} \exp\left(\frac{j2\pi l k}{M}\right) \mathbf{Q}_l \mathbf{e}_p \right) \quad (B.0.3)
 \end{aligned}$$

Or:

$$MMSE = \frac{\sigma^2}{M} \sum_{l=0}^{M-1} \left(\mathbf{e}_p^* \mathbf{Q}_l \mathbf{e}_p + \mathbf{e}_p^* \mathbf{Q}_l \sum_{k \in F_B(p)} \mathbf{f}_p(k) \exp\left(\frac{-j2\pi l k}{M}\right) \right) = \frac{\sigma^2}{M} \sum_{l=0}^{M-1} \mathbf{e}_p^* \mathbf{Q}_l \mathbf{F}_l^p \quad (B.0.4)$$

Using Parseval's theorem (A.0.4) can be written as:

$$MMSE = \sigma^2 \mathbf{e}_p^* (\mathbf{v}_0 + \sigma^2 \sum_{k \in F_B(p)} \mathbf{v}_k \mathbf{f}_p(k)) \quad (B.0.5)$$

Appendix C

Chapter 6 Proofs

C.0.1 Proof of Lemma 1

We can always write $\mathbf{G}$ as the sum of its projections on a subspace $\mathbf{V}$ and its orthogonal subspace $\tilde{\mathbf{V}}$ (in the space of M dimensional vectors)

$$\mathbf{T} = \mathbf{V}\Phi + \tilde{\mathbf{V}}\tilde{\Phi} = \mathbf{T}_{\parallel} + \mathbf{T}_{\perp} \quad (\text{C.0.1})$$

where Φ is a $K \times K$ matrix and $\tilde{\Phi}$ is a $(M - K) \times K$ matrix. $\mathbf{T}_{\parallel}$ and $\mathbf{T}_{\perp}$ are the projections of $\mathbf{T}$ on two orthogonal subspaces spanned by $\mathbf{V}$ and $\tilde{\mathbf{V}}$, respectively. We have $\tilde{\mathbf{V}}^* \mathbf{H}^* \mathbf{R}^{-1} \mathbf{H} \tilde{\mathbf{V}} = \mathbf{0} \implies \|\mathbf{R}^{-1/2} \mathbf{H} \tilde{\mathbf{V}}\| = 0 \implies \mathbf{R}^{-1/2} \mathbf{H} \tilde{\mathbf{V}} = \mathbf{0}$. Since $\mathbf{R}$ is full rank, we have $\mathbf{H} \tilde{\mathbf{V}} = \mathbf{0}$ ($\tilde{\mathbf{V}}$ is in the right null space of $\mathbf{H}$) or

$$\mathbf{H} \mathbf{T}_{\perp} = \mathbf{0} \quad (\text{C.0.2})$$

Postmultiplying (6.2.7) by $\mathbf{F}_{\perp}$ we obtain

$$\frac{1}{T^2} \sum_k \mathbf{T}_k^* \mathbf{H}_k^* \mathbf{G}_k^* \mathbf{G} \mathbf{H} \mathbf{T}_{\perp} + \mu \mathbf{T}^* \mathbf{T}_{\perp} = \frac{1}{T} \mathbf{G} \mathbf{H} \mathbf{T}_{\perp}. \quad (\text{C.0.3})$$

Using (C.0.2) and (C.0.3) we obtain the desired result ($\mathbf{T}_{\perp} = \mathbf{0}$).

C.0.2 Proof of Lemma 2

Postmultiplying (6.2.6) by $\mathbf{GH}$ and premultiplying (6.2.7) by $\mathbf{HT}$, we get

$$\frac{1}{T^2} \mathbf{HT} \sum_k \mathbf{T}_k^* \mathbf{H}_k^* \mathbf{G}_k^* \mathbf{GH} + \frac{1}{T} \mathbf{RG}^* \mathbf{GH} = \frac{1}{T} \mathbf{HTGH} \quad (\text{C.0.4})$$

$$\frac{1}{T^2} \mathbf{HT} \sum_k \mathbf{T}_k^* \mathbf{H}_k^* \mathbf{G}_k^* \mathbf{GH} + \mu \mathbf{HTT}^* = \frac{1}{T} \mathbf{HTGH}. \quad (\text{C.0.5})$$

Subtracting (C.0.4) from (C.0.5) yields

$$\frac{1}{T} \mathbf{RG}^* \mathbf{GH} = \mu \mathbf{HTT}^* \quad (\text{C.0.6})$$

