Loughborough University

University Library

Author/Filing Title SHARMA, V.

Class Mark T

Please note that fines are charged on ALL overdue items.

0403818958

Multiuser Spatial Diversity Techniques for Wireless Communication Systems

Thesis submitted to Loughborough University in candidature for the degree of Doctor of Philosophy.

Vimal Sharma

Advanced Signal Processing Group
Loughborough University
2008

Loughborough University
Pilkington Library

Date 23/10/09

Class T

Acc No. 04038 18958

Abstract

Multiple antennas at the transmitter and receiver, formally known as multiple-input multiple-output (MIMO) systems have the potential to either increase the data rates through spatial multiplexing or enhance the quality of services through exploitation of diversity. In this thesis, the problem of downlink spatial multiplexing, where a basestation (BS) serves multiple users simultaneously in the same frequency band is addressed. Spatial multiplexing techniques have the potential to make huge saving in the bandwidth utilization. We propose spatial diversity techniques with and without the assumption of perfect channel state information (CSI) at the transmitter. We start with proposing improvement to signal-to-leakage ratio (SLR) maximization based spatial multiplexing techniques for both flat fading and frequency selective channels. For orthogonal frequency division multiplexing (OFDM), even for a frequency selective channel, the channel in each frequency bin would appear as flat, hence beamformers developed for frequency flat fading channels can be applied in each frequency bin for a frequency selective OFDM system. However, for spatial multiplexing of TDMA systems with frequency selective channels, we propose a novel approach based on uplink-downlink duality. A general framework based on channel shortening filters has been considered. A better link performance is achieved as compared to equalization based spatial multiplexing techniques.

The spatial diversity techniques proposed in the first phase of the thesis require nearly perfect CSI at the transmitter, which is in general not possible. We therefore

focus on developing spatial diversity techniques that are resilient to errors in the CSI. Several techniques based on advanced convex optimization theory have been proposed to obtain robust solution by incorporating the possibility of CSI errors in the design. A robust SLR based downlink beamformer using worst-case performance optimization has been proposed. Having demonstrated outstanding performance, the problem of robust downlink beamforming using semi-definite programming (SDP) has been addressed. Two novel schemes using worst-case performance optimization and positive semidefinite (PSD) constraints have been proposed. Finally, a robust downlink beamformer with per-antenna power constraint has also been developed.

Acknowledgements

I would like to express my gratitude to my supervisor Dr. Sangarapillai Lambotharan for his continuous support and interest into my research. He provided support regardless of time and place. I am grateful for him to provide me with great motivation, enthusiasm and technical insight.

I also wish to thank my co-supervisor Professor Jonathon A. Chambers for providing me with motivation and encouragement to do well. I would also like to thank Dr. Andreas Jakobsson for having some really useful discussion in the early phases of my PhD. The discussion not only resulted in new results but most importantly provided me with the right direction to continue my research.

I am extremely thankful to Professor Alex. B. Gershman for providing me with the once in a life time opportunity to visit his department at Darmsatdt University, Germany. Also thanks to my colleagues Imran Wajid, Michael Ruebsamen and Haihua Chen for providing me both technical insight in my research and making my stay in Germany a memorable experience. Special thanks to Mrs. Marlis Gorecki for helping me in organizing this trip.

I would be naive not to thank Cardiff University and all the admin staff there for ensuring that I was provided with up-to date equipment and facilities needed to conduct my research and making the process of transferring to Loughborough University smooth and non-distractive.

Special thanks to all the people who made me feel like home in Cardiff University

and Loughborough University especially Clive for ensuring that I am never lost and providing the technical support in the initial phases of my first year, Kianoash for providing all the useful discussions that not only encouraged me but at time provided new ideas to conduct my research. Tariq for providing me with the first step in my research and also always having an answer for a question. Thanks to Cumanan and Ranaji for encouraging me in the final stages and being friends they have been. Also to all other colleagues in the lab including Min, Andrew, Layteen, Ricky, Mohsen for making it a friendly place to work in.

I have no words to thank my family for putting the faith in me and providing me with unconditional support throughout my time as a student.

Statement of Originality

The original contributions are on the proposal of downlink spatial diversity techniques with and without the assumption of perfect channel state information at the transmitter. The novelty of the contributions are supported by two submitted journals and seven published conference papers. The contributions are summarized as follows:

In Chapter 4, a novel iterative scheme to the problem of downlink beamforming in a multiuser MIMO system has been proposed based on maximizing the signal-to-leakage ratio. The performance of the proposed scheme has been studied for both flat-fading and frequency selective fading channels. The results have been published in [1,2]:

- V. Sharma and S. Lambotharan, "Multiuser downlink MIMO beamforming using an iterative optimization approach", in Proc. *IEEE Vehicular Technology Conference Fall, Montreal, Canada*, 25-28 Sep, 2006.

- V. Sharma and S. Lambotharan, "Interference suppression in multiuser downlink MIMO beamforming using an iterative optimization approach", in Proc. *European Signal Processing Conference, Florence, Italy*, 4-8 Sep, 2006.

In Chapter 5, a channel shortening based spatial multiplexing scheme for multiuser MISO systems based on uplink-downlink duality has been proposed for frequency selective channel environments. The work has been publication in [3,4]:

- V. Sharma and S. Lambotharan, "Space-time channel shortening based spatial

multiplexing schemes for multiusers with multipath channels," in Proc. *IEEE Asilomar 2007, California, USA*, Nov., 2007.

- V. Sharma and S. Lambotharan, "Space-time channel shortening based spatial multiplexing techniques using uplink-downlink duality," in Proc. *IEEE Vehicular Technology Conference 2008 Spring, Singapore*, May, 2008.

and has been submitted for possible publication in [5]:

- V. Sharma and S. Lambotharan, "Space-time channel shortening based spatial multiplexing schemes for frequency selective multiuser channels using uplink-downlink duality," submitted to *IEEE Trans. Sig. Process.*, April, 2008.

In Chapter 6, robust solutions to the problem of downlink beamforming under the assumption that only an erroneous channel state information is available at the transmitter have been proposed. We also derived closed form expressions for two important regularization parameters. Results have been published for both flat fading and frequency selective fading channels in [6, 7]:

- V. Sharma and S. Lambotharan,"Robust multiuser beamformers in MIMO-OFDM systems", in Proc. *IEEE International Conference on Communications Systems, Singapore*, 30 Oct - 2 Nov, 2006.

- V. Sharma, S. Lambotharan and A. Jakabsson, "Robust transmit multiuser beamforming using worst case performance optimization," in Proc. *IEEE Vehicular Technology Conference, Spring, Singapore*, May, 2008.

In Chapter 7, two novel techniques for the problem of downlink beamforming in a MISO system based on worst-case performance optimization have been proposed. The conventional downlink beamforming based on worst-case performance optimization violates the positive-semidefinite constraints on the channel covariance matrices. Our

proposed solutions not only outperform this technique, but also preserve the positive semidefinite constraints. The results have been published in [8]:

- V. Sharma, Imran Wajid, Alex B. Gershman, Haihua Chen, and Sangarapillai Lambotharan, "Robust downlink beamforming using positive semi-definite covariance constraints," in Proc. *IEEE Workshop of Smart Antennas, Darmstadt, Germany*, Feb., 2008.

and will be submitted for a possible publication in [9]:

- V. Sharma, I. Wajid, A. B. Gershman, H. Chen, and S. Lambotharan "Robust downlink beamforming using worst-case performance optimization and positive semi-definite covariance constraints," to be submitted to *IEEE Trans. Sig. Process.*

In Chapter 8, a robust solution to the problem of downlink beamforming with per-antenna power constraints has been proposed.

Contents

List of Tables xi

List of Figures xiv

Acronyms xvii

Notational Conventions xviii

1 Introduction 1
- 1.1 The Wireless Communications Channel . . . 2
- 1.2 Multiple Antennas . . . 4
  - 1.2.1 Gains and Features of MIMO System . . . 5
  - 1.2.2 Tradeoffs Between Gains . . . 8
- 1.3 Multiuser Downlink System . . . 9
  - 1.3.1 Frequency Selective Channels . . . 10
  - 1.3.2 Non-Perfect Channel State Information . . . 11
- 1.4 Figures of Merit . . . 12
  - 1.4.1 Mean Square Error . . . 12
  - 1.4.2 Signal to Interference-plus-Noise Ratio . . . 12
  - 1.4.3 Bit Error Rate . . . 12
- 1.5 Thesis Outline . . . 13

2 Literature Review 17
- 2.1 Single User Spatial Diversity Techniques . . . 18
  - 2.1.1 Receiver Antenna Diversity . . . 18
  - 2.1.2 Transmit Antenna Diversity . . . 21
- 2.2 Joint Transceiver Design . . . 25
- 2.3 Multiuser Spatial Diversity Techniques . . . 27
  - 2.3.1 MIMO-MAC and MIMO-BC . . . 27
  - 2.3.2 Single Receive Antenna . . . 28
  - 2.3.3 Multiple Receive Antennas . . . 30
- 2.4 Summary . . . 33

3 Convex Optimization Theory 34
- 3.1 Why Convex? . . . 35
- 3.2 Basic Optimization Concepts . . . 36

3.2.1 Convex Sets 36
3.2.2 Convex Cones 37
3.2.3 Convex Functions 37
3.3 Convex Optimization Problems 38
3.3.1 Art of Using Convex Optimization 39
3.4 Canonical Optimization Problems 40
3.4.1 Linear Program 40
3.4.2 Quadratic Programming 41
3.4.3 Second Order Cone Programming 42
3.4.4 Semidefinite Programming 42
3.5 Duality and KKT Conditions 43
3.6 Robust Convex Optimization 46
3.7 Interior Point Methods 48
3.8 Techniques based on Convex Optimization 49
3.8.1 Multiuser beamforming 49
3.8.2 Uplink-Downlink Duality via Lagrangian Duality 51
3.9 Robust Techniques 54
3.9.1 Worst-Case Robust Beamforming 54
3.9.2 Robust Multiuser Downlink Beamforming 56
3.10 Summary 57

4 Downlink Beamforming based on Maximizing SLR 58
4.1 System Model 60
4.2 Algorithms 61
4.2.1 The Signal-to-Leakage Ratio Algorithm 62
4.2.2 The Proposed Algorithm 64
4.3 Numerical Examples 67
4.4 Frequency-Selective Channels 73
4.5 MIMO-OFDM System Model 73
4.6 Algorithms and Simulation Results 76
4.7 Summary 78

5 Diversity Techniques for Frequency Selective Channels 80
5.1 Downlink Spatial Multiplexing System Model 82
5.1.1 Downlink MMSE based STEQ and STCS filters 87
5.2 Problem Statement 89
5.2.1 Criterion 1 (C1): Max-Min Fairness 89
5.2.2 Criterion 2 (C2): Min Power 90
5.3 Virtual Uplink Model 90
5.3.1 Uplink MMSE based STEQ and STCS filters 93
5.4 Effective Channels in the Uplink and the Downlink 95
5.5 Duality 99
5.6 Algorithms 105
5.6.1 C1: Max-Min Fairness 105
5.6.2 C2: Min Power 107

5.7 Simulation Results 108
5.8 Summary 113

6 Robust Downlink Beamfoming based on Maximizing SLR 114
6.1 System Model 116
6.2 Algorithms 119
6.2.1 Non-Robust Design 119
6.2.2 Robust Design 120
6.2.3 Diagonal Loading 122
6.3 Simulation Results 123
6.3.1 Example I - Mean Feedback in MIMO Systems 125
6.3.2 Example II - Feedback in MIMO-OFDM 126
6.4 Summary 131

7 Robust Beamforming with PSD Constraints 132
7.1 Problem Formulation 134
7.2 Worst-Case Robust Beamforming 138
7.3 Robust Beamforming with PSD Constraints 140
7.3.1 The First Method 144
7.3.2 The Second Method 145
7.3.3 Relationship Between the Regularization Parameters 147
7.4 Joint Beamforming and BS Assignment 148
7.5 Problem Formulation 149
7.6 Simulation Results 153
7.6.1 Single BS 153
7.6.2 BS Assignment 155
7.7 Summary 159

8 Robust Beamforming with Per Antenna Power Constraints 161
8.1 Problem Formulation 162
8.2 Robust Design 165
8.3 Power Constraints Per Group of Antennas 168
8.4 Simulation Results 168
8.5 Summary 169

9 Conclusions and Future Work 171
9.1 Summary 171
9.2 Future Work 173

A Worst-Case Performance Optimization 175

Bibliography 189

List of Tables

1.1 System configurations and diversity order achieved. 7

4.1 Code rate and generating polynomials for the coders used in simulations. 78

5.1 Vectors and matrices used in the chapter 83

6.1 Gain (dB) of the robust algorithm over the non-robust algorithm to achieve a BER of 10^{-2}. 125

6.2 Experimental values of $\bar{\epsilon}_1$ and $\bar{\epsilon}_2$ used in the simulations. 128

List of Figures

1.1 Multiple antenna system models. 5
1.2 Block diagram of a multiuser downlink system. 9

2.1 Receiver diversity in a SIMO system. 19
2.2 Performance of receive diversity technique (MRC) for various number of antennas. 21
2.3 Performance of transmit diversity technique (transmit-MRC) for various number of antennas. 23
2.4 Performance of MIMO beamforming. 26

3.1 Convex sets (balls) and non-convex sets (spheres). 37

4.1 The BER performance for all the users is plotted as a function of the SNR for a MU-MIMO system with $N_T = 6$ transmit antennas and $K = 5$ users, each equipped with $N_{R_i} = 3, i = 1, 2, \ldots, 5$ receive antennas. 69
4.2 The average BER performance is plotted for SLR and the proposed algorithm for various iterations as a function of SNR for a MU-MIMO system with $N_T = 6$ transmit antennas and $K = 5$ users, each equipped with $N_{R_i} = 3, i = 1, 2, \ldots, 5$ receive antennas. 70
4.3 SINR outage is plotted for conventional beamforming, SLR and the proposed algorithm at the 10th iteration at a SNR of 10dB for a MU-MIMO system with $N_T = 6$ transmit antennas and $K = 5$ users, each equipped with $N_{R_i} = 3, i = 1, 2, \ldots, 5$ receive antennas. 71
4.4 SINR outage is plotted for conventional beamforming, SLR and the proposed algorithm at the 20th iteration at a SNR of 5dB for a MU-MIMO system with $N_T = 6$ transmit antennas and $K = 5$ users, each equipped with $N_{R_i} = 3, i = 1, 2, \ldots, 5$ receive antennas. 72
4.5 OFDM system model. 72
4.6 The block diagram of a MU-MIMO OFDM transmitter with N_T transmitting antennas. 76
4.7 The block diagram of a MU-MIMO OFDM receiver for the ith user with N_{R_i} receive antennas. 76
4.8 The block diagram of a MU-MIMO encoder. 77

4.9 Normalized throughput for SLR and the proposed algorithm at the 10th iteration for various coding rates presented in Table 4.1. 78

5.1 Downlink spatial multiplexing system model. 84
5.2 Downlink MMSE ST filter design 88
5.3 Virtual uplink model. 91
5.4 Uplink MMSE ST filter design. 93
5.5 Result for P1, balanced SINR targets within the available total power. 111
5.6 Result for P1, balanced MSE targets within the available total power. 111
5.7 Result for P2, minimum total power required to achieve the given SINR targets. 112
5.8 Result for P1, BER vs SNR comparison. 112

6.1 Frobenius norm of the error introduced in the CSI due to feedback delay. 118
6.2 Accurate and inaccurate frequency response of a frequency selective channel. 118
6.3 The average BER performance for all the users is plotted as a function of the SNR for a MU-MIMO system with $N_T = 6$ transmit antennas and $K = 3$ users, each equipped with $N_{R_i} = 3$ receive antennas. . . . 127
6.4 The average BER performance for all the users for the case of mean feedback is plotted as a function of the SNR for a MU-MIMO system with $N_T = 6$ transmit antennas and $K = 3$ users, each equipped with $N_{R_i} = 3$ receive antennas. 127
6.5 The average BER performance for all the users is plotted as a function of the SNR for a MU-MIMO OFDM system with $N_T = 6$ transmit antennas and $K = 3$ users, each equipped with $N_{R_i} = 3$ receive antennas. Here the multipath channel length N_h is fixed to 3, but the block size N_B is changed from 4 to 32. 129
6.6 The average BER performance for all the users is plotted as a function of the SNR for a MU-MIMO OFDM system with $N_T = 6$ transmit antennas and $K = 3$ users, each equipped with $N_{R_i} = 3$ receive antennas. Here the block size N_B is fixed to 8, but the channel length N_h is varied from 3 to 6. 130

7.1 Downlink system model. 150
7.2 Total transmitted power versus the angular separation ϕ. 156
7.3 Histogram of the constraints versus the normalized constraint value for $\epsilon = 0.2$ and $\phi = 7°$. 156
7.4 Histogram of the constraints versus the normalized constraint value for $\epsilon = 0.15$ and $\phi = 6.5°$. 157
7.5 Total transmitted power versus minimum required SINR for values of ϵ in the range $\{0.05, \ldots, 0.2\}$ and $\phi = 7°$ 157
7.6 Total transmitted power versus the angular separation ϕ. 159
7.7 Total transmitted power versus the angular separation ϕ. 160

8.1 Total transmitted power versus target SINRs. 170

8.2 Probability density function for the attained SINR targets when $\epsilon = 0.05$. 170

Acronyms

AWGN	Additive White Gaussian Noise
BD	Block Diagonal
BER	Bit Error Rate
BS	Basestation
CCI	Co-Channel Interference
CDF	Cumulative Distribution Function
CDMA	Code-Division Multiple Access
CP	Cyclic Prefix
CSI	Channel State Information
DFT	Discrete Fourier Transform
DPC	Dirty Paper Coding
FDD	Frequency Division Multiplexing
GSM	Global System for Mobile Communication
IBI	Inter-Block Interference
IDFT	Inverse Discrete Fourier Transform
IID	Independently Identically Distributed
ISI	Intersymbol Interference
IUI	Inter-User Interference
KKT	Karush-Kuhn-Tucker
LP	Linear Program/Programming
MIMO	Multiple-Input Multiple-Output

MIMO-BC	MIMO Broadcast Channel
MIMO-MAC	MIMO Multiple Access Channel
MISO	Multiple-Input Single-Output
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
MRC	Maximal Ratio Combining
MS	Mobile Station
MSE	Mean Square Error
MU-MIMO	Multiuser Multiple-Input Multiple-Output
MUI	Multiuser Interference
MVDR	Minimum Variance Distortionless Response
OFDM	Orthogonal Frequency Division Multiplexing
PSD	Positive Semidefinite
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
QCQP	Quadratically Constrained Quadratic Program/Programming
QoS	Quality of Service
QP	Quadratic Program/Programming
QPSK	Quadrature Phase-Shift Keying
SDMA	Space Division Multiple Access
SDP	Semidefinite Program/Programming
SDR	Semidefinite Relaxation
SER	Symbol Error Rate
SIMO	Single-Input Multiple-Output
SINR	Signal-to-Interface plus Noise Ratio
SISO	Single-Input Single-Output
SLR	Signal to Leakage Ratio

SNR	Signal to Noise Ratio
SOCP	Second-Order Cone Program/Programming
ST	Space-time
STBC	Space-Time Block Codes/Coding
STCS	Space-time Channel Shortening
STEQ	Space-time Equalization
SVD	Singular Value Decomposition
TDD	Time Division Multiplexing
TDMA	Time-Division Multiple Access
TIR	Target Impulse Response
UDD	Uplink-Downlink Duality
V-BLAST	Vertical-Bell Labs Layered Space-Time Architecture
WLAN	Wireless Local Area Network
ZMCSCG	Zero Mean Circularly Symmetric Complex Gaussian

Notational Conventions

$\mathbf{A}, \mathbf{a}, a$	Throughout the thesis boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, and italics denote scalars.
$\{\mathbf{A}\}^T, \{\mathbf{A}\}^H$	Transpose and Hermitian transpose of matrix $\mathbf{A}$ respectively.
$\{*\}$	Convolution.
$\mathrm{E}\{\cdot\}$	Statistical expectation.
$\lvert\cdot\rvert$	Absolute value.
$\mathrm{Re}\{\cdot\}$,$\mathrm{Im}\{\cdot\}$	Real and imaginary part.
$\lVert\mathbf{a}\rVert_2$	Euclidean norm of vector $\mathbf{a}$: $\sqrt{\mathbf{a}^H\mathbf{a}}$.
$\lVert\mathbf{A}\rVert_F$	Frobenius norm of matrix $\mathbf{A}$: $\sqrt{\mathrm{Tr}\{\mathbf{A}^H\mathbf{A}\}}$.
$\{\cdot\}^\dagger$	Pseudo-inverse.
$\mathrm{diag}\{\mathbf{a}\}$	Diagonal matrix with vector $\mathbf{a}$ on its main diagonal.
$\mathbf{A} \succeq 0$	$\mathbf{A}$ is positive semidefinite.
$\mathbf{I}, \mathbf{0}$	Identity and all zeros matrices.
$\mathrm{rank}\{\mathbf{A}\}$	Rank of the matrix $\mathbf{A}$.
$\mathrm{null}\{\mathbf{A}\}$	Vectors in the null space of matrix $\mathbf{A}$.
$\mathrm{Tr}\{\mathbf{A}\}$	Trace of $\mathbf{A}$.
$\mathcal{P}_{\max}\{\mathbf{A}\}, \mathcal{P}_{\min}\{\mathbf{A}\}$	Eigenvector corresponding to the largest and smallest eigenvalue of matrix $\mathbf{A}$ respectively.
$\mathcal{P}_{\mathrm{gen}}\{\mathbf{A}, \mathbf{B}\}$	Eigenvector corresponding to the largest generalized eigenvalue of the matrix pair $\mathbf{A}$ and $\mathbf{B}$.
$\lambda_{\max}\{\mathbf{A}\}$	Maximum eigenvalue of matrix $\mathbf{A}$.
$\sim$	Distributed according to.
$\mathcal{CN}(\mathbf{m}, \mathbf{C})$	Complex circularly symmetric Gaussian vector distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$.
$\mathbf{A}^{1/2}$	Hermitian square root of the Hermitian matric $\mathbf{A}$, i.e., $\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{A}$.
$\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_{++}$	The set of real, nonnegative real, positive real numbers respectively.

$\mathbb{R}^{n\times m}$, $\mathbb{C}^{n\times m}$	The set of $n \times m$ matrices with real- and complex-valued entries.
$\mathbb{S}^n$	The set of Hermitian $n \times n$ matrices $\mathbb{S}^n \triangleq \{\mathbf{A} \in \mathbb{C}^{n\times n} \vert \mathbf{A} = \mathbf{A}^H\}$.
$\mathbb{S}^n_+$	The set of Hermitian positive semidefinite $n \times n$ matrices $\mathbb{S}^n_+ \triangleq \{\mathbf{A} \in \mathbb{C}^{n\times n} \vert \mathbf{A} = \mathbf{A}^H \geq \mathbf{0}\}$.
$\mathbb{S}^n_{++}$	The set of Hermitian positive definite $n \times n$ matrices $\mathbb{S}^n \triangleq \{\mathbf{A} \in \mathbb{C}^{n\times n} \vert \mathbf{A} = \mathbf{A}^H > \mathbf{0}\}$.
sup, inf	Supremum(lowest upper bound) and infimum (highest upper bound).
∇f, $\nabla^2 f$	Gradient and Hessian of function f.
$\cap$, $\cup$	Intersection and union.
dom f	Domain of function f.
$\{\cdot\}^*$	Optimal value.
λ,$\boldsymbol{\lambda}$	Lagrange multiplier (vector) associated with inequality constraint(s).
ν,$\boldsymbol{\nu}$	Lagrange multiplier (vector) associated with equality constraint(s).
N_T,N_{R_i}	Number of antennas at transmitter and i^{th} receiver.
K	Number of users.
N_h	Length of the channel impulse response.
N_f	Length of the space-time filter impulse response.
N_b	Length of the target impulse response.
N_c	Number of carriers.
N_{CP}	Length of the cyclic prefix.
N_B	Number of carriers in block (to be quantized).
$\{\cdot\}^{\text{UL}}$, $\{\cdot\}^{\text{DL}}$	Quantities in the uplink and downlink.
s, $\mathbf{s}$	Data symbol and vector.
$\mathbf{w}$	Complex beamformer (transmit) vector.
x, $\mathbf{x}$	Transmit signal sample and vector.
$\mathbf{g}$	Complex beamformer (receive) vector.
$\mathbf{h}, \mathbf{H}$	Complex channel gain vector and matrix.
n,$\mathbf{n}$	Additive white Gaussian noise sample and vector.
y, $\mathbf{y}$	Received signal sample and vector.
$\mathbf{R}$	True channel covariance matrix.
$\hat{\mathbf{R}}$	Known channel covariance matrix.
Δ	Unknown uncertainty (mismatch) matrix for the channel covariance matrix.

Chapter 1

Introduction

Multiuser wireless communication systems such as the wireless cellular networks for e.g. global system for mobile communications (GSM) [10] or the wireless local area network (WLAN) for e.g. IEEE 802.11 (Wi-Fi) [11] impose several challenges such as providing higher data rates, mitigating interference and maintaining a level of required quality of service (QoS). Furthermore, the ever increasing demand for simpler and lighter mobile units [1] has resulted in a need to move the receiver complexity to the basestation (BS) in a cellular network. The transceiver has to pre-process the data prior to transmission to combat the phenomena of the physical wireless channel and inter-user interference (IUI) to support the users in the system. The pre-processing at the transceivers may require the knowledge of channel state information (CSI). In practice, the CSI is fed back from receiver terminals. The CSI available at the transceiver would normally be in error due to feedback delay, feedback error, quantization, etc.

Several advanced signal processing techniques exist which assume perfect CSI at the transmitter. However, in practice, this is an unrealistic assumption. The performance of these algorithm is known to degrade as the quality of CSI available at the transmitter worsens. Therefore, in the presence of such errors, *robust* techniques

[1] These refer to the end user product such as the mobile phones, laptops, palmtops, etc.

are needed which will incorporate for the errors in the CSI into the design.

The aim of this thesis is to provide an extensive literature review of fundamentals and the key topics in multiuser downlink communication systems and to propose several transceiver optimization techniques to improve the performance of wireless communications systems.

1.1 The Wireless Communications Channel

The wireless communication channel as a medium poses several challenges for high-speed communications. A signal propagating through the wireless communication channel suffers from various different impediments before arriving at the receiver. Moreover, *noise* [2] and *interference* [3] also impair the quality of the received signal. The receiver applies several advanced signal processing techniques [4] to overcome the effect of these phenomena in order to construct the original signal.

Path loss is one of the main impediment. It results in the reduction in power density of the transmitted signal as it propagates through the medium and it may be due to effects such as free-space loss, refraction, diffraction, reflection and absorption. The terrain contours, urban or rural environments, foliage, distance between the transmitter and the receivers also influence path loss.

On the other hand, the wireless propagation channel also induces *fading*, in which the received signal exhibits random fluctuations in signal level. These fluctuations are induced as a result of the superposition of the multiple copies of the original transmitted signal [5], each traversing a different path and experiencing differences in attenuation, delay and phase shift. Fading is usually modeled as a time-varying

[2]This refers to Johnson Nyquist noise and its amplitude is assumed to have a Gaussian probability density function

[3]This mainly refers to signals arriving form undesired users in the system.

[4]These techniques may be applied at the transmitter to pre-remove the effect of channel, interference etc.

[5]This phenomenon is known as the multipath.

random change in the amplitude and phase of the transmitted signal and the terms *slow* and *fast* fading refer to the rates at which the magnitude and phase change occur. Slow fading is caused by events such as *shadowing*, where a large obstruction such as large building obscures the main signal path between the transmitter and the receiver. Fast fading however, is caused due to the scattering and reflection off objects between the transmitter and receiver.

In a multipath channel, the transmitted signal arrives through multiple paths (different gains) and hence arriving at the receiver at different times (different delays). This causes time dispersion of the transmitted signal. A measure of this time dispersion is called the channel delay spread τ_{max} [12, 13]. This delay spread causes frequency selective fading which can be characterized by the coherence bandwidth of the channel $B_c \approx 1/\tau_{max}$ [12, 13]. It measures the frequency bandwidth over which the channel remains correlated. The channel is known as *frequency selective* if the coherence bandwidth of the channel is less than the signal bandwidth else the channel is known as *frequency flat.*

Scatterers and transmitter/receiver mobility induce time variations in the propagation channel and this time variation is characterized by doppler spread F_d i.e. a signal of frequency v_c hertz spreads over a finite spectral bandwidth ($v_c \pm v_{max}$) [12,13]. The time selective fading is characterized by the coherence time which is inversely proportional to the doppler spread and is given as $T_c \approx 1/F_d$ [12, 13]. This gives the measure of time over which the propagation channel remains correlated i.e. measures how fast the channel changes in time - the larger the coherence time, the slower the channel fluctuations.

As mentioned earlier, a signal propagating in a multipath channel experiences different gains over different paths before arriving at the receiver through different angle of arrivals (AOA). Hence, these paths add up differently at different point in space i.e. the spatial location of the antenna. This effect is known as space-

selective fading. It mainly depends on the angle spread σ_θ of the arriving paths and is characterized by the coherence distance, $D_c \approx 1/\sigma_\theta$ [12, 13] and it refers to the distance over which the channel impulse response remains strongly correlated. Techniques including beamforming and spatial multiplexing could be exploited using the space selective behavior as it induces channel variations from antenna to antenna. The larger the angle spread - the shorter the coherence distance.

The transceiver thus needs to be equipped with appropriate signal processing techniques to overcome these channel induced effects. Moreover, in a multiuser environment, where users either compete or corporate for the available resources [6], the transceiver must be able to optimally allocate to the most suitable user or optimally share the resources among all users in the system.

1.2 Multiple Antennas

Wireless systems continue to strive for even higher data rates with limited bandwidth, power and complexity. The use of multiple antennas at both the transmitter and the receiver has opened up a new domain that can be exploited to provide significant increase in channel capacity. Pioneering work by [14–16] predicted remarkable increase in spectral efficiency which leads to an explosion in the research activity to characterize the capacity limits of multiple-input multiple-output (MIMO) channels for both point-to-point (single user) and multiuser systems. An excellent overview of Shannon capacity for MIMO system for both single and multiuser has been provided in [17].

Space-time processing is the core idea in MIMO systems. In such a system the time is complemented with the spatial dimension of multiple spatially distributed antennas, hence it is generally viewed as an extension to a popular technology known as *smart antennas*. See [12, 13, 18] for a detailed overview of the recent advances in

[6] This could be time, frequency, code etc.

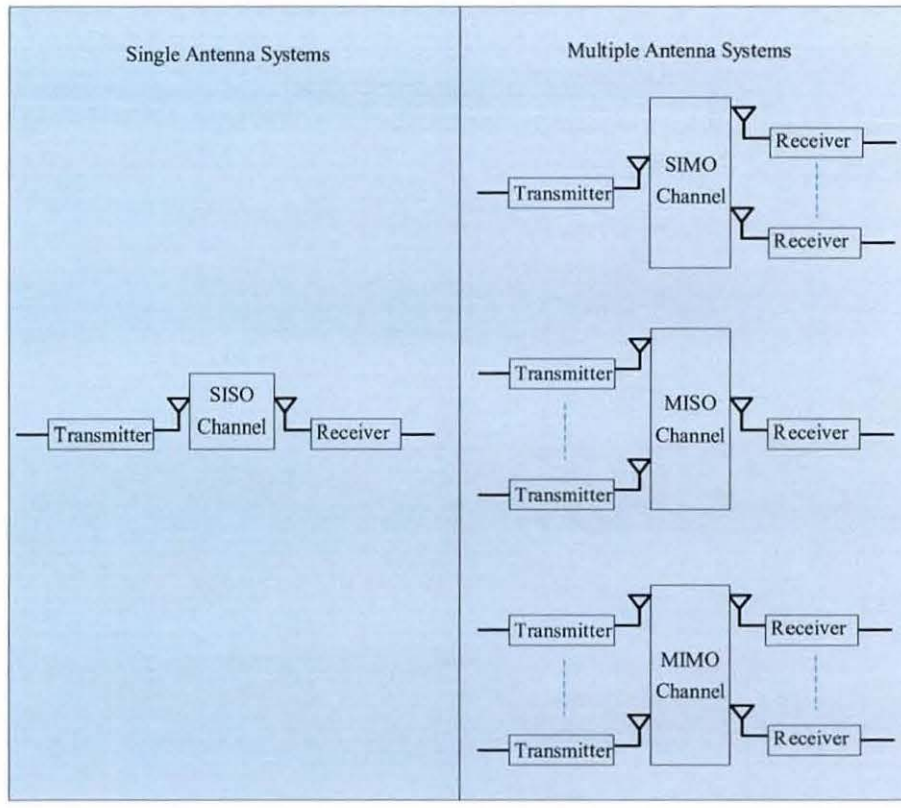

Figure 1.1: Multiple antenna system models.

aspects of space-time processing MIMO communications.

1.2.1 Gains and Features of MIMO System

A key feature of MIMO systems can be summarized as its ability to use the multipath propagation in the favor of the user [19]. It uses fading and multipath delay spread for multiplying data rates. The success of MIMO is largely due to the fact that it enhances the system performance with no extra spectrum, only hardware complexity is added. In an attempt to contribute and to improve the performance of MIMO communications systems in particular multiuser MIMO systems, we will investigate and propose various techniques.

We begin by looking at different configuration of a multiple antennas systems. Fig. 1.1 presents the different configurations for the multiple antenna systems, which in-

cludes single-input single-output (SISO), single-input multiple-output (SIMO), multiple-input single-output (MISO) and MIMO systems. Using multiple antenna arrays at the transmitter, receiver or both gives rise to improvement in the system performance. This increase in system performance is due to gains that antenna arrays are able to achieve in MIMO systems.

Array gain or beamforming gain can be achieved by increasing the average signal to noise ratio (SNR) at the receiver through the process of coherently combing the signals which arrive at the receiver antennas with different amplitudes and phases. This process increases the average received signal power which is proportional to the number of receiver antennas. Hence array gain offers improved coverage and QoS. Beamforming gain can be easily characterized as a shift in the BER curve (plotted against the received power per antenna) due to the gain in SINR. See [19, 20] for further details on array gain in MIMO systems.

Due to channel fluctuations the signal suffers from fading and to combat this fading, diversity techniques are exploited. Both transmit and receive diversity are exploited for multiple antenna systems. Receive antenna diversity can be used in SIMO channels [21]. On the other hand, transmit diversity can be achieved in MISO channels with or without the knowledge of channel at the transmitter. See [22,23] for transmit diversity schemes based on space-time block codes (STBC) in the absence of channel knowledge at the transmitter. Whereas in MIMO channel a combination of receive and transmit diversity can be achieved. Diversity is characterized by a number of independently fading branches, more formally known as diversity order. Table 1.2.1 lists the diversity order achieved for a given system configuration, assuming that all channels from transmit antenna to receive antenna and vice versa fade independently. Diversity gain provides improvement in link reliability and this gain can be easily characterized as the increase in the BER (plotted against received power per antenna on a logarithmic scale) slope in the low BER region.

Table 1.1: System configurations and diversity order achieved.

System Configuration	Diversity Order
SIMO	Number of receive antennas (N_R).
MISO	Number of transmit antennas (N_T).
MIMO	Product of number of receive and transmit antennas ($N_T \times N_R$).

Spatial multiplexing offers for the same bandwidth and power consumption a linear increase in the transmission rate (i.e. capacity). Unlike array and diversity gain spatial multiplexing is only possible in MIMO systems [14,16,24]. An an example, let us consider a system with N_T transmit and N_R receive antennas. In such a system, a bit stream can be multiplexed into N_T bit streams with a rate of $1/N_T$, modulated and transmitted simultaneously through the N_T transmit antennas, which can then be demultiplexed and demodulated using $N_R \geq N_T$ receive antennas and combined together to form the original signal stream. Similarly in a multiuser scenario, spatial multiplexing can used in both uplink and downlink, also known as space division multiple access (SDMA). In uplink multiple users (K) can simultaneously transmit data through the same channel to a BS equipped with multiple receive antennas $N_T \geq K$. Similarly in downlink, a BS can transmit data to multiple users using spatial filtering, which allows users to decode their signals independently. Spatial multiplexing allows increase in capacity and is proportional to the number of antennas at the BS and the number of users. An example of such a transmission scheme over a MIMO is often referred to as V-BLAST (Vertical-Bell Labs Layered Space-Time Architecture) [25].

Co-channel interference (CCI) arises due to reuse of the frequency in wireless channels [12,26]. Using multiple antennas the effect of CCI can be mitigated, however it requires knowledge of CSI for both desired signal and the co-channel signals. For multiuser downlink MIMO systems, with the knowledge of CCI for the channels between the BS and the users, techniques such as multiuser beamforming can be

used to increase network capacity and to reuse the network resources. See [27] for an example of interference mitigation in multiuser MIMO systems.

1.2.2 Tradeoffs Between Gains

In a multi-antenna system both beamforming and diversity gain may be achieved simultaneously by coherently combining the received signal. Beamforming gain requires CSI at the transmitter and is independent of the channel statistics whereas diversity gain may be achieved independently of the CSI, however, requires the statistical behavior of the channel. Thus, there is no tradeoff between beamforming and diversity gain.

On the other hand, beamforming and multiplexing gain have a fundamental tradeoff. Since maximum beamforming requires only the maximum singular value of the channel may be used [28], whereas to achieve maximum multiplexing a subset of channel singular values (sub-channels) are used based on water-filling solutions [29].

Finally, although diversity and multiplexing gain can be achieved simultaneously, they share a fundamental tradeoff in terms of how much of each gain can be extracted in a communication system [30]. It has been shown that diversity-multiplexing tradeoff is essentially the tradeoff between the error probability and the data rate of the system.

The deployment of multiple antennas at the transceiver allows us to exploit these gains, which may be achieved fully simultaneously or share a tradeoff. In this thesis, we consider a wireless system where the transceiver is equipped with an array of multiple antennas, however, each user in the system could either have single antenna or multiple antennas.

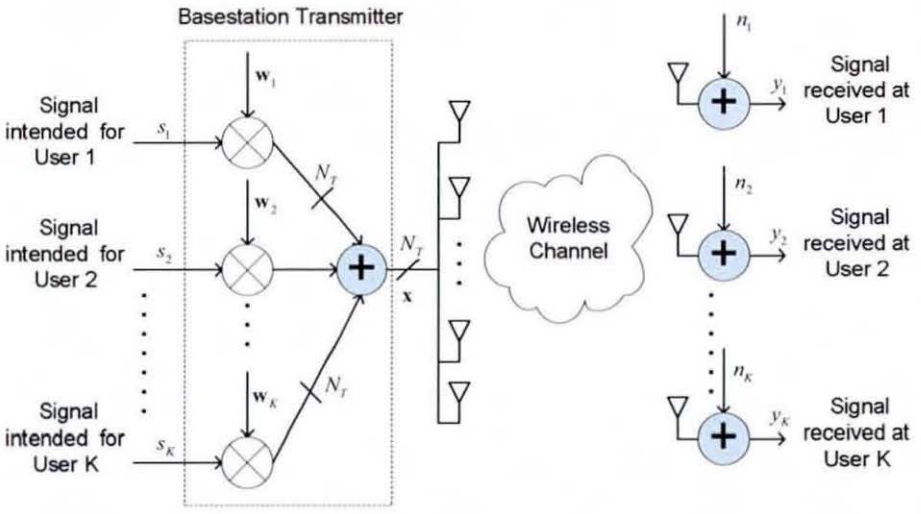

Figure 1.2: Block diagram of a multiuser downlink system.

1.3 Multiuser Downlink System

In this thesis we solely study the problem of spatial multiplexing, or in particular downlink beamforming. Let us begin by looking at the basics of this problem. Fig 1.2. depicts a basic block diagram of a multiuser downlink system. In such a system, a BS equipped with N_T antennas simultaneously transmits independent data streams to K decentralized users. Each user could have either single or multiple antennas. To do this, the BS multiplexes the data intended for different users and transmits a vector $\mathbf{x}(t) \in \mathbb{C}^{N_T \times 1}$ at time t as

$$\mathbf{x}(t) = \sum_{j=1}^{K} \mathbf{w}_j s_j(t), \tag{1.1}$$

where $\mathbf{w}_j \in \mathbb{C}^{N_T \times 1}$ and s_j are the beamforming vector and the data symbol intended for the jth user. The problem of downlink beamforming is to design a set of beamforming vectors and perform spatial multiplexing, so that a pre-defined threshold on QoS [7] is achieved for each user. Assuming single antenna users, the signal received by the jth user at time t is given as

[7] This could be signal-to-interference plus noise ratio, mean square error, bit error rate, capacity etc.

$$y_j(t) = \mathbf{h}_j\mathbf{x}(t) + n_j(t), \tag{1.2}$$

where $\mathbf{h}_j \in \mathbb{C}^{1\times N_T}$ is the complex (flat fading) channel vector between the BS and the jth user and n_j is additive white Gaussian noise (AWGN) at j^{th} user with variance σ_n^2.

1.3.1 Frequency Selective Channels

If the channels seen between the BS and the users are frequency selective, space-time filters are required to perform spatial multiplexing at the transmitter. Defining $\mathbf{W}_j = [\mathbf{w}_j(0), \mathbf{w}_j(1), \ldots, \mathbf{w}_j(N_f - 1)]$ as the space-time filter for the jth user, where N_f is the filter length, the transmit signal from the BS at time t can be written as

$$\mathbf{x}(t) = \sum_{j=1}^{K}\sum_{n=0}^{N_f-1} \mathbf{w}_j(n)s_j(t-n). \tag{1.3}$$

Assuming that the channel has an impulse response of length N_h and users are equipped with a single antenna, the signal received by the jth user can be written as

$$y_j(t) = \sum_{m=0}^{N_h-1} \mathbf{h}_j(m)\mathbf{x}(t-m) + n_j(t), \tag{1.4}$$

where $\mathbf{h}_j(m) \in \mathbb{C}^{1\times N_T}$ is the complex channel vector for the mth delay between the BS and the jth user. In this thesis we provide an extensive literature review and propose various techniques on the beamformer design for flat fading channels and space-time filter design for frequency selective fading channels assuming perfect CSI is available at the transmitter.

1.3.2 Non-Perfect Channel State Information

A wireless communication system may employ time division multiplexing (TDD) or frequency division multiplexing (FDD) as its duplexing scheme. TDD uses a single frequency to transmit signals in both the downstream (forward link) and upstream (reverse link) directions. In TDD, the *reciprocity principle* is normally exploited to obtain the CSI at the transmitter. However, a FDD based system uses different frequencies for downstream and upstream transmissions. Thus, the reciprocity principle in general cannot be exploited in FDD. As a consequence, the CSI is estimated by the receiver on the forward link and then fed back through a feed back channel from the receiver to the transmitter on the reverse link. This, however, requires additional bandwidth and power, but only a limited amount of resources are allocated to such a feed back channel. As a result, the CSI available at the transmitter is generally in error, these may arise due to feedback error, feedback delay, quantization, estimation errors and so on. In the presence of such errors, robust techniques are generally needed to incorporate for the CSI errors. Let us define a channel covariance matrix as $\mathbf{R} = \mathrm{E}\{\mathbf{h}^H\mathbf{h}\} \in \mathbb{S}_+^{N_T \times N_T}$. Throughout the thesis, we will use the following model

$$\mathbf{R}_{\text{Actual}} = \mathbf{R}_{\text{Known}} + \Delta, \tag{1.5}$$

where $\mathbf{R}_{\text{Actual}}$ is the actual (error-free) channel covariance matrix, $\mathbf{R}_{\text{Known}}$ is the known channel covariance matrix at the transmitter and Δ is the unknown uncertainty matrix, where its Frobenius norm has assumed to be bounded above by a known constant, i.e. $\|\Delta\|_F \leq \varepsilon$. Such an upper bound can be obtained in practise using a priori statistical information. In this thesis we will provide an extensive literature review on robust beamforming techniques which incorporate for the CSI errors in the design and propose novel techniques using advanced convex optimization theory.

1.4 Figures of Merit

In this section we discuss the common measures of performance that will be readily used throughout this thesis.

1.4.1 Mean Square Error

The mean square error (MSE) is defined as

$$\mathrm{MSE} \triangleq \mathrm{E}\{|\hat{x} - x|^2\} \tag{1.6}$$

where $\hat{x}$ corresponds to the estimation of the transmitted symbol x and is bounded by $0 < \mathrm{MSE} \leq 1$. An MSE in the vicinity of 0 corresponds to a good estimate of the transmitted symbol, thus systems are usually optimized to minimize the MSE.

1.4.2 Signal to Interference-plus-Noise Ratio

In any communication system the received signal may be written as $y = x + i + n$, where x corresponds to the desired signal component, n corresponds to interference and i corresponds to noise. Thus, the signal to interference-plus-noise ratio (SINR) for a communication system is defined as

$$\mathrm{SINR} \triangleq \frac{\mathrm{E}\{|x|^2\}}{\mathrm{E}\{|i|^2\} + \mathrm{E}\{|n|^2\}} \tag{1.7}$$

and is bounded by $0 < \mathrm{SINR} < \infty$. The higher the SINR the better it is for a communication link, hence the systems are usually optimized to maximize the SINR.

1.4.3 Bit Error Rate

Bit error rate (BER) is known to be the ultimate performance measure for a digital communication system and it is defined as the bit error probability. Assuming $i + n$

in (1.7) is Gaussian distributed, then the symbol error probability P_e can be defined as [31]

$$P_e = \alpha \mathcal{Q}\left(\sqrt{\beta}\text{SINR}\right) \tag{1.8}$$

where $\mathcal{Q}$ is the $\mathcal{Q}$-function defined as $\mathcal{Q}(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\lambda^2/2} d\lambda$ [31], α and β are constants that depend on the signal constellation [31]. The BER can be approximate from P_e as

$$\text{BER} \approx P_e/k \tag{1.9}$$

where $k = \log_2 M$ is the number of bits per symbol and M is the constellation size. The BER is bounded by $0 < \text{BER} \leq 0.5$. A BER of 0.5 corresponds to an outage probability of 1, thus normally system are usually optimized to minimize the BER.

1.5 Thesis Outline

Numerical optimization techniques has played a major role in various research disciplines over the last few decades. In particular, convex optimization has emerged as an important tool and found applications in almost all fields including signal processing, communications, finance, etc. It is a general view that once a problem is put into a convex optimization form, it can be naturally solved due to the advancements in the numerical optimization techniques. Chapter two outlines the important concepts to understand the fundamentals of convex optimization theory. The chapter describes the most generic classes of convex problems, namely, linear programming (LP), quadratic programming (QP), second-order cone programming (SOCP) and semidefinite programming (SDP). The chapter also lays the foundations to Lagrange

optimization and Karush-Kuhn-Tucker (KKT) conditions, which allow the convex problems to be solved both analytically (if possible) and numerically. Finally, the interior-point methods are introduced. These are the backbone of all the conic optimization software available in the market, for example, CVX [32], SeDuMi [33], etc.

Various diversity techniques including time, frequency, polarization and space diversity are introduced in chapter three. An extensive literature review on both receiver and transmitter spatial diversity techniques is presented. The particular attention is on spatial diversity techniques at the transmitter, as they form the core of this thesis. Spatial diversity at the transmitter can be broken down into techniques requiring CSI, such as beamforming and techniques which do not require CSI, such as STBC.

Chapter four proposes a downlink beamforming technique based on maximizing signal to leakage ratio (SLR). Herein we propose an iterative scheme where instead of maximizing the SLR at the frontend of the receiver, the SLR at the output of the receiver is maximized. We show high gain in performance as the iterations are increased, however this gain vanishes as the noise floor is reached. Simulations results are presented for both flat fading and frequency selective channels, where we employ orthogonal-frequency division multiplexing (OFDM) at the transmitter.

The problem of separating multiple users transmitting data in the uplink is considerably easy as compared to the downlink. This is because, for the uplink the weight vectors can be designed individually for each user. Thus, it is desirable to establish a link between the uplink and downlink paradigms. Chapter five discusses this link which is more formally known as the uplink-downlink duality. The duality has been well established for flat fading channels, however we extend this to frequency selective channels. In such a scenario, the beamformers are replaced by space-time filters, which could be viewed as multiple-tap beamformers. Finally, we propose partial

equalization or channel shortening based spatial multiplexing schemes. In this approach, the effective channel is shortened to a desired target impulse response (TIR). The performance of channel shortening based space-time filters is compared to that of a full equalization based scheme.

In chapter six, we propose a robust counterpart to the SLR based beamformer using worst-case performance optimization. Analytical expressions are derived for the diagonal loading parameters, which appear in the cost function for the SLR based robust beamforming. Simulation results confirm the superior performance of the robust solution over the non-robust solution for a flat fading environment. An application of this scheme for frequency selective channels using OFDM is also presented. Here, we assume that instead of feeding back CSI for all N_c frequency bins, the CSI is quantized over $N_B < N_c$ blocks and fed back to the transmitter.

Chapter seven builds upon the well known technique of robust downlink beamforming using worst-case performance optimization. The worst-case performance optimization violates the positive semidefinite (PSD) constraints on the channel covariance matrices. This is an immediate consequences of negative diagonal loading in the cost function. The chapter builds upon this technique by incorporating for the semidefinite constraints. Two different SDP based methods are proposed, which are approximated into convex problems using semidefinite relaxation (SDR). The resulting designs are shown to be less conservative as compared to the worst-case performance optimization which does not incorporate for the PSD constraints.

The problem of conventional downlink beamforming can be expressed as minimizing the total sum power subject to satisfying some QoS constraints. However, in practice each antenna or a sub-group of antennas may be equipped with their own individual power amplifiers. In chapter eight, we discuss the problem of downlink beamforming with per-antenna power constraints and QoS constraints. We propose a robust solution based on worst-case performance optimization.

Finally conclusions are drawn in chapter nine and a brief outline of the possible future research directions is also provided.

Chapter 2

Literature Review

Multiple-input multiple-output (MIMO) communication system, where multiple antennas are employed at both the transmitter and the receiver, has proven to be an extremely promising technique in enhancing data rates of a wireless communication system without requiring additional spectrum. Moreover, multiple antennas in MIMO systems can be used to achieve diversity [22], beamforming [34] or spatial multiplexing [25].

Random fluctuations of signals, for example in time, frequency and space is the main impairment in wireless channels [12]. In order to improve reception at the receiver, diversity techniques are used, as they provide the receiver with multiple independent look of the signal. These multiple looks are more formally known as *diversity branches.* The probability of all the branches being in fade together decreases as the number of diversity branches increases. Therefore, with a high probability, the receiver can detect data reliably as the probability of at least one branch or link with a good signal increases.