Premultiplying (C.0.6) by $\mathbf{H}^* \mathbf{R}^{-1}$ ¹, results in

$$\frac{1}{T} \mathbf{H}^* \mathbf{G}^* \mathbf{GH} = \mu \mathbf{H}^* \mathbf{R}^{-1} \mathbf{HTT}^*. \quad (\text{C.0.7})$$

LHS of (C.0.7) is Hermitian so RHS must be Hermitian. Using Hermitian property of $\mathbf{H}^* \mathbf{R}^{-1} \mathbf{HTT}^*$ and (6.3.5) we obtain $\mathbf{\Lambda} \mathbf{\Phi} \mathbf{\Phi}^* = \mathbf{\Phi} \mathbf{\Phi}^* \mathbf{\Lambda}$. Since $\mathbf{\Lambda}$ is diagonal, this implies that $\mathbf{\Phi} \mathbf{\Phi}^*$ must be diagonal, assuming that the eigenvalues are distinct.²

C.0.3 Proof of Diagonality of Φ

By substituting (6.3.7) in (6.3.12) the MSE expression can be written as

$$MSE = tr \left(\sum_k \int_{-\frac{1}{2T}}^{\frac{1}{2T}} [\mathbf{\Phi}_k^* \mathbf{\Lambda}_k \mathbf{\Phi}_k + \mathbf{I}]^{-1} df \right) \quad (\text{C.0.8})$$

From linear algebra [92]:

Theorem 2: If $\mathbf{P}$ is a positive definite matrix and S is an index set, then³

$$\mathbf{P}^{-1}(S) \geq [\mathbf{P}(S)]^{-1} \quad (\text{C.0.9})$$

¹ $\mathbf{R}$ is assumed to be invertible

²If the eigenvalues are not distinct, define $\tilde{\mathbf{\Lambda}} + \delta \mathbf{\Lambda}$, where $\delta \mathbf{\Lambda}$ is a perturbation matrix to make $\tilde{\lambda}_{ii}$ distinct. ϕ_{ii} in (6.3.20) is a continuous function of $\mathbf{\Lambda}$, which implies that if $\lim \delta \mathbf{\Lambda} \rightarrow 0$, $\phi_{ii}(\tilde{\mathbf{\Lambda}}) \rightarrow \phi_{ii}(\mathbf{\Lambda})$.

³ $\mathbf{A} > \mathbf{B}$ if and only if $\mathbf{A} - \mathbf{B}$ is positive definite.

where the LHS is the principal submatrix of $\mathbf{P}^{-1}$ defined by index set S while the RHS is the inverse of corresponding submatrix of P . Inequality in (C.0.9) is satisfied with equality if and only if $\mathbf{P}$ has zero elements out of index set S . From Theorem 2 we can easily conclude that for a positive definite matrix $\mathbf{P}$, $[P_{ii}]^{-1} \geq P_{ii}^{-1}$ with equality when $\mathbf{P}$ is diagonal. Therefore one can write

$$\text{tr}([\Phi^* \Lambda \Phi + \mathbf{I}]^{-1}) \geq \sum_{i=1}^K [\Phi^* \Lambda \Phi + \mathbf{I}]_{ii}^{-1} \quad (\text{C.0.10})$$

The lower bound in (C.0.10) is achieved when $\Phi^* \Lambda \Phi = \mathbf{D}_1$ is a diagonal matrix. We know $\Phi^* \Phi = \mathbf{D}_2$ is also diagonal. From these two equations have $\mathbf{D}_2 \Lambda \Phi = \Phi \mathbf{D}_1$. This implies that Φ must be diagonal, assuming diagonal elements of $\mathbf{D}_2 \Lambda$ and $\mathbf{D}_1$ are distinct. If they are not distinct using a similar argument as the one used in Lemma 1, we can infer that Φ is diagonal. From the above discussion we conclude that, setting Φ diagonal achieves the lower bound of (C.0.8) and does not restrict our optimization search.

C.0.4 Proof of (6.3.14)

Substituting (6.3.4) and (6.3.5) in (C.0.7), we get

$$\frac{1}{T} \Lambda \Phi \mathbf{Y}^* \mathbf{Y} \Phi^* \Lambda = \mu \Phi \Phi^* \Lambda \quad (\text{C.0.11})$$

since Φ is diagonal, (C.0.11) can be further simplified to

$$\Phi [\frac{1}{T} \Lambda \mathbf{Y}^* \mathbf{Y} - \mu \mathbf{I}] \Phi^* = \mathbf{0} \quad (\text{C.0.12})$$