Wireless channels in general exhibit all or a combination of space-time-frequency fading. The actual diversity captured by the receiver depends on the inherent diversity that is available in the channel, the coding and the modulation scheme used for

transmission, and the receiver design itself. This inherent diversity depends on the size of the codeword used, the number of transmit antennas N_T, the number of receiver antennas N_R, the coherence bandwidth B_c and the coherence time T_c [12,13]. There are many ways to obtain diversity. The utilization of time/frequency diversity incurs at the expense of time (in the case of time diversity) and bandwidth (in case of frequency diversity) due to introduction of redundancy. Spatial diversity is an attractive alternative that does not sacrifice time or bandwidth. Moreover, it provides array gain or increased average SNR [12,13]. This, however may incur at an expense of additional hardware complexity. However in comparison to system resources such as the time and bandwidth, hardware complexity is diminutive. The exact nature of the scheme that extracts spatial diversity depends on the antenna configuration (SIMO, MISO or MIMO).

2.1 Single User Spatial Diversity Techniques

2.1.1 Receiver Antenna Diversity

Receiver diversity requires multiple receiver antennas that are spaced far enough apart so that the channel between each transmit and receive antenna can be assumed uncorrelated. As shown in Fig.2.1 the received signal is combined using a weight vector $\mathbf{w} \in \mathbb{C}^{N_R \times 1}$. The central idea in receiver diversity techniques is to design the weight vector $\mathbf{w}$ to reduce the effect of fading and to maximize SNR. Let us consider a SIMO system shown in Fig. 2.1. The input-output relation for the channel can then be expressed as

$$y_i = \sqrt{E_s} h_i s + n_i, \quad i = 1, 2, ..., N_R, \tag{2.1}$$

where y_i is the received signal at the ith receiver antenna, s is the transmit symbol

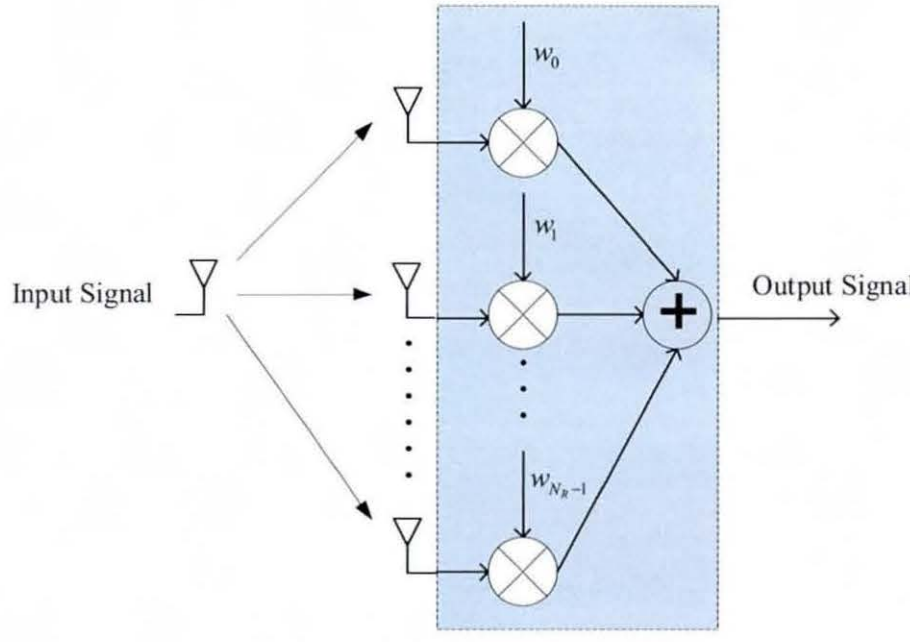

Figure 2.1: Receiver diversity in a SIMO system.

with unit variance, E_s is the average symbol energy in each link and n_i is the additive white Gaussian noise (AWGN) with zero mean and variance σ_n^2. The channel gains $h_i, i = 1, 2, ..., N_R$, are assumed to be zero-mean complex Gaussian random variables with unity variance, and the gains are independent across the antennas. Assuming that the complex channel gains are perfectly known to the receiver, the receiver performs *maximal ratio combining* (MRC) [31, 35] to maximize the SNR as

$$z = \sqrt{E_s}\mathbf{h}^H\mathbf{h}s + \mathbf{h}^H\mathbf{n} \tag{2.2}$$

where $\mathbf{h} \in \mathbb{C}^{N_R \times 1}$,$\mathbf{n} \in \mathbb{C}^{N_R \times 1}$ and $\mathbf{w} = \mathbf{h}^H$ is used, which is effectively a *match filter* and known to be optimal for single-user scenario. The post-MRC SNR is then given by

$$\text{SNR} = \rho \sum_{i=1}^{N_R} |h_i|^2, \tag{2.3}$$

where $\rho = E_s/\sigma_n^2$ is the average SNR per receive antenna. Assuming that the receiver employs a Maximum Likelihood (ML) detector, the symbol error rate (SER) is given by [36]

$$P_e \approx \bar{N}_e Q\left(\sqrt{\frac{\text{SNR}d_{min}^2}{2}}\right), \tag{2.4}$$

where $\bar{N}_e$ and d_{min} are the nearest neighbors and the minimum distance of the constellation used. Combining (2.3) and (2.4), it has been shown in [31] that the average SER $\bar{P}_e$ is given by

$$\bar{P}_e \leq \bar{N}_e\left(\frac{1}{1+\rho d_{min}^2/4}\right)^{N_R} \leq \bar{N}_e\left(\frac{\rho d_{min}^2}{4}\right)^{-N_R}. \tag{2.5}$$

Hence, a diversity order equal to the number of receiver antennas N_R is achieved, since in the absence of diversity ($N_R = 1$), $\bar{P}_e \leq \bar{N}_e(\rho d_{min}^2/4)^{-1}$. Diversity affects the slope of the SER vs SNR curve on a log-log scale. The magnitude of the slope equals the diversity order N_R. This is demonstrated in 2.2, where we plot SER as a function of SNR for a fading link, with BPSK modulation for $N_R = 1, 2$, and 4. Moreover, the average postprocessing SNR is given by

$$\bar{\text{SNR}} = N_R.\text{SNR}.$$

Hence, in addition to the increased diversity order, the average SNR is also enhanced by a factor of N_R over a single-receive antenna case due to the array gain. We have seen that full diversity and array gain are achieved using receiver diversity. Other popular receiver diversity techniques include *selection combining* [13] and *equal gain combining* [13].

Employing multiple antennas at the receiver adds to the complexity and cost of the receiver. Therefore, to keep the receiver structure simple, transmit diversity is

preferred.

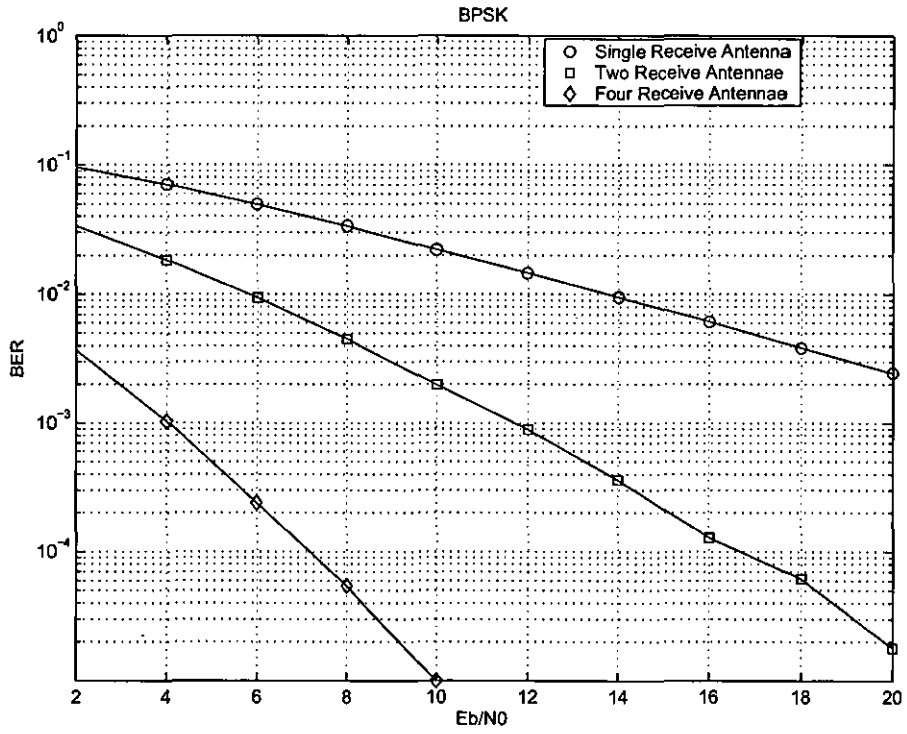

Figure 2.2: Performance of receive diversity technique (MRC) for various number of antennas.

2.1.2 Transmit Antenna Diversity

In transmitter diversity techniques, the signal is pre-processed or pre-coded prior to transmission [22, 23]. Transmit diversity techniques can be classified into two main categories: techniques for which the channel state information (CSI) is available at the transmitter and techniques for which the CSI is not available at the transmitter.

Channel Unknown to the transmitter

A well known and fully developed transmitter diversity scheme is *space-time block coding* (STBC), originally proposed by Alamouti [22] for a MISO (two transmit and one receive antenna) flat fading channels and later extended to frequency-selective

channels in [37, 38]. The Alamouti scheme extracts a diversity order of 2 (full N_T diversity).

Channel Known to the transmitter

Transmit - MRC

A well known technique for transmitter diversity is known as transmit maximal ratio combining (transmit-MRC) [39,40]. The signal is transmitted from each antenna after being weighted appropriately, so that the signals arrive in phase at the receiver antenna and add coherently. The signal at the receiver can be written as

$$y = \sqrt{\frac{E_s}{N_T}}\mathbf{h}\mathbf{w}s + n, \tag{2.6}$$

where y is the received signal, $\mathbf{w} \in \mathbb{C}^{N_T \times 1}$ is the MRC weight vector, $\mathbf{h} \in \mathbb{C}^{1 \times N_T}$ is the channel gain vector and n is zero mean circularly symmetric complex Gaussian (ZMCSCG) noise with variance σ_n^2. The weight vector is designed such that $\|\mathbf{w}\|_2^2 = N_T$, so that the average total power of the transmitted signal is 1. The weight vector $\mathbf{w}$ that maximizes the SNR is given by [39]

$$\mathbf{w} = \sqrt{N_T}\frac{\mathbf{h}^H}{\sqrt{\|\mathbf{h}\|_F^2}}. \tag{2.7}$$

The SNR at the receiver is given by

$$\text{SNR} = \|\mathbf{h}\|_2^2 \rho \tag{2.8}$$

where $\rho = E_s/\sigma_n^2$ is the average SNR at the receiver antenna. Assuming, that the separation between the antennas is greater than the coherence bandwidth and rich scattering environment, it follows that the average probability of symbol error in the high SNR regime is upper-bounded by

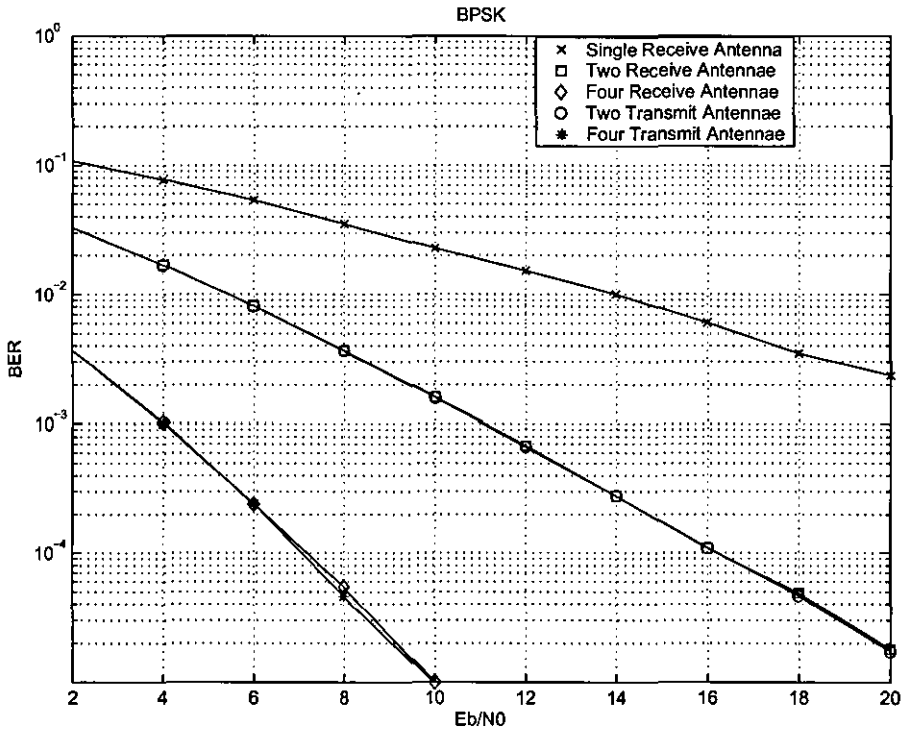

Figure 2.3: Performance of transmit diversity technique (transmit-MRC) for various number of antennas.

$$\bar{P}_e \leq \bar{N}_e \left(\frac{\rho d_{min}^2}{4} \right)^{-N_T}. \tag{2.9}$$

Thus transmit-MRC delivers a diversity order of N_T in the presence of independent and identically distributed (IID) Rayleigh fading. The average received SNR at the receiver defined as

$$\bar{\text{SNR}} = N_T \rho,$$

and is improved by a factor of N_T over a SISO link. Hence when perfect CSI is available to the transmitter, transmit-MRC will deliver array gain and diversity gain. A MISO system with N_T transmit antennas, using transmit-MRC, has the same performance as SIMO system with the same number of receiver antennas (employing receive-MRC). This is illustrated in Fig. 2.3, where we plot SER as a function of SNR, with BPSK modulation and $N_T = 1, 2$, and 4.

MIMO Beamforming

In a MIMO system, to extract spatial diversity a technique known as dominant eigenmode transmission or beamforming can be used. In such a system, the signal is weighted with a weight vector and transmitted through an array of antennas. The received signal vector is then given by

$$\mathbf{y} = \sqrt{\frac{E_s}{N_T}}\mathbf{H}\mathbf{w}s + \mathbf{n}, \tag{2.10}$$

where $\mathbf{y} \in \mathbb{C}^{N_R \times 1}$ is the received signal vector, $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ is the channel matrix, $\mathbf{w} \in \mathbb{C}^{N_T \times 1}$ is the weight vector and $\mathbf{n} \in \mathbb{C}^{N_R \times 1}$ is a spatially white ZMCSCG noise vector. Again $\mathbf{w}$ must be designed such that $\|\mathbf{w}\|_2^2 = N_T$ to maintain the required total transmitted energy. The receiver then weights and sums the received signal at each of the antennas according to

$$z = \mathbf{g}^H\mathbf{y}, \tag{2.11}$$

where $\mathbf{g} \in \mathbb{C}^{N_R \times 1}$ is the receiver weight vector. The SNR at the receiver is then given by

$$\text{SNR} = \frac{\|\mathbf{g}^H\mathbf{H}\mathbf{w}\|_2^2}{N_T\|\mathbf{g}\|_2^2}\rho. \tag{2.12}$$

where $\rho = E_s/\sigma_n^2$ is the SNR per receiver antenna. In order to maximize the SNR, we perform singular value decomposition (SVD) of the channel matrix $\mathbf{H}$ as

$$\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^H. \tag{2.13}$$

It can be verified that the SNR is maximized when $\mathbf{w}$ and $\mathbf{g}$ are the left and right singular vectors corresponding to the largest singular value σ_{max} of $\mathbf{H}$. Hence the input-output relation for the channel reduces to

$$z = \sqrt{E_s}\sigma_{max}s + n \tag{2.14}$$

where n is ZMCSCG noise with variance σ_n^2 and $\sigma_{max} = \sqrt{\lambda_{max}}$, where λ_{max} is the maximum eigenvalue of $\mathbf{HH}^H$. The SNR at the receiver is then given by

$$\text{SNR} = \lambda_{max}\rho. \tag{2.15}$$

The $\bar{P}_e$ for a system employing dominant eigenmode transmission may be upper and lower bounded in the high SNR regime by [12]

$$\bar{N}_e\left(\frac{\rho d_{min}^2}{4min(N_T, N_R)}\right)^{-N_T N_R} \geq \bar{P}_e \geq \bar{N}_e\left(\frac{\rho d_{min}^2}{4}\right)^{-N_T N_R}. \tag{2.16}$$

The above result implies that the SER slope must maintain a slope of magnitude $N_T N_R$, as a function of SNR (on a log-log scale). Hence, we can conclude that the dominant eigenmode transmission extracts a full diversity order of $N_T N_R$ [12]. This is depicted in Fig. 2.4, where we plot SER as a function of SNR, with BPSK modulation for a SISO, MISO using transmit-MRC with $N_T = 2$ and MIMO using eigenmode transmission with $N_T = 2$ and $N_R = 2$.

2.2 Joint Transceiver Design

Joint transceiver design deals with the problem of jointly designing the transmit and receive filters. This problem is commonly known as linear precoding at the transmitter and equalization at the receiver. Here CSI is assumed to be at both ends of the link. The design goal is to select an optimal pair of linear transformations $\mathbf{F}$ (precoder) and $\mathbf{G}$ (decoder) of blocks of the transmit symbols and receive samples, respectively, that operate *jointly* and *linearly* on the time and space dimensions. In all designs, the paradigm of linear precoding/decoding exploits the channel eigendecomposition

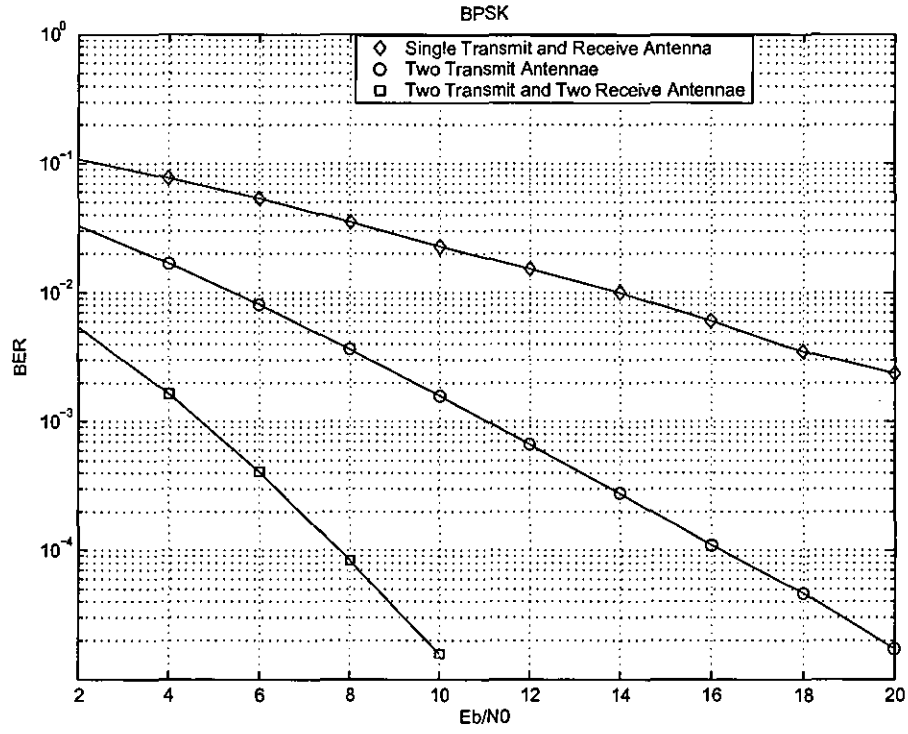

Figure 2.4: Performance of MIMO beamforming.

in constructing the optimal **F**,**G**. The distinct solutions are characterized by how the power is loaded on each channel eigenfunction. The precoder designs capitalize on the available knowledge about the channel by investing the available power wisely in each dimension.

In literature various joint design of linear precoders and decoders according to a variety of design criterion have been considered. Some of these are, [41–43] and [44], where the objective to minimize was defined as sum of the mean square error (MSE) of the channel substreams was under an average power constraint. In [45], the above criterion was generalized by using a weighted sum of the substream MSEs. A maximum SINR criterion with a ZF constraint was proposed in [44]. The original complicated design problem is greatly simplified in these criterion because the channel turns out to be diagonalized by the transmit-receive processing. Several other approaches based on minimizing the determinant of the MSE structure were proposed in [43]. A unified framework for the problem of joint transmit and receiver beamforming for multicar-

rier channels based on convex optimization was presented in [46,47]. Some interesting results for the joint transceiver design for MIMO channel shortening were presented in [48].

2.3 Multiuser Spatial Diversity Techniques

So far, we have looked at a single user wireless system. However, spatial diversity techniques can be used to transmit signal simultaneously to multiple users and we will refer to this system as MU-MIMO. MU-MIMO has been used in many applications such as satellite communications and cellular systems. We will refer to the forward link MU-MIMO channel as MIMO broadcast channel (MIMO-BC) and the reverse link MU-MIMO channel as MIMO multiple access channel (MIMO-MAC).

2.3.1 MIMO-MAC and MIMO-BC

In downlink beamforming a BS simultaneously transmits data to multiple users without compromising the available radio spectrum. The users may be equipped with single antenna, and hence have no ability for spatial discrimination. If the users are equipped with multiple antennas, they could perform some type of interference suppression. However in both cases, we have a MIMO system. Much of the research in MU-MIMO has been focused on the uplink scenario, also known as MIMO-MAC. In such a scenario a BS equipped with multiple antennas can separate the signals arriving from several different users. However, the focus here in this thesis is on downlink MU-MIMO systems, also known as MIMO-BC.

The main issue in MIMO downlink transmission is to maximize the received signal power for each user while minimizing interference leaked to all other users. The solution for achieving this goal is subject to the availability of CSI at the transmitter. CSI is crucial in multiuser scenario interference suppression needs to be performed

prior to transmission.

Consider a BS equipped with N_T antennas simultaneously transmits data to K users each equipped with N_{R_i} antennas (single antenna receiver means $N_{R_i} = 1, \forall i$). The flat fading channel between BS and the ith user is given by

$$\mathbf{H}_i = \begin{bmatrix} h_i^{1,1} & h_i^{1,2} & \dots & h_i^{1,N_T} \\ h_i^{2,1} & h_i^{2,2} & \dots & h_i^{2,N_T} \\ \vdots & \vdots & \ddots & \vdots \\ h_i^{N_{R_i},1} & h_i^{N_{R_i},2} & \dots & h_i^{N_{R_i},N_T} \end{bmatrix} \tag{2.17}$$

Let m_j denote the number of data streams intended for user j. Usually the number m_j depends on the data rate for user j, the total transmit power, the achievable signal-to-interface plus noise ratio (SINR) and the number of receiver and transmitter antennas. The BS wishes to transmit a vector $\mathbf{s}_j \in \mathbb{C}^{m_j \times 1}$ to the jth. Prior to transmission vectors $\mathbf{s}_j, j = 1, 2, \dots, K$ are mapped to a vector $\mathbf{x}$ of size $N_T \times 1$ through a linear mapping as

$$\mathbf{x} = \sum_{j=1}^{K} \mathbf{W}_j \mathbf{s}_j \tag{2.18}$$

where $\mathbf{W}_j = [\mathbf{w}_{1_j}, \mathbf{w}_{2_j}, \dots, \mathbf{w}_{m_j}]$ and $\mathbf{w}_{i_j}$ correspond to the transmit beamformer for the ith symbol of the jth user. We can write the signal received by the jth user as

$$\mathbf{y}_j = \mathbf{H}_j \mathbf{W}_j \mathbf{s}_j + \sum_{k=1, k \neq j}^{K} \mathbf{H}_j \mathbf{W}_k \mathbf{s}_k + \mathbf{n}_j \tag{2.19}$$

where $\mathbf{n}_j$ is assumed to be spatially white noise with $\mathrm{E}\{\mathbf{n}_j(t)\mathbf{n}_j^H(t)\} = \mathbf{I}$.

2.3.2 Single Receive Antenna

In this section we assume that the receivers are equipped with single antenna only. Hence the receivers are unable to perform any type of interference suppression. The

transmitter must perform pre-coding in order to pre compensate for the interference at the receiver terminals. Note, in this case $m_j = 1, j = 1, 2, \ldots, K$.

Channel Inversion

This is the simplest techniques of all and is based on channel inversion. This technique imposes constraints that all the interference terms are zero. In order for this scheme to work, the number of transmitting antennas should be greater than the number of receiving antennas i.e. $N_T \geq K = N_R$ [49, 50]. The precoder at the transmitter should perform

$$\mathbf{x} = \mathbf{H}^{\dagger}\mathbf{s} = \mathbf{H}^{H}(\mathbf{H}\mathbf{H}^{H})^{-1}\mathbf{s} \tag{2.20}$$

where $\mathbf{s} = [s_1, s_2, \ldots, s_K] \in \mathbb{C}^{K\times 1}$ is the signal vector for all K users and $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_K] \in \mathbb{C}^{K\times N_T}$ is the channel matrix for all K users. Channel inversion ideally cancels all the interference so that each user will see only the signal transmitted to it (except white noise) as

$$y_j = s_j + n_j. \tag{2.21}$$

However in [51] it has been shown that this technique does not provide linear capacity growth with $\min(N_T, N_R)$ that should be achievable in the MU channel. This is because for an ill-conditioned channel matrix, at least one of the singular values of $(\mathbf{H}\mathbf{H}^{H})^{-1}$ is very large and therefore requires a large normalization factor which drastically reduces the SNR at the receivers.

Regularized Channel Inversion

The problems in channel inversion simply arise due to the stringent requirement that all the interference be zero. We could allow a limited amount of interference

i.e. regularizing the inverse in the zero-forcing (ZF) filter above. This technique is known as regularized channel inversion. Recently such a technique based on minimum mean square error (MMSE) criterion was proposed in [51]. This is achieved at the transmitter as follows

$$\mathbf{x} = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H + \alpha\mathbf{I})^{-1}\mathbf{s} \tag{2.22}$$

where, $\alpha = K/P$ is the regularization parameter and P is the total transmit power. This technique maximizes the SINR and results in linear increase in the capacity.

Vector Modulo Pre-coding

Even with regularized channel inversion there is still a significant gap in the performance offered by regularization and the sum capacity bound. Techniques based on so called dirty paper coding (DPC) [52, 53] have shown to approach sum capacity (in some cases it could even achieve it). DPC techniques generally employ non-traditional techniques such as non-linear coding and high dimensional lattices, and they are often difficult to implement. For these techniques, the symbol stream itself is coded rather than a spatial filter (beamformer) to mitigate the inter-user interference. However the focus of this thesis is on linear processing techniques and we will not discuss techniques based on DPC any further. For further reading, the reader is referred [52–54].

2.3.3 Multiple Receive Antennas

A natural extension to MU-MIMO would be to employ multiple antennas at the receivers. This allows transmission of parallel data streams to multiple users and enables the receivers to have some degree of freedom to suppress interference. Transmission of parallel data streams to multi-antenna receivers has been accomplished by

for example BLAST [12] for single user systems. However, the remaining sections focus on multiple users.

Channel Block Diagonalization

A block diagonalization (BD) based method has been provided in [55–58]. It removes interuser interference, but the receiver needs to perform some type of spatial demultiplexing to separate and decode individual data steams sent to it. The ultimate goal for this is to find $\mathbf{W}$ so that

$$\bar{\mathbf{H}}\bar{\mathbf{W}} = \begin{bmatrix} \mathbf{M}_1 & & \\ & \ddots & \\ & & \mathbf{M}_K \end{bmatrix} \tag{2.23}$$

Assuming that up to N_{R_j} data streams are transmitted to user j, then $\mathbf{M}_j$ in the above equation is $N_{R_j} \times N_{R_j}$. In case if $m_j \leq N_{R_j}$ some of the columns of $\mathbf{M}_j$ would be zero. An example of such a technique based on sum-capacity has been provided in [58]. Define an extended channel matrix $\tilde{\mathbf{H}}_j$ for the jth user which excludes the channel matrix for user j as

$$\tilde{\mathbf{H}}_j = [\mathbf{H}_1^T \ldots \mathbf{H}_{j-1}^T \mathbf{H}_{j+1}^T \ldots \mathbf{H}_K^T]^T \tag{2.24}$$

Perform SVD of $\tilde{\mathbf{H}}_j$

$$\tilde{\mathbf{H}}_j = \tilde{\mathbf{U}}_j \tilde{\Sigma}_j \begin{bmatrix} \tilde{\mathbf{V}}_j^{(1)} & \tilde{\mathbf{V}}_j^{(0)} \end{bmatrix}^H \tag{2.25}$$

where $\tilde{\mathbf{V}}_j^{(0)}$ holds $(N_T - L_j)$ singular vectors that are in the null space of $\tilde{\mathbf{H}}_j$ and L_j is rank of $\{\tilde{H}_j\}$. Therefore $\tilde{\mathbf{V}}_j^{(0)}$ are the candidate for the beamforming matrix $\mathbf{W}_j$ for user j. If the number of null spaces is more than the number of data streams that user j can support, a linear combination of the null spaces must be used to form $\mathbf{W}_j$.

Coordinated Tx/Rx beamforming

BD algorithms have a stringent constraint on the number of transmitting antennas N_T. In this section we look at methods where such a constraint can be relaxed, however we should have $N_T \geq \sum_{j=1}^{K} m_j$ i.e. no smaller than the total number of data streams to be transmitted. If this is not the case, spatial multiplexing must be augmented with other multiple access techniques such as time and frequency multiplexing. A strategy to optimally group best K users for spatial multiplexing in a given time-frequency slot has been provided in [59].

Let us consider the case where $m_j = 1$ (i.e. single data stream per user). Then each users symbol (scalar here) is multiplexed using its corresponding beamforming vector. The signal at the receiver is decoded using $\mathbf{g}_j$ as

$$\begin{aligned} \bar{x}_j = \mathbf{g}_j^H \mathbf{y}_j &= \sum_{k=1}^{K} \mathbf{g}_j^H \mathbf{H}_j \mathbf{w}_k s_k + \mathbf{g}_j^H \mathbf{n}_j \\ &= \sum_{k=1}^{K} \bar{\mathbf{h}}_j^H \mathbf{w}_k s_k + \bar{n}_j \end{aligned} \tag{2.26}$$

where $\bar{\mathbf{h}}_j^H = \mathbf{g}_j^H \mathbf{H}_j$ represents the effective channel from the transmit array to the output of the receiver beamformer $\mathbf{w}_j$, and $\bar{n}_j = \mathbf{g}_j^H \mathbf{n}_j$ represents the noise at the output of the receive beamformer $\mathbf{g}_j$. Defining $\bar{\mathbf{H}}^H = [\bar{\mathbf{h}}_1, \bar{\mathbf{h}}_1, \ldots, \bar{\mathbf{h}}_K]$, we can rewrite the above equation as

$$\bar{\mathbf{x}} = \bar{\mathbf{H}}\mathbf{W}\mathbf{s} + \bar{\mathbf{n}} \tag{2.27}$$

where $\bar{\mathbf{x}} = [\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_K]^T$ and $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_K]$. Assume that the receiver uses the conjugate of its transmit weights for downlink reception. Also assume that BS knows the CSI and interference generated for each user in the system. Hence BS can determine each user's beamformer. In turns out that what ever the design

criterion is at the receiver, its optimal value will depend on the beamformers $\mathbf{W}$. On the other hand the choice of $\mathbf{W}$ depends on $\bar{\mathbf{H}}$, which in turns depends on the receiver beamformers $\mathbf{g}_j$. This interdependency of $\mathbf{g}_j$ and $\mathbf{W}$ suggests use of iterative methods. Examples of such iterative method based on uplink-downlink duality in terms of SINR can be found in [34]. These duality results were extended to systems where the receivers are also equipped with multiple antennas in [60], where the duality is proven for the sum-MSE regions attained in the uplink-downlink paradigm.

2.4 Summary

In this chapter we carried out an extensive literature survey on spatial diversity techniques. In particular, we looked at the receiver diversity, transmitter diversity and joint transmitter and receiver diversity techniques. Multiuser diversity techniques have also been presented. A brief introduction to convex optimization based diversity techniques has also been provided. In particular the focus was on beamformer design using SDP and SOCP. Application of Lagrange duality in this context has also been discussed.

Chapter 3

Convex Optimization Theory

Mathematical optimization plays a major role in the engineering research community including signal processing, communications and so on [61]. Unconstrained optimization such as the well known *least squares* and constrained optimization such as *linear programming* have been around for many years and have been widely exploited [62]. A new general class of mathematical optimization, known as *convex optimization* has emerged as a sturdy candidate for constrained optimization in the last decade or so, with new applications of convex optimizations constantly being reported from almost every area of engineering, including signal processing, communications, control. circuit design, information theory, economics, computer science etc. See [61] and the references therein.

Convex optimization theory allows for a wider range of mathematical optimization to be solved efficiently. This is due to the fact, *interior point methods* developed in the 1980s to solve linear programs, can be generalized to solve other convex optimization problems [63]. In a nutshell, convex problems can be solved optimally either in closed form using *Lagrange duality* or numerically using interior point methods [61]. As a consequence, it is typically said, that once a problem is expressed in a convex form, it has been solved. However a major drawback in engineering is that most of the

problems are not convex when directly formulated. Hence the challenge in using convex optimization is to recognize and formulate the problem in convex form.

3.1 Why Convex?

In engineering, a vast number of the design problems may be cast as a constrained optimization problem of the form [61]

$$\begin{aligned} \min_{\mathbf{x}} \quad & f_0(\mathbf{x}) \\ \text{s.t.} \quad & f_i(\mathbf{x}) \leq 0 \qquad i = 1, \ldots, m \\ & h_i(\mathbf{x}) = 0 \qquad i = 1, \ldots, p \end{aligned} \tag{3.1}$$

where $\mathbf{x}$ is the optimization variable of the problem. The function f_0,f_i and h_i are the cost, inequality constraints, and equality constraints. However, there are a number of potential hurdles, which effectively make the problem (3.1) quite tedious to solve. These impediments can be classified as [61]

1. The dimension n of the optimization variable may be very large.
2. The domain of the problem may be riddled with local optima.
3. The problem might not be feasible.
4. Stopping criteria available may be arbitrary.
5. The algorithms might have poor convergence rates.
6. There might be a problem with numerical accuracy, which could cause the minimization problem to stop all together or wander.

The first three problems could be easily dealt with provided that the f_i are all convex and h_i are affine [64]. If this is the case, the problem is convex and any local optima, is in fact, a global optimum; feasibility can be determined unambiguously, and accurate stopping criterion can be obtained using the principle of *duality* [61]. The problem of convergence and numerical sensitivity were still potential problems, until in the late '80's and '90's, researchers showed that if f_i, in addition to convexity, satisfied a property known as *self-concordance*, the problems of convergence and numerical sensitivity may be dealt with using interior point methods [63, 65]. Interestingly, a large number of functions used in engineering satisfy the self-concordance property, as a consequence, a large number of convex optimization in engineering can now be solved with great efficiency.

3.2 Basic Optimization Concepts

In this section, we introduce some basic optimization concepts that would be readily used throughout this thesis.

3.2.1 Convex Sets

A convex set $\mathcal{S} \in \mathbb{R}^n$ is defined mathematically as follows [61]

$$\theta \mathbf{x} + (1-\theta)\mathbf{y} \in \mathcal{S}, \qquad \forall \theta \in [0,1] \text{ and } \mathbf{x}, \mathbf{y} \in \mathcal{S}. \tag{3.2}$$

In simple words, (3.2) is interpreted as, a set $\mathcal{S}$ is convex if for any two points $\mathbf{x}, \mathbf{y} \in \mathcal{S}$, the line segment between these two points is also in $\mathcal{S}$. For example, the ball $\mathcal{S} = \{\mathbf{x} \mid \|\mathbf{x}\| \leq \varepsilon\}$ is convex, however a sphere $\mathcal{S} = \{\mathbf{x} \mid \|\mathbf{x}\| = \varepsilon\}$ is not a convex set, since the line segment joining any two points is no longer on the sphere. Fig. 3.1 shows a plot of convex balls and non-convex spheres in 2-D for various value of ε. In general, convex sets have non-empty interior i.e. they must have solid body with no

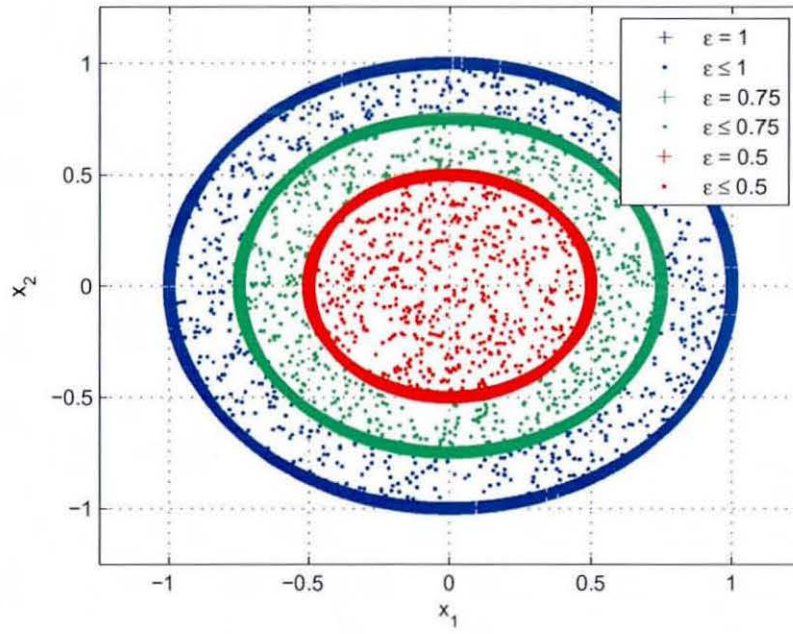

Figure 3.1: Convex sets (balls) and non-convex sets (spheres).

holes. Other examples of convex sets include ellipsoids, polyhedron, and so on.

3.2.2 Convex Cones

A set $\mathcal{K}$ is said to be a convex cone, if for each $\mathbf{x} \in \mathcal{K}$ and each $\alpha \geq 0, \alpha\mathbf{x} \in \mathcal{K}$ and is convex [61] i.e.

$$\theta\mathbf{x} + (1-\theta)\mathbf{y} \in \mathcal{K}, \qquad \forall\theta \geq 0 \text{ and } \mathbf{x}, \mathbf{y} \in \mathcal{K}. \tag{3.3}$$

Examples of convex cone include the nonnegative orthant $\mathbb{R}^n_+$, the positive semidefinite matrix cone $\mathcal{K} = \mathbb{S}^n_+ = \{\mathbf{X} \mid \mathbf{X} \text{ symmetric and } \mathbf{X} \succeq 0\}$, and so on.

3.2.3 Convex Functions

A function $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ is said to be convex if for any two points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ [61]

$$f(\theta\mathbf{x} + (1-\theta)\mathbf{y}) \leq \theta f(\mathbf{x}) + (1-\theta) f(\mathbf{y}), \qquad \forall\theta \in [0,1]. \tag{3.4}$$

f is concave, if $-f$ is convex. The convexity of a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ can also be characterized by its ∇f and Hessian $\nabla^2 f$. We know, that the gradient yields a first order Taylor approximation at $\mathbf{x}_0$:

$$f(\mathbf{x}) \simeq f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) \tag{3.5}$$

We have the following *first-order condition*: f is convex if and only if for all $\mathbf{x}, \mathbf{x}_0 \in \mathbf{dom} f$, $f(\mathbf{x}) \geq f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0)$ i.e., the first order approximation of f is a *global underestimator* [61]. Recalling, that the Hessian of $f, \nabla^2 f$, yields a second order Taylor series expansion around $\mathbf{x}_0$:

$$f(\mathbf{x}) \simeq f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \nabla^2 f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \tag{3.6}$$

We have the following necessary and sufficient *second order conditions*: a twice differentiable function f is convex if and only if for all $\mathbf{x} \in \mathbf{dom}, \nabla^2 f(\mathbf{x}) \succeq 0$, i.e. its Hessian is positive semidefinite on its domain [61]. Thus, for example a linear function is always convex, while a quadratic function $\mathbf{x}^T\mathbf{P}\mathbf{x} + \mathbf{a}^T\mathbf{x} + b$ is convex if and only if $\mathbf{P} \succeq 0$.

3.3 Convex Optimization Problems

Mathematically, a convex problem can be written in standard form as [61,64]

$$\begin{aligned} \min_{\mathbf{x}} \quad & f_0(\mathbf{x}) \\ \text{s.t.} \quad & f_i(\mathbf{x}) \leq 0 \qquad i = 1, \ldots, m \\ & h_i(\mathbf{x}) = 0 \qquad i = 1, \ldots, p \end{aligned} \tag{3.7}$$

where the vector $\mathbf{x} \in \mathbb{R}^n$ is the *optimization variable* of the problem. The function

f_0 is the *objective function* or *cost function.* The functions $f_i, i = 1, 2, \ldots, m$ are convex functions and the functions $h_i, i = 1, 2, \ldots, p$ are linear functions [1]. The inequalities $f_i(\mathbf{x}) \leq 0$ are called the *inequality constraints* and equalitities $h_i(\mathbf{x}) = 0$ are called the *equality constraints.*

The *domain* of the optimization problem (3.7) is the set of points for which the objective and the constraints are defined and is denoted as

$$D = \bigcap_{i=0}^{m} \mathrm{dom} f_i \cap \bigcap_{i=0}^{p} \mathrm{dom} h_i \tag{3.8}$$

Problem (3.7) is said to be *feasible* if there exists a point $\mathbf{x} \in D$ that satisfies all the constraints $f_i(\mathbf{x}) \leq 0$ and $h_i(\mathbf{x})$, the problem is said to be *non-feasible* otherwise. The *optimal value* or the *solution* of the optimization problem is achieved at the optimal point $\mathbf{x}^\star$ if and only if it has the smallest objective among all feasible points i.e. for any feasible point $\mathbf{z} \in D$, $f_0(\mathbf{z}) \geq f_0(\mathbf{x}^\star)$.

3.3.1 Art of Using Convex Optimization

The key or the art of using convex optimization is to formulate non convex problems into convex problems. Unfortunately its not systematic to formulate a non-convex problem into one that is convex, rather it is an art which can only be learned by means of examples. There are two main ways to formulate problems into convex form [47, 61]

- Firstly, by using *change of variables*, a non-convex problem can be easily formulated into a convex problem which is equivalent to the original problem. For e.g. let us consider minimizing the ℓ_2-norm of a vector i.e. $\min \|\mathbf{w}\|_2$, using a change of variable as $\mathbf{W} = \mathbf{w}\mathbf{w}^H$, we can now minimize the trace of the new variable $\mathbf{W}$ i.e. $\min \mathrm{trace}\{\mathbf{W}\}$, which is equivalent to minimizing the ℓ_2-norm

[1] A function $h : \mathcal{R}^n \rightarrow \mathcal{R}$ is linear if it satisfies $h(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha h(\mathbf{x}) + \beta h(\mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{R}^n$ and all $\alpha, \beta \in \mathcal{R}$

of $\mathbf{w}$. Thus change of variables has immediate advantages in transforming a non-convex problem into an equivalent convex problem.

- Secondly, by *relaxing the problem* i.e. removing some of the constraints. This technique is sufficient as long as both the non-convex problem and its formulated convex problem are equivalent i.e. have the same set of optimal solutions (related by some mapping). A famous example of this technique is semidefinite relaxation (SDR), where a non-convex constraint restricting the rank of the optimization variable matrix may be dropped. As shown in the example above, we could minimize trace of the matrix $\mathbf{W}$ which is equivalent of minimizing the ℓ_2-norm of the vector $\mathbf{w}$. However, with change of variable we introduce an additional constraint, $\text{rank}\{\mathbf{W}\} = 1$. This constraint is non-convex, hence makes the whole optimization problem non-convex. Later, in the thesis, we will see that even after dropping this constraint the resulting optimization problem (which is now convex) returns a rank 1 matrix. However, it should be said that this may not necessary hold for all problems.

3.4 Canonical Optimization Problems

In this section, we provide the most general form of canonical optimization problem formulations, which is extremely useful in practice and for which efficient software packages are available. Once a problem is cast into one of these forms, the problem can be considered as essentially solved [61].

3.4.1 Linear Program

The most simplest of these, is a linear program (LP), where the objective and the constraint functions are all affine. A general LP has the form

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c}^T\mathbf{x} + d \\ \text{s.t.} \quad & \mathbf{G}\mathbf{x} \preceq \mathbf{h} \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned} \tag{3.9}$$

where $\mathbf{G} \in \mathbb{R}^{m\times n}$ and $\mathbf{A} \in \mathbb{R}^{p\times n}$.

3.4.2 Quadratic Programming

Secondly, we have a quadratic program (QP), this is where the objective function is quadratic, and the constraint function are affine. A QP has the form

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{x}^T\mathbf{P}\mathbf{x} + \mathbf{q}^T\mathbf{x} + r \\ & \mathbf{G}\mathbf{x} \preceq \mathbf{h} \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned} \tag{3.10}$$

where $\mathbf{P} \in \mathbb{S}^n_+$, $\mathbf{G} \in \mathbb{R}^{m\times n}$, and $\mathbf{A} \in \mathbb{R}^{p\times n}$. In a QP, a convex quadratic function is minimized over a polyhedron. QP include LP as a special case, this may be obtain by setting $\mathbf{P} = 0$ in the objective of (3.10). A variation of QP, is a quadratically constrained quadratic program (QCQP), here both the objective and the constraints are quadratic. This has the form

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{x}^T\mathbf{P}_0\mathbf{x} + \mathbf{q}_0^T\mathbf{x} + r_0 \\ & \mathbf{x}^T\mathbf{P}_i\mathbf{x} + \mathbf{q}_i^T\mathbf{x} + r_i \leq 0, \qquad i = 1, 2, \ldots, m \\ & \mathbf{A}\mathbf{x} = \mathbf{b}, \end{aligned} \tag{3.11}$$

where $\mathbf{P}_i \in \mathbb{S}^n_+, i = 1, 2, \ldots, m$. In a QCQP, we minimize a convex quadratic function over a feasible region that is the intersection of ellipsoids. Similarly to a QP, setting $\mathbf{P}_i = 0, i = 1, 2, \ldots, m$ in the constraints of (3.11) we obtain an LP.

3.4.3 Second Order Cone Programming

A second order cone program (SOCP) can be written as

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{f}^T\mathbf{x} \\ \text{s.t.} \quad & \|\mathbf{A}_i\mathbf{x} + \mathbf{b}_i\|_2 \leq \mathbf{c}_i^T\mathbf{x} + \mathbf{d}_i \qquad i = 1, 2, \ldots, m \\ & \mathbf{F}\mathbf{x} = \mathbf{g} \end{aligned} \tag{3.12}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the optimization variable, $\mathbf{A}_i \in \mathbb{R}^{n_i \times n}$, and $\mathbf{F} \in \mathbb{R}^{p \times n}$. The first constraint in (3.12) is known as a second order cone constraint, since it requires the affine function $(\mathbf{A}\mathbf{x} + \mathbf{b}, \mathbf{c}^T\mathbf{x} + \mathbf{d})$ to lie in the second-order cone in $\mathbb{R}^{k+1}$. Setting $\mathbf{c}_i = 0, i = 1, 2, \ldots, m$ and squaring both sides of the constraints, we obtain a QCQP. Similarly, if $\mathbf{A}_i = 0, i = 1, 2, \ldots, m$, then the SOCP reduces to a LP. SOCP's are more general than both QCQP's and LP's.

3.4.4 Semidefinite Programming

The most general of all the form is a semidefinite program (SDP). This subsumes linear, quadratic and second-order cone programming. A SDP can be written as,

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{c}^T\mathbf{x} \\ \text{s.t.} \quad & \mathbf{x}_1\mathbf{F}_1 + \mathbf{x}_2\mathbf{F}_2 + \ldots + \mathbf{x}_n\mathbf{F}_n + \mathbf{G} \preceq \mathbf{0} \\ & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned} \tag{3.13}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the optimization variable and $\mathbf{G}, \mathbf{F}_0, \mathbf{F}_1, \ldots, \mathbf{F}_n \in \mathbb{S}^{k \times k}$ are symmetric matrices, and $\mathbf{A} \in \mathbb{R}^{p \times n}$. The inequality constraints in (3.13) are also known as linear matrix inequality (LMI). A SDP simplifies to a LP if the matrices $\mathbf{G}, \mathbf{F}_1, \ldots, \mathbf{F}_n$ are all diagonal.

So far we outlined the basic structure of the most commonly used form of the canonical optimization problem. However, it must be noted that not all optimization problems will have one of the above structures, namely a LP, QP, QCQP, SOCP or a SDP. This effectively means the readily available software for solving convex problems might not be useful and custom code (software) might be needed to solve the problem. In this case, one may employ the ellipsoid, subgradient or cutting plane methods, which offer exact stopping criteria and only need gradient information. On the other hand, if Hessian information is also available one may employ interior-point methods, which offer faster convergence [61].

3.5 Duality and KKT Conditions

The *Lagrangian* $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ for the original (primal) problem in (3.7) is defined as the objective function augmented with a weighted sum of the constraint functions. This can be written as

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}), \tag{3.14}$$

where λ_i and ν_i are the *Lagrange multipliers* associated with the ith inequality $f_i(\mathbf{x}) \leq 0$ and equality $h_i(\mathbf{x}) = 0$ constraints respectively.

The objective $f_0(\mathbf{x})$ in (3.7) is known as the *primal objective* and the optimization variable $\mathbf{x}$ is termed the *primal variable*. Lagrange multiplier vectors $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ associated with the problem (3.7) are known as the *dual variables* and the *dual objective* or the *dual function* $g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ is defined as the minimum value of the Lagrangian

over $\mathbf{x}$: for $\boldsymbol{\lambda} \in \mathbb{R}^m$, $\boldsymbol{\nu} \in \mathbb{R}^p$

$$g(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \inf_{\mathbf{x} \in \mathcal{D}} \left(f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}) \right). \tag{3.15}$$

The dual function is concave even when the original problem is not convex, since it is pointwise infimum of a family of affine functions of $(\boldsymbol{\lambda}, \boldsymbol{\nu})$ [61]. We say that $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ are *dual feasible* if $\boldsymbol{\lambda} \geq 0$ and $g(\boldsymbol{\lambda}, \boldsymbol{\nu})$ is finite i.e. $g(\boldsymbol{\lambda}, \boldsymbol{\nu}) > -\infty$.