Appendix D

Peak Power of the Output Signal

Let $y_i(m), m = 0, \dots, M - 1$ and $Y_i(l), l = 0, \dots, M - 1$ be the m -th time domain sample and the corresponding l -th frequency domain sample of the transmitted signal from antenna i respectively. Then one can write:

$$Y_i(l) = \sum_{k=1}^K t_{ik} A_k(l) \quad (\text{D.0.1})$$

where $[\mathbf{T}_l]_{ij} = t_{ij}, l = 1, \dots, M - 1$. The corresponding time domain output can be written as:

$$\begin{aligned}
y_i(m) &= \frac{1}{M} \sum_{l=0}^{M-1} Y_i(l) \exp\left(\frac{j2\pi lm}{M}\right) \\
&= \frac{1}{M} \sum_{l=0}^{M-1} \sum_{k=1}^K t_{ik} A_k(l) \exp\left(\frac{j2\pi lm}{M}\right) \\
&= \frac{1}{M} \sum_{l=0}^{M-1} \sum_{k=1}^K t_{ik} \sum_{m'=0}^{M-1} a_i(m') \exp\left(\frac{-j2\pi lm'}{M}\right) \exp\left(\frac{j2\pi lm}{M}\right) \\
&= \frac{1}{M} \sum_{k=1}^K t_{ik} \\
&\quad \sum_{l=0}^{M-1} \sum_{m'=0}^{M-1} a_k(m') \exp\left(\frac{-j2\pi lm'}{M}\right) \exp\left(\frac{j2\pi lm}{M}\right) \\
&= \sum_{k=1}^K t_{ik} a_k(m)
\end{aligned} \tag{D.0.2}$$

Therefore, the FFT and IFFT cancel each other and the transmitted signals are equivalent to the single-carrier signals, and share their moderate PAPR properties (in comparison to OFDM).

Appendix E

Proof of Independence of Noise at Different Detection Stages

It can be easily seen that ν_i $i = 1, \dots, P-1$ is a complex random variable with mean zero and variance $2||\mathbf{w}_1||^2\sigma^2$. Assuming ZF criteria $\mathbf{w}_i$ has to satisfy the following conditions:

$$\begin{aligned}\mathbf{w}_i^* \mathbf{h}_i &= 1, i = 1, \dots, P \\ \mathbf{w}_i^* \mathbf{h}_j &= 0, j = i + 1, \dots, P.\end{aligned}\tag{E.0.1}$$

i.e. $\mathbf{w}_i$ is perpendicular to the subspace spanned by $\mathbf{h}_{j+1}, \dots, \mathbf{h}_P, j \geq i$. The post-detection SNR for the i -th detected component of $\mathbf{a}$ is [11]:

$$\rho_i = \frac{E\{|a_i|^2\}}{2\sigma^2||\mathbf{w}_i||^2}\tag{E.0.2}$$

where $E\{.\}$ denotes expectation. The optimum vector $\mathbf{w}_i$ is the minimum norm vector satisfying (E.0.1). This minimum norm vector lies in the subspace generated by $\mathbf{h}_i, \dots, \mathbf{h}_P$ [80] (Th.2, sec 3.10). Vector $\mathbf{w}_i \forall i < j$ is orthogonal to the subspace

generated by $\mathbf{h}_j, \dots, \mathbf{h}_P$ and since, the optimum vector $\mathbf{w}_j$ lies in this subspace we therefore conclude that $\mathbf{w}_i$ is orthogonal to $\mathbf{w}_j \forall i > j$.

or on the other hand:

$$\begin{aligned} \mathbb{E}\{\nu_i \nu_j^*\} &= \mathbb{E}\{\mathbf{w}_i^* \mathbf{n} \mathbf{n}^* \mathbf{w}_j\} = \mathbf{w}_i^* \mathbb{E}\{\mathbf{n} \mathbf{n}^*\} \mathbf{w}_j \\ &= \mathbf{w}_i^* \sigma^2 \mathbf{I}_P \mathbf{w}_j \end{aligned} \tag{E.0.3}$$

where $\mathbf{I}_P$ is $P \times P$ identity matrix. Since $\mathbf{w}_i$ and $\mathbf{w}_j$ are orthogonal, we have $\mathbb{E}\{\nu_i \nu_P^*\} = 0$. ν_i and ν_P are zero-mean random variables so they must be independent.

Acknowledgment

I would like to thank Professor S. Loyka who inspired me to generalize this proof.

Bibliography

- [1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, Mar. 1998.
- [2] E. T. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. on Telecom.*, vol. 10, pp. 585–595, Nov/Dec 1999.
- [3] A. Pualraj and B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Mag.*, vol. 14, pp. 49–82, Nov. 1997.
- [4] J. H. Winters, "Optimum combining in digital mobile radio with cochannel interferences," *IEEE J. Select. Areas Commun.*, vol. 2, pp. 528–539, July 1984.
- [5] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.
- [6] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication:performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.
- [7] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communication:performance results," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 451–460, Mar. 1999.

- [8] S. L. Ariyavistakul, "Turbo space-time processing to improve wireless channel capacity," *IEEE Trans. Commun.*, vol. 48, pp. 1073–1083, Aug. 1999.
- [9] G. J. Foschini, "Layered-space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech J.*, pp. 41–59, Autumn 1996.
- [10] G. J. Foschini, D. C. Chizhik, M. J. Gans, C. P. Papadias, and R. A. Valenzuela, "Analysis and performance of some basic space-time architecture," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 303–320, Apr. 2003.
- [11] P. W. Wolnisky, G. J. Foschini, G. D. Golden, and R. Valenzuela, "V-BLAST: An architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. ISSE'98*, Pisa, Italy, 1998, pp. 295–300.
- [12] G. J. Foschini, G. D. Golden, R. Valenzuela, and P. W. Wolnisky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1841–1852, Nov. 1999.
- [13] C. Z. W. Hassel and J. Thompson, "A comparison of detection algorithms including BLAST for wireless communication using multiple antennas," in *Proc. PIMRC'2000*, London, UK, 2000, pp. 698–703.
- [14] W. J. Choi, R. Negi, and J. M. Cioffi, "Combined ML and DFE decoding for the V-BLAST system," in *Proc. WCNC*, New Orleans, USA, 2000, pp. 1243–1248.
- [15] W. K. Wai, C. Y. Tsui, and R. S. Cheng, "A low complexity architecture of the V-BLAST system," in *Proc. WCNC'2000*, Chicago, USA, 2000, pp. 310–314.
- [16] X. Li, A. Lozano, and G. J. Foschini, "Reduced complexity detection algorithm for systems using multi-element arrays," in *Proc. Globecom*, San Francisco, USA, 2000, pp. 1072–1076.

- [17] A. Bhargave, R. J. P. Figueiredo, and T. Eltoft, "A detection algorithm for the V-BLAST system," in *Proc. Globecom*, San Antonio, USA, 2001, pp. 494–498.
- [18] S. Haykin, *Adaptive Filter Theory*. New Jersey: Prentice Hall, 2002.
- [19] B. Hassibi, "An efficient square-root algorithm for BLAST," in *Proc. ICASSP*, Istanbul, Turkey, 2000, pp. 737–740.
- [20] J. Benesty, Y. Huang, and J. Chen, "A fast recursive algorithm for optimum sequential signal detection in a BLAST system," *IEEE Trans. Signal Processing*, vol. 51, pp. 1722–1730, July 2003.
- [21] W. Zha and S. D. Blostein, "Modified decorrelating decision-feedback detection of blast space-time systems," in *Proc. ICC*, New York City, NY, Apr. 2002, pp. 335–339.
- [22] E. Biglieri, G. Taricco, and A. Tulino, "Decoding space-time codes with Blast architecture," *IEEE Trans. Signal Processing*, vol. 50, pp. 2547–2552, Oct. 2002.
- [23] V. Tarokh, H. J. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: Performance results," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 451–460, Mar. 1999.
- [24] D. Falconer and G. J. Foschini, "Theory of minimum mean-square-error QAM systems employing decision feedback equalization," *Bell Labs Tech J.*, pp. 1821–1849, Dec. 1973.
- [25] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney, "MMSE decision-feedback equalizers and coding-part I: Equalization results," *IEEE Trans. Commun.*, vol. 43, pp. 2582–2594, Oct. 1995.
- [26] —, "MMSE decision-feedback equalizers and coding-part II: Equalization results," *IEEE Trans. Commun.*, vol. 43, pp. 2595–2604, Oct. 1995.