The dual function $g(\boldsymbol{\lambda}, \boldsymbol{\nu})$ serves as a lower bound on the optimal value f^* of the problem (3.7) [61]. For any feasible set $(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu})$:

$$\begin{aligned} f_0(\mathbf{x}) &\geq f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{x}) \\ &\geq \inf_{\mathbf{z} \in \mathcal{D}} \left(f_0(\mathbf{z}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{z}) + \sum_{i=1}^{p} \nu_i h_i(\mathbf{z}) \right) \\ &= g(\boldsymbol{\lambda}, \boldsymbol{\nu}) \end{aligned} \tag{3.16}$$

where we have used the fact that $f_i(\mathbf{x}) \leq 0$ and $h_i(\mathbf{x}) = 0$ for any feasible $\mathbf{x}$ and $\lambda_i \geq 0$ for any feasible λ_i in the first inequality. Thus, for a feasible set $(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu})$, we have

$$\min_{\mathbf{x}} f_0(\mathbf{x}) \geq \max_{\boldsymbol{\lambda}, \boldsymbol{\nu}} g(\boldsymbol{\lambda}, \boldsymbol{\nu}). \tag{3.17}$$

Duality gap is the measure of the difference between the primal objective $f_0(\mathbf{x})$ and the dual objective $g(\boldsymbol{\lambda}, \boldsymbol{\nu})$. We say, *weak duality* holds, if (3.17) is satisfied with strict inequality and if (3.17) is satisfied with equality *strong duality* holds.

The best lower bound on the original problem may be obtained solving the following optimization problem

$$\begin{aligned} \max_{\boldsymbol{\lambda},\boldsymbol{\nu}} \quad & g(\boldsymbol{\lambda},\boldsymbol{\nu}) \\ \text{s.t.} \quad & \boldsymbol{\lambda} \geq 0 \end{aligned} \tag{3.18}$$

Problem (3.18) is commonly known as the *Lagrange dual problem* and is always a convex optimization problem, since the objective to be maximized $g(\boldsymbol{\lambda},\boldsymbol{\nu})$ is always concave and the constraint is convex. This holds, regardless of whether or not the primal problem (3.7) is convex [61].

Note, as mentioned earlier, that mathematical optimization problem normally suffers from arbitrary stopping criterion. However, the above results from the Lagrange dual problem provides a non-heuristic stopping criterion. This is simply because, that a primal-dual feasible point $(\mathbf{x},(\boldsymbol{\lambda},\boldsymbol{\nu}))$ localizes the optimum solution in the interval defined by the duality gap i.e. $f^* \in [g(\boldsymbol{\lambda},\boldsymbol{\nu}), f_0(\mathbf{x})]$. If $g(\boldsymbol{\lambda},\boldsymbol{\nu}) = f_0(\mathbf{x})$, then the duality gap is zero, and both the primal and the dual variables are at the optimal solution. Let us denote the primal optimum variable as $\mathbf{x}$ and dual optimum variable as $(\boldsymbol{\lambda}^*,\boldsymbol{\nu}^*)$. Since $\mathbf{x}^*$ minimizes $L(\mathbf{x},\boldsymbol{\lambda}^*,\boldsymbol{\nu}^*)$ over $\mathbf{x}$, the gradient of $L(\mathbf{x},\boldsymbol{\lambda}^*,\boldsymbol{\nu}^*)$ must vanish at $\mathbf{x}^*$, i.e.,

$$\nabla f_0(\mathbf{x}^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(\mathbf{x}^*) + \sum_{i=1}^{p} \nu_i^* \nabla h_i(\mathbf{x}^*) = 0. \tag{3.19}$$

Thus we have

$$f_i(\mathbf{x}^*) \leq 0 \qquad i = 1,2,\ldots,m, \tag{3.20}$$

$$h_i(\mathbf{x}^*) = 0 \qquad i = 1,2,\ldots,p, \tag{3.21}$$

$$\lambda_i^* \geq 0 \qquad i = 1,2,\ldots,m, \tag{3.22}$$

$$\lambda_i^* f_i(\mathbf{x}^*) = 0 \qquad i = 1,2,\ldots,m. \tag{3.23}$$

Collectively, the conditions (3.19)-(3.23) are known as the Karush-Kuhn-Tucker (KKT) conditions for optimality [61]. Conditions (3.20) and (3.21) represent primal feasibility of $\mathbf{x}^*$. Condition (3.22) represent dual feasibility and the condition (3.23) signifies the *complementary slackness* for the primal and dual inequality constraint pair: $f_i(\mathbf{x}) \leq 0$ and $\lambda_i \geq 0$.

KKT conditions, in general are necessary but not sufficient for optimality. However, for convex optimization problems, KKT conditions are also sufficient [61]. KKT conditions reduces to the well known stationary conditions $\nabla f_i(\mathbf{x}^*) = 0$, in the absence of constraints, i.e. a minimum must be obtained at the point where the gradient of f_0 vanishes. However, in the presence of constraints, the optimal solution is attained at a KKT point $\mathbf{x}^*$, which, together with some dual feasible vector $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ satisfies the KKT conditions (3.19)-(3.23). KKT conditions have proven to be very useful in practice to obtain solutions analytically (when possible).

3.6 Robust Convex Optimization

Robust optimization models in mathematical programming has received much attention recently see [66–68]. Here we will review some of these models and their extensions. Consider a convex optimization of the form

$$\begin{aligned} \min_{\mathbf{x}} \quad & f_o(\mathbf{x}) \\ \text{s.t.} \quad & f_i(\mathbf{x}) \leq 0, \quad i = 1, 2, \ldots, m \end{aligned} \tag{3.24}$$

where each f_i is convex. In many engineering design applications, the data defining the constraint and the objective functions may be inexact, corrupted by noise, or may fluctuate around with time around a nominal value. In the application of multiuser beamforming such errors may arise due to imperfect CSI at the transmitter, receiver

or both. Recalling from §1.3.2, that we model the estimate of the channel covariance matrix as (1.5), where Δ represents the errors (uncertainties) in the CSI. The way these uncertainties in the CSI are modeled in the channel covariance estimate, the robust techniques may be classified into two categories, *Bayesian* (or stochastic) and the *Maximin* (or worst-case) approach [47,61].

- The Bayesian approach considers that the statistics of the error is known and in such cases the traditional approach simply solves (3.24) by using the nominal value (or mean) of the data. However, the stochastic design only guarantees a certain system performance averaged over the data that could have caused the current estimated data. No guarantee can be given in terms of the instantaneous performance. An example of such an approach can be found in [69,70] respectively, where a multi-antenna transmitter was designed to maximize the mean SNR and mean BER assuming errors due to Gaussian noise and quantization errors.

- The maximin approach on the other hand, considers that the errors belongs to a predefined uncertainty region, and the final objective is the optimization of the worst system performance for any error in this region. In practice the assumption of error being bounded is satisfied with high probability. Let us denote the set of perturbed functions parameterized by δ : $f_i(\mathbf{x};\delta)$, with δ taken from an uncertainty set $\mathcal{S}$. Then a feasible robust solution $\mathbf{x}$ is the one that satisfies $f_i(\mathbf{x};\delta) \leq 0, \quad \forall\delta \in \mathcal{S}$ or, equivalently $\max_{\delta\in\mathcal{S}} f_i(\mathbf{x};\delta) \leq 0$. We can write such an optimization problem as

$$\begin{aligned} \min_{\delta\in\mathcal{S}} \quad & \max f_o(\mathbf{x};\delta) \\ \text{s.t.} \quad & f_i(\mathbf{x};\delta) \leq 0, \quad \forall\delta \in \mathcal{S}, \quad i = 1,2,\ldots,m \end{aligned} \tag{3.25}$$

the parameters to be in an uncertainty region, for example an ellipsoid. Some examples include [74–76, 105] where the presumed steering vector is modelled as

$$\mathbf{s} = \hat{\mathbf{s}} + \mathbf{e} \tag{3.38}$$

where $\mathbf{e}$ is the noise vector and it is assumed to have norm bounded i.e $\|\mathbf{e}\|_2 \leq \epsilon$. Then the robust solution could be obtained by imposing a good response in all directions within the uncertainty region i.e.,

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}^H \mathbf{R} \mathbf{w} \\ \text{s.t.} \quad & |\mathbf{w}^H c| \geq 1 \quad \forall c = \hat{\mathbf{s}} + \mathbf{e}, \quad \|\mathbf{e}\| \leq \epsilon \end{aligned} \tag{3.39}$$

Such a problem is a semi-infinite nonconvex quadratic problem and needs to be simplified. The single constraint $\min_{\|\mathbf{e}\| \leq \epsilon} |\mathbf{w}^H(\hat{\mathbf{s}} + \mathbf{e})| \geq 1$ is equivalent to the original semi-infinite set of constraints and then, by applying the triangle and Cauchy-Schwarz inequalities along with $\|\mathbf{e}\| \leq \epsilon$, the following is obtained

$$|\mathbf{w}^H \hat{\mathbf{s}} + \mathbf{w}^H \mathbf{e}| \geq |\mathbf{w}^H \hat{\mathbf{s}}| - |\mathbf{w}^H \mathbf{e}| \geq |\mathbf{w}^H \hat{\mathbf{s}}| - \epsilon \|\mathbf{w}\| \tag{3.40}$$

where the lower bound is indeed achieved if $\mathbf{e}$ is proportional to $\mathbf{w}$ with a phase such that $\mathbf{w}^H \mathbf{e}$ has opposite direction of $\mathbf{w}^H \hat{\mathbf{s}}$ [105]. Now, since $\mathbf{w}$ admits any arbitrary rotation without affecting the cost, $\mathbf{w}^H \hat{\mathbf{s}}$ can be forced to be real and nonnegative. The problem can be finally formulated in the convex form as

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}^H \mathbf{R} \mathbf{w} \\ \text{s.t.} \quad & \mathbf{w}^H \hat{\mathbf{s}} \geq 1 + \epsilon \|\mathbf{w}\| \\ & \mathrm{Im}\{\mathbf{w}^H \hat{\mathbf{s}}\} = 0 \end{aligned} \tag{3.41}$$

A similar problem was considered in [105] for a general-rank signal model.

3.9.2 Robust Multiuser Downlink Beamforming

In [78] the SDP based solution presented in (3.29), was extended to the case where only an erroneous estimate of the CSI is available at the transmitter using worst-case performance optimization (see §3.6). The final optimization problem which can be cast into a SDP using SDR is given as follows

$$\begin{aligned} \min_{\mathbf{W}_i} \quad & \sum_{i=1}^{K} \mathrm{tr}(\mathbf{W}_i) \\ \text{s.t.} \quad & \mathrm{tr}((\mathbf{R}_i - \epsilon_i \mathbf{I})\mathbf{W}_i) - \gamma_i \sum_{j \neq i} \mathrm{tr}((\mathbf{R}_i + \epsilon_i \mathbf{I})\mathbf{W}_j) \geq \gamma_i \sigma_i^2, \\ & \mathbf{W}_i \geq 0, \quad i = 1, 2, \ldots, K \end{aligned} \tag{3.42}$$

where ϵ_i is norm of the uncertainty matrix (see §1.3.2). The problem of robust downlink beamforming under CSI errors was also presented using an approach of probabilistic constrained based optimization. Here as appose to optimizing the system for the worst-case the system is optimized for a given probability. An SDP based solution was recently proposed in [106].

$$\begin{aligned}
&\min_{\mathbf{W}_i} \sum_{i=1}^{K} \mathrm{tr}\{\mathbf{W}_i\} \\
&\text{s.t.} \quad \mathrm{tr}\{(\mathbf{R}_i - c_i\mathbf{I})\mathbf{W}_i - \gamma_i \sum_{j=1, j\neq i}^{K} \mathrm{tr}\{(\mathbf{R}_i + c_i\mathbf{I})\mathbf{W}_j\} \geq \gamma_i \sigma_i^2, \\
&\mathbf{W}_i \geq 0, \qquad i = 1, \ldots, K.
\end{aligned} \tag{3.43}$$

where $c_i = \sqrt{2}\sigma_e \mathrm{erf}^{-1}(2p_i - 1)$ and σ_e^2 is the variance of the elements of the channel uncertainty matrix, $\mathrm{erf}\{\cdot\}$ denotes the Gaussian error function and p_i is a pre-defined non-outage probability.

3.10 Summary

A brief overview of convex optimization theory has been provided in this chapter. Basic concepts and tools of convex optimization that are readily used in this thesis were introduced. The most generic forms of canonical optimization problems namely LPs, QPs, SOCPs and SDPs were presented. The concepts of Lagrange duality and KKT conditions were also discussed. Robust convex optimization, has also been discussed with special emphasis on worst-case performance optimization. Interior point methods, although out of the scope of this thesis, were included for the completeness of convex optimization theory. Finally, robust techniques based on convex optimization theory were discussed, where the CSI available at the transmitter is assumed to be in error.

Chapter 4

Multiuser Downlink Beamforming based on Maximizing Signal-to-Leakage Ratio

Multiuser multiple-input multiple-output (MU-MIMO) systems have gained a considerable amount of interest in recent years due to their potential for providing high capacity, increasing diversity and interference suppression [12]. This, as established in Chapter 1, is possible due to the gains that could be achieved from a MIMO system, without additional spectrum.

Recent research in MU-MIMO systems is aimed at developing techniques which allow users to efficiently share the scarce spectrum [107]. However, such systems generally suffer from co-channel interference (CCI) induced due to frequency re-use, and multi-user interference (MUI), as a consequences of multiple users access the same frequency simultaneously. Thus, techniques are generally required for the suppression of interference so that the per user capacity in a multiuser environment hindered by both CCI and MUI, should be closer to the capacity of a single user system [16].

The focal point of this chapter is on spatial diversity techniques in a downlink

wireless communication system, where a basestation (BS) could simultaneously serve multiple users without compromising available radio spectrum [59]. This requires the BS to pre-compensate interference so that a particular user in the cell will not see the signals that are meant to be transmitted to other users. It is also possible for the BS to perform beamforming to suppress MUI to end users and to maximize overall capacity.

In an attempt to suppress MUI, several techniques have been proposed [34,58,59, 83,108–110]. One technique is to pre-process the signal at the BS so that MUI will be completely cancelled at the receiver for each user. Two such methods known as "block-diagonalization" and "successive optimization" have been proposed in [58]. However, both these methods require the number of transmitting antennas to be greater than the sum of all receiving antennas of all users. Another approach proposed in [109] makes use of space-time block codes (STBC) to design a unitary precoder to cancel the CCI. Once again this method requires a large number of antennas. A closed form solution is presented in [110] which is based on maximizing a lower bound for the product of signal-to-interference plus noise ratio (SINR). The algorithm achieves good performance but again it requires the number of transmitting antennas to be greater than the number of receiving antennas. All these schemes provide superior performance, however they impose a restriction on the number of transmit antennas to be greater than the number of antennas of all users combined.

An iterative algorithm based on uplink-downlink duality was presented in [34], where the global optimum for the downlink beamforming is obtained for the case of a single antenna at the receivers. However, in this chapter, we adopt a signal to leakage ratio (SLR) criterion proposed in [108], but propose various techniques to improve the performance further. Even though, this family of algorithms is not supported by any known optimality criteria, such as SINR or minimum mean square error (MMSE), we considered this criterion for its simplicity.

According to the approach in [108], the transmit weight vector for the ith user will be determined by maximizing the transmit power to the ith user while minimizing the interference (leakage) caused to all other users. However, instead of considering the interference at the output of the array of antennas of each user, we considered the interference at the output of the beamformer of each user. The rational behind this method is that the BS knows the set of beamformers that each user will eventually use, hence it can take advantage of this in the design process. We demonstrated the performance of the proposed method could be further improved by designing the transmit weight vectors using an iterative optimization approach.

4.1 System Model

Consider a downlink MU-MIMO system consisting of one BS with N_T transmit antennas communicating with K users each having N_{R_i} receive antennas. A block diagram is shown in Fig. 1.2, where $s_i(t)$ denotes the signal for the ith user at time t. The signal $s_i(t)$ is then multiplied by a beamformer weight vector $\mathbf{w}_i(t)$ before being transmitted over a multiuser channel. Hence, the $N_T \times 1$ transmitted signal vector at time t is given by

$$\mathbf{x}(t) = \sum_{i=1}^{K} \mathbf{w}_i s_i(t). \tag{4.1}$$

It is assumed that the data $s_i(t)$ and the beamformer weights $\mathbf{w}_i$ are normalized so that

$$\mathrm{E}|s_i(t)|^2 = 1 \quad \text{and,}$$

$$\|\mathbf{w}_i\|_2^2 = 1, \quad k = 1, 2, \ldots, K,$$

The $N_T \times 1$ signal vector $\mathbf{x}(t)$ is then transmitted over a multiuser channel. As-

suming the channel is frequency non-selective (i.e. flat), the received signal vector $\mathbf{y}_i(t)$ for the ith user at time t is written as

$$\mathbf{y}_i(t) = \mathbf{H}_i \sum_{j=1}^{K} \mathbf{w}_j s_j(t) + \mathbf{n}_i(t), \tag{4.2}$$

where $\mathbf{n}_i(t)$ is the additive white Gaussian noise (AWGN) vector at the ith user with variance σ_i^2. The channel $\mathbf{H}_i$ is assumed to be block fading. Assuming the ith user employs N_{R_i} antennas, the $N_{R_i} \times N_T$ channel matrix can be written as

$$\mathbf{H_i} = \begin{bmatrix} h_i^{(1,1)} & h_i^{(1,1)} & \dots & h_i^{(1,N_T)} \\ h_i^{(2,1)} & h_i^{(2,2)} & \dots & h_i^{(2,N_T)} \\ \vdots & \vdots & \ddots & \vdots \\ h_i^{(N_{R_i},1)} & h_i^{(N_{R_i},2)} & \dots & h_i^{(N_{R_i},N_T)}, \end{bmatrix} \tag{4.3}$$

where $h_i^{(p,m)}$ denote the channel coefficient between the mth transmit and p^{th} receive antennas for user i. We assume that the receiver for user i knows its own channel state information (CSI), $\mathbf{H}_i$, perfectly and feeds it back to the BS without any errors.

4.2 Algorithms

In the remaining sections, we will drop the time index t for notational simplicity. Hence, we can rewrite equation (4.2) as

$$\mathbf{y}_i = \underbrace{\mathbf{H}_i \mathbf{w}_i s_i}_{\text{Signal of Interest}} + \underbrace{\sum_{j=1, j\neq i}^{K} \mathbf{H}_i \mathbf{w}_j s_j}_{\text{Interference}} + \underbrace{\mathbf{n}_i}_{\text{Noise}}, \tag{4.4}$$

where the second term quantifies the interference caused to user i from all other users. The aim is to mitigate this interference for all users. For simplicity, we assume

that the estimate of s_i for the ith user is based on a *maximum ratio combining* (MRC)[1] [108], i.e.

$$s_i \triangleq \frac{\mathbf{w}_i^H \mathbf{H}_i^H \mathbf{y}_i}{\|\mathbf{H}_i \mathbf{w}_i\|_2^2}, \tag{4.5}$$

where $\frac{\mathbf{w}_i^H \mathbf{H}_i^H}{\|\mathbf{H}_i \mathbf{w}_i\|_2^2}$ denotes the MRC receiver. Then an estimate of the transmit symbol s_i for the ith user, denoted as $\tilde{s}_i$ can be written as

$$\tilde{s}_i = s_i + \frac{\mathbf{w}_i^H \mathbf{H}_i^H \sum_{j=1, j\neq i}^{K} \mathbf{H}_i \mathbf{w}_j s_j}{\|\mathbf{H}_i \mathbf{w}_i\|_2^2} + \frac{\mathbf{w}_i^H \mathbf{H}_i^H \mathbf{n}_i}{\|\mathbf{H}_i \mathbf{w}_i\|_2^2}, \tag{4.6}$$

and the output SINR for user i would be given by [108]

$$\text{SINR}_i = \frac{\|\mathbf{H}_i \mathbf{w}_i\|_2^2}{\sigma_i^2 + \frac{\sum_{j=1, j\neq i}^{K} \|\mathbf{w}_i^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{w}_j\|_2^2}{\|\mathbf{H}_i \mathbf{w}_i\|_2^2}}. \tag{4.7}$$

The power of the desired signal in (4.4) is given by $\|\mathbf{H}_i \mathbf{w}_i\|_2^2$. Similarly, the interference caused by the ith user to the jth user is given by $\|\mathbf{H}_j \mathbf{w}_i\|_2^2$. The quantity, called *leakage* for user i, is the total power leaked from this user to all other users and is defined in [108] as

$$\sum_{j=1, j\neq i}^{K} \|\mathbf{H}_j \mathbf{w}_i\|_2^2. \tag{4.8}$$

4.2.1 The Signal-to-Leakage Ratio Algorithm

Given a fixed transmit power for each user, the weight vectors $\mathbf{w}_i$, $i = 1, 2, \ldots, K$, are designed such that the SLR is maximized for every user [108]

[1]The receiver may employ other diversity combining techniques such as selection combining, switched combining or equal gain combining techniques see chapter 3 for further details.

$$\begin{aligned} \max_{\mathbf{w}_i} \quad & \frac{\|\mathbf{H}_i\mathbf{w}_i\|_2^2}{\sum_{j=1,j\neq i}^{K}\|\mathbf{H}_j\mathbf{w}_i\|_2^2} \\ \text{s.t.} \quad & \|\mathbf{w}_i\|_2^2 = 1, \qquad i = 1,2,\ldots,K. \end{aligned} \tag{4.9}$$

By denoting $\tilde{\mathbf{H}}_i = [\mathbf{H}_i^H ... \mathbf{H}_{i-1}^H \mathbf{H}_{i+1}^H ... \mathbf{H}_K^H]^H$ as an extended channel matrix that excludes the channel $\mathbf{H}_i$, the SLR for user i can be written as

$$\text{SLR}_i = \frac{\|\mathbf{H}_i\mathbf{w}_i\|^2}{\|\tilde{\mathbf{H}}_i\mathbf{w}_i\|^2} = \frac{\mathbf{w}_i^H\mathbf{H}_i^H\mathbf{H}_i\mathbf{w}_i}{\mathbf{w}_i^H\tilde{\mathbf{H}}_i^H\tilde{\mathbf{H}}_i\mathbf{w}_i}. \tag{4.10}$$

The above equation can be solved using the Rayleigh-Ritz quotient result [111]

$$\frac{\mathbf{w}_i^H\mathbf{H}_i^H\mathbf{H}_i\mathbf{w}_i}{\mathbf{w}_i^H\tilde{\mathbf{H}}_i^H\tilde{\mathbf{H}}_i\mathbf{w}_i} \leq \lambda_{max}(\mathbf{H}_i^H\mathbf{H}_i, \tilde{\mathbf{H}}_i^H\tilde{\mathbf{H}}_i), \tag{4.11}$$

where λ_{max} is the largest generalized eigenvalue of the matrix pair $\mathbf{H}_i^H\mathbf{H}_i$ and $\tilde{\mathbf{H}}_i^H\tilde{\mathbf{H}}_i$. The equality holds only if $\mathbf{w}_i$ is proportional to the generalized eigenvector corresponding to the largest generalized eigenvalue, i.e.

$$\mathbf{w}_i \propto \mathcal{P}_{\text{gen}}(\mathbf{H}_i^H\mathbf{H}_i, \tilde{\mathbf{H}}_i^H\tilde{\mathbf{H}}_i), \tag{4.12}$$

where $\mathcal{P}_{\text{gen}}\{\mathbf{A},\mathbf{B}\}$ returns the eigenvector corresponding to the largest generalized eigenvalue of matrix pair $\mathbf{A}$ and $\mathbf{B}$. The proportionality constant is chosen to normalize $\mathbf{w}_i$ to unity. If $\tilde{\mathbf{H}}_i^H\tilde{\mathbf{H}}_i$ is invertible, then the generalized eigenvalue problem reduces to

$$\lambda_{max}(\mathbf{H}_i^H\mathbf{H}_i, \tilde{\mathbf{H}}_i^H\tilde{\mathbf{H}}_i) = \lambda_{max}((\tilde{\mathbf{H}}_i^H\tilde{\mathbf{H}}_i)^{-1}(\mathbf{H}_i^H\mathbf{H}_i)), \tag{4.13}$$

and $\mathbf{w}_i$ is the eigenvector corresponding to the largest eigenvalue of

$$(\tilde{\mathbf{H}}_i^H\tilde{\mathbf{H}}_i)^{-1}(\mathbf{H}_i^H\mathbf{H}_i).$$

4.2.2 The Proposed Algorithm

The method proposed in [108] considers the interference present at the output of the array of antennas of each user in the design process. However, we observed consideration of the interference present at the beamformer output instead of the output of the array of antennas substantially improves the overall bit error rate (BER) performance. This is possible in the design, as the BS knows the beamformer vectors that will be eventually used by all users, as it knows the forward channel of all users. The proposed design is based on an iterative optimization approach.

First Iteration

In the first iteration the SLR considered for user 1 will be same as that in [108]

$$\mathrm{SLR}_1 = \frac{\mathbf{w}_1^H \mathbf{H}_1^H \mathbf{H}_1 \mathbf{w}_1}{\mathbf{w}_1^H \tilde{\mathbf{H}}_1^H \tilde{\mathbf{H}}_1 \mathbf{w}_1}. \tag{4.14}$$

Similarly to (4.12) the solution to maximizing (4.14) is given by the Rayleigh-Ritz quotient result. Then the beamformer weight for user 1 is given by

$$\mathbf{w}_1^1 = \mathcal{P}_{\mathrm{gen}}\{\mathbf{H}_1^H \mathbf{H}_1, \tilde{\mathbf{H}}_1^H \tilde{\mathbf{H}}_1\}. \tag{4.15}$$

In order to compute the weight vector for user 2 we use the fact that the BS knows the beamformer weight vector for user 1. Hence the channel from the signal s_2 at the BS to the output of the beamformer of user 1 can be written as $\theta_1^{\mathbf{H}} \mathbf{H}_1 \mathbf{w}_2$, where $\theta_1 = \mathbf{H}_1 \mathbf{w}_1$ is the required beamformer for user 1. Therefore, the interference power caused by user 2 to user 1 can be written as $\mathbf{w}_2^{\mathbf{H}} \mathbf{H}_1^{\mathbf{H}} \theta_1 \theta_1^{\mathbf{H}} \mathbf{H}_1 \mathbf{w}_2$ instead of $\mathbf{w}_2^{\mathbf{H}} \mathbf{H}_1^{\mathbf{H}} \mathbf{H}_1 \mathbf{w}_2$. We therefore, replace the term $\mathbf{H}_1^H \mathbf{H}_1$ in the denominator of (4.10) with a rank 1 matrix $\mathbf{H}_1^{\mathbf{H}} \theta_1 \theta_1^{\mathbf{H}} \mathbf{H}_1 = \mathbf{H}_1^{\mathbf{H}} \mathbf{R}_1 \mathbf{H}_1$.

We can now define the SLR for user 2 as

$$\mathrm{SLR}_2 = \frac{\mathbf{w}_2^H \mathbf{H}_2^H \mathbf{H}_2 \mathbf{w}_2}{\mathbf{w}_2^H (\mathbf{H}_1^H \mathbf{R}_1 \mathbf{H}_1 + \sum_{j=3}^{K} \mathbf{H}_j^H \mathbf{H}_j) \mathbf{w}_2}. \tag{4.16}$$

Similarly, we can generalize (4.16) for the ith user in the first iteration as

$$\mathrm{SLR}_i = \frac{\mathbf{w}_i^H \mathbf{H}_i^H \mathbf{H}_i \mathbf{w}_i}{\mathbf{w}_i^H \underbrace{\left(\sum_{j=1}^{i-1} \mathbf{H}_j^H \mathbf{R}_j \mathbf{H}_j + \sum_{j=i+1}^{K} \mathbf{H}_j^H \mathbf{H}_j \right)}_{\mathbf{P}_i} \mathbf{w}_i}. \tag{4.17}$$

Therefore, in the first iteration, the beamforming weight vector for users excluding user 1 are computed according to the pseudocode described in algorithm 1.

Algorithm 1 Iterative Multiuser Downlink Beamforming based on Maximizing Signal-to-Leakage Ratio - 1st Iteration

1. Using the weight vector for user 1 to $i-1$ in the 1st iteration, construct the $\mathbf{P}_i$ matrix defined in (4.17).
2. `if`
   - $\mathrm{rank}(\mathbf{P}_i) \geq N_T$;
     where $\mathrm{rank}\{\cdot\}$ denotes the rank of the matrix.
3. `then`
   - $\mathbf{w}_i^1 = \mathcal{P}_{\max}\{\mathbf{P}_i\}$;
     where $\mathcal{P}_{\max}\{\cdot\}$ returns the eigenvector corresponding to the largest eigenvalue of the matrix.
4. `else`
   - $\mathbf{w}_i^1 = \mathrm{null}\{\mathbf{P}_i\}$;
     where $\mathrm{null}\{\cdot\}$ returns the eigenvectors that are in the null space of the matrix.

Other Iterations

After the first iteration we have a set of beamformer weight vectors for all users. We can now use these weight vectors to carry out further iterations. Carrying out these

iterations will force the beamformer weight vectors to converge to a set of weight vectors which will result in further reduction of CCI. In the lth iteration, the SLR for user i is determined as

$$\text{SLR}_i^l = \frac{\mathbf{w}_i^{(l)H}\mathbf{H}_i^H\mathbf{H}_i\mathbf{w}_i^{(l)}}{\mathbf{w}_i^{(l)H}\underbrace{\left(\sum_{j=1,j\neq i}^{K}\mathbf{H}_j^H\mathbf{H}_j\mathbf{w}_j^{(l-1)}\mathbf{w}_j^{(l-1)H}\mathbf{H}_j^H\mathbf{H}_j\right)}_{\mathbf{Q}_i^l}\mathbf{w}_i^l}. \tag{4.18}$$

where $\mathbf{w}^{(l-1)}$ is the weight vector obtained in the previous iteration i.e. $l-1$. If the total number of antennas N_T at the BS is not less than the total number of users K, the beamformer weight vector for the ith user will be computed according to

$$\mathbf{w}_i^l = \text{null}\{\mathbf{Q}_i^l\}. \tag{4.19}$$

For $\mathbf{P}_i$ in (4.17) and $\mathbf{Q}_i^l$ in (4.18), it is possible to have a null space of dimension greater than one. In this case, the beamformer weight vector should be chosen as a linear combination of all null vectors. The determination of the optimum combination in the sense of maximizing power transferred to the desired user is also an eigenvector problem. The linear combination coefficients are given by the eigenvector (denoted by $\mathbf{g}_i$) corresponding to the largest eigenvalue of $(\mathbf{A}_i^H\mathbf{H}_i^H\mathbf{H}_i\mathbf{A}_i)$, where $\mathbf{A}_i$ is a matrix containing all the null vectors of $\mathbf{P}$ or $\mathbf{Q}$ for user i, and $\mathbf{H}_i$ is the channel matrix of user i. The weight vector for user i is obtained as

$$\mathbf{w}_i^l = \mathbf{A}_i\mathbf{g}_i. \tag{4.20}$$

The proposed iterative algorithm has been summarized in algorithm 2.

Algorithm 2 Iterative Multiuser Downlink Beamforming based on Maximizing Signal-to-Leakage Ratio

1. INITIALIZE $l = 1$ and $l_{\max}$.

   1st Iteration

   - Compute the downlink weight vector for the all users $\mathbf{w}_i^1$, $i = 1, 2, \ldots, K$ in the 1st iteration using the pseudocode in algorithm 1.
   - if
     - $\text{rank}\{\text{null}\{\mathbf{P}_i\}\} > 1$.
     - Update weight vector for the ith user using (4.20).

   Other Iterations

   - REPEAT
     - $l = l + 1$.
     - Using the weight vector of all users in the $(l-1)$th iteration $\mathbf{w}_i^{l-1}$, $i = 1, 2, \ldots, K$ construct the $\mathbf{Q}_i^l$ matrix defined in (4.18).
     - if
       * $\text{rank}\{\text{null}\{\mathbf{Q}_i^l\}\} > 1$.
       * Update weight vector for the ith user using (4.20).
     - Compute the downlink weight vector for the ith user in the lth iteration using (4.19).
   - UNTIL $l = l_{\max}$.

2. Use $\mathbf{w}_i^{l_{\max}}$ as the beamformers for transmission.

4.3 Numerical Examples

We considered a MU-MIMO system with one BS equipped with N_T antennas and K users each equipped with N_{R_i} antennas. The data symbols are generated using quadrature phase-shift keying (QPSK) modulation. The total transmitted power per symbol period across all transmit antennas is normalized to unity. The entries of channel $\mathbf{H}$ are zero mean independent and identically-distributed (IID) Gaussian random variables with unity variance and generated independently for each transmission symbol. The noise is zero mean and spatially and temporally uncorrelated,

i.e.

$$\mathrm{E}\left\{\mathbf{n}_i\mathbf{n}_i^H\right\} = \sigma_i^2\mathbf{I}_{N_{R_i}}, \text{ and}$$

$$\mathrm{E}\left\{\mathrm{Tr}(\mathbf{H}_i\mathbf{H}_i^H)\right\} = N_{R_i}N_T.$$

Fig. 4.1, depicts the difference between SLR and the proposed algorithm for the first iteration only. The result shows the BER performance of all the users using both SLR and the proposed algorithm. We have considered the case with $N_T = 6$ transmitting antennas and $K = 5$ users each with $N_{R_i} = 3$ receiving antennas. It can be seen that the SLR produces the same BER performance for all users as expected. On the other side, we note that the proposed algorithm produces different BER performance for all users. This is due to the fact that when the weights for user i are obtained by maximizing the signal to leakage ratio, it tends to reduce the leakage to all the other users. However, since we use the weight vectors of users 1 to $i-1$ in the design of weight vector of user i, users 1 to $i-1$ are more likely to benefit in terms of interference suppression rather than users $i+1$ to K. In order to gain from this effect we are encouraged to carry out further iterations. Also carrying out further iterations ensures that the average BER of all users is the same and therefore guarantees the same quality of service (QoS) for all users.

Similarly, in Fig. 4.2, the results presented consider the case with $N_T = 6$ transmitting antennas and $K = 5$ users each equipped with $N_{R_i} = 3$ receiving antennas. But unlike Fig. 4.1, in Fig. 4.2 the average BER of all users is depicted for various number of iterations. We note that the performance is greatly improved when the number of iterations is increased. But the performance converges roughly around 20 iterations and further iterations have marginal improvement on the BER performance.

To understand the proposed algorithm better, we look at the SINR outage (or cumulative distribution function (CDF)), which is plotted to show and compare the distribution of the SINR achieved at the output of the receiver. Figs. 4.3 and 4.4

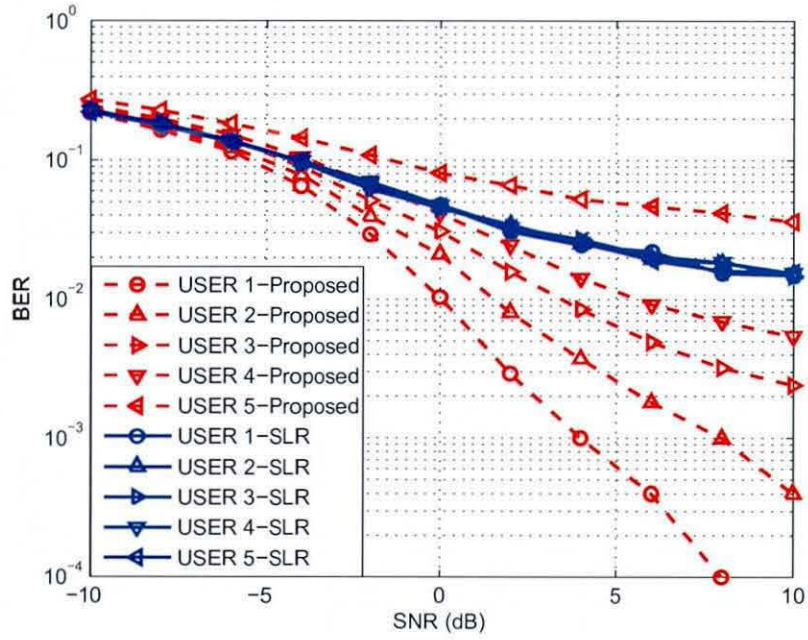

Figure 4.1: The BER performance for all the users is plotted as a function of the SNR for a MU-MIMO system with $N_T = 6$ transmit antennas and $K = 5$ users, each equipped with $N_{R_i} = 3, i = 1, 2, \ldots, 5$ receive antennas.

show the SINR outage for the proposed algorithm as compared to the SINR outage of [108] and the conventional single user beamforming solution [112]

$$\mathbf{w}_i \propto \mathcal{P}_{\max}\left(\mathbf{H}_i^H \mathbf{H}_i\right). \tag{4.21}$$

In Fig. 4.3 the proposed algorithm achieves SINR of larger than 20dB for 80% of the channel realizations at an SNR of 10 dB. Similarly in Fig. 4.4 the proposed algorithm achieves SINR of larger than 15dB for 90% of the channel realizations at an SNR of only 5 dB. Whereas both SLR and conventional beamforming have relatively poor outage performances.

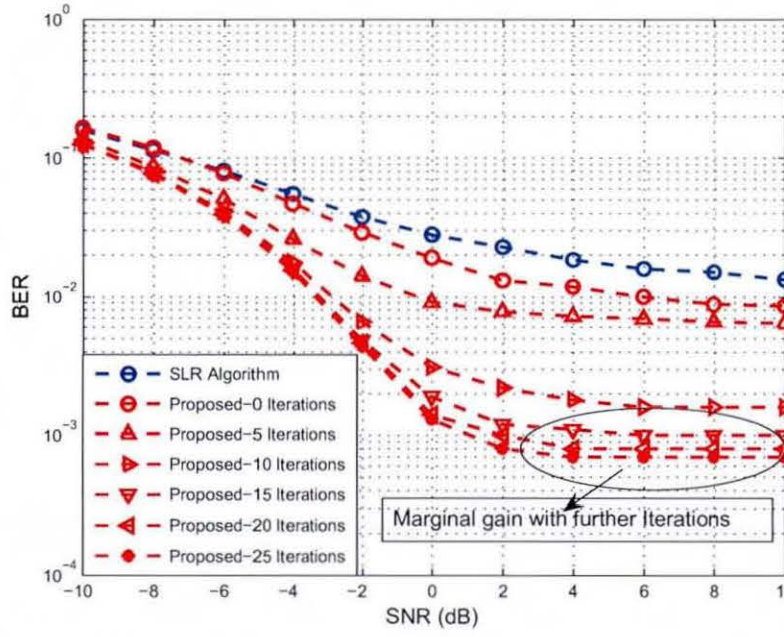

Figure 4.2: The average BER performance is plotted for SLR and the proposed algorithm for various iterations as a function of SNR for a MU-MIMO system with $N_T = 6$ transmit antennas and $K = 5$ users, each equipped with $N_{R_i} = 3, i = 1, 2, \ldots, 5$ receive antennas.

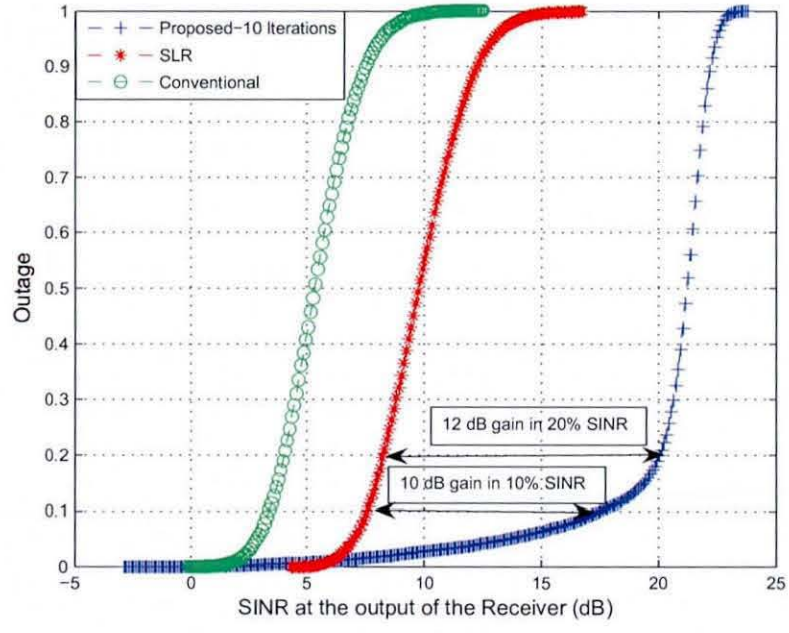

Figure 4.3: SINR outage is plotted for conventional beamforming, SLR and the proposed algorithm at the 10th iteration at a SNR of 10dB for a MU-MIMO system with $N_T = 6$ transmit antennas and $K = 5$ users, each equipped with $N_{R_i} = 3, i = 1, 2, \ldots, 5$ receive antennas.

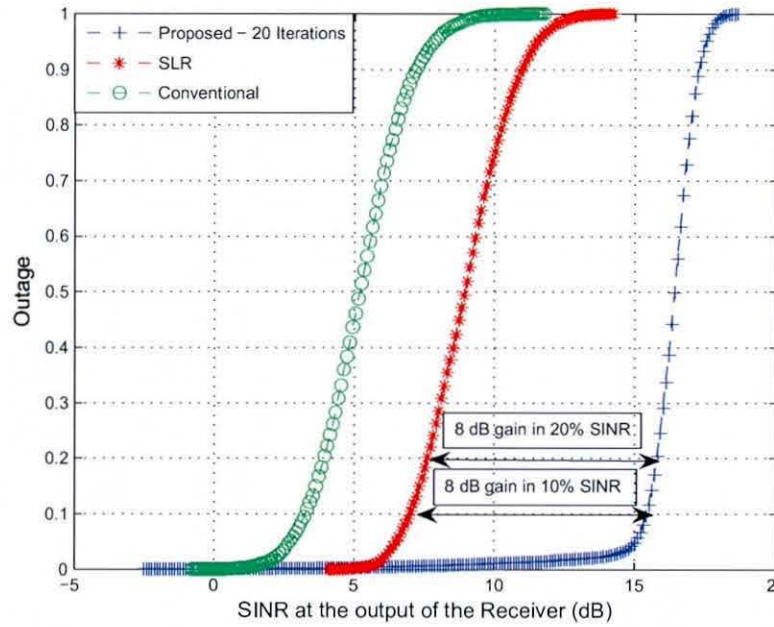

Figure 4.4: SINR outage is plotted for conventional beamforming, SLR and the proposed algorithm at the 20th iteration at a SNR of 5dB for a MU-MIMO system with $N_T = 6$ transmit antennas and $K = 5$ users, each equipped with $N_{R_i} = 3, i = 1, 2, \ldots, 5$ receive antennas.

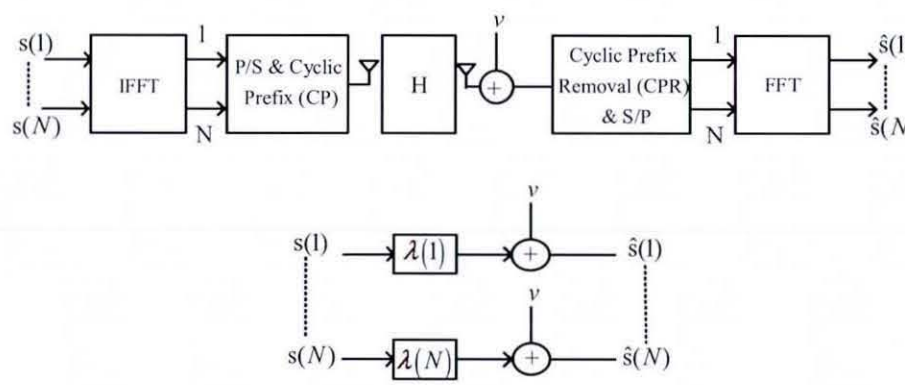

Figure 4.5: OFDM system model.

4.4 Frequency-Selective Channels

The performance of the above algorithms was tested over flat-fading channels. The algorithm can be extended to a frequency selective channels by employing techniques such as orthogonal frequency division multiplexing (OFDM). An OFDM system model is shown in Fig. 4.5. OFDM is a digital multi-carrier modulation scheme, where a large number of closely-spaced orthogonal sub-carriers are used to transmit data. The data are divided into several parallel data streams or channels, one for each sub-carrier. Each sub-carrier is modulated with a conventional modulation scheme (such as quadrature amplitude modulation (QAM) or phase shift keying (PSK)) at a low symbol rate, maintaining total data rates similar to conventional single-carrier modulation schemes in the same bandwidth [113]. OFDM transforms a frequency selective channel into parallel flat fading sub-channels, where N is the number of carriers. Therefore, the techniques developed for flat fading channels, may be simply extended to frequency selective channels using OFDM. In [2] we applied the algorithm to a frequency selective channel model and used various error control codes to evaluate system performance.

4.5 MIMO-OFDM System Model

We consider the downlink of a MU-MIMO OFDM system with a BS equipped with N subcarriers and N_T transmit antennas. There are K geographically dispersed users, each equipped with N_{R_i} receive antennas. Fig. 4.6 and Fig. 4.7 represent the block diagram for the transmitter and receiver at each user for a MU-MIMO-OFDM system. A block diagram of the MU-MIMO encoder is shown in Fig. 4.8, where $s_i(k)$ denotes the transmitted data (modulated symbol) intended for user i on the kth tone. The signal $s_i(k)$ is then multiplied by a beamformer weight vector $\mathbf{w}_i(k)$, where $\mathbf{w}_i(k)$ is the beamforming vector for user i for the kth tone. Hence, the $N_T \times 1$ signal vector

for the kth tone is given by

$$\mathbf{x}(k) = \sum_{i=1}^{K} \mathbf{w}_i(k) s_i(k). \tag{4.22}$$

It is assumed that the data $s_i(k)$ and the beamformer weights $\mathbf{w}_i(k)$ are normalized so that

$$\mathrm{E}\{|s_i(k)|^2\} = 1, \quad \|\mathbf{w}_i(k)\|^2 = 1, \quad \forall i$$

Stacking the vectors $\mathbf{x}(k)$ for $k = 0, 2, \ldots, N-1$ into a matrix of size $N_T \times N$ to form an OFDM block of transmit signal vectors that is to be transmitted over the MIMO channel is given by

$$\mathbf{X} = [\mathbf{x}(0), \mathbf{x}(1), \ldots, \mathbf{x}(N-1)], \tag{4.23}$$

where each row vector of X of size $1 \times N$ is the data vector to be transmitted over the mth transmit antenna. Before being transmitted, the data vector is modulated by an inverse Fourier transform (IDFT) into an OFDM symbol vector $\mathbf{y}_m(t)$. Then a cyclic prefix (CP) of length N_{CP} is appended to $\mathbf{y}_m(t)$. These operation may be written as

$$y_m(t) = \frac{1}{N} \sum_{k=0}^{N-1} \mathbf{X}(m,k) e^{\frac{2j\pi kt}{N}} \tag{4.24}$$

$$\hat{\mathbf{y}}_m = [y_m(N - N_{\text{CP}}), \ldots, y_m(N-1) y_m(0), \ldots, y_m(N-1)]^T \tag{4.25}$$

The OFDM symbol vector $\hat{\mathbf{y}}_m(t)$ is then transmitted through the mth antenna over a frequency selective multiuser channel of order N_h. To avoid inter-block interference (IBI), the guard interval is chosen to satisfy $N_{\text{CP}} \geq N_h - 1$. Assuming that the channel impulse response is invariant during the entire block interval, the signal received at the pth antenna of the ith user is given by

$$z_{p,i}(t) = \sum_{m=1}^{N_t} \hat{y}_m(t) * h_i^{p,m}(t) + n_{p,i}(t) \tag{4.26}$$

where $h_i^{p,m}(t)$ is the complex channel gain between the mth transmit and p^{th} receive antenna of the ith user. $n_{p,i}(t)$ is AWGN present at the pth receive antenna of the ith user. At the receiver, the CP is first removed (CPR) and then an N-point discrete Fourier transform (DFT) is performed to yield the demodulated signal vector $\tilde{\mathbf{y}}_{p,i}(k)$. This operation may be written as

$$\hat{\mathbf{z}}_{p,i} = [z_{p,i}(N_{\text{CP}}), \ldots, z_{p,i}(N_{\text{CP}}+1), \ldots, z_{p,i}(N+N_{\text{CP}}-1)]^T \tag{4.27}$$

$$\tilde{y}_{p,i}(k) = \sum_{k=0}^{N-1} \hat{z}_{p,i}(t) e^{\frac{-2j\pi kt}{N}} \tag{4.28}$$

As shown in Fig. 4.5, we can model the frequency selective channel as a collection of N parallel flat fading channels. Therefore, the received signal vector over the kth tone $\tilde{\mathbf{y}}_i(k)$ (which can also be written as $\sum_{p=1}^{N_{R_i}} \tilde{\mathbf{y}}_{p,i}(k)$), for the ith user can be written as

$$\tilde{\mathbf{x}}_i(k) = \mathbf{H}_i(k) \sum_{i=1}^{K} \mathbf{w}_i(k) s_i(k) + \tilde{\mathbf{n}}_i(k), \tag{4.29}$$

where $\hat{\mathbf{n}}_i(k)$ is zero mean circularly symmetric complex gaussian (ZMCSCG) noise vector with variance σ_i^2. The channel matrix $\mathbf{H}_i(k)$ represents the frequency response of the channel for user i for the kth tone. Assuming the ith user employs N_{R_i} antennas, the $N_{R_i} \times N_T$ channel matrix for the kth tone can be written as

$$\mathbf{H}_i(k) = \begin{bmatrix} h_i(k)^{(1,1)} & \ldots & h_i(k)^{(1,N_T)} \\ \vdots & \ddots & \vdots \\ h_i(k)^{(N_{R_i},1)} & \ldots & h_i(k)^{(N_{R_i},N_T)} \end{bmatrix} \tag{4.30}$$

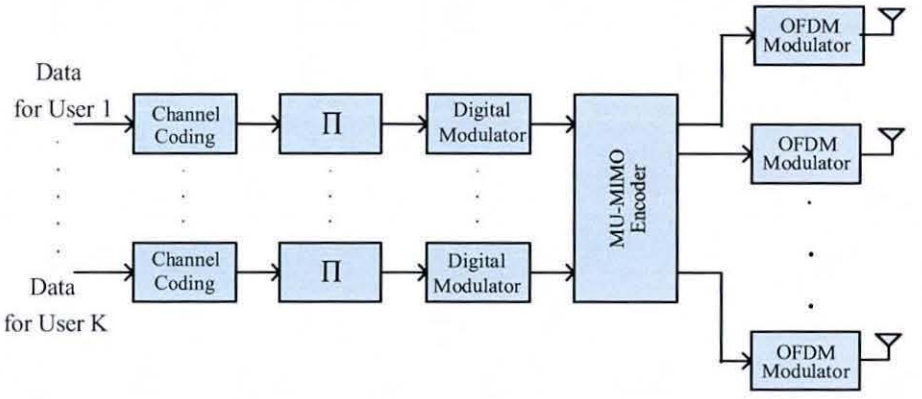

Figure 4.6: The block diagram of a MU-MIMO OFDM transmitter with N_T transmitting antennas.