- [27] S. U. H. Qureshi, "Adaptive equalization," *Proceedings of the IEEE*, vol. 73, pp. 1349–13385, Sept. 1985.
- [28] S. Verdu, "Minimum probability of error for asynchronous gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, pp. 744–765, Mar. 1998.
- [29] S. Roy and D. Falconer, "Optimum infinite-length MMSE multi-user decision-feedback space-time processing in broadband cellular radio," *Wireless Personal Communications*, vol. 27, Oct. 2003.
- [30] N. A-Dhahir and A. H. Sayed, "The finite-length multi-input multi-output MMSE-DFE," *IEEE Trans. Signal Processing*, vol. 48, pp. 1–17, Oct. 2000.
- [31] A. Duel-Hallen, "Equalizer for multiple input/multiple output channels and pam systems with cyclostationary input sequence," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 630–639, Apr. 1992.
- [32] C. Tidestav, M. Sternad, and A. Ahlen, "Reuse within a cell-interference rejection or multiuser detection," *IEEE Trans. Commun.*, vol. 47, pp. 1511–1522, Oct. 1999.
- [33] C. Tidestav, A. Ahlen, and M. Sternad, "Realizable MIMO decision feedback equalizers: Structure and design," *IEEE Trans. Signal Processing*, vol. 49, pp. 121–133, Jan. 2001.
- [34] S. L. Ariyavisitakul, J. H. Winters, and I. Lee, "Optimum space-time processor with dispersive interference: Unified analysis and required filter span," *IEEE Trans. Commun.*, vol. 47, pp. 1073–1083, July 1999.
- [35] A. Duel-Hallen, "A family of multiuser decision-feedback detectors for asynchronous code-division," *IEEE Trans. Commun.*, vol. 43, pp. 421–434, Feb/Mar/Apr 1995.

- [36] A. Lozano and C. Papadias, "Layered space-time receivers for frequency-selective wireless channels," *IEEE Trans. Commun.*, vol. 50, pp. 65–73, Jan. 2002.
- [37] X. Zhu and R. D. Murch, "Layered space-time equalization for wireless MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1189–1203, Nov. 2003.
- [38] A. Duel-Hallen and C. Heegard, "Delayed decision-feedback sequence estimation," *IEEE Trans. Commun.*, vol. 37, pp. 428–436, May 1989.
- [39] A. M. Tehrani, "Space-time equalization in wireless systems with multiple transmit/receive array antennas," Ph.D. dissertation, Stanford University, Department of Electrical, & Computer Engineering, Nov. 2000.
- [40] D. Falconer, T. Matsumoto, M. Ran, A. Springer, and P. Zhu. (2003, May) A mixed OFDM plus single-carrier mode air interface. [Online]. Available: WWWRF/WG4/Subgroups on new Air interfaces
- [41] J. Proakis, *Digital Communications*. McGraw-Hill, 1995.
- [42] D. Falconer, S. Ariyavisitakul, A. Benyamin-seeyar, and B. Edison, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 4, pp. 58–66, Apr. 2002.
- [43] R. Kalbasi, R. Dinis, D. Falconer, and A. H. Banihashemi, "Hybrid time-frequency layered space-time receivers for severe time-dispersive channels," in *Proc. IEEE SPAWC*, Lisbon, Portugal, 2004.
- [44] Z. Wang, X. Ma, and G. B. Giannakis, "OFDM or single-carrier block transmissions?" *IEEE Trans. Commun.*, vol. 52, pp. 380–394, Mar. 2004.
- [45] *Broadband Radio Access networks (BRAN); HIPERLAN Type2 Technical Specification part I-physical layer*, ETSI DTS/BRAN030003-1, Oct. 1999.

- [46] L. J. Cimini, "Analysis and simulation of digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. 33, pp. 665–675, July 1985.
- [47] A. Gusmao and R. D. N. Esteves, "On frequency-domain equalization and diversity combining for broadband wireless communications," *IEEE Trans. Commun.*, vol. 51, pp. 1029–1033, July 2003.
- [48] R. Dinis and A. Gusmao, "A class of nonlinear signal-processing schemes for bandwidth-efficient ofdm transmission with low envelope fluctuation," *IEEE Trans. Commun.*, vol. 52, pp. 2009–2018, Nov. 2004.
- [49] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, vol. 33, pp. 100–109, Feb. 1995.
- [50] A. Gusmo, R. Dinis, J. Conceio, and N. Esteves, "Comparison of two modulation choices for broadband wireless communications," in *Proc. IEEE VTC Spring*, Tokyo, Japan, May 2000, pp. 1300–1305.
- [51] N. Benvenuto and S. Tomasin, "On the comparison between OFDM and single carrier with DFE using a frequency domain feedforward filter," *IEEE Trans. Commun.*, vol. 50, pp. 947–955, June 2002.
- [52] G. Kadel, "Diversity and equalization in frequency domain-a robust and flexible receiver technology for broadband mobile communication systems," in *Proc. IEEE VTC*, Phoenix, USA, 1997, pp. 894–898.
- [53] M. V. Clark, "Adaptive frequency-domain equalization and diversity combining for broadband wireless communications," *IEEE Trans. Commun.*, vol. 16, pp. 1358–1359, Oct. 1998.