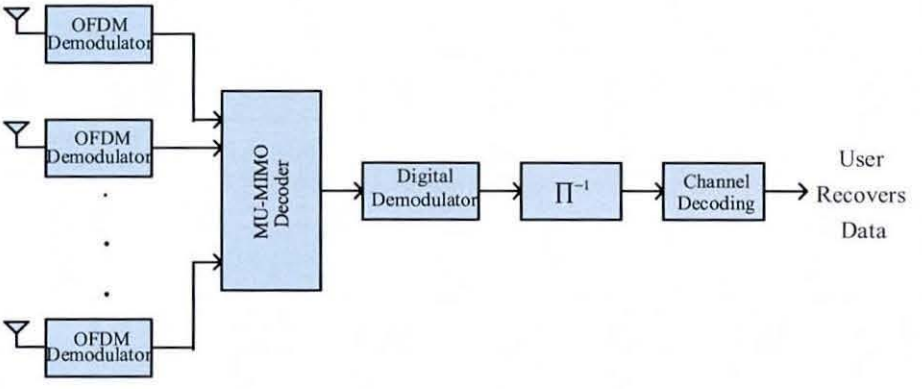

Figure 4.7: The block diagram of a MU-MIMO OFDM receiver for the ith user with N_{R_i} receive antennas.

where $h_i^{m,p}(k)$ denote the channel gain between the mth transmit and p^{th} receive antennas, for user i. Here, we assume that the receiver for user i has access to accurate CSI, $\mathbf{H}_i$.

4.6 Algorithms and Simulation Results

The algorithms presented in section 4.2.1 and 4.2.2 can be simply extended to frequency selective channel using the MIMO-OFDM system presented in section 4.4.

We considered a MU-MIMO OFDM system with one BS equipped with $N_T = 6$ antennas and $K = 5$ users each equipped with $N_{R_i} = 3$ antennas. The binary bits are generated randomly for each user and are coded using convolutional codes and

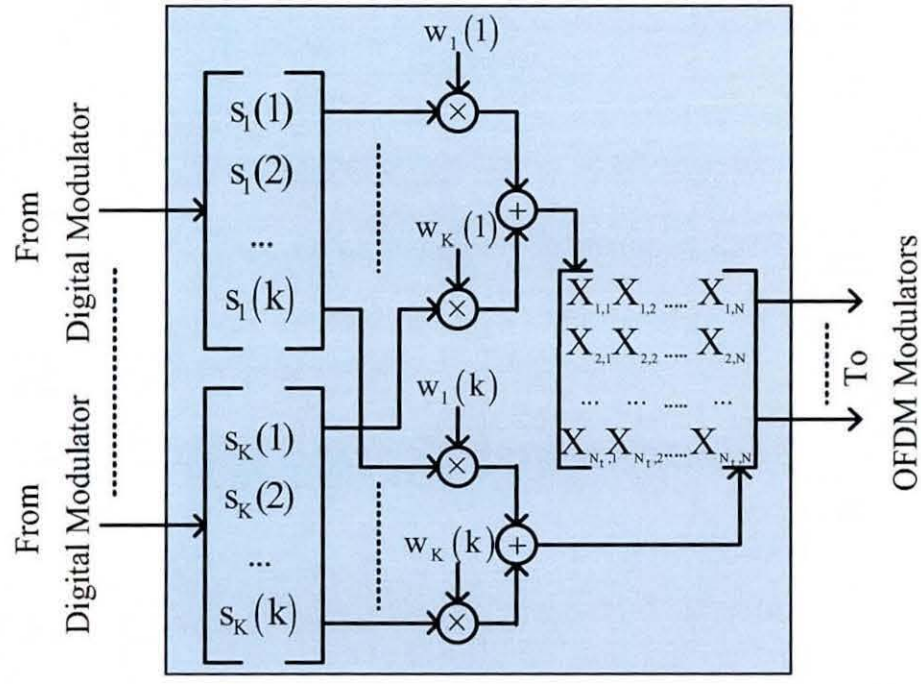

Figure 4.8: The block diagram of a MU-MIMO encoder.

randomly interleaved before being mapped into QPSK symbols. The total transmitted power per symbol period across all transmit antennas is normalized to unity. The channel taps between the pth receive and m^{th} transmit antenna $h^{p,m}$ are ZMCSCG random variables with variance $\sigma_h^2 = 1$. The noise is also ZMCSCG and spatially and temporally uncorrelated, i.e.

$$\mathrm{E}\{\mathbf{n}_i\mathbf{n}_i^H\} = \sigma_i^2\mathbf{I}_{N_{R_i}}$$

We have assumed that the BS is equipped with $N = 64$ subcarriers and the channel length between the mth transmit and p^{th} receive antenna is $N_h = 3$ and hence we use a CP of length $N_{\text{CP}} = 3$ in the OFDM modulator. In the next simulation we incorporate forward error correction and employ various convolutional codes, see Table 4.1 for code rates and generating polynomials. In Fig. 4.6, the normalized throughput of the proposed algorithm is shown at the 10th iteration as compared to the throughput of the SLR algorithm. We note that the proposed algorithm achieves peak throughput

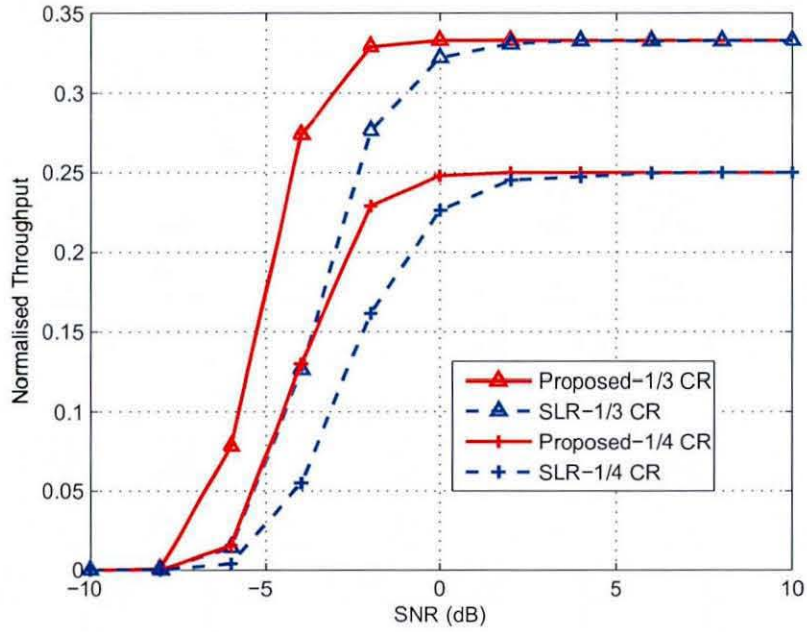

Figure 4.9: Normalized throughput for SLR and the proposed algorithm at the 10th iteration for various coding rates presented in Table 4.1.

at lower SNR, which would be even lower if further iterations are employed and the throughput is higher than the SLR algorithm for all the four different convolutional codes presented.

Table 4.1: Code rate and generating polynomials for the coders used in simulations.

Coding Rate	Generating Polynomials			
	G_1	G_2	G_3	G_4
$\frac{1}{3}$	(1100101)	(1011011)	(0000111)	-
$\frac{1}{4}$	(1000)	(1111)	(1001)	(1011)

4.7 Summary

We proposed enhancements to a recently proposed SLR design for multiuser beamformers for both flat fading and frequency selective channel environments. Our method explicitly considered the interference present at the beamformer output in-

stead of the interference present at the output of an array of antennas of any user. We demonstrated a significant improvement in the BER performance using the proposed modifications. To further increase the performance, we proposed an iterative optimization approach which also guarantees a lower error floor and equal BER performance for all users.

Chapter 5

Multiuser Spatial Diversity Techniques for Frequency Selective Channels using Uplink Downlink Duality

There has been considerable research interest in spatial multiplexing schemes, due to their potential to significantly improve spectral efficiency of a wireless communication systems [25, 114]. In particular joint *beamforming* and *power control* techniques have emerged as a suitable candidate for improving spectral efficiency at the cost of relatively increased complexity [34, 59, 80–82, 115].

In this chapter we consider the problem of downlink multiuser spatial multiplexing and power control techniques for frequency selective environment. Frequency selective channels in a multiuser environment could introduce both intersymbol interference (ISI) and interuser interference (IUI), and we propose space-time (ST) pre-processors at the transmitter so that both ISI and IUI at the multiuser terminal are mitigated. For frequency flat channels a beamformer could be employed at the transmitter to

mitigate IUI [34, 59, 60, 80–82, 115]. However, for frequency selective channels, ST filters are required at the transmitter to mitigate both IUI and ISI. We propose ST equalization (STEQ) and ST channel shortening (STCS) based schemes at the transmitter. A multiuser multiplexing based on pre-equalization is aimed at minimizing both ISI and IUI present at the user terminals. In this case the user terminal at the receiving end is not expected to perform any equalization. In contrast, a pre-channel shortening based spatial multiplexing scheme aims to mitigate IUI, but performs only a partial equalization so that the signal received at a user terminal will not have any contributions from other users. The received signal may contain controlled amount of ISI defined by the length of the target impulse response (TIR) of the channel shortening filter, i.e. the channel between the transmitter and each user terminal will be frequency selective. However the length of the effective channel will be shorter than the original channel. In this case, the receiver terminal will also need to perform equalization such as maximum likelihood detection. However, since the length of the effective channel can be controlled at the transmitter, the complexity of the receiver equalization can be controlled in the ST processor design. When the target length of the channel shortening filter is set to one, the proposed scheme will be identical to STEQ at the transmitter.

Previous work known in the context of downlink beamforming based on *uplink-downlink duality* (UDD) [34, 59, 60, 115] can be extended to frequency selective channels using orthogonal frequency division multiplexing (OFDM). In OFDM a cyclic prefix (CP) of length $N_h - 1$, where N_h is the length of the channel impulse response, is usually appended to eliminate ISI. The redundancy introduced by CP is increased as the impulse response of the channel increases. However, the proposed scheme employs time domain channel shortening filters so that OFDM with reduced CP length could be used. Moreover, the proposed scheme does not confine to OFDM schemes, and other radio access schemes such as time-division multiple access (TDMA) and code-

division multiple access (CDMA) could also benefit from the proposed approaches.

The designs for STEQ and STCS at the receiver (uplink) are well established [38, 116–120], however there has been little work performed on the design of STCS and STEQ at the transmitter (for downlink) [3]. This is due to the fact that ST filters in the uplink can be optimized independently for each user without taking into consideration of the ST filters of other users. However, in the downlink, the problem is difficult to solve as the ST filters of all users need to be optimized jointly. Here, we show that the problem of optimizing power and ST filters in the downlink is equivalent to solving a virtual uplink problem. Here we extend the results of UDD known for flat fading channels to design STCS and STEQ filters for a frequency selective environment. We also takes into consideration the complexity of the receiver which is controlled by the length of the TIR. Simulation results show that the STCS based design provides a superior performance in terms of minimizing the total transmit power over a STEQ based design.

5.1 Downlink Spatial Multiplexing System Model

A downlink spatial multiplexing system based on a ST filter at the transmitter is shown in Fig. 5.1, where $\mathbf{x}(k) = [x_1(k), x_2(k), \ldots, x_K(k)]^T \in \mathbb{C}^{K\times 1}$ denotes the signal vector to be transmitted to K users at time k. The signal component $x_j(k), j = 1, 2, \ldots K$ denotes the data symbol intended for the jth user. The signal transmitted from the ith antenna ($i = 1, 2, \ldots, N_T$) at time k can be described as

$$r_i^{\text{DL}}(k) = \sum_{j=1}^{K} \sum_{n=0}^{N_F-1} w(n)^{j,i} q_j^{1/2} x_j(k-n)$$

where $w^{j,i}(n)$ is the gain of the nth tap ($n = 0, 1, \ldots, N_f - 1$) of the filter at the ith transmit antenna and designed for the jth user, q_j is the power allocated to the jth user, i.e. the jth element of the power allocation vector $\mathbf{q} = [q_1, q_2, \ldots, q_K]^T \in$

Table 5.1: Vectors and matrices used in the chapter

Vector / Matrix	Dimension
$\mathbf{x}(k)$	$\mathbb{C}^{K\times 1}$
$\mathbf{x}^k_{k-(N_f-1)}$	$\mathbb{C}^{KN_f\times 1}$
$\mathbf{x}^k_{k-(N_h-1)}$	$\mathbb{C}^{KN_h\times 1}$
$\mathbf{x}^k_{k-(N_f+N_h-1)}$	$\mathbb{C}^{K(N_f+N_h-1)}$
$\mathbf{r}^k_{k-(N_h-1)}$	$\mathbb{C}^{KN_h\times 1}$
$\mathbf{y}^k_{k-(N_f-1)}$	$\mathbb{C}^{KN_h\times 1}$
$\mathbf{q}$	$\mathbb{R}^{K\times 1}$
$\mathbf{p}$	$\mathbb{R}^{K\times 1}$
$\mathbf{Q}$	$\mathbb{R}^{K\times K}$
$\mathbf{P}$	$\mathbb{R}^{K\times K}$
$\hat{\mathbf{Q}}$	$\mathbb{R}^{KN_f\times KN_f}$
$\hat{\mathbf{P}}$	$\mathbb{R}^{KN_h\times KN_h}$
$\bar{\mathbf{Q}}$	$\mathbb{R}^{K(N_f+N_h-1)\times K(N_f+N_h-1)}$
$\bar{\mathbf{P}}$	$\mathbb{R}^{K(N_f+N_h-1)\times K(N_f+N_h-1)}$
$\mathbf{W}(n)$	$\mathbb{C}^{N_T\times K}$
$\hat{\mathbf{W}}$	$\mathbb{C}^{N_T\times KN_f}$
$\mathbf{w}_j$	$\mathbb{C}^{N_TN_f\times 1}$
$\bar{\mathbf{W}}$	$\mathbb{C}^{N_TN_h\times K(N_f+N_h-1)}$
$\mathbf{H}(m)$	$\mathbb{C}^{N_T\times K}$
$\hat{\mathbf{H}}$	$\mathbb{C}^{K\times N_TN_h}$
$\mathbf{h}_j$	$\mathbb{C}^{N_TN_h}$
$\bar{\mathbf{H}}$	$\mathbb{C}^{N_TN_f\times K(N_f+N_h-1)}$
$\boldsymbol{\beta}$	$\mathbb{R}^{K\times K}$
$\mathbf{b}_j$	$\mathbb{C}^{1\times N_b}$
$\tilde{\mathbf{b}}_j$	$\mathbb{C}^{1\times K(N_f+N_h-1)}$
$\mathbf{B}$	$\mathbb{C}^{K\times N_b}$
$\mathbf{h}^{\mathrm{UL}}_{\mathrm{eff},j}$	$\mathbb{C}^{1\times K(N_f+N_h-1)}$
$\mathbf{h}^{\mathrm{DL}}_{\mathrm{eff},j}$	$\mathbb{C}^{1\times K(N_f+N_h-1)}$
$\mathbf{h}^{\mathrm{UL}}_{\mathrm{win},j}$	$\mathbb{C}^{1\times K(N_f+N_h-1)}$
$\mathbf{h}^{\mathrm{UL}}_{\mathrm{wall},j}$	$\mathbb{C}^{1\times K(N_f+N_h-1)}$
$\mathbf{h}^{\mathrm{DL}}_{\mathrm{win},j}$	$\mathbb{C}^{1\times K(N_f+N_h-1)}$
$\mathbf{h}^{\mathrm{DL}}_{\mathrm{wall},j}$	$\mathbb{C}^{1\times K(N_f+N_h-1)}$
$\bar{\mathbf{H}}_j$	$\mathbb{C}^{N_TN_f\times K(N_f+N_h-1)}$
$\bar{\mathbf{H}}_{\mathrm{win}}$	$\mathbb{C}^{N_TN_f\times K(N_f+N_h-1)}$
$\bar{\mathbf{H}}_{\mathrm{wall}}$	$\mathbb{C}^{N_TN_f\times K(N_f+N_h-1)}$
$\bar{\mathbf{W}}_j$	$\mathbb{C}^{N_TN_h\times K(N_f+N_h-1)}$
$\bar{\mathbf{W}}_{\mathrm{win}}$	$\mathbb{C}^{N_TN_h\times K(N_f+N_h-1)}$
$\bar{\mathbf{W}}_{\mathrm{wall}}$	$\mathbb{C}^{N_TN_h\times K(N_f+N_h-1)}$
Γ^{UL}	$\mathbb{R}^{K\times K}$
Γ^{DL}	$\mathbb{R}^{K\times K}$
Υ^{UL}	$\mathbb{R}^{K\times K}$
Υ^{DL}	$\mathbb{R}^{K\times K}$
$\boldsymbol{\sigma}$	$\mathbb{R}^{K\times 1}$
$\mathbf{p}_{ext}$	$\mathbb{R}^{(K+1)\times 1}$
Φ^{UL}	$\mathbb{R}^{(K+1)\times(K+1)}$

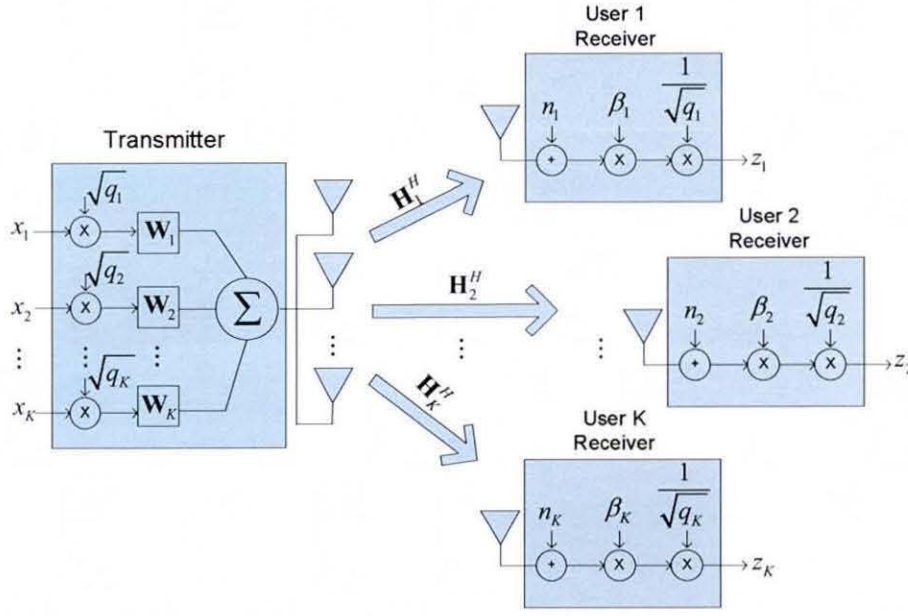

Figure 5.1: Downlink spatial multiplexing system model.

$\mathbb{R}^{K\times 1}$. Stacking $r_i^{\mathrm{DL}}(k)$ into a vector $\mathbf{r}^{\mathrm{DL}}(k) = [r_1(k), r_2(k), \ldots, r_{N_T}(k)]^T \in \mathbb{C}^{N_T\times 1}$ and defining $\mathbf{Q} = \mathrm{diag}\{[q_1, q_2, \ldots, q_K]\} \in \mathbb{R}^{K\times K}$, we could write the transmitted signal as

$$\begin{bmatrix} r_1^{\mathrm{DL}}(k) \\ r_2^{\mathrm{DL}}(k) \\ \vdots \\ r_{N_T}^{\mathrm{DL}}(k) \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{W}(0) & \mathbf{W}(1) & \ldots & \mathbf{W}(N_f-1) \end{bmatrix}}_{\hat{\mathbf{W}}} \underbrace{\begin{bmatrix} \mathbf{Q}^{1/2} & 0 & \ldots & 0 \\ 0 & \mathbf{Q}^{1/2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \mathbf{Q}^{1/2} \end{bmatrix}}_{\hat{\mathbf{Q}}^{1/2}}$$

$$\underbrace{\begin{bmatrix} \mathbf{x}(k) \\ \mathbf{x}(k-1) \\ \vdots \\ \mathbf{x}(k-(N_f-1)) \end{bmatrix}}_{\mathbf{x}_{k-(N_f-1)}^{k}}$$

$$\mathbf{r}^{\mathrm{DL}}(k) = \sum_{n=0}^{N_f-1} \mathbf{W}(n)\mathbf{Q}^{1/2}\mathbf{x}(k-n) = \hat{\mathbf{W}}\hat{\mathbf{Q}}^{1/2}\mathbf{x}_{k-(N_f-1)}^{k} \tag{5.1}$$

where $\hat{\mathbf{W}} \in \mathbb{C}^{N_T \times KN_f}$ denotes the ST filter at the transmitter, $\hat{\mathbf{Q}} \in \mathbb{R}^{KN_h \times KN_h}$ is a block diagonal matrix with $\mathbf{Q}$ as its block, and the the matrix

$$\mathbf{W}(n) = \begin{bmatrix} w^{1,1}(n) & w^{1,2}(n) & \dots & w^{1,K}(n) \\ w^{2,1}(n) & w^{2,2}(n) & \dots & w^{2,K}(n) \\ \vdots & & \ddots & \vdots \\ w^{N_T,1}(n) & w^{N_T,2}(n) & \dots & w^{N_T,K}(n) \end{bmatrix}$$

contains the 'snap-shot' filter parameters for the nth delay. The ST filter for j^{th} user can be written in a vector form as

$$\begin{aligned} \mathbf{w}_j &= \left[(\mathbf{W}(0)^{1:N_T,j})^T \ (\mathbf{W}(1)^{1:N_T,j})^T \ \dots \ (\mathbf{W}(N_f-1)^{1:N_T,j})^T \right]^T \\ &= \left[(\mathbf{w}_j(0))^T \ (\mathbf{w}_j(1))^T \ \dots \ (\mathbf{w}_j(N_f-1))^T \right]^T . \end{aligned} \tag{5.2}$$

Let β_j denote the norm of $\mathbf{w}_j$ before normalization and normalize ST filter of each user such that $\|\mathbf{w}_j\|_2 = 1$. The vector $\mathbf{r}_k^{\text{DL}}$ is then transmitted over a frequency selective channel. The signal received at the jth user at time k can be described as,

$$y_j^{\text{DL}}(k) = \sum_{i=1}^{N_T} \sum_{m=0}^{N_h-1} (h^{i,j}(m))^* r_i^{\text{DL}}(k-m) + n_j(k)$$

where $h^{i,j}(m)$ is the gain of the mth path ($m = 0, 1, \dots, N_h - 1$) of the channel between the ith transmit antenna and the jth user and $n_j(k)$ denotes additive white Gaussian noise (AWGN) at the jth user receive with variance σ_n^2. The received signal $y_j^{\text{DL}}(k)$ for all users (only for the purpose of designing the transmitter filters) can be stacked into a vector $\mathbf{y}^{\text{DL}}(k)$ as

$$
\begin{aligned}
\mathbf{y}^{\mathrm{DL}}(k) &= \sum_{m=0}^{N_h-1} \mathbf{H}^H(m)\mathbf{r}^{\mathrm{DL}}(k-m) + \mathbf{n}(k) \\
\begin{bmatrix} y_1^{\mathrm{DL}}(k) \\ y_2^{\mathrm{DL}}(k) \\ \vdots \\ y_{n_T}^{\mathrm{DL}}(k) \end{bmatrix} &= \underbrace{\begin{bmatrix} (\mathbf{H}(0))^H & (\mathbf{H}(1))^H & \dots & (\mathbf{H}(N_h-1))^H \end{bmatrix}}_{\hat{\mathbf{H}}^H} \\
&\quad \underbrace{\begin{bmatrix} \mathbf{W}(0) & \dots & \mathbf{W}(N_f-1) & 0 & \dots & 0 \\ 0 & \mathbf{W}(0) & \dots & \mathbf{W}(N_f-1) & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \mathbf{W}(0) & \dots & \mathbf{W}(N_f-1) \end{bmatrix}}_{\bar{\mathbf{W}}} \\
&\quad \underbrace{\begin{bmatrix} \hat{\mathbf{Q}}^{1/2} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{Q}}^{1/2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \hat{\mathbf{Q}}^{1/2} \end{bmatrix}}_{\bar{\mathbf{Q}}^{1/2}} \underbrace{\begin{bmatrix} \mathbf{x}(k) \\ \mathbf{x}(k-1) \\ \vdots \\ \mathbf{x}(k-(N_f-N_h-1)) \end{bmatrix}}_{\mathbf{x}_{k-(N_f-N_h-1)}^k} \\
&= \hat{\mathbf{H}}^H \bar{\mathbf{W}} \bar{\mathbf{Q}}^{1/2} \mathbf{x}_{k-(N_f-N_h-1)}^k \qquad (5.3)
\end{aligned}
$$

where $\hat{\mathbf{H}}^H \in \mathbb{C}^{K \times N_T N_h}$ denotes the downlink channel matrix between the BS and the users, $\bar{\mathbf{W}} \in \mathbb{C}^{N_T N_h \times K(N_f+N_h-1)}$ is a convolutional matrix consisting of the ST filter coefficients $\hat{\mathbf{W}}$. We assume the power allocation matrix is fixed over a block period, so that $\bar{\mathbf{Q}} \in \mathbb{R}^{K(N_f+N_h-1) \times K(N_f+N_h-1)}$ is a block diagonal matrix consisting of the power allocation matrix $\hat{\mathbf{Q}}$ as its block element.The matrix

$$\mathbf{H}(m) = \begin{bmatrix} h^{1,1}(m) & h^{1,2}(m) & \dots & h^{1,K}(m) \\ h^{2,1}(m) & h^{2,2}(m) & \dots & h^{2,K}(m) \\ \vdots & & \ddots & \vdots \\ h^{N_T,1}(m) & h^{N_T,2}(m) & \dots & h^{N_T,K}(m) \end{bmatrix}$$

contains the 'snap-shot' channel parameters of the mth path. The jth user employs a simple receiver denoted by $q_j^{-1/2}\beta_j$. The estimate $z_j^{\text{DL}}(k)$ of the data $x_j(k)$ transmitted to the jth user is written as

$$\begin{aligned} z_j^{\text{DL}}(k) &= q_j^{-1/2}\beta_j \mathbf{h}_j^H \mathbf{y}^{\text{DL}}(k) \\ &= \frac{\beta_j}{q_j^{1/2}}\left(\mathbf{h}_j^H \bar{\mathbf{W}}\bar{\mathbf{Q}}^{1/2}\mathbf{x}_{k-(N_F-N_h-1)}^k + n_j(k)\right) \end{aligned} \tag{5.4}$$

where $\mathbf{h}^j$ is the channel vector between the BS and the jth user, defined in (5.5).

$$\begin{aligned} \mathbf{h}_j &= \left[\ (\mathbf{H}(0)^{1:N_T,j})^T\ (\mathbf{H}(1)^{1:N_T,j})^T\ \dots\ (\mathbf{H}(N_h-1)^{1:N_T,j})^T\ \right]^T \\ &= \left[\ (\mathbf{h}_j(0))^T\ (\mathbf{h}_j(1))^T\ \dots\ (\mathbf{h}_j(N_h-1))^T\ \right]^T \end{aligned} \tag{5.5}$$

5.1.1 Downlink MMSE based STEQ and STCS filters

Fig. 5.2 shows the structure of the minimum mean square error (MMSE) STCS filter for the downlink. The STCS filter for the jth user $\mathbf{w}_j$ is designed to transform the channel impulse response $\mathbf{h}_j$ of the jth user to the TIR vector

$$\mathbf{b}_j = \left[\ b_j(0)\ \ b_j(1)\ \ \dots\ \ b_j(N_b-1)\ \right] \in \mathbb{C}^{1\times N_b},\ j = 1, 2, \dots, K,$$

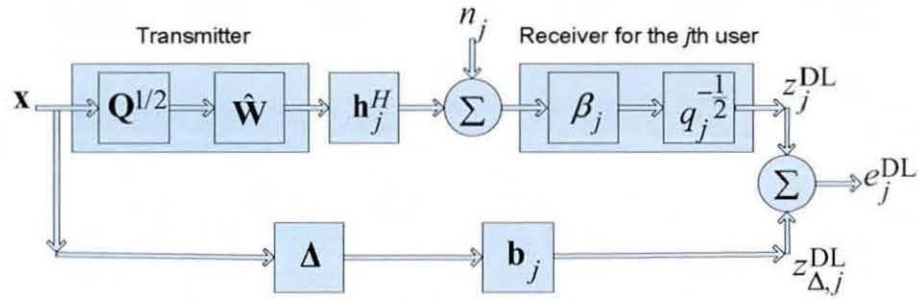

Figure 5.2: Downlink MMSE ST filter design

where N_b is the TIR length. Define an augmented TIR vector for the jth user as (5.6), where Δ is the equalization delay in the range $\Delta = 0, 1, \ldots, N_h + N_f - N_b - 1$ and $s = N_h + N_f - N_b - \Delta - 1$. The first and the last all-zero vectors are to account for the suppression of $N_h - N_b - 1$ taps with equalization delay Δ, and the remaining all zeros vectors correspond to the the signals coming from the other users to be suppressed. The (channel shortening) equalization error $e_j^{\text{DL}}(k)$ for the jth user, is the difference between the output signal $z_j^{\text{DL}}(k)$ and the target output $z_{\Delta,j}^{\text{DL}}(k)$

$$\begin{aligned} e_j^{\text{DL}}(k) &= z_{\Delta,j}^{\text{DL}}(k) - z_j^{\text{DL}}(k) \\ &= \tilde{\mathbf{b}}_j^H \mathbf{x}_{k-(N_f+N_h-1)}^k - z_{\Delta,j}^{\text{DL}}(k). \end{aligned}$$

$$\tilde{\mathbf{b}}_j = \left[\mathbf{0}_{1\times N_T\Delta} | b_j(0)\ \mathbf{0}_{1\times N_T} b_j(1)\ \mathbf{0}_{1\times N_T} \ldots b_j(N_b-1)\ \mathbf{0}_{1\times N_T} | \mathbf{0}_{1\times s} \right] \in \mathbb{C}^{1\times K(N_f+N_h-1)} \tag{5.6}$$

Let us now calculate the mean square error (MSE) $\varepsilon_j^{\text{DL}}$ for the jth user in the downlink as

$$\begin{aligned}\varepsilon_j^{\mathrm{DL}} &= \mathrm{E}\left\{\left(e_j^{\mathrm{DL}}(k)\right)\left(e_j^{\mathrm{DL}}(k)\right)^*\right\} \\ &= \frac{\beta_j^2}{q_j}\mathbf{h}_j^H\bar{\mathbf{W}}\bar{\mathbf{Q}}\bar{\mathbf{W}}\mathbf{h}_j - \frac{2\beta_j}{\sqrt{q_j}}\mathbf{h}_j^H\bar{\mathbf{W}}\bar{\mathbf{Q}}\tilde{\mathbf{b}}_j + \frac{\beta_j^2}{q_j}\sigma_j^2 + \tilde{\mathbf{b}}_j^H\tilde{\mathbf{b}}_j\end{aligned} \tag{5.7}$$

It should be noted that (5.7) is coupled with K channel shortening filters and the transmit power allocation $\mathbf{q}$. Hence it is quite difficult to optimize these powers and filters jointly. Note, when N_b is set to 1, we obtain a STEQ based design, which is also coupled in K ST filers and the power allocation $\mathbf{q}$. In the next section we will show that this problem is equivalent to solving a virtual uplink problem using UDD theorem.

5.2 Problem Statement

We wish to solve the downlink spatial multiplexing problem for the following optimization criteria:

5.2.1 Criterion 1 (C1): Max-Min Fairness

$$\begin{aligned}\min_{\hat{\mathbf{W}},\boldsymbol{\beta},\mathbf{q},\mathbf{B}} \quad & \max_{1\le j\le K} \varepsilon_j^{\mathrm{DL}} \\ \text{s.t.} \quad & \|\mathbf{q}\|_1 \le P_{max}\end{aligned} \tag{5.8}$$

where $\varepsilon_j^{\mathrm{DL}}$ is the downlink MSE of the jth user, $\hat{\mathbf{W}}$ is the ST filter, $\mathbf{q}$ is the power allocation vector, $\mathbf{B} = [\mathbf{b}_1^T, \mathbf{b}_2^T, \ldots, \mathbf{b}_K^T]^T$ is the TIR of all users, $\boldsymbol{\beta} = \mathrm{diag}\{[\beta_1, \beta_2, \ldots, \beta_K]\}$ is a diagonal matrix with the norms of ST filters and $P_{\max}$ is the maximum possi-

ble transmission power. Here we minimize the maximum MSE among all users in the system, subject to a total power constraint. Later we will show that this design effectively balances the MSE of all users.

5.2.2 Criterion 2 (C2): Min Power

$$\begin{aligned} \min_{\hat{\mathbf{W}},\boldsymbol{\beta},\mathbf{q},\mathbf{B}} \quad & \|\mathbf{q}\|_1 \\ \text{s.t.} \quad & \varepsilon_j^{\mathrm{DL}} \leq \xi_j^{\mathrm{DL}} \\ & j = 1, \ldots, K \end{aligned} \tag{5.9}$$

where ξ_j^{DL} is the MSE target for the jth user. Here we minimize the total transmission power subject to constraints imposed on the required quality of services.

5.3 Virtual Uplink Model

Fig.5.3 shows a virtual uplink model. This model is obtained by switching the roles of the transmitter and the receiver from the downlink. It is assumed that the quantities $\hat{\mathbf{W}}$, $\hat{\mathbf{H}}$, $\mathbf{B}$, $\boldsymbol{\beta}$ and σ_j^2 are the same as in the downlink, however the power allocation vector $\mathbf{p}$ may differ from the downlink power allocation $\mathbf{q}$. Here K users transmit simultaneously to a single BS. The signal transmitted by the jth user in the uplink is given by

$$r_j^{\mathrm{UL}}(k) = p_j^{1/2} x_j(k).$$

Collecting the signal $x_j(k)$ for $j = 1, 2, \ldots, K$ into vector $\mathbf{x}(k) \in \mathbb{C}^{K\times 1}$ and defining $\mathbf{P} = \mathrm{diag}\{[\mathbf{p}\}$, where $\mathbf{p} = [p_1, p_2, \ldots, p_K]^T$, is the power allocation vector in the uplink, the signal received by the BS from K users can be written as

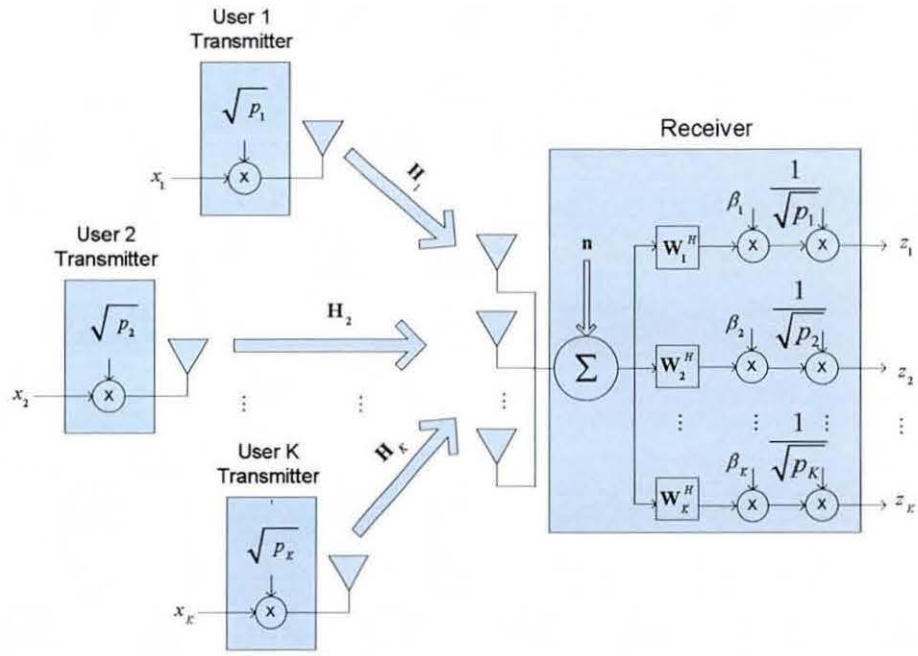

Figure 5.3: Virtual uplink model.

$$\begin{bmatrix} y_1^{\mathrm{UL}}(k) \\ y_2^{\mathrm{UL}}(k) \\ \vdots \\ y_{N_T}^{\mathrm{UL}}(k) \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{H}(0) & \mathbf{H}(1) & \dots & \mathbf{H}(N_h-1) \end{bmatrix}}_{\hat{\mathbf{H}}} \underbrace{\begin{bmatrix} \mathbf{P}^{1/2} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{P}^{1/2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{P}^{1/2} \end{bmatrix}}_{\hat{\mathbf{P}}^{1/2}}$$

$$\underbrace{\begin{bmatrix} \mathbf{x}(k) \\ \mathbf{x}(k-1) \\ \vdots \\ \mathbf{x}(k-(N_h-1)) \end{bmatrix}}_{\mathbf{x}_{k-(n_h-1)}^{k}} + \begin{bmatrix} n_1(k) \\ n_2(k) \\ \vdots \\ n_{N_T}(k) \end{bmatrix}$$

$$\mathbf{y}^{\mathrm{UL}}(k) = \sum_{m=0}^{N_h-1} \mathbf{H}(m)\mathbf{P}^{1/2}\mathbf{x}(k-m) + \mathbf{n}(k) = \hat{\mathbf{H}}\hat{\mathbf{P}}^{1/2}\mathbf{x}_{k-(N_h-1)}^{k} + \mathbf{n}(k) \quad (5.10)$$

where $\hat{\mathbf{H}} \in \mathbb{C}^{N_T \times KN_h}$ denotes the channel matrix in the uplink, $\hat{\mathbf{P}} \in \mathbb{R}^{KN_h \times KN_h}$

is a block diagonal matrix with $\mathbf{P}$ as its block. The BS uses a ST filter to separate the users. The receiver for the jth user is given by $p_j^{-1/2}\beta_j\mathbf{w}_j^H$. To process the channel output $\mathbf{y}^{\text{UL}}(k)$ with the filter $p_j^{-1/2}\beta_j\mathbf{w}_j^H$, $\mathbf{y}^{\text{UL}}(k)$ should be accumulated into a regressor of length $N_T N_f$

$$\underbrace{\begin{bmatrix} \mathbf{y}(k) \\ \mathbf{y}(k-1) \\ \vdots \\ \mathbf{y}(k-(N_f-1)) \end{bmatrix}}_{\mathbf{y}^k_{k-(N_f-1)}} = \underbrace{\begin{bmatrix} \mathbf{H}(0) & \dots & \mathbf{H}(N_h-1) & 0 & \dots & 0 \\ 0 & \mathbf{H}(0) & \dots & \mathbf{H}(N_h-1) & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \mathbf{H}(0) & \dots & \mathbf{H}(N_h-1) \end{bmatrix}}_{\bar{\mathbf{H}}}$$

$$\underbrace{\begin{bmatrix} \hat{\mathbf{P}}^{1/2} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \hat{\mathbf{P}}^{1/2} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \hat{\mathbf{P}}^{1/2} \end{bmatrix}}_{\bar{\mathbf{P}}^{1/2}} \underbrace{\begin{bmatrix} \mathbf{x}(k) \\ \mathbf{x}(k-1) \\ \vdots \\ \mathbf{x}(k-(N_f-N_h-1)) \end{bmatrix}}_{\mathbf{x}^k_{k-(N_f-N_h-1)}}$$

$$+ \underbrace{\begin{bmatrix} \mathbf{n}(k) \\ \mathbf{n}(k-1) \\ \vdots \\ \mathbf{n}_{k-(N_f-1)} \end{bmatrix}}_{\mathbf{n}^k_{k-(N_f-1)}}$$

$$\mathbf{y}^k_{k-(N_f-1)} = \bar{\mathbf{H}}\bar{\mathbf{P}}^{1/2}\mathbf{x}^k_{k-(N_f-N_h-1)} + \mathbf{n}^k_{k-(N_f-1)} \tag{5.11}$$

$\bar{\mathbf{H}} \in \mathbb{C}^{N_T N_f \times K(N_f+N_h-1)}$ is a convolutional matrix consisting of channel coefficients $\hat{\mathbf{H}}$. Assuming the power allocation matrix is fixed over a block period, then $\bar{\mathbf{P}} \in \mathbb{R}^{K(N_f+N_h-1)\times K(N_f+N_h-1)}$ is a block diagonal matrix consisting of power allocation matrix $\hat{\mathbf{P}}$ as its block element. The estimate $z_j^{\text{UL}}(k)$ of data $x_j(k)$ transmitted by the

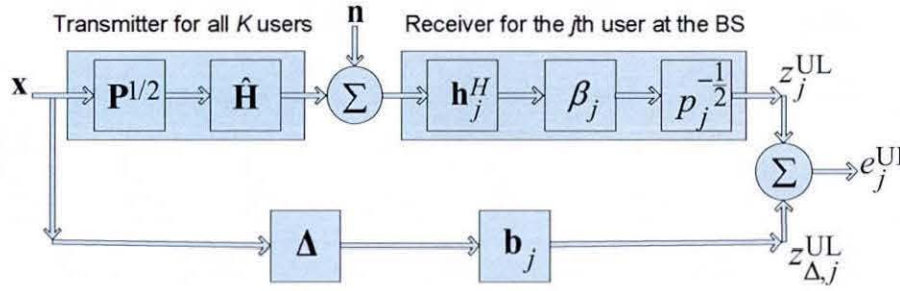

Figure 5.4: Uplink MMSE ST filter design.

j^{th} user at the output of the BS receiver is given by

$$\begin{aligned} z_j^{\text{UL}}(k) &= p_j^{-1/2}\beta_j \mathbf{w}_j^H \mathbf{y}_{k-(N_f-1)^k} \\ &= \frac{\beta_j}{\sqrt{p_j}} \mathbf{w}_j^H \left(\bar{\mathbf{H}}\bar{\mathbf{P}}^{1/2}\mathbf{x}_{k-(N_h+N_F-1)}^k + \mathbf{n}_{k-(Nf-1)}^k \right), \end{aligned} \tag{5.12}$$

where $\mathbf{w}_j \in \mathbb{C}^{N_T N_f \times 1}$ denotes the ST filter for the jth user in the uplink defined in (5.2).

5.3.1 Uplink MMSE based STEQ and STCS filters

Fig.5.4 shows the structure of the MMSE STCS filter in the uplink. We define an equalization error $e_j^{\text{UL}}(k)$ for the jth user as

$$\begin{aligned} e_j^{\text{UL}}(k) &= z_{\Delta,j}^{\text{UL}}(k) - z_j^{\text{UL}}(k) \\ &= \tilde{\mathbf{b}}_j^H \mathbf{x}_{k-(N_h+N_f-1)}^k - z_{\Delta,j}^{\text{UL}}(k) \end{aligned}$$

and determine the corresponding MSE $\varepsilon_j^{\text{UL}}$ as,

$$\begin{aligned}\varepsilon_j^{\mathrm{UL}} &= \mathrm{E}\left\{\left(e_j^{\mathrm{UL}}(k)\right)\left(e_j^{\mathrm{UL}}(k)\right)^*\right\} \\ &= \frac{\beta_j^2}{p_j}\mathbf{w}_j^H\left(\bar{\mathbf{H}}\bar{\mathbf{P}}\bar{\mathbf{H}}+\sigma_j^2\mathbf{I}\right)\mathbf{w}_j - \frac{2\beta_j}{\sqrt{p_j}}\mathbf{w}_j^H\bar{\mathbf{H}}\bar{\mathbf{P}}\tilde{\mathbf{b}}_j + \tilde{\mathbf{b}}_j^H\tilde{\mathbf{b}}_j. \qquad (5.13)\end{aligned}$$

We note that unlike (5.7), (5.13) is uncoupled as it is a function of the channel shortening filter $\mathbf{w}_j$ and the power allocation p_j for the jth user. Hence these variables can be optimized independently for each user. Differentiating (5.13) with respect to $\mathbf{w}_j$ and equating the result to zero, we obtain [119, 120]

$$\mathbf{w}_{j,\text{opt}} = \frac{\sqrt{p_j}}{\beta_j}\left(\bar{\mathbf{H}}\bar{\mathbf{P}}\bar{\mathbf{H}}^H+\sigma_n^2\mathbf{I}\right)^{-1}\bar{\mathbf{H}}\bar{\mathbf{P}}^{1/2}\tilde{\mathbf{b}}_j. \qquad (5.14)$$

Defining

$$\mathbf{R} = \left(\mathbf{I} - \bar{\mathbf{P}}^{1/2}\bar{\mathbf{H}}^H\left(\bar{\mathbf{H}}\bar{\mathbf{P}}\bar{\mathbf{H}}^H+\sigma_j^2\mathbf{I}\right)^{-1}\bar{\mathbf{H}}\bar{\mathbf{P}}^{1/2}\right), \qquad (5.15)$$

and substituting (5.14) into (5.13), we obtain $\varepsilon_j^{\mathrm{UL}}$ in terms of the TIR vector $\tilde{\mathbf{b}}_i$ [119, 120]

$$\begin{aligned}\varepsilon_j^{\mathrm{UL}} &= \tilde{\mathbf{b}}_j^H\mathbf{R}\tilde{\mathbf{b}}_j \\ &= \mathbf{b}_j^H\mathbf{R}_j\mathbf{b}_j \qquad (5.16)\end{aligned}$$

where $\mathbf{R}_j$ contains the parts of the matrix $\mathbf{R}$ selected by the nonzero part of the $\tilde{\mathbf{b}}_j$. To further minimize the MSE we need to optimize (5.16) subject to a constraint $\mathbf{b}_j^H\mathbf{b}_j = 1$ i.e.,

$$\begin{aligned} \min_{\mathbf{b}_j} \quad & \varepsilon_j^{\mathrm{UL}} \\ \text{s.t.} \quad & \mathbf{b}_j^H \mathbf{b}_j = 1. \end{aligned} \tag{5.17}$$

The solution to (5.17) is obtained using eigendecompostion, where $\mathbf{b}_{j,\mathrm{opt}}$ is given by [119, 120]

$$\mathbf{b}_{j,\mathrm{opt}} = \mathcal{P}_{\min}\Big(\mathbf{R}_j\Big) \tag{5.18}$$

where, $\mathcal{P}_{\min}(\cdot)$ returns the eigenvector corresponding to the smallest eigenvalue. The algorithm to compute $\mathbf{w}_{j,\mathrm{opt}}$, $\mathbf{b}_{j,\mathrm{opt}}$ and the optimum equalization delay $\Delta_{j,\mathrm{opt}}$ is summarized in algorithm 3. Note we use the above scheme to obtain results for STEQ based method by setting $N_b = 1$.

5.4 Effective Channels in the Uplink and the Downlink

From previous section, we note that ST filters in the uplink are optimized independently for each user. Therefore instead of solving C1 or C2 in the downlink, it is easier to solve the equivalent versions of C1 or C2 in the uplink. To obtain the equivalent formulations of C1 and C2 in the uplink, we first need to establish a duality between the downlink and uplink. To facilitate this we go through the following definitions. The effective channel seen at the BS in the uplink from the jth user is given by

$$\mathbf{h}_{\mathrm{eff},j}^{\mathrm{UL}} = \mathbf{w}_j^H \bar{\mathbf{H}}.$$

Algorithm 3 MMSE Channel Shortening Design

1. INITIALIZE
   - $\Delta_{\max} = N_h + N_F - N_b - 1$.
   - For a given power allocation vector $\mathbf{p}$, construct $\bar{\mathbf{P}}$.
   - For given uplink CSI $\hat{\mathbf{H}}$, construct $\bar{\mathbf{H}}$.
   - Compute $\mathbf{R}$ using (5.15).
2. FOR $j = 1:1:K$
   - Initialize $\text{SINR}_{\Delta,j,\text{opt}} = -\infty$, $\bar{\mathbf{P}}_{\text{win},j} = \mathbf{0}$, $\bar{\mathbf{P}}_{\text{wall},j} = \mathbf{0}$, $\bar{\mathbf{H}}_j = \mathbf{0}$, $\bar{\mathbf{H}}_{\text{win},j} = \mathbf{0}$, $\bar{\mathbf{H}}_{\text{wall},j} = \mathbf{0}$, and $\mathbf{R}_j = \mathbf{0}$.
   - Set $\bar{\mathbf{H}}_j(:, j : K : \Delta_{\max})) = \bar{\mathbf{H}}((:, j : K : \Delta_{\max}))$
3. FOR $\Delta = 0:1:\Delta_{\max}$
   - $\mathrm{t} = (K\Delta) + (j : K : KN_b)$
   - $\mathbf{R}_j = \mathbf{R}(t,t)$.
   - Compute $\mathbf{b}_j$ using (5.18).
   - Construct $\tilde{\mathbf{b}}_j$ using (5.6).
   - Compute $\mathbf{w}_j$ using (5.14).
   - $\bar{\mathbf{P}}_{\text{win},j}(t,t) = \bar{\mathbf{P}}(t,t)$
   - $\bar{\mathbf{P}}_{\text{wall},j} = \bar{\mathbf{P}} - \mathbf{P}_{\text{win}}$
   - $\bar{\mathbf{H}}_{\text{win},j}(:,t) = \mathbf{H}(:,t)$
   - $\bar{\mathbf{H}}_{\text{wall},j} = \bar{\mathbf{H}} - \mathbf{H}_{\text{win}}$
   - Compute $\text{SINR}_{\Delta,j} = \frac{\mathbf{w}_j^H \mathbf{H}_{\text{win},j} \mathbf{P}_{\text{win},j} \mathbf{H}_{\text{win},j}^H \mathbf{w}_j}{\mathbf{w}_j^H (\mathbf{H}_{\text{wall},j} \mathbf{P}_{\text{wall},j} \mathbf{H}_{\text{wall},j}^H + \sigma_j^2 \mathbf{I}) \mathbf{w}_j}$
     - if $(\text{SINR}_{\Delta,j} > \text{SINR}_{\Delta,j,\text{opt}})$
       * $\mathbf{w}_{j,\text{opt}} = \mathbf{w}_j$
       * $\mathbf{b}_{j,\text{opt}} = \mathbf{b}_j$
       * $\tilde{\mathbf{b}}_{j,\text{opt}} = \tilde{\mathbf{b}}_j$
       * $\Delta_{j,\text{opt}} = \Delta$
       * $\text{SINR}_{\Delta,j,\text{opt}} = \text{SINR}_{\Delta,j}$
     - end if
4. END FOR Δ
5. END FOR j

This can also be written as,

$$\mathbf{h}_{\text{eff},j}^{\text{UL}} = \mathbf{w}_j^H \bar{\mathbf{H}}_j + \sum_{i=1, i\neq j}^{K} \mathbf{w}_j^H \bar{\mathbf{H}}_i$$

where $\bar{\mathbf{H}}_j$ defined in (5.20) is the channel convolutional matrix of the jth user augmented channel matrix $\tilde{\mathbf{H}}_j$ defined in (5.19).

$$\begin{aligned} \tilde{\mathbf{H}}_j &= [\mathbf{0}_{N_t\times(K-j)}\mathbf{h}_j(0)\mathbf{0}_{N_t\times(K-j)} \ \ldots \ \mathbf{0}_{N_t\times(K-j)}\mathbf{h}_j(N_h-1)\mathbf{0}_{N_t\times(K-j)}] \\ &= [\tilde{\mathbf{H}}_j(0) \ \tilde{\mathbf{H}}_j(1) \ \ldots \ \tilde{\mathbf{H}}_j(N_h-1)] \end{aligned} \tag{5.19}$$

$$\bar{\mathbf{H}}_j = \begin{bmatrix} \tilde{\mathbf{H}}_j(0) & \tilde{\mathbf{H}}_j(1) & \ldots & \tilde{\mathbf{H}}_j(N_h-1) & 0 & \ldots & 0 \\ 0 & \tilde{\mathbf{H}}_j(0) & \tilde{\mathbf{H}}_j(1) & \ldots & \tilde{\mathbf{H}}_j(N_h-1) & \ldots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \tilde{\mathbf{H}}_j(0) & \tilde{\mathbf{H}}_j(1) & \ldots & \tilde{\mathbf{H}}_j(N_h-1) \end{bmatrix}. \tag{5.20}$$

We write $\bar{\mathbf{H}}_j$ in terms of window $\bar{\mathbf{H}}_{\text{win},j}$ and wall part $\bar{\mathbf{H}}_{\text{wall},j}$ as follows [116],

$$\bar{\mathbf{H}}_j = \bar{\mathbf{H}}_{\text{win},j} + \bar{\mathbf{H}}_{\text{wall},j},$$

where

$$\bar{\mathbf{H}}_{\text{win},j} = [\mathbf{0}_{N_tN_F\times\Delta} \ \bar{\mathbf{H}}_j(:, \Delta+1:\Delta+N_b) \ \mathbf{0}_{N_tN_F\times K(\Delta_{\max}-\Delta)}]$$

is the window part and

$$\bar{\mathbf{H}}_{\text{wall},j} = \bar{\mathbf{H}}_j - \bar{\mathbf{H}}_{\text{win},j}$$

is the wall part. We therefore write the window and the wall part of the effective

channel as

$$\mathbf{h}_{\text{win},j}^{\text{UL}} = \mathbf{w}_j^H \bar{\mathbf{H}}_{\text{win},j}$$

$$\mathbf{h}_{\text{wall},j}^{\text{UL}} = \mathbf{w}_j^H \bar{\mathbf{H}}_{\text{wall},j} + \sum_{i=1, i\neq j}^{K} \mathbf{w}_j^H \bar{\mathbf{H}}_i.$$

We can rewrite the effective channel $\mathbf{h}_{\text{eff},j}^{\text{UL}}$ as

$$\begin{aligned} \mathbf{h}_{\text{eff},j}^{\text{UL}} &= \mathbf{h}_{\text{win},j}^{\text{UL}} + \mathbf{h}_{\text{wall},j}^{\text{UL}} \\ &= \mathbf{w}_j^H \bar{\mathbf{H}}_{\text{win},j} + \mathbf{w}_j^H \bar{\mathbf{H}}_{\text{wall},j} + \sum_{i=1, i\neq j}^{K} \mathbf{w}_j^H \bar{\mathbf{H}}_i. \end{aligned} \tag{5.21}$$

Similar to uplink, effective channel seen by the user in the downlink can be written as

$$\begin{aligned} \mathbf{h}_{\text{eff},j}^{\text{DL}} &= \mathbf{h}_{\text{win},j}^{\text{DL}} + \mathbf{h}_{\text{wall},j}^{\text{DL}} \\ &= \mathbf{h}_j^H \bar{\mathbf{W}}_{\text{win},j} + \mathbf{h}_j^H \bar{\mathbf{W}}_{\text{wall},j} + \sum_{i=1, i\neq j}^{K} \mathbf{H}_j^H \bar{\mathbf{W}}_i, \end{aligned} \tag{5.22}$$

where $\bar{\mathbf{W}}_j$ defined in (5.24) is the ST filter convolutional matrix of the jth user augmented ST filter $\tilde{\mathbf{W}}_j$ defined in (5.23) and

$$\bar{\mathbf{W}}_{\text{win},j} = [\mathbf{0}_{N_t N_F \times K\Delta} \ \bar{\mathbf{W}}_j(:, \Delta+1 : \Delta+N_b) \ \mathbf{0}_{N_t N_F \times K(\Delta_{\max}-\Delta)}]$$

is the window part and

$$\bar{\mathbf{W}}_{\text{wall},j} = \bar{\mathbf{W}}_j - \bar{\mathbf{W}}_{\text{win},j},$$

is the is the wall part.

$$\begin{aligned}\tilde{\mathbf{W}}_j &= [\mathbf{0}_{N_t\times(K-j)}\mathbf{w}_j(0)\mathbf{0}_{N_t\times(K-j)} \quad \dots \quad \mathbf{0}_{N_t\times(K-j)}\mathbf{w}_j(N_F-1)\mathbf{0}_{N_t\times(K-j)}] \\ &= [\tilde{\mathbf{W}}_j(0) \;\; \tilde{\mathbf{W}}_j(1) \;\; \dots \;\; \tilde{\mathbf{W}}_j(N_F-1)] \end{aligned} \tag{5.23}$$

$$\bar{\mathbf{W}}_j = \begin{bmatrix} \tilde{\mathbf{W}}_j(0) & \tilde{\mathbf{W}}_j(1) & \dots & \tilde{\mathbf{W}}_j(N_F-1) & 0 & \dots & 0 \\ 0 & \tilde{\mathbf{W}}_j(0) & \tilde{\mathbf{H}}_j(1) & \dots & \tilde{\mathbf{W}}_j(N_F-1) & \dots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \tilde{\mathbf{W}}_j(0) & \tilde{\mathbf{W}}_j(1) & \dots & \tilde{\mathbf{W}}_j(N_F-1) \end{bmatrix}. \tag{5.24}$$

It should be noted that,

$$\|\mathbf{h}_{\text{eff},j}^{\text{DL}}\|_2^2 = \|\mathbf{h}_j^H\bar{\mathbf{W}}\|_2^2 = \|\mathbf{h}_{\text{eff},j}^{\text{UL}}\|_2^2 = \|\mathbf{w}_j^H\bar{\mathbf{H}}\|_2^2. \tag{5.25}$$

5.5 Duality

Lemma 1: *The equality $\varepsilon_j^{UL} = \varepsilon_j^{DL}$ for $j = 1, 2, \dots, K$ holds if and only if, similar SINR targets $\gamma_1, \gamma_2, \dots, \gamma_K$ can be achieved in both the uplink, with power allocation vector* $\mathbf{p}$, *and in the downlink, with the power allocation vector* $\mathbf{q}$, *for fixed* $\hat{\mathbf{W}}$, $\boldsymbol{\beta}$ *and* $\mathbf{B}$.

Proof: Using the definitions from section 5.4, it can be shown that the MSE expressions in (5.13) and (5.7) can be written as

$$\varepsilon_j^{\mathrm{UL}} = \frac{\beta_j^2}{p_j}\left(\sum_{i=1}^{K} p_j \|\mathbf{w}_j^H \bar{\mathbf{H}}_i\|_2^2 + \sigma_j^2 \underbrace{\|\mathbf{w}_j\|_2^2}_{1}\right) - 2\|\mathbf{w}_j^H \bar{\mathbf{H}}_{\mathrm{win},j}\|_2^2 + 1 \tag{5.26}$$

$$\varepsilon_j^{\mathrm{DL}} = \frac{\beta_j^2}{q_j}\left(\sum_{i=1}^{K} q_j \|\mathbf{h}_j^H \bar{\mathbf{W}}_i\|_2^2 + \sigma_j^2\right) - 2\|\mathbf{h}_j^H \bar{\mathbf{W}}_{\mathrm{win},j}\|_2^2 + 1. \tag{5.27}$$

To prove Lemma 1 we show that $\varepsilon_j^{\mathrm{UL}} = \varepsilon_j^{\mathrm{DL}}$. From (5.26) and (5.27) we note only the first term in (5.26) differs from the first term in (5.27), as the second terms in (5.26) and (5.27) are equivalent from the relationship in (5.25) and the third and the final terms are also trivially equivalent. Hence, we are left to prove that the first terms in (5.26) and (5.27) equivalent. Using the definitions from section 5.4, we write the uplink and downlink SINR respectively in terms of the effective channel as follows:

$$\mathrm{SINR}_j^{\mathrm{UL}} = \frac{p_j \|\mathbf{w}_j^H \bar{\mathbf{H}}_{\mathrm{win},j}\|_2^2}{p_j \|\mathbf{w}_j^H \bar{\mathbf{H}}_{\mathrm{wall},j}\|_2^2 + \sum_{i\neq j} p_i \|\mathbf{w}_j^H \bar{\mathbf{H}}_i\|_2^2 + \sigma_j^2 \underbrace{\|\mathbf{w}_j\|_2^2}_{1}} \tag{5.28}$$

$$\mathrm{SINR}_j^{\mathrm{DL}} = \frac{q_j \|\mathbf{h}_j^H \bar{\mathbf{W}}_{\mathrm{win},j}\|_2^2}{q_j \|\mathbf{h}_j^H \bar{\mathbf{W}}_{\mathrm{wall},j}\|_2^2 + \sum_{i\neq j} q_i \|\mathbf{h}_j^H \bar{\mathbf{W}}_i\|_2^2 + \sigma_j^2}. \tag{5.29}$$

Assuming SINR targets achieved in the uplink with power allocation vector $\mathbf{p}$ are the same as the SINR targets achieved in the downlink with power allocation vector $\mathbf{q}$ for fixed $\mathbf{W}$, $\boldsymbol{\beta}$ and $\mathbf{B}$, we can state (5.28) and (5.29) are equal to each other:

$$\begin{aligned} &\frac{p_j \|\mathbf{w}_j^H \bar{\mathbf{H}}_{\mathrm{win},j}\|_2^2}{p_j \|\mathbf{w}_j^H \bar{\mathbf{H}}_{\mathrm{wall},j}\|_2^2 + \sum_{i\neq j} p_i \|\mathbf{w}_j^H \bar{\mathbf{H}}_i\|_2^2 + \sigma_j^2} \\ = &\frac{q_j \|\mathbf{h}_j^H \bar{\mathbf{W}}_{\mathrm{win},j}\|_2^2}{q_j \|\mathbf{h}_j^H \bar{\mathbf{W}}_{\mathrm{wall},j}\|_2^2 + \sum_{i\neq j} q_i \|\mathbf{h}_j^H \bar{\mathbf{W}}_i\|_2^2 + \sigma_j^2}. \end{aligned} \tag{5.30}$$

Multiplying the LHS and RHS of (5.30) by $\frac{\beta_j^2}{p_j}$ and $\frac{\beta_j^2}{q_j}$ respectively, we obtain

$$= \frac{\dfrac{\beta_j^2 \|\mathbf{w}_j^H \bar{\mathbf{H}}_{\text{win},j}\|_2^2}{\beta_j^2 \|\mathbf{w}_j^H \bar{\mathbf{H}}_{\text{wall},j}\|_2^2 + \frac{\beta_j^2}{p_j} \sum_{i \neq j} p_i \|\mathbf{w}_j^H \bar{\mathbf{H}}_i\|_2^2 + \frac{\beta_j^2}{p_j}\sigma_j^2}}{\dfrac{\beta_j^2 \|\mathbf{h}_j^H \bar{\mathbf{W}}_{\text{win},j}\|_2^2}{\beta_j^2 \|\mathbf{h}_j^H \bar{\mathbf{W}}_{\text{wall},j}\|_2^2 + \frac{\beta_j^2}{q_j} \sum_{i \neq j} q_i \|\mathbf{h}_j^H \bar{\mathbf{W}}_i\|_2^2 + \frac{\beta_j^2}{q_j}\sigma_j^2}}. \tag{5.31}$$

From (5.25) we note that the terms on the numerator of both LHS and RHS of (5.31) are equal. Hence to prove that the LHS and RHS satisfy the equality, we need to prove that the denominators of both LHS and RHS of (5.31) are equal to each other. Equating the denominators on both sides of (5.31), we obtain,

$$\begin{aligned} & \beta_j^2 \|\mathbf{w}_j^H \bar{\mathbf{H}}_{\text{wall},j}\|_2^2 + \frac{\beta_j^2}{p_j} \sum_{i \neq j} p_i \|\mathbf{w}_j^H \bar{\mathbf{H}}_i\|_2^2 + \frac{\beta_j^2}{p_j}\sigma_j^2 \\ = \; & \beta_j^2 \|\mathbf{h}_j^H \bar{\mathbf{W}}_{\text{wall},j}\|_2^2 + \frac{\beta_j^2}{q_j} \sum_{i \neq j} q_i \|\mathbf{h}_j^H \bar{\mathbf{W}}_i\|_2^2 + \frac{\beta_j^2}{q_j}\sigma_j^2. \end{aligned} \tag{5.32}$$

Using definitions

$$\|\mathbf{w}_j \bar{\mathbf{H}}_{\text{wall},j}\|_2^2 = \|\mathbf{w}_j \bar{\mathbf{H}}_j\|_2^2 - \|\mathbf{w}_j \bar{\mathbf{H}}_{\text{win},j}\|_2^2$$

and

$$\|\mathbf{h}_j \bar{\mathbf{W}}_{\text{wall},j}\|_2^2 = \|\mathbf{h}_j \bar{\mathbf{W}}_j\|_2^2 - \|\mathbf{h}_j \bar{\mathbf{W}}_{\text{win},j}\|_2^2,$$

we rewrite (5.32) as

$$\begin{aligned}
&\frac{\beta_j^2}{p_j}\sum_{i=1}^{K} p_i \|\mathbf{w}_j^H \bar{\mathbf{H}}_i\|_2^2 - \beta_j^2 \|\mathbf{w}_j^H \bar{\mathbf{H}}_{\text{win},j}\|_2^2 + \frac{\beta_j^2}{p_j}\sigma_j^2 \\
= \; &\frac{\beta_j^2}{q_j}\sum_{i=1}^{K} q_i \|\mathbf{h}_j^H \bar{\mathbf{W}}_i\|_2^2 - \beta_j^2 \|\mathbf{h}_j^H \bar{\mathbf{W}}_{\text{win},j}\|_2^2 + \frac{\beta_j^2}{q_j}\sigma_j^2.
\end{aligned} \tag{5.33}$$

From (5.25) we note that

$$\beta_j^2 \|\mathbf{w}_j^H \bar{\mathbf{H}}_{\text{win},j}\|_2^2 = \beta_j^2 \|\mathbf{h}_j^H \bar{\mathbf{W}}_{\text{win},j}\|_2^2$$

hence cancelling out these terms from both sides of (5.33) we obtain

$$\begin{aligned}
&\frac{\beta_j^2}{p_j}\left(\sum_{i=1}^{K} p_i \|\mathbf{w}_j^H \bar{\mathbf{H}}_i\|_2^2 + \sigma_j^2\right) \\
= \; &\frac{\beta_j^2}{q_j}\left(\sum_{i=1}^{K} q_i \|\mathbf{h}_j^H \bar{\mathbf{W}}_i\|_2^2 + \sigma_j^2\right)
\end{aligned} \tag{5.34}$$

Comparing (5.34) with the first terms in (5.26) and (5.27), we conclude that,

$$\varepsilon_j^{\text{UL}} = \varepsilon_j^{\text{DL}}. \tag{5.35}$$

□

We now use the uplink-downlink duality along with Lemma 1 to show that uplink and downlink have the same MSE achievable region.

Theorem 1. *With* $\hat{\mathbf{W}}$, $\boldsymbol{\beta}$, $\mathbf{B}$, *and total power* P_{max}, *both links have the same MSE achievable region under a total power constraint.*

Proof: We can collect the uplink SINR values in (5.28) for $j = 1, 2, \ldots K$ in a vector as in (5.36). Equation (5.36) can be rearranged to (5.37). Let us define

matrices and Υ^{UL} and Γ^{UL} as in (5.38), (5.39) and a vector $\boldsymbol{\sigma} = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_K^2]^T$.

$$\begin{bmatrix} \gamma_1^{\mathrm{UL}} \\ \gamma_2^{\mathrm{UL}} \\ \vdots \\ \gamma_K^{\mathrm{UL}} \end{bmatrix} = \begin{bmatrix} \frac{p_1 \|\mathbf{w}_1^H \bar{\mathbf{H}}_{\mathrm{win},1}\|_2^2}{p_1 \|\mathbf{w}_1^H \bar{\mathbf{H}}_{\mathrm{wall},1}\|_2^2 + \sum_{i \neq 1} p_i \|\mathbf{w}_1^H \bar{\mathbf{H}}_i\|_2^2 + \sigma_1^2} \\ \frac{p_2 \|\mathbf{w}_2^H \bar{\mathbf{H}}_{\mathrm{win},2}\|_2^2}{p_2 \|\mathbf{w}_2^H \bar{\mathbf{H}}_{\mathrm{wall},2}\|_2^2 + \sum_{i \neq 2} p_i \|\mathbf{w}_2^H \bar{\mathbf{H}}_i\|_2^2 + \sigma_2^2} \\ \vdots \\ \frac{p_K \|\mathbf{w}_K^H \bar{\mathbf{H}}_{\mathrm{win},K}\|_2^2}{p_K \|\mathbf{w}_K^H \bar{\mathbf{H}}_{\mathrm{wall},K}\|_2^2 + \sum_{i \neq K} p_i \|\mathbf{w}_K^H \bar{\mathbf{H}}_i\|_2^2 + \sigma_K^2} \end{bmatrix} \tag{5.36}$$

$$\begin{bmatrix} \frac{p_1 \|\mathbf{w}_1^H \bar{\mathbf{H}}_{\mathrm{win},1}\|_2^2}{\gamma_1^{\mathrm{UL}}} \\ \frac{p_2 \|\mathbf{w}_2^H \bar{\mathbf{H}}_{\mathrm{win},2}\|_2^2}{\gamma_2^{\mathrm{UL}}} \\ \vdots \\ \frac{p_K \|\mathbf{w}_K^H \bar{\mathbf{H}}_{\mathrm{win},K}\|_2^2}{\gamma_K^{\mathrm{UL}}} \end{bmatrix} = \begin{bmatrix} p_1 \|\mathbf{w}_1^H \bar{\mathbf{H}}_{\mathrm{wall},1}\|_2^2 + \sum_{i \neq 1} p_i \|\mathbf{w}_1^H \bar{\mathbf{H}}_i\|_2^2 + \sigma_1^2 \\ p_2 \|\mathbf{w}_2^H \bar{\mathbf{H}}_{\mathrm{wall},2}\|_2^2 + \sum_{i \neq 2} p_i \|\mathbf{w}_2^H \bar{\mathbf{H}}_i\|_2^2 + \sigma_2^2 \\ \vdots \\ p_K \|\mathbf{w}_K^H \bar{\mathbf{H}}_{\mathrm{wall},K}\|_2^2 + \sum_{i \neq K} p_i \|\mathbf{w}_K^H \bar{\mathbf{H}}_i\|_2^2 + \sigma_K^2 \end{bmatrix} \tag{5.37}$$

$$\Upsilon^{\mathrm{UL}} = \begin{bmatrix} \|\mathbf{w}_1^H \bar{\mathbf{H}}_{\mathrm{wall},1}\|_2^2 & \|\mathbf{w}_2^H \bar{\mathbf{H}}_1\|_2^2 & \ldots & \|\mathbf{w}_K^H \bar{\mathbf{H}}_1\|_2^2 \\ \|\mathbf{w}_1^H \bar{\mathbf{H}}_2\|_2^2 & \|\mathbf{w}_2^H \bar{\mathbf{H}}_{\mathrm{wall},2}\|_2^2 & \ldots & \|\mathbf{w}_K^H \bar{\mathbf{H}}_2\|_2^2 \\ \vdots & \vdots & \ldots & \vdots \\ \|\mathbf{w}_1^H \bar{\mathbf{H}}_K\|_2^2 & \|\mathbf{w}_2^H \bar{\mathbf{H}}_K\|_2^2 & \ldots & \|\mathbf{w}_K^H \bar{\mathbf{H}}_{\mathrm{wall},K}\|_2^2 \end{bmatrix} \tag{5.38}$$

$$\Gamma^{\mathrm{UL}} = \mathrm{diag}\left\{ \left[\frac{\|\mathbf{w}_1^H \bar{\mathbf{H}}_{\mathrm{win},1}\|_2^2}{\gamma_1}, \frac{\|\mathbf{w}_2^H \bar{\mathbf{H}}_{\mathrm{win},2}\|_2^2}{\gamma_2}, \ldots, \frac{\|\mathbf{w}_K^H \bar{\mathbf{H}}_{\mathrm{win},K}\|_2^2}{\gamma_K} \right] \right\} \tag{5.39}$$

Using (5.38), (5.39) and $\boldsymbol{\sigma}$, we can write (5.37) as

$$\Gamma^{\mathrm{UL}} \mathbf{p} = \Upsilon^{\mathrm{UL}} \mathbf{p} + \boldsymbol{\sigma}. \tag{5.40}$$

Solving (5.40) for the uplink power allocation $\mathbf{p}$, we obtain

$$\mathbf{p} = \left(\Gamma^{\mathrm{UL}} - \Upsilon^{\mathrm{UL}} \right)^{-1} \boldsymbol{\sigma} \tag{5.41}$$

Similarly one can characterize the downlink power allocation to achieve the same set of SINR targets as

$$\mathbf{q} = \left(\Gamma^{\mathrm{DL}} - \Upsilon^{\mathrm{DL}}\right)^{-1} \boldsymbol{\sigma} \tag{5.42}$$

where, Υ^{DL} and Γ^{DL} are defined in (5.43) and (5.44) respectively. Using the relationship in (5.25), we can rewrite (5.42) as

$$\Upsilon^{\mathrm{DL}} = \begin{bmatrix} \|\mathbf{h}_1^H \bar{\mathbf{W}}_{\mathrm{wall},1}\|_2^2 & \|\mathbf{h}_2^H \bar{\mathbf{W}}_1\|_2^2 & \dots & \|\mathbf{h}_K^H \bar{\mathbf{W}}_1\|_2^2 \\ \|\mathbf{h}_1^H \bar{\mathbf{W}}_2\|_2^2 & \|\mathbf{h}_2^H \bar{\mathbf{W}}_{\mathrm{wall},2}\|_2^2 & \dots & \|\mathbf{h}_K^H \bar{\mathbf{W}}_2\|_2^2 \\ \vdots & \vdots & \dots & \vdots \\ \|\mathbf{h}_1^H \bar{\mathbf{W}}_K\|_2^2 & \|\mathbf{h}_2^H \bar{\mathbf{W}}_K\|_2^2 & \dots & \|\mathbf{h}_K^H \bar{\mathbf{W}}_{\mathrm{wall},K}\|_2^2 \end{bmatrix} \tag{5.43}$$

$$\Gamma^{\mathrm{DL}} = \mathrm{diag}\left\{ \left[\frac{\|\mathbf{h}_1^H \bar{\mathbf{W}}_{\mathrm{win},1}\|_2^2}{\gamma_1} \frac{\|\mathbf{h}_2^H \bar{\mathbf{W}}_{\mathrm{win},2}\|_2^2}{\gamma_2} \dots \frac{\|\mathbf{h}_K^H \bar{\mathbf{W}}_{\mathrm{win},K}\|_2^2}{\gamma_K} \right] \right\} \tag{5.44}$$

$$\mathbf{q} = \left(\Gamma^{\mathrm{UL}} - (\Upsilon^{\mathrm{UL}})^T\right)^{-1} \boldsymbol{\sigma} \tag{5.45}$$

Now to complete the proof, we will show that the total transmit powers given by $\|\mathbf{p}\|_1$ and $\|\mathbf{q}\|_1$ are identical

$$\begin{aligned} \|\mathbf{p}\|_1 &= \mathbf{1}_K^T \mathbf{p} = \mathbf{1}_K^T \left(\left(\Gamma^{\mathrm{UL}} - \Upsilon^{\mathrm{UL}}\right)^{-1} \right)^T \boldsymbol{\sigma} \\ &= \mathbf{1}_K^T \left(\left(\Gamma^{\mathrm{UL}} - (\Upsilon^{\mathrm{UL}})^T\right)^{-1} \boldsymbol{\sigma} = \mathbf{1}_K^T \mathbf{q} = \|\mathbf{q}\|_1, \end{aligned} \tag{5.46}$$

where power allocation $\mathbf{q}$ achieves the same SINR targets $\gamma_1, \gamma_2, \dots, \gamma_K$ in the downlink with the same total power i.e. $\|\mathbf{q}\|_1 \leq P_{\max}$. Hence, we have shown that

with the same $\hat{\mathbf{W}}$, $\mathbf{B}$, $\boldsymbol{\beta}$ and appropriate power allocation $\mathbf{p}$, we attain specified MSE targets $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_K$ in the uplink. These targets can also be achieved in the downlink with the same $\hat{\mathbf{W}}$, $\mathbf{B}$, $\boldsymbol{\beta}$ and power allocation $\mathbf{q}$ with the total power constraint $\|\mathbf{p}\|_1 = \|\mathbf{q}\|_1 \leq P_{\max}$.

□

5.6 Algorithms

According to the UDD results, both the uplink and the downlink share the same normalized MSE regions for a given total power constraint. This effectively means that we can obtain solutions to the downlink optimization problems C1 and C2, by optimizing the MSE of the equivalent uplink problems.

5.6.1 C1: Max-Min Fairness

Firstly we consider problem based on C1 in the downlink. The immediate consequence of Lemma 1 and Theorem 1 is that the solution to C1 can be obtained by solving the virtual uplink problem which can be written as

$$\begin{aligned} \min_{\hat{\mathbf{W}},\boldsymbol{\beta},\mathbf{p},\mathbf{B}} \quad & \max_{1\leq j\leq K} \varepsilon_j^{\mathrm{UL}} \\ \text{s.t.} \quad & \|\mathbf{p}\|_1 \leq P_{\max}. \end{aligned} \tag{5.47}$$

For each power allocation the optimum MMSE based ST filters have the structure as in (5.14). The MMSE achieved with such a STCS filter is related to the maximum SINR as follows

$$\varepsilon_j^{\mathrm{UL,min}} = \frac{1}{1 + \mathrm{SINR}_j^{\mathrm{UL,max}}}. \tag{5.48}$$

Thus, instead of minimizing the maximum MSE, we can equivalently maximize the minimum SINR. We therefore write (5.47) as

$$\begin{aligned} \max_{\hat{\mathbf{W}},\boldsymbol{\beta},\mathbf{p},\mathbf{B}} \quad & \min_{1\leq j\leq K} \mathrm{SINR}_j^{\mathrm{UL}} \\ \text{s.t.} \quad & \|\mathbf{p}\|_1 \leq P_{\max}. \end{aligned} \tag{5.49}$$

To solve this we follow the work in [34] and [121]. Letting $\mathbf{p}_{opt}$ be the global maximizer of optimization problem (5.49), and assuming fixed $\hat{\mathbf{W}}$, $\boldsymbol{\beta}$, $\mathbf{B}$ while optimizing (5.49) over $\mathbf{p}$, $\mathbf{p}_{opt}$ can be characterized as

$$\mathbf{p}_{ext} = \mathcal{P}_{max}\{\Phi\} \tag{5.50}$$

where $\mathcal{P}_{max}\{\cdot\}$ returns the eigenvector corresponding to the maximum eigenvalue of a matrix, $\mathbf{p}_{ext} = [\mathbf{p}_{opt} \quad 1]^T$ is an extended power allocation vector and Φ is an extended coupling matrix defined as

$$\Phi^{\mathrm{UL}} = \begin{bmatrix} (\Gamma^{\mathrm{UL}})^{-1}\Upsilon^{\mathrm{UL}} & (\Gamma^{\mathrm{UL}})^{-1}\boldsymbol{\sigma} \\ \frac{1}{P_{\max}}\mathbf{1}^T(\Gamma^{\mathrm{UL}})^{-1}\Psi^{\mathrm{UL}} & \frac{1}{P_{\max}}\mathbf{1}^T(\Gamma^{\mathrm{UL}})^{-1}\boldsymbol{\sigma} \end{bmatrix}. \tag{5.51}$$

$\mathbf{p}_{opt}$ is constructed from $\mathbf{p}_{ext}$ such that $[\mathbf{p}_{ext}]_{K+1} = 1$. See [34] and [121] for a formal proof. We now obtain the optimum $\hat{\mathbf{W}}$, $\mathbf{B}$ and $\boldsymbol{\beta}$ for the optimization problem in (5.49) for a given power allocation vector $\mathbf{p}$ using Algorithm 3. We also obtain the optimum solution to (5.49) for fixed $\hat{\mathbf{W}}$, $\mathbf{B}$ and $\boldsymbol{\beta}$ from (5.50). The immediate consequence of this allows us to optimize one variable while keeping other fixed in an iterative manner. It can be shown that C1 is strictly monotonically increasing in $P_{\max}$ and converges to the global optimum, (see [34] for proof). Problem C1 is commonly known as the fairness problem due to the identical SINR or MMSE achieved by all users at the optimum solution, as shown in the simulation results. The solution to

the optimization problem (5.49) is presented in Algorithm 4.

5.6.2 C2: Min Power

Similarly to C1, C2 can be written into a virtual uplink problem as

$$\begin{aligned} \min_{\hat{\mathbf{W}},\boldsymbol{\beta},\mathbf{p},\mathbf{B}} \quad & \|\mathbf{p}\|_1 \\ \text{s.t.} \quad & \varepsilon_j^{\text{UL}} \leq \xi_j^{\text{UL}} \\ & j = 1, \ldots, K. \end{aligned} \tag{5.52}$$

It follows that using (5.48) we replace the MSE constraints with SINR constraints. Hence (5.52) can be written as

$$\begin{aligned} \min_{\hat{\mathbf{W}},\boldsymbol{\beta},\mathbf{p},\mathbf{B}} \quad & \|\mathbf{p}\|_1 \\ \text{s.t.} \quad & \text{SINR}_j^{\text{UL}} \leq \gamma_j^{\text{UL}} \\ & j = 1, \ldots, K, \end{aligned} \tag{5.53}$$

where γ_j^{UL} are the SINR targets required by j^{th} user. Problem (5.53) is related to the problem (5.49) since setting $P_{\max} = P_{\min,\text{opt}}$, where $P_{\min,\text{opt}} = \|\mathbf{p}\|_1$ is the optimum solution of (5.53), in (5.49) would give the same solution as (5.53). Additional degree of freedom to minimize the total transmission power is achieved if $\frac{1}{\lambda_{\max}(\Phi)} > 1$, where $\lambda_{\max}\{\cdot\}$ returns the maximum eigenvalue of a matrix as shown in [34]. Hence Algorithm 4 can be adopted, starting off with the same iterations as in Algorithm 4. If $\frac{1}{\lambda_{\max}(\Phi)} < 1$, as iterations goes to infinity, the problem is infeasible and initial conditions must be relaxed e.g. dropping some users [34]. Once $\frac{1}{\lambda_{\max}(\Phi)} \geq 1$ is satisfied, the global solution to (5.53) can be found by changing the power control policy in

the subsequent iterations, i.e. we allocate power such that the SINR constraints are satisfied with equality. This power allocation as shown previously, is given by (5.41). It can be shown that C2 is strictly monotonically increasing in $\|\mathbf{p}\|_1$ and is shown to converge to the global optimum (see [34] for proof). Full algorithm to solve (5.53) is presented in Algorithm 5.

Algorithm 4 Op1: Max-Min Fairness

1. Initialize $n = 0$, $P_{\max}$, $\mathbf{p}_0$, η which controls the required accuracy and compute the ST filters $\hat{\mathbf{W}}_0$, TIR $\mathbf{B}_0$ and β_0 using Algorithm 3.

   Uplink Channel

2. REPEAT
   - n = n + 1
   - For given $\hat{\mathbf{W}}_{n-1}$, $\mathbf{B}_{n-1}$, β_{n-1} find $\mathbf{p}_n$ by solving (5.50).
   - Update the uplink ST filters $\hat{\mathbf{W}}_n$, TIR $\mathbf{B}_n$ and β_n for power allocation $\mathbf{p}_n$ using Algorithm 3.
   - Compute the uplink SINR values for the given $\mathbf{p}_n$, $\hat{\mathbf{W}}_n$, $\mathbf{B}_n$ and β_n.
3. UNTIL $\{\max_{1\leq j\leq K}\text{SINR}_j - \min_{1\leq j\leq K}\text{SINR}_j\} \leq \eta$.

   Downlink Channel

   - Compute the downlink power allocation $\mathbf{q}$ using (5.45), which should achieve the same SINR targets as in uplink with the same total power $\|\mathbf{q}\|_1 = P_{\max} = \|\mathbf{p}_n\|_1$.

5.7 Simulation Results

A system comprising of a transmitter with $N_t = 4$ transmitting antennae and $K = 3$ single antenna users is considered. We generate independent, unity power, transmit symbols which are scaled according to the power requirements prior to transmission. The frequency selective channel is fixed for each data block and is assumed to have an impulse response of length $N_h = 12$. It is however changed between blocks according

Algorithm 5 Op2: Min Power

1. Initialize $n = 0$, power allocation vector $\mathbf{p}_0$, $P_{\min,0}$, η which controls the required accuracy and compute the ST filters $\mathbf{W}_0$, TIR $\mathbf{B}_0$ and $\boldsymbol{\beta}_0$ using Algorithm 3.

   Uplink Channel

2. REPEAT
   - Assuming the problem is feasible, i.e. $\frac{1}{\lambda_{\max}(\Phi)} > 1$.
   - $n = n + 1$
   - For given ST filters $\hat{\mathbf{W}}_{n-1}$, TIR $\mathbf{B}_{n-1}$ and $\boldsymbol{\beta}_{n-1}$, find $\mathbf{p}_n$ by solving (5.41).
   - Update $P_{\min,n} = \|\mathbf{p}_n\|_1$.
   - Update the ST filters $\hat{\mathbf{W}}_n$, TIR $\mathbf{B}_n$ and $\boldsymbol{\beta}_n$ for power allocation vector $\mathbf{p}_n$ using Algorithm 3.
   - Compute the uplink SINR values for the given $\mathbf{p}_n$, $\mathbf{W}_n$, $\mathbf{B}_n$ and $\boldsymbol{\beta}_n$.
3. UNTIL $\{P_{\min,n} - P_{\min,n-1}\} \leq \eta$.

   Downlink Channel

   - Compute the downlink power allocation $\mathbf{q}$ using (5.45), which should achieve the same SINR targets as in uplink with the same total power $\|\mathbf{q}\|_1 = P_{\min,n} = \|\mathbf{p}_n\|_1$.

to a zero mean complex Gaussian distribution. The noise is zero mean i.i.d complex circularly symmetric AWGN with variance $\sigma_n^2 = 0.01$. For a fair comparison between the performance of the channel shortening based spatial multiplexing scheme with that of an equalization based scheme, the lengths of the equalizer and the channel shortening filter are assumed to be the same, $N_F = (N_h - 1)K = 33$.

Fig. 5.5 depicts the performance of both the channel shortening and the full equalization based multiplexing schemes in terms of achievable balanced SINR targets computed using (5.29). Note that the balanced SINR targets are plotted, i.e. these are the SINR values achieved by all users in the system. The results indicate that the channel shortening for a given total power achieves a higher balanced SINR level as compared to full equalization.

For clarity in Fig. 5.6, we have plotted the equivalent balanced MSE targets achieved by all users in the system. The MSE values have been computed using (5.7). Similarly as before, we note that STCS filters achieve lower balanced MSE targets as compared to the STEQ filters. It should also be noted from Fig. 5.5 and Fig. 5.6 that the relationship used to transform the MSE optimization problem into a SINR optimization based problem, (stated in (5.48)), holds.

We now perform a simulation for C2, where we wish to minimize the total transmit power in order to attain some pre-defined SINR targets. We assumed that all users have the same SINR targets. Fig. 5.7 depict performance in terms of the required transmitter power for both channel shortening and full equalization at the transmitter. The results in Fig. 5.7 indicate that full equalization at the transmitter requires relatively more total power as compared to channel shortening based scheme for various TIR lengths, in order to achieve identical SINR targets. Results of both problems indicate that the channel shortening based spatial multiplexing scheme has the ability to provide a superior performance. This is because the channel shortening filter has an extra degree of freedom to suppress IUI, by relaxing the requirement on ISI.

We carry out further simulation for C1 to compare the BER performance in a CDMA based system. We considered a DS-CDMA scheme with spreading factor 32. All three users use the same spreading code to study the spatial multiplexing performance. For channel shortening based multiplexer, we use Rake receivers for each user to coherently combine the remaining unequalized paths. The BER results depicted in Fig. 5.8 confirm the SINR and the MSE advantages seen in Fig. 5.5 and Fig. 5.6 respectively, i.e. the channel shortening based spatial multiplexing scheme attains relatively better BER performance than a complete equalization based multiplexing scheme.

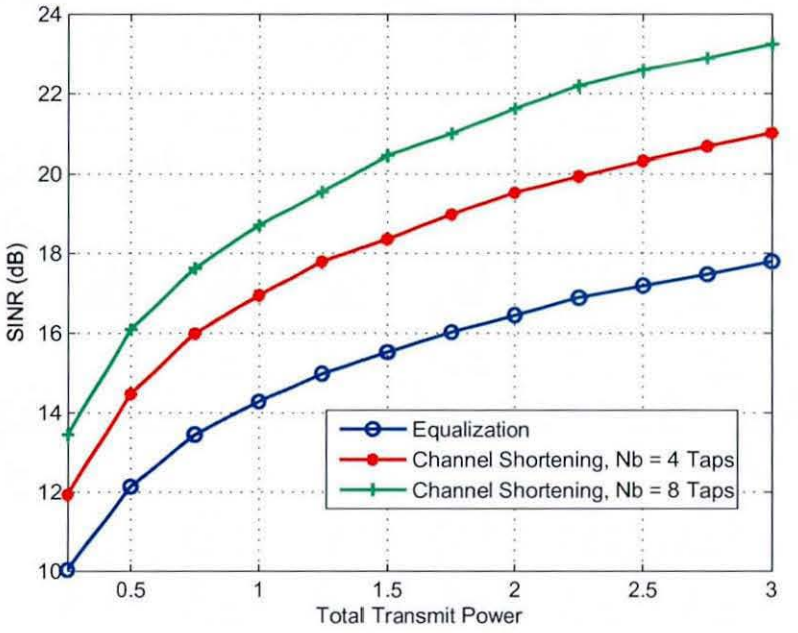

Figure 5.5: Result for P1, balanced SINR targets within the available total power.

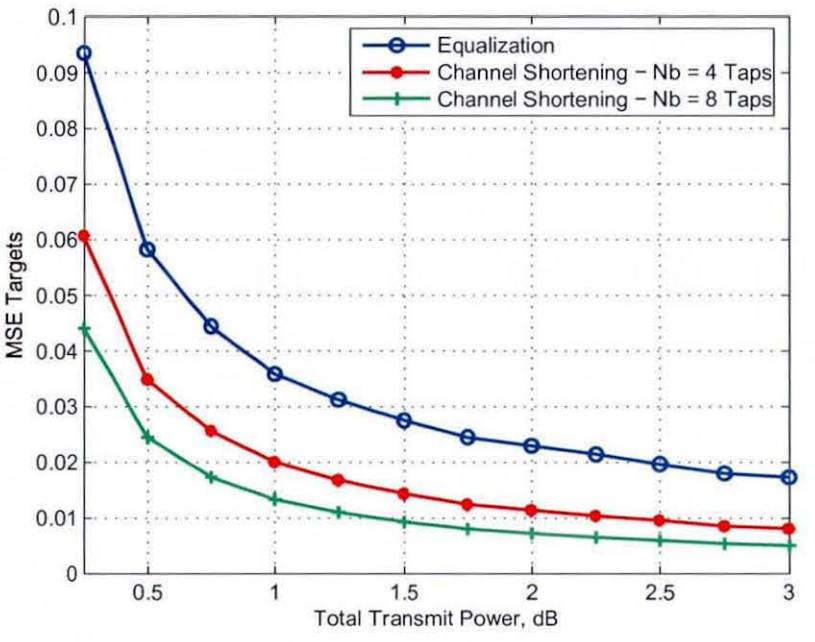

Figure 5.6: Result for P1, balanced MSE targets within the available total power.

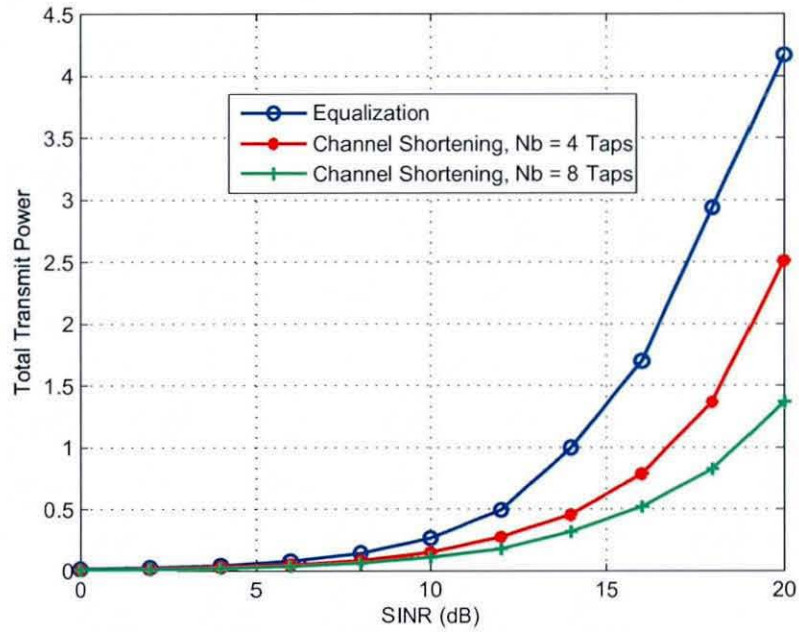

Figure 5.7: Result for P2, minimum total power required to achieve the given SINR targets.

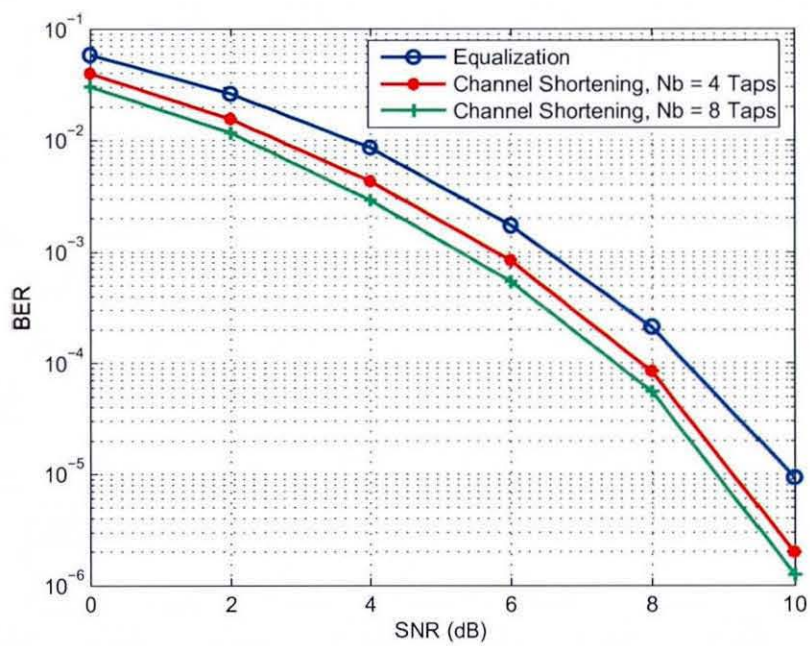

Figure 5.8: Result for P1, BER vs SNR comparison.

5.8 Summary

Spatial multiplexing schemes based on channel shortening and full equalization at the transmitter has been proposed. The schemes are based on the results of UDD and solves the problem of multiuser downlink beamforming but over frequency selective channels. Algorithmic solutions for two problems with different optimization criterion have been proposed. Firstly, we proposed solutions to the problem of satisfying the QoS constraints of all users with a total power constraint, this as we saw from the simulation results, balances the SINR or the MSE targets of all users. The solution to this problem allocates all the available power to achieve the QoS targets. Thus, secondly, we proposed solutions to the problem of minimizing the total transmit power subject to satisfying the QoS constraints of all users. We showed that for both optimization problems the channel shortening based spatial multiplexing provides better performance than full equalization based schemes. This is due to the fact that channel shortening based design is able to exploit the degree of freedom it has in terms of suppressing the ISI, whereas the equalization based design has a strict requirement of canceling the ISI completely.

Chapter 6

Robust Downlink Beamforming based on Maximizing Signal-to-Leakage Ratio

Extensive research has been conducted on multiple-input multiple-output (MIMO) systems due to their potential for providing high capacity, increased diversity and mitigating interference in multi-user (MU) scenarios. Conventional techniques focus on the receiver for mitigating distortions such as channel impairment and interference. However recent interests have been shifted for optimizing the transmitters in an attempt to keep the receiver complexity at low. The transmitter diversity can also be exploited to form multiuser multiplexing. In this chapter, we focus on a transmitter optimization technique based on spatial diversity in a downlink wireless communication system [59], where a basestation (BS) could simultaneously serve multiple users without compromising the available radio spectrum. To achieve this the BS pre-compensates for the interference allowing users in the cell to maximize their signal power and to reduce interference. The BS can also perform beamforming to suppress multiuser interference (MUI) to end users and to maximize overall system capacity.

Several transmit beamforming techniques have been proposed in the recent literature [1,2,34,58,83,108–110], most of them however, require nearly perfect knowledge of the channel at the transmitter. But due to imperfections, the channel state information (CSI) available at the transmitter will always be somewhat in error. These imperfections mainly arise due to time variations of the channels, feedback delay, quantization of CSI etc. The performance degrades substantially due to these imperfections. Hence, the motivation here is to design techniques which will incorporate such imperfections and still provide good performance. Some good examples are the recent advances in robust beamforming techniques [6,74–76,122]. Most of these techniques model the uncertainties as an unknown parameter which is bounded by a known norm based on some a-priori knowledge. The problem is then generally converted into a constrained optimization program and solved for worst case performance optimization both analytically and numerically. One such example is proposed in [74] for general rank beamforming, where the received signal and noise plus interference covariance matrices are assumed to be in error.

Here we will build on the recent results on the robust capon beamformer [122] and design a robust multiuser beamformer which will be less sensitive to the expected CSI errors. A signal to leakage ratio (SLR) based optimization criterion is adopted [108] to design the robust multiuser beamformer. This optimization criterion is chosen because it provides a closed form solution as opposed to the iterative solutions obtained for signal-to-interference and noise ratio (SINR) criterion [34]. However similar robust techniques could also be applied to SINR based methods. In the SLR approach the transmit weight vector for the ith user is determined by maximizing the transmit power to the ith user while minimizing the interference (leakage) caused to all other users. Here, we will extend this model by incorporating imperfect CSI by explicitly modelling the uncertainties and adopt a worst-case performance optimization criterion as proposed in [74] to design the proposed robust transmitter.

To model the CSI errors, we consider the errors are randomly generated according to a Gaussian distribution. Later we also discuss two practical examples, where errors in the CSI may arise due to imperfections. In the first example, the transmitter estimates the CSI using the available feedback information based on Bayesian estimation theory. Such an estimate is sensitive to the feedback delay and hence introduces error which is an increasing function of the feedback delay, see as an example, Fig. 6.1. We therefore apply the proposed technique to demonstrate robustness against such errors and perform a simulation under this general setup to show a significant improvement in performance over the non-robust method. In the latter example, we consider an OFDM based MU-MIMO system (see Fig. 4.5 for the baseband representation of an OFDM system, which transforms a frequency selective channel into N parallel flat fading channels), where multi-user multiplexing is required in each frequency bin. It requires feedback of equivalent MIMO CSI in each frequency for all users. This could result into excessive feedback overhead. One way to reduce this feedback is to exploit the correlation of feedback information in adjacent frequency bins. For example frequency bins could be divided into a number of groups and the average (mean) CSI for each group can be fed back. This has been illustrated in Fig. 6.2. This will however result into inaccurate CSI for each frequency bin, hence could reduce the performance significantly. We therefore propose to use robust MIMO multi-user beamforming technique to mitigate effect of CSI error.

6.1 System Model

Consider a downlink MU-MIMO system consisting of one BS with N_T transmit antennas communicating with K users, each having N_{R_i} receive antennas. A block diagram is shown in Fig. 1.2, where $\mathbf{s} = [s_1(t), \ldots, s_K(t)]$ denote the transmitted signal vector whose elements are assumed to be independent and identically distributed

with unity variance, i.e. $\mathrm{E}\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}$. The signal vector $\mathbf{s}(t)$ is then multiplied by a normalized beamformer matrix $\mathbf{W} = [\mathbf{w}_1(t), \ldots, \mathbf{w}_K(t)]$ with $\|\mathbf{w}_k\|^2 = 1$, before being transmitted over an multiuser channel. Hence, the $N_T \times 1$ transmitted signal vector at time t is given by

$$\mathbf{x}(t) = \sum_{k=1}^{K} \mathbf{w}_k s_k(t) = \mathbf{W}(t)\mathbf{s}(t) \tag{6.1}$$

The signal vector $\mathbf{x}(t)$ is then transmitted over an multiuser channel. Assuming that the channel is frequency non-selective, the received signal vector $\mathbf{y}_i(t)$ for the ith user at time t, can be written as

$$\mathbf{y}_i(t) = \mathbf{H}_i(t)\mathbf{x}(t) + \mathbf{n}_i(t) \tag{6.2}$$

where $\mathbf{n}(t)$ is an additive white Gaussian noise (AWGN) vector. The matrix $\mathbf{H}_i(t)$ is assumed to be block fading. Assuming the ith user employs N_{R_i} antennas, the $N_{R_i} \times N_T$ channel matrix can be written as

$$\mathbf{H}_i = \begin{bmatrix} h_i^{(1,1)} & \ldots & h_i^{(1,N)} \\ \vdots & \ddots & \vdots \\ h_i^{(N_{R_i},1)} & \ldots & h_i^{(N_{R_i},N_T)} \end{bmatrix} \tag{6.3}$$

where, $h_i^{p,m}$ denote the channel coefficient between the mth transmit and p^{th} receive antennas, for user i. Here, we assume that the receiver for user i has access to accurate CSI, $\mathbf{H}_i$. However, given the reasons mentioned previously, we will allow for imperfect CSI at the transmitter.