- [54] D. Falconer and F. R. Magee, "Adaptive channel memory truncation for maximum-likelihood sequence estimation," *Bell Labs Tech J.*, pp. 1541–1562, Nov. 1973.
- [55] S. L. Ariyavisitakul and L. J. Greenstein, "Reduced-complexity equalization techniques for broadband wireless channels," *IEEE J. Select. Areas Commun.*, vol. 1, pp. 5–15, Jan. 1997.
- [56] S. Roy and D. Falconer, "Multi-user decision-feedback space-time processing with partial cross-feedback connectivity," in *Proc. IEEE VTC Spring*, Tokyo, Japan, May 2000.
- [57] J. Tubbax, L. Vanderperre, S. Donnay, and M. Engels, "Single-carrier communications using decision-feedback equalization for multiple antennas," in *Proc. ICC'03*, Anchorage, Ak , USA, May 2003, pp. 2321–2325.
- [58] *Universal Mobile Telecommunications Systems (UMTS); Selection Procedures for the Choice of Radio Transmission of the UTMS*, ETSI Technical Report TR101 112v3.1.0, Nov. 1997.
- [59] N. A. Zervos and I. Kalet, "Optimized decision feedback equalization versus optimized orthogonal frequency division multiplexing for high-speed data transmission over the local cable network," in *Proc. ICC'89*, Boston, MA , USA, June 1989, pp. 1080–1085.
- [60] Y. Li, J. Winters, and N. R. Sollenberger, "MIMO-OFDM for wireless communications: Signal detection with enhanced channel estimation," *IEEE Trans. Commun.*, vol. 50, pp. 1471–1477, Sept. 2002.
- [61] R. Dinis, A. Gusmo, and N. Esteves, "On broadband block transmission over strongly frequency-selective fading channels," in *Proc. Wireless 2003*, vol. 1, Calgary, Canada, July 2003, pp. 261–269.

- [62] N. Benvenuto and S. Tomasin, "Block iterative DFE for single carrier modulation," *IEE Electronic Letters*, vol. 38, pp. 1144–1145, sept. 2002.
- [63] T. Abe and T. Matsumoto, "Space-time turbo equalization in frequency selective mimo channels," *IEEE Trans. Veh. Technol.*, vol. 52, pp. 469–482, May 2003.
- [64] R. Kalbasi, R. Dinis, D. Falconer, and A. H. Banihashemi, "Layered space-time receiver for single-carrier transmission with iterative frequency-domain equalization," in *Proc. IEEE VTC Spring*, Milan, Italy, May 2004.
- [65] —, "An iterative frequency-domain layered space-time receiver for SDMA systems with single-carrier transmission," in *Proc. IEEE ICASSP*, Montreal, Canada, May 2004.
- [66] R. Dinis, R. Kalbasi, D. Falconer, and A. H. Banihashemi, "Iterative layered space-time receivers for single-carrier transmission over time-dispersive channels," *IEEE Commun. Lett.*, vol. 8, pp. 579–581, sep 2004.
- [67] *Channel models for HIPERLAN/2 in different scenarios*, ETSI Std. 3ERI085B, 1998.
- [68] A. M. Chan and G. W. Wornell, "A class of block-iterative equalization for intersymbol interference channels: Fixed channel results," *IEEE Trans. Commun.*, vol. 49, pp. 1966–1976, Nov. 2001.
- [69] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, vol. 42, pp. 3221–3231, Dec. 1994.
- [70] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, pp. 2198–2206, Dec. 2001.

- [71] A. Scaglione, P. Stoica, G. B. Giannakis, and H. Sampath, "Optimal design for space-time linear precoders and decoders," *IEEE Trans. Signal Processing*, vol. 50, pp. 1051–1064, May 2002.
- [72] T. Berger and D. Tufts, "Optimum pulse amplitude modulation part:I transmitter-receiver design and bounds from information theory," *IEEE Trans. Inform. Theory*, vol. 13, pp. 196–208, Apr. 1967.
- [73] T. Ericson, "Optimum pam filters are always band limited," *IEEE Trans. Inform. Theory*, vol. 19, pp. 570–573, July 1973.
- [74] K. Lee and D. P. Petersen, "Optimal linear coding for vector channels," *IEEE Trans. Commun.*, vol. 24, pp. 1283–1290, Dec. 1976.
- [75] H. Malvar and D. Staelin, "Optimal pre- and postfiltering for multichannel signal processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 287–289, Feb. 1988.
- [76] J. Salz, "Digital transmission over cross-coupled linear channels," *AT & T Tech J.*, pp. 1147–1159, July, Aug. 1985.
- [77] B. Song and J. A. Ritcey, "Joint pre and postfilter design for spatial diversity equalization," *IEEE Trans. Signal Processing*, vol. 45, pp. 276–279, Jan. 1997.
- [78] R. Kalbasi, D. D. Falconer, and A. H. Banihashemi, "Optimum mimo systems are bandlimited: Optimal and suboptimal approaches," in *Proc. of the Wireless 2003 Conference*, Calgary, Canada, July 2003.
- [79] —, "Optimum space-time processors under symbol-rate sampling are bandlimited," accepted for publication in *IEEE Trans. on SP*.
- [80] D. G. Luenberger, *Optimization By Vector Space Methods*. NY: John Wiley & Sons, 1968.