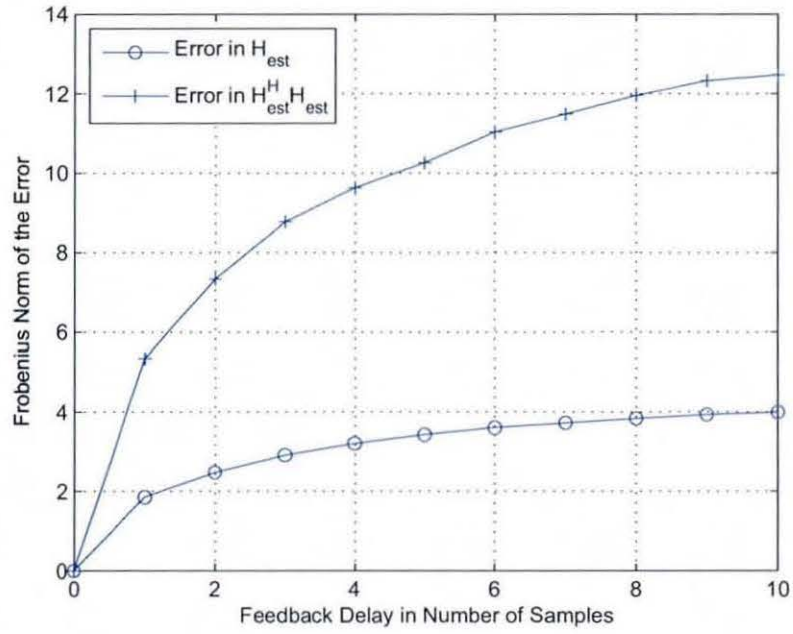

Figure 6.1: Frobenius norm of the error introduced in the CSI due to feedback delay.

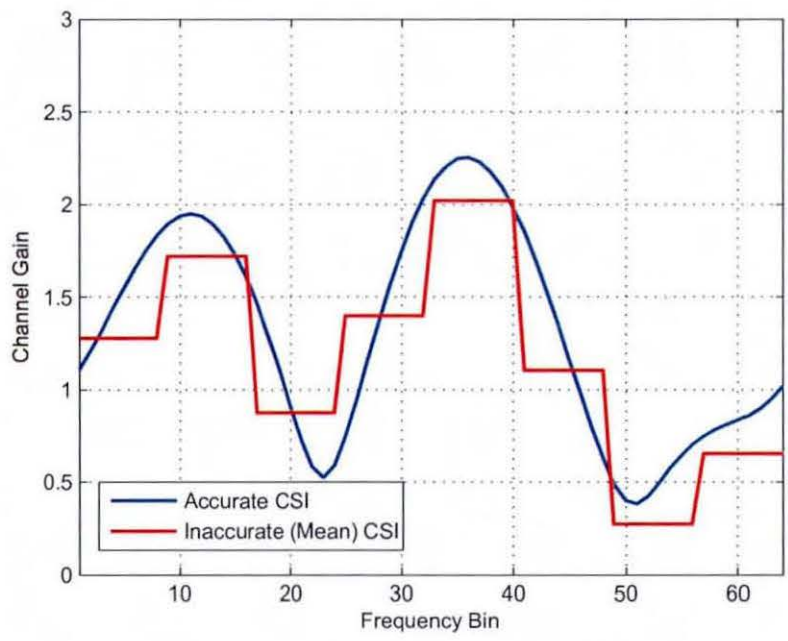

Figure 6.2: Accurate and inaccurate frequency response of a frequency selective channel.

6.2 Algorithms

For notational simplicity, let us drop the time index t and proceed to rewrite (6.2) as

$$\mathbf{y}_i = \mathbf{H}_i\mathbf{w}_i s_i + \sum_{k=1,k\neq i}^{K} \mathbf{H}_i\mathbf{w}_k s_k + \mathbf{n}_i, \tag{6.4}$$

where the second term quantifies the interference caused to user i from all other users. The aim is to mitigate this interference for all users. The power of the desired signal in (6.4) is given by $\|\mathbf{H}_i\mathbf{w}_i\|^2$. Similarly, the interference caused by the ith user to the kth user is given by $\|\mathbf{H}_k\mathbf{w}_i\|^2$. The total power leaked from this user to all other users, the so called leakage for user i, is defined as [108]

$$\sum_{k=1,k\neq i}^{K} \|\mathbf{H}_k\mathbf{w}_i\|^2 \tag{6.5}$$

6.2.1 Non-Robust Design

Given a fixed transmit power for each user, the weight vectors $\mathbf{w}_i$, $i = 1, 2, \ldots, K$, are designed such that the signal-to-leakage noise ratio (SLR) is maximized for every user [108]

$$\mathbf{w}_i^{opt} = \arg\max_{w_i} \underbrace{\frac{\|\mathbf{H}_i\mathbf{w}_i\|^2}{\sum_{k=1,k\neq i}^{K} \|\mathbf{H}_k\mathbf{w}_i\|^2}}_{\text{SLR for user } i} \quad \text{s.t.} \quad \|\mathbf{w}_i\|^2 = 1, \forall i \tag{6.6}$$

The solution to the above equation is given by the Rayleigh-Ritz quotient result [111]

$$\mathbf{w}_i^{opt} = \mathcal{P}_{\max}\left\{(\tilde{\mathbf{H}}_i^H\tilde{\mathbf{H}}_i)^{-1}(\mathbf{H}_i^H\mathbf{H}_i)\right\}, \tag{6.7}$$

where $\mathcal{P}_{\max}\{\cdot\}$ returns the principal eigenvector of the matrix, that is the eigenvector corresponding to its maximal eigenvalue. Also $\tilde{\mathbf{H}}_i = [\mathbf{H}_i^H ... \mathbf{H}_{i-1}^H \mathbf{H}_{i+1}^H ... \mathbf{H}_K^H]^H$

is an extended channel matrix that excludes $\mathbf{H}_i$ only.

6.2.2 Robust Design

Now, let us proceed to consider the case where the CSI available at the transmitter is imperfect. Then, we can write the CSI available at the transmitter as

$$\mathbf{H}_{i_p} = \mathbf{H}_{i_a} + \mathbf{\Delta}_{i_H}, \tag{6.8}$$

where the presumed MIMO channel matrix is denoted by $\mathbf{H}_{i_p}$ and the actual channel matrix is denoted by $\mathbf{H}_{i_a}$. Here, $\mathbf{\Delta}_i$ is the unknown matrix mismatch. These mismatches may occur due to quantization errors, erroneous feedback, feedback delay and variations of the channel. In simulation results we will consider the case where these mismatches arises due to feedback delay. For simplicity let us define $\mathbf{R}_i = \mathbf{H}_i^H \mathbf{H}_i$ and $\tilde{\mathbf{R}}_i = \tilde{\mathbf{H}}_i^H \tilde{\mathbf{H}}_i$. Hence, using (6.8) we can write,

$$\mathbf{R}_{i_p} = \mathbf{R}_{i_a} + \mathbf{\Delta}_1, \quad \text{and} \quad \tilde{\mathbf{R}}_{i_p} = \tilde{\mathbf{R}}_{i_a} + \mathbf{\Delta}_2 \tag{6.9}$$

where the presumed matrices are denoted by $\mathbf{R}_{i_p}$ and $\tilde{\mathbf{R}}_{i_p}$ respectively and the actual matrices are denoted by $\mathbf{R}_{i_a}$ and $\hat{\mathbf{R}}_{i_a}$. Here, $\mathbf{\Delta}_1$ and $\mathbf{\Delta}_2$ are the unknown matrix mismatches. In the presence of these mismatches the output SLR can be written as

$$\mathrm{SLR}_i = \frac{\mathbf{w}_i^H \mathbf{R}_{i_p} \mathbf{w}_i}{\mathbf{w}_i^H \tilde{\mathbf{R}}_{i_p} \mathbf{w}_i} \tag{6.10}$$

Let us assume that the mismatch matrices $\mathbf{\Delta}_1$ and $\mathbf{\Delta}_2$ are bounded in their norm by some constants as

$$\|\mathbf{\Delta}_1\|_F^2 \leq \epsilon_1, \quad \text{and} \quad \|\mathbf{\Delta}_2\|_F^2 \leq \epsilon_2 \tag{6.11}$$

where ϵ_1 and ϵ_2 represent the radius of the assumed uncertainty sphere. To provide robustness to such norm-bounded mismatches, we use the result of [74]. The beamformer weight vector is obtained by maximizing the worst-case output SLR. This corresponds to the following optimization problem

$$\max_{\mathbf{w}_i} \min_{\Delta_1,\Delta_2} \frac{\mathbf{w}_i^H(\mathbf{R}_{i_a}+\mathbf{\Delta}_1)\mathbf{w}_i}{\mathbf{w}_i^H(\tilde{\mathbf{R}}_{i_a}+\mathbf{\Delta}_2)\mathbf{w}_i} \tag{6.12}$$

This problem can be written as

$$\max_{\mathbf{w}_i} \frac{\min_{\|\Delta_1\|_F\leq\epsilon_1} \mathbf{w}_i^H(\mathbf{R}_{i_a}+\mathbf{\Delta}_1)\mathbf{w}_i}{\max_{\|\Delta_2\|_F\leq\epsilon_2} \mathbf{w}_i^H(\tilde{\mathbf{R}}_{i_a}+\mathbf{\Delta}_2)\mathbf{w}_i} \tag{6.13}$$

To solve (6.13), we note that [74]

$$\min_{\|\Delta_1\|_F\leq\epsilon_1} \mathbf{w}_i^H(\mathbf{R}_{i_a}+\mathbf{\Delta}_1)\mathbf{w}_i = \mathbf{w}_i^H(\mathbf{R}_{i_a}-\epsilon_1\mathbf{I})\mathbf{w}_i \tag{6.14}$$

$$\max_{\|\Delta_2\|_F\leq\epsilon_2} \mathbf{w}_i^H(\tilde{\mathbf{R}}_{i_a}+\mathbf{\Delta}_2)\mathbf{w}_i = \mathbf{w}_i^H(\tilde{\mathbf{R}}_{i_a}+\epsilon_2\mathbf{I})\mathbf{w}_i \tag{6.15}$$

where the worst-case mismatch matrices $\mathbf{\Delta}_1$ and $\mathbf{\Delta}_2$ are given by

$$\mathbf{\Delta}_1 = -\epsilon_1\frac{\mathbf{w}_i\mathbf{w}_i^H}{\|\mathbf{w}_i\|^2}, \quad \text{and} \quad \mathbf{\Delta}_2 = \epsilon_2\frac{\mathbf{w}_i\mathbf{w}_i^H}{\|\mathbf{w}_i\|^2} \tag{6.16}$$

Therefore, the optimization problem is reduced to

$$\max_{\mathbf{w}_i} \frac{\mathbf{w}_i^H(\mathbf{R}_{i_a}-\epsilon_1\mathbf{I})\mathbf{w}_i}{\mathbf{w}_i^H(\tilde{\mathbf{R}}_{i_a}+\epsilon_2\mathbf{I})\mathbf{w}_i} \tag{6.17}$$

Note the error bound ϵ_1 has to be smaller than the maximal eigenvalue of $\mathbf{R}_{i_a}$ [74]. Therefore, the parameter ϵ which is smaller than the maximal eigenvalue of

$\mathbf{R}_{i_a}$ has to be chosen. A simple interpretation of this condition is that the allowed uncertainty in the signal covariance matrix should be sufficiently small. Clearly the structure of the problem now is similar to that of the problem before. Using this fact, the solution can be expressed in the following form

$$\mathbf{w}_i^{\text{rob}} = \mathcal{P}_{\max}\left\{(\tilde{\mathbf{R}}_{i_a} + \epsilon_2\mathbf{I})^{-1}(\mathbf{R}_{i_a} - \epsilon_1\mathbf{I})\right\} \tag{6.18}$$

6.2.3 Diagonal Loading

In order to solve (6.18) and choose the values of parameters ϵ_1 and ϵ_2 we analyze the statistics of ϵ_1 and ϵ_2. We assumed circularly symmetric white Gaussian noise components for the elements of the MIMO channel $\mathbf{H}_i$ and the uncertainty matrix $\mathbf{\Delta}_i$. To do this, we first derive the expressions for calculating the expected value of these norms. From (6.9) we see that

$$\begin{aligned} \mathbf{\Delta}_1 &= \mathbf{R}_{i_a} - \mathbf{R}_{i_p} & \mathbf{\Delta}_2 &= \tilde{\mathbf{R}}_{i_a} - \tilde{\mathbf{R}}_{i_p} \\ \|\mathbf{\Delta}_1\|_F^2 &= \text{tr}\{\mathbf{\Delta}_1^H\mathbf{\Delta}_1\} & \|\mathbf{\Delta}_2\|_F^2 &= \text{tr}\{\mathbf{\Delta}_2^H\mathbf{\Delta}_2\} \end{aligned} \tag{6.19}$$

We may simplify the expressions for Frobenius norm of $\|\mathbf{\Delta}_1\|_F^2$ and $\|\mathbf{\Delta}_2\|_F^2$ in (6.19) as shown below.

For a MIMO channel matrix of size $N_{R_i} \times N_T$, the matrix $\Delta_i^H\Delta_i$ $(i = 1, 2)$ will be of dimension $N_T \times N_T$. We note that the expected values of all the diagonal elements of $\Delta_i^H\Delta_i$ are equal and also the expected values of all the non-diagonal elements are also the same. It can then be shown that the expected values of the diagonal and the non-diagonal elements are given by (6.20) and (6.21) respectively, as shown overleaf, where D refers to the diagonal elements and ND refers to the non-diagonal elements. Hence the Frobenius norm of Δ_i is given by (6.22). Expressions in equation (6.23) and (6.24), directly follow from this result, where N_T is the number of transmitting antennas, N_{R_i} is the number of receiving antennas, $\hat{N}_{R_i} = N_{R_i}(K-1)$ and K is

the total number of users in the system. Furthermore, σ_H^2 and σ_Δ^2 are the variance of elements of channel matrix $\mathbf{H}$ and the misadjustment matrix $\mathbf{\Delta}$ respectively. In the simulation, we examine the effect of ϵ_1 and ϵ_2 by choosing various factors of $\bar{\epsilon_1} = \mathrm{E}\|\Delta_1\|^2$ and $\bar{\epsilon_2} = \mathrm{E}\|\Delta_2\|^2$ for the error bound ϵ_1 and ϵ_2.

$$\mathrm{E}\{\|\Delta_i\|^2\}_{\mathrm{D}} = \mathrm{E}\left\{ \|\sum_{j=1}^{N_{R_i}} \Delta_j^H h_j h_j^H \Delta_j + \Delta_j^H \Delta_j\|^2 \right\} \tag{6.20}$$

$$\mathrm{E}\{\|\Delta_i\|^2\}_{\mathrm{ND}} = \mathrm{E}\left\{ \|\sum_{j=1}^{N_{R_i}} h_j^H h_{j+1} + h_j^H \Delta_{j+1} + \Delta_j^H \Delta_{j+1}\|^2 \right\} \tag{6.21}$$

$$\mathrm{E}\{\|\Delta_i\|_F^2\} = (N)\mathrm{E}\{\Delta_i^H \Delta_i\}_{\mathrm{D}} + (N^2 - N\ )\mathrm{E}\{\|\Delta_i\|^2\}_{\mathrm{ND}} \tag{6.22}$$

$$\begin{aligned} \mathrm{E}\left\{\|\mathbf{\Delta}_1\|_F^2\right\} &= N_T\left\{2N_{R_i}\sigma_H^2\sigma_\Delta^2 + 2N_{R_i}\sigma_\Delta^4 + \{N_{R_i}^2 - N_{R_i}\}\sigma_\Delta^4\right\} \\ &\quad + \{N_T^2 - N_T\}\left\{2N_{R_i}\sigma_H^2\sigma_\Delta^2 + N_{R_i}\sigma_\Delta^4\right\} \end{aligned} \tag{6.23}$$

$$\begin{aligned} \mathrm{E}\left\{\|\mathbf{\Delta}_2\|_F^2\right\} &= N_T\left\{2\hat{N_{R_i}}\sigma_H^2\sigma_\Delta^2 + 2\hat{N_{R_i}}\sigma_\Delta^4 + \{\hat{N_{R_i}}^2 - \hat{N_{R_i}}\}\sigma_\Delta^4\right\} \\ &\quad + \{N_T^2 - N_T\}K\left\{2\hat{N_{R_i}}\sigma_H^2\sigma_\Delta^2 + \hat{N_{R_i}}\sigma_\Delta^4\right\} \end{aligned} \tag{6.24}$$

6.3 Simulation Results

We considered a MU-MIMO system with one BS equipped with $N_T = 6$ antennas and $K = 3$ users each equipped with $N_{R_i} = 3$ antennas. The data symbols are generated using QPSK modulation. The total transmitted power per symbol period across all transmit antennas is normalized to unity.

The entries of channel matrix $\mathbf{H}$ and the mismatch matrix $\mathbf{\Delta}$, are zero mean Gaussian random variables with variances $\sigma_H^2 = 1$ and $\sigma_\Delta^2 = 0.005$ respectively. The Rayleigh fading channel coefficients are generated independently for each transmission symbol. The noise is zero mean and spatially and temporally uncorrelated, i.e.

$$\mathrm{E}\{\mathbf{v}_i\mathbf{v}_i^H\} = \sigma_i^2\mathbf{I}_{N_{R_i}}, \quad \text{and} \quad \mathrm{E}\{\mathrm{tr}(\mathbf{H}_i\mathbf{H}_i^H)\} = N_{R_i}N_T$$

We choose the expected values of $\bar{\epsilon_1} = 1.0412$ and $\bar{\epsilon_2} = 1.4734$ obtained from (6.23) and (6.24) for ϵ_1 and ϵ_2 in (6.18) to calculate the robust multiuser beamforming vector. The rational behind this is examined later. Figure (6.3) depicts the difference between non-robust SLR and the proposed robust algorithm. The result shows the average BER performance of all the users using both the robust and the non-robust multi user beamforming algorithm. It can be seen that robust algorithm provides a gain of 4dB over the non-robust algorithm at a BER of 10^{-2}. Table 6.1 shows the gain in performance for different values of ϵ_1 and ϵ_2 as compared with the non-robust algorithm. We note from the previous section that increasing the values of ϵ could result in a negative definite matrix, $(\mathbf{R}_a - \epsilon_1\mathbf{I})$ in (6.18). Therefore ϵ_1 has be less than the largest eigenvalue of the matrix $\mathbf{R}_a$, as explained in the previous section as well as in [74]. Hence, it is very important to choose the value of ϵ_1 and ϵ_2 appropriately. We performed a set of simulations for various values of ϵ_1 and ϵ_2 as a factor of their expected values and tabulated the performance gain in Table 6.1. It appears from table 6.1 that choosing the error bounds ϵ_1 and ϵ_2 as their expected values $\bar{\epsilon_1}$ and $\bar{\epsilon_2}$ provides a satisfactory result. In practice, it may be possible to obtain these expected values using a priori knowledge of the channel variations, feedback delay and quantization errors.

In the next section we discuss two practical application of the proposed robust beamformers.

Table 6.1: Gain (dB) of the robust algorithm over the non-robust algorithm to achieve a BER of 10^{-2}.

Case	$0.0\bar{\epsilon_1}$	$0.2\bar{\epsilon_1}$	$0.4\bar{\epsilon_1}$	$0.6\bar{\epsilon_1}$	$0.8\bar{\epsilon_1}$	$1.0\bar{\epsilon_1}$
$0.0\bar{\epsilon_2}$	0	0	0.2	0.2	0.2	0
$0.2\bar{\epsilon_2}$	2	2	2	2	3	2.5
$0.4\bar{\epsilon_2}$	2	3	3	3	4	3
$0.6\bar{\epsilon_2}$	3.5	3.5	3	3	4	4
$0.8\bar{\epsilon_2}$	3.5	4	4	4	4	4
$1.0\bar{\epsilon_2}$	3.5	4	4	4	4	4

6.3.1 Example I - Mean Feedback in MIMO Systems

In this example we assume that the transmitter only has access to imperfect channel feedback, see [123] for details on partial feedback. The channel $\mathbf{H}(t)$ is assumed to be Rayleigh fading and is generated using the following AR(1) random process with a forgetting factor ρ [123],

$$\mathbf{H}(t) = \rho\mathbf{H}(t-1) + \hat{\rho}\mathbf{W}(t) \tag{6.25}$$

where, parameter ρ is obtained from the channel profile as follows,

$$\rho = J_0\left(\frac{2\pi T v_{mb}}{\lambda}\right) \tag{6.26}$$

where $J_0(.)$ denotes the zeroth-order bessel function of the first kind, and parameters T, v_{mb}, and λ denote the duration of each data frame (or the interval between two consecutive feedback), mobile speed and carrier wavelength respectively. The entries of $\mathbf{W}(t)$ are assumed to independent and identically distributed (i.i.d) circularly symmetric Gaussian, each with standard deviation $\hat{\rho} = \sqrt{(1-\rho^2)}$. Assuming that the transmitter observes $\hat{\mathbf{H}}(t) = \mathbf{H}(t-d)$ at the output of the feedback channel, conditioned on $\hat{\mathbf{H}}(t)$, the distribution of $\mathbf{H}(t)$ can be obtained as $\mathbf{H} \sim \mathcal{N}(\mu, \alpha\mathbf{I})$ [124],

where,

$$\begin{aligned}
\mu = E(\mathbf{H}|\hat{\mathbf{H}}) &= E(\mathbf{H}) + \frac{cov(\hat{\mathbf{H}}, \mathbf{H})}{var(\hat{\mathbf{H}})}(\hat{\mathbf{H}} - E(\hat{\mathbf{H}})) \\
&= \rho^d \mathbf{H}(t-d) \qquad (6.27) \\
\alpha = var(\mathbf{H}|\hat{\mathbf{H}}) &= var(\mathbf{H}) - \frac{cov^2(\hat{\mathbf{H}}, \mathbf{H})}{var(\hat{\mathbf{H}})} \\
&= (1 - \rho^{2d}) \qquad (6.28)
\end{aligned}$$

Hence the transmitter uses $\bar{\mathbf{H}}(t) = \alpha^d \mathbf{H}(t-d)$ as an estimate of $\mathbf{H}(t)$. The error between the actual CSI and its estimate is then given by:

$$\begin{aligned}
\mathrm{E}(t) &= \mathbf{H}(t) - \rho^d \mathbf{H}(t-d) \\
&= \rho^d \mathbf{H}(t) + \sum_{i=0}^{d-1} \rho^i \mathbf{W}(t-i) - \rho^d \mathbf{H}(t-d) \\
&= \sum_{i=0}^{d-1} \rho^i \mathbf{W}(t-i) \qquad (6.29)
\end{aligned}$$

where, the distribution of $\mathrm{E}(t)$ is given by $\mathrm{E} \sim \mathcal{N}(0, (1-\rho^{2d})\mathbf{I}))$.

Fig. 6.4 shows the BER performance of both robust and non-robust schemes for the mean feedback scenario. Here a feedback delay of one data frame is assumed i.e. $d = T = 1$ and $\rho = 0.99$. Similarly as before the robust beamformer outperforms the non-robust scheme. However in this scenario, the diagonal loading parameters have been simply chosen as 5% of $\|\mathbf{R}_{i_a}\|_2$ and $\|\tilde{\mathbf{R}}_{i_a}\|_2$ for ϵ_1 and ϵ_2 respectively.

6.3.2 Example II - Feedback in MIMO-OFDM

In this example, we investigate the application of the robust beamformers to MU-MIMO OFDM systems. In this case multiuser multiplexing is required to be performed for each sub-carriers at the transmitter. Ideally this would require CSI feedback for each sub-carrier. Since this will result into excessive feedback overhead, an

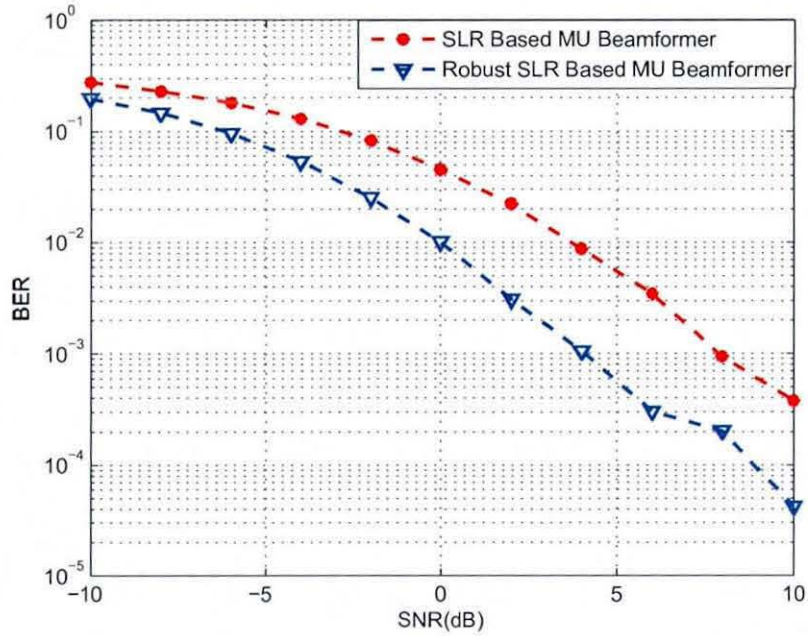

Figure 6.3: The average BER performance for all the users is plotted as a function of the SNR for a MU-MIMO system with $N_T = 6$ transmit antennas and $K = 3$ users, each equipped with $N_{R_i} = 3$ receive antennas.

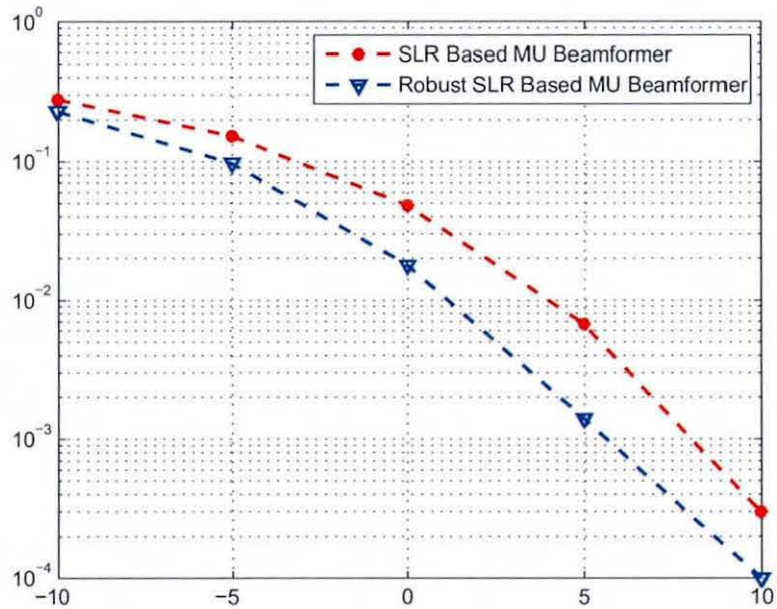

Figure 6.4: The average BER performance for all the users for the case of mean feedback is plotted as a function of the SNR for a MU-MIMO system with $N_T = 6$ transmit antennas and $K = 3$ users, each equipped with $N_{R_i} = 3$ receive antennas.

Table 6.2: Experimental values of $\bar{\epsilon}_1$ and $\bar{\epsilon}_2$ used in the simulations.

N_c	N_h	N_B	$\bar{\epsilon}_1$	$\bar{\epsilon}_2$	N_c	N_h	N_B	$\bar{\epsilon}_1$	$\bar{\epsilon}_2$
64	3	4	1.32	1.88	128	3	8	1.38	1.96
64	3	8	2.82	4.07	128	4	8	1.91	2.70
64	3	16	5.85	8.72	128	5	8	2.46	3.58
64	3	32	9.64	14.66	128	6	8	3.04	4.45

attractive solution is to feedback CSI for a group of adjacent sub-carriers instead of feeding back CSI in each sub-carrier. Hence we propose to feedback the mean CSI, obtained by averaging the CSI over a number of adjacent sub-carriers (here we refer to them as a block). Hence the difference between the mean CSI and the CSI of the individual sub-carrier should be considered as error in the available CSI at the transmitter. This error will increase as the block size N_B or the channel length N_h is increased. We therefore apply our robust techniques to this scenario and analyze the performance for various block lengths N_B and channel lengths N_h.

Fig. 6.5, depicts the performance of both robust and non-robust algorithms based on the mean CSI feedback for various block lengths. We considered $N_c = 64$, $N_h = 3$, and $N_B = 4, 8, 16$ and 32. We note that the BER performance degrades as the block length is increased from 4 to 32 but the robust algorithm is able to outperform the non-robust algorithm in all cases.

Similarly in Fig. 6.6, we fixed the block length N_B to 8, but analyzed the performance for various channel lengths $N_h = 3$ to 6. Again we see the BER performance degrades as the channel length increases but the robust algorithm is able to outperform the non robust-algorithm.

Table 6.2 shows the values for ϵ_1 and ϵ_2 used in the above simulations. The values refer to the Frobenius norm of the mismatch matrices in (6.11). We chose these values based on the computation of mean ϵ_1 and ϵ_2 as discussed before.

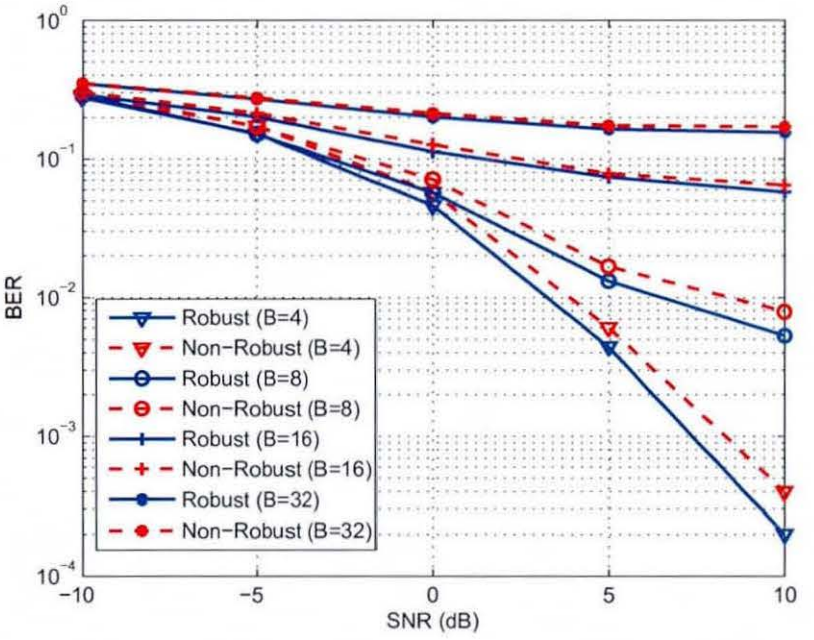

Figure 6.5: The average BER performance for all the users is plotted as a function of the SNR for a MU-MIMO OFDM system with $N_T = 6$ transmit antennas and $K = 3$ users, each equipped with $N_{R_i} = 3$ receive antennas. Here the multipath channel length N_h is fixed to 3, but the block size N_B is changed from 4 to 32.

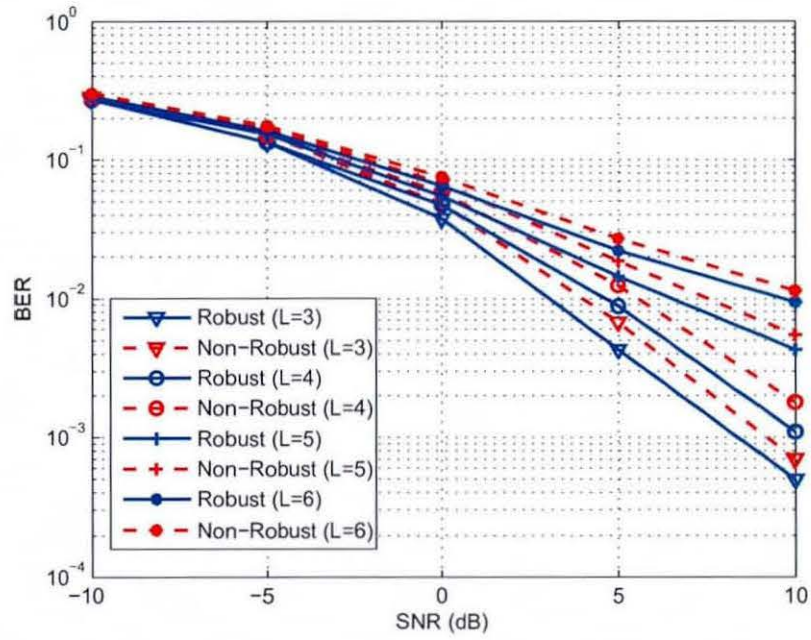

Figure 6.6: The average BER performance for all the users is plotted as a function of the SNR for a MU-MIMO OFDM system with $N_T = 6$ transmit antennas and $K = 3$ users, each equipped with $N_{R_i} = 3$ receive antennas. Here the block size N_B is fixed to 8, but the channel length N_h is varied from 3 to 6.

6.4 Summary

A robust transmit beamforming scheme has been proposed for a MU-MIMO system. The proposed scheme incorporates for the errors in the CSI feedback from the receiver to the transmitter. The misadjustments have been modelled by Gaussian noise components and the Frobenius norm of the error matrix has been assumed to be bounded above by a known parameter based on some a-priori knowledge. Errors in the CSI due to feedback delay have been considered. Simulation results generated for a Rayleigh fading environment confirm that the proposed algorithm outperform the non-robust algorithm over a wide range of SNR. We also demonstrated importance of using this robust technique in a couple of practical scenarios. In the first example we investigated the effect of feedback delay which resulted into a considerable amount of errors in the CSI available at the transmitter. In the second example, we applied the robust technique in a MIMO-OFDM scenario where limitation on feedback information could result into considerable amount of errors in the CSI.

Chapter 7

Robust Downlink Beamforming with Positive Semidefinite Covariance Constraints

Throughput of a multiuser wireless system is limited and dictated by inter-user interference (IUI) and system resources such as bandwidth and transmit power. Spatial multiplexing based strategies significantly improve spectral efficiency of a multiuser wireless system by exploiting spatial characteristics of the propagation channel. In particular, the downlink beamforming based spatial multiplexing techniques have proven to be effective to mitigate IUI while minimizing the transmitter power, thus improving the spectral efficiency of a multiuser wireless system.

In this chapter, firstly we consider the problem of joint multiuser downlink beamforming and power control in a single-cell environment, where independent data streams are transmitted from a single multi-antenna BS to several decentralized single antenna receivers, under the assumption that only an erroneous channel state information (CSI) is available at the transmitters. Later, we extend the proposed schemes along with simulation results to the problem of joint multiuser downlink beamform-

ing, power control and BS assignment in a multi-cell environment. In this scheme independent data streams are transmitted from multiple BSs to several decentralized single antenna receivers. We also assume the CSI available at the BSs is in error.

Several downlink beamforming methods within the context of single BS and perfect channel state information (CSI) have been developed in [34, 78, 80–82, 125, 126]. The problem of joint optimal beamforming and power control, for achieving a specific set of target SINRs while minimizing the total transmitted power, has been studied in [34, 78].

The scheme in [127] assumed perfect CSI at the transmitter. However in practical situations, perfect CSI may not be available and the performance of the transmitter beamforming techniques of [34, 78, 80–82, 125, 126] could severely degrade as the quality of the available CSI reduces. Typically, the estimates of CSI are made available to the transmitter through feedback channels from the receivers. The CSI estimates are normally in error due to quantization, feedback delay, channel dynamics and estimation errors. Therefore, robust techniques are required to mitigate the effect of CSI errors. Such robust transmit beamformer design based on worst-case performance optimization has been proposed in [77] and [79]. In these designs, the Frobenius norm of mismatches between the presumed and the actual downlink channel covariance matrices have been assumed to be bounded above by a known constant. The beamformer performance is then optimized for the worst-case mismatch. However, the worst-case performance optimization-based transmit beamformers of [77] and [79] ignore the positive semidefinite (PSD) constraints on the mismatched downlink channel covariance matrices. As a result, these techniques can be overly conservative in practical scenarios.

As opposed to worst-case performance optimization, another popular approach is to use a probabilistic constraint based performance optimization. In [106], the problem of downlink beamforming has been solved using a probabilistic constraint

based optimization. In this approach, the worst-case constraints are replaced by more flexible probabilistic constraints i.e. the constraints are satisfied with a specific probability. The statistics of the mismatches between the presumed and the actual downlink channel covariance matrices are assumed to be known. Unfortunately the probabilistic based approach adopted in [106] violates the PSD constraints on the mismatched downlink channel covariance matrices. The proposed worst-case performance optimization based scheme in this scheme could also be extended to a probabilistic constraint based performance optimization with semidefinite constraints.

In this chapter, we propose two novel robust transmit beamforming and BS assignment techniques within the framework of worst-case designs, but preserving the aforementioned PSD constraints on the mismatched downlink channel covariance matrices. The resulting design corresponds to a non-convex optimization problem but can be approximated by a convex SDP problem using semi-definite relaxation (SDR). The complexity of the design based on the first approach is comparable to that of the worst-case design of [77]. Our second technique has a slightly higher complexity than the worst-case design of [77] as it resorts to an algorithm involving several SDP iterations. However the proposed second technique offers a superior performance than the first technique. Simulation results demonstrate that for an imperfect CSI scenario, both of our proposed methods outperform the conventional robust transmit beamformer of [77].

7.1 Problem Formulation

Let us consider a system with a single BS transmitter with N_T antennas and K decentralized single-antenna receivers as shown in Fig.1.2. The BS transmits a vector $\mathbf{x}(t) \in \mathbb{R}^{N_T \times 1}$ at time t as

$$\mathbf{x}(t) = \sum_{i=1}^{K} s_i(t)\mathbf{w}_i$$

where $\mathbf{w}_i$ $(i = 1, \ldots, K)$ are the $N_T \times 1$ complex beamforming vectors (to be determined) and $s_i(t)$ is the data symbol intended for the ith user. For simplicity, we assume that all $s_i(t)$ are uncorrelated and have the same normalized power $\mathrm{E}\{|s_i(t)|^2\} = 1$. The vector $\mathbf{x}(t)$ is then transmitted over a flat-fading quasi-static channel. The received signal at the ith user is then given by

$$y_i(t) = \mathbf{h}_i \mathbf{x}(t) + n_i(t)$$

where $\mathbf{h}_i$ is the $1 \times N_T$ downlink channel vector for the ith user and $n_i(t)$ zero-mean circularly symmetric complex Gaussian (ZMCSCG) noise with variance σ_i^2. The SINR of the ith user is given by

$$\mathrm{SINR}_i = \frac{\mathbf{w}_i^H \mathbf{R}_i \mathbf{w}_i}{\sum_{j=1, j\neq i}^{K} \mathbf{w}_j^H \mathbf{R}_i \mathbf{w}_j + \sigma_i^2} \tag{7.1}$$

where

$$\mathbf{R}_i = \mathrm{E}\{\mathbf{h}_i^H \mathbf{h}_i\} \in \mathbb{S}_+^{N_T \times N_T}$$

is the downlink channel covariance matrix for the ith user. A meaningful problem formulation is to minimize the total transmit power subject to user QoS constraints [80]. This optimization problem can be written as [78], [80]

$$\begin{aligned} &\min_{\mathbf{w}_i} \sum_{i=1}^{K} \mathbf{w}_i^H \mathbf{w}_i \\ &\text{s.t.} \quad \frac{\mathbf{w}_i^H \mathbf{R}_i \mathbf{w}_i}{\sum_{j=1, j\neq i}^{K} \mathbf{w}_j^H \mathbf{R}_i \mathbf{w}_j + \sigma_i^2} \geq \gamma_i, \qquad i = 1, \ldots, K \end{aligned} \tag{7.2}$$

or equivalently,

$$\min_{\mathbf{w}_i} \sum_{i=1}^{K} \mathbf{w}_i^H \mathbf{w}_i$$
$$\text{s.t.} \quad \mathbf{w}_i^H \mathbf{R}_i \mathbf{w}_i - \gamma_i \sum_{j=1, j\neq i}^{K} \mathbf{w}_j^H \mathbf{R}_i \mathbf{w}_j - \gamma_i \sigma_i^2 \geq 0, \qquad i = 1, \ldots, K \tag{7.3}$$

where γ_i is the minimal acceptable SINR for the ith user. This is a quadratic optimization problem with quadratic non-convex constraints. Such problems could be NP-complete in general, which makes it much difficult to solve directly. However, it has been shown in [77], that this particular problem inherits a structure that makes it possible to find the global optimum efficiently. In [78], an approach based on semidefinite programming which is able to attain the optimum solution efficiently has been proposed. Using a change of variable $\mathbf{W} = \mathbf{w}\mathbf{w}^H$ and using the property of the trace operator $\text{Tr}\{\mathbf{AB}\} = \text{Tr}\{\mathbf{BA}\}$ [128], it follows that $\mathbf{w}^H\mathbf{R}\mathbf{w} = \text{Tr}\{\mathbf{RW}\}$. Thus, introducing a change of variables, the problem in (7.3) can be written as,

$$\min_{\mathbf{W}_i} \sum_{i=1}^{K} \text{Tr}\{\mathbf{W}_i\}$$
$$\text{s.t.} \quad \text{Tr}\{\mathbf{R}_i \mathbf{W}_i\} - \gamma_i \sum_{j=1, j\neq i}^{K} \text{Tr}\{\mathbf{R}_i \mathbf{W}_j\} - \gamma_i \sigma_i^2 \geq 0, \tag{7.4}$$
$$\mathbf{W}_i \succeq 0, \quad \text{rank}\{\mathbf{W}_i\} = 1, \qquad i = 1, \ldots, K,$$

where additional constraint $\mathbf{W}_i \succeq 0$ is required to ensure the positive semidefiniteness of the optimization variable and $\text{rank}\{\mathbf{W}_i\} = 1$ is required to ensure that the problem (7.4) is equivalent to the problem in (7.2). Thus, if the optimal solution to (7.4) has $\text{rank}\{\mathbf{W}_i\} = 1, i = 1, 2, \ldots, K$, then it is also the optimal solution to the original problem (7.2). However, the constraint, $\text{rank}\{\mathbf{W}_i = 1\}, i = 1, 2, \ldots, K$, is non-convex, which effectively makes the whole problem in (7.4) non-convex. Relaxing

the constraint on the rank of $\mathbf{W}_i$, would give a semidefinite optimization problem with an optimal cost which provides a lower bound for the original problem in (7.2) [77]. This technique is formally known as the semidefinite relaxation (SDR) or Lagrangian relaxation, since it is the Lagrangian dual of the original problem [129, 130]. Problem (7.4), without the rank constraint belongs to a class of SDP and can be efficiently solved using interior point methods. Freely available softwares SeDuMi [33] or CVX [32] can be used for this purpose.

Interestingly, it has recently been shown that the problem in (7.2) can be cast into a second-order cone program (SOCP). As an arbitrary phase rotation of all the weight vectors does not change the SINR i.e., if $\mathbf{w}_i$ is optimum, $\mathbf{w}_i e^{j\phi_i}, i = 1, 2, \ldots, K$ will also satisfy the optimality. Thus, we can restrict ourself to the case where $\mathbf{h}_i^h \mathbf{w}_i$ is real valued i.e. $\text{imag}\{\mathbf{h}_i \mathbf{w}_i\} = 0$. Taking this into consideration, we can recast the SINR constraint in a SOCP framework as [87]

$$\left(1 + \frac{1}{\gamma_i}\right) |\mathbf{h}_i \mathbf{w}_i|^2 \geq \left\| \begin{matrix} \mathbf{h}_i \mathbf{W} \\ \sigma_i \end{matrix} \right\|_2^2, \quad i = 1, 2, \ldots, K \tag{7.5}$$

where $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_K]$. Thus we can write the problem in (7.2) in a SOCP as

$$\begin{aligned} \min_{\mathbf{w}_i} \quad & \sum_{i=1}^{K} \mathbf{w}_i^H \mathbf{w}_i \\ \text{s.t.} \quad & \sqrt{\left(1 + \frac{1}{\gamma_i}\right)} |\mathbf{h}_i \mathbf{w}_i| \geq \left\| \begin{matrix} \mathbf{h}_i \mathbf{W} \\ \sigma_i \end{matrix} \right\|_2, \quad i = 1, 2, \ldots, K \end{aligned} \tag{7.6}$$

It can be verified that both (7.4) and (7.6) attain the same optimum solution with an arbitrary phase shift.

7.2 Worst-Case Robust Beamforming

When the available CSI at the transmitter is imperfect, the performance of the non-robust approach in (7.2) will severely degrade. Robust modification has therefore been proposed in [77], where the true channel covariance matrices are assumed to be

$$\mathbf{R}_i = \hat{\mathbf{R}}_i + \mathbf{\Delta}_i, \tag{7.7}$$

where $\hat{\mathbf{R}}_i$ is the available estimate of the covariance matrix and $\mathbf{\Delta}_i$ is the uncertainty matrix whose Frobenius norm is bounded above by a known constant ϵ_i, i.e.

$$\|\mathbf{\Delta}_i\|_F \leq \epsilon_i. \tag{7.8}$$

Recall from Chapter 2, that (7.8) in 2-D defines a ball with a radius equal to ϵ. The basic idea behind worst-case performance optimization is to identify the worst possible error and to optimize the cost function for this worst possible error. This effectively ensures that the system would perform satisfactorily for the whole range of errors within the bound in (7.8). Problem (7.2), under the worst-case performance optimization framework, can be written as

$$\begin{aligned} &\min_{\mathbf{w}_i} \sum_{i=1}^{K} \mathbf{w}_i^H \mathbf{w}_i \\ &\text{s.t.} \quad \min_{\|\Delta_i\|_F \leq \epsilon_i} \frac{\mathbf{w}_i^H(\mathbf{R}_i + \Delta_i)\mathbf{w}_i}{\sum_{j=1, j\neq i}^{K} \mathbf{w}_j^H(\mathbf{R}_i + \Delta_i)\mathbf{w}_j + \sigma_i^2} \geq \gamma_i, \qquad i = 1, \ldots, K. \end{aligned} \tag{7.9}$$

Here, we wish to minimize the total transmit power subject to the constraint that the worst-case (or the minimum) SINR is above a pre-defined threshold. In [74, 77] it has been shown that the worst case SINR occurs when the numerator in the SINR equation attains its maximum with respect to Δ_i and the denominator attains its

maximum with respect to Δ_i. Thus, problem (7.9) can be written as,

$$\begin{aligned} &\min_{\mathbf{w}_i} \sum_{i=1}^{K} \mathbf{w}_i^H \mathbf{w}_i \\ &\text{s.t.} \quad \frac{\min_{\|\Delta_{1,i}\|_F \le \epsilon_i} \mathbf{w}_i^H (\mathbf{R}_i + \Delta_{1,i}) \mathbf{w}_i}{\max_{\|\Delta_{2,i}\|_F \le \epsilon_i} \sum_{j=1, j\neq i}^{K} \mathbf{w}_j^H (\mathbf{R}_i + \Delta_{2,i}) \mathbf{w}_j + \sigma_i^2} \ge \gamma_i, \qquad i = 1, \ldots, K. \end{aligned} \tag{7.10}$$

where separate uncertainty matrices $\Delta_{1,i}$ and $\Delta_{2,i}$ are used in the numerator and denominator of the QoS constraints in (7.10). Hence we need to determine [74]

$$\begin{aligned} \min_{\|\Delta_{1,i}\|_F} \quad & \mathbf{w}_i^H (\mathbf{R}_i + \Delta_{1,i}) \mathbf{w}_i \\ \text{s.t.} \quad & \|\Delta_{1,i}\|_F \le \epsilon_i, \end{aligned} \tag{7.11}$$

$$\begin{aligned} \max_{\|\Delta_{2,i}\|_F} \quad & \mathbf{w}_j^H (\mathbf{R}_i + \Delta_{1,i}) \mathbf{w}_j \\ \text{s.t.} \quad & \|\Delta_{2,i}\|_F \le \epsilon_i. \end{aligned} \tag{7.12}$$

The solutions to the subproblems (7.11) and (7.12) can be obtained analytically using Lagrange multipliers as in (7.13) and (7.14) respectively. See Appendix A for the proof.

$$\mathbf{w}_i^H (\mathbf{R}_i - \epsilon_i \mathbf{I}) \mathbf{w}_i \tag{7.13}$$

$$\mathbf{w}_j^H (\mathbf{R}_i + \epsilon_i \mathbf{I}) \mathbf{w}_j. \tag{7.14}$$

Thus, as shown in [77], the robust worst-case performance optimization problem

can be approximated as

$$\begin{aligned} &\min_{\mathbf{w}_i} \sum_{i=1}^{K} \mathbf{w}_i^H \mathbf{w}_i \\ &\text{s.t.} \quad \frac{\mathbf{w}_i^H(\mathbf{R}_i - \epsilon_i \mathbf{I})\mathbf{w}_i}{\sum_{j=1, j\neq i}^{K} \mathbf{w}_j^H(\mathbf{R}_i + \epsilon_i \mathbf{I})\mathbf{w}_j + \sigma_i^2} \geq \gamma_i, \qquad i = 1, \ldots, K. \end{aligned} \tag{7.15}$$

The above problem can be cast into a SDP and solved efficiently using interior-point methods [77]. Unfortunately, the PSD constraints on the matrices $\hat{\mathbf{R}}_i + \mathbf{\Delta}_i$ have been ignored in (7.15). This is due to the negative diagonal loading, in the term $\mathbf{w}_i^H(\mathbf{R}_i - \epsilon_i \mathbf{I})\mathbf{w}_i$. As a consequence, the matrix $\hat{\mathbf{R}} - \epsilon_i \mathbf{I}$ will not be positive definite, , when ϵ_i is larger than the smallest eigenvalue of $\hat{\mathbf{R}}_i$. Hence, the parameter ϵ_i less than the smallest eigenvalue of $\hat{\mathbf{R}}_i$ has to be chosen [122]. This can be interpreted as restricting the error in the channel covariance matrices to be sufficiently small. This, however might not be true in all practical scenarios. Therefore, to satisfy the PSD constraints and allowing for large errors on the channel covariance matrices, one must enforce the PSD constraints into the worst-case design. Another important observation is that worst-case design in practice are too conservative, as the worst-case errors could occur possibly with only a very low probability.

In the next section, we build on the the framework of worst-case performance optimization based robust beamforming, by adding PSD constraints on the channel covariance matrices. We propose two novel methods for the worst-case robust beamforming with PSD constraints. Unlike the worst-case design, the proposed methods not only satisfy the PSD constraints but also are less conservative.

7.3 Robust Beamforming with PSD Constraints

In this section, we take into consideration of the PSD constraints and develop a

less conservative approach for the worst-case transmit beamforming. Our objective is to solve the following problem:

$$\begin{aligned}
&\min_{\mathbf{w}_i} \sum_{i=1}^{K} \mathbf{w}_i^H \mathbf{w}_i \\
&\text{s.t.} \quad \min_{\|\Theta_i\| \le \epsilon_i} \frac{\mathbf{w}_i^H (\hat{\mathbf{R}}_i + \mathbf{\Delta}_i) \mathbf{w}_i}{\sum_{j=1, j \ne i}^{K} \mathbf{w}_j^H (\hat{\mathbf{R}}_i + \mathbf{\Delta}_i) \mathbf{w}_j + \sigma_i^2} \ge \gamma_i \\
&\qquad \hat{\mathbf{R}}_i + \mathbf{\Delta}_i \ge 0, \qquad i = 1, \ldots, K.
\end{aligned} \tag{7.16}$$

Note an additional constraint $\hat{\mathbf{R}}_i + \mathbf{\Delta}_i \ge 0$, as compared to worst-case design in (7.2), has been introduced to enforce the positive semidefiniteness on the channel covariance matrices. Instead of solving (7.16) directly, we follow the approach of [77] and [79], and approximate this problem into

$$\begin{aligned}
&\min_{\mathbf{w}_i} \sum_{i=1}^{K} \mathbf{w}_i^H \mathbf{w}_i \\
&\text{s.t.} \quad \frac{\min_{\|\check{\mathbf{\Delta}}_i\| \le \epsilon_i} \mathbf{w}_i^H (\hat{\mathbf{R}}_i + \check{\mathbf{\Delta}}_i) \mathbf{w}_i}{\max_{\|\check{\mathbf{\Delta}}_i\| \le \epsilon_i} \sum_{j=1, j \ne i}^{K} \mathbf{w}_j^H (\hat{\mathbf{R}}_i + \check{\mathbf{\Delta}}_i) \mathbf{w}_j + \sigma_i^2} \ge \gamma_i \\
&\qquad \hat{\mathbf{R}}_i + \check{\mathbf{\Delta}}_i \ge 0, \quad \hat{\mathbf{R}}_i + \check{\mathbf{\Delta}}_i \ge 0, \quad i = 1, \ldots, K
\end{aligned} \tag{7.17}$$

where separate uncertainty matrices $\check{\mathbf{\Delta}}_i$ and $\check{\mathbf{\Delta}}_i$ are used in the numerator and denominator of the QoS constraints in (7.17), respectively. Such an approximation strengthens the original QoS constraints of (7.16), [79, 122]. Introducing auxiliary variables ξ_i, we can rewrite (7.17) as

$$
\begin{aligned}
&\min_{\mathbf{w}_i,\xi_i} \sum_{i=1}^{K} \mathbf{w}_i^H \mathbf{w}_i \\
&\text{s.t.} \quad \min_{\|\check{\mathbf{\Delta}}_i\| \le \epsilon_i} \mathbf{w}_i^H (\hat{\mathbf{R}}_i + \check{\mathbf{\Delta}}_i)\mathbf{w}_i \ge \xi_i \\
&\qquad \xi_i - \max_{\|\check{\mathbf{\Delta}}_i\| \le \epsilon_i} \gamma_i \sum_{j=1, j\ne i}^{K} \mathbf{w}_j^H (\hat{\mathbf{R}}_i + \check{\mathbf{\Delta}}_i)\mathbf{w}_j \ge \gamma_i \sigma_i^2 \\
&\qquad \hat{\mathbf{R}}_i + \check{\mathbf{\Delta}}_i \ge 0, \;\; \hat{\mathbf{R}}_i + \check{\mathbf{\Delta}}_i \ge 0, \;\; \xi_i \ge 0, \qquad i = 1, \ldots, K.
\end{aligned} \tag{7.18}
$$

Let us split the problem into subproblems by separately establishing worst-case for the first and second constraints in (7.18).

The sub-problem associated with the second constraint in (7.18) can be written as

$$
\begin{aligned}
&\max_{\check{\mathbf{\Delta}}_i} \sum_{j=1, j\ne i}^{K} \mathbf{w}_j^H (\hat{\mathbf{R}}_i + \check{\mathbf{\Delta}}_i)\mathbf{w}_j \\
&\text{s.t.} \quad \hat{\mathbf{R}}_i + \check{\mathbf{\Delta}}_i \ge 0, \qquad \|\check{\mathbf{\Delta}}_i\| \le \epsilon_i, \quad i = 1, \ldots, K.
\end{aligned}
$$

The maximum objective value of (7.19) is given in [74] as

$$
\sum_{j=1, j\ne i}^{K} \mathbf{w}_j^H (\hat{\mathbf{R}}_i + \epsilon_i \mathbf{I})\mathbf{w}_j. \tag{7.19}
$$

where the related PSD constraints of (7.17) are automatically satisfied.

Let us now consider the sub-problem corresponding to the first constraint in (7.18). If we solve this problem without taking into account the PSD constraints of (7.17), the resulting solution could lead to a negative diagonal loading [74]. As a result, the related PSD constraints will be violated. Therefore, to enforce the PSD constraints, let us write $\hat{\mathbf{R}}_i = \mathbf{Q}_i^H \mathbf{Q}_i$ and model the uncertainties in $\mathbf{Q}_i$. In particular, $\mathbf{Q}_i$ can be the training data matrix used to estimate the downlink covariance matrix $\mathbf{R}_i$.

Defining $\bar{\mathbf{\Delta}}_i$ as a norm-bounded uncertainty matrix in $\mathbf{Q}_i$, we can rewrite the sub-problem in the first constraint of (7.18) as

$$\begin{aligned} &\min_{\bar{\mathbf{\Delta}}_i} \mathbf{w}_i^H(\mathbf{Q}_i + \bar{\mathbf{\Delta}}_i)^H(\mathbf{Q}_i + \bar{\mathbf{\Delta}}_i)\mathbf{w}_i \\ &\text{s.t.} \quad \|\bar{\mathbf{\Delta}}_i\| \leq \eta_i \end{aligned} \tag{7.20}$$

where η_i is some known bound on $\bar{\mathbf{\Delta}}_i$. To solve (7.20), we use the following lemma.

Lemma 1: If the mismatch is sufficiently small so that

$$\eta_i\|\mathbf{w}_i\| \leq \|\mathbf{Q}_i\mathbf{w}_i\| \tag{7.21}$$

then the solution to (7.20) is given by

$$\bar{\mathbf{\Delta}}_{i,*} = -\frac{\eta_i\mathbf{Q}_i\mathbf{w}_i\mathbf{w}_i^H}{\|\mathbf{w}_i\|\|\mathbf{Q}_i\mathbf{w}_i\|}$$

and the minimum value of the objective function is given by

$$(\|\mathbf{Q}_i\mathbf{w}_i\| - \eta_i\|\mathbf{w}_i\|)^2.$$

Proof: Using the triangle and the Cauchy-Schwarz inequalities along with the constraint $\|\bar{\mathbf{\Delta}}_i\| \leq \eta_i$, we get

$$\begin{aligned} \|(\mathbf{Q}_i + \bar{\mathbf{\Delta}}_i)\mathbf{w}_i\| &\geq \|\mathbf{Q}_i\mathbf{w}_i\| - \|\bar{\mathbf{\Delta}}_i\mathbf{w}_i\| \\ &\geq \|\mathbf{Q}_i\mathbf{w}_i\| - \|\bar{\mathbf{\Delta}}_i\|\|\mathbf{w}_i\| \\ &\geq \|\mathbf{Q}_i\mathbf{w}_i\| - \eta_i\|\mathbf{w}_i\|. \end{aligned} \tag{7.22}$$

It can be easily shown that if $\eta_i\|\mathbf{w}_i\| \leq \|\mathbf{Q}_i\mathbf{w}_i\|$, then

$$\|(\mathbf{Q}_i + \bar{\mathbf{\Delta}}_{i,*})\mathbf{w}_i\| = \|\mathbf{Q}_i\mathbf{w}_i\| - \eta_i\|\mathbf{w}_i\|. \tag{7.23}$$

This completes the proof of Lemma 1. □

Using (7.21) and Lemma 1, the first constraint in (7.18) can be written as

$$(\|\mathbf{Q}_i\mathbf{w}_i\| - \eta_i\|\mathbf{w}_i\|)^2 \geq \xi_i. \tag{7.24}$$

Using (7.19) and (7.24), the problem in (7.18) can be simplified as

$$\begin{aligned} &\min_{\mathbf{w}_i,\xi_i} \sum_{i=1}^{K} \mathbf{w}_i^H\mathbf{w}_i \\ &\text{s.t.} \quad (\|\mathbf{Q}_i\mathbf{w}_i\| - \eta_i\|\mathbf{w}_i\|)^2 \geq \xi_i \\ &\qquad \xi_i - \gamma_i \sum_{j=1,j\neq i}^{K} \mathbf{w}_j^H(\hat{\mathbf{R}}_i + \epsilon_i\mathbf{I})\mathbf{w}_j \geq \gamma_i\sigma_i^2 \\ &\qquad \xi_i \geq 0, \quad i = 1,\ldots,K. \end{aligned} \tag{7.25}$$

In the next two subsections, we present two different SDR-based methods to solve the problem in (7.25).

7.3.1 The First Method

Using the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} (\|\mathbf{Q}_i\mathbf{w}_i\| - \eta_i\|\mathbf{w}_i\|)^2 &= \|\mathbf{Q}_i\mathbf{w}_i\|^2 + \eta_i^2\|\mathbf{w}_i\|^2 - 2\eta_i\|\mathbf{Q}_i\mathbf{w}_i\|\|\mathbf{w}_i\| \\ &\geq \|\mathbf{Q}_i\mathbf{w}_i\|^2 + \eta_i^2\|\mathbf{w}_i\|^2 - 2\eta_i\|\mathbf{Q}_i\|\|\mathbf{w}_i\|^2. \end{aligned} \tag{7.26}$$

Using (7.26), the first constraint in (7.25) can be modified as

$$\|\mathbf{Q}_i\mathbf{w}_i\|^2 + \eta_i^2\|\mathbf{w}_i\|^2 - 2\eta_i\|\mathbf{Q}_i\|\|\mathbf{w}_i\|^2 \geq \xi_i. \tag{7.27}$$

Introducing a new variable

$$\mathbf{W}_i = \mathbf{w}_i\mathbf{w}_i^H$$

and using the property

$$\mathrm{Tr}\{\mathbf{w}_i^H\hat{\mathbf{R}}_i\mathbf{w}_i\} = \mathrm{Tr}\{\hat{\mathbf{R}}_i\mathbf{w}_i\mathbf{w}_i^H\} = \mathrm{Tr}\{\hat{\mathbf{R}}_i\mathbf{W}_i\}$$

we rewrite (7.27) as

$$\begin{aligned}
&\mathrm{Tr}\{\hat{\mathbf{R}}_i\mathbf{W}_i\} + \eta_i^2\mathrm{Tr}\{\mathbf{W}_i\} - 2\eta_i\|\mathbf{Q}_i\|\mathrm{Tr}\{\mathbf{W}_i\} \geq \xi_i \\
&\mathbf{W}_i \geq 0, \quad \mathrm{rank}(\mathbf{W}_i) = 1.
\end{aligned} \tag{7.28}$$

Dropping the non-convex rank-one constraint in (7.28) and reformulating the objective function and the second constraint of (7.25) in terms of the new variable $\mathbf{W}_i$, we obtain the following problem formulation

$$\begin{aligned}
&\min_{\mathbf{W}_i,\xi_i} \sum_{i=1}^{K} \mathrm{Tr}\{\mathbf{W}_i\} \\
&\text{s.t.} \quad \mathrm{Tr}\{\hat{\mathbf{R}}_i\mathbf{W}_i\} + \eta_i^2\mathrm{Tr}\{\mathbf{W}_i\} - 2\eta_i\|\mathbf{Q}_i\|\mathrm{Tr}\{\mathbf{W}_i\} \geq \xi_i \\
&\qquad \xi_i - \gamma_i \sum_{j\neq i}^{K} \mathrm{Tr}\{(\hat{\mathbf{R}}_i + \epsilon_i\mathbf{I})\mathbf{W}_j\} \geq \gamma_i\sigma_i^2 \\
&\qquad \mathbf{W}_i \geq 0, \quad \xi_i \geq 0, \quad i = 1,\ldots,K.
\end{aligned} \tag{7.29}$$

This is a convex SDP problem that can be solved by means of, for example, interior-point algorithms [?, 32, 33, 61, 131].

7.3.2 The Second Method

Let us express the first constraint in (7.25) as

$$\|\mathbf{Q}_i\mathbf{w}_i\| \geq \eta_i\|\mathbf{w}_i\| + \sqrt{\xi_i} \tag{7.30}$$

Squaring both sides, we obtain,

$$\|\mathbf{Q}_i\mathbf{w}_i\|^2 \geq \eta_i^2\|\mathbf{w}_i\|^2 + \xi_i + 2\eta_i\sqrt{\xi_i}\|\mathbf{w}_i\|. \tag{7.31}$$

Using a new variable $\mathbf{W}_i = \mathbf{w}_i\mathbf{w}_i^H$, we express the constraint (7.31) as

$$\begin{aligned} &\mathrm{Tr}\{\hat{\mathbf{R}}_i\mathbf{W}_i\} - \eta_i{}^2\mathrm{Tr}\{\mathbf{W}_i\} - \xi_i \geq 2\eta_i\zeta_i \\ &\mathbf{W}_i \geq 0, \quad \mathrm{rank}(\mathbf{W}_i) = 1. \end{aligned} \tag{7.32}$$

where $\zeta_i = \sqrt{\xi_i}\sqrt{\mathrm{Tr}\{\mathbf{W}_i\}}$. Using the SDR approach again, we drop the rank-one constraint in (7.32), and reformulate the objective function and the second constraint of (7.25) in terms of the new variable $\mathbf{W}_i$. This results in the following optimization problem:

$$\begin{aligned} &\min_{\mathbf{W}_i,\xi_i} \sum_{i=1}^{K} \mathrm{Tr}\{\mathbf{W}_i\} \\ &\text{s.t.} \quad \mathrm{Tr}\{\hat{\mathbf{R}}_i\mathbf{W}_i\} - \eta_i^2\mathrm{Tr}\{\mathbf{W}_i\} - \xi_i \geq 2\eta_i\zeta_i \\ &\qquad \xi_i - \gamma_i \sum_{j\neq i}^{K} \mathrm{Tr}\{(\hat{\mathbf{R}}_i + \epsilon_i\mathbf{I})\mathbf{W}_j\} \geq \gamma_i\sigma_i^2 \\ &\qquad \mathbf{W}_i \geq 0, \quad \xi_i \geq 0, \quad i = 1, \ldots, K. \end{aligned} \tag{7.33}$$

However, the first constraint in (7.33) is still non-convex due to the term $2\eta_i\zeta_i$ on the right hand side. Therefore, we resort to an iterative scheme to solve (7.33). We

find $\mathbf{W}_{i,l}$ and $\xi_{i,l}$ in the lth iteration by means of solving the following optimization problem

$$\begin{aligned}
&\min_{\mathbf{W}_{i,l},\xi_{i,l}} \sum_{i=1}^{K} \mathrm{Tr}\{\mathbf{W}_{i,l}\} \\
&\text{s.t.} \quad \mathrm{Tr}(\hat{\mathbf{R}}_i\mathbf{W}_{i,l}) - \eta_i^2\mathrm{Tr}\{\mathbf{W}_{i,l}\} - \xi_{i,l} \geq 2\eta_i\zeta_{i,l-1} \\
&\qquad \xi_{i,l} - \gamma_i \sum_{j\neq i}^{K} \mathrm{Tr}\{(\hat{\mathbf{R}}_i + \epsilon_i\mathbf{I})\mathbf{W}_{j,l}\} \geq \gamma_i\sigma_i^2 \\
&\qquad \mathbf{W}_{i,l} \geq 0, \quad \xi_{i,l} \geq 0, \quad i = 1,\ldots,K
\end{aligned} \tag{7.34}$$

where $\zeta_{i,l-1} = \sqrt{\xi_{i,l-1}}\sqrt{\mathrm{Tr}\{\mathbf{W}_{i,l-1}\}}$ is the solutions obtained in the $(l-1)$th iteration.

Similar to (7.29), each iteration of (7.34) belongs to a class of convex SDP problems which can be solved using modern convex optimization tools [32,33,131].

The solutions to the problems in (7.29) and (7.34) may not be rank-one in general. In such cases, the so-called *randomization techniques* [132], [133] should be used to obtain approximate solutions to the original problem in (7.16). However, our simulation results show that we always obtain a rank-one solution for $\mathbf{W}_i$. Therefore, we retrieve the beamforming weight vectors $\mathbf{w}_i$ from $\mathbf{W}_i$ using the principal eigenvector corresponding to its single non-zero eigenvalue.

7.3.3 Relationship Between the Regularization Parameters

Let us now obtain an approximate relationship between the parameters ϵ_i and η_i. We know that

$$(\hat{\mathbf{R}}_i + \check{\mathbf{\Delta}}_i) = (\mathbf{Q}_i + \bar{\mathbf{\Delta}}_i)^H(\mathbf{Q}_i + \bar{\mathbf{\Delta}}_i). \tag{7.35}$$

Expanding (7.35), we can write

$$\check{\Delta}_i = \mathbf{Q}_i^H \bar{\Delta}_i + \bar{\Delta}_i^H \mathbf{Q}_i + \bar{\Delta}_i^H \bar{\Delta}_i. \tag{7.36}$$

Taking the Frobenius norms on both sides of (7.36) and using the triangle and Cauchy-Schwarz inequalities, we obtain

$$\begin{aligned} \|\check{\Delta}_i\| &= \|\mathbf{Q}_i^H \bar{\Delta}_i + \bar{\Delta}_i^H \mathbf{Q}_i + \bar{\Delta}_i^H \bar{\Delta}_i\| \\ &\leq \|\mathbf{Q}_i^H \bar{\Delta}_i\| + \|\bar{\Delta}_i^H \mathbf{Q}_i\| + \|\bar{\Delta}_i^H \bar{\Delta}_i\| \\ &\leq \|\mathbf{Q}_i\|\|\bar{\Delta}_i\| + \|\bar{\Delta}_i\|\|\mathbf{Q}_i\| + \|\bar{\Delta}_i^H \bar{\Delta}_i\| \\ &\leq 2\eta_i\|\mathbf{Q}_i\| + \eta_i^2, \end{aligned} \tag{7.37}$$

and, therefore

$$\epsilon_i \leq 2\eta_i\|\mathbf{Q}_i\| + \eta_i^2. \tag{7.38}$$

Using (7.38), for any given ϵ_i we can compute an approximate value for η_i by converting the inequality in (7.38) to an equality.

7.4 Joint Beamforming and BS Assignment

So far, we considered a single BS transmitter serving multiple decentralized single antenna users. In these schemes, the assignment of mobile stations (MSs) to the BSs has been assumed to be known i.e. which MS has been connected to which BS is known. The assignment is performed in a decentralized manner, where a MS is assigned to a BS to which it has the best radio link in terms of channel gains. It is also possible to optimally assign BSs. The problem of joint optimal beamforming and power control has been extended to BS assignment in [127, 134]. The scheme

in [134] exploits the uplink-downlink duality to obtain an iterative solution. The BS assignment problem is solved for an equivalent uplink problem and the same assignment is used in the downlink as a solution to the problem of joint beamforming and power control. In [127], the authors proposed a semidefinite programming (SDP) based solution. In [127] the antenna array of each BS is considered as part of a single giant BS. The problem of joint downlink beamforming and power control with this single giant BS is then cast into a convex SDP and optimal beamforming vectors are obtained for this giant BS [127]. Although each MS is allowed to be connected to multiple BSs, the optimal solution turns out that each MS is connected only to a single BS [127]. Here, we will briefly outline the system model for a joint beamforming and BS assignment problem. The two novel methods developed in the previous sections can also be applied to this problem as illustrated in the simulation results.

7.5 Problem Formulation

Let us consider a system with N BS transmitters each equipped with N_T antennas and K decentralized single-antenna mobile stations (MSs) in a flat fading channel environment as shown in Fig. 7.1. We wish to optimally design the downlink beamformers and assign each MS to a best possible BS. Each BS could serve multiple MSs, however the optimum solution turns out to be that each MS is connected only to a single BS. Let us denote the BS assigned to the ith user as $\kappa(i)$ and define $\mathcal{S}(n) = \{i : \kappa(i) = n\}$ as a set of indices of mobiles that have been assigned to the nth BS. The BS n transmits a $N_T \times 1$ signal vector at time t to a set of mobiles $\mathcal{S}(n)$ as

$$\mathbf{x}_n(t) = \sum_{i \in \mathcal{S}(n)} \mathbf{w}_i s_i(t), \tag{7.39}$$

where $\mathbf{w}_i \in \mathbb{C}^{N_T \times 1}$ are the complex beamforming vectors to be determined and

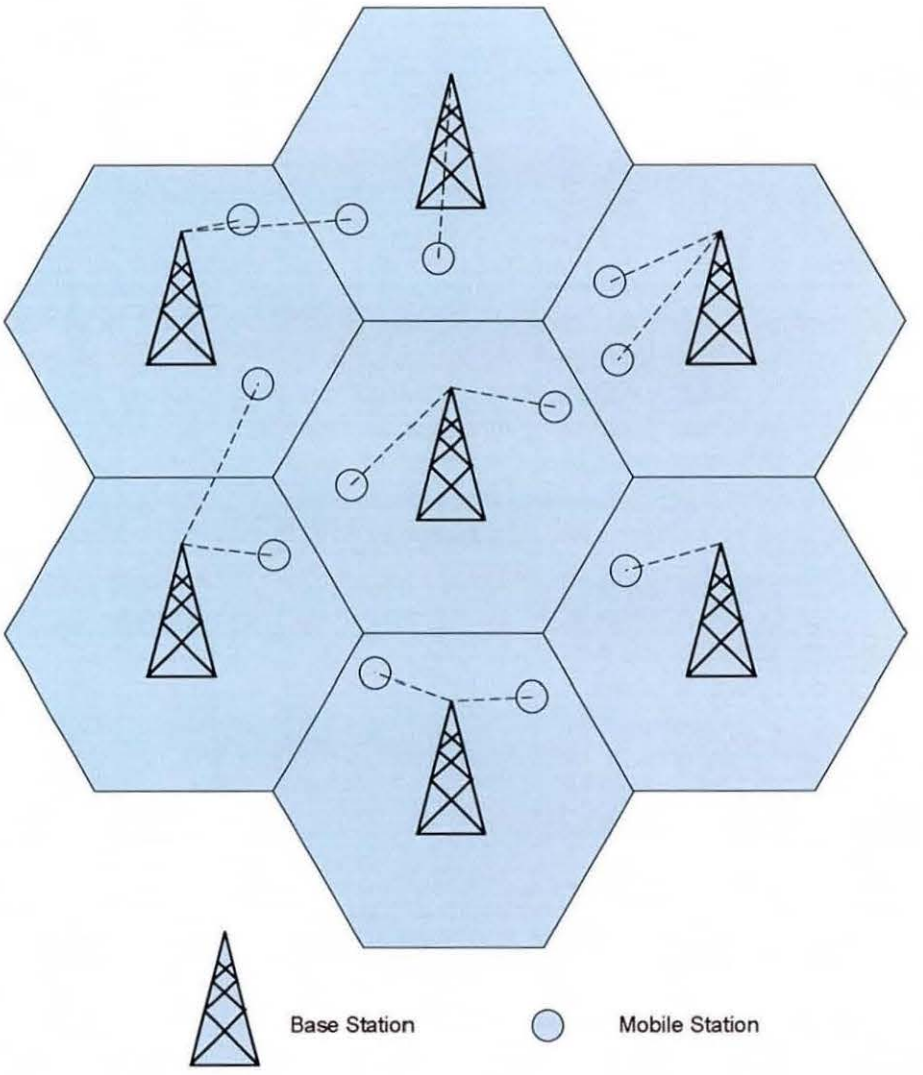

Figure 7.1: Downlink system model.

$s_i(t)$ is the data symbol intended for the ith user. We assume that all $s_i(t)$ are uncorrelated and have the same normalized power $\mathrm{E}\{|s_i(t)|^2\} = 1$. The vector $\mathbf{x}_n(t)$ is then transmitted from the nth BS over frequency flat channels. The signal received by the ith user is given by

$$y_i(t) = \sum_{n=1}^{N} \mathbf{h}_{i,n}\mathbf{x}_n(t) + n_i(t), \tag{7.40}$$

where $\mathbf{h}_{i,n} \in \mathbb{C}^{1\times N_T}$ is the complex channel vector between the ith user and the nth BS, $n_i(t)$ is the ZMCSCG noise with variance σ_i^2. Since i^{th} user is connected to the BS $\kappa(i)$, the SINR of the ith user can be written as

$$\mathrm{SINR}_i = \frac{\mathbf{w}_i^H \mathbf{R}_{i,\kappa(i)} \mathbf{w}_i}{\sum_{j=1, j\neq i}^{K} \mathbf{w}_j^H \mathbf{R}_{i,\kappa(j)} \mathbf{w}_j + \sigma_i^2}, \tag{7.41}$$

where

$$\mathbf{R}_{i,n} = \mathrm{E}\{\mathbf{h}_{i,n}^H \mathbf{h}_{i,n}\} \in \mathbb{S}_+^{N_T\times N_T}, \qquad \begin{aligned} i &= 1, 2, \ldots, K \\ n &= 1, 2, \ldots, N \end{aligned}$$

is the downlink channel covariance matrix between the nth BS and the ith user. Let us first assume the sets $\mathcal{S}(n), n = 1, 2, \ldots, N$ are known, i.e. the indices $\kappa(i)$ are known. This BS assignment can be performed in a decentralized manner, where a MS is assigned to a BS to which it has the best link in terms of channel gain. Now we wish to design the beamformers $\mathbf{w}_i$ for all K users. A meaningful problem formulation is to minimize the total transmit power from all the BSs subject to satisfying each user's QoS [80]. This optimization problem can be written as [78], [80]

$$\min_{\mathbf{w}_i} \sum_{i=1}^{K} \mathbf{w}_i^H \mathbf{w}_i$$
$$\text{s.t.} \quad \frac{\mathbf{w}_i^H \mathbf{R}_{i,\kappa(i)} \mathbf{w}_i}{\sum_{j=1, j\neq i}^{K} \mathbf{w}_j^H \mathbf{R}_{i,\kappa(j)} \mathbf{w}_j + \sigma_i^2} \geq \gamma_i, \qquad i = 1, \ldots, K \tag{7.42}$$

where γ_i is the minimum acceptable SINR for the ith user. In [78], it has been shown that the non-convex problem (7.42) can be relaxed into a convex problem using SDR and can be solved efficiently using SDP. Nevertheless, the problem formulation in (7.42) itself is non-optimal as the indices $\kappa(i)$ are assumed to be known or are determined in a decentralized manner. The optimum solution to downlink beamforming and BS assignment could be obtained if the problem is optimized over both the beamforming vectors $\mathbf{w}_i$ and the BS assignment $\kappa(i)$. In [127] it has been shown that, a conceptually interesting way to solve this is to relax the problem in (7.42) and allow all BSs to simultaneously transmit to all MSs. This is possible by considering the antenna array of each BS as part of the antenna array of a single giant BS. This problem can be formulated as designing a single weight vector

$$\bar{\mathbf{w}}_i = \begin{bmatrix} (\mathbf{w}_{i,1})^T & (\mathbf{w}_{i,2})^T & \ldots & (\mathbf{w}_{i,N})^T \end{bmatrix}^T, \tag{7.43}$$

for transmission over a single channel covariance matrix

$$\bar{\mathbf{R}}_i = \begin{bmatrix} \mathbf{R}_{i,1} & \mathbf{0}_{N_t \times N_t} & \ldots & \mathbf{0}_{N_t \times N_t} \\ \mathbf{0}_{N_t \times N_t} & \mathbf{R}_{i,2} & \ldots & \mathbf{0}_{N_t \times N_t} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N_t \times N_t} & \mathbf{0}_{N_t \times N_t} & \ldots & \mathbf{R}_{i,N} \end{bmatrix}, \tag{7.44}$$

for the ith user. We assume that $\mathrm{E}\{\mathbf{h}_{i,n_1}^H \mathbf{h}_{i,n_2}\} = 0$ for all $n_1 \neq n_2$ and $\mathbf{0}_{N_t \times N_t}$ denotes a matrix of all zeros. Using $\bar{\mathbf{w}}_i$ and $\bar{\mathbf{R}}_i$, we can write the SINR for the ith

user as,

$$\mathrm{SINR}_i = \frac{\bar{\mathbf{w}}_i^H \bar{\mathbf{R}}_i \bar{\mathbf{w}}_i}{\sum_{j=1, j\neq i}^{K} \bar{\mathbf{w}}_j^H \bar{\mathbf{R}}_i \bar{\mathbf{w}}_j + \sigma_i^2} \tag{7.45}$$

Now we wish to design the beamformers $\bar{\mathbf{w}}_i$ for all K users. This optimization problem can be written as [78], [80]

$$\begin{aligned} &\min_{\bar{\mathbf{w}}_i} \sum_{i=1}^{K} \bar{\mathbf{w}}_i^H \bar{\mathbf{w}}_i \\ &\text{s.t.} \quad \frac{\bar{\mathbf{w}}_i^H \bar{\mathbf{R}}_i \bar{\mathbf{w}}_i}{\sum_{j=1, j\neq i}^{K} \bar{\mathbf{w}}_j^H \bar{\mathbf{R}}_i \bar{\mathbf{w}}_j + \sigma_i^2} \geq \gamma_i \\ &\quad\; i = 1, \ldots, K \end{aligned} \tag{7.46}$$

According to this new problem formulation, a MS can be served by multiple BSs, however, as we will see later (and as in [127]), the optimality will enforce each MS to choose only a single BS. The problem in (7.46) can be approximated to a convex SDP using SDR, [78] and [127].

The problem formulation in (7.46) assumes perfect CSI at the transmitter. We considered the problem of joint downlink beamforming and BS assignment with imperfect CSI at the transmitter using worst-case performance optimization with PSD constraints in [9]. The two novel schemes developed in §7.2 were also extended to the problem of robust beamforming and BS assignment in [9].

7.6 Simulation Results

7.6.1 Single BS

In the first section of simulations we consider the scenario used in [78], where the authors consider a single linear transmit antenna array of $N_T = 8$ sensors spaced half

a wavelength apart. There are $K = 3$ single-antenna users. One user is located at $\theta_1 = 10°$ relative to the array broadside and the other two are located at $\theta_{2,3} = 10° \pm \phi$, where ϕ is varied from 0° to 10°. It is assumed that the users are surrounded by a large number of local scatterers corresponding to a spread angle of $\sigma_\theta = 2°$, as seen from the BS. In such a scenario, the channel covariance matrix takes the form [78]

$$[\mathbf{R}_i]_{k,l} = e^{j\pi(k-l)\sin\theta_i} e^{-\frac{(\pi(k-l)\sigma_\theta \cos\theta_i)^2}{2}}. \tag{7.47}$$

The noise at each user is assumed to be additive white Gaussian with variance $\sigma_i^2 = 1$ $(i = 1, \ldots, K)$. The same SINR threshold $\gamma_i = \gamma$ for all $i = 1, \ldots, K$ is used for all the users. For each channel covariance matrix $\hat{\mathbf{R}}_i$ $(i = 1, \ldots, K)$, the corresponding error matrix $\mathbf{\Delta}_i$ has been uniformly and randomly generated in a sphere centered at zero with the radius ϵ_i. For simplicity we assume that $\epsilon_i = \epsilon$ for all $i = 1, \ldots, K$. In all the examples, we compare the proposed two methods to the robust and non-robust schemes presented in [78] and [77].

Fig. 7.2 depicts the total transmitted power versus the angular separation ϕ for ϵ that is varied in the interval $[0.05, \ldots, 0.2]$ where the corresponding values of η_i are calculated using (7.38). The value of γ used in this figure is 5 dB. We note that the power consumption of all robust beamformers increases as the value of ϵ is increased. This is not unexpected as with an increment in ϵ, the uncertainty region grows and the quality of available CSI degrades. However, we can see that the proposed robust techniques have better transmitted power requirements than the worst-case robust technique of [77] for the complete range of ϵ. Note that, although our first method satisfies the PSD constraints, it is only able to offer minimal gain in terms of power over the worst-case design of [77]. However, our second method is able to offer a substantial gain in terms of the transmitted power. This is especially true in the case when the user angle separation is low.

Figs. 7.3 and 7.4 depict histograms of the number of constraints versus the nor-

malized constraint value, ζ_i, which is defined as

$$\zeta_i = \frac{1}{\gamma_i \sigma_i^2} \mathbf{w}_i^H \mathbf{R}_i \mathbf{w}_i - \frac{1}{\sigma_i^2} \sum_{j=1, j \neq i}^{K} \mathbf{w}_j^H \mathbf{R}_i \mathbf{w}_j. \tag{7.48}$$

Note that if $\zeta_i \geq 1$, then the corresponding constraint is satisfied; otherwise it is violated.

In Figs. 7.3 and 7.4, we consider a set of scenarios with $\{\epsilon = 0.2, \phi = 7°\}$ and $\{\epsilon = 0.15, \phi = 6.5°\}$, respectively. The value of γ used in these figures is 5 dB. We can see from these two figures that the non-robust technique of [78] violates almost 50% of the constraints. The robust technique in [77] satisfies all the constraints, but it is obviously too conservative and $\zeta_i \gg 1$ for all constraints (i.e. constraints are over-satisfied). Although the proposed techniques also over-satisfy the constraints, they do it in a lesser extent than the approach in [77]. This is especially true for the second method.

Fig. 7.5 depicts the total transmitted power needed to achieve a given set of SINR thresholds $\gamma = \{3\text{dB}, \ldots, 8\text{dB}\}$. The value of ϵ is varied in the range $\{0.05, \ldots, 0.2\}$ and the angular separation used is $\phi = 7°$. Once again the corresponding values of η_i are calculated using (7.38). We can see a similar trend as we saw in Fig. 7.2, that the performance of the beamformers worsens when the value of ϵ is increased. However, the proposed beamformers perform better than the worst-case robust beamformer of [77] for a range of SINR targets.

7.6.2 BS Assignment

In this section, we consider a scenario with multiple BSs, where each BS is equipped with a linear array of antennas ($N_T = 8$) spaced half a wavelength apart. For each user in the system, we generate a set of N channel covariance matrices using (7.47). The noise at each user is assumed to be additive white Gaussian with variance $\sigma_i^2 = 1$

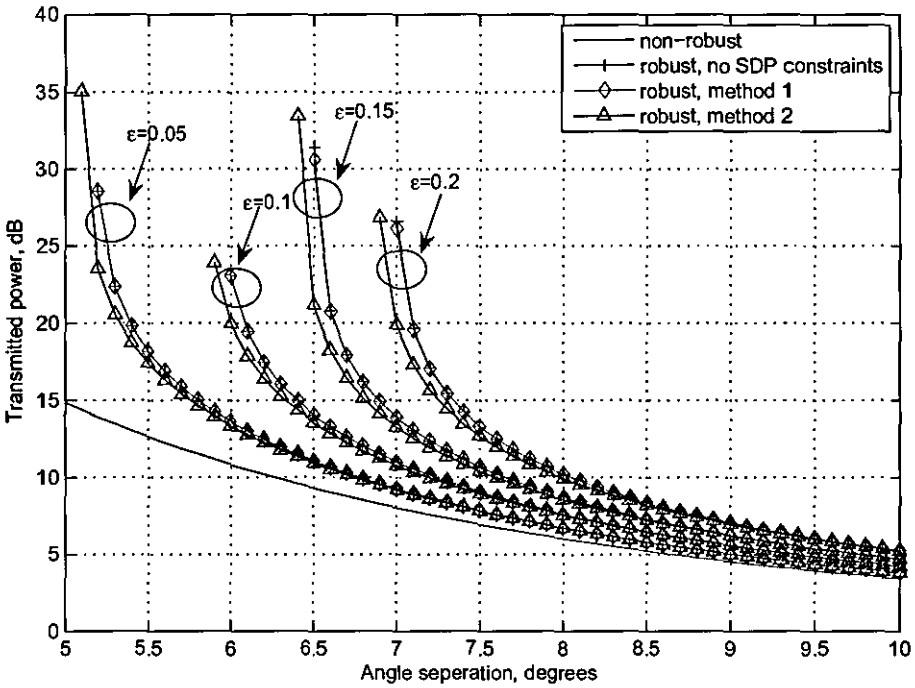

Figure 7.2: Total transmitted power versus the angular separation ϕ.

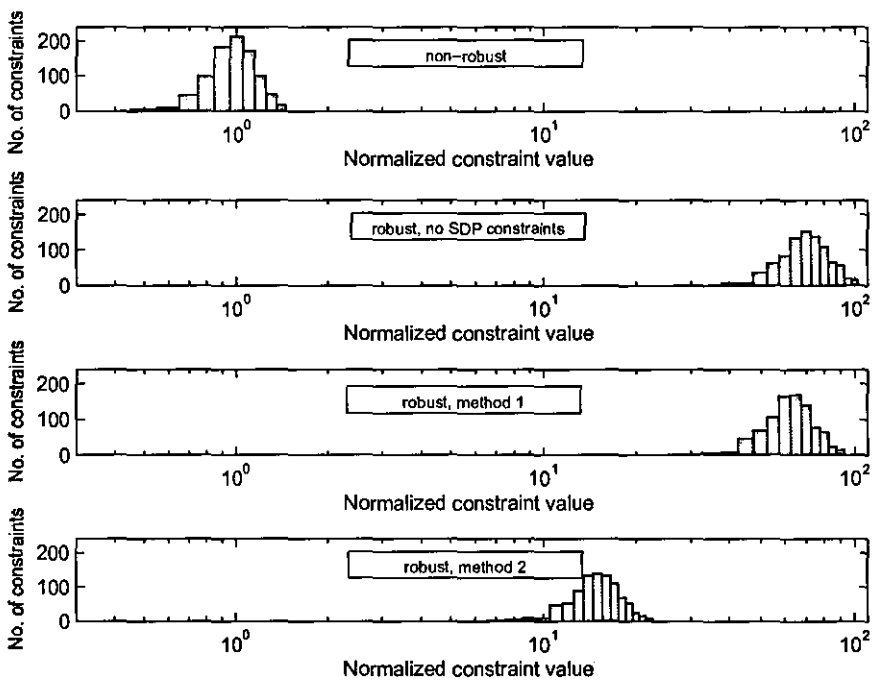

Figure 7.3: Histogram of the constraints versus the normalized constraint value for $\epsilon = 0.2$ and $\phi = 7°$.

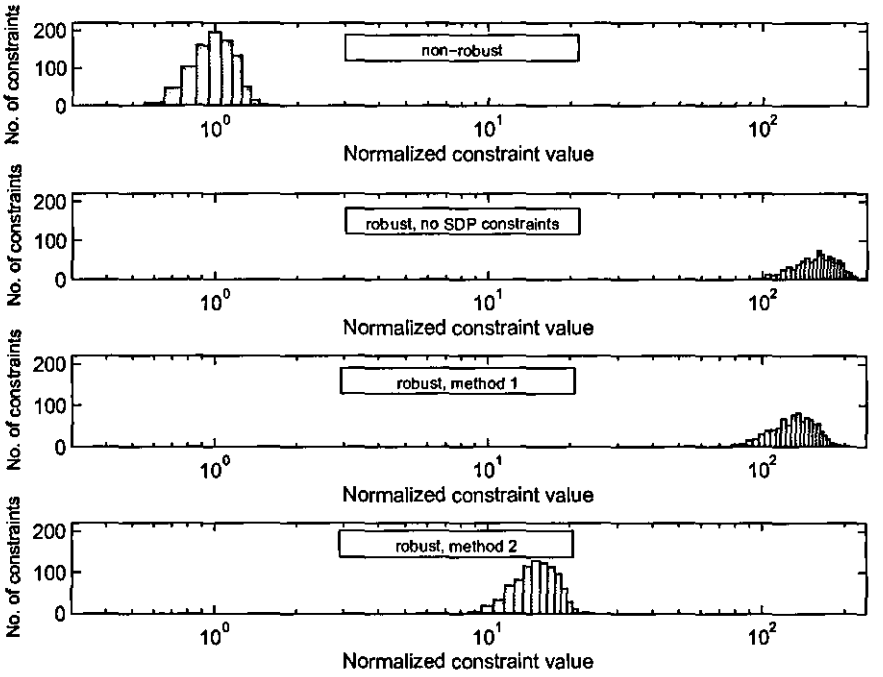

Figure 7.4: Histogram of the constraints versus the normalized constraint value for $\epsilon = 0.15$ and $\phi = 6.5°$.

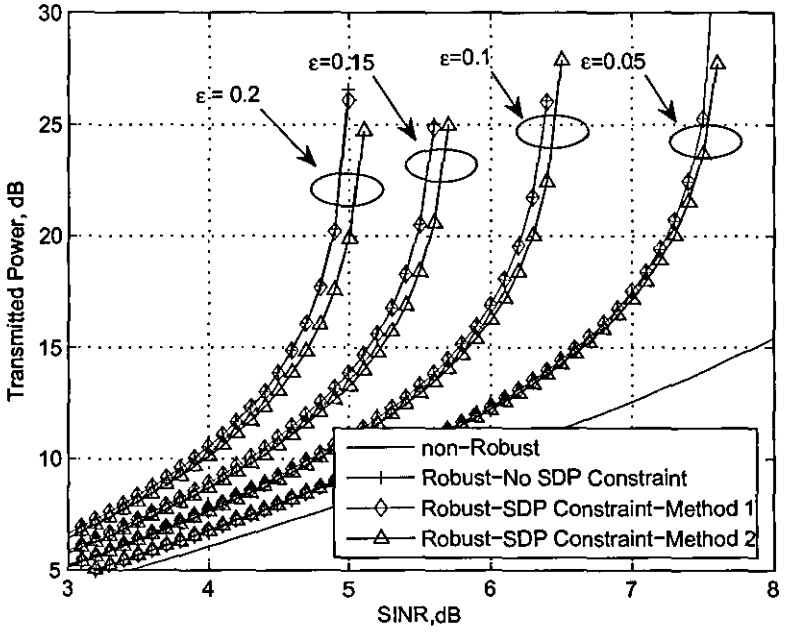

Figure 7.5: Total transmitted power versus minimum required SINR for values of ϵ in the range $\{0.05, \ldots, 0.2\}$ and $\phi = 7°$

$(i = 1, \ldots, K)$. The same SINR threshold $\gamma_i = \gamma$ for all $(i = 1, \ldots, K)$ is used for all the users and $\epsilon_i = \epsilon$ for all $(i = 1, \ldots, K)$. We then jointly solve the problem of beamforming and BS assignment. Note as mentioned earlier, each user is assigned only one BS, whereas one BS might have more than one users assigned to it.

In the first simulation we consider a scenario with 2 BSs. There are $K = 3$ single-antenna users. The first user is kept fixed and located at $\theta_{1,1} = 10°$,$\theta_{1,2} = 30°$ relative to the array broadside of BS 1 and BS 2 respectively. The other two users relative to user 1 and the two BSs are located at $\theta_{2,k} = \theta_{1,k} + \phi$, $\theta_{3,k} = \theta_{1,k} - \phi$, $(k = 1, 2)$, where ϕ is varied from 0° to 15°. An angular spread of 2° is considered. Fig. 7.6 depicts the total transmitted power from the assigned BSs versus the angular separation ϕ for ϵ that is varied in the interval $[0.05, \ldots, 0.2]$ where the corresponding values of η_i are calculated using (7.38). The value of γ used in this figure is 5 dB. Similarly as before, we note that the proposed robust techniques have a better transmit power requirement than the worst-case robust techniques of [77] for a complete range of ϵ. We note that the transmit power needed to achieve SINR targets of 5dB with BS assignment is comparably lower than power needed in a single BS scenario. This is simply because there are additional degrees of freedom to minimize the transmit power due to the availability of extra BSs.

In the second simulation, we consider a scenario with $N = 3$ BSs and $K = 3$ single-antenna users. The first user is kept fixed and is located at $\theta_{1,1} = 15o$, $\theta_{1,2} = 35°$ and $\theta_{1,3} = 60°$ relative to the array broadside of BS 1, BS 2 and BS 3 respectively. The other two users relative to user 1 and the three BSs are located at $\theta_{2,k} = \theta_{1,k} + \phi$, $\theta_{3,k} = \theta_{1,k} - \phi$, $(k = 1, 2)$, where ϕ is varied from 0° to 15°. An angular spread of 2° is considered. Fig. 7.7 depicts the total transmitted power from the assigned BSs versus the angular separation ϕ for γ that is varied in the interval $[1, \ldots, 5]$dB and ϵ is kept fixed at 0.1 for the whole range of γ. The proposed robust techniques offer a better transmit power requirement than the worst-case robust techniques of [77] for

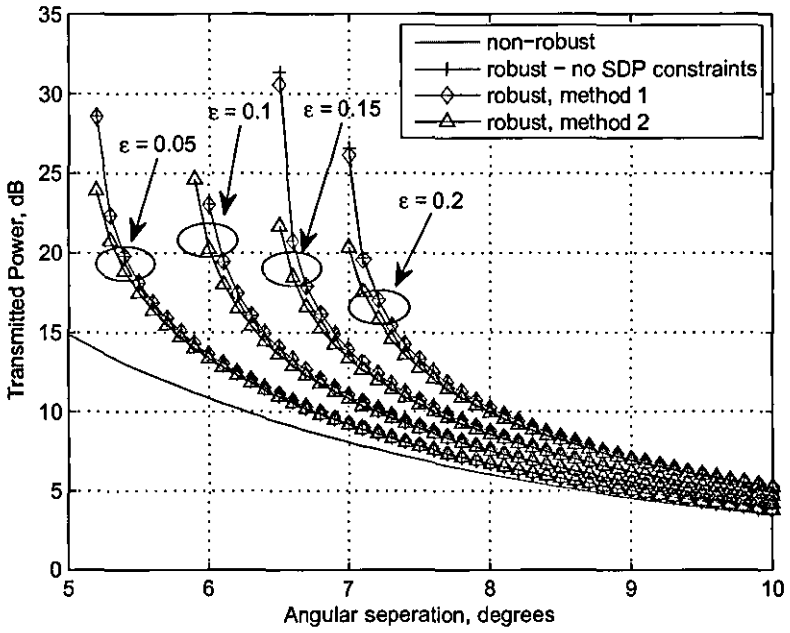

Figure 7.6: Total transmitted power versus the angular separation ϕ.

a complete range of γ.

7.7 Summary

In this chapter, we discussed the problem of robust downlink beamforming under a worst-case performance optimization framework. We highlighted the fact that the worst-case based designs violate the positive semidefinite constraints on the channel covariance matrices and are highly conservative. We therefore, developed a worst-case based optimization technique by incorporating the PSD constraints. The proposed techniques use semi-definite relaxation to approximately convert the original beamforming problems into a convex semi-definite programming problem. The proposed techniques have shown to outperform the conventional transmit beamforming techniques.

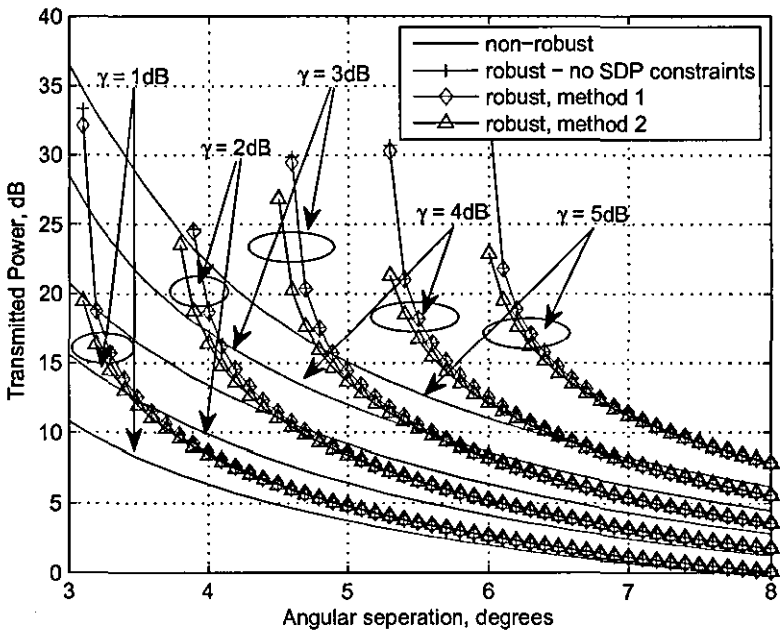

Figure 7.7: Total transmitted power versus the angular separation ϕ.

Chapter 8

Robust Downlink Beamforming with Per Antenna Power Constraints

Spatial multiplexing techniques significantly improve the spectral efficiency of wireless communication systems. In particular downlink beamforming has proven to be a simple, yet very efficient technique to improve spectral efficiency of a wireless communication system. Here we consider the downlink beamforming problem in a multi-user environment, where a basestation (BS) is equipped with multiple transmitting antennas, simultaneously transmits independent data streams to multiple single-antenna receivers.

The problem of conventional downlink beamforming can be defined as minimizing the total transmit power subject to satisfying the quality of services (QoS) such as the mean square error (MSE) or the signal-to-interference plus noise ratio (SINR) for all users in the system. Various advanced algorithms for conventional downlink beamforming have been proposed in [34, 77, 80, 81, 123] and references therein. In a practical system, each antenna of the transmitter may be equipped with its own power

amplifier and thus a more meaningful problem formulation is to minimize the total transmit power subject to a per antenna power constraint and the QoS constraints. In [87,135] several advanced algorithms for downlink beamforming with per antenna power constraints under the assumption that perfect CSI at the transmitter have been proposed.