- [81] M. L. Honig, P. Crespo, and K. Steiglitz, "Suppression of near- and far-end crosstalk by linear pre- and post-filtering," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 614–629, Aug. 1992.
- [82] Y. Li and Z. Ding, "A simplified approach to optimum diversity combining and equalization in digital data transmission," *IEEE Trans. Commun.*, vol. 43, pp. 2285–2288, Aug. 1995.
- [83] R. G. Gallager, *Information Theory and Reliable Transmission*. New York: Wiley, 1968.
- [84] M. L. Honig and U. Madhow, "Optimization of transmitter pulses for two-user data communications," in *Proc. IEEE ISIT'93*, San Antonio, Texas, USA, Jan. 1993, pp. 372–372.
- [85] B. R. Petersen and D. D. Falconer, "Minimum mean square equalization in cyclostationary and stationary interference-analysis and subscriber line calculations," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 931–940, Aug. 1991.
- [86] G. D. Golden, J. E. Mazo, and J. Salz, "Transmitter design for data transmission in the presence of data-like intereferer," *IEEE Trans. Commun.*, vol. 43, pp. 614–629, Feb./Mar. 1995.
- [87] N. Al-Dhahir and J. Cioffi, "Finite length vs. infinite-length mmse-dfe: The connection," in *Proc. of 27 Asilomar Conference*, Pacific Grove, CA, USA, Oct. 1993, pp. 677–681.
- [88] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, pp. 357–363, Mar. 1998.
- [89] H. Sampath, H. Bolcskei, and A. Pualraj, "Pre-equalization for MIMO wireless channels with delay spread," in *Proc. IEEE VTC Fall*, Boston, MA, USA, May 2000.

- [90] R. F. H. Fischer, C. Windpassinger, A. Lampe, and J. B. Huber, "Tomlinson-Harashima precoding in space-time transmission for low-rate backward channels," in *Proc. IEEE Int. Zurich Seminar on Broadband Commun. (IZS02)*, Zurich, Switzerland, Feb. 2002.
- [91] —, "Space-time transmission using tomlinson-harashima precoding," in *Proc. 4ITG Conf. Source and Channel Coding*, Berlin, Germany, Jan. 2002.
- [92] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge Univ. Press, 1985.
- [93] H. Bolcskei, D. Gesbert, and A. Pualraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, pp. 225–234, Feb. 2002.
- [94] D. Gesbert, H. Bolcskei, D. A. Gore, and A. Paulraj, "Outdoors MIMO wireless channels: Models and performance prediction," *IEEE Trans. Commun.*, vol. 50, pp. 1926–1934, Dec. 2002.
- [95] D. Shiu, G. J. Foschini, M. J. Gans, and J. M. Khan, "Fading correlation and its effects on the capacity of multielement antenn systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.
- [96] A. Hjorunges and J. A. D. Gesbert, "Precoding for sapce-time block codes in (non-)Kronecker correlated MIMO channels," in *Proc. Eur. Signal Proc. Conf.*, Vienna, Austria, Sept. 2004, pp. 515–518.
- [97] J. Salz and J. H. Winters, "Effect of fading correlation on adaptive arrays in digital mobile radio," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 1049–1057, Nov. 1994.
- [98] K. I. Pedersen, J. B. Andersen, J. P. Kermol, and P. E. Mogensen, "A stochastic multiple-input-multiple-output radio channel model for evaluation of space-time

coding algorithm,” in *Proc. IEEE VTC Fall*, Boston, MA, USA, May 2000, pp. 893–897.