However, in practical situations perfect CSI may not be available at the transmitter and the performance of the transmit beamforming methods could degrade as the quality of the available CSI reduces [34,77,80,81,87,123,135]. Typically, only an estimate of CSI is available at the transmitter through a feedback from the receiver. This estimate is normally in error due to quantization, feedback delay, estimation errors etc. Therefore, robust techniques are required to take into account the CSI errors.

In this chapter, we propose a robust solution to this problem of downlink beamforming with per antenna power constraints and QoS constraints under the assumption that only erroneous CSI is available at the transmitter. We solve this problem using worst-case performance optimization. We assume that the mismatches between the presumed and the actual downlink channel covariance matrices are norm-bounded by a known constant. The beamformer performance is optimized for the worst-case mismatch. We demonstrate the proposed robust solution attains the QoS targets with probability one while the non-robust scheme attains the target QoS only with probability 0.5.

8.1 Problem Formulation

Let us consider a wireless system with one BS and K decentralized single antenna users. The BS transmits a multiplexed vector $\mathbf{x}(t) \in \mathbb{C}^{N_T \times 1}$ to K users at time t as

$$\mathbf{x}(t) = \sum_{j=1}^{K} \mathbf{w}_j s_j(t), \tag{8.1}$$

where $\mathbf{w}_j \in \mathbb{C}^{N_T \times 1}$ is the complex beamforming vector for the jth user to be determined and $s_j(t)$ denotes the data symbol intended for the jth user. We assume that all $s_j(t)$ are uncorrelated and have the same normalized power $\mathrm{E}\{|s_j(t)|^2\} = 1$. The signal received by the jth user can be written as

$$\begin{aligned} y_j(t) &= \mathbf{h}_j \mathbf{x}(t) + n_j \\ &= \underbrace{\mathbf{h}_j \mathbf{w}_j s_j(t)}_{\text{Signal}} + \underbrace{\sum_{j \neq i}^{K} \mathbf{h}_j \mathbf{w}_i s_i(t)}_{\text{Interference}} + \underbrace{n_j}_{\text{Noise}} \end{aligned} \tag{8.2}$$

where $\mathbf{h}_j \in \mathbb{C}^{1 \times N_T}$ denotes the flat fading channel vector between the jth user and the BS and n_j denotes the zero-mean additive white Gaussian noise (AWGN) with variance σ_n^2. From (8.2), the signal-to-interference plus noise ratio (SINR) for the jth user can be defined as

$$\mathrm{SINR}_j = \frac{\mathbf{w}_j^H \mathbf{R}_j \mathbf{w}_j}{\sum_{i \neq j}^{K} \mathbf{w}_i^H \mathbf{R}_j \mathbf{w}_i + \sigma_n^2}, \tag{8.3}$$

where $\mathbf{R}_j = \mathrm{E}\{\mathbf{h}_j^H \mathbf{h}_j\}$. A meaningful way to solve the problem of downlink beamforming with per-antenna power and QoS constraints is to solve the following optimization problem [135]

$$\begin{aligned}
\min_{\alpha} \quad & \alpha \sum_{n=1}^{N_T} \lambda_n \\
\text{s.t.} \quad & \left[\sum_{j=1}^{K} \mathbf{w}_j \mathbf{w}_j^H\right]_{n,n} \leq \alpha \lambda_n \qquad n = 1, 2, \ldots, N_T \\
& \frac{\mathbf{w}_j^H \mathbf{R}_j \mathbf{w}_j}{\sum_{i \neq j}^{K} \mathbf{w}_i^H \mathbf{R}_j \mathbf{w}_i + \sigma_n^2} \geq \gamma_j \\
& j = 1, 2, \ldots, K.
\end{aligned} \tag{8.4}$$

where γ_j is the minimum acceptable QoS for the jth user and λ_n is the maximum transmission power for the nth transmit antenna. The solutions to the above problem using second-order cone programming (SOCP) has been proposed in [87]. They propose elegant iterative algorithms based on the Lagrangian dual of the SOCP of the optimization problem in (8.4). However, the above problem formulation is not optimal from a system designer point of view, since it minimizes the the maximum power margin over all the antennas. Thus, a more appropriate problem formulation is to solve the following optimization problem which is aimed at minimizing the total sum power [135]

$$\begin{aligned}
\min_{\mathbf{w}_j} \quad & \sum_{j=1}^{K} \mathbf{w}_j^H \mathbf{w}_j \\
\text{s.t.} \quad & \left[\sum_{j=1}^{K} \mathbf{w}_j \mathbf{w}_j^H\right]_{n,n} \leq \lambda_n \qquad n = 1, 2, \ldots, N_T \\
& \frac{\mathbf{w}_j^H \mathbf{R}_j \mathbf{w}_j}{\sum_{i \neq j}^{K} \mathbf{w}_i^H \mathbf{R}_j \mathbf{w}_i + \sigma_n^2} \geq \gamma_j \\
& j = 1, 2, \ldots, K.
\end{aligned} \tag{8.5}$$

However, the above problem formulation is not optimal from a system designer

point of view, since it minimizes the the maximum power margin over all the antennas. Thus, a more appropriate problem formulation is to solve the problem in (8.5), which is aimed at minimizing the total sum power.

However the solutions proposed in both [87] and [135] do not incorporate for the CSI errors. Hence, the performance of the design in [87] and [135] degrades as the quality of the available CSI worsens. In the next section we propose a robust solution to (8.5) based on worst-case performance optimization.

8.2 Robust Design

Let us assume that the true channel covariance matrices $\mathbf{R}_j$ are given as

$$\mathbf{R}_j = \hat{\mathbf{R}}_j + \mathbf{\Delta}_j,$$

where $\hat{\mathbf{R}}_j$ is the known covariance matrix at the transmitter and $\mathbf{\Delta}_j$ is the unknown uncertainty matrix which is assumed to be bounded above as

$$\|\mathbf{\Delta}_j\|_F \leq \eta_j.$$

Using the framework of worst-case performance optimization, we wish to solve the following optimization problem

$$\begin{aligned}
\min_{\mathbf{w}_j} \quad & \sum_{j=1}^{K} \mathbf{w}_j^H \mathbf{w}_j \\
\text{s.t.} \quad & \left[\sum_{j=1}^{K} \mathbf{w}_j \mathbf{w}_j^H\right]_{n,n} \leq \lambda_n \qquad n = 1, 2, \ldots, N_T \\
& \min_{\|\mathbf{\Delta}_j\|_F \leq \eta_j} \frac{\mathbf{w}_j^H(\hat{\mathbf{R}}_j + \mathbf{\Delta}_j)\mathbf{w}_j}{\sum_{i \neq j}^{K} \mathbf{w}_i^H(\hat{\mathbf{R}}_j + \mathbf{\Delta}_j)\mathbf{w}_i + \sigma_n^2} \geq \gamma_j, \quad j = 1, 2, \ldots, K. \qquad (8.6)
\end{aligned}$$

To strengthen the QoS constraints in (8.6), we introduce two separate uncertainty matrices, $\breve{\mathbf{\Delta}}_j$ and $\breve{\mathbf{\Delta}}_j$ in the numerator and the denominator of the SINR constraints respectively [105]. The resulting problem can be written as

$$\begin{aligned}
\min_{\mathbf{w}_j} \quad & \sum_{j=1}^{K} \mathbf{w}_j^H \mathbf{w}_j \\
\text{s.t.} \quad & \left[\sum_{j=1}^{K} \mathbf{w}_j \mathbf{w}_j^H\right]_{n,n} \leq \lambda_n \qquad n = 1, 2, \ldots, N_T \\
& \frac{\min_{\|\breve{\mathbf{\Delta}}_j\|_F \leq \eta_j} \mathbf{w}_j^H (\hat{\mathbf{R}}_j + \breve{\mathbf{\Delta}}_j) \mathbf{w}_j}{\max_{\|\breve{\mathbf{\Delta}}_j\|_F \leq \eta_j} \sum_{i \neq j}^{K} \mathbf{w}_i^H (\hat{\mathbf{R}}_j + \breve{\mathbf{\Delta}}_j) \mathbf{w}_i + \sigma_n^2} \geq \gamma_j, \quad j = 1, 2, \ldots, K. \quad (8.7)
\end{aligned}$$

Problem (8.7) can be further written as

$$\begin{aligned}
\min_{\mathbf{w}_j, \zeta_j} \quad & \sum_{j=1}^{K} \mathbf{w}_j^H \mathbf{w}_j \\
\text{s.t.} \quad & \left[\sum_{j=1}^{K} \mathbf{w}_j \mathbf{w}_j^H\right]_{n,n} \leq \lambda_n \qquad n = 1, 2, \ldots, N_T \\
& \min_{\|\breve{\mathbf{\Delta}}_j\|_F \leq \eta_j} \mathbf{w}_j^H (\hat{\mathbf{R}}_j + \breve{\mathbf{\Delta}}_j) \mathbf{w}_j - \zeta_j \geq 0 \\
& \zeta_j - \max_{\|\breve{\mathbf{\Delta}}_j\|_F \leq \eta_j} \gamma_j \sum_{i \neq j}^{K} \mathbf{w}_i^H (\hat{\mathbf{R}}_j + \breve{\mathbf{\Delta}}_j) \mathbf{w}_i + \gamma_j \sigma_n^2 \geq 0, \quad j = 1, 2, \ldots, K. \quad (8.8)
\end{aligned}$$

where ζ_j is an auxiliary variable. Let us first solve sub-problems in the second and third constraints of (8.8) and the optimum solution is given by (8.9) and (8.10) respectively [77, 105].

$$\begin{aligned}
\min_{\|\breve{\mathbf{\Delta}}_j\| \leq \eta_j} \quad & \mathbf{w}_j^H (\hat{\mathbf{R}}_j + \breve{\mathbf{\Delta}}_j) \mathbf{w}_j \\
= \quad & \sum_{i \neq j}^{K} \mathbf{w}_i^H (\hat{\mathbf{R}}_j - \eta_j \mathbf{I}) \mathbf{w}_i. \quad (8.9)
\end{aligned}$$

$$\begin{aligned}\max_{\|\breve{\Delta}_j\|\leq\eta_j} \quad & \sum_{i\neq j}^{K} \mathbf{w}_i^H(\hat{\mathbf{R}}_j + \breve{\Delta}_j)\mathbf{w}_i \\ = \quad & \sum_{i\neq j}^{K} \mathbf{w}_i^H(\hat{\mathbf{R}}_j + \eta_j\mathbf{I})\mathbf{w}_i. \end{aligned} \tag{8.10}$$

Substituting (8.9) and (8.10) into (8.8), and rewriting the objective and the constraints in the new variable $\mathbf{W}_j = \mathbf{w}_j\mathbf{w}_j^H$ and using the property $\mathrm{Tr}\{\mathbf{w}_j^H\hat{\mathbf{R}}_j\mathbf{w}_j\} = \mathrm{Tr}\{\hat{\mathbf{R}}_j\mathbf{W}_j\}$, we can write the final optimization based on worst-case performance optimization as

$$\begin{aligned}\min_{\mathbf{W}_j} \quad & \sum_{j=1}^{K} \mathrm{tr}\{\mathbf{W}_j\} \\ \text{s.t.} \quad & \left[\sum_{j=1}^{K} \mathbf{W}_j\right]_{n,n} \leq \lambda_n \qquad n = 1,2,\ldots,N_T \\ & \mathrm{tr}\left\{(\hat{\mathbf{R}}_j - \eta_j\mathbf{I})\mathbf{W}_j\right\} - \zeta_j \geq 0 \\ & \zeta_j - \gamma_j \sum_{i\neq j}^{K} \mathrm{tr}\left\{(\hat{\mathbf{R}}_j + \eta_j\mathbf{I})\mathbf{W}_i\right\} + \gamma_j\sigma_n^2 \geq 0 \\ & \mathbf{W}_j \geq 0, \quad \mathrm{rank}\{\mathbf{W}_j\} = 1, \quad j = 1,2,\ldots,K. \end{aligned} \tag{8.11}$$

The non-convex rank constraint is dropped using semidefinite relaxation (SDR) and the resulting optimization, belongs to a class of SDP, and can be solved using interior-point methods [32, 33]. The solutions to the problem in (8.11) may not be rank-one in general. In such cases, the so-called *randomization techniques* [132, 133] should be used to obtain approximate solutions to the original problem. However, our simulation results show that we always obtain a rank-one solution for $\mathbf{W}_i$. Therefore, we retrieve the beamforming weight vectors $\mathbf{w}_i$ from $\mathbf{W}_i$ using the principal eigenvector corresponding to its single non-zero eigenvalue.

8.3 Power Constraints Per Group of Antennas

The problem formulations may be extended to, constraints on group of antenna. Let us suppose the N antennas have been partitioned into L sets as $\mathcal{N}_1, \mathcal{N}_2, \ldots, \mathcal{N}_L$ and the power of each set $\mathcal{N}_l$ is restricted to P_l. In this case, the problem formulation is defined as

$$\begin{aligned} \min_{\mathbf{w}_j} \quad & \sum_{j=1}^{K} \mathbf{w}_j^H \mathbf{w}_j \\ \text{s.t.} \quad & \sum_{i \in \mathcal{N}_l} \left[\sum_{j=1}^{K} \mathbf{w}_j \mathbf{w}_j^H \right]_{i,i} \leq P_l \qquad n = 1, 2, \ldots, N_T \\ & \frac{\mathbf{w}_j^H \mathbf{H}_j \mathbf{w}_j}{\sum_{i \neq j}^{K} \mathbf{w}_i^H \mathbf{H}_j \mathbf{w}_i + \sigma_n^2} \geq \gamma_j, \quad j = 1, 2, \ldots, K. \end{aligned} \tag{8.12}$$

The robust design proposed in section 8.2, can be extended to the problem formulation in (8.12).

8.4 Simulation Results

We consider a wireless system with 1 BS equipped with $N_T = 8$ antennas and $K = 3$ single antenna users. We generate independent, unity power, transmit symbols which are scaled according to the power requirements prior to transmission. The frequency flat channel is fixed for each data block, however it is changed between blocks according to a zero mean complex Gaussian distribution. The noise is zero mean, complex circularly symmetric AWGN with variance $\sigma_n^2 = 1$. The maximum transmission power from each antenna is fixed to $\lambda_n = 1, n = 1, 2, \ldots, N_T$.

Fig. 8.1 depicts the required transmission power of the robust and the non-robust schemes for a wide range of SINR targets. For simplicity, identical SINR thresholds $\gamma_j = \gamma$ for all $j = 1, 2, \ldots, K$ has been used. The results have been averaged over

100 random channels. The value of ϵ_j $(j = 1, 2, \ldots, K)$ has been set to 0.2. As can be seen from the results, the robust scheme requires more transmission power as compared to the non-robust scheme. This result is not unexpected nor discouraging. This is because robust scheme always consider a worst scenario and in an attempt to achieve the target SINR for the worst possible error in the CSI, it draws more power. However the advantage of the robust scheme is demonstrated in the SINR profile in Fig. 8.1.

Fig. 8.2 depicts the probability density function for the attained SINR targets for the robust and non-robust schemes. The results have been generated using 2000 Monte-Carlo runs with $\gamma_j = 5\text{db}$ $(j = 1, 2, \ldots, K)$. For each channel covariance matrix $\hat{\mathbf{R}}_j (j = 1, 2 \ldots, K)$, the elements of the corresponding error matrix $\mathbf{\Delta}_j (j = 1, 2 \ldots, K)$ has been uniformly and randomly generated in a sphere centered at zero with the radius $\eta_j = 0.05 (j = 1, 2, \ldots, K)$. We note that the non-robust scheme violates the SINR constraint 50% of the time, where as the robust scheme scheme attains the SINR targets all the time.

8.5 Summary

We proposed a robust solution to the problem of downlink beamforming with constraints on per antenna power and quality of services. The proposed solution has been solved using worst-case performance optimization, where we assumed that the Frobenius norm of the mismatch matrices between the actual and the presumed channel covariance matrices upper bounded above by a known constant. The resulting problem is non-convex, however it is approximated as a convex problem using SDP and SDR. Simulation results confirm the superior performance of the robust scheme over the non-robust scheme.

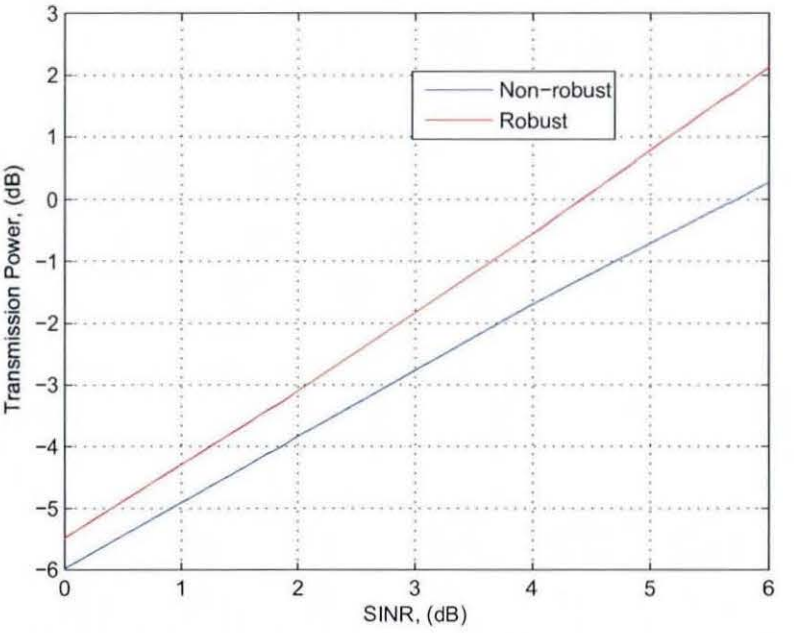

Figure 8.1: Total transmitted power versus target SINRs.

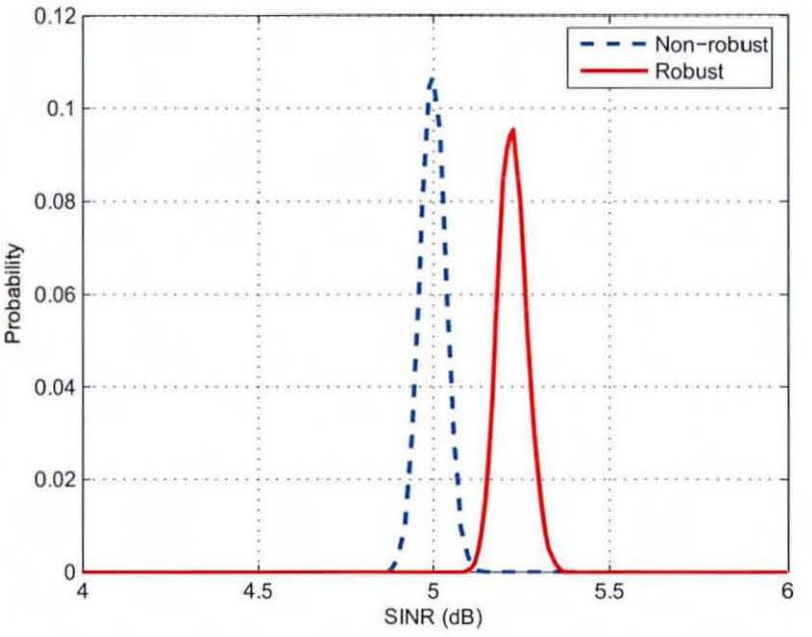

Figure 8.2: Probability density function for the attained SINR targets when $\epsilon = 0.05$.

Chapter 9

Conclusions and Future Work

9.1 Summary

Simultaneous transmission of signals to multiple users in the same frequency band within a downlink spatial multiplexing setup has been studied. Novel spatial diversity techniques have been proposed for the cases of perfect and imperfect channel state information at the transmitters. The proposed techniques covered a wide range of access schemes such as TDMA, CDMA and OFDM and various channel profiles such as flat fading and frequency selective fading. The core of the contribution was on the proposal of convex optimization based beamformer design to tackle the problem of imperfect CSI at the transmitter. In particular, the contributions can be summarized as follows:

1. An iterative method based on spatial multiplexing for a MU-MIMO downlink system using SLR criterion has been proposed. The results were presented for both flat fading and frequency selective fading channels.

2. A Channel shortening based spatial multiplexing scheme based on uplink-downlink duality has been proposed. Simulation results demonstrated the power efficiency of channel shortening based spatial multiplexing scheme over full equalization

based design.

3. A robust spatial multiplexing technique based on SLR criterion has been proposed. Results were presented for both flat fading and frequency selective MIMO channels. An application of the robust design to an OFDM scenario with quantization error has also been presented.

4. Two novel algorithms for downlink beamforming based on worst-case performance optimization and SDP constraints have been proposed. The proposed algorithms have been demonstrated to be power efficient and to satisfy the PSD constraints as opposed to the conventional robust downlink beamforming techniques based on worst-case performance optimization.

5. A robust solution to the problem of downlink beamforming with per antenna (as opposed to conventional beamforming, which has sum power constraint) and QoS constraint has been proposed. Simulations results confirm the superiority of the robust design over the non-robust design.

In chapter 4, we proposed extensions to downlink beamforming techniques based on maximization of SLR criterion. Assuming that the BS has a priori knowledge of the receiver beamformers and the forward channel of all users, we proposed an iterative optimization approach, where beamformer weight vectors obtained in the $\{n-1\}^{\text{th}}$ iteration are used to optimize the beamformers in the nth iteration. The proposed solutions provided a superior performance in terms of BER and outage probability.

In chapter 5, we proposed a channel shortening based spatial multiplexing scheme using uplink-downlink duality. We proved that the uplink-downlink duality theorem known for the flat fading channels holds for frequency selective channels. We firstly proposed an algorithm, which jointly optimizes the space-time filters and the power allocation for each user, in order to satisfy the QoS constraints of all users under a total power constraint. We then proposed an algorithm, where the space-time filters and

powers for all users are optimized jointly to minimize the total power while achieving a predefined set of QoS constraints. Simulation results confirmed the power efficiency of channel shortening based spatial multiplexing schemes over full equalization based spatial multiplexing schemes.

In chapter 6, we proposed a robust counterpart to the SLR based downlink beamforming algorithm discussed in chapter 4. We assumed that the CSI known at the transmitter is in error and modelled uncertainties in the channel covariance matrices using a convex ball, with a pre-defined radius. We used the framework of worst-case performance optimization to obtain a robust solution. The solution to this problem also turns out to be a generalarized eigenvalue problem, however with regularization factors in the numerator and denominator of the SLR criterion. We also provided closed form expressions for these regularization factors.

In chapter 7, we proposed two novel algorithms, which were built upon on the conventional worst-case performance optimization for downlink beamforming. The worst-case performance optimization violates the PSD constraints on the channel covariance matrices. Thus, the solution is over conservative as it is not restricted to a set of PSD matrices. We, therefore, propose two novel schemes which incorporate for the PSD constraints. Simulations results confirm the superiority of the proposed solutions over the worst-case based design which violates the PSD constraints.

In chapter 8, we studied the problem of downlink beamforming with per antenna power and QoS constraints. We investigated the problem under the worst-case performance optimization framework, and proposed a robust SDP based solution. Simulation results confirm superior performance of the proposed approach.

9.2 Future Work

The possible future directions include

1. Robust Spatial Multiplexing Techniques for Frequency Selective Channels

   - Our preliminary work on uplink-downlink duality based spatial multiplexing techniques can be extended to take into account possible errors in the channel state information. The validity of the uplink-downlink duality theorem for imperfect CSI should be verified analytically and worst case optimization based design can be developed for both complete equalization and channel shortening based spatial multiplexers.

2. Robust Joint Transceiver Design for Multiuser Multiplexing

   - The robust algorithms described in Chapter 6, 7 and 8, considered multiple antennas at the transmitter and single antenna for the user terminals. When user terminals also employ multiple antennas, the optimality would require the spatial multiplexers at the transmitter and the receiver filters be jointly optimized. There has been some work on the joint transceiver design for point to point MIMO systems, however, joint transceiver design for multi-user MIMO multiplexing is relatively a more challenging problem, because the optimum filters need to be designed jointly for all users together with optimum power allocation. Moreover, the problem is even more difficult if errors in the CSI are also considered. This poses a very challenging and interesting problem to be looked upon in future.

Appendix A

Worst-Case Performance Optimization

Here, we derive solutions to the problems

$$\begin{aligned} \min_{\boldsymbol{\Delta}_{1,i}} \quad & \mathbf{w}_i^H(\mathbf{R}_i + \boldsymbol{\Delta}_{1,i})\mathbf{w}_i \\ \text{s.t.} \quad & \|\boldsymbol{\Delta}_{1,i}\|_F \leq \epsilon_i, \end{aligned} \tag{A.1}$$

$$\begin{aligned} \max_{\boldsymbol{\Delta}_{2,i}} \quad & \mathbf{w}_j^H(\mathbf{R}_i + \boldsymbol{\Delta}_{2,i})\mathbf{w}_j \\ \text{s.t.} \quad & \|\boldsymbol{\Delta}_{2,i}\|_F \leq \epsilon_i. \end{aligned} \tag{A.2}$$

From the linearity of the objective function of (A.1) in the variable $\boldsymbol{\Delta}_{1,i}$, it follows that the inequality constraint $\|\boldsymbol{\Delta}_{1,i}\|_F \leq \epsilon_i$, may be replaced with the equality constraint $\|\boldsymbol{\Delta}_{1,i}\|_F = \epsilon_i$ or equivalently $\|\boldsymbol{\Delta}_{1,i}\|_F^2 = \epsilon_i^2$. Following this, the Lagrangian for (A.1) can be written as

$$\mathrm{L}(\mathbf{\Delta}_{1,i}, \lambda_i) = \mathbf{w}_i^H(\mathbf{R}_i + \mathbf{\Delta}_{1,i})\mathbf{w}_i + \lambda(\|\mathbf{\Delta}_{1,i}\|_F^2 - \epsilon_i^2) \tag{A.3}$$

Differentiating the Lagrangian (A.3), with respect to $\mathbf{\Delta}_{1,i}$ and equating the result to zero, we obtain

$$\frac{\partial \mathrm{L}}{\partial \mathbf{\Delta}_{1,i}} = \mathbf{w}_i\mathbf{w}_i^H + 2\lambda_i\mathbf{\Delta}_{1,i} = 0. \tag{A.4}$$

Hence, we obtain,

$$\mathbf{\Delta}_{1,i} = -\frac{1}{2\lambda_i}\mathbf{w}_i\mathbf{w}_i^H \tag{A.5}$$

Substituting (A.5) into the constraint $\|\mathbf{\Delta}_{1,i}\|_F^2 = \epsilon_i^2$, we obtain

$$\epsilon_i^2 = \frac{1}{4\lambda_i^2}(\|\mathbf{w}_i\|_2^2)^2 \tag{A.6}$$

Taking into account that by definition $\lambda \geq 0$, from (A.5), we obtain

$$\lambda_i = \frac{\|\mathbf{w}_i\|^2}{2\epsilon_i} \tag{A.7}$$

Substituting (A.7) into (A.5), we obtain the optimum $\mathbf{\Delta}_{1,i}$ as

$$\mathbf{\Delta}_{1,i}^* = -\epsilon_i\frac{\mathbf{w}_i\mathbf{w}_i^H}{\|\mathbf{w}_i\|_2^2} \tag{A.8}$$

Subsituting (A.8) into the cost of (A.1), we obtain the minimum objective as

$$\mathbf{w}_i^H(\mathbf{R}_i - \epsilon_i\mathbf{I})\mathbf{w}_i \tag{A.9}$$

Following, the same steps us above, and taking into consideration similar arguments, one may obtain the optimum $\mathbf{\Delta}_{2,i}$ as

$$\Delta_{2,i}^{*} = \epsilon_i \frac{\mathbf{w}_j \mathbf{w}_j^H}{\|\mathbf{w}_j\|_2^2}, \tag{A.10}$$

and maximum objective to the subproblem (A.2), which are given by

$$\mathbf{w}_j^H (\mathbf{R}_i + \epsilon_i \mathbf{I}) \mathbf{w}_j. \tag{A.11}$$

The change in sign in the loading is introduced as the fact that (A.2) is a maximization problem, where as (A.1) is a minimization problem.

Bibliography

[1] V. Sharma and S. Lambotharan, "Interference suppression in multiuser downlink MIMO beamforming using an iterative optimization approach," in *Proc. EUSIPCO, Florence, Italy*, Sep. 2006.

[2] V. Sharma and S. Lambotharan, "Multiuser downlink MIMO beamforming using an iterative optimization approach," in *Proc. IEEE VTC, Montreal, Canada*, Sep. 2006.

[3] V. Sharma and S. Lambotharan, "Space time channel shortening based spatial multiplexing schemes for multiusers with multipath channels," in *Proc. IEEE Asilomar, California, USA*, pp. 1700–1704, Nov. 2007.

[4] V. Sharma and S. Lambotharan, "Space-time channel shortening based spatial multiplexing techniques using uplink-downlink duality," in *Proc. IEEE Vehicular Technology Conference, Spring, Singapore*, May 2008.

[5] V. Sharma and S. Lambotharan, "Space-time channel shortening based spatial multiplexing schemes for frequency selective multiuser channels using uplink-downlink duality," *submitted to IEEE Trans. on Sig. Process.*, April 2008.

[6] V. Sharma and S. Lambotharan, "Robust multiuser beamformers in MIMO-OFDM systems," in *Proc. IEEE ICCS, Singapore*, Oct.-Nov. 2006.

[7] V. Sharma, S. Lambotharan, and A. Jakabsson, "Robust transmit multiuser beamforming using worst case performance optimization," in *Proc. IEEE Vehicular Technology Conference, Spring, Singapore*, May 2008.

[8] V. Sharma, I. Wajid, A. B. Gershman, H. Chen, and S. Lambotharan, "Robust downlink beamforming using positive semi-definite covariance constraints," in *Proc. IEEE Workshop on Smart Antennas, Darmstadt, Germany*, Feb. 2008.

[9] V. Sharma, I. Wajid, A. B. Gershman, H. Chen, and S. Lambotharan, "Robust downlink beamforming using worst-case performance optimization and positive semi-definite covariance constraints," *to be submitted to IEEE Trans. Sig. Process.*

[10] S. M. Redl, M. K. Weber, and M. W. Oliphant, *An Introduction to GSM.* U.K.: Artech House Publishers, 1998.

[11] P. Roshan and J. Leary, *802.11 Wireless LAN Fundamentals.* Cisco Press, 2004.

[12] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications.* United Kingdom: Cambridge University Press, 2003.

[13] H. Bolcskei, D. Gesbert, C. B. Papadias, and A. J. van der Veen, *Space-Time Wireless Systems.* United Kingdom: Cambridge University Press, 2006.

[14] G. Foschini, "Layered space-time architecture for wireless communications in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, pp. 41–59, 1996.

[15] J. Winters, "On the capacity of radio communication systems with diversity in a rayleigh fading environment," *IEEE J. Select. Areas Commun.*, vol. 5, pp. 871–878, June 1987.

[16] I. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. Tel.*, vol. 10, pp. 585–595, November/December 1999.

[17] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 684–702, June 2003.

[18] A. B. Greshman and N. D. Sidiropoulos, *Space-Time Processing for MIMO Communications.* United Kingdom: John Wiles and Sons Ltd, 2005.

[19] D. Gesbert, M. Shafi, D. S. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," *IEEE J. Select. Areas Commun.*, vol. 21, pp. 281–301, April 2003.

[20] J. B. Andersen, "Array gain and capacity for known random channels with multiple element arrays at both ends," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2172–2178, Nov 2000.

[21] W. Jakes, *Microwave Mobile Communications.* New York: Wiley, New York, 1974.

[22] S. Alamouti, "A simple transmit diversity technique for wireless commmunications," *IEEE J. Sel. Areas. Comm*, pp. 1451–1458, October 1998.

[23] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block coding for high rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, pp. 744–765, March 1998.

[24] A. Paulraj and T. Kailath, "Increasing capacity in wireless broadcast systems using distributed transmission/directional reception." US Patent 5 345 599, 1994.

[25] G. J. Foschini, G. D. Golden, P. W. Wolniansky, and R. A. Valenzuela, "Simplified processing for wireless communication at high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Select. Areas Commun. Wireless Commun. Series*, vol. 17, pp. 1841–1852, Nov 1999.

[26] S. Catreux, P. F. Driessen, and L. J. Greenstein, "Simulation results for an interference-limited multiple-input multiple-output cellular system," *IEEE Commun. Letters*, vol. 4, no. 11, pp. 334–336, 2000.

[27] D. Popescu and C. Rose, "Multiuser MIMO systems and interference avoidance," in *Proc. IEEE Acoustics, Speech, and Signal Processing*, vol. 4, (Hongkong), pp. 828–831, Apr 2003.

[28] J. B. Anderson, "Array gain and capacity for known random channels with multiple element arrays at both ends," *IEEE Journal on Selected Areas in Communications*, vol. 18, pp. 2172–2178, Nov. 2000.

[29] I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. on Telecommunications*, vol. 10, pp. 585–595, Nov.-Dec 1999.

[30] L. Zheng and D. N. C. Tse, "Diversty and multiplexing: A fundamental tradeoff in multiple antenna systems," *IEEE Trans. Inform. Theory*, pp. 1073–1096, May 2003.

[31] J. G. Proakis, *Digital Communications.* New York: McGraw Hill, 1995.

[32] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming." avialable at http://www.stanford.edu/boyd/cvx/V.1.0RC3, Feb. 2007.

[33] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optim. Meth. Software*, vol. 11-12, pp. 625–653, Aug. 1999.

[34] M. Schubert and H. Boche, "Solution of multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Vehicular Tech.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[35] V. Poor and G. Wornell, *Wireless Communications: Signal Processing Perspectives.* Upper Saddle River, NJ: Prentice Hall, 1998.

[36] J. Cioffi, "Class reader for ee379a-digital communication.." URL http://www.stanford.edu/class/ee379a., 2002.

[37] E. Lindskog and A. Paulraj, "A transmit diversity scheme for channels with intersymbol interference," in *Proc. ICC*, pp. 307–311, 2000.

[38] N. Al-Dhahir, "Single-carrier frequency-domain equalization for space-time block-coded transmission over frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 5, no. 7, pp. 304–306, 2001.

[39] T. Lo, "Maximal ratio transmisson," *IEEE Tran. Comm.*, vol. 47, pp. 1458–1461, Oct. 99.

[40] X. Feng and C. Leung, "A new optimal transmit and receive diversity scheme.," in *Proc. IEEE PACRIM*, vol. 2, (Victoria, Cananda), pp. 538–541, Aug. 2001.

[41] K. H. Lee and D. P. Petersen, "Optimal linear coding for vector channels," *IEEE Trans. Commun.*, vol. 24, p. 12831290, Dec. 1976.

[42] J. Salz, "Digital transmission over cross-coupled linear channels," *At&T Tech. J.*, vol. 64, p. 11471159, July-Aug 1985.

[43] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (mimo) transmission systems," *IEEE Trans. Commun.*, vol. 42, p. 32213231, Dec. 1994.

[44] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers part i: Unification and optimal designs," *IEEE Trans. Sig. Process.*, vol. 47, p. 19882006, July 1999.

[45] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for mimo channels using the weighted mmse criterion," *IEEE Trans. Comms.*, vol. 49, pp. 2198–2206, Dec. 2001.

[46] D. P. Palomar, J. M. Coffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans Sig. Process.*, vol. 51, pp. 2381–2401, Sep. 2003.

[47] D. P. Palomar, *A unified framework for communications through MIMO channels.* PhD thesis, Technical University Catalonia (UPC), Barcelona, Spain, 2003.

[48] C. Toker, S. Lambotharan, and J. A. Chambers, "Joint transceiver design for MIMO channel shortening," *IEEE Trans. Sig. Process.*, vol. 55, pp. 3851–3866, July 2007.

[49] T. Haustein, C. von Helmolt, E. Jonvieck, V. Jungnickel, and V. Pohl, "Performance of mlmo systems with channel inversion.," in *Proc. IEEE VTC.*, May 2002.

[50] D. Gerlach and A. Pualraj, "Adaptive transmitting antenna arrays with feedback," *IEEE Sig. Proc. Letters*, vol. 1, pp. 150–152, Oct. 1994.

[51] C. B. Peel, B. M. Hochwald, and A. Swindlehurst, "A vector-perturbation technique for near-capacity multi-antenna multi-user communication-part I: channel inversion and regularization," *IEEE Trans. Comm.*, vol. 53, pp. 195–202, Jan. 2003.

[52] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439–441, May 1983.

[53] C. B. Peel, "On "writing on dirty paper"," *IEEE Sig. Process. Mag.*, vol. 20, pp. 112–113, May 2003.

[54] B. M. Hochwald, C. B. Peel, and A. Swindlehurst, "A vector-perturbation technique for near-capacity multi-antenna multi-user communication-part II: Perturbation," *IEEE Trans. Comm.*, vol. 53, pp. 537–544, March 2005.

[55] A. Bourdoux and N. Khaled, "Joint Tx-Rx optimization for MIMO-SDMA based on a null-space constraint," in *Proc. IEEE VTC*, Sep. 2002.

[56] M. Rim, "Multi-user downlink beamforming with multiple transmit and receive antennas," *Elecironics Letters*, vol. 38, pp. 1725–1726, Dec. 2002.

[57] R. Choi and R. Murch, "A transmit preprocessing technique for multiuser mimq systems using a decomposition approach," *IEEE Trans. Wireless Comm.*, vol. 3, pp. 20–24, Jan. 2004.

[58] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Sig. Process.*, vol. 52, pp. 461–471, Feb. 2004.

[59] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Harrdt, "An introduction to the multiuser MIMO downlink," *IEEE Comm. Mag.*, vol. 42, pp. 60–67, Oct. 2004.

[60] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE Transceiver Optimization for Multiuser MIMO Systems: Duality and Sum-MSE Minimization," *IEEE Trans. Sig. Process.*, vol. 55, no. 11, pp. 5436–5446, Sep. 2007.

[61] S. Boyd and L. Vandenberghe, *Convex Optimization.* United Kingdom: Cambridge Univewrsity Press, 2004.

[62] D. G. Luenberger, *Linear and Nonlinear Programming.* Reading, Massachusetts: Addison-Wesley Publishing Co., 1984.

[63] Y. Nesterov and A. Nemirovski, "Interior-point polynomial algorithms in convex programming," *SIAM Stud. Appl. Math.*, vol. 13, 1994.

[64] D. G. Luenberger, *Optimization by Vector Space Methods.* New York: Wiley, 1969.

[65] Y. Ye, *Interior Point Algorithms. Theory and Analysis.* John Wiley & Sons, 1997.

[66] A. Ben-Tal and A. Nemirovskii, "Robust convex optimization," *Math. Oper. Res.*, vol. 23, pp. 769–805, 1998.

[67] L. E. Ghaoui, F. Oustry, and H. Lebret, "Robust solutions to uncertain semidefinite programs," *SIAM J. Optimization*, vol. 9, no. 1, pp. 33–52, 1998.

[68] A. Ben-Tal, L. E. Ghaoui, and A. Nemirovskii, "Robust semidefinite programming," *Handbook of Semidefinite Programming*, March 2000.

[69] A. Narula, M. J. Lopez, M. D.Trot, and G. W. Wornell, "Efficient use of side information in multiple-antenna data transmission over fading channels," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1423–1436, Oct. 1998.

[70] A. Wittneben, "Optimal predictive tx combining diversity in correlated fading for microcellular mobile radio applications," in *Proc. IEEE Global Telecommunications Conference (GLOBECOM 1995)*, pp. 13–17, Nov. 1995.

[71] Y. Yong, S. A. Vorobyov, and A. B. Gershman, "Robust linear receivers for multiaccess space-time block-coded mimo systems: A probabilistically constrained approach," *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 1560–1570, Aug. 2006.

[72] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proc. IEEE*, vol. 73, pp. 433–481, Mar. 1984.

[73] S. Verdú and V. Poor, "On minimax robustness: A general approach and applications," *IEEE Trans. Inf. Theory*, vol. 30, pp. 328–340, Mar. 1984.

[74] S. Shahbazpanahi, A. B. Gershman, Z. Q. Luo, and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans on Sig. Process.*, vol. 51, pp. 2257–2269, Sep. 2003.

[75] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. on Sig. Process.*, vol. 53, pp. 1684–1696, May 2005.

[76] J. Li, P. Stoica, and Z. Wang, "On robust capon beamforming and diagonal loading," *IEEE Trans. on Sig. Process.*, vol. 51, pp. 1702–1715, July 2003.

[77] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," *Handbook of Antennas in Wireless Communications*, 2001.

[78] M. Bengtsson and B. Ottersten, "Optimal downlink beamforming using semidefinite optimization," in *Proc. 37th Allerton Conf. Comm. Control and Comp.*, Sept. 1999.

[79] M. Biguesh, S. Shahbazpanahi, and A. B. Gershman, "Robust downlink power control in wireless cellular systems," *EURASIP J. Wireless Communications and Networking*, pp. 261–272, Dec. 2004.

[80] F. Rashid-Farrokhi, L. Tassiulas, and K. J. R. Liu, "Transmit beamforming and power control for cellular wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1437–1450, Oct. 1998.

[81] F. Rashid-Farrokhi, L. Tassiulas, and K. J. R. Liu, "Joint optimal power control and beamforming in wireless networks using antenna arrays," *IEEE J. Sel. Areas Commun.*, vol. 46, no. 10, pp. 1313–1324, Oct. 1998.

[82] E. Visotsky and U. Madhow, "Optimum beamforming using transmit antenna arrays," in *Proc. IEEE Veh. Technol. Conf.*, vol. 1, pp. 851–856, 1999.

[83] M. Bengtsson, "A pragmatic approach to multiuser spatial multiplexing," in *Proc. IEEE Sensor Array and Multichannel Sig. Process. Workshop*, (Rosslyn, USA), pp. 130–134, Aug. 2002.

[84] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed mimo receivers," *IEEE Trans. Signal Process.*, vol. 54, pp. 161–176, Jan. 2006.

[85] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra Applicat.*, vol. 284, pp. 193–228, Nov. 1998.

[86] M. Schubert and H. Boche, "Iterative multiuser uplink and downlink beamforming under SINR constraints," *IEEE Trans. Signal. Process.*, vol. 53, pp. 2324–2334, Jul. 2005.

[87] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, pp. 2646–2660, June 2007.

[88] L. E. Brennan, J. D. Mallet, and I. S. Reed, "Adaptive arrays in airborne mti radar," *IEEE Trans. Antennas Propagation*, vol. 24, pp. 607–615, Sept. 1976.

[89] H. Cox, "Resolving power and sensitivity to mismatch of optimum array processors," *J. Acoust. Soc. Amer.*, vol. 54, pp. 771–758, 1973.

[90] J. L. Krolik, "The performance of matched-field beamformers with mediterranean vertical array data," *IEEE Trans. Signal Processing*, vol. 44, pp. 2605–2611, Oct. 1996.

[91] A. B. Gershman, V. I. Turchin, and V. A. Zverev, "Experimental results of localization of moving underwater signal by adaptive beamforming," *IEEE Trans. Signal Processing*, vol. 43, pp. 2249–2257, Oct. 1995.

[92] E. Y. Gorodetskaya, A. I. Malekhanov, A. G. Sazontov, and N. K. Vdovicheva, "Deep-water acoustic coherence at long ranges: Theoretical prediction and effects on large-array signal processing," *IEEE J. Ocean. Eng.*, vol. 24, pp. 156–171, Apr. 1999.

[93] A. B. Gershman, E. Nemeth, and J. F. Bhme, "Experimental performance of adaptive beamforming in a sonar environment with a towed array and moving interfering sources," *IEEE Trans. Sig. Process.*, vol. 48, pp. 246–250, Jan. 2000.

[94] J. Capon, R. J. Greenfield, and R. J. Kolker, "Multidimensional maximum-likelihood processing for a large aperture seismic array," *Proc. IEEE*, vol. 55, pp. 192–211, Feb. 1967.

[95] Y. Kameda and J. Ohga, "Adaptive microphone-array system for noise reduction," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1391–1400, Dec. 1986.

[96] L. C. Godara, "Application of antenna arrays to mobile communications. ii. beam-forming and direction-of-arrival considerations," *Proc. IEEE*, vol. 85, pp. 1195–1245, Aug. 1997.

[97] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-10, pp. 853–863, Nov. 1974.

[98] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 27–34, Jan. 1982.

[99] E. K. Hung and R. M. Turner, "A fast beamforming algorithm for large arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-19, pp. 598–607, July 1983.

[100] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, pp. 397–401, July 1998.

[101] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*. New York: Wiley, 1980.

[102] D. D. Feldman and L. J. Griffiths, "A projection approach to robust adaptive beamforming," *IEEE Trans. Sig. Process.*, vol. 42, pp. 867–876, Apr. 1994.

[103] H. Cox, R. M. Zeskind, and M. H. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1365–1376, Oct. 1987.

[104] A. B. Gershman, "Robust adaptive beamforming in sensor arrays," *Int. J. Electron. Commun.*, vol. 53, pp. 305–314, Dec. 1999.

[105] S. A. Vorobyov, A. B. Gershman, and Z. Q. Luo, "Robust adaptive beamforming using worst-case performance optimization," *IEEE Trans. Signal Process.*, vol. 51, pp. 313–324, Feb. 2003.

[106] B. K. Chalise, S. Shahbazpanahi, A. Czylwik, and A. B. Gershman, "Robust downlink beamforming based on outage probability specifications," *IEEE Trans. Wireless Comms.*, vol. 6, pp. 3498–3503, Oct. 2007.

[107] A. Alexiou and M. Haardt, "Smart antenna technologies for future wireless systems: trends and challenges," *IEEE Commun. Mag.*, vol. 42, pp. 90–97, Sept. 2004.

[108] A. Tarighat, M. Sadek, and A. H. Sayed, "A multiuser beamforming scheme for downlink MIMO channels based on maximizing signal-to-leakage ratios," in *Proc. IEEE ICASSP05*, (Philadelphia, USA), pp. 1129–1132, 2005.

[109] R. Chen, J. G. Andrews, and R. W. Heath, "Multiuser spacetime block coded MIMO system with unitary downlink precoding," in *Proc. IEEE ICC*, (Paris, France), pp. 2689–2693, June 2004.

[110] K. K. Wong, R. D. Murch, and K. B. Letaief, "Performace enhancement of multiuser MIMO wireless communication systems," *IEEE Trans. Comms.*, vol. 50, pp. 1960–1970, Dec. 2002.

[111] G. H. Golub and C. F. V. Loan, *Matrix Computations (3rd edition).*

[112] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications.* U.K.: Cambridge University Press, 2003.

[113] A. Pandharipande, "Principles of OFDM," *IEEE Potentials*, vol. 21, pp. 16–19, Apr-May 2002.

[114] G. Raleigh and J. M. Cioffi, "Spatial-temporal coding for wireless communications," *IEEE Trans. Commun.*, vol. 46, pp. 357–366, 1998.

[115] P. Vishwanath and D. Tse, "Sum capacity of the multiple antenna gaussian broadcast channel and uplink-downlink duality," *IEEE. Trans. Inf. Theory.*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

[116] P. J. M. Melsa, R. C. Younce, and C. E. Rohrs, "Impulse response shortening for discrete multitone transceivers," *IEEE Trans. Communs.*, vol. 44, no. 12, pp. 1662–1672, Dec. 1996.

[117] N. Al-Dhahir and J. M. Cioffi, "A bandwidth-optimized reduced-complexity equalized multicarrier transceivers," *IEEE Trans. Commun.*, vol. 45, pp. 948–956, Aug. 1997.

[118] G. Arslan, B. L. Evans, and S. Kiaei, "Optimum channel shortening for discrete multitone transceivers," in *Proc. IEEE ICASSP*, vol. 5, pp. 2965–2968, June 2000.

[119] D. Daly, C. Heneghan, and A. D. Fagan, "Minimum mean-sqaured error impulse response shortening for discrete multitone transceivers," *IEEE Trans. Sig. Process.*, vol. 52, no. 1, pp. 301–306, Jan. 2004.

[120] C. Toker, S. Lambotharan, and J. A. Chambers, "Joint spatial and temporal channel shortening techniques for frequency selective fading MIMO channels," *IEE. Proc. Commun.*, vol. 152, no. 1, pp. 89–94, Feb. 2005.

[121] W. Yang and G. Xu, "Optimal downlink power assignment for smart antenna systems," in *Proc. IEEE ICASSP*, May 1998.

[122] J. Li and P. Stoica, *Robust Adaptive Beamforming.* United Kingdom: Wiley Series in Telecommunications and Signal Processing, 2006.

[123] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2632–2639, Sep. 2001.

[124] S. M. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory.* Englewood Cliffs, N.J.: Prentice-Hall, 1993.

[125] C. Farsakh and J. A. Nossek, "Spatial covariance based downlink beamforming in an sdma mobile radio system," *IEEE Trans. Communs.*, vol. 46, pp. 1497–1506, Nov. 1998.

[126] H. Boche and M. Schubert, "A new approach to power adjustment for spatial covariance based downlink beamforming," in *Proc. ICASSP 01*, pp. 2957–2960, April 2001.

[127] M. Bengtsson, "Jointly optimal downlink beamforming and base station assignment," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 2961–2964, May 2001.

[128] G. Strang, *Linear Algebra and Its Applications, 3rd Ed.* Thomson Learning, Inc, 1988.

[129] K. Anstreicher and H. Wolkowicz, "On Lagrangian relaxation of quadratic matrix constraints," *SIAM Journal on Matrix Analysis and Applications*, vol. 22, no. 1, pp. 41–55, 2001.

[130] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Review*, vol. 38, no. 1, pp. 49–95, 1996.

[131] J. Löfberg, "Yalmip: A toolbox for modelling and optimization in matlab," in *Proc. CACSD*, 2004.

[132] S. Zhang, "Quadratic maximization and semidefinite relaxation," *Math. Program.*, vol. 87, pp. 453–465, 2000.

[133] P. Tseng, "Further results on approximating nonconvex quadratic optimization by semidefinite programming relaxation," *SIAM J. Optim*, vol. 14, pp. 268–283, July 2003.

[134] F. R. Farrokhi, K. J. R. Liu, and L. Tassiulas, "Downlink power control and base station assignment," *IEEE Comms. Letters*, vol. 1, pp. 102–104, July 1997.

[135] Y. K. Yazarel and D. Aktas, "Downlink beamforming under individual SINR and per antenna power constraints," in *Proc. IEEE Pacific Rim conference on Communications, Computers and Signal Processing*, pp. 422–425, Aug. 2007.