- [99] S. Loyka and F. Gagnon, “Performanmce analysis of the V-BLAST algorithm: An analytical aproach,” *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1326–1337, 2004.
- [100] S. H. Nam, O. Shin, and K. B. Lee, “Transmit power allocation for a modified V-BLAST system,” *IEEE Trans. Commun.*, vol. 52, pp. 1074–1079, July 2004.
- [101] N. Wang and S. D. Blostein, “Minimum BER power allocation for MIMO spatial multiplexing systems,” in *Proc. ICC’05*, Seoul, Korea, May 2005.
- [102] R. Kalbasi, D. D. Falconer, and A. H. Banihashemi, “Optimal power allocation and transmitter optimization in a two-transmitter V-BLAST with error propagation,” in *Proc. of the 2003 Canadian Workshop on Information Theory*, Waterloo, Canada, May 2003, pp. 92–95.
- [103] —, “Optimum power allocation for a V-BLAST system with two antennas at the transmitter,” accepted for publication in *IEEE Comm. Letters*.
- [104] J. Coon, J. Siew, M. Beach, A. N. Nix, S. Armour, and J. McGeen, “A comparison of MIMO-OFDM and MIMO-SCFDE in WLAN environment,” in *Proc. GLOBECOM*, San Fransisco, USA, Dec. 2002, pp. 3296–3201.
- [105] Y. G. Li, “Simplified channel estimation for OFDM systems with multiple transmit antennas,” *IEEE Trans. Wireless Commun.*, vol. 1, pp. 67–75, Jan. 2002.
- [106] Y. G. li, N. Seshadri, and S. Ariyavisitakul, “Channel estimation for OFDM systems with transmit diversity in mobile wireless channels,” vol. 17, pp. 461–471, Mar. 1999.

- [107] M. V. Eyuboglu, "Detection of coded modulation signals on linear severely distorted channels using decision-feedback noise prediction with interleaving," *IEEE Trans. Commun.*, vol. 36, pp. 401–409, Apr. 1998.
- [108] F. Zhao, G. Mathew, and B. Farhang-Boroujeny, "Techniques for minimizing error propagation in decision feedback detectors for recording channels," *IEEE Trans. Magn.*, vol. 37, pp. 592–602, Jan. 2001.
- [109] J. E. Smee and N. C. Beaulieu, "Error-rate evaluation of linear equalization and decision feedback equalization with error propagation," *IEEE Trans. Commun.*, vol. 46, pp. 656–665, May 1998.
- [110] Z. Tiani, "Mitigating error propagation in decision-feedback equalization for multiuser CDMA," *IEEE Trans. Commun.*, vol. 52, pp. 525–529, Apr. 2004.
- [111] R. A. Kenedy, B. D. O. Anderson, and R. R. Bitmead, "Tight bounds on the error probabilities of decision feedback equalizers," *IEEE Trans. Commun.*, vol. 35, pp. 1022–1028, Oct. 1987.
- [112] M. Jin, B. Farhang-Boroujeny, G. Mathew, and K. C. Indukumar, "A novel fast approach for estimating error propagation in decision feedback detectors," vol. 19, pp. 668–676, Apr. 2001.
- [113] M. Reuter, J. C. Allen, J. R. Zeidler, and R. C. North, "Mitigating error propagation effects in a decision feedback equalizer," *IEEE Trans. Commun.*, vol. 49, pp. 2028–2041, Nov. 2001.
- [114] P. Monsen, "Adaptive equalization of the slow fading channel," *IEEE Trans. Commun.*, vol. 22, pp. 1064–1075, Aug. 1974.
- [115] D. L. Duttwiler, J. E. Mazo, and D. G. Messerschmitt, "An upper bound on the error probability in decision-feedback equalization," *IEEE Trans. Inform. Theory*, vol. 20, pp. 490–497, July 1974.

- [116] M. Chiani, "Introducing erasures in decision-feedback equalization to reduce error propagation," *IEEE Trans. Commun.*, vol. 45, pp. 757–760, July 1997.
- [117] A. Fertner, "Improvement of bit-error-rate in decision feedback equalizer by preventing decision-error propagation," *IEEE Trans. Signal Processing*, vol. 46, pp. 1872–1877, July 1998.
- [118] T. Abe, S. Tomisato, and T. matsumoto, "Space-time turbo equalization in frequency-selective mimo channels," *IEEE Trans. Veh. Technol.*, vol. 52, pp. 469–475, May 2003.
- [119] D. Falconer and S. L. Ariyavisitakul, "Broadband wireless using using single carrier and frequency domain equalization," in *Proc. IEEE Wireless Personal Multimedia Communications*, Honolulu, HI, USA, Oct. 2002.
- [120] D. Falconer and M. Muck, "Feasibility of multi-bandwidth transmission in future wireless systems," in *Proc. IST-Summit*, Dresden, Germany, June 2005